

# Technical Design Report

The *BABAR* Collaboration

March, 1995

This document and the material and data contained therein, was developed under sponsorship of the United States Government. Neither the United States nor the Department of Energy, nor the Leland Stanford Junior University, nor their employees, nor their respective contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any liability or responsibility for accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use will not infringe privately owned rights. Mention of any product, its manufacturer, or suppliers shall not, nor is intended to, imply approval, disapproval, or fitness for any particular use. A royalty-free, nonexclusive right to use and disseminate same for any purpose whatsoever, is expressly reserved to the United States and the University.

---

# The BABAR Collaboration

**LAPP Annecy, *Annecy-le-Vieux, France***

D. Boutigny, Y. Karyotakis, S. Lees-Rosier, P. Petitpas

**INFN, Sezione di Bari and Università di Bari, *Bari, Italy***

C. Evangelista, A. Palano

**Beijing Glass Research Institute, *Beijing, China***

G. Chen, Y.T. Wang, O. Wen

**Institute of High Energy Physics, *Beijing, China***

Y.N. Guo, H.B. Lan, H.S. Mao, N.D. Qi, W.G. Yan, C.C. Zhang, W.R. Zhao, Y.S. Zhu

**University of Bristol, *Bristol, UK***

N. Dyce, B. Foster, R.S. Gilmore, C.J.S. Morgado

**University of Bergen, *Bergen, Norway***

G. Eigen

**University of British Columbia, *Vancouver, British Columbia, Canada***

C. Goodenough, C. Hearty, J. Heise, J.A. McKenna

**Brunel University, *London, UK***

T. Champion, A. Hasan, A.K. McKemey

**Budker Institute of Nuclear Physics, *Novosibirsk, Russia***

A.R. Buzykaev, V.N. Golubev, V.N. Ivanchenko, S.G. Klimenko, E.A. Kravchenko, G.M. Kolachev, A.P. Onuchin, V.S. Panin, S.I. Serednyakov, A.G. Shamov, Ya.I. Skovpen, V.I. Telnov

**California Institute of Technology, *Pasadena, California, USA***

D.G. Hitlin, J. Oyang, F.C. Porter, M. Weaver, A.J. Weinstein, R. Zhu

**University of California, Davis, *Davis, California, USA***

F. Rouse

**University of California, IIRPA, *La Jolla, California, USA***

A.M. Eisner, M. Sullivan, W. Vernon, Y.-X. Wang

**University of California, Irvine, Irvine, California, USA**

K. Gollwitzer, A. Lankford, M. Mandelkern, G. McGrath, J. Schultz,  
D. Stoker, G. Zioulas

**University of California, Los Angeles, Los Angeles, California, USA**

K. Arisaka, C. Buchanan, J. Kubic, W. Slater

**University of California, San Diego La Jolla, California, USA**

V. Sharma

**University of California, Santa Barbara, Santa Barbara, California, USA**

D. Bauer, D. Caldwell, A. Lu, H. Nelson, J. Richman, D. Roberts, M. Witherell, S. Yellin

**University of California, Santa Cruz, Santa Cruz, California, USA**

J. DeWitt, D. Dorfan, A.A. Grillo, C. Heusch, R.P. Johnson, E. Kashigin, S. Kashigin, W. Kroeger,  
W. Lockman, K. O'Shaughnessy, H. Sadrozinski, A. Seiden, E. Spencer

**Carleton University and CRPP<sup>†</sup>, Ottawa, Ontario, Canada**

K. Edwards, D. Karlen, M. O'Neill<sup>†</sup>

**University of Cincinnati, Cincinnati, Ohio, USA**

S. Devmal, B.T. Meadows, A.K.S. Santha, M.D. Sokoloff

**University of Colorado, Boulder, Colorado, USA**

A. Barker, B. Broomer, E. Erdos, W. Ford, U. Nauenberg, H. Park, P. Rankin, J. Roy, J.G. Smith

**Colorado State University, Fort Collins, Colorado, USA**

J. Harton, R. Malchow, M. Smy, H. Staengle, W. Toki, D. Warner, R. Wilson

**Technische Universität Dresden, Institut für Kern- und Teilchenphysik,**

*Dresden, Germany*

J. Brose, G. Dahlinger, P. Eckstein, K.R. Schubert, R. Schwierz, R. Seitz, R. Waldi

**Joint Institute for Nuclear Research, Dubna, Russia**

A. Bannikov, S. Baranov, I. Boyko, G. Chelkov, V. Dodonov, Yu. Gornushkin, M. Ignatenko,  
N. Khovansky, Z. Krumstein, V. Malyshev, M. Nikolenko, A. Nozdrin, Yu. Sedykh, A. Sissakian,  
Z. Silagadze, V. Tokmenin, Yu. Yatsunenکو

---

**University of Edinburgh, Edinburgh, UK**

K. Peach, A. Walker

**INFN, Sezione di Ferrara, Ferrara, Italy**

L. Piemontese

**Laboratori Nazionali di Frascati dell' INFN, Frascati, Italy**

R. Baldini, A. Calcaterra, R. De Sangro, I. Peruzzi (also Univ. Perugia), M. Piccolo, A. Zallo

**INFN, Sezione di Genova and Università di Genova, Genova, Italy**

A. Buzzo, R. Contri, G. Crosetti, P. Fabbri, S. Farinon, R. Monge, M. Olcese, R. Parodi, S. Passaggio, C. Patrignani, M.G. Pia, C. Salvo, A. Santroni

**University of Iowa, Iowa City, Iowa, USA**

U. Mallik, E. McCliment, M.-Z. Wang

**Iowa State University, Ames, Iowa, USA**

H.B. Crawley, A. Firestone, J.W. Lamsa, R. McKay, W.T. Meyer, E.I. Rosenberg

**Northern Kentucky University, Highland Heights, Kentucky, USA**

M. Falbo-Kenkel

**University of Lancaster, Lancaster, UK**

C.K. Bowdery, A.J. Finch, F. Foster

**Lawrence Berkeley Laboratory, Berkeley, California, USA**

G.S. Abrams, D. Brown, T. Collins, C.T. Day, S.F. Dow, F. Goosen, R. Jacobsen, R.C. Jared, J. Kadyk, L.T. Kerth, I. Kipnis, J.F. Kral, R. Lafever, R. Lee, M. Levi, L. Luo, G.R. Lynch, M. Momayezi, M. Nyman, P.J. Oddone, W.L. Pope, M. Pripstein, D.R. Quarrie, J. Rasson, N.A. Roe, M.T. Ronan, W.A. Wenzel, S. Wunduke

**Lawrence Livermore National Laboratory, Livermore, California, USA**

O. Alford, J. Berg, R.M. Bionta, A. Brooks, F.S. Dietrich, O.D. Fackler, M.N. Kreisler, M.A. Libkind, M.J. Mugge, T. O'Connor, L. Pedrotti, X. Shi, W. Stoeffl, K. van Bibber, T.J. Wenaus, D.M. Wright, C.R. Wuest, R.M. Yamamoto

**University of Liverpool, Liverpool, UK**

J.R. Fry, E. Gabathuler, R. Gamet, A. Muir, P. Sanders

**University of London, Imperial College of Science, Technology and Medicine,**  
*London, UK*

P. Dornan, A. Duane, L. Moneta, J. Nash, D. Price

**University of London, Queen Mary & Westfield College,** *London, UK*

D.V. Bugg, P.F. Harrison, I. Scott, B. Zou

**University of London, Royal Holloway & Bedford New College,** *Egham, Surrey, UK*

Y. Gao, M.G. Green, D.L. Johnson, E. Tetteh-Lartey

**University of Louisville,** *Louisville, Kentucky, USA*

C.L. Davis

**McGill University,** *Montréal, Quebec, Canada*

D. Britton, R. Fernholz, D. MacFarlane, P. Patel, C. Smith, B. Spaan, J. Trischuk

**University of Manchester,** *Manchester, UK*

J. Allison, R. Barlow, G. Lafferty, K. Stephens

**University of Maryland,** *College Park, Maryland, USA*

C. Dallapiccola, M. Foucher, H. Jawahery, A. Skuja

**University of Massachusetts,** *Amherst, Massachusetts, USA*

J. Button-Shafer, J.-J. Gomez-Cadenas, S.S. Hertzbach, R.R. Kofler, M.G. Strauss

**Massachusetts Institute of Technology,** *Cambridge, Massachusetts, USA*

R.F. Cowan, M.J. Fero, R.K. Yamamoto

**INFN, Sezione di Milano and Università di Milano,** *Milano, Italy*

M. Calvi, C. Cattadori, R. Diaferia, F. Lanni, C. Matteuzzi, F. Palombo, A. Sala, T. Tabarelli

**University of Mississippi,** *Oxford, Mississippi, USA*

M. Boone, S. Bracker, L. Cremaldi, K. Gounder, R. Kroeger, J. Reidy, D. Summers

**Université de Montréal,** *Montréal, Quebec, Canada*

G. Beaudoin, M. Beaulieu, B. Lorazo, J.P. Martin, P. Taras, V. Zacek

**Mount Holyoke College,** *South Hadley, Massachusetts, USA*

H. Nicholson, C.S. Sutton

**INFN, Sezione di Napoli and Università di Napoli,** *Napoli, Italy*

N. Cavallo, L. Lista, S. Mele, P. Parascandolo, C. Sciacca

---

**University of Notre Dame, Notre Dame, Indiana, USA**

J.M. Bishop, N.N. Biswas, N.M. Cason, J.M. LoSecco, A.H. Sanjari, W.D. Shephard

**Oak Ridge National Laboratory/Y-12, Oak Ridge, Tennessee, USA**

F.S. Alsmiller, R.G. Alsmiller, Jr., T.A. Gabriel, J.L. Heck

**LAL Orsay, Orsay, France**

D. Breton, R. Cizeron, S. Du, A.-M. Lutz, J.M. Noppe, S. Plaszczynski, M.-H. Schune,  
E. Torassa, K. Truong, G. Wormser

**INFN, Sezione di Padova and Università di Padova, Padova, Italy**

F. Dal Corso, M. Morandin, M. Posocco, R. Stroili, C. Voci

**Ecole Polytechnique Palaiseau, LPNHE, Palaiseau, France**

L. Behr, G. Bonneaud, P. Matricon, G. Vasileiadis, M. Verderi

**LPNHE des Universités Paris 6 et Paris 7, Paris, France**

M. Benayoun, H. Briand, J. Chauveau, P. David, C. De La Vaissiere, L. Del Buono, J.F. Genat,  
O. Hamon, P. Leruste, J. Lory, J.-L. Narjoux, B. Zhang

**INFN, Sezione di Milano and Università di Pavia, Pavia, Italy**

P.F. Manfredi, V. Re, V. Speziali, F. Svelto

**University of Pennsylvania, Philadelphia, Pennsylvania, USA**

L. Gladney

**INFN, Sezione di Pisa, Università di Pisa<sup>†</sup> and Scuola Normale Superiore<sup>‡</sup>,  
Pisa, Italy**

G. Batignani<sup>†</sup>, S. Bettarini, F. Bosi, U. Bottigli<sup>†</sup>, M. Carpinelli, F. Costantini<sup>†</sup>,  
F. Forti, D. Gambino, M. Giorgi<sup>†</sup>, A. Lusiani<sup>†</sup>, P.S. Marrocchesi, M. Morganti<sup>†</sup>,  
G. Rizzo, G. Triggiani<sup>†</sup>, J. Walsh

**Prairie View A&M University, Prairie View, Texas, USA**

M. Gui, D.J. Judd, K. Paick, D.E. Wagoner

**Princeton University, Princeton, New Jersey, USA**

C. Bula, C. Lu, K.T. McDonald

**INFN, Istituto Superiore di Sanità, Roma, Italy**

C. Bosio

**INFN, Sezione di Roma and Università “La Sapienza,” Roma, Italy**

F. Ferroni, E. Lamanna, M.A. Mazzoni, S. Morganti, G. Piredda, R. Santacesaria

**Rutgers University, Rutgers, New Jersey, USA**

P. Jacques, M. Kalelkar, R. Plano, P. Stamer

**Rutherford Appleton Laboratory, Chilton, Didcot, UK**P.D. Dauncey, J. Dowdell, B. Franek, N.I. Geddes, G.P. Gopal, R. Halsall,  
J.A. Lidbury, V.J. Perera**CEA, DAPNIA, CE-Saclay,<sup>1</sup> Gif-sur-Yvette, France**R. Aleksan, P. Besson, T. Bolognese, P. Bourgeois, A. de Lesquen, A. Gaidot, L. Gosset,  
G. Hamel de Monchenault, P. Jarry, G. London, M. Turluer, G. Vasseur, C. Yeche, M. Zito**Shanghai Institute of Ceramics (SICCAS), Shanghai, China**

J.R. Jing, P.J. Li, D.S. Yan, Z.W. Yin

**University of South Carolina, Columbia, South Carolina, USA**

M.V. Purohit, J. Wilson

**Stanford Linear Accelerator Center, Stanford, California, USA**D. Aston, R. Becker-Szendy, R. Bell, E. Bloom, C. Boeheim, A. Boyarski, R.F. Boyce,  
D. Briggs, F. Bulos, W. Burgess, R.L.A. Cottrell, D.H. Coward, D.P. Coupal, W. Craddock,  
H. DeStaebler, J.M. Dorfan, W. Dunwoodie, T. Fieguth, D. Freytag, R. Gearhart, T. Glanzman,  
G. Godfrey, G. Haller, J. Hewett, T. Himel, J. Hoefflich, W. Innes, C.P. Jessop, W.B. Johnson,  
H. Kawahara, L. Keller,<sup>2</sup> M.E. King, J. Krebs, P. Kunz, W. Langeveld, E. Lee, D.W.G.S. Leith,  
V.G. Lüth, H. Lynch, H. Marsiske, T. Mattison, R. Melen, K. Moffeit, L. Moss, D. Muller,  
M. Perl, G. Oxoby, M. Pertsova, H. Quinn, B.N. Ratcliff, S.F. Schaffner, R.H. Schindler,  
S. Shapiro, C. Simopolous, A.E. Snyder, E.J. Soderstrom, J. Vav'ra, S. Wagner, D. Walz,  
R. Wang, J.L. White, W. Wisniewski, N. Yu**Stanford University, Stanford, California, USA**

P. Burchat, R. Zaliznyak

---

<sup>1</sup>Subject to approval of funding agency.<sup>2</sup>Retired

**Academia Sinica, Taipei, Taiwan**

H.-Y. Chau, M.-L. Chu, S.-C. Lee

**University of Texas at Dallas, Richardson, Texas, USA**

J.M. Izen, X. Lou

**INFN, Sezione di Torino and Università di Torino, Torino, Italy**

F. Bianchi, D. Gamba, G. Giraudo, A. Romero

**INFN, Sezione di Trieste and Università di Trieste, Trieste, Italy**

L. Bosisio, R. Della Marina, G. Della Ricca, B. Gobbo, L. Lanceri, P. Poropat

**TRIUMF, Vancouver, British Columbia, Canada**

R. Henderson, A. Trudel

**Tsinghua University, Beijing, China**

Y.P. Kuang, R.C. Shang, B.B. Shao, J.J. Wang

**Vanderbilt University, Nashville, Tennessee, USA**

R.S. Panvini, T.W. Reeves, P.D. Sheldon, M.S. Webster

**University of Victoria, Victoria, British Columbia, Canada**

M. McDougald, D. Pitman

**University of Wisconsin, Madison, Wisconsin, USA**

H.R. Band, J.R. Johnson, R. Prepost, G.H. Zapalac

**York University, Toronto, Ontario, Canada**

W. Frisken

Babar and the distinctive likeness are trademarks of  
Laurent de Brunhoff and are used with his permission.

Copyright © Laurent de Brunhoff  
All Rights Reserved



---

---

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Detector Overview</b>	<b>7</b>
2.1	Introduction . . . . .	7
2.2	General Design Considerations . . . . .	8
2.2.1	Acceptance . . . . .	8
2.2.2	Charged Track Resolution and Multiple Scattering . . . . .	9
2.2.3	Photon Efficiency and Resolution . . . . .	11
2.2.4	Identification of Hadrons . . . . .	12
2.2.5	Interaction with the Accelerator . . . . .	13
2.2.6	Considerations of Cost and Schedule . . . . .	14
2.3	The Experimental Design . . . . .	14
2.4	Detector Optimization . . . . .	20
2.4.1	Impact of Particle Identification . . . . .	20
2.4.2	Optimization of Performance versus Cost . . . . .	21
2.5	Detector Performance . . . . .	25
2.5.1	Vertex Detector . . . . .	25
2.5.2	Drift Chamber . . . . .	28
2.5.3	Particle Identification . . . . .	30
2.5.4	Electromagnetic Calorimeter . . . . .	32
2.5.5	Muon and Neutral Hadron Detector . . . . .	37
2.5.6	Electronics, Trigger, and Data Acquisition . . . . .	39

2.6	Physics Performance . . . . .	39
2.6.1	Acceptance and Mass Resolution for Decays to Charged Particles . . . . .	40
2.6.2	Separation between $B$ Decay Vertices . . . . .	41
2.6.3	$\pi^0$ Efficiency and Resolution . . . . .	42
2.6.4	Lepton Identification . . . . .	43
2.6.5	Charged Hadron Identification . . . . .	43
2.7	Performance for Non- $CP$ Physics . . . . .	44
2.7.1	Other $B$ Physics . . . . .	45
2.7.2	Detector Issues for Charm Physics . . . . .	46
2.7.3	Detector Issues for Tau Physics . . . . .	46
2.7.4	Detector Issues for Two-Photon Physics . . . . .	48
<b>3</b>	<b>Physics with BABAR</b> . . . . .	<b>51</b>
3.1	Physics Context . . . . .	51
3.2	Simulation Tools . . . . .	55
3.2.1	ASLUND . . . . .	55
3.2.2	GEANT Simulation—BBSIM . . . . .	56
3.3	Studies of $B^0 \rightarrow J/\psi$ Modes . . . . .	57
3.3.1	$B^0 \rightarrow J/\psi K_S^0$ . . . . .	57
3.3.2	$B^0 \rightarrow J/\psi K_L^0$ . . . . .	58
3.3.3	$B^0 \rightarrow J/\psi K^{0*}$ . . . . .	60
3.4	Studies of $B^0$ to Double Charm Modes . . . . .	61
3.4.1	$B^0 \rightarrow D^+ D^-$ . . . . .	61
3.4.2	$B^0 \rightarrow D^{*+} D^{*-}$ . . . . .	63
3.4.3	$B \rightarrow DD^*$ . . . . .	64
3.5	Studies of $B \rightarrow \pi\pi$ Modes . . . . .	64
3.5.1	$B^0 \rightarrow \pi^+ \pi^-$ . . . . .	65

---

3.5.2	$B^0 \rightarrow \pi^0\pi^0$ . . . . .	67
3.5.3	$B^+ \rightarrow \pi^+\pi^0$ Decays . . . . .	68
3.6	$B^0 \rightarrow \rho^\pm\pi^\mp$ . . . . .	68
3.7	Tagging Modes . . . . .	72
3.7.1	Kaon Tags . . . . .	73
3.7.2	Lepton Tags . . . . .	74
3.7.3	Other Tags . . . . .	76
3.7.4	Combined Tagging . . . . .	76
3.8	Estimate of $CP$ -Angle Measurement . . . . .	77
3.8.1	Method of Calculation . . . . .	77
3.8.2	Tagging Modes . . . . .	78
3.8.3	$CP$ Modes . . . . .	79
3.9	CKM Matrix Determination . . . . .	81
3.9.1	$V_{cb}$ . . . . .	82
3.9.2	$V_{ub}$ . . . . .	84
3.9.3	$V_{td}$ . . . . .	85
3.10	Rare $B$ Decays . . . . .	85
3.10.1	$B \rightarrow \tau\nu$ . . . . .	86
3.10.2	$B \rightarrow X_s\ell^+\ell^-$ . . . . .	87
3.10.3	$B \rightarrow X_s\nu\bar{\nu}$ . . . . .	87
3.11	$CP$ Asymmetries in Charged $B$ Decays . . . . .	87
3.12	Charm Physics . . . . .	88
3.12.1	$D^0\bar{D}^0$ Mixing . . . . .	88
3.12.2	$CP$ Violation in Charm Decays . . . . .	89
3.13	Tau Physics . . . . .	90
3.14	Two-Photon Physics . . . . .	91

---

3.15	Summary . . . . .	92
<b>4</b>	<b>Vertex Detector</b>	<b>97</b>
4.1	Vertex Detector Requirements . . . . .	97
4.1.1	Resolution . . . . .	97
4.1.2	Acceptance . . . . .	98
4.1.3	Efficiency . . . . .	98
4.1.4	Radiation Tolerance . . . . .	99
4.1.5	Reliability . . . . .	99
4.2	Vertex Detector Overview . . . . .	100
4.2.1	Choice of Technology . . . . .	100
4.2.2	Detector Layout . . . . .	100
4.2.3	Electronic Readout . . . . .	102
4.2.4	Mechanical Support . . . . .	104
4.3	Detector Performance Studies . . . . .	105
4.3.1	Resolution . . . . .	106
4.3.2	Pattern Recognition . . . . .	112
4.3.3	Solid Angle Coverage . . . . .	114
4.4	Silicon Detectors . . . . .	115
4.4.1	Requirements . . . . .	115
4.4.2	Silicon Detector Design . . . . .	117
4.4.3	Fanout Circuit Design . . . . .	121
4.4.4	R&D on Detectors and Fanouts . . . . .	124
4.5	Electronic Readout . . . . .	124
4.5.1	Introduction . . . . .	124
4.5.2	Readout Chip . . . . .	125
4.5.3	Hybrid Design . . . . .	130

---

4.5.4	Data Transmission . . . . .	132
4.5.5	Baseline Design . . . . .	133
4.5.6	Power Supplies . . . . .	135
4.5.7	Electronics R&D . . . . .	135
4.6	Mechanical Support and Assembly . . . . .	136
4.6.1	IR Constraints . . . . .	137
4.6.2	Module Assembly . . . . .	137
4.6.3	Detector Assembly and Installation . . . . .	139
4.6.4	Detector Placement and Survey . . . . .	142
4.6.5	Detector Monitoring . . . . .	144
4.6.6	R&D Program . . . . .	146
4.7	Services, Utilities, and ES&H Issues . . . . .	147
4.7.1	Services and Utilities . . . . .	147
4.7.2	ES&H Issues . . . . .	147
<b>5</b>	<b>Drift Chamber</b>	<b>151</b>
5.1	Physics Requirements and Performance Goals . . . . .	151
5.2	Tracking Chamber Overview . . . . .	153
5.3	Projected Performance . . . . .	155
5.4	Drift System Design . . . . .	158
5.4.1	Cell Design . . . . .	158
5.4.2	Layer Arrangement . . . . .	159
5.4.3	Total Channel Count . . . . .	162
5.4.4	Cell Studies . . . . .	166
5.4.5	Gain Variations . . . . .	170
5.4.6	Electrostatic Forces and Stability . . . . .	171
5.4.7	Pattern Recognition Studies . . . . .	171

5.5	Gas Choice and Properties . . . . .	172
5.6	Mechanical Design . . . . .	175
5.6.1	Endplates . . . . .	175
5.6.2	Inner Wall . . . . .	176
5.6.3	Outer Wall . . . . .	177
5.6.4	Joints . . . . .	178
5.6.5	R&D Program on Structural Components . . . . .	179
5.6.6	Wires . . . . .	180
5.6.7	Feedthroughs . . . . .	181
5.6.8	Stringing . . . . .	183
5.6.9	Endplate Connections . . . . .	185
5.7	Front-End Electronics . . . . .	187
5.8	High Voltage System . . . . .	188
5.9	Calibration and Monitoring . . . . .	189
5.9.1	Calibration . . . . .	189
5.9.2	Slow Controls . . . . .	189
5.9.3	Monitoring . . . . .	190
5.10	Integration . . . . .	191
5.10.1	Overall Geometry and Mechanical Support . . . . .	191
5.10.2	Cable Plant and Utilities Routing . . . . .	191
5.10.3	Access . . . . .	192
5.10.4	Integration Aspects of the Gas System . . . . .	193
5.10.5	Installation and Alignment . . . . .	193
5.11	R&D Using Prototype Drift Chambers . . . . .	194
5.11.1	Prototype I Chamber . . . . .	194
5.11.2	Prototype II . . . . .	195

---

5.11.3	Other R&D Efforts . . . . .	195
<b>6</b>	<b>Particle Identification</b>	<b>197</b>
6.1	Physics Requirement and Performance Goals . . . . .	197
6.1.1	Introduction . . . . .	197
6.1.2	<i>B</i> Flavor Tagging . . . . .	197
6.1.3	Exclusive <i>B</i> decays . . . . .	199
6.1.4	Summary of Requirements . . . . .	200
6.2	Particle Identification Overview . . . . .	200
6.2.1	The DIRC . . . . .	201
6.2.2	The ATC . . . . .	203
6.3	Projected Performance . . . . .	204
6.3.1	Simulation Based on Prototype Results . . . . .	204
6.3.2	Pattern Recognition in the DIRC . . . . .	207
6.3.3	Particle Identification in the DIRC . . . . .	210
6.3.4	Particle Identification in the ATC . . . . .	212
6.3.5	Performance Requirements for Track Reconstruction . . . . .	212
6.3.6	Effects of Backgrounds on PID Detector Performance . . . . .	214
6.4	The DIRC Detector . . . . .	216
6.4.1	DIRC Mechanical Design . . . . .	216
6.4.2	Photodetectors and Readouts . . . . .	223
6.4.3	Laser Flasher Monitoring and Calibration System . . . . .	225
6.4.4	Integration Issues . . . . .	226
6.4.5	Research and Development Program . . . . .	228
6.5	The ATC Detector . . . . .	232
6.5.1	ATC Mechanical Design . . . . .	232
6.5.2	Photodetectors and Readouts . . . . .	234

---

6.5.3	Monitoring and Calibration System . . . . .	237
6.5.4	Integration Issues . . . . .	237
6.5.5	Research and Development Program . . . . .	238
<b>7</b>	<b>Electromagnetic Calorimeter</b>	<b>245</b>
7.1	Physics Requirements and Performance Goals . . . . .	245
7.1.1	Physics Processes Influencing Performance Goals . . . . .	245
7.1.2	Summary of Performance Targets . . . . .	248
7.2	Calorimeter Overview . . . . .	249
7.2.1	Technology Choice . . . . .	249
7.2.2	Description of the Calorimeter . . . . .	250
7.2.3	Readout Chain and Trigger . . . . .	253
7.2.4	Review of Options . . . . .	254
7.3	Projected Calorimeter Performance . . . . .	256
7.3.1	Contributions to Photon Resolution and Efficiency . . . . .	256
7.3.2	Modeling . . . . .	259
7.3.3	Expected Performance for Photons . . . . .	260
7.3.4	Expected Performance for $\pi^0$ s . . . . .	261
7.3.5	$e/\pi$ Separation . . . . .	264
7.3.6	Performance and Cost Optimization . . . . .	266
7.4	Crystal Subassemblies and Readout . . . . .	269
7.4.1	Photodiode Readout . . . . .	270
7.4.2	Light Collection . . . . .	274
7.4.3	Light Yield Measurements . . . . .	276
7.4.4	Reliability of Inaccessible Readout Components . . . . .	278
7.5	Calibration . . . . .	280
7.5.1	Requirements and Ingredients . . . . .	280

---

7.5.2	Energy Calibration with Beam Events . . . . .	281
7.5.3	Source Calibration . . . . .	282
7.6	Mechanical Support Structure . . . . .	283
7.6.1	Design Considerations . . . . .	283
7.6.2	Barrel Fabrication, Assembly, and Installation . . . . .	288
7.6.3	Endcap Fabrication, Assembly, and Installation . . . . .	296
7.7	Optimization and Prototype Studies . . . . .	299
7.8	Crystal Procurement Issues . . . . .	301
7.8.1	Radiation Hardness . . . . .	301
7.8.2	Quality Control and Testing . . . . .	303
<b>8</b>	<b>Muon and Neutral Hadron Detector</b>	<b>307</b>
8.1	Physics Requirements and Performance Goals . . . . .	307
8.2	Detector Overview . . . . .	312
8.2.1	The Iron Structure . . . . .	312
8.2.2	The Active Detector Choice . . . . .	315
8.3	Projected Performance . . . . .	316
8.3.1	Muon Identification . . . . .	316
8.3.2	Muon Tagging . . . . .	322
8.3.3	$K_L^0$ Detection . . . . .	324
8.4	Detector Design and R&D . . . . .	328
8.4.1	Chamber Construction and Assembly . . . . .	328
8.4.2	System Layout . . . . .	330
8.5	Gas System . . . . .	335
8.5.1	Gas Composition and Flow Rates . . . . .	335
8.5.2	Mixer . . . . .	336
8.5.3	Distribution . . . . .	336

---

8.6	Front-End Electronics and High Voltage . . . . .	336
8.7	Final Assembly, Installation, and Monitoring . . . . .	337
<b>9</b>	<b>Magnet Coil and Flux Return</b>	<b>341</b>
9.1	Physics Requirements and Performance Goals . . . . .	341
9.2	Overview . . . . .	342
9.2.1	Description of Key Interfaces . . . . .	344
9.3	Summary of Projected Magnet Performance . . . . .	346
9.3.1	Central Field Magnitude and Coil Performance . . . . .	346
9.3.2	Shielding of Forward Q2 . . . . .	347
9.3.3	Flux Return . . . . .	347
9.4	Superconducting Solenoid . . . . .	350
9.4.1	Magnetic Design . . . . .	350
9.4.2	Cold Mass Design . . . . .	352
9.4.3	Quench Protection and Stability . . . . .	354
9.4.4	Cold Mass Cooling . . . . .	358
9.4.5	Cryostat Design . . . . .	359
9.4.6	Coil Assembly and Transportation . . . . .	361
9.5	Cryogenic Supply System and Instrumentation . . . . .	362
9.6	Flux Return . . . . .	364
9.6.1	Overview . . . . .	364
9.6.2	Barrel Flux Return Description . . . . .	364
9.6.3	End Door Description . . . . .	368
9.6.4	Options and Detailed Design Issues . . . . .	371
9.6.5	Procurement, Fabrication, Assembly, and Schedule . . . . .	371
<b>10</b>	<b>Electronics</b>	<b>375</b>

---

10.1	Overview . . . . .	375
10.1.1	Introduction . . . . .	375
10.1.2	Front-End . . . . .	377
10.1.3	Trigger . . . . .	378
10.1.4	Data Acquisition . . . . .	380
10.2	Silicon Vertex Detector . . . . .	382
10.2.1	Requirements . . . . .	382
10.2.2	Overview . . . . .	382
10.2.3	Data Format . . . . .	384
10.2.4	Description of the Readout Module . . . . .	384
10.3	Drift Chamber . . . . .	385
10.3.1	Requirements . . . . .	385
10.3.2	Preamplifiers . . . . .	386
10.3.3	Digitizer Options . . . . .	387
10.3.4	Implementing the All-FADC Scheme . . . . .	388
10.3.5	Calibration . . . . .	391
10.3.6	Research and Development . . . . .	391
10.4	DIRC . . . . .	394
10.4.1	Requirements . . . . .	394
10.4.2	Baseline Design . . . . .	395
10.4.3	Research and Development . . . . .	398
10.4.4	High Voltage System . . . . .	400
10.4.5	Low Voltage Power Supplies and Control Systems . . . . .	400
10.5	Aerogel Threshold Cherenkov Counter (ATC) . . . . .	401
10.5.1	Requirements and Overview . . . . .	401
10.5.2	Front-End Electronics . . . . .	401

10.5.3	Readout . . . . .	402
10.6	Calorimeter . . . . .	402
10.6.1	Requirements . . . . .	402
10.6.2	Overview . . . . .	403
10.6.3	The Preamplifier Card . . . . .	406
10.6.4	Digitizing Board . . . . .	407
10.6.5	Readout Module . . . . .	411
10.6.6	Electronic Calibration . . . . .	413
10.6.7	Monitoring and Control . . . . .	413
10.6.8	Research and Development . . . . .	414
10.7	Muon System Electronics . . . . .	414
10.7.1	Electronics Requirements . . . . .	414
10.7.2	General Architecture . . . . .	415
10.7.3	Implementation . . . . .	416
10.7.4	Time Measurements . . . . .	419
10.7.5	Time Calibration . . . . .	419
10.7.6	Monitoring . . . . .	420
10.7.7	Diagnostics . . . . .	420
10.8	Trigger Requirements and Background Rates . . . . .	420
10.8.1	Introduction . . . . .	420
10.8.2	Requirements . . . . .	421
10.8.3	Backgrounds and Trigger Rates . . . . .	423
10.8.4	Definitions of Trigger Filter Concepts . . . . .	424
10.8.5	Performance of Some Simple Triggers . . . . .	426
10.9	Level 1 Trigger . . . . .	428
10.9.1	Overview . . . . .	428

---

10.9.2	Drift Chamber Trigger . . . . .	432
10.9.3	Calorimeter Trigger . . . . .	435
10.9.4	Global Trigger . . . . .	436
10.9.5	Simulation . . . . .	437
10.9.6	Trigger System . . . . .	438
10.10	Data Acquisition . . . . .	441
10.10.1	Introduction . . . . .	441
10.10.2	Requirements . . . . .	442
10.10.3	Architectural Overview . . . . .	445
10.10.4	Readout Crates . . . . .	449
10.10.5	Event-Data Flow Control . . . . .	452
10.10.6	Event Assembly . . . . .	456
10.10.7	Event Distribution . . . . .	460
10.10.8	Data Integrity . . . . .	462
10.10.9	Research and Development . . . . .	462
10.11	Level 2 Trigger . . . . .	463
10.11.1	Filter . . . . .	463
10.11.2	Implementation . . . . .	464
10.11.3	Conclusion . . . . .	465
10.12	Global Support Electronics . . . . .	466
10.12.1	Detector Monitoring and Control . . . . .	466
10.12.2	Data Monitoring Support . . . . .	468
10.12.3	Diagnostic Support . . . . .	468
10.12.4	Calibration Support . . . . .	469
<b>11</b>	<b>Computing</b> . . . . .	<b>475</b>
11.1	Requirements . . . . .	475

11.1.1	Operational Requirements . . . . .	476
11.1.2	Technical Requirements . . . . .	477
11.2	Overview . . . . .	482
11.2.1	Chosen Technologies . . . . .	482
11.2.2	Overall Description . . . . .	483
11.3	Computing Model . . . . .	485
11.3.1	Architectural Model . . . . .	485
11.3.2	Operational Model . . . . .	488
11.3.3	Baseline Cost Model . . . . .	493
11.4	Software Environment . . . . .	498
11.4.1	Software Methodologies and Languages . . . . .	498
11.4.2	Development Environment . . . . .	500
11.4.3	Data Model . . . . .	501
11.4.4	Code Management and Distribution . . . . .	503
11.4.5	Databases . . . . .	506
11.4.6	Graphics . . . . .	508
11.5	Online System . . . . .	510
11.5.1	Introduction . . . . .	510
11.5.2	Logical View of the Online System . . . . .	512
11.5.3	Subsystem View . . . . .	514
11.5.4	Operator View . . . . .	515
11.5.5	Implementation View . . . . .	516
11.6	Reconstruction and Analysis Framework . . . . .	517
11.6.1	Scope . . . . .	517
11.6.2	Conceptual Overview . . . . .	518
11.6.3	Toolkits . . . . .	519

---

11.6.4	Framework . . . . .	519
11.6.5	Application Builder . . . . .	519
11.6.6	Standard Applications . . . . .	520
11.6.7	Job Submission . . . . .	520
11.6.8	Bulk Production . . . . .	520
11.7	Computing Support . . . . .	521
11.7.1	Introduction . . . . .	521
11.7.2	Collaboration Support . . . . .	521
11.7.3	Infrastructure Support . . . . .	522
11.8	Integration Issues . . . . .	523
11.9	System Responsibilities and Management . . . . .	524
<b>12</b>	<b>Interaction Region and Backgrounds</b>	<b>529</b>
12.1	PEP-II Design . . . . .	530
12.1.1	Parameters . . . . .	530
12.1.2	Interaction Region Components . . . . .	531
12.1.3	Background Implications . . . . .	532
12.1.4	Operating Modes . . . . .	534
12.2	Tools . . . . .	535
12.2.1	QSRAD . . . . .	535
12.2.2	Decay TURTLE . . . . .	535
12.2.3	OBJEGS . . . . .	536
12.2.4	GEANT . . . . .	536
12.2.5	GELHAD . . . . .	538
12.3	Background Sources . . . . .	538
12.3.1	Synchrotron Radiation . . . . .	538
12.3.2	Lost Particles . . . . .	541

12.3.3	Hadrons . . . . .	544
12.3.4	Luminosity Backgrounds . . . . .	547
12.3.5	Other Sources . . . . .	550
12.4	Background Rates and Detector Responses . . . . .	552
12.4.1	Vertex Detector . . . . .	553
12.4.2	Drift Chamber . . . . .	558
12.4.3	Particle ID . . . . .	560
12.4.4	Calorimeter . . . . .	561
12.4.5	Muons . . . . .	562
12.5	Physics Impact . . . . .	562
12.6	Summary . . . . .	563
<b>13</b>	<b>Safety</b>	<b>567</b>
13.1	Introduction . . . . .	567
13.2	Detector Safety Overview . . . . .	567
13.3	Beam Pipe and Support Barrel . . . . .	568
13.4	Vertex Detector . . . . .	569
13.5	Drift Chamber . . . . .	569
13.6	Particle Identification . . . . .	570
13.6.1	DIRC . . . . .	570
13.6.2	Aerogel . . . . .	571
13.7	Electromagnetic Calorimeter . . . . .	571
13.8	Muon and Neutral Hadron Detector . . . . .	572
13.9	Magnet Coil and Flux Return . . . . .	573
13.10	Generic Hazards . . . . .	574
<b>14</b>	<b>Facilities, Assembly, Access and Integration</b>	<b>577</b>

---

14.1	Facilities . . . . .	577
14.2	Detector Coordinate System . . . . .	579
14.3	Structural Support of Systems . . . . .	579
14.4	Installation Overview . . . . .	580
14.5	Detector Component Installations . . . . .	580
14.5.1	IR-2 Detector Hall Preparation . . . . .	580
14.5.2	Coil and IFR Installation . . . . .	582
14.5.3	Barrel Calorimeter Installation . . . . .	585
14.5.4	DIRC Tube Assembly Installation . . . . .	586
14.5.5	Drift Chamber Installation . . . . .	587
14.5.6	Aerogel/Forward Calorimeter Installation . . . . .	587
14.5.7	Vertex Assembly Installation . . . . .	588
14.5.8	Backward End Plug Installation . . . . .	589
14.5.9	Electronics House . . . . .	589
14.5.10	Detector Transport System . . . . .	589
14.5.11	Final Position of Detector in Beam Line . . . . .	589
14.5.12	Service Space . . . . .	590
14.6	Detector Maintenance Access . . . . .	590
14.7	Detector Integration . . . . .	591
14.7.1	Assembly Clearances . . . . .	592
14.7.2	Good Neighbor Policy . . . . .	593
14.7.3	Common Services . . . . .	593
<b>15</b>	<b>Collaboration Issues and Project Management</b>	<b>597</b>
15.1	Membership . . . . .	597
15.2	Collaboration Council . . . . .	599
15.3	Spokesperson . . . . .	599

15.4	Executive Board . . . . .	599
15.5	Technical Board . . . . .	600
15.6	Finance Review Committee . . . . .	600
15.7	Communications . . . . .	600
15.8	Construction Responsibilities . . . . .	601
<b>16</b>	<b>Cost and Schedule</b>	<b>605</b>
16.1	Introduction . . . . .	605
16.2	Project Cost . . . . .	605
16.3	Schedule . . . . .	608
16.4	Detector Systems . . . . .	609
16.4.1	Vertex Detector . . . . .	609
16.4.2	Tracking Chamber . . . . .	609
16.4.3	Particle Identification . . . . .	610
16.4.4	CsI(Tl) Calorimeter . . . . .	611
16.4.5	Flux Return Instrumentation . . . . .	612
16.4.6	Magnet . . . . .	612
16.4.7	Electronics . . . . .	612
16.4.8	Computing . . . . .	613
16.4.9	Management, Installation and Integration . . . . .	615

---

---

# Introduction

**I**n the thirty years that have elapsed since the surprising discovery of  $CP$  violation in  $K_L^0$  meson decay, the Standard Model of elementary particle physics has evolved considerably. The Standard Model with three quark generations accommodates  $CP$  violation in quark decays in an elegant and economical manner. Through one imaginary parameter and three real numbers, the Standard Model completely characterizes the weak couplings of the six quarks as expressed in the Cabibbo-Kobayashi-Maskawa matrix. We have little evidence, however, that this explanation is correct.

Indeed, there are hints from cosmological arguments that, in fact, it is not.  $CP$  violation is a key piece in the puzzle of the matter dominance in the universe. The excess of matter over antimatter is not an easily explained property of the universe. In any theory with baryon-number-violating processes (for example, most grand unified theories) the universe at very early times contains equal (thermal) populations of quarks and antiquarks. How and when the currently observed imbalance develops is one of the key puzzles of cosmology. Attempts to explain this with only Standard Model  $CP$  violation have so far not been successful, leading to the suggestion that perhaps  $CP$  violation physics is the Achilles heel of the Standard Model. Certainly it merits detailed exploration.

The study of  $CP$ -violating asymmetries in the decays of  $B^0$  mesons to  $CP$  eigenstates promises, at last, to provide a test of the Standard Model explanation of  $CP$  violation. More generally, the measurement of these  $CP$  asymmetries can provide us with a series of unique, and uniquely stringent, consistency tests of the quark sector of the Standard Model. These tests, which will be carried out by the *BABAR* detector at PEP-II, will provide us with perhaps the best chance to challenge the Standard Model in new and quantitatively precise ways. This challenge has come about through the interplay of a series of experimental and theoretical events.

While it had been understood for several years that the measurement of  $CP$ -violating asymmetries in  $B^0$  decay could lead to important tests of the CKM matrix [Car81], the experiments seemed beyond reach. The discovery of a surprisingly long  $b$ -quark lifetime by MAC [Fer83] and Mark II [Loc83] at PEP in 1983, together with an unexpectedly large  $B^0$ - $\bar{B}^0$  mixing by UA1 [Alb87b] and ARGUS [Alb87a] in 1987 made it possible to contemplate making these measurements. It soon became clear [SNO89] that the richest and most straightforward approach involved experiments at a variety of  $e^+e^-$  machines,

either in the  $\Upsilon(4S)$  region, in the PEP/PETRA continuum region, or at the  $Z^0$  pole. Experiments at hadron machines can also address this physics. While hadronic experiments potentially have larger samples of  $B$  mesons, they face substantial trigger and combinatoric problems [HER94]. The most favorable  $e^+e^-$  experimental situation, that is, the one producing the smallest statistical error with the least integrated luminosity, is the asymmetric storage ring first proposed by Oddone [Odd87]. This machine would boost the decaying  $B^0$  mesons in the laboratory frame, allowing existing vertex measuring technology to measure the time order of  $B^0 - \bar{B}^0$  decay pairs, even with the short  $B$  meson flight distance. This, however, would require event samples that are obtainable only with significant advances in storage ring luminosity.

PEP-II promises to provide the required luminosity [SLA91], initially  $3 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ , ultimately  $10^{34}$ , with asymmetric  $\Upsilon(4S)$  production at a  $\beta\gamma$  of 0.56 (9 GeV electrons on 3.1 GeV positrons). The  $B\bar{B}$  production rate will be 3 Hz at the initial luminosity, ultimately rising to 10 Hz. The experimental challenge is then to provide high efficiency, high resolution exclusive state reconstruction in a situation new to the  $e^+e^-$  world: a center-of-mass in motion in the laboratory. The boost is not extreme: at PEP-II,  $90^\circ$  in the center-of-mass is at  $65^\circ$  in the laboratory. Nonetheless, the challenges this boost poses for the detector are novel, and not easily met.

The measurement of  $CP$ -violating asymmetries in  $B^0$  meson decay will be a major breakthrough. It will, however, be only the first installment in a rich program that will ultimately provide us with very stringent tests of the Standard Model. The key to these tests is the measurement of the *unitarity triangle* relations between the CKM matrix elements. The ultimate precision of the unitarity triangle tests depends not only on the  $CP$ -violating asymmetries, which measure the angles of the unitarity triangle, but on improved measurements of the triangle sides as well. The series of measurements that is uniquely possible at an asymmetric  $e^+e^-$  storage ring operated at the  $\Upsilon(4S)$  resonance will both measure the angles for the first time and substantially improve the measurement of the triangle sides. They will, further, provide unique tests of underlying theoretical assumptions and measurements of experimental systematic uncertainties. Should discrepancies with Standard Model predictions emerge, the program can also pinpoint particular extensions of the Standard Model, be they supersymmetry, left-right models, or other specific models.

While the search for  $CP$  violation and for possible contradictions with the Standard Model CKM picture of its origin is the most exciting physics challenge for  $BABAR$ , it is by no means the only physics opportunity to be exploited. The asymmetric configuration and the resultant separation of  $B$  vertices makes possible searches for rare  $B$  decays, such as  $b \rightarrow u$  transitions, in a very clean environment. The large production rate for charm, either directly in continuum  $c\bar{c}$  events or indirectly from  $B$  decays, provides opportunities for charm physics. Similarly, aspects of tau physics can readily be studied here with large data samples in a clean environment. These opportunities are discussed in some greater detail in Chapter 3.

A detector designed to carry out this ambitious experimental program must be capable of high efficiency reconstruction of extremely rare exclusive final states in the presence of combinatoric and accelerator-generated backgrounds. This requirement places a premium on solid angle coverage, charged particle momentum resolution and species identification, and photon resolution and detection efficiency. The design of the *BABAR* detector attempts to balance these requirements against one another and against cost and schedule imperatives.

Soon after the approval of PEP-II in October of 1993, efforts began to form a major international collaboration to build and operate a new detector at PEP-II's single interaction region. Time to build a detector is short, and there has been a clear convergence of thinking [SLA93, BEL94, HEL92] about new detector designs. The SLAC directorate, with the approval of the Scientific Policy Committee, invited a group of experimentalists to submit a Letter of Intent and a Technical Design Report. The primary aim of this group is the detailed study of *CP*-violating asymmetries in  $B^0$  meson decay, with a wide variety of other *B*, charm,  $\tau$ , and two-photon physics that will also be accessible at new levels of precision. An inaugural meeting was held in December of 1993. The Collaboration drafted a governance document, chose a name and its leadership, and entered into the detailed design of the *BABAR* detector.

A detailed *Letter of Intent* [Bab94] was submitted to the SLAC Experimental Advisory Committee (EPAC) in June of 1994. The EPAC recommended that the *BABAR* Collaboration be approved to proceed to a Technical Design Report (TDR). This recommendation was accepted by the SLAC Director, who asked for a TDR by the end of 1994. This document is that report.

The detector described in the *Letter of Intent* was a conceptual design, which in some instances contained alternative technology choices for several detector subsystems. The design discussed herein is substantially more detailed. It contains specific choices for the subsystem technologies, and the design of these subsystems has been optimized to provide a balance between cost and adequate physics performance.

The planning of the PEP-II project has been done in sufficient detail that the Laboratory is confident that it can be constructed in five years, starting in October of 1993. In designing *BABAR*, we have taken as an important constraint that a detector capable of measuring *CP*-violating asymmetries in  $B^0$  meson decay be operational on a timescale matched to that for building PEP-II.

The *BABAR* Collaboration management team is functioning, with a Spokesman, Deputy Spokesman, Technical Coordinator, and a Project Engineer. System managers are in place, as well. Should the EPAC recommend approval on the basis of this document, the TDR, together with a detailed Work Breakdown Structure, cost estimate, critical path schedule, and Management Plan, will be submitted for approval by the DOE and the funding agencies of the non-US institutions. With the establishment of Memoranda of Understanding (MOUs)

detailing the specific responsibilities of each collaborating group, the construction of the *BABAR* detector will commence in earnest.

The following Chapter describes the proposed detector, discusses some of the design decisions and their physics basis and reviews the performance expected component by component. Chapter 3 presents a physics overview, including results of simulation studies for the key *CP* violation study channels as well as an overview of other physics opportunities. Detailed descriptions of each detector subsystem are then provided in individual chapters. Finally there are chapters that summarize general detector issues, such as accelerator-induced backgrounds, safety, installation and integration, management, budget and schedule.

---

## References

---

- [Alb87a] H. Albrecht *et al.* (ARGUS Collaboration), *Phys. Lett.* **192B**, 245 (1987).
- [Alb87b] C. Albejar *et al.* (UA1 Collaboration), *Phys. Lett.* **186B**, 247 (1987).
- [BEL94] BELLE Collaboration, “Letter of Intent for A Study of  $CP$  Violation in  $B$  Meson Decays,” submitted to the TRISTAN Program Advisory Committee (1994).
- [Car81] A. Carter and A.I. Sanda, *Phys. Rev. Lett.* **45**, 952 (1980); *Phys. Rev.* **D23**, 1567 (1981);  
I. Bigi and A.I. Sanda, *Nucl. Phys.* **B193**, 85 (1981); *Nucl. Phys.* **B281**, 41 (1987).
- [Fer83] E. Fernandez *et al.* (MAC Collaboration), *Phys. Rev. Lett.* **51**, 1022 (1983).
- [HEL92] “HELENA, A Beauty Factory at Hamburg,” DESY 92/41 (1992).
- [HER94] HERA-B Proposal, DESY-PRC 94/02 (1994).
- [Loc83] N. Lockyer, *et al.* (Mark II Collaboration), *Phys. Rev. Lett.* **51**, 1316 (1983).
- [Odd87] P. Oddone, in Proceedings of the UCLA Workshop: Linear Collider  $B\bar{B}$  Factory Conceptual Design, edited by D. Stork (World Scientific, 1987) p. 243.
- [SLA91] “An Asymmetric  $B$  Factory Based on PEP: Conceptual Design Report,” edited by M. Zisman, SLAC-372, LBL PUB-5303 (1991).
- [SLA93] “Status Report on the Design of a Detector for the Study of  $CP$  Violation at PEP-II at SLAC,” SLAC-419 (1993).
- [SLA94] “Letter of Intent for the Study of  $CP$  Violation and Heavy Flavor Physics at PEP-II,” SLAC-443 (1994).
- [SNO89] “Report of the  $B$  Physics Group: 1. Physics and Techniques,” G. Feldman *et al.*, in Proceedings of the 1988 DPF Summer Study (Snowmass, Colorado) p. 561 (1989).



---

---

# Detector Overview

---

## 2.1 Introduction

---

The primary physics goal of the *BABAR* experiment is the systematic study of *CP* asymmetries in decays of the  $B^0$  and  $\bar{B}^0$  into *CP* eigenstates, as discussed in Chapter 3. The secondary goals are to explore the wide range of other *B* physics, charm physics,  $\tau$  physics, two-photon physics, and  $\Upsilon$  physics that becomes accessible with the high luminosity of PEP-II. The design of the detector is optimized for the *CP* studies, but also serves well for the other physics opportunities.

The critical experimental objectives to achieve the required sensitivity for *CP* measurements are:

- To reconstruct the decays of  $B^0$  mesons into a wide range of exclusive final states with high efficiency and low background;
- To tag the flavor of the other *B* meson in the event with high efficiency and purity; and
- To measure the relative decay time of the two *B* mesons.

Studies carried out over the last few years at SLAC [SLA91, SLA93] and for *B* factory proposals from other sites [CLE93, KEK92, HEL92] have led to a fairly complete picture of the performance requirements that the detector systems must meet. In the *Letter of Intent* [SLA94], we outlined a conceptual detector design that would satisfy these requirements. In this document, we describe the technical design of the *BABAR* detector and report the results of simulation studies.

In this chapter, we list the general design considerations, describe the detector design, give details of the detector performance, and discuss the capability of the detector to do the important physics available at PEP-II.

## 2.2 General Design Considerations

---

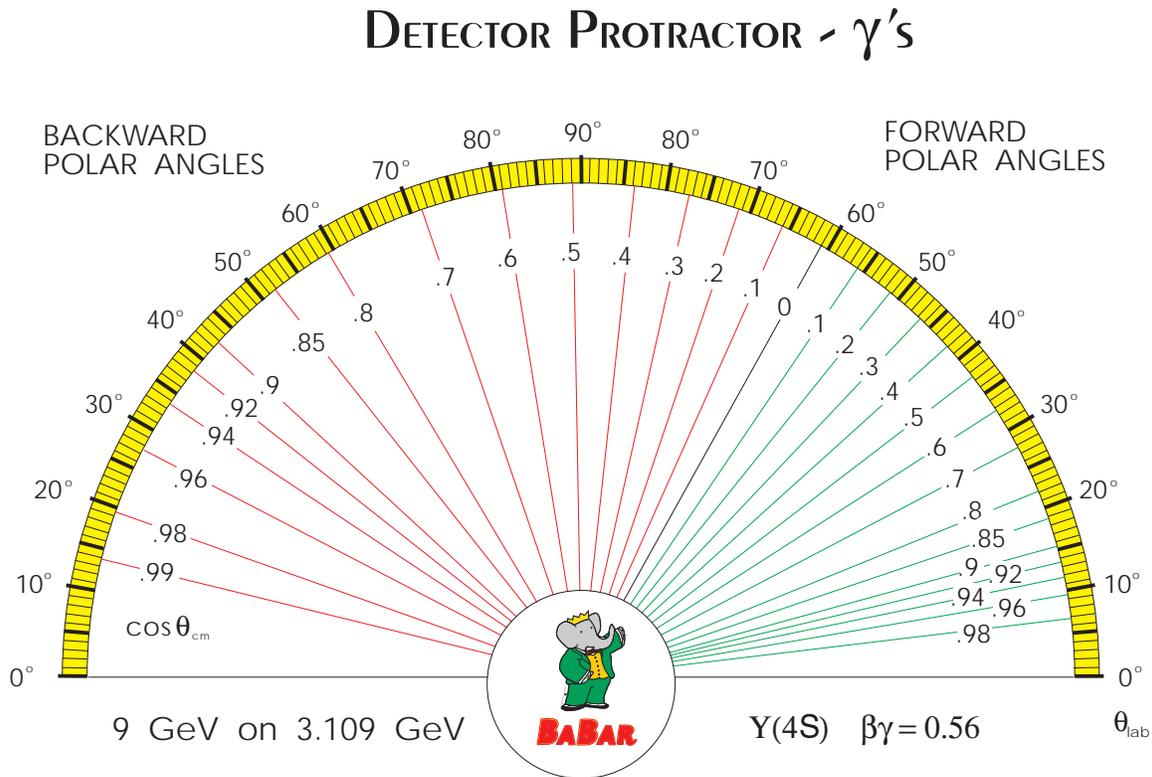
### 2.2.1 Acceptance

The crucial  $CP$  asymmetry measurements at PEP-II are made using events in which one  $B^0$ , usually decaying into a  $CP$  eigenstate, is completely reconstructed, and the other  $B$  in the event is tagged as a  $B^0$  or  $\bar{B}^0$ , using either a lepton or charged kaon. The  $CP$  decay modes of interest generally have branching ratios below  $10^{-4}$ , and reconstructing them requires observing anywhere from two to six charged particles and often one or more  $\pi^0$ s. The charged particles must be detected with very good momentum resolution, precise vertex information, and clean particle identification. To achieve good efficiency for these rare events, it is important that the full performance of the detector should extend over a very large solid angle. This is made slightly more difficult by the energy boost, which folds one-half of the particles into the region  $\cos \theta_{lab} > 0.5$ . Figure 2-1 shows the relation between center-of-mass and laboratory polar angles for photons with the designed boost of  $\beta\gamma = 0.56$ . It should be noted, however, that the boost causes a separation of the two  $B$  vertices, which provides an additional reduction in background from continuum events.

Figure 2-2 shows the detection efficiency for the decay  $B^0 \rightarrow J/\psi K_s^0 \rightarrow e^+e^-\pi^+\pi^-$  as a function of forward polar angle coverage, which is one critical parameter in achieving good acceptance. The acceptance for charged particles extends to  $\cos \theta \sim 0.96$  in the forward direction and  $\cos \theta \sim -0.87$  in the backward direction. In both cases, the limit is set by how close the active coverage of the vertex detector can come to the edge of the first machine dipole magnet (B1). These angular limits correspond to a geometric acceptance of about 57% for observing all four charged tracks. For the detectors outside the vertex detector, the acceptance is designed to extend at least to these limits, so there is no further loss of geometric acceptance.

Another critical parameter is the minimum momentum cutoff in the acceptance for charged and neutral particles. The momentum distribution for the particles from  $B$  decays at rest is quite soft: the average momentum for a pion from the process studied in Figure 2-2 is about  $1.8 \text{ GeV}/c$ , but the momentum range for tagging kaons extends below  $0.3 \text{ GeV}/c$ . The pion from the process  $B^0 \rightarrow D^{*+}D^{*-}$ ,  $D^{*+} \rightarrow D^0\pi^+$  has a typical momentum of around  $130 \text{ MeV}/c$ . The energy spectrum for photons is even softer, since each photon receives on average half of the energy of the neutral pion which created it. It is, therefore, necessary to maintain good efficiency and resolution for charged tracks and for photons at low momentum. The tracking efficiency cuts off at  $p_t \sim 60 \text{ MeV}/c$ , and the efficiency for photons at  $E \sim 20 \text{ MeV}$ .

It is particularly important to identify electrons and muons cleanly over as wide a range of angle and momentum as possible. The primary leptons from  $B$  decays, which are used

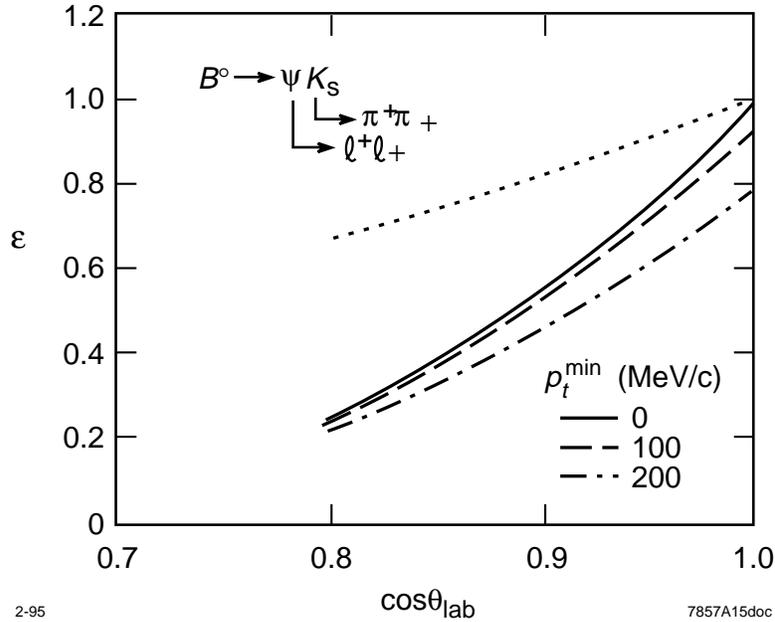


**Figure 2-1.** Protractor showing the relationship between center-of-mass and laboratory polar angles for photons at  $\beta\gamma = 0.56$ .

to tag the flavor of the  $B$ , have momentum as low as about  $1 \text{ GeV}/c$ . Secondary leptons coming from the sequence  $B \rightarrow DX$ ,  $D \rightarrow \ell Y$  provide additional tagging information, if it is possible to identify them cleanly. Thus, the efficiency of the lepton tag is limited by the minimum momentum at which electrons and muons can be cleanly identified, which is less than  $500 \text{ MeV}/c$  for the  $B\text{A}\text{B}\text{A}\text{R}$  detector.

### 2.2.2 Charged Track Resolution and Multiple Scattering

As a result of the characteristically low momentum of particles produced by  $B$  mesons decaying at rest, the errors on charged particle track parameters are usually dominated by multiple Coulomb scattering rather than by the intrinsic spatial resolution of the tracking chamber. For example, the resolution in measuring the longitudinal ( $z$ ) position of the  $B$  decay vertex, and therefore the decay time, is determined primarily by the amount of material before the first measurement in the vertex detector and by the radius of the beam



**Figure 2-2.** Detection efficiency for the decay  $B^0 \rightarrow J/\psi K_S^0 \rightarrow e^+ e^- \pi^+ \pi^-$  as a function of forward polar angle coverage for different cuts on the minimum transverse momentum accepted. The dotted line shows the efficiency as a function of backward polar angle coverage.

pipe. This leads to a strategy of making the first  $z$  measurement as close as possible to the beam pipe. It is also important to put the second layer as close as possible with little intervening material for the fraction of tracks in which there is no useful hit in the first layer.

Multiple Coulomb scattering also dominates the angle and momentum resolutions for most tracks. Good momentum resolution requires a continuous tracking volume filled with a mixture of gas and wires that has a long track length and long average radiation length. Another way to improve momentum resolution is to increase the magnetic field, although one must be careful not to compromise the acceptance for low momentum particles. For most of the momentum range useful for  $B$  meson studies, the angle measurements are made primarily in the vertex detector, before the original angle information is degraded by multiple scattering.

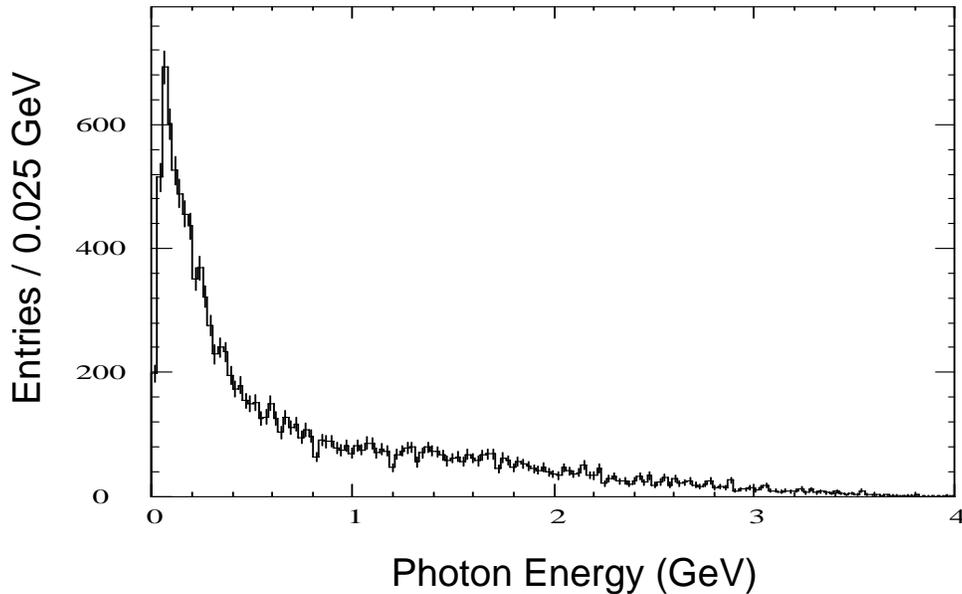
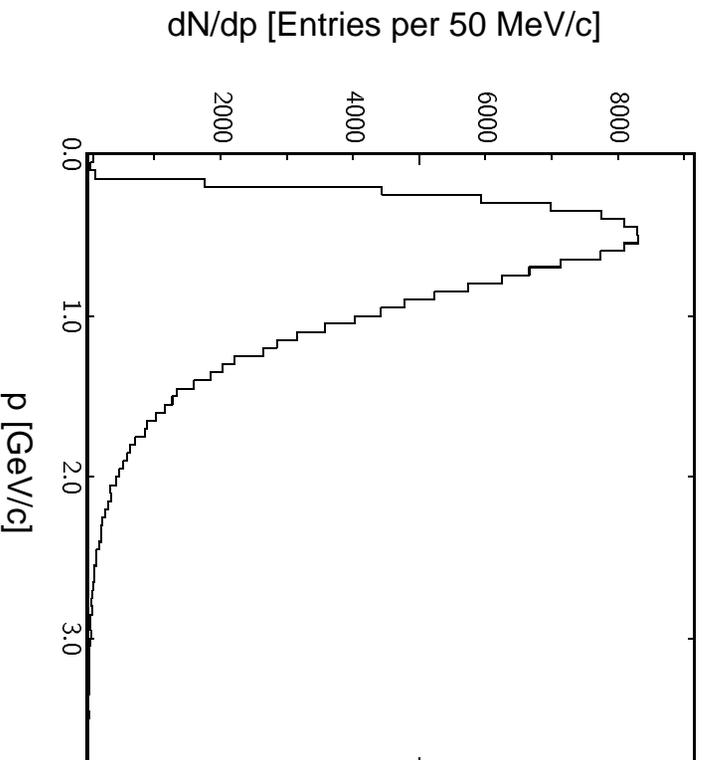


Figure 2-3. Energy spectrum for photons from the process  $B^0 \rightarrow \rho^\pm \pi^\mp$ ,  $\rho^\pm \rightarrow \pi^\pm \pi^0$ .

### 2.2.3 Photon Efficiency and Resolution

The low momentum range determines the detector requirements for photon detection even more than for charged particles. Very good energy and angular resolution in the electromagnetic calorimeter for photons in the energy range from 20 MeV to 5 GeV are important. Figure 2-3 shows the energy spectrum of  $\gamma$ s from the process  $B^0 \rightarrow \rho^\pm \pi^\mp$ ,  $\rho^\pm \rightarrow \pi^\pm \pi^0$ , which is a very important  $CP$  mode. The  $\pi^0$  spectrum from the modes  $D^{*+} D^{*-}$ ,  $J/\psi K^{*0}$ , and  $J/\psi K_s^0$  is softer than that from  $\rho\pi$ , while the spectrum from  $\pi^0\pi^0$  is harder. This leads to the choice of a CsI calorimeter with fairly fine segmentation. Good resolution at high energy requires that the crystals be sufficiently long so that the rear leakage is negligible. Low-energy resolution requires that the electronic noise be kept low.

To achieve the required physics performance, one must also keep the number of radiation lengths of material before the calorimeter to a minimum. The major sources of such material are the particle identification device, and, in the endcap region, the drift chamber endplate and associated hardware. In the barrel region, much of the material is close to the front face of the calorimeter. Material in this location is the least troublesome, since both electrons from a photon conversion often reach the calorimeter.



**Figure 2-4.** *Momentum spectrum for kaons from  $B$  decays, which are used for tagging the flavor of the  $B$ .*

## 2.2.4 Identification of Hadrons

An essential requirement for the  $CP$  physics is to identify kaons for tagging with high efficiency and low probability of wrong sign tags. As can be seen in Figure 2-4, most of the kaons from the process  $B \rightarrow DX$ ,  $D \rightarrow KY$  have momentum less than 1 GeV/ $c$ . There is a tail at higher momentum, however, which occurs at forward angles.

The other requirement is to separate pions from kaons in decays such as  $B^0 \rightarrow \pi^+\pi^-(K^+\pi^-)$ , as well as in charmed meson and  $\tau$  decays. At the extreme decay angle accepted by the experiment, one pion from  $B^0 \rightarrow \pi^+\pi^-$  decay has a momentum of  $\sim 4.2$  GeV/ $c$  at  $\cos\theta_{lab} = 0.95$ , and the other has a momentum of  $\sim 1.5$  GeV/ $c$  at  $\cos\theta_{lab} = -0.69$ . Although kinematic fitting provides some help in discriminating between these two-body decay modes, achieving complete separation requires a Cherenkov detector for particle identification.

These physics requirements lead to a choice of particle identification system which combines good  $dE/dx$  information from the drift chamber to identify kaons with momenta below 0.7 GeV/ $c$  and a Cherenkov device in the barrel and forward endcap regions for higher

momenta, giving good  $K/\pi$  separation up to about 4 GeV/ $c$ . There are no two-body decay pions in the backward endcap region, so there is no reason to supplement  $dE/dx$  there.

### 2.2.5 Interaction with the Accelerator

PEP-II represents a new type of  $e^+e^-$  collider and is, therefore, an ambitious project. Highest priority is being given to making the machine reliable and easy to operate with consistently high luminosity and low background. Machine-design choices for the interaction region impact the detector design. A support tube containing the interaction region magnets, the masks to reduce beam-related backgrounds, the beam pipe, and the vertex detector, is envisioned to run through the interaction region at a radius of  $\sim 20$  cm. This is the best method for achieving precise and reproducible relative alignments of the machine components within the detector. While the support tube introduces  $\sim 0.005X_0$  of extra material into the fiducial volume of the detector, the effect on resolution is rather small. The angles and impact parameters are measured in the vertex detector, inside the support tube. The transverse momentum is measured primarily in the drift chamber, in which there is very little material for multiple scattering. Over the momentum range of charged particles from  $B$  decay, the degradation in momentum resolution due to the support tube is always less than 5%. Clearly, the prospect of better machine performance and reproducible alignment of the vertex detector outweighs these consequences for tracking resolution.

The accelerator magnets closest to the interaction point also have consequences for the detector design. The need to separate the two beams before the first parasitic crossing necessitates dipole magnets 21 cm from the collision point; the need to focus the beams to a very small size necessitates having quadrupoles, which are also placed completely within the detector volume. The dipole magnets complicate the vertex detector mechanical design because they occupy much of the region in which the readout electronics and mechanical support would be most conveniently located. The length of the detector solenoid magnet is also limited, because if it were to become too large, it would be necessary to shield a machine quadrupole (Q2) that contains iron, causing severe complications for the detector and accelerator, and making access to inner detector components more difficult. There has been a continuing joint effort of physicists and engineers working on the accelerator and detector to understand these conflicts and to arrive at solutions which serve the common goal of obtaining the best physics possible.

Because of the unprecedented beam currents expected during PEP-II operation, there has been a great deal of attention given to backgrounds in the various detectors. Detailed simulations of the synchrotron radiation and lost beam-particle backgrounds have been refined continuously since the earliest days of the machine design and have had substantial impact on the design of the interaction region. The anticipated background levels, inflated

by a reasonable safety factor, are used to set detector parameters such as radiation hardness of silicon electronics and bandwidth of the data acquisition system.

### 2.2.6 Considerations of Cost and Schedule

Although the primary considerations in designing the detector are physics requirements and detector capabilities, the process should and does involve consideration of cost. Often, one must choose among different ways to obtain comparable performance that have very different costs. As in most such detectors, the calorimeter is the most expensive single system. Because the calorimeter cost increases with volume, every effort has been made to minimize calorimeter volume as much as possible within the constraints imposed by physics performance.

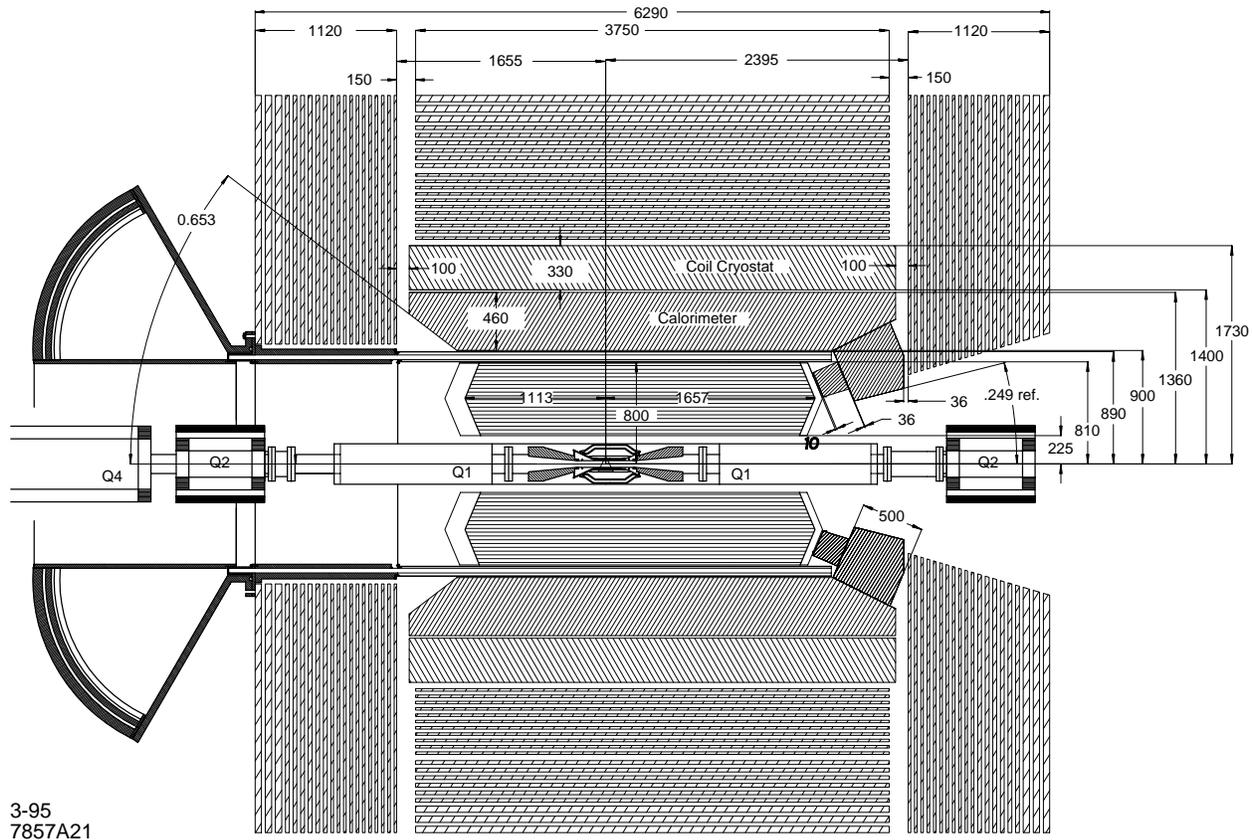
The schedule for PEP-II construction encompasses a six-month commissioning run starting in the fall of 1998. Detector commissioning off the beam line should commence at the same time, so that the detector is ready for installation into the interaction region at the end of the machine commissioning period. This represents an extremely tight schedule on which to design, engineer, construct, and assemble an experimental apparatus as large and complex as *BABAR*. In considering detector technologies and alternate experimental designs, schedule constraints have been an important criterion.

## 2.3 The Experimental Design

---

The *BABAR* design is shown in Figures 2-5 and 2-6. It consists of a silicon vertex detector, a drift chamber, a particle identification system, a CsI electromagnetic calorimeter, and a magnet with an instrumented flux return. The superconducting solenoid is designed for a field of 1.5 T, and the flux return is instrumented for muon identification and coarse hadron calorimetry. All of these detectors operate with good performance for laboratory polar angles between  $17^\circ$  to  $150^\circ$ , corresponding to the range  $-0.95 < \cos \theta_{cm} < 0.87$ . The vertex detector is mounted inside the support tube along with the beam pipe, the first accelerator dipole, and quadrupole magnets. In this section, we give a brief description of the detector; in Section 2.5, we summarize the performance of the detector subsystems.

The detector coordinate system is defined with  $+z$  in the boost (high-energy beam) direction and  $+y$  in the vertical direction. The high-energy beam travels clockwise around PEP-II, so the  $+x$  direction is away from the ring center. The coordinate system origin is the nominal collision point, which is offset in the  $-z$  direction from the geometrical center of the detector magnet. Although the beams collide with each other head-on, they are separated

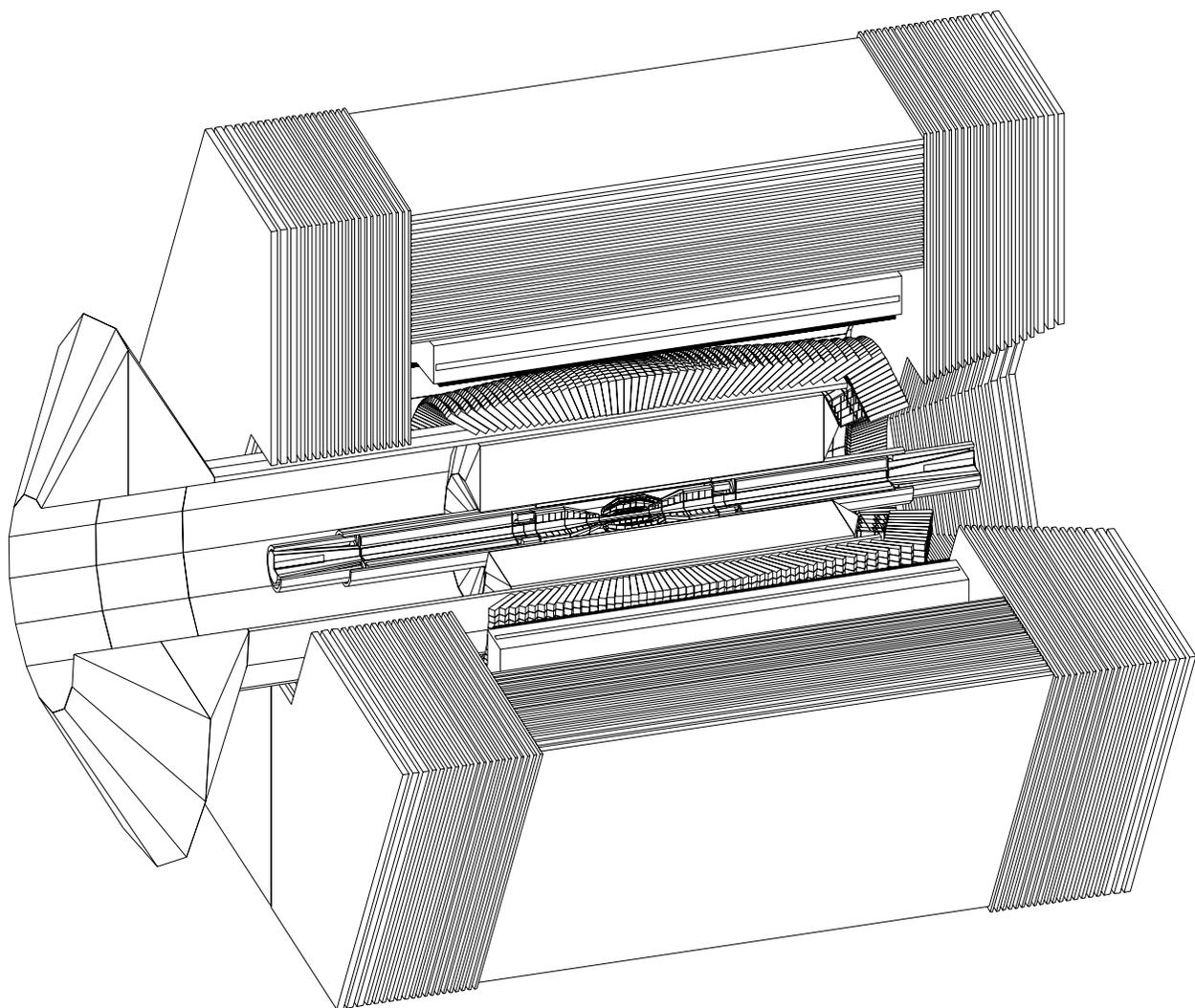


3-95  
7857A21

Figure 2-5. Cross-sectional view of the detector.

while still inside the detector magnet field. The detector is rotated 20 mrad relative to the beam direction (around the  $y$  axis) to minimize the resulting orbit distortions. The  $z$  direction thus corresponds to the magnetic field direction, and deviates slightly from the boost direction.

The tracking system in *BABAR* consists of the vertex detector and a drift chamber. The vertex detector is the only tracking device inside the 20 cm radius of the support tube. It is used to measure precisely both impact parameters for charged tracks ( $z$  and  $r - \phi$ ); these measurements are used to determine the difference in decay times of the two  $B^0$  mesons. The vertex detector also provides the measurements of production angles, given the momentum information from the drift chamber. Finally, charged particles with  $p_t$  between  $\sim 40$  MeV/ $c$  and  $\sim 100$  MeV/ $c$  are tracked only with the vertex detector, which must therefore provide good pattern recognition. To serve these various functions, the vertex detector requires excellent spatial resolution, low multiple scattering, small segmentation, and reasonably good resistance to radiation.



**Figure 2-6.** *Three-dimensional view of the detector.*

The vertex detector consists of five layers of double-sided silicon strip detectors. The inner three layers are in a barrel geometry with detectors parallel to the beam pipe. The outer two layers combine barrel detectors in the central region with wedge detectors forward and backward. The entire vertex detector contains about  $0.94\text{ m}^2$  of silicon and about 150,000 readout channels. This is somewhat larger than silicon vertex detectors built to date, although it is substantially smaller than those designed for LHC. The size is determined by its dual function as a vertex detector and as the only tracking device within a radius of 20 cm.

The second component of the tracking system is the drift chamber, which is used primarily to achieve excellent momentum resolution and pattern recognition for charged particles with  $p_t > 100\text{ MeV}/c$ . It also supplies information for the charged track trigger and a measurement of  $dE/dx$  for particle identification. The chamber extends in radius from 22.5 cm, just outside the support tube, to 80 cm.

For most particles of interest at PEP-II, the optimum momentum resolution is achieved by having a continuous tracking volume with a minimum amount of material to cause multiple scattering. By using a helium-based gas mixture with low-mass wires and a magnetic field of 1.5 T, very good momentum resolution can be obtained. The forward edge of the chamber is situated 1.66 m from the interaction point, which makes it possible to obtain reasonable momentum resolution down to the limit of forward acceptance, 300 mr.

Two design options were considered for the drift chamber: a conventional layout with four axial and six stereo superlayers, each consisting of four individual layers; and an all-stereo arrangement of forty layers with alternating U and V stereo angles. The momentum and angular resolutions for the entire tracking system were found to be comparable for the two chambers. The axial/stereo chamber makes it more straightforward to design a charged particle trigger in hardware which has a fairly sharp cutoff in  $p_t$ . The axial/stereo arrangement has been adopted as the baseline design for the drift chamber.

The chamber is designed to minimize the amount of material in front of the particle identification and calorimeter systems in the heavily populated forward direction. The readout electronics are mounted only on the backward end of the chamber, and the endplates are designed as truncated cones probably to be made of carbon fiber.

As stated above, there are two primary goals for the particle identification system. One is to identify kaons for tagging beyond the momentum range well-separated by  $dE/dx$ . The other is to identify pions from few-body decays, such as  $B^0 \rightarrow \pi^+\pi^-$  or  $B^0 \rightarrow \rho\pi$ . A new detector technology is required to meet these goals, and in the barrel region, a DIRC (Detector of Internally Reflected Cherenkov radiation) is used. Cherenkov light produced in  $1.75 \times 3.5\text{ cm}^2$  quartz bars is transferred by total internal reflection, while preserving the angle, to a large water tank outside the backward end of the magnet. The light is observed by an array of photomultiplier tubes at the outside of the tank, where images governed by the Cherenkov

angle are formed. A mirror at the forward end of the bars reflects the forward-going light, preserving the angle information. This arrangement provides at least 4 standard deviation  $\pi/K$  separation up to almost the kinematic limit for particles from  $B$  decays.

In the forward region, for  $\cos\theta > 0.90$ , there are two layers of aerogel threshold Cherenkov (ATC) counters, one of higher density ( $n = 1.055$ ) with a kaon threshold at  $\sim 1.6$  GeV/ $c$ , and the other a lower density ( $n = 1.0065$ ) for high-momentum identification. This provides good tagging efficiency in the forward region, as well as allowing particle identification for pions from  $B^0 \rightarrow \pi^+\pi^-$ . In the backward region, no particle identification is needed other than  $dE/dx$ . The tagging kaons are quite slow there, and for any  $B^0 \rightarrow \pi^+\pi^-$  event with a pion in the backward endcap region, the other pion is outside the limiting forward acceptance.

The electromagnetic calorimeter must have superb energy resolution down to very low photon energies. This is provided by a fully projective CsI(Tl) crystal calorimeter, which has excellent energy and angular resolution and retains high detection efficiency at the lowest relevant photon energies. The calorimeter consists of a cylindrical barrel section with inner radius of 90.5 cm and a conical forward endcap. The barrel calorimeter contains 5880 trapezoidal crystals; the forward endcap calorimeter contains 900 crystals. The crystal length varies from  $17.5X_0$  in the forward endcap to  $16X_0$  in the backward part of the barrel; the typical transverse size is  $\sim 4.8 \times 4.8$  cm<sup>2</sup> at the front face.

Each crystal is read out by two independent silicon photodiodes. Electronic noise and beam-related backgrounds dominate the resolution at low photon energies, while shower leakage from the rear of the crystals dominates at higher energies. As mentioned above, minimizing the calorimeter volume is an important consideration in choosing the detector geometry; in the present design, it is  $\sim 5.9$  m<sup>3</sup>.

To achieve very good momentum resolution without increasing the tracking volume and therefore the calorimeter cost, it is necessary to have a field of 1.5 T. The magnet is therefore of superconducting design, with an inner radius of 1.40 m for the coil dewar and a cryostat length of 3.85 m. The total thickness of the cryostat and coil is 0.25–0.4 interaction lengths, which is thin enough for reasonable efficiency in detecting neutral hadrons that traverse it. The magnet is similar to many operating detector magnets, so the engineering and fabrication should be straightforward. The nonstandard features are segmentation of the iron for an Instrumented Flux Return (IFR), and the complications caused by the DIRC readout in the backward region.

The IFR is designed to identify muons with minimum momentum around 0.5 GeV/ $c$  and to detect neutral hadrons. The magnet flux return iron is divided into 20 layers, and the thickness increases from 2–5 cm. Between most of the iron absorber layers are 3 cm-wide gaps with Resistive Plate Chambers (RPCs), which serve as the active detectors. A system with 17 layers of RPCs in the barrel and 16 in the endcap has been costed. With this configuration, it is possible to identify muons down to 0.5 GeV. As an additional benefit,

the IFR serves as a coarse hadron calorimeter, making it possible to detect  $K_L^0$ s. The RPCs represent a proven technology which adapts well to the *BABAR* geometry.

The high data rate at PEP-II requires a data acquisition system which is more advanced than those used at present  $e^+e^-$  experiments. The bunch crossing period of 4.2 ns is so short that the interactions are effectively continuous, as in fixed-target experiments. The rate of all processes to be recorded at the design luminosity of  $3 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$  is about 100 Hz. Simulations of machine backgrounds show hit rates of about 100 kHz per layer in the drift chamber and about 140 MHz in the first silicon layer. The goal is to operate with negligible deadtime even if the backgrounds are 10 times higher than present estimates, an environment which might develop, especially early in the life of the experiment.

In the electronics architecture adapted for *BABAR*, the experiment can take data in parallel with processing and transporting the information from previous events. In addition, a unified architecture for all detector systems minimizes the overall cost while enhancing the reliability of the system. The memory buffers that traditionally exist at the event-building stage are replaced with large storage capacity immediately after digitization. This allows long latencies in the data collection, which in turn decouples data acquisition from readout of the data.

The Level 1 trigger uses trigger primitives provided by special front-end interfaces from the drift chamber and electromagnetic calorimeter. The trigger decisions are generated with a fixed  $9.5 \mu\text{s}$  latency at a maximum rate of 10 kHz at 10 times the nominal machine backgrounds. The Level 1 trigger is required for readout of the vertex detector because of the high hit rates from machine backgrounds. An optional Level 2 trigger could be used to reduce the number of triggers before data are extracted and transmitted. A final software Level 3 trigger running in commercial processors reduces the number of events to a number that can be written to archival storage.

The computing loads are also larger than those for previous experiments at  $e^+e^-$  colliders, although they are within what is handled by existing hadron experiments. The wide geographical distribution of the collaboration adds an additional consideration in the design of the computing system. Although much of the computing system will necessarily be installed at the SLAC site, there will also be regional centers to provide more convenient access to the data for the widely distributed institutes that comprise the *BABAR* Collaboration. Finally, there will be significant computing installations at home institutions with network links to the regional centers and the SLAC site.

The online requirement for CPU is 3000 MIPS, assuming a 3 kHz rate input to the Level 3 farm. The total CPU power needed for reconstruction, Monte Carlo, DST analysis, and interactive analysis is about 17,000 MIPS. The storage requirements for *BABAR* are approximately 100 Tbytes/yr of tape storage. In addition, about 2 Tbytes of disk storage is needed for a staging space for physics analysis. Aggregate network capacity for offline tasks of more

than 200 Mbytes/s is required, dominated by the reconstruction passes and multiple reads through the data for DST creation and analysis.

A common computing environment based on clusters of distributed UNIX workstations and industry-standard networks is being developed to handle this massive computing demand. Flexible solutions are being chosen to take advantage of the evolution that the technologies will inevitably undergo during the *BABAR* construction schedule.

## 2.4 Detector Optimization

---

In the process of arriving at the *BABAR* design described above, the collaboration considered a wide range of detector technologies and geometry options. The motivations for some of the important choices made early in the design process are described in the *Letter of Intent*. Since that time, two important decisions were made which have a large impact on the design of the entire experiment. These were the choices of hadron identification technology and the optimization of the various detector dimensions, especially the outer radius of the main tracking chamber.

### 2.4.1 Impact of Particle Identification

At the time of the *Letter of Intent*, *BABAR* had under consideration three types of Cherenkov detectors to extend hadron identification beyond the momentum region covered by  $dE/dx$  information from the drift chamber. The design and engineering of the entire detector depended on choice of technology for particle identification. The DIRC occupied less radial space than either a fast RICH or ATC counters, but the DIRC required penetration of the backward endcap region for its optical readout. Thus integration engineering for the experiment could not proceed before this decision was made.

A comprehensive review of the different technologies looked at performance obtained by prototypes, simulation of expected physics performance, impact on the rest of the experiment, and prospects for meeting the tight construction schedule. The collaboration selected the DIRC solution. Because it uses the smallest radial space of the three candidate solutions, about 10 cm including mechanical supports and clearance space, it minimizes the cost of the CsI calorimeter. It does, however, complicate the engineering of the backward endcap and the access to the detector. Detailed studies indicate that access is possible to the end of the drift chamber in approximately one eight-hour shift. This is essential to avoid long downtimes which could compromise the factory-like performance of the machine and detector.

Since the DIRC is made of long barrel staves, it does not provide coverage in the very forward direction. This hole in the acceptance is filled by the two layers of aerogel threshold counters. The layers have indices of refraction which allow  $K/\pi$  separation at momenta up to about 4.3 GeV/ $c$ . The DIRC plus aerogel, after allowing for small gaps between quartz modules, cover about 95% of the acceptance of the tracking devices.

### 2.4.2 Optimization of Performance versus Cost

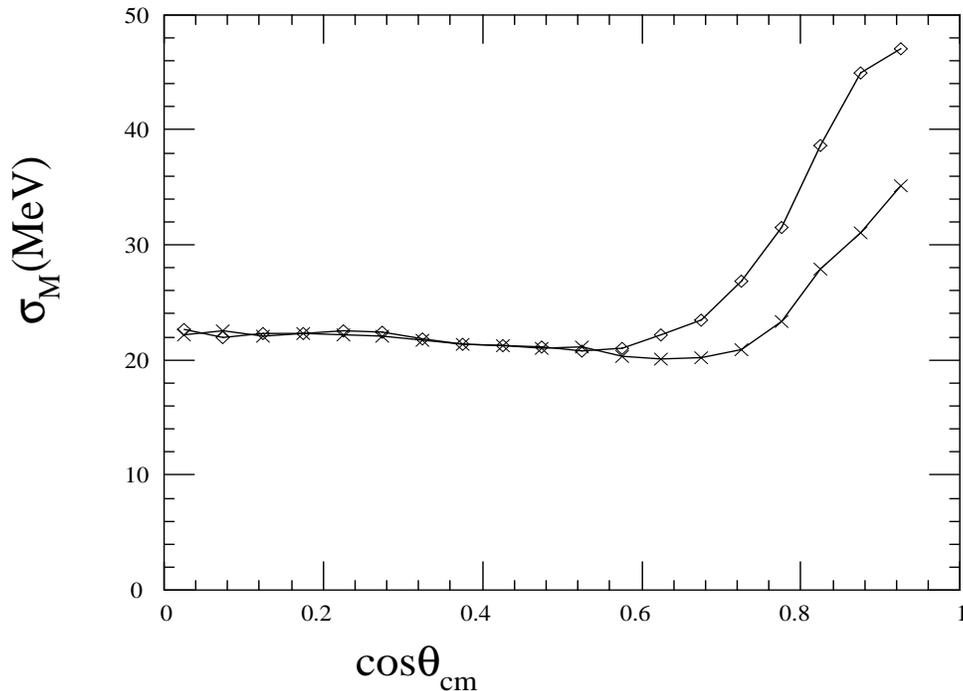
The optimization of the *BABAR* detector geometry is tightly constrained by the accelerator design and by cost considerations. The configuration of the interaction region, in particular the dimensions and locations of the  $B1$  magnets, determine the angular acceptance in both forward and backward directions. The resulting center of mass acceptance of  $-0.95 < \cos \theta_{cm} < 0.87$  is reasonably close to the preferred symmetric acceptance.

The forward length of the detector is determined primarily by the need to maintain good momentum resolution and efficiency for charged tracks over this entire acceptance. In conventional solenoidal detectors, the resolution in the forward direction is significantly degraded because the particles exit the drift chamber at a relatively small radial distance from the beam. This is unacceptable at PEP-II, where the fraction of particles going forward is greater and their momentum is higher. In order for tracks at 300 mr to pass through 20 of the 40 drift chamber layers, corresponding to a radius of 52 cm, the active volume of the drift chamber extends to a distance of 166 cm from the interaction point.

Figure 2-7 shows the mass resolution for  $B^0 \rightarrow \pi^+\pi^-$  as a function of the  $\cos \theta_{cm}$  of the faster pion. There is a clear break at  $\cos \theta_{cm} = 0.7$ , where the pions begin to exit the chamber without traversing all 40 layers. The fraction of events with degraded resolution and the average degradation would increase if the forward length of the drift chamber were shortened.

This effect can be seen even when the resolution is averaged over all decay angles. Figure 2-8 shows the mass resolutions for  $B^0 \rightarrow \pi^+\pi^-$  and  $D^+ \rightarrow K^-\pi^+\pi^+$  as a function of the forward length of the drift chamber. The  $B^0 \rightarrow \pi^+\pi^-$  resolution in particular increases significantly as the chamber is shortened. Probably more important is the fact that the track-finding efficiency would suffer for forward tracks traversing fewer than 20 layers, especially with high rates in the inner layers from machine backgrounds. The solid angle coverage of the DIRC in the center of mass also depends on the forward length of the detector. In the backward direction, the drift chamber endplate is located a distance of 111 cm from the interaction point, reflecting the lower momentum of the tracks.

The detector length beyond the forward drift chamber endplate is strongly constrained by the necessity to minimize the fringe field of the detector solenoid at the location of the

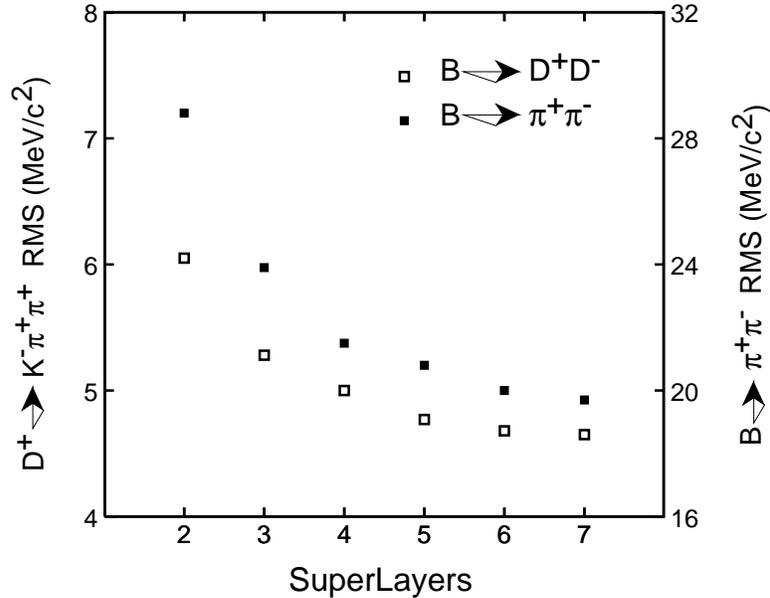


**Figure 2-7.** Mass resolution for  $B^0 \rightarrow \pi^+\pi^-$  as a function of the  $\cos\theta_{cm}$  of the forward-going pion. The lower curve is for the chosen design, and the upper curve is for a forward length of the drift chamber which is shortened by 44 cm, corresponding to only four superlayers at 300 mr.

quadrupole Q2 beam line magnet. There is just enough room for two layers of ATC counters and the forward calorimeter. The resulting detector length, combined with the calorimeter depth, then determines the barrel-endcap calorimeter interface.

The largest single element in the detector cost is the CsI calorimeter, and much effort has gone into minimizing its cost. The length has been reduced to the minimum compatible with the demands on forward tracking just discussed. The other primary degree of freedom is the choice of the outer radius of the drift chamber. The present radius of 80 cm is the minimum compatible with good performance of the tracking system. Combined with the inner radius of 22.5 cm, this small outer radius puts significant demands on the charged particle tracking detectors to achieve adequate performance, both in momentum resolution and in pattern recognition, within the limited space available.

Figure 2-9 shows the mass resolutions for  $B^0 \rightarrow \pi^+\pi^-$  and  $D^+ \rightarrow K^-\pi^+\pi^+$ , plotted as a function of the outer radius. The  $B^0 \rightarrow \pi^+\pi^-$  mass resolution degrades by about 10% for each 4 cm change in the outer radius; there are significant effects in the  $D^+$  resolution as well. The effects on the robustness of pattern recognition may be greater, especially if

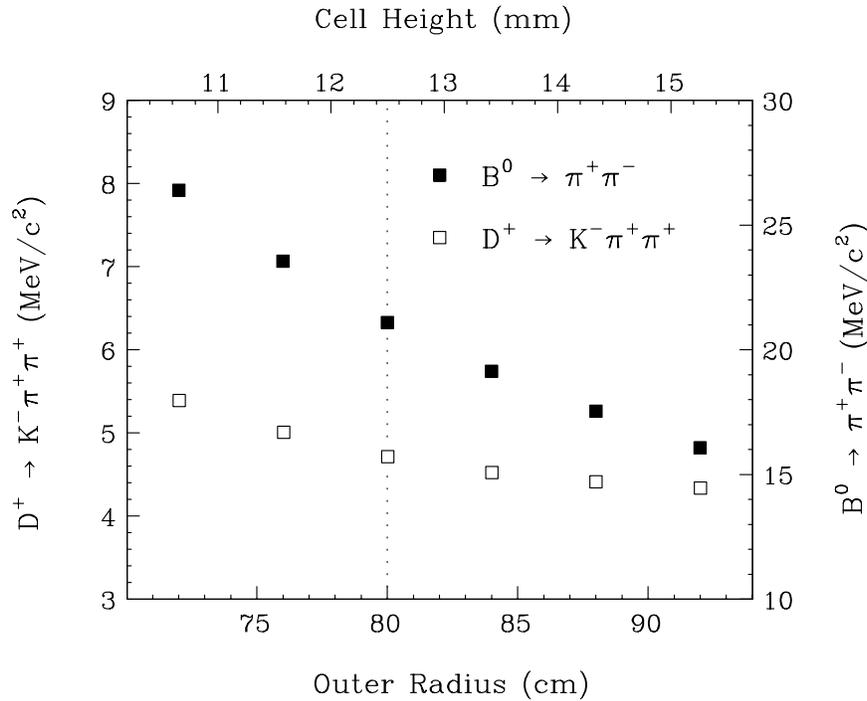


**Figure 2-8.** Mass resolutions for  $B^0 \rightarrow \pi^+\pi^-$  and  $D^+ \rightarrow K^-\pi^+\pi^+$  as functions of the forward length of the drift chamber. A superlayer corresponds to four layers, and the present design has tracks at 300 mr crossing five superlayers.

machine backgrounds are high, although they are harder to quantify. Experience indicates that pattern recognition with high background rates requires at least 40 layers in the drift chamber; if radial space were not at such a premium, more layers would be desired. Given this minimum number of layers, the only way to reduce the drift chamber radius further is to reduce the cell size. The cell height of 12.5 mm in the present design is about the minimum practical size both for mechanical integrity of the endplate and for uniformity of the drift cell.

The second major reduction in calorimeter cost came from removing about one radiation length from the depth of the calorimeter given in the *Letter of Intent* design, averaged over the full calorimeter. This does affect the resolution for high energy showers, and some effort was made to leave enough length in the forward endcap calorimeter to reduce the serious impact of shower leakage there. The mass resolution for the benchmark mode  $B^0 \rightarrow \pi^0\pi^0$  changes by about 10% for every 0.5 radiation length variation in this length. Significant but smaller effects are observed through the many physics analyses using  $\gamma$ s and  $\pi^0$ s, such as  $B^+ \rightarrow \pi^+\pi^0$ .

Additional savings on the calorimeter cost were achieved at some cost in physics performance. The total number of crystals has been reduced from about 10,000 in the LOI to the present 6780. This reduces costs for both crystal production and electronics. A crystal in the center



**Figure 2-9.** Mass resolutions for  $B^0 \rightarrow \pi^+ \pi^-$  and  $D^+ \rightarrow K^- \pi^+ \pi^+$ , plotted as functions of the outer radius of the drift chamber. Also shown is the corresponding cell size.

of the barrel subtends an angle of about 5 mrad. The angular resolution is approximately proportional to the width of the crystal, and the  $\pi^0$  mass resolution is proportional to the angular resolution. Further reduction of the segmentation would also increase the fraction of showers which contain energy from nearby particles.

A backward endcap calorimeter was designed to sit behind the drift chamber endcap and inside the DIRC bars. It would lie outside the angular acceptance of the vertex detector, and would subtend about 1.5% of full solid angle in the center of mass. The physics motivations were two-photon physics and a slight improvement of the hermeticity of the detector. As part of the cost optimization, the backward endcap calorimeter was removed from the initial detector design. Space has been left for addition of this item at a later date to enhance the physics capability of the detector.

The efforts to minimize the drift chamber radius and length also minimize the cost of the solenoidal magnet. Although the magnet flux return design has 20 iron layers, a minimal system of 16–17 layers could be instrumented with RPCs at the start of the experiment to save production costs. The magnet cost would not be reduced if there were fewer iron layers.

In the vertex detector, the one saving that was seriously considered was removal of the middle layer, reducing the total number of layers to four. Because much of the cost for any vertex detector lies in detector development, electronics development, mechanical engineering, and prototypes, the cost savings turned out to be rather small. Given this fact, the special need in this experiment for excellent efficiency of the vertex detector, and the severe limits on the access to the detector, led to the choice of a five-layer design. The extra layer provides redundancy that will be needed if channels or sections of the detector fail over the expected long-running periods without access for repairs.

The *BABAR* detector should be capable of carrying out the rich physics program available at PEP-II for a period of ten years or more. The detector design described in this document is the result of a continuing process of balancing this requirement with a hard realism about the need to match funds that are available to the collaborators over a rather short period of time. As a result of this process, the physics performance has suffered measurably, but not yet enough to severely limit the future of the experiment. Other steps were considered to reduce the cost further, but they would have severely compromised the physics performance for the duration of *BABAR*.

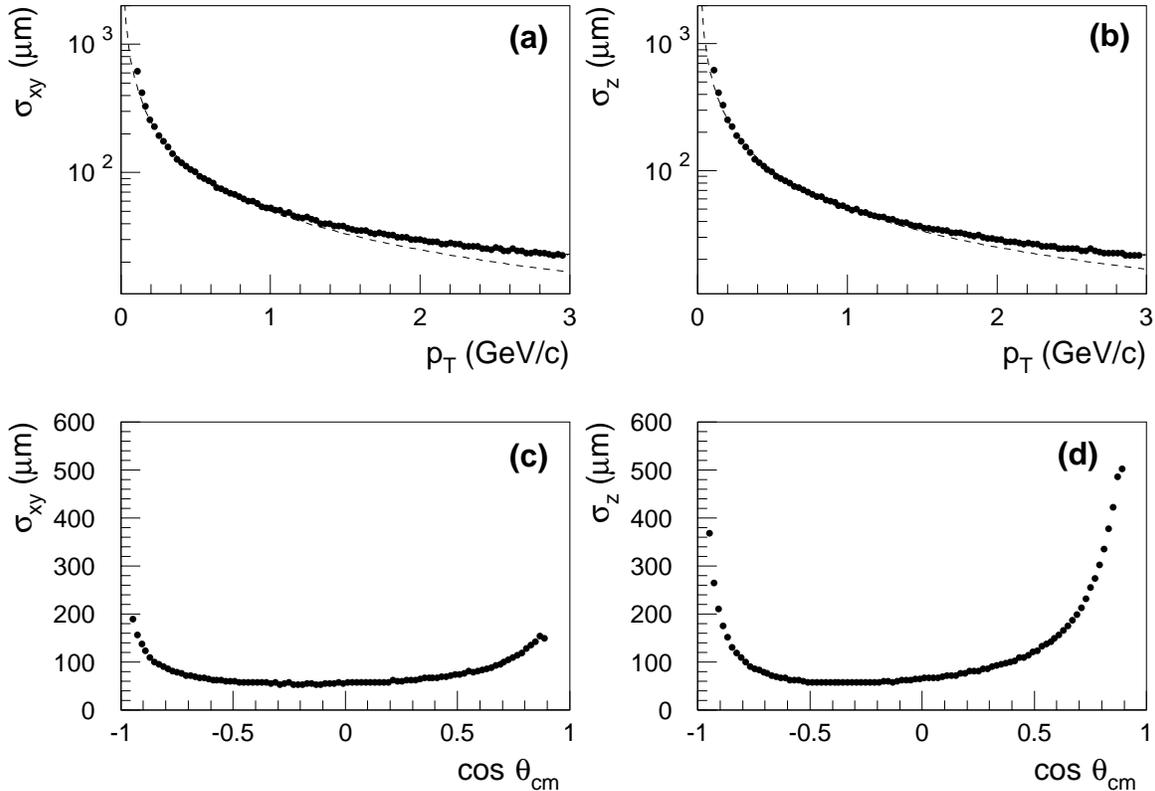
## 2.5 Detector Performance

---

In this section, we describe the performance that is projected for each of the detector subsystems. We give only the most important results; more detailed discussions appear in the individual system chapters. Where possible, we show the polar angle dependence in the center-of-mass frame. The primary performance characteristics are efficiency and resolution. The luminosity required to obtain a given sensitivity is inversely proportional to the efficiency and directly proportional to  $(1+B/S)$ , where  $B/S$  is the ratio of background events to signal events. The ratio  $B/S$  varies widely for different decay processes, but is proportional to the resolution in the invariant  $B$  mass. The background also increases with the mass resolution for intermediate states, such as a  $D^+$  or  $\pi^0$ . A summary of the major parameters of all the detector subsections is given in Table 2-2.

### 2.5.1 Vertex Detector

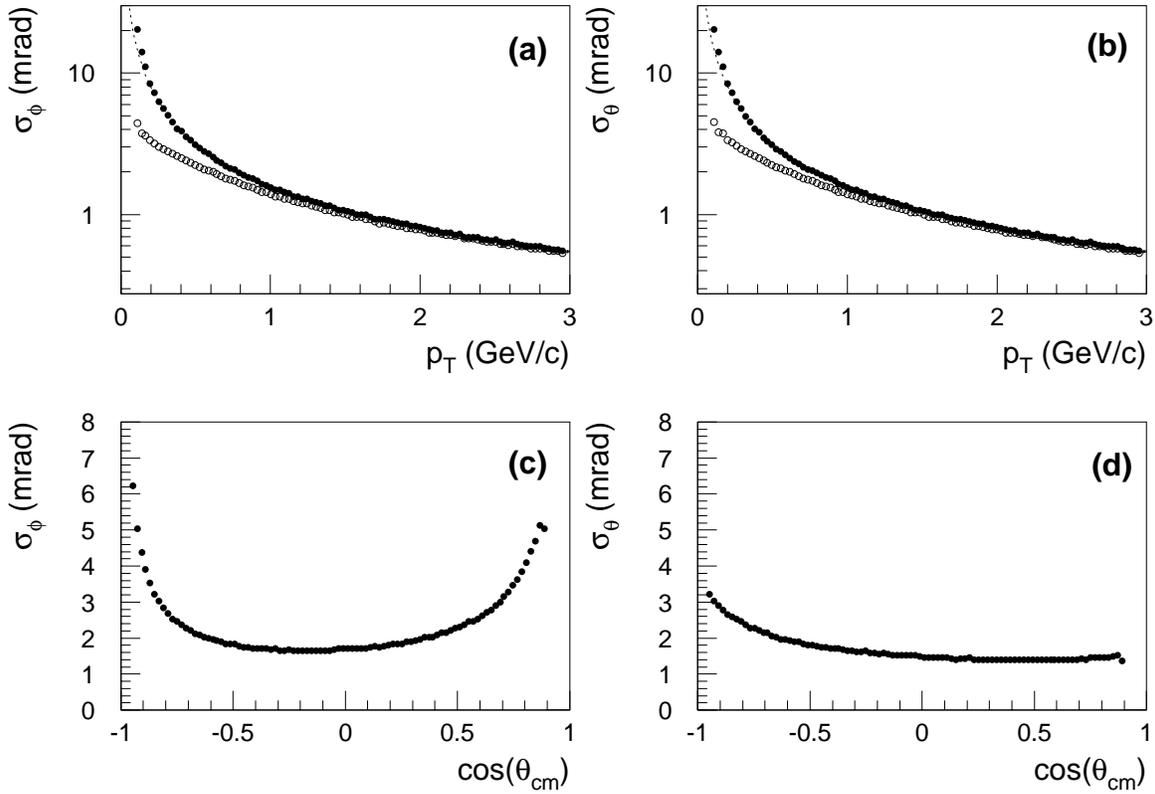
The vertex detector information dominates the measurement of the track impact parameters, both along and perpendicular to the beam direction. Background rejection using vertex information improves with better impact parameter resolution. Figure 2-10 shows the resolution for both impact parameters as a function of  $p_t$ . These resolutions are close to the lower limits that are imposed by multiple scattering in the beam pipe and the active



**Figure 2-10.** Resolutions for impact parameters at the vertex: (a)  $\sigma_{xy}$  as a function of  $p_t$  at  $\cos\theta_{lab} = 0$ ; (b)  $\sigma_z$  as a function of  $p_t$  at  $\cos\theta_{lab} = 0$ ; (c)  $\sigma_{xy}$  as a function of  $\cos\theta_{cm}$  for  $p_{cm} = 1 \text{ GeV}/c$ ; and (d)  $\sigma_z$  as a function of  $\cos\theta_{cm}$  for  $p_{cm} = 1 \text{ GeV}/c$ . For (a) and (b), the dashed curve represents  $50 \mu\text{m}/p_t$ .

silicon. At  $\cos\theta_{lab} = 0$ , which is  $\cos\theta_{cm} = -0.5$ , the resolution for a single track is about  $50 \mu\text{m}/p_t \oplus 15 \mu\text{m}$  in both dimensions. (All such momenta are assumed to be in  $\text{GeV}/c$ .)

The second set of plots in Figure 2-10 shows the dependence of resolution on  $\cos\theta_{cm}$  at constant  $p_{cm} = 1 \text{ GeV}/c$ . One can use  $1/p$  scaling to determine the error for other momenta. The  $xy$  resolution is below  $100 \mu\text{m}$  over the range  $-0.85 < \cos\theta_{cm} < 0.70$ . The angular dependence is more symmetric in the center of mass than one might expect; although the tracks traverse more material at forward laboratory angles, they also have higher momentum in the laboratory frame. The  $z$  resolution is similar to the  $xy$  resolution for most of the angular range but becomes larger as the track angle approaches the  $z$  axis. The two resolutions are approximately related by the expression  $\sigma_z \sim \sigma_{xy}/\sin\theta_{lab}$ . Determining the resolution in vertex separation between two  $B$  decays requires a full physics simulation; results from this study are discussed in Section 2.6.2.



**Figure 2-11.** Resolutions for angle measurements: (a)  $\sigma_\phi$  as a function of  $p_t$  at  $\cos\theta_{lab} = 0$ ; (b)  $\sigma_\theta$  as a function of  $p_t$  at  $\cos\theta_{lab} = 0$ ; (c)  $\sigma_\phi$  as a function of  $\cos\theta_{cm}$  for  $p = 1$  GeV/c; and (d)  $\sigma_\theta$  as a function of  $\cos\theta_{cm}$  for  $p = 1$  GeV/c. For (a) and (b), the dashed curve corresponds to  $1.6 \text{ mr}/p_t$ , and the open circles correspond to the resolutions using a vertex constraint, as described in the text.

The two angles for charged tracks are also determined primarily in the vertex detector, assuming that the curvature is measured in the drift chamber. Figure 2-11 shows the resolution in both angles as a function of  $p_t$  and polar angle. At  $\cos\theta_{lab} = 0$ , the resolution is about  $1.6 \text{ mr}/p$  in both  $\theta$  and  $\phi$ . The angular resolutions are quite good over most of the acceptance of the experiment, good enough that the momentum error dominates the mass resolution in almost all cases. For momenta less than about  $0.5 \text{ GeV}/c$ , the angles are determined better by using a vertex constraint constructed using the faster tracks. Figure 2-11 shows the effects of such a constraint, assuming a reasonable error on the reconstructed vertex of  $100 \mu\text{m}$ . In particular, for slow pions coming from the process  $D^{*+} \rightarrow \pi^+ D^0$ , this gives an angular resolution of about  $4 \text{ mr}$ .

The angular acceptance for the entire experiment is determined by the vertex detector, which is limited by machine components. In the forward region, great care has been taken to reduce

the space between the outer surface of the B1 magnet and the region covered by active silicon to about 1 cm. This results in coverage down to about  $17^\circ$ , with a slight  $\phi$  variation due to the hexagonal structure. To obtain such tight clearance in the forward direction, the services for the beam pipe are concentrated in the backward region. In addition, the mechanical support for the silicon is somewhat larger there, to make installation and precise alignment more tractable. As a result, the acceptance extends backward to a polar angle of about  $150^\circ$ . The resulting acceptance extends from  $-0.87 < \cos \theta_{lab} < 0.96$ , which transforms into  $-0.95 < \cos \theta_{cm} < 0.87$ .

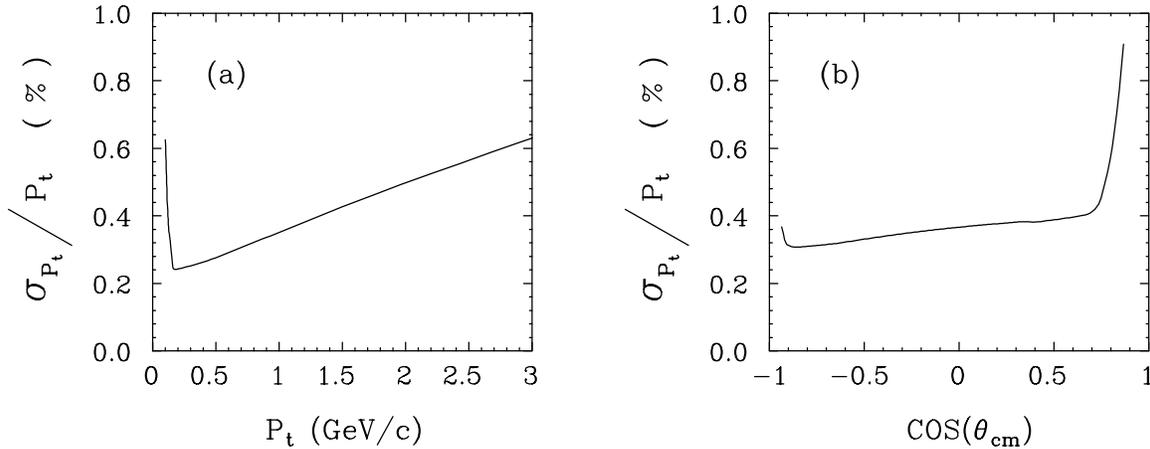
Because the vertex detector provides crucial tracking information for all charged tracks, it is important to achieve almost perfect efficiency within the acceptance of the experiment. For tracks with transverse momentum greater than about  $100 \text{ MeV}/c$ , the tracking efficiency in the drift chamber alone is very high, as it is in existing experiments. The efficiency of yielding a good vertex detector track matching a track already identified in the drift chamber is at least 96%, even under conditions of high machine background. Tracks with transverse momentum between  $40 \text{ MeV}/c$  and about  $100 \text{ MeV}/c$  are reconstructed primarily in the silicon tracker. Detailed studies of pattern recognition, including levels of machine background and silicon inefficiency higher than expected, show a track-finding efficiency of about 88% for these soft tracks.

Access to the vertex detector will be difficult. It will require removal of the support tube and the machine components in the interaction region, a process which will occur rarely under normal running conditions. For this reason, the vertex detector is being designed to be extremely reliable and robust. This includes having redundant readout paths, for example. A more important consequence of the robustness requirement is the decision to include five layers of double-sided silicon detectors. This provides some insurance that in case of steady loss of channels over a long run, or an unlikely loss of an entire readout module, the track efficiency stays high even for low momentum tracks.

## 2.5.2 Drift Chamber

To achieve the required sensitivity for the  $CP$  measurements with  $BABAR$ , it is necessary to measure all charged tracks within the acceptance of the detector with excellent precision and high efficiency. The angular acceptance is within  $17^\circ < \theta < 150^\circ$ ; the interesting tracks for  $B$  physics are in the range  $60 \text{ MeV}/c < p_t < 2.5 \text{ GeV}/c$ , while for charm and tau physics, some tracks are of higher momentum.

The main tracking chamber provides the precise measurement of  $p_t$  needed for good mass resolution on exclusive  $B$  decays. The full tracking system, consisting of vertex detector and drift chamber, provides very good pattern recognition capability for charged tracks even at 10 times the nominal machine background. The pulse height information from drift



**Figure 2-12.**  $p_t$  resolution at (a)  $\theta_{lab} = 90^\circ$  as a function of  $p_t$ , and (b)  $p_{cm} = 1.0$  GeV/c as a function of  $\cos \theta_{cm}$ .

chamber signals is used to measure the mean ionization loss ( $dE/dx$ ), which can be used to separate kaons from pions well at low momentum, and to provide some discrimination at high momentum. Finally, the drift chamber information is used to construct a tracking trigger, one of the two major triggers for the experiment, over 92% of the full solid angle in the center of mass.

The primary performance parameter for the tracking chamber is, therefore, the momentum resolution for charged particles with  $p_t$  between 100 MeV/c and 2.5 GeV/c. The value of 100 MeV/c is set by the inner radius of the chamber, 22.5 cm, and the need to get enough measurements inside the chamber for good determination of the momentum. The upper limit is the kinematic limit for particles resulting from  $B$  decays.

Figure 2-12(a) shows the  $p_t$  resolution as a function of  $p_t$  for charged particles at  $\cos \theta_{lab} = 0$ . The resolution can be parameterized by  $\sigma(p_t)/p_t \sim [0.21\% + 0.14\% \times p_t]$  from 0.2 GeV/c to the kinematic limit for  $B$  decays, 2.6 GeV/c. The constant term is the contribution of multiple scattering, and its low value is due to the use of a helium-based gas in the chamber. The values here are expected to be slightly optimistic because they assume perfect efficiency and pattern recognition within the chamber. The resolution degrades below a momentum of 180 MeV/c because the path length is shortened in the chamber, but it is still quite good down to about 100 MeV/c.

The tracking chamber covers the complete polar angle range allowed by the beam-line components,  $-0.87 < \cos \theta_{lab} < 0.96$ . The offset of the chamber center from the interaction point in the boost direction results in a  $p_t$  resolution which is fairly independent of  $\cos \theta_{cm}$  except at very forward angles, as shown in Figure 2-12(b). The resolution increases for tracks with  $\cos \theta_{cm} > 0.72$  which exit the end of the tracking chamber before they reach the

maximum radius. For tracks with  $p_t$  less than  $0.2 \text{ GeV}/c$  and at forward laboratory angles,  $\sigma_p/p$  is significantly greater than  $\sigma_{p_t}/p_t$  because the contribution from the error on the polar angle measurement becomes large. The acceptance and mass resolution for important exclusive  $B$  decays are discussed in Section 2.6.1.

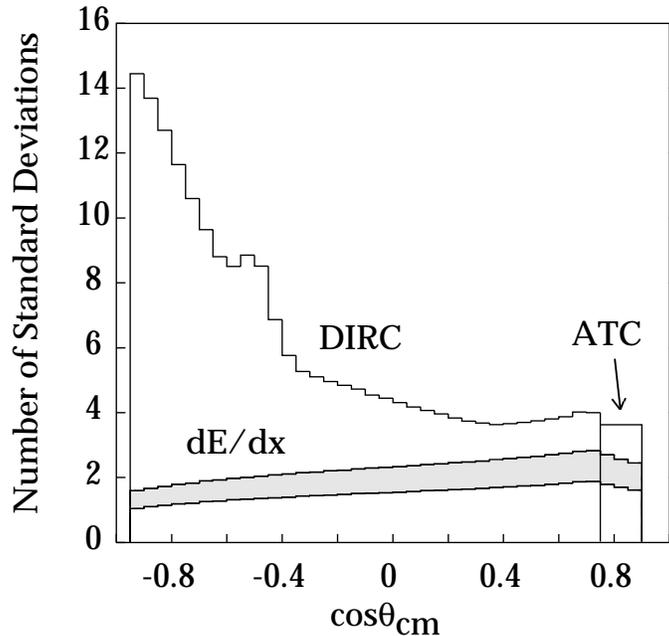
The  $dE/dx$  measurements with the designed chamber provide at least  $3\sigma$   $K/\pi$  separation at momenta up to  $0.7 \text{ GeV}/c$ , which covers a substantial fraction of the kaons used for tagging, as one can see from Figure 2-4. This includes most of the small number of tagging kaons in the backward endcap region, where no other hadron identifier is used. For pions with a momentum of  $3 \text{ GeV}/c$  from the decay  $B^0 \rightarrow \pi^+\pi^-$ , the ionization measurements may provide  $K/\pi$  separation at about the  $2\sigma$  level.

### 2.5.3 Particle Identification

There are two important benchmarks of performance for the particle identification system. One is the ability to separate  $B^0 \rightarrow \pi^+\pi^-$  from  $B^0 \rightarrow K^+\pi^-$  and the other is the effective efficiency for kaon tagging. Because the momentum range for each of these depends on polar angle, one must consider their performances separately at different angles.

Most of the pions from  $B^0 \rightarrow \pi^+\pi^-$  fall in the barrel region, in which the DIRC is providing the hadron identification. The number of photoelectrons observed by the DIRC for pions from this source is  $\sim 40$  at the backward end of the DIRC, reducing to about  $\sim 20$  at  $\cos\theta_{lab} = 0.0$ , and increasing again to  $\sim 50$  at the forward end of the DIRC. Of course, the momentum of the forward pions is greater, which makes the  $K/\pi$  separation more difficult. Figure 2-13 shows the number of standard deviations of  $K/\pi$  separation for pions from this decay mode as a function of  $\cos\theta_{cm}$ . It drops from complete separation for backward pions to a separation of about 4 standard deviations at  $\cos\theta_{cm} = 0.0$ . It is then constant for forward angles up to  $\cos\theta_{cm} = 0.7$ , which corresponds to the forward edge of the DIRC. Two estimates of  $K/\pi$  separation from  $dE/dx$  are also shown for reference: they correspond to optimistic and pessimistic estimates of performance.

For  $B^0 \rightarrow \pi^+\pi^-$  events with both pions in the acceptance of the tracking system, the fraction of pions which miss the DIRC is about 11%; 4% due to cracks between the quartz bars in azimuth, and about 7% in the forward endcap region. All of the  $B^0 \rightarrow \pi^+\pi^-$  events with a pion in the backward endcap region have their other pions striking the B1 magnet, so there is no acceptance lost by not having high momentum hadron identification there. In the forward endcap, the low-density ATC layer achieves  $K/\pi$  separation to  $4.3 \text{ GeV}/c$ . In both barrel and forward endcap regions, the information from the Cherenkov detectors is combined with  $dE/dx$  and kinematics to separate  $B^0 \rightarrow \pi^+\pi^-$ . The performance using all three techniques is discussed in Section 2.6.5.



**Figure 2-13.**  $K/\pi$  separation for pions from  $B^0 \rightarrow \pi^+\pi^-$  as a function of  $\cos\theta_{cm}$ , with full particle identification (DIRC and ATC), and with  $dE/dx$  only. The shaded region delineates two  $dE/dx$  curves. The upper curve represents the separation obtained using a modified version of a  $dE/dx$  program from Va’vra *et al.*, as described in Chapter 5. The lower curve includes a degradation of the resolution by 50% to better account for observed experiences with some large drift chambers in magnetic detectors.

Kaon tagging is important for all of the  $CP$  studies to be done with  $BABAR$ . The figure of merit is the effective tagging efficiency,  $\varepsilon(1 - 2w)^2$ , where  $\varepsilon$  is the fraction of events tagged and  $w$  is the fraction of those in which the tag is of the wrong sign. Table 2-1 shows the effect of the DIRC on the kaon tagging efficiency. With perfect hadron identification and the  $BABAR$  acceptance, assuming the loss of kaons decaying in flight, the effective tagging efficiency using charged kaons alone would be about 22%. Much of the kaon spectrum shown in Figure 2-4 is beyond the momentum region in which  $dE/dx$  information is useful. The effective tagging efficiency for events with charged kaons is 14%, if one uses optimistic  $dE/dx$  information for kaon tagging, combining it with information from lepton tagging. Adding the information from the DIRC and ATC detectors, the effective tagging efficiency rises to 21%, close to the limit of perfect kaon tagging. The main uncertainties in this number are due to the uncertainties in the kaon yield from  $B$  decays.

In the forward region, the refractive index of the ATC counters is chosen specifically to optimize the tagging efficiency. Two layers of ATC counters are used, one with  $n = 1.055$  corresponding to a kaon threshold of 1.8 GeV/c, and the other with  $n = 1.0065$  for momenta

Information Available	$\epsilon(1 - 2w)^2(\%)$
Perfect Identification at Production	27
Perfect Identification, with Decays	22
$dE/dx$ Only	14
DIRC, ATC, and $dE/dx$	21

**Table 2-1.** *The effective kaon tagging efficiency. The top line corresponds to every hadron being identified correctly at production. For the line below, every hadron is identified correctly except those which decay before the DIRC. For the last two lines, the efficiency includes the effect of using lepton tagging information in events with kaons.*

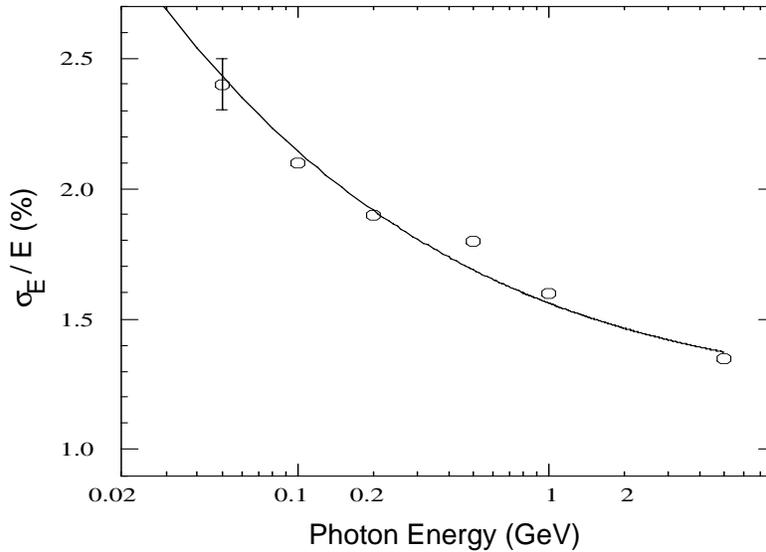
up to 4.3 GeV/c. Pions with momentum greater than about 0.7 GeV/c are detected with high efficiency, which means reasonable overlap with the  $dE/dx$  measurements in the drift chamber.

Traditionally in ring-imaging Cherenkov systems, pattern recognition or association of photoelectrons with charged tracks can be a difficult problem. It is a measure of the robustness of the DIRC that good performance is observed in full GEANT simulations with effectively no pattern recognition. Every photoelectron with proper timing is used as a candidate for each track, and all possible values of the Cherenkov angle ( $\theta_c$ ) are calculated. When all of these many values of  $\theta_c$  are plotted in a histogram for a given track, a peak at the true Cherenkov angle stands above the flat background, even in busy events. The ability to measure the mean value  $\langle\theta_c\rangle$  is reduced very little from the case with no background. If the number of photoelectrons  $N_{pe}$  were reduced due to lower-than-expected transmission or photocathode efficiency, the main effect on the measurement would be the fact that  $\delta(\langle\theta_c\rangle) \propto 1/\sqrt{N_{pe}}$ .

Another important use of the particle identification detectors is to extend the region of good lepton tagging to lower momentum. The DIRC gives  $e/\pi$  separation of at least 4 standard deviations for  $p_{lab} < 0.7$  GeV/c, giving a second contribution to electron identification in a region in which the calorimeter information is somewhat less clear. This is a useful region for tagging electrons, especially at backward angles in the center of mass. There is also some  $\mu/\pi$  separation which complements the IFR muon identification, again having the greatest effect for muons with  $\cos\theta_{cm} < 0$ .

## 2.5.4 Electromagnetic Calorimeter

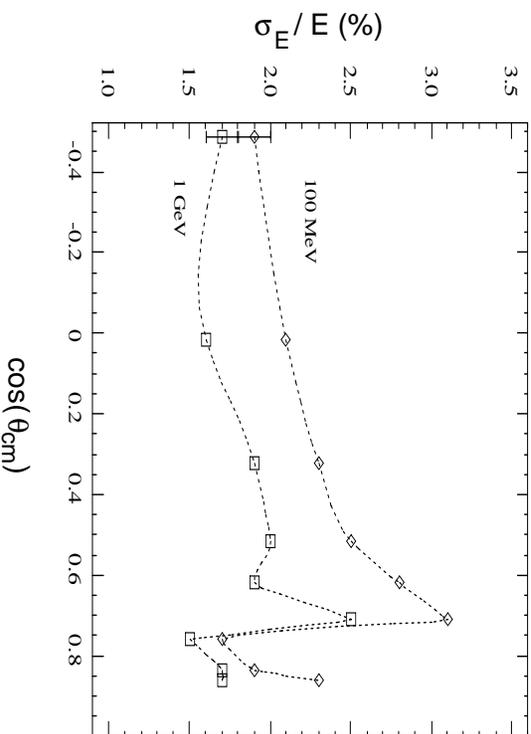
Achieving the desired sensitivity for measuring asymmetries in  $CP$  decay modes such as  $J/\psi K_S^0$ ,  $J/\psi K^{*0}$ ,  $D^{*+}D^{*-}$ , and  $\rho^\pm\pi^\mp$  requires observing  $\pi^0$ s with very high efficiency and good resolution. This is the physics goal which drives the performance requirements of the



**Figure 2-14.** Energy resolution at  $\cos\theta_{cm} = 0$  as a function of photon energy. The resolution is defined as  $FWHM/2.36$ . The error bar on the first point indicates a typical uncertainty in determining the resolution. The solid line shows the target energy resolution.

electromagnetic calorimeter. The calorimeter is also used for electron identification and to supplement the IFR information in identifying muons and  $K_L^0$ s, as well as providing information for the neutral trigger.

For decay modes such as  $B^0 \rightarrow \rho^\pm \pi^\mp$ ,  $\rho^\pm \rightarrow \pi^\pm \pi^0$ , the  $\pi^0$  momentum ranges up to about 2.5 GeV/c in the center of mass, leading to the energy spectrum for photons in the laboratory shown in Figure 2-3. From this distribution and others for additional modes, one sees that the photon energy region of interest for  $B$  physics is from about 20 MeV to about 5 GeV in the laboratory. The  $B$  mass resolution for this and other modes using  $\pi^0$ s is dominated by the photon energy resolution. Figure 2-14 shows the energy resolution as a function of photon energy at  $\cos\theta_{cm} = 0.0$ . The energy resolution ( $FWHM/2.36$ ) is about 2.1% at 100 MeV and about 1.6% at 1 GeV. Included in these numbers is the effect of material before the calorimeter, which at a photon energy of 100 MeV degrades the resolution by 10–15% and reduces the efficiency by 12–22%. Figure 2-16 shows the amount of material between the IP and the various detector subsystems, up to the calorimeter. The amount of that material is about  $0.23X_0/\sin\theta$  for the barrel, of which 75% is in the DIRC, and about  $0.35X_0$  for the forward endcap. The resolution meets the target energy resolution for photons at this angle,  $\sigma_E/E = 1\%/E(\text{GeV})^{\frac{1}{4}} \oplus 1.2\%$ . The constant term arises from leakage, inter-calibration errors, and nonuniformity of light collection.

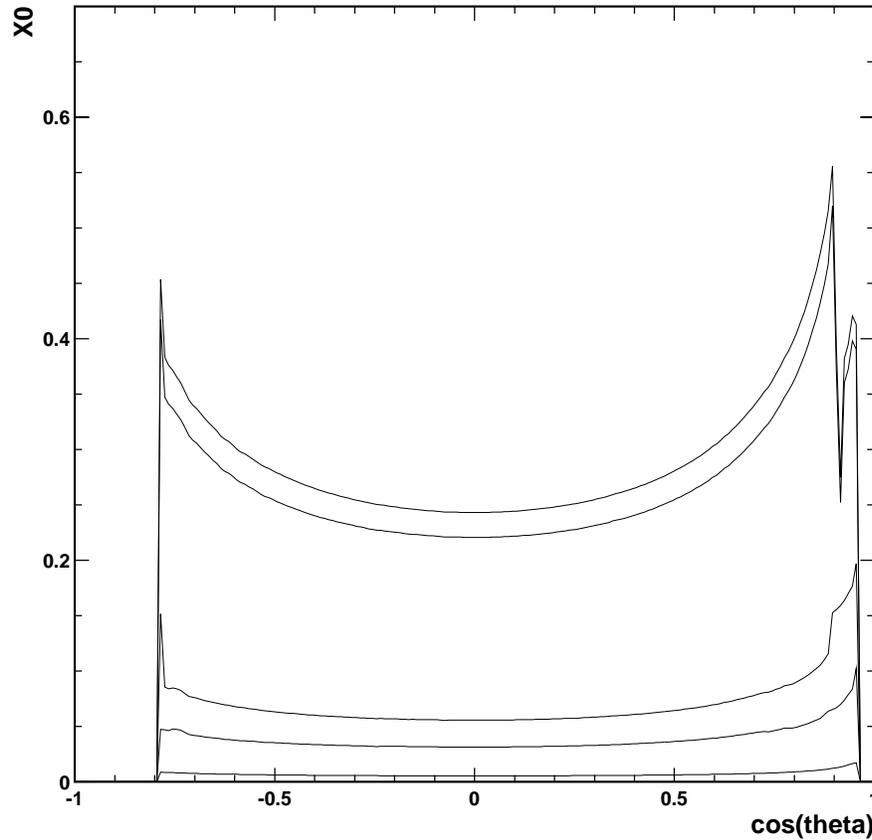


**Figure 2-15.** Energy resolution as a function of  $\cos\theta_{cm}$  for  $E_{\gamma lab} = 0.1$  GeV and 1.0 GeV. The error bars on the first points indicate typical uncertainty in determining the resolution.

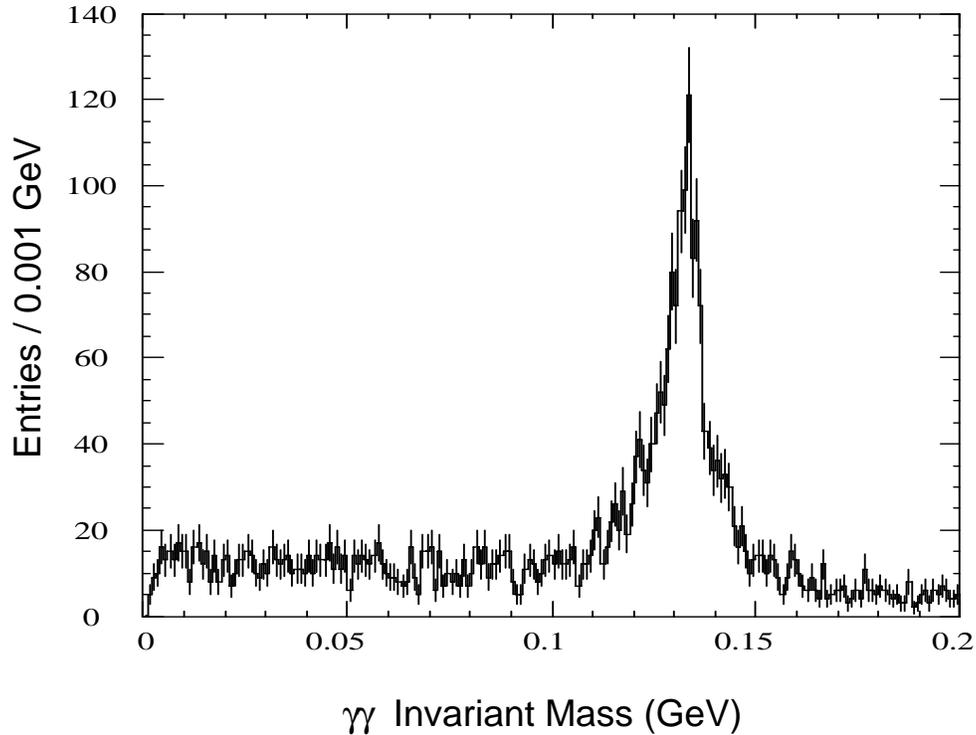
Figure 2-15 shows the energy resolution as a function of  $\cos\theta_{cm}$  for photons with energies of 0.1 and 1.0 GeV in the laboratory. The resolution at  $\cos\theta_{cm} = 0.5$  is 2.5%(2.0%) for photons of energy 100 MeV(1 GeV). The resolution at angles more forward than that increases because of leakage of the shower from the uncovered sides of crystals in the staggered arrangement. The left edge of the plot is  $\cos\theta_{cm} < -0.5$ , which corresponds to  $90^\circ$  in the laboratory; the resolution remains approximately flat at more backward angles.

Another important performance parameter is the efficiency for  $\pi^0$ s; however, this depends subtly on details of the data treatment in specific analyses. This efficiency is determined by a cut on the  $\gamma\gamma$  invariant mass, the value of which depends on the background level for a particular physics process. Figure 2-17 shows the  $\gamma\gamma$  mass spectrum for  $\pi^0$ s from the process  $B^0 \rightarrow \rho^\pm\pi^\mp$ . The resolution function is a Gaussian with  $\sigma \sim 5$  MeV, plus an additional tail at low  $m(\gamma\gamma)$ . The Gaussian part of the  $m(\gamma\gamma)$  resolution function is dominated by the angular resolution of the calorimeter for  $\pi^0$ s with momentum greater than about 1 GeV/c. Material in front of the calorimeter is responsible for the low mass tail. In Section 2.6.3, the  $\pi^0$  efficiency from full physics simulations of a couple of physics processes is discussed.

The angular resolution for photons in the barrel region is given by the approximate formula  $\sigma_\theta = 3\text{mr}/\sqrt{E(\text{GeV})} \oplus 2\text{mr}$ , which is consistent with scaling from experience with the CLEO-II calorimeter. The angular resolution is determined by the transverse crystal size and the average distance to the interaction point.



**Figure 2-16.** Amount of material between the IP and the various detector subsystems, measured in radiation lengths, as a function of  $\cos \theta_{lab}$ . The five curves, from bottom to top, give the material up to and including the beam pipe, the vertex detector, the drift chamber, the particle-identification system, and the aluminum shield in front of the calorimeter. The aerogel geometry used in making this figure was in an early stage of development and does not correspond to that described in Chapter 6.

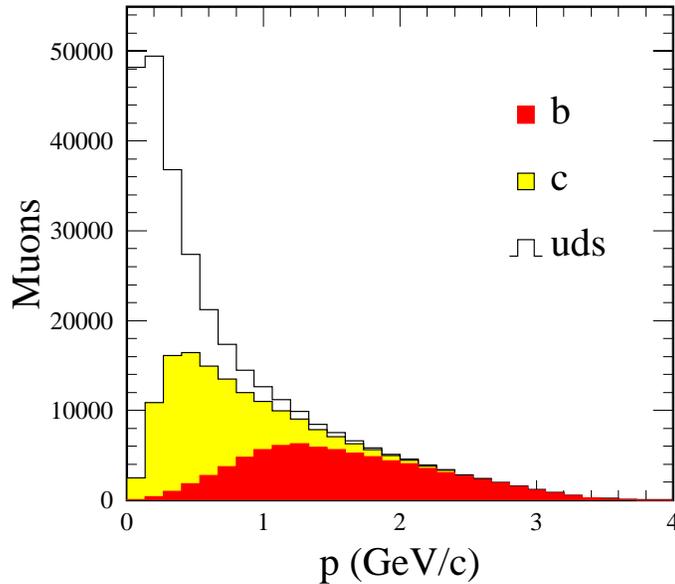


**Figure 2-17.**  $\gamma\gamma$  invariant mass spectrum from the process  $B^0 \rightarrow \rho^\pm \pi^\mp$ .

The barrel calorimeter covers the range  $-0.80 < \cos \theta_{lab} < 0.89$ ; the forward endcap calorimeter extends the forward coverage to  $\cos \theta_{lab} = 0.97$ . Showers at least two counters in from the edge are contained well enough for good measurement. In the center of mass, this corresponds to containment of any shower in the range  $-0.92 < \cos \theta_{cm} < 0.87$ .

Another parameter that determines the efficiency of the calorimeter is the minimum detectable energy for photon showers. The incoherent electronic noise per crystal is expected to be about 150 keV, which adds up to 750 keV for a typical cluster of 25 crystals. This is low enough that the minimum detectable energy should be determined primarily by backgrounds from the event and from beam interactions, and is expected to be 10–20 MeV.

The final measure of calorimeter performance is the ability to identify electrons. Simulations and experience with CLEO-II indicate a very high efficiency for identifying electrons down to a momentum of 500 MeV/ $c$ . The probability for a pion to fake an electron is about  $1 \times 10^{-3}$  at high momentum.



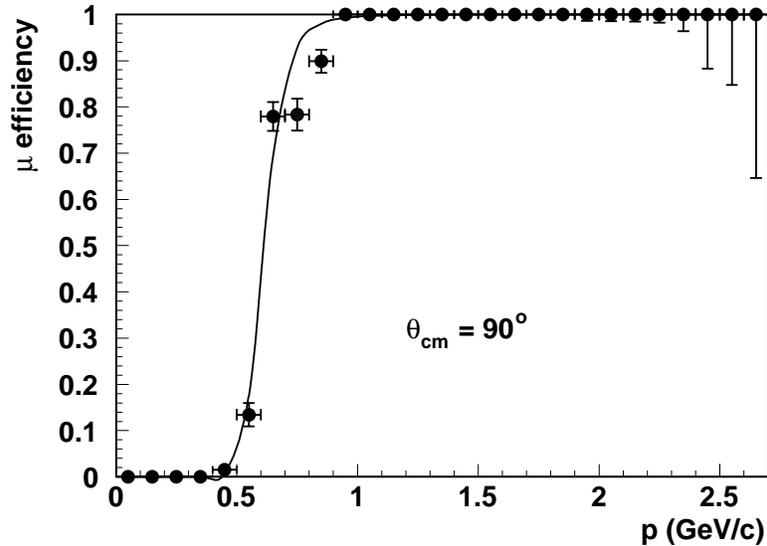
**Figure 2-18.** Momentum spectrum of muons in the laboratory. The three sources are primary, secondary, and continuum muons.

### 2.5.5 Muon and Neutral Hadron Detector

The primary experimental goal of the Instrumented Flux Return (IFR) is to reduce the lower momentum limit for cleanly identifying muons from the value of  $1.4 \text{ GeV}/c \times \csc \theta_{lab}$  that can be obtained with an unsegmented iron absorber to about  $0.6 \text{ GeV}/c \times \csc \theta_{lab}$ . This increases the efficiency for tagging the flavor of a  $B$  meson substantially, and it also increases the size of the lepton sample for studies of semileptonic decays. Figure 2-18 shows the momentum spectrum of muons, both primary and secondary.

Figure 2-19 shows the efficiency for identifying muons produced at  $\cos \theta_{cm} = 0$  as a function of laboratory momentum. The efficiency is about 50% for muons of 650 MeV/c momentum and rises to an efficiency of 90% for muons of 800 MeV/c momentum. The fraction of pions decaying before the IFR at a momentum of  $\sim 1 \text{ GeV}/c$  is about 4%, but by comparing measured momentum with range in the iron, one can reduce this background substantially. In the region below about 600 MeV/c, there is supplemental information from the DIRC which helps identify muons. The efficiency cutoff as a function of  $p_{lab}$  is shown in Figure 2-19 and remains approximately the same as the angle varies.

The thickness of the iron layers is graded from 2 cm on the inside to 5 cm on the outside. The finer segmentation of the innermost layers is chosen to identify muons of the lowest possible momentum. Below 1.0 GeV/c in the center of mass, the muons are predominantly secondary,



**Figure 2-19.** Efficiency for identifying muons with  $\cos\theta_{cm} = 0$  as a function of  $p_{lab}$ .

*i.e.*, from semileptonic decays of the charm product of a  $B$  decay. In events in which one  $B$  decay is completely reconstructed, it is possible to use information about the isolation of the lepton to help separate primary from secondary muons.

Although the IFR has been designed and optimized with muons as the primary goal, the importance of detecting the  $K_L^0$  from  $B^0 \rightarrow J/\psi K_L^0$  has been kept in mind. The goal is not to obtain an accurate measurement of the  $K_L^0$  energy, but simply to identify the  $K_L^0$  and to measure its angle reasonably well. For one set of cuts, the efficiency for detecting the  $K_L^0$  from  $B^0 \rightarrow J/\psi K_L^0$  is 19%, with a signal/background ratio of 3.5. The most important background process is  $B^0 \rightarrow J/\psi K^*$ ,  $K^* \rightarrow K_L^0 \pi$  in which the  $K^*$  cannot be vetoed.

A third use of the IFR is to detect whether there is any hadronic energy which has escaped the electromagnetic calorimeter. This is useful, for example, in the study of exclusive  $b \rightarrow u$  semileptonic decays, for which a serious background comes from continuum charm events which contain leptons. In suppressing these backgrounds, it is helpful to know if there is substantial missing energy, and if so, in what direction it lies. Although this has not been studied specifically for  $BABAR$ , it is known to be helpful in such studies now done at CLEO.

### 2.5.6 Electronics, Trigger, and Data Acquisition

The primary performance measure of a trigger system is its efficiency for benchmark physics processes. The charged track trigger requires at least two tracks in the drift chamber, one with  $p_t > 0.21 \text{ GeV}/c$  and another with  $p_t > 0.13 \text{ GeV}/c$ . The neutral trigger requires two showers in the electromagnetic calorimeter, both with reconstructed energy deposits above a threshold that is efficient for muons. The orthogonality of the requirements allows good cross calibration of trigger efficiency. With no further restrictions, the total trigger rate is simulated to be about 8 kHz at 10 times nominal background, which is greater than the specification of 2 kHz. The rate can be reduced further rather simply by introducing a  $p_t$  cut on the track with the largest transverse momentum and an energy cut on the largest energy deposit.

The efficiency for events such as  $B^0 \rightarrow \pi^+\pi^-$ ,  $\bar{B}^0 \rightarrow \mu X$  is simulated to be close to 100% for any reasonable variation of the proposed trigger. There are so many charged tracks and photons with high momentum in such events that the efficiency is very robust.

Tau events with 1+1 prongs and low-mass two-photon production are more demanding of the trigger, and therefore drive the detailed trigger logic. Efficiencies larger than about 70% are expected for events with  $\tau^+ \rightarrow e^+\nu\bar{\nu}$ ,  $\tau^- \rightarrow \mu^-\bar{\nu}\nu$ . The crucial performance parameter for such events is not the efficiency itself, but the systematic error in determining the efficiency. Since the efficiencies for  $\tau$  events are dominated by the angular acceptance of the experiment, they should be measurable with great precision.

Reasonable thresholds for the trigger described above are  $p_t > 0.6 \text{ GeV}/c$  for at least one track in the charged particle trigger and  $E > 0.6 \text{ GeV}$  for at least one cluster in the neutral trigger. With these thresholds, the rate is about 400 Hz at nominal backgrounds and 1400 Hz at 10 times nominal. Thus, the proposed trigger architecture with Level 1 and Level 3 meets the design requirements and has very high efficiency. The optional Level 2 ensures the robustness of the design. When more information is available on such parameters as the Level 3 latency and event-building rate, it will be known whether the Level 2 trigger is needed.

## 2.6 Physics Performance

---

The previous section contains the results of a series of simulation studies of particular detector systems, often based on simulations of a single track or shower. In this section, we complete the picture by discussing aspects of the detector performance which require full physics simulations to study properly, or require combining information from multiple detector systems. The physics modes studied are those benchmarks described in Chapter 3.

Detector	Technology	Dimensions	Performance
SVT	Double-sided Silicon Strip	5 Layers $r = 3.2 - 14.4$ cm $-0.87 < \cos \theta < 0.96$	$\sigma_z = \sigma_{xy} = 50 \mu\text{m}/p_t \oplus 15 \mu\text{m}$ $\sigma_\phi = \sigma_\theta = 1.6 \text{ mrad}/p_t$
DC	Small Cell Drift Chamber	40 Layers $r = 22.5 - 80.0$ cm $-111 < z < 166$ cm	$\sigma(p_t)/p_t = [0.21\% + 0.14\% \times p_t]$
PID	DIRC	$1.75 \times 3.5 \text{ cm}^2$ quartz $-0.84 < \cos \theta < 0.90$	$N_{pe} = 20 - 50$ $\gtrsim 4\sigma K/\pi$ separation for all $B$ decay products
PID	ATC	$n=1.0065, 1.055$ $0.916 < \cos \theta < 0.955$	$N_{pe} = 10$ $\pi/K$ separation up to $4.3 \text{ GeV}/c$
CAL	CsI(Tl)	$16 - 17.5 X_0$ $\sim 4.8 \times 4.8$ cm crystals	$\sigma_E/E = 1\%/E(\text{GeV})^{\frac{1}{4}} \oplus 1.2\%$ $\sigma_\theta = 3 \text{ mrad}/\sqrt{E(\text{GeV})} \oplus 2 \text{ mrad}$
MAG	Superconducting Segmented Iron	IR=1.40 m L = 3.85 m	$B = 1.5 \text{ T}$
IFR	RPC	16–17 Layers	$\epsilon_\mu > 90\%$ for $p_\mu > 0.8 \text{ GeV}/c$

**Table 2-2.** Parameter summary (all angles are in the laboratory).

### 2.6.1 Acceptance and Mass Resolution for Decays to Charged Particles

The best measures of the overall acceptance of  $B_{\text{BABAR}}$  are the efficiencies for multiparticle  $CP$  final states. The efficiency found for the process  $B^0 \rightarrow J/\psi K_s^0$ ,  $J/\psi \rightarrow \ell^+ \ell^-$ ,  $K_s^0 \rightarrow \pi^+ \pi^-$  is 57%, which corresponds to about 87% per track, including geometric acceptance. The polar angle acceptance set at the vertex detector by the B1 accelerator magnets is about 91%, so this represents rather little additional loss for lepton identification,  $K_s^0$  finding, and so forth.

Another benchmark channel is  $B^0 \rightarrow D^{*+} D^{*-} \rightarrow \pi^+(K^-\pi^+)\pi^-(K^+\pi^-)$ , which emphasizes acceptance for charged particles with very low momentum. The efficiency for measuring each slow bachelor pion is about 87%, including the geometric acceptance of 91%. The ability to track slow pions in the vertex detector is crucial in maintaining this high an efficiency. The total efficiency for reconstructing a  $B^0$  in this mode is about 40%.

Good mass resolution is needed to reduce the combinatoric backgrounds, and for all-charged decay modes, this is dominated by the momentum resolution. The primary physics benchmark for momentum resolution is the mass resolution for the mode  $B^0 \rightarrow \pi^+ \pi^-$ , in which the

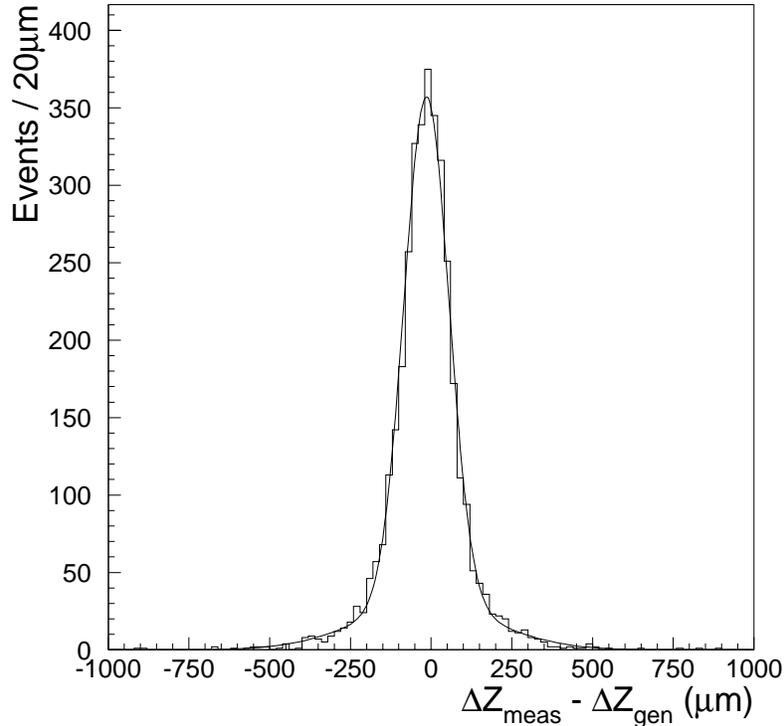
pions are at the kinematic limit. It has the added feature of a large background peak from  $B^0 \rightarrow K^+\pi^-$  separated by only about  $43 \text{ MeV}/c^2$  in mass, making the mass resolution even more critical than for other modes. Simulation shows a mass resolution of about  $21 \text{ MeV}/c^2$  where this value does not yet take into account the effect of missed hits and fake hits in the drift chamber. This resolution is determined by the tracking length in the drift chamber ( $80 - 22 = 58 \text{ cm}$ ), the magnetic field, the intrinsic resolution of each measurement, and the number of radiation lengths within the chamber volume.

The key to reducing backgrounds for  $B$  decay modes such as  $B^0 \rightarrow D^{*+}D^{*-}$  is good resolution on the mass difference  $\delta m = m(D\pi) - m(D)$ . Because the pion is slow in the  $D^*$  rest frame, the resolution on this mass difference is typically dominated by the angular resolution of the slow pion. Simulation shows a resolution of about  $\pm 0.4 \text{ MeV}$  in  $\delta m$ , which compares very well with existing experiments. This is achievable by using the production point and a precise vertex detector measurement near the beam pipe, which reduces the effect of the multiple scattering on the angle measurement.

### 2.6.2 Separation between $B$ Decay Vertices

The most important role of the vertex detector is to determine the separation between the two  $B$  decay vertices along the  $z$  axis, which is needed to measure the  $CP$  asymmetries. For  $CP$  studies in modes with little background, this does not place a very stringent requirement on intrinsic position resolution. The average separation of the two  $B$  vertices is  $250 \mu\text{m}$ , and the asymmetry peaks at  $\Delta z \approx 550 \mu\text{m}$ . The degradation in the measurement of the  $CP$  asymmetries due to imperfect vertex resolution is less than 10% if the separation is measured with an error of  $125 \mu\text{m}$ . Figure 2-20 shows the resolution in  $\Delta z \equiv z_{CP} - z_{tag}$  for the case of  $B^0 \rightarrow \pi^+\pi^-$  with a primary lepton tag. The expected distribution for  $\Delta z$  can best be characterized as a narrow Gaussian of about  $70 \mu\text{m}$  which contains 80% of the events and a wider gaussian of about  $220 \mu\text{m}$  which contains the rest. The resolution in  $\Delta z$  for this and other modes is determined primarily by the error on the  $z$  position of the tagging vertex.

For measurements of  $CP$  decay modes with small backgrounds, the vertex resolution is sufficiently good that there is little degradation of the asymmetry measurement compared to the case of perfect resolution. For other modes, especially the modes used to measure  $\sin 2\alpha$ , the vertex separation of the two  $B$  decays has an additional importance, since it is used to suppress the dominant backgrounds. The  $\Delta z$  distribution of the background is a gaussian centered at zero with  $\sigma \sim 100 \mu\text{m}$ , where the width is determined by the resolution. For all of these modes, the effective background level at  $\Delta z \approx 250 \mu\text{m}$  is very sensitive to the resolution  $\sigma(\Delta z)$ , and the  $CP$  measurement improves as this resolution decreases. The precision achievable is limited by multiple scattering, which sets a natural scale for the point resolution in the inner layers.



**Figure 2-20.** Resolution in  $\Delta z \equiv z_{CP} - z_{tag}$  for the case of  $B^0 \rightarrow \pi^+\pi^-$  with a primary lepton tag.

### 2.6.3 $\pi^0$ Efficiency and Resolution

Full physics simulations were done to measure the efficiency for finding  $\pi^0$ s embedded in full events. For the process  $B^0 \rightarrow J/\psi K_s^0$ ,  $J/\psi \rightarrow \ell^+\ell^-$ ,  $K_s^0 \rightarrow \pi^0\pi^0$ , the efficiency is 25%, or about 0.44 times the efficiency obtained for the same process with  $K_s^0 \rightarrow \pi^+\pi^-$ . This corresponds to an efficiency for each  $\pi^0$  of about 60% each, including geometric acceptance. Since the typical  $\pi^0$  momentum is slightly less than 1 GeV/c, this result demonstrates the importance of high detection efficiency for photons of modest energy.

The material in front of the barrel calorimeter is dominated by the DIRC quartz bar and associated mechanical supports. The low-mass tail shown in Figure 2-17 is mostly due to pairs produced in the material before the calorimeter, which are not included in the shower energy. The net effect of all such material is to reduce the efficiency for  $\pi^0$ s in the  $\rho^\pm\pi^\mp$  mode by about 15% and to degrade the mass resolution by 10%.

For most  $B$  decay modes with at least one neutral pion, the energy resolution of the calorimeter dominates the mass resolution. The benchmark decay analogous to  $B^0 \rightarrow \pi^+\pi^-$

is  $B^0 \rightarrow \pi^0\pi^0$ ; not only is it an important physics mode, but the resolution is closely related to that for such modes as  $B^+ \rightarrow \pi^+\pi^0$ . The energy resolution averaged over the entire detector leads to a mass resolution for  $B^0 \rightarrow \pi^0\pi^0$  of about 64 MeV (FWHM/2.36). The resolution function is, of course, asymmetric with a low-mass tail. The background from the continuum is larger than the signal for reasonable estimates of the branching ratio, so the success of this analysis depends critically on maintaining the quoted energy resolution.

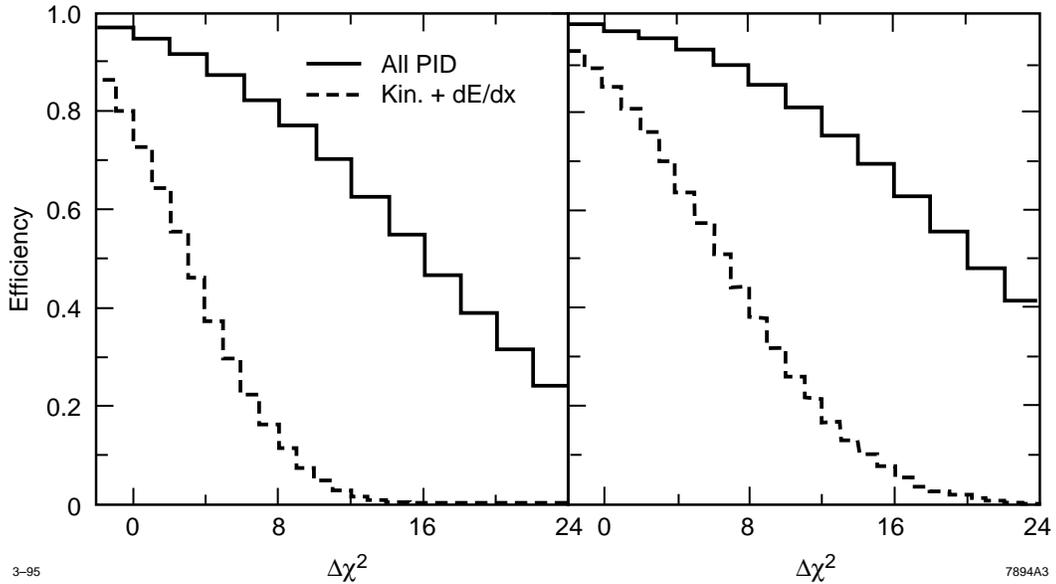
### 2.6.4 Lepton Identification

The identification of high-energy leptons from  $B^0$  decays such as those coming from  $B^0 \rightarrow J/\psi K_s^0$  is straightforward, and the efficiency is very high. Tagging leptons are more challenging, however, since the spectrum of useful leptons extends to rather low momentum, as seen in Figure 2-18. A full study shows an effective tag efficiency for primary leptons of 13.2%, assuming a semileptonic branching fraction of 10.5%. Besides the branching fraction, the factors which determine the tagging efficiency are the primary and secondary lepton spectra, the polar angle acceptance, the efficiency at low momentum, and the probability of pions decaying. Good identification efficiency is obtained for low momentum electrons using the calorimeter and the DIRC or ATC counters; for low momentum muons with the IFR and the DIRC. This contributes to additional tagging efficiency using secondary leptons.

### 2.6.5 Charged Hadron Identification

The bulk of charged hadron identification is done with the particle identification detectors, the DIRC and ATC counters. In Figure 2-13, we showed the  $\pi/K$  separation obtained for the important decay mode  $B^0 \rightarrow \pi^+\pi^-$  using the DIRC and  $dE/dx$  information. To gain full separation from the  $K^+\pi^-$  decay, both of these and kinematics will be used.

Figure 2-21 shows the efficiency for accepting a  $B^0 \rightarrow \pi^+\pi^-$  event as a function of  $\Delta\chi^2$ , the difference in  $\chi^2$  between a  $\pi^+\pi^-$  assignment and a  $K^+\pi^-$  assignment, for two different assumptions about drift chamber performance. The  $\chi^2$  represents the consistency with a given hypothesis using either kinematics and  $dE/dx$  alone or including the particle identification detectors. A reasonable measure of good suppression of the competing mode is  $\Delta\chi^2 > 9$ . Using optimistic calculations of drift chamber performance (right-hand plot), and requiring that  $\Delta\chi^2 > 9$  from  $dE/dx$  and kinematics alone, one gets an efficiency of about 32%. On the other hand, adding the information from the DIRC and ATC counters, the corresponding efficiency is about 86%. The left-hand plot shows the same information for an estimate of resolution in momentum and  $dE/dx$  which is based on operating experience with CLEO. In this case, the efficiencies for  $\Delta\chi^2 > 9$  are about 7% and 75%, depending on whether the PID



**Figure 2-21.** The efficiency for accepting a  $B^0 \rightarrow \pi^+\pi^-$  event as a function of  $\Delta\chi^2 = \chi^2(\pi^+\pi^-) - \chi^2(K^+\pi^-)$ , using pessimistic (left) and optimistic (right) estimates of drift chamber performance. In calculating the  $\chi^2$  for the two hypotheses, kinematics and  $dE/dx$  are used for the dashed histograms, while for the solid histograms all information is used.

detectors are being used. The DIRC and ATC are even more important if the drift chamber performance does not quite meet its design goals.

The other important analysis measuring the  $CP$  asymmetry parameter  $\sin 2\alpha$ , that of the modes  $B^0 \rightarrow \rho^\pm\pi^\mp$ , is somewhat less demanding in hadron identification. The penguin decay  $B^0 \rightarrow K^+\rho^-$  is expected to have a small branching ratio, even compared to  $B^0 \rightarrow \rho^\pm\pi^\mp$ . The branching ratio for  $B^0 \rightarrow K^{*+}\pi^-$  is not that small, but the separation, both by kinematics and by identification of the charged particles, is somewhat easier, due to the lower average particle momenta. The continuum background, which is relatively more important for this mode, is also significantly suppressed by good  $K/\pi$  separation.

## 2.7 Performance for Non- $CP$ Physics

While operating at the  $\Upsilon(4S)$ , the  $BABAR$  experiment will be used for a broad program of  $B$  physics, charm physics, tau physics, and two-photon physics. The high luminosity of PEP-II ensures that, in all of these areas, it will be possible to do physics well beyond present experiments. Although the success of the  $CP$  program is the primary goal in the design of

*BABAR*, the needs of the other parts of the physics program have been kept in mind. In most cases, the performance of the detector can be estimated reliably by modest extrapolation from the sensitivity of present CLEO-II analyses.

Fortunately, the main detector requirements for all of these physics goals are the same as for the *CP* program. There are only a couple cases in which additional constraints were imposed by these other physics areas. Most notable of these is in the trigger design, which is more demanding for tau and two-photon physics than it is for bottom or charm physics. Most importantly, none of the other physics is seriously compromised by the choices we have made for *B* physics.

### 2.7.1 Other *B* Physics

The *BABAR* detector is optimally designed for all *B* physics, not just the measurement of *CP* asymmetries. The detector requirements for other *B* physics are included in the requirements for the *CP* physics. The events have the same characteristics, and most of the interesting *B* physics involves complete reconstruction of either one or both *B* decays.

An example is the decay  $B^+ \rightarrow \rho^0 \ell^+ \nu$ , which is a prime candidate for measurement of  $|V_{ub}|$ . The basic detector requirements are good lepton identification and good efficiency for detecting all particles in the event, which are also important for all *CP* studies. The dominant backgrounds in the existing study of this mode are from  $c\bar{c}$  events in which a fast lepton from one charm particle is combined with two pions from the charm particle in the opposite hemisphere. At PEP-II, these two charm vertices are separated by  $\sim 600 \mu\text{m}$  transversely and  $\sim 600 \mu\text{m}$  longitudinally, so it should be rather straightforward to separate them from the signal events in which all three tracks come from the same vertex. In addition, one uses the missing momentum in the event to “reconstruct” the neutrino. Missing neutral hadrons dominate the resolution for this reconstruction, so the ability to detect such hadrons should improve the information and help to reduce backgrounds.

A more challenging task is to measure the decay  $B^+ \rightarrow \tau^+ \nu_\tau$ , which is the best way to measure  $f_B$ . The analysis requires completely reconstructing a large sample of  $B^-$  decays, and then looking to see if the remaining particles are consistent with  $B^+ \rightarrow \tau^+ \nu_\tau$ ,  $\tau^+ \rightarrow \ell^+ \nu_\ell \bar{\nu}_\tau$ . An obvious background for this decay is a semileptonic decay in which the hadronic particles are not detected. Again, a very high efficiency for charged particles, photons, and neutral hadrons is required to obtain the optimal suppression of backgrounds. The largest gap in the detector acceptance is the 300 mr cone in the forward direction. Detailed studies are needed to determine whether that loss of acceptance is an important source of background, and if so, whether some form of veto detector could reduce the loss. (See the discussion below concerning the similar decay  $D_s^+ \rightarrow \mu^+ \bar{\nu}_\mu$ .)

### 2.7.2 Detector Issues for Charm Physics

The detector requirements for doing charm physics are almost identical to those for  $B$  physics. The clearest indication of this is the success that CLEO-II has had in charm physics. The main differences at PEP-II will be the increased luminosity and the boost along the beam direction. As shown in the previous section, the acceptance of  $BABAR$  is very good, and the performance is fairly uniform even when studied as a function of center-of-mass polar angle.

The  $BABAR$  vertex detector (and the one soon to be installed in CLEO) represents a major improvement for charm physics. For example, the study of the important semileptonic decay  $D^+ \rightarrow \bar{K}^{*0} \ell^+ \bar{\nu}$  is presently limited in CLEO because of combinatoric backgrounds. Because of the long  $D^+$  lifetime, the  $D^+$  travels an average of  $\sim 700 \mu\text{m}$  at PEP-II, which will make it possible to remove most combinations which include particles from the production vertex. The vertex detector will be essential in separating doubly Cabibbo-suppressed decays from mixing in the analysis of such decays as  $D^0 \rightarrow K^+ \pi^-$ ,  $K^+ \rho^-$ . The proper time distribution for the mixed events is proportional to  $t^2 e^{-t/\tau}$ , which makes it possible to fit the time distribution and extract the mixing component.

One of the charm physics topics sure to be still interesting at the time that  $BABAR$  starts taking data is the precise measurement of the  $D_s^+$  decay constant in the decay mode  $D_s^+ \rightarrow \mu^+ \bar{\nu}_\mu$ . This decay is barely seen above background in present CLEO data. With greater integrated luminosity and vertex information to help reduce backgrounds, a much better measurement should be possible. One background that is difficult to remove comes from the semileptonic decay  $D^+ \rightarrow K_L^0 \mu^+ \nu_\mu$ , in which the  $K_L^0$  is undetected. The IFR should be useful in reducing such a background.

### 2.7.3 Detector Issues for Tau Physics

The luminosity of PEP-II will make it possible to reach unprecedented sensitivity for rare tau decay modes, especially neutrinoless modes, in which kinematic constraints keep backgrounds low. In addition, the enhanced capabilities of  $BABAR$  relative to CLEO-II (and ARGUS) will make it possible to perform precision measurements of tau decay properties in common decay modes. The design of the experiment, although optimized for  $B$  physics, is very well-suited to the important tau physics. Most of the general design considerations for tau physics are the same as those for  $B$  physics: high efficiency for charged particles and photons of fairly low momentum, lepton and hadron identification, good geometric acceptance, *etc.* It is primarily in the trigger that the tau physics is more demanding, and the trigger is therefore designed with that in mind.

In contrast to the case for  $b\bar{b}$  and  $q\bar{q}$  events, which typically contain many tracks, tau pairs have low multiplicity, and thus an efficient trigger becomes a challenge. High efficiency is not the problem for these measurements, however; precise knowledge of the efficiency is. Because of the redundant, orthogonal, low-multiplicity triggers used in *BABAR*, these efficiencies should be high and precisely measured. It is therefore unlikely that the trigger efficiency will limit the systematic error on measurements of common tau decays.

The typical separation between the two tau decays is  $\sim 400 \mu\text{m}$  transverse to the  $z$  axis and  $\sim 300 \mu\text{m}$  along the  $z$  axis. The precise vertex resolution of *BABAR* will make it possible to exploit this separation on an event-by-event basis to separate multiprong tau pairs from  $q\bar{q}$  backgrounds.

The tau neutrino mass measurement is an analysis which is particularly sensitive to backgrounds which fake multiprong tau decays. To suppress these backgrounds, one usually tags the event with a lepton from the other tau. In that case, the major background is from  $c\bar{c}$  events, where, *e.g.*, one  $D$  decays to  $K_L^0 \ell^+ \nu_\ell$ , and there are no fragmentation particles. Because charmed particles are also long-lived, flight paths are not so useful to suppress this background. The IFR permits the detection of  $K_L^0$ s, however, making this background easier to veto.

The excellent lepton and hadron identification of *BABAR* is also of great help for tau physics. The IFR provides efficient  $\mu/\pi$  separation at momenta down to  $\sim 0.5 \text{ GeV}/c$ . Improved discrimination between the  $\mu\nu\nu$  and  $\pi\nu$  decay modes of the tau is of great help in many important tau analyses. Efficient  $\pi/K$  separation over almost all of the kinematic range is important for studies of the  $K\pi\pi$  and  $KK\pi$  systems.

In making precise tau branching ratio measurements, typically both tau decays are fully reconstructed in each event. The error obtainable depends on the ability to model the efficiency of the detector for full reconstruction of many particles. This in turn relies on the ability of *BABAR* to detect, with very high efficiency, all tracks and photons within the geometric acceptance, especially at low momentum. The emphasis placed on high efficiency for low energy photons and low momentum pions for  $B$  physics serves to optimize the detector for this tau physics as well.

To suppress backgrounds from tau decay modes containing an extra  $\pi^0$  or  $K_L^0$ , one typically rejects events containing extra showers. Because of the soft energy spectrum for photons from tau decays, it is important to detect photons cleanly with energies as low as 50 MeV. Thus, it is important to remove sources of spurious low-energy showers, including extra showers from hadronic interactions in the calorimeter, which are the hardest to model. The IFR, operating as a “tail-catcher,” may be useful to help in identifying such showers. In addition, the transverse segmentation of the calorimeter crystals can help a great deal in distinguishing such showers from true photons.

## 2.7.4 Detector Issues for Two-Photon Physics

The high luminosity of PEP-II will extend the study of exclusive two-photon physics (*i.e.*, the reaction  $e^+e^- \rightarrow e^+e^-\gamma^*\gamma^* \rightarrow e^+e^-X$ ) from the present 2 GeV up to at least 5 GeV in mass. In addition to the gains caused by the much higher luminosity, the *BABAR* detector will improve on the existing, very successful CLEO-II design in several ways, including superior calorimetry at low energies, better particle identification at high momenta, precise vertex measurement, and a very flexible data acquisition and triggering system. This will greatly improve both resonance searches and detailed QCD tests. Most of this two-photon physics can be done with a detector optimized for the *CP*-violation physics goals. However, some of the interesting  $\gamma\gamma$  processes have significant impacts on the trigger and on the backward endcap coverage.

Several two-photon reactions result in final states with only two charged hadrons, sometimes with rather small transverse momenta. The challenge for the trigger is to accept these events without introducing a large rate from beam backgrounds or cosmic rays. Redundant triggers using either two calorimeter clusters or two drift chamber tracks (or combinations of these) with variable thresholds should suffice for this purpose. In addition, there are interesting all-neutral final states (*e.g.*,  $\pi^0\pi^0$ ) that can also be accepted using the calorimeter trigger. High efficiency can be maintained for final state masses greater than 2 GeV/ $c^2$ , while modest efficiency at lower mass preserves the ability to overlap with present two-photon measurements.

There are two single-tagged  $\gamma\gamma$  reactions which are quite sensitive to the minimum backward detection angle. The first of these, the production of spin-1 resonances, gives a prime tool for identifying hybrid mesons ( $1^{-+}$ ) and studying charmonium  $\chi$  ( $1^{++}$ ) states, since one can uniquely distinguish between spin-1 and even-spin production at low  $Q^2$  [where  $Q^2 = 4E_{beam}E_e \sin^2 \frac{\theta}{2}$ , and  $E_e(\theta)$  is the energy (polar angle) of the detected positron]. The angular distribution of the scattered leptons is strongly peaked at small polar angles, and the beam asymmetry folds some of the scattered positrons out to detectable regions in the backward endcap. The second of these reactions, the study of single-tagged exclusive hadron production at high  $Q^2$ , allows testing of some very basic QCD predictions in a regime ( $Q^2 > 3 \text{ GeV}^2$ ) which has never been studied before. In both of these cases, the final-state hadrons are concentrated primarily in the barrel and forward regions so the only issue is detection of the scattered positrons with energies in the range  $1 < E < 3 \text{ GeV}$ . This is a primary motivation for the backward endcap calorimeter. Although for financial considerations it is not included in the present design, we hope to be able to add it later to upgrade the capability of the detector for two-photon physics.

---

## References

---

- [CLE93] CLEO Collaboration, “Detector for a  $B$  Factory,” Cornell Report CLNS-91-1047-REV (1993).
- [HEL92] “HELENA, A Beauty Factory at Hamburg,” DESY 92/41 (1992).
- [KEK92]  $B$  Physics Task Force, “Progress Report on Physics and Detector of KEK Asymmetric  $B$  Factory,” KEK Report KEK-92-3 (1992).
- [SLA91] “Workshop on Physics and Detector Issues for a High-Luminosity Asymmetric  $B$  Factory at SLAC,” SLAC-373 (1991).
- [SLA93] “Status Report on the Design of a Detector for the Study of  $CP$  Violation at PEP-II at SLAC,” SLAC-419 (1993).
- [SLA94] “Letter of Intent for the Study of  $CP$  Violation and Heavy Flavor Physics at PEP-II,” SLAC-443 (1994).



---



---

# Physics with BABAR

---

## 3.1 Physics Context

---

The study of  $CP$  violation in  $B$  decays is the primary physics objective of  $BABAR$  [SLA89]. The aim is to achieve a sufficient number of independent determinations of the parameters of the Cabibbo-Kobayashi-Maskawa (CKM) matrix to overdetermine those quantities. This then provides a test of the Standard Model interpretation of  $CP$  violation and, perhaps, a window into physics beyond the Standard Model [Nir93]. This chapter summarizes the essential physics points of this endeavor and briefly discusses other physics that can be done with  $BABAR$ . The measurement of  $CP$  violation asymmetries puts stringent requirements on the detector, so optimizing it for this purpose means that it is also well suited to study most of the other physics of interest in this energy regime.

The two mass eigenstates of the neutral  $B$  meson system can be written as:

$$\begin{aligned} |B_L\rangle &= p|B^0\rangle + q|\bar{B}^0\rangle, \\ |B_H\rangle &= p|B^0\rangle - q|\bar{B}^0\rangle, \end{aligned} \quad (3.1)$$

where  $H$  and  $L$  stand for Heavy and Light, respectively. Defining  $M \equiv (M_H + M_L)/2$ ,  $\Delta M \equiv M_H - M_L$ , and neglecting the tiny difference in width between  $B_H$  and  $B_L$ , the decay widths satisfy  $\Gamma_H = \Gamma_L \equiv \Gamma$ . In this approximation, mixing in the  $B_d^0$  system is governed by a single phase:

$$\left(\frac{q}{p}\right)_{B_d^0} = \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} = e^{2i\phi_M}. \quad (3.2)$$

The amplitudes for decays into a  $CP$  eigenstate  $f$  are

$$A \equiv \langle f | \mathcal{H} | B^0 \rangle, \quad \bar{A} \equiv \langle f | \mathcal{H} | \bar{B}^0 \rangle. \quad (3.3)$$

Let us define

$$r_{CP}(f) \equiv \frac{q}{p} \frac{\bar{A}}{A}. \quad (3.4)$$

The time-dependent rates for initially pure  $B^0$  or  $\bar{B}^0$  states to decay into a final  $CP$  eigenstate at time  $t$  can then be written as

$$\begin{aligned} \Gamma(B_{phys}^0(t) \rightarrow f) &= |A|^2 e^{-\Gamma t} \times \left( \frac{1 + |r_{CP}(f)|^2}{2} \right. \\ &\quad \left. + \frac{1 - |r_{CP}(f)|^2}{2} \cos(\Delta M t) - \text{Im } r_{CP}(f) \sin(\Delta M t) \right), \end{aligned} \quad (3.5)$$

$$\begin{aligned} \Gamma(\bar{B}_{phys}^0(t) \rightarrow f) &= |A|^2 e^{-\Gamma t} \times \left( \frac{1 + |r_{CP}(f)|^2}{2} \right. \\ &\quad \left. - \frac{1 - |r_{CP}(f)|^2}{2} \cos(\Delta M t) + \text{Im } r_{CP}(f) \sin(\Delta M t) \right). \end{aligned} \quad (3.6)$$

The time-dependent  $CP$  asymmetry

$$a_f(t) = \frac{\Gamma(B^0(t) \rightarrow f) - \Gamma(\bar{B}^0(t) \rightarrow f)}{\Gamma(B^0(t) \rightarrow f) + \Gamma(\bar{B}^0(t) \rightarrow f)} \quad (3.7)$$

is given by

$$a_f(t) = \frac{(1 - |r_{CP}(f)|^2) \cos(\Delta M t) - 2 \text{Im } r_{CP}(f) \sin(\Delta M t)}{1 + |r_{CP}(f)|^2}. \quad (3.8)$$

This result holds for a  $CP$ -even final state, while for  $CP$ -odd states there is an additional minus sign in  $r_{CP}(f)$ . When only a single amplitude with a given weak decay phase  $\phi_D$  dominates the decay, one has

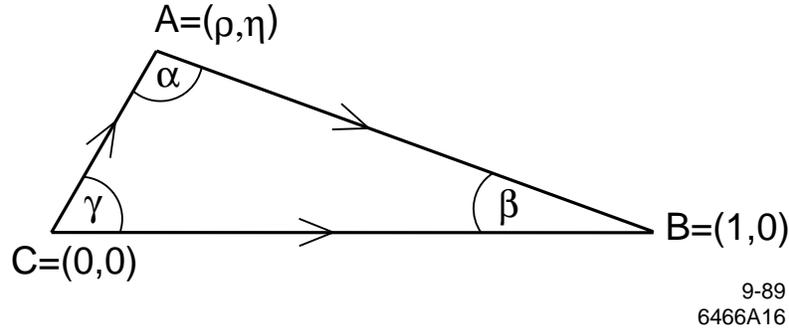
$$\frac{\bar{A}}{A} = e^{-2i\phi_D}. \quad (3.9)$$

Since  $\text{Im } r_{CP}(f) = \sin 2(\phi_M - \phi_D)$ , Equation 3.8 simplifies to

$$a_f(t) = -\sin 2(\phi_M - \phi_D) \sin(\Delta M t). \quad (3.10)$$

In an  $e^+e^- B$  Factory, the initial  $B^0$  and  $\bar{B}^0$  are produced in a coherent  $B^0\bar{B}^0$  state and remain in this state until one of the particles decays. If one  $B$  decays to a flavor-tagging mode and the other decays to a  $CP$ -study mode, the event can be used to reconstruct the time dependence of the asymmetry. In this case, the time  $t$  in the equations above is the time between the tagging decay and the  $CP$ -study-mode decay. The tagging decay may be the later decay, in which case  $t < 0$ . Thus, the time-integrated  $CP$  asymmetry vanishes if  $|r_{CP}(f)| = 1$ , which makes essential a measurement of the time dependence. The asymmetric machine configuration and an accurate vertex determination make this possible.

The Standard Model predictions for  $CP$  violation are most often presented in terms of the unitarity constraints for the three-generation CKM matrix. If one assumes the Standard



**Figure 3-1.** *The Unitarity Triangle.*

Model, one can derive relationships between the mixing phases and the weak phases of various contributions to the decay amplitudes and the angles of a triangle, generally called the Unitarity Triangle. Unitarity of the CKM matrix requires, among others, the relationship

$$V_{tb}V_{td}^* + V_{cb}V_{cd}^* + V_{ub}V_{ud}^* = 0. \quad (3.11)$$

The three complex quantities  $V_{id}V_{ib}^*$  can be represented as a triangle in the complex plane. The three angles of this triangle are labeled

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad (3.12)$$

$$\beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad (3.13)$$

$$\gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \quad (3.14)$$

Figure 3-1 shows the unitarity triangle, as it is usually drawn, rescaled by the side  $V_{cd}V_{cb}^*$ . This makes the base of the triangle real and of unit length. The apex of the triangle is then the point  $(\rho, \eta)$  in the complex plane, in the notation introduced by Wolfenstein [Wol85] for the parameters of the CKM matrix.

To overdetermine the unitarity triangle, experiments must fix as many of the parameters as possible that give the sides,  $|V_{tb}V_{td}^*|/|V_{cb}V_{cd}^*|$  and  $|V_{ub}V_{ud}^*|/|V_{cb}V_{cd}^*|$ , and the angles,  $\alpha$ ,  $\beta$ , and  $\gamma$ . Defining  $\lambda = V_{us} = \sin\theta_{Cabibbo}$ , the quantities  $V_{td} = \lambda$ ,  $V_{ud} = \cos\theta_{Cabibbo}$ , and  $V_{tb} = 1$  are already well-determined up to terms of  $O(\lambda^4)$ . The value of  $V_{cb}$  will be well-determined by the study of semileptonic  $B$  decays at CLEO II, with possible further improvement from  $BABAR$ . Thus, the primary aim of the  $B$  Factory must be to measure accurately the magnitudes of  $V_{ub}$ ,  $V_{td}$ , and the  $CP$ -violating asymmetries in the decay modes  $B_d \rightarrow J/\psi K_S^0$ ,  $B_d \rightarrow J/\psi K_L^0$ ,  $B_d \rightarrow J/\psi K^*$ ,  $B_d \rightarrow D^+D^-$ ,  $B_d \rightarrow D^{*+}D^{*-}$ , and  $B_d \rightarrow \pi^+\pi^-$ ,  $B_d \rightarrow \rho\pi$ . Determination

Quark Process	$B_d$ Mode	$\phi_M - \phi_D$
$b \rightarrow c\bar{c}s$	$J/\psi K_S^0$ $J/\psi K_L^0$ $J/\psi K^{0*}$	$\beta$
$b \rightarrow c\bar{c}d$	$D^+D^-$ $D^{+*}D^{-*}$ $D^{\pm*}D^{\mp}$	$\beta$
$b \rightarrow u\bar{u}d, d\bar{d}d$	$\pi^+\pi^-$ $\rho^\pm\pi^\mp$ $a_1^\pm\pi^\mp$	$\alpha$

**Table 3-1.**  $CP$  modes and Standard Model asymmetry predictions for  $B_d$  decays.

of CKM elements will be discussed below. First, we concentrate on the time-dependent  $CP$  asymmetries.

Table 3-1 shows the Standard Model relationship between the angles of the unitarity triangle and the predicted  $arg(r_{CP}(f)) = 2(\phi_M - \phi_D)$ , assuming a single weak decay amplitude dominates in each of these cases. In addition to pure  $CP$  eigenstates, other channels with  $CP$ -self-conjugate quark content can be analyzed for  $CP$ -violating contributions. These include mixed spin states such as  $J/\psi K^*$  and vector-scalar mixtures such as  $\rho\pi$  and  $D^*D$  where the various possible charge assignments have both  $CP$  and isospin relationships [Ale91, Ale93].

For the  $b \rightarrow c\bar{c}s$  decays, which give channels that measure the angle  $\beta$ , we are in a fortunate situation. Up to corrections of order  $A\lambda^4$ , the penguin diagrams have the same weak phase as the tree diagrams. Thus, these channels have an unambiguous relationship between the measured asymmetry and the angle  $\beta$ . For channels that measure the angle  $\alpha$ , the situation is not quite as good, although still satisfactory. In this case, QCD penguin diagrams are expected to give contributions on the order of perhaps 5–20% of the tree amplitude, and they have a different weak phase. In addition, the  $Z^0$ -mediated penguin diagram also contributes at the level of perhaps 5% of the tree amplitude, with the same weak phase as the QCD penguins [Des95]. Isospin-based analysis can select an isospin channel that has no QCD-penguin contribution [Gro90]. The uncertainty in  $\sin 2\alpha$  extracted from the asymmetry in this channel due to the  $Z^0$ -penguin contributions is then not more than 0.05%. This is less than the expected experimental uncertainty. Furthermore, an isospin-based analysis will give a measure of the total penguin contributions, and hence, can be used to tighten the bounds on the possible  $Z^0$ -penguin effects. In addition, measurement of the time dependence of the asymmetry will allow us to fit the  $\cos(\Delta Mt)$  behavior as well as the  $\sin(\Delta Mt)$  in Equation 3.8. This then gives a further constraint on the magnitude of penguin contributions.

Isospin analyses can probably be carried out most effectively for the  $\rho\pi$  channels, since all three channels decay to  $\pi^+\pi^-\pi^0$  and so have similar detection efficiencies. This channel has the additional advantage that the interference regions between the different charge combinations of  $\rho\pi$  give further information, allowing a determination of  $\sin 2\alpha$  and  $\cos 2\alpha$  which resolves the ambiguity between  $\alpha$  and  $90 - \alpha$  [Qui93]. For the  $\pi\pi$  case, the isospin analysis requires a measurement of the rates for  $B^0 \rightarrow \pi^0\pi^0$  and  $B^+ \rightarrow \pi^+\pi^0$ .

The simulation estimates for the accuracy of  $\alpha$  measurements given below, assume that penguin contributions are sufficiently small that the measured asymmetries can be used directly to determine  $\alpha$  without isospin analysis. The isospin analysis can be used to check this assumption and to obtain a penguin-corrected value. The degradation in accuracy due to this multichannel analysis is not expected to be large.

## 3.2 Simulation Tools

---

Detailed simulation work has been performed to determine the efficiency of the proposed detector to make the necessary measurements of the channels of interest and to ascertain that backgrounds can be simultaneously reduced to an acceptable level. Studies of tagging modes have also been carried out. Two different Monte Carlo simulation packages, ASLUND and BBSIM, were used for this purpose. ASLUND is a fast parametric Monte Carlo, while BBSIM is a detailed simulation based on the GEANT package [CER93]. Both are designed to be flexible, so that variations in detector design can be investigated. Brief descriptions of each are given below. Where not otherwise noted, branching ratio information is taken from the 1994 PDG compilation [PDG94].

### 3.2.1 ASLUND

ASLUND consists of two major components: the JETSET 7.3 [Sjo93] event generator configured for the PEP-II environment and a parametric simulation of the detector. Other generators (*e.g.*, KORALB [Jad90]) are available using the interface routine BEGET [Wri94]. The parametric simulation includes charged tracking, particle identification, and calorimetry.

Charged particle tracking is simulated using the TRACKERR package [Inn93] to estimate the fully correlated error matrix, which is then used to smear the track parameters. The main drift chamber, the silicon vertex detector, and any intervening material are modeled. The TRACKERR input format is used to specify detector geometry, measurement precision, and material distribution. The combination of ASLUND and TRACKERR has been compared to BBSIM for the mode  $B^0 \rightarrow \pi^+\pi^-$  and found to agree within 15% for the predicted  $B^0$  mass width.

Subdetector	Simulation Level
Beam Line, Beam Pipe, and Support Tube	Geometry
Silicon Vertex Detector	Digitizations
Drift Chamber	Hits
DIRC	Digitizations
ATC	Geometry
CsI Calorimeter	Hits
Coil and Instrumented Flux Return	Digitizations

**Table 3-2.** *BABAR* subdetectors modeled in BBSIM.

Simulation of particle identification ( $\pi/K/p$ ) information is available for  $dE/dx$ , DIRC, and aerogel threshold counters (ATCs). The geometrical configuration of these devices can also be specified using the TRACKERR format, and performance parameters can be varied to assess the sensitivity to detector capabilities.

The geometry of the calorimeter, and any material in front of it can also be specified in the TRACKERR format. Photons converted into  $e^+e^-$  pairs in the material in front of the calorimeter are tracked to determine whether or not they will deposit any energy in the calorimeter. Energy loss in material and acceptance losses due to the curl-up in the magnetic field are taken into account in determining the calorimeter signal. A detailed parameterization of the calorimeter response based on GEANT simulations has been developed for some analyses, *e.g.*,  $\rho\pi$  and  $DD$ . Analogous to the other devices discussed above, energy and angular resolutions can be varied by modifying simple parameters.

### 3.2.2 GEANT Simulation—BBSIM

Simulation packages for event generation (BEGET), detector response, and subsequent analysis (BBSIM) have been written to aid in the design and optimization of the *BABAR* detector.

The generation of an event from an  $e^+e^-$  collision and the subsequent loading of decay products into GEANT banks is handled by the BEGET [Wri94] package. The generators currently available include  $\Upsilon(4S)$  decays and hadronic continuum events via JETSET [Sjo93] and  $\tau^+\tau^-$  decays via KORALB [Jad90]. BEGET is capable of overlaying background hits originating from beam-gas collisions on simulated events.

The BBSIM package is based on the CERN detector description and simulation tool, GEANT [CER93]. The latter consists of packages to construct the detector geometry using a set of 15 different volume types; to step charged and neutral particles through the detector,

Cut	$\pi^+\pi^-$	$\pi^0\pi^0$
Geometry and Lepton ID Inefficiency	0.82	0.82
$J/\psi$ Invariant Mass From 0.14 to 0.44 MeV/c	1.00	1.00
Combined $J/\psi$ Efficiency	0.82	0.82
Geometry, Efficiency and Track/Photon Cuts	0.78	0.55
$\pi^0$ Association Cuts	-	0.89
$\pi^0$ Mass Cuts	-	0.93
$K_s^0$ Momentum $0.9 \leq p \leq 2.9$ MeV/c	-	1.00
$K_s^0$ Invariant Mass	0.94	1.00
Combined $K_s^0$ Efficiency	0.73	0.45
$B^0$ Momentum From 0.14 to 0.44 MeV/c	0.99	0.99
$B^0$ Candidate Mass	1.00	0.96
Overall $B^0$ Reconstruction Efficiency	0.59	0.35

**Table 3-3.** Summary of efficiencies for  $B^0$  reconstruction in  $B^0 \rightarrow J/\psi K_s^0$ .

simulating the full variety of interactions with the detector; to define, register, and digitize the Monte Carlo track hits (typically track positions and directions); and to display the detector components, particle trajectories, and track hits. The BBSIM framework consists of a driver routine to invoke GEANT; a database facility, DBIN, to define detector geometry, materials and media; and a set of subpackages, one per detector subsystem, to define the subdetector geometry, register track hits, simulate the detector response, store the results, and perform subsequent analysis. The subsystems included in BBSIM, and the corresponding levels of simulation, are shown in Table 3-2.

### 3.3 Studies of $B^0 \rightarrow J/\psi$ Modes

#### 3.3.1 $B^0 \rightarrow J/\psi K_s^0$

The analysis of this channel and a study of the background were performed using the ASLUND Monte Carlo. This work is described in more detail in Reference [Har95]. The results are based on an assumed branching ratio of  $0.5 \times 10^{-3}$ , calculated using isospin from the measured  $B^+ \rightarrow J/\psi K^+$  rate [PDG94].

The efficiencies obtained are summarized in Table 3-3. The analysis proceeded as follows.

- Candidate  $J/\psi \rightarrow \ell^+\ell^-$  events were selected. Lepton identification efficiencies (calculated using BBSIM) and a  $J/\psi$  mass cut were applied, and  $K_s^0$ s were reconstructed

in both the charged and neutral decay modes. In the charged mode, the invariant mass was required to be within  $\pm 25 \text{ MeV}/c^2$  of the nominal  $K_s^0$  mass. To help suppress the combinatorial background in the neutral mode, various cuts were made on the consistency of the four photons having a common point of origin while also being consistent with the two  $\pi^0$  sub-masses, and the  $K_s^0$  mass and lifetime. The invariant mass was required to be within  $\pm 40 \text{ MeV}/c^2$  of the nominal  $K_s^0$  mass. Note that reconstruction of the neutral  $K_s^0$  decays requires the calorimeter to reconstruct the  $\pi^0$ s.

- Candidate  $J/\psi$  and  $K_s^0$  events were combined to form  $B^0$  candidates. A cut was applied on the  $B^0$  momentum in the  $\Upsilon(4S)$  rest frame (kinematically, this is  $\sim 340 \text{ MeV}/c$ ). The invariant mass of the  $B^0$  candidate was required to be between 5.20 and 5.36  $\text{GeV}/c^2$ .

Potential backgrounds to this decay mode include those from  $B^0 \rightarrow J/\psi X$ , cascade semi-leptonic decays ( $b \rightarrow cl^- \bar{\nu} \rightarrow sl^+ l^- \nu \bar{\nu}$ ), and continuum production of quark-antiquark pairs. Large numbers of each were generated and passed through the analysis described above. The fractions of events surviving all the cuts are summarized in Table 3-4. Also shown are the relative production rates compared to signal events, given by

$$\text{normalization} = \frac{\sigma_{BG} \times B(BG)}{\sigma_{b\bar{b}} \times B(B^0 \rightarrow \text{channel})}.$$

The predicted upper limit on the number of background events per reconstructed signal  $B^0$  is given for each source. No background events survived the cuts in these simulations. All backgrounds together contribute no more than 6%. These levels are considered negligible for the extraction of  $\sin 2\beta$  from the decay asymmetry.

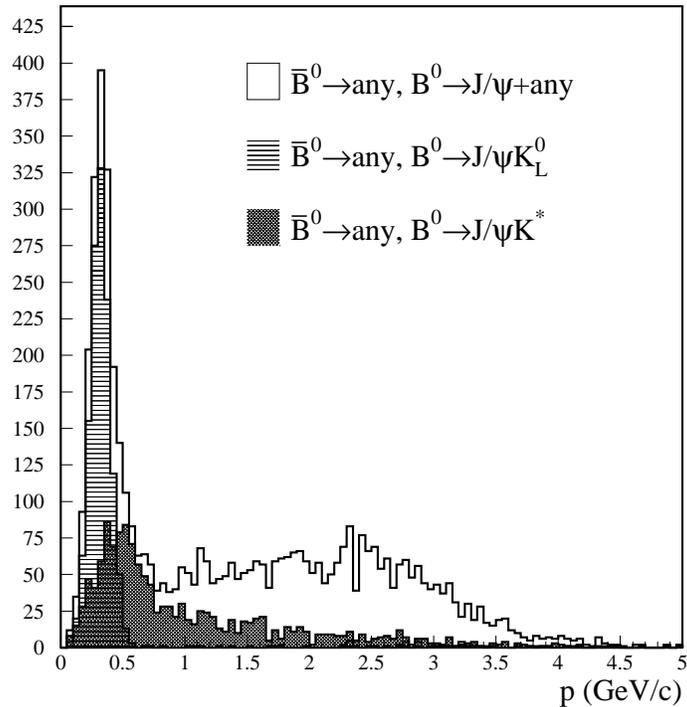
### 3.3.2 $B^0 \rightarrow J/\psi K_L^0$

The branching ratio for this channel is the same as that for  $J/\psi K_s^0$ , but the expected asymmetry has the opposite sign. Thus, high-efficiency detection of  $K_L^0$ s can provide an alternate measurement of  $\beta$  and improve the overall accuracy with which this parameter is determined.

A study of the mode  $B^0 \rightarrow J/\psi K_L^0$  has been carried out [Wri94b]. Since the  $K_L^0$  momentum is not well-measured, the  $K_L^0$  direction and the  $J/\psi$  momentum are combined, and a  $B$ -mass constraint is applied to determine  $p_{rest}$ , the momentum of the  $B$  in the  $\Upsilon(4S)$  rest frame. The resulting distribution for  $p_{rest}$  is shown in Figure 3-2, in which only events from the  $\Upsilon(4S)$  are shown; continuum backgrounds in this channel are negligible, as with other  $J/\psi$  channels. A cut  $p_{rest} \leq 0.42 \text{ GeV}/c$  is made to eliminate false combinations.

Channel	Efficiency	Normalization	# per observed signal event
Charged Mode search:			
$J/\psi K_S^0 \rightarrow l^+ l^- \pi^+ \pi^-$	0.59	1.00	1.00
$J/\psi X$	$7 \times 10^{-4}$	$3.9 \times 10^1$	$4.6 \times 10^{-2}$
Semileptonic	$8 \times 10^{-7}$	$7.8 \times 10^2$	$1 \times 10^{-3}$
$q\bar{q}$ continuum	$1.3 \times 10^{-7}$	$8.1 \times 10^4$	$1.8 \times 10^{-2}$
Neutral Mode search:			
$J/\psi K_S^0 \rightarrow l^+ l^- \pi^0 \pi^0$	0.35	1.00	1.00
$J/\psi X$	$5 \times 10^{-5}$	$7.8 \times 10^1$	$1.1 \times 10^{-2}$
Semileptonic	$1 \times 10^{-7}$	$1.6 \times 10^3$	$4.7 \times 10^{-4}$
$q\bar{q}$ continuum	$1 \times 10^{-7}$	$1.6 \times 10^5$	$4.7 \times 10^{-2}$

**Table 3-4.** Background contributions for the  $J/\psi K_S^0$  channel.



**Figure 3-2.** Momentum distribution of  $B^0$  ( $p_{rest}$ ) from  $\Upsilon(4S) \rightarrow B\bar{B}$  in the  $\Upsilon(4S)$  rest frame, including the effect of the nominal  $e^+e^-$  beam smearing and  $K_L^0$  angular resolution.

Sample	Reconstruction Efficiency	# Reconstructed	Background From $K^*$	Other Background
Low $K^*$	0.25	544	31	74
High $K^*$	0.08	168	65	117
Combined	0.33	712	96	191

**Table 3-5.** Efficiencies for  $B^0$  reconstruction and sensitivity to  $\alpha$  for  $B^0 \rightarrow J/\psi K_L^0$ .

One sees that a substantial background which survives this cut comes from  $B^0 \rightarrow J/\psi K^*$  with  $K^* \rightarrow K_L^0 \pi^0$ . Since the asymmetry for  $J/\psi K^*$  is opposite in sign to that for  $J/\psi K_L^0$ , these events contribute to a dilution of the measured asymmetry in the same fashion as do wrong-sign tags. In order to obtain an accurate estimate of the error on the asymmetry, the data is divided into two samples, one with a low  $K^*$  background and the other with a higher background. The two samples are then treated independently in determining  $\beta$  sensitivity. The high  $K^*$  background sample is defined as follows: any pion in the event is combined with the  $K_L^0$  to define  $K^*$  candidates. The  $K_L^0$  momentum is recalculated under this assumption along with a  $B$  mass constraint. Any event in which a candidate  $K^*$  has a mass within  $\pm 30$  MeV/ $c^2$  of the true  $K^*$  mass is assigned to the high background sample. Events with no  $K^*$  candidates are assigned to the low background sample. Table 3-5 shows the results for the separate samples and the combined sensitivity for this channel.

### 3.3.3 $B^0 \rightarrow J/\psi K^{0*}$

Only the neutral decay mode of the  $K^{0*}$  is useful for  $CP$ -violation measurements. The analysis of this channel and a study of the background were performed using the ASLUND Monte Carlo and the published  $B^0 \rightarrow J/\psi K^{0*}$  branching ratio of  $1.6 \times 10^{-3}$  [PDG94]. The analysis proceeded in a manner similar to that for  $B \rightarrow J/\psi K_S^0$  described above. First,  $J/\psi$  and  $K_S^0$  candidates were formed as above, where only the charged decay mode of the  $K_S^0$  was used in this case. Candidate  $\pi^0$ s were obtained from pairs of calorimeter clusters, assuming that the photons originated at the nominal interaction position of the experiment. All combinations of  $\pi^0$  and  $K_S^0$  candidates were made, and a mass cut applied to the resulting  $K^{0*}$  candidates. All mass cuts and the various subefficiencies of each stage of the reconstruction are summarized in Table 3-6.

Cut	Efficiency
Geometry and Lepton ID Inefficiency	0.82
$3.06 < J/\psi$ Mass $< 3.14$ GeV/c <sup>2</sup>	1.00
Combined $J/\psi$ Efficiency	0.82
Geometry and $E_{\gamma min} = 25$ MeV/c <sup>2</sup>	0.73
$0.120 < \pi^0$ Mass $< 0.150$ GeV/c <sup>2</sup>	0.95
Combined $\pi^0$ Efficiency	0.69
Geometry and Track Cuts	0.77
$0.468 < K_s^0$ Mass $< 0.528$ GeV/c <sup>2</sup>	0.92
Combined $K_s^0$ Efficiency	0.71
Combined $K^{0*}$ Sub-efficiencies	0.50
$0.73 < K^{0*}$ Mass $< 1.05$ GeV/c <sup>2</sup>	0.98
Overall $K^{0*}$ Efficiency	0.50
Combined $B^0$ Sub-efficiencies	0.40
$B^0$ Momentum in $\Upsilon$ Rest Frame	0.99
$B^0$ Candidate Mass	0.99
Overall $B^0$ Reconstruction Efficiency	0.39

**Table 3-6.** *Cuts and Efficiencies for  $B^0$  Reconstruction in  $B^0 \rightarrow J/\psi K^{0*}$ .*

## 3.4 Studies of $B^0$ to Double Charm Modes

### 3.4.1 $B^0 \rightarrow D^+ D^-$

A branching fraction of  $6 \times 10^{-4}$  is used for this channel and is obtained by averaging the measured branching fractions for  $B^+$  and  $B^0$  to decay to  $DD_S$  and multiplying by  $\sin^2 \theta_C$ .

Only those  $D^+$  modes with no more than one  $\pi^0$  and a branching fraction greater than about 1% are considered. Only those  $K_s^0$ s in the  $\pi^+\pi^-$  decay mode are included. The channels used are shown in Table 3-7 with branching ratios taken from Reference [PDG94]. The mass resolutions quoted are vertex constrained. The efficiency for each mode in Table 3-7 includes cuts on particle identification for each track and a cut on  $\chi^2$  probability for the reconstructed  $D$  vertices at 2%. Mass cuts of  $\pm 3\sigma_{m_D}$  are made on each  $D$  meson candidate.

Pairs of  $D$  meson candidates are combined to form  $B$  candidates. Each  $B$  candidate is required to have a total momentum between 180 and 440 MeV/c in the  $\Upsilon(4S)$  rest frame. A vertex and kinematic fitting package is then used to fit the entire  $B$  event, constraining the intermediate  $D$  masses and vertices and the two  $D$ s to a single vertex. The  $\chi^2$  probability

Mode	Branching Fraction	Single $D$ Efficiency	$\sigma_{m_D}$ (MeV)
$K^-\pi^+\pi^+$	0.091	0.67	4.3
$K^-\pi^+\pi^+\pi^0$	0.064	0.42	6.9
$K_S^0\pi^+$	0.014	0.59	5.0
$K_S^0\pi^+\pi^0$	0.049	0.38	9.1
$K_S^0\pi^-\pi^+\pi^+$	0.035	0.41	3.7
$K_S^0\pi^-\pi^+\pi^+\pi^0$	0.027	0.26	5.5

**Table 3-7.** Efficiency and mass resolutions for  $D^+$  reconstruction.

$B \rightarrow D^+ D^- \rightarrow$	Branching Fraction ( $\times 10^{-3}$ )	Efficiency	$\sigma_{m_B}$ (MeV)	$\sigma_{z_B}$ ( $\mu\text{m}$ )	# of Events in $30 \text{ fb}^{-1}$	S/B
$D^+ \rightarrow K^-\pi^+\pi^+$ $D^- \rightarrow K^+\pi^-\pi^-$	8.3	0.45	5.0	47.0	80.0	1.8
$D^+ \rightarrow K^-\pi^+\pi^+$ $D^- \rightarrow K^+\pi^-\pi^-\pi^0$	11.6	0.24	6.5	54.0	61.0	0.7
$D^+ \rightarrow K_S^0\pi^+$ $D^- \rightarrow K^+\pi^-\pi^-$	2.5	0.38	5.3	58.0	14.0	> 3.0

**Table 3-8.** Efficiency, mass and vertex resolution, signal, and background for three of the  $D$  decay mode combinations used in the  $B \rightarrow D^+ D^-$  reconstruction.

for this fit is required to be greater than 2%. The reconstructed mass is required to lie within  $\pm 20 \text{ MeV}/c^2$  of the  $B^0$  mass. For events passing the above cuts, a  $B$  mass constraint is added to determine the  $B$  decay point. There are many combinations of the  $D$  decay modes listed in Table 3-7, each with different efficiencies and resolutions. Table 3-8 gives the efficiencies and resolutions for three of the 21 combinations. The complete set of numbers can be found in Reference [Cou95].

Potential sources of background are from other decays of the  $\Upsilon(4S)$  and continuum quark-antiquark production. Table 3-8 includes an estimate of signal-to-background ratios ( $S/B$ ) for the modes discussed. Channels with two or more  $\pi^0$ s in the final state tended to have large backgrounds and were, therefore, not included in this analysis. Although each combination of  $D$  decay modes is analyzed separately, one can sum over all modes to obtain a total reconstructable branching fraction of 0.044, an average acceptance of 0.27, and a final event sample of 259 events with a background of 351 events.

Mode	Branching Fraction	Single $D^0$ Efficiency	$\sigma_{m_{D^0}}$ (MeV)
$K^- \pi^+$	0.040	0.77	5.2
$K^- \pi^+ \pi^0$	0.138	0.48	8.8
$K_s^0 \pi^+ \pi^-$	0.026	0.48	4.0
$K_s^0 \pi^+ \pi^- \pi^0$	0.049	0.29	6.8
$K^- \pi^+ \pi^+ \pi^-$	0.081	0.55	3.7
$K^- \pi^+ \pi^+ \pi^- \pi^0$	0.043	0.34	5.5

**Table 3-9.** Efficiency for reconstruction of  $D^0$ s produced in  $B \rightarrow D^{*+} D^{*-}$  decays.

### 3.4.2 $B^0 \rightarrow D^{*+} D^{*-}$

A branching fraction of  $1.5 \times 10^{-3}$  is used for this mode which is 2.5 times larger than that for  $D^+ D^-$ . This factor is the typical increase for vector-vector decay modes over comparable scalar-scalar channels.

The only decay chain considered is that in which each  $D^{*\pm}$  decays to  $D^0 \pi^\pm$  with a branching fraction of  $68 \pm 1.6\%$ . The efficiency for finding the slow pions has been studied in more detailed simulations of the vertex detector and found to be better than 95% over the momentum range of interest above  $p_t = 60 \text{ MeV}/c$  for  $\theta > 300 \text{ mr}$ . The branching fraction, efficiency, and mass resolution for the  $D^0$  modes are given in Table 3-9.

In reconstructing  $D^*$  mesons, all  $D^0$  candidates are combined with charged pions and a cut made at  $\pm 3\sigma$  on the  $D^* - D^0$  mass difference. Pairs of  $D^*$  candidates are then fit for the  $B$  mass, constraining the vertex and mass of the intermediate  $D^*$  and  $D^0$ . Table 3-10 gives the efficiency, resolutions, and background levels for three of the 21 combinations. The complete set of results can be found in Reference [Cou95]. Although each combination of  $D^0$  decay modes is analyzed separately, one can sum over the modes to obtain a total reconstructable branching fraction of 0.125, an average acceptance of 0.15, and a final event sample of 473 events with a background of 21 events.

In computing  $CP$  reach,  $D^{*+} D^{*-}$  channels are assumed to be dominated by a single  $CP$  state. If this turns out not to be the case, a partial-wave analysis will need to be performed to disentangle the  $CP = +$  and  $CP = -$  states. This will decrease the effectiveness of these channels by a factor that depends on the fraction of each  $CP$  state present.

$B \rightarrow D^{*+}D^{*-} \rightarrow$	Branching Fraction ( $\times 10^{-3}$ )	Efficiency	$\sigma_{m_B}$ (MeV)	$\sigma_{z_B}$ ( $\mu\text{m}$ )	# of Events in $30 \text{ fb}^{-1}$	S/B
$D^0 \rightarrow K^- \pi^+$ $D^0 \rightarrow K^+ \pi^-$	1.6	0.41	5.6	53.0	16.0	> 2.0
$D^0 \rightarrow K^- \pi^+$ $D^0 \rightarrow K^+ \pi^- \pi^0$	11.0	0.21	5.8	65.0	58.0	> 7.0
$D^0 \rightarrow K_S^0 \pi^+ \pi^-$ $D^0 \rightarrow K^+ \pi^-$	2.1	0.29	4.3	55.0	10.0	> 1.2

**Table 3-10.** Efficiency, mass and vertex resolution, signal, and background for three of the  $D$  decay mode combinations used in the  $B \rightarrow D^{*+}D^{*-}$  reconstruction.

### 3.4.3 $B \rightarrow DD^*$

While this mode is not a  $CP$  eigenstate,  $CP$ -violating asymmetries can be measured using both  $D^{*+}D^-$  and  $D^{*-}D^+$  data [Ale91]. A full simulation of these modes has not been carried out, but efficiencies and backgrounds can be estimated using the two previous studies. The decays  $B \rightarrow D^{*-}D_s^+$  and  $B \rightarrow D^-D_s^{*+}$  have been observed with branching ratios of  $1.2 \pm 0.6\%$  and  $2.1 \pm 1.5\%$ , respectively [PDG94]. Rescaling by  $\sin^2 \theta_C$  gives reasonable estimates for the  $DD^*$  rates. Using the acceptances determined in the two previous sections we predict a total reconstructable branching fraction of 0.085, an average acceptance of 0.168, and a final event sample of 557 events.

## 3.5 Studies of $B \rightarrow \pi\pi$ Modes

The measurement of  $\sin 2\alpha$  from a  $CP$ -violating asymmetry in the  $B^0 \rightarrow \pi^+\pi^-$  mode requires the selection of rather clean samples of  $\pi^+\pi^-$  as well as  $\pi^\pm\pi^0$  and  $\pi^0\pi^0$  events [Ale92]. The latter are needed for an isospin analysis to estimate the size of the penguin contribution to the total decay amplitude and to extract  $\sin 2\alpha$  from the  $CP$  asymmetry for the case in which the QCD-penguin amplitudes are large [Gro90]. Unfortunately, the branching ratios for charmless two-body  $B$  decays into either pions or kaons are expected to be very small, on the order of  $10^{-5}$  or less [Bro93]. Therefore, these channels suffer from a potentially large background due to continuum quark-antiquark pair production, despite a clean signature. Simple cuts strongly reduce the continuum background and allow one to enrich the sample in two-body  $B^0$  decays. In addition, the branching ratio of  $B \rightarrow K\pi$  modes may be comparable to  $B \rightarrow \pi\pi$ . In order to disentangle the  $\pi^+\pi^-$  component of the signal from  $K^\pm\pi^\mp$  events, a very good particle identification system operating up to high momentum (4–4.5 GeV/c) is needed. This

is one of the main features of the *BABAR* detector. Studies for all of these channels have been performed using both ASLUND and BBSIM. Because of the large sample of Monte-Carlo events needed, the continuum background was studied using ASLUND only. The branching ratio for  $B \rightarrow \pi^+\pi^-$  is taken to be  $1.2 \times 10^{-5}$ , which is based on the measurement from CLEO [Bat93a].

### 3.5.1 $B^0 \rightarrow \pi^+\pi^-$

Despite apparent simplicity, analysis of this mode, dealing with rejection factors at the  $10^{-5}$  to  $10^{-6}$  level, requires a precise knowledge of detector systematics. Strong track quality cuts will be needed; these are not included in the present study. Software tools (full detector simulation and pattern recognition packages) are being developed and will be used to refine the selection cuts and to give more accurate estimates of signal efficiencies and background rejection.

#### Reduction of Continuum Background

The reduction of the continuum background is achieved by cuts applied in the  $\Upsilon(4S)$  center-of-mass frame. The first two cuts exploit the fact that the  $B$  mesons are produced almost at rest, and thus the pions are almost monoenergetic, in the  $\Upsilon(4S)$  rest frame. The third uses the fact that continuum background events are mostly jet-like, whereas  $B^0\bar{B}^0$  decays are not. A cut on particle identification ( $\Delta\chi^2$ ) gives a significant background reduction. This occurs because the sample of background tracks hard enough to masquerade as a  $B \rightarrow \pi\pi$  event is enriched in kaons and protons.

Charged particles of momentum  $p^*$  in the  $\Upsilon(4S)$  center of mass are selected in a very narrow range:  $2.35 \leq p^*(\pi^\pm) \leq 2.95$  GeV/ $c$ . This momentum window contains virtually 100% of reconstructed pions (or kaons) from two-body decays. The probability that a continuum charged particle lies in this momentum window is less than 2%. The probability that a continuum event contains two such particles with opposite electric charge is  $3.3 \times 10^{-3}$ .

Any selected pair of opposite sign particles forms a  $B^0$  candidate. The  $B^0$  candidate momentum  $p^*(B)$  in the  $\Upsilon(4S)$  center of mass frame is required to lie in the 150–500 MeV/ $c$  range. This cut selects pairs with large opening angles. To be conservative, we have used a rather large window for the  $B^0$  mass cut. The signal-to-background ratio (S/N) is a strong function of the angle  $\theta_{sph}$  between the B-decay axis and the sphericity axis formed by the remaining tracks in the event. The background is strongly peaked at  $\cos\theta_{sph} = \pm 1$  while the distribution of the signal in this variable is flat. We divide the data into 3 bins ( $< 0.7$ ,  $0.7-0.9$ , and  $0.9-1.0$ ) in  $|\cos\theta_{sph}|$  in order to optimize the  $CP$ -sensitivity in this channel. The

Cut	$B^0 \rightarrow \pi^+\pi^-$	$q\bar{q}$
$\pi^+\pi^-$ Acceptance and $p^*(\pi^\pm)$ in [2.35, 2.95]	0.862	$3.3 \times 10^{-3}$
0 : Preselected Events (Normalization)	1.000	1.000
1 : $p^*(B)$ in [0.150, 0.500]	0.994	0.246
2 : $\Delta\chi^2 \geq 0$	0.972	0.275
1 + 2 + $ \cos\theta_{sph}  < 0.7$	0.538	$5.0 \times 10^{-6}$
1 + 2 + $0.7 <  \cos\theta_{sph}  < 0.9$	0.166	$6.0 \times 10^{-6}$
1 + 2 + $0.9 <  \cos\theta_{sph}  < 1.0$	0.094	$3.9 \times 10^{-5}$
Overall Efficiency per Event Generated	0.798	$5.0 \times 10^{-6}$

**Table 3-11.** Efficiencies of background rejection cuts for the  $B^0 \rightarrow \pi^+\pi^-$  mode.

particle identification requirement on  $\Delta\chi^2$  is explained in the following section. The results are summarized in Table 3-11.

## Particle Identification

The rejection of  $K\pi$  events from the sample is a crucial step in measuring the correct asymmetry. The criterion used combines information from kinematics (mass resolution),  $dE/dx$ , and the DIRC particle identification system. The quality of all three pieces of information is essential to obtaining good background rejection without significant loss of signal.

Combining all kinematic and PID information leads to a  $\chi^2$  test with up to five degrees of freedom. In reality, within the acceptance of BABAR, one particle lies outside of the DIRC acceptance  $\sim 10\%$  of the time. Thus, the number of measurable degrees of freedom differs from event to event. We define a discriminating variable  $\Delta\chi_i^2$  as the average  $\chi^2$  difference between the null hypothesis ( $\pi\pi$ ) and any one of the alternate hypotheses. The five possible terms are a kinematic term for the pair,  $\Delta\chi_{kin}^2$ , and two possible particle identification terms for each member of the pair,  $\Delta\chi_{j\,dE/dx}^2$  and  $\Delta\chi_{j\,DIRC}^2$ . We retain only the hypothesis giving  $\chi^2 \geq 0$ . Simulation samples of  $B^0 \rightarrow \pi\pi$  and  $B^0 \rightarrow K\pi$  are used to determine  $\alpha$ , the probability of rejecting a true  $\pi\pi$  event, and  $\beta$ , the probability of accepting an event with a  $K$ , as functions of the minimum  $\Delta\chi^2$  needed to reject the alternate hypotheses. The results obtained with ASLUND are listed in Table 3-12. One can see from this table that for a significance  $\alpha = 5\%$ , obtained for  $\Delta\chi_{crit}^2 \simeq 4$ ,  $\beta$  can be maintained below 1%. The simulation's results for the  $dE/dx$  and momentum resolutions are based on ideal performance of the drift chamber and are probably optimistic. They do not include the single cell efficiencies, track overlaps, background rate effects, polar angle dependence of the  $dE/dx$  response, or a myriad of other factors which need to be addressed by an operating

$\min \Delta\chi^2$	0	4	9	0	4	9
	$1 - \alpha$			$\beta$		
Kinematics	86	51	12	9.7	1.8	—
$dE/dx$	75	36	6	12	1.5	—
Kin. + $dE/dx$	92	80	53	4.7	1.1	—
DIRC	74	69	58	1.8	—	—
Kin. + $dE/dx$ + DIRC	98	95	87	2.4	0.8	—

**Table 3-12.** Probabilities for selecting a  $\pi^+\pi^-$  event ( $1 - \alpha$ ) and a  $K^+\pi^-$  event ( $\beta$ ), for several minimum values  $\Delta\chi^2$ , based on an ASLUND simulation.

experiment. We note that neither CLEO nor ARGUS have been able to achieve reliable  $K/\pi$  separation for high momentum tracks with this method.

## Results

The  $\pi^+\pi^-$  events are separated into three classes for analysis of  $CP$  sensitivity [Sny95]. Class 1 contains events with  $\cos\theta_{sph} < 0.7$ , class 2 contains those with  $0.7 \leq \cos\theta_{sph} \leq 0.9$ , and class 3 is all other events. Using the results listed in Table 3-11 and assuming a branching ratio of  $1.2 \times 10^{-5}$  [Bat93a], one can reach a background to signal ratio ( $B/S$ ) of around 2 with a large efficiency  $\epsilon = 54\%$ . While the backgrounds in classes 2 and 3 are higher, inclusion of these events still improves the overall sensitivity.

The effect of the difference between the signal and the background in  $\Delta z$  the axial separation of the two  $B$  decays, further reduces the impact of continuum backgrounds. This is discussed below.

### 3.5.2 $B^0 \rightarrow \pi^0\pi^0$

The reconstruction of  $\pi^0$ s from calorimeter data is discussed in Chapter 7. The resolution obtained on the  $\pi^0$  mass using BBSIM is  $8.3 \text{ MeV}/c^2$ . This leads to an efficiency of 84% for  $\pi^0$  reconstruction, restricting  $\pi^0$  candidates to the  $[100 - 170 \text{ MeV}/c^2]$   $\gamma\gamma$  mass window. The same mass window is chosen for background studies with ASLUND.

The results obtained with ASLUND for this mode are summarized in Table 3-13. We assume a branching ratio for  $B \rightarrow \pi^0\pi^0$  of  $5 \times 10^{-6}$ . After selection of the  $\pi^0\pi^0$  pair, the same kinematic cuts as in the  $\pi^+\pi^-$  modes are applied. Using BBSIM, the resolution on the reconstructed  $B^0$  mass in  $\pi^0\pi^0$  is found to be  $49.7 \text{ MeV}/c^2$ , dominated by the errors on the

Cut	$B^0 \rightarrow \pi^0\pi^0$	$q\bar{q}$
$\pi^0\pi^0$ Acceptance and $p^*(\pi^0)$ in [2.350, 2.950]	0.849 (0.772)	$0.6 \times 10^{-3}$
0 : Preselected Events (Normalization)	1.000	1.000
1 : $p^*(B)$ in [0.150, 0.500]	0.989	0.263
2 : Cosine Angle w.r.t. Sphericity Axis $\leq 0.8$	0.770	0.115
3 : Cosine Angle w.r.t. Sphericity Axis $\leq 0.7$	0.661	0.071
1 + 2	0.765	0.032
1 + 2 + $m(\pi\pi)$ in [5.120, 5.440]	0.754	0.015
1 + 3	0.656	0.019
1 + 3 + $m(\pi\pi)$ in [5.120, 5.440]	0.648	$6.5 \times 10^{-3}$
Overall Efficiency per Event Generated	0.550 (0.518)	$4.0 \times 10^{-6}$

**Table 3-13.** Efficiencies of background rejection cuts for the  $B^0 \rightarrow \pi^0\pi^0$  mode. The numbers quoted in parentheses are obtained when combinatorial background in signal events is suppressed.

measurement of photon energies. The cut on the  $\pi^0$  masses, together with a  $3\sigma$  cut on the  $B^0$  mass ( $5.120 < m < 5.440 \text{ GeV}/c^2$ ), yields an efficiency of 55% with a B/S of less than 4 to 1. These simulations have been made using a somewhat different detector geometry than given in this *Technical Design Report*, but the conclusions are still valid.

### 3.5.3 $B^+ \rightarrow \pi^+\pi^0$ Decays

In addition to the  $\pi^0\pi^0$  mode, the isospin analysis requires results from  $B^+ \rightarrow \pi^+\pi^0$  decays. A treatment similar to that described above has been made for this mode, and the results are presented in Table 3-14. An efficiency of 53% is achieved with a B/S of less than 3.5 to 1 using an assumed branching ratio of  $1.2 \times 10^{-5}$ .

## 3.6 $B^0 \rightarrow \rho^\pm\pi^\mp$

The branching ratio for  $B^0 \rightarrow \rho^+\pi^-$  is taken to be  $(f_\rho/f_\pi)^2 \times BR(B^0 \rightarrow \pi^+\pi^-) = 2.9 \times 10^{-5}$ . The value  $(f_\rho/f_\pi)^2 = 2.4$  is determined from  $\tau \rightarrow \rho\nu$  and  $\tau \rightarrow \pi\nu$  data [PDG94]. This number is higher than that used in the *Letter of Intent* because recent LEP data on  $\tau \rightarrow \rho\nu$  has caused a significant increase in the best fit value. The corresponding rate for  $B \rightarrow \pi^+\rho^-$  is more model dependent; we use  $R = (B^0 \rightarrow \pi^+\rho^-)/(B^0 \rightarrow \rho^+\pi^-) = 1$ . Estimates for

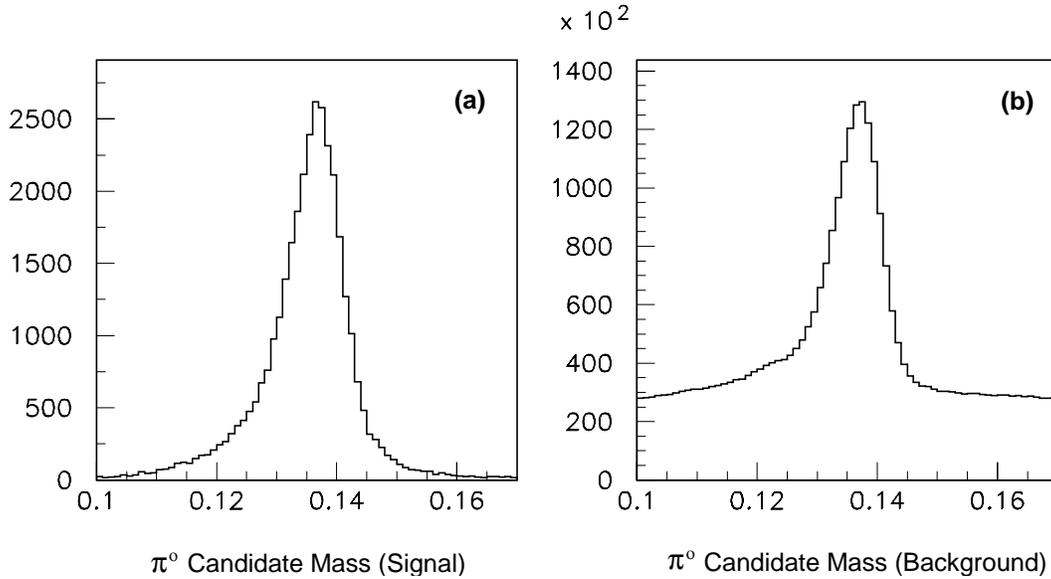
Cut	$B^+ \rightarrow \pi^+ \pi^0$	$q\bar{q}$
$\pi^+ \pi^0$ Acceptance and $p^*(\pi)$ in [2.35, 2.95]	0.863 (0.803)	$3.8 \times 10^{-3}$
0 : Preselected Events (Normalization)	1.000	1.000
1 : $p^*(B)$ in [0.150, 0.500]	0.987	0.252
2 :  Cosine Angle w.r.t. Sphericity Axis  $\leq 0.8$	0.756	0.120
3 :  Cosine Angle w.r.t. Sphericity Axis  $\leq 0.7$	0.650	0.069
4 : $\Delta\chi^2 \geq 0$	0.942	0.487
1 + 2 + 4	0.717	0.013
1 + 2 + 4 + $m(\pi\pi)$ in [5.150, 5.410]	0.713	$4.5 \times 10^{-3}$
1 + 3 + 4	0.617	$6.2 \times 10^{-3}$
1 + 3 + 4 + $m(\pi\pi)$ in [5.150, 5.410]	0.614	$2.3 \times 10^{-3}$
Overall Efficiency per Event Generated	0.529 (0.506)	$8.5 \times 10^{-6}$

**Table 3-14.** Efficiencies of background rejection cuts for the  $B^+ \rightarrow \pi^+ \pi^0$  mode. The numbers quoted in parentheses are obtained when combinatorial background in signal events is suppressed.

the ratio range from a high of 15 [Ale91] to a low of about 1/12 [Dea93]. The sensitivity of  $\delta\alpha$  to this ratio varies as  $\sqrt{(R+1)/2R}$ .

An analysis of this channel and a study of the background were performed using the ASLUND Monte Carlo. The calorimeter performance was determined using BBSIM and was parameterized for the study. Only the backgrounds due to production of quark-antiquark pairs in the continuum have been considered, as they are expected to be the dominant source. Other  $\rho\pi$  channels decay to the same final state,  $\pi^+ \pi^- \pi^0$ , and will be subject to similar backgrounds. Eventually, all three possible final states will need to be studied to allow a full isospin analysis. However, it is sufficient to show that the detector is well-suited to this study, which has the potential to provide the best measurement of  $\sin 2\alpha$ . Both calorimetry and particle identification are important in the reconstruction of  $\rho^\pm \pi^\mp$  and the reduction of the continuum backgrounds. This study assumes small penguin contributions and neglects the interference between the channels.

The mass distributions of photon pairs obtained using the ASLUND Monte Carlo are shown in Figures 3-3(a) and (b) for the signal and light-quark continuum events, respectively. The low-mass tail of the  $\pi^0$  peak is caused mainly by incomplete containment of showers as defined by the clustering algorithm, and possibly by longitudinal leakage from the back of the calorimeter. The peak of the signal distribution is offset slightly to the right of the nominal  $\pi^0$  mass as a consequence of the combination of the low-energy tail and the resolution parameterization in ASLUND, which maintains the mean energy of clusters at the correct value.



**Figure 3-3.** Reconstructed two-photon mass for (a) signal and (b) light-quark background events.

Candidate  $B^0$ s were reconstructed by combining  $\rho^\pm$  candidates and all charged tracks of the opposite sign. Further cuts on the  $B^0$ -candidate momentum in the rest frame of the  $\Upsilon(4S)$ , and on its reconstructed mass, were made. These cuts have a dramatic effect on the backgrounds from all sources. The effects of all the above cuts on the signal are summarized in Table 3-15.

The fraction of continuum events that contribute a  $B^0$  candidate to the background at this stage is  $5.4 \times 10^{-3}$ . Assuming a branching ratio of  $5.8 \times 10^{-5}$  for  $B^0 \rightarrow \rho^\pm \pi^\mp$ , we find that the continuum background dominates the signal by a factor of  $\sim 140:1$ . A number of further cuts were made to reduce continuum backgrounds:

- All  $B^0$  candidates were required to have a vertex within 5 cm of the nominal interaction point in the  $z$  direction. This excluded tracks from conversions and nonprompt decays with no appreciable effect on the signal.
- The two charged tracks from a  $B^0$  decay to  $\rho^\pm \pi^\mp$  were required to form a good vertex in three dimensions. In charm background events, the two charged tracks involved may arise from different charm decays and hence be well-separated in space. Candidates were therefore required to have a vertex  $\chi^2$  per degree of freedom less than 5.
- In the decay  $B \rightarrow \rho^\pm \pi^\mp$ , the  $\rho$  is 100% polarized, so that the angle of the decay axis with respect to its line of flight has a  $\cos^2 \theta_{decay}$  distribution. Background events,

Cut	Signal Efficiency
Geometry and $E_{min}(\gamma) = 25 \text{ MeV}$	0.76
$111 \leq m(\pi^0) \leq 160 \text{ MeV}/c^2$	0.98
Combined $\pi^0$ Efficiency	0.74
Geometry and $p_{t,min}(\pi) = 60 \text{ MeV}$	0.91
$0.37 \leq m(\rho) \leq 1.17 \text{ GeV}/c^2$	1.00
Combined $\rho^{+-}$ Efficiency	0.91
Geometry and $p_{t,min} = 60 \text{ MeV}$	0.96
$B^0$ Momentum in $\Upsilon$ Rest Frame	1.00
$B^0$ Candidate Mass	1.00
Overall $B^0$ Reconstruction Efficiency	0.65

**Table 3-15.** Efficiencies of  $B^0$  reconstruction for  $\rho\pi$  modes.

however, should show no such correlation. A cut was placed at  $|\cos\theta_{decay}| \geq 0.25$  of the angle between the two directions in the  $\rho$  rest frame.

- More severe cuts were made on the  $B^0$ -candidate momentum in the  $\Upsilon(4S)$  rest frame, and on its invariant mass.
- As in the  $\pi^+\pi^-$  case, the light-quark background events have a pronounced jet structure, with the identified  $B$ -decay tracks lying within the jets. The thrust (or sphericity) axis of the remaining particles in the event is strongly correlated with the  $B$ -decay direction in background events and uncorrelated in signal events. Data are divided into six bins in  $|\cos\theta_{thrust}|$  and all data is used to achieve maximum  $CP$ -sensitivity. Results for three of these bins are shown in the Table 3-16. For the best region,  $|\cos\theta_{thrust}| \leq 0.5$  a  $B/S$  of less than  $3/2$  is achieved with kaon tags, while with lepton tags this ratio is less than 0.15.

The efficiencies (and the relative tagging efficiencies) of the above cuts for signal and background are summarized in Table 3-16 along with the efficiencies for initial reconstruction, the branching ratios and cross sections, and finally, the overall background contamination per reconstructed  $B$  meson. Unbracketed numbers correspond to kaon tags, bracketed ones to lepton tags. A multivariate analysis has also been performed, following the lines of Reference [Ale91], and leads to similar results.

Cut	Signal	Light Quark	Charm
Reconstructions and Preliminary Cuts	0.65	$6.70 \times 10^{-3}$	$3.20 \times 10^{-3}$
$ Z_{vert}  < 5 \text{ cm}$	1.00	1.00	1.00
Chi-squared per d.o.f. of Vertex	0.98	0.92	0.68
$ \cos\theta_{decay}  > 0.25$	0.98	0.85	0.87
$0.14 \leq p(B^0) \leq 0.44 \text{ GeV}/c$	0.98	0.42	0.42
$5.20 \leq m(B^0) \leq 5.36 \text{ GeV}/c^2$	0.95	0.31	0.30
Tagging Efficiency w.r.t. Signal	1.0(1.0)	0.47(0.049)	1.4(0.065)
One candidate per event	1.0	0.77	0.68
Efficiency for $0.0 <  \cos\theta_{thrust}  < 0.5$	0.29	$1.13(0.117) \times 10^{-5}$	$0.69(0.031) \times 10^{-5}$
Efficiency for $0.7 <  \cos\theta_{thrust}  < 0.8$	0.058	$1.65(0.170) \times 10^{-5}$	$1.98(0.089) \times 10^{-5}$
Efficiency for $0.9 <  \cos\theta_{thrust}  < 1.0$	0.058	$16.0(1.65) \times 10^{-5}$	$13.5(0.611) \times 10^{-5}$

**Table 3-16.** Example of background rejection cuts for  $\rho\pi$  modes. Unbracketed numbers correspond to kaon tags (class b), bracketed ones to lepton tags (class a). Data samples in higher  $|\cos(\theta_{thrust})|$  bins have higher backgrounds.

### 3.7 Tagging Modes

Measurements of  $CP$  asymmetries and  $B^0$  or  $B_s$  mixing require the determination of the  $B$  flavor. At *BABAR*, events are studied in which a coherent  $B^0\bar{B}^0$  state produces one  $B$  meson decaying into a mode under investigation for  $CP$  violation, while the flavor of the second  $B$  is reconstructed (or tagged) from its decay products. Any particle with a distinct flavor quantum number provides a possible tag. Typically this information is in the form of a lepton charge, a  $D$  charge, or a  $K$  charge.

The observed asymmetry  $a_{obs}(t)$  can be related to the expected asymmetry,  $a_{theory}$ , by

$$a_{obs}(t) = D \cdot a_{theory}(t) \quad (3.15)$$

with a dilution factor  $|D| < 1$ , where  $D = (1 - 2w)$ , and  $w$  is the wrong-tag probability. (The further dilution effect of smearing due to limited time resolution is discussed in the following section.) Different types of tagging information in a single event are combined, leading to an event-by-event definition of the appropriate dilution factor.

Details of the most recent tagging studies are discussed in References [Wal95] and [Pla95]. The results are similar. The numbers reported here are taken from the first of these studies, which uses a Monte Carlo simulation based on a version of the event generator JETSET 7.4 [Sjo93] tuned to reproduce ARGUS data [Alb94b]. This study builds on earlier work which identified a number of effective tagging strategies [Jaf94, Wal94a, Pia94]. The strategy is to

use all possible information in each event to assign the appropriate flavor and dilution factor, and then to extract asymmetries from a two-dimensional fit of the data against time and effective dilution. The detector acceptance, momentum and energy resolution, and particle identification are simulated with ASLUND.

A flavor estimator that allows us to combine all tagging information in an event is assigned as follows. For each event, find all possible tagging identifiers. For each identifier, calculate the quantities  $L_i$ ,  $Q_i$ , and  $D_i$  and define  $B_i = L_i \cdot D_i \cdot Q_i$ . The likelihood ratio  $L_i$ , which estimates whether the tagging track is correctly identified as a particle of type  $i$ , is given by

$$L_i = \frac{e^{-\chi_i^2/2}}{\sum_{j=e,\mu,\pi,K,p} e^{-\chi_j^2/2}}, \quad (3.16)$$

where  $\chi_i^2$  is the sum of the  $\chi^2$  from each available source of information for particle type  $i$ . The flavor charge  $Q_i$  for a track of type  $i$  is defined to be +1 for a  $B^0$  right-sign tag and -1 for the opposite sign. Right sign is defined below for each class of decays. The appropriate dilution factor for  $D_i$  for this particular tag type and event characteristics is then calculated from the simulation. Eventually, these numbers will also be checked by the study of events in which both  $B$  mesons decay to tagging modes. For each event a combined flavor estimator is defined as follows: for events with more than one possible tagging identifier, the  $B_i$  are added like relativistic velocities ( $v/c$ ) to give a  $\hat{B}$ , such that  $-1 \leq \hat{B} \leq 1$ ; for events with only one identifier,  $\hat{B} = B_1$ . If the single estimators were statistically independent,  $\hat{B}$  would be equal to the effective dilution  $D_{eff}$  for that event. In fact, they are not fully independent, but one can measure  $D_{eff}(\hat{B})$  using doubly tagged modes [Wal95]. A cut on  $\hat{B}$  provides a simple way to select against poor tags using all the information in the event. A two-dimensional fit to the asymmetry against time and  $\hat{B}$  does not require such a cut and thus has slightly higher tagging performance. Other studies [Pla95] use a slightly different approach to combining all possible tagging information, directly defining an effective wrong tag fraction  $x_{eff}$  for each event rather than introducing  $\hat{B}$ .

### 3.7.1 Kaon Tags

The sign of a kaon's charge is a signature of its strangeness. The principal sources of  $K$ s from  $b$ -quark decay are:

- $b \rightarrow cW^-; c \rightarrow W^+s, s \Rightarrow K^-$ , (a)
- $b \rightarrow XW^-; W^- \rightarrow \bar{c}s, s \Rightarrow K^-$  and/or  $\bar{c} \rightarrow W^+\bar{s}, \bar{s} \Rightarrow K^+$ , (b)
- $b \rightarrow cW^-; W^- \rightarrow \bar{u}s, s \Rightarrow K^-$  and/or  $c \rightarrow XW^+, W^+ \rightarrow u\bar{s}, \bar{s} \Rightarrow K^+$ , (c)
- $b \rightarrow Xs\bar{s}; s \Rightarrow K^-$  and/or  $\bar{s} \Rightarrow K^+$ , (d)
- $b \rightarrow sq\bar{q}$  (*penguin*);  $s \Rightarrow K^-$ . (e)

We call the  $K$  from the dominant process (a) the right-sign kaon. Thus,  $Q_K = +1$  for a  $K^+$  and  $-1$  for  $K^-$ . Processes (b), (c), and (d) yield both right-sign and wrong-sign kaons, typically with a second strange particle in the event. Doubly Cabibbo-suppressed and penguin processes (e) give tiny contributions. Events with multiple like-sign kaons can occur since the processes listed above are not mutually exclusive. Wrong-sign particles contribute to the wrong-tag probability, but as the list above shows, other characteristics of the event can be used to reduce the wrong-tag fraction. Aside from the presence or absence of a second  $K$  in the event and its charge, the wrong-tag fraction also depends somewhat on the kaon momentum in the  $\Upsilon(4S)$  rest frame.

Kaons are identified via  $dE/dx$  by the DIRC in the barrel, and by aerogel Cherenkov counters in the forward endcap. This information is combined into a likelihood. Tracks with  $L_K > 0.5$  are counted as charged kaons in assigning event classes. All  $K$ s with  $L_K > 0.1$  contribute to the determination of  $\hat{B}$ . In this simulation, kaons which decay before reaching the DIRC do not contribute. This is a conservative assumption, since  $dE/dx$  information or the geometric signature of a kaon decay in flight may in fact lead to the recovery of good kaon identification in some cases.

On the basis of the simulations described above, the performance of all kaon tags (with or without leptons) is  $\epsilon_K D_{eff}^2 = 0.21$ . Here,  $\epsilon_K$  is the kaon tagging efficiency, that is, the fraction of generic  $B^0$  decays for which a kaon tag is available.

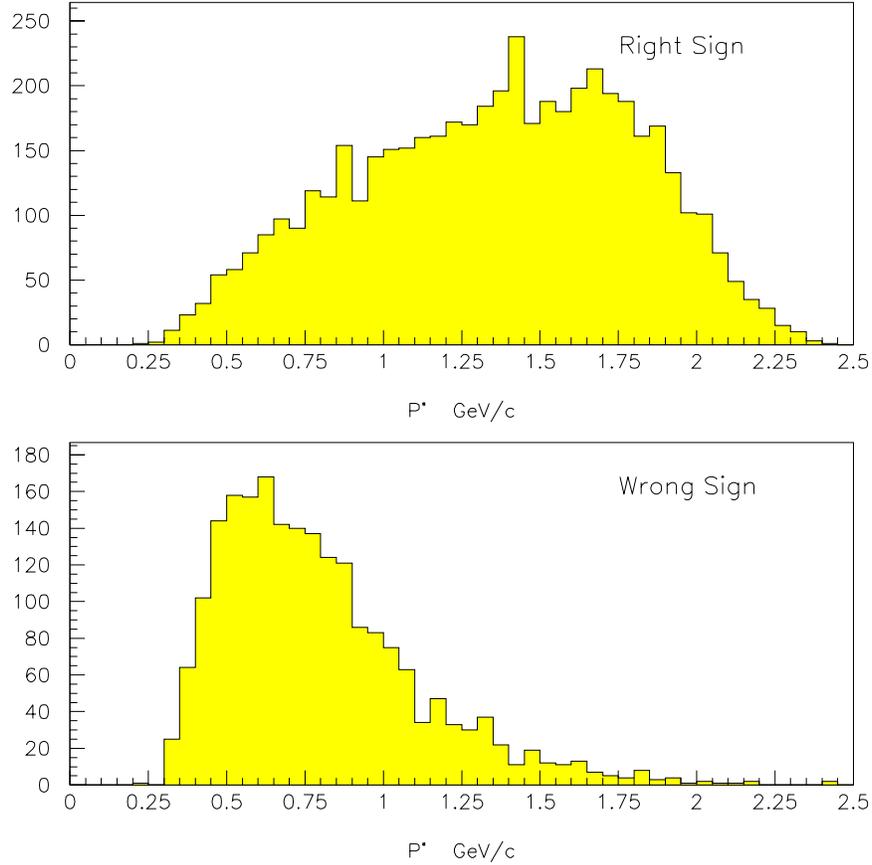
### 3.7.2 Lepton Tags

Electrons and muons in  $b$  decays come from several sources:

$$\begin{aligned}
 b &\rightarrow XW^-; W^- \Rightarrow l^-, (0.105e + 0.105\mu), & W^- \rightarrow \tau^- \Rightarrow l^- (\sim 0.005e + 0.005\mu), & \text{(a)} \\
 b &\rightarrow Xc; c \rightarrow XW^+, W^+ \Rightarrow l^+ (\sim 0.09e + 0.09\mu), & & \text{(b)} \\
 b &\rightarrow \bar{c}s; \bar{c} \rightarrow XW^-W^- \Rightarrow l^- (\sim 0.01e + 0.01\mu), & & \text{(c)} \\
 b &\rightarrow \bar{c}s; \bar{c}c \rightarrow J/\psi \text{ or } J/\psi', J/\psi \Rightarrow l^+l^- (\sim 0.003e + 0.003\mu). & & \text{(d)}
 \end{aligned}$$

The lepton from a semileptonic  $B$  decay, process (a), is defined as a right-sign lepton. Process (d) is a tiny source of spurious tags when one or the other lepton is missed. Each decay path leads to a distinct distribution in the lepton momentum, the number and invariant mass of the other accompanying particles, and the distribution of the remaining energy with respect to the lepton direction [Jaf94, Pia94]. All such information is used, when available, to determine the wrong-tag fraction. For example, Figure 3-4 shows the distributions of right- and wrong-sign leptons as a function of momentum,  $p^*$ , in the  $\Upsilon(4S)$  rest frame. At low momentum, wrong-sign leptons dominate and can contribute usefully to tagging.

Electrons are identified in the electromagnetic calorimeter and at low momenta by the DIRC and  $dE/dx$  measurements. Muons are identified in the IFR supplemented by DIRC at low



**Figure 3-4.** *Distribution of momentum in the  $\Upsilon(4S)$  rest frame for leptons. Right-sign leptons are  $l^+$  for  $B^0$  decays and  $l^-$  for  $\bar{B}^0$ .*

momenta. In low momentum regions, where the separation from hadrons is not perfect, each candidate is weighted by its likelihood, and cuts at  $L_e > 0.7$  and  $L_\mu > 0.7$  are applied.

We take the semileptonic branching fraction of the  $B^0$  meson to be 10.5%. Simulations of the lepton tagging performance give  $\epsilon_l D_{eff}^2 = 0.13$ . Here,  $\epsilon_l$  is the lepton tagging efficiency, that is, the fraction of signal events with an identified lepton tag. The performance number includes a cut  $|\hat{B}| > 0.4$ . A two-dimensional fit to the asymmetry, binning the data in time and in  $\hat{B}$ , gives slightly higher tagging performance,  $\epsilon_l D_{eff}^2 = 0.16$ .

Class	Selection Efficiency ( $\epsilon$ )	Efficiency ( $\hat{B} \geq 0.1$ )	Performance ( $\epsilon D_{eff}^2$ )
A	0.20	.19	0.13
B	0.80	.66	0.21
All	1.0	0.86	0.34

**Table 3-17.** Performance figures for lepton and kaon tags and for combined tagging using all events.

### 3.7.3 Other Tags

In addition to lepton and kaon tags, charged pions have a significant flavor correlation at the high end of the momentum spectrum due to two-body decays with a charged  $\pi$  or  $\rho$  meson, and at the low end due to pions from  $D^{*+} \rightarrow \pi^+ D^0$  [Sny90]. Additional  $D$  tags can be formed from neutral  $K$ s. Strange baryons, though rare, provide a small contribution to tagging. These effects are included in the analysis presented here. Ongoing studies continue to identify further ways of reconstructing events that can lead to a flavor assignment with a non-negligible  $D_{eff}$ . All such information will be used to maximize tagging performance.

### 3.7.4 Combined Tagging

All available information is used to calculate a combined tagging performance factor  $\hat{B}$  for each event, thus creating a two-dimensional distribution in time and the effective dilution  $D_{eff}(\hat{B})$  (or equivalently  $x_{eff}$ ) to be used in fitting for the asymmetry. We divide tagged events into two classes, class A being those events with a lepton which has a high probability of being a primary lepton from semileptonic  $B$  decay, and class B being all other events. This separation is necessary because the vertex resolution for the tagging decay is different for the two cases, and hence the contribution to a  $CP$ -asymmetry measurement must be determined separately for each case. More precisely, class A includes events containing either: (1) a lepton with  $p^* > 1.4$  GeV/ $c$ ; or (2) a lepton with  $p^* > 0.5$  GeV/ $c$ , and no other charged tracks or a total energy of less than 0.5 GeV within the same hemisphere in the  $\Upsilon(4S)$  rest frame. The results are given in Table 3-17. The efficiency values include misidentified particles, but the effective dilution factors correct for this. The quantity labeled “performance” is defined to be the sum over bins with  $\hat{B} \geq 0.1$  of  $\epsilon D_{eff}^2$ , where  $\epsilon$  is the fraction of signal events in the bin. If the lepton information alone were used for class A events, the performance would be only 0.125; the higher combined performance listed in the table shows the impact of the additional information on these events.

## 3.8 Estimate of $CP$ -Angle Measurement

---

### 3.8.1 Method of Calculation

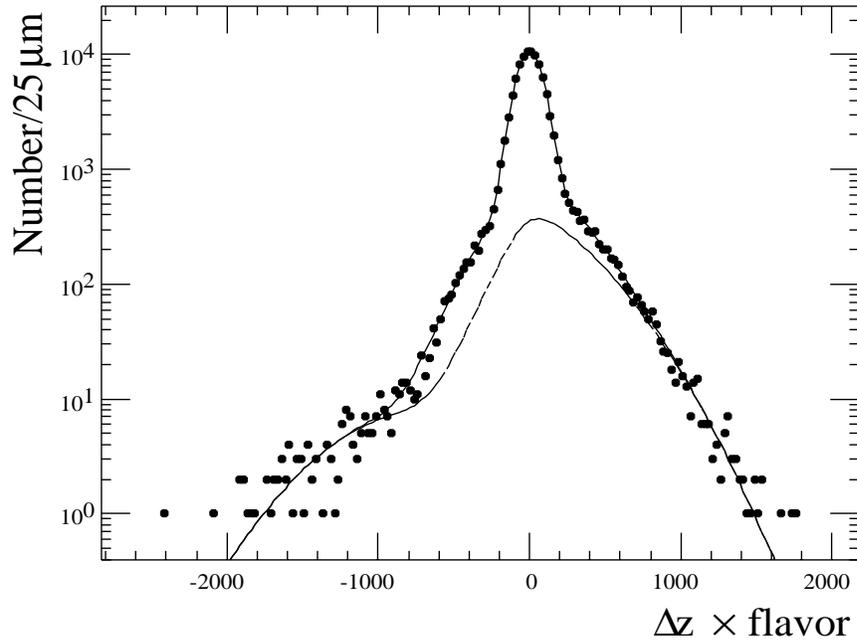
Using the approximation that both  $B$ s have a boost equal to  $\beta$ , namely the boost of the  $\Upsilon(4S)$ , the quantity  $\Delta z = z_{cp} - z_{tag}$  is related to the time difference between the two  $B$  decays by:

$$t = \frac{\Delta z}{\beta\gamma}.$$

Based on the simulations described above, each  $CP$  analysis channel (loosely called a  $CP$  mode in the following) and each tagging class, are characterized by efficiency, background (or dilution for tags), and  $z$  resolution. For each  $CP$  mode and tagging class, the distribution in time and  $b$  flavor is regenerated from these parameters. The program CPEXTRACT [Sny94] is then used to perform a binned maximum likelihood fit to this data and to make an estimate of the error on the asymmetry for a sample corresponding to  $30 fb^{-1}$  of  $\Upsilon(4S)$  data. This fit includes the dependence on time, tagging type, and background. Errors are quoted for zero  $CP$  asymmetry, since our sensitivity to small asymmetries is the quantity of interest. More details of the calculation are given in Reference [Sny95].

The  $\Delta z$  resolution of background events (such as those that occur in the  $\pi^+\pi^-$  and  $\rho\pi$  analyses) is treated separately in CPEXTRACT. This enhances our ability to discriminate against these backgrounds since the background events are produced at  $\Delta z = 0$ , and the resolution on  $\Delta z$  is better for background events than for signal plus tag events. The resolutions on  $\Delta z$  for  $\rho\pi$  backgrounds can be parameterized by two Gaussian distributions, with 90% of the events in a narrow component ( $\sigma_{narrow} = 80 \mu\text{m}$ ) and the remaining events in the wide component ( $\sigma_{wide} = 280 \mu\text{m}$ ). Furthermore, since the expected asymmetry behaves as  $\sin(K\Delta z)$ , where  $K = \delta M/\beta\gamma c$ , the impact of background on the fitted asymmetry parameters is minimized by correct treatment of the  $\Delta z$  dependence. This is illustrated by Figure 3-5 which shows the  $\Delta z$  distribution for  $B \rightarrow \rho\pi$  with a class B (mostly kaon) tag, an assumed asymmetry of 0.866 ( $\beta = 30^\circ$ ), and an assumed background-to-signal ratio of 9. As can be seen, the background drops significantly by the time the asymmetry develops. The improvement in  $CP$ -resolution obtained by explicitly fitting for the  $\Delta z$  dependence of the background is shown in Figure 3-6.

Here, wrong-sign tags are assumed to have the same  $\Delta z$  distribution as right-sign tags. This is a pessimistic assumption because, in general, the wrong-sign  $\Delta z$  distribution will be significantly broader, which will tend to reduce the adverse effects that they have on the time-dependent asymmetry. In the final analysis, they must be handled correctly, or a significant systematic error will be introduced.

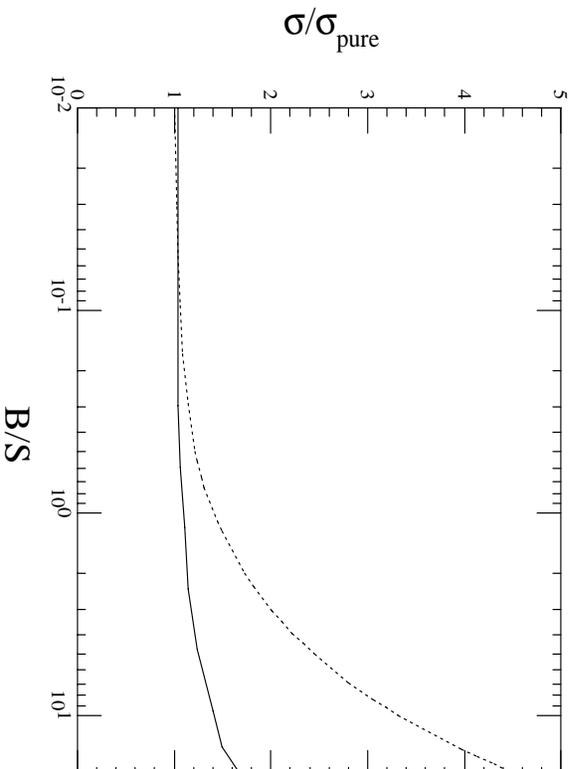


**Figure 3-5.** Distribution of  $\Delta z$  for the  $CP$ -state  $\rho\pi$  using class B tag and an assumed background-to-signal ratio of 9.

For lepton tagged events, there is a strong dependence of the  $\Delta z$  resolution on polar angle. This is accounted for by the simple expedient of dividing the direct lepton tag into two polar angle bins and treating each bin as an independent tag. Further refinement of this treatment can no doubt be used to improve accuracy in the actual data analysis.

### 3.8.2 Tagging Modes

For tagging efficiencies, we use the numbers given in Table 3-17 above. For both class A and class B tags, we estimate the  $z$  coordinate of the tag by fitting a common vertex to all tracks except those associated with the  $CP$ -state and those with transverse impact parameter greater than 1 mm. This last effectively excludes the decay products of kaons, lambdas, *etc.* Since the  $z$  resolution depends strongly on the direction of the tagging track [Cha95], the tags are divided into seven bins based on the tangent of the dip angle ( $|\tan\theta_{dip}|$ ) in order to optimize our  $CP$ -sensitivity. The numbers used in this document were obtained from a high statistics ASLUND simulation. Details can be found in [Wea95]. BBSIM simulations [Cha95] confirm these results for lepton tags.



**Figure 3-6.** Ratio of  $\sigma_{cp}$  in the presence of background to  $\sigma_{cp}$  obtained without background as function of the background-to-signal ( $B/S$ ) ratio. The solid curve corresponds to fitting for the  $\Delta z$  dependencies of the background and the signal. The dashed curve results if  $\Delta z$  information is not exploited.

### 3.8.3 $CP$ Modes

The parameters required to determine sensitivity (branching fractions, number of events expected, backgrounds, and  $z_{CP}$  resolutions) for each  $CP$ -mode under consideration are compiled in Table 3-18. These results are collected from the studies described above. The expected number of events before tagging is based on  $30fb^{-1}$  collected on the  $\Upsilon(4S)$ . The  $z$  resolutions for each mode are characterized by three numbers:  $\sigma_{narrow}$ ,  $\sigma_{wide}$ , and the ratio of the area of the narrow Gaussian to the wide one. For  $D^+D^-$ , many decay modes contribute, and events are grouped into two classes, four-prong and six-prong decays. A single Gaussian then gives an adequate parameterization of the  $z$  resolution of  $D^+D^-$  events in each class. Similarly, for  $D^{*+}D^{*-}$  decays, events are classified by the number of  $\pi^0$ s in the final state, and then a single Gaussian suffices to characterize the  $z$  resolution of each class.

For  $J/\psi K_L$  the background consist of two components: a  $CP$  symmetric part (the first number) and a  $CP$  anti-symmetric part given in parenthesis. For all modes with background the number in square brackets gives the degradation in  $\sigma_{cp}$  that results from the background. Our procedure of estimating the  $CP$ -violating asymmetry by dividing the data into samples with varying background levels and fitting explicitly to the  $\Delta z$  distribution of the background

Mode	Branching Fraction	Usable Sample	Efficiency	Reconstructed Number	Background	$z$ Resolution
$J/\psi K_S^0$	$0.5 \times 10^{-3}$	2160	0.41	1106	negligible	36/111/16
$J/\psi K_L^0$	$0.5 \times 10^{-3}$	2160	0.33	712	191(96)[1.3]	36/111/16
$J/\psi K^*$	$1.6 \times 10^{-3}$	788	0.39	307	negligible	37/155/10
$D^+ D^-$	$6 \times 10^{-4}$	993	0.25	248	109 [1.02]	52(6-prongs) 75(4-prongs)
$D^{*+} D^{*-}$	$7 \times 10^{-4}$	3130	0.155	485	62 [1.005]	43 (0 $\pi^0$ ) 51 (1 $\pi^0$ ) 69 (2 $\pi^0$ )
$\pi^+ \pi^-$	$1.2 \times 10^{-5}$	432	0.80	346	1957 [1.11]	34/161/4.4
$\rho^\pm \pi^\mp$	$5.8 \times 10^{-5}$	2088	0.56	1162	33420 [1.18]	36/240/24

**Table 3-18.** Simulated efficiencies, backgrounds, and vertex resolutions for  $CP$  modes. The usable sample reflects the visible branching ratio and the  $30 \text{ fb}^{-1}$  sample size. Reconstructed number of events does not include tagging efficiencies. The background number given is for the total sample; in the analysis the data are divided into bins of varying background to minimize the adverse impact on  $CP$  resolution. The factor by which background actually degrades the resolution is given in square brackets.

when estimating the  $CP$ -violating asymmetry gives a substantial improvement in sensitivity over a naive treatment which uses a factor of  $\sqrt{1 + B/S}$ .

The resolution on the  $CP$ -violating phase ( $\phi_M - \phi_D$ ) for each  $CP$  state is given in Table 3-19. The combined sensitivity for  $\sin 2\alpha$  and  $\sin 2\beta$  is estimated to be  $\pm 0.059$  and  $\pm 0.085$ , respectively, for  $30 \text{ fb}^{-1}$ , or one year's data at nominal luminosity. The channels  $J/\psi K^*$ ,  $D^* D$ ,  $D^* D^*$ ,  $\rho\pi$ , and  $a_1\pi$ , which are not pure  $CP$  eigenstates, are assumed to be dominated by a single  $CP$  eigenstate. If this is not the case, the resolution obtained with these modes will be reduced. An accuracy on the  $CP$  phase comparable to the  $DD$  modes can be expected. The results in this table differ in many ways from those given in the *Letter of Intent*. The differences arise from two major sources. First, our knowledge of the branching ratio to the modes of interest has improved somewhat, resulting in changes in these key input numbers. Second, more complete simulation work, including backgrounds, has been carried out for most of the channels included in Table 3-19, resulting in changed efficiencies and signal-to-background ratios. The number given for the  $a_1\pi$  channel is an estimate based on the work of Reference [Ale91]. The number obtained in that paper has been rescaled, reflecting the integrated luminosity assumed here, using the tagging efficiency expected for this detector, and correcting the assumed branching fraction for  $B \rightarrow \pi\pi$ . For the  $D^* D$  channel, the estimate is based on the simulations of the  $DD$  and  $D^* D^*$  channels and the measured  $B \rightarrow D_s D^*$  branching ratios.

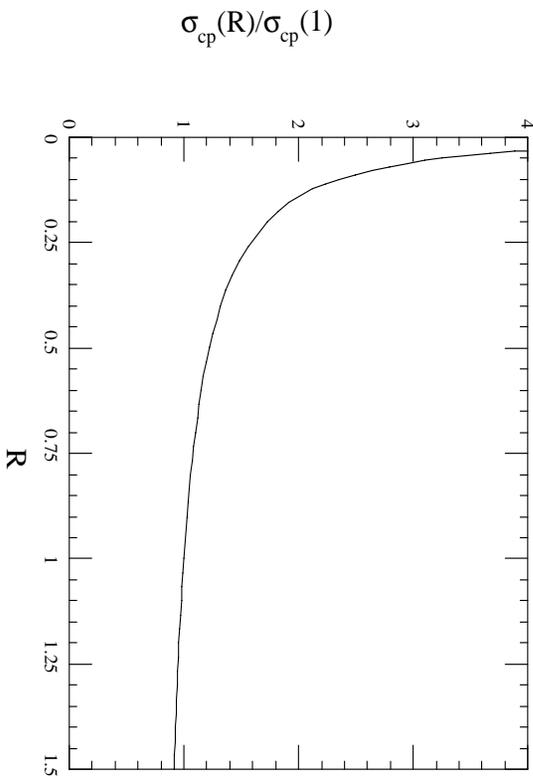
CP State	Br	$\phi_M - \phi_D$	A Tags	B Tags	Combined
$J/\psi K_S^0$	$0.5 \times 10^{-3}$	$\beta$	0.15	0.13	0.098
$J/\psi K_L^0$	$0.5 \times 10^{-3}$		0.25	0.21	0.16
$J/\psi K^*$	$1.6 \times 10^{-3}$		0.29	0.25	0.19
$D^+ D^-$	$6 \times 10^{-4}$	$\beta$	0.32	0.27	0.21
$D^{*+} D^{*-}$	$7 \times 10^{-4}$		0.24	0.20	0.15
$D^{*\pm} D^\mp$	$8 \times 10^{-4}$				0.15*
Combined		$\beta$			0.059
$\pi^+ \pi^-$	$1.2 \times 10^{-5}$	$\alpha$	0.29	0.27	0.20
$\rho\pi$	$5.8 \times 10^{-5}$		0.16	0.16	0.11
$a_1\pi$	$6 \times 10^{-5}$				0.24*
Combined		$\alpha$			0.085

**Table 3-19.** Error on the measurement of  $(\phi_M - \phi_D)$  from various channels for  $30 \text{ fb}^{-1}$ . The A-tag sample is chiefly direct leptons, and the B sample is all other tags. The numbers marked \* ( $a_1\pi$  and  $D^*D$ ) are estimates; all others are based on simulations.

The accuracy achievable for the  $\rho\pi$  channel depends on the ratio  $R = (B^0 \rightarrow \rho^+\pi^-) / (B^0 \rightarrow \rho^-\pi^+)$  [Ale91, Ale93]. The results given in Table 3-19 assume  $R = 1$ . The dependence of the sensitivity on this assumption is shown in Figure 3-7.

### 3.9 CKM Matrix Determination

The values of the CKM matrix elements determine the lengths of the sides of the unitarity triangle. Accurate measurements of these elements are thus an important ingredient in overconstraining the triangle and probing the source of  $CP$  violation. Furthermore, the values of the CKM elements are fundamental input parameters within the Standard Model and hence, must be measured as precisely as possible; better knowledge of these parameters may provide some insight into their origin. The present status of these elements is summarized in Figure 3-8 in the Wolfenstein parameterization, where the shaded area is that allowed in the Standard Model for the apex of the unitarity triangle in the  $(\rho, \eta)$  plane. The parameters used for the plot of the allowed region for the unitarity triangle are taken from the draft of a Review of Modern Physics article by P. Burchat and J. Richman [Ric95]. The two directly-measured CKM parameters are  $|V_{cb}| = 0.040 \pm 0.003$  and  $|V_{ub}/V_{cb}| = 0.076 \pm 0.026$ . The value of the parameter that describes  $B_d^0$ - $\bar{B}_d^0$  mixing is  $\Delta m(B^0) = x_d/\tau_{B^0} = 0.43 \pm 0.06 \text{ ps}^{-1}$ . The top quark mass evaluated at the relevant scale for this mixing is  $m_t = 165 \pm 15 \text{ GeV}$ . The hadronic matrix elements obtained in lattice calculations are  $f_B\sqrt{B_B} = 200 \pm 40 \text{ MeV}$ , and  $B_K = 0.75 \pm 0.05$ . The resulting errors are added in quadrature, and the lines shown



**Figure 3-7.** *Dependence of the error on the CP asymmetry angle  $\alpha$  on the assumed value for the ratio of decay widths  $R = \Gamma(B^0 \rightarrow \rho^+ \pi^-) / \Gamma(B^0 \rightarrow \rho^- \pi^+)$ , normalized to the case  $R = 1$ .*

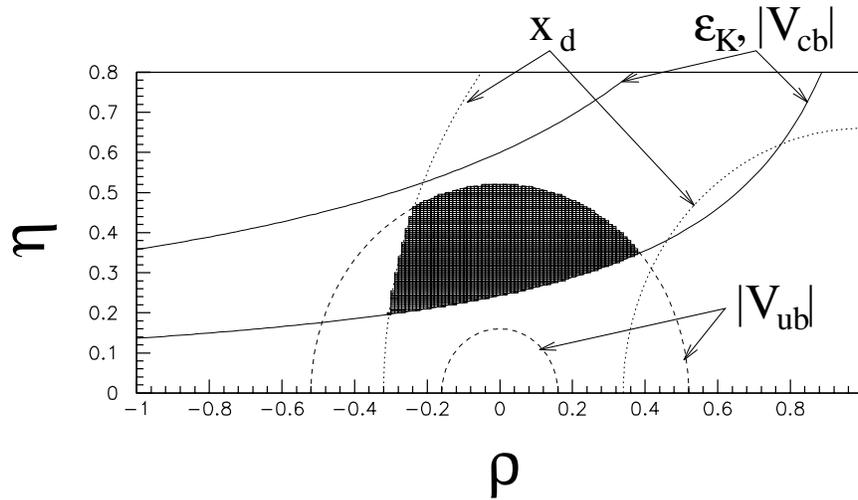
in the plot correspond to  $\pm 1.5\sigma$  limits. The overlapping limits result in the following set of allowed ranges for the unitary angles:

$$\begin{aligned}
 -0.59 &\leq \sin 2\alpha \leq 1.0 \\
 0.29 &\leq \sin 2\beta \leq 0.88 \\
 -1.0 &\leq \sin 2\gamma \leq 1.0.
 \end{aligned}
 \tag{3.17}$$

It is important to remember that this picture represents the Standard Model and can be dramatically altered if new physics is present. Examples include new contributions to  $\epsilon$  or  $B_d^0$ - $\bar{B}_d^0$  mixing, even if no new sources of CP violation are present. It is equally important to note that, since the range of values of parameters such as  $B_K$  and  $B_B$  is constrained by theory rather than experiment, the overlap region is a best estimate rather than a 90% confidence level result.

### 3.9.1 $V_{cb}$

Values for  $V_{cb}$  are extracted from the investigation of inclusive and exclusive semileptonic  $B$  decays. CLEO and LEP have determined  $V_{cb}$  at the 7.5% level [Ric95], depending on the



**Figure 3-8.** Constraints in the Standard Model in the  $\rho\eta$  plane. The shaded area corresponds to that allowed for the apex of the rescaled unitarity triangle.

extraction technique, and  $B_{\text{BABAR}}$  can expect to reduce these errors to the few percent level due to the expected large data sample and efficient reconstruction techniques. The inclusive semileptonic branching fraction  $B_{SL}$  can be determined from:

1. Measurement of the inclusive single-lepton momentum spectrum. This technique yields significant data samples, but the procedure used to fit the observed spectrum to the expected shape for primary and secondary leptons from  $B$  and charm decay, respectively, introduces a large model dependence.
2. Charge and angular correlations in dilepton events. This offers less model dependence, as the measured correlations can be used to separate the primary and secondary lepton spectra, instead of relying on theory.
3. Separate measurement of  $B_{SL}$  for charged and neutral  $B$  meson decay. Here, one  $B$  in the event must be reconstructed in order to tag the charge of the other; this will require the large tagged event sample available to  $B_{\text{BABAR}}$ .

Determination of  $V_{cb}$  from  $B_{SL}$  via technique (1) at CLEO and LEP is already dominated by the theoretical error, while method (2) still offers room for improvement on the systematic and statistical errors. In addition, it is important to make better measurements of  $B_{SL}$  for the charged and neutral  $B$  mesons separately. These are used to determine the ratio of  $B$  lifetimes, which in turn are used in interpreting exclusive semileptonic and hadronic decay rates. In these measurements, one  $B$  in the event is reconstructed in order to tag the charge of the other; this will benefit from the large tagged event sample available to  $B_{\text{BABAR}}$ .

Exclusive semileptonic decays offer a reliable model-independent determination of  $V_{cb}$  within the framework of heavy quark effective theory (HQET). The heavy quark symmetry allows the normalization of the  $q^2$ -dependent hadronic form factors with good precision at zero recoil for the charm hadron system. This technique is best suited for the process  $B \rightarrow D^* \ell \nu_\ell$  as the leading corrections to the HQET result arise only at higher order,  $1/m_Q^2$ . Due to the presence of the neutrino, reconstruction of exclusive semileptonic channels is more difficult than for hadronic decays; at present, the  $B \rightarrow D^* \ell \nu_\ell$  mode is the most precisely measured. One technique at the  $\Upsilon(4S)$  is to determine the missing momentum distribution recoiling against the  $D^* \ell$  system. The zero recoil configuration for the charm hadrons results in a reduction of phase space for the decay (as the  $B$  and  $D^*$  mesons have approximately equal velocities), yielding a statistically limited data sample, and requires good detection efficiency for low momentum pions arising from the  $D^*$  decay. Measurements of other exclusive modes, such as  $D \ell \nu_\ell$  and  $D^{**} \ell \nu_\ell$ , will also be precisely studied at BABAR.

### 3.9.2 $V_{ub}$

An accurate determination of  $V_{ub}$  is important to test closure of the unitarity triangle. Data near the end-point region of the lepton momentum spectrum in inclusive semileptonic  $B$  decays have established that  $V_{ub}$  is nonzero. However, converting the measured rate into a value of  $V_{ub}$  introduces substantial errors. This conversion is highly model dependent due to the small phase space available at the end-point, the large uncertainties in the calculation of the rates for the resonant modes, and the relative sizes of the contributions of resonant and nonresonant modes in this region. The subtraction of background from the small data sample is an additional large source of error. The present experimental error on the ratio  $|V_{ub}|/|V_{cb}|$  is comparable to the theoretical uncertainty, and thus new, less model dependent techniques for extracting this CKM element are necessary.

Measurement of exclusive semileptonic decays ( $B \rightarrow X_u \ell \nu_\ell$ , where  $X_u = \pi, \rho, \text{ or } \omega$ ) requires a large data sample and should provide an improved determination of  $V_{ub}$ . Such decays have yet to be observed, but model predictions indicate that the rate should be distributed over many exclusive channels, with no dominant modes. The theoretical uncertainties are expected to be lower than in the inclusive case. Further reductions in the theoretical errors can be obtained via measurements of the form factor  $q^2$  distributions in  $c \rightarrow d$  transitions such as  $D \rightarrow X_u \ell \nu_\ell$ . Given adequate statistics and understanding of experimental systematics from vertex separation and  $\pi/K$  identification, such form factor measurements can be made at BABAR.

An alternative method [Bar90] of extracting  $V_{ub}$  from semileptonic  $B$  decays is to measure the invariant mass spectrum of the final state hadrons *below* the charm hadron threshold,

*i.e.*,  $m_X < m_D$ . A study of this method at  $BABAR$  has yet to be made, but the advantages of separate vertexing of the two  $B$ s will probably be important for this technique to be feasible.

### 3.9.3 $V_{td}$

It has recently been shown [Atw94, Gol94, Des94] that the proposed method of obtaining the ratio  $|V_{td}/V_{ts}|^2$  from the ratio of branching fractions  $B(B \rightarrow \rho\gamma)/B(B \rightarrow K^*\gamma)$  is not valid due to the potentially large long-distance contributions to these exclusive  $B$  decay modes. The sizes of these long-distance contributions could be determined with a measurement of exclusive radiative charm decays, such as  $D \rightarrow \rho\gamma$ , which are expected to have no short-distance contamination. Separate measurements of  $B^- \rightarrow \rho^-\gamma$  and  $B^0 \rightarrow \rho^0\gamma$  as well as  $B_s \rightarrow \phi\gamma, K^{*0}\gamma$  may also provide an estimate of the long-distance rates.

A measurement of  $B_s^0$ - $\bar{B}_s^0$  mixing could also yield a value for this ratio of CKM elements as

$$\frac{x_s}{x_d} = \frac{m_{B_s} \eta_{QCD}^{B_s} B f_{B_s}^2 |V_{ts}|^2}{m_{B_d} \eta_{QCD}^{B_d} B f_{B_d}^2 |V_{td}|^2} = \xi_s^2 \frac{1}{\lambda^2 [(1 - \rho)^2 + \eta^2]}. \quad (3.18)$$

The factor  $\xi_s$ , which multiplies the ratio of CKM elements, measures the amount of  $SU(3)$  breaking effects and is estimated to be 1.3. Determination of  $f_{B(s)}$  from the decay  $B_{(s)} \rightarrow \tau\nu_\tau$  will greatly reduce the theoretical uncertainty in this mixing ratio. The allowed region of the  $\rho\eta$  plane, determined above in the Standard Model, yields  $x_s = 10$ –57. A precise measurement of  $B_s$  mixing requires a substantial run on the  $\Upsilon(5S)$ . It is needed in order to obtain a clean determination of  $|V_{td}|/|V_{ts}|$ . If  $V_{ts}$  is relatively large, a more accurate value of  $|V_{td}/V_{ts}|^2$  could be obtained from a measurement of  $\Delta, /$ , for the  $B_s$  meson [Bro95].

## 3.10 Rare $B$ Decays

---

Rare  $B$  decays are an important testbed of the Standard Model. They offer a complementary strategy in the search for new physics by probing the indirect effects of new interactions in higher order processes. In particular, the probing of loop-induced couplings can provide tests of the detailed structure of the Standard Model at the level of radiative corrections, where the Glashow-Iliopoulos-Maiani cancellations are important. The recent observation of radiative  $B$  decays by CLEO has already provided bounds on the CKM ratio  $|V_{ts}/V_{cb}|$  as well as powerful constraints on new physics. The high luminosity and separated  $B$  vertices available at PEP-II and improved particle identification capabilities of  $BABAR$  will be needed to allow accurate measurements of many rare modes. No full detector study has yet been

made for any of these channels. However, *BABAR* is designed to give as complete information on all tracks as can reasonably be achieved, and hence, is well-suited to extend our knowledge of these modes.

### 3.10.1 $B \rightarrow \tau\nu$

The decay  $B \rightarrow \tau\nu$  is of interest because the branching ratio will yield a value of the product  $f_B^2|V_{ub}|^2$ . The general expression for the decay width for the decay  $B \rightarrow \ell\nu$  is

$$\Gamma(B^+ \rightarrow l^+\nu) = \frac{G_F^2}{8\pi} f_B^2 m_B m_l^2 |V_{ub}|^2 \left(1 - \frac{m_l^2}{m_b^2}\right)^2,$$

where  $G_F$  is the Fermi constant,  $f_B$  is the  $B$ -meson coupling constant,  $m_B$  is the mass of the  $B$  meson,  $m_l$  is the lepton mass, and  $V_{ub}$  is the CKM-matrix element. For the decay  $B \rightarrow \tau\nu$ , the branching ratio is about  $8 \times 10^{-5}$ .

Leptonic decays of the  $B$  meson are expected to be best identified by fully reconstructing one charged  $B$  meson (the tagging  $B$ ) in events in which the only observed particle coming from the decay of the second  $B$  meson is a charged lepton. The separation of the decay vertices of the two  $B$  mesons, along with the excellent mass resolution of the detector, allows an unambiguous identification of the tagging  $B$  meson. The problem then is to reject decays of the nontagging  $B$  mesons which are not purely leptonic but in which only a single lepton is identified in the detector. A classic example of this type of decay is the semileptonic decay  $B \rightarrow \ell K_L^0 \nu$ , in which neither the  $\nu$  nor the  $K_L^0$  are detected. The *BABAR* detector needs to have excellent muon and  $K_L^0$  detection and rejection capabilities in order to veto these types of decays with high efficiency.

For the decay  $B \rightarrow \mu\nu$ , the branching ratio is about  $4 \times 10^{-7}$ , a factor of 200 smaller than that for  $B \rightarrow \tau\nu$ . However,  $B \rightarrow \mu\nu$  decays should be much easier to detect and discriminate from backgrounds because they are two-body decays, and hence the muons are close to monoenergetic. Further Monte Carlo simulation work needs to be done before a prediction can be made of the level of sensitivity of the *BABAR* detector to  $B \rightarrow \ell\nu$  decays and to determine whether  $\tau\nu$  or  $\mu\nu$  searches give better sensitivity. However, it is already clear that the general design of the detector and the separate vertexing of the two  $B$ s make *BABAR* competitive with any other experiment for this search.

### 3.10.2 $B \rightarrow X_s \ell^+ \ell^-$

This decay receives short-distance contributions from electromagnetic and  $Z^0$  penguin diagrams as well as  $W$  box diagrams, and long-distance contributions from the process  $B \rightarrow K^{(*)} \psi^{(\prime)}$  followed by  $\psi^{(\prime)} \rightarrow \ell^+ \ell^-$  and from  $c\bar{c}$  continuum intermediate states. The short distance contributions lead to the inclusive branching fractions  $B(B \rightarrow X_s \ell^+ \ell^-) = (15, 7, 2) \times 10^{-6}$  for  $\ell = e, \mu, \tau$ , and will likely be observed before PEP-II is operational. However, the best method of separating the long and short distance contributions, and observing any small deviations from the Standard Model, is to measure the various kinematic distributions associated with the final-state lepton pair, such as the lepton pair invariant mass distribution, the lepton pair forward-backward asymmetry, and the tau polarization asymmetry. Measurement of these distributions requires the high statistics samples available at PEP-II. Here, the long-distance contributions dominate only in the  $m(\ell^+ \ell^-)$  regions near the  $J/\psi$  and  $\psi'$  resonances, and observations away from these peaks cleanly separate the short distance physics. The values of the forward-backward and tau polarization asymmetries are large for currently favored values of the top quark mass around  $175 \text{ GeV}/c^2$  [Abe94]. Measurement of all three kinematic distributions would provide a unique determination of the sign and magnitude of each contributing short distance operator, and hence, would provide a stringent test of the Standard Model.

### 3.10.3 $B \rightarrow X_s \nu \bar{\nu}$

The decay  $B \rightarrow X_s \nu \bar{\nu}$  proceeds through  $Z^0$ -penguin and  $W$ -box diagrams yielding an inclusive branching fraction of  $5 \times 10^{-5}$ . Measurement of this process would help to determine the size of  $Z^0$ -penguin effects in rare  $B$  hadronic decays, such as  $B \rightarrow \pi\pi$ . A search for such decays will focus on events with large missing  $E_T$  opposite a reconstructed  $B$ . High integrated luminosity, good detector hermeticity, and efficient  $K_L$  identification for background rejection are important for this mode. The search will require at least  $10^5$  reconstructed  $B$  mesons. The reduction in continuum background problems offered by the separate vertexing of the two  $B$ s in an event will be a significant advantage for this channel.

## 3.11 *CP* Asymmetries in Charged *B* Decays

---

*CP* asymmetries in charged  $B$  decays are another important and as-yet-unexplored region of physics. While these asymmetries offer no direct measurement of CKM parameters, unlike the neutral  $B$  case, their observation is nonetheless essential to our full understanding of the physics of  $B$  decays. Any observed asymmetry in charged  $B$  decays is a direct *CP*-violation

effect and would, for example, rule out the superweak alternative theory of  $CP$  violation. These asymmetries require interference between two amplitude contributions that have both different weak phases and different strong phases. In the Standard Model they are expected to be largest in a channel such as  $K\pi$  in which tree amplitudes are CKM suppressed relative to penguin contributions, and hence the relative strengths of the two types of terms are comparable. Unknown strong phases make even this statement a crude estimate rather than a rigorous prediction. However, if factorization proves to be a good approximation, then, within the Standard Model, this expectation for  $K\pi$  becomes quite reliable. Of course, the total rate for such a channel is also suppressed, so it remains to be seen whether it will be easier to measure a small asymmetry in a relatively high rate channel or a larger asymmetry in a rare decay mode.

There are aspects of detector design of particular importance. Just as with the neutral channels, issues such as particle identification for distinguishing  $K^\pm\pi^0$  from  $\pi^\pm\pi^0$  decays, are crucial to any such search. Clearly, one also needs to be able to detect  $\pi^0$ s and other calorimeter signatures to perform such studies. Likewise, good detector resolution is needed to allow cuts that eliminate light-quark-produced backgrounds without excessive deterioration in signal. Thus, a detector designed to perform well in the neutral  $B$  case is also well-adapted for a study of charged  $B$  decays and a search for possible  $CP$ -violating effects in these channels.

## 3.12 Charm Physics

---

$BABAR$ 's expected high  $D$  meson production rates, good time resolution between decays, particle identification systems, and near hermetic efficiency should make high sensitivity searches into  $D\bar{D}$  mixing, rare decay physics, and  $CP$  violation in the charm sector feasible.

### 3.12.1 $D^0\bar{D}^0$ Mixing

Within the Standard Model,  $D^0\bar{D}^0$  mixing is expected to be very small. The amount of mixing is governed by  $x_{mix} = \Delta m/\Gamma$ , and  $y_{mix} = \Delta\Gamma/\Gamma$ , where  $\Delta m$  is the mass difference between  $D^0 - \bar{D}^0$  states, and  $\Gamma$  is the decay rate. In the Standard Model, estimates for  $x_{mix}$  and  $y_{mix}$  are in the range  $10^{-3}-10^{-5}$  [Liu94]. If mixing is enhanced, this would be a sign of new physics. Unless the  $D\bar{D}$  mixing rate is exceptionally large, it will not be significantly tested in the near future.  $BABAR$  can play an important role in placing a limit on the experimental sensitivity of  $D\bar{D}$  mixing at about  $x_{mix} \sim 10^{-2} - 10^{-3}$ , a range of interest for some nonstandard model predictions [Hal92].

The experimental limitations on mixing sensitivity come primarily from two sources, doubly-Cabibbo-suppressed decays (DCSD), which mimic mixing in pure hadronic channels, and experimental backgrounds. The mixing term behaves as  $\sin(\Delta mt)$ . Thus, with small  $x_{mix}$ , one looks for mixing with optimum sensitivity at or beyond two decay lifetimes. Even after a few lifetimes, one is sensitive to fluctuations in the tails of the DCSD rate. In addition, experimental backgrounds can result from combinatorial counting, poor vertex reconstruction, particle misidentification, and confusion with other long-lived decay products. Such backgrounds can be reduced by vertex separation cuts, particle identification requirements, and rejection of mass reflections, *etc.*

Either resonant or continuum production of  $D^*$ s can be used for mixing studies in  $BABAR$ . Commonly used decay chains such as  $D^{*+} \rightarrow D^0 \pi^+$ , then  $D^0 \rightarrow K^- \pi^+$ ,  $K^- \pi^- \pi^+ \pi^-$ ,  $K^- l \bar{\nu}$ , and even  $K^- \pi^+ \pi^0$ , require measurement of the sign of the  $\pi$  from the  $D^*$  decay to define the initial  $D^0$  flavor. In the boosted system of  $BABAR$ , this  $\pi$  usually has enough momentum to be reconstructed with high efficiency. The good particle identification properties of  $BABAR$  are important in analyzing such channels. The  $BABAR$  vertex detector will have adequate proper time resolution to resolve the decay time development of the DCSD from that of decay after mixing.

With sufficiently high statistics samples, searches via semileptonic decays may ultimately provide better mixing limits. Semileptonic decay channels have no DCSD contributions, but experimental backgrounds due to incomplete reconstruction of the fully leptonic final state can dominate the measurement error.

### 3.12.2 $CP$ Violation in Charm Decays

$CP$ -violating effects are also predicted to be very small in the charm sector, although some have argued that they may be significantly enhanced without requiring physics beyond the Standard Model [Gol89, Buc93]. Again, with high reconstruction efficiencies for charm decays and efficient tagging mechanisms,  $BABAR$  will be sensitive to observation of  $CP$ -violating effects in the charm sector.

In order for  $CP$  violation to occur, there must exist two transition amplitudes to a final-state  $f$  which interfere with nonzero relative phase. In the Standard Model, neutral  $D$  decay  $CP$ -violating asymmetries from mixing are suppressed by the low mixing rate and are not expected to be seen. However, direct  $CP$ -violating effects in Cabibbo-suppressed modes, due to interference between the suppressed tree amplitude and an unsuppressed penguin amplitude, may be as high as a few times  $10^{-3}$  [Buc93, Fry93]. Strong candidates for searches include  $D^+ \rightarrow \bar{K}^{*0} K^+$ ,  $D_s^+ \rightarrow K^{*+} \eta'$  decays, and  $D^0 \rightarrow K^+ K^-$ . Limits set by current experiments [Fra94] have not tested the  $A_{CP} \sim 10^{-3}$  range, and  $BABAR$  should be able to make major contributions.

### 3.13 Tau Physics

---

Tau physics is now precision physics. There are no longer any obvious discrepancies between experimental tau physics and Standard Model predictions: lepton universality has been confirmed to better than 1%; world average exclusive branching fractions sum to within 1% of 100%. Further progress in tau physics will come from high precision, detailed studies of both leptonic and semihadronic tau decays, studies of tau spin structure, and searches for rare or forbidden decays. Taus are still a rich field of physics, and the potential still exists for uncovering physics beyond the Standard Model.

The cross section for tau pairs at  $E_{cm} = 10.58$  GeV is 0.91 nb, so an integrated luminosity sample of  $30 \text{ fb}^{-1}$  will contain  $27 \times 10^6$  tau pairs. (The current total world sample is on the order of  $6 \times 10^6$  tau pairs, dominated by CLEO.) With such a large sample, significant contributions can be made to all of the topics discussed below. Many studies will be dominated by systematic errors, so it is important to critically evaluate the strengths and weaknesses of *BABAR* for addressing these issues.

Tau physics is more difficult at *BABAR*, than at LEP because: at *BABAR* there will be significant backgrounds from  $q\bar{q}$  events and two-photon physics, which require relatively hard, mode-dependent cuts to isolate a signal;  $\pi^0$ s are softer and therefore more difficult to detect reliably; and electrons and muons are softer and are thus more difficult to distinguish from pions. However, the extremely large data sample will make it possible to trade statistics for systematics, for example, in the use of clean tags such as  $\tau \rightarrow e\nu\nu$ , so that improved systematic errors can be achieved. More importantly, the scope of interesting physics goals will be more focused; *BABAR* will concentrate on some important measurements that can be done very well with its large data sample. Some of the topics that will be of interest in the twenty-first century, to which *BABAR* can make a significant contribution, are discussed below.

- Second-class currents (which violate isospin) have not yet been observed in tau decay. The decay  $\tau^- \rightarrow \eta\pi^-\nu_\tau$  is predicted to proceed at the  $10^{-5}$  level. The decay  $\tau^- \rightarrow \omega\pi^-\nu_\tau$  occurs predominantly through the vector current; the decay via the axial-vector current, which violates isospin, can be observed by analyzing the angular distribution.
- Detailed studies of tau decays to three pseudoscalar mesons ( $\pi\pi\pi\nu_\tau$ ,  $K\pi\pi\nu_\tau$ , and  $KK\pi\nu_\tau$ ) will yield information on the Weiss-Zumino anomaly, isospin violation,  $SU(3)_f$  breaking,  $K_1$  mixing, and other aspects of these hadronic currents.
- Detailed studies of the spectral functions in decays to  $3\pi\nu_\tau$ ,  $4\pi\nu_\tau$ ,  $5\pi\nu_\tau$ , and  $6\pi\nu_\tau$ , in comparison with data from  $e^+e^-$  annihilation, permit precise tests of CVC and PCAC.

- Studies of the spectral functions, carefully separated into vector and axial vector components, and strange and nonstrange components, permit tests of Weinberg, DMO, and other sum rules.
- Precision measurements of the Michel parameters  $\rho$ ,  $\eta$ ,  $\xi$ , and  $\delta$  in leptonic tau decays and of the parity-violating neutrino helicity in semihadronic decays test the V–A Lorentz structure of the  $\tau - W^- - \nu_\tau$  coupling, and are thus sensitive to the presence of scalar (charged Higgs) or V+A ( $W_R$ ) currents. The spin-dependent Michel parameters  $\xi$  and  $\delta$  can be measured by exploiting the well-understood (in QCD) correlations between tau spins in  $e^+e^-$  annihilation through the analysis of momentum and angle correlations between the decay products of the two taus in the event.
- One can search for  $CP$  violation in the leptonic sector in a variety of ways. A nonzero electric dipole moment will increase the total  $e^+e^- \rightarrow \tau^+\tau^-$  cross section and angular distribution and affect the energy distribution of the  $\pi^-$  in the decay  $\tau^- \rightarrow \pi^-\nu_\tau$ . It will also produce nonzero expectation values for  $CP$ -odd correlation tensors between the decay products of the two taus in the event.
- In tau decays, final states containing no neutrinos violate lepton family number. Limits on branching fractions for these processes are at the few  $10^{-6}$  level. Some models, with massive Majorana neutrinos or mass-dependent couplings, predict branching ratios near this level. It is thus of interest to push these limits down by another order of magnitude.
- The current best direct bound on the tau neutrino mass is around  $25 \text{ MeV}/c^2$ . Constraints from Big Bang nucleosynthesis and SN-1987A exclude long-lived Dirac tau neutrinos below 15 keV. However, if the tau neutrino has a mass between 1 and  $25 \text{ MeV}/c^2$  and a lifetime in the range  $10^3$ – $10^6$  seconds, it can evade nucleosynthesis bounds, contribute to dark matter, and help explain how type II supernovae explode. It is thus of great interest to search for a tau neutrino mass in this range.

## 3.14 Two-Photon Physics

---

Monte Carlo studies [Bau92] have shown that a high-luminosity  $B$  Factory will extend the study of exclusive two-photon physics from the present  $2 \text{ GeV}/c^2$  up to at least  $5 \text{ GeV}/c^2$  in mass. This will likely reveal many of the exotic bound-states now being sought, such as glueballs,  $q\bar{q}g$  hybrids, and four-quark resonances. Measurements of the two-photon couplings of light quark and charmonium resonances will continue to be very important for understanding quark dynamics. Furthermore, a real challenge to perturbative QCD predictions of exclusive hadron production from two-photon interactions can be made by

Untagged Reactions	Published Data	Expected $B$ Factory ( $100 \text{ fb}^{-1}$ )
$\gamma\gamma \rightarrow \eta' \rightarrow \pi^+\pi^-\gamma$	2000	300,000
$\gamma\gamma \rightarrow \iota(1.42 \text{ GeV}) \rightarrow \bar{K}K\pi$	0	600
$\gamma\gamma \rightarrow \eta_c \rightarrow \text{all modes}$	10	30000
$\gamma\gamma \rightarrow \pi^+\pi^- (W > 3 \text{ GeV})$	10	800

Single-Tagged Reactions	Published Data	Expected $B$ Factory ( $100 \text{ fb}^{-1}$ )
$\gamma\gamma^* \rightarrow \eta' \rightarrow \pi^+\pi^-\gamma$	300	6000
$\gamma\gamma^* \rightarrow 1^{++}(1.42 \text{ GeV}/c^2) \rightarrow \bar{K}K\pi$	30	1100
$\gamma\gamma^* \rightarrow \eta_c \rightarrow \text{all modes}$	1	3000
$\gamma\gamma^* \rightarrow \pi^+\pi^- (W > 3 \text{ GeV}/c^2)$	—	100

**Table 3-20.** Comparison of current published event samples for two-photon physics with the number of events expected at BABAR with  $100 \text{ fb}^{-1}$  of luminosity.

obtaining statistics in the  $> 3 \text{ GeV}/c^2$  mass range. Finally, there is a strong incentive to study single-tagged two-photon reactions (where one of the scattered  $e^\pm$ s is detected) since, when one of the photons is far from the mass shell, spin-1 resonances can be isolated (especially  $J^{PC} = 1^{-+} q\bar{q}g$  hybrid states) and unique QCD predictions can be tested.

Table 3-20 summarizes Monte Carlo event rate predictions for  $100 \text{ fb}^{-1}$  of data at PEP-II, along with present world average data samples. Clearly, at least a factor of 30 improvement in statistics can be made in all of these exclusive processes. Furthermore, new channels will be available, with a high-quality detector, which have not been explored in the past. Most of this can be accomplished with little impact on the optimal detector for studying  $CP$  violation. The main requirement is a very flexible trigger which can allow detection of low-mass and low-multiplicity final states with balanced transverse momenta. However, study of the single-tagged reactions require detection of low-angle scattered positrons in the backward direction down to polar angles of  $300 \text{ mrad}$ . Those topics will need to wait for a detector upgrade which adds a backward calorimeter.

### 3.15 Summary

The variety of physics accessible at PEP-II demands a general-purpose detector, as is typical in colliding beam experiments. The detector discussed here has been optimized for  $CP$ -violation studies. However, since so many modes contribute to such studies, this process automatically leads to a detector which can address a much wider range of physics issues. The set of topics presented in this chapter is in no way intended to be an exhaustive list

of all the physics studies possible with this detector, but rather an overview showing that a rich program is expected. The *BABAR* detector will undertake all possible aspects of this program.

## References

---

- [Abe94] F. Abe *et al.*, *Phys. Rev. Lett.* **73**, 2667 (1994).
- [Alb94b] H. Albrecht *et al.* (ARGUS Collaboration), *Z. Phys.* **C62**, 371 (1994).
- [Ale91] R. Aleksan, I. Dunietz, B. Kayser, and F. LeDiberder, *Nucl. Phys.* **B361**, 161 (1991).
- [Ale92] R. Aleksan, A. Gaidot, and G. Vasseur, “Probing the Penguin Contribution in  $B \rightarrow \pi\pi$  at an Asymmetric  $B$  Meson Factory,” DAPNIA/SPP 92–19 (1992).
- [Ale93] R. Aleksan, L. Oliver, O. Pene, and J.C. Raynal, *Phys. Lett.* **B317**, 123 (1993).
- [Atw94] D. Atwood, B. Blok, and A. Soni, SLAC–PUB–6635 (1994).
- [Bar90] V. Barger, C.S. Kim, and R.J.N. Phillips, *Phys. Lett.* **B251**, 629 (1990).
- [Bat93a] M. Battle *et al.*, *Phys. Rev. Lett.* **71**, 3922 (1993).
- [Bau92] D.A. Bauer, SLAC–400, 608 (1992); also SLAC–373, 173 (1992).
- [Bro93] T.E. Browder *et al.*, “A Review of Hadronic and Rare  $B$  Decays,” CLNS 93/1261 (1993).
- [Bro95] T.E. Browder and S. Pakvasa, University of Hawaii Report UH–511–814–95 (1995).
- [Buc93] F. Buccella *et al.*, *Phys. Lett.* **B302**, 319 (1993).
- [CER93] GEANT Detector Description and Simulation Tool, Version 3.15, CERN Program Library W5103, CERN (1993).
- [Cha95] J. Chauveau and P. David, “Reconstruction of the Vertex in Lepton-tagged  $B$  Mesons,” *BABAR Note # 224* (1995).
- [Cou95] D. Coupal, “Study of  $B^0 \rightarrow D^+ D^-$  and  $B^0 \rightarrow D^{*+} D^{*-}$  in *BABAR*,” *BABAR Note # 211* (1995).
- [Dea93] A. Deandrea, N. Di Bartolomeo, R. Gatto, and G. Nardulli, *Phys. Lett.* **B318**, 549 (1993).
- [Des94] N.G. Deshpande, X.G. He, and J. Trampetic, University of Oregon Report OITS–564 (1994).
- [Des95] N.G. Deshpande and X.G. He, *Phys. Rev. Lett.* **74**, 26 (1995).

- [Fra94] P.L. Frabetti *et al.*, *Phys. Rev.* **D50**, 2953 (1994).
- [Fry93] J.R. Fry and T. Ruf, “Prospects of Studying  $CP$  Violation and Mixing in  $D$  Decays at a  $\tau$ -charm Factory,” Third Workshop on the  $\tau$ -charm Factory, Marbella, Spain (1993).
- [Gol89] M. Golden and B. Grinstein, *Phys. Lett.* **B222**, 501 (1989).
- [Gol94] E. Golowich and S. Pakvasa, University of Massachusetts Report UMHEP-411 (1994).
- [Gro90] M. Gronau and D. London, *Phys. Rev. Lett.* **65**, 3381 (1990).
- [Hal92] L. Hall and S. Weinberg, *Phys. Rev.* **D48**, 979 (1992).
- [Har95] P. Harrison, “Study of  $B \rightarrow J/\psi K_s^0$ ,  $J/\psi K^*$  and  $\rho\pi$ ,” *BABAR Note # 218* (1995).
- [Inn93] W. Innes, “TRACKERR, A Program for Calculating Tracking Errors,” *BABAR Note # 121* (1993).
- [Jad90] S. Jadach and Z. Was, “KORALB 2.1: An Upgrade with Tauola Library of Tau Decays,” *Comput. Phys. Comm.* **64**, 267 (1991).
- [Jaf94] D.E. Jaffe, F. Le Diberder, M.H. Schune, and G. Wormser, “Treatment of Weighted Events in a Likelihood Analysis of  $B_s$  Oscillations or  $CP$  Violation,” *BABAR Note # 132* (1994).
- [Liu94] T. Liu, “The Future of High-Sensitivity Charm Experiments,” in Proceedings of the CHARM2000 Workshop, edited by D. Kaplan and S. Kwan (Fermilab, Batavia, Illinois), pp 7–9 (1994).
- [Nir93] For a Review of Standard Model Predictions on  $CP$  Violation in  $B$  Decays see Y. Nir and H.R. Quinn, *Ann. Rev. Nucl. Sci.* **42**, 211 (1992), and references contained therein.
- [PDG94] L. Montanet *et al.* (Particle Data Group), *Phys. Rev.* **D50** (1994).
- [Pia94] M.G. Pia, “ $B$  Tagging with Muons,” *BABAR Note # 192* (1994).
- [Pla95] S. Plaszczynski, M.-H. Schune, and G. Wormser, “An Improved Tagging for the *BABAR* Experiment,” *BABAR Note # 209* (1995).
- [Qui93] H.R. Quinn and A. Snyder, *Phys. Rev.* **D48**, 2139 (1993).
- [Ric95] J.D. Richman and P.R. Burchat, “Lepton and Semileptonic Decays of Charm and Bottom Hadrons,” submitted to *Rev. of Mod. Phys.*

- [Sjo93] T. Sjöstrand, “PYTHIA 5.7 and JETSET 7.4 Physics and Manual,” CERN-TH 7112/93 (1993).
- [SLA89] For earlier studies see SLAC-353 (1989), SLAC-400 (1992), and SLAC-419 (1993).
- [Sjo93] T. Sjostrand, “JETSET 7.4,” CERN-TH 7111/93, (1993).
- [Sny90] A. Snyder and S. Wagner, “An Alternate Method of  $B$ -Tagging for  $CP$  Violation Studies,” *BABAR Note # 25* (1990).
- [Sny94] A. Snyder, “ $CP$  Extraction Tool,” *BABAR Note # 188* (1994).
- [Sny95] A. Snyder, “ $CP$  Reach Calculation of the TDR,” *BABAR Note # 219* (1995).
- [Wal94a] R. Waldi, “Flavor Tagging of  $B^0$  Decays: Using All Available Information,” *BABAR Note # 190* (1994).
- [Wal95] R. Waldi, “Flavor Tagging Studies for the *BABAR* TDR,” *BABAR Note # 204* (1995).
- [Wea95] M. Weaver, “Tag Mode Vertex Resolution Studies,” *BABAR Note # 227* (1995).
- [Wol85] L. Wolfenstein, *Phys. Rev.* **D3**, 2381 (1985).
- [Wri94a] D. Wright, “BEGET: The  $B$  Factory Event Generator,” *BABAR Note # 149* (1994).
- [Wri94b] D. Wright, “ $K_L^0$  identification in  $B^0 \rightarrow J/\psi K_L^0$  Decays,” *BABAR Note # 201* (1994).

---

---

# Vertex Detector

---

## 4.1 Vertex Detector Requirements

---

The major motivation for building *BABAR* is to find and precisely measure *CP* violation in the decays of neutral *B* mesons. To this end, one needs to determine for each event the time interval between the two *B*-meson decays. This is accomplished by reconstructing the two primary decay vertices, which is the main task of the vertex detector. For particles with low transverse momenta ( $p_t < 100 \text{ MeV}/c$ ), which may not be reconstructed in the drift chamber, the vertex detector should also provide full tracking information. Due to this capability to perform stand-alone tracking, the *BABAR* vertex detector is known as the silicon vertex tracker (SVT).

Below, we will discuss the requirements for the SVT and give a concise overview, including performance studies. A detailed discussion of all design aspects completes this chapter.

### 4.1.1 Resolution

Without a measurement of the *B* decay vertex, no useful *CP* asymmetries can be extracted at the  $\Upsilon(4S)$ . Therefore, the most important function of the SVT is the determination of the *B* decay positions, especially along the beam direction. The measurement of *CP* asymmetries does not place a very stringent requirement on the intrinsic position resolution; there is less than a 10% loss in precision in the asymmetry measurement due to imperfect vertex resolution if the separation between the *B* vertices  $\Delta z$  is measured with a resolution of approximately one half the mean separation, which is  $250 \mu\text{m}$  at PEP-II [Sny94]. This translates into a single vertex resolution of better than about  $80 \mu\text{m}$  both for *CP* eigenstates and for tagging final states. Vertex resolution in this range is readily achieved with silicon strip detectors.

We can benefit from additional precision over that required for the measurement of  $\Delta z$ . Pattern recognition, vertex reconstruction, and, in particular, background rejection will all be improved by better intrinsic position resolution, which in turn will improve the efficiency

and purity of the samples used for the  $CP$  asymmetry measurements. Therefore, our aim is to achieve the best precision practical. Constraints which limit the achievable resolution include the total number of readout electronics channels, mechanical constraints, and considerations of cost and schedule. The multiple scattering in the beam pipe and in the silicon itself sets a lower limit on the useful intrinsic resolution, corresponding to a point resolution of about 10–15  $\mu\text{m}$  for measurements made close to the IP and 30–40  $\mu\text{m}$  for the outer layers [For94].

The vertex detector also dominates in the determination of the track angles. The angle measurements should be good enough that they do not significantly contribute to the experimental uncertainty in the measurements of track impact parameters or the invariant mass of combinations of tracks. This requirement does not impose a more stringent limit on point resolution than that determined through consideration of the issues discussed above. The tracking resolution needed for matching tracks to the particle identification system and the calorimeter is mostly provided by the drift chamber.

### 4.1.2 Acceptance

The coverage of the SVT must be as complete as technically feasible, given the location of the B1 magnets below  $17.2^\circ$  (300 mr) with respect to the beam line in both the forward and backward directions. To maximize the space available for the SVT and its readout electronics in the boost direction, machine components such as cooling manifolds and vacuum flanges are located in the backward region. Therefore, it is only possible to extend the backward coverage to within  $30^\circ$  of the beam line. In the  $\Upsilon(4S)$  center-of-mass, this corresponds to  $-0.95 < \cos \theta_{cm} < 0.87$ .

There should be as little material as possible within the active tracking volume. Special attention must be paid to minimizing the amount of material between the IP and the first measurement in order to reduce multiple scattering. The beam pipe itself contributes  $0.006X_0$  at normal incidence. Material located beyond the inner layers does not significantly degrade the measurement of track impact parameters, but does affect the performance of the overall tracking system and leads to increased photon conversions in the active region.

### 4.1.3 Efficiency

Our goal is to achieve close-to-perfect track reconstruction efficiency within the active volume of the tracking detectors when information from both the drift chamber and the SVT is used. The pattern recognition capabilities of the combined tracking system must be robust enough to tolerate background levels up to 10 times nominal, where nominal is defined for a 1 nTorr vacuum pressure in the beam pipe within  $\pm 30$  m of the IP. Low momentum particles that do

not traverse many drift chamber planes, such as many of the charged pions from  $D^*$  decays, must be reconstructed in the SVT alone. For this category of tracks, with  $p_t$  less than 100 MeV/ $c$ , we want to achieve reconstruction efficiencies of at least 80–90%. The SVT must also be efficient for particles such as  $K_s^0$ s that decay within the active volume. Together, these determine the number of measurements along a track and the necessary single-hit efficiency.

The impact parameter resolutions are determined by the precision of the measurement closest to the IP. The performance of the inner vertex layers must therefore be optimized for both high efficiency and good point resolution. It follows that redundancy for the first measurement is an important design requirement.

#### 4.1.4 Radiation Tolerance

The expected machine-related backgrounds set the requirements for the data transmission bandwidth and radiation resistance of all components located close to the interaction region [Lev94]. As described in Chapter 12, the expected dose in the innermost layer of the SVT averages about 33 krad/yr. The radiation is highly nonuniform in azimuth, peaking in the bend plane of the accelerator with a local maximum of up to 240 krad/yr over a small region covering approximately  $6^\circ$  in azimuth. At the location of the front-end electronics for Layer 1 the maximum dose is 110 krad/yr, with an average value of 47 krad/yr. Detector and front-end electronics are specified to be able to withstand at least 10 times the annual radiation dose. The readout electronics must be fabricated with radiation-hard technologies, and special attention must be paid to the sensitivity of the detector performance to radiation.

#### 4.1.5 Reliability

The SVT is mounted inside a support tube of radius 20 cm, which also supports and aligns the machine elements closest to the IP. Access to the SVT is not possible without a major shutdown involving removal of the support tube from the detector. The reliability requirements for the SVT are therefore more stringent than usual for such a device, with implications for engineering design at all levels. The detector layout must provide redundant measurements wherever possible; the electronic readout must be very robust and have redundancy built in for critical data and control lines; and the functionality of all components must not be compromised by exposure to the expected radiation levels. The detector monitoring and interlock system must serve as a safeguard against catastrophic failure in the event of a component malfunction or a simple human error.

## 4.2 Vertex Detector Overview

---

### 4.2.1 Choice of Technology

The SVT design is based on double-sided silicon microstrip detectors. The characteristics of this technology that make it attractive for the *BABAR* detector are: high precision for measuring the location of charged particles, tolerance to high background levels, and reduction in mass made possible through double-sided readout. The process for the fabrication of double-sided silicon detectors is now mature enough to be employed in a large-scale application and to meet the performance standards outlined above.

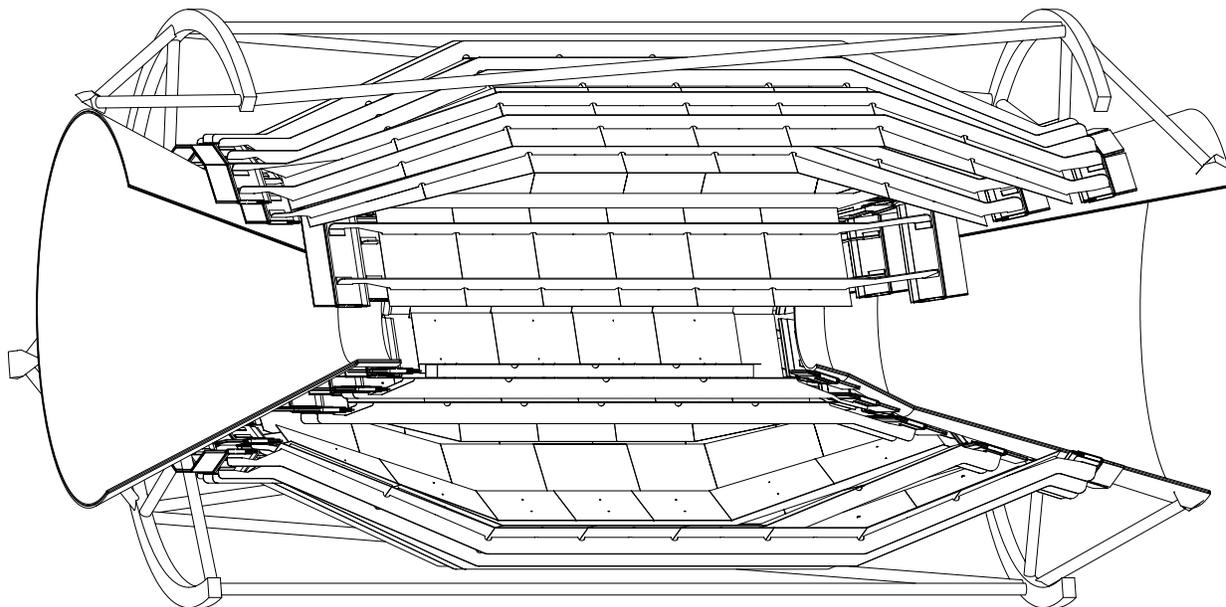
### 4.2.2 Detector Layout

The SVT will provide five measurements, in each of two orthogonal directions, of the positions of all charged particles with polar angles in the region  $17.2^\circ < \theta < 150^\circ$ . A three-dimensional cut-away view of the SVT is shown in Figure 4-1. Each of the three inner layers has six detector modules, arrayed azimuthally around the beam pipe, while the outer two layers consist of 16 and 18 detector modules, respectively. A side view of the detector is shown in Figure 4-2, and an end view is shown in Figure 4-3.

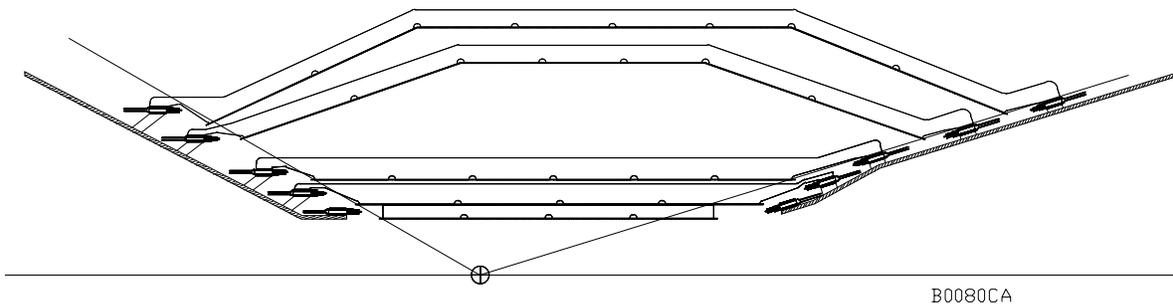
The inner detector modules are traditional barrel-style structures, while the outer detector modules employ a novel arch structure in which the detectors are electrically connected across an angle. The bends in the arch modules minimize the area of silicon required to cover the solid angle and also avoid very large track incident angles.

In order to satisfy the requirement of minimizing material in the detector acceptance region, one of the main features of the SVT design is the mounting of the readout electronics entirely outside the active detector volume. There is a 1 cm space between the 300 mr stay-clear and the B1 magnet in the forward direction; all of the forward electronics are mounted here. In the backward direction, there is space below about 500 mr. In both directions, space is very tight, and the electronic and mechanical designs are closely coupled in the narrow region available.

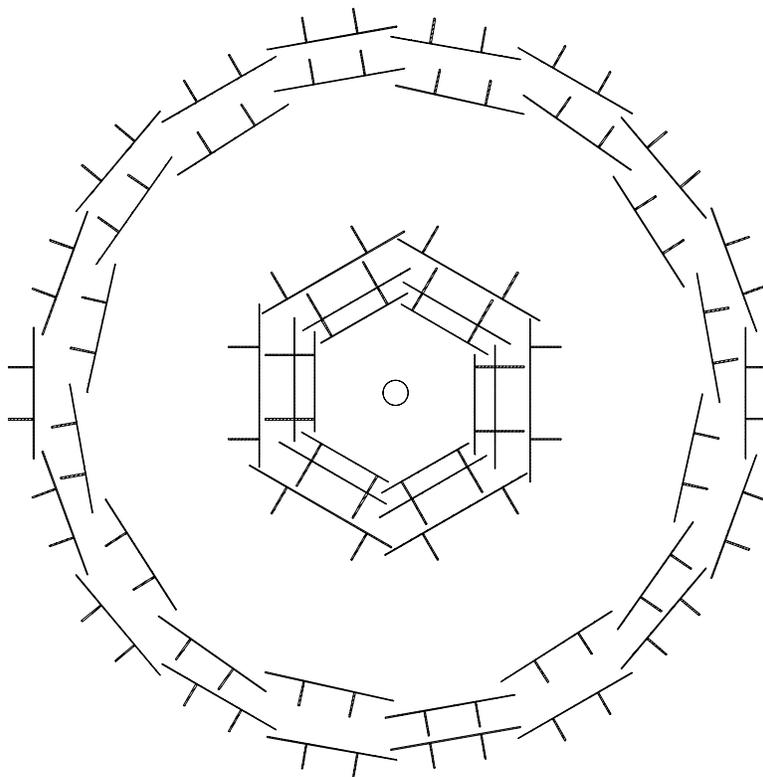
The layout specifications for this five-layer design are given in Table 4-1. The strips on the two sides of the rectangular detectors in the barrel regions are oriented parallel ( $\phi$  strips) or perpendicular ( $z$  strips) to the beam line. In the forward and backward regions of the two outer layers, the angle between the strips on the two sides of the trapezoidal detectors is approximately  $90^\circ$ , and the  $\phi$  strips are tapered. Floating strips are used to improve the position resolution for near-perpendicular angles of incidence; the capacitive coupling



**Figure 4-1.** *Three-dimensional cutaway view of the SVT.*



**Figure 4-2.** *Cross-sectional view of the SVT in a plane containing the beam axis.*



**Figure 4-3.** *Cross-sectional view of the SVT in a plane perpendicular to the beam axis. The lines perpendicular to the detectors represent structural beams.*

between the floating strip and the neighboring strips results in increased charge sharing and better interpolation. For larger incident angles, the wider readout pitch minimizes the degradation in resolution that occurs because of the limited track path length associated with each strip. These issues are discussed in more detail in Section 4.3.

The design has a total of 340 silicon detectors of seven different types. The total silicon area in the SVT is  $0.94 \text{ m}^2$ , and the number of readout channels is  $\sim 150,000$ .

### 4.2.3 Electronic Readout

As emphasized above, all readout electronics are located outside the active volume, below 300 mr in the forward direction and below about 500 mr in the backward region. To accomplish this,  $\phi$  strips on the forward or backward half of a detector module are electrically connected with wire bonds. This results in total strip lengths associated with a single readout channel of up to  $\sim 14 \text{ cm}$  in the inner three layers and up to  $\sim 24 \text{ cm}$  in the outer two layers.

Quantity	Layer 1	Layer 2	Layer 3	Layer 4a	Layer 4b	Layer 5a	Layer 5b
Radius (mm)	32	40	54	120	127	140	144
Wafers/Module	4	4	6	7	7	8	8
Modules/Layer	6	6	6	8	8	9	9
Silicon Area (cm <sup>2</sup> )	457	683	1072	1506	1582	2039	2082
Overlap in $\phi$ (%)	2.4	1.8	1.8	4.0	4.0	2.0	2.0
Readout pitch ( $\mu\text{m}$ ):							
$\phi$	50	50	50	65–100		65–100	
$z$	100	100	100	200		200	
Floating Strips:							
$\phi$	—	—	—	1		1	
$z$	1	1	1	1		1	
Intrinsic Resolution ( $\mu\text{m}$ ):							
$\phi$	10	10	10	10–12		10–12	
$z$	12	12	12	25		25	
R.O. Section/Module	4	4	4	4		4	
ICs/R.O. Section	6	8	10	4		4	
Readout Channels	18432	24576	30720	32768		36864	
Strip Length (mm):							
$\phi$	95	115	136	177/223	186/232	232/241	241
$z$	39	48	64	35–52	35–52	35–52	35–52
$z$ Ganging:							
Forward $\times 2$	34%	18%	7%	82%	88%	71%	60%
Forward $\times 3$						29%	40%
Backward $\times 2$	34%	18%	7%	82%	71%	60%	60%
Backward $\times 3$				18%	29%	40%	40%

**Table 4-1.** Parameters of the SVT layout. See text for more detail on the meanings of the different quantities. The intrinsic resolution is calculated at  $90^\circ$  track incidence assuming  $S/N = 20 : 1$ . The  $z$ -ganging numbers represent the percentage of detector strips connected to one other strip ( $\times 2$ ) or two other strips ( $\times 3$ ).

The signals from the  $z$  strips are brought to the readout electronics using fanout circuits consisting of conductive traces on a thin flexible insulator (for example, copper traces on Kapton). The traces are wire-bonded to the ends of the  $z$  strips. To read out all the  $z$  strips with the same number of electronics channels as the  $\phi$  strips, some  $z$  strips will be electrically connected, or ganged, together. The length of the  $z$  strips is much shorter, typically 5 cm in the inner layers and either 10 or 15 cm in the outer layers where there is either  $\times 2$  or  $\times 3$  ganging.

Front-end signal processing is performed by ICs mounted on hybrid circuits that distribute power and signals, and thermally interface the ICs to the cooling system. The signals from the readout strips, after amplification and shaping, are compared to a preset threshold. The time interval during which they exceed the threshold (time over threshold, or TOT) is an analog variable related to the charge induced on the strip. This time interval is digitally recorded prior to readout. Unlike the ordinary peak-amplitude measurement at the shaper output, the TOT technique has a nonlinear input-to-output relationship which is approximately logarithmic. This is an advantage since it compresses the dynamic range and allows one to achieve good position resolution and large dynamic range with a minimum number of bits.

TOT readout is less complicated than a linear analog readout system, resulting in less development time, lower cost, and smaller space and power requirements. At the same time, it is a significant improvement over a purely digital (hit/no-hit) readout, and provides sufficient analog resolution for position interpolation, time-walk correction, and background rejection. The readout IC is expected to be about 8–9 mm long and to dissipate no more than 2.0 mW per channel. The total power that will be generated by the SVT readout is  $\sim 300$  W.

For each channel with a signal above threshold, the strip number within the readout module, the time of arrival of the signal, and the digital value of the TOT will be read out. There are four readout sections per detector module, where the module is divided in half along  $z$ , and the  $\phi$  and  $z$  strips are grouped together separately. The data from one-half of a detector module will be transmitted from the hybrid on a flexible cable to a transition card located approximately 40 cm away, where the signals are converted to transmission via conventional cables.

#### 4.2.4 Mechanical Support

The silicon detectors and the associated readout electronics are assembled into mechanical units called detector modules. Each module contains several silicon detectors glued to low-mass beams constructed of carbon and Kevlar fiber-epoxy laminates. The beams are attached to the hybrid electronic circuits at each end. A ceramic or aluminum substrate for the hybrid provides precise mechanical mounting surfaces and is the heat sink for the electronics.

The inner layer is supported from the second layer; the detector modules from Layers 1 and 2 are glued together with rigid beams, forming sextants which are then mounted from the support cones in the forward and rear directions. Each detector module of the third, fourth, and fifth layer is mounted on the support cones independently of the other modules. In the fourth and fifth layers, there are two different types of modules in each layer, an inner one, labeled  $a$ , and an outer one, labeled  $b$ , occupying slightly different radial positions. Thus there are seven different types of detector modules.

The support cones are double-layered carbon-fiber structures which are mounted from the B1 magnets. Cooling water flows between the two carbon-fiber layers around aluminum mounts which protrude through the outer surface of the cone. Mounting pins in the hybrid structure provide the alignment between the modules and the aluminum mounts in the cone, and thermal contact is made to provide cooling for the front-end electronics located on the hybrid. The support cones are divided to allow the vertex detector to be assembled in two halves and then mounted on the B1 magnets by clam-shelling the pieces together.

The stiffness of the overall structure is provided by a very low mass space frame, constructed of carbon-fiber tubes, connecting the forward and backward support cones. It consists of rings at each end held rigid by 12 struts spanning the length of the detector. The rings are connected to the support cones by an additional series of 12 struts at each end. All material is carbon-fiber laminate. Preliminary finite element calculations show that this structure meets the tolerances for rigidity. The motivation for this space frame stems mainly from the possible relative motion of the two B1 magnets during the assembly procedure. Cooling water, power, and signal lines are routed along the B1 magnets to points outside the active region where manifolds for the cooling water and drivers for the electronics are located.

## 4.3 Detector Performance Studies

---

Cost, engineering complexity, space, and cooling requirements each constrain performance. In this section, we discuss how the SVT performance has been optimized within these constraints. Design and performance studies have considered the following:

- the impact of the number and locations of detector layers on track parameter resolution and pattern recognition;
- the effect of different readout schemes for recording charge information on intrinsic resolution, background rejection, and ultimately, on pattern recognition; and
- optimization of the readout pitch, taking into account the required resolution, the number of readout channels, and, in the case of the outer layers, the signal loss that occurs when the readout pitch becomes very large.

### 4.3.1 Resolution

#### Intrinsic Resolution

We have studied the effects on intrinsic position resolution of the strip pitch, readout pitch, and threshold levels, as well as various schemes for recording pulse information. The deposition of charge in a silicon strip detector was simulated by a Monte Carlo program, taking into account the effects of Landau fluctuations in energy loss, diffusion, channel-to-channel gain variations (5%), and noise. The simulation is in good agreement with experimental data on the energy loss in silicon, and with experimental measurements of position resolution in strip detectors.

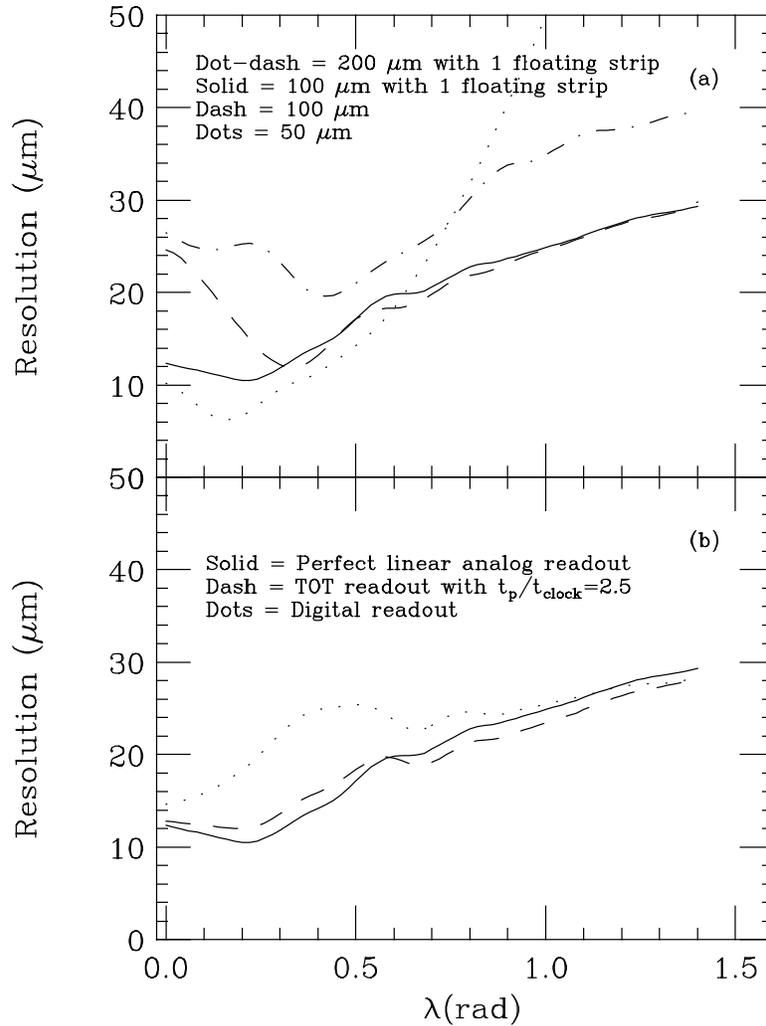
The intrinsic position resolution was studied as a function of the track incident angle in the plane normal to the strips. Charged particles most often cross the detectors at close to normal incidence in the plane perpendicular to the  $\phi$  strips; however, the angle of incidence of tracks in the plane perpendicular to the  $z$  strips extends to 300 mr from the beam line, or  $\lambda = \pi/2 - 0.3 \approx 1.3$  rad, where  $\lambda$  is the dip angle measured relative to normal incidence. The resolution degrades significantly at large dip angles, especially for small readout pitch. This is due to inefficiency in the readout for strips with small signals deposited by tracks at grazing angles.

Figure 4-4(a) shows the intrinsic resolution for 300  $\mu\text{m}$ -thick silicon as a function of the incident angle  $\lambda$  for four different strip and readout pitches. Strip and readout pitch coincide in the absence of floating strips; floating strips, when present, increase capacitive charge-sharing and thereby improve the point resolution. The simulation assumed a noise level of 1200  $e^-$ , corresponding to a signal-to-noise ratio of about 20 for perpendicular tracks, and the threshold was set to 4 times the noise. The algorithm employed to determine the position uses the digital centroid for the central strips and applies a correction based on the relative amount of charge in the two edge strips of a cluster:

$$X = X_0 + 0.5 \cdot p \cdot \frac{Q_1 - Q_N}{Q_1 + Q_N},$$

where  $X_0$  is the digital centroid (*i.e.*, geometrical center) of strips 2 through  $N - 1$ ;  $Q_1(Q_N)$  is the charge deposited on the first(last) hit strip, and  $p$  is the readout pitch.

As seen in Figure 4-4(a), the resolution degrades significantly as the angle of incidence increases for 50  $\mu\text{m}$  strip and readout pitch. With 100  $\mu\text{m}$  strip and readout pitch, the resolution does not degrade as much with larger incident angles but, as expected, the resolution at normal incidence is about a factor of 2 worse than that for 50  $\mu\text{m}$  pitch. On the other hand, 100  $\mu\text{m}$  readout pitch with one floating strip gives a resolution equivalent to 50  $\mu\text{m}$  readout pitch without floating strips for angles smaller than 45° ( $\lambda < 0.8$ ). Above 45°, it is equivalent to 100  $\mu\text{m}$  strip pitch. Therefore, floating strips allow us to retain



**Figure 4-4.** (a) Intrinsic resolution for a 300  $\mu\text{m}$ -thick silicon detector as a function of the dip angle for various strip configurations, assuming perfect analog readout. (b) Same as (a), assuming 100  $\mu\text{m}$  readout pitch and one floating strip, for various readout techniques.

good resolution for small incident angles with fewer readout channels and less degradation in resolution for large incident angles. The upper curve in Figure 4-4(a) shows that for a  $200\ \mu\text{m}$  readout pitch with one floating strip, the resolution is about  $25\ \mu\text{m}$  at normal incidence. This configuration is used for the  $z$  strips in layers 4 and 5.

We have also studied the effect of the nonlinearity and quantization error of a TOT scheme compared both to an ideal analog readout and to a digital readout in which no information is recorded about the signal size [Roe94b]. The position is determined using the algorithm described above, but using the TOT, rather than the charge, to weight the edge strips.

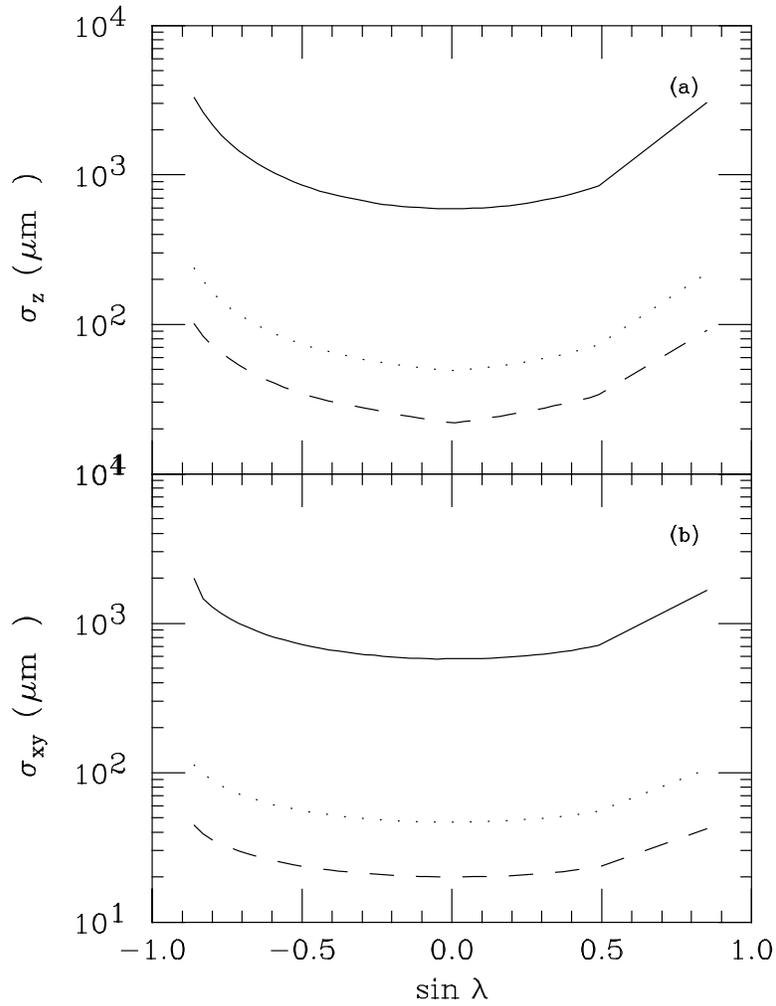
The resolution is shown as a function of track incident angles in Figure 4-4(b). The signal shaping was assumed to have a peaking time 2.5 times longer than the period of the clock used in the digitization of the pulse length for this simulation. This value is consistent with that expected in the inner vertex layers, where fast shaping is necessary to cope with larger backgrounds. The resolution for a TOT readout scheme is practically the same as ideal analog readout. Digital readout, which provides hit/no-hit information only, shows much poorer resolution in comparison.

In addition to the advantage of improved resolution, analog readout may be useful for background rejection of non-MIP signals, for pattern recognition techniques which correlate pulse height with angle, and for correction of the time-walk which will affect the time-stamp information generated by the readout chip.

### Track Parameter Resolution

The track parameter resolutions described in Chapter 2 were calculated using an analytical technique based on the method of Billoir [Bil84] which can be applied to arbitrary detector geometries in the program TRACKERR [Inn93]. Figure 4-5 shows the resolutions for the track impact parameter along the beam ( $\sigma_z$ ) and in the plane perpendicular to the beam ( $\sigma_{xy}$ ) for various momenta. The resolution is dominated by multiple scattering and not by the intrinsic position resolution of the silicon detectors, except for the highest momentum tracks from  $B$  decays.

A Monte Carlo analysis was performed to determine how the intrinsic resolution and the number of detector layers affect the track parameter resolution. The study included the effect of a catastrophic failure in Layer 1, the most important for vertex resolution. The intrinsic resolution was varied on the first two layers in the range  $5\text{--}25\ \mu\text{m}$  and on the outer two layers in the range  $10\text{--}50\ \mu\text{m}$ . The point resolution for Layer 3 was kept fixed at  $10\ \mu\text{m}$ . The quoted resolutions are for normally incident tracks; these were degraded as the dip angle  $\lambda$  increased, so that  $\sigma = \sigma_0 \cdot (1 + 1.5\lambda)$ .



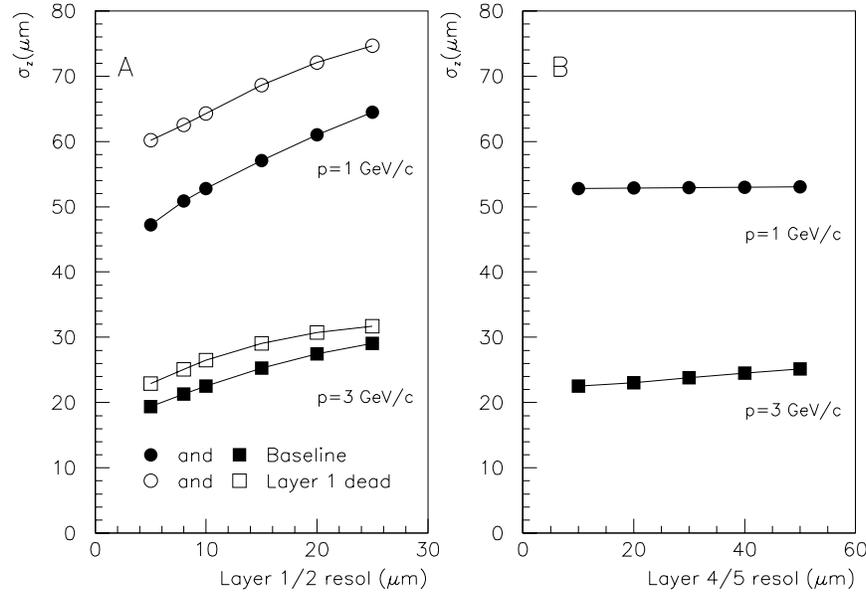
**Figure 4-5.** (a) Impact parameter resolution along the beam ( $\sigma_z$ ) and (b) in the plane perpendicular to the beam ( $\sigma_{xy}$ ) vs. dip angle for three different momenta: 100 MeV/c (solid line), 1.0 GeV/c (dotted line), and 3.0 GeV/c (dashed line).

In addition to the baseline geometry with five layers, we considered a four-layer layout which was obtained by removing the third layer. The basic conclusions of these studies are summarized as follows (see Reference [For94] for a full discussion):

- The angle and momentum resolution are essentially independent of the intrinsic resolution.
- The  $z$  impact parameter, on the other hand, does depend on the intrinsic resolution as shown in Figure 4-6(a) (the  $xy$  impact parameter has a similar behavior). In changing the intrinsic resolution on Layers 1 and 2 from 10 to 20  $\mu\text{m}$ , respectively,  $\sigma_z$  degrades by about 15–20% for 1 GeV/ $c$  and by almost 40% for 3 GeV/ $c$  tracks. For soft tracks ( $< 100$  MeV/ $c$ ), no dependence on the intrinsic resolution is observed because multiple scattering dominates. Most tracks produced in  $B_{\text{ABAR}}$  will be soft ( $< 500$  MeV/ $c$ );  $B$  vertices, however, will be determined by the hardest tracks in the event.

The requirement that the track parameter resolution be dominated by multiple scattering for most tracks therefore leads us to choose the best intrinsic resolution technically feasible in Layers 1 and 2. Within the constraints imposed by the mechanical, electronic, and silicon detector performance, the best intrinsic resolution achievable in the inner layers is  $\sim 10$   $\mu\text{m}$  in  $\phi$  (employing 50  $\mu\text{m}$  readout pitch) and  $\sim 12$   $\mu\text{m}$  in  $z$  (employing 100  $\mu\text{m}$  pitch with one floating strip). A larger pitch in  $z$  reduces the amount of ganging required and provides good resolution for tracks at small crossing angles.

- If for some reason Layer 1 is not usable, the tracking resolution is degraded, as shown in Figure 4-6(a). The degradation in the  $z$ -impact parameter resolution is 19% at 3 GeV/ $c$  and 22% at 1 GeV/ $c$ ; thus, Layer 2 is an effective backup in case of a failure in Layer 1.
- The resolution on Layers 4 and 5 does not significantly affect any of the track parameters (Fig 4-6(b)), and the choice of the readout pitch in the outer layers is determined by other considerations (*i.e.*, noise, occupancy, and detector design considerations).
- Given the relatively soft momentum spectrum expected in  $B_{\text{ABAR}}$ , we were concerned that the material in the third layer (whose main motivation is pattern recognition) could degrade the tracking resolution. TRACKERR studies have shown [For94] that the effect of removing Layer 3 is in fact small (at most 5%) for high momentum tracks and negligible for soft tracks.



**Figure 4-6.**  $z$ -impact parameter resolution as a function of the point resolution (a) in Layers 1 and 2 for the baseline design and for the case of a Layer 1 failure, and (b) in Layers 4 and 5. All values are at  $\theta \approx 90$  deg for 1 GeV/c and 3 GeV/c tracks.

### Vertex Resolution

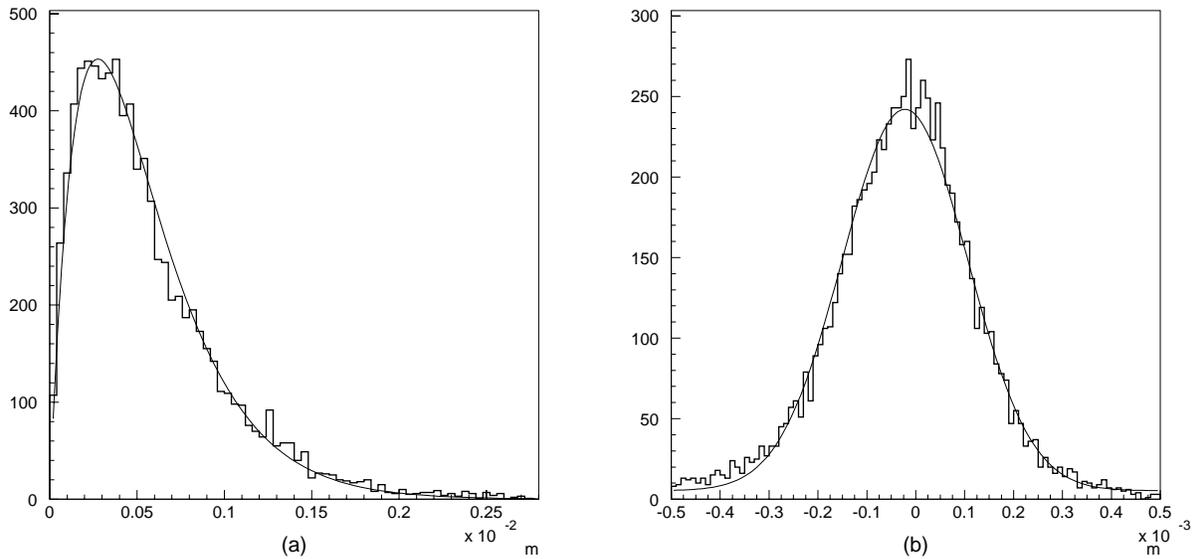
The  $z$  vertex resolution has been calculated using a parameterized Monte Carlo simulation for three  $B$  decay modes:  $B^0 \rightarrow \pi^+\pi^-$ ,  $B^0 \rightarrow J/\psi K_S^0$ , and  $B^0 \rightarrow D^+D^-$ . For each decay, the distribution of the difference between the reconstructed and generated  $B^0$  vertex positions in the  $z$  direction is fit with a sum of two Gaussian functions. The results are shown in Table 4-2. The results for  $B \rightarrow \pi^+\pi^-$  have been checked with a full GEANT-based Monte Carlo simulation, which gave reasonable agreement for both  $\sigma_n$  and  $\sigma_w$ . In all cases, the expected rms resolution (obtained by taking the weighted sum of  $\sigma_n$  and  $\sigma_w$ ) is better than the  $80 \mu\text{m}$  required for  $CP$  violation studies.

The  $z$  vertex resolution for the tagging  $B$  has also been studied with both parameterized and GEANT-based Monte Carlo simulations. The resolution depends upon which type of tag one employs. For direct leptons, we find  $\sigma_n \approx 60 \mu\text{m}$ , while for kaon tags,  $\sigma_n \approx 75 \mu\text{m}$ , with about 70–80% of the tags contained in the narrow Gaussian.

Good vertex resolution in the plane perpendicular to the beam pipe is also important. This information will help in associating tracks from  $B$  and  $D$  decays to the correct secondary or tertiary vertices, thus reducing combinatorial backgrounds. Figure 4-7(a) shows the distribution in the  $xy$  plane of the distance between two  $D$  vertices in the decay  $B^0 \rightarrow D^+D^-$ .

Mode	$\sigma_n$ ( $\mu\text{m}$ )	$\sigma_w$ ( $\mu\text{m}$ )	$f_n$ (%)	rms
$\pi^+\pi^-$	$27.5 \pm 0.4$	$85.8 \pm 3$	$84.0 \pm 0.4$	$36.8 \pm 0.6$
$\Psi K_s$	$42.1 \pm 0.7$	$168 \pm 7$	$88.0 \pm 0.2$	$57 \pm 1.0$
$D^+D^-$	$46.5 \pm 0.6$	$159 \pm 10$	$94.0 \pm 0.1$	$53.3 \pm 0.8$

**Table 4-2.** A two-Gaussian fit to the  $z$  vertex resolution for various  $CP$  eigenstates;  $\sigma_n$  and  $\sigma_w$  are the rms of the narrow and wide Gaussians, respectively, and  $f_n$  is the fraction of the area contained in the narrow Gaussian. The rms is the weighted sum of  $\sigma_n$  and  $\sigma_w$ .

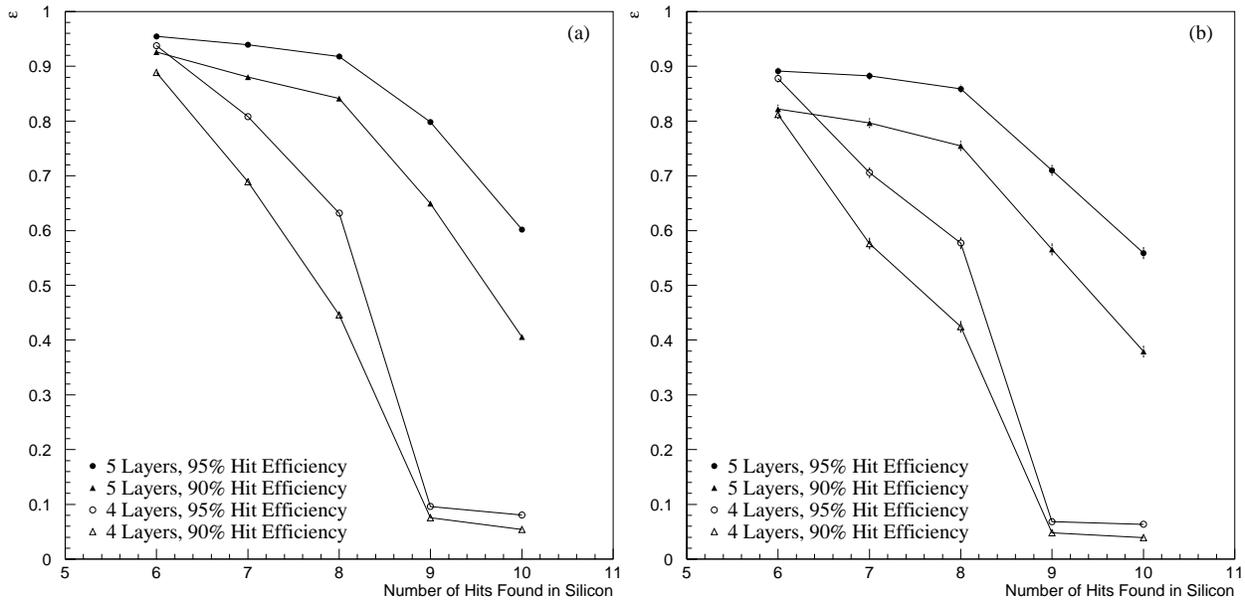


**Figure 4-7.** (a) Separation between  $D$  vertices in the  $xy$  plane. (b) Resolution on the separation between the  $D$  vertices in the  $x-y$  plane. Note the different scales on the abscissa.

The distribution is fit to a function of the form  $A \cdot x e^{-x/\lambda}$  and yields a value of  $\lambda = 276 \mu\text{m}$ . For comparison, the expected resolution on the separation between the two  $D$  decay points in the  $x-y$  plane is plotted in Figure 4-7(b) for the decay  $D \rightarrow K\pi\pi$ . The rms resolution is about  $125 \mu\text{m}$ .

### 4.3.2 Pattern Recognition

The pattern recognition capabilities of the  $BABAR$  vertex detector have been studied using a GEANT simulation with a detailed model of the detector geometry. The simulation includes



**Figure 4-8.** Track finding efficiency for tracks from generic  $B$  decays: (a) all tracks, (b) tracks measured in the SVT only ( $p_t < \sim 90$  MeV/c).

detector overlaps, inactive regions near the edges of the wafers, strip inefficiency, background hits, and the effects of ganging the  $z$  strips. A method based on Kalman filtering techniques was used to locate tracks in the SVT. For tracks that went a sufficient distance into the drift chamber (corresponding to  $p_t > 90$  MeV/c), the results of a fit to the drift chamber track segment were used as a starting point for the silicon track finding. Hits in the SVT that were best associated with each of these track segments were added to the track. Once this was finished, these hits were removed from the SVT and a vertex-only pattern recognition algorithm was used on the remaining hits to find low- $p_t$  tracks. This vertex-only algorithm was based on the same Kalman filtering technique but used combinations of hits in the outer layers of the SVT to generate the initial track parameters.

All tracks were required to cross five layers of silicon. The efficiency for finding tracks was computed by requiring a minimum number of silicon hits per track and by requiring each track fit to pass a fairly loose  $\chi^2$  cut ( $\chi^2/N_{dof} < 10$ ). A single hit in the silicon was defined as either a  $\phi$  or a  $z$  hit, so that a track going through five layers of silicon would generate ten hits. Six silicon hits was considered to be the minimum required in order for the number of degrees of freedom for a fit to the silicon-only tracks to be  $\geq 1$ . The track finding efficiency for all tracks from generic  $B$  decays is plotted in Figure 4-8(a) as a function of the number of hits required. For this simulation, two geometries were studied: the baseline five-layer SVT, and a four-layer version in which Layer 3 was removed. The single hit efficiency was set to either 90% or 95%, and additional hits from beam-induced backgrounds at a level corresponding to 10 times nominal were overlaid on the hits from the  $B^0\bar{B}^0$  event. The

efficiency for the five-layer geometry remains high for up to eight required hits per track, while for the four-layer device, the efficiency is high only for the six-hit requirement.

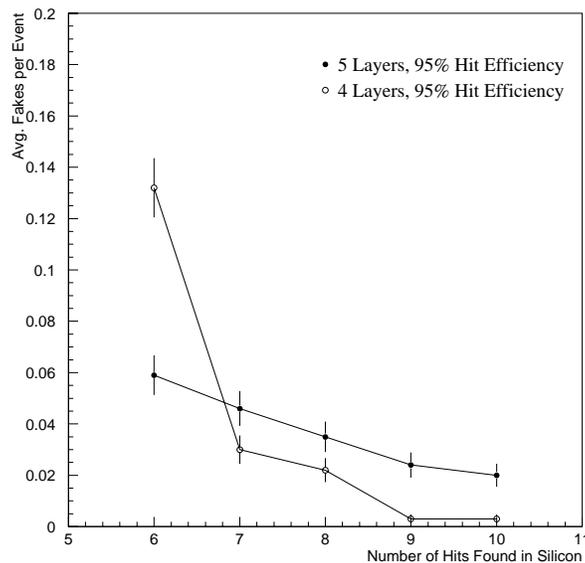
In Figure 4-8(b), the efficiency is plotted for the same two geometries, but only for tracks which do not traverse enough layers in the drift chamber to be reliably found there. The overall efficiency is somewhat lower, and the same general behavior is observed, namely, a plateau for six, seven, or eight required hits in the five-layer detector and a sharply falling efficiency for more than six required hits in the four-layer version. The algorithm employed for finding tracks using the vertex detector only is quite sensitive to the detector efficiency, as seen in the figure. For comparison, in a perfect detector with 100% efficiency, the silicon-only track-finding efficiency would be 95% for the five-layer detector when requiring seven hits. Both the silicon-only and the combined silicon plus drift chamber track-finding efficiencies are quite insensitive to the background level.

The frequency at which fake tracks are found is also important to quantify. In this study, a fake track was defined as a track which had less than half of its hits from a single real track. A loose impact parameter cut was imposed, rejecting any tracks originating more than 5 cm from the nominal beamspot in  $z$  or more than 2 cm in  $xy$ . The average number of fake tracks per event is shown in Figure 4-9 as a function of the number of hits found in the silicon. This measurement was also performed assuming a background level that was 10 times nominal. In order to achieve reasonable efficiency for good tracks in the four-layer geometry, one cannot ask for more than six hits to be found per track. Therefore, the relevant rate in the four-layer geometry is 0.13 fake tracks per event. The five-layer geometry provides a relative reduction in the fake rate of a factor of  $\sim 3$  if one requires seven or eight hits on a track, without any significant reduction in track-finding efficiency. Approximately half of the fake tracks have  $p_t$  (as measured in the silicon) large enough to generate a corresponding track segment in the drift chamber. Most of these fakes should therefore disappear when attempts are made to match them to drift chamber hits.

### 4.3.3 Solid Angle Coverage

As discussed above, the SVT must have high efficiency over the region  $17.2^\circ < \theta < 150^\circ$ . The stay-clear at  $17.2^\circ$  (300 mr) in the forward region does not allow greater coverage, and beam-line components in the backward region prevent extension below  $\sim 500$  mr.

The geometry of the SVT in the inner layers is hexagonal, so the solid angle coverage varies as a function of azimuth [Roe94a]. To meet the solid angle coverage requirements, the first two layers are extended along  $+z$  so that a track emitted at 300 mr from  $z = 0$ , at any azimuthal angle, is in the SVT acceptance. There is then some coverage down to 245 mr for tracks in certain azimuthal regions.



**Figure 4-9.** Fake track rate in the SVT at 10 times nominal background rate.

In the outer layers it is not possible to extend the detectors beyond the 300 mr stay clear. However there are many more staves per layer, and coverage extends to between 300 and 310 mr everywhere.

In the backward direction, there is a similar variation of coverage with azimuthal angles. At the minimum radii of the inner layers there is nominal coverage to within  $30^\circ$  of the beam line. In the center-of-mass, the SVT covers  $-0.95 < \cos \theta_{cm} < 0.87$ , so the acceptance in the backward direction is actually greater than that in the forward direction in this frame.

## 4.4 Silicon Detectors

The SVT will be constructed from double-sided, AC-coupled silicon strip detectors. These solid state devices are a technically mature solution to the requirements the SVT must meet to provide precise, highly segmented, robust tracking near the interaction point. The detailed requirements which the detectors must meet are discussed below.

### 4.4.1 Requirements

**Readout Strip Efficiency.** The silicon detectors must maintain high single-point efficiency in order to achieve the requirements given in Section 4.1 for high overall track reconstruction efficiency and good tracking resolution. Loss of efficiency can occur from

intrinsic strip inefficiencies, from bad interconnections, or from faulty electronics channels. Intrinsic strip inefficiencies can occur due to production defects which result in strips with unacceptably large leakage currents, from accidents during assembly causing the strip to be physically damaged, or from a breakdown in the AC-coupling capacitor. The latter problem is referred to as a pinhole and is due to a small hole in the oxide separating the implant from the metal readout strip above it. Pinholes can occur during fabrication, or they can be generated later on. Understanding and controlling the level of pinholes is one of the primary concerns in our program of silicon detector R&D.

Our goal is to achieve an overall single detector strip failure rate of less than 1%. Data from a large production of double-sided DC-coupled detectors (ALEPH) show that 60–70% can be achieved with a maximum inefficiency of 1%. On this basis, we expect that a 50% yield can be achieved for double-sided AC-coupled detectors while maintaining similar standards.

**Point Resolution.** As described in Section 4.3, we have determined from Monte Carlo simulations [For94] that the intrinsic point resolution should be  $15\ \mu\text{m}$  or better in both  $z$  and  $\phi$  for the inner layers. These are the point resolutions for tracks at near-normal incidence. As the angle between the track and the plane normal to the strip increases, the resolution degrades. We require the resolution to degrade by no more than a factor of approximately 3 for angles up to  $75^\circ$  ( $\lambda \sim 1.3$ ) from normal.

**Radiation Hardness.** A further requirement is that the quoted resolution values hold up to an integrated dose of  $\sim 2\ \text{Mrad}$  of ionizing radiation (electromagnetic in origin). This requirement leads to the use of AC-coupled detectors in order to avoid the problems associated with direct coupling of the large leakage currents which can occur at such large doses. It also has implications in the choice of the biasing scheme.

**Minimum Mass.** To achieve good vertex resolution, it is especially important to minimize the material up to and including the first measurement. This requirement, and the need to provide precise vertexing in both  $z$  and  $\phi$ , leads to the choice of double-sided detectors. We plan to use  $300\ \mu\text{m}$ -thick silicon wafers, which are a standard choice and present acceptable handling properties. While it may be possible to go to  $200\ \mu\text{m}$ -thick silicon, the gain is only by the square-root of the total material before the first measurement (including the  $0.006X_0$  in the beam pipe), while the fabrication and handling will be much more problematic.

### 4.4.2 Silicon Detector Design

From the above requirements and from the discussion in Sections 4.1–4.3, we have arrived at the detector specifications and design parameters which are described in this section. A more complete discussion can be found in Reference [Bat94b].

**Substrate and implant type.** The wafers will be n-type, with a resistivity in the range 4–8 k $\Omega$  cm, corresponding to a depletion voltage of 40 to 80 V. These values seem to be a reasonable compromise between the need to have a low depletion voltage and the need to avoid type inversion in the presence of radiation damage.

We will employ  $p^+$  strips on the junction side and  $n^+$  strips on the ohmic side, with  $p^+$ -blocking implants in between; see Figure 4-10 for a cross-sectional view. This choice has proven to be a reliable technology [Bat94a] which is directly available without extensive R&D.

**Coupling to preamplifier.** The strips are connected to the preamplifiers through a decoupling capacitor. AC coupling prevents the amplifier from integrating the leakage current with the signal; handling high leakage currents due to radiation damage imposes an additional burden on the preamplifier design and has other undesirable operational implications.

The value of the decoupling capacitance must be much larger than the total strip capacitance, which is as large as 35 pF. It is possible to use DC-coupled detectors with external capacitors on a separate chip [Bat92]; however, with an electronics pitch of 50  $\mu$ m, we would need a  $\sim$ 12 mm-long chip to get a  $C_{AC} = 180$  pF decoupling capacitor. These dimensions are prohibitively large given the limited space available in the hybrid region. There would also be a significant increase in the number of wirebonds. Capacitors which are integrated on the detectors are the most compact solution, minimizing the number of wirebonds and yielding a value for the decoupling capacitance of 30–80 pF/cm [Ton94], depending on the implant width.

**Bias resistors.** The bias resistors must be between 4 and 20 M $\Omega$ . The lower limit is determined by two factors. The noise has a  $1/\sqrt{R_B}$  dependence, and if several strips are ganged together, the effective resistance is correspondingly decreased. Another factor is the requirement that, for floating strips, the product  $R_B \cdot C_{TOT}$  must be much larger than the amplifier peaking time (100–400 ns) to allow for capacitive charge partition. The upper limit (20 M $\Omega$ ) is dictated by the allowable potential drop due to the strip leakage current, which

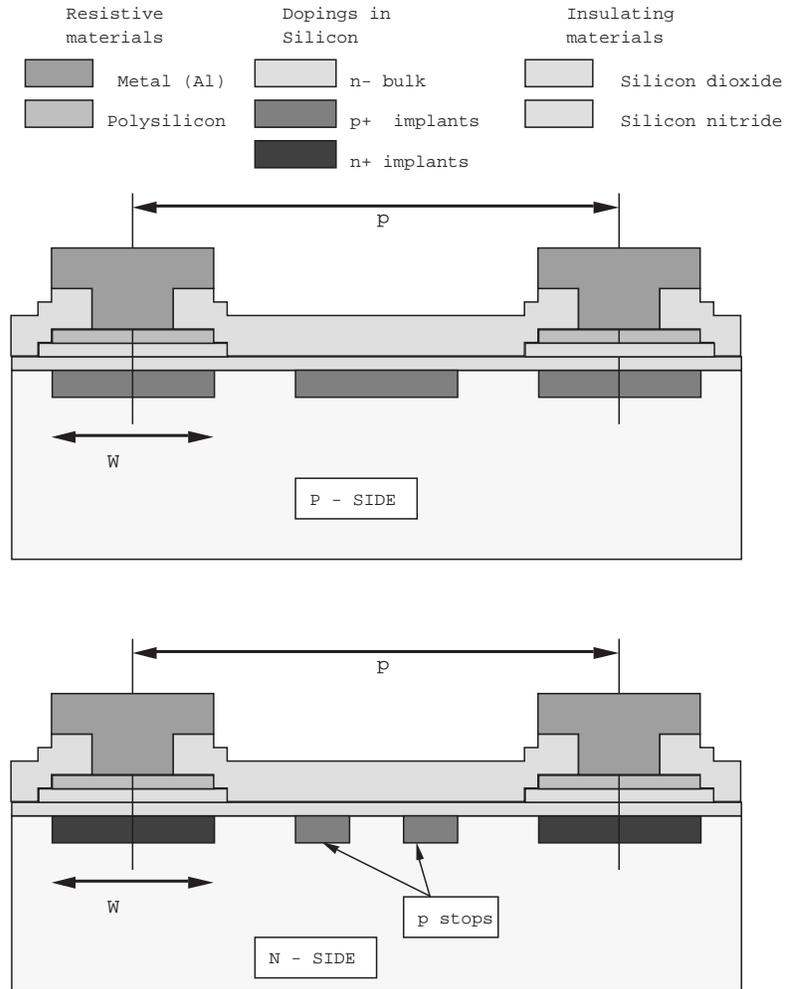


Figure 4-10. Artist's conception of a silicon detector cross section.

is taken to be 100 nA at maximum. A good target value is 8 M $\Omega$ . A final requirement is that the bias resistor be quite stable for the expected radiation doses.

To meet these requirements, we plan to use polysilicon bias resistors. In prototype CMS detectors [Ton94], values for the sheet resistance of polysilicon of 40 k $\Omega/\square$  were achieved, and 50 k $\Omega/\square$  is feasible. Thus, it is possible to fabricate an 8 M $\Omega$  resistor with a 6  $\mu\text{m}$ -wide, 960  $\mu\text{m}$ -long polysilicon resistor. With a suitable shaping of the polysilicon line, the space required by the resistor will be 480  $\mu\text{m}$  for a 25  $\mu\text{m}$  implant pitch.

For the junction side, punch-through biasing [Hol89] could also be chosen if it is confirmed to be sufficiently radiation hard. Punch-through biasing requires less space, and it can also reduce the parallel noise on the junction side, since the dynamic resistance is much higher at low currents (70 M $\Omega$  at 1 nA, decreasing to a few M $\Omega$  at 100 nA).

Considering the space needed to accommodate the biasing resistors and to gracefully degrade the electric field close to the edge with a guard ring structure, we specify the active region of the detectors to be 1.4 mm smaller than the physical dimensions (700  $\mu\text{m}$  on each edge).

**Optimization of  $z$  and  $\phi$  readout strips.** A major issue is which side of the detector (junction or ohmic) should read which coordinate ( $z$  or  $\phi$ ). The capacitance, and consequently, the noise is smaller on the junction side than on the ohmic side, and the strip pitch on the junction side can be 25  $\mu\text{m}$ , while on the ohmic side, it is limited to about 50  $\mu\text{m}$  because of the  $p$ -stop implant. For these reasons and because the  $z$  vertex measurement is more important from the point of view of physics, we use the junction side for the  $z$  strips on the inner layers. The better performance of the junction side also helps compensate for the additional resistance and capacitance imposed by the longer  $z$  fanout circuit.

In order to maintain acceptable signal-to-noise ratios for tracks at large dip angles, we employ a 100  $\mu\text{m}$  readout pitch for these  $z$  strips with one floating strip interleaved between every two readout strips. We have considered using a wider readout pitch, for example, 200  $\mu\text{m}$  for the very forward and backward regions in order to increase the signal at large dip angles. However, this would involve yet another detector design, and based on our present estimates of achievable electronic noise, it does not appear to be necessary.

Acceptable resolution can be obtained for the  $\phi$  strips on the inner layers using the ohmic side. Two solutions are possible; either a 50  $\mu\text{m}$  readout pitch without floating strips, since there is no room for them on the ohmic side, or a 100  $\mu\text{m}$  readout pitch with one floating strip. Either solution is feasible, and they should give roughly equivalent position resolution for single tracks. Double-track resolution is better for the first solution, and the noise contribution due to detector leakage currents is doubled in the latter solution. Therefore, preference goes to a 50  $\mu\text{m}$  readout pitch without floating strips. Although this choice has twice as many readout channels, the cost implications are not very important because the

Detector Type	$z$ -readout Side			$\phi$ -readout Side		
	$C_{int}$ (pF/cm)	$C_{AC}$ (pF/cm)	$R_{series}$ ( $\Omega$ /cm)	$C_{int}$ (pF/cm)	$C_{AC}$ (pF/cm)	$R_{series}$ ( $\Omega$ /cm)
I	1.3	40	7	2.8	40	7
II	1.3	40	7	2.8	40	7
III	1.3	40	7	2.8	40	7
IV	1.5	80	3.5	1.3	40	7
V	1.5	80	3.5	1.3	40	7
VI	1.5	80	3.5	1.3	30	9.2
VII	1.5	80	3.5	1.3	37	7.5

**Table 4-3.** *Electrical parameters for the different detector types.*

electronics cost is dominated by the development effort and consequently the per channel incremental cost is not significant.

On the outer layers, the  $\phi$  strips are quite long (up to 24 cm), and their capacitance can become very high on the ohmic side. The position resolution is not as important in these layers, and the maximum track crossing angle is about  $45^\circ$  for the  $z$  strips. Therefore, we choose instead to put the  $\phi$  strips on the junction side in order to optimize the signal-to-noise ratio for these very long strips. A  $100\ \mu\text{m}$  readout strip pitch with one floating strip is foreseen in the rectangular sections, decreasing to about  $65\ \mu\text{m}$  at the ends of the trapezoidal detectors.

For the design of  $\phi$  strips on the trapezoidal detectors, experimental investigations are needed to decide whether to keep the width of the implants constant, the size of the gaps constant, or to maintain a constant ratio of the gap to the implant width. In theory, a constant ratio between width and pitch is favored because most electrical parameters (*i.e.*, depletion voltage and interstrip capacitance) approximately scale with this quantity and therefore would remain constant along the length of the wedge.

For the  $z$  strips in the outer layers, the ohmic side is chosen with a readout pitch of  $200\ \mu\text{m}$  and one floating strip. The larger readout pitch is employed in order to minimize the ganging of  $z$  strips.

The geometrical layout of the seven different types of wafers, including strip pitches and number of floating strips, is summarized in Table 4-1 (Section 4.2). In Table 4-3, the interstrip capacitance, coupling capacitance, and series resistance per unit length are summarized for each detector type.

Det. Type	I	II	III	IV	V	VI	VII
$z$ Length (mm)	47.7×	57.6×	45.5×	46.0×	55.0×	68.1×	63.4×
$\phi$ Width (mm)	39.8	49.4	65.4	52.5	52.5	35.0–44.0	44.0–52.5
Junction on	$z$	$z$	$z$	$\phi$	$\phi$	$\phi$	$\phi$
Layer 1	24	—	—	—	—	—	—
Layer 2	—	24	—	—	—	—	—
Layer 3	—	—	36	—	—	—	—
Layer 4a	—	—	—	24	0	16	16
Layer 4b	—	—	—	8	16	16	16
Layer 5a	—	—	—	9	27	18	18
Layer 5b	—	—	—	0	36	18	18
Total	24	24	36	41	79	68	68
Total w/Spares	28	28	42	46	88	76	76
Batches	3	3	4	4	8	7	7

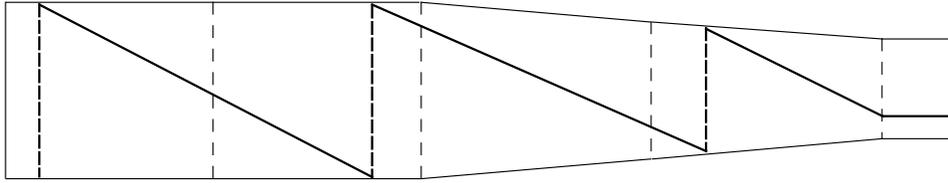
**Table 4-4.** Shape, multiplicity, and number of batches for the different detector types, with one spare module per module type.

**Wafer sizes and quantities.** Table 4-4 summarizes the shapes and number of detectors for the baseline design. The quoted multiplicities refer to the installed modules plus one spare per module type. The number of fabrication batches quoted assumes a yield of 50%, *i.e.*, 12 good detectors per batch of 24 wafers.

Using the numbers in Table 4-4, we see that the current design employs seven different types of detectors (*i.e.*, seven sets of masks) and needs 36 fabrication batches, for a total of 340 installed detectors. Having so many types of detectors complicates both the design and production phases, especially for prototype and spare production. A reduction in the number of detector types would be most welcome; however, this represents the minimum which we have been able to achieve in our present baseline design.

### 4.4.3 Fanout Circuit Design

The SVT front-end electronics will be located outside of the active tracking volume to minimize the amount of material crossed by particles within the acceptance. As a consequence, the signals from both the  $z$  and  $\phi$  microstrips must be carried to the front-end chips by flexible fanout circuits. While the  $\phi$  fanout circuits are just one-to-one connections a few centimeters long, the  $z$  fanouts are more complicated. They must extend the full length of the detector modules in order to connect all the  $z$  strips and to provide interconnections, or ganging, in cases where the number of available readout channels exceeds the number of  $z$



**Figure 4-11.** Schematic drawing of fanout connection to  $z$  strips showing ganging.

strips. Figure 4-11 depicts the concept of ganging. Two or three  $z$  strips (dashed lines) are connected in series by the fanout traces (solid lines), bringing the signals to the end of the detector module where they can be wire-bonded to the front-end electronics IC.

Table 4-5 gives a detailed list of the required fanout circuits with their geometrical features. There are four possible types of fanout circuits per layer ( $z/\phi$  and forward/backward), though in many cases, the forward (F) and backward (B) circuits are identical. The number of input readout strips is shown together with the number of front-end electronic channels available in the corresponding readout sections. The number of  $z$  strips always exceeds the corresponding number of readout channels: therefore ganging is required. The number of  $\phi$  strips is often smaller than the corresponding number of available channels. In such cases, some of the readout channels are not used. Typical pitches at the input and output bonding pads are given; the line pitches can be smaller in critical places. The total number of fanout circuits is 208; the number of different circuit layouts is 28.

Similar circuits have already been realized by industry as very thin, high precision Kapton flex cables [Lev] for the L3 Silicon Microvertex Detector [DiB95] and the Aleph VDET200 detector [Bag94]. Alternative solutions exist but are not suited to our design. Double-metal fanouts integrated on the detectors exhibit much larger parasitic capacitances. A thin glass substrate has interesting properties, but in our case flexible circuits are needed to route the signals to the readout electronics.

The base material for the  $z$  fanout of Reference [DiB95] is  $50\ \mu\text{m}$ -thick type-E kapton, plated with a  $0.25\ \mu\text{m}$  nickel adhesion layer followed by  $2.5\ \mu\text{m}$  of electroplated copper. High precision masks were realized by electron beam photolithography with a typical geometrical resolution of  $1\ \mu\text{m}$ . The required conducting lines were chemically etched to a width of about  $11\ \mu\text{m}$ , starting from line widths of about  $20\ \mu\text{m}$  on the masks. The pitch of the lines could be as small as  $33\ \mu\text{m}$ . Bonding pads were electroplated with  $2\ \mu\text{m}$  nickel followed by  $1\ \mu\text{m}$  of high-purity amorphous gold.

This technology is satisfactory both for the mechanical and the electrical properties. The average radiation thickness of the fanout circuit is only  $2.3 \times 10^{-4} X_0$ , and the degradation of the signal-to-noise ratio due to the fanout contribution to the capacitance and resistance

Layer	Fanout Type	Length (cm)	Number of Readout		Typical Pitch at		Number of Circuits
			Strips	Channels	Input( $\mu\text{m}$ )	Output ( $\mu\text{m}$ )	
1	$z$ , F+B	12.5	950	768	100	50	12
	$\phi$ , F+B	3.0	768	768	50	50	12
2	$z$ , F+B	14.5	1150	1024	100	50	12
	$\phi$ , F+B	3.0	960	1024	50	50	12
3	$z$ , F+B	15.6	1360	1280	100	50	12
	$\phi$ , F+B	2.0	1280	1280	50	50	12
4a	$z$ , F	19.7	885	512	200	50	8
	$z$ , B	24.3	1115	512	200	50	8
	$\phi$ , F+B	2.0	512	512	65	50	16
4b	$z$ ,F	20.6	930	512	200	50	8
	$z$ ,B	24.2	1160	512	200	50	8
	$\phi$ , F+B	2.0	512	512	65	50	16
5a	$z$ , F	25.2	1160	512	200	50	9
	$z$ , B	25.1	1205	512	200	50	9
	$\phi$ , F+B	2.0	512	512	65	50	18
5b	$z$ , F+B	26.1	1205	512	200	50	18
	$\phi$ , F+B	2.0	512	512	65	50	18

**Table 4-5.** Summary of fanout circuit characteristics.

is acceptable. For example, the ALEPH Kapton fanout circuits contribute 0.6 pF/cm to the interstrip capacitance and 0.1 pF/cm to the strip-to-back capacitance. An acceptable production yield (about 50%) could be achieved [DiB95] by allowing a maximum of 1.5% of optically detected defects (shorts or breaks in the lines), most of which could be repaired.

One potential drawback of this approach is that the additional wirebonds required to connect the fanout can result in a loss of channels due to faulty connections if great care is not taken in assembly. These aspects must be further investigated in 1995. It is also necessary to perform more accurate measurements and calculations of interstrip and back capacitance for the fanout circuits. In addition, mechanical and electrical tests will assess possible limits in the sharpness of bending. Other substrates, such as Upilex, may offer superior characteristics and better production yield; these are currently being investigated.

#### 4.4.4 R&D on Detectors and Fanouts

Although the technology is rather stable and reliable, some R&D and prototyping must be done before freezing the design by the end of 1995. We plan to submit a prototype batch of AC-coupled, double-sided detectors early in 1995. The mask design will include several structures, including a detector which is almost full size ( $\sim 15 \text{ cm}^2$ ) for use in a test beam in the summer of 1995, as well as several test structures to measure capacitances and leakage currents for various strip pitches and implant widths. In addition, a small prototype wedge detector and various test devices for process characterization and radiation studies will be included. Prototype fanouts will also be employed in this test beam.

Following the first prototype submission, there will be a second submission consisting of two batches, with one full-size detector per batch. These detectors will be put into a later test beam together with prototype readout electronics.

### 4.5 Electronic Readout

---

#### 4.5.1 Introduction

The SVT readout electronics must:

- Process the signals that appear on the readout strips with the aim of digitizing the charge;
- Temporally correlate the detector signals with an externally generated Level-1 trigger;
- Format the vertex data, which consists of the channel ID and the digital values of the charge and time-stamp, and transmit them to the DAQ system upon receipt of a read command; and
- Remain functional up to 10 times nominal background with graceful degradation at higher background levels.

The charge information is retained for several reasons. One important reason is to correct for comparator time-walk; the pulse arrival time must be measured offline as accurately as possible in order to minimize the number of background hits associated with an event. A second reason is to improve, by interpolation, the position measurement of hits for which there is charge-sharing between adjacent readout strips. Finally, charge information may also be used to reject background hits and to correlate  $\phi$ - and  $z$ -side hits.

The readout electronics are subdivided into readout sections. A readout section services the  $\phi$  or  $z$  strips for one-half of a detector module (either the forward or backward end). The vertex readout electronics consists of a monolithic readout chip; a hybrid circuit, which provides the mechanical support for the chips and the substrate for connections to the chip; a transition card, which interfaces the front-end readout to the DAQ system; and ancillary units, such as power supplies and detector bias supplies. The readout chip, the hybrid, the transition card, and the power supplies will be described in the remainder of this section.

For further information on the choices that underlie the design of the readout electronics and the relevant architectural details, the reader is referred to the relevant *BABAR* notes [Lev94, Roe94b, Joh95a, Joh95b].

## 4.5.2 Readout Chip

### Functional Overview

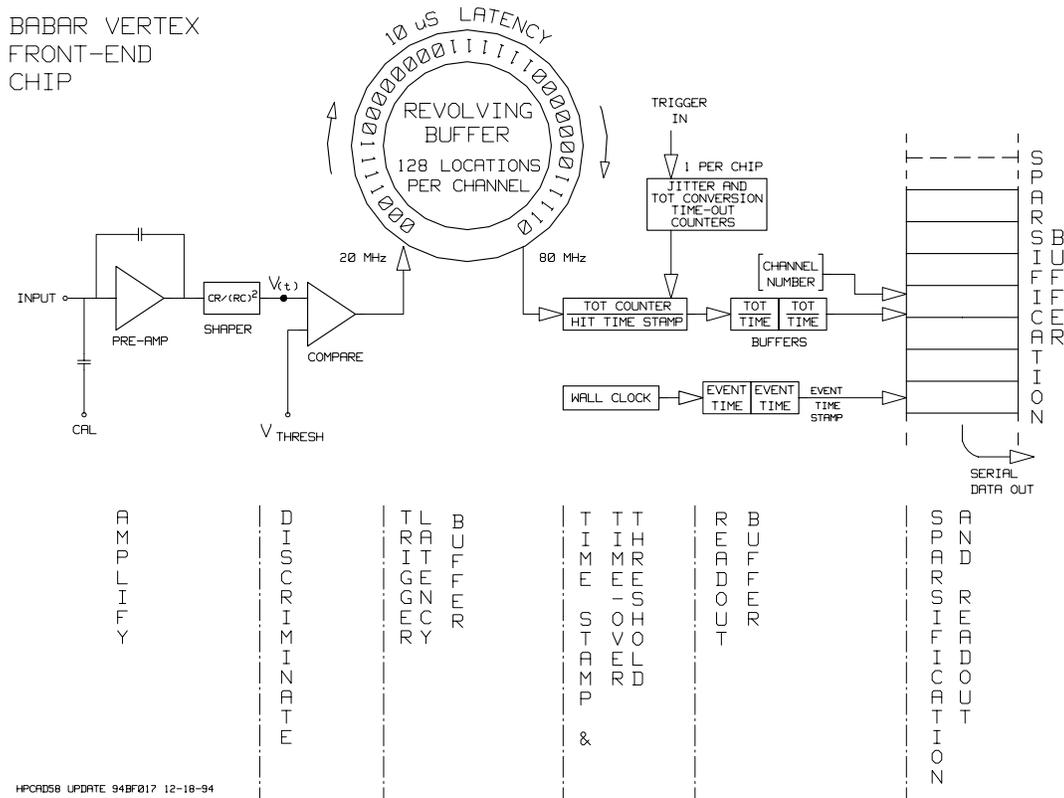
The readout chip must amplify and digitize the input signals for all channels in parallel. In addition, the signals must be buffered for the duration of the Level-1 trigger latency. For those hits selected by the trigger, the digitized data must be buffered until a read command is received. Upon receipt of the read command, only signals which exceed a pre-defined threshold should be read out. Since the time between beam crossings will be very small (4.2 ns), none of the operations of the chip can be restricted to occur between beam crossings. Furthermore, the maximum expected Level-1 trigger rate is too high to allow data acquisition to be suspended during readout operation. Therefore, during normal operation of the chip the analog front-end is actively acquiring data while digitization, buffering and readout occur.

### Requirements

The requirements for the readout chip are summarized below. For more information, see Reference [Joh95a].

1. Mechanical requirements:
  - Number of channels per chip: 128;
  - Chip size: total width  $< 6.2$  mm, and total length  $< 9$  mm; and
  - Channel-to-channel pitch:  $50 \mu\text{m}$ .
2. Operational requirements:

- Operating temperature: maximum 40°C;
  - Radiation tolerance: up to 2 Mrad (ten years at 200 krad/yr); and
  - Power dissipation: not to exceed 2 mW/channel.
3. Dynamic range: Minimum input charge is 0.75 fC (0.20 MIP), and maximum is 19 fC (5 MIP); circuit must accept either positive or negative signals. It is desirable to increase the dynamic range to 38 fC (10 MIP) if feasible.
  4. Analog resolution: 5 bits on a linear scale are required to achieve the required resolution of 0.25 MIP for the above dynamic range, while a minimum of 3 bits are necessary for a nearly optimal position measurement with range compression in a nonlinear system.
  5. Time-stamp resolution: 100 ns time-stamp resolution is required for the inner layers, increasing to a maximum of 400 ns in the outer layers to match the pulse shaping times.
  6. Noise performance: The noise must not exceed 800  $e^-$  rms at a risetime of 100 ns with an external capacitance of 10 pF, an external resistance of 100  $\Omega$ , and a leakage current of 100 nA. It must not exceed 1200  $e^-$  rms at a risetime of 400 ns with an external capacitance of 50 pF, an external resistance of 200  $\Omega$ , and a leakage current of 200 nA. (These specifications refer to the inner and outer layers of the SVT, respectively.)
  7. Maximum data rate: Simulations show that the lost-particle backgrounds dominate the overall rates; at nominal background levels, the maximum hit rate per strip is about 300 kHz.
  8. Deadtime limits: The maximum total deadtime of the system must not exceed 10% at a 10 kHz trigger rate and at 10 times the nominal expected background rate.
  9. Double-pulse resolution: The maximum inefficiency from pulse overlap must be less than 3%.
  10. Trigger specifications: The trigger has a nominal latency of 9.5  $\mu$ s, a maximum jitter of 0.5  $\mu$ s, and the minimum time between triggers is 1.5  $\mu$ s. The maximum Level 1 trigger rate is nominally 2 kHz, but a conservative upper limit of 10 kHz is assumed.
  11. Output data format:
    - only hits above threshold are transmitted (sparse readout); and
    - the data packet should include a synchronization code, chip ID, time stamp, data, status word, and a trailer.
  12. Calibration requirements: Provision must be made for injecting a known amount of charge at the inputs of an arbitrary, remotely selectable set of channels.



**Figure 4-12.** Block diagram of the silicon vertex detector readout chip.

- Failure recovery: It is desirable to reduce the risk of failure by employing conservative design methods and building in redundancy where possible.

## Readout Chip Implementation

The chip described at a block diagram level in Figure 4-12 is intended for implementation in a  $0.8\ \mu\text{m}$  CMOS radiation hard process and has been designed to comply with the requirements discussed above [Joh95b]. Each chip consists of 128 channels together with additional circuitry that is common to all channels. The chip operation is synchronized by an external clock, which is frequency-divided on the chip to obtain synchronization of different operations at different frequencies.

Each channel consists of an analog section, a comparator with a setable threshold ( $V_{th}$ ), a trigger latency buffer, a counter, and a back-end buffering section. The analog section consists of a charge-sensitive preamplifier followed by a  $\text{CR}(\text{RC})^2$  shaping amplifier, which implements a second-order semi-Gaussian impulse response. The shaping has a twofold

	$t_{peak}$ (ns)	Length (cm)	p/n	Cap. (pf)	ENC( $e^-$ )	Signal Loss
Layers 1–3 $z$	100	5	p	7.7	400	10%
Layer 3 $\phi$	200	15	n	35.5	730	29%
Layer 5 $z$ , x3	400	15	n	29.3	630	10%
Layer 5 $\phi$	400	25	p	24.7	590	33%

**Table 4-6.** *Equivalent noise charge for various strip configurations.*

purpose: to improve the signal-to-noise ratio (S/N) over that obtained at the preamplifier output, and to provide a signal shape suitable for the TOT processing in the next section of the analog circuitry. The equivalent noise charge (ENC) after the shaper depends on the front-end amplifier/shaper characteristics and on the characteristics of the detector to which it is connected. Table 4-6 illustrates the ENC, which is calculated using a detailed SPICE model of both the analog front-end and the detector, for a variety of strip lengths found in the SVT. The simulation assumed an equivalent noise at the input transistor of  $1.9 \text{ nV}/\sqrt{\text{Hz}}$  and a post-irradiation leakage current of  $1 \mu\text{A}/\text{cm}^2$ , corresponding to approximately 1.5 Mrad of ionizing radiation. The signal loss is the fraction of the signal which is lost to the back of the detector and therefore will not be collected at the preamplifier. There is additional charge lost to the neighboring strips; however, this is typically smaller and will be seen, provided it is above threshold.

The function of a channel can be understood by following the signal path through it. The fast detector signal with charge  $Q$  appears at the output of the shaping amplifier with a time dependence given by:

$$V(x) = k \cdot Q \cdot x^2 e^{-2x}, \quad (4.1)$$

where  $k$  (in Volts/Coulomb) is the charge sensitivity of the preamplifier-shaping amplifier combination, and  $x = t/t_p$  is the time normalized to the peaking time  $t_p$ . The value of  $t_p$  is selectable by an external chip control in order to comply with noise, deadtime, and double-pulse-resolution requirements.

The shaped pulse is compared with the preset threshold  $V_{th}$ , and if it exceeds  $V_{th}$ , it produces a TOT at the comparator output. The TOT is an analog variable which carries information about  $Q$ . The relationship between  $Q$  and TOT is nonlinear and is not far from a logarithmic dependence. This is desirable because it provides good resolution for small  $Q$  and reduced resolution for large  $Q$ , effectively compressing the range into fewer bits.

A simulation was performed to evaluate the ratio of the pulse width to the sampling period which was required to achieve good position resolution using TOT information [Roe94b]. The position resolution as a function of the crossing angle between the track and the silicon strip detector was evaluated for several different assumptions: perfect linear analog readout, digital readout (hit/no-hit information only), and TOT with variable  $t_{peak}/t_{clock}$ . For the

TOT simulation, a CR-RC<sup>2</sup> shaper was assumed. The results are shown in Figure 4-4; it is observed that TOT with  $t_{peak}/t_{clock} = 2.5$  is essentially as good as a perfect linear analog readout.

It should be pointed out that an ENC-based noise specification makes sense so long as it is restricted to the linear section consisting of the preamplifier and shaping amplifier and refers, for example, to the threshold at which the resulting pulse is discriminated. For the accuracy of the TOT-based charge measurement, a different S/N must be defined:

$$(S/N)_{TOT} = TOT / \langle \sigma \rangle, \quad (4.2)$$

where  $\langle \sigma \rangle$  is the rms, noise-induced dispersion on the TOT. As a first approximation,  $(S/N)_{TOT} = (Q - V_{th}/k)/ENC$ , where  $k$  is the gain of the preamplifier/shaper combination in volts/Coulomb.

The TOT information lends itself to a straightforward digitization, for instance, by comparison with the period of a reference clock. TOT digitization and trigger delay compensation are performed by the revolving buffer of Figure 4-12, in which the status of the TOT comparator is stored under the supervision of a write pointer operating at the frequency of the reference clock. The number of locations in the revolving buffer is given by the maximum expected trigger delay, taking into account the trigger latency, trigger jitter, and the front-end chip jitter, divided by the period of the writing clock.

The readout operation is initiated when a trigger is received, beginning with the generation of the event time stamp, obtained by latching the contents of a 16-bit counter (one per chip) which continuously counts the write clock (the time stamp counter in Figure 4-12). A timed sequence ensues, controlled by the trigger-jitter counter and the TOT full-scale counter, both of which are common to all channels in the chip. The read pointer for the revolving buffer is set at the position corresponding to the earliest possible data in the buffer, taking into account the fixed trigger latency, the trigger jitter, and the position of the write pointer at the instant of the trigger arrival.

The read pointer is controlled by a clock at a higher frequency than the write pointer in order to minimize deadtime, because no new triggers can be received while the read pointer is active. The data associated with the trigger are retrieved from the revolving buffers by detecting a zero-to-one transition in the revolving buffer. When a zero-to-one transition is found, three actions take place:

- A hit flag is generated;
- The TOT counter is enabled to count the number of ones in the buffer; and
- A hit time stamp is latched.

The number of bits counted corresponds to the digital value of the TOT and hence, of the charge  $Q$ . Based on simulation of the TOT technique [Roe94b], it has been determined that an adequate resolution for position measurement is obtained with between 3 and 5 bits, so 4 bits are envisaged to represent the digital value of the TOT. The hit information in the individual channels consists of 10 bits; five of them represent the hit time stamp, four represent the TOT value, and one is for the hit flag.

The back-end section of each channel includes additional buffers into which the hit data can be stored while awaiting readout. If there is no backlog of data to be read out, the hit data are completed by adding the information about the channel number, which requires 7 additional bits, and transferred into a special buffer where sparsification takes place. Removing the hit flag, there are 16 bits per hit channel which are then to be transmitted upon receipt of a data transmit command.

### 4.5.3 Hybrid Design

#### Functional Overview

The hybrid is composed of several components which are hybridized into one electrical unit [Col94]. These include the high-density interconnect (HDI), the front-end chips and other components that are mounted on the HDI, the thermal interface between the chips, and the water-cooling system, and the cable or tail which connects the HDI to the transition card.

The hybrid is connected to the silicon detectors by flexible fanout circuits which bring the detector signals to the front-end chips. The hybrid is mechanically mounted onto the fiber composite beams which provide structural rigidity for the detector modules. The mounting of the detector modules onto the cooling/support cones is accomplished with pins and screws which are inserted from inside the cone through the aluminum mounts of the cone. The screws pass through the thermally conductive substrate of the hybrid and into the rigid mechanical structure of the module. This mechanical support thus provides the thermal contact that allows heat from the ICs to flow into the cooling water circulating around the aluminum mounts. From this brief description of the functionality of the hybrid, it is apparent that the hybrid design is complicated by the fact that the requirements imposed by the mechanical support, the electronic readout, and the cooling must all be met in a self-consistent solution. The physical space available for the hybrid is very limited, imposing yet another constraint on the design.

### High-Density Interconnect and Thermal Interface

The HDI is a simple fanout circuit on which the readout chips are mounted. It provides connections for each chip to the data and control lines and to the low voltage supplies and grounds. It may also include some passive filtering and any external components required by the front-end chip. It is important that the layout of the HDI minimize crosstalk and noise, especially on the analog power and ground planes. The HDI must also provide connections to the detector for the detector biasing voltages and provide a local, low-impedance connection from the  $\phi$ -side to  $z$ -side preamplifiers for the detector current return path. There must also be a low-impedance thermal path from the front-end chips to the aluminum mounts which protrude from the cooling/support cone; this thermal path will either be through the HDI itself or through a substrate to which it is laminated.

There are five different HDI layouts required, one for each of the five layers. The same HDI layout is used for both  $\phi$ - and  $z$ -side readout within a layer. There are several options for the implementation of the HDI. It may be a flexible kapton circuit laminated to an aluminum heat sink, or a thick-film ceramic circuit fabricated from a thermally conductive material such as AlN or BeO. Thin-film circuitry on a ceramic substrate is also an option. At present, these design options are being studied, and an optimum choice in terms of performance, ease of fabrication, and cost is being determined.

The thermal path from the front-end chips to the heat sink has been evaluated with Finite Element Analysis (FEA) simulations, assuming a flexible kapton circuit laminated to an aluminum substrate. A simple model was built and tested experimentally, and the FEA calculation was found to agree well with the measured temperatures. Having validated the simulation, a more detailed FEA calculation was performed, taking into account the precise dimensions of the hybrid and all of the layers in the multilayer kapton circuit. The calculations indicate a temperature drop of about 15°C from the IC to the cooling water. Given the maximum chip operating temperature of about 40°C, this is more than adequate. The design is being further refined, and more experimental tests are planned.

### Hybrid Tail

The tail is the connection between the HDI and the transition card (described below). It is envisioned as a flexible multilayer circuit which will have power and ground planes and one or more signal planes. It will be fabricated separately from the HDI itself and will be connected to two HDIs by means of bump-bonds, solder, or other permanent connection. The connection to the transition card will utilize a mateable connector. The biggest constraint on the design of the tail is the limited space available for cables from the inner SVT layers to pass between the cooling/support cone and the outer detector modules. This restricts the width of the cable to less than about 0.75 cm, and the thickness to less than about 0.3 mm.

The trace impedance, which is about  $30\ \Omega$ , must be matched to the drivers and receivers at either end.

#### 4.5.4 Data Transmission

##### Data Transmission Requirements

The main requirements for the data transmission system are given below.

1. Data rate: between 60 and 80 Mbits/s/readout section;
2. Power consumption (inside the detector volume): below 5 W/readout section;
3. Radiation resistance (inside the detector volume): the system must be functional within the specified performance after being exposed to a 100 krad integrated dose of ionizing radiation; and
4. Error rate: less than  $10^{-11}$ /bit. This would imply roughly 1 bit error in every  $10^6$  events transmitted.

##### Functional Overview

Two different designs have been studied for transmitting data from the SVT front-end readout IC to the DAQ system and also to send clock, trigger and control signals from the DAQ system to the readout IC. The two designs differ primarily in where the transition between electrical and optical transmission occurs.

In the first design which was considered, the transition from electrical to optical transmission was made on transition cards which sit inside the support tube at the outboard end of the B1 magnets. A flexible circuit carries electrical signals between the HDI and the transition card, and optical fibers connect the transition card to the DAQ system. This option has the advantage of avoiding noise pick-up in the 5–6 m cable run from the transition card to the outside of the detector, and reducing the volume of cable in the space-constrained area inside of the support tube. However, it requires customized packaging and use of radiation-hardened opto-electrical components on the transition card which are very costly.

In the second design, electrical transmission is used from the HDI all the way outside of the detector to a point where crates can be mounted to support a commercial optical transmission system. The transition cards inside of the support tube are still needed; however they would have little or no active circuitry. Because this option can make use of commercial fiber optic

systems, the cost is reduced and it is possible to use optical transmission systems common to other sub-detectors. The opto-electric system is also accessible for repair and does not need to be radiation hard.

The transition cards are required in both designs in order to provide a place where the hybrid tail can terminate. In the second option, if the impedance of the hybrid tail and the transmission cables extending beyond the transition card can be matched sufficiently well, all active components can be removed from the transition card thus improving the reliability of the system inside the support tube where access is limited.

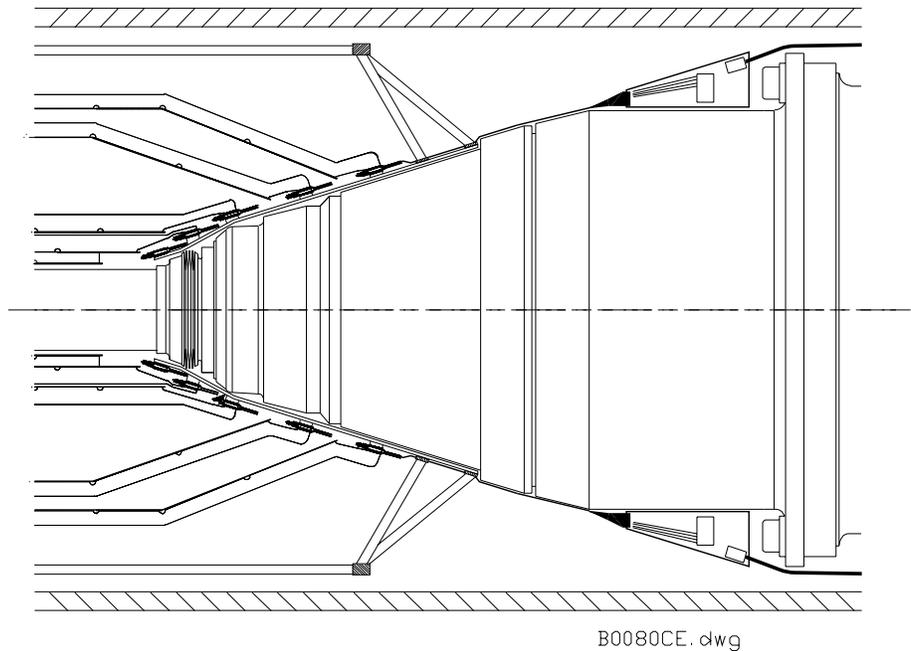
After studying the space budget for cabling beyond the transition card and the feasibility of the readout ICs transmitting signals electrically out of the support tube area, the second option has been chosen as the baseline design.

#### 4.5.5 Baseline Design

The baseline design for transmitting data out of the detector starts with the readout ICs transmitting differential signals on pairs of traces up the hybrid tail to the transition card. The signals are there transferred to conventional paired cables for transmission to an interface card outside of the detector. Likewise, clock and control signals coming from the interface cards will be transmitted as differential signals on conventional cable to the transition card inside the support tube and then transferred onto the hybrid tail to the HDI and the readout ICs.

The hybrid tail is still under development, and its precise electrical characteristics are still to be determined. The design of the hybrid tail is constrained by mechanical requirements and limited by the available technology for producing flexible cables. If the impedance of the hybrid tail can be closely matched to that of available cabling, the signal transmission through the transition card will be only passive. If the impedance mismatches are too great, commercially available bipolar components have been identified to receive and re-transmit signals on the transition card. Both the design of the transmitting and receiving circuits on the readout ICs and the choice of components for the transition card have been made to maintain the consistent differential, low power, low noise transmission signals. The identified bipolar components for the transition card will meet the speed, power and radiation requirements of the system.

The transition cards will not only provide the physical connection between hybrid tail and conventional cable and any necessary impedance matching for signal transmission, it will also provide necessary filtering for all the DC voltages bussed to the detectors and the readout ICs. There will be one transition card for each hybrid tail. They will be located inside the support tube, approximately 20 cm from the outside layer of the SVT. They are mounted



**Figure 4-13.** Side view of the connections from the readout electronics to the transition card, which is a triangularly shaped circuit mounted from the B1 magnet.

from a cooling ring which is attached to the far end of the B1 magnet. See Figure 4-13 for a side view showing the connections from the readout electronics to the triangularly shaped transition card. A multi-conductor cable bundle for each transition card will carry all necessary power, sense and data lines from the transition card through the support tube to a point outside the detector.

The interface cards will provide transition between electrical signals and optical signals. The exact location for these cards has not been determined but they will be located approximately 5–6 m from the interaction point, outside the detector. Eight data lines will be multiplexed onto one gigabit fiber optic link for transmission to the DAQ system. Likewise, clock and control signal coming from the DAQ on a fiber will be fanned out and transferred to electrical cables going into the detector. The exact level of multiplexing between electrical signals and optical is made to match bandwidths and partitioning of DAQ functions.

### 4.5.6 Power Supplies

The power supply system for the SVT must provide the voltage sources with adequate current levels for two purposes: biasing the silicon detectors, and supplying power to the front-end chip, transition card, and remote opto-electronic circuits. The low-noise electronic readout for the silicon vertex detector imposes stringent requirements on the power supplies, and they must be carefully designed to prevent common-mode noise from entering the system.

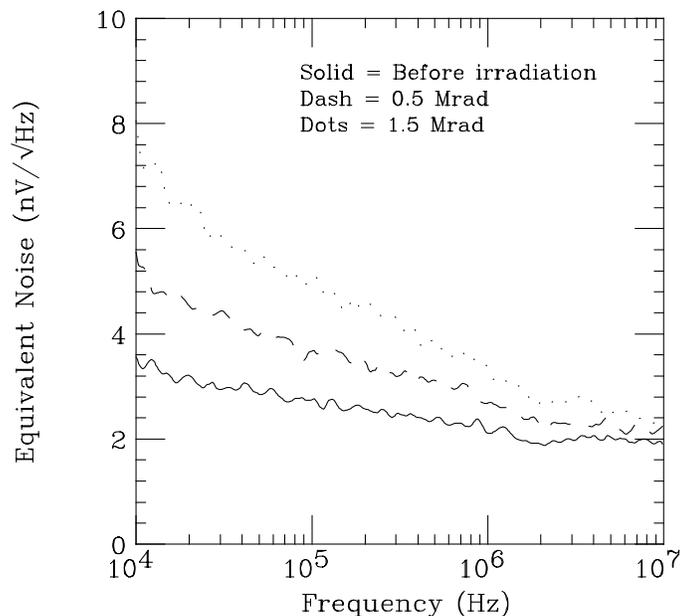
Both supply subsystems must be built up from modular, fully floating units without galvanic connection between the utility side and the load side, with the two sides electrostatically shielded. Preference is generically given to continuous-type voltage regulators in view of their lower noise levels. However, tests are foreseen to determine whether smaller switching regulators with better power yield can also provide satisfactory performance.

The power supplies must also feature remotely controlled, digitally preset voltages. Optical connection of the voltage controls is recommended. The power supplies must be capable of sweeping the output voltage throughout the nominal range as specified below.

The specifications for the two different supply subsystems are as follows. For the detector bias supply, the nominal full scale voltage is 80 V per module, with a setting resolution of 7 bits and maximum current of 0.1 A. Ripple must be less than 10 mV peak to peak, and noise must be less than 2 mV rms. For the electronics power supply, the nominal full scale voltage is 6 V per module, with a setting resolution of 8 bits and maximum current of 100 A. Ripple must be less than 2 mV peak to peak, and noise must be less than 500  $\mu$ V rms.

### 4.5.7 Electronics R&D

The R&D to support the development of the electronic readout system described above has already begun. Characterization of radiation-hard CMOS processes is well underway. Test structures have been submitted to Honeywell and to UTMC to characterize their radiation-hard processes. This work was begun as part of an R&D program for SDC. The UTMC circuits have already been evaluated, and the work is continuing with irradiation and testing of the Honeywell chips. Honeywell has an 0.8  $\mu$ m, triple-metal process which is attractive both for its small minimum feature size and for the third metal layer which is useful for minimizing the circuit size and may also prove useful for shielding. Initial results with this process are promising. Figure 4-14 gives comparison of the noise performance of the Honeywell process before and after irradiation, plotted versus frequency. For the 100 ns shaping time which is foreseen for the inner detector layers, the corresponding frequency is about 3 MHz. At this frequency, the noise increases by about 15% after 500 krad and by



**Figure 4-14.** Noise levels for Honeywell transistor before and after irradiation.

about 30% after 1.5 Mrad of ionizing radiation. Further tests up to a total integrated dose of 5 Mrad are planned.

A first round of test chips on the TOT concept has been submitted and is being tested in early 1995. These test chips will be used to evaluate the TOT technique and to measure the extent to which digital activity on the chip during data acquisition can contribute to front-end noise. Work on a first full-scale prototype of the TOT front-end chip is underway, and the first submission is expected in the spring of 1995. This chip will be operated on a test bench in the fall of 1995 with prototype *BABAR* detectors and will then go into a test beam.

For the hybrid, a mechanical prototype is foreseen early in 1995 as a means of qualifying a vendor. This will be followed by a working electronic prototype in mid-1995, which is designed to work together with the first prototype TOT chip.

## 4.6 Mechanical Support and Assembly

---

An overview of the SVT mechanical support is provided in Section 4.2. In this section we provide a more detailed account of the constraints on the mechanical design due to

the accelerator components near the IP and describe the details of the detector assembly, installation, survey, and monitoring.

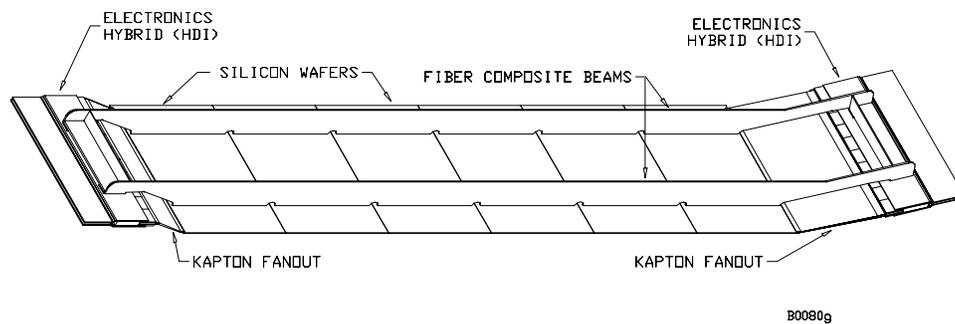
### 4.6.1 IR Constraints

The support structure design and configuration of the SVT is dictated by the configuration and assembly procedure of the machine components near the interaction point, as well as by the SVT geometry. The close spacing of the bunches in PEP-II (4.2 ns) and the desire to avoid parasitic crossings dictate the need for a pair of permanent dipole magnets located about 20 cm from the IP on either side. These B1 magnets occupy most of the region below  $17.2^\circ$  (300 mr). In order to minimize the mass inside the active tracking volume, it is desirable to mount all of the electronics below the 300 mr line. In the forward direction, this requires that electronics, cooling, cabling, and support be confined in a volume one centimeter thick around the B1 magnet. The use of this small space below the 300 mr line must be carefully coordinated with the needs of the accelerator. In the backward direction, coverage to within  $30^\circ$  of the beam pipe leaves room for some machine components which are moved out of the forward region. The solid angle coverage of the SVT is therefore restricted to the region  $17.2 < \theta < 150^\circ$ .

A further constraint on the SVT is imposed by the PEP-II support tube, which is a tube of radius 20 cm that aligns the final magnets closest to the IP. The central part of the support tube is a thin carbon-fiber structure, while the outer sections are stainless steel. The SVT must be supported inside the carbon-fiber support tube and therefore cannot be mounted from the drift chamber endplates. In the baseline design, the SVT is supported from the B1 magnets. Since the SVT must be installed with both B1 magnets in place, it must be assembled in two halves and then clam-shelled around the beam pipe. The assembly and alignment of the beam pipe, B1 magnets, Q1 magnets, and SVT inside the support tube will take place in a staging area away from the interaction hall. The entire assembly will then be transported and installed in the interaction hall. This procedure is reversed in order to gain access to the SVT for repair.

### 4.6.2 Module Assembly

The SVT is constructed from detector modules, each of which is mechanically and electrically independent of the other modules. Each module consists of silicon wafers bonded to fiber composite beams, with a high density interconnect (HDI) electronic hybrid at each end. The HDIs are electrically connected to the silicon strips by means of flexible circuits and are mechanically supported by the fiber composite beams. The entire module assembly is a rigid



**Figure 4-15.** *Detector module from Layer 3, consisting of six silicon detectors wirebonded together and read out at each end.*

structure that can be tested and transported with relative ease. A drawing of a detector module from Layer 3 is shown in Figure 4-15.

Assembly of the detector modules begins with the preparation of the necessary parts. The silicon detectors must be fully tested, including a long-term test under full bias voltage. The fanout circuits will be optically inspected and single strip tested for shorts/opens. The readout hybrids must be assembled and tested, starting with the HDI circuit, the front-end chips, any additional passive components, and the hybrid support. Finally, the completed beams, which provide mechanical stiffness, must be inspected to ensure they meet specifications. These individual parts will be fabricated at different institutions from which they can easily be shipped to the institutions at which the module assembly is carried out. The hybrids will be retested after shipment.

The assembly of the inner barrel-shaped modules and the outer arch-shaped modules is necessarily different. However, there are common steps. Generally, the procedure is as follows:

1. The  $z$  and  $\phi$  fanouts are glued to the detectors and wire-bonded to the strips. The ganging bonds between  $\phi$  strips are performed.
2. The silicon detectors and readout hybrids are held on a suitable fixture and aligned relative to each other. The fanouts are glued to the hybrids and wire-bonded to the input channels of the readout ICs. Electrical tests, including an infrared laser strip scan, are performed and the detector-fanout assemblies (DFAs) are visually inspected.
3. The final assembly stage is different for different layers. For modules of Layers 1 and 2, the DFA is bonded to the fiber composite beams with appropriate fixtures ensuring alignment between the mounting surfaces on the HDI and the detectors. The module is again tested. A module from Layer 1 and Layer 2 are then joined together by gluing

the beams on the top of Layer 1 to the bottom of Layer 2. The combined structure is called a sextant module. For Layer 3, the DFA is bonded to the fiber composite beams as with Layers 1 and 2.

For the modules of Layers 4 and 5, the DFA is held in a suitable fixture and bent at the corners of the arch and at the connection to the HDI. The module is tested. The fiber composite beams are bonded to the module with fixtures assuring alignment between the silicon detectors and the mounting surfaces on the HDI. This procedure has been performed with real detectors that have been successfully operated in a test beam.

Once completed, these detector modules are extremely rigid devices that can be stored and subjected to long-term testing. They are then shipped to the location for final assembly and installation of the detector.

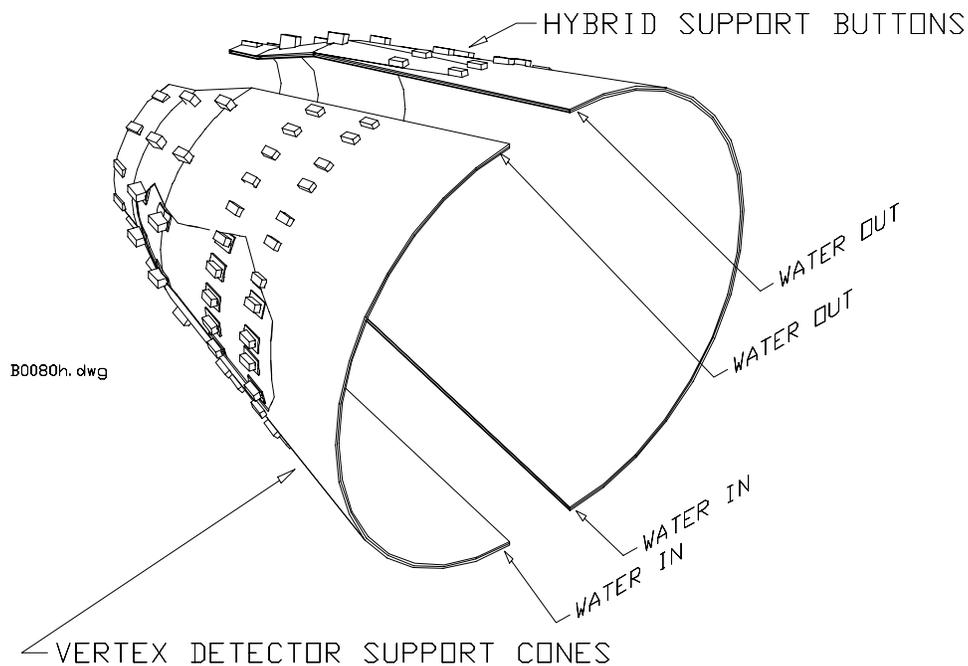
### 4.6.3 Detector Assembly and Installation

#### Half-Detector Assembly

The detector is assembled in halves in order to allow the device to be clam-shelled around the beam pipe. The detector modules are supported at each end by cooling/support cones constructed from double-walled carbon-fiber laminates. Cooling water circulates between the two carbon-fiber shells and around aluminum mounting pieces which protrude through the outer shell. The cones are split along a vertical plane and have alignment pins and latches that allow them to be connected together around the B1 magnets. See Figure 4-16 for a drawing of the cooling cone. The two carbon-fiber support cones are mechanically connected by a low-mass carbon-fiber space frame.

During the half-detector assembly, the two half-cones will be held in a fixture which holds them in precise relative alignment. The detector modules are then mounted to the half-cones at each end. A manipulating fixture holds the detector modules during this operation and allows for well-controlled positioning of the module relative to the half-cones. Pins located in the aluminum mounting pieces provide precise registration of the modules, which are then screwed down. Accurate alignment of the mounting with respect to the silicon wafers is achieved by a pair of mating fixtures. One is a dummy module and the other simulates the mating surfaces on the cone. These fixtures are constructed together and mate perfectly. One is used to verify the machining of the aluminum mounting pieces on the cones. The other is used to position the mounting points on the HDI during the assembly of the modules as mentioned above.

The connection between the module and the cone (called the foot) provides for accurate and reproducible alignment of the module and conduction of heat from the HDI heat sink to the



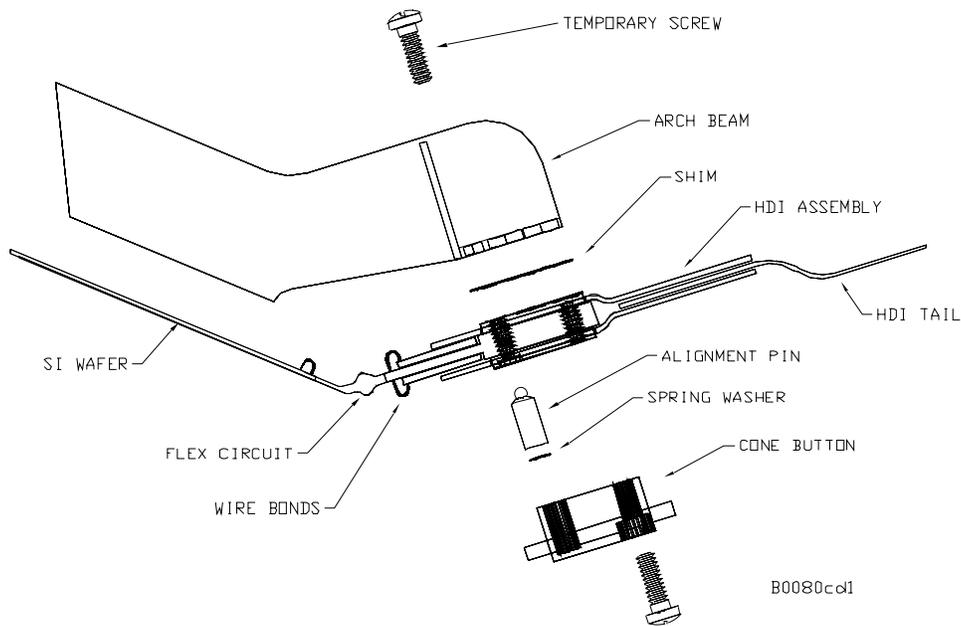
**Figure 4-16.** Schematic view of the cooling/support cone. The cone is constructed in two halves and clam-shelled together. Water circulating between the two carbon-fiber shells cools the readout electronics, which are mounted on aluminum pieces protruding through the outer shell.

cooling water circulating between the cone layers. A detail of the foot region, which contains the readout electronics and the mounting pieces, is shown in Figure 4-17.

After verification of the alignment, the connection between the HDI and the support beams is permanently glued. The glue joint allows for the correction of small errors in the construction of either the cones or the modules. After the beam is glued, the module may be removed and remounted on the cone as necessary. The design of the foot allows this glue joint to be cleaved and remade should major repair of the module be required. After each detector module is mounted, it is electrically tested using a laser scan to verify its functionality. As each layer is completed, it is optically surveyed and the data are entered into a database. Finally, the two half-cones are connected together with the space-frame, resulting in a completed half-detector assembly.

### Mount to B1 Magnets and Transport to IR

When the two half-detector assemblies are complete, they are brought to the staging area where both B1 magnets and one of the Q1 magnets have been assembled onto the beam



**Figure 4-17.** Exploded view of foot showing mounting pins and connection to the detector module.

pipe. Fixtures are employed to hold the cones as they are brought together and clam-shelled around the beam pipe. The two half-detector assemblies are mated, and the latches between them are closed. The cables from the HDIs are routed to the transition cards, which are mounted in cooling manifolds at the ends of the B1 magnets. The entire detector is then thoroughly tested, and an optical survey is performed. After the survey is complete, the support tube is slid over the entire assembly and mated to the Q1 magnet. The final Q1 magnet is then installed. At this point, the assembly is relatively rigid and can be transported to the interaction hall and installed in the accelerator.

The detector assembly as described above forms a rigid structure as long as the cones and space frame are connected together. This structure is supported on the B1 magnets. During transport of the support tube assembly to the interaction hall, it is possible for the magnets to have as much as a 1 mm relative motion [Bow95]. This motion is reversible, and they will return to their original alignment when installed in the accelerator. In addition, differential thermal expansion may affect the relative alignment of the magnets during periods in which the temperature is not controlled, and relative motion of the magnets and the beam pipe may occur should there be seismic activity.

The support of the detector from the magnets must allow for this motion without placing stress on the silicon wafers. In addition, the position relative to the IR must be reproducible when installed in the accelerator. These constraints are met by mounting the support cones

on a pair of gimbal rings. One gimbal ring connects the forward cone to the B1 magnet so as to constrain its center in  $x$ ,  $y$ , and  $z$ , while allowing rotation about the  $x$  and  $y$  axes. A second set of gimbal rings supports the cone in the backward direction in a similar manner, with an additional sleeve that allows both for motion along  $z$  and rotation about the  $z$  axis, relative to the B1 magnets.

### Installation of Complete Assembly into Detector

The clearances between the vertex detector and the beam tube and B1 magnets are on the order of 1 to 3 mm. During transport of the support tube assembly, the critical clearances must be monitored in real time to ensure that no accidental damage to the detector occurs.

In its final position, the support tube assembly will be supported from the ends. Thus it is necessary that it is always supported only from the ends during installation. One possible installation scenario employs beams attached to the ends of the Q1 magnets which support the entire assembly from the ends. The beams are threaded through the drift chamber, resting on wheeled guides or tracks, which have been previously aligned with respect to the *BABAR* detector. Once the support tube is through and supported at both ends, the temporary assembly beams are removed.

## 4.6.4 Detector Placement and Survey

### Placement Accuracy

The SVT must provide spatial resolutions on the order of  $10\ \mu\text{m}$ . Final locations of each of the wafers relative to each other and to the IR will be determined by track survey. This requires a certain degree of overlap of the modules within a layer. There must be overlap in  $z$  as well as  $\phi$ , so as to accurately locate the  $z$  positions of the wafers in a single module with respect to each other. These requirements are discussed in more detail below.

Mechanical tolerances and measurements must be such that the process of track survey converges in a reasonable time. Placement of the wafers within a module should be within  $25\ \mu\text{m}$  and optically surveyed to a few  $\mu\text{m}$ . Placement of a module relative to other modules should be on the order of  $100\ \mu\text{m}$  and be optically surveyed to within  $25\ \mu\text{m}$ .

Stability of the positions should be such that resurvey with tracks is rarely required. This leads to the requirement that the relative positions of the various wafers be stable to the  $5\ \mu\text{m}$  level over long periods of time (months or more). The position of the entire detector structure with respect to the IR can be followed more easily, so that variations on the order of a day can be tracked. However, longer time constants are desirable.

Stability of the detector components at the  $5\ \mu\text{m}$  level requires a stable operating temperature. Preliminary calculations for the thermal expansion of the entire structure predict on the order of  $0.5\ \mu\text{m}/^\circ\text{C}$  over the length of the active region of the detector. If the temperature inside the support tube is maintained at  $\pm 1^\circ\text{C}$ , thermal expansion will not be a problem.

### Survey with Tracks

Previous experiments have shown that a silicon vertex detector is best aligned using track data together with self-consistency constraints, because the point resolution in silicon is much more precise than that from a drift chamber. The most powerful alignment constraint comes from the overlap between wafers adjacent in  $\phi$ . Another important constraint comes from muon-pair events, where knowledge of the mass, boost, and common origin of the two-track system provides a means of linking hits on opposite sides of the vertex detector. A more detailed discussion of this subject can be found in Reference [Bro95].

The SVT alignment is parameterized as a pair of nested transformations for global and local coordinates. The global alignment of the SVT describes the position and orientation of each detector layer, considered as a rigid body, relative to the drift chamber. Previous experiments have shown that a good global alignment helps speed the convergence of the local alignment and allows a simple, low-statistics check of time-dependent effects such as detector motion. The procedure depends completely on the drift chamber tracking and simply requires consistency between SVT hits and the extrapolated drift chamber track positions. Systematic effects limit the accuracy of the global alignment to  $\sim 50\ \mu\text{m}$ . Assuming an accuracy of  $\sim 2\ \text{mm}$  for the drift chamber track extrapolation in the  $z$  direction, this implies that a global alignment can be performed with  $\sim 2000$  tracks. Further assuming two useful high momentum tracks per event, we find that 1000 hadronic events will suffice to perform a global alignment.

The local alignment describes the position and orientation of each silicon wafer relative to its nominal position in the layer. The local alignment of the SVT derives most of its statistical precision from the active area overlap between  $\phi$ -adjacent wafers. The overlap directly constrains three of the six geometric degrees of freedom of each wafer, namely those that define its position in the nominal wafer plane. A further strong constraint on the wafer translation in the out-of-plane direction comes from combining the overlap information for all wafers in a given layer. This is essentially a circumference constraint, using the fact that the size of the wafers is precisely known.

The number of events needed to perform the local SVT alignment can be roughly estimated. In the following, we assume that each wafer is aligned separately for six geometric degrees of freedom, and we further assume four useable medium-momentum tracks per event. We require that the alignment should not contribute significantly (less than 5%) to the SVT

Layer	$\Omega$	# of Wafers	# of Tracks/Wafer	# of Hadronic Events
1	2.4%	24	80	20K
2	1.8%	24	80	27K
3	1.8%	36	80	40K
4	4.0% (4.8%)	112	40	28K (24K)
5	2.0% (1.9%)	144	40	74K (77K)

**Table 4-7.** Fractional overlaps ( $\Omega$ ) of the SVT layers and the resulting number of hadronic events required for an accurate alignment. The values in parentheses for Layers 4 and 5 correspond to the wedge modules.

point resolution, implying that each wafer position must be known to better than 30% of its average point resolution. The overlap solid angle and number of wafers are different for each layer and are given in Table 4-7 together with the number of events needed. Layer 5 is seen to set the limit, requiring  $\sim 75$ K hadronic events to be precisely aligned. This is mainly because of the large number of wafers in this layer.

Overlaps do not constrain the relative positions of the different layers, nor the relative positions of the different annuli which comprise the layers. An annulus is here defined as the set of all wafers in a layer having the same  $z$  position. The relative positions of these sets of wafers are best obtained using  $e^+e^- \rightarrow \mu^+\mu^-$  events; we require an alignment precision of 30% of the intrinsic resolution for each annulus of a given layer. To estimate the approximate number of muon pairs required, we assume that the three parameters defining the muon trajectory have been measured with essentially no error on one side of the SVT, and this trajectory is compared to hits on the opposite side. The tracks are presumed to have 2.5 GeV/ $c$  momentum, and the material of each layer is presumed to be 0.5% of a radiation length on average. This gives the resolution of the track extrapolation due to multiple scattering as presented in Table 4-8, which is seen to dominate the total (extrapolation  $\oplus$  intrinsic) resolution. Layer 5 is again seen to require the most events to be accurately aligned, namely 7000 muon pairs or the equivalent of 27,000 hadronic events. Thus the overlaps set the overall minimum number of events needed for alignment.

## 4.6.5 Detector Monitoring

### Position Monitoring Systems

Although the final placement of the silicon wafers will be measured and monitored with charged particles which traverse the silicon detector and drift chamber, two displacement monitoring systems will be designed to measure relative changes in the position of the silicon

Layer	Track $\sigma$	Hit $\sigma$	# of Annuli	# of Muon Pair Events
1	25 $\mu\text{m}$	10 $\mu\text{m}$	4	225
2	30 $\mu\text{m}$	10 $\mu\text{m}$	4	325
3	50 $\mu\text{m}$	10 $\mu\text{m}$	6	1300
4	80 $\mu\text{m}$	30 $\mu\text{m}$	7	4000
5	100 $\mu\text{m}$	30 $\mu\text{m}$	8	7000

**Table 4-8.** *Properties of the different SVT layers, and the number of muon-pair events these imply for an accurate alignment.*

detector with respect to the machine elements and the support tube. One displacement monitoring system will be used to monitor relative positions during transportation of the support tube with the silicon detector inside it and during data taking. This system consists of either capacitive displacement monitors or LED-photodiode reflection monitors which are sensitive to relative displacements between the silicon detectors and the machine components such as the beam pipe, magnets, and support tube.

In addition, a laser system will monitor displacements of the outer layer of detectors with respect to the drift chamber during data taking. Given that the SVT layers are not mounted on the same support as the drift chamber, it is possible that motion between the two will occur. To monitor this motion, short infrared laser pulses are brought in with fiber optics (*e.g.*, 50  $\mu\text{m}$  core diameter) which are attached to the drift chamber. The laser light shines through small holes in the support tube and reaches the outer layers of the silicon detector. The resulting signals are read out with the normal silicon detector readout system.

Displacement monitoring systems based on capacitive sensors and laser pulses have been used successfully by many experiments in the past [Acc94, Bin93, Cac92, Bre91]. The two systems are complementary in technology and in their sensitivity to different kinds of displacements. For example, the L3 experiment [Acc94] attains a resolution of a few microns with the laser monitoring system for transverse displacements. With their capacitive monitoring device, they attain a resolution of 1–2  $\mu\text{m}$  in the radial direction and 5–10  $\mu\text{m}$  in the transverse direction.

### Radiation Monitoring

To protect the silicon detector system against potentially damaging beam losses, and to monitor the total radiation dose that the detectors and electronics receive, silicon diodes will be installed close to the beam pipe in the vicinity of the SVT. If the radiation dose exceeds a certain threshold, a beam-dump signal will be sent to the PEP-II control room. This sort of radiation protection system is already used in all of the LEP experiments. The silicon

diodes will be read out with various gains so that the dynamic range of the full set of sensors covers quiet running and potentially damaging conditions.

#### 4.6.6 R&D Program

The following R&D projects are planned before the design of the SVT mechanical configuration is finalized.

**Cable.** Prototypes of the cable from the hybrid to the transition card will be constructed. This will allow proof of the details of cable routing and mechanical robustness. It will also allow the electrical properties to be measured to verify simulations.

**Hybrid.** Realistic mechanical modules of the high-density interconnect (HDI) are required. The HDI is a critical element both in the cooling of the electronics and the mounting of the detector modules. Models will be tested for heat transfer capability and for module mounting schemes.

**Inner Layer Sextant.** A full-scale mock up of the inner layer sextant will be constructed. It will be used to verify thermal stability calculations and to investigate the effects of nonuniform beam pipe cooling. It will also be used to test and practice assembly techniques.

**Arch Modules.** Full-scale mock ups of the arch detector modules will also be constructed and used with the prototype cones to verify cooling and mounting techniques.

**Cones and Space Frame.** A set of prototype cones and a space frame will be built to provide realistic tests of cooling, mechanical rigidity, and thermal stability. In addition, they will be used to design assembly fixtures and test assembly techniques for mounting modules onto cones.

**Full-Scale Model of the IR.** A model of the B1 magnets and the beam pipe near the IP will be constructed. This will aid in identifying interference problems and verifying mounting schemes. It will also provide a test bed for the design of various installation fixtures.

---

## 4.7 Services, Utilities, and ES&H Issues

---

### 4.7.1 Services and Utilities

The vertex detector requires the following services, which must be brought inside the support tube to a location near the outboard end of the B1 magnet.

**Data and Control Lines.** Approximately 624 wire pairs are required on each side of the detector to service the 52 transition cards. These consist of 2 clock, 2 control and 2 data links per readout section providing redundancy for each signal. The interface cards will multiplex these into 26 optical links for data and 26 optical links for control on each side of the detector.

**Power.** The readout ICs will require three low-voltage power supplies (two analog, one digital), and the transition card will require at least one power supply. This amounts to four power and return cables per readout section, or 416 supplies and returns for each half of the detector. Each line carries only a few watts, except for the transition card supply which could be a few tens of watts. In addition to these low voltage supplies, we require 52 detector bias supplies per detector half. The bias voltage ranges from 40–80 V, and the detector draws very little current. The power supplies will all be specially procured to the vertex detector specifications in order to control electronic noise.

**Cooling Water.** The readout electronics and transition cards will be water cooled. Two sets of water connections will be required for each cone (since each cone is constructed from two halves), and one water connection for each transition card support. The cooling water will be supplied by a special low volume chiller system dedicated to the vertex detector.

**Dry Air or Nitrogen.** The vertex detector requires a dry, stable environment, and dry air or nitrogen from each side is planned.

### 4.7.2 ES&H Issues

There are very few ES&H issues which impact the construction, assembly, and operation of the SVT. The detector bias voltages can exceed 50 V and therefore qualify as high voltage, though they are extremely low current. There is a potential for water leaks from the cooling

systems. This water would be confined to the volume inside the support tube and would pose a danger primarily to the vertex detector itself. The laser monitoring system would employ Class 1 lasers which pose essentially no danger.

## References

---

- [Acc94] M. Acciarri *et al.*, CERN-PPE/94-122, to be published in *Nucl. Instr. Methods A*.
- [Bag94] G. Bagliesi *et al.*, “Test Beam Results from Prototypes of the Upgraded ALEPH Vertex Detector,” ALEPH Note 94-069 (1994).  
P. Coyle *et al.*, “ALEPH Vertex Detector,” presented at the Sixth Pisa Meeting on Advanced Detectors, Elba, Italy (1994).
- [Bat92] G. Batignani *et al.*, “Experience with the ALEPH Silicon Vertex Detector,” *Nucl. Instr. Methods A* **315**, 121–124 (1992).
- [Bat94a] G. Batignani *et al.*, “Double-Sided Silicon Strip Detectors in Pisa,” presented at the Sixth Pisa Meeting on Advanced Detectors, Elba, Italy (1994).
- [Bat94b] G. Batignani *et al.*, “Specifications of the Silicon Detectors for the Vertex Detector,” *BABAR Note # 196* (1994).
- [Bil84] P. Billoir, *Nucl. Instr. Methods* **225**, 352 (1984).
- [Bin93] N. Bingefors *et al.*, *Nucl. Instr. Methods A* **328**, 447 (1993).
- [Bow95] G. Bowden, “IP Beam Pipe/Vertex Detector Clearance” (1995) (unpublished).
- [Bre91] A. Breakstone *et al.*, *Nucl. Instr. Methods A* **305**, 39 (1991).
- [Bro95] D. Brown, “*BABAR* Vertex Detector Alignment Requirements,” *BABAR Note # 148* (1995).
- [Cac92] M. Caccia *et al.*, *Nucl. Instr. Methods A* **315**, 143 (1992).
- [Col94] T. Collins, “*BABAR* Hybrid Design Considerations,” *BABAR Note # 222* (1994).
- [DiB95] D. DiBitonto *et al.*, “Ultra-thin, High Precision Flex Cable for the L3 Silicon Microvertex Detector,” to be submitted to *Nucl. Instr. Methods A*.
- [For94] F. Forti, “TRACKERR Studies for Optimization of Vertex Detector Resolution,” *BABAR Note # 195* (1994).
- [Hol89] P. Holl *et al.*, *IEEE Trans. Nucl. Sci.* **36**, N. 1, 251 (1989).
- [Inn93] W. Innes, “TRACKERR, A Program for Calculating Tracking Errors,” *BABAR Note # 121* (1993).
- [Joh95a] Front-End Working Group, “Requirements Specifications for Silicon Vertex Detector Readout Chip,” edited by R. Johnson, *BABAR Note # 213* (1995).

- [Joh95b] Front-End Working Group, “Target Design Specifications for Silicon Vertex Detector Readout Chip,” edited by R. Johnson, *BABAR Note # 214* (1995).
- [Lev] Max Levy Autograph, Inc., Philadelphia, PA (USA).
- [Lev94] M. Levi, “The Impact of Backgrounds on the Silicon Architecture and Detector Trigger,” *BABAR Note # 136* (1994).
- [Roe94a] N. Roe and L. Kerth, “Vertex Detector Solid Angle Coverage,” *BABAR Note # 160* (1994).
- [Roe94b] N. Roe, “Silicon Strip Resolution for Time-Over-Threshold Readout,” *BABAR Note # 161* (1994).
- [Sny94] A. Snyder, “The Effect of Vertex Cuts on CP Reach,” *BABAR Note # 177* (1994).
- [Ton94] G. Tonelli *et al.*, “Development of Double-Sided  $\mu$ strip Detectors for High Radiation Environment,” presented at the *Sixth Pisa Meeting on Advanced Detectors*, Elba, Italy (1994).

---



---

## Drift Chamber

---

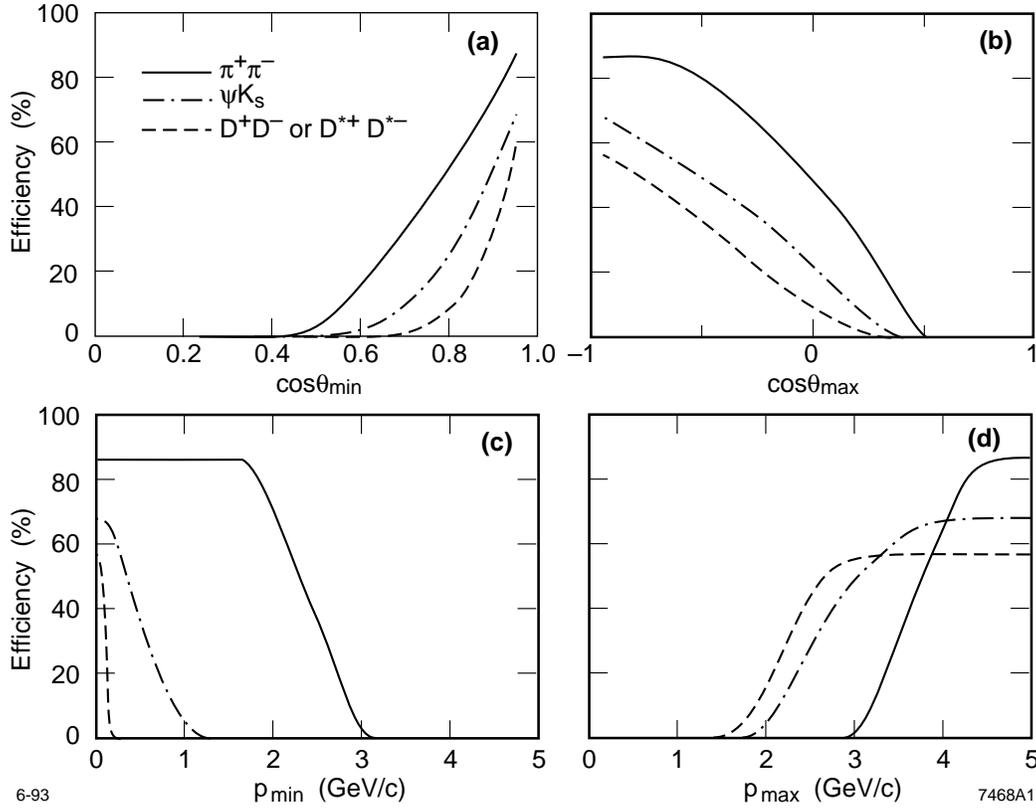
### 5.1 Physics Requirements and Performance Goals

---

The goals of the *BABAR* experiment require that exclusive final states from  $B^0$  decays be reconstructed efficiently and with high resolution. This places stringent demands on the performance of the main tracking chamber. It must provide maximal solid angle coverage, good momentum resolution at all momenta, and efficient reconstruction of tracks as low as 100 MeV/ $c$ . In addition, the tracking chamber provides one of the two major triggers for the experiment. These requirements are met through the use of a small-cell, low-mass drift chamber which, in addition to providing excellent momentum resolution for low-momentum tracks, minimally degrades the performance of the calorimeter and particle identification device.

Figure 5-1 shows the integrated efficiency for a variety of all-charged  $B$  decay modes as a function of the momentum and polar angle acceptance. The channels  $B^0 \rightarrow \pi^+\pi^-$ ,  $B^0 \rightarrow J/\psi K_S^0$  with  $K_S^0 \rightarrow \pi^+\pi^-$ , and  $B^0 \rightarrow D^+D^-$  with  $D^+ \rightarrow K^-\pi^+\pi^-$  have been used in the upper two plots in Figure 5-1 to demonstrate the importance of the forward angle acceptance limit. For example, a factor of 6.5 would be lost in the  $D^+D^-$  efficiency if good momentum and particle identification information were not available beyond  $\cos\theta = 0.8$ . Thus the drift chamber must cover the polar angle range down to the beam-line components in the forward direction, namely 300 mr or  $\cos\theta < 0.955$ . The backward angle requirement is considerably less stringent. The high- and low-momentum requirements are illustrated in the lower two plots of Figure 5-1; the channel  $B^0 \rightarrow D^{*+}D^{*-}$ , with its subsequent cascade decay  $D^{*+} \rightarrow \pi^+D^0$  producing a soft pion, places the greatest demand on the reconstruction of low momentum tracks. The ability to find and reconstruct charged particles down to 100 MeV/ $c$  is clearly required for good efficiency in this mode and in other channels containing one or more  $D^{*+}$  mesons.

In order to reconstruct exclusive final states with minimal background, the chamber must provide excellent momentum resolution. For tracks with  $p$  above 1 GeV/ $c$ , such as those arising from the decay  $B^0 \rightarrow \pi^+\pi^-$ , a resolution of  $\sigma_{p_t}/p_t \simeq 0.3\% \times p_t$  is anticipated. For the momentum range 0.1 to 1.0 GeV/ $c$ , relevant for tracks from the decay  $B^0 \rightarrow D^+D^-$ ,  $\sigma_{p_t}/p_t \simeq 0.3\text{--}0.4\%$  should also be achievable. The expected resolution on the reconstructed



**Figure 5-1.**  $B$  reconstruction efficiency for charged particle detection from  $B^0 \rightarrow \pi^+\pi^-$  (solid),  $B^0 \rightarrow J/\psi K_S^0$ ,  $J/\psi \rightarrow \ell^+\ell^-$ ,  $K_S^0 \rightarrow \pi^+\pi^-$  (dash-dot), and  $B^0 \rightarrow D^+D^-$ ,  $D^+ \rightarrow K^-\pi^+\pi^+$  for the angular distributions and  $B^0 \rightarrow D^{*+}D^{*-}$ ,  $D^{*+} \rightarrow \pi^+D^0$ ,  $D^0 \rightarrow K^-\pi^+$  for the momentum distributions (dashed). Each set of curves is shown as a function of: (a) minimum detectable  $\theta$ ; (b) maximum detectable  $\theta$ ; (c) minimum detectable momentum; and (d) maximum detectable momentum.

$B^0$  mass for these decays and the dependence on various drift chamber detector parameters is given in Section 5.3.

At PEP-II, the average charged particle momentum is less than 1 GeV/c. For most particles, the major limitation on track parameter resolution in the drift chamber will be multiple scattering. In order to minimize this effect, the chamber will be made of low-mass materials. A thin inner wall facilitates matching drift chamber and vertex detector tracks, improves the contribution of the high-precision measurement in the outer layer of the SVT to the  $p_t$  resolution, and minimizes backgrounds due to photon conversions in the chamber wall. The material in the outer walls and endplates of the chamber must also be kept to a minimum

in order to reduce the impact of the drift chamber on calorimeter and particle identification performance.

The drift chamber is expected to participate in the detector trigger by providing one of two complementary trigger streams. The drift chamber must furnish a charged track trigger with a latency of less than  $9.5 \mu\text{s}$  and a maximum jitter of about  $0.5 \mu\text{s}$ . The latency constraint results from the buffering in the silicon vertex detector, which therefore requires a strobe at less than 10 kHz. The drift chamber will trigger on one and one-half tracks, *i.e.*, one track traversing the full chamber radius together with one track traversing at least half the chamber radius.

The main tracking chamber serves several other functions as well. For low momentum particles, the chamber will provide particle identification through ionization loss ( $dE/dx$ ) measurements. For the helium-based gas mixtures under consideration, a resolution of just under 7% should be attainable for the  $dE/dx$  measurement, allowing  $\pi/K$  separation up to 700 MeV/ $c$ . This capability is complementary to the DIRC in the barrel region, but is essential in the extreme backward direction where no dedicated particle identification device is foreseen.

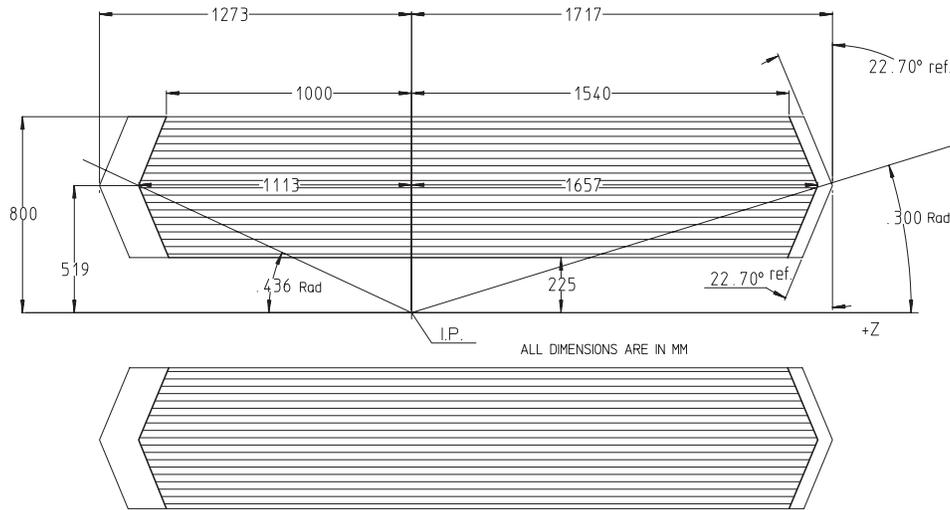
The drift chamber also allows the reconstruction of secondary vertices, such as decays of  $K_S^0$ s outside the silicon detector volume. For this purpose, the chamber should be able to measure not only the transverse coordinate, but also the longitudinal position of tracks with good ( $\sim 1$  mm) resolution. Good  $z$  resolution also aids in matching drift chamber and silicon tracks, and in projecting tracks to the DIRC and calorimeter.

Finally, the chamber must be operational with a margin of safety at the expected backgrounds, which are predicted to be  $\sim 5$  kHz/cell for the innermost layers. For a five-year period of operation, the integrated charge is expected to be 0.025 C/cm for sense wires in the innermost drift chamber layers, where variations in  $z$  and  $\phi$  have been included. Details are described in Chapter 12.

## 5.2 Tracking Chamber Overview

---

In order to meet the requirements outlined above, a small-cell drift chamber has been chosen for the *BABAR* detector. Low- $Z$  materials, including a helium-based gas and aluminum wires, will be used in the design. The chamber, illustrated in Figure 5-2, is 280 cm in length and occupies the radial space between 22.5 cm and 80 cm. It is bounded by the support tube at its inner radius and the particle identification device at its outer radius. Forty layers of wires will fill this volume, providing up to 40 spatial and ionization loss measurements for charged particles with transverse momentum greater than 180 MeV/ $c$ . Two alternatives have been considered for the wire and layer arrangement: a baseline axial/stereo chamber containing

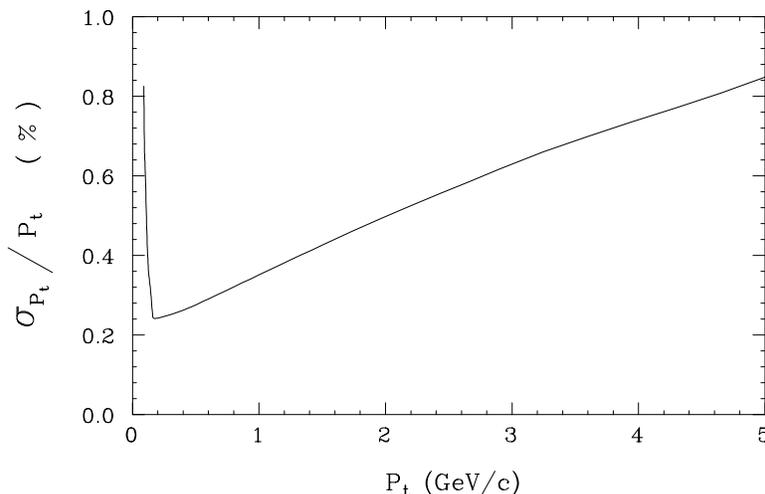


**Figure 5-2.** Tracking chamber geometry showing the conical endplates. The chamber is offset in  $z$  from the interaction point and extends 1.66 m in the forward direction and 1.11 m in the backward direction. A ten-superlayer structure is indicated. The drift chamber envelope includes a 16 cm space at the backward endcap for readout cards, cables, and an rf-shield. The forward endcap includes a 6 cm space for feedthroughs and an rf-shield.

four superlayers of axial wires and six superlayers of stereo wires, and an alternative all-stereo chamber containing only layers with positive and negative stereo-angle wires. The wire arrangements are discussed in detail in Section 5.4. For an outer radius of 80 cm, there are about 7000 sense wires of  $20\ \mu\text{m}$ -diameter gold-plated tungsten and about 45,000 field wires of  $55\ \mu\text{m}$ -diameter gold-plated aluminum. The exact ratio of the number of field to sense wires is still under study.

A chamber length of 166 cm in the forward direction has been chosen. This ensures sufficient coverage for forward-going tracks to avoid compromising the invariant mass resolution, while at the same time does not threaten the electrical stability of the chamber or lead to a prohibitively expensive calorimeter. In the backward direction, a length of 111 cm means that particles with polar angles down to  $435\ \text{mr}$  penetrate as far as the mid-plane of the chamber, which yields good tracking and  $dE/dx$  information.

The endplates, which carry an axial load of 2400 kg, are designed as truncated cones, with the outer part inclined at  $22.7^\circ$  to the vertical. The conic shape provides a substantial increase in mechanical strength and allows particles to enter the forward particle identification device at close to normal incidence. In the forward direction, the tip of the endplate cone lies on the  $300\ \text{mr}$  line. The inner sections of the endplates will be conically shaped as well. This decreases the wire lengths, permitting larger stereo angles. A detailed comparison of endplate shapes is given in a later section.



**Figure 5-3.** Momentum resolution expected for pions at  $\theta = 90^\circ$  for the five-layer silicon vertex detector and the 40-layer drift chamber in a 1.5 T magnetic field. For the drift chamber, an averaged radiation length of 340 m and a single-cell spatial resolution of  $140 \mu\text{m}$  have been assumed. Both the axial/stereo and all-stereo chamber designs give the resolution shown in the figure.

The inner cylindrical wall of the drift chamber, which sits 15 mm outside the  $0.005X_0$  support tube, acts as a gas seal and is not foreseen to provide structural support. The goal is therefore to minimize the inner wall material and the resultant contribution to multiple scattering. To that end, the inner wall will be made out of carbon fiber.

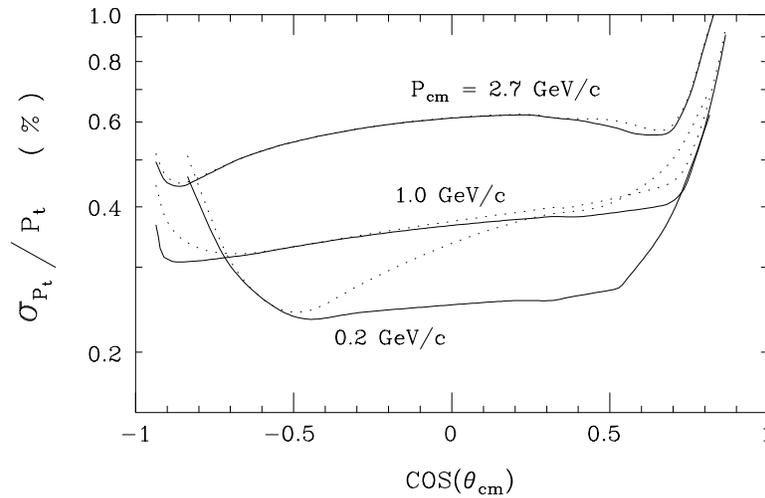
The outer cylindrical wall bears the wire load from the endplates. It is desirable to make this wall as thin as possible, so that the performance of the calorimeter is not degraded. Carbon fiber will therefore be used for the outer wall as well as the endplates.

The chamber will be filled with a helium-based gas at or near atmospheric pressure. In order to minimize the amount of material before the forward endcap, all readout electronics and HV distribution will be placed on the backward endplate.

## 5.3 Projected Performance

---

The expected transverse momentum resolution for the five-layer silicon vertex detector and the 40-layer drift chamber in a 1.5 T magnetic field is shown in Figures 5-3 and 5-4. The

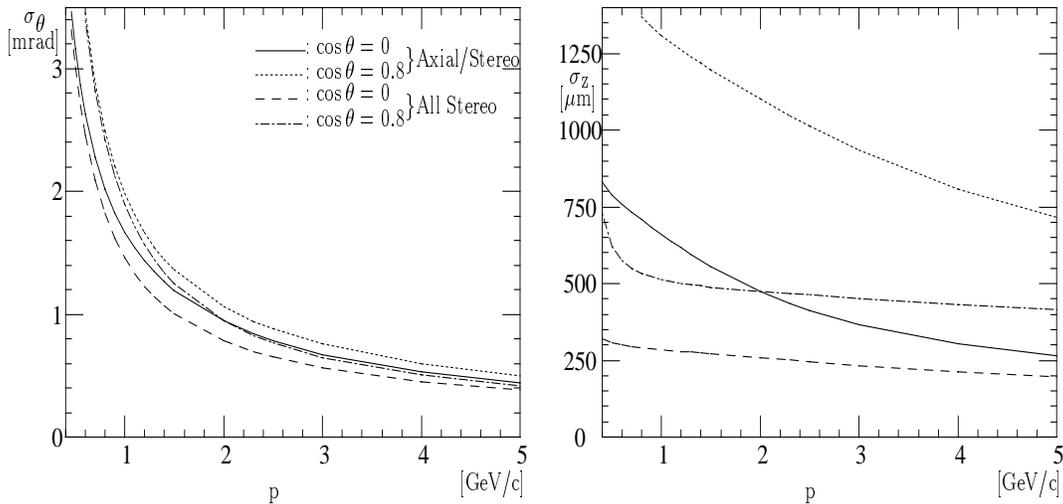


**Figure 5-4.** *Transverse momentum resolution vs. center-of-mass angle for pions produced with three different center-of-mass momenta. The dotted curves show the total momentum resolution  $\sigma_p/p$  assuming a  $100\ \mu\text{m}$  vertex constraint.*

large, low-mass tracking volume provides excellent resolution<sup>1</sup>. For simplicity, a radiation length of 340 m has been used in the calculations: the helium-based gas represents an 800 m radiation length, while the wires, when averaged over space, present a slightly larger amount of material compared to the gas. The inclusion of the effects of discrete wires would result in better chamber performance than that shown. However, systematic effects which would be encountered when reconstructing tracks from actual data have not been included. Such calculations therefore represent the optimal performance of the tracking system, and have been used in a comparative way when exploring various design choices.

Figure 5-4 shows that nearly symmetric forward-backward coverage has been achieved in the center-of-mass frame for momenta above 200 MeV/c. For a He:C<sub>4</sub>H<sub>10</sub> 80:20 gas mixture, with the material from the wires and gas averaged together and with a single-cell spatial resolution of 140  $\mu\text{m}$ , the expected resolution varies from 0.2% for low momentum tracks to 0.5% for high momentum tracks in the central region. The resolution degrades considerably for low-angle tracks, which exit the chamber through the endplates. At low momenta, the total momentum resolution  $\sigma_p/p$  is dominated by the polar angle uncertainty caused by multiple scattering in the silicon. This uncertainty can be significantly reduced by using the primary vertex to impose a constraint on the track. Conservatively assuming the vertex resolution to be 100  $\mu\text{m}$  yields the dotted  $\sigma_p/p$  curves shown in Figure 5-4. At momenta above 1 GeV/c, however,

<sup>1</sup>The calculations were performed using the TRACKERR [Inn93] package, which employs the Kalman-filter technique to analytically propagate the error matrix. The simulation assumes an average 140  $\mu\text{m}$  single cell resolution and multiple scattering contributions from the drift chamber gas and wires.



**Figure 5-5.** Expected angular resolution (left) and expected  $z$  resolution, both at the DIRC, for two different polar angles and for the two chamber configurations.

the error in the transverse component dominates the total momentum resolution. The  $z$  position and polar angle resolutions at the DIRC are shown in Figure 5-5 for various momenta and polar angles. These estimates were obtained using a modified version TRACKERR to simulate the effect of propagating errors outward in the drift chamber to the DIRC radius.

In optimizing the performance of the tracking system, a number of global design parameters have been examined. Neglecting material, the transverse momentum resolution is directly proportional to the single-cell spatial resolution and inversely proportional to the square root of the number of measurements, to the magnetic-field strength, and to the square of the measured arc length in the transverse plane. However, for most momenta encountered in this experiment, multiple scattering in the material within the tracking volume is also an important factor, dominating the resolution below  $0.5 \text{ GeV}/c$ . The range of mass resolutions which could be provided by the tracking system is illustrated in Table 5-1. It would be tempting to enlarge the radius of the drift chamber beyond 80 cm. However, this is a very expensive direction for design change, since it also drives the volume of CsI required for the calorimeter. Removing the PEP-II support tube is an equally expensive way to improve tracking resolution, since it leads to a much higher risk of reduced luminosity from the machine. This has left magnetic field strength as the most cost-effective way to obtain suitable  $B$  mass resolution for  $CP$  studies.

Outer Radius (cm)	Support-Tube	Gas-Wires	Resolution ( $\mu\text{m}$ )	$B$ Field (Tesla)	Mass Resolution ( $\text{MeV}/c^2$ )	
					$B^0 \rightarrow \pi^+\pi^-$	$B^0 \rightarrow D^+D^-$
80	Yes	He-Al	140	<u>1.0</u>	$30.7 \pm 0.3$	$5.69 \pm 0.09$
<b>80</b>	<b>Yes</b>	<b>He-Al</b>	<b>140</b>	<b>1.5</b>	<b><math>20.8 \pm 0.2</math></b>	<b><math>4.70 \pm 0.08</math></b>
80	<u>No</u>	He-Al	140	1.5	$17.1 \pm 0.1$	$4.37 \pm 0.08$
<u>95</u>	Yes	He-Al	140	1.5	$14.0 \pm 0.1$	$4.07 \pm 0.07$
<u>95</u>	<u>No</u>	He-Al	140	1.5	$12.5 \pm 0.1$	$3.73 \pm 0.06$
80	Yes	<u>Ar-Cu</u>	140	1.5	$25.2 \pm 0.2$	$5.76 \pm 0.09$
80	Yes	He-Al	<u>210</u>	1.5	$26.0 \pm 0.2$	$5.09 \pm 0.09$

**Table 5-1.** The effect of chamber parameters on mass resolution for  $B \rightarrow \pi^+\pi^-$  and  $B \rightarrow D^+D^-$ . The entries in bold represent the parameters of the baseline design. The underlined values denote a variation with respect to the baseline. The cases with no support tube are unrealistic to the extent that the effects of a finite thickness for the drift chamber inner wall are not considered.

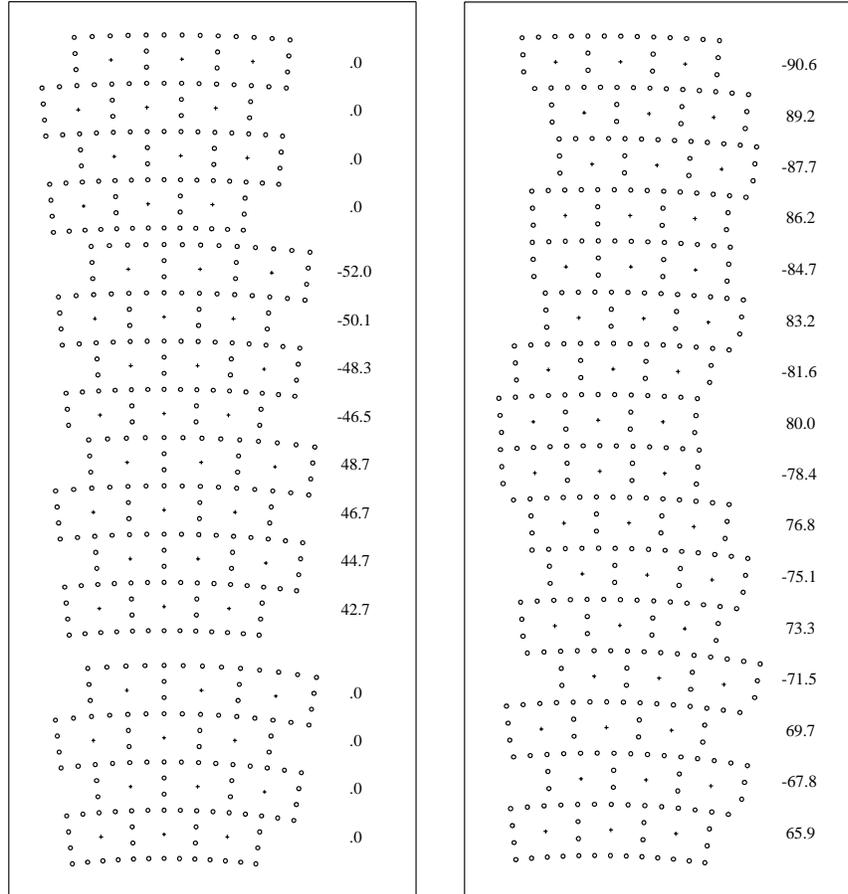
## 5.4 Drift System Design

### 5.4.1 Cell Design

The drift chamber uses small cells arranged in 40 concentric layers about the axis of the chamber. We have chosen a rectangular cell shape, approximately 13 mm by 19 mm along the radial and azimuthal directions, respectively. In the baseline arrangement, drift cells are arranged in superlayers with either an axial or stereo orientation.

Each cell has one sense wire surrounded by a rectangular grid of field wires, as shown in Figure 5-6. The sense wires are made of gold-coated tungsten,  $20\ \mu\text{m}$  in diameter, tensed with a weight of 50 g. The deflection due to gravity is  $120\ \mu\text{m}$  at the mid-length of the longest wires. Wire lengths vary from 2.6 m to 2.8 m, due to the conic shape of the end plates. Approximately +1.8 kV is applied to the sense wires with the field wires at ground potential, giving an avalanche gain of approximately  $5 \times 10^4$ .

The field wires are gold-coated aluminum with a diameter of  $55\ \mu\text{m}$ . This diameter is sufficiently large to keep the electric field on the wire surface below 20 kV/cm, in order to avoid whisker growth on the wires. A tension of 53 g is applied to the field wires to match the gravitational sag of the sense wires. This tension is roughly one-half the yield tensile strength of the wire.



**Figure 5-6.** Layers of rectangular cells grouped into axial-stereo (left plot), and all stereo (right plot) configurations. The axial-stereo case has four layers of cells per superlayer and an extra layer of field wires at each axial-stereo boundary to enclose the cells. In the all-stereo case, the layers alternate between positive and negative stereo angles. The stereo angles (mr) of the layers are also shown for each case.

### 5.4.2 Layer Arrangement

Two different configurations of wire layers are being considered (Figure 5-6). The baseline configuration has both axial (A) and stereo (U and V) superlayers, while an optional layout having only stereo layers that alternate with positive and negative angles from layer to layer has also been studied.

### Axial-Stereo Layer Arrangement

The baseline version of the chamber layout contains ten superlayers with four layers per superlayer. The layers within each superlayer have the same cell count, with cells staggered by half a cell from one layer to the next. This provides a constant cell pattern at all azimuthal angles, which allows local segment finding and left-right ambiguity resolution within a superlayer. The stereo angles of the superlayers alternate between axial (A) and stereo (UV) pairs, in the order AUVAUVAUVA, as shown in Figure 5-6. The difference between the minimum and maximum cylindrical radius of a stereo-layer is known as its stereo sagitta. For each UV pair of superlayers, a stereo sagitta of 5.5 mm is chosen at the innermost layer. This sets the stereo angle for the inner layer, while stereo angles in the other layers of the UV pair are determined by requiring that the wires have the same twist angle (the change in azimuthal angle over the length of the wire) to preserve the cell shapes along the wires. The tangent of the stereo angle varies as the square root of the layer radius within the UV superlayers.

The charged track trajectories in the magnetic field appear circular in the axial views. Three of the five track parameters (the curvature  $\kappa$ , azimuthal angle  $\phi$ , and distance to origin  $r_0$ ) can be determined from a circle fit in the axial view, while the  $z$  position and polar angle  $\theta$  are determined using the stereo layers. The hardware trigger performs a fast transverse momentum determination using only the axial layers for tracks traversing a significant portion of the chamber. The cell widths for layers in the innermost superlayer vary by  $\pm 7\%$ , requiring separate time-distance relationships for each layer. This variation decreases at larger radii. Because of the hyperboloidal geometry of stereo layers, the radial space between layers at an axial-stereo boundary varies along the wire, so an extra enclosing layer of field wires is added at the boundaries to provide cell shape uniformity. These additional field layers introduce 5 mm of dead space at six axial-stereo boundaries. A bias voltage of 230 V is applied to the corner field wires in the enclosing layers to sweep out the ions in the dead regions. Table 5-2 gives the radii and stereo angles of all the layers.

The 5 mm enclosing space and the 5.5 mm stereo sagitta result in 10.5 mm radial regions of free space at three radial locations on the end plates, which are otherwise fully populated with feedthroughs. This space may be used to our advantage for gas ports and mechanical stand-offs for cable supports.

### All-Stereo Layer Arrangement

The all-stereo optional layout also has 40 layers, but with alternating positive and negative stereo angles. The cell width is nearly constant in all layers, with the cell count increasing by five from one layer to the next layer outward in radius. This constant cell size throughout the whole chamber allows (to first order) a single time-distance relationship to be used for

Layer	Number of cells	Radius at $z = 0$ (mm)	Radius at $z = 1380$ (mm)	Twist (mrad)	Stereo (mrad)	Cell Width (mm)
1	90	241.8	241.8	.0	.0	16.9
2	90	254.4	254.4	.0	.0	17.8
3	90	266.9	266.9	.0	.0	18.6
4	90	279.5	279.5	.0	.0	19.5
5	108	297.0	302.6	386.0	42.0	17.3
6	108	309.3	315.2	386.0	43.8	18.0
7	108	321.7	327.7	386.0	45.5	18.7
8	108	334.0	340.3	386.0	47.3	19.4
9	126	347.3	352.9	-356.8	-45.3	17.3
10	126	359.7	365.5	-356.8	-47.0	17.9
11	126	372.0	378.0	-356.8	-48.6	18.6
12	126	384.4	390.6	-356.8	-50.2	19.2
13	144	408.2	408.2	.0	.0	17.8
14	144	420.8	420.8	.0	.0	18.4
15	144	433.3	433.3	.0	.0	18.9
16	144	445.9	445.9	.0	.0	19.5
17	162	463.4	469.0	308.7	52.2	18.0
18	162	475.8	481.6	308.7	53.6	18.5
19	162	488.3	494.1	308.7	55.0	18.9
20	162	500.7	506.7	308.7	56.4	19.4
21	180	513.7	519.3	-293.1	-54.9	17.9
22	180	526.2	531.9	-293.1	-56.2	18.4
23	180	538.6	544.4	-293.1	-57.6	18.8
24	180	551.0	557.0	-293.1	-58.9	19.2
25	198	574.6	574.6	.0	.0	18.2
26	198	587.2	587.2	.0	.0	18.6
27	198	599.7	599.7	.0	.0	19.0
28	198	612.3	612.3	.0	.0	19.4
29	216	629.8	635.4	264.7	60.7	18.3
30	216	642.3	648.0	264.7	61.9	18.7
31	216	654.8	660.5	264.7	63.1	19.0
32	216	667.2	673.1	264.7	64.3	19.4
33	234	680.1	685.7	-254.7	-63.0	18.3
34	234	692.6	698.3	-254.7	-64.2	18.6
35	234	705.1	710.8	-254.7	-65.3	18.9
36	234	717.6	723.4	-254.7	-66.5	19.3
37	252	741.0	741.0	.0	.0	18.5
38	252	753.6	753.6	.0	.0	18.8
39	252	766.1	766.1	.0	.0	19.1
40	252	778.7	778.7	.0	.0	19.4

**Table 5-2.** Radii and stereo angles of all layers in the axial-stereo chamber.

all the cells. The allowed stereo sagitta of 15 mm determines the stereo angles for each layer, for fixed length wires. The stereo angles range from 66 to 115 mr, from the inner to the outer layer. The layer arrangement is shown in Figure 5-6. Table 5-3 gives the radii and stereo angles of all layers. Alternative arrangements with as few as 36 layers could be considered in this layout. Likewise, consecutive layers could be arranged as superlayers with the same cell count and stereo angle, by sacrificing a small amount of cell uniformity.

Each stereo view has 20 equally spaced layers in the baseline design. In this geometry, tracks are found in each of the views using local-continuity algorithms. For tracks with large curvature, or where overlaps occur, this geometry may be advantageous. Given the projected resolutions in azimuthal angle  $\phi$  and curvature  $\kappa$ , the probability of confusion in matching the two views is small due to the relatively low density of tracks in an  $\Upsilon(4S)$  event. Cell widths in this design vary over the length of the chamber by a maximum of  $\pm 3\%$  for the innermost layer, decreasing towards larger radii. The cell height in the radial direction is constant.

### 5.4.3 Total Channel Count

The total channel count in either of the internal layout schemes described above depends on the choice for the total number of layers in the chamber, the channel increment from one layer to the next (or from one superlayer to the next), and the number of channels assigned to the first layer of the chamber. The number of options for these three design parameters is quite limited, as shown in Table 5-4.

The choice of 40 layers is mandatory for an axial-stereo layout. This allows four axial superlayers to complete the primary function of providing a  $p_t$  measurement, along with three pairs of UV stereo superlayers (with just 12 layers per view). Four layers per superlayer allow three-out-of-four majority logic for triggering and offline segment finding. In principle, the all-stereo design would allow a reduction to as few as 36 layers; this would then be comparable to the ARGUS drift chamber, but with only two rather than three views.

The cell aspect ratio, *i.e.*, ratio of width to height, is defined by the choice of the increment in the cell count,  $\Delta_i$ , from one layer to the next in proceeding radially outward. In the case of a superlayer design, the cell count is kept fixed for four and is then increased by  $4\Delta$ . The simplest designs keep  $\Delta_i = \Delta$  fixed over the entire chamber, leading to a uniform cell size. Choosing  $\Delta = 6$ , as has been used for both ARGUS and CLEO II, produces a nearly square cell in cross section. This is well-suited to the low average momentum in  $\Upsilon(4S)$  decays, which results in large angles of incidence on the drift cells with respect to the radial direction. The nearly circular isochrones which dominate a large fraction of the drift-cell response for a square cross section are optimal for handling such arbitrary entrance angles.

Layer	Number of cells	Radius at $z = 0$ (mm)	Radius at $z = 1380$ (mm)	Twist (mrad)	Stereo (mrad)	Cell Width (mm)
1	80	242.2	258.7	718.2	65.8	19.0
2	85	255.5	272.0	-700.2	-67.5	18.9
3	90	268.8	285.3	683.5	69.2	18.8
4	95	282.1	298.6	-667.9	-70.8	18.7
5	100	295.5	312.0	653.4	72.4	18.6
6	105	308.8	325.3	-639.7	-74.0	18.5
7	110	322.1	338.6	626.9	75.5	18.4
8	115	335.4	351.9	-614.8	-77.0	18.3
9	120	348.8	365.3	603.4	78.5	18.3
10	125	362.1	378.6	-592.6	-79.9	18.2
11	130	375.4	391.9	582.4	81.4	18.1
12	135	388.7	405.2	-572.7	-82.8	18.1
13	140	402.1	418.6	563.4	84.1	18.0
14	145	415.4	431.9	-554.6	-85.5	18.0
15	150	428.7	445.2	546.2	86.8	18.0
16	155	442.0	458.5	-538.2	-88.1	17.9
17	160	455.4	471.9	530.5	89.4	17.9
18	165	468.7	485.2	-523.1	-90.7	17.8
19	170	482.0	498.5	516.0	91.9	17.8
20	175	495.3	511.8	-509.2	-93.1	17.8
21	180	508.7	525.2	502.7	94.4	17.8
22	185	522.0	538.5	-496.4	-95.6	17.7
23	190	535.3	551.8	490.3	96.7	17.7
24	195	548.6	565.1	-484.5	-97.9	17.7
25	200	562.0	578.5	478.8	99.1	17.7
26	205	575.3	591.8	-473.4	-100.2	17.6
27	210	588.6	605.1	468.1	101.4	17.6
28	215	601.9	618.4	-463.0	-102.5	17.6
29	220	615.3	631.8	458.1	103.6	17.6
30	225	628.6	645.1	-453.3	-104.7	17.6
31	230	641.9	658.4	448.7	105.7	17.5
32	235	655.2	671.7	-444.2	-106.8	17.5
33	240	668.6	685.1	439.8	107.9	17.5
34	245	681.9	698.4	-435.6	-108.9	17.5
35	250	695.2	711.7	431.5	110.0	17.5
36	255	708.5	725.0	-427.5	-111.0	17.5
37	260	721.9	738.4	423.6	112.0	17.4
38	265	735.2	751.7	-419.8	-113.0	17.4
39	270	748.5	765.0	416.1	114.0	17.4
40	275	761.8	778.3	-412.5	-115.0	17.4

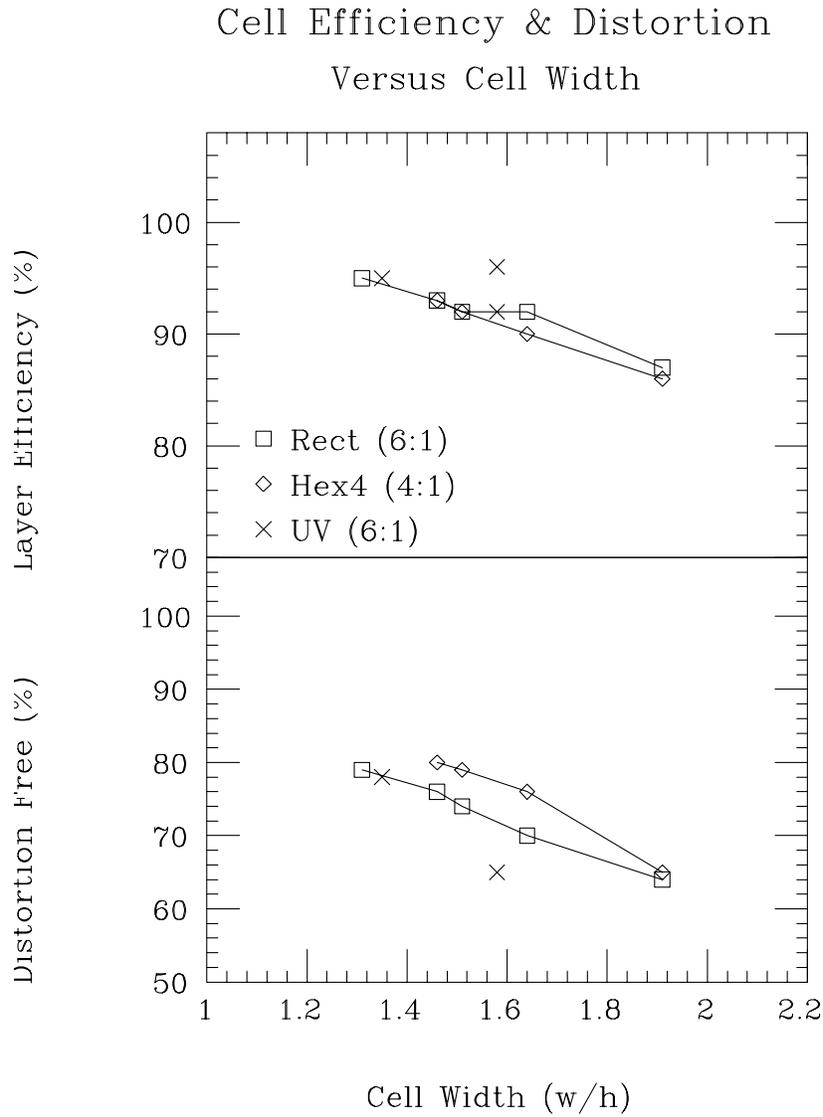
**Table 5-3.** Radii and stereo angles of all layers in the all-stereo chamber.

$N_{\text{layers}}$	$N_1$	$\Delta$	$w/h$ Ratio	w (cm)	Total Channels	Comments
40	90	(18)/4	1.343–1.550	1.69–1.95	6840	Axial-Stereo Design
	80	5	1.308–1.429	1.74–1.90	7100	All-Stereo Design
40	72	4	1.571	2.15	6000 (5760)	Ideal
	90	5	1.257	1.72	7500 (7200)	
	108	6	1.047	1.43	9000 (8640)	
36	65	4	1.571	2.39	4860	All-stereo Option All-stereo Option
	82	5	1.257	1.91	6102	
	98	6	1.047	1.59	7308	
36	60	6	1.047	1.88	5940	ARGUS
	51	96	1.047	1.50	12240	CLEO II

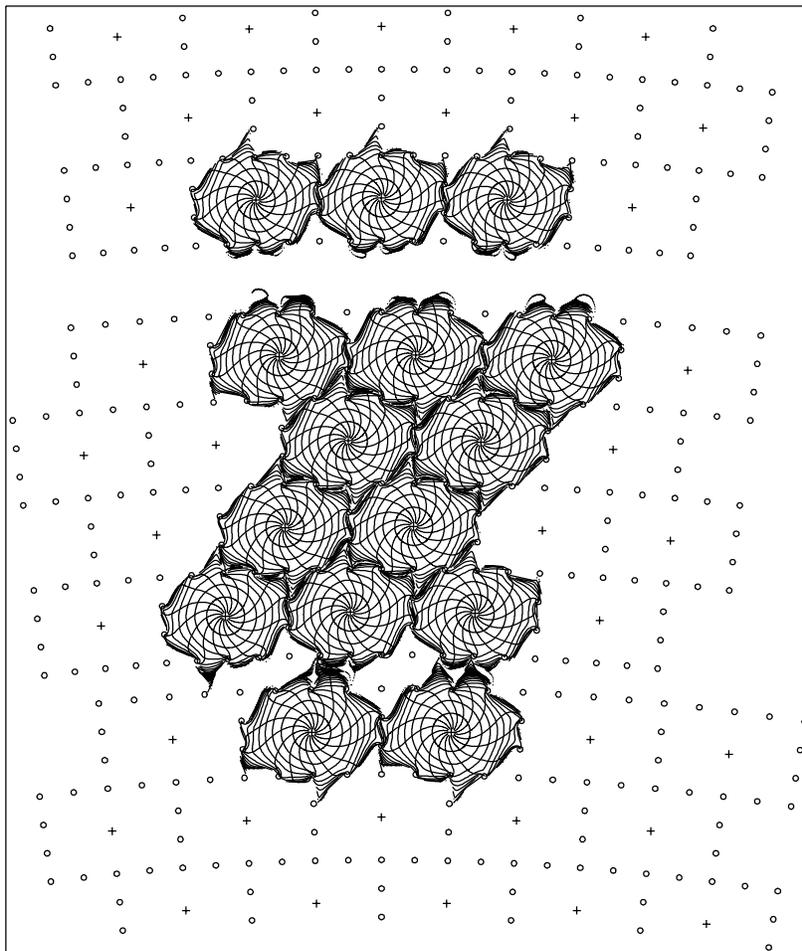
**Table 5-4.** Total channel counts for the baseline drift chamber designs, compared with various alternative solutions with different numbers of layers ( $N_{\text{layers}}$ ), cell aspect ratios ( $w/h$ ) as derived from possible layer-to-layer cell increments ( $\Delta$ ), or first-layer channel counts ( $N_1$ ). The total channel counts in parentheses are for an axial-stereo design, with a four-layer constant-channel superlayer structure.

The present designs for the  $BABAR$  chamber compromise on the choice of layer increment in two ways: (1) by choosing to vary  $\Delta$  slightly from one layer to the next, and (2) by deviating significantly from the ideal square-cell design. Drift cell simulations have shown that increasing the  $w/h$  ratio beyond  $\sim 1.65$  begins to result in performance degradation. The drift cells in either of our design layouts are distinctly rectangular, with the width being 30–55% larger than the height. These are already close to the maximum feasible for good performance in the  $\Upsilon(4S)$  environment, as shown in Figure 5-7, leaving little room for additional reduction in channel count.

The last parameter to be considered is the number of cells in the first layer of the chamber. Here, the robustness of the inner layers of the chamber with respect to occupancy is of primary concern. The design choice for  $BABAR$  lies midway between ARGUS and CLEO II, despite the fact that circulating beam currents in PEP-II will be about 50 times larger than at DORIS II. The ARGUS value was chosen to accommodate the charged multiplicity in  $\Upsilon(4S)$  decays. At  $BABAR$ , considerable effort has been made to understand and control beam-related backgrounds. However, the potential for higher rates has led to a conservative choice of  $N_1 = 80$  or 90 in our design.



**Figure 5-7.** Layer efficiencies (upper) and distortion free regions (lower) as a function of the cell width. Layer efficiencies account for cases where 2 adjacent cells in a layer contain information. Distortion free is defined to be cases with less than  $100\ \mu\text{m}$  corrections for angle of incidence.

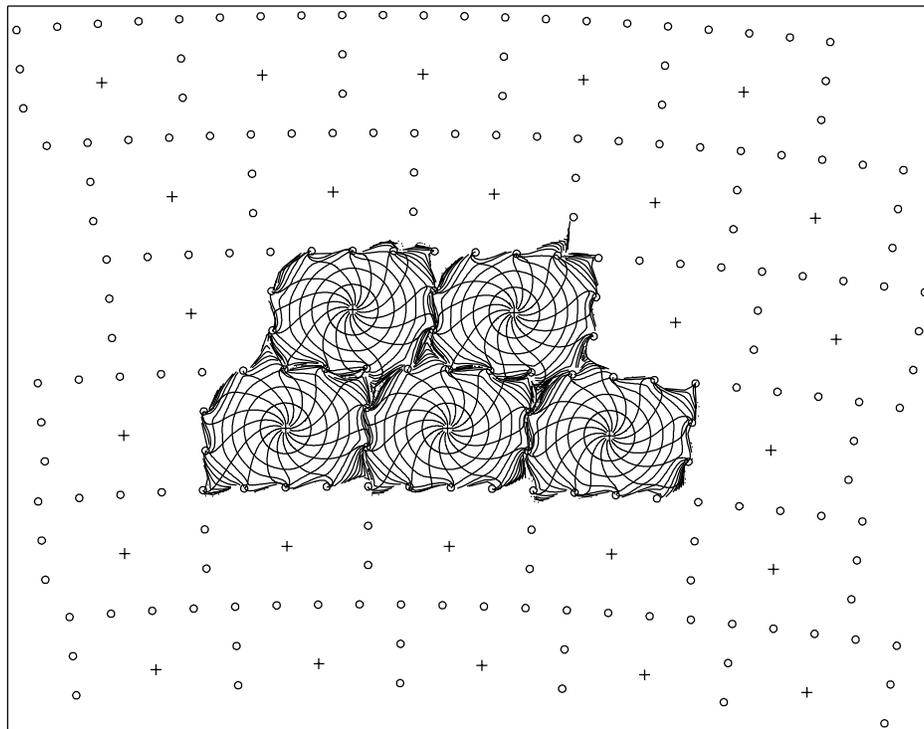


**Figure 5-8.** Isochrones for rectangular cells in superlayers. An endplate view is shown of the stereo layers on either side of a four-layer axial superlayer, with enclosing field wires. A 230 V potential on the cell's corner wires in the enclosing layers clears out the ions between these layers. The plots show equal drift time contours at 100 ns intervals.

#### 5.4.4 Cell Studies

A simulation package (DCSIM), which models drift chamber responses to charge deposition for given wire configurations, has been used to study various cell properties, including drift time isochrones, time-distance relationships, distortions, and gain variations.

Isochrones for the rectangular cells in axial-stereo superlayers are shown in Figure 5-8. A similar plot for the all-stereo design is shown in Figure 5-9. The cells have circular contours near the sense wire but become distorted near the field wires. In the superlayer case, a

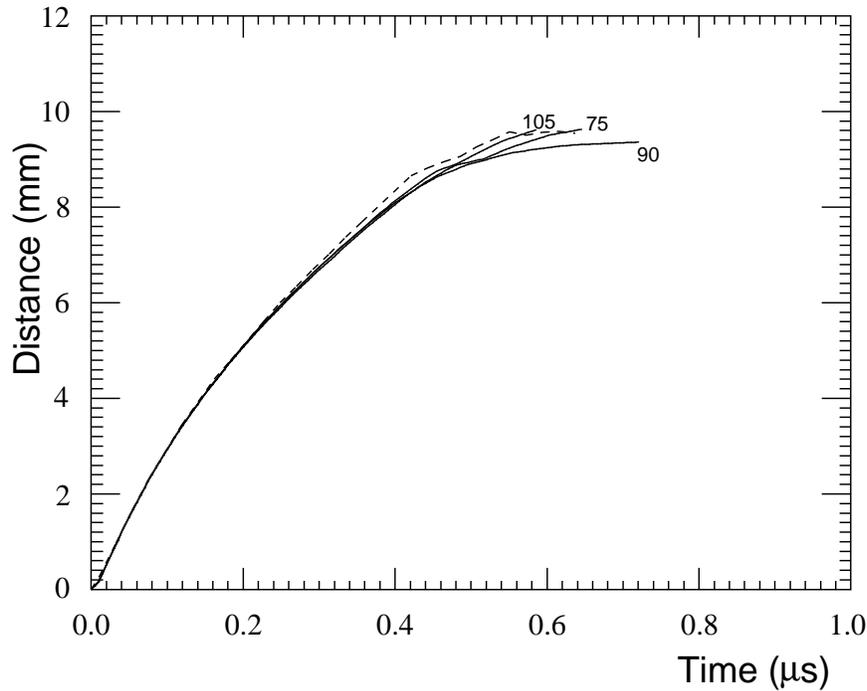


**Figure 5-9.** *Isochrones for all-stereo cells. The plots show equal drift time contours at 100 ns intervals.*

potential of +230 V at the cell corners in the enclosing layers clears out the electrons in the dead space between these enclosing layers.

Although isochrones show drift times of electrons anywhere in the cell, they do not easily describe the collection of electrons originating along tracks. The time-distance relationship of tracks in small cells has been simulated and compared with measurements made with the Prototype I chamber. The simulation program generates tracks and computes the average drift time of electrons over a short track segment centered about the minimum arrival point. The length of a short segment used in the simulation was made equal to the mean free path between primary electrons. Figure 5-10 shows the agreement between the simulated time-distance curves and measurements in Prototype I with an 80:20 mixture of He:C<sub>4</sub>H<sub>10</sub> gas.

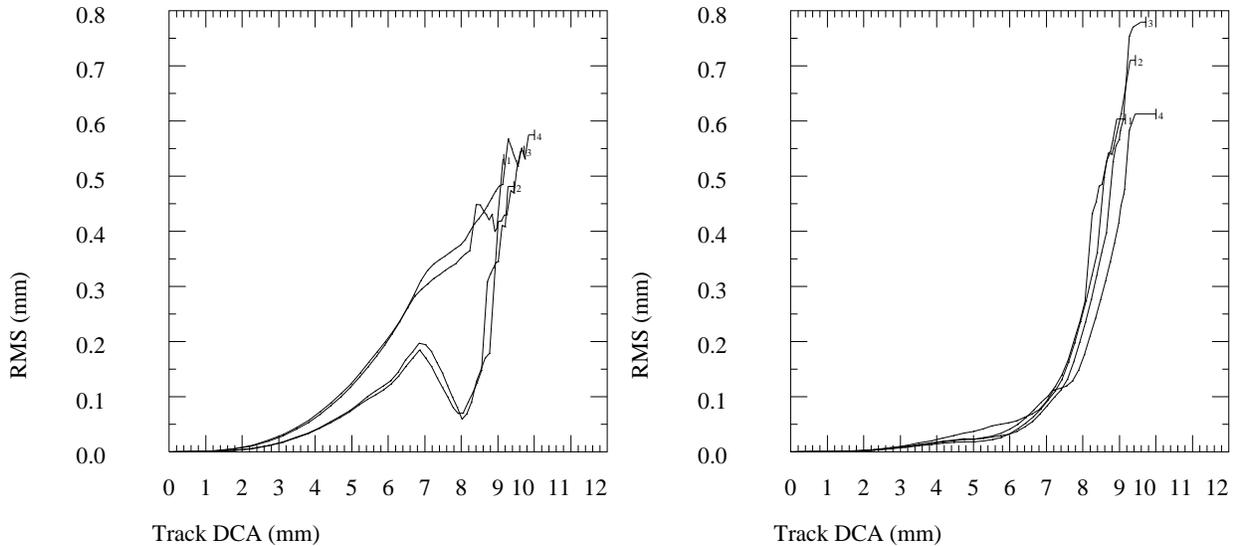
These simulated time-distance relationships have then been used to study other cell properties. The time-distance dependence on the track entrance angle has been calculated, along with the distortions which would arise if a single time-distance function were used over a range of track angles. For each drift time in a cell the track distance from the wire at various



**Figure 5-10.** *Simulated and measured time-distance functions. The solid curves are the simulated time-distance curves for track angles of 75, 90, and 105°, where 90° is along the radial direction. Measurements in a prototype chamber with no magnetic field are shown by the dashed curve.*

track angles is calculated, and the average and rms spreads in distance over a range of track angles are determined. The rms spread is a measure of the distortion that a single time-distance function would have due to the effects of track angles in the cell. Figure 5-11 shows the rms spread in distance over track angles within  $\pm 30^\circ$  of the radial direction for the four layers within a superlayer. The left plot has all field wires at ground potential, while the right plot has compensated voltages of +230 V at cell corners on enclosing layers and -120 V on the pairs of wires between the sense wires. The distortions are smaller in the latter case. One can compare distortions for various cell geometries by recording the distance of closest approach (DCA) at which the rms variation is  $100 \mu\text{m}$ . For the rectangular cell shown, this is at 7.4 mm for vertical tracks, which is 78% of the 9.5 mm geometric half-width of the cell. Table 5-5 shows the distance of closest approach values at the distortion limit of  $100 \mu\text{m}$  for various cell and layer geometries.

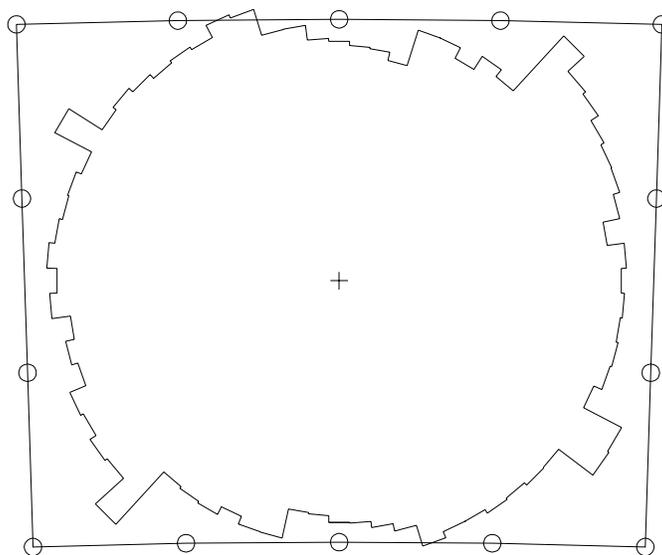
The spatial resolution has been determined for those parts of the cell in which a sufficient number of electrons is collected within the leading edge of the signal, defined to be the case in which two electrons arrive within 10 ns of the theoretical minimal drift time. The DCA



**Figure 5-11.** Distortions in a rectangular cell vs. the track's distance of closest approach. The distortions are calculated over a range of track angles with a single time-distance relationship. The four curves show distortions in the four layers in a superlayer in the axial-stereo design, for tracks which are within  $\pm 30^\circ$  of the radial direction (*i.e.*,  $p_t > 350 \text{ MeV}/c$  in an 80 cm drift chamber). The left plots are for the case with all field wires at ground voltage, while the right plot has  $+230 \text{ V}$  on the cell corners along the enclosing layers and  $-120 \text{ V}$  on each pair of field wires between the sense wires.

Cell Geometry	Cell Size $d_r \times d_z$ (mm $\times$ mm)	Distortion % of cell (see text)	Performance % of cell (see text)
Rectangular 6:1	15.0 $\times$ 20.2	78.	73.
Square 5:1	15.0 $\times$ 15.4	86.	71.
Square 4:1	15.0 $\times$ 15.4	76.	69.
Hexagonal 5:1	15.0 $\times$ 20.2	80.	74.

**Table 5-5.** Time-distance distortions and resolution performance of cells. The distortion column gives the percentage of the cell width that has less than  $100 \mu\text{m}$  of rms variation in the time-distance relationship. The performance column gives the percentage of the cell width that produces signals with at least two primary electrons collected within the first 10 ns of the pulse shape. These values are for tracks within  $\pm 30^\circ$  of the radial direction in an 80:20 He:C<sub>4</sub>H<sub>10</sub> gas mixture and a 1.5 T magnetic field.



**Figure 5-12.** *Spatial resolution performance in a rectangular cell, as described in the text. The jagged contour depicts track angles and impact parameters at which an average of at least two primary clusters will contribute to the first 10 ns of the pulse shape.*

has been calculated at each track angle for regions which satisfy this criterion. Tracks with larger DCA values generally have shorter collection regions within the 10 ns window, and therefore poorer resolutions, while tracks with smaller DCA values have better resolutions. A contour plot of these DCA performance limits as a function of the track angle is shown in Figure 5-12 for rectangular cells. The ratio of the area within the contour and the geometric area of the cell is a figure of merit for the cell. Table 5-5 shows these values for various cell geometries with the number of field wires per sense wire shown (6:1). The 5:1 square cell has the least distortion. The rectangular and hexagonal cells are somewhat worse. The square cell (4:1 field wire to sense wire ratio), having only one field wire between sense wires, has the most distortion.

### 5.4.5 Gain Variations

Gain variation as a function of  $z$  position has been studied using DCSIM, which computes charge densities on sense wires. The gain changes within cells are based on the approximate relation that a 1% change in the charge on the wire results in a 20% change in avalanche gain. In the axial-stereo design, the charge on the wire varies due to the varying cell widths within superlayers and the varying radial size between enclosing layers at boundary cells. The largest gain variation of boundary cells is found to be  $\pm 3\%$ , while the largest gain

variation between an inner layer and an outer layer in a superlayer is  $\pm 4\%$ . In the all-stereo case, the  $\pm 3\%$  increase in cell width at the ends of the wire length leads to a gain variation of less than  $\pm 2\%$  along the wire length. Gain changes due to field wires' phase locations relative to the sense wires from cell to cell in a layer are miniscule in the all-stereo case.

### 5.4.6 Electrostatic Forces and Stability

Electrostatic forces and stability limits have been simulated using DCSIM. The accuracy of the program has been verified with observed wire deflections in the 100-wire test chamber, Prototype I. For this purpose, the 2.5 m-long sense and field wires were tensioned with only 9 g so that wire instability could be observed at nominal high voltage. For the *BABAR* chamber, at the nominal stringing tension, the predicted deflection due to electrostatic forces is  $160\ \mu\text{m}$  for the worst case wire, and the instability point is far from the operating point. Therefore, for the present cell configuration, it appears that there is no problem with electrostatic deflections or stability.

### 5.4.7 Pattern Recognition Studies

#### Axial-Stereo Case

In similar chambers in other detectors, pattern recognition is performed by finding track segments within each superlayer, and then linking segments to form tracks. The left-right ambiguity resolution is done during the segment finding. The axial segments are linked to form circles in the  $x$ - $y$  plane, while the stereo segments provide the  $z$  and  $\theta$  coordinates. After segments have been linked, a full three-dimensional fit is made to all the wire hits. Part of this pattern finding and fitting has been done for the prototype chamber. Some losses may arise at lower momentum due to the reduced efficiency for segment finding within a superlayer or difficulties in segment linking for tracks with significant curvature, particularly for those particles which do not originate from the interaction point.

#### All-Stereo Case

Several studies concerning pattern recognition and track reconstruction in an alternating-layer all-stereo chamber have been made in order to investigate potential limitations of having only two views. In particular, two independent approaches to the pattern recognition problem have achieved efficiencies between 95% and 99% (depending on method and backgrounds) for finding tracks despite the noncircular projections in the stereo views. The absence

of superlayers also raises a concern about unresolvable left-right ambiguities in a single view. This has been shown to occur for less than 0.5% of tracks from generic  $\Upsilon(4S)$  events, and appears to be further reduced by a factor of 2 by the careful choice of the azimuthal offsets of each layer. Most of the ambiguities that do occur are expected to be resolved by information from the other view. Finally, the possible mismatching of track projections found in the two views has been investigated using a parameterized approach [Bri93] and with full pattern recognition and fitting. Both studies indicate that genuine mispairing of projections from opposite views happens in about 0.5% of all cases. With extreme backgrounds, and depending on the matching criteria, mismatches for up to 2.8% of tracks have been found. Most of these, however, arise from the inclusion of small segments of an additional track, or fake track, in one view; the actual contribution to track-finding inefficiency is estimated to be less than 1%.

## 5.5 Gas Choice and Properties

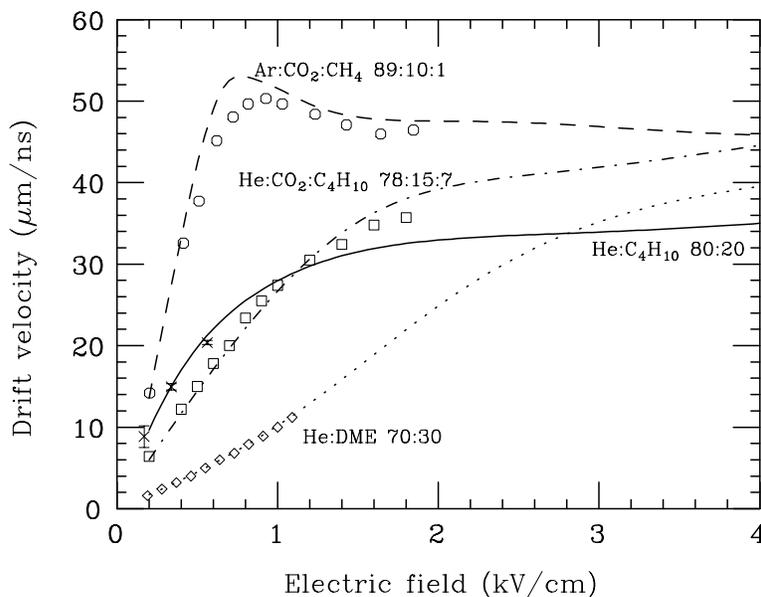
---

The choice of gas for the drift chamber is driven primarily by the needs to reduce the total amount of material, minimize multiple scattering for low momentum tracks, and to operate efficiently in a 1.5T magnetic field. These requirements are quite well satisfied by mixtures of helium and hydrocarbons. In particular, mixtures with 10–30% of various hydrocarbons afford a small Lorentz angle, good resolution, and low multiple scattering. In Table 5-6, the properties of the baseline gas choices are shown, along with other possibilities which have been considered for use in the *BABAR* Drift Chamber. The drift velocities and Lorentz angles are determined with the Boltzmann integration code MAGBOLTZ [Bia89]. The  $dE/dx$  calculations are performed with a modified version of a program from Va'vra *et al.* [Vav82]. The table shows that the helium mixtures under consideration have a radiation length more than five times larger than that of HRS gas, Ar:CO<sub>2</sub>:CH<sub>4</sub> 89:10:1, a commonly used argon-based mixture. Figure 5-13 shows the calculated drift velocity *vs.* electric field for four of the gases in the table. The helium-based gases are not saturated, but lead to better performance than typical argon mixtures, since the smaller Lorentz angle results in a more uniform distance-time relationship.

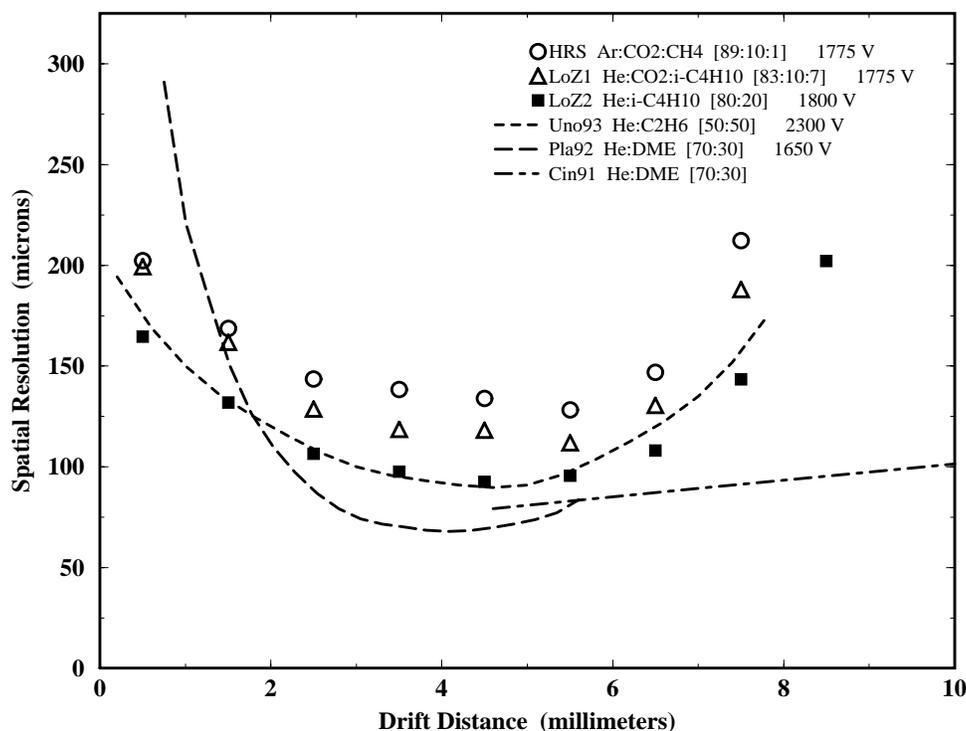
The two gases which have received the most attention are an 83:10:7 mixture of He:CO<sub>2</sub>:C<sub>4</sub>H<sub>10</sub> and an 80:20 mixture of He:C<sub>4</sub>H<sub>10</sub>. These gases and the benchmark argon mixture, HRS gas, have been studied in a prototype chamber, which is described in Section 5.11. The results for the spatial resolution obtained in these studies are summarized in Figure 5-14. Both helium-based gases provide good spatial resolution. The 80:20 He:C<sub>4</sub>H<sub>10</sub> mixture is our preference based on measured spatial resolution and simulated  $dE/dx$  resolution, although there are concerns regarding the safety implications if this proves to be a flammable gas. We have also studied whether these gases exhibit gain changes or high-voltage breakdown

Gas Mixture	Ratio	$X_0$ (m)	Primary Ions/cm	$v_d$ ( $\mu\text{m}/\text{ns}$ )	$\theta_L$ (deg)	$dE/dx$ Resol. (%)	$K/\pi$ Separation	
							$p$ for $3\sigma$ (MeV/c)	# of $\sigma$ at 2.6 GeV/c
Ar:CO <sub>2</sub> :CH <sub>4</sub>	89:10:1	124	23.6	49	52	7.3	665	2.4
He:C <sub>2</sub> H <sub>6</sub>	50:50	686	23.1	31	45	6.6	720	2.1
He:DME	70:30	723	22.4	6	8	6.7	720	2.1
He:C <sub>3</sub> H <sub>8</sub>	70:30	733	24.1	24	36	6.5	730	2.2
He:C <sub>4</sub> H <sub>10</sub>	80:20	807	21.2	22	32	6.9	710	2.1
He:CO <sub>2</sub> :C <sub>4</sub> H <sub>10</sub>	83:10:7	963	13.8	19	26	8.5	660	1.7

**Table 5-6.** Properties of various gas mixtures at atmospheric pressure and 20°C. The drift velocity ( $v_d$ ) and Lorentz angle ( $\theta_L$ ) are given for an electric field of 600 V/cm with no magnetic field and with 1.5 T, respectively. The  $dE/dx$  resolution is calculated for a minimum-ionizing particle. Also listed are the momenta below which there is at least  $3\sigma$   $K/\pi$  separation, and the  $K/\pi$  separation at 2.6 GeV/c.



**Figure 5-13.** Calculated and measured drift velocities as a function of electric field for zero magnetic field. The calculations use the code of Biagi (see text). The solid curve and crosses [Pla92] are for the 80:20 He:C<sub>4</sub>H<sub>10</sub> mixture; dot-dashed curve and squares [Boy92] are for 78:15:7 He:CO<sub>2</sub>:C<sub>4</sub>H<sub>10</sub>; the dotted curve and diamonds [Cin91] are for 70:30 He:DME; and the dashed curve and circles [Boy92] are for (89:10:1) Ar:CO<sub>2</sub>:CH<sub>4</sub> (HRS gas).



**Figure 5-14.** The prototype drift chamber results for HRS gas, He:CO<sub>2</sub>:C<sub>4</sub>H<sub>10</sub>, and 80:20 He:C<sub>4</sub>H<sub>10</sub>. Points represent data from the Prototype I chamber; curves are the results of studies performed elsewhere.

after a period of irradiation. The aging studies were performed with a small proportional counter and an <sup>55</sup>Fe source. All of the helium-based gases in Table 5-6 were tested except for the DME mixture. The isobutane and propane mixtures showed negligible aging. The 50:50 He:C<sub>2</sub>H<sub>6</sub> and 78:15:7 He:CO<sub>2</sub>:C<sub>4</sub>H<sub>10</sub> mixtures both had some gain loss, equivalent to ~30% for an accumulated charge of 1 C/cm. These studies indicate that all of these gases are appropriate for our purposes since the expected charge accumulation over the lifetime of BABAR should be under 0.1 C/cm. Further long-term aging studies are planned.

Angle (Degrees) Outer/Inner	Thickness (mm)	Deflection (mm)	Maximum Stress (MPa)
0/20	6.5	36.2	154
5/20	6.5	18.3	75
10/20	6.5	6.6	50
20/20	6.5	2.2	32
30/20	6.5	1.2	24
20/20	4.7	3.4	51

**Table 5-7.** Finite-element calculations of deflections and stress for various endplate shapes. A 2400 kg load was evenly applied between 22.5 and 80 cm and supported at the outer radius. The endplate is supported at the outer radius only, and the maximum deflection occurs at the inner radius.

## 5.6 Mechanical Design

### 5.6.1 Endplates

The chamber endplates must support the wire load of  $\sim 2400$  kg yet be of negligible thickness in radiation lengths to minimize photon conversion before the calorimeter. The offset of the chamber and the beam energy asymmetry at PEP-II lead to boosted events such that photons with  $\cos \theta_{cm} > 0.67$  exit through the forward endplate, while only those with  $\cos \theta_{cm} < -0.91$  exit through the rear endplate. The two endplates will be identical, but the readout and high-voltage connections will be installed on the backward endplate only.

The choice of carbon fiber as the structural material and the application of the double cone geometry shown in Figure 5-2 allow for extremely thin endplates. The material proposed is an intermediate modulus graphite fiber IM7 [IM7] coupled with a toughened epoxy (977-2) or a cyanate ester resin system (954-2). The mechanical advantage of the double-cone shape is illustrated in Table 5-7 which shows material thicknesses, deflections under wire load, and maximum stress for various endplate configurations. The entries in the angle column represent the angles of the outer and inner cones with respect to a flat endplate—thus 0/0 would be a flat end plate, and 20/20 is closest to the  $22.7^\circ$  symmetric cone illustrated in Figure 5-2 as the baseline design. The scale for acceptable deflections is set by the 9 mm elongation of the aluminum field wires under the 53 g stringing tension. The factored limit stress in carbon fiber is estimated at 61 MPa, which includes a de-rating factor of 3 for the holes, and a safety factor of 3.

The double-conical endplate provides substantial mechanical advantages without seriously degrading the forward tracking length, or introducing too large an angle for axial drilling of the feedthrough holes. The radius of curvature at the cone tip (the transition between the inner and outer cone) has been limited to 50 mm in order to reduce the stress associated with sharper angles. The curvature also allows better laminate properties in the transition area by reducing fiber breakage due to expansion and contraction during the cure and allows the hole pattern to be drilled without any discontinuity.

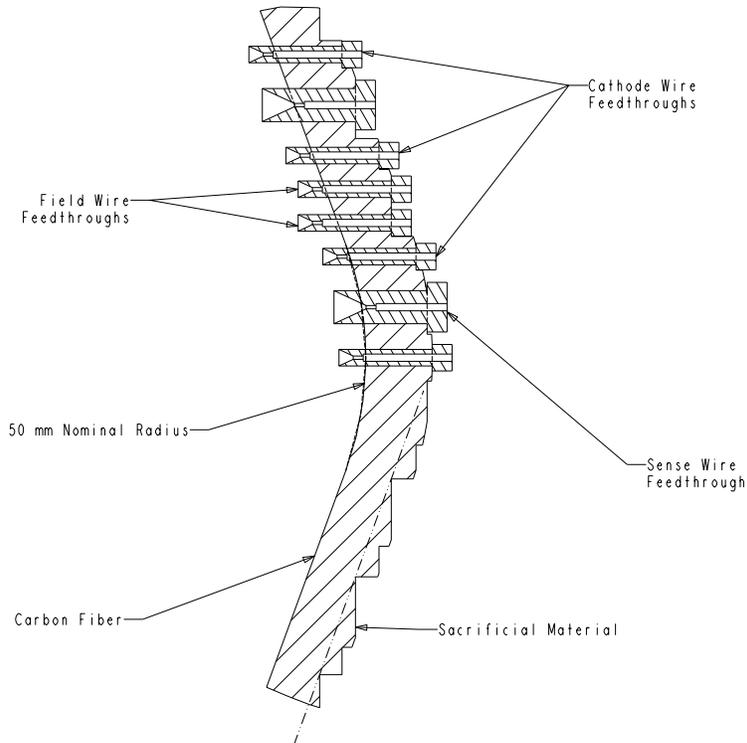
The endplate thickness of 6.5 mm corresponds to 2.8% of a radiation length; the feedthroughs (1.2%), rf-shield (0.6%), and cable support spider (0.4%) roughly double the material thickness of the forward endplate to about  $0.065X_0$ . Cables, preamplifiers, and HV boards add an additional  $\sim 0.1X_0$  to the backward endplate. The endplates will be laid-up on a mandrel, covered with rubber cure pads, then bagged and cured in an autoclave. Both of the resin systems under consideration are cured at 350°F under a pressure of 100 psi. A 3 mm sacrificial layer of nonstructural material will be applied to the outsides of the endplates. A considerable fraction of this layer is then removed as flat annuli are machined for each wire layer as shown in Figure 5-15. These surfaces allow machining tools to make a perpendicular approach to the surface when drilling the feedthrough holes and provide a registration surface for the feedthroughs.

The holes for the field wire feedthroughs will be approximately 2.5 mm in diameter, while the sense wire holes will be 5–7 mm in diameter, depending on the choice of restringing options currently under consideration. The holes will be drilled axially, which simplifies the drilling process, the hole survey, and connections to the electro-mechanical boards. The wires exit from the feedthroughs at the appropriate stereo angle, and the wire position accuracy is therefore weakly coupled to the insertion depth of the feedthrough. The depth of insertion is controlled by the machined annuli, which act as registration surfaces.

### 5.6.2 Inner Wall

For good  $p_t$  resolution, it is essential to minimize the material between the silicon vertex detector and the drift chamber. The inner wall of the chamber lies immediately outside the  $0.005X_0$  support tube and represents a small amount of material by comparison. The inner wall is not required to carry any of the wire load but must support any differential pressure between the inside and outside of the chamber.

An engineering case study has been performed [Smi94], which assumed that the inner wall must withstand a test pressure of 2 kPa (20 mbar), and the results are summarized in Table 5-8. While beryllium would provide a minimal amount of material, this option would be very expensive and unlikely to be feasible with the required thickness of less than 0.5 mm. Therefore, the inner wall will be constructed from 0.4 mm of the same carbon fiber



**Figure 5-15.** Cross section through the endplate cone tip showing the feedthrough seating. Narrow annuli are milled to seat feedthroughs in the cathode layers and alternate with wider annuli containing the sense wire feedthroughs and, in the case of a 6:1 cell design, pairs of field shaping wires.

as the rest of the chamber. An rf-shield will be required which will contribute an additional 0.1% of a radiation length, and concerns about helium permeability may lead to a Mylar layer as a gas seal.

### 5.6.3 Outer Wall

The outer wall bears the axial wire load between the endplates. The large circumference allows this load to be supported by as little as a 1.6 mm-thick ( $0.006X_0$ ) carbon-fiber tube. In order to be robust against local impacts, a 3.2 mm-thick ( $0.012X_0$ ) option is being pursued as more realistic. This is still a small amount of material in comparison to the  $0.18X_0$

Material	Thickness (mm)	% of Radiation Length
Radel [RAD]	2.45	0.82
Vitrex PEEK 450G [VIT]	2.19	0.73
Carbon Fiber IM7/977-2 [IM7]	0.40	0.15
Beryllium	0.32	0.09

**Table 5-8.** *Inner wall thickness required to withstand a test differential pressure of 20 mbar, determined from Roark's Formulae [You89] assuming no stiffeners and ends held circular.*

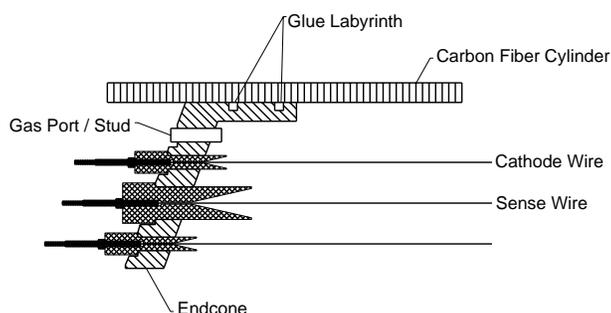
contributed by the DIRC, which is located between the outer wall of the tracking chamber and the calorimeter.

During operation, the chamber will be maintained at a constant temperature. However, during transport, downtime, and other anomalous situations such as magnet quenches, relatively large temperature fluctuations might be encountered; in such cases, thermal expansion can lead to significant changes in wire tension. An aluminum outer wall would provide a perfect thermal match to aluminum field wires but lead to an increase in the tungsten sense wires' tension of about 0.5% per °C. The thermal expansion coefficient of carbon fiber is well matched to that of the sense wires but leads to a decrease in the aluminum field wire tension of 0.8% per °C. The sign of this effect and the fact that the field wires rather than the sense wires are at risk make carbon fiber preferable to an aluminum shell.

The outer tube will be constructed of the same material as the endplates. An aluminum foil will be adhesively bonded to the outer surfaces to provide rf-shielding. The inner surface may be treated with a carbon spray to ensure adequate electrical conduction to drain charge. The low moisture absorption characteristic of the carbon fiber is expected to eliminate any concerns about humidity control.

#### 5.6.4 Joints

The inner and outer walls will be installed after the chamber wires have been strung. The outer joint is within the physics acceptance, so every effort must be made to minimize the amount of material. Both joints must occupy minimal radial space to preserve the maximum possible tracking length. The outer wall will be attached using a double-labyrinth gluing technique that has been used on several previous chambers. A conceptual sketch of the outer joint is shown in Figure 5-16. For gluing, the chamber is turned vertically and epoxy is injected via many syringes into the lower labyrinth and eventually flows to the upper.



**Figure 5-16.** *Conceptual design for the outer-wall/endplate joint.*

The glue injection ceases once the upper recess is filled and glue starts to appear at small air-escape holes.

The inner joint is not within the acceptance region of the detector, so the amount of material is not critical. A removable joint, using an O-ring gas seal, will be used. If at some future date the support tube were removed, this would allow the option of replacing the inner wall with a lower-mass alternative.

The inner and outer joint regions must also accommodate gas fittings and attachment points, the latter being used both for supporting the chamber and for pretensioning the endplates prior to stringing. A ring of studs will be embedded in the carbon fiber endplates at the outer and inner radii during the fabrication process. The studs alternate in an appropriate ratio with gas fittings. The advantage of embedding during fabrication is that none of the structural fibers are broken.

### 5.6.5 R&D Program on Structural Components

The proposal to use carbon-fiber laminates as the main structural component of the chamber requires a thorough research and development program. Such materials are widely used for applications such as satellite structures [Bra94], in which extreme physical conditions are coupled with zero access for maintenance. For high-energy physics applications, the long radiation length of carbon fiber (about 25 cm) is an additional attraction. The R&D program will examine basic material properties and fabrication techniques using small samples, or coupons, followed by the evaluation of increasingly substantial engineering models which address specific areas of concern, leading eventually to the construction of a full-sized prototype endplate. In parallel with this process, theoretical models will be developed to allow the calculation of properties using finite element analysis and classical lamination theory. The interaction between these parallel developments should lead to a complete

mechanical specification and an integrated finite element analysis model of the entire chamber prior to the construction of the full-sized prototype.

At the coupon level, investigations will include evaluation of basic material properties with standard tests (American Standard Test Methods) such as tensile strength (ASTM D3039), humidity absorption (ASTM D560), and adhesive tensile lap strength (ASTM D1002). Design-specific tests such as the evaluation of outgassing effects on chamber aging and the helium permeability of the proposed laminates will be performed. Extensive drilling tests can also be carried out at this stage in order to investigate hole accuracy, ovality, breakout, tool-bit wear, lubricant absorption, procedure elapsed time, and quality assurance techniques.

Engineering models will be developed to address the joint design with respect to strength, gas-seal, and assembly procedure. Model tubes will be constructed to verify compression and buckling calculations. The insertion, seating, and gas-seal of the feedthroughs will be investigated, and a full segment of the endplate will be constructed to confirm the laminate lay-up and cure procedures.

After the engineering model tests and the laminate design have been completed, and after a manufacturing technique, a machining strategy, and quality assurance methods have been developed, a full-sized prototype endplate will be constructed and partially drilled to verify the entire fabrication procedure. Such a prototype may subsequently be further drilled to investigate deflection properties and could be shipped to the stringing site for evaluation of the proposed prestressing and load transfer technique.

### 5.6.6 Wires

Transverse momentum resolution is dominated by multiple scattering up to the highest momenta of interest. This has led to the design of a chamber with low-mass field wires and a helium-based gas. Aluminum has a relatively long radiation length and is a good field wire candidate. Bare aluminum is not suitable, however, because oxidation forms an insulating layer on the wire surface leading eventually to electrical discharges. Previous chambers have used gold-coated aluminum, but thin uniform layers of gold have been hard to achieve. We plan to use 55  $\mu\text{m}$ -diameter aluminum wire with a 0.50  $\mu\text{m}$  gold coating. The contribution of the gold to the total radiation length is roughly equal to that of the aluminum. Unfortunately, gold coatings this thin tend to flake off and provide nonuniform coverage. Other alternatives, such as silicon carbide wires and nickel-coated aluminum wires, have been investigated but proved unacceptable.

Aluminum wires also suffer creep, causing wires to lose tension over time. With careful choice of the wire type, this can be limited to an acceptable level of less than 0.1% of the wire length per year after an initial rapid increase in length over the first few hundred hours.

The relative softness of aluminum means that considerable care must be taken to design a reliable crimping procedure.

For the sense wires, 20  $\mu\text{m}$  tungsten is proposed, which would contribute approximately one-half the material contained in gold-plated aluminum field wires. Although tungsten-rhenium has recently become popular as a wire material, its use has been ruled out for the *BABAR* drift chamber due to the  $\sim 50\%$  increase in resistance, which would degrade the signal-to-noise performance of the chamber. Some consideration is being given to a larger 30  $\mu\text{m}$  wire to improve signal preservation of signal amplitude from the forward end of the chamber. This would increase the sense wire material by a factor of 2.25, making it comparable to the field wires.

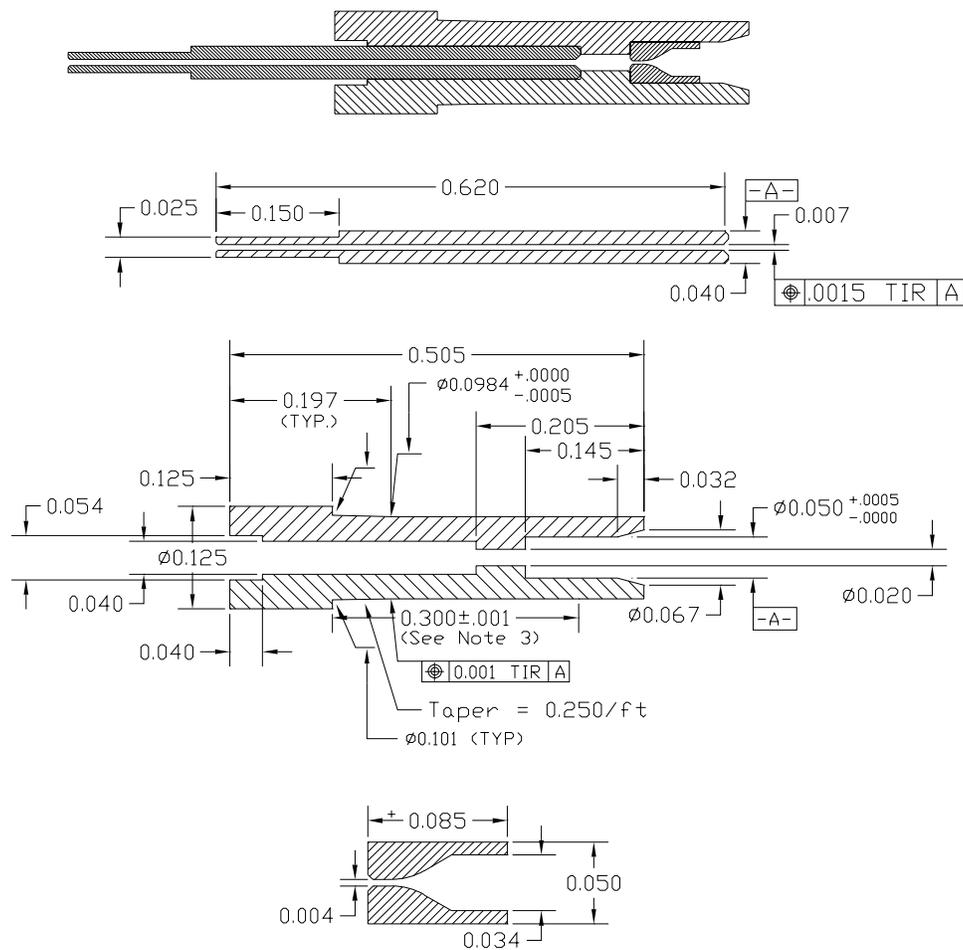
Wire aging and cross-talk tests will be performed, and a full understanding of the mechanical and electrical wire/feedthrough/endplate interface will be sought using a full-length prototype chamber, Prototype II.

### 5.6.7 Feedthroughs

Feedthroughs insulate field and sense wires electrically from the endplates and provide for accurate wire placing. They must be inexpensive to manufacture, must provide a gas seal, and must not react chemically with the chamber gas.

A preliminary field-wire feedthrough design is shown in Figure 5-17. The body of the feedthrough is a single piece of injection-molded plastic with a 2.5 mm precision outer diameter and a length of about 15 mm. Wire positioning is provided by a separate brass insert containing a 100  $\mu\text{m}$ -diameter precision hole. This has a 1 mm bend radius on the feedthrough exit end to avoid kinking stereo wires. A metal insert (brass) was chosen to provide low electric field density at the interface with the plastic body, avoiding discharge problems in the insulator. Delrin or celenex plastic insulation is presently being considered.

The wire is secured in the feedthrough using a gold-plated aluminum crimp pin of a design similar to that used successfully in several TRIUMF-built drift chambers. The TRIUMF pins were actually made of brass, which has the advantage of providing a more rigid contact for connecting the electro-mechanical boards, but provides for a narrower tolerance for crimping wires than a softer material such as aluminum. Good results in terms of wire breakage have been obtained at SLD where aluminum wires were crimped in softer aluminum crimp pins. The small diameter region of the crimp pin is the crimp locator, while the large diameter region is used for electrical connections. This separate-function design is expected to reduce a crimp-pin breakage problem observed at SLD, where the mechanical connections were made in the same region as the crimp.



**Figure 5-17.** Field-wire feedthrough design, showing crimp pin, Delrin sleeve and precision brass insert for wire location.

Wire positioning is provided by accurate centering of the  $100\ \mu\text{m}$  locator hole in the brass insert of the feedthrough. Measurements of 1000 feedthroughs manufactured using a similar design for the Beijing Energy Spectrometer (BES) showed that the concentricity of the precision hole with respect to the feedthrough outer diameter was  $11\ \mu\text{m}$ . The feedthrough itself is required to have an outer-diameter tolerance of  $10\ \mu\text{m}$ , as was achieved with the BES feedthroughs, and must be fitted tightly in the endplate hole.

The gas seal is provided by a slight (1:48) taper at the back of the feedthrough, which may be adjusted to provide the desired degree of press fit into the endplate. The crimp pin is press fit in the plastic body and may be sealed with epoxy if necessary.

The sense-wire feedthrough design will have a larger (5–7 mm) outer diameter and a longer body (25 mm). This provides both thicker insulating walls and a longer projection from the endplates to better shield the HV from the grounded endplate. The precision inserts will be identical to those used for the field wires, and a brass/copper crimp will be used.

Feedthrough prototypes are expected to be available in early 1995. Planned tests include dimensional tolerance checks both at the time of molding and over long term to study stability, as well as checks of the electrical properties of the feedthroughs. The Prototype II chamber will use the new feedthroughs to test their performance in a real chamber environment, including electrical isolation and gas leakage tests.

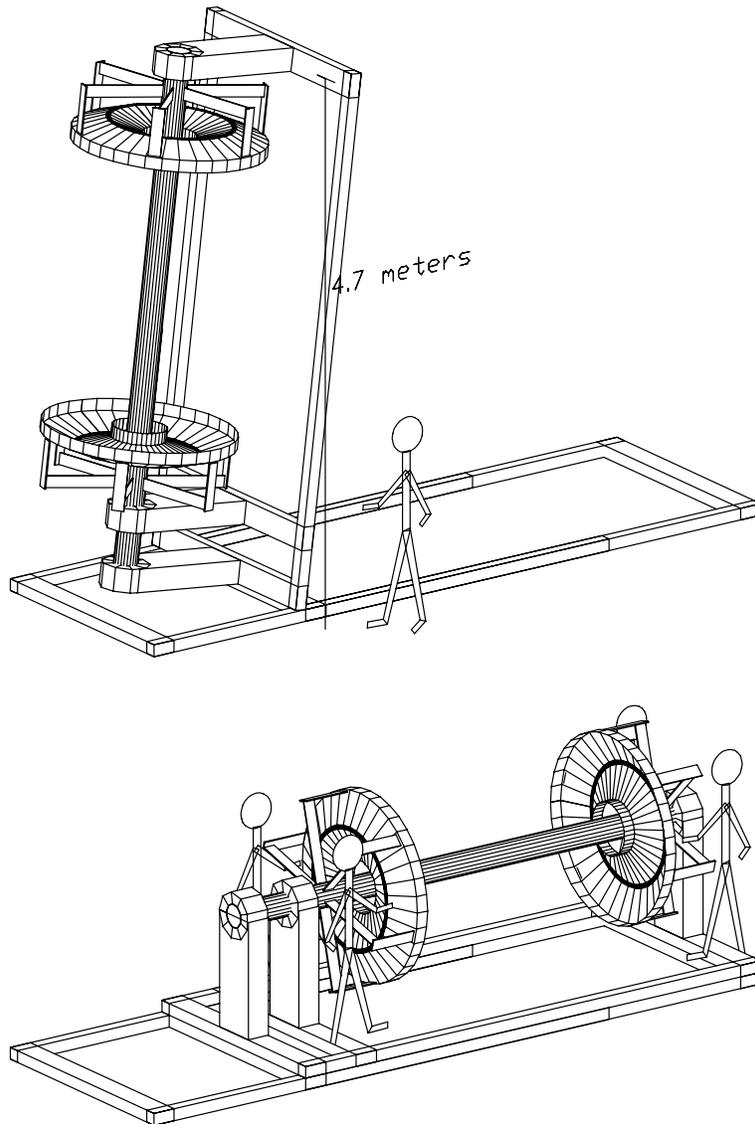
### 5.6.8 Stringing

The chamber will likely be strung in a large clean-room at TRIUMF in Vancouver. The  $7.82 \times 11.45$  m, Class 10,000 clean-room has a ceiling structure containing lighting fixtures and 18 distributed HEPA-filter/air intake blowers. The stringing stand proposed for the chamber would be about  $2.0 \times 4.7$  m and would require a ceiling height of about 6.0 m when rotated into the vertical position. This would necessitate raising the ceiling of the TRIUMF clean room by about 3.0 m. Temperature control ( $\pm 2^\circ\text{C}$ ) will also be installed.

The chamber will be strung horizontally without the outer or inner cylindrical shells in place. Each endplate will be attached from its outer edge to a central axle of the stringing stand by means of external support spiders. The endplates will be preloaded at their inner radii, and possibly at intermediate positions as well. All loads are transferred to the central axle by the spiders. Stringers will be able to work between and under this support structure, while the interior region of the chamber is completely unobstructed.

Two teams of three people (or two people and a robot) each may work simultaneously as the chamber is strung from the inner radius out. The wire is inserted through one endplate by a stringer, transported to the other endplate either by the robot or by the third stringer, and is then inserted through the appropriate hole for the second stringer. Both stringers thread the wire onto a feedthrough/crimp-pin assembly and insert the assembly into the endplates. When the first stringer has crimped one end of the wire, the second stringer may tension and crimp the other. A robot would offer a smooth method of transporting the wire and would ensure that the correct hole was used. Depth perception is notoriously bad when looking at a field of strung wires, and a robot would help to avoid stringing errors. The chamber can also be kept cleaner by enclosing the space between the endplates, including the robots, in a very clean tent within the outer clean room.

Wires which fail during the stringing process will be restrung with the chamber rotated into a vertical position, aligned at the appropriate stereo angle. A magnetized needle will be



**Figure 5-18.** *The stringing and pretensioning stand for chamber assembly, showing the normal horizontal position for wire insertion with two stringing teams and a vertical position for wire replacement.*

lowered through the top hole and captured by a similar needle inserted into the lower hole. The stringing stand illustrated in Figure 5-18 has a dual support at one end, allowing the chamber to be cantilevered for the installation of the inner and outer shells. The inner wall supports no axial load and is installed after stringing to avoid accumulation of dust.

### 5.6.9 Endplate Connections

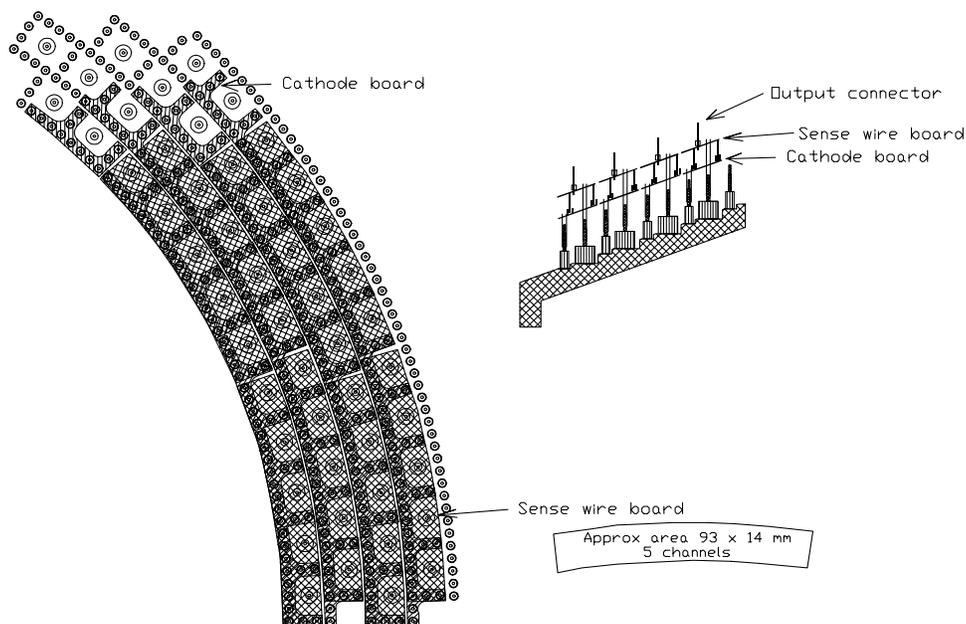
#### Electro-Mechanical Boards

The electro-mechanical boards form the interface between the feedthroughs and the electronics. They perform the following functions:

- Provide a ground plane for the cathode wires at both ends of the chamber. Insulating feedthroughs are being used for the field wires to allow tension measurements to be performed and broken wires to be located. In the axial-stereo layout, some of the cathode wires will be biased at a few hundred volts, which also requires insulated feedthroughs.
- Provide high voltage to the sense wires. This will be done on the backward endplate only.
- Provide collection of signals from sense wires and distribution of calibration pulses to a standard preamplifier card. All signals will be extracted from the backward endplate. No cables will be present on the forward endplate.

Thin cathode boards are mounted on both ends of the chamber and separate sense wire boards are mounted on the backward endplate only. The cathode boards are copper-coated, printed circuit boards approximately 0.25 mm thick, providing a low-resistance, low-inductance connection. A good quality ground is essential to avoid cross-talk between sense wires arising from induced transient signals on neighboring field wires (Figure 5-19). The boards are mounted parallel to the endplate and connect to the cathode wire feedthrough pins. Each cathode board is connected to neighboring boards with removable jumpers for the ground and bias voltages.

The  $\sim 0.25$  mm-thick sense-wire circuit board is mounted on the backward endplate parallel to the cathode boards but separated by approximately 6 mm to avoid high voltage (HV) breakdown. The connection to the sense-wire crimp pins is made through the large holes in the cathode boards as illustrated in in Figure 5-19. At the forward endplate, the sense-wire crimp pins will be capped to prevent corona discharges into the air. The sense wire boards will provide a standard connection for the preamplifier cards, for ease and safety of



**Figure 5-19.** One concept for endplate cathode boards for grounding field wires and HV/signal distribution boards. In this example, the cathode boards connect together 12 cathode wires and are jumpered together to form the ground plane. The sense wire boards mount over the ground plane and provide a standardized connection between the preamplifiers and five sense wires.

installation and replacement. The cathode boards and sense wire boards will rarely, if ever, be removed.

The sense wire boards will also include a HV bus, as well as the isolating resistors and capacitors that couple to the sense wires. Jumpers will connect adjacent boards in the  $\phi$ -direction, allowing distribution of high voltage. This scheme allows the HV to be applied without the preamplifier cards attached, which offers greater flexibility and safety during testing or troubleshooting.

### Preamplifier Cooling

At a constant voltage, the gas gain depends strongly on the gas density, which is proportional to pressure and inversely proportional to temperature. Therefore, for good  $dE/dx$  resolution, both the pressure and temperature must be kept constant or at least be well-monitored. While pressure is always uniform across the chamber, temperature differences at the chamber surface may generate differences within the gas volume. The total heat generated by the preamplifiers is on the order of 35 W and may be removed by flowing nitrogen gas.

## Cables

The coaxial HV cables and ribbon-style signal cables will be attached to the backward endplate only. A typical HV cable as considered here has a 3 mm outer diameter and corresponds to about 6% of a radiation length. Approximately 220 cables will be routed radially over the endplates to the various HV distribution points. The volume of HV cables is sufficiently small that only a single layer will be formed at the outer radius of the chamber.

The ribbon cables supply the preamplifiers with power, carry test lines, and the preamplifier outputs. A typical unshielded PVC coated ribbon cable is 33 mm wide, 2 mm thick, and corresponds to 0.75% of a radiation length. The cables are stacked six deep, corresponding to 4.5% of a radiation length at the outer radius.

A low-mass structure is required to support the cables over the endplate. The anchor points will be the studs at the inner and outer wall joints. Cables are supported on a spider-like structure which will be constructed of three rings, one each at the inner radius, the cone tip and the outer radius, connected by spokes. In the axial-stereo case, intermediate connections to the endplate are possible at the three hole-free superlayer transition regions. Approximately 6 cm of clearance is required between the endplate and the support to allow for card replacement.

## 5.7 Front-End Electronics

---

The front-end electronics provide amplification and shaping of the signals for optimal timing and pulse height ( $dE/dx$ ) resolution. We have identified parameters of the drift chamber which will impact the design of the front-end electronics. For example, the combination of wire resistance and terminating resistor is being optimized for best signal/noise behavior. The shaping time and required dynamic range depend on the drift gas selected and the mode of drift time measurement (TDC, FADC) and are the subject of simulations. In order to optimize power and noise performance, the on-chamber, front-end electronics will consist of a bipolar ASIC chip serving several channels. The front-end electronics are mounted on PC boards, which connect directly to the feedthrough pins on the backward endplate.

The amplified signals are transmitted about 30 m to the readout cards using shielded twisted pairs. The signals are differential to minimize the emission of electromagnetic interference and to allow for common mode rejection at the ends of the transmission lines. They are sent to a readout card containing discriminators and pulse-height digitizers based on a deadtimeless FADC system. The discriminator outputs are used by the trigger segment finder logic, and the FADC samples are stored continuously in a buffer memory. When a trigger is received, the FADC samples are read from the dual port buffer memory, and the

drift time and charge information is extracted from the FADC raw data and transmitted to the event builder. The sampling clock will probably run at 79 MHz, a submultiple of the 238 MHz machine frequency. This will allow the suppression of recurrent machine-related noise and yield a timing resolution on the order of 2 ns, which is sufficient to match the intrinsic position resolution of the chamber (about 125  $\mu\text{m}$ ). Further details of this readout system can be found in Chapter 10.

## 5.8 High Voltage System

---

The high voltage (HV) system will provide  $\sim 2$  kV to the 7000 sense wires while minimizing the impact of HV problems, such as sparking in the chamber, broken wires, and failed HV capacitors. Forty HV supplies will be used to drive approximately 220 HV cables. This number ensures that a very low current is drawn even with high background levels. A sensitive trip level will protect against wire breakage due to sparking. Forty supplies will also provide enough current to ramp the HV in less than ten seconds, thereby minimizing detector downtime during frequent injections. A possible implementation would involve using the LeCroy 1454 system, consisting of a mainframe interfaced to VME or CAMAC which houses four HV cards, each having 12 channels capable of providing up to 3 kV and 2.5 mA.

In the example illustrated in Figure 5-19, a set of five sense wires receives voltage from a single HV-distribution card located on the backward endplate. The HV card contains a HV bus with jumpers to adjacent cards, and an isolation resistor and one or two decoupling capacitors for each wire. The resistor values for the sense wires will be approximately 0.5 M $\Omega$  per wire, ensuring that the voltage drop across the resistor is  $\leq 1$  V. The induced current in the first sense-wire layer has been estimated to be approximately 50 nA for the nominal background levels. Two HV capacitors mounted in series would mean that a single failure (short) would not affect the performance of the chamber; however, space limitations on the endplate may not allow this solution. The capacitance value will be selected on the basis of signal propagation measurements and calculations.

Each HV cable supplies four such HV distribution cards; the jumpers are used to distribute the voltage between cards. If there is a HV problem, 32 sense wires ( $\sim 0.3\%$  of the total) will be affected until the next access to the drift chamber endplate.

---

## 5.9 Calibration and Monitoring

---

### 5.9.1 Calibration

The timing characteristics and the gain of the electronics chain will be calibrated with a pulser system. The input of every preamplifier channel is connected via a coupling capacitor and isolating resistor to a calibration bus. When a voltage step is applied to the bus, a defined charge is injected into the front-end of the preamplifier, and the whole electronics chain can be calibrated. The matching of these calibration capacitors on bipolar ASICs is excellent. It has been shown to be advantageous to connect neighboring channels to different calibration buses. It is planned that every chip will be serviced by four buses, and that these buses will be distributed to a limited area of the drift chamber, similar to the case of the high voltage and preamplifier power supplies.

### 5.9.2 Slow Controls

The slow control system will monitor the operation of the drift chamber, the high voltages, power supplies, and test pulse distribution. The front-end hardware can be monitored with the help of suitable online histograms. However, it is essential that the proper hardware diagnostic tools exist in order to detect the failure of a particular component as soon as possible. The histogram observation method frequently implies rather substantial delays in discovering a component failure. A hardware sensor can set off an alarm instantly when, for example, a power supply fails.

The slow control system is designed to respond to expert commands from a local console, generate programmed sequences in response to a command issued from the run control, and take corrective actions in the event of an abnormal situation (voltage trips, *etc.*). This requires that the system have its own local processor, a set of readout and control cards to communicate with the appropriate sensors and control devices, and a link to a local area network (LAN) to receive the run control commands.

The slow controls also generate alarm messages to the central error reporting system as well as status reports. Most of these messages can be handled through the LAN, but some alarms—the hazardous gas monitors, for example—will act on specified hardware interlocks. A link is also required to a database containing the mapping for the physical channels, the calibration constants, and, if required, some voltage or current limits for a channel.

**High Voltages.** The sense wires, nominally at  $\sim 2$  kV, will be monitored for the current they draw. The slow control system provides fast ramp-up and ramp-down of voltages. The field wires, maintained at ground potential, do not require monitoring except in the case of small bias voltages for the axial-stereo design.

**Low Voltages and Currents.** The slow control system will also be used to monitor the voltages and currents in the preamplifiers and the local CAMAC, VME, and VXI crates. This will be achieved with slow 64-channel ADCs having individually programmable gains and differential inputs. About 150 channels are required for the front-end crate power supplies, and about 450 to monitor each preamplifier card, assuming 32 channels and two voltages per card.

**Temperature, Pressure, and Gas Flow.** The same slow ADCs will be used to monitor ambient temperatures and pressures at various locations and the response from a gas gain monitor. Air flow switches will be handled with parallel input registers.

All of the above controls can be realized using commercial CAMAC modules controlled by a VME processor via a standard A2 CAMAC controller and a VME CAMAC branch driver.

**Slow Control via Data Monitoring.** The performance of all cells will be monitored locally by histogramming of hits in each cell, track segment density, amplitude and time distributions, and other relevant distributions. These histograms are inspected either visually or, in some cases, automatically if the shape can be compared to a template histogram.

### 5.9.3 Monitoring

The slow controls described above provide one means of monitoring gas properties such as drift velocity as a function of electric field, the gas gain at the operating point of the chamber, the composition and purity of the gas, and the level of contaminants (such as  $O_2$  and  $H_2O$ ), which could influence the operation of the chamber. In addition, the quality of incoming premixed gas bottles will be monitored by small test chambers before being introduced to the gas system.

The gas flow and the relative pressures and temperatures must be monitored at various locations for proper operation of the chamber, since the drift time and gas gain depend on the gas density. The chamber pressure is expected to follow the ambient pressure, which fluctuates by  $\pm 4\%$  for this geographical region. These pressure differences will affect the chamber gain and drift velocity. The gain can be corrected by applying a scale factor to

the pulse height measurements. The drift velocity poses a more difficult problem, since the gas is nonsaturated. Corrections to the track position measurements depend on the distance from the sense wire, and range by at least a factor of 5 over the cell. The SLD experiment measured a 2% change in drift velocities due to ambient pressure fluctuations, which were corrected to 0.2%. The effects of the pressure on the drift velocity are being investigated using the Prototype I chamber. Simulations will be used to determine the best parameterization for the correction.

Monitoring some gas properties inside the chamber would be also advantageous. A laser calibration system is being considered for measurement of the drift time in the chamber volume by ejecting electrons into drift cells through photoemission from field wires using a high-intensity UV laser. The laser beam is transmitted to the chamber via optical fibers. A beam port into the chamber volume will be situated near the midpoint of the outside wall. This arrangement is compact, with few optical elements, such as a small mirror and lens, required to direct the beam transverse to the field wires. The fiber would lie along the outside surface of the chamber with the beam directed into the chamber at  $90^\circ$  by a  $45^\circ$  mirror. The feasibility of this approach will be determined as part of the prototype program.

## 5.10 Integration

---

### 5.10.1 Overall Geometry and Mechanical Support

The mechanical envelope for the drift chamber is shown in Figure 5-2. The envelope includes the space needed for endplates and front-end electronics in addition to the active detector volume. It does not include the volume needed for mechanical supports or cable routing.

The method for supporting the drift chamber is still under study. In one scenario, the chamber is held from the accelerator pedestal in the backward direction and the support tube column in the forward direction. This arrangement allows the support tube to be repositioned (by remote-control cams at each end) or removed independently of the drift chamber. The supports are fastened to the lower edge of the inner radius of each endplate. In another scenario, the chamber is held at its outer radius by the barrel IFR.

### 5.10.2 Cable Plant and Utilities Routing

The cables and utilities that must be brought into the drift chamber from outside the detector include signal, high voltage, and preamplifier power cables; calibration and monitoring

signals; drift gas; and preamplifier cooling gas. All connections are made to the backward endplate, except for gas lines and monitor signals which will be at both ends.

The cross-sectional area and mass are dominated by the signal and HV cables. The cable routing scheme has been laid out assuming 9100 cables of 2.3 mm diameter (signal, monitoring, calibration, and preamplifier power) and 220 cables of 5 mm diameter (HV). The 2.3 mm is approximately the size of RG-174; the actual cables used are likely to be smaller and fewer. Including an additional safety factor of 2, the total cross sectional area reserved for cables is 866 cm<sup>2</sup>. The cables and gas lines will be brought into the detector along 2.5 cm-deep cable trays that cover approximately 70% of the inner radius of the DIRC support tube (*i.e.*, at the outer radius of the drift chamber).

The total cable run to the electronics room is approximately 35 m. If thicker cable is needed to preserve signal quality, a transition box will be located close to the detector.

### 5.10.3 Access

Access to the drift chamber will be needed to replace failed HV capacitors or preamplifiers, to fix gas leaks, and to replace broken wires. None of this is anticipated to be a frequent occurrence. For example, similar front-end electronics have been used on OPAL chambers for seven years without access. In that time, only one of 400 cards has failed. This reliability is achieved both by design—including diode transient suppressors, for example—and by a thorough program of pre-installation testing. A small, randomly distributed number of failed preamplifier cards will have negligible impact on the detector performance. Two HV capacitors will be used in series on each sense wire to improve the reliability of the HV distribution.

A broken wire is a more significant problem than a failed preamplifier. The proposed R&D program will ensure that deformation of the feedthrough and wire due to crimping is understood and will establish the allowed range of crimping force. The crimping tools will be repeatedly calibrated throughout the chamber stringing to ensure operation within this range. With careful attention to such details and a testing period of several months following stringing, experiments such as SLD and BNL-787 have built drift chambers with large numbers of aluminum wires without incurring a single wire break during operation. Tests performed at SLAC indicate that broken wires tend to curl up at the feedthrough rather than fall into the rest of the chamber.

Although regular access will not be required, access procedures have been developed for both ends of the drift chamber. In the forward direction, the endcap-instrumented flux return is vertically split and can be withdrawn along rails at an angle relative to the beam line.

Temporary scaffolding may then be installed, giving very good access to the drift chamber endplate.

In the backward direction, the DIRC support cylinder and water tank create a tunnel containing the Q1, Q2, and Q4 magnets, pumps and masks, flux return iron, and the drift chamber cable trays. For approximately 70% of the azimuthal angle, the return iron can be withdrawn on rails mounted on the support cylinder. After clearing the DIRC water tank, they are then removed by crane.

The time needed to achieve access will be one shift for either endplate. The reassembly time will be longer by approximately one shift for the forward endplate because of the extra time needed to realign the forward calorimeter.

#### 5.10.4 Integration Aspects of the Gas System

The drift chamber has two gas systems, the drift gas (He:C<sub>4</sub>H<sub>10</sub>, for example) and the dry nitrogen used to remove the 35 W produced by the preamplifiers. Work is currently underway to establish the flammability of each of the drift gases under study. If necessary, the rf-shield at each endplate can be used to make a second gas seal. This volume would then be flushed by nitrogen to ensure that a flammable concentration of gas cannot accumulate due to a leaking feedthrough. (In the backward direction, nitrogen is also the cooling gas.) In particular, the nitrogen gas distribution will be designed to ensure that a flammable concentration cannot accumulate between the feedthroughs and the printed circuit boards mounted on them. The exhaust gas and a gas sample pumped from the endplate region will both be monitored for oxygen and isobutane concentration. Fire-retardant cables are required in any event as a fire safety measure for the detector.

A large leak of helium could damage the DIRC or aerogel photomultiplier tubes. In addition to the precautions listed above, the photomultiplier tubes may need to be located in a gas-tight volume.

#### 5.10.5 Installation and Alignment

The drift chamber is installed after the magnet, IFR, calorimeter, and DIRC. The machine components from Q2 to Q4 are absent, but the accelerator pedestal is installed in the backward direction. The drift chamber is inserted from the forward end along a beam passing through the chamber to the accelerator pedestal. The forward support tube column is then installed and the drift chamber connected to its permanent supports.

There is an assembly tolerance of 15 mm between the chamber inner radius and the accelerator support tube outer radius. This clearance also allows the remote movement of the tube, which may be as much as 5 mm.

An optical alignment will be performed after the accelerator support tube has been installed in order to place the drift chamber to within 1 mm, and to align the axis of the chamber to within a few mrad. The final, precision alignment will be extracted from data.

## 5.11 R&D Using Prototype Drift Chambers

---

Simulation and prototyping efforts are underway in the areas of cell geometry, orientation, layer packing, pattern recognition, and track reconstruction. Several stand-alone data analysis packages are being used, and there have been recent advances on the incorporation of the drift chamber into the BBSIM GEANT package. Prototype efforts are continuing at SLAC, with the current prototype, Prototype I, and further prototyping efforts are being planned.

### 5.11.1 Prototype I Chamber

A small-cell prototype chamber has been built and is being used to study electrostatic stability, feedthrough design, front-end electronics, cell crosstalk, and wire and gas properties. The prototype is 2.5 m in length, with flat aluminum endplates, 72 anode wires of 20  $\mu\text{m}$ -diameter gold-coated tungsten, and 512 cathode wires of 55  $\mu\text{m}$ -diameter unplated aluminum 5056. The wires are arranged into three superlayers of hexagonal cells with the center superlayer positioned at a 50 mr stereo angle.

Studies of suitable helium-based drift gases are continuing with Prototype I, using cosmic ray particles. To date, the three component gas mixture He:CO<sub>2</sub>:C<sub>4</sub>H<sub>10</sub> (83:10:7) and the two component gas mixture He:C<sub>4</sub>H<sub>10</sub> (80:20) have been investigated. Preliminary results are included in Figure 5-14, along with measurements from previous gas studies [Cin91, Pla92, Uno93]. Plans are underway to study other gas mixtures, including He:C<sub>3</sub>H<sub>8</sub> and He:C<sub>4</sub>H<sub>10</sub> with lower partial pressures of isobutane. The possibility of including CF<sub>4</sub> as an additive to counter aging is being considered.

Prototype I has also been the test bed for a study of the achievable position resolution using a FADC sampling system and a low-density helium-based gas with pronounced discrete clustering. Data from four anodes, sampled with 8-bit resolution at 250 MHz ( $\simeq f_{\text{accel}}$ ), have been artificially compressed to 125 MHz ( $\simeq f_{\text{accel}}/2$ ) and 83.3 MHz ( $\simeq f_{\text{accel}}/3$ ), and several algorithms for extracting the timing information have been applied to the signals. The results are steadily improving under refinement of these algorithms and currently show

single wire position resolutions of  $127\ \mu\text{m}$  (at 125 MHz) and  $134\ \mu\text{m}$  (at 83.3 MHz) for the gas mixture He:C<sub>4</sub>H<sub>10</sub> (80:20). These should be compared to the single wire resolution of  $109\ \mu\text{m}$  obtained with a discriminator and TDC (500 ps bin size). The timing algorithms have yet to be fully optimized, and so further improvements in these values are possible. In early 1995, the preamplifiers of Prototype I will be replaced by ones of a better design, and all 72 channels will be instrumented with 100 MHz FADCs. The Prototype I chamber will then provide full tracking with FADCs.

### 5.11.2 Prototype II

A second full-length prototype, Prototype II, will be constructed in 1995 and will incorporate all design decisions made to date. It will have sloped carbon fiber endplates, the new crimp-style feedthroughs, and gold-coated aluminum cathode wires.

The preamplifier electronics for this chamber will be furnished by UCSC, in the form of an ASIC with adjustable gain, rise time, and shaping time. This preamplifier will allow precise determination of the characteristics of the final preamplifier design in accordance with the chamber cell response and the timing algorithm used. Cell shape, aspect ratio, layer arrangement, and the segmentation of high voltage, power, and calibration signals will reflect the decisions of the group with respect to the internal chamber geometry.

### 5.11.3 Other R&D Efforts

Depending upon the availability of a test beam, a somewhat smaller test chamber could be used for  $dE/dx$  studies. It would also be desirable to use such a chamber for studies of tracking,  $dE/dx$ , and Lorentz angles in helium-based gas mixtures in the presence of a magnetic field.

Studies on the aging of the drift gas in a high rate environment are underway at TRIUMF and at LBL. The most likely gas candidates, as identified in the Prototype I studies, will be subjected to long term aging tests. The usefulness of CF<sub>4</sub> as a gas additive to counter aging will be investigated as well.

## References

---

- [Bia89] The MAGBOLTZ Simulation Code was kindly supplied by S.F. Biagi of the University of Liverpool. Some results from this code are discussed in S.F. Biagi, *Nucl. Instr. Methods* **A283**, 716 (1989).
- [Boy92] A. Boyarski, D. Briggs, and P. Burchat, *Nucl. Instr. Methods* **A323**, 267 (1992).
- [Bra94] D. Brashford *et al.*, “The Use of High Stiffness Material and Dimensionally Stable Materials in Spacecraft Applications,” in Proceedings of the Int. Workshop on Advanced Materials for High Precision Detectors, CERN 94-07, p. 9 (1994).
- [Bri93] D. Britton and E. Hyatt, “Pattern Recognition in an All Stereo Drift Chamber,” *BABAR Note # 102* (1993).
- [Cin91] V. Cindro *et al.*, *Nucl. Instr. Methods* **A309**, 411 (1991).
- [IM7] Manufactured by ICI Fiberite, Tempe, AZ.
- [Inn93] W. Innes, “TRACKERR, A Program for Calculating Tracking Errors,” *BABAR Note # 121* (1993).
- [Pla92] S. Playfer *et al.*, *Nucl. Instr. Methods* **A315**, 494 (1992).
- [RAD] Manufactured by Dupont.
- [Smi94] C. Smith, McGill Internal Technical Memo, McG-BF-TM-0007, (1994).
- [Uno93] S. Uno *et al.*, *Nucl. Instr. Methods* **A330**, 55 (1993).
- [Vav82] J. Va’vra, L. Roberts, D. Freytag, and P. Clancey, *Nucl. Instr. Methods* **A203**, 109 (1982). The code has been modified to use the number of primary ions in the calculations.
- [VIT] Manufactured by ICI Advanced Materials, Exton, PA.
- [You89] W.C. Young, “Roark’s Formulas for Stress and Strain,” 6th Edition, McGraw Hill, (1989).

---

---

# Particle Identification

---

## 6.1 Physics Requirement and Performance Goals

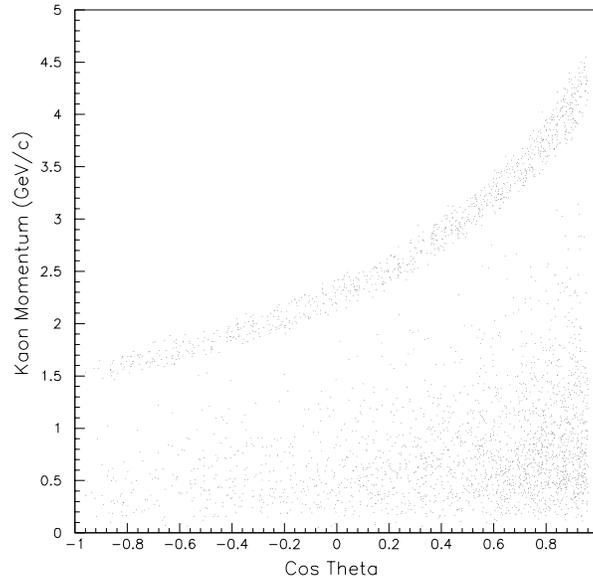
---

### 6.1.1 Introduction

**E**xcellent particle identification for hadrons and leptons over a large range of solid angle and momentum is an essential requirement for meeting the physics objectives of *BABAR*. In particular, measurements of  $CP$  violation require particle identification, both to reconstruct exclusive final states and to tag the quark content of the other  $B$  in the event. Information from the drift chamber, calorimeter, and the instrumented flux return can be used to identify most of the leptons and many of the hadrons. However, these systems are not sufficient to distinguish charged pions from kaons with momenta greater than about  $0.7\text{ GeV}/c$ , or protons above  $1.3\text{ GeV}/c$ , as is required to obtain efficient tagging and event reconstruction. To meet these requirements, dedicated particle identification will be provided by a combination of Cherenkov counters: a ring-imaging detector in the barrel region (the DIRC, for Detection of Internally Reflected Cherenkov light) and a threshold detector in the forward endcap region (the ATC, for Aerogel Threshold Cherenkov counter) with two indices of refraction. No dedicated particle identification in the backward region is needed.

### 6.1.2 $B$ Flavor Tagging

The flavor of  $B$  mesons will primarily be tagged with charged kaons and leptons: a  $b$ -quark decay leads to a  $K^-$  and/or a direct lepton  $\ell^-$ , and a  $\bar{b}$ -quark decay leads to a  $K^+$  and/or a direct lepton  $\ell^+$ .  $B$ -flavor tagging thus relies on the identification of kaons and leptons in an environment of inclusive  $B$  decays with large multiplicities of low-momentum daughters, where relative abundances of pions, kaons, and protons are approximately 7:1:0.2. Because of the boost ( $\beta\gamma = 0.56$ ), the average  $\pi(K)$  momentum depends on the polar angle, and



**Figure 6-1.** *Inclusive momentum spectrum for kaons as a function of polar angle in the laboratory frame. The banded region is an enhanced population of kaons from the decay  $B \rightarrow K\pi$ .*

ranges from about 0.3(0.45) GeV/ $c$  in the backward direction to about 0.75(1.1) GeV/ $c$  in the forward direction.

The effective tagging efficiency is defined as  $\epsilon_{eff} = \epsilon_{tag} \times (1 - 2w)^2$ , where  $\epsilon_{tag}$  is the fraction of tagged  $B$  events, and  $w$  is the fraction of tagged events which are incorrect. The effective efficiency is limited by physics itself, for example, by wrong-sign contamination from Cabibbo-suppressed decays or from kaon or lepton pairs. In the case of kaon tagging, even assuming perfect particle identification and no kaon decays in flight, the maximum tagging rates and minimum wrong-tag fractions which can be achieved within the  $BABAR$  acceptance are 36.3% and 6.6%, respectively. This leads to a maximum effective efficiency of 27.3%. Kaon decays result in a further 20% reduction of efficiency. The loss of efficiency introduced by the PID system must be small in comparison. Given the relative abundances noted above, a small misidentification rate for pions is particularly crucial.

The measurement of ionization in the drift chamber will provide  $\pi/K$  separation better than  $3\sigma$  up to 0.7 GeV/ $c$ , but will be of limited assistance in the relativistic rise region at higher momenta. As a consequence, the  $dE/dx$  measurement is not itself sufficient, since 48% of all kaons would remain untagged, as demonstrated by the momentum distribution shown in Figure 6-1.

Kaon tagging thus requires dedicated PID systems that achieve high efficiency and purity. Furthermore, because the forward region subtends a large fraction of the center-of-mass solid angle, tagging efficiency is sensitive to the forward PID coverage. In the proposed design, the DIRC covers 86.5% of the full solid angle, down to  $\theta = 445$  mr, and, combined with  $dE/dx$ , provides an effective tagging efficiency of 18.5%. The very forward part ( $300 < \theta < 412$  mr) covered by the ATC represents 5% of the full solid angle, and a dedicated PID system in this region, in combination with  $dE/dx$ , increases the effective tagging efficiency by 3.4% to 19.1%. To make this contribution, the forward aerogel needs to be able to tag kaons up to 1.5–2 GeV/ $c$  through the use of an aerogel layer with an index of refraction  $n = 1.055$ .

In the case of lepton tagging, the  $dE/dx$  measurement ensures electron identification over the full momentum spectrum. However, muon identification by range measurement in the flux return is efficient only above about 0.7 GeV/ $c$ . At lower momenta, the DIRC, primarily designed for its  $\pi/K$  separation capabilities, will also provide  $\mu/\pi$  separation, complementary to the IFR.

### 6.1.3 Exclusive $B$ decays

The search for  $CP$ -violating modes, such as  $B^0 \rightarrow \pi^+\pi^-$  or  $B^0 \rightarrow \rho\pi$ , and the measurement at much higher momentum of  $V_{ub}$  through charmless  $B$  decays, places greater demands on the performance of the PID system than those required for tagging. The exclusive branching ratios of interest are very small ( $10^{-5}$ ), and  $CP$  violation studies are made even more challenging by the need to tag the other  $B$  in  $\Upsilon(4S) \rightarrow B^0\bar{B}^0$  decays. While much of the background from the continuum can be suppressed by simple kinematic cuts, the reduction of correlated backgrounds from  $B^0 \rightarrow \pi^+K^-$ ,  $B^0 \rightarrow \rho K$ , and  $B^0 \rightarrow K^*(892)\pi$  demands that pions be positively identified, with a minimum of kaon misidentification, in the high-momentum range up to 3–4.5 GeV/ $c$ . Non- $B$  physics will also be demanding in terms of kaon identification. A challenging example is the study of rare strange decays of the tau lepton in which the momentum spectrum is very hard (up to 8 GeV/ $c$ ).

An example of a detailed analysis of a two-body mode requiring high-momentum particle identification is given in Chapter 3. A study of the decay  $B^0 \rightarrow \pi^+\pi^-$  is reported which relies on the combination of DIRC and ATC to suppress the correlated  $K^+\pi^-$  background. These ASLUND studies show that  $dE/dx$  and kinematics, while providing significant capability for identification of the  $B^0 \rightarrow \pi^+\pi^-$  exclusive channel, are not sufficient to guarantee both high efficiency and low background. Requiring a reasonable separation between the competing mass hypotheses, *e.g.*, by demanding that the  $\chi^2$  for the  $\pi^+\pi^-$  assignment exceeds alternatives by at least four, quickly reduces acceptances from 92% to 80% in the case of idealized particle tracking and  $dE/dx$  performance; accounting for reasonable degradation in the actual running detector lowers these estimates to unacceptable levels. The addition of

the DIRC/ATC system provides a clear margin of safety. The acceptance for  $\pi^+\pi^-$  remains at the 95% level given this and even more stringent background suppression requirements.

### 6.1.4 Summary of Requirements

A dedicated PID system is required both for  $B$ -flavor tagging and to identify exclusive final states. These two purposes place different demands on the system: tagging requires positive identification of kaons and leptons with minimum pion background, mostly in the low-momentum range; measurement of exclusive decays requires positive tagging of pions with minimum kaon misidentification, mostly in the high-momentum range, resulting in the requirement of highly efficient  $\pi/K$  separation over the full spectrum.

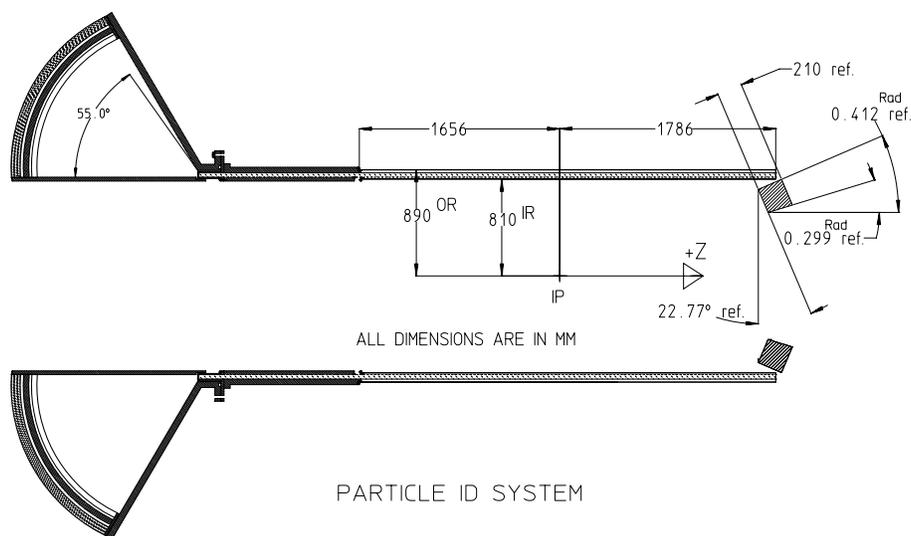
These requirements also place constraints on the detectors and material located inside the PID system, in order to benefit fully from the intrinsic performance capability. Key issues include the accuracy of track extrapolation to the DIRC and a low level of backscattering and  $\delta$ -ray backgrounds in the threshold ATC counter.

Likewise, the PID system must not significantly degrade the performance of the detectors located outside; both the physical volume occupied and material budget in radiation lengths are crucial in this respect. Of primary concern is the importance of calorimetry in measuring exclusive decays such as  $B^0 \rightarrow \pi^0\pi^0$  and  $B^0 \rightarrow \rho\pi$ , for which gamma detection and mass resolution are essential.

## 6.2 Particle Identification Overview

---

To meet the requirements described above, a system containing two types of Cherenkov detectors has been incorporated into  $BABAR$ . The barrel region ( $25.5^\circ < \theta < 147^\circ$ ) is covered by a DIRC [Rat92], which provides good performance over the whole momentum range while occupying only a thin radial region. An ATC covers the forward region ( $17.1^\circ < \theta < 23.6^\circ$ ) augmenting the kaon tagging coverage and providing  $\pi/K$  separation up to  $4.3 \text{ GeV}/c$ . The boundary between the systems was chosen to maximize the acceptance for particle identification within the constraints of magnet length and calorimeter position and angle. Both systems maintain low mass in order to minimize their effect on low-momentum photon calorimetry. The layout of the main components of the particle identification system is shown in Figure 6-2.

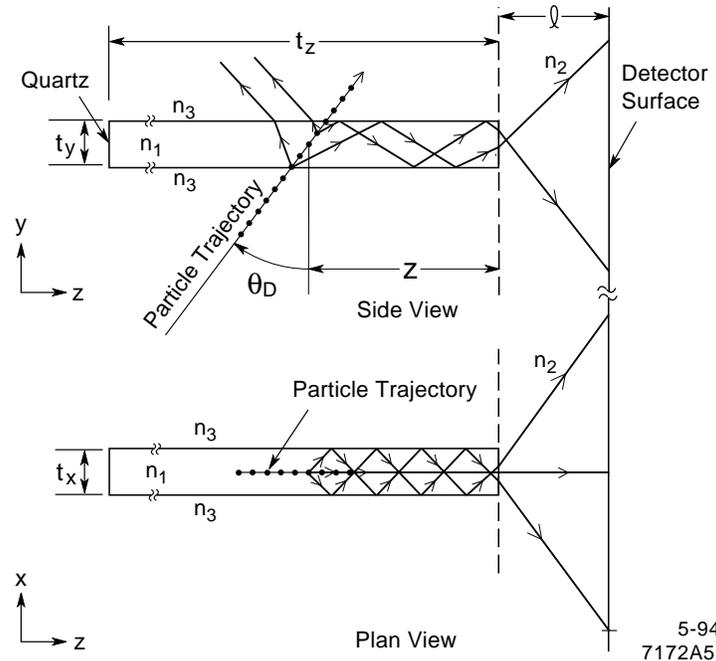


**Figure 6-2.** Elevation view of the PID system geometry showing the barrel, DIRC, and forward ATC systems.

### 6.2.1 The DIRC

A charged particle traversing a DIRC quartz bar with velocity  $\beta$  in a medium of refractive index  $n$  produces Cherenkov light if  $n\beta \geq 1$ . The DIRC Cherenkov radiators are 4.7 m-long rectangular quartz bars oriented parallel to the  $z$  axis of the detector. Through internal reflections, the Cherenkov light from the passage of a particle through the DIRC is carried to the ends of the bar as shown schematically in Figure 6-3. At the readout end, photons are detected outside the magnetic field region by conventional photomultiplier tubes (PMTs); a reflecting mirror at the other end of the bar returns forward-going light to the PMTs. For each track, the Cherenkov image expands from the end of the source bar across a standoff region filled with water to a toroidal surface of closely packed PMTs.

The DIRC uses as a radiator 156 quartz bars arranged in a 12-sided polygon around the beam line. This maximizes azimuthal coverage, simplifies construction, and minimizes edge effects. For sufficiently fast charged particles, some part of the Cherenkov radiation cone emitted by the particle ( $\theta_c(E) = \cos^{-1}[1/\beta n]$ , with  $n = 1.474$ ) is captured by internal reflection in the bar and transmitted to the photon detector array located at the backward end of the detector. (Forward-going light is first reflected from a mirror located on the end of the bar.) The high optical quality of the quartz preserves the angle of the emitted Cherenkov light. The measurement of this angle, in conjunction with knowing the track angle and momentum from the drift chamber, allows a determination of the particle velocity. An advantage of the DIRC for an asymmetric collider is that the high momentum tracks are boosted forward,



**Figure 6-3.** Schematic of a single radiator bar of a DIRC counter. The particle trajectory is shown as a connected line of dots; representative trajectories of Cherenkov photons are shown by lines with arrows.

which causes a much higher light yield than for particles at normal incidence. This is due to two effects: the longer path length in the quartz and a larger fraction of the produced light being internally reflected in the bar.

Each quartz bar is 1.75 cm thick, 3.5 cm wide, and 470 cm long, and is constructed by gluing together four shorter bars. The total radial thickness of the DIRC, including quartz thickness, sagitta from the polygonal shape, mechanical supports, and a 1 cm clearance on each side, is 10 cm. This material represents  $0.18X_0$  at normal incidence. An effort has been made to minimize both the radial thickness and the amount of material, since these increase the radius and cost of the barrel calorimeter while degrading its performance for soft photons.

The photon detector consists of 13,400 conventional 1.125-inch-diameter phototubes. They are organized in a close-packed array at a distance of 120 cm from the end of the radiator bars. The phototubes, together with modular bases, are located in a gas-tight volume as protection against helium leaks from the drift chamber.

The photo-detection surface approximates a partial cylindrical section in elevation and a toroid when viewed from the end. The standoff region has reflecting surfaces along the inner

and outer radii to reduce the number of phototubes required. The opening angle between these surfaces is still under review but is taken to be  $55^\circ$  at this time. To maintain good photon transmission for all track dip angles, the standoff region is filled with water. The water seal occurs at a quartz window that is glued to the end of the bar assemblies. Most of the mass of the standoff box structure is high permeability steel, which provides adequate magnetic shielding for the phototubes.

The DIRC radiators are supported on the central support cylinder, which is cantilevered from the strong support tube. Alternative designs that do not require a cantilevered support for the central support cylinder are also under consideration. Both the strong support tube and the standoff box are supported by a yoke directly to the barrel IFR. Since the DIRC bars penetrate the backward endcap return iron, the DIRC mechanical design interacts strongly with other structures at the backward end of the detector (Section 6.4.1).

A single-bar DIRC using 47 phototubes with an air standoff has demonstrated this detector concept (Section 6.4.5). A larger prototype with a realistic water standoff region will shortly be tested in a particle beam. It will be capable of imaging almost all Cherenkov patterns on a track-by-track basis. This prototype and other laboratory measurements will provide a detailed test of the optical quality of the quartz, glue joint, and mirrored surfaces, and establish the performance (single photoelectron response, quantum efficiency, and timing resolution) using approximately 500 PMTs. The prototype will also be used to test front-end electronics.

### 6.2.2 The ATC

The forward PID detector is a silica aerogel Cherenkov counter covering the polar angle range  $17.1^\circ < \theta < 23.6^\circ$ . Both kaon tagging and pion identification are achieved with two refractive indices,  $n \simeq 1.0065$  and  $n \simeq 1.055$ .

There are two key components in the ATC system: the aerogel radiator and the readout device. Candidates for both are commercially available from well-established suppliers. Several suppliers of aerogel with a polymer gel structure have been identified; material with the appropriate density range, transmittance, purity, and anticipated stability has been obtained in small quantities.

Aerogel with good optical quality is produced by Airglass, Lockheed, Jet Propulsion Laboratory (JPL), Lawrence Livermore National Laboratory (LLNL), Aerojet, and the Budker Institute of Nuclear Physics (BINP) at Novosibirsk. Caltech's Space Radiation Laboratory regularly uses large aerogel blocks from Airglass in balloon-based Cherenkov detectors. Lockheed's production facility can, in principle, produce enough high quality aerogel for our needs in 10 months [Men94]. The aerogel produced by JPL is hydrophobic, has good

optical transmittance [Tso94], and has been extensively tested in our beam measurements. A two-step process for aerogel production has been developed by LLNL and is being transferred to the commercial company Aerojet [Hru94].

In order to collect Cherenkov photons efficiently inside the solenoid, the photon detector must work in magnetic fields up to 1.5 T, have a large UV-sensitive area with high quantum efficiency, and have low noise. The baseline design uses bialkali fine mesh photomultiplier tubes (FM PMT) from Hamamatsu with a newly developed high quantum efficiency photocathode. Hamamatsu's FM PMTs with conventional bialkali photocathodes have been widely used in high-energy experiments [Gil88] for many years.

The average thickness in front of the endcap calorimeter is 9.5% of a radiation length. However, the present design places phototubes in the active region in front of the DIRC. The implications of this for DIRC and calorimeter performance need further study.

## 6.3 Projected Performance

---

### 6.3.1 Simulation Based on Prototype Results

The barrel DIRC and forward aerogel detectors are being studied using the simulation programs ASLUND and BBSIM. The former, based on fast parameterizations of the detector responses, allows studies requiring high statistics, and the GEANT-based BBSIM program provides a tool for simulating whole events more realistically. In the latter, other subdetectors in *BABAR* are defined with materials and geometries consistent with their present designs. This provides a fairly realistic source for secondary processes representing event-related backgrounds. Primaries are tracked as well as the secondary products they produce, such as  $\delta$ -rays, pair production, hadron scattering, backsplash, and decays. PID detector geometries are defined to a high level of detail, and simulation of response is tuned to the prototype test results (Section 6.4.5).

#### Barrel DIRC Detector

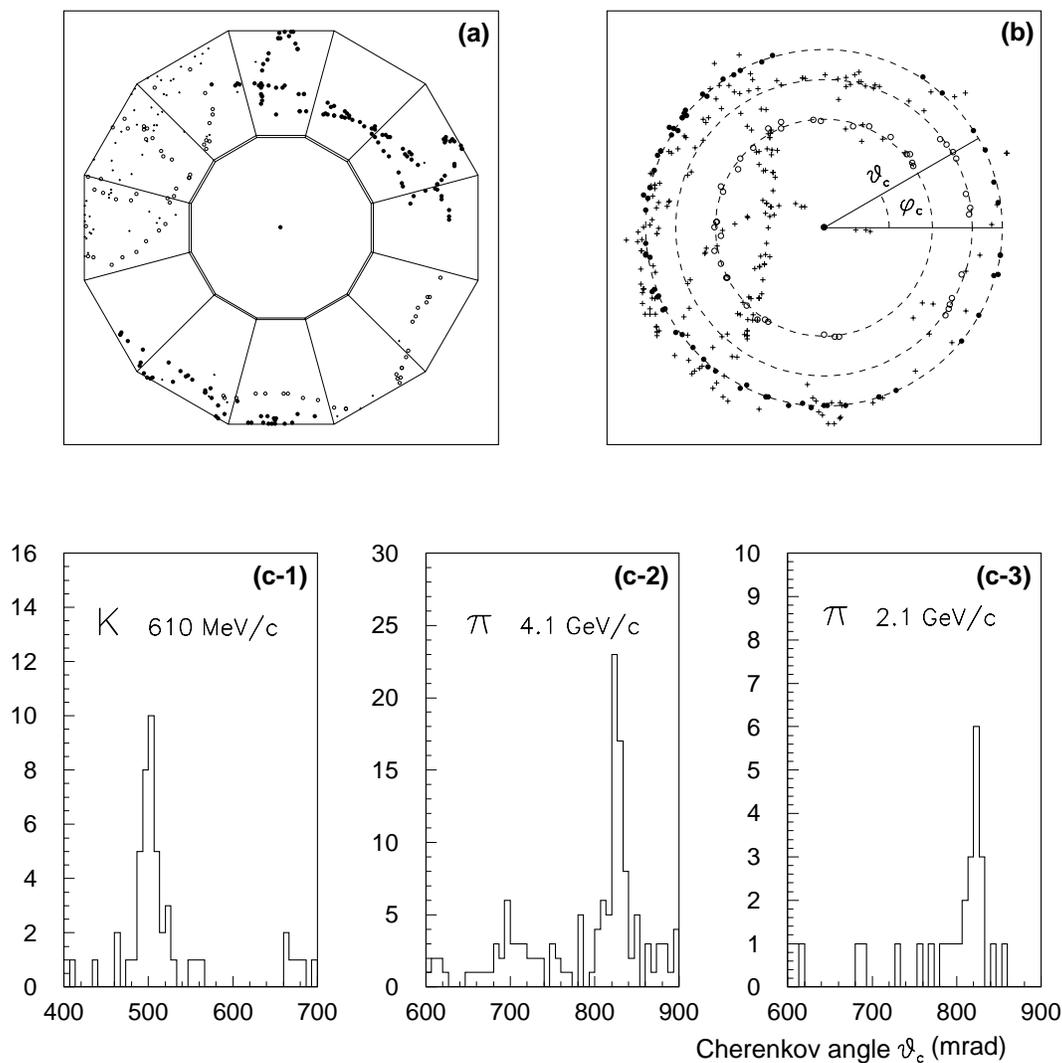
**ASLUND Simulation.** The behavior of the DIRC has been simulated in the ASLUND program; much of it is calculated analytically, and part is done approximately. The parameters for some of the approximations are adjusted to agree with the results of a stand-alone simulation of the DIRC. This program includes the propagation of photons to phototubes placed on a toroidal detector surface and the calculation of errors in the measured angles; it accounts for absorption in the quartz, in the water, and at surfaces; it calculates the  $\pi/K$

separation taking into account all ambiguities. Some aspects, such as the 12-sided geometry and the wavelength dependence of the absorption, are done in an approximate way. This simulation ignores background effects from  $\delta$ -rays and other tracks, so its estimates are somewhat optimistic. In the future, GEANT simulations that include background effects will be used to determine the parameters used in the ASLUND simulation.

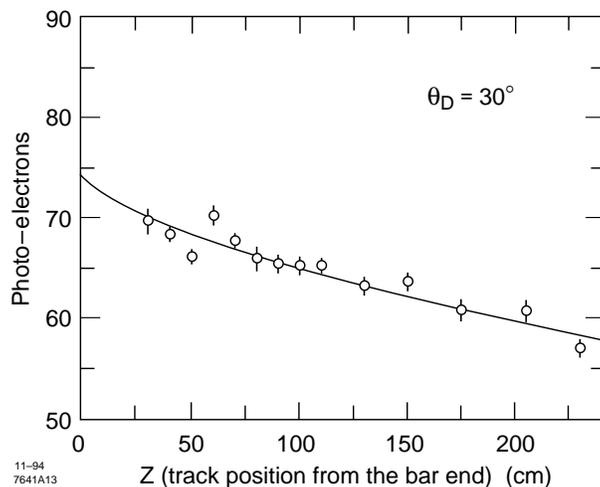
**BBSIM Simulation.** The barrel DIRC has been implemented in BBSIM with a realistic geometry. The detector description is parameterized in an easily modifiable form to create a description of the detector close to the engineering drawings. The active length of 178.6 cm in the forward direction provides an angular coverage to  $\cos\theta = 0.91$ , and 2 cm gaps between the modules result in a 4% loss in azimuth. The radiator represents  $0.18X_0$  at normal incidence. All physics parameters, such as absorption lengths, refractive indices, and overall photoelectron yield, are tuned to the prototype test results.

The generation and propagation of Cherenkov photons are fully simulated. After tracking is complete, all particles crossing the DIRC with momenta above Cherenkov threshold are processed through a customized simulation. Photons are generated with a Poisson distribution along the curving track trajectory through the radiator. Dependence of the refractive index  $n(E)$  of the radiator on photon energy is used to generate photons emitted at the Cherenkov angle  $\theta_c(E)$ . Photons produced in the DIRC bars are propagated until they reach the phototube detector surface. Light transport includes total internal reflection inside the quartz bars, diffraction at the quartz/water interface, and possible reflections from the mirrors at the forward end of the bar or within the standoff tank. Losses due to absorption in the bars or the water are included as functions of the photon energy. Photon absorption at the mirrors, scattering by surface imperfections, reflection of light at the quartz-water boundary (including the effects of photon polarization and angle of incidence), and the quantum efficiency of the PMT photocathode are also taken into account. Photons which strike the sensitive area of a PMT are recorded with digitized spatial and time information. A typical event with one  $B$  decaying to an arbitrary channel and the second  $B$  to  $\pi^+\pi^-$  has been projected onto the phototube detector plane in Figure 6-4(a). The time window for background in the simulation is 15 ns. Four charged tracks reach the DIRC:  $\pi(4.1)$ ,  $\pi(2.1)$ ,  $K(0.61)$ , and  $\pi(0.25)$ ; the momenta, in GeV/ $c$ , are indicated in parentheses.

Measurements made in cosmic ray Prototype-I tests (Section 6.4.5) have been used to tune the absorption coefficients and the reflection efficiencies at the bar surfaces and mirrors. The resulting comparison between predicted and measured numbers of photoelectrons at various positions and angles are in excellent agreement, as illustrated in Figure 6-5.



**Figure 6-4.** (a) PMT hits projected onto the x-y plane. The particles producing the hits have distinctive markers: Cherenkov images (conic sections) from the  $B \rightarrow \pi^+\pi^-$  decays are shown as solid circles, hits from the two other charged tracks as open circles, and from the secondary tracks as '+'. (b) PMT hits projected in Cherenkov angle space with the four tracks superimposed at the center. The hits from the two pions from the  $B$  decay overlap at the largest radius,  $\phi_c \approx 820$  mrad. The '+' marker in this case corresponds to assignment ambiguities. (c) Cherenkov angle projections for three of the tracks and the associated ambiguities.



**Figure 6-5.** Simulated photon yield (solid curve) compared to track position dependence of the photoelectron yield at a dip angle of  $\theta_D = 30^\circ$  as measured (circles) in Prototype-I using cosmic rays.

### Forward Aerogel Detector

The forward aerogel PID simulation in ASLUND is based on test beam results and the measured single photoelectron response of a FM PMT. The photoelectron yield of the phototube is described by a Poisson distribution with an excess noise factor of 2. The mean of the distribution is determined by the length of the trajectory through the aerogel block, where the full length at normal incidence corresponds to 10 photoelectrons. The output of the simulation is the pulse height in photoelectrons reflecting the single photoelectron pulse-height distribution of the FM PMT.

A simulation of the forward PID system has been developed in BBSIM. The geometrical description will correspond to the present design of the ATC with two layers and include a realistic simulation of the mechanical structure fabricated in carbon or fiberglass, as well as two rings of FM PMTs located on the inside and outside of the two aerogel layers. The geometry is fully parameterized for easy testing of options.

#### 6.3.2 Pattern Recognition in the DIRC

The complex event of Figure 6-4(a) illustrates that the Cartesian coordinates of the hits are not the easiest representation with which to do pattern recognition, *i.e.*, to assign hits to tracks. The most appropriate change of variable appears to be from Cartesian coordinates

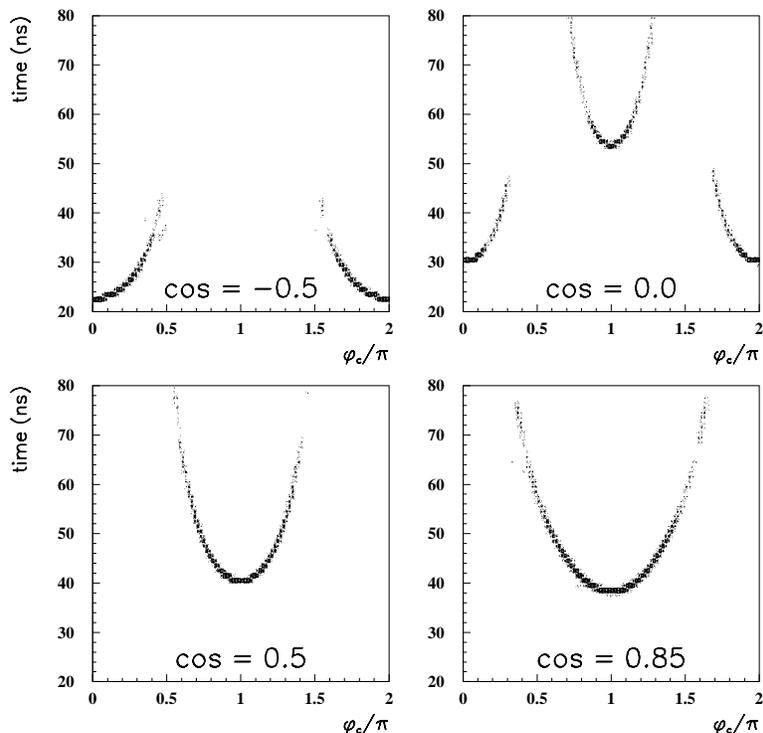
to Cherenkov angle  $\theta_c$ . In the representation in which the track segment defines origin and dip angle, the matching hits cluster at the radius corresponding to the particle species. This not only leads directly to a measurement of Cherenkov angle, but also obviates the need for sophisticated pattern recognition (see Figure 6-4(c)).

**Cherenkov Angle Measurement and Mass Hypothesis Testing.** Photons are presumed to originate from the centers of curved track segments. They are seen by a PMT as an image on the far side of the plane interface between the standoff tank and the quartz bar. A ray can be traced back from the PMT to the center of the quartz bar with a single refraction at the quartz-water interface. The angle between this ray and the track segment is the measured value of  $\theta_c$ . The azimuthal angle  $\phi_c$  is also determined in this manner. The optical distance (corrected for refractive index) from track to PMT is used to estimate the arrival time of the PMT signal.

Up to a 16-fold ambiguity exists in the determination of  $\theta_c$  because the number of bounces off the quartz walls and the mirrors is unknown. Only one solution is correct in each case, provided the photon is truly associated with the given track. Ambiguities can be eliminated when the corresponding photon path yields unphysical values for reflection, refraction, or Cherenkov emission angle. Timing information will also be used to reject solutions when the computed and measured time of arrival for a photon at the PMT are inconsistent. On average, only two solutions remain with a timing cut of  $\pm 7.5$  ns. Tighter timing cuts may be available, depending on the choice of phototube and readout system.

For each track, the distribution of the Cherenkov angles of all potentially associated photons shows a clear peak, corresponding to the parent particle type, above a background of wrong track-hit associations (Figure 6-4(c)). The peak position and the number of photoelectrons are then compared to that expected for each particle type and a mass assignment made on the basis of an hypothesis test. Consistency between angles and timing information will also be checked using the expected relationship between  $\theta_c$  and  $\phi_c$  for a hit and the propagation time for the photon to reach the PMT (Figure 6-6).

**Intrinsic Resolution.** The full DIRC pattern recognition algorithm has been applied to single pion samples to estimate the intrinsic resolution of the Cherenkov angle  $\theta_c$ . The single photoelectron  $\theta_c$  resolutions are approximately 9.5 mr, largely independent of track momentum and dip angle. The number of associated photoelectrons per track is given in Table 6-1. In this study, the availability of precise timing information was not taken into account. The intrinsic resolution of the DIRC detector is obtained assuming that the  $N_{PE}$  individual measurements of  $\theta_c$  are uncorrelated.



**Figure 6-6.** Measured time at the PMT face as a function of the photon Cherenkov azimuthal angle  $\phi_c$  for several track polar angles. The track momentum is  $4 \text{ GeV}/c$ , and the value  $\phi_c = \pi$  is defined for a photon whose initial direction projected onto the bar plane is towards the mirror.

Momentum ( $\text{GeV}/c$ )	Sine of Particle Dip Angle			
	-0.5	0	0.5	0.85
0.500	39	22	25	—
0.800	34	31	28	61
1.000	36	35	29	58
2.000	39	38	31	61
4.000	41	41	32	60

**Table 6-1.** Mean value of the number of detected associated photoelectrons as a function of pion momentum and sine of the dip angle. The effective packing fraction, *i.e.*, the ratio of active to total surface, is assumed to be 53% for this estimate.

Momentum (GeV/c)	Sine of Particle Dip Angle			
	-0.5	0	0.5	0.85
0.500	7.1	2.9	7.0	—
0.800	4.7	2.6	5.1	4.6
1.000	4.0	2.5	4.2	4.4
2.000	2.8	2.2	2.9	2.7
4.000	2.2	2.1	2.4	2.0

**Table 6-2.** Resolution on the  $\theta_c$  (mr) measurement per pion from BBSIM. Charged tracks are extrapolated from the vertex to the DIRC bars.

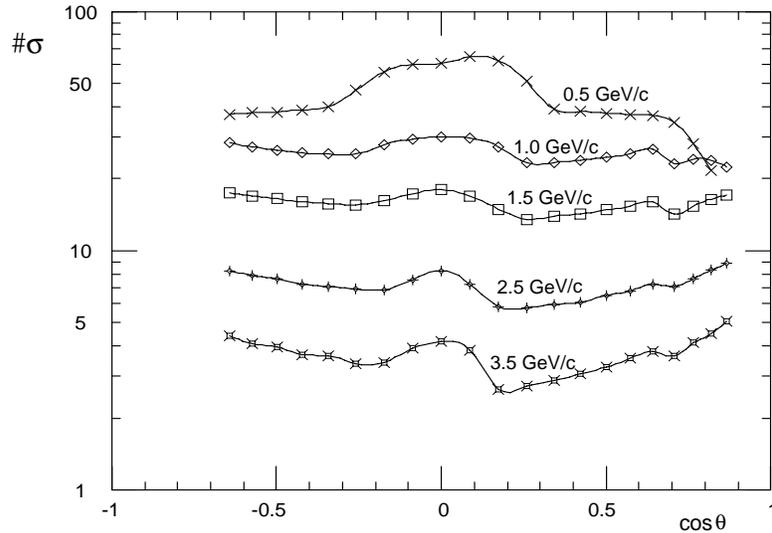
**Effective Resolution.** The DIRC performance is not dependent on intrinsic qualities alone. At low momentum, the resolution of the Cherenkov angle measurement is mainly determined by resolution of the track extrapolation, in addition to multiple scattering in the DIRC itself.

Using the reconstruction algorithm which is currently available in BBSIM, track parameters are obtained at the point of closest approach to the IP by a progressive fitting procedure using drift chamber and silicon vertex detector hits. Therefore, the procedure for swimming the track from the primary vertex to the DIRC includes angular errors coming from Coulomb scattering in the support tube and the inner wall of the drift chamber. As a result, the angular errors at the DIRC are significantly overestimated by BBSIM, especially in azimuth. Table 6-2 gives the resulting effective  $\theta_c$  resolution per track. A procedure with better accuracy would be to use a Kalman filter fit to extrapolate track parameters from the outer drift chamber layers to the DIRC. Track extrapolation errors would then include only the effect of the material between the last drift chamber hit and the quartz bar, which is equivalent to  $0.02X_0$  at normal incidence. From preliminary results, the quoted errors on the extrapolated angle should be reduced by 10 to 40% using such a technique. A constrained fit to the Cherenkov image may further improve upon the errors.

### 6.3.3 Particle Identification in the DIRC

The ultimate performance of the DIRC is limited by the Cherenkov angle difference between particle species. In a quartz radiator (with  $n = 1.474$ ), the difference in Cherenkov angle between a pion and a kaon is fairly large at low momenta but is as small as 6.5 mr at 4 GeV/c. The same difference occurs between a muon and a pion at 770 MeV/c.

Figure 6-7 shows the  $K/\pi$  separation as a function of track momentum and  $\cos\theta$ , obtained with the ASLUND simulation of DIRC performance. The separation is very clear in the



**Figure 6-7.** Predicted  $K/\pi$  separation performance of the DIRC, quoted in terms of the number of standard deviations, vs.  $\cos\theta$ , for different momenta.

momentum region that is most interesting for tagging kaons in the barrel, 0.5 to 1.5 GeV/c, which is why the kaon tagging efficiency is so close to the ideal case. The separation is always  $\gtrsim 4\sigma$  for particles within the kinematic limits for  $B$  decays, as can be seen in Figure 2-13.

Table 6-3 shows similar information on  $\pi/K$  separation, but for both ASLUND and BBSIM simulations of single track samples. This comparison is meant to illustrate the range of performance estimates obtained from current simulations. The ASLUND results in column (a) are known to be somewhat optimistic due to a lack of background, but on the other hand do not incorporate potential improvements brought about by a constrained fit to the Cherenkov image. The BBSIM results in column (b) include the effects from a realistic simulation of background hits but overestimate the contribution due to the track angle uncertainty because of the extrapolation of errors from the vertex and also do not incorporate a constrained fit. The actual  $\pi/K$  separation is probably somewhat better than the ASLUND result, but the comparison between the two simulations is an instructive measure of the reliability of predictions from our present understanding. It is clear that good  $\pi/K$  separation is achievable in most of the phase space of the asymmetric collider.

BBSIM as well as ASLUND studies show that a  $\mu/\pi$  separation with greater than  $2\sigma$  can be achieved by the DIRC at momenta below about 500 MeV/c. The separation power is somewhat spoiled by the track direction uncertainty due to multiple scattering. However, a separation of more than two standard deviations up to 500 MeV/c may be of great interest for tagging purposes and is likely to improve with more accurate track fits and extrapolations.

Momentum GeV/c	Sine of Particle Dip Angle							
	-0.5		0		0.5		0.85	
	a	b	a	b	a	b	a	b
0.500	38	68	60	167	38	69	—	—
0.800	35	35	40	63	33	32	24	36
1.000	27	26	30	42	25	25	23	24
2.000	11.0*	9.3*	12.2	11.7	9.7	8.9	9.7	9.6
4.000	2.9*	3.0*	3.1*	3.0*	2.3*	2.7*	3.9	3.3

**Table 6-3.**  $\pi/K$  separation in standard deviations as a function of track momentum and the sine of the dip angle from a simulation using (a) ASLUND and (b) BBSIM. The \* entries are outside the kinematic region populated by  $B$  decays.

### 6.3.4 Particle Identification in the ATC

Projected particle identification capabilities of the aerogel system are based on beam tests. The  $n = 1.0085$  aerogel counter's response to  $3.5 \text{ GeV}/c$  positive pions and protons (below Cherenkov threshold) was measured (Figure 6-8). If a cut is made at 55 ADC counts, corresponding to 1.5 photoelectrons, one obtains a 97% efficiency for pion detection with a 3% misidentification of below threshold particles. Similar separation between pion and kaons was also obtained at  $3.5 \text{ GeV}/c$  after subtraction of the background due to electronic noise, wrongly tagged pions, and kaon decays.

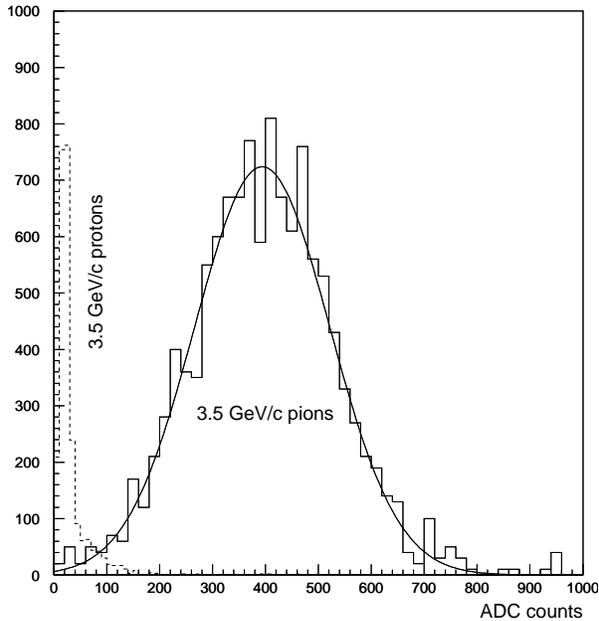
For the current design, the Monte Carlo simulation tuned with our test beam results shows that a  $\beta=1$  charged particle passing through the aerogel produces a signal of more than ten photoelectrons in both the low and high refraction index aerogel blocks. With a ten photoelectron signal, the aerogel PID system will give  $\pi/K$  separation up to  $\simeq 4.3 \text{ GeV}/c$  [Oya94], as shown in Figure 6-8.

### 6.3.5 Performance Requirements for Track Reconstruction

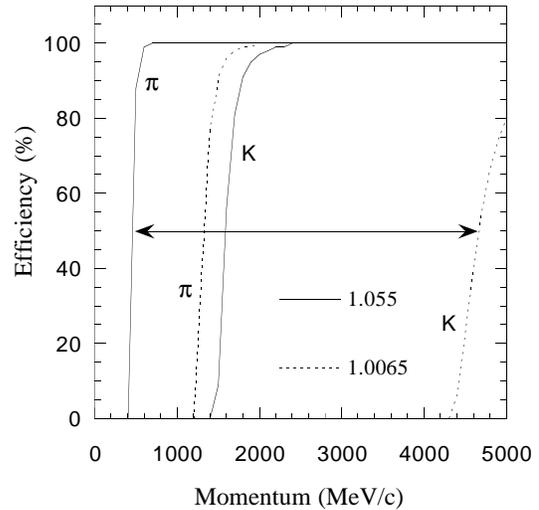
In a simple model, the measurement error of the Cherenkov angle per track in the DIRC can be written

$$\sigma_{\theta_c}^2 = \left( \frac{\sigma_{PE}}{\sqrt{N_{PE}}} \right)^2 + \sigma_{trk}^2 + \sigma_{ext}^2,$$

where  $\sigma_{PE}/\sqrt{N_{PE}}$  is the intrinsic resolution,  $\sigma_{trk}^2$  is the track angular error at the last drift chamber layer, and  $\sigma_{ext}^2$  is error on extrapolation to the DIRC due to multiple scattering in the outer wall of the drift chamber and in the DIRC supports. The study of errors due



**Figure 6-8.** Responses of a prototype 1.0085 aerogel counter to 3.5 GeV/c pions (solid line) and to the below threshold protons (dashed line) in a 1.3 T magnetic field.



**Figure 6-9.** Threshold curve for pions and kaons as a function of momentum for a two-index aerogel system.

to multiple scattering in the quartz bar itself, which contribute to  $\sigma_{PE}$  but are not fully uncorrelated, requires further work. In principle, these effects can be reduced by fitting the Cherenkov image.

Clearly, it would be desirable to maintain the track error contributions, due to the resolutions on the projected angles  $\theta$  and  $\phi$ , below the intrinsic resolution (*i.e.*, below 1.0 mr) at least for large momenta at which the  $\pi/K$  Cherenkov angle difference is small. In this context, it is important to minimize the amount of material between the drift chamber and the DIRC radiator. For example, increasing the amount of material to 8% of a radiation length at a  $60^\circ$  dip angle (*i.e.*, twice the present design) would lead to more than 1 mr uncertainty for a 4 GeV/c track due to multiple scattering. The present design of the drift chamber and DIRC support achieves this goal, as shown in Figure 5-5.

Background Source	Primary Track-Equivalent/Event
Event-Related Background	4.7
Beam-Gas EM (10×nom.bkg)	0.60
Radiative Bhabha	0.11
PMT Noise	0.020
Cosmic Rays	0.00016
Others	Negligible

**Table 6-4.** Backgrounds in primary track-equivalent units (33 photoelectrons per track) for a readout time of 50 ns.

### 6.3.6 Effects of Backgrounds on PID Detector Performance

#### Effects of Backgrounds on DIRC Response

Several background sources can affect the performance of the DIRC, such as event-related backgrounds, synchrotron radiation photons, beam-gas interactions, cosmic rays, phototube dark current BBSIM has been used to simulate the detector response. Each random background source has been integrated over an assumed 50 ns readout time, and the probability of a double hit in the same PMT in one event has been studied for readout times of 50 ns and 100 ns.

**Event Related Backgrounds.** Based on a BBSIM simulation, a typical  $\Upsilon(4S) \rightarrow B^0 \bar{B}^0$  event produces an average of 345 photoelectrons detected in the DIRC. Of these, 190 are attributed to an average of 5.7 primary tracks which lie within the DIRC acceptance and have momenta above Cherenkov threshold. Thus, each primary particle contributes 33 photoelectrons. The other hits are due to secondary tracks produced in the detector, mostly in the form of backsplash from the electromagnetic calorimeter, that have much softer momentum spectra than the primary tracks. Their random angular distribution gives an event-related background in the DIRC, which manifests itself in the angular coordinate system used for pattern recognition as a flat contribution under the Cherenkov peaks for primary particles. This background can be subtracted easily.

**Other Backgrounds.** The machine is a source of backgrounds from lost particles, radiative Bhabhas, and other backgrounds as discussed in Chapter 12. The largest source of machine-related background is the beam-gas electromagnetic term as shown in Table 6-4. Other backgrounds such as synchrotron radiation, cosmic rays, and PMT dark current have very small or negligible impact on the DIRC response.

**Multihit Probability.** On average, 175 PMTs are hit by the 190 photons associated with measurable tracks in a *BABAR* event, giving a double-hit probability of  $\sim 8\%$ . Including hits due to event-related background, the number of PMTs with a double hit increases to 16 and the probability becomes 9%. If beam-gas interactions are included at 10 times nominal background, the double-hit probability is 9.2% for a read-out time of 50 ns and 9.5% for a readout time of 100 ns.

### Effects of Backgrounds on ATC Response

Since each cell is read out by a PMT placed outside the Cherenkov radiator, charged particles passing through a PMT pose little problem for particle identification in the ATC. One background source is shower leakage from the calorimeter. For incident photons, this is generally not a problem, as no track will point to that cell; one can actually use the photons to determine the size of the effect. Backsplash from electromagnetic showers consists mostly of photons and electrons below Cherenkov threshold. The photons cause a problem only to the extent that they convert in the material of the PID system.

Another background is  $\delta$ -ray production, since a below-threshold kaon can be misidentified as a pion by knocking out an electron with momentum above Cherenkov threshold. Both an analytic calculation [Oya94, Gro92] and a detailed GEANT simulation [Shi94] have been used to study the effect. The results are in good agreement with measurements in beam tests. In the forward direction, the study shows that  $\sim 1\%$  of 4 GeV/ $c$  kaons may be misidentified as pions due to  $\delta$ -ray production in the drift chamber wall and backsplash from the calorimeter. With  $\sim 97\%$  detection efficiency for pions, the system still provides  $\sim 4\sigma$   $\pi/K$  separation at 4 GeV/ $c$ . Though  $\delta$ -ray production and backsplash cause about a 2–3% probability of misidentifying a low momentum ( $p < 1.2$  GeV/ $c$ ) kaon as a pion, they do not affect  $B \rightarrow \pi\pi$  and  $B \rightarrow \pi K$  separation since both daughters from these decays have  $p > 1.5$  GeV/ $c$ . In the worst case, such  $K \rightarrow \pi$  misidentification results in a  $B$  tagging inefficiency of less than 3%.

Synchrotron radiation photons are not expected to cause problems. Their energies lie between 4 and 100 keV, and cannot produce any above threshold (1 MeV for  $n = 1.06$ ) electrons. Their wavelengths are also not in the photocathode-sensitive region.

Beam-gas background should not be a problem either with a 100 ns integration time (the PMT's pulse width is a few tens of ns). The average number of hits per microsecond in the whole CsI calorimeter is about 0.8. Therefore, fewer than 0.08 hits per 100 ns are expected over the 144 aerogel cells, corresponding to  $\leq 0.0005$  hits per cell during one readout cycle.

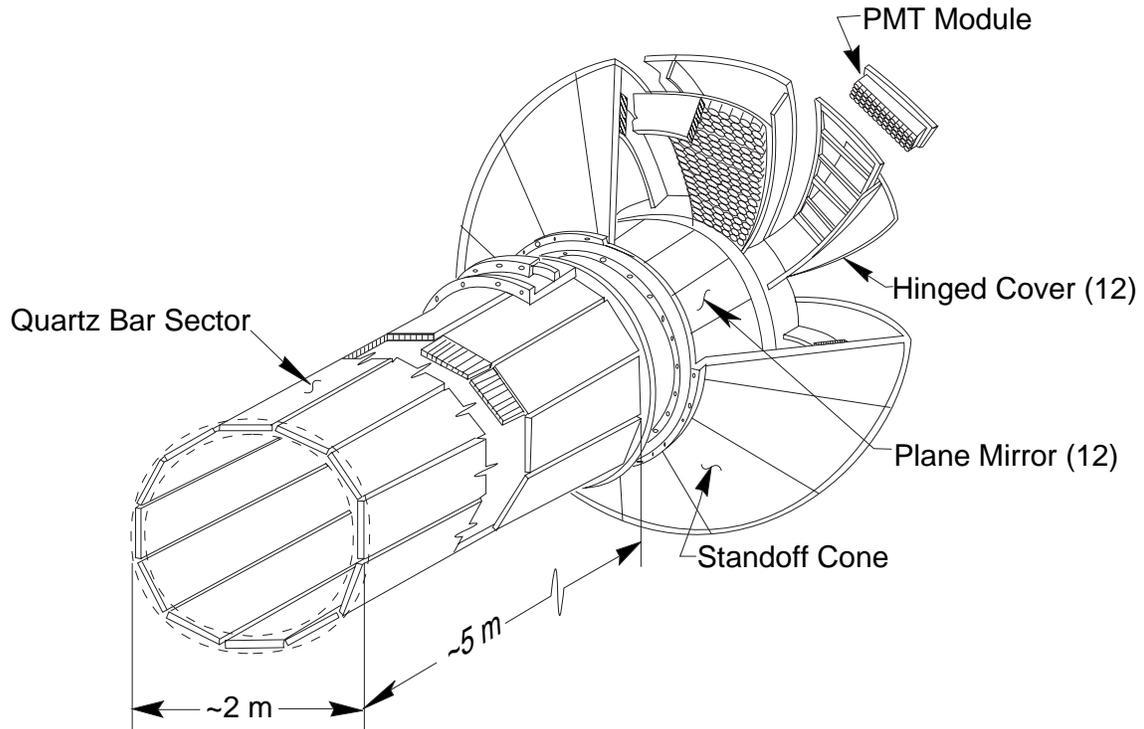


Figure 6-10. Mechanical elements of the DIRC.

## 6.4 The DIRC Detector

### 6.4.1 DIRC Mechanical Design

#### Principles of Operation

The major mechanical elements of the DIRC, shown schematically in Figures 6-10 and 6-11, include the following.

**Quartz Radiator Bars and Sectors.** The bars must have small attenuation loss for  $\lambda \geq 300$  nm. The surfaces must be accurately rectilinear and smooth to preserve light angle and intensity after many reflections. The choice of bar thickness and width involves trade-offs among light intensity, angular resolution, material ahead of the calorimeter, and cost. The bars are assembled into 12 separately covered and mounted sectors of 13 bars each. The bars of a sector are united at the water boundary by a glued quartz window which is sealed by an O-ring to the assembly flange.

**Figure 6-11.** *Schematic of the DIRC support system.*

**Support Tubes and Assembly Flange.** The central support tube is an aluminum honeycomb between cylindrical skins which holds the sectors within the sensitive volume of the detector. Its material thickness in radiation lengths will be small compared to the quartz bars. The central support tube will be cantilevered from the strong support tube. This steel tube, axially aligned with the central support tube, also holds the quartz sectors. It is mounted inside the backward IFR and is the central structural element of the DIRC, strong enough to support the beam line elements and the backward poletip. The steel tube is supported independently by the barrel IFR. The assembly flange is a short aluminum section rigidly connecting the strong support tube to the standoff region. Its functions are to reduce the magnetic flux coupled to the PMTs, to seal the sector windows, and to support and seal the inner and outer radius boundaries of the standoff volume.

**Standoff Region and Mirrors.** A water-filled volume between the windows at the ends of the quartz sectors and the PMTs permits the Cherenkov image to expand. The volume is defined by a cylindrical inner boundary and a conical outer boundary that nearly meet at the sector windows, plus a back toroidal surface covered by the PMTs. The walls are of high-quality magnetic steel to eliminate the effects of stray magnetic fields on the PMTs. Pure water is inexpensive, gives plenty of light transmission, and matches quartz fairly well in index of refraction and dispersion. There are 24 flat mirrors azimuthally aligned with the sectors. Half are mounted on the inner radius cylinder, the others on the cone at a  $55^\circ$  angle to the beam line. The choice of this angle is a compromise between the number of PMTs needed and ambiguities in the Cherenkov images and is still under review. To minimize mirror distortion from stress, a neutrally buoyant honeycomb structure is proposed.

**Phototubes.** The PMTs need good single photoelectron response and good quantum efficiency for  $\lambda \geq 300$  nm. The PMT faces are immersed directly in water, covering the entire back toroidal surface in a close-packed geometry. The module geometry for the PMTs has not been finalized. If individual tubes (or very small modules) are used, the toroidal steel PMT support plate (Figure 6-12), located at a radius somewhat larger than that of the PMTs, can accommodate the necessary access holes without sacrificing either mechanical structure or magnetic shielding at the PMT face. Steel cover plates that would limit access will be avoided.

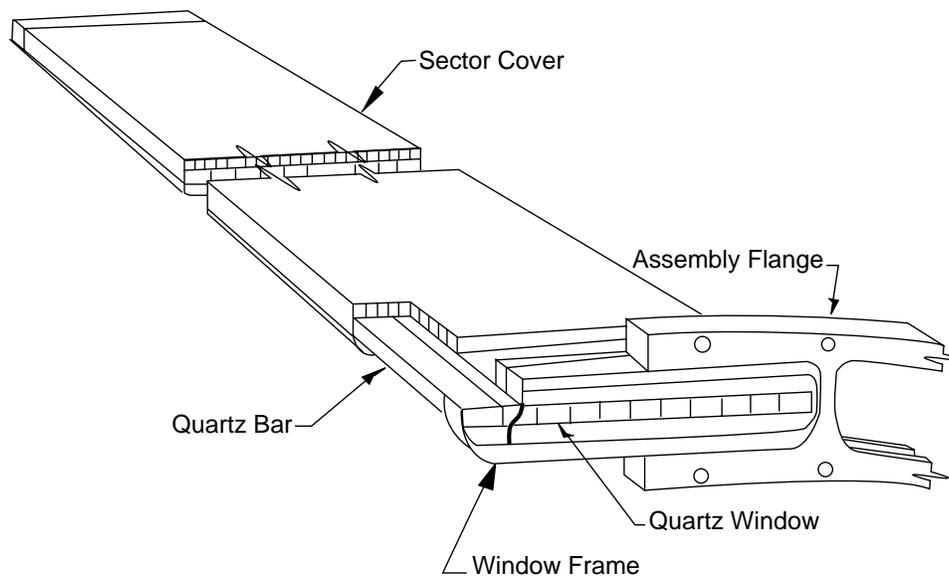
### Quartz

The quartz bars will be made from bulk natural fused quartz (*e.g.*, VITREOSIL 055). This material has an index of refraction of 1.474, good light transmission for  $\lambda \geq 300$  nm, is reasonably radiation hard, and is relatively free of bubbles and inclusions. This bulk material will be used to produce bar pieces of  $117.5 \times 3.5 \times 1.75$  cm<sup>3</sup>, which will be glued to form 156 full-length bars. These dimensions were chosen with cost, photoelectron yield, angular resolution, and minimization of material in front of the calorimeter in mind. The bars need to be extremely rectilinear to preserve the Cherenkov image, and well polished to preserve the light after many reflections. We have set up an optical laboratory for quality control and characterization of all the optical properties of the quartz bars. Cosmic ray tests, as described in Section 6.4.5, indicate that commercially produced bars meet the requirements.

The decision to glue short bars together to form a full-length radiator is driven by the excessive cost of tooling needed to cut and polish longer bars at standard optical industry shops. An epoxy with excellent optical properties and sufficient mechanical strength to join the short bars together has been selected and tested. Alignment fixtures and gluing jigs will be used in the process. Radiation hardness of the quartz, the epoxy, and its long term properties will be verified. Since the bars are read out in one direction, mirrors will be needed on the other end. We will use glass mirrors coated on the front surface with aluminum and silicon dioxide. For maximum reflection, the mirror will not be coupled optically to the bar end since many photons are internally reflected.

### Sectors and Covers

There are several reasons that the quartz bars are arranged in azimuthal subassemblies or sectors. First, for structural reasons, it is necessary to allow azimuthal gaps for ribs to couple the inside and outside radii of the DIRC support tubes. Second, a water-tight seal is needed at the junction with the standoff region. Third, the optics of the standoff region work best to maintain the resolution and minimize reconstruction ambiguities if the mirrors are flat and aligned with the sectors. These conditions tend to drive the sector multiplicity down.



**Figure 6-12.** *Schematic view of bars assembled into a mechanical and optical sector.*

On the other hand, they increase the radial width of the DIRC, not only because the sagitta is proportional to the square of the sector width, but also because the stiffness of the sector covers is a sensitive function of the width. For the parameters considered here, the radial budget is: sagitta, 3 cm; quartz and gaps, 2 cm; sector covers, 2 cm; and support tubes, 1 cm; for a total of 8 cm. The minimal azimuthal gap, between the inside corners of quartz bars in neighboring sectors, is 1 cm, which is sufficient to accommodate side-by-side O-rings. The average coverage loss between sectors is then approximately 2 cm, less than 5% of the total circumference.

The bars are assembled into a mechanical and optical unit or sector on a long optical table. At the end near the standoff region, each bar is glued to a common 3 mm-thick quartz window. The window area is somewhat larger than that of the bars so that light can emerge unimpeded at large angles from all parts of the sector. Because the bars are close together with sides parallel, care must be taken to avoid optical crosstalk. For this reason, the bars are separated at the support points by thin, small-area plastic wafers aluminized on both sides. To seal the sector at the standoff region, an aluminum window frame is sealed by O-rings; first to the back of the window and then to the assembly flange as the final part of sector installation. Some details are shown schematically in Figure 6-12.

In our present design, each bar is supported at 0.6 m intervals, 0.3 m from the glue joints. In this geometry, each joint is normally at a point of maximal (but very small) stress, but it is subjected to zero additional stress due to a misalignment of one of the nearest supports.

For the part of the sector inside the sensitive region of the detector,  $\sim 3$  m in length, the periodic bar supports are mounted to top and bottom plates of low density honeycomb, 1 cm thick by 42 cm wide, which are coupled together by thin strips on the sides. The periodic supports are strengthened by a transverse stiffener, which is a few centimeters wide axially. The bars do not touch the honeycomb or its stiffeners directly but are spaced below by an aluminized 1 mm-diameter fiber and clamped on top by a threaded screw, aluminized on the tip. Transverse support is supplied by stressing the spaced coupling strips on the sides, so that the cover supports the sector independent of orientation. The part of the sector outside the sensitive region of the detector,  $\sim 2$  m in length, has top and bottom cover plates of solid aluminum.

To protect the glue joints during extreme (*e.g.*, seismic) axial motion, each bar is compressed at its own weight ( $< 100$  kg/sector) by spring loading the front surface mirror at the non-readout end of each sector. The force is returned to the window by the tension in the thin skin (0.1 mm stainless steel) outer layer of the sector cover. Because the covered sectors can easily be made gas tight, the internal atmosphere is readily controlled by a nitrogen purge.

### Support Tubes and Assembly Flange

Figures 6-10 and 6-11 illustrate the concepts discussed in this section. Very different structures are needed to support the sectors in the sensitive (central support tube) and insensitive (strong support tube) parts of the detector.

The central support tube must be as thin as feasible. The problem is made somewhat difficult by the long distance from the major structural elements of the DIRC. To avoid coupling this relatively fragile assembly simultaneously to structurally independent ends of the detector, one of the proposed designs uses a cantilever. The major structural elements are aluminum: 1 mm-thick skins at inner and outer radii; 3 mm-thick circumferential ribs slotted for the sectors and spaced every 0.6 m; and 2 mm-thick axial ribs every  $30^\circ$  linking the circumferential ribs. The average material thickness of the support tube is  $\leq 0.04X_0$ . Carbon-fiber support structures will be studied in order to further reduce the radiation-length thickness. Studies using a finite element analysis show that this design will carry the required load of 17 kg per meter per sector. The structure is predicted to deflect a few millimeters when loaded, with buckling the nearest failure mode. The load limit can be increased by distributing the forces at the corners of the ribs. Inside the central region, but beyond the maximum angle of the barrel calorimeter, the central support tube is a heavy aluminum structure like the steel strong support tube to which it is joined. This design reduces the required length of the low density cantilever.

The strong support tube, located inside the IFR, requires steel for strength and for returning magnetic flux from the poletip. To accommodate whatever solution is adopted for beam line

support, the strong support tube is designed to support, in addition to the DIRC, the backward beam line elements and the poletip, which feels an axial force of 300 tons at 1.5 T. Given septa of 1 cm width every 30° to couple the inner and outer skins, careful analysis shows that the tube is strong enough for this worst case. Unlike the central support tube, the strong support tube is not limited to a radial width of 8 cm. At the outside radius, the strong support tube is attached by a flange to the yoke supported by the barrel IFR. The inner radius matches that of the assembly flange and the cylinder of the standoff region, which are discussed below. The preferred fabrication method for the strong support tube is to cut slots for the sectors into annular plates of convenient thickness. These are pinned for alignment and held together by as many as 24 axial bolts (12 each on the inside and outside radii).

The 15 cm-deep aluminum assembly flange is rigidly supported from the back of the strong support tube and has two major functions. First, it breaks the magnetic circuit between the standoff region and the poletip, IFR, and strong support tube; and second, it provides the seal and supports for the sectors and standoff regions. Based on the magnetic simulation done to date, it is known that the field at the PMTs is acceptable with the gap; whether it is essential has still to be determined. If it is not, this assembly flange could be eliminated, saving 15 cm in the length of the quartz bars. The strong support tube could handle the major mechanical functions, which would save a significant amount of radial space.

### Standoff Region

The standoff volume is filled with approximately six tonne of water, contained by an inner radius cylinder, an outer radius cone at 55°, and a toroidal back PMT support plate. They are of high quality magnetic steel; simulations using 5 cm-thick steel (permeability  $\sim 10^4$ ) gave  $B \leq 0.05$  T at the PMTs with a central field of 1.5 T. Tests on conventional 29 mm-diameter PMTs showed no effect at this field, independent of orientation. It is expected that further optimization through simulation will lead to reduced steel thickness in some regions. In any case, it is essential that the joints between the PMT plate, the cone, and the cylinder provide good flux transfer. With 5 cm-thick walls, the total weight is  $\sim 15$  tonne. The water reservoir is mounted below the standoff region, so the forces do not change significantly when water is transferred. We plan to use de-ionized reverse-osmosis water, which has an attenuation length of 10 m at  $\lambda = 300$  nm. The water will recirculate slowly through a filter and under a UV lamp, and the attenuation will be monitored offline. In addition, there will be an emergency pressure relief valve, an expansion reservoir for temperature changes, and a dam to channel water in case of leaks.

**Figure 6-13.** *Schematic of a possible PMT module assembly design.*

## **Mirrors**

The mirrors in the standoff region are needed to minimize the required number of PMTs; they have the same azimuthal symmetry as the sectors. The  $0^\circ$  dip angle mirror is essential because half the light from each sector emerges from the quartz bar in a direction which is pointed towards a smaller radius. The mirrors could be aligned from within the standoff volume, although it would be convenient to use sealed adjustments for outside control. The mirrors on the conical surface are presently assumed to be placed at a  $55^\circ$  angle to reduce the number of PMTs.

The mirror design is complicated by the need for immersion in water for a long period of time. If a front surface coating is used, it must be covered for protection. A likely possibility is to use aluminum with a coating of either silicon dioxide or a series of dielectric layers. An alternative is to coat the back of a sheet of thin UV-transmitting quartz, which is then sealed against water. Mirror “coupons” could be immersed in the water circuit and removed periodically to measure the degradation of reflectivity with time. The mirrors will be made neutrally buoyant by using a hexcell support structure to minimize the distortions.

## **PMTs and Modules**

Approximately 13,400 PMTs of 29 mm diameter will be needed to cover the toroidal surface in a close-packed array. The packing fraction depends upon design details but will be  $\sim 87\%$ . Typically, the PMTs are only sensitive inside a diameter of 25 mm. To recover this lost region, each PMT will carry a reflecting cone with a depth of 4 mm and radial width 2 mm

to collect most of the light hitting the cone at dip angles less than or equal to  $30^\circ$ . The individual PMTs will be mounted into modules that can be joined together to provide a leak-tight interface between the water of the standoff tank and the phototube array. Several different schemes are now under evaluation. One of these is shown in Figure 6-13.

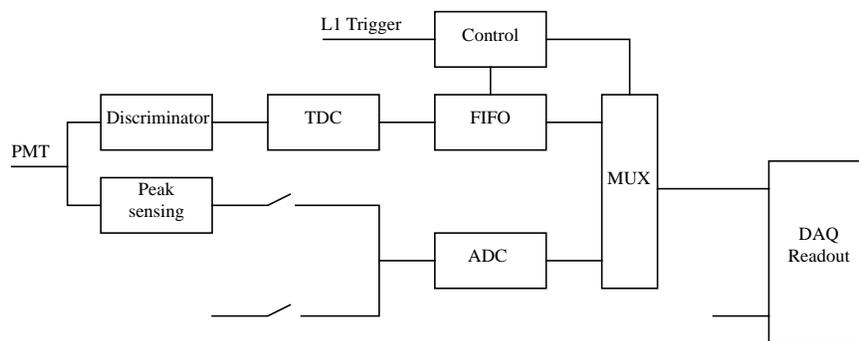
## 6.4.2 Photodetectors and Readouts

### DIRC Readout

In principle, the Cherenkov photon arrival time contains information about the Cherenkov angle. This information becomes competitive with the position information obtained from the array when the time resolution approaches 100 ps for a single photoelectron hit. If cost were not an issue, modern instrumentation and very good PMTs could provide time resolution at this level. However, it is substantially less expensive to determine the angle information from the position on the PMT array plane and not attempt such precise timing. Timing information, however, is still a powerful tool for distinguishing signal from background. For each track, the dominant background comes from the event itself, which occurs over a time window of about 50 ns. Thus, to gain a significant factor in signal to noise on these backgrounds requires a timing resolution of a few nanoseconds. The readout electronics will provide a timing resolution on the order of 1 ns. Although the pulse height carries no additional information for a single hit, it is useful for calibration purposes. We propose implementing a system which offers the option of measuring the pulse height during dedicated calibration runs.

Figure 6-14 is a block diagram of the readout chain for a single PMT. During normal physics running, the time is recorded in a FIFO whenever the signal at the output of the PMT crosses a preset threshold. The FIFO is large enough to store the maximum data for the duration of the Level 1 trigger latency,  $9.5 \mu\text{s}$ . Upon receiving a Level 1 trigger signal, the data in the FIFO falling within a  $\pm 0.5 \mu\text{s}$  window around the trigger are sent to the DAQ readout module where they are buffered before being shipped over the FDDI network to the Level 3 farm for further processing. The baseline design calls for the front-end electronics to be located near the PMTs, with 64 channels per card and one fiber-optic link to the DAQ readout modules for each group of 960 channels. The VME-based DAQ readout modules, located in the electronics house, will receive four optical fibers each. The baseline design of this system and a description of continuing research are presented in Chapter 10.

**Photodetectors and Bases.** The criteria of greatest importance for the DIRC photosensors are high quantum efficiency, maximal coverage of the photon image plane, adequate gain, good timing resolution ( $\sim 1$  ns), low noise, and low cost. The performance of candidate



**Figure 6-14.** Block diagram of the readout chain for a single PMT. The FIFO is capable of storing data for the duration of the Level 1 trigger latency.

PMTs from three leading manufacturers has been tested: Hamamatsu R268, Philips XP2982, and EMI 9124A. These PMTs are all of similar design: 29 mm in diameter with bialkali photocathodes. As a result of these investigations, the Hamamatsu R268 has been chosen for the forthcoming prototype with 500 tubes. These tubes are, therefore, included in the DIRC baseline design. The market will be researched again before the final choice of PMTs is made.

Both timing and pulse height information were obtained for the three tube types using a prompt Cherenkov light signal. This was produced by a  $\beta$  particle emitted from a  $^{106}\text{Ru}$  source traversing a thin quartz plate. The source was sufficiently weak that nearly all detected signals were the result of single photoelectrons. The absolute gain was derived from the measured pulse height distribution and from the known ADC charge sensitivity. The time distributions were obtained with the same system. A very fast TDC start signal was generated by stopping the  $\beta$  particle in a scintillator. Gaussian fits to these time distributions were used to extract their rms time spread. The noise (dark current) and the photocathode response, were measured. (The latter is defined to be the mean PMT pulse height divided by the gain when the photocathode is uniformly illuminated.) These phototube characteristics, which measure photon detection ability, have been compared for the three tubes using the same light source operated at the same intensity.

The results of the measurements are collected in Table 6-5. The voltages given in the table are the highest voltages at which the PMTs may be safely operated according to criteria which are believed conservative. The dark current can vary from tube to tube and is shown only to indicate that it is small enough to be of no concern. Except for the time resolution, the performance results favor the R268 tube. The improvement in time resolution that would come from using the Philips tube does not compensate for the loss in response.

Tube	HV (kV)	Gain ( $10^8$ )	Noise Rate (kHz)	Pulse Height Resol. $\sigma$ /PH(%)	Time Resol. (ns)	Relative Response
R268	1.6	1.1	2.6	42	1.2	1.6
XP2982	2.0	1.2	3.4	38	0.55	1.0
9124A	2.0	5.7	2.2	20	1.3	1.3

**Table 6-5.** Comparison of the three tested PMTs.

The details of both the electrical and mechanical design of the PMT base depend on the particular device chosen and the mounting scheme. However, neither of these choices poses particularly difficult problems. A base composed of two printed circuit boards has been developed for the current choice of PMT (Hamamatsu R268) in the 500 PMT DIRC Prototype-II. An alternate scheme [Hub92] using a voltage multiplier is also being studied.

The HV feed to 24 individual bases will be provided by a single PC board, serviced by one or two HV supply channels. The same board may also be used to route the signal lines onto a single 24-pair ribbon cable (as with the prototype) or to signal processing electronics on the same board. This board can be mounted directly behind the PMT support.

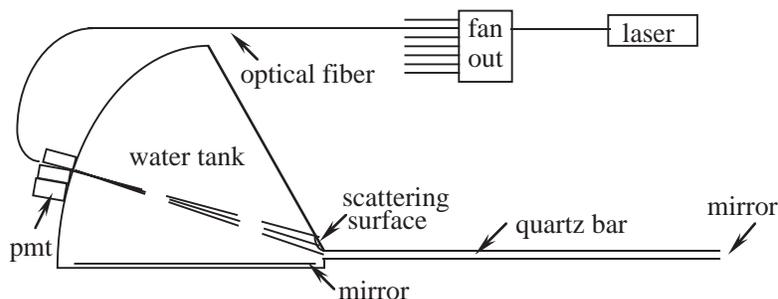
The PMTs are located in the standoff box. Their backs and bases will be easily accessible for repair and/or replacement.

### 6.4.3 Laser Flasher Monitoring and Calibration System

The DIRC laser flasher system will be used to measure the gain and relative quantum efficiencies of each tube and to monitor the optical performance of the quartz bars, glue joints, mirrors, and windows. The conceptual design for this system uses one or more pulsed nitrogen lasers as the light source. The wavelength, 337 nm, is characteristic of Cherenkov photons. The pulse duration is shorter than the timing resolution of the DIRC. The number of lasers will be determined by their repetition rate and the desired trigger rate.

The laser light will be distributed by optical fibers. Two methods are under study:

- Locate the fiber between photomultiplier tubes (see Figure 6-15). The light will be directed towards a scattering surface located above the bar window.
- Locate the fibers at the junction of the quartz bar exit and the standoff box with one fiber directed along the bar towards its far side and the other directed into the water through a scattering surface. (The use of LEDs or Xenon flashers in place of lasers is also under study.)



**Figure 6-15.** *Conceptual design of the DIRC laser flasher system. The 337 nm light is distributed by four fibers per sector and reflected from a diffuse scattering surface, providing a reasonably uniform distribution of single photons to all phototubes.*

Each of the fibers of a sector will be a different length to provide a timing difference that can be used to distinguish pulses from different fibers.

The scattering surfaces will give a reasonably uniform illumination of all phototubes. The light level will be adjusted so that the probability per pulse of any particular PMT being struck by a photon is small. This ensures that essentially all hits will be due to single photoelectrons.

The gain of each PMT will be derived from the single photoelectron pulse height spectrum; the number of hits is a measure of the stability of the relative quantum efficiency.

Light will travel the length of the bar and reflect off the far mirror before striking the PMT. The ratio of the number of these hits to direct hits is proportional to the transmission of the bar, glue joints, window, and the reflectivity of the mirror. The short pulse duration of the light source is needed so that these hits can be distinguished on the basis of timing.

It is anticipated that a flasher calibration will take approximately 10 minutes and be performed once per week. The laser trigger will be used also as a level 1 detector trigger, with a partition for DIRC data only.

#### 6.4.4 Integration Issues

##### Offline DIRC Assembly and Test

A set of twelve aluminum boxes will be built to simulate the sector boxes containing the quartz bars. These will be dimensionally correct and loaded to the weight of a real sector. These boxes will be preassembled into the central support tube where clearance checks and modifications can, if needed, be made. These boxes will also serve to align the light section

of the central support tube with the heavy aluminum section. At this point, the central support tube is rigidly held off the floor by a spider and framework. The central support tube and the strong support tube are mated into an accurate cylinder 5 m long.

The sector boxes will all be installed at the top of the support tube; the support tube will be rotated for each new sector box. We begin by removing a dummy aluminum box and bringing up a sector on an adjustable assembly strongback. The box is precisely aligned with the slot and moves in on rollers pulled by a line. Compressed air flow around the sides reduces friction. After installation, the sector will be sealed to the assembly flange and the next insertion will take place.

The standoff region, which weighs approximately 15 tonne, could be preassembled and moved by crane. A major advantage of this approach is that it would allow an inspection of the PMTs and electronics before moving onto the beam line. At this stage, the weight of the device is supported by the floor. The 24 mirrors in the standoff are installed and aligned, after which the PMT support plate is bolted to the cone. The PMTs and electronics are installed and a dummy assembly flange mounted. The water reservoir and other devices can be attached for calibration, monitoring, and possibly cosmic ray tests.

### Online DIRC Installation and Test

The support tube and sector assembly is inserted into the detector, supported by a yoke directly from the barrel IFR. After surveying, shims are used to align the central support tube with the drift chamber. The standoff structure is bolted directly to the assembly flange and sealed. The procedures described above are repeated for inspection and testing and upon successful completion, the beam line elements, poletip, and utilities are placed inside the DIRC tunnel.

### Cabling and Access

The signals from the 13,400 DIRC phototubes are carried by ribbon wax or twisted pair cables to the front-end electronics, located in nearby crates. They are multiplexed onto a small number of optical fibers for transmission to the electronics hut. The cable plant is dominated by the  $\sim 900$  HV cables. There are, in addition, 144 optical fibers for calibration and additional cables for monitoring and charge-injection calibration. The cable routing is straightforward and presents no interference to other systems.

All active components of the DIRC are located in the standoff box and can be accessed by removing the electronics covers—no other detector subsystems must be moved. Access is expected to be required infrequently, since the failure of a small number of isolated channels will have no impact on the physics performance. No access is required to the radiators.

## 6.4.5 Research and Development Program

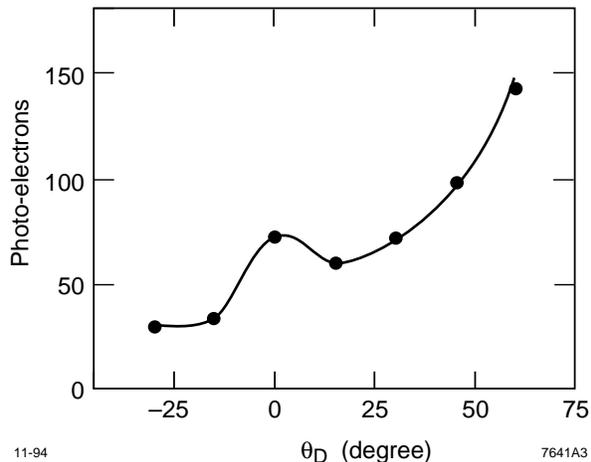
### Prototype-I: Proof of Principle

The goal of the conceptual prototype (Prototype I) was to demonstrate the basic proof of principle of the DIRC [Ast94]. Because no detector like the DIRC had been built before, this prototype needed to address many fundamental issues, and be easily modified to respond to new issues that arose during the studies. The detector was placed in a cosmic ray beam hardened to 1 GeV/ $c$  by an iron stack. A series of scintillation counters served as the trigger, and a set of straw tube chambers was used for tracking in the angular resolution studies. The quartz bar used initially was  $120 \times 4.7 \times 1.7 \text{ cm}^3$ . It was supported at a few contact points inside a light-tight box. Later, two bars were glued together to make a bar 240 cm in length. A mirror was attached to the non-readout end in order to make studies at an effective distance of 4 m from the readout end in the reverse configuration. For the photoelectron yield and attenuation length studies, a Burle 8850 51 mm PMT was glued to the end of the bar. For the angular resolution measurements, an array of 47 EMI 9124A 29 mm PMTs was placed at varying distances from the end of the bar with air filling the standoff region.

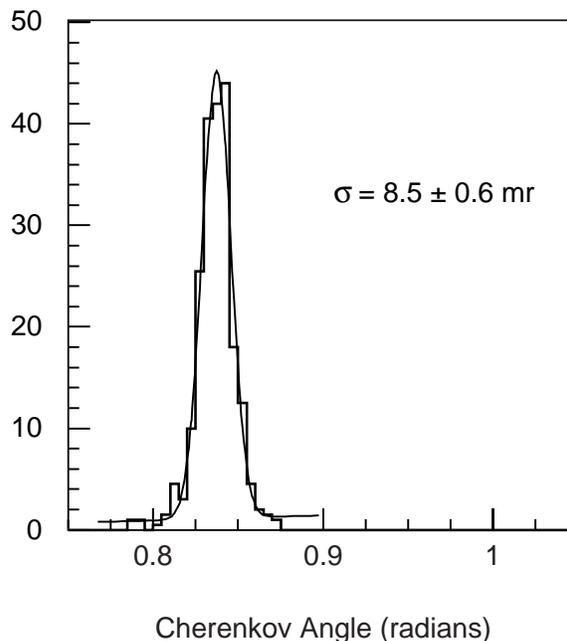
The photoelectron yield was studied as a function of dip angle and position along the bar. The results of one of these measurements are shown in Figure 6-16, for a position in the middle of a bar. These results are consistent with the Monte Carlo prediction, shown as a solid line. The behavior at large dip angles demonstrates a major advantage of the DIRC at an asymmetric collider in that the highest momentum tracks, which are forward-going at large dip angles, produce large numbers of photoelectrons. The measured yield with the PMT directly glued to the bar does not represent the number expected in the real DIRC at the same distance and angle, since there are additional losses from the packing of the array tubes, water in the standoff region, and transition effects.

There are also length-dependent losses from attenuation in the quartz, both from light absorption in the bulk material, and from losses at the surface. To study these effects, a series of measurements was made with both the 120 and 240 cm bars, in both forward and mirror reflected configurations at a constant dip angle. As expected, when characterized by a simple exponential, the measured attenuation lengths become somewhat longer as the path lengths increase. However, with this readout system, they are approximately 10% per meter, consistent with a bulk transmission length of greater than 10 m, and an internal reflection coefficient of more than 0.9995. This performance, which is completely adequate for the DIRC needs of *BABAR*, is shown in Figure 6-5.

The single photoelectron angular resolution was studied for a variety of dip angles, positions of the beam along the bar, and standoff distances in air. Major contributors to the resolution are source and detector size, and chromatic dispersion. Lesser contributions come from the momentum spread in the beam, multiple scattering, tracking, and reconstruction errors



**Figure 6-16.** Observed photoelectron yield from a 1.7 cm-thick DIRC bar at a track position 60 cm from the photodetector end as a function of the track dip angle  $\theta_D$ . The solid line is a Monte Carlo simulation; the statistical errors are smaller than the data points. There is a scale error of 3% due to calibration uncertainty.



**Figure 6-17.** Single Cherenkov photon angle distribution at  $\theta_D = 30^\circ$ , a standoff distance of 90 cm, and a track position 205 cm from the end of a 240 cm-long bar. The line shows a fit to the data obtained with a Gaussian plus polynomial form.

in the telescope. A typical result is shown in Figure 6-17. The measurement resolution of  $8.5 \pm 0.6$  mr is in good agreement with Monte Carlo simulations for the Prototype-I configuration, but differs from the resolution which will be obtained with the actual DIRC.

In addition to the cosmic ray tests, a test in a proton beam was performed using a quartz bar identical to the one described above. A PMT was attached directly at each end of the bar with a silicone optical joint, and both amplitude and timing information were recorded. Proton beams of  $\beta = 0.8$  and  $0.95$  traversed the quartz bar at various positions and angles along the quartz length. The number of photoelectrons detected by the PMT, towards which the emitted Cherenkov light propagates by internal reflection, was measured to be 65 for  $\beta = 0.95$  at a  $30^\circ$  angle of incidence. The observed absorption was only 3% per meter. The fact that the number of photoelectrons was slightly smaller and the absorption length longer than in the cosmic test is thought to be due to the different absorption properties of the optical couplers between the bar and PMT.

A significant amount of light was also detected by the opposite PMT. The direction and speed of propagation of this light has been measured, and two components have been clearly identified. The first corresponds to reflection from the other PMT of the normal Cherenkov light. In this setup, its fraction amounted to 8% of the emitted light.

The second component cannot be produced by Cherenkov emission of the incident particle since it travels in the wrong direction, with a low longitudinal propagation speed that does not depend on the incident track angle, and is seen even below the Cherenkov propagation threshold. This second component, produced isotropically in the bar, was present in approximately 30% of the events and corresponded to about four photoelectrons (with an exponentially falling amplitude spectrum). In *BABAR*, this component will lead to an extra background of 2 to 3 photoelectrons per track, if a simple extrapolation holds, and therefore will not degrade the DIRC performance.

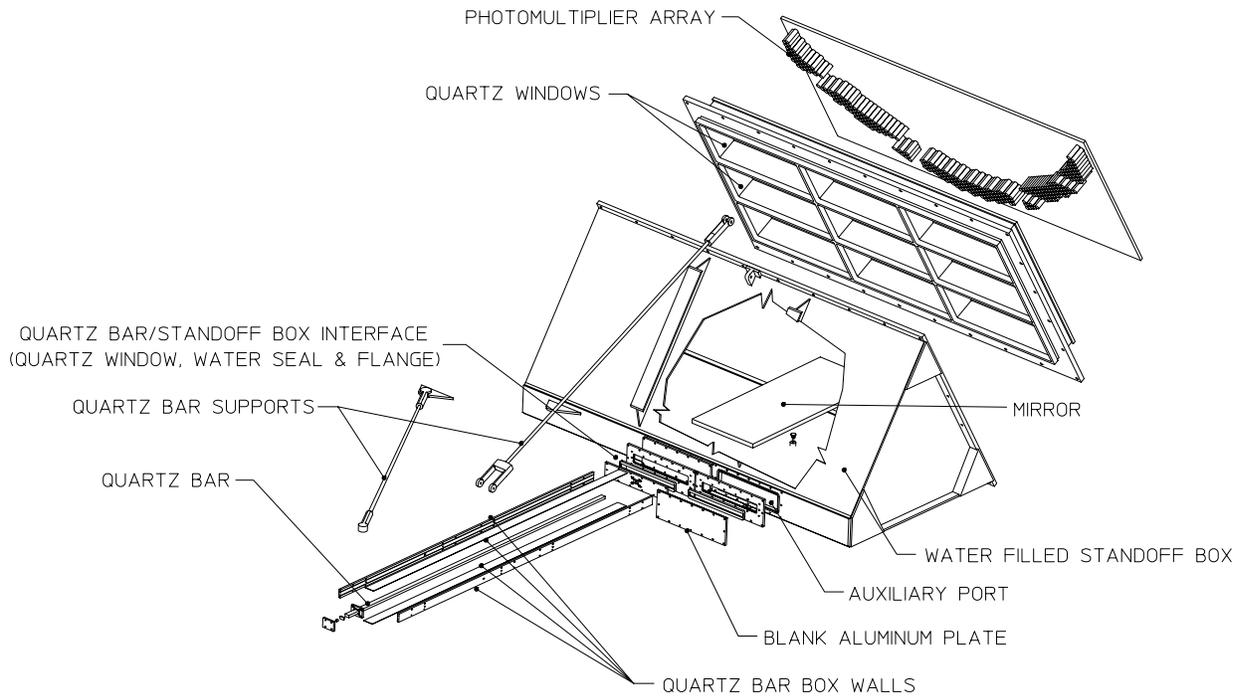
The conceptual prototype test program, which is now completed, was essential as the proof-of-principle test, demonstrating that the DIRC can produce enough photoelectrons with sufficient angular resolution for all track angles and positions to provide good  $\pi/K$  separation for *BABAR*.

### Prototype-II: Large-Scale Engineering Prototype

After the proof of principle, the next important step is a proof of viability for a large, complete detector system. To that end construction is nearing completion on a large scale prototype (Prototype-II), which approximates a full size section of the eventual configuration. It provides a realistic model to test many of the principal engineering and fundamental system performance issues. In particular, it will allow the observation of a nearly complete Cherenkov image on a track-by-track basis and the direct measurement of the resolution, in mass or  $\Delta\beta$ , also on a track-by-track basis. Moreover, because of its size, the prototype provides a realistic assessment of the impact of the mechanical structure on the performance of the final device. Examples of components which may affect performance are glue joints between bars, quartz windows, water-filled standoff box, mirrors, PMTs, supports, and interface.

A cut-away isometric view of Prototype-II is shown in Figure 6-18. The prototype consists of two  $1.67 \times 4.6 \times 120$  cm quartz bars glued together to form a composite bar 240 cm long. This bar is coupled to a water-filled standoff region with one mirror inside and 500 PMTs (29 mm in diameter) on the back plane of the standoff region for photodetection. The phototubes are read out into standard modular electronics.

The dimensions of the standoff box are 2.4 m wide by 1.2 m on each edge, with a  $60^\circ$  angle between each of the large surfaces. There are two possible ports at which the quartz bar can be connected to the standoff box, one centered and one off center. The latter port allows



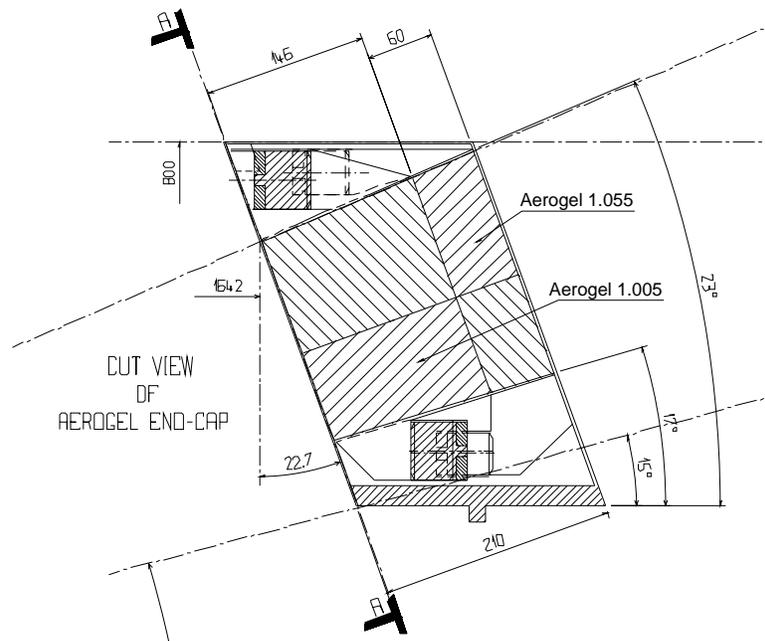
**Figure 6-18.** *Isometric view of the Prototype-II detector.*

the study of the extreme end regions of the Cherenkov conic images outside the acceptance of the back plane when the quartz bar is in the center port. The quartz bar is connected to the port by an interface consisting of a quartz window sealed in a watertight frame bolted to the standoff box. The port and interface are the same size as in the final configuration, which will accommodate a complete sector of quartz bars.

For simplicity, the back surface is a plane, instead of a curved toroidal shape, in the final configuration. Because of the limited number of PMTs for the prototype, they must be specially configured on the back plane for each beam dip angle to subtend the specific band of light corresponding to the Cherenkov image produced at that angle. To provide maximum flexibility for this, the PMTs are separated from the water volume by a 2 cm-thick quartz window barrier, made up of nine window panes as shown.

Construction of Prototype-II at LBL is scheduled for completion at the end of February, 1995. It will then be installed in a hardened cosmic ray telescope at SLAC for tests from March through early May. It will then be shipped to CERN for extended tests starting in early June of 1995, with beams of pions, kaons, and protons below about 4 GeV/c.

Future upgrades of this prototype will include lengthening the quartz bar to 480 cm, attaching bar mirrors, adding more bars side by side (up to 12) to make up a quartz bar sector as in



**Figure 6-19.** *R-θ* view of the ATC detector

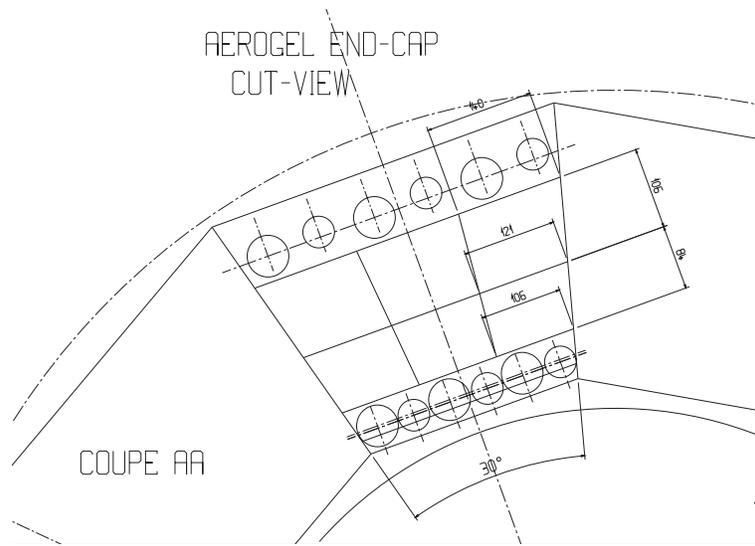
the final configuration, adding more PMTs, testing various calibration schemes and different electronics. The versatility of this prototype construction also permits future tests of different PMT mounting and interface schemes with the water volume, such as having the PMT front faces immersed in the water.

## 6.5 The ATC Detector

### 6.5.1 ATC Mechanical Design

The forward PID system consists of 144 blocks of aerogel arranged in two rings that are two layers deep, as shown in Figures 6-19 and 6-20. The front layer, 146 mm long, has a refractive index of  $\sim 1.006$ , and the back layer, 60 mm long, has an index of  $\sim 1.055$ . The detector will be built in two C-shaped pieces, as shown in Figure 6-21, as is the endcap electromagnetic calorimeter, to which it is attached.

Each half endcap is composed of:



**Figure 6-20.** *R- $\phi$  view of one ATC module.*

- Six modules, subtending  $30^\circ$  in  $\phi$ , each housing six aerogel blocks of two indices. The block dimensions are shown in Figure 6-20. The modules, shown in Figure 6-21, are made of a light composite material covered by a water-resistant PTFE Teflon with a reflectance of about 95% over the region 300–700 nm. PTFE teflon has been used extensively in our beam tests. Any diffuse reflector used for the aerogel wrapping must be of low density and high reflectivity. It must remain stable over a long period. Other candidates are Millipore paper and Kodak paint. All are hydrophobic and commercially available. Reflectivities of these materials are under study.
- Each aerogel block is read out by a Hamamatsu fine mesh photomultiplier through a diffusing box located on the top and bottom of the module, as shown in Figure 6-19. Phototubes of 52 and 35 mm diameter are used for the 1.0065 and 1.055 blocks, respectively. The effect of the material in the phototube on the performance of the DIRC and calorimeter is being studied.
- The modules are glued on a half cone skin made of 2 mm-thick composite material. This front skin supports kapton high voltage and signal cables.
- The entire front section, equipped with readout and aerogel blocks, is mounted inside a conical backing structure arranged as a half-shell in azimuth. This support provides rigidity for each half endcap, supports cables and services, prevents light leaks between cells, and provides the mechanical connection to the supporting calorimeter endcap.

Generated heat is removed by blowing nitrogen at the top and bottom of the detector around all tubes. The nitrogen flow also protects the aerogel against humidity.

The size of the aerogel blocks is a compromise between minimizing the number of readout channels and maximizing the number of photoelectrons. A low-density block of transverse size  $10 \times 10$  cm has been demonstrated in a test beam to have acceptable performance. A block of this size at 170 cm from the interaction point corresponds to a  $3.2^\circ$  angle. Only 2.1% of kaons overlap with another charged particle within this angular range. This fraction drops to 0.4% for a  $1.6^\circ$  cone, corresponding to the case in which two particles fall in the same block, whatever the impact point. Larger block sizes also lead to poor and nonuniform photon collection. Simulations show that going from a  $10 \times 10$  cm<sup>2</sup> to a  $10 \times 20$  cm<sup>2</sup> block reduces the average number of photoelectrons by a factor of 3. For these reasons, a  $10 \times 10$  cm transverse block dimension has been adopted for the baseline design.

### 6.5.2 Photodetectors and Readouts

In order to collect Cherenkov photons efficiently inside the solenoid, the photon detector must work in magnetic fields of up to 1.5 T, have a large UV sensitive area with high quantum efficiency, and have low noise. The baseline design uses 19-stage fine mesh photomultiplier tubes (FM PMT) from Hamamatsu with a newly developed high quantum efficiency bialkali photocathode. Hamamatsu's FM PMTs with a conventional bialkali photocathode have been widely used in high-energy experiments [Gil88]. The newly developed FM PMT gives much higher quantum efficiency (see Figure 6-22). Using the new FM PMT,  $\sim 1.7$  times as many Cherenkov photons were observed from a prototype aerogel counter in our beam test at the KEK PS facility than were observed with an older version.

Our recent studies using a spectrometer magnet at SLAC [Gea94] confirm that a conventional 16-stage FM PMT has a gain of  $4 \times 10^5$  at 1.5 T, if tilted at  $45^\circ$  with respect to the field [Kic93, Suz86, Ham94]. At  $\sim 5^\circ$  the relative gain drops by factor of  $\sim 200$  for 1.35 T, as shown in Figure 6-23. Without the magnetic field, the FM PMT has a gain on the order of  $10^7$ , an intrinsic excess noise factor (ENF) of  $\sim 2$  due to the inefficiency of the mesh dynode structure, and negligible electronic noise. When operated in a magnetic field, the FM PMT's gain drops substantially, its effective ENF increases due to the influence of the field on electron trajectories in the tubes, and the overall electronic noise increases because an amplifier is needed to increase the signal. As a consequence, even though the FM PMT gives the highest gain at a  $45^\circ$  angle to the magnetic field, an optimal operation angle must be determined for single photon response. A series of tests have been performed which indicate that a 2" FM PMT tilted  $\sim 5^\circ$  with respect to the field gives a reasonable photoelectron yield along with an optimal pedestal-to-signal separation for the aerogel counter readout.

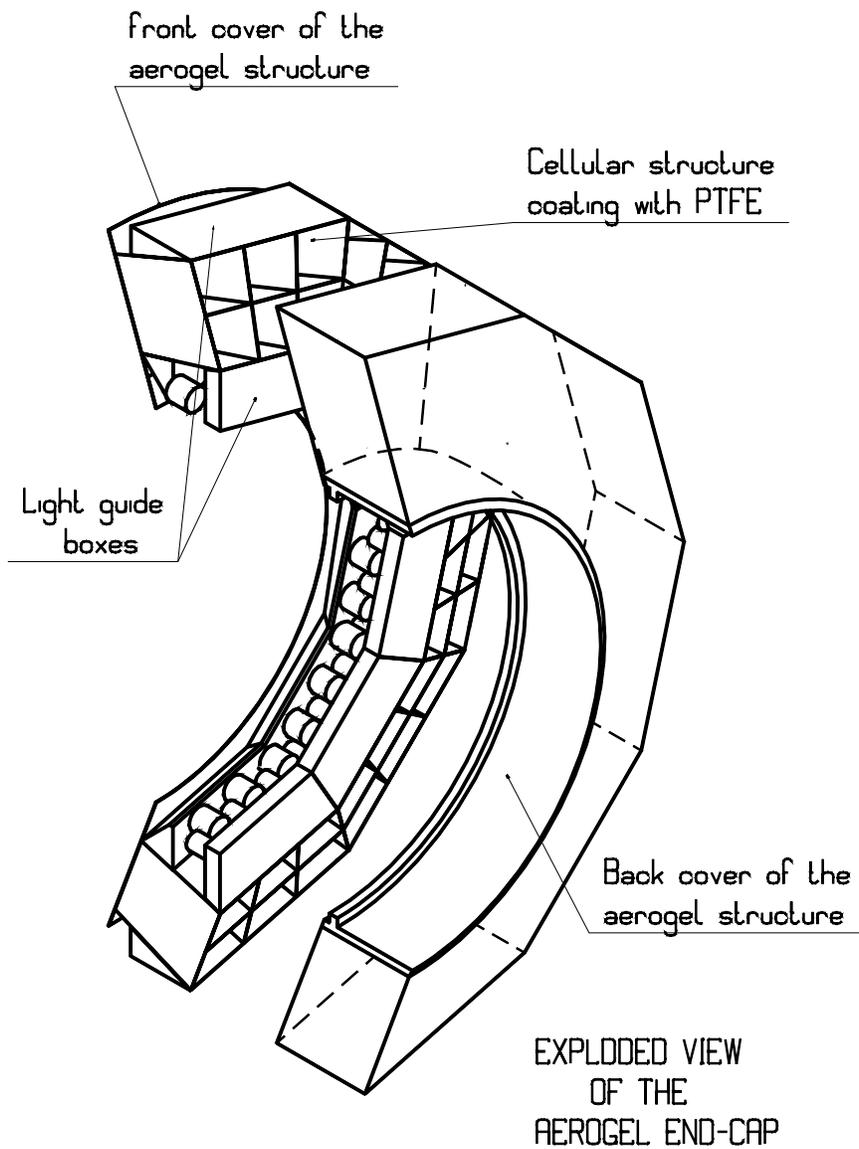
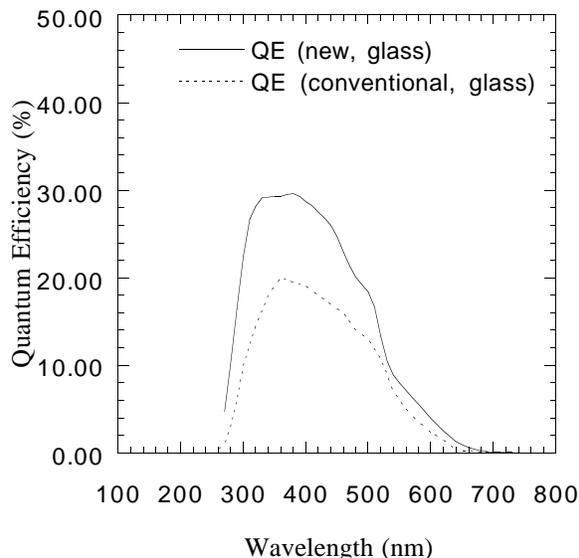
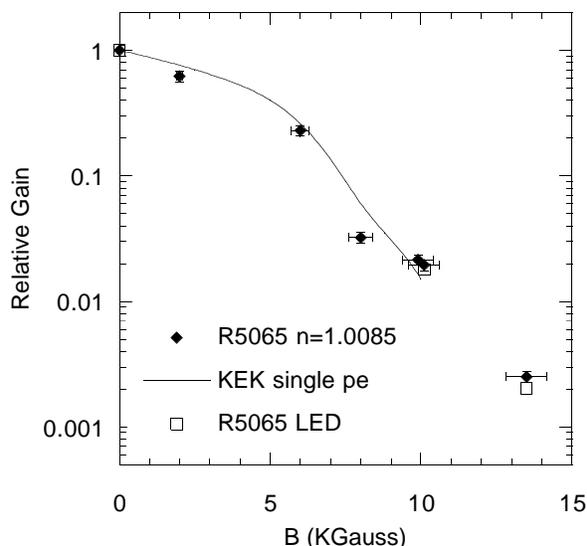


Figure 6-21. Three-dimensional view of one sector of the forward ATC.



**Figure 6-22.** Comparison of the quantum efficiency of the newly developed fine mesh PMT (solid line) vs. a conventional PMT (dashed line). Both PMTs have glass windows.

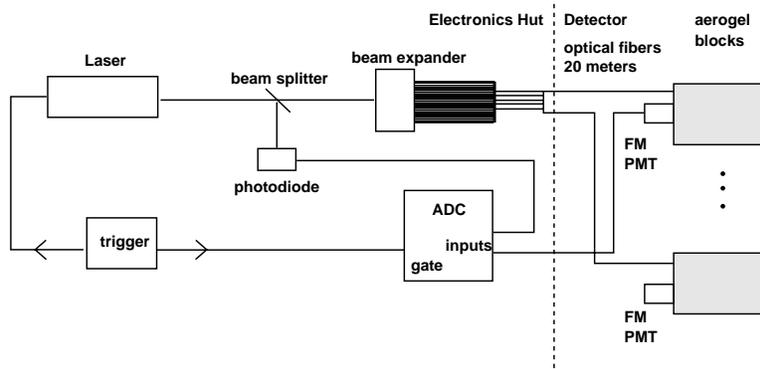


**Figure 6-23.** Measured relative gain of a fine mesh PMT operated in magnetic fields up to 1.35 T.

More compact readout devices, such as hybrid photodiodes (HPDs) and microchannel plates (MCP PMTs) are also being considered. The characteristics of HPDs and MCP PMTs, and their performances in magnetic fields, have been extensively studied by the aerogel PID group and by others at KEK, DESY [Kic93, Suz86], and CERN [Des94a]. An outline of their characteristics is provided here. More detailed descriptions can be found in Section 6.5.5.

The HPD has the advantage of being a  $ENF \sim 1$  photon detector, while its overall effective electronic noise is on the order of a few photoelectrons. Its low gain,  $\sim 3 \times 10^3$ , requires the use of a charge-sensitive amplifier. However, the overall effective electronic noise has been reduced to less than or equal to 1 photoelectron in a recent test. The improvement in the signal-to-noise ratio of the HPD comes from using a photodiode with a thicker depletion layer, applying more voltage in the vacuum gap, and reducing the thickness of the inactive layer in front of the photodiode [Des94b]. Improving the HPD's performance has been one of our major R&D efforts.

Newly developed GaAsP micro-channel plate photomultiplier tubes (MCP PMT) from In-tevac have quantum efficiencies of  $\sim 10\%$  at 350 nm and  $\sim 40\%$  at 500–600 nm [Edg92], providing a better average collection efficiency for Cherenkov light than FM PMTs. The



**Figure 6-24.** *Conceptual design of the ATC laser flasher system.*

MCP PMT has a gain on the order of  $10^6$ , an ENF measured to be  $\sim 1.4$ , and is immune to magnetic fields [Kic93]. The Novosibirsk Institute of Semiconductor has successfully produced an InGaAsP MCP PMT with a quantum efficiency of  $\sim 50\%$  at 600 nm [Alp94].

### 6.5.3 Monitoring and Calibration System

A laser-driven optical fiber will be attached to each aerogel block to monitor the stability of the aerogel and photodetector on a short term basis. The system will use a pulsed nitrogen laser, which has a 337 nm wavelength. The laser beam is optically split into two beams: one is monitored by a photodiode to determine the stability of the laser pulse, and the second is expanded and coupled into an optical fiber which is attached to the back of each aerogel block (Figure 6-24). The long-term calibration of the system will be performed using Bhabha events, dimuons, and pions from  $K_S^0$  decays.

### 6.5.4 Integration Issues

#### ATC Assembly, Installation and Test

After positioning and gluing the six modules of each endcap segment, the cells are filled with the aerogel blocks, and photodetectors are mounted inside the diffusing boxes. Bases and pre-amplifiers are also mounted during this operation. The detector is closed by attaching the back cover of the aerogel structure. Each half endcap is assembled and transported independently. Special tools have to be developed for positioning and gluing the modules.

Each ATC endcap is mounted on the front face of the endcap calorimeter, and aligned with respect to it, before installing the two detectors.

### Cabling and Access

The front-end electronics for the ATC are situated on the detector, and for any access to this part, the ATC and the electromagnetic endcap calorimeter have to be removed. After dismantling the back cover of the aerogel structure, access to all electronics and aerogel blocks is possible.

Cables from the detector to the DAQ, which is located outside the *BABAR* detector, run between the beam pipe and the inner radius of the ATC.

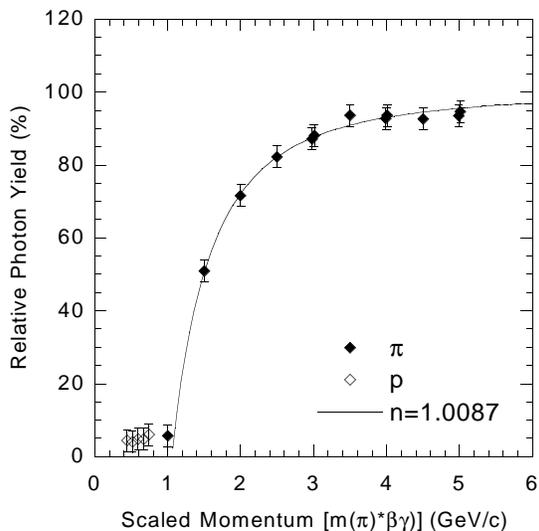
## 6.5.5 Research and Development Program

### Prototype Beam Test Results

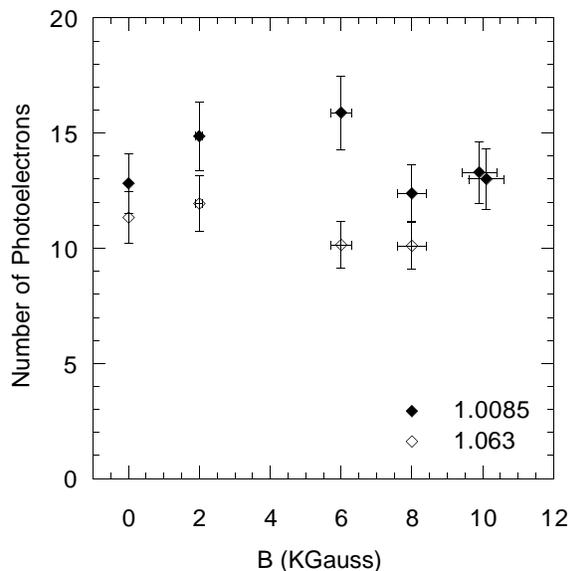
Proof-of-principle requirements have been met through several beam tests of prototype aerogel counters read out with fine mesh photomultipliers in a strong magnetic field at CERN and KEK. More than 10 photoelectrons have been observed with both  $n = 1.0085$  and  $n = 1.063$  aerogel counters. Threshold curves have been mapped out, and  $4\sigma$  separation between below threshold particles and above threshold particles has been obtained. In magnetic fields up to 1.0 T, 13 photoelectrons were measured when a 5 GeV/ $c$  charged pion traversed a masked 3" fine mesh PMT with a UV window and a sensitive diameter of 44 mm.

Both  $n = 1.0085$  and  $n = 1.063$  aerogel counters with sizes of  $10 \times 10 \times 10$  cm and  $10 \times 10 \times 4$  cm, respectively, were wrapped with eight layers of 1.5 mil PTFE covered with a thin layer of aluminized mylar and placed into the beam. Minimum ionizing particles were selected by four scintillation trigger counters and identified by two CO<sub>2</sub> gas Cherenkov counters operated at about 3 atm. A lead brick with a thickness of 2 cm, corresponding to  $3.6X_0$ , was placed behind the aerogel counter, between the fourth trigger counter and a veto counter, to veto electrons as well as to simulate backsplash from the calorimeter.

Light yields of 9.2 photoelectrons were obtained by reading out the  $n = 1.0085$  counter with a 2" RCA Quantacon PMT (C31000M); 10.2 photoelectrons were obtained from the  $n = 1.063$  aerogel counter by a masked RCA Quantacon PMT with effective sensitive diameter of 27 mm. The momentum-dependence of the light yield demonstrates that both aerogel counters behaved as threshold Cherenkov devices, as shown in Figure 6-25. The measured refractive indices  $1.0087 \pm 0.004$  and  $1.061 \pm 0.005$  were in good agreement with the refractive indices provided by the aerogel manufacturer, 1.0085 and 1.063, respectively.



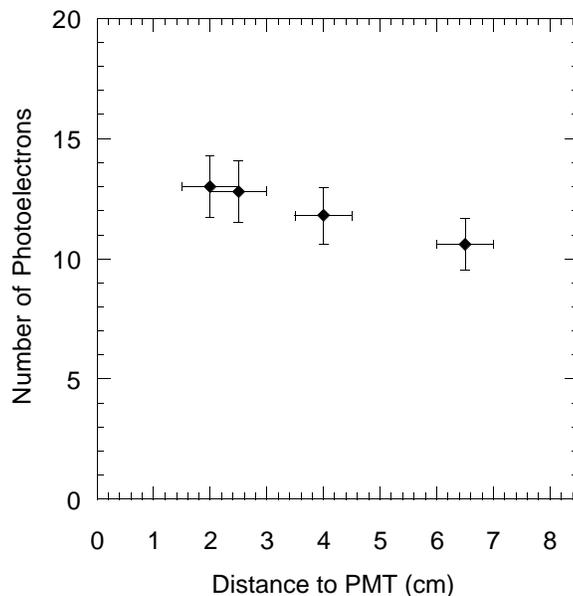
**Figure 6-25.** Measured threshold curves of the aerogel Cherenkov counters.



**Figure 6-26.** Photoelectron yield of the aerogel counters read out by fine mesh PMT's in the magnetic field.

The aerogel detectors were also read out by fine mesh PMTs operated in magnetic fields up to 1.0 T. The setup is very similar to the beam tests described above, except that the aerogel counters were sandwiched between two  $10 \times 10 \times 10$  mm trigger counters placed inside the magnet to have a better determination of the path of charged particles in the field. The 1.06 aerogel counter used in this beam test is the same as the one used in the previous test, while the dimensions of 1.0085 aerogel counter were changed to 12 cm long and  $10 \times 8.5$  cm in transverse dimensions. The photoelectron yields of  $n = 1.0085$  and  $n = 1.063$  aerogel counters at different magnetic fields are summarized in Figure 6-26.

Another beam test of the prototype 1.0085 aerogel counter was carried out at the KEK PS test beam facility. Sixteen photoelectrons were measured for 3.5 GeV/c charged pions using a newly developed fine mesh PMT with a glass window and an effective sensitive diameter of 44 mm in a 1.3 T field. This corresponds to a light yield of 18 photoelectrons for  $\beta = 1$  particles. A clear separation of the above and below threshold particles was obtained, Figure 6-8. The uniformity of the light collection efficiency was also studied in the  $n = 1.0085$  aerogel counter by running the beam at various transverse positions across the counter. The average light yield is plotted as a function of the distance from the beam to the PMT in Figure 6-27. The measured response is determined to change by less than 12% over a 4 cm distance.



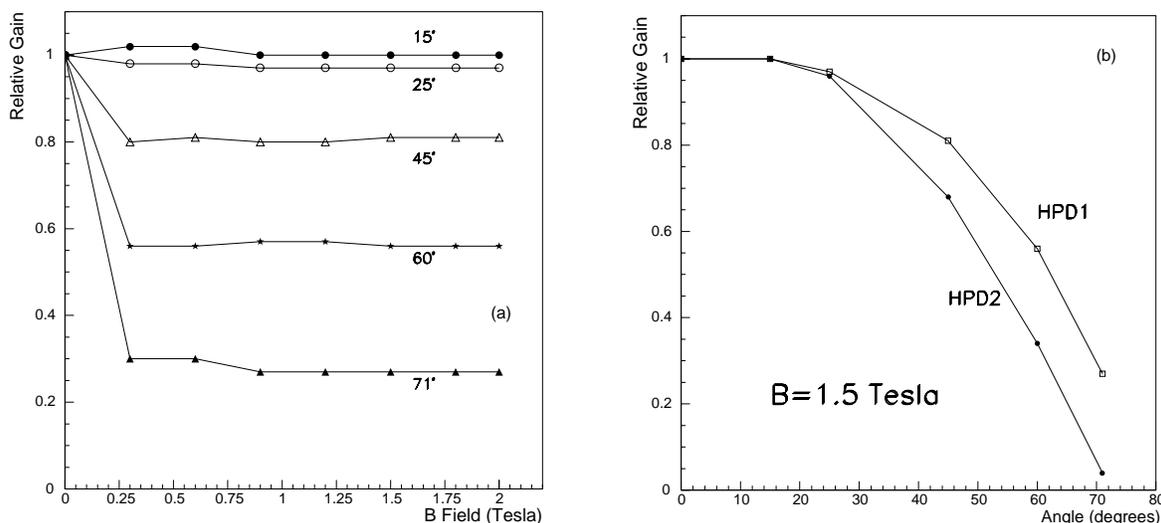
**Figure 6-27.** *Uniformity of the light collection efficiency.*

### The HPD option

The proof of principle that an aerogel-counter-based PID system can work efficiently in the *BABAR* environment, namely that more than ten photoelectrons are read out in a magnetic field of 1.5 T has been obtained using Fine Mesh PMTs. There are, however, good reasons to continue the investigation of other photon detector candidates. Among them, the most promising seems to be the Hybrid Photodiode (HPD) because of its substantial immunity to magnetic field effects, its wide angular acceptance and its extreme compactness (25 mm total length).

The HPD tube consists of a normal photocathode coupled to an anode made out of a silicon photodiode separated by a thin vacuum gap. An incoming photon is converted into a free electron at the photocathode, accelerated toward the silicon diode, and stopped in the depletion volume. A gain of about 3000 is achieved for a 15 kV electric field. The charge is collected on the anode and read out on a charge preamplifier.

**Results of measurements in a magnetic field.** Several 1" diameter HPD tubes have been tested in a magnetic field. The tubes were operated at 8 kV. A simple setup allowed variation of the angle between the HPD axis and the field direction with a precision of about 1–2 degrees. A green LED attached to the HPD was used as the source of light.



**Figure 6-28.** (a) HPD response versus magnetic field for a tube with a 5.6 mm gap; (b) HPD response versus angle with the field for two tubes with different gaps.

A scan in angle was performed, at values between  $0^\circ$  and  $71^\circ$  and at different values of the field between 0 and 2 T [Cal95]. The results of these measurements are given in Figure 6-28. From these measurements it is clear that HPDs show no loss of signal in magnetic field at  $0^\circ$  angle, and a small decrease at higher angles ( $\sim 30\%$  at  $45^\circ$ – $50^\circ$ , depending on the individual device).

To find the optimal parameters for an HPD tube to be used as an aerogel photon detector we intend to perform a series of laboratory tests in close contact with the manufacturer. A potential of about 15 kV will be used to operate the tubes in order to increase the gain and thereby reduce the importance of the electronic noise. However, it has yet to be demonstrated that 15 kV is possible without extending the gap between photocathode and anode beyond 5–8 mm and without any induced noise on the diode and the preamplifiers. Other parameters, such as diode thickness (relevant for noise) and window thickness, must also be optimized. We expect to receive the final tube design by the spring of 1995; in the meantime, several prototypes that are already available will be tested in the laboratory, using LEDs and aerogel cosmic rays tests, to evaluate progress.

### Tests at CERN

In 1995 we plan a test of all elements needed for the final choice of the detector configuration using a beam at CERN. In particular, aerogel blocks from different suppliers and a variety of photon detectors and preamplifiers will be studied. Based on these results, we plan to

build a full ( $30^\circ$ ) section of the ATC module at the beginning of 1996 and test it in a beam of pions, kaons, and protons to study all aspects of the detector performance in a realistic environment.

## References

---

- [Alp94] V. Alperovich *et al.*, *Nucl. Instr. Methods* **A340**, 429, (1994).
- [Ast94] D. Aston *et al.*, “Test of a Conceptual Prototype of a Total Internal Reflection Cherenkov Imaging Detector (DIRC) with Cosmic Muons,” SLAC-PUB-6731, (1994).
- [Ben91] B. Bencheikh, R. DeSalvo, W. Hao, C. Xu, and K. You, “A Simple Light Detector Gain Measurement Technique,” CERN-LAA-HC/91-007, (1991). CERN-PPE/91-143, (1991).
- [Cal95] M. Calvi *et al.*, “Hybrid Photodiode Performance in High Magnetic Field,” *BABAR Note # 208* (1995)
- [Des94a] R. DeSalvo *et al.*, CERN-LAA/HC/94-04 (1994); CERN-PPE/93-101 (1993), DEP Scientific Technical Information (1993); *DIRC R&D Plans*, Internal Note, Paris (1994).
- [Des94b] R. DeSalvo, private communication.
- [Edg92] J.P. Edgecumbe *et al.*, Proceedings of SPIE (1992).
- [Gea94] Magnet was provided and operated by R. Gearhart, SLAC.
- [Gen92] P. Bailly and J.-F. Genat, “A 100 Picosecond Resolution, 6 Microsecond Full Scale Multihit Time Encoder in CMOS Technology,” Proc. of the Third Int. Conf. on Electronics for Future Colliders, Chestnut Ridge, 57 (1993).  
See also: J.-F. Genat, “High Resolution Digital Time to Digital Converters,” Proc. of the First Int. Conf. on Electronics for Future Colliders, Chestnut Ridge, (1991).
- [Gil88] T.T. Giles *et al.* (CLEO-II), *Nucl. Instr. Methods* **A252**, 41 (1988); S. Ahmad *et al.* (STAR), *Nucl. Instr. Methods* **A330**, 416 (1993).
- [Gro92] J.E. Grove *et al.*, *Nucl. Instr. Methods* **A314**, 495 (1992).
- [Ham94] Hamamatsu, private communication.
- [Hru94] L. Hrubesh, LLNL, private communication.
- [Hub92] L. Hubbeling, “Large Photomultiplier Systems—A New Approach,” CERN-Preprint ECP 92-10, (1992).
- [Kic93] H. Kichimi *et al.*, *Nucl. Instr. Methods* **A325**, 451 (1993).

- [Men94] D. Mendez, Lockheed, private communication.
- [Oya94] J. Oyang, *BABAR Note # 137* (1994).
- [Pur93] Purschke *et al.*, “A New High Voltage Supply for Large Photomultiplier Systems,” International Conf. on Electronics for Future Colliders, CERN Preprint, (1993)
- [Rat92] B. Ratcliff, *BABAR Note # 92*; P. Coyle *et al.*, “The DIRC Counter: A New Type of Particle Device Identification for the *B* Factory,” *Nucl. Instr. Methods* **A343**, 292 (1994).
- [Shi94] X. Shi, *BABAR Note # 142* (1994). “Status Report on the Design of a Detector for the Study of *CP* Violation at PEP-II at SLAC,” SLAC-419 (1993).
- [Suz86] S. Suzuki *et al.*, *IEEE Trans. Nucl. Sci.* **NS-33**, 377 (1986); H. Saito *et al.*, *Nucl. Instr. Methods* **A270** (1988) 319; J. Janoth *et al.*, DESY 93-119 (1993).
- [Tso94] Dr. Tsou, JPL, private communication.

---



---

# Electromagnetic Calorimeter

---

## 7.1 Physics Requirements and Performance Goals

---

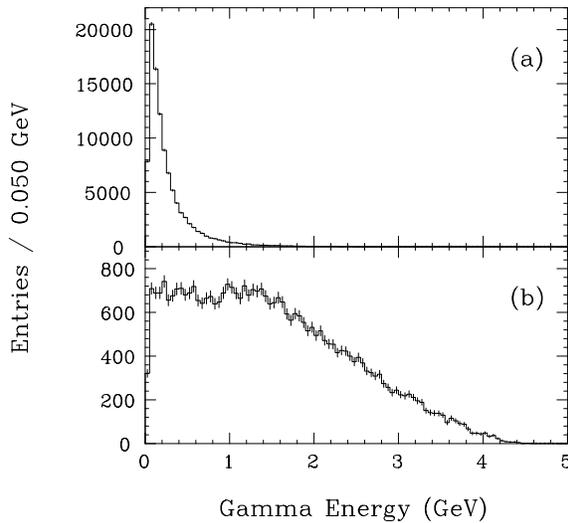
The introduction of high resolution, highly efficient CsI calorimeters has already led to major advances in the areas of conventional  $B$  meson, charmed meson, charmed baryon, and  $\tau$  physics at the  $\Upsilon(4S)$  [Sch94]. CsI calorimetry is also ideally suited to the primary physics goals of the  $BABAR$  detector, which require the tagging and reconstruction of  $B$  events containing  $CP$  eigenstate decays. The need to reconstruct  $CP$  eigenstates containing one or more  $\pi^0$  decays makes excellent electromagnetic calorimetry essential to the success of the experiment. High efficiency for low-energy photons is required to offset the small branching fractions typical of all  $CP$  eigenstates, and to make it possible to reconstruct final states containing several  $\pi^0$ s. Good energy and angular resolution improve the  $\pi^0$  and  $B$  mass resolutions, in turn improving the signal-to-background ratios of these rare decays. By facilitating lepton identification ( $e/\pi$  and  $e/\mu$  separation), the calorimeter also provides one of the tags needed for every  $CP$  analysis.

### 7.1.1 Physics Processes Influencing Performance Goals

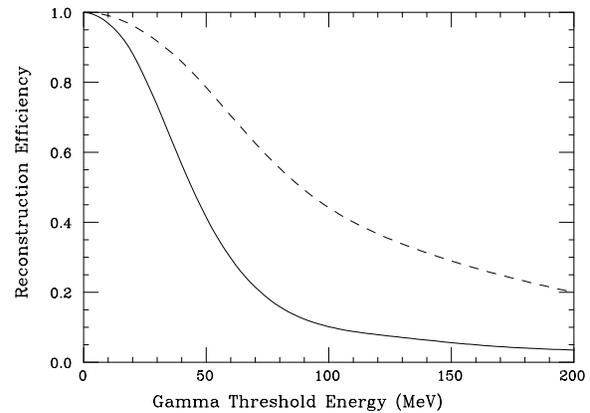
#### Energy and Angular Resolution

The understanding of  $CP$  violation in the  $B$  meson system requires the reconstruction of final states such as  $B^0 \rightarrow J/\psi K_S^0$ ,  $B^0 \rightarrow J/\psi K^{*0}$ ,  $B^0 \rightarrow \rho^\pm \pi^\mp$ ,  $B^0 \rightarrow D^+ D^-$ ,  $B^0 \rightarrow D^{*+} D^{*-}$ ,  $B^0 \rightarrow \phi \pi^0$ , and  $B^0 \rightarrow K_S^0 K_S^0$ . These channels either contain a direct  $\pi^0$  or can be reconstructed through a decay chain containing one or more daughter  $\pi^0$ s. For example, the decay  $K^{*0} \rightarrow K_S^0 \pi^0$  occurs about one-sixth of the time (which is one-half of its decays to  $CP$  eigenstates), and many  $D^\pm$  and  $D^{*\pm}$  decays with large branching fractions contain  $\pi^0$ s.

Generic  $B$  decays contain an average of 5.5 charged tracks and 5.5 photons, with about 50% of the photon energies below 200 MeV (Figure 7-1(a)). Because the  $B$  meson is moving relatively slowly in the lab frame, the photon energy resolution, rather than the angular resolution, determines the  $B^0$  mass resolution [Ale94b]. This is true even for the kinematically most



**Figure 7-1.** Photon energy spectrum in (a) generic  $B$  decays and (b)  $B^0 \rightarrow \pi^0\pi^0$  events.



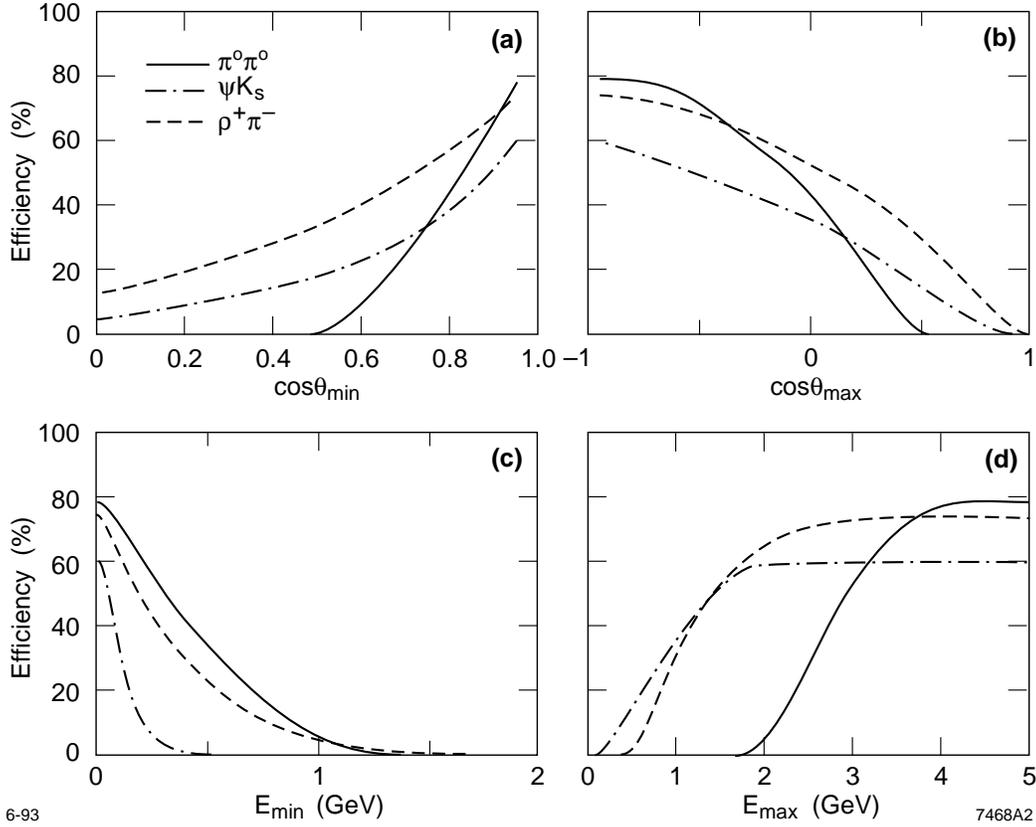
**Figure 7-2.**  $\pi^0$  (dashed line) and  $B$  meson reconstruction efficiency (solid line) vs. photon energy threshold for generic  $B$  decays. (Taken from reference [HEL92]. Efficiencies exclude geometric acceptance.)

extreme channel of interest,  $B^0 \rightarrow \pi^0\pi^0$  (used to untangle penguin contributions to  $B^0 \rightarrow \pi^+\pi^-$ ), in which both  $\pi^0$ s have momenta above 1 GeV/ $c$ , and photon energies extend up to 4.5 GeV (Figure 7-1(b)).

### Photon Efficiency and Solid Angle Coverage

Sensitivity to low-energy photons is critical for efficient  $\pi^0$  detection. The  $\pi^0$  and  $B$  meson reconstruction efficiency for generic  $B$  decays, as a function of the minimum detectable photon energy, is shown in Figure 7-2. Reconstruction efficiency is also sensitive to solid angle coverage. The reconstruction efficiency, as a function of both the minimum and maximum measured photon energy and the minimum and maximum detectable polar angle, is shown in Figure 7-3 for (1)  $B^0 \rightarrow \pi^0\pi^0$ , producing four relatively high-energy photons, (2)  $B^0 \rightarrow J/\psi K_S^0$ ,  $K_S^0 \rightarrow \pi^0\pi^0$ , producing four relatively soft photons, and (3)  $B^0 \rightarrow \rho^\pm\pi^\mp$ , producing only two photons. Losses in the forward direction (Figure 7-3(a)) are not as severe as for low-multiplicity all-charged modes, though  $\pi^0\pi^0$  is somewhat worse than  $\pi^+\pi^-$ . In the backward direction (Figure 7-3(b)) the demands on solid angle coverage are more modest.

Lepton tagging requires the calorimeter to reject pions misidentified as electrons to a level of  $\sim 10^{-3}$ .



**Figure 7-3.** Dependence of  $B$  reconstruction efficiency on (a) minimum detectable  $\theta$ ; (b) maximum detectable  $\theta$ ; (c) minimum detectable energy; and (d) maximum detectable energy for decay photons from  $B^0 \rightarrow \pi^+\pi^-$  (solid),  $B^0 \rightarrow J/\psi K_S^0$ ,  $J/\psi \rightarrow \ell^+\ell^-$ ,  $K_S^0 \rightarrow \pi^0\pi^0$  (dash-dot), and  $B^0 \rightarrow \rho^\pm\pi^\mp$ ,  $\rho^\pm \rightarrow \pi^\pm\pi^0$  (dash).

The requirements placed on the calorimeter by non- $CP$  physics are generally similar to those from  $CP$  physics. One exception is the need for efficient tagging of outgoing leptons in two-photon events. A backward endcap, which would contribute little to overall  $B$  reconstruction efficiency, would be important for this purpose. Another exception occurs in low-multiplicity final states that do not arise from  $B$  decays (*e.g.*, Bhabha scattering, and charm and tau decays), for which the energy range can extend up to the higher beam energy. For these processes, leakage, and hence detector depth, is the primary consideration.

### 7.1.2 Summary of Performance Targets

These physics requirements lead to a calorimeter design based on quasi projective CsI(Tl) crystals over the entire available solid angle.

The target energy resolution for photons at a polar angle of  $90^\circ$  is:

$$\frac{\sigma_E}{E} = \frac{1\%}{\sqrt[4]{E(\text{GeV})}} \oplus 1.2\%. \quad (7.1)$$

The constant term arises from front and rear leakage ( $\leq 1.0\%$ ), inter-calibration errors (0.25%) and light collection nonuniformity ( $\leq 0.5\%$ ). This expression does not include electronic noise, which is estimated to be  $\leq 150$  keV per crystal in the proposed readout system, because this contribution is negligible for typical clusters of 16–25 crystals, even at low energies. The resolution degrades at small polar angles, as discussed in subsequent sections, as a consequence of the staggered arrangement of crystals.

For comparison, this performance represents a considerable improvement over that of the CLEO-II calorimeter. A number of factors contribute to the improvement: the use of projective geometry in the forward endcap, a reduction of back leakage and losses in inactive material in front of and between the crystals, and reduction of electronic noise (allowing more crystals to be summed). In addition, the usable solid-angle coverage has been increased.

The angular resolution is determined by the transverse crystal size and average distance to the interaction point. The target angular resolution is:

$$\sigma_\theta = \frac{3 \text{ mr}}{\sqrt[2]{E(\text{GeV})}} \oplus 2 \text{ mr}. \quad (7.2)$$

As in CLEO-II, the minimum detectable energy for photons is about 10–20 MeV. It is expected to be largely determined by beam- and event-related backgrounds in the calorimeter, since the intrinsic CsI(Tl) efficiency should be close to 100%, even at energies as low as a few MeV.

The solid angle coverage of the calorimeter extends from a polar angle of 250 mr in the forward direction ( $\cos\theta = 0.97$ ) to  $\pi - 653$  mr in the backward direction ( $\cos\theta = -0.80$ ). A backward endcap has been designed to cover the angular range down to  $\pi - 560$  mr, primarily for two-photon tagging. It is not included in the baseline detector, but is a candidate for a future upgrade.

The performance of the calorimeter is summarized in Table 7-1.

Parameter	<i>BABAR</i> Design Performance
$\frac{\sigma_E}{E}$ (Stochastic Term) at 90°	1%
$\frac{\sigma_E}{E}$ (Constant Term) at 90°	1.2%
$\sigma_\theta$ at 1 GeV at 90°	5 mr
Efficiency at 20 MeV at 90°	85%
Efficiency at 100 MeV at 90°	95%
$\pi/e$ Rejection at 500 MeV/ $c$	$\text{few} \times 10^{-3}$
Minimum Detectable Energy	10–20 MeV
Electronic Noise/Crystal	$\leq 150$ keV

**Table 7-1.** Target performance for the *CsI(Tl)* calorimeter.

## 7.2 Calorimeter Overview

### 7.2.1 Technology Choice

The considerations leading to the choice of *CsI(Tl)* for the *BABAR* calorimeter are discussed in References [Ale94a] and [Hea94]. The *BABAR* physics goals require an electromagnetic calorimeter with excellent photon energy resolution and efficiency at low energies ( $\leq 100$  MeV). The calorimeter sits inside the solenoid and thus requires a readout system that operates in a magnetic field. These requirements, and considerations of cost and engineering impact upon other detector subsystems and the accelerator, lead to the choice of a *CsI(Tl)* crystal calorimeter with photodiode readout. The relevant properties of *CsI(Tl)* are summarized in Table 7-2.

The radiation environment at PEP-II is more severe than that at existing  $e^+e^-$  colliders because of much higher beam currents. Given the radiation hardness of *CsI* crystals, the calculated background rates during luminosity running, discussed in Chapter 12, imply a wide safety margin. However, in actual storage ring operations, experience shows that most of the radiation exposure comes during beam injection and machine studies. Scaling typical CESR operation to the beam currents expected at PEP-II gives a radiation dose of about 1.5 krad/yr at a radius of 45 cm and 0.5 krad/yr at 100 cm [Blu86]. The PEP-II masking system has been designed with graded apertures to prevent the deposition of large amounts of radiation in the interaction region due to accidental beam loss. Nonetheless, it is desirable

Properties	CsI	CsI(Tl)	CsI(Na)
Radiation Length (cm)	1.86	1.86	1.86
Absorption Length (cm)	34.2	34.2	34.2
Light Yield (Photons/MeV $\times 10^3$ )	2–10	50–60	38–44
Light Yield Temperature Coef. (%/°C)	0.1	0.1	0.1
Molière Radius (cm)	3.8	3.8	3.8
Peak Emission (nm)	320	565	420
Lower Wavelength Cutoff (nm)	260	320	300
Refractive Index at Emission Maximum	1.95	1.79	1.84
Decay Time (ns)	10	940	630
Density (g/cm <sup>3</sup> )	4.51	4.51	4.51
Hygroscopic	slight	slight	weak

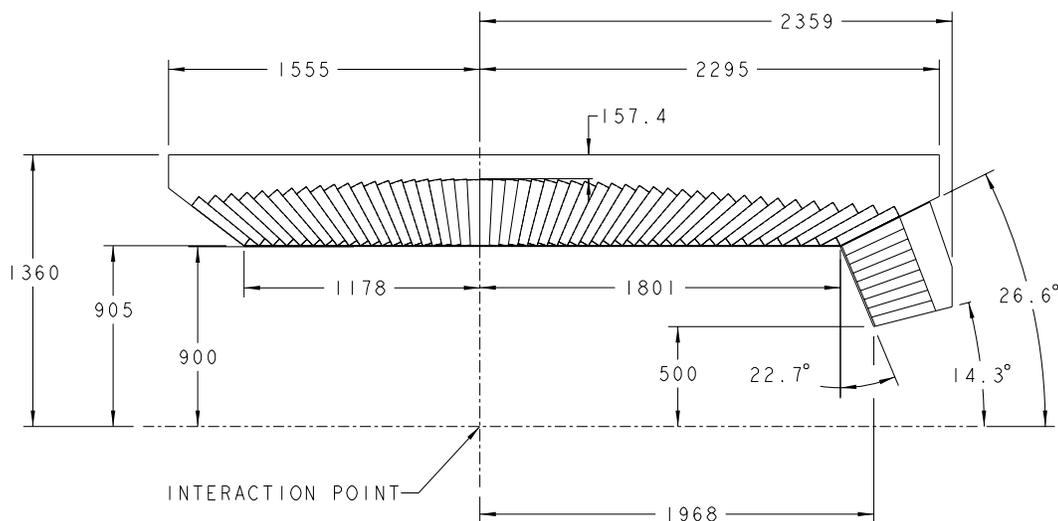
**Table 7-2.** *Properties of pure and doped CsI.*

to use detector components that are intrinsically radiation hard and can survive doses of tens of krad over the lifetime of the experiment.

## 7.2.2 Description of the Calorimeter

The baseline design of the CsI(Tl) calorimeter consists of a cylindrical barrel section and a forward conic endcap, as shown in Figure 7-4. Radially, the barrel is located outside the particle ID system and within the magnet cryostat. It weighs 24.2 metric tonne and is supported off each end of the coil at two points by eccentric pins. The barrel has an inner radius of 90 cm and an outer radius of 135.6 cm. It is located asymmetrically about the interaction point, extending, at its inner radius, 117.8 cm in the backward direction and 180.1 cm in the forward direction. The barrel covers a solid angle corresponding to  $-0.80 \leq \cos \theta \leq 0.89$  in the laboratory frame (Figure 7-4).

The barrel is built out of 250  $\mu\text{m}$ -thick carbon fiber composite (CFC) compartments that house individual crystals (Figure 7-5). At 90° to the beam, a compartment extends 43 cm radially. Aluminum support elements bonded to the back of each compartment allow them to be grouped into modules three crystals wide and seven crystals long, and to carry loads to a cylindrical strongback structure, which is fixed to the coil. By supporting the crystal at its rear, minimal material is placed in front of the crystals. The front material consists of



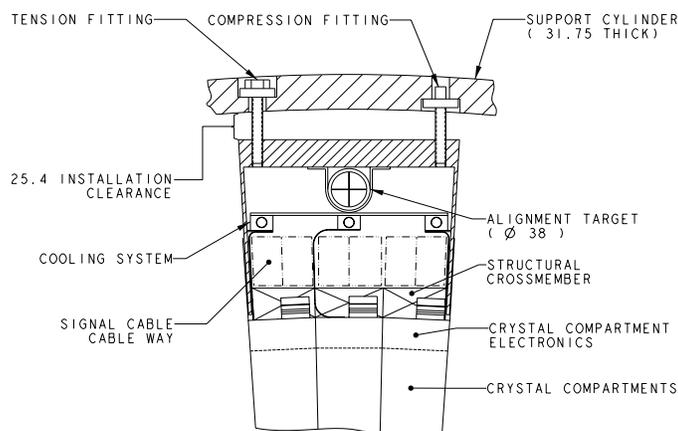
**Figure 7-4.** Side view showing dimensions (in mm) of the calorimeter barrel and forward endcap.

two 1 mm-thick cylinders of aluminum, separated by foam, which provide a gas seal and rf shielding; cooling, cables, and services are located at the back of each module.

The forward endcap is a conic section, with front and back surfaces tilted at  $22.7^\circ$  with respect to the vertical, which conforms to the drift chamber endplates, minimizing the distance to them. The endcap weighs about 4 metric tonne. It is supported off the coil rather than the flux return end doors, and is precisely aligned to the barrel to minimize the gap between the barrel and forward endcap. To further reduce the gap's effect on photon efficiency, it is nonprojective. The forward endcap starts at an inner radius of 50.0 cm from the beam line, making an angle of 250 mrad with respect to the beam. It covers the solid angle from  $\cos \theta = 0.97$  to  $\cos \theta = 0.89$  in the lab.

The forward endcap is composed of two identical monolithic sections consisting of load-bearing outer containers filled with a honeycomb of 250  $\mu\text{m}$ -thick CFC compartments, each containing one crystal package. The forward endcap is segmented vertically into two pieces, each of which can be retracted separately using a set of removable rails installed after the doors are opened. This allows fast access to the barrel end region.

Material in front of the calorimeter is kept to a minimum in order to achieve the required performance. In the barrel, the major source of such material is the DIRC ( $\sim 0.18X_0$ ). Including the beam pipe, the vertex detector, and the drift chamber, the total amount of material represents  $\sim 0.23X_0$  at normal incidence. The endplate of the drift chamber and the forward PID system contribute most of the material in front of the forward endcap. The endplate constitutes  $\sim 0.065X_0$ . The aerogel represents an average of  $\sim 0.095X_0$ . However,



**Figure 7-5.** Side view of typical barrel module and strongback.

the ATC material is lumped: the aerogel itself represents  $\sim 0.035X_0$ , with the balance concentrated in the phototubes.

In both the barrel and endcap, each crystal is wrapped with a diffuse, reflecting material on its sides, with a reflector on its front face. Diode readouts and the preamplifier packages are located at the rear of each crystal box. Each crystal is held loosely in its box by a rear stiffening ring that also acts to locate and support the diode/preamplifier package. Cables for diode bias, preamplifier power and cooling, and calibration signals enter the back of the compartment. The crystal and electronics are shielded by a foil liner and a metal cap, respectively.

There are 5880 crystals in the barrel, arranged in 49 rows of distinct sizes, each having 120 identical crystals in azimuth. The forward endcap has a total of 900 crystals, made up from nine distinct radial rows ( $3 \times 120, 3 \times 100, 3 \times 80$ ) arranged to give approximately the same crystal dimensions everywhere.

The crystals are tapered along their length with trapezoidal cross sections. The average area of the front faces of the crystals is  $4.8 \times 4.7$  cm, while the back face area is  $6.1 \times 6.0$  cm. They vary in length in  $0.5X_0$  steps from  $17.5X_0$  in the forward part of the barrel to  $16.0X_0$  in the backward part, with  $17.5X_0$  in the forward endcap. The barrel and endcap have total crystal volumes of  $5.2 \text{ m}^3$  and  $0.7 \text{ m}^3$ , respectively. Table 7-3 summarizes the crystal counts.

To minimize the loss of tracks that traverse inactive material between crystals, the crystals are arranged to be slightly nonprojective in  $\theta$  with respect to the interaction point.

	Barrel	Forward Endcap
Number of Crystals ( $17.5X_0$ )	840	900
Number of Crystals ( $17.0X_0$ )	840	0
Number of Crystals ( $16.5X_0$ )	840	0
Number of Crystals ( $16.0X_0$ )	3360	0
Total Number of Crystals	5880	900
Total Volume ( $\text{m}^3$ )	5.2	0.7
Number of Crystals/Modules ( $\phi$ )	120/40	120, 100, 80
Number of Crystals/Modules ( $\theta$ )	49/7	9
Module Organization	$3_\phi \times 7_\theta$	Two C-sections

**Table 7-3.** Summary of crystal counts and sizes in the barrel and endcap.

### 7.2.3 Readout Chain and Trigger

The electronic readout chain is shown schematically in Figure 10-13 and described in detail in Chapter 10. Two independent photodiodes view the scintillation light from each crystal. Each diode has an independent dual-range preamplifier with calibration input. This two-fold redundancy is based on reliability issues presented in Section 7.4.4. Shaped, amplified signals are carried differentially to the end of the detector where they are input to ADCs which sample the data at a rate between 3 MHz and 5 MHz. The signals are digitized to an 18-bit effective dynamic range.

Digitized signals are passed out of the detector on fiber optic cables to the DAQ receiver modules, where calibration and waveform processing is performed. The data are then stored in video RAM. Trigger sums of  $\sim 25$  crystals are formed digitally on the DAQ cards. These data are passed along to the trigger for the full trigger decision.

The barrel digitization electronics is housed in 20 bays at the ends of the barrel. Forward endcap electronics is housed behind the endcap in six bays. Fiber optic cables are routed through channels between the barrel and endcap flux returns. Table 7-4 summarizes the channel organization and required space for the on-detector readout electronics.

	Barrel	Forward Endcap
Crystals	5880	900
Photodiodes	11760	1800
Preamplifier Boards	5880	900
Cables to ADC	11760	1800
Channels/ADC Board	25	25
ADC Boards/Bay	12	6
ADC Bays	20	6
Number of Fibers to DAQ (cm <sup>2</sup> )	240 (4)	36 (1)

**Table 7-4.** *Summary of front-end electronics components.*

## 7.2.4 Review of Options

The baseline calorimeter has been described above. A number of variations have been evaluated, with some remaining under consideration. They are described briefly below.

### Monolithic versus Modular Support System

Monolithic and modular support structures have been considered. Monolithic support structures, such as the CFC system used in the L3 BGO system, are fully integrated systems that combine strength and low mass. The major drawbacks to such assemblies are the increased complexity of fabrication and the high cost, arising in part from the need to incorporate proper compensation for the deformation of the structure when it is loaded.

Modular systems require more material to achieve the required strength and resistance to deformation. They are less expensive because each unit is small and hence easier to fabricate, handle, and assemble. Fabrication and assembly can occur at many locations simultaneously.

The *BABAR* calorimeter structure is a hybrid, with a modular system in the barrel and a monolithic system in the endcap. The relatively small size of the endcap should alleviate the problems of fabrication anticipated for a larger monolithic system, while reducing installation and access problems.

### Longitudinal Segmentation of Crystals

Segmentation of crystals along their length (at about  $3X_0$  from the front) to provide information on longitudinal shower development has been considered. Some improvement in  $e/\pi$  separation is indicated by Monte Carlo studies, but this is judged not to have a significant impact on the physics performance. Since the cost of implementation is prohibitive, this option is no longer being considered.

### Crystal Cross Section: Trapezoidal Versus Hexagonal

While trapezoidal cross sections are conventional for crystal calorimeters, the use of hexagonal cross sections may offer certain advantages. In the endcap, an arrangement of hexagonal crystals on a spherical bounding surface can be designed that allows a single crystal size to be used, eliminating the problems associated with progressively smaller crystals of trapezoidal cross section and with abrupt transitions across superlayer boundaries.

Laboratory measurements indicate that for hexagonal crystals, about 25% more light is collected than for trapezoidal crystals. The response along the crystal is also more uniform [Jes94]. This might allow the acceptance criteria for crystal light yield and uniformity to be relaxed, leading to decreased costs. The crystal cost depends on the growth technique adopted; this is lower in some cases for hexagonal crystal shapes. However, the decreased growing cost is largely balanced by the increased cost of cutting and polishing six side surfaces instead of four.

Engineering design considerations make the cost and complexity of the support structures for hexagonal crystals greater than those for trapezoidal crystals. Complications also arise with hexagonal crystals at the interface of the barrel with the forward endcap, where it is vital to minimize gaps and inactive material.

Overall, we have not been convinced of any cost advantage in using hexagonal crystals. Therefore, trapezoidal crystals have been adopted in the baseline design for both barrel and endcap. The study of hexagonal crystals for use in the endcap will continue, subject to physics, engineering, and cost considerations, since the endcap is not on the same critical path as the barrel.

### Direct Photodiode Versus Wavelength Shifter/Photodiode Readout

Two options for light collection and readout are presently under study. They are large area photodiodes directly attached to the crystal back face, and smaller-area photodiodes affixed to the edges of a wavelength-shifting plastic plate covering the crystal back face. Both

techniques have produced results consistent with the desired noise performance. The cost, reliability, and reproducibility of each method are currently being evaluated (Section 7.4.1).

## 7.3 Projected Calorimeter Performance

---

### 7.3.1 Contributions to Photon Resolution and Efficiency

A number of effects beyond shower statistics contribute to the calorimeter resolution, with different effects dominating in different energy regions. Some of these contributions are linked, in that reducing the effect of one may facilitate improving another. The effects include:

1. Fluctuations in energy loss caused by leakage out of the front and the rear, by losses out the side due to the staggered geometry, and by losses in the inter-crystal material.
2. Fluctuations in transverse energy spread, with cluster sizes optimized to include effects of noise and background.
3. Scintillation light collection efficiency.
4. Light collection nonuniformity, coupled to fluctuations in shower position.
5. Incoherent electronic noise in the readout device.
6. Coherent noise and pickup.
7. Digitizer resolution.
8. Calibration of the energy scale, including crystal inter-calibration and the effects of time and temperature.
9. Material *en route* to the calorimeter. This affects efficiency as well as resolution.
10. Beam-related backgrounds.

In the high-energy regime, rear leakage (1) dominates the energy resolution. Fluctuations in energy loss (1), along with contributions from nonuniform light collection and calibration uncertainties (4) and (8), are expected to yield the constant term shown in Equation (7.1). Transverse shower spread and choice of reconstruction algorithms (2) also contribute at high energy, but become especially important to the energy resolution in the intermediate-energy range.

The low-energy regime is characterized by the importance of noise (5) and (6), and of beam-related backgrounds (10). Such backgrounds, especially if well above nominal levels, can also indirectly affect resolution through the need to reduce the number of crystals summed into a cluster, or the need to reduce the shaping time for the readout from the optimal value for maximizing the signal-to-electronic-noise ratio. Electronic noise is determined by the properties of the photodiode and preamplifier/shaper. Its value in units of energy is inversely proportional to the amount of light collected per unit of energy deposited (3).

Photon efficiency is primarily affected by the amount and distribution of material in front of the calorimeter (9), *i.e.*, in other detector systems, and also by losses (1) due to geometric edge effects.

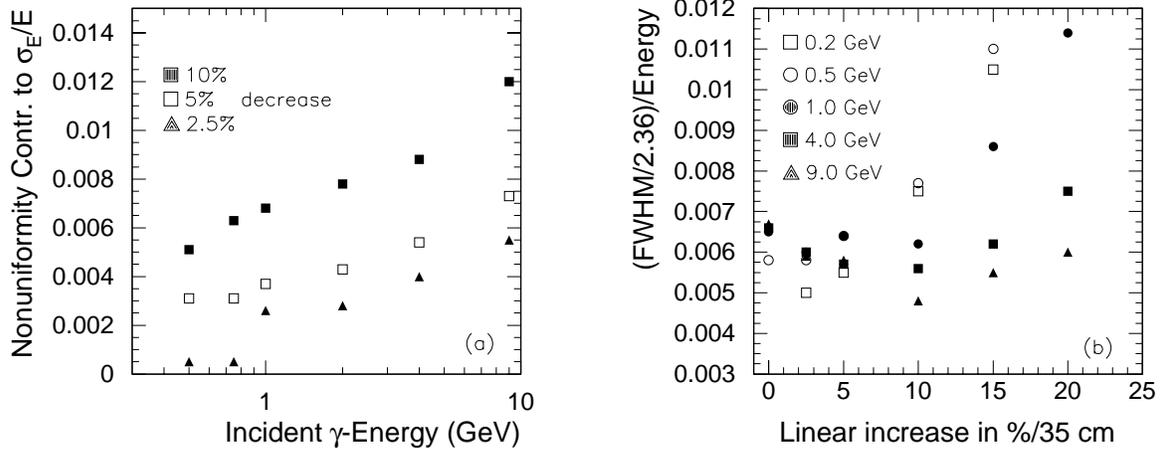
Photon angular resolution is given by position resolution divided by the distance to the interaction point. Even for the best possible algorithm, fluctuations in shower shape (transverse energy spread) will limit the position resolution. The transverse crystal size at the rear should be less than two Molière radii, but returns diminish when that size is comparable to a Molière radius.

### Contribution to Energy Resolution from Crystal Nonuniformity

The effect of nonuniform light output along a CsI crystal has been studied using a stand-alone GEANT simulation. An untapered, 35 cm-long crystal with large lateral dimensions ( $40 \times 40$  cm) is divided along its length into 100 slices. A nonuniform response is imposed as a weight on the energy deposited in the slices for each event.

A study of linearly decreasing and increasing efficiencies for light collection along the crystal for photon energies between 200 MeV and 9 GeV has been done. The contribution to the energy resolution  $\sigma_E/E$  due to crystal nonuniformity is shown in Figure 7-6(a) for three different linearly decreasing weight functions (10%, 5%, and 2.5% over the full length of 35 cm). Nonuniformities with decreases up to 5% over the full crystal length lead to an energy resolution contribution which meets the requirement stated in Section 7.1.2.

Linearly increasing weight functions lead to changes in the energy resolution that cannot be presented in the same way; for photon energies above 1 GeV the energy resolution improves. Figure 7-6(b) shows the resolution  $\sigma_E/E$  as a function of the increase over 35 cm. The result for higher energies can be understood as a compensation for leakage of the electromagnetic shower at the rear end of the crystal. A similar behavior has been observed in a Monte Carlo study for the CLEO-II calorimeter [Blu86].



**Figure 7-6.** Energy resolution for nonuniform light output along a CsI crystal of 35 cm length. (a) Contribution to  $\sigma_E/E$  from light output decreasing towards the crystal end (2.5%, 5%, 10% over the crystal length). (b)  $\sigma_E/E$  for light output increasing towards the crystal end. Note that the crystal dimensions used for this study are much larger than baseline lengths and transverse leakage is ignored.

### Contribution to Energy Resolution from Beam Backgrounds

Simulations indicate that beam-related background is almost entirely caused by beam-gas interactions or by off-axis beam particles striking machine elements near the interaction region. Synchrotron radiation photons are a negligible source of background [Zis91]. The energy distribution of the shower remnant photons has a median of  $\sim 500$  keV. About 311 depositions per microsecond with a total energy of 0.33 GeV are expected in the barrel, while in the forward endcap, about 20 depositions per microsecond with a total energy of 0.07 GeV are expected. There are on average 0.7 (0.4) depositions per microsecond above 20 MeV in the barrel (forward endcap) that can produce fake photons in reconstructed events. The effect of beam backgrounds on photon energy resolution has been studied. Lost particle backgrounds corresponding to 1 times nominal and 10 times nominal levels in  $1 \mu\text{s}$  windows were generated together with 100 MeV photons in the barrel and forward endcap. Comparison with the zero background case shows contributions, in quadrature, to the resolution  $\sigma = FWHM/2.36$  of 0.55 MeV from 1 times nominal background and 1.05 MeV from 10 times nominal background (*i.e.*, 0.55% and 1.05%, respectively, at 100 MeV).

### 7.3.2 Modeling

The expected calorimeter performance has been studied using a GEANT Monte Carlo simulation of the full detector. The GEANT model of the calorimeter contains a description of the barrel with crystal lengths varying in  $0.5X_0$  steps from  $17.5X_0$  in the forward part to  $16X_0$  in the backward part, with  $17.5X_0$  crystals in the forward endcap. The model includes  $500\ \mu\text{m}$  of carbon fiber composite as structural material between the crystals and a 5 mm gap at the barrel/forward endcap interface.

The full detector model includes a DIRC particle ID device in the barrel and an aerogel device in the forward endcap. They are modeled according to the description given in Chapter 2. The DIRC represents about  $0.18X_0$  at normal incidence. Together with the beam pipe, the silicon vertex detector, and the drift chamber, this totals  $0.23X_0$  in the barrel region at normal incidence. In the forward region, the materials incorporated in the detector model, as represented in Figure 2-16, is different from that described in the balance of the Technical Design Report. Thus, in the forward region, material in front of the endcap calorimeter totals about  $0.35X_0$ .

Photons are reconstructed by forming clusters of contiguous crystals with energies above 0.5 MeV. In studies of the performance for single photons, their energies are determined by summing up the energy deposited in an area corresponding to  $5 \times 5$  crystals at  $90^\circ$ . In studies involving  $\pi^0$ s, the cluster energy is determined as the sum of the energies of all contiguous crystals making up the cluster. To deal with overlapping showers, the cluster is searched for additional local energy maxima, which are then treated as seeds for subclusters. The subcluster energies are determined in an iterative way by dividing up the energy in shared crystals according to the seed energy as it develops and adding it to the seed. In both methods, an incoherent electronic noise contribution of 150 keV per crystal is included. Since the distribution of deposited energy typically shows a non-Gaussian tail on the low side, the resolution is defined as  $\sigma = \text{FWHM}/2.36$ . In order to allow for effects not included in the GEANT simulation, such as nonuniform light collection and calibration uncertainties, 0.5% is added in quadrature to  $\sigma_E/E$ .

Efficiency for single photons is defined as the fraction of events with a ratio of measured over generated photon energy,  $R = E/E_0$ , above an  $E_0$ -dependent cut value:

$$RCUT(E_0) = 1 - 0.05 \times \frac{\sigma_E/E(E_0)}{\sigma_E/E(E_0^{max})}, \quad (7.3)$$

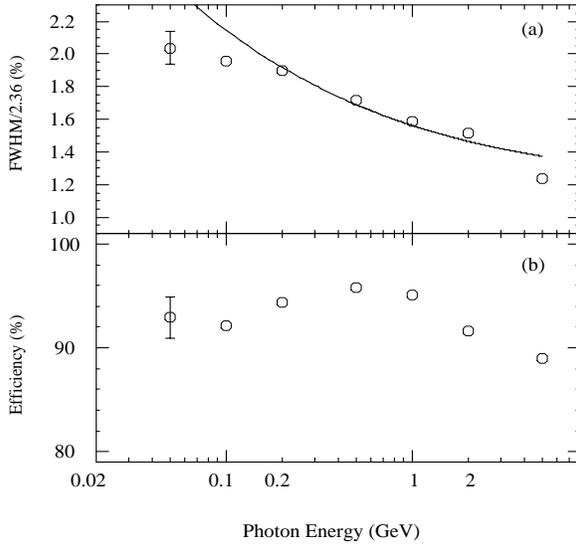
where  $\sigma_E/E(E_0)$  is the target energy resolution (Equation 7.1) at energy  $E_0$ , and  $E_0^{max} = 5\ \text{GeV}$ ; therefore,  $RCUT(5\ \text{GeV}) = 0.95$ , and  $RCUT(100\ \text{MeV}) = 0.92$ . (Note that in calculating  $R$ , the measured photon energy distribution has been calibrated so that its peak value equals  $E_0$ .)

### 7.3.3 Expected Performance for Photons

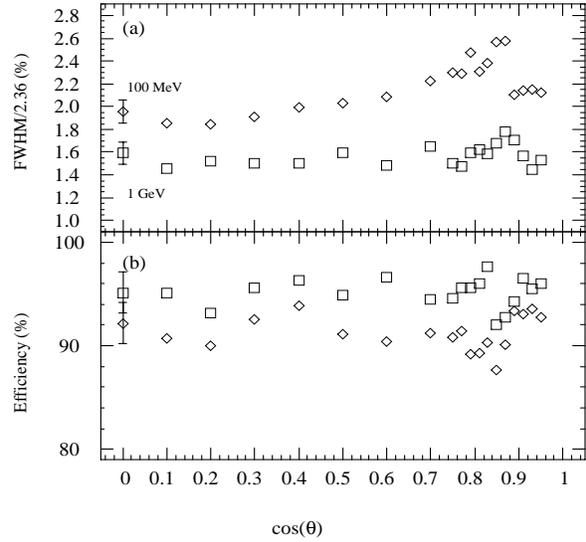
Initially, the performance of the calorimeter has been studied without any material in front of it. Figure 7-7 shows energy resolution and efficiency, as defined above, *vs.* energy, for single photons with polar angles near  $90^\circ$ . Also shown (solid line) is the target energy resolution (Equation 7.1). The predicted resolution, based on  $16.5X_0$  crystals near  $90^\circ$ , is just above the target resolution. The sawtooth crystal layout at the inner radius of the barrel calorimeter causes some performance degradation towards small polar angles. Figure 7-8 shows energy resolution and efficiency *vs.* polar angle for photons with energies of 100 MeV and 1 GeV. One can see a 40% (20%) worsening of resolution for 100 MeV (1 GeV) photons at  $\cos\theta$  values near the end of the barrel. This effect is due to increased leakage, mostly out of the uncovered crystal side faces. This degradation is probably acceptable over the limited solid angle involved. (It could be ameliorated by decreasing the crystals' transverse dimensions (front-side leakage) and increasing their length (rear-side leakage). Studies have shown that simply filling in the sawtooth pattern degrades efficiency and uniformity of light collection significantly [Spi94]).

A photon that converts in the material between the main tracking volume and the calorimeter can either be lost entirely or so degraded that it will fail to lead to a satisfactory  $\pi^0$  mass. While energy loss by the conversion  $e^+$  and  $e^-$  is a contributing factor, even more important is their deflection by the solenoidal field, so that one or both of the leptons may not reach the barrel calorimeter. For the endcap, in contrast, the conversion leptons will spiral around the field lines and eventually reach the calorimeter close to the intersection point of the projected photon. Important considerations besides the amount of material traversed include the proximity of that material to the calorimeter and whether the particle ID detector can “flag” the conversion [Eis93]. The effect of material in front of the calorimeter has been simulated using the full GEANT model of all detector systems. The results for energy resolution and efficiency are shown in Figure 7-9. Note that there is only a small effect on resolution since most photons pass through the material unaffected. However, those that interact experience a significant energy loss, resulting in a reduction in efficiency of 2-32% at 100 MeV and 6-12% at 1 GeV, depending on  $\cos\theta$ . More details can be found in Reference [BaB94]; the effect on  $\pi^0$  reconstruction is presented below.

The detailed reconstruction algorithms needed to determine the positions of showers are still being developed. A crude algorithm being used at the moment is based on a corrected energy-weighted center-of-gravity method. Due to the finite granularity of the calorimeter, the resulting shower positions are biased towards the crystal centers. A correction, derived from Monte Carlo, is applied in an attempt to remove this bias. The resulting resolution in polar angle  $\theta$  is shown in Figure 7-10 as a function of energy, together with a parameterization of the CLEO-II angular resolution. It is slightly better than the CLEO-II resolution, as a result of our lower electronic noise, smaller transverse crystal sizes, and longer distances to



**Figure 7-7.** (a) Energy resolution and (b) efficiency (as defined in the text) vs. energy for photons with polar angles near  $90^\circ$  and  $16.5X_0$  crystals without any material in front of the calorimeter. The solid line shows the target resolution (Equation 7.1).

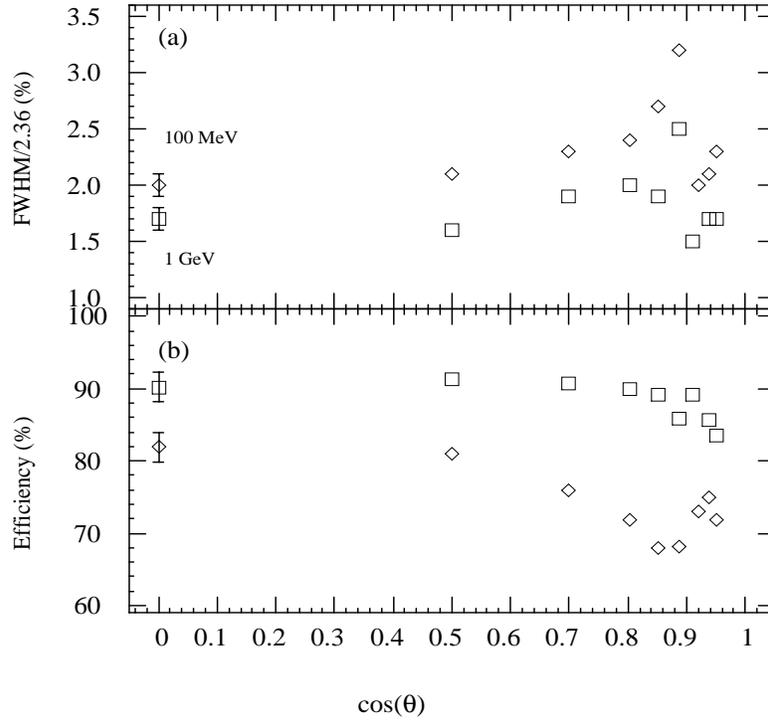


**Figure 7-8.** (a) Energy resolution and (b) efficiency vs.  $\cos\theta_{cm}$  for 100 MeV and 1 GeV photons without any material in front of the calorimeter.

many of the crystals. Figure 7-11 shows  $\theta$  resolution vs. polar angle for different photon energies. Note that the angular resolution is comparable for showers in the barrel and in the endcap since, in contrast to the case at CLEO, the endcap crystals are also arranged in quasi-projective geometry.

### 7.3.4 Expected Performance for $\pi^0$ s

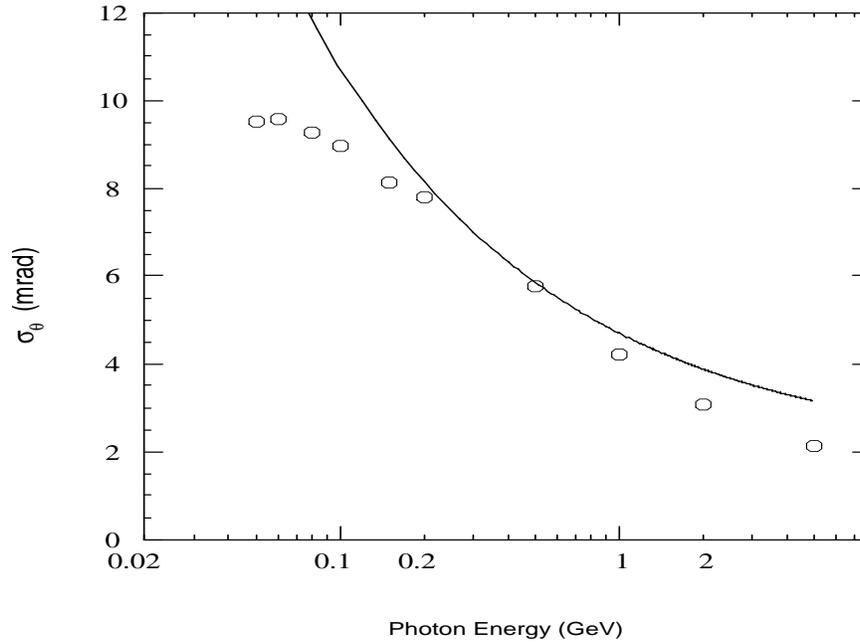
Most photons detected in the calorimeter arise from  $\pi^0$ s. To study  $\pi^0$  mass resolution and efficiency and the effect of material in front of the calorimeter, single, monoenergetic  $\pi^0$ s have been simulated using the GEANT model of the full detector. Figure 7-12 shows  $\pi^0$  mass resolution and efficiency vs.  $\pi^0$  momentum for  $\pi^0$ s with  $\cos\theta_{cm} = 0$ . Mass resolution is again defined as  $\sigma = \text{FWHM}/2.36$ ; efficiency is obtained using a  $\pm 3\sigma$  cut around the  $\pi^0$  mass. Also indicated is the performance without any material in front for  $\pi^0$  momenta of 100 MeV/c and 1 GeV/c. As is the case for single photons, the inactive material has very little effect on the resolution, but significantly degrades the efficiency.



**Figure 7-9.** (a) Energy resolution and (b) efficiency for 100 MeV and 1 GeV photons vs. polar angle with the full detector model.

In addition, several benchmark processes relevant for measuring  $CP$  asymmetries have been used to study  $\pi^0$  mass resolution and efficiency:  $B^0 \rightarrow J/\psi K^{*0}$ ,  $K^{*0} \rightarrow K_S^0 \pi^0$ ;  $B^0 \rightarrow \rho^\pm \pi^\mp$ ,  $\rho^\pm \rightarrow \pi^\pm \pi^0$ ; and  $B^0 \rightarrow \pi^0 \pi^0$ . The  $\pi^0$ s from those processes cover a wide range of momentum, with  $B^0 \rightarrow J/\psi K^{*0}$  yielding the softest  $\pi^0$ s and  $B^0 \rightarrow \pi^0 \pi^0$  the hardest. Figure 7-13(a) shows the  $\gamma\gamma$  invariant mass spectrum in the decay  $B^0 \rightarrow J/\psi K^{*0}$ , with no material in front of the calorimeter. (Note that the  $K_S^0$  is not allowed to decay to  $\pi^0 \pi^0$  in this study, and the  $J/\psi$  decays into lepton pairs only.) The  $\pi^0$  mass resolution and efficiency are  $\sigma_{m_{\pi^0}} = 4.0 \text{ MeV}/c^2$  and 74%, respectively. Also shown (Figure 7-13(b)) is the mass spectrum when the detector material is included in front. Again, the additional material leads to a relatively small degradation in resolution to  $\sigma_{m_{\pi^0}} = 4.5 \text{ MeV}/c^2$ , but causes a very significant low-mass tail which results in a drop in efficiency to 59%. The results for all three channels are summarized in Table 7-5.

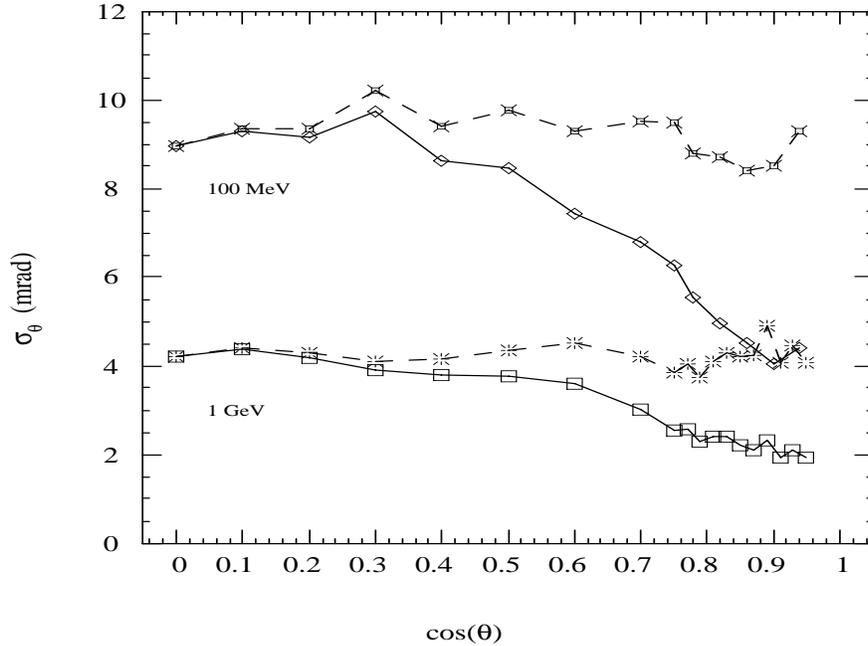
In the all-neutral channel  $B^0 \rightarrow \pi^0 \pi^0$ , we achieve a  $B^0$  mass resolution of  $\sigma_{m_B} = 62 \text{ MeV}/c^2$  (Figure 7-15) and an overall reconstruction efficiency of 62%, using a  $\pi^0$ -mass cut of  $110 \leq m_{\gamma\gamma} \leq 160 \text{ MeV}/c^2$  and a  $B$ -meson mass cut of  $5000 < m_{\pi^0 \pi^0} < 5350 \text{ MeV}/c^2$  (Figure 7-14).



**Figure 7-10.** Resolution in polar angle vs. energy for photons with polar angles near  $90^\circ$ . The solid line shows a parameterization of the measured CLEO-II resolution in the barrel.

$B$ Decay Channel	$\sigma_{m_{\pi^0}}$ (MeV/ $c^2$ )	Efficiency (%)
$B^0 \rightarrow J/\psi K^{*0}$	4.5	59
$B^0 \rightarrow \rho^\pm \pi^\mp$	5.5	65
$B^0 \rightarrow \pi^0 \pi^0$	8.3	72

**Table 7-5.**  $\pi^0$  mass resolution and efficiency in various  $B$  decay channels, modeled with the complete detector material in front of the calorimeter.

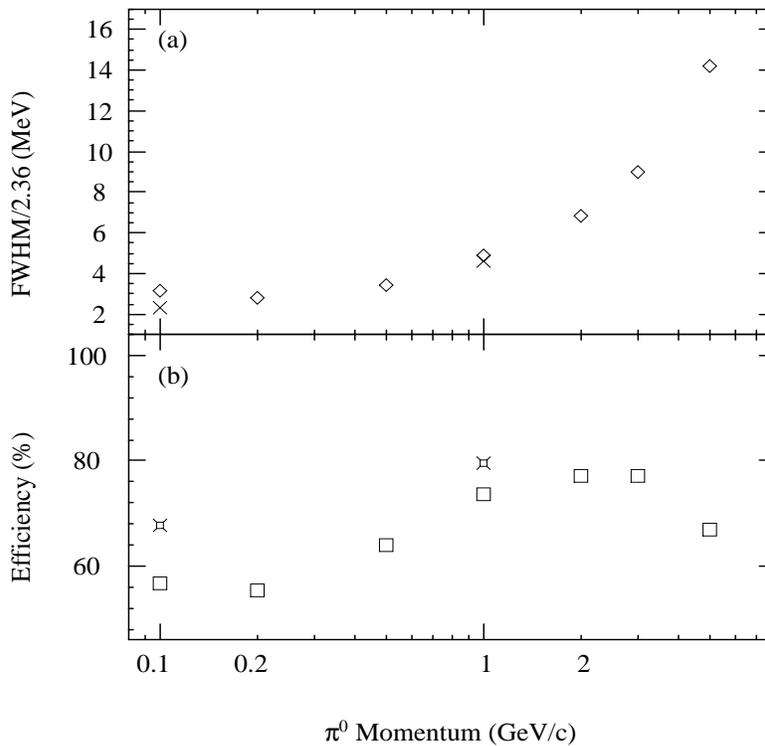


**Figure 7-11.** Resolution in polar angle vs. polar angle for 100 MeV and 1 GeV photons. The points connected by the dashed lines show the resolution with the distance effect removed.

### 7.3.5 $e/\pi$ Separation

In the *BABAR* calorimeter, electrons are totally absorbed, while muons and some pions, having momenta above a few hundred MeV/ $c$ , penetrate to the Instrumented Flux Return (IFR), depositing only a fraction of their energy.

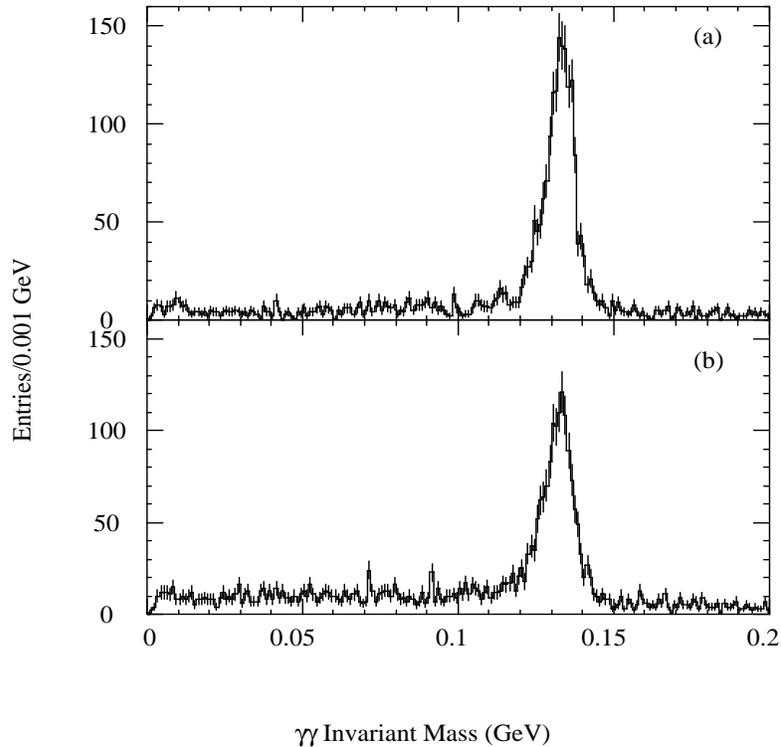
Since the drift chamber provides a precise momentum measurement, electrons can be identified using energy-momentum matching. In addition, electromagnetic showers are usually narrower than hadronic showers, so the lateral shower profile can be used to further distinguish electrons from pions. The results of such a study [Zhu95] using a GEANT simulation of the detector, with electrons of momenta in the range 500 MeV/ $c$  to 5 GeV/ $c$  produced at polar angles of  $90^\circ$ , indicate that the probability of misidentifying a pion as an electron is  $\sim 10^{-3}$  for an electron efficiency of  $\sim 95\%$  (Table 7-6).



**Figure 7-12.** (a) Mass resolution and (b) efficiency (as defined in the text) for single, mono-energetic  $\pi^0$ s vs.  $\pi^0$  momentum with the full detector model ( $\pi^0$ s are generated with  $\cos \theta_{cm} = 0$ ). The crosses indicate the performance for 100 MeV/c and 1 GeV/c  $\pi^0$ s with no material in front of the calorimeter.

Momentum	Events	Electron	$\pi^+$	$\pi^-$
	Generated	1000	5000	5000
0.5 GeV/c	Surviving	926	3	7
1.0 GeV/c		968	1	1
5.0 GeV/c		972	0	1

**Table 7-6.** Results of a GEANT study of  $e/\pi$  separation.

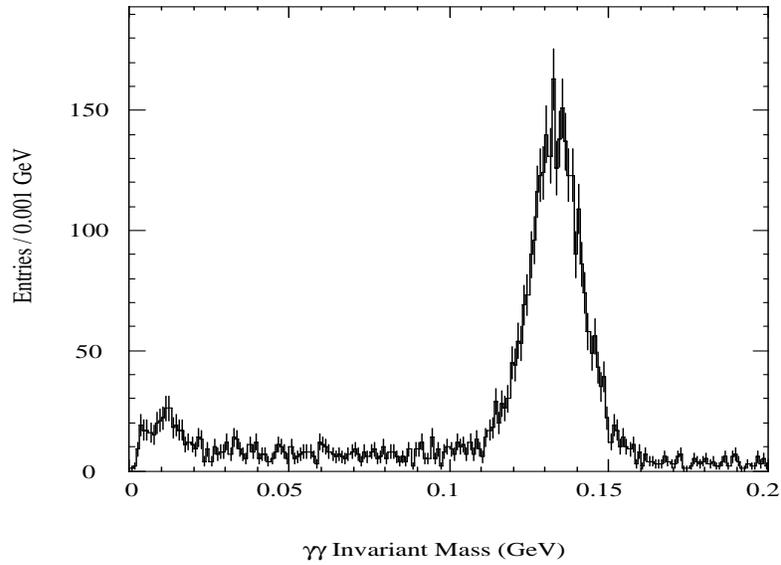


**Figure 7-13.**  $\gamma\gamma$  invariant mass distribution in the process  $B^0 \rightarrow J/\psi K^{*0}$ ,  $K^{*0} \rightarrow K_S^0 \pi^0$ ,  $K_S^0 \rightarrow \pi^+ \pi^-$  for (a) no material in front of the calorimeter and (b) the full detector model.

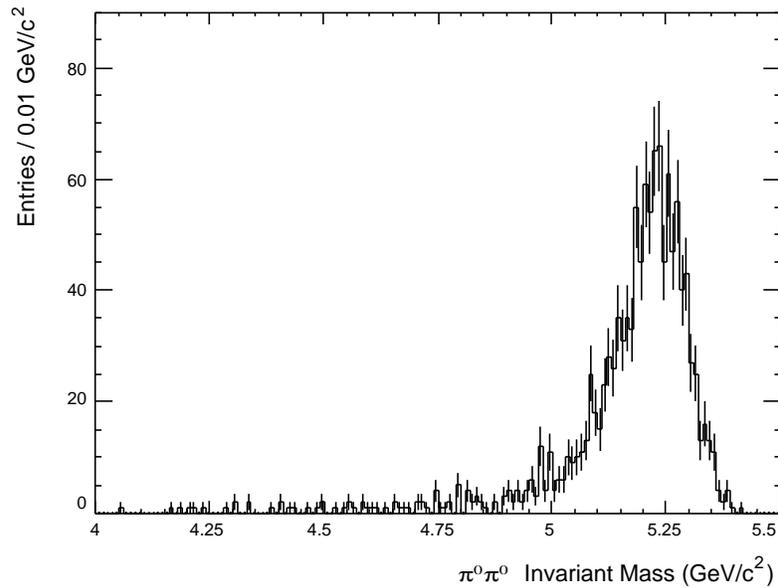
### 7.3.6 Performance and Cost Optimization

Taking into account the correlation between maximal photon energy and polar angle caused by the boost of the center-of-mass system, crystal lengths can be varied as a function of polar angle without significantly affecting the calorimeter performance in  $CP$ -relevant  $B$  decay modes. The process  $B^0 \rightarrow \pi^0 \pi^0$  has been used as a benchmark. Crystal lengths are chosen to keep the corresponding energy resolution (Figure 7-16) close to the target resolution at all polar angles. This mode has also been used to check the effect of shorter crystals on the  $B$  mass resolution. For a calorimeter in which all crystals are shortened by  $1X_0$ , one obtains  $\sigma_{m_B} = 73 \text{ MeV}/c^2$  and an overall reconstruction efficiency of 57%. Such a performance is significantly worse than that of the baseline design.

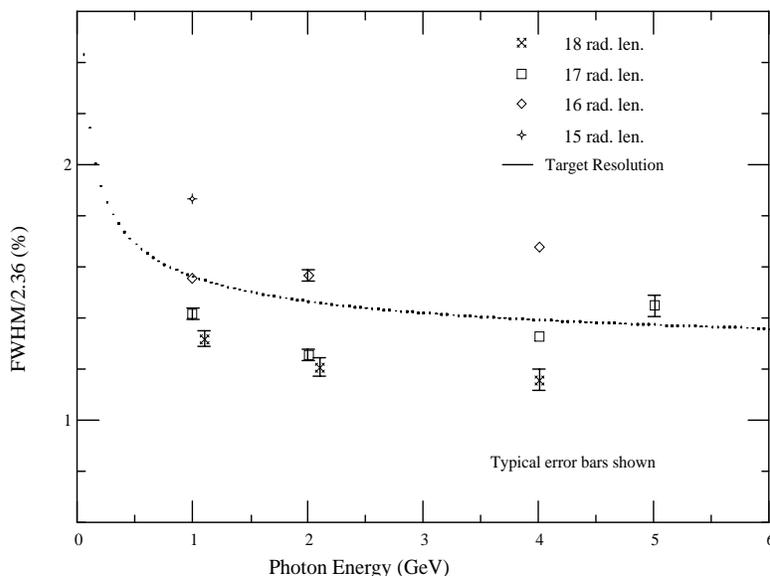
Using the center-of-gravity method described above, angular resolution has been studied as a function of transverse crystal size (see Figure 7-17). One finds a rather weak dependence, especially at high energies where there is significant energy sharing between the crystals in



**Figure 7-14.**  $\gamma\gamma$  invariant mass distribution in the process  $B^0 \rightarrow \pi^0\pi^0$ , using the full detector model.



**Figure 7-15.**  $\pi^0\pi^0$  invariant mass distribution where  $110 < m_{\gamma\gamma} < 160 \text{ MeV}/c^2$ , in the process  $B^0 \rightarrow \pi^0\pi^0$ , using the full detector model.



**Figure 7-16.** Energy resolution vs. energy for a number of crystal lengths and for photons with polar angle near  $90^\circ$ . The solid line shows the target resolution (Equation 7.1).

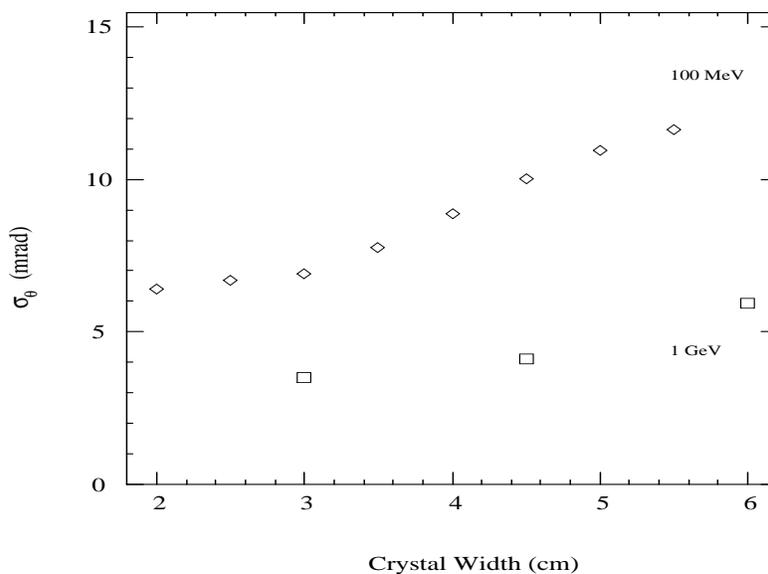
Material Thickness	$\sigma_E/E$ (%)	Efficiency (%)
0 $\mu\text{m}$	1.6	82
500 $\mu\text{m}$	2.1	79
800 $\mu\text{m}$	2.3	78

**Table 7-7.** Energy resolution and efficiency for 100 MeV photons with  $\cos\theta_{cm} = 0$  for different thicknesses of the inter-crystal material.

the cluster. The angular resolution for our choice of transverse crystal size is consistent with the target (Equation 7.2).

The effect of the CFC structural material and crystal wrappings on energy resolution and efficiency is given in Table 7-7 for photons with energies of 100 MeV and polar angles around  $60^\circ$ . Note that the baseline design has an amount of inter-crystal material equivalent to about 800  $\mu\text{m}$  of CFC. As can be seen, the performance degrades slowly for material thicknesses greater than 500  $\mu\text{m}$  of CFC equivalent.

Making the inter-crystal cracks nonprojective does not help with the loss of resolution shown in Table 7-7, but it does allow recovery of much of the lost efficiency. This has been studied in detail for the 800  $\mu\text{m}$  case. In both  $\theta$  and  $\phi$ , most of the loss is recovered with a nonprojective angle of 15 mr (about 6 times the angle subtended by the crack). The baseline design includes



**Figure 7-17.** Resolution in polar angle vs. transverse crystal size for 100 MeV and 1 GeV photons.

15 mr of nonprojectivity in  $\theta$  only, and relies on misalignment to produce a smaller amount of nonprojectivity in  $\phi$ . Even without the latter, the photon efficiencies are within 2% of the zero-material values; this loss is small compared to that resulting from material in front of the calorimeter.

## 7.4 Crystal Subassemblies and Readout

---

A crystal subassembly consists of a CsI(Tl) crystal covered with a suitable reflective material wrapped in a thin aluminum foil and a readout package located at the rear. The readout package consists of silicon PIN photodiodes closely coupled to low-noise, charge-sensitive preamplifiers, both enclosed in a metal cap. (The exact crystal dimensions are given in Section 7.6; for a detailed description of the front-end electronics refer to Chapter 10.) All components are selected to have a uniform and high level of light collection, together with minimal noise contribution from the readout electronics, in order to achieve optimal energy resolution.

### 7.4.1 Photodiode Readout

#### Signal-to-Noise Considerations

The requirement for the readout system is that at the lowest photon energies of interest ( $\sim 20$  MeV), the electronic noise must not dominate the resolution. In the absence of electronic noise, the resolution at these energies is expected to be on the order of 500 keV, mostly due to shower leakage fluctuations. The electronic noise contribution to the cluster energy resolution is required to be smaller than 500 keV. This results in an upper limit of about 150 keV per crystal, assuming that less than a dozen crystals are summed into a cluster at such low photon energies.

Electronic noise has two components, an incoherent one which derives from thermal and shot noise in the readout, and a coherent one which comes predominantly from pickup. Experiments using arrays of CsI(Tl) similar to the one proposed here have achieved incoherent noise performances of 250–500 keV per crystal at room temperature [Ake92, Kub92]. Noise performance at the 150 keV/crystal level should be achievable with improvements in average crystal light output, surface preparation, light collection, photodiode packaging, low-noise charge-sensitive preamplifiers, and shaper amplifiers. A target at this level has been set for the readout system. This meets the resolution requirements and allows the use of radioactive  $\gamma$  sources for crystal monitoring (with beams off). Note that lower noise per crystal implies improved energy resolution because a larger number of crystals per shower can be summed.

Shielding is needed for protection against coherent pickup noise. The entire calorimeter is enclosed in a metallic Faraday housing. The crystal housings are metallized to reduce cross coupling, and the diode-to-preamplifier connection uses shielded cable of minimum length. Both the diode and the preamplifier are encapsulated in a metal cap for electromagnetic shielding.

The equivalent electronic noise energy (ENE) for a crystal with a given readout system can be calculated as

$$\text{ENE} = \frac{\text{ENC}}{L(\text{pe}/\text{MeV})}, \quad (7.4)$$

where ENC is the equivalent noise charge in number of electrons, and  $L(\text{pe}/\text{MeV})$  is the light yield in electrons per MeV of energy deposition. Equation 7.4 indicates that achievement of low ENE requires not only low intrinsic noise of the electronics (ENC), but also a highly efficient photodetector coupled with maximal light output of the crystal. The resulting light yield requirements are detailed in Section 7.4.3.

Each crystal is read out by two independent photodiode-preamplifier-shaper channels [Hal95]. The two channels are then summed before digitization. The redundancy is due to reliability

requirements (Section 7.4.4), and the noise per crystal is the incoherent sum of the noise of the two channels.

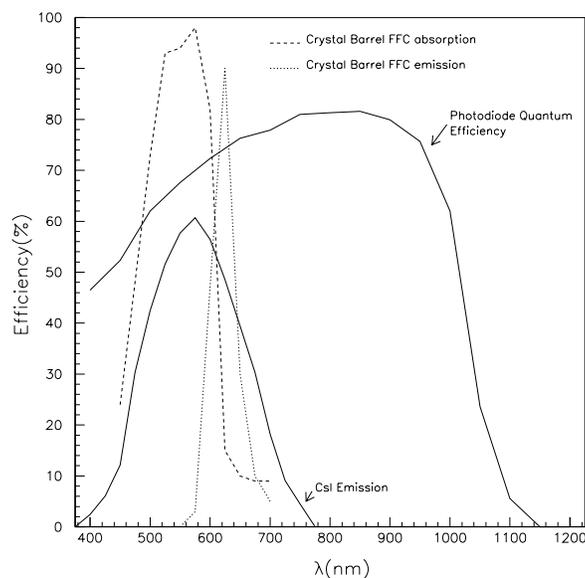
The principal contributions to the electronic noise are the shot noise from the dark current of the diode and the thermal noise of the preamplifier [Gro84]. The shaping time affects the relative contributions of the two components. Two possible readout schemes, one using two Hamamatsu S3588-03 photodiodes ( $3.4 \times 30 \text{ mm}^2$  active area) mounted on a wavelength shifter (WLS) and the other using two Hamamatsu S2744-03 photodiodes ( $10 \times 20 \text{ mm}^2$  active area) directly affixed to the crystal, are being considered. A preamplifier-shaper hybrid for both options has been developed and tested. The equivalent noise charge per crystal for the WLS option is  $620 e^-$  and for the direct option  $680 e^-$ , using a  $2 \mu\text{s}$  shaping time [Jes95a]. Note that a slight cooling of the diodes may further reduce the noise. The dark current of the photodiodes is halved by a reduction in temperature of as little as  $5\text{--}7^\circ\text{C}$ . Studies of noise *vs.* temperature are presently underway.

### Wavelength Shifter/Photodiode Readout

Typical silicon PIN photodiodes (*e.g.*, Hamamatsu S3588-03, S2744-03) show a broad spectral response peaking at wavelengths around 960 nm. They have a quantum efficiency of about 75%. The photon spectrum of CsI(Tl), on the other hand, has its peak emission around 560 nm. In addition, there is an order of magnitude mismatch between the light-emitting area of the crystal and the sensitive area of the photodiode. Thus, the purpose of the WLS is twofold: to shift the wavelength of the CsI(Tl) scintillation light to the more sensitive range of the photodiode, and to work as a flux concentrator by collecting the light onto the sensitive area of small, relatively inexpensive photodiodes through total internal reflection.

Fluorescent dyes are suitable materials for wavelength shifters [Lor86]. They show a large difference in absorption and emission wavelengths and re-emit the fluorescent light isotropically. The absorption of the dye should match the emission of CsI(Tl), and the re-emission of fluorescent light must take place at wavelengths as long as possible, with good internal quantum efficiency, small luminescent decay time, and low light loss. The overlap of the absorption and re-emission spectra of the wavelength shifter should be minimal in order to prevent self-absorption of luminescent light.

Typical wavelength shifters consist of thin plastic tiles doped with dye concentrations of a few hundred ppm [Lor86, Fis87]. The dyes used so far (BBQ, Y7, BASF Lumogen F red 339) have Stokes shifts of only  $\sim 50 \text{ nm}$  [Ake92, Kam83] and therefore show a large self-absorption of about 50% [Kei70]. Figure 7-18 shows the absorption and re-emission spectra of a wavelength shifter similar to that used by the Crystal Barrel [Ake92]. Transmission as high as 30% has been achieved with this material.

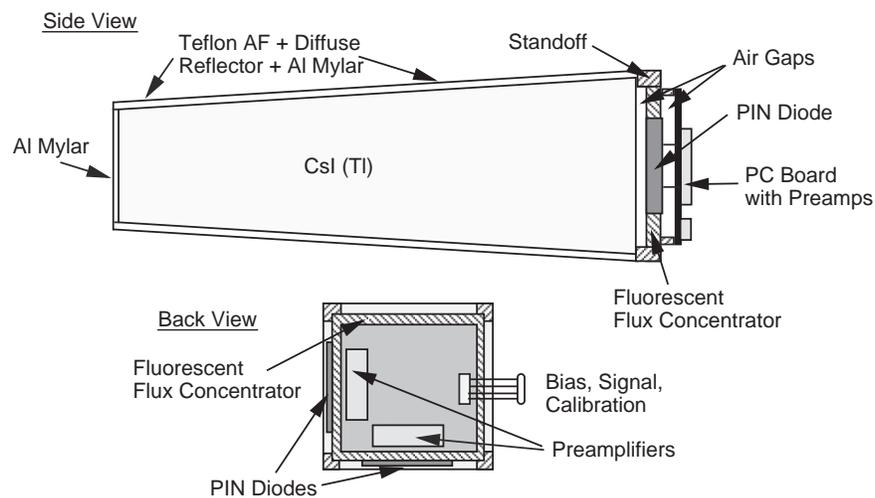


**Figure 7-18.** Absorption and fluorescence re-emission spectra of the Crystal Barrel wavelength shifter using BASF Lumogen F red 339 dye. Overlaid are the emission spectrum of CsI(Tl) and the photosensitivity of the Hamamatsu S3588 photodiode.

Fluorescent dyes that are used for Nd:YAG pumped dye lasers (*e.g.*, Styryl 8, Styryl 9M) absorb around 550 nm, re-emit above 750 nm, and have nearly no self-absorption. Acrylic material, such as PMMA, is used to carry the dye. These dyes must, however, be dissolved in a polar co-polymer before adding to the PMMA. Whether these dyes have high enough quantum efficiency and the optimal concentration for high absorption and high quantum efficiency are being investigated.

Figure 7-19 shows the proposed WLS setup with two PIN diodes. The WLS covers almost the entire rear face of the crystal. The re-emitted light is concentrated along the WLS side faces by total internal reflection and read out using small-area rectangular photodiodes (*e.g.*, Hamamatsu S3588-03,  $30 \times 3.4 \text{ mm}^2$ ). The diodes are attached to the WLS with Kodak HE-80 epoxy or Cargille Meltmount. The thickness of the WLS is primarily chosen to match the area of the photodiode. For redundancy, and in order to minimize light losses, two photodiodes are glued to two adjacent side faces of the WLS. To increase the light collection efficiency, a white diffuse reflector is placed behind the wavelength shifter, with an air gap to maintain total internal reflection inside the WLS. The other edges are polished and coated with white diffuse reflecting paint.

The light yield for a variety of crystals has been measured with the WLS/diode option [Jes95a]. For an  $18X_0$  tapered crystal with a front face  $4.5 \times 4.5 \text{ cm}^2$  and rear face  $5.8 \times 5.8 \text{ cm}^2$ , wrapped in three layers of  $38 \mu\text{m}$ -thick Teflon, a light yield of 4600 *pe*/MeV has been measured. The



**Figure 7-19.** *CsI(Tl) readout assembly (not to scale).*

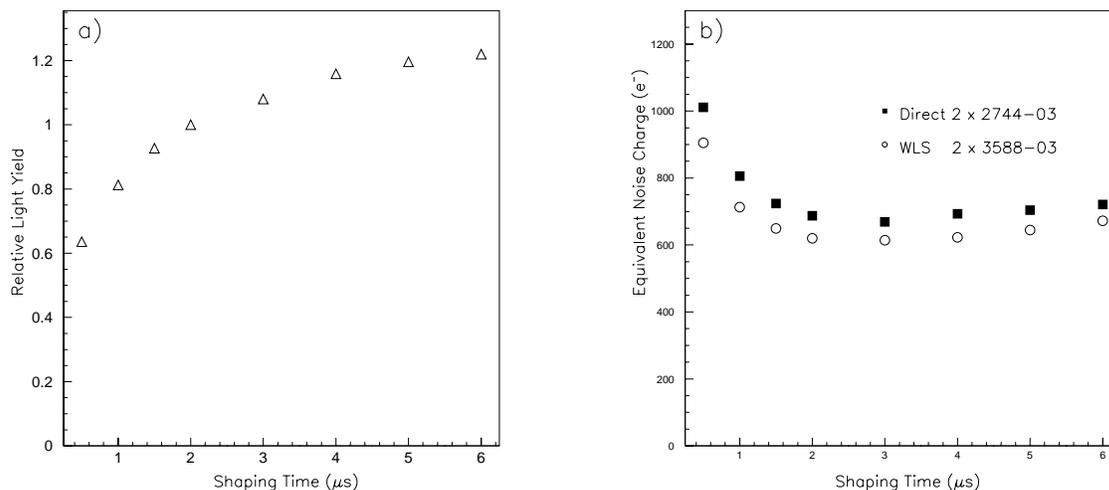
equivalent noise charge for the two diodes combined is  $620 e^-$ , which implies an equivalent noise energy of 135 keV per crystal. For comparison, the Crystal Barrel Collaboration has achieved an equivalent noise energy of 220 keV with a similar wavelength shifter geometry and one diode [Ake92].

Tests of possible radiation damage or aging effects of the WLS and the epoxy used for the WLS/photodiode bond must be completed. Radiation damage tests will be carried out using  $^{60}\text{Co}$  sources. There will also be long-term tests at normal and elevated temperatures.

### Direct Photodiode Readout

Photodiodes can, alternatively, be placed directly on the back of the CsI(Tl) crystal. To collect a sufficient amount of scintillation light, two large-area photodiodes (*e.g.*, Hamamatsu S2744-03,  $10 \times 20 \text{ mm}^2$ ) must be used. The balance of the back surface is covered with a diffuse reflector. This technique is more sensitive than the WLS readout to variations in light production and collection because the back surface is not fully covered by detectors. Direct readout with four small-area (Hamamatsu S1723,  $10 \times 10 \text{ mm}^2$ ) photodiodes is used in the CLEO-II calorimeter, which yields an equivalent noise energy of 500 keV [Kub92].

A variety of crystals have been measured using two Hamamatsu S2744-03 ( $20 \times 10 \text{ mm}^2$ ) diodes [Jes95a]. Using the same  $18X_0$  crystal as in the WLS test previously described, we measure a light yield of  $5960 \text{ } pe/\text{MeV}$ . The equivalent noise charge of the two diodes combined is  $680 e^-$  to give an equivalent noise energy of 114 keV per crystal. For comparison, the BELLE Collaboration has achieved 150 keV in similar tests [BEL94].



**Figure 7-20.** (a) Relative light yield vs. shaping time and (b) noise vs. shaping time for the direct and WLS readout schemes.

## Conclusions

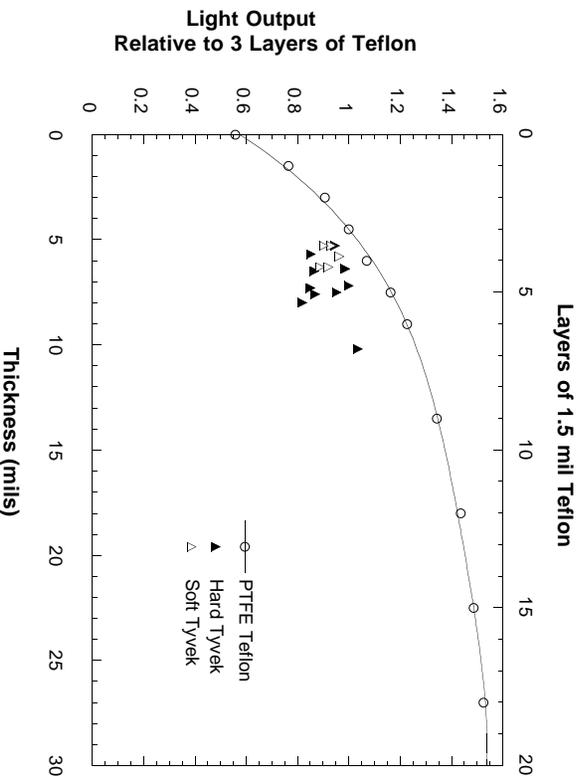
In bench tests, the WLS and direct readout options have achieved comparable performance. Further improvements in performance are anticipated with new WLS dyes and diode cooling. Once noise equivalents of about 150 keV per crystal using WLS readout are routinely achieved, this solution will be adopted. The overall complexity and reliability of such a readout system is comparable to that of direct diode coupling but has the advantage of substantially lower cost.

Note that all tests used a shaping time of 2  $\mu\text{s}$ . A lower value may be needed in order to reduce beam-related backgrounds [Lev94]. Figure 7-20 shows how to scale the light yield and noise for different values of the shaping time.

## 7.4.2 Light Collection

### Surface Preparation, Wrapping, Coating, and Tuning

Each crystal is cut and polished by the manufacturer. The light output of the crystal is required to be uniform along its length. This is traditionally accomplished by roughening the crystal surface selectively to reduce locally the reflected light and then wrapping the crystal in multiple layers of a suitable reflector, such as PTFE Teflon. The highest light yield is obtained by the use of a high-reflectivity white diffuse reflector with a small airgap



**Figure 7-21.** *Light output vs. number of layers of PTFE Teflon and DuPont Tyvek. (In each case, there is one cover layer of aluminum foil.)*

between the reflector and the crystal surface. The light yield increases with the number of Teflon layers used to wrap the crystal (Figure 7-21).

One potential wrapping scheme for the crystals uses four layers of PTFE Teflon, each layer being  $38 \pm 5 \mu\text{m}$  thick. This yields  $\sim 58\%$  of the available light with a total Teflon thickness of  $152 \pm 10 \mu\text{m}$ . To provide electronic shielding, each crystal is wrapped in addition with  $25 \mu\text{m}$  of aluminum foil.

A more efficient, automated approach employing Tyvek, a spunbonded olefin paper produced by DuPont, is under investigation. Tyvek has a high diffuse reflectivity similar to Teflon for the same thickness but is considerably more robust. A further advantage over Teflon is that Tyvek can be printed with a pattern of dots to selectively absorb light, thereby producing the required tuning pattern in an automated fashion. It is also possible to bond the  $25 \mu\text{m}$  aluminum shielding foil directly to the Tyvek.

Coating CsI crystals with clear Teflon AF as a sealant to reduce the effects of humidity on the crystal surface is also being evaluated. Teflon AF has a low index of refraction compared to CsI and transmission measurements at the peak wavelength of scintillation emission ( $560 \text{ nm}$ ) indicate that the amount of light leaving the end of the crystal increases by about 5% compared to an uncoated crystal. A 2.54 cm-diameter by 2.54 cm-height CsI cylinder with diamond-polished end faces was coated with Teflon AF, and its transmission was measured. The crystal was then placed in a room-temperature chamber with 100%

humidity for three days. No change was observed in the transmission of the crystal after this exposure. The radiation hardness of Teflon coatings has been studied, and no change in optical properties for exposures of 40 krad of  $^{60}\text{Co}$   $\gamma$ -rays has been observed.

Water-based cross-linkable fluorocarbon (WXF) offers the same performance as Teflon AF but with a simpler dipping or brushing application and a lower temperature curing step. WXF can be applied in thicker coatings as well.

Both Teflon AF and WXF can be loaded with white  $\text{TiO}_2$  powder to provide a highly reflective, diffuse coating. We are studying the application of a clear coating (to mimic a low index of refraction air layer) followed by a white diffuse coating (to mimic a wrapping of PTFE film, for example). This coating system offers greater control of thickness than is obtainable with wrappings.

### Uniformity Specifications

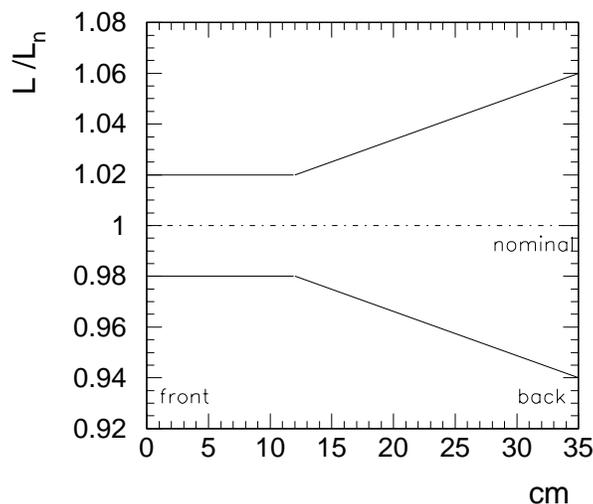
The simulation results of Section 7.3.1 allow specification of light output uniformity. The nonuniformity contribution must be well below the target energy resolution (Equation 7.1). Requiring a contribution to  $\sigma_E/E$  below 0.5% translates into an integral variation of the light output between  $-6\%$  and  $+6\%$  from the front end to the back end of the crystal, as shown in Figure 7-22. Variations near shower maximum are the most critical. However, short distance variations of 2% in the front third of the crystals are tolerable (Figure 7-22). A procedure for measuring the light output uniformity is described in Section 7.4.3.

### 7.4.3 Light Yield Measurements

A number of crystals having dimensions similar to our baseline design were obtained from potential vendors and the light yields measured [Jes95a]. The readout techniques, shaping times, polishing, wrapping, and preparation are close to that of the baseline design.

The tests were made with a wrapping of three layers of PTFE Teflon film (each  $38\ \mu\text{m}$  thick) with one layer of aluminum foil outside. The light yield has been measured in three ways:

1. A 3.4 mm-thick WLS and two  $30 \times 3.4\ \text{mm}^2$  S3588-03 diodes was epoxied to adjacent edges. Two standard calorimeter preamps were used, with the WLS positioned 0.5 mm from the back face of the crystal.
2. Two  $10 \times 20\ \text{mm}^2$  S2744-03 diodes were directly coupled to the crystal using optical grease. Two standard calorimeter preamps were used.



**Figure 7-22.** Upper limits for the light output nonuniformity in a 35 cm-long CsI(Tl) crystal, assuming no more than 0.5% contribution to  $\sigma_E/E$  for all photon energies up to 5 GeV.

3. A two-inch photomultiplier tube (Hamamatsu R669) with an extended-red multi-alkali photocathode (peak sensitivity at 650 nm) was positioned 0.5 mm from the crystal back face. The light yield is measured relative to that from a small calibration crystal, since this will be the procedure used for vendor specifications.

For the photodiode readout, the output signal is shaped using a  $2\ \mu\text{s}$  shaping time. Figure 7-20 shows how the relative light yield varies with shaping time. The absolute calibration of electron number per ADC channel for the diode is obtained by using a 59.5 keV x-ray from a  $^{241}\text{Am}$  source, taking into account that 3.62 eV is required to produce one electron-hole pair in silicon.

### Light Yield Specification

The light yield of CsI(Tl) crystals is directly related to the resolution and the energy-equivalent electronic noise. To ensure that the calorimeter can achieve the required performance, every crystal must satisfy a light yield specification supplied to the crystal manufacturers.

The light yield is sensitive to the thallium iodide concentration in the crystal, the light attenuation length, the shape and size of the crystal, the surface preparation, the wrapping, and the position of the source which introduces the scintillation. It is necessary to provide a practical and economical measurement procedure for the vendor. Therefore, the light yield

Vendor	Front Dim. (cm <sup>2</sup> )	Rear Dim. (cm <sup>2</sup> )	Length (cm)	LY Rel. to Std. (%)	LY Direct (pe/MeV)	LY WLS (pe/MeV)
Kharkov	25	25	34	32	6340	4200
Kharkov	34	20	34	32	5960	4600
Crismatec	28	18	23	27	4820	3650
Crismatec (hex)	28	18	22	31	5300	4250
Horiba	28	18	23	30	8600	5560
Horiba (hex)	28	18	23	50	9800	6800
Polsyscine	34	20	34	35	6500	4550

**Table 7-8.** Light yield (LY) measurements (as described in the text) for full-sized crystals [Jes95a]. All crystals have a square cross section except where indicated.

specification is given in terms of light yield relative to a small (*e.g.*, 2.54 cm diameter  $\times$  2.54 cm height) standard CsI(Tl) crystal, measured using  $\gamma$  lines from a radioactive source such as <sup>22</sup>Na, and a red-sensitive PMT. Light yield for a full-sized crystal can be defined as the average of a set of measurements taken at different source positions along the length of the crystal; such measurements are also used in determining the crystal's light output uniformity.

Assuming the WLS readout option, with 620  $e^-$ 's ENC for each crystal, a light yield of 620/0.15 MeV  $\approx$  4100 photoelectrons/MeV is required to obtain 150 keV incoherent noise. From Table 7-8, a relative light yield of  $\geq 30\%$  is therefore required from each wrapped crystal.

#### 7.4.4 Reliability of Inaccessible Readout Components

The barrel crystals and their front-end readout up to the ADC boards are inaccessible, barring disassembly of the detector. It is anticipated that this will not take place until ten years after the commissioning of the detector. Endcap crystals are relatively accessible and can be repaired on a much shorter time scale. We therefore require that no barrel crystal should fail in ten years. Achieving this reliability requires screening for defective parts before assembly and establishing an adequate natural lifetime for the remaining components.

The components most likely to fail are the connectors, bonds, and electronic components. Mechanical and thermal stresses and chemical aging may all be contributory factors. A stable thermal and humidity environment can reduce these effects.

A burn-in technique is being developed to screen all components, while random samples are collected to estimate the natural lifetimes by accelerated aging tests. Screening involves stressing components to remove those subject to infant mortality.

To establish the natural lifetime of components is difficult. It is empirically observed that, in the limit of a large number of components, each with a natural lifetime much greater than the operational lifetime, the failure rate follows an exponential distribution with a time constant  $1/\text{MTBF}$ , the mean time between failures. This time is typically  $\geq 100$  years for electronics components, implying that a joint MTBF in excess of 60,000 years must be established if one requires less than one failure in 6000 crystals in 10 years. To establish this MTBF with certainty requires testing 6000 components for 10 years, which is not feasible. There are three techniques available to estimate the MTBF:

1. Assume aging is proportional to some power of temperature  $T^\alpha$  (the Arrhenius rate reaction formula). Test a number of components at high temperature and scale to the operating temperature. The power  $\alpha$  can only be estimated from similar studies.
2. Use published formulae for estimating the failure rates of components based on experience from, *e.g.*, the military [Mil90].
3. Extrapolate failure rates from similar experiments such as CLEO-II, which operates 32,000 photodiodes and preamplifiers under similar conditions.

Adopting method (3), Table 7-9 lists cumulative diode and preamplifier failures and noisy diodes over five years of operation. The initial 48 diode failures are attributed to connector failures or insufficient screening. Considering all 48 plus the 14 noisy diodes as failures, the preamplifier appears far more likely to fail than the diode.

Using Poisson statistics, at the 95% confidence limit, one finds  $\text{MTBF}_{\text{diode}} \geq 1954$  years,  $\text{MTBF}_{\text{preamp}} \geq 832$  years and combined,  $\text{MTBF}_{\text{diode+preamp}} > 643$  years. The WLS option uses diodes of similar area and manufacture, implying 93 failures per 6000 crystals after 10 years if a single diode/preamp readout is chosen. Table 7-10 shows the number of failing crystals as redundancy is increased, assuming  $\text{MTBF}_{\text{diode+preamp}} > 643$  years. Twofold redundancy yields 308 crystals with one dead channel and two crystals with two dead channels after 10 years [Jes95b]. A crystal with one dead channel has half the light yield and  $1/\sqrt{2}$  the electronic noise in ENC, hence  $\sqrt{2}$  higher equivalent noise energy.

Time (Months)	Diode Failures	Noisy Diodes	Preamplifier Failures
0	48	0	17
5	48	2	58
12	48	7	85
24	48	12	100
32	48	12	112
38	48	12	119
50	48	13	128
60	48	14	142

**Table 7-9.** Cumulative diode and preamp failures for the CLEO-II calorimeter from 1989 to 1994.

Failed Channels	1 Diode 1 Preamp	2 Diodes 2 Preamps	3 Diodes 3 Preamps	4 Diodes 4 Preamps
0	5907	5815	5725	5635
1	93	184	271	356
2		1	4	9
3			0	0
4				0

**Table 7-10.** Number of failed channels on a crystal for various levels of redundancy, after 10 years, assuming the MTBF of (diode+preamp) is greater than 643 years.

## 7.5 Calibration

### 7.5.1 Requirements and Ingredients

Several methods will be used in calibrating:

1. Electronics calibration will be done at least weekly with beams off. The  $\sim 2000$  events needed per channel can be collected in several minutes. Calibration will entail the injection of charge at the preamplifier input, with controllable amplitude, phase (with respect to digitization), and pattern, using the system described in Chapter 10, in order to determine the detailed response function for each channel of electronics.

2. Bhabha events will be used for tracking overall gains. Enough data for this can be obtained on a timescale of well under a day.
3. Photons and electrons of known energy, as well as minimum ionizing tracks, will be used to set the energy scale (gain), including inter-channel variation of crystal light output and optical properties of the photodiode (WLS) readout.
4. Radioactive  $\gamma$  sources will provide low-energy point(s) for the response of individual crystals. Energy deposited in the front of a single crystal can be compared to Bhabhas in order to track optical changes. The source spectrum is a direct monitor of noise in energy units.

Items (2) and (4) also allow us to track changes over time of the entire optical path, including possible radiation damage or environmental changes. CLEO-II has achieved a calibration accuracy of  $\pm 0.2\%$  at  $\sim 5$  GeV increasing to  $\pm 1\%$  at 25 MeV, and a time-dependent gain stability of under  $\pm 0.25\%$  [Kub92]. *BABAR* can at least equal this performance and improve on the accuracy at low energies because lower electronic noise allows the use of radioactive sources. Although it is not part of the baseline design, a fiber optic distribution of light pulses is still under consideration for monitoring crystals.

The calibration system will be used to set initial gains for all channels to within  $\sim 10\%$ , to provide approximate real-time calibration factors for forming trigger sums and to provide online diagnostics.

### 7.5.2 Energy Calibration with Beam Events

At design luminosity, the total rate of Bhabha coincidences (above 300 mr) is  $\sim 92$  Hz, with at least 28 Bhabha  $e^\pm$  per hour incident on each crystal. Since the mean fraction of energy in the peak crystal is  $\sim 0.7$ , at least half of these events should be usable for calibration with either an iterative procedure or a fit. The goal of  $\leq 0.25\%$  precision requires about 40 usable events in a crystal, which can be achieved in under three hours. For polar angles backwards of  $133^\circ$ , Bhabha singles must be used. A special trigger can prescale such events as a function of polar angle to obtain the sample.

The crystal response *vs.* energy must be mapped using photons of known or constrained energy. The process  $e^+e^- \rightarrow \gamma\gamma$  provides maximum-energy data at each polar angle. In  $e^+e^-\gamma$  and  $\gamma\gamma\gamma$  final states, the energy of the lower energy photon can be computed from event topology. Radiative Bhabha yields have been calculated for a minimum separation of  $15^\circ$  between the photon and each lepton. At design luminosity, the usable yield of photons above 1 GeV ranges from 10 to 300 per hour per GeV per  $10^\circ$  of polar angle, varying across photon phase space. Lower energy photons can be obtained from  $\pi^0 \rightarrow \gamma\gamma$  decays; on the

order of one  $\pi^0$  per  $B\bar{B}$  event is expected to be usable. These are the methods by which CLEO-II presently covers the photon energy range of interest.

High-energy minimum ionizing tracks deposit most of their  $dE/dx$  loss, averaging  $\sim 200$  MeV, in individual crystals, providing a useful high-statistics comparison to Bhabhas. The yield from the  $\mu^+\mu^-$  final state alone is at least 0.7 per hour per crystal, but non-interacting  $\pi^\pm$  with momentum greater than  $2.1$  GeV/ $c$ , which traverse  $\leq 2/3$  of a crystal in azimuth, will also be useful.

### 7.5.3 Source Calibration

Radioactive  $\gamma$  sources are included in the baseline design for several reasons. First, since the energy is deposited at the front of a crystal, comparison to Bhabhas will help identify the origin of a change in crystal gain, *e.g.*, radiation damage or other change in light collection. Second, a source spectrum provides a direct monitor of noise in energy units. Third, sources are the easiest way to make initial gain settings. Fourth, sources provide low-energy points for the responses of individual crystals to actual photons. Fifth, integrated sources do allow continual system checks after installation and whenever the accelerator is not running. Source energies of 2 MeV or higher are desirable to allow for cases of bad electronic noise and to better mimic photons of physics interest.

Source runs will be performed with beams off using a special processing mode which involves a search for signal waveforms in video RAM memory and the extraction of their amplitudes. A source activity of  $0.01$   $\mu\text{C}$  for one crystal implies 1000 usable events in ten seconds. An actual source run will be longer, due to processing limitations, but still well under 20 minutes.

Three options are under study:

1. Mounting  $0.01$   $\mu\text{C}$   $^{60}\text{Co}$  sources ( $t_{1/2} \approx 5$  years) at the front face of each crystal and using the 2.51 MeV sum line. For adequate efficiency, this may entail inserting a source at the end of a fine needle into a very narrow hole drilled several centimeters into the front of the crystal. The feasibility of this approach is being evaluated.
2. Pneumatically inserting several  $100$   $\mu\text{C}$  sources at  $z = 0$  just outside the support tube or within the PID enclosure. A good candidate is  $^{208}\text{Tl}$ , which has a 2.614 MeV  $\gamma$  line, driven by an  $\alpha$  decay chain from  $^{228}\text{Th}$  ( $t_{1/2} \sim 2$  years) or  $^{232}\text{U}$  ( $t_{1/2} \sim 100$  years). Encapsulation prevents  $\alpha$  particles or radon from escaping.
3. Circulating about  $1000$   $\text{cm}^3$  of liquid, activated in a shielded neutron source located outside the detector, through fine tubes running in front of all crystals. Choices include:
  - (a) A fluorine-containing fluid, with a metastable state of  $^{16}\text{O}$  via  $^{19}\text{F}$  ( $n,\alpha$ )  $^{16}\text{N}$ ;

	Barrel ( cm)	Forward Endcap ( cm)
Inner Gas Seal	0.5	3.0
Crystal Box Extension	0.1	0.7
Inward Dimension Tol.	0.6	0.6
Crystal Length	29.76–32.55	32.55
Outward Dimension Tol.	0.6	0.6
WLS/Diode Package+Spacers	0.4	0.4
Gap to Preamplifier	0.1	0.1
Preamplifier/Header	2.6	2.6
Cooling Manifold	1.5	1.2
Module Insertion Fixture/Cables	2.5	
Module Strongback	1.6	
Alignment Jacks	2.5	0.8
Aluminum Strongback /Frame	3.2	2.0
Total	46.0	44.5

**Table 7-11.** Thickness of the calorimeter assembly, barrel, and endcap.

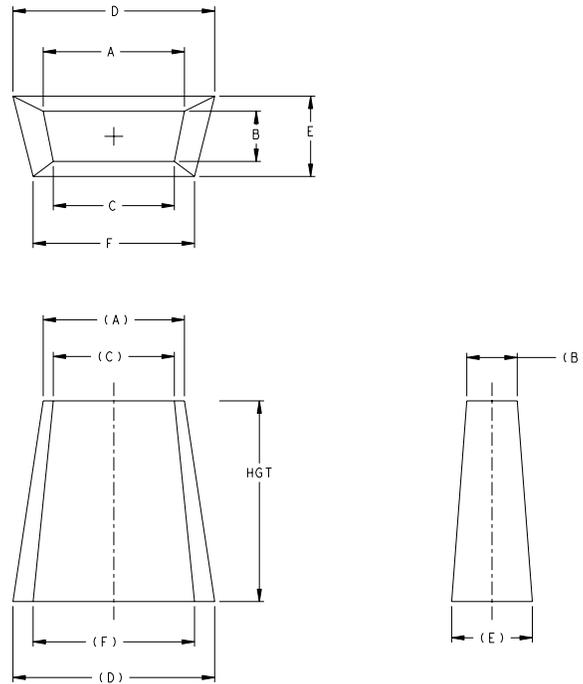
neutrons of  $\sim 4\text{-}6$  MeV from a Po-Be source are appropriate. The decay provides a 6.13 MeV  $\gamma$  line, with  $t_{1/2}$  of only 7sec. This choice has been assessed as posing no significant radiation safety hazard. (b) Water, with  $^{16}\text{O}$  (n,p)  $^{16}\text{N}$  leading to the same  $\gamma$  line. Activation could be by 14 MeV neutrons from a DT source. By using saltwater and a moderated neutron source, one could additionally provide 2.75 MeV and 1.37 MeV  $\gamma$  lines (with  $t_{1/2}$  of 15 hours) via  $^{23}\text{Na}$  (n, $\gamma$ )  $^{24}\text{Na}$ .

## 7.6 Mechanical Support Structure

---

### 7.6.1 Design Considerations

The mechanical structure of the *BABAR* calorimeter is designed to minimize inactive material between individual crystals, while providing hermeticity at the barrel-endcap interface. Access, testing, repair, and reliability are addressed in the design of the overall structure, as are the crystal loading procedures, and installation and system integration. RF shielding and cooling of electronics, cable routing, and humidity control considerations are also discussed in the following sections. Table 7-11 details the radial dimensions of the barrel and forward endcap.



**Figure 7-23.** *Crystal dimensions (Tables 7-12 and 7-13).*

Since the detector is located in a seismically active region, safety standards require that the calorimeter be designed to withstand accelerations of up to 2 g horizontally and 1 g vertically, without significant damage, both during the assembly period and after installation.

**Crystal Loading, Sizes, and Tolerances.** CsI(Tl) is grown in polycrystalline form and is a soft, deformable material that cannot support substantial loads. Therefore, the structure must support individual crystals without transferring the load through neighboring crystals.

The crystal sizes and counts for the barrel and endcap are given in Tables 7-12 and 7-13. The columns in the tables give the dimensions of the crystal edges as shown in Figure 7-23.

Crystals are required to fit loosely into their barrel or endcap compartments. Crystal dimensional tolerances are summarized in Figure 7-24. The tolerance on crystal transverse dimensions ( $\pm 225 \mu\text{m}$ ) allows the cost of cutting and polishing to be minimized. Consequently, an additional radial space of 0.5 cm is required, as indicated in Table 7-11.

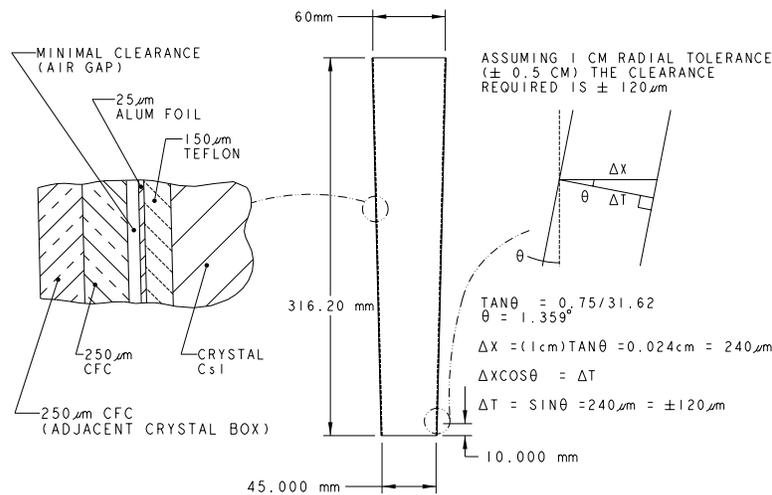
**Wall Material and Thickness, and Inter-Crystal Material.** Properties of structural materials under consideration for the compartment walls are given in Table 7-14.

$\theta$ Row	Number Needed	Volume (cc)	A (cm)	B (cm)	C (cm)	D (cm)	E (cm)	F (cm)	Height (cm)
1	120	882.5	4.956	4.773	4.736	5.779	5.523	5.524	32.55
2	120	888.4	4.953	4.773	4.736	5.812	5.559	5.559	32.55
3	120	894.6	4.950	4.773	4.736	5.846	5.595	5.594	32.55
4	120	901.0	4.947	4.773	4.736	5.880	5.633	5.631	32.55
5	120	907.6	4.943	4.773	4.736	5.916	5.672	5.669	32.55
6	120	914.4	4.939	4.773	4.736	5.952	5.712	5.708	32.55
7	120	921.3	4.935	4.773	4.736	5.988	5.753	5.748	32.55
8	120	896.7	4.930	4.773	4.736	5.994	5.766	5.759	31.62
9	120	903.5	4.925	4.773	4.736	6.030	5.808	5.800	31.62
10	120	910.4	4.919	4.773	4.736	6.065	5.850	5.841	31.62
11	120	917.3	4.913	4.773	4.736	6.101	5.893	5.882	31.62
12	120	924.4	4.907	4.773	4.736	6.136	5.936	5.924	31.62
13	120	931.4	4.900	4.773	4.736	6.171	5.980	5.965	31.62
14	120	938.3	4.892	4.773	4.736	6.204	6.023	6.007	31.62
15	120	910.6	4.884	4.773	4.736	6.197	6.028	6.009	30.69
16	120	916.9	4.876	4.773	4.736	6.226	6.069	6.048	30.69
17	120	922.9	4.867	4.773	4.736	6.254	6.108	6.086	30.69
18	120	928.7	4.857	4.773	4.736	6.279	6.147	6.123	30.69
19	120	934.2	4.847	4.773	4.736	6.302	6.184	6.158	30.69
20	120	939.3	4.837	4.773	4.736	6.322	6.220	6.191	30.69
21	120	943.9	4.826	4.773	4.736	6.339	6.253	6.222	30.69
22	120	911.3	4.814	4.773	4.736	6.306	6.237	6.204	29.76
23	120	914.6	4.802	4.773	4.736	6.315	6.264	6.228	29.76
24	120	917.2	4.790	4.773	4.736	6.320	6.287	6.248	29.76
25	120	919.2	4.778	4.773	4.736	6.321	6.307	6.266	29.76
26	120	920.6	4.765	4.773	4.736	6.318	6.322	6.279	29.76
27	120	921.2	4.752	4.773	4.736	6.310	6.334	6.289	29.76
28	120	882.4	4.740	4.773	4.736	6.299	5.896	6.294	29.76
29	120	882.4	4.740	4.773	4.736	6.299	5.896	6.294	29.76
30	120	921.2	4.752	4.773	4.736	6.310	6.334	6.289	29.76
31	120	920.6	4.765	4.773	4.736	6.318	6.322	6.279	29.76
32	120	919.2	4.778	4.773	4.736	6.321	6.307	6.266	29.76
33	120	917.2	4.790	4.773	4.736	6.320	6.287	6.248	29.76
34	120	914.6	4.802	4.773	4.736	6.315	6.264	6.228	29.76
35	120	911.3	4.814	4.773	4.736	6.306	6.237	6.204	29.76
36	120	907.5	4.826	4.773	4.736	6.293	6.208	6.177	29.76
37	120	903.2	4.837	4.773	4.736	6.277	6.176	6.147	29.76
38	120	898.4	4.847	4.773	4.736	6.258	6.142	6.115	29.76
39	120	893.4	4.857	4.773	4.736	6.236	6.106	6.081	29.76
40	120	888.0	4.867	4.773	4.736	6.212	6.068	6.045	29.76
41	120	882.3	4.876	4.773	4.736	6.185	6.029	6.008	29.76
42	120	876.5	4.884	4.773	4.736	6.157	5.990	5.971	29.76
43	120	870.5	4.892	4.773	4.736	6.127	5.949	5.932	29.76
44	120	864.4	4.900	4.773	4.736	6.096	5.909	5.893	29.76
45	120	858.3	4.907	4.773	4.736	6.064	5.868	5.854	29.76
46	120	852.2	4.913	4.773	4.736	6.031	5.827	5.815	29.76
47	120	846.1	4.919	4.773	4.736	5.998	5.787	5.776	29.76
48	120	840.0	4.925	4.773	4.736	5.965	5.747	5.737	29.76
49	120	834.1	4.930	4.773	4.736	5.931	5.708	5.699	29.76

**Table 7-12.** Crystal dimensions and counts for the barrel. Dimensions A through F are defined in Figure 7-23.

$\theta$ Row	Number Needed	Volume (cc)	A (cm)	B (cm)	C (cm)	D (cm)	E (cm)	F (cm)	Height (cm)
1	80	741.0	4.298	4.670	3.944	4.995	5.426	4.583	32.55
2	80	802.4	4.641	4.670	4.289	5.394	5.428	4.985	32.55
3	80	863.6	4.982	4.670	4.632	5.792	5.430	5.386	32.55
4	100	739.6	4.247	4.670	3.970	4.941	5.432	4.618	32.55
5	100	788.3	4.519	4.670	4.244	5.257	5.433	4.937	32.55
6	100	836.9	4.790	4.670	4.517	5.573	5.433	5.256	32.55
7	120	737.7	4.209	4.670	3.984	4.898	5.434	4.636	32.55
8	120	778.4	4.437	4.670	4.214	5.162	5.433	4.903	32.55
9	120	819.1	4.665	4.670	4.445	5.427	5.432	5.171	32.55

**Table 7-13.** Crystal dimensions and counts for the endcap. Dimensions A through F are defined in Figure 7-23.



**Figure 7-24.** Crystal and wrapping tolerances.

Wall Materials	Density (g/cm <sup>3</sup> )	dE/dx (MeV/g/cm <sup>2</sup> )	X <sub>0</sub> (g/cm <sup>2</sup> )
Carbon Fiber	1.68	1.90	45.2
Glass Fiber	1.70	1.87	33.0
Aluminum	2.70	1.62	24.0

**Table 7-14.** Structural-material properties (assumes 40% epoxy and 60% fiber loading for fiber materials).

CFC with 250  $\mu\text{m}$  wall thickness has been chosen for crystal compartments because it has the best combination of strength and minimum material ( $X_0$ ). There are two such walls between crystals, resulting in a total amount of structural material between crystals of  $2 \times 0.09\% X_0$ . Mechanical tolerances of the CFC boxes that form the support structure can be made small because the boxes are formed on accurate mandrels and assembled in precision trough fixtures. (The fabrication technique in the endcap places a different set of requirements on the dimensions.)

Inactive material between crystals also includes reflective wrappings, shielding foils, and airgaps due to mechanical tolerances. This adds another  $2 \times 0.07\% X_0$ , resulting in a total of  $2 \times 0.16\% X_0$  between crystals.

**Nonprojectivity.** To minimize unmeasured energy due to inactive material between crystals, the crystals are arranged to be slightly nonprojective in  $\theta$  and projective in  $\phi$  with respect to the IP. The degree of nonprojectivity is set by the requirement that straight-line tracks cannot be completely contained in the gaps between individual crystals. Simulations show that degradation of reconstruction efficiency begins for photons that traverse more than one fourth of the gap length before entering a crystal. To prevent this degradation, a fixed  $\theta$  offset of 15 mr is added to the projective angle for a given crystal and this offset is applied to each crystal in the barrel and endcap. The crystals then project to a point that varies between 2 mm and 5 cm from the IP along the beam line. To minimize the number of unique crystal shapes, the crystal gap is centered projectively at  $90^\circ$  with respect to the beam line. The front faces of the crystals are uniform in width and perpendicular to the crystal axis. In  $\phi$ , the crystals are arranged so that the center lines of their compartment gaps project to the beam axis.

The crystal row of the last barrel and first endcap at the forward barrel/endcap interface require special treatment because of the larger gap in this region. They will be shaped so that the barrel/endcap gap is sufficiently nonprojective.

**RF Shielding of the Electronics.** The barrel and endcap are constructed as separate Faraday cages. The inner rf-shields are provided by a laminate of two layers of 1 mm-thick aluminum (each about four skin-depths at 100 kHz) separated by a thin foam core. The interface region between the inner and outer faces of the barrel and forward endcap are tied together with flexible metallic fingers to shield the 3 mm gap. Shielding for inter-crystal cross-talk and rf noise from the preamplifiers and photodiodes of an individual crystal is provided by aluminum enclosures attached to the  $25\ \mu\text{m}$  aluminum wrap of the crystal.

**Environmental Control for Crystals and Electronics.** The calorimeter will be kept at an appropriate constant temperature in order to keep the leakage current of the photodiodes below 2 nA, to maintain stable preamplifier performance, and to keep a constant CsI(Tl) scintillation light yield.

The two preamplifiers for each crystal generate about 100–250 mW, giving a total heat load inside the rf-shield of about 2.5 kW. Heat flow calculations on a model crystal assembly with the conservative assumption of 500 mW total power generated at the electronics indicate that active cooling is required. Thermal control to better than  $\pm 1^\circ\text{C}$  is achieved by flowing coolant into the calorimeter along circumferential manifolds into pipes running axially along the crystal rows. Each crystal's electronic shield is attached by braid to the manifold. The cooling system output is cycled through an external heat exchanger. Digitizing electronics located in shielded enclosures at the end of the barrel calorimeter housing and the outer rear of the endcap are cooled using the same system.

The outer surfaces of the calorimeter form a gas enclosure for humidity control. Dry nitrogen flowing through the volume at a slow rate and slight overpressure is used to maintain a dry atmosphere to prevent degradation of the crystal surfaces.

**Signal and Control Routing.** Signal and control cables are passed through the Faraday/environmental shield using sealed bulkheads containing signal feedthroughs. Signal cable lengths between preamplifiers and bulkheads are minimized, with the longest run being approximately 1.5 m. A standard 12-pair shielded cable will be used.

Digitizing electronics cards are attached to the feedthrough connectors in the bulkheads. Fiber-optic cabling from the digitizer cards of both the barrel and forward endcap is routed outward between the barrel and endcap flux return iron.

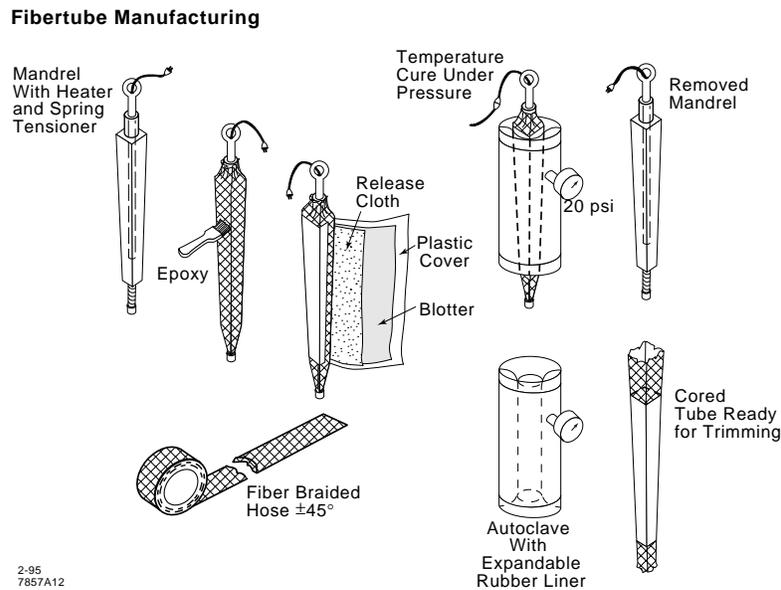
**Access and Component Replacement.** Individual barrel crystals and their readout packages are not accessible once the calorimeter has been installed in the solenoid. However, the modular design facilitates testing and replacement of diode/preamp packages before this installation. The endcap can be removed for internal maintenance. All calorimeter digitizer cards can be readily accessed for repair by opening the endcap doors.

## 7.6.2 Barrel Fabrication, Assembly, and Installation

### Barrel System

In order to avoid the transmission of loads through the crystals, the barrel cylinder is divided into independently supported and aligned modules which have little or no mechanical contact with each other. Their light weight limits deflections. The structure is composed of CFC elements between crystals in which minimum material is necessary, and aluminum plates in which structural strength is required and more material can be tolerated.

The basic units are the CFC compartments (boxes) that house individual crystals. The CFC boxes are combined into modules containing  $3 \times 7$  crystals, each weighing approximately 80 kg. Figure 7-27 shows the basic module structure. The individual compartments are stiffened at their rear (outer) ends with a ring that also corrects for tolerances in the crystal placement and locks the crystal longitudinally into position in their boxes. The boxes are joined to a strongback made from thin aluminum side plates (tabs) and an extruded aluminum U-channel fastened to the rear of the modules. The module strongbacks transmit loads to a large cylindrical aluminum tube with stiffening flanges at each end, into which the modules are inserted and aligned. Both the module strongback and the cylindrical tube have access ports to allow for assembly, alignment, cabling, and repair prior to insertion into



**Figure 7-25.** Schematic of tube fabrication by a wet-layup technique.

the magnet coil. After insertion into the coil, the tube transmits its load to the coil and flux return through a two-point mounting scheme on each end flange.

## Module Fabrication

The choice of module size is made by considering ease of handling (without requiring large cranes or special fixtures within dry-rooms), access to elements within, size and complexity of fabrication fixtures, and the deflections and strengths of the elements. The delivery schedule of each crystal type, and its influence on the assembly schedule, is also considered. There are 280 modules in the barrel, 40 in azimuth and 7 along the beam axis.

Each CFC box is fabricated using a conventional wet (epoxy-resin) layup technique, as illustrated in Figure 7-25. This technique allows reproducible dimensions of the inside box. The final wall thickness is the only variable dimension. For each crystal size, an aluminum master mandrel is precision machined to the size of a wrapped crystal (*i.e.*, crystal dimension plus  $\sim 175 \mu\text{m}$ ). The mandrel contains a heating and thermostatic control element for curing the epoxy. The mandrel blanks may be fabricated by casting.

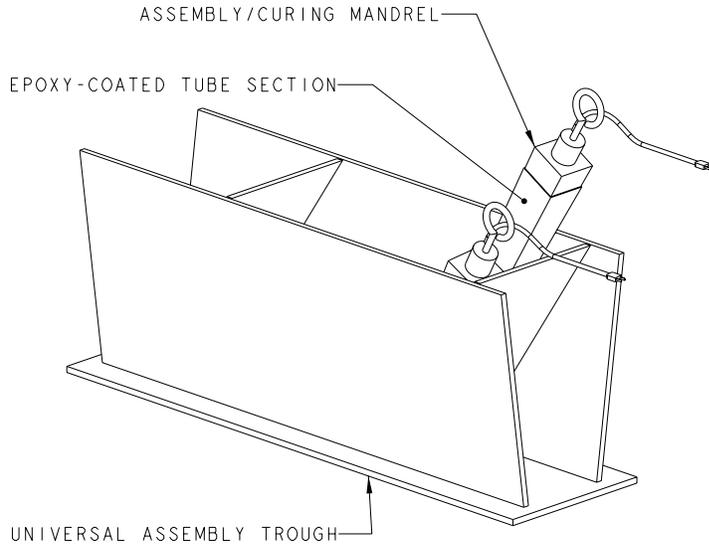
To fabricate a CFC box, the mandrel is first sprayed with a release agent such as Teflon. A woven tube, or sock, of CFC is stretched over the mandrel. The weave is chosen at  $\pm 45\text{--}60^\circ$  to accommodate the taper of the crystal. Socks will be custom fabricated to a finished

thickness of  $250\ \mu\text{m}$ . These contain 400 strands per inch of  $25\ \mu\text{m}$  filament. The mandrel will be covered with a  $25\ \mu\text{m}$  aluminum foil electronic shield before placing the sock over the mandrel. The sock is next uniformly impregnated with thermal-setting epoxy resin. The mandrel and sock are wrapped together in peel-ply, bleeder cloth, and a protective polyvinyl sheet. The layup is cured under 2 atm pressure applied uniformly by a simple fixture at  $40^\circ\text{C}$  using a heater in the mandrel. After the  $\sim 2$  hour cure, the tube is removed from the mandrel and trimmed to shape and length. The mandrel and tooling are reused. Tests performed over the last three years (Figure 7-28) using CFC boxes of similar weave fabricated with this technique and with several commercial epoxy systems indicate that the dimensional creep of such boxes is acceptable. Small factories can be set up at various collaborating institutions to fabricate the 5880 elements for the calorimeter barrel.

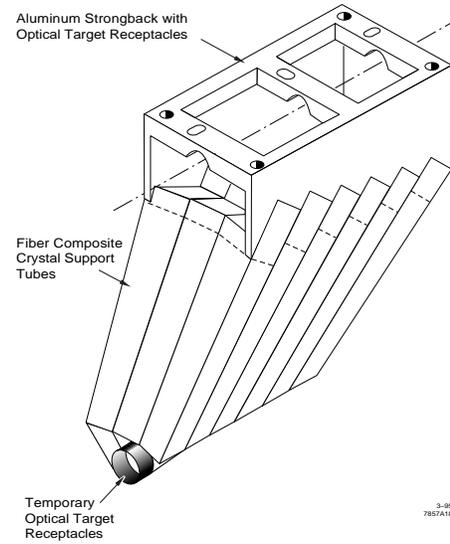
There are seven distinct module types along the length of the barrel, for which 40 modules in  $\phi$  (plus one or two spares each) are required. The boxes are joined into  $3 \times 7$  arrays for each of these module types by constructing an aluminum trough fixture with inside dimensions corresponding to the finished outside dimensions (symmetric wedge in  $\phi$  and a three-faceted wedge on each  $\theta$  end (Figure 7-27) of the desired unit. The trough fixture has stops along each of the 21 crystal axes (Figure 7-26) to set the depth. To lay up a module, a mandrel is inserted into each of the 21 boxes. The exterior surfaces of the boxes are roughened and wetted with epoxy. The tubes and mandrels are then forced into the trough, in rows of three, until all 21 spaces are occupied. By seating the ends at the depth stops, the epoxy is forced out automatically, accommodating the irregular shapes of the outer walls of the boxes while preserving the inside dimensions. The mandrel heaters cure the unit. Once cured, the mandrels are removed, leaving the module in the trough. Any excess epoxy is cleaned off, and the perimeter edges of the two outer rows of boxes are trimmed to their final shape.

The next module manufacturing step is the stiffening of the boxes and attachment of aluminum support elements at the rear of the box array. The trough and unfinished module are moved to a dry-room area so that each compartment can be loaded with its wrapped CsI crystal. An example of a suitable facility is Building 109 at SLAC, which contains an existing, high-quality 2–3% humidity dry-room. This room will be enlarged and refurbished with a high performance lithium-chloride-based dehumidifier to allow crystal and module storage, assembly, and electronic testing in the same area. The original dehumidifier for the building will be refurbished and used for the adjacent temporary assembly area, and as an emergency backup system.

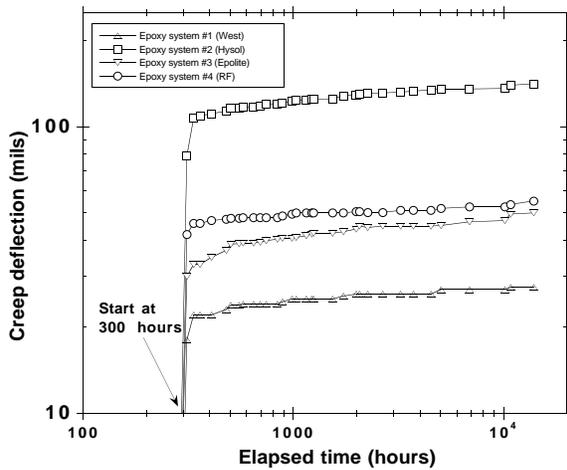
To complete a module, each crystal is pushed forward into its compartment. The taper of the crystal module and the fabrication tolerance ensure that the crystals will sit loosely within the box ( $\pm 0.5$  cm from nominal radial position). An aluminum enclosure (Figure 7-29) is inserted into the top and glued to the inside of the box with a small amount of conducting epoxy so that the crystal's foil wrap remains in good electrical contact with the enclosure. When all crystals are in place, the 1.27 cm-thick rear extruded aluminum U-channel strongback



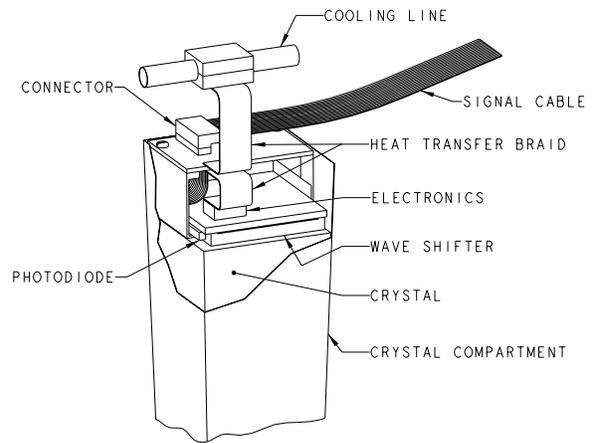
**Figure 7-26.** Trough fixture for combining tubes into module subassemblies.



**Figure 7-27.** Schematic of a barrel module.



**Figure 7-28.** Creep tests of glass-fiber woven boxes prepared by a wet-layup technique.



**Figure 7-29.** Detailed side view of a module showing crystal, shielded readout package, cooling, and cabling.

assembly is screwed and bonded to the side supports. Bridging elements are bonded in place to stiffen the module.

A set of wire alignment target holders are epoxied to each module as shown in Figure 7-27. Their position is set by the trough, allowing exact and reproducible placement of the targets.

### Module Completion and Testing

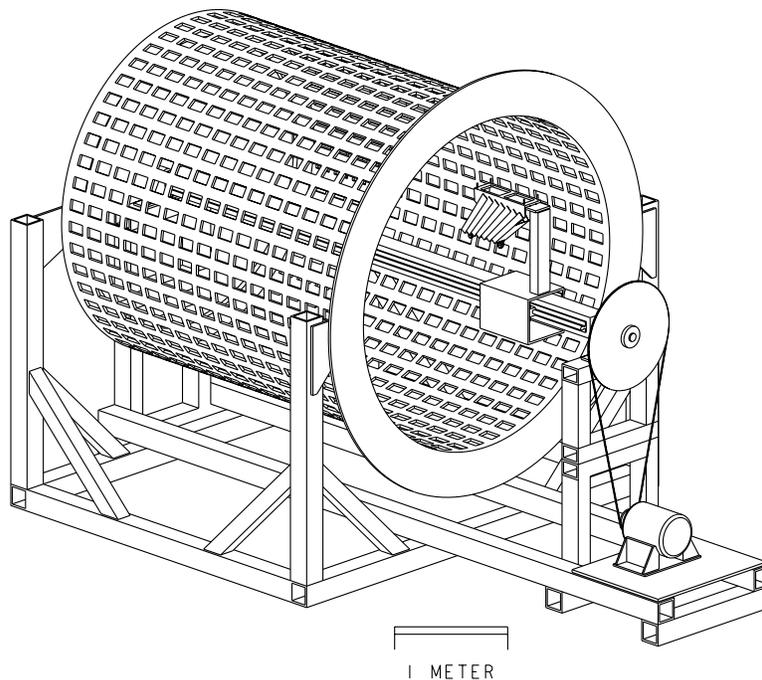
When the module is completed and removed from its trough, it is placed into dry-room storage until electronics installation. The electronics assembly involves the addition of a wavelength shifter/diode package and preamplifier board to each compartment in a module. These elements attach to the aluminum enclosure and are spring loaded to press against the back face of the crystal. A lip machined into the aluminum enclosure maintains a uniform air gap between the WLS and the crystal. If the direct diode readout method is selected, crystals have their PIN diodes pre-attached (and pre-tested) before being inserted into the module. The preamplifier is connected by soldering the diode leads through plated vias in the preamplifier board, and the preamplifier card is strain-relieved to the enclosure cover. The stamped cover of 1.5 mm thickness is screwed down to the aluminum enclosure, making contact all around the perimeter. Cables enter through the top by means of a ribbon connector. The finished module may be tested and returned to dry storage prior to insertion onto the strongback cylinder.

### Barrel Assembly

The assembly of the entire barrel and endcap calorimeter takes place adjacent to the dry-room in SLAC Building 109. A temporary dry-room ( $5 \times 5 \times 5$  m) is constructed with unistrut, wire mesh, and inner and outer plastic sheet walls. Walls are sealed to an epoxy-painted concrete floor and dry air blown between them. This temporary assembly area is connected to the storage/testing dry-room by a short interlocked passageway through which modules can be carried from storage.

The assembly fixture is shown in Figure 7-30. The large cylindrical strongback is first mounted to a support frame at the two points on each of the end-flanges that will eventually be used for support by the coil. By mounting it in this way, the strongback can be transferred to the coil supports without affecting the alignment of the modules. Alignment of individual modules compensates for strongback and module distortions that occur during the loading procedure.

A module is moved to the assembly area after electronics testing. The strongback of each module is mounted vertically on a movable module insertion arm on an assembly fixture (Figure 7-30). The arm is rotated to the proper  $\phi$  orientation and moved axially into



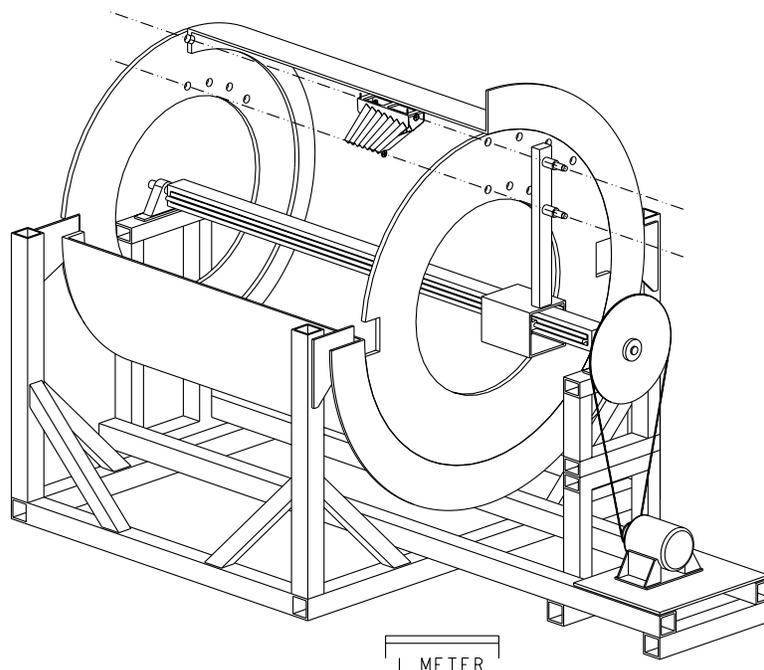
**Figure 7-30.** Barrel assembly fixture for inserting and mounting modules in the strongback.

place within the barrel. The module is then rough-mounted to the strongback. Initially, all modules are retracted radially about 2.5cm from nominal, allowing space for adjacent modules to be installed. When all modules are loaded, the strongback cylinder will be in its fully deformed state. The alignment mechanism is then used to align each module into the correct radius and orientation.

The alignment procedure is accomplished by placing an accurately machined template on the ends of the strongback cylinder and iteratively aligning the pre-mounted wire alignment targets with the template in each of the 40  $\phi$  module rows. Alignment telescopes are mounted on the module installation arm for simplicity. The procedure is shown schematically in Figure 7-31. Two or more crews can work simultaneously to align separate rows of modules, and the alignment procedure itself does not add significant deformations to previously aligned modules. Thus, a single alignment pass should be sufficient.

### Finite Element Analysis

A preliminary finite element analysis has been performed both on the strongback cylinder loaded with dummy modules and on a few representative individual modules. This approach



**Figure 7-31.** *Barrel alignment.*

is suggested by the independent suspension of the modules in the cylinder, allowing separate calculation of module deflections. The cylinder analysis determines an appropriate cylinder thickness, taking into account the amount of material removed for module access ports, and the required sizes of the forward and backward flanges. The forward(backward) end flange used in the cylinder analysis is 2.5 cm(5.0 cm)-thick and 36 cm(20 cm)-wide.

The strengthening tab thickness used in the model of a module is 0.64 cm. The analysis assumes for simplicity that the crystal boxes are made of 250  $\mu\text{m}$ -thick aluminum sheets (the modulus of elasticity of aluminum is about the same as the quasi-isotropic modulus of CFC). While we will not have enough layers of carbon fiber to make the composite behave isotropically, the steep pitch of the sock fibers makes the box stiffer than aluminum in the bending-strain direction. The loading of the crystal boxes is assumed to be concentrated where the box walls meet at right angles because the box material is so thin that it will have little mid-span bending stiffness. Under these conditions, the loading shifts to these corners where the boxes are stiffest.

The module model includes a matrix of 1.0 mm-thick aluminum electronics enclosures assembled behind the crystals. These enclosures also serve to keep each thin CFC crystal box square at this end. Additional stiffeners are included in modules near the end of the cylinder, where the center of gravity of the crystals is farther from the cylinder wall. Stiffer support

Element	Deflection (mm)
Module in Horizontal (end)	0.40
Module in Horizontal (center)	0.42
Module in Vertical	0.04
Strongback Cylinder (unloaded)	0.04
Strongback Cylinder (loaded)	0.67
End Flanges (cylinder loaded)	0.04

**Table 7-15.** *Static deflections encountered in the barrel.*

is required in this region, but there is also more room for material without obstructing electronics and cable ways.

Static deflections of the system are summarized in Table 7-15. For example, when all modules are installed, the 3.175 cm-thick strongback cylinder is deformed by  $\sim 670 \mu\text{m}$  at the center of the cylinder. The inner faces of modules mounted in the horizontal plane undergo the maximum deflection expected, about  $420 \mu\text{m}$  relative to the cylinder, or 1.1 mm total vertical deflection relative to the cylinder end flange mounting points. However, this 1.1 mm deflection is effectively removed during the final alignment of the crystal modules.

### Testing and Closure: Cabling, Cooling, Front and Back Panels

When modules are first installed in the strongback cylinder, an *in situ* electronic test is performed to ensure that there are no problems with individual crystals that require removal and replacement of a module. Once the final alignment is done, the cooling manifolds and cooling tubes to each box are installed, followed by the final cable runs. Plumbing for dry-nitrogen gas circulation is installed as is the inner rf/environmental cylinder. In order to keep all plumbing electrically isolated from the cylinder, nonmetallic couplers are used on the outer connections to manifolds.

A sufficiently long system test is then performed to find any remaining problems in the electronics. Next, the strongback cylinder access ports are closed with sufficiently thick sheet metal to complete the Faraday cage and gas seal.

**Barrel Installation into the Coil.** The fully aligned, cabled, and tested assembly is carried to the assembly area for the coil and flux return in the detector hall. A system test is once again performed to check that no damage occurred during transport to the hall. Installation into the detector is discussed in Chapter 14.

### 7.6.3 Endcap Fabrication, Assembly, and Installation

#### Endcap Overview

The design of the endcap is quite different conceptually from that of the barrel. Once installed, the barrel will remain unopened, possibly for the lifetime or the experiment, whereas the endcap must be capable of rapid mounting and demounting while maintaining precision mating to the barrel. The mechanical structure is sophisticated, involving a high precision, rigid outer shell with minimal material between the crystals.

The forward endcap is a conic section, with front and back surfaces tilted at  $22.7^\circ$  to the vertical to match the drift chamber endplate, and is built in two monolithic pieces to enable rapid demounting for access to inner parts of the detector. It is supported off the solenoid coil and precisely aligned with the calorimeter barrel, and has been designed to minimize both the material and the air gap between the two. The geometry is almost projective, with the crystal axes pointing to an axial position 5 cm from the interaction point. The total weight of the endcap is approximately 4 tonne.

In order to preserve optimum light collection and both spatial and energy resolution, similar-sized crystals have been used throughout. This results in a layout of nine rings of trapezoidal crystals, grouped in three super-rings of 120, 100 and 80 crystals. The total number of crystals is 900, arranged in 20-fold symmetry.

#### Crystal Honeycomb Container

Each crystal package, consisting of crystal, readout electronics and cooling circuit, is held rigidly in a separate compartment of the honeycomb structure, which is glued inside a load-bearing container, as shown in Figure 7-32. Structural rigidity is obtained from the solid aluminum inner wall of the conic section, reinforced at the rear with a semicircular flange. Bonded to this inner wall is the front face of the box, formed from two aluminum skins of 1 mm thickness separated by a nomex core to give additional rigidity. These last two surfaces also form part of the Faraday cage enclosing the barrel and endcap. The faces of the container that mate to each other or to the barrel are made from 1 mm carbon fiber composite (CFC) and serve mainly to bond the honeycomb structure of crystal boxes. A semicircular flange is bonded onto the rear of the thin outer wall of the conic section, both to maintain its shape and form part of the structural support.

The honeycomb compartments are individually made in the form of tapered boxes of CFC of wall thickness  $250\ \mu\text{m}$ . The technique involves stretching a CFC sock over a mandrel, coating it with epoxy, and curing it under pressure. The honeycomb is built up from 20 identical, wedge-shaped modules, each consisting of 45 boxes. Assembly takes place in a precision-

dimensioned aluminum trough. Mandrels are inserted into the boxes, the outer surfaces of which were previously roughened and coated with epoxy, and the whole cured under pressure. Ten modules are epoxied into each segment of the endcap, using a combination of mandrels and simple fixtures to ensure bonding to the walls of the container.

### Structural Assembly

Each of the two containers is prepared for transport, or assembly, by installing radial rails, fitted with cooling circuits and pre-cabled trays, between the inner and outer semicircular flanges, and clamping the walls of each crystal compartment in the outer super-ring to a rail and tensioning them. The rails form a hub at the center, giving the necessary rigidity to circumferential shearing.

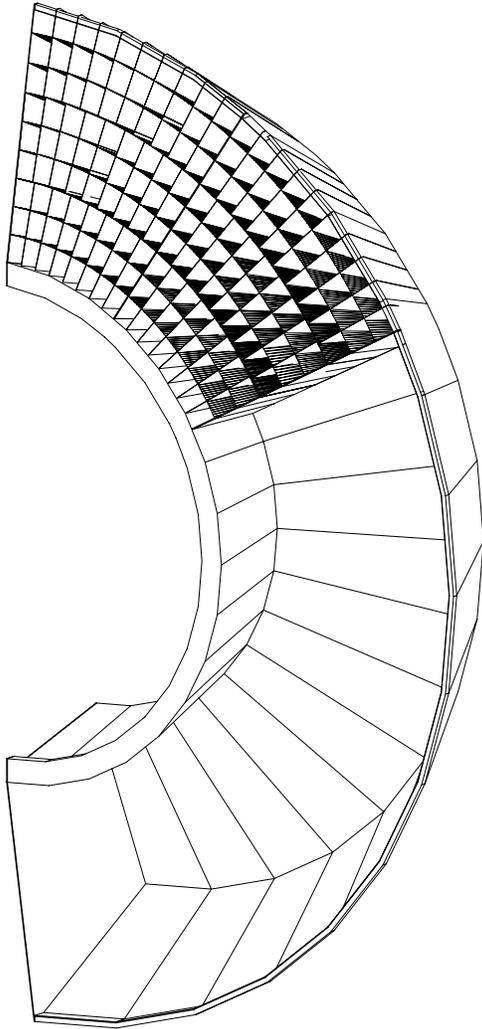
During crystal loading, the container is held vertically in an assembly jig which supports it from the front by means of the two rear flanges and the solid aluminum inner wall, shown in Figure 7-33. As each radial rail is removed in turn, crystal packages are loosely fitted into the compartments to rest on pre-measured styrofoam pads and secured with a square ring. The readout and preamplifier electronics unit for each crystal, pre-mounted with its heatsink on a cap, is screwed to the ring.

Heat sinks, electrical shielding and connections are secured to the rail and the outer crystal compartments clamped and tensioned. A final electrical test is performed on each group of crystals installed. The half ring support structure, with pre-mounted electrically shielded ADC housings (Figure 7-34), is bolted onto the outer flange of the container, and electrical connections made. The aluminum support plate is now bolted onto the support structure at its outer radius and the hub formed by the rails at its inner radius, and cover plates bolted over the access holes to complete the Faraday cage. Two assembly frames are required in order that each segment of the endcap may be completely prepared for installation.

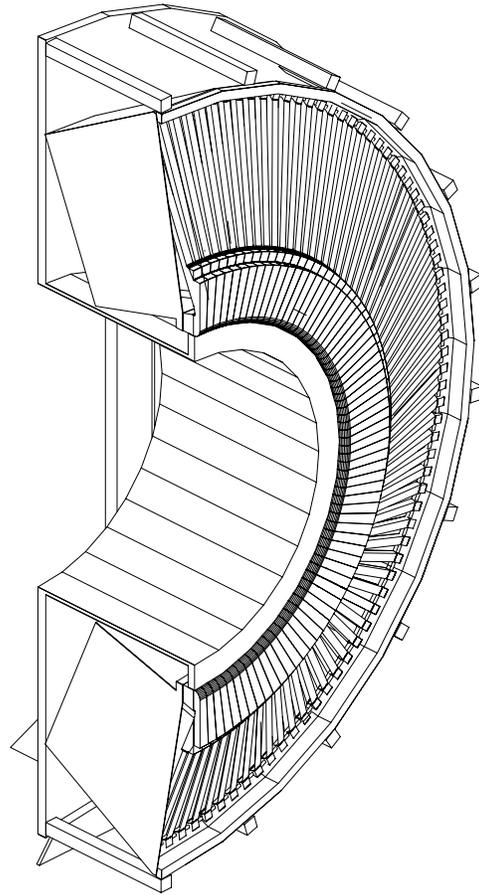
### Prototype Analysis

Stress analysis has been performed for individual crystal compartments, a module of 45 such compartments, and one segment of the endcap comprising a honeycomb of ten modules, using finite-element modelling of the loaded structures. The results indicate a rather uniform pattern of stresses and deformations with maximum values of 10 MPa and 0.1 mm respectively.

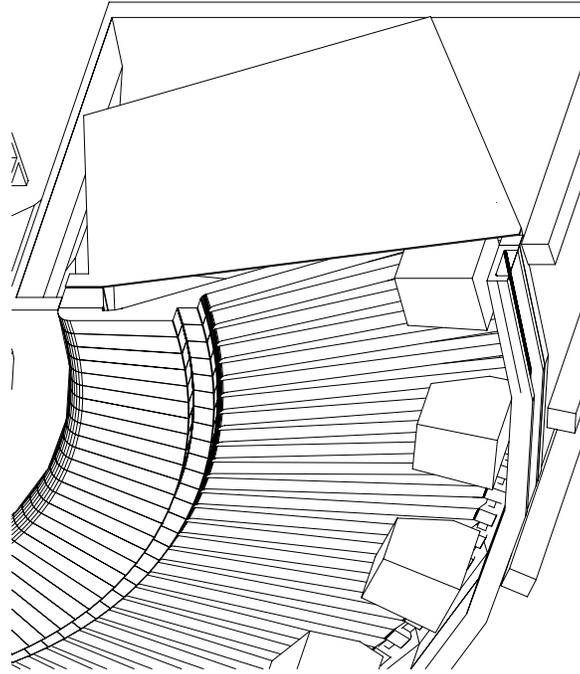
Because of the precision required in mating the barrel and endcap, it may be necessary to build a mechanical pre-production prototype consisting of a half-endcap, install it in the assembly frame and load it with weighted blocks having similar properties to CsI. Measurements taken with strain gauges might then be compared with predictions from the



**Figure 7-32.** *Endcap load-bearing container.*



**Figure 7-33.** *Section of endcap container showing assembly jig and rails.*



**Figure 7-34.** *Detail of endcap support structure showing ADC housings.*

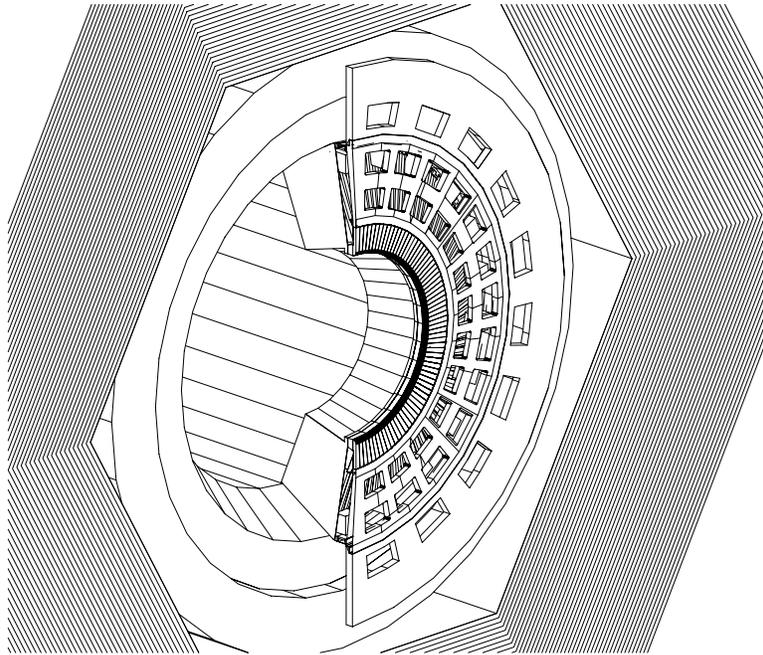
finite-element analysis, and differences fed back into the model to enable modification of the final endcap design.

**Endcap Installation into the Coil.** The endcap halves are carried to the assembly area in the detector hall. A system test is done to verify that the endcap has not been damaged in transport. Installation of the endcap halves is discussed in Chapter 14. Figure 7-35 shows the barrel and half of the endcap installed in the detector.

## 7.7 Optimization and Prototype Studies

---

Optimization and prototype studies consist of testing and refining individual calorimeter components as they are designed and built, as well as system tests. The system tests will consist of two beam tests. The first will primarily examine the optical elements and front-end electronics, while the second will feature a full system prototype. It is anticipated that the first beam test will occur in the fall of 1995, and that the second will occur prior to full production, in the spring of 1996. Potential beam test sites in the US and Europe are



**Figure 7-35.** *Side view of mated barrel and half of endcap.*

being evaluated. Test-site selection criteria include the availability of low-energy beams to measure the effect of electronic noise on resolution and the availability of mixed beams of electrons and pions to test  $e/\pi$  separation and the effect of hadronic split-offs.

In the first beam test, it is anticipated that we will use an array of  $5 \times 5$  crystals that have been tuned and wrapped in the manner planned for the actual calorimeter crystals. The optical elements to be tested consist of a photodiode and wavelength shifter readout, and the front-end electronics consists of the preamplifier/shaper amplifier package. To study their performance, these components will be tested in a realistic environment. The boxes used to hold each crystal will be made of the same materials as are planned for the actual mechanical support structure. The front-end electronics and its shielding will be those planned for the calorimeter. While not the final production version, an electronics cooling system will be implemented to obviate the effects of heating on the photodiode and electronics noise performance. Peak-sensing ADCs with a 12–13 bit resolution will most likely be used, rather than a prototype ADC. We will also not attempt to prototype the trigger and data acquisition system for this beam test. Aside from studying the front-end electronics, this test is an opportunity to gain experience with handling crystals and with the effects of the mechanical support and shielding materials. We anticipate results showing the effects of electronics noise and interstitial material on position and energy resolution.

The second test will consist of a nearly full system prototype, prior to production. One or two prototype mechanical modules will be utilized. This will give us further experience with mechanical fabrication techniques and with the installation of crystals and front-end electronics into the modules. However, tests to verify the mechanical strength and viability of the support structure will be conducted separately as part of the development of the structure. Based on the results of the first beam test, an improved, final production version of the front-end electronics and shielding will be implemented, as will a final version of the cooling system and cable layout. It is anticipated that prototype ADCs featuring a 17–18 bit dynamic range, though not necessarily a production version, will be utilized along with DAQ boards. The second beam test will give us a realistic check of the entire readout chain and shielding and of the functionality of the support structure before production begins.

## 7.8 Crystal Procurement Issues

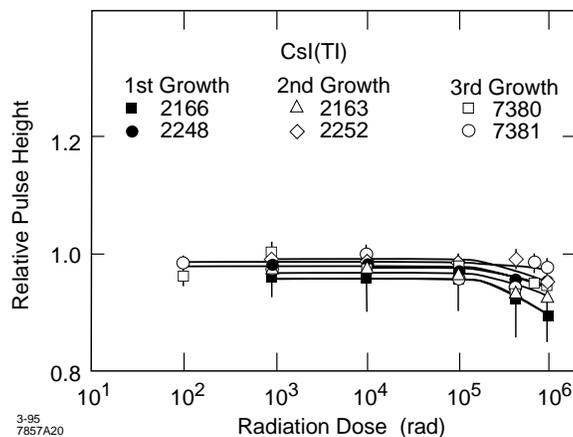
---

### 7.8.1 Radiation Hardness

The primary source of radiation for the CsI calorimeter is showers caused by beam-gas interactions or off-axis particles striking machine elements at or near the IR, not synchrotron radiation. The typical energies of the shower particles are in the MeV range. The expected dose level for the barrel calorimeter is  $\sim 1$  krad/yr, and for the forward endcap  $\sim 10$  krad/yr.

At  $\sim 1$  MeV, the photon attenuation length in CsI is  $\sim 5$  cm, or  $2.7X_0$ . The radiation dose is concentrated in the front 10–15 cm of the crystal. However, scintillation light produced deep in the crystal is also affected by radiation damage, as radiation damage results in the formation of color centers or absorption bands [Hol88], and our readout system collects light reflecting from the front and side crystal surfaces. Measurements of a small CsI(Tl) crystal under radiation indicate that CsI(Tl) is relatively radiation resistant (Fig. 7-36). However, larger samples from various vendors suffer a degradation in light yield [Hit92] when irradiated, perhaps due to impurities in large CsI(Tl) samples.

Since impurities in the CsI(Tl) salt may increase the likelihood for radiation damage in CsI(Tl) crystals, we will attempt to establish a technique for growing radiation-hard CsI(Tl) crystals by carefully controlling the purity of the salt. We will establish a radiation-hardness specification for crystal procurement as a final check of the finished crystals. We will investigate the role of impurities in CsI(Tl) salt by measuring crystal samples doped with known impurities. We will study the role of crystal processing in the development of radiation-hard crystals. To determine trace elements and defects in the crystal, material characterization studies and micro-structural analyses will be performed. We will also study the light yield, scintillation, transmittance, and absorption of these CsI(Tl) crystals. If



**Figure 7-36.** Relative light yield of small ( $2.54 \times 2.54 \text{ cm}^2$ ) Q&S CsI(Tl) crystals vs. radiation dose, measured with an avalanche photodiode with  $1.275 \text{ MeV}$   $\gamma$ s from a  $^{22}\text{Na}$  source.

correlations can be established between radiation hardness and an absence of impurities in the salt and/or the crystal growing process, this understanding can be incorporated into the growth of the crystals produced for the *BABAR* calorimeter.

Radiation hardness specifications for crystal procurement will be established to check the crystals supplied by vendors. A high dose test will be performed on small ( $2.54 \text{ cm}$  cylinders) samples from production boules. A sample will receive a short-term dose of  $20 \text{ krad}$ , after which its light output will be measured. We require that the light output diminish by less than  $20\%$ , and that it recover within five days to within  $5\%$  of the light output measured prior to irradiation. Large crystals may be tested for uniformity of transmission as well as light output as a function of time.

In addition to the high-dose radiation measurements, we will carry out long-term ( $\sim 12$  months), low-dose ( $\sim 1\text{-}3 \text{ rad/d}$ ) radiation studies using a modest number of production crystals from each supplier, selected after a consistent production quality has been established. These crystals will be placed in an environment similar to the one planned for the calorimeter (enclosure with dry-nitrogen flow, such that the relative humidity is  $\leq 3\%$  at room temperature) and irradiated from the front with a  $^{137}\text{Cs}$  source. The gain and resolution of these crystals and a set of control crystals kept in the same type of environment but without radiation exposure will be monitored throughout the irradiation period. This test will help us distinguish radiation damage effects from aging effects.

### 7.8.2 Quality Control and Testing

Obtaining the desired energy resolution from the calorimeter will require strict control over the quality of the 6780 crystals, which will be of different types and from different manufacturers.

As the crystals are produced by the manufacturer, the following tests will be applied, and unsatisfactory crystals will be rejected.

- The dimensions of the crystal will be measured to ensure that they lie within the required mechanical tolerances;
- Each crystal will be visually inspected for flaws;
- The light output from a standard source at several positions along the polished and wrapped crystal will be measured using a standard readout system; this will measure the overall light yield and the variation with position; and
- Samples of the melt will be analyzed for trace impurities.

These tests will be carried out at the point of manufacture by the manufacturer working in close liaison with representatives from the collaboration calorimeter group. The exact details of the tests (which are still under study) will be specified as part of the purchase contract.

Immediately prior to assembly at the experiment, the crystals will be tested again, to confirm that the properties have not changed due to damage in transit or other causes.

Each accepted crystal will be assigned a unique identification number and its properties stored in a database. This will enable us to keep track of any batch-to-batch variations in the properties of crystals that may emerge during the course of the experiment.

## References

---

- [Ake92] E. Aker *et al.*, *Nucl. Instr. Methods* **A321**, 69 (1992).
- [Ale94a] R. Aleksan *et al.*, “Report of the *Ad Hoc* Calorimeter Committee,” *BABAR Note # 157* (1994).
- [Ale94b] R. Aleksan *et al.*, “Resolution Studies Using GEANT,” *BABAR Note # 152* (1994).
- [BaB94] “Letter of Intent for the Study of *CP* Violation and Heavy Flavor Physics at PEP-II,” SLAC-443 (1994).
- [BEL94] BELLE Collaboration, “Letter of Intent for a Study of *CP* Violation in *B* Meson Decays” (1994).
- [Blu86] E. Blucher *et al.*, *Nucl. Instr. Methods* **A249**, 201 (1986).
- [Eis93] A.M. Eisner, “The Effect of Material *en route* to the EM Calorimeter,” *BABAR Note # 110* (1993).
- [Fis87] F. Fischer *et al.*, *Nucl. Instr. Methods* **A257**, 512 (1987).
- [Gro84] D. Groom, “Silicon Photodiode Detection of Bismuth Germanate Scintillation Light,” *Nucl. Instr. Methods* **A219**, 141 (1984).
- [Hal95] G. Haller, D. Freytag, and J. Hoefflich, “Proposal for an Electronics System for the *BABAR* Calorimeter,” *BABAR Note #184* (1994).
- [Hea94] C. Hearty, “Performance Comparison between CsI and LKr Calorimeters for PEP-II,” *BABAR Note # 141*, (1994).
- [HEL92] “HELENA, A Beauty Factory in Hamburg,” DESY Report DESY 92/041 (1992).
- [Hit92] D. Hitlin and G. Eigen, in Proceedings of the Int. Workshop on Heavy Scintillators for Scientific and Industrial Applications, edited by F. Nataristefani *et al.* (Editions Frontieres, 1992), p. 467.
- [Hol88] I. Holl *et al.*, *IEEE Trans. Nucl. Sci.* **35**, 105 (1988).
- [Jes94] C. Jessop *et al.*, “Report on the Performance of Hex *vs.* Trapezoidal Crystals,” *BABAR Note # 215* (1994).
- [Jes95a] C. Jessop *et al.*, “Development of the Front-End Readout for the *BABAR* CsI Calorimeter,” *BABAR Note # 216* (1995).

- [Jes95b] C. Jessop *et al.*, “Reliability Issues for *BABAR* CsI Calorimeter,” *BABAR Note # 217* (1995).
- [Kam83] T. Kamon *et al.*, *Nucl. Instr. Methods* **A213**, 261 (1983).
- [Kei70] G. Keil *et al.*, *Nucl. Instr. Methods* **87**, 111 (1970).
- [Kub92] Y. Kubota *et al.*, *Nucl. Instr. Methods* **A320**, 66 (1992).
- [Lev94] M. Levi, “Impact of Backgrounds on the Calorimeter Resolution,” *BABAR Note # 151* (1994).
- [Lor86] E. Lorenz *et al.*, *Nucl. Instr. Methods* **A239**, 235 (1986).
- [Mil90] U.S. Military Handbook, “Reliability Prediction of Electronic Equipment” (1990).
- [Sch94] R. Schindler, “Experimental Highlights of *B* meson Production and Decay,” in L. Montanet *et al.*, *Physical Review* **D50**, 1601 (1994).
- [Spi94] A. Spitkovsky and F.C. Porter, “Study of Light Transport in Different CsI Crystal Geometries,” *BABAR Note # 146* (1994).
- [Zhu95] R. Zhu, “ $e/\pi$  Separation in the *BABAR* CsI Calorimeter,” *BABAR Note # 151* (1995).
- [Zis91] “An Asymmetric *B* Factory Based on PEP,” ed. by M. Zisman, SLAC-372 (1991).



---

---

# Muon and Neutral Hadron Detector

---

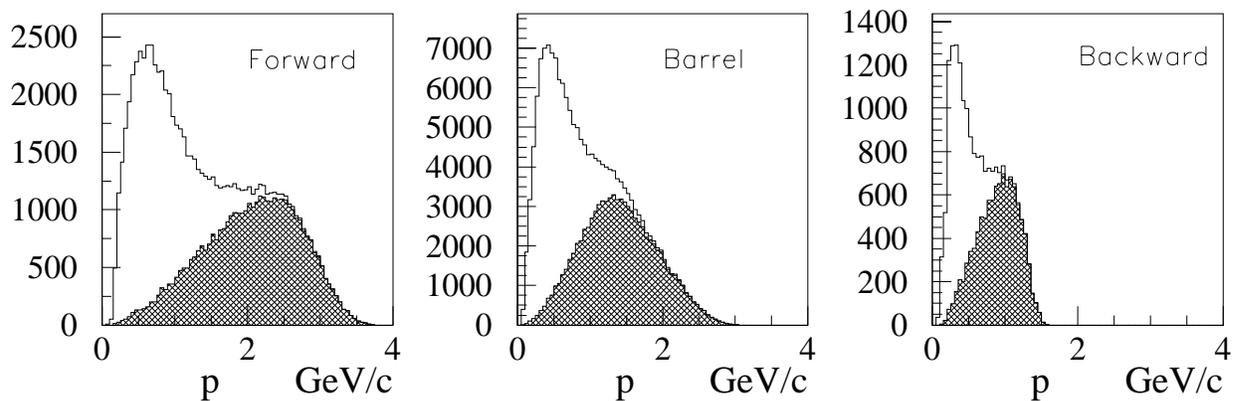
## 8.1 Physics Requirements and Performance Goals

---

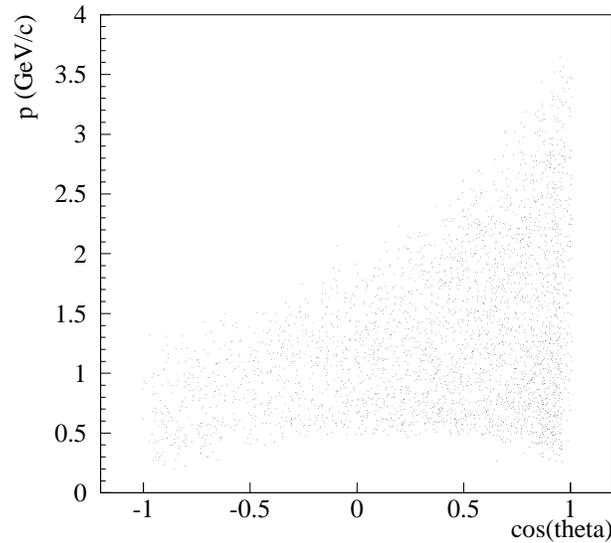
**M**uon identification and neutral hadron detection are provided by the Instrumented Flux Return (IFR), which makes use of the large iron structure needed as the magnet return yoke. Excellent muon identification over the widest possible momentum range and the ability to detect  $K_L^0$ s are distinctive features of the *BABAR* detector.

At the  $\Upsilon(4S)$ , muons are produced mostly in semileptonic decays, either directly from the  $B$  mesons or from the cascade  $D$ s. The sign of the charge determines the  $b$  or  $c$  flavor of the parent meson, thus providing a clean tag for the  $CP$ -asymmetry measurements.

For all particle species, the asymmetry of the machine results in a strong correlation between momentum and angle. The absolute population and the shape of the muon momentum spectra are quite different in the backward, central, and forward regions of the detector, as can be seen in Figure 8-1.



**Figure 8-1.** Momentum distribution of direct and cascade muons in (a) the forward endcap, (b) the barrel, and (c) the backward endcap regions. The shaded part of the histogram refers to direct decays.

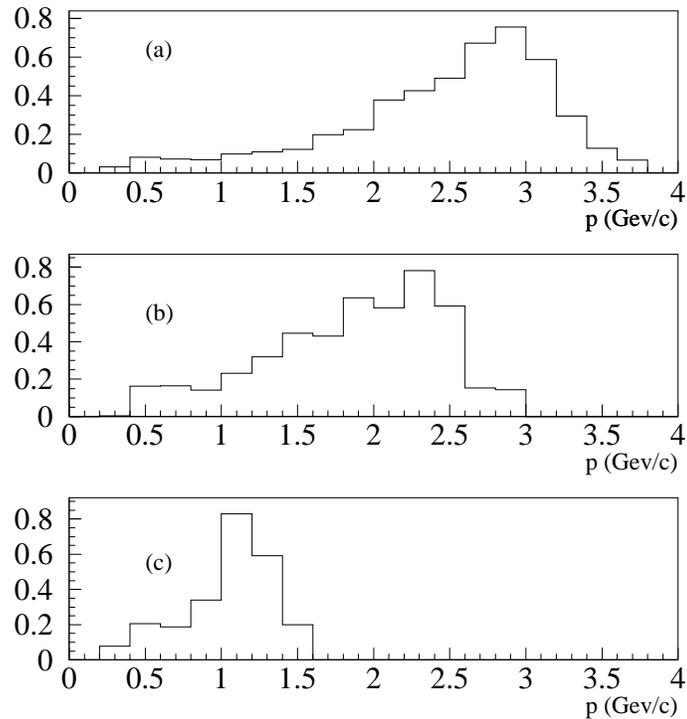


**Figure 8-2.** Correlation between momentum and  $\cos\theta$  for particles reaching the IFR; different momentum cut-offs for barrel and endcaps can be seen.

The main goal for the IFR detector is to achieve the highest practical tagging efficiency. About 18% of all  $B$  decays contain at least one muon in the region covered by the  $BABAR$  detector; these are roughly divided into 8% with a direct muon, 8% with a cascade muon, 1% with a muon from a  $\tau$  decay, and 1% dimuon events. High momentum muons ( $p \geq 1.2 \text{ GeV}/c$ ) are mostly direct decays, while muons from  $D$  decay peak at  $500 \text{ MeV}/c$ . A gain of almost a factor of two in tagging efficiency can be achieved if one can also detect low momentum (well below  $1 \text{ GeV}/c$ ) muons and correctly assign them to  $B$  or cascade  $D$  decays. Detection of low momentum muons will also improve the measurement of  $|V_{bc}|$  by reducing the systematic error in the extrapolation to low momenta.

The muon momentum range to be covered by the IFR therefore extends up to a few  $\text{GeV}/c$ ; the lower end of the range is fixed by magnetic bending in the barrel region ( $\sim 450 \text{ MeV}/c$ ) and by energy losses in the inner detectors ( $\sim 250 \text{ MeV}/c$ ) in the two endcap regions, as shown in Figure 8-2.

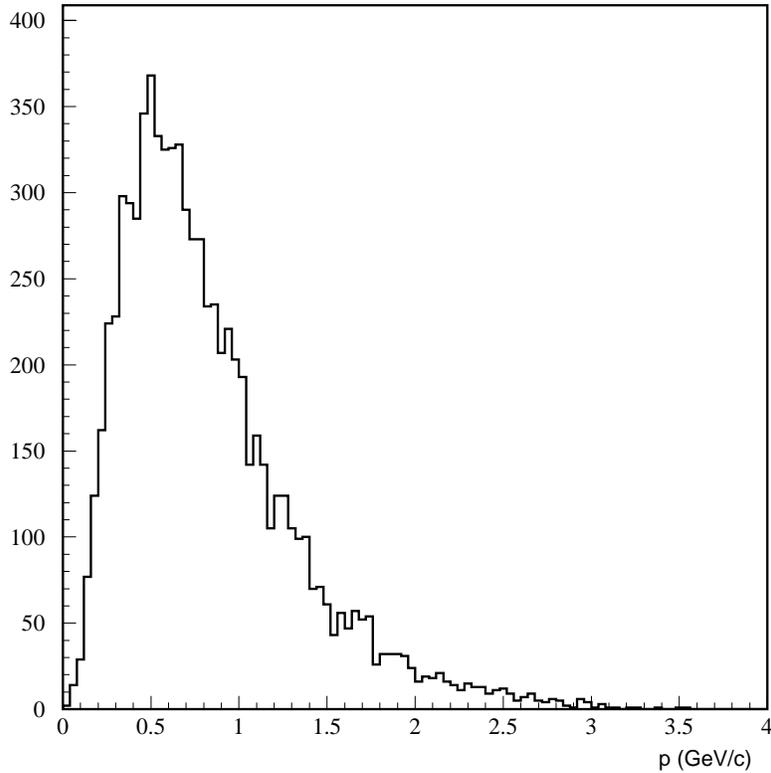
Charged tracks found in the central drift chamber will be matched to tracks in the IFR. Identification as muons or hadrons will result from detailed analysis of the hit patterns in the active detectors. The ratio of muons to hadrons in the three sections of the detectors is shown in Figure 8-3. The maximum tolerable level of hadron misidentification will depend on the physics topics under study; for tagging purposes, it has been shown that the hadron contamination in the muon sample should not exceed the 15% level. This translates into an upper limit of 5% for the probability that a hadron will be identified as a muon.



**Figure 8-3.** Muon to hadron ratio as a function of momentum in the three sections of the detectors: (a) forward endcap, (b) barrel, and (c) backward endcap.

The IFR also constitutes a hadron calorimeter; it will allow the detection of  $K_L^0$ s and other neutral hadrons. A measurement of the  $K_L^0$  energy is not foreseen due to the limited calorimetric resolution. In any event, such a measurement produces no real improvement in event reconstruction. The hits induced by the charged secondaries produced in the  $K_L^0$  interaction will be used to detect  $K_L^0$ s. (About 70% of  $K_L^0$ s interact before reaching the IFR.) The  $K_L^0$  direction is then inferred from the location of the energy deposition. The ability to identify  $K_L^0$ s and reconstruct their directions creates the opportunity to use the decay  $B^0 \rightarrow J/\psi K_L^0$  for  $CP$ -violation studies. This could provide an event sample comparable to the benchmark decay  $B^0 \rightarrow J/\psi K_S^0$  and act as a systematic check on any observed  $CP$  asymmetry, since  $K_L^0$  and  $K_S^0$  are opposite  $CP$  states. The parent  $B$  can be reconstructed from the four-vector of the  $J/\psi$ , measured in a leptonic decay, and the direction of the  $K_L^0$ .

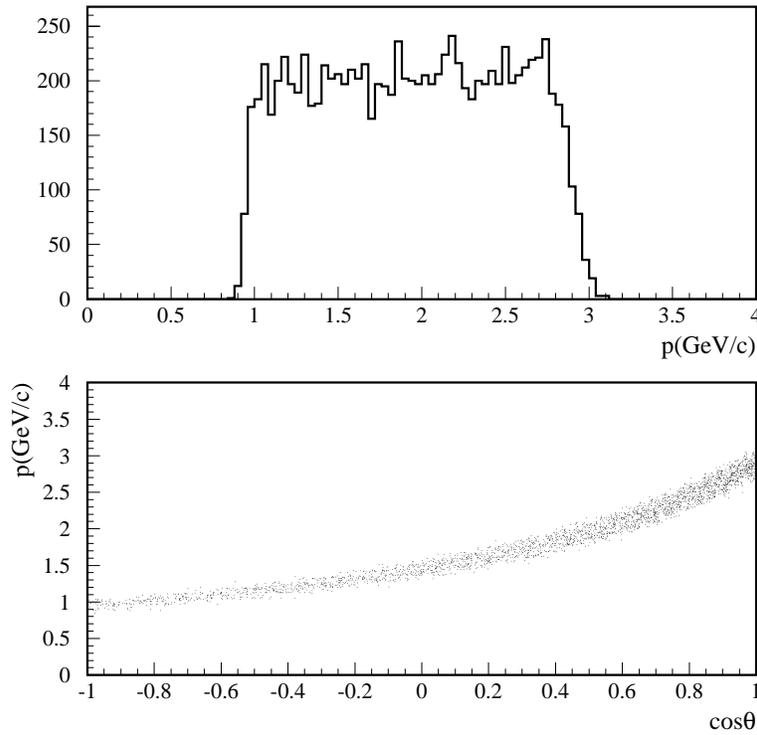
The momentum spectrum for  $K_L^0$ s produced in  $B\bar{B}$  events peaks at 500 MeV/c (Figure 8-4).  $K_L^0$ s from  $B^0 \rightarrow J/\psi K_L^0$  are more energetic, as can be seen in Figure 8-5(a). Since this case is a two-body decay, the momentum distribution is practically flat, and extends from 1 to 3 GeV/c. The  $K_S^0$  detection efficiency averaged over this momentum range is  $\sim 50\%$ ; this is the benchmark with which to compare the  $K_L^0$  detection efficiency. The measurement of the  $K_L^0$  direction is enough to identify the  $B^0 \rightarrow J/\psi K_L^0$  decay, given the strong correlation



**Figure 8-4.** Momentum distribution for all  $K_L^0$ s produced in  $B^0\bar{B}^0$  events.

between momentum and angle. The spread can be seen in Figure 8-5(b) and is due to the  $B$ 's momentum distribution. However, a few significant physics backgrounds must also be handled. Decays of both charged and neutral  $B$  mesons to  $J/\psi K^*(892)$ , in which the  $K^*(892)$  decays to a  $K_L^0\pi$  pair, produce events that can directly mimic the desired signal with branching ratios slightly larger than the signal. The momentum distribution and the momentum angle correlation of the  $K_L^0$ s produced in this process are shown in Figure 8-6. These backgrounds can be suppressed by identifying the extra pion and reconstructing the  $K^*(892)$  mass. Events in the charged case,  $B^+ \rightarrow J/\psi K^{*+}$ ,  $K^{*+} \rightarrow K_L^0\pi^+$ , may be further suppressed by rejecting events in which the charged pion can be associated with an isolated  $J/\psi$  vertex.

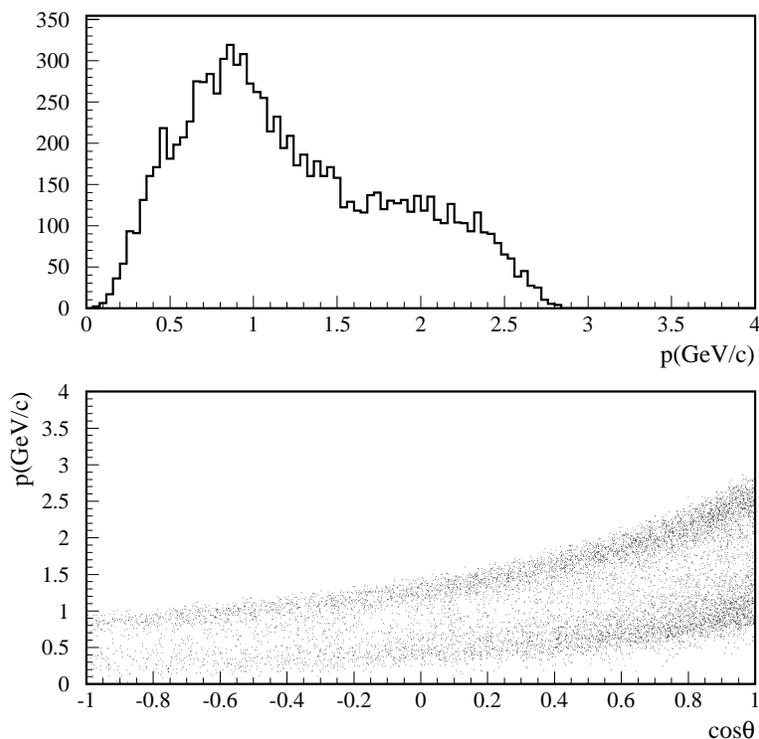
To improve the signal to background ratio, it is essential to have good resolution in the measurement of the  $K_L^0$  direction in addition to an efficient reconstruction of the  $K_L^0$ . Figure 8-7 shows the dependence of the signal and of the background, before any other cut is applied, on the angular resolution of the detector. The  $K_L^0$  detection efficiency improves steadily with a finer iron segmentation, as shown later in this chapter.



**Figure 8-5.** The momentum distribution and the momentum vs. angle correlation for all  $K_L^0$ s produced in  $B^0 \rightarrow J/\psi K_L^0$  decays.

$K_L^0$  detection is useful in channels for which the  $K_L^0$  direction alone is sufficient to reconstruct the momentum of the parent  $B$ . In particular, several  $D$  decays can be fully reconstructed using charged tracks in a separate vertex and the direction of the  $K_L^0$ . For example, we have tested the possibility of reconstructing the decay  $D^- \rightarrow K_L^0 \pi^+ \pi^- \pi^-$ . Decays of  $B^0$  to  $D^- \pi^+$  have been generated in which the  $D^-$  is then allowed to decay to  $K_L^0 \pi^+ \pi^- \pi^-$ . The three-pion  $D^-$  decay should allow a good vertex reconstruction. The  $K_L^0$  is then reconstructed by measuring its direction in the IFR system and recomputing its momentum under the assumption that it comes from a  $D^-$  decay. Figure 8-8(a) shows the resulting  $K_L^0$  momentum distribution together with the fraction interacting in the IFR (shaded area). Figure 8-8(b) shows the resulting  $D^- \pi^+$  effective mass for those events which can be reconstructed. Requiring in addition a recoil  $B^0$  (Figure 8-8(c)) should allow this decay mode to be detected with little background.

In addition to exclusive event reconstruction, the IFR can be used to veto events with missing hadronic energy, in particular background events for the decay  $B \rightarrow \tau \nu_\tau$ ; when used as a veto, it is obviously essential for the detector to have as complete a solid angle coverage as practical and the highest possible efficiency.



**Figure 8-6.** *The momentum distribution and momentum vs. angle correlation of the  $K_L^0$ s produced in  $B \rightarrow J/\psi K^*(892)$ .*

## 8.2 Detector Overview

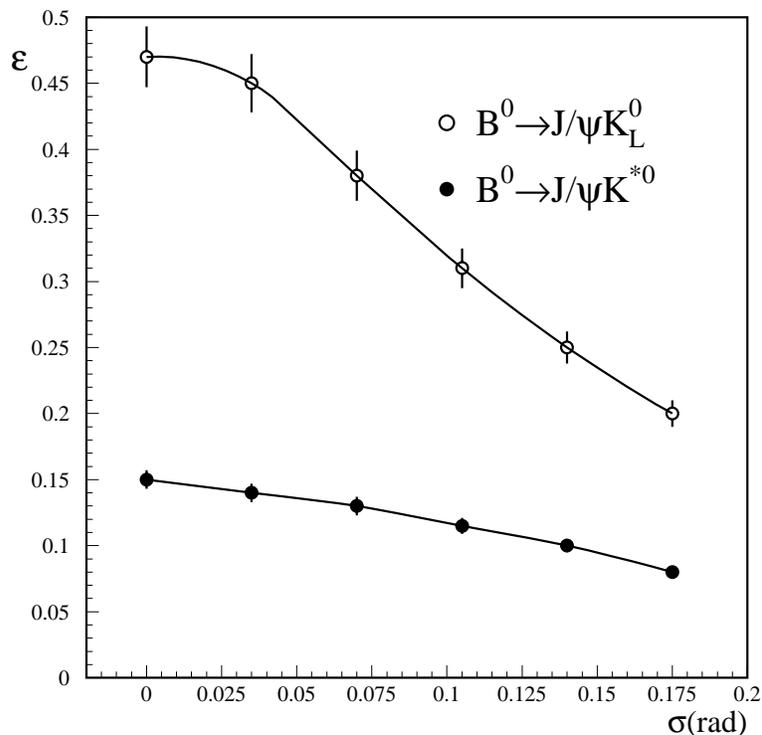
---

### 8.2.1 The Iron Structure

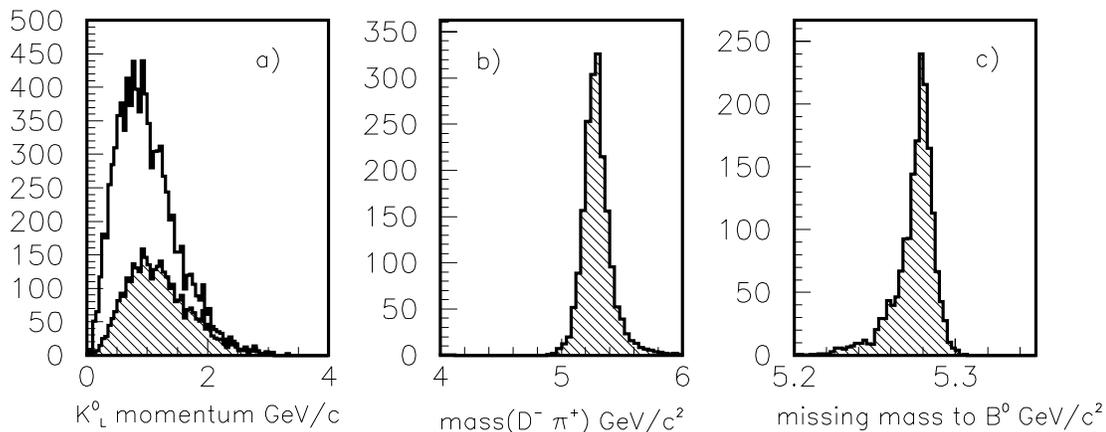
The muon and neutral hadron detector uses as an absorber the flux return iron of the 1.5 T superconducting solenoid; this massive steel structure also serves as the support for the inner detectors.

The design of the iron plates has to take into account both the requirements dictated by its use as hadron absorber and muon filter and the complex mechanical problems caused by the need to provide access to the inner detectors, all subject to the space constraints imposed by the machine components and the experimental hall.

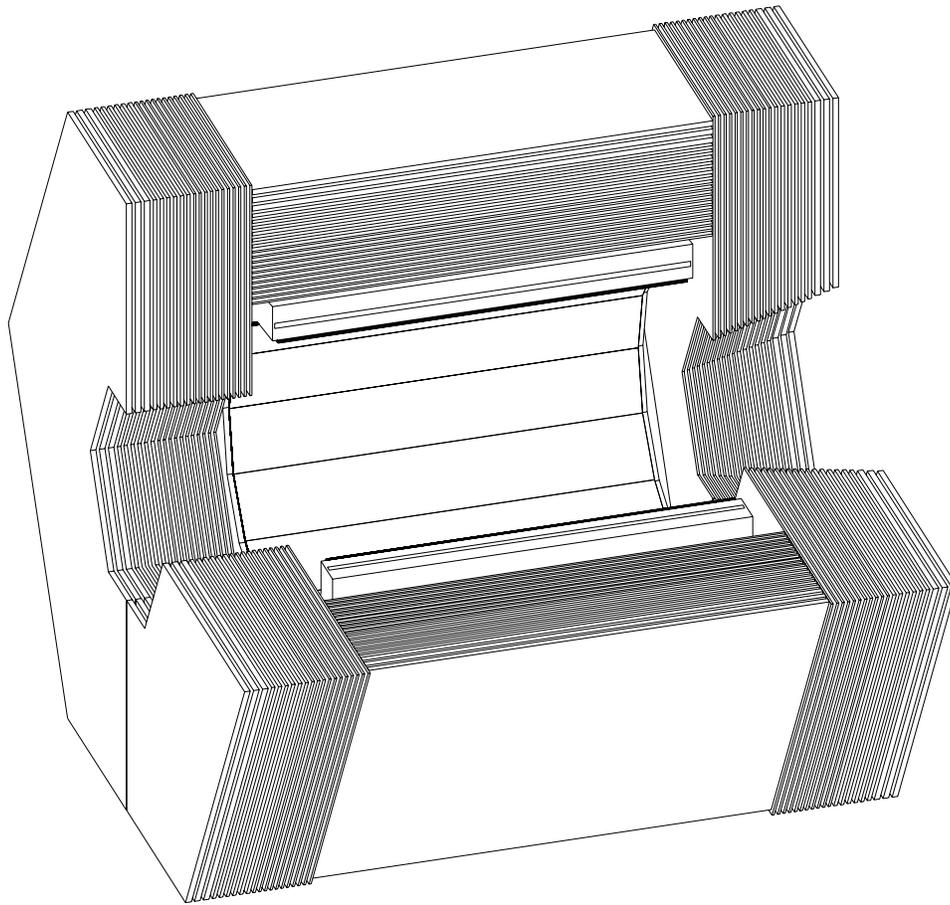
The IFR consists of three main components: the barrel and the backward and forward endcaps. The barrel extends radially from 1.78 to 3.01 m and is divided into sextants; the length of each sextant is 3.75 m, and the width varies from 1.88 to 3.23 m. Each endcap



**Figure 8-7.** Signal and background efficiencies for reconstructing  $B^0 \rightarrow J/\psi K_L^0$  as functions of the angular resolution in the  $K_L^0$  direction.



**Figure 8-8.** (a) Reconstructed  $K_L^0$  momentum from  $D^- \rightarrow K_L^0 \pi^+ \pi^- \pi^-$ . The shaded area represents the fraction interacting in the IFR. (b)  $D^- \pi^+$  effective mass distribution. (c) Missing mass to the reconstructed  $B^0 \rightarrow D^- \pi^+$ .



**Figure 8-9.** *The IFR detector.*

consists of hexagonal plates, divided vertically into two parts to allow opening of the detector and has a central hole for the beam components and the magnetic shields.

The plates are tied together by side steel plates whose thickness (5 cm) reduces the area available for the active detector. The dead area is concentrated at the joining of the hexagonal blocks and accounts for a small loss in solid angle, approximately 7%. The endcap gaps are divided into three by stiffeners needed to prevent bending due to the magnetic forces and to limit the stress due to mechanical and potential seismic loads. A view of the detector is shown in Figure 8-9.

The iron segmentation is dictated by the physics requirements of identifying low-momentum muons and detecting  $K_L^0$ s. The plate thickness is graded from 2 cm for the inner plates to 5 cm for the outer ones, based on simulation studies which will be described below. The structure consists of 11 plates 2 cm thick followed by six plates 3 cm thick and three plates

5 cm thick. The total thickness is 55 cm and is the result of a compromise between the cost, which scales with the volume, and the need to reduce hadron punchthrough. Most of the simulations, however, have been done with a total thickness of 60 cm of iron (this has been the reference design).

The fraction of  $B$ -decay hadrons reaching the IFR that will fail to be absorbed in 55 cm of iron is 1.3% overall: 0.2% in the backward region, 1.0% in the barrel, and 1.7% in the forward endcap.

### 8.2.2 The Active Detector Choice

The IFR detectors will cover a surface of approximately 2500 m<sup>2</sup>; they will be inserted in the gaps between the iron plates where access and replacement will be very difficult, perhaps impossible. It is, therefore, necessary that the technique chosen be simple and reliable as well as low in cost. Two technologies were considered, one using Plastic Streamer Tubes (PSTs) and the other using Resistive Plate Counters (RPCs). RPC technology was chosen for its superior flexibility and performance and larger active area. It was originally developed by R. Santonico and collaborators [San81, San88]; the technology is now well past its development stage, having been successfully used in several experiments [Accel, Cosmi]. A large ( $\sim 600$  m<sup>2</sup>) system is presently in use in the L3 [L3RPC] experiment at CERN, and much larger systems are now being designed for the LHC experiments [LHC94].

As will be discussed in more detail in Section 8.4.1, the proposed RPC is essentially a gas gap at atmospheric pressure enclosed between two 2 mm-thick Bakelite (phenolic polymer) plates with bulk resistivity on the order of  $10^{11}$   $\Omega$ -cm. These electrodes are coated on their outer faces, with thin graphite layers connected to high voltage ( $\sim 8$  kV) and to ground, respectively. A crossing charged particle produces a quenched spark that produces signals on external pick-up electrodes.

Active R&D is ongoing in *BABAR* and other groups to select a nonflammable and environmentally safe gas mixture.

Several RPC modules will be joined together to fill each IFR layer; on the two sides, planes of orthogonal readout strips will be attached for signal readout, so that a space point will be measured for each hit. The design of the strips is not finalized yet, pending further studies of the granularity needed. The total number of digital readout channels calculated on the basis of a 3 cm pitch is  $\sim 40,000$ . Using data from the L3 chambers and Monte Carlo simulations we will investigate whether a coarser granularity will suffice and whether better performance can be obtained by grading the pitch with the distance from the IP. The construction of the strip planes, *i.e.*, the gluing of pre-made aluminum strips on a PVC support, is considerably

faster than that of the chambers, so this item is not on the critical path. We defer final decisions until after the completion of more sophisticated Monte Carlo simulations.

Each layer will be inserted into the iron gaps as a single object using appropriate tooling; readout electronics cards will then be positioned near the gap border where they will be readily accessible for testing and maintenance; the last 20 cm of each gap will therefore be made 10 mm wider, to make room for the cards and cables.

## 8.3 Projected Performance

---

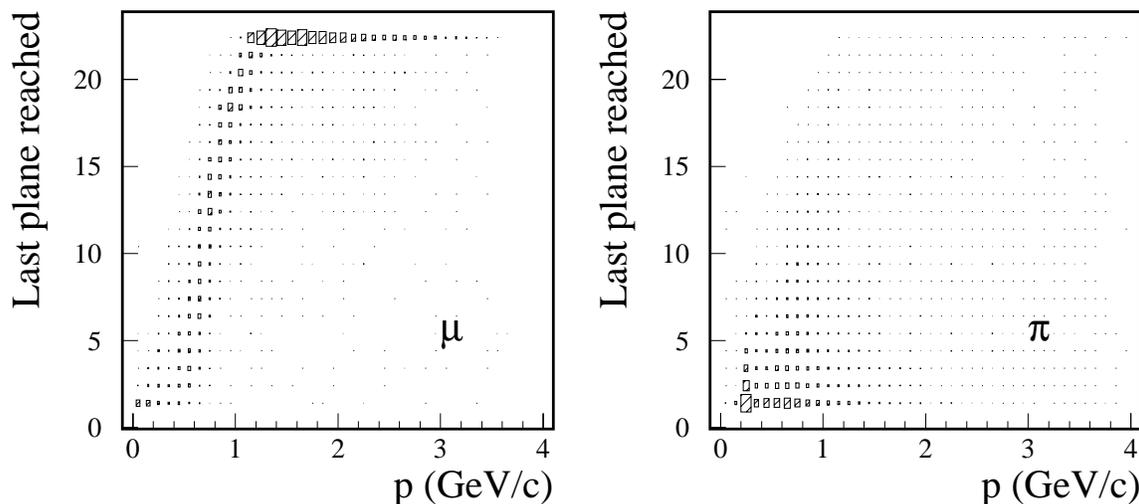
### 8.3.1 Muon Identification

Muons are positively identified if they penetrate all layers of the iron. Non-penetrating muons (below 1.1–1.5 GeV/ $c$ , depending on the incidence angle) can be identified from the ionization losses in the iron. Pion separation is achieved through a combination of range and hit pattern cuts; its effectiveness increases with segmentation.

Following the program outlined in the *Letter of Intent*, several simulation studies have been performed in order to determine the IFR configuration with optimal performance for cost. In order to study a large number of segmentation configurations, as a starting point we simulated, using the GEANT implementation of the *BABAR* detector geometry, an IFR geometry consisting of 61 detector layers 1.2 cm thick alternating with 60 plates of iron of 1 cm thickness. The ratio of iron to detector thickness was chosen to be the same as in the *Letter of Intent* geometry, in which 24 iron planes of 2.5 cm were alternated with 25 detector layers of 3 cm. Using this method, all configurations with iron thicknesses that are multiples of 1 cm can be reproduced by switching off the undesired planes at the analysis level. We have studied many possible configurations, including both uniform and graded, as a function of radial distance from the beam, thicknesses. The minimum plate thickness studied is 2 cm; plates thinner than 2 cm are not considered practical for mechanical reasons.

The BEGET generator [Wri94] is used to produce  $B^0\bar{B}^0$  events in which the  $B^0$  can decay into any channel (branching ratios from [PDG94] are used), and the  $\bar{B}^0$  becomes invisible to the detector. In this way,  $B^0$  tagging can be studied, assuming that the  $CP$  eigenstate decay of the  $\bar{B}^0$  is completely reconstructed.

Figure 8-10 shows the distribution of the last detector layer reached as a function of momentum for muons and pions. Muons with momentum below  $\sim 1.3$  GeV/ $c$  stop in the IFR, and most of the pions interact in the iron.

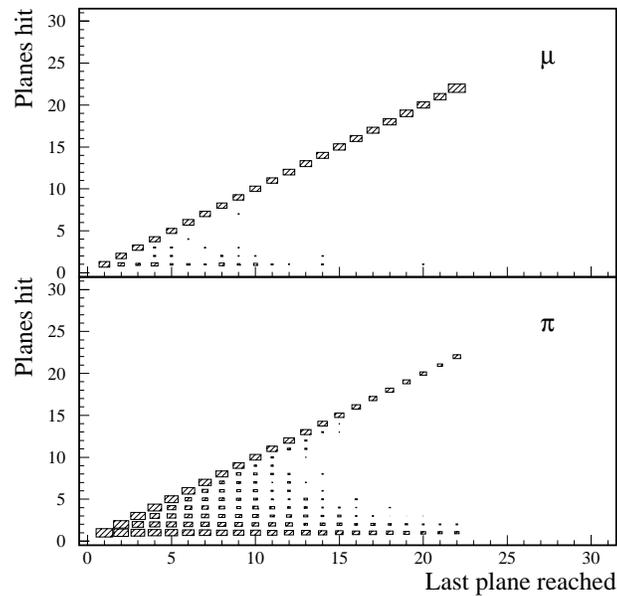


**Figure 8-10.** Distribution of the last detector plane reached as a function of momentum for  $\mu$  and  $\pi$ .

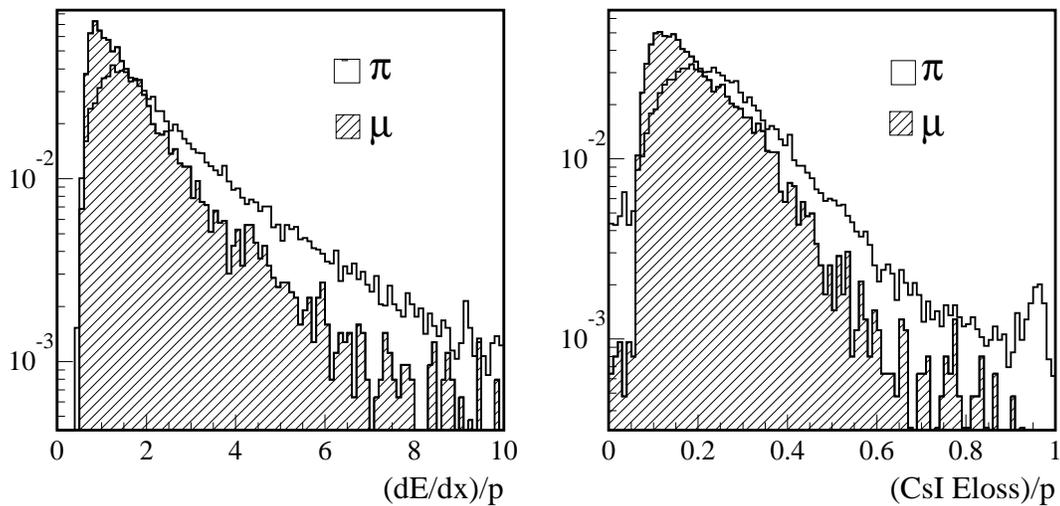
A criterion which is particularly useful for distinguishing muons from hadrons is based on the number of planes hit versus the number of planes traversed [Calc94]. A muon track typically has one hit in each layer, while hadrons can interact strongly in the iron, generating a hadron shower with several particles producing multiple hits per layer. It is also possible that only neutrals are present, and this will cause the absence of hits in one or more planes, even for an ideal detector. In general, for pion tracks the number of planes hit is smaller than the number of planes traversed, while muon tracks have hits in almost all of the planes they traverse. The difference in the behavior of muons and pions in this respect can be seen in Figure 8-11, which shows the number of planes hit versus number of planes traversed. A cut on the difference between these numbers gives a powerful pion rejection.

In a further development of this idea [Lis94], a maximum likelihood function can be built from a detailed study of the number of hits in each plane for muons and for hadrons, as a function of momentum. This can be used to determine a combined probability for a track to be a muon or a hadron. We believe that this technique, still under development, will be the one actually used to identify muons.

Energy losses in the drift chamber and CsI calorimeter have slightly different distributions for muons and pions at low energy (Figure 8-12), and this difference can be used to reduce pion contamination at low momenta.



**Figure 8-11.** Number of planes hit vs. number of planes spanned. The size of the boxes is proportional to the log of the number of entries. 100% detector efficiency is assumed.



**Figure 8-12.** Energy loss in the inner tracker (average per hit) and in the CsI calorimeter. Pions and muons are normalized to equal area.

Several Monte Carlo analyses have been performed with GHEISHA and FLUKA as hadron shower generators. The variables used to discriminate muons are the following:

- Number of planes hit;
- Traversed material (Fe cm equivalent);
- Number of planes spanned versus number of planes hit;
- Number of consecutive missed planes along the track;
- Likelihood function for number of hits in each layer;
- Tracking residuals; and
- Energy deposition in the inner detectors.

Several sets of cuts depending on momentum and polar angle have been used for these variables, in order to increase efficiency while keeping hadron contamination as low as possible.

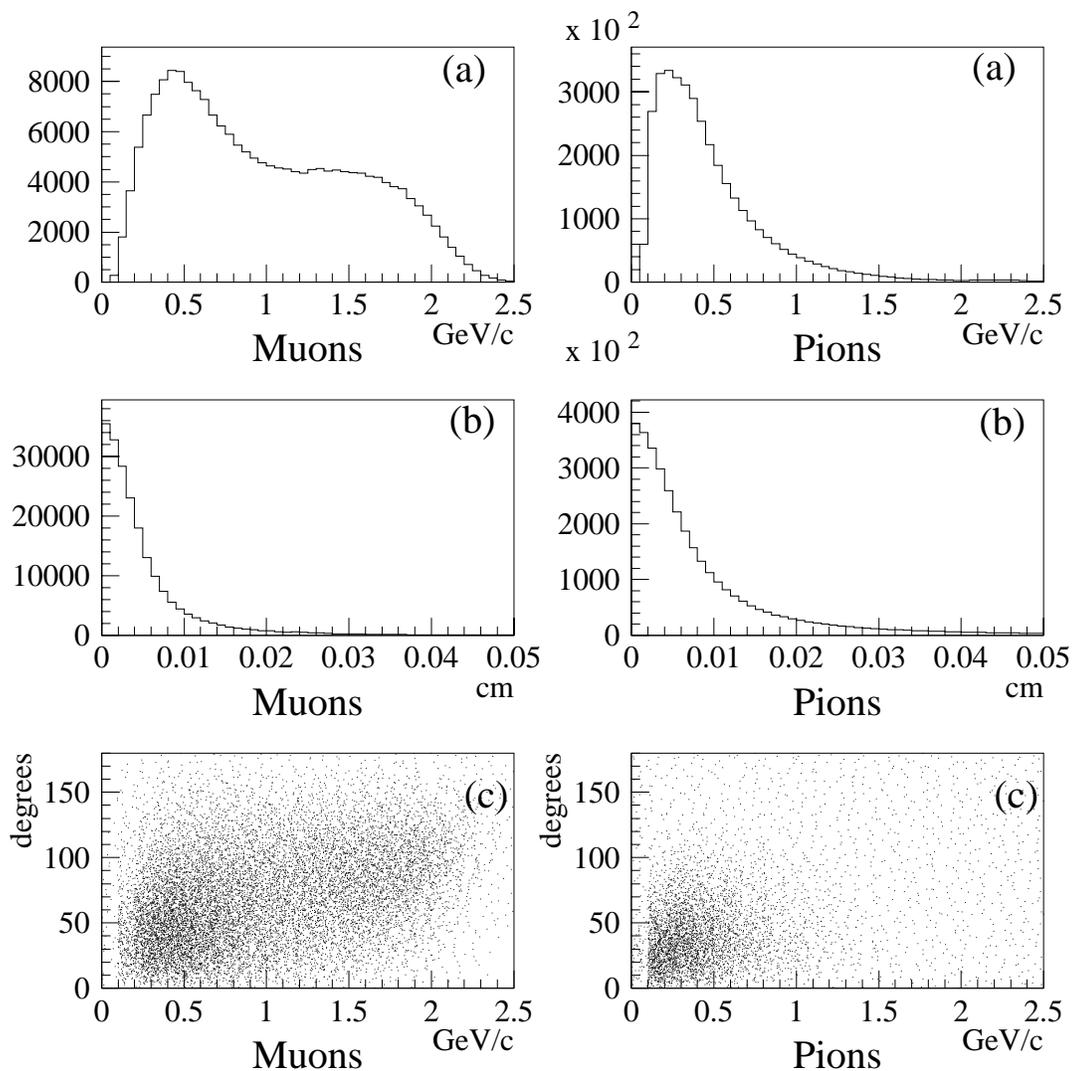
Misidentified hadrons assigned to the muon sample result in a dilution of the tagging efficiency, and introduce a background to the study of semileptonic decays and other channels containing muons. Different analyses, however, will require different tuning of these cuts. A typical example is the reconstruction of  $J/\psi \rightarrow \mu^+\mu^-$  in which we can safely use loose identification criteria because of strong kinematic constraints.

Flavor tagging, on the other hand, requires a clean muon sample because the effectiveness of the tag depends strongly on the tagging purity. In this case, the background to the muon sample from misidentified hadrons can be further reduced by studying the kinematic features of the two classes of tracks [Pia94]. A set of discriminant variables, such as those shown in Figure 8-13, can be identified and a probability density function derived for each variable. From the probability distributions associated with each discriminant variable, one can calculate the probability that a track is a muon as follows [Jaf94]:

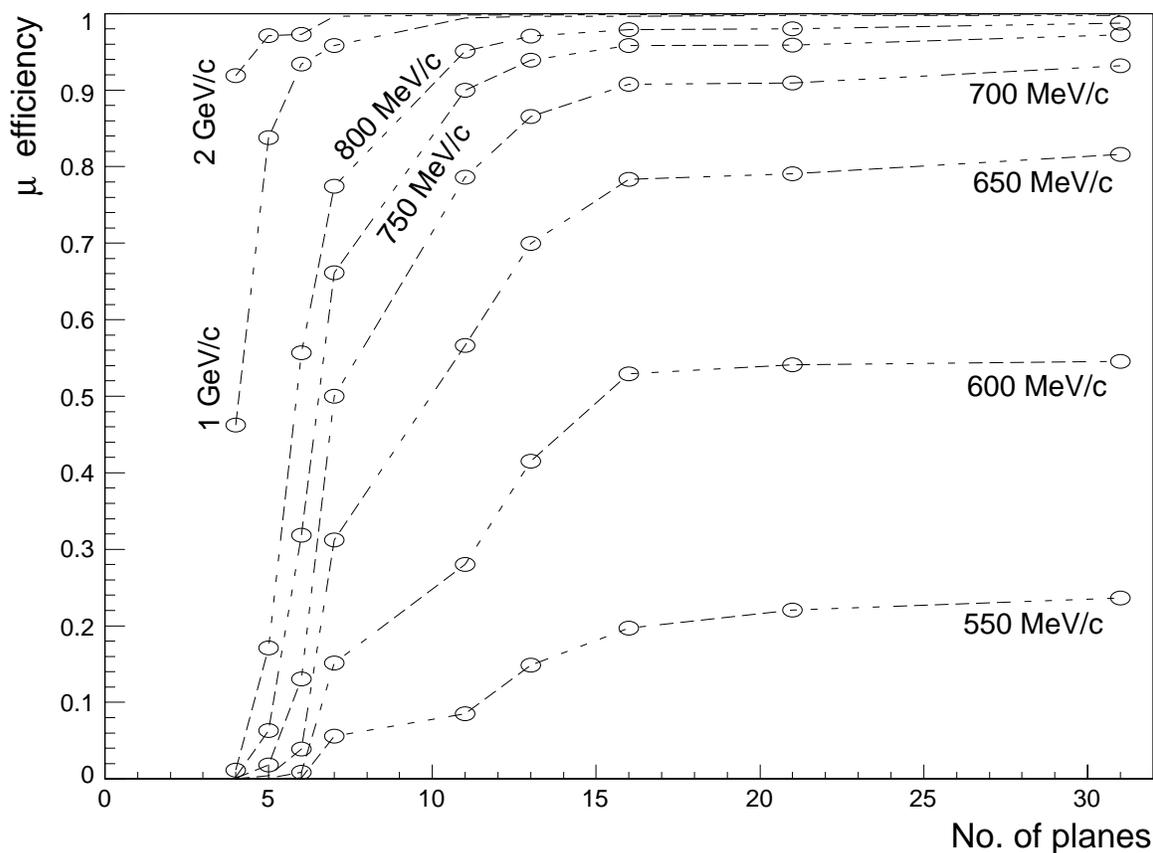
$$y_{eff}^{\mu} = 1 - \frac{\eta \prod g_i^{\pi}}{(1 - \eta) \prod g_i^{\mu} + \eta \prod g_i^{\pi}}, \quad (8.1)$$

where  $g_i^{\mu}$  is the probability associated with the  $i^{th}$  discriminant variable for the particle to be a muon,  $g_i^{\pi}$  is the similar probability for the particle to be a pion, and  $\eta$  is the fraction of pions in the total muon-pion sample. A simple cut in  $y_{eff}^{\mu}$  has been found to correctly identify  $\sim 95\%$  of the muons, with a rejection ratio of 15 : 1 for pions.

Configurations with uniform absorber segmentation have been studied with steel plate thicknesses ranging from 2 cm to 20 cm, and 31 to 4 detector layers, respectively. The efficiency



**Figure 8-13.**  $\mu/\pi$  discriminant variables: (a) center-of-mass momentum, (b) distance from the beam line, and (c) isolation, *i.e.*, the opening angle from the closest track.



**Figure 8-14.** Muon efficiency as a function of the number of detector planes for fixed momentum value. The simulation is based on FLUKA (total iron thickness of 60 cm). Passage through a detector plane is simply counted with perfect efficiency; no simulation of RPC information is incorporated. Inner detector cuts are not applied.

for muon identification at a given hadron misidentification probability (4%) as a function of the iron segmentation is shown in Figure 8-14 for several momentum ranges.

Reducing the number of planes clearly degrades the efficiency and worsens the signal to background ratio at low momenta and, to a lesser extent, at high momenta. It is also evident from the figure that increasing the segmentation up to 30 iron plates does not improve the muon efficiency significantly for penetrating muons.

These results are obtained with a set of simplifying assumptions that detection efficiency (both geometric and for hits) is 100% and that each GEANT energy deposition gives exactly one hit in the detector. Further studies to refine this analysis are underway; it is, however, safe to assume that for a realistic detector, we will have to rely more on likelihood methods

than on imposing cuts in the hit pattern. The loss in performance for configurations with a smaller number of detector planes will therefore be significant.

### 8.3.2 Muon Tagging

The muon identification efficiency and pion contamination obtained in these full Monte Carlo studies have been parameterized and used in an ASLUND analysis to determine the effective tagging efficiency as a function of the detector segmentation. The effective efficiency is obtained after false tags have been taken into account. False tags may arise from:

- Muons mislabeled as direct or cascade (100% wrong tag);
- Muons coming from a source other than  $B$  or  $D$  decay ( $\sim 50\%$  wrong tag); and
- Hadrons misidentified as pions ( $\sim 50\%$  wrong tag).

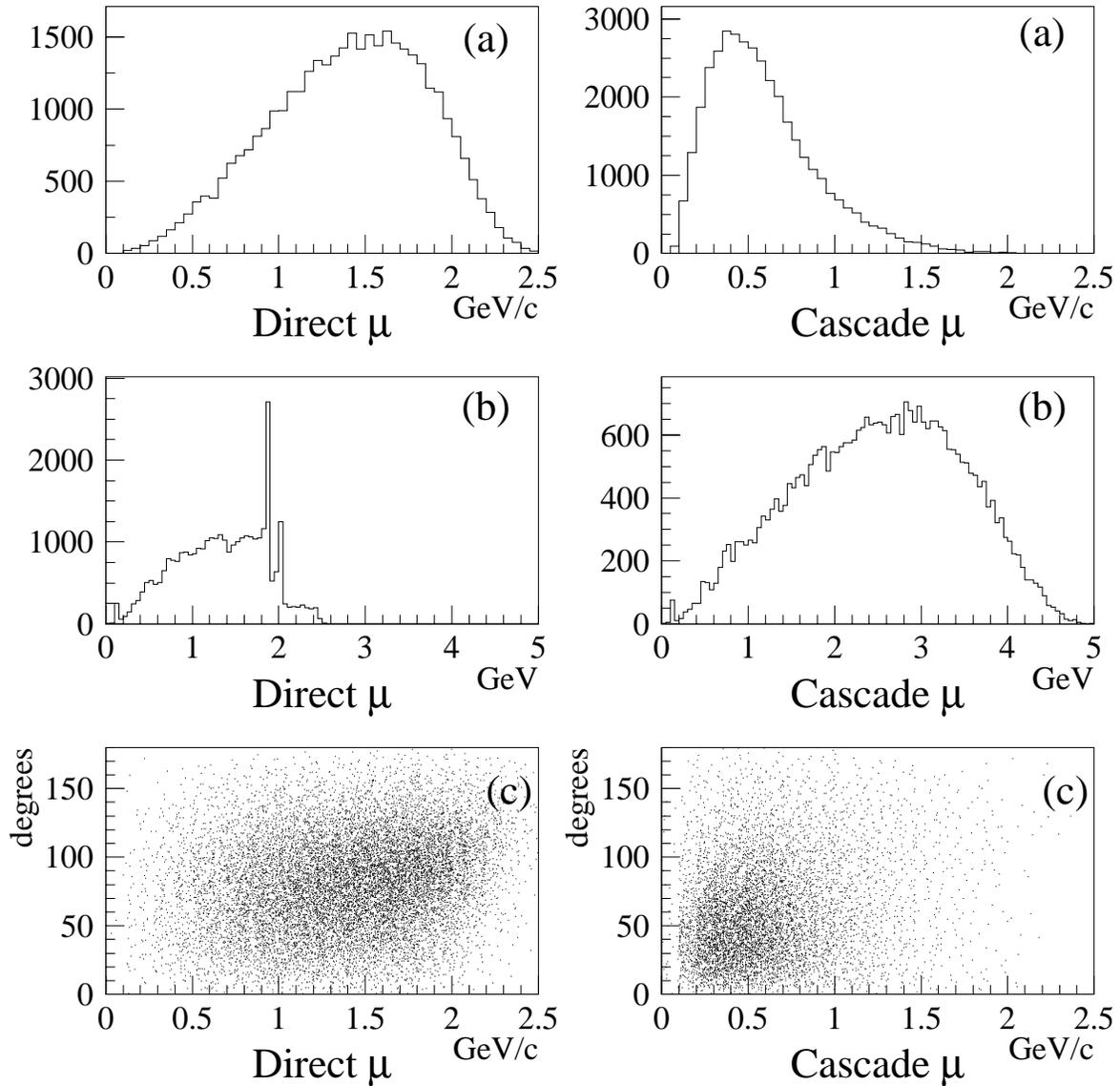
A sample of  $1.2 \times 10^6$  Monte Carlo  $B$  decays was generated in order to study the characteristics of the direct *vs.* cascade muons and to choose discriminant variables. The kinematic method used to distinguish muons from pions was applied to distinguish direct from cascade muons. The variables found to provide the highest discriminant power were identified as the center-of-mass momentum, the recoiling invariant mass, and the isolation; their distributions are shown in Figure 8-15 for both classes of muons. The probabilities that a track belongs to either class of muons can be expressed as:

$$y_{eff}^{direct} = 1 - \frac{\eta^{cascade} \prod g_i^{cascade}}{(1 - \eta^{cascade}) \prod g_i^{direct} + \eta^{cascade} \prod g_i^{cascade}}, \quad (8.2)$$

$$y_{eff}^{cascade} = 1 - \frac{\eta^{direct} \prod g_i^{direct}}{(1 - \eta^{direct}) \prod g_i^{cascade} + \eta^{direct} \prod g_i^{direct}}, \quad (8.3)$$

where the  $g_i$  are the probabilities associated with the  $i^{th}$  discriminant variable belonging to either class, and  $\eta^{direct}$  and  $\eta^{cascade}$  are the fractions of direct and cascade muons, respectively, in the total muon sample. A simple cut in the  $y_{eff}$  variables has been found to be very effective for assigning muons to the correct class.

The effective tagging efficiency reaches almost 9% for a configuration with uniform 2 cm segmentation and decreases as the steel plate thickness increases; using the graded 20-plane configuration that we propose, the figure is 8.7%; while for a 16-plane configuration, the efficiency drops below 8%.



**Figure 8-15.** Direct and cascade  $\mu$  discriminant variables: (a) center of mass momentum, (b) recoiling invariant mass, and (c) isolation, *i.e.*, the opening angle from the closest track.

### 8.3.3 $K_L^0$ Detection

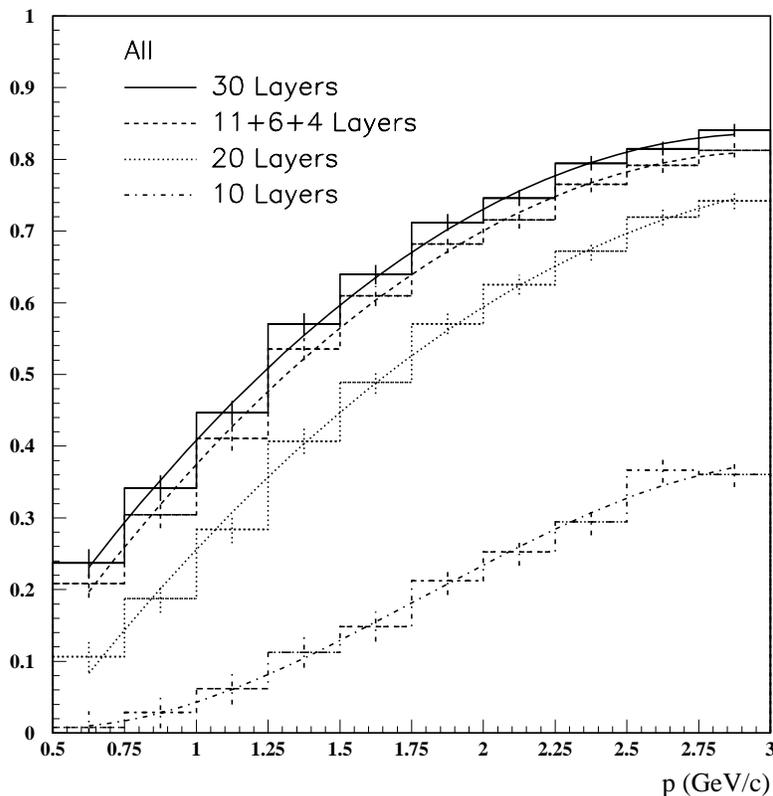
The two main issues for reconstructing the decay  $B^0 \rightarrow J/\psi K_L^0$  are the identification of IFR hits belonging to the  $K_L^0$  and the determination of the  $K_L^0$  direction from those hits. Since the  $K_L^0$ s are detected via the charged secondaries from their interaction or decay, a full Monte Carlo simulation that includes hadronic interactions is required to produce a realistic hit pattern. The detector response has been simulated with BBSIM. The FLUKA hadronic simulation package has been used because the  $K_L^0$  cross sections below 1 GeV/c in GHEISHA do not agree with the experimental data [Say68]. Full detection efficiency in both barrel and endcaps was assumed, and hit coordinates were generated assuming a 3 cm readout pitch.

A study of the effect of iron segmentation on the detection efficiency has been performed by generating single particle  $K_L^0$  events [Bal94]. Figure 8-16 shows results for the  $K_L^0$  detection efficiency as a function of momentum for four different configurations of iron plate thicknesses. Here, the  $K_L^0$  signature is taken to be at least four IFR plane hits. Detection efficiency for  $K_L^0$ s improves with segmentation; thin plates are required, especially in the inner part of the detector, where most  $K_L^0$ s start or continue to interact. However, too coarse a segmentation in the outer part of the detector also decreases the efficiency for  $K_L^0$ s that start interacting in the IFR, for which the angular resolution is much better. Figure 8-16 shows that one can obtain a sizeable improvement in the efficiency by adopting a nonuniform iron segmentation; a 21-layer nonuniform configuration has much better performance than a 20-layer uniform segmentation and is almost as good as a 30-layer one.

A study of the  $K_L^0$  identification from the IFR hits, employing pattern recognition, has been made [Wri94b]. A sample of  $B\bar{B}$  events, in which one  $B$  decays to  $J/\psi K_L^0$  and the other to any mode, has been generated. The  $K_L^0$  identification process consists of clustering the hits and eliminating clusters that are near the extrapolation of charged tracks from the central drift chamber. The remaining clusters are due mostly to neutral particles ( $K_L^0$ s, neutrons, and photons from  $\pi^0$ s) and charged hadrons. The composition of the clustered hits in the IFR is given in Table 8-1. About 80% of  $K_L^0$ s have at least one hit in the IFR. Requiring four or more hits reduces the fraction to 68%. Neutral clusters that happen to be near a charged track are sometimes eliminated. Currently, the best algorithm reduces the  $K_L^0$  efficiency from 68% to 56% due to these accidental overlaps.

While there is a correlation between the number of detected IFR hits and the momentum of the  $K_L^0$  (Figure 8-18), the distribution is quite wide and depends on whether the  $K_L^0$  interacted before reaching the IFR. This  $K_L^0$  momentum measurement is too crude to be suitable for reconstructing and cutting on the  $B$  invariant mass.

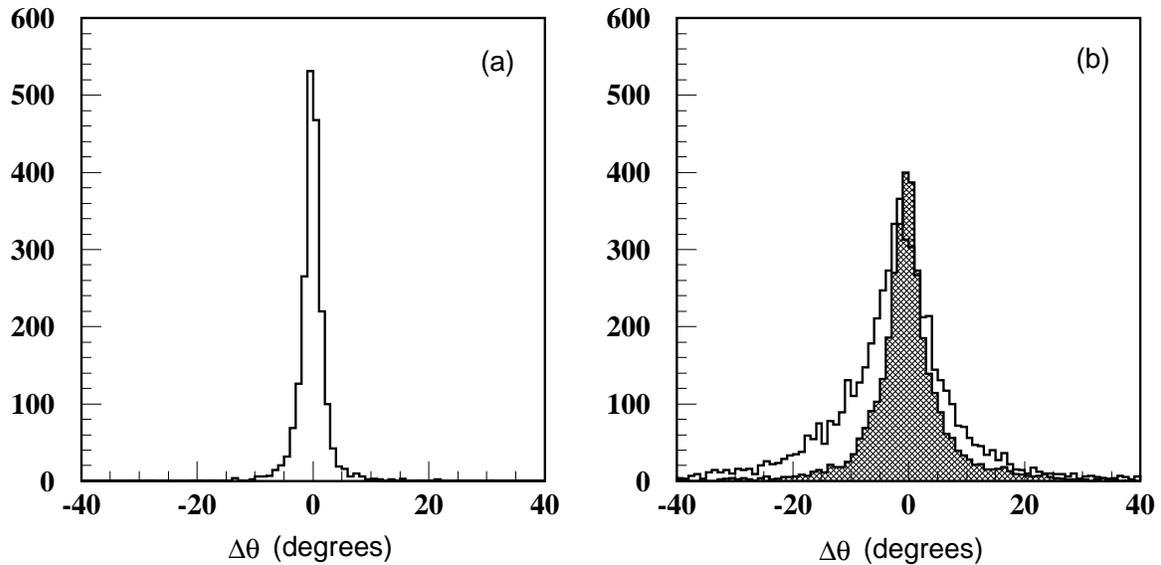
The effectiveness of background rejection cuts depends on the angular resolution of the measured  $K_L^0$  direction. For the  $K_L^0$ s that interact or decay before reaching the IFR, which make up  $\sim 70\%$  of the sample, the angular resolution is worsened, since the detectable



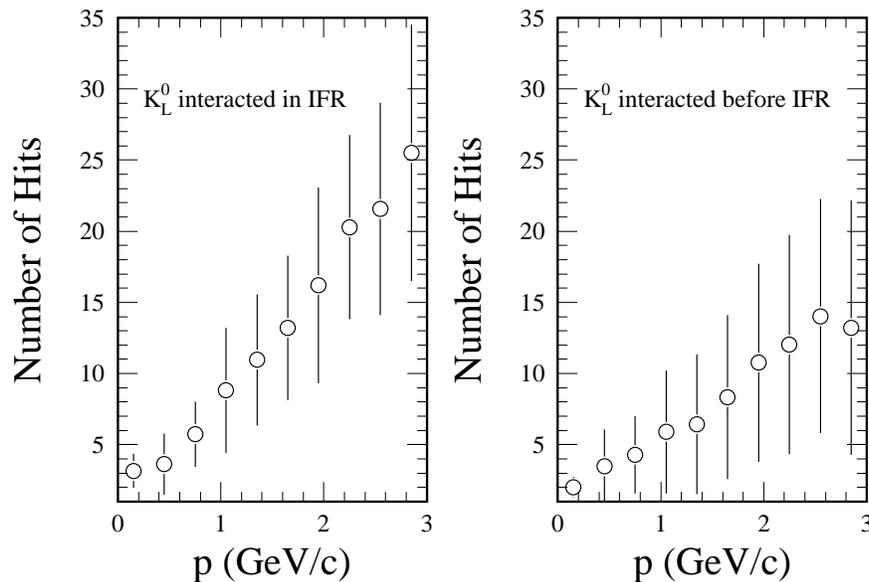
**Figure 8-16.** Efficiency for detecting  $K_L^0$ s as a function of momentum. The different sets of points represent different segmentations of the flux return iron. A solution with 21 active planes with graded separation is compared.

Criteria	$K_L^0$ eff.	Sample Composition			
		$CP$ $K_L^0$	Other $K_L^0$	Other Neutral	Other Charged
> 3 IFR Hits	68%	21%	4%	9%	66%
After Cuts	56%	77%	11%	4%	8%

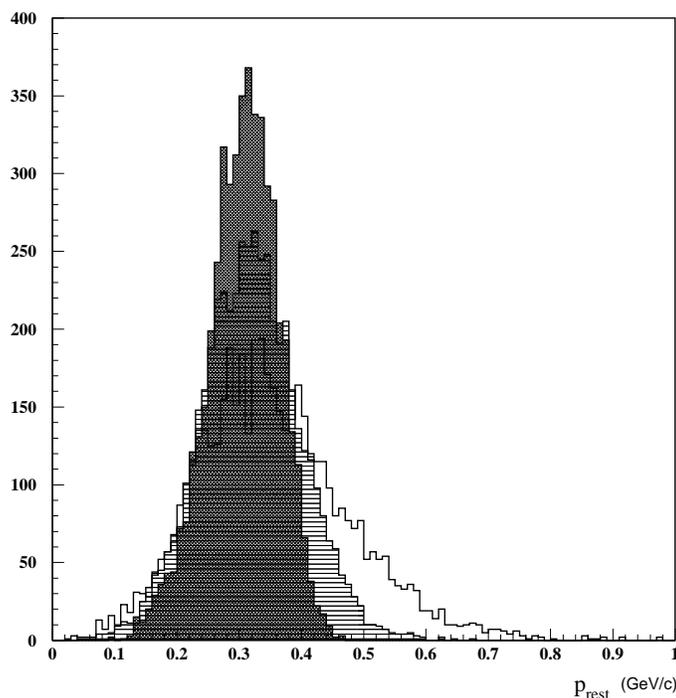
**Table 8-1.** Composition of  $B^0 \rightarrow J/\psi K_L^0$  events after pattern recognition. The second line indicates the composition after a cut that removes clusters near charged tracks.



**Figure 8-17.** (a) Angular resolution for  $K_L^0$  from  $B^0 \rightarrow J/\psi K_L^0$  determined by averaging the unsmeared IFR hit positions of  $K_L^0$ s that do not interact until reaching the IFR. (b) Angular resolution with (black) and without (open) the inclusion of a detector layer inside the coil for the  $K_L^0$ s that interacted before the IFR.



**Figure 8-18.** Average number of IFR hits as a function of momentum for  $K_L^0$  in simulated  $\Upsilon(4S) \rightarrow (B^0 \rightarrow J/\psi K_L^0)(\bar{B}^0 \rightarrow X)$  events. The error bars represent one standard deviation of the spread in the data.



**Figure 8-19.** Reconstructed momentum of the  $B$  meson in the  $\Upsilon(4S)$  rest frame ( $p_{\text{rest}}$ ) for  $B^0 \rightarrow J/\psi K_L^0$ . The narrower distribution includes only the smearing effects due to the beam energy spread and finite momentum resolution. The other two histograms show the widening due to the  $K_L^0$  angular resolution for the cases in which a detector layer inside the coil is included (dashed) or not (open).

secondaries must travel an appreciable distance and pass through additional material before reaching an active IFR detector. For example, the angular resolution for  $K_L^0$ s that interact before reaching the IFR is about a factor of 2 worse than for those that do not interact before reaching the IFR, as shown in Figure 8-17. However, the inclusion of a detector layer inside the coil significantly reduces this smearing effect, as shown in Figure 8-17(b). Therefore, such a detector has been included in our baseline design.

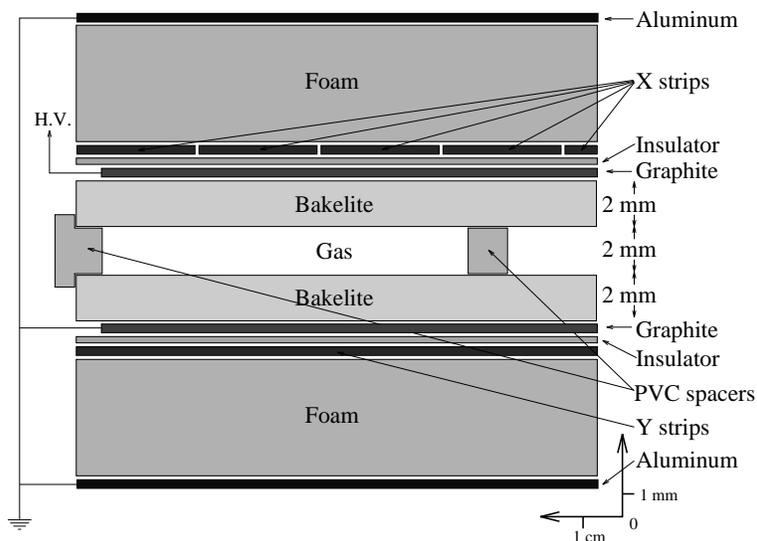


Figure 8-20. Schematic representation of RPC components.

## 8.4 Detector Design and R&D

### 8.4.1 Chamber Construction and Assembly

RPCs, shown schematically in Figure 8-20, have been chosen on the basis of their high intrinsic and geometrical efficiency, low cost, robustness, and flexibility in segmentation. We are planning to use the technology [San81, San88] developed by R. Santonico *et al.*, which is well-proven in various experiments, accelerators [Accel], and for cosmic ray physics [Cosmi], and whose use is foreseen in the first-level muon trigger systems designed for LHC experiments at CERN [LHC94]. The knowledge gained in years of R&D has already been transferred to industry, and the production capabilities are now well-matched to the large detector area required by *BABAR*.

As shown in Figure 8-20, the electrode plates are made of 2 mm-thick Bakelite (phenolic polymer) with a volume resistivity of  $10^{11}$ – $10^{12}$   $\Omega$ -cm, painted on the external surfaces with graphite of high surface resistivity ( $\sim 100$  k $\Omega$ /square) and covered by two 300  $\mu$ m PVC insulating films. The two graphite layers are connected to high voltage ( $\sim 8$  kV) and ground, respectively. Gap surfaces are treated with linseed oil, which strongly enhances RPC performance (low noise, high efficiency). Planarity of the two electrodes is assured by PVC spacers (0.8 cm<sup>2</sup>) located on a 10 cm-square grid in the sensitive volume. The active volume is filled with an argon-based gas mixture at atmospheric pressure.

A charged particle crossing the chamber produces a quenched spark which produces signals on external pickup electrodes. There is active R&D in *BABAR* and in other groups to select a nonflammable and environmentally safe gas mixture. The discharge generated by an ionizing particle is quenched by:

- UV photon absorption by hydrocarbon molecules;
- Capture of outer electrons by electronegative additives, reducing the size of the discharge; and
- Switching off the electric field around the discharge point.

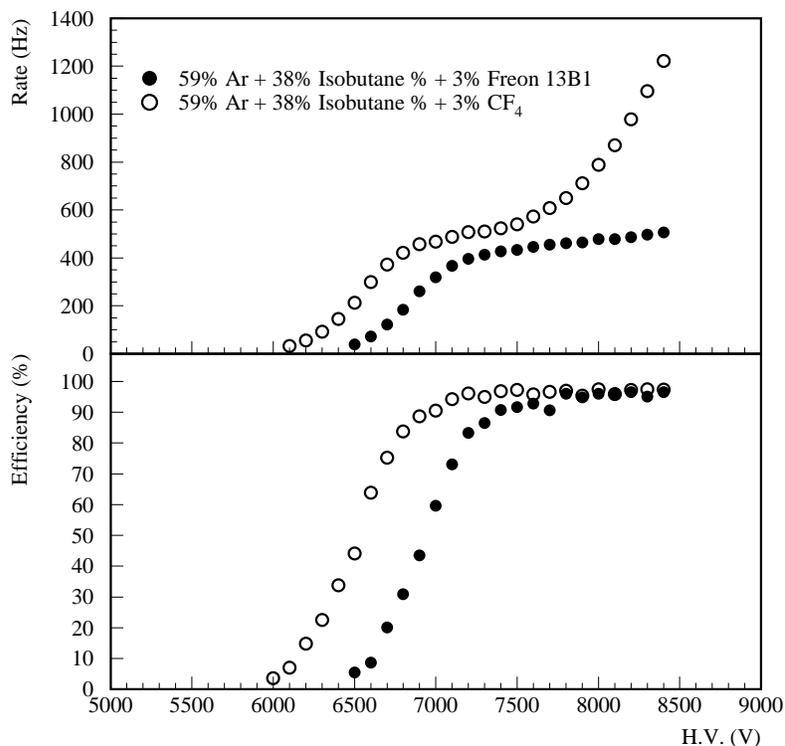
In fact, the duration of the discharge is on the order of 10 ns, while the relaxation time of the resistive electrodes is  $t = \rho\epsilon \simeq 10^{-2}$  s, so during the discharge the electrode plates behave like insulators, and a small area around the impact point shows a deadtime on the order of  $t$ .

Induced pulses are collected on two pickup planes, 29 mm wide and separated by 2 mm, made of aluminum strips glued onto plastic foils located on either side of the chamber. Strips run in two orthogonal directions to provide two-dimensional information. The induced charge is on the order of 100 pC, and the pulse has a rise time of 2 ns and a duration of 10 ns. The low event rate in the IFR, the small size ( $\sim 3$  mm<sup>2</sup>), and the short deadtime ( $\sim 10$  ms) of the area affected by each quenched spark mean that there is no efficiency problem. Stochastic noise depends on Bakelite plate resistivity and on the operating voltage; a value well below 1 kHz/m<sup>2</sup> can easily be reached.

In the upper part of Figure 8-21, the signal rate is shown as a function of high voltage for a standard mixture (Ar 59%, Isobutane 38%, Freon 13B1 3%), and with Freon 14 (CF<sub>4</sub>) instead of Freon 13B1. The Bakelite bulk resistivity is  $10^{11}$   $\Omega$ -cm in this test, and a clean cosmic ray plateau with negligible background can be seen with Freon 13B1. In the case of Freon 14, the plateau starts  $\sim 300$  V earlier, but the single rate goes up with high voltage. In any case, the singles rate is not a problem for the IFR. Assuming, for example, a singles rate of 1 kHz/m<sup>2</sup> due only to noise, we would have three random hits in a 1  $\mu$ s time window around the event time for the full  $\sim 2500$  m<sup>2</sup> IFR. If the fast timing capability of the RPC is used, by means of TDC measurements, a factor of 100 reduction can be achieved.

The efficiency plateaus for the two gas mixtures cited above are shown in the lower part of Figure 8-21. With the standard mixture, a timing resolution of  $\sigma \simeq 1$  ns is reached. As reported in Section 8.5, work is presently under way at LLNL and Naples to optimize RPC performance, with the goal of using nonflammable and environmentally acceptable gases.

Taking into account the dead areas of frames and spacers, the RPC efficiency has been measured to be higher than 96%; this value can be brought closer to 100% if, for some

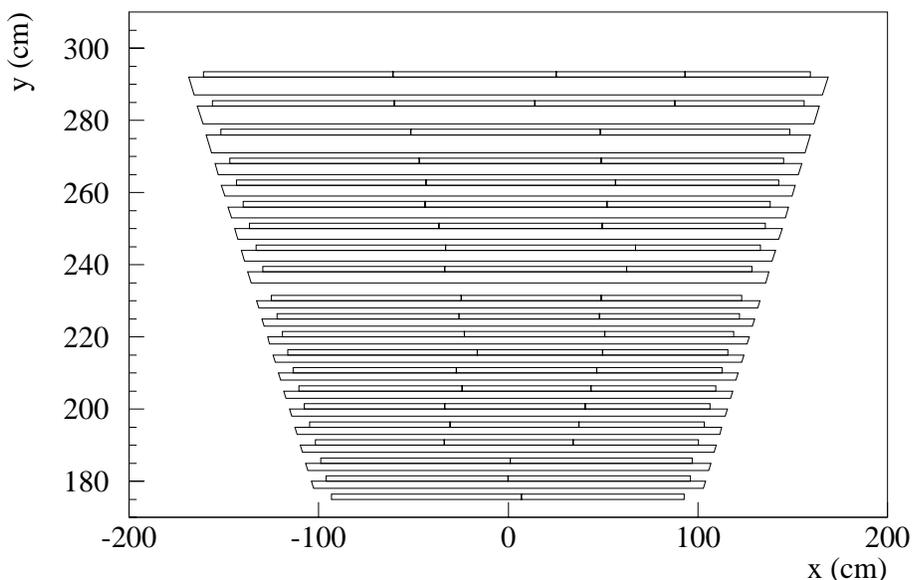


**Figure 8-21.** Singles counting rate (cosmic rays and random noise) and efficiency plateaus for the two gas mixtures 59% Ar, 38% Isobutane, and either 3% Freon 13B1 or 3% CF<sub>4</sub>.

detector layer, a two-gap approach is chosen, as has been done in the L3 [L3RPC] experiment at LEP, with the two spacer grids staggered by 5 cm.

### 8.4.2 System Layout

The production plant [Plant] that we plan to use, following the excellent experience with the L3 RPCs, has developed construction tooling which allows a maximum chamber size of 1 m by 2.2 m. Within these limits, modules can be built in practically any shape. Although it is possible to modify the tooling in order to increase the maximum width and/or length, we have studied a layout of the chambers in the gaps within the present constraints. We will carefully consider the cost effectiveness of decreasing the number of modules versus retooling. Most of the cost will be based on the number of modules built, not on the total surface covered; on the other hand, major modifications to the present tooling will increase the unit cost of the chambers.



**Figure 8-22.** Detailed layout of a barrel sextant. The angular coverage is in excess of 98% and is obtained with chambers of only five different widths. The extra space at each side will be used by steel plates that will hold the iron together.

## Barrel

In the six barrel sections, all the layers have the same length. The iron plates are 375 cm long, so all modules are 185 cm long and are joined in pairs; 2.5 cm is left at each end to allow for connectors, cables, and gas piping. The width of the iron gaps varies from 188 cm for the innermost layer to 323 cm for the outermost. The width of the modules therefore varies in order to ensure maximum geometric coverage. If the maximum width of each module is fixed at the 100 cm that the present production facility allows, it is necessary to use two adjacent modules in the inner layers and four in the outer ones. The total number of modules in each gap is therefore four (Layers 1 to 3), six (Layers 4 to 19), or eight (Layers 20 to 21), given the split in the  $z$  direction.

For each of the six barrel sections, 124 modules are needed, for a total of 744. Given such a large number of modules, it is important to find a layout which uses a minimum number of different sizes. This will translate into a cost and time saving not only in the production process but also in the subsequent steps. All of the quality control, testing, and assembly procedures will be greatly simplified if only a few basic shapes are needed. On the other hand, it is important to maximize the geometric coverage; as much as possible of the gap surface in each layer has to be covered.

Layer #	Radius (cm)	Gap width	Chan. 1 (cm)	Chan. 2 (cm)	Chan. 3 (cm)	Chan. 4 (cm)	Total Width
1	175	188.4	100	86			186
2	198	194.1	96	96			192
3	203	199.9	100	96			196
4	208	205.7	68	68	66		202
5	213	211.5	74	68	66		208
6	218	217.2	74	74	66		214
7	223	223.0	86	68	66		220
8	228	228.0	86	74	66		226
9	233	234.6	100	66	66		232
10	238	240.3	96	74	68		238
11	243	246.1	96	74	74		244
12	248	251.9	100	74	74		248
13	256	254.0	100	86	66		252
14	262	260.9	96	96	66		258
15	268	275.0	100	86	86		272
16	274	281.9	100	96	86		278
17	280	288.8	96	100	86		286
18	286	295.8	100	96	96		292
19	294	305.0	100	100	100		300
20	302	314.2	96	74	74	68	312
21	310	316.3	96	86	66	66	314

**Table 8-2.** Active module dimensions for barrel coverage.

For this optimization, we have relied on a computer calculation in which the gap widths are introduced with the constraints that each should be filled at least to the 98% level with 2 cm free in order to allow some clearance for the installation. With these criteria and a maximum module width of 100 cm, solutions are found starting with a minimum of six different sizes. One of these solutions is shown in Figure 8-22, and the relevant parameters listed in Table 8-2. In this layout, there are 744 barrel RPC modules in six different widths (Table 8-3).

The iron is segmented into 20 planes of graded thickness, from 2 to 5 cm as discussed in the previous subsections; the total iron thickness is 55 cm, and air gap between plates is 3 cm. Each barrel section is split into two parts to keep the weight within the crane capability available in the IR. Taking into account mechanical tolerances for joining the the two parts, the total IFR thickness is 120 cm. The total RPC area in the barrel is 1320 m<sup>2</sup>. In each layer, four, six, or eight modules are joined, and the strip planes glued on before insertion in

Width (cm)	Number of Modules
100	144
96	156
86	108
74	132
68	72
66	132
Number of RPCs:	744

**Table 8-3.** *Number of different modules needed for barrel coverage.*

the iron; if the strip pitch is 3 cm, there will be a total of approximately 12,000 longitudinal and 16,000 transverse strips.

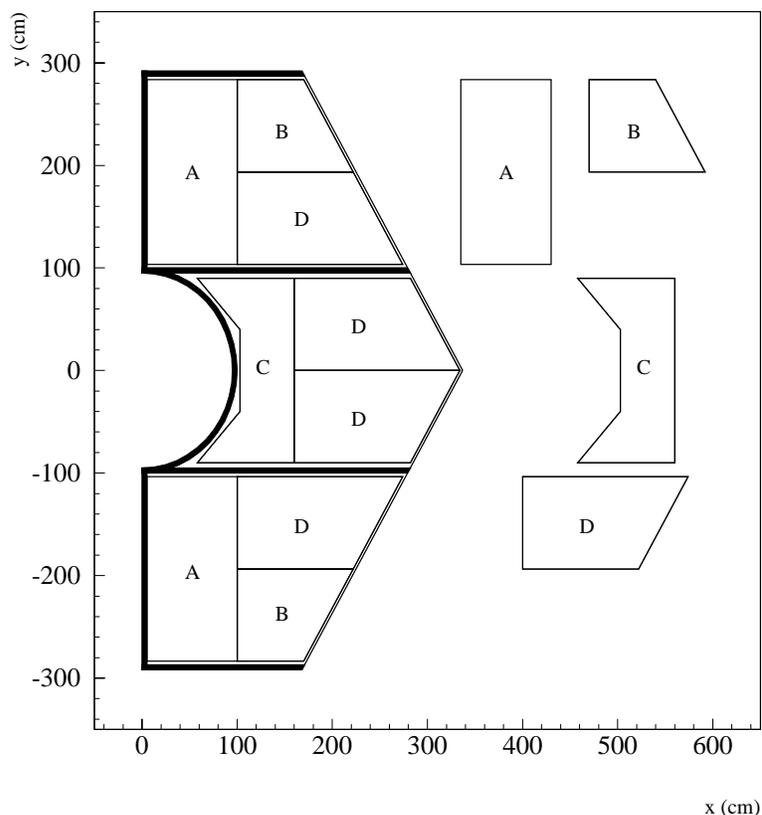
### Endcaps

The endcap geometry presents a bigger challenge for defining an optimum RPC module layout, since the chambers have irregular shapes, and the presence of stiffeners between plates imposes severe constraints on chamber size and shape. The plates all have the same dimensions apart from the hole around the beam line, which differs for each of the forward endcap gaps.

Each endcap is divided vertically into halves to allow access to the inner detector without interfering with the machine components; Figure 8-23 shows the typical layout of a half layer: the maximum width and length of the modules are 100 and 220 cm, respectively, as for the barrel layout. This scheme assumes that two horizontal stiffeners 5 cm thick are sufficient to withstand the magnetic forces of a field up to 1.5 T, as discussed in Chapter 9.

All the backward endcap plates are of the same size; the inner plates of the forward endcap have a slightly smaller hole. The position of the horizontal stiffeners is the same for all plates, at 95 cm above and below the beam line, so that the layout of the chambers is the same for all plates. A few specially shaped modules of slightly different sizes have to be used in the innermost layers of the forward endcap to obtain maximum coverage.

The total number of endcap chambers is 720: 160 will be rectangular, 95 cm by 180 cm sides; 160 smaller trapezoids, 320 larger trapezoids, 80 special shapes, in three sizes. The area of plate is approximately 27 m<sup>2</sup>, for a total of 1100 m<sup>2</sup> for the total RPC area required for the two endcaps.

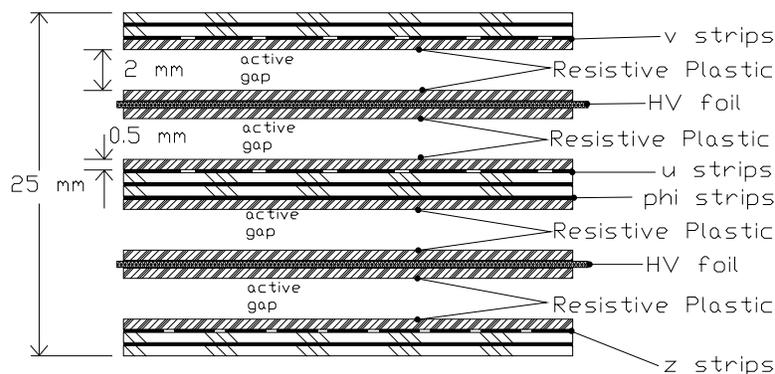


**Figure 8-23.** *Layout of the RPC modules in the endcaps.*

### Inner Cylindrical Layer

A double-gap, resistive-plastic RPC, as shown in Figure 8-24, will be used to accommodate the largest number of coordinate planes in the 3 cm allotted space inside the solenoid. Each double-gap RPC has a single high-voltage plane made of aluminum foil laminated between two pieces of resistive plastic. The ground planes on either side of each high voltage plane have the readout strips attached to their back sides. The resulting four readout planes encode  $z$ ,  $\phi$ ,  $u$ , and  $v$  coordinates with 2 cm resolution. The thickness of the entire sandwich is 2.5 cm, leaving 0.5 cm of insertion tolerance.

The inner RPC layers will be laminated together on a cylindrical mold of the final radius. The final laminate will be self supporting in its cylindrical shape. The bottom of the cylinder along the beam axis will have an open slit to accommodate the rail that supports the inner detectors. Inner RPC strips that intersect this opening ( $z$ ,  $u$ , and  $v$  strips) will be brought out to the ends of the cylinder on ribbon cable. The inner RPC front-end electronics are located on the end of the cylinder.



**Figure 8-24.** Inner cylindrical RPC design.

Although double gap RPCs have been successfully operated, cylindrical RPCs have never been constructed. Therefore, we will perform limited R&D on the cylindrical RPC prior to construction. This R&D will involve constructing and operating a small planar version of the inner RPC multi-gap configuration, followed by the construction and operation of a partial segment of the cylindrical inner RPC.

## 8.5 Gas System

### 8.5.1 Gas Composition and Flow Rates

The traditional RPC gas mix of 59% argon, 37% isobutane, and 4% Freon 13B1 is undesirable in two respects. Isobutane is flammable and Freon 13B1, an ozone-destroying substance, may prove difficult to obtain over the lifetime of *BABAR*. We are currently performing R&D at LLNL and Naples on nonflammable RPC gas mixes that use quenching agents besides F13B1. Both sites have precision gas mixers to prepare candidate mixtures that are fed to multiple RPCs, which are tested for efficiency, noise, cluster size, and time resolution with cosmic ray triggers. Preliminary results are encouraging. We have obtained efficient RPC operation with a mixture of 37% isobutane and no freon. We have also demonstrated efficient RPC operation using much smaller concentrations of isobutane (less than 8%) with F13B1. Preliminary tests show that the environmentally acceptable Freons 116 and 23, as well as SF<sub>6</sub>, have promising quenching properties. More evaluation of these and other gases is planned. Therefore, we plan to run the IFR RPCs with less than 10% isobutane and with 4–10% of the benign Freons 116, 23, 14, and SF<sub>6</sub>. The exact proportions of these gases must be experimentally determined to provide the highest efficiencies and maximal position

resolution, at the possible expense of time resolution. Typical steady state operation of the IFR will require about two volume changes of gas per week, corresponding to a flow rate of  $\sim 2$  l/min, while startup operations are expected to require about one volume change per day corresponding to a flow rate of  $\sim 6$  l/min.

### 8.5.2 Mixer

The three gases will be stored outside and transported to the mixing station through heated metal tubing. The three gases flow continuously through the mixer, passing through a filter, relief valve, pneumatic shut-off valve, regulator, and flow controller. The flow controller settings determine the mixed gas composition. A separate flow meter monitors the resulting flow rate for each gas.

### 8.5.3 Distribution

The mixed gas will be transported between the mixing station and the detector through a one-inch diameter metal tubing. Supply manifolds on the detector will distribute the gas in parallel to 14 separate groups of chambers (six barrel sextants plus four quadrants on each endcap). Within each group, chambers in a given layer will be connected in series, while the different layers will be connected in parallel. Matching return manifolds will collect the gas for venting at a remote location. A blower will be used to reduce chamber over-pressure due to the back pressure of the vent line.

## 8.6 Front-End Electronics and High Voltage

---

The front-end electronics for the IFR will be located directly in the iron gaps of the flux return as close as possible to the detectors; the relatively high number of channels foreseen ( $\sim 40,000$ ) requires that a certain amount of on-board processing be done in order to minimize the size of the outgoing cable plant. The electronic requirements for the detector are quite easy to fulfill—no pulse height data acquisition is foreseen. A single bit per struck strip has to be recorded and sent to the data acquisition system. The excellent timing properties of RPCs could also be exploited, and given the very low occupancy this subsystem is expected to have, a relatively small number of TDCs (one per layer) would be sufficient. The front-end electronics boards will include a very simple discriminator. The pulses originating from the RPC are quite large (200 mV) and do not have a large amplitude jitter, so a single transistor would be enough to standardize them and convert to TTL levels. Connection of

the strips to the boards will use twisted and flat cables, at least one set of strips (either longitudinal or transverse) will be connected to the electronics through a very short cable path so that minimal deterioration of the timing characteristics will occur. High voltage will be individually fed to each detector (a typical configuration in the barrel will include between four and eight detectors per plane) through series resistors ( $\sim 20\text{ M}\Omega$ ); small additional series resistors will be used as fuses so malfunctioning modules can be remotely disconnected from the HV supplies. Provision will be made for measuring and recording individual currents drawn from the detectors.

## 8.7 Final Assembly, Installation, and Monitoring

---

The major components of the IFR detectors (the Bakelite modules, readout strips, and the front-end electronics boards) will be manufactured at various institutions and shipped to SLAC for final assembly. Approximately 750 Bakelite modules each for the barrel and the endcaps will be manufactured in the RPC production plant in Italy. As they are manufactured, the modules will be shipped to Naples, Genoa, and Frascati for acceptance testing with cosmic ray muons with the accepted modules shipped to SLAC. The shipments will be staggered over an 18-month period with approximately 140 modules being shipped every three months. At SLAC, the modules will be processed in batches of  $\sim 12$  per week. Each module will be outfitted with a single, wide (20 cm) readout strip and inserted into a 12-chamber cosmic ray tower where they will be tested for high voltage plateau and gas integrity for approximately one day. After tests, the readout strip will be removed, and the modules corresponding to a given IFR layer will be placed side by side on a pallet, then joined at the seams with plastic film to form a single IFR detector plane. Full-length readout strips will be laminated across the modules of the detector plane. The front-end electronics cards will be mounted on the edge of the detector, which will be accessible from the open end of the barrel. For  $\phi$  planes, the 4 m-long  $\phi$  readout strips will be directly soldered to the front-end electronics cards added to the ends. On the  $z$  planes, the  $z$  strips terminate at the edge of the detector plane that is buried inside the IFR, so ribbon cable will be attached to the ends of the strips and folded at right angles to bring the signals out to the front-end electronics cards. The completed detector plane, on its pallet, will be inserted into a cosmic ray tower for detailed tests of the efficiency of each channel, using components of the *BABAR* data acquisition system. After final testing, the completed detector plane will be transported to the IR on its pallet. Each plane will be hoisted into position and slid off the pallet into the appropriate gap between the IFR plates. Spacers, wedged into the gap, will prevent the detectors from sliding within the gap. During operation, the efficiency of the IFR detector planes will be monitored by reconstructing muon pairs.

## References

---

- [Accel] A. Antonelli *et al.* (FENICE Collaboration), *Nucl. Instr. Methods* **A337**, 34 (1993);  
E. Petrolo *et al.* (WA 92 Collaboration), *Nucl. Instr. Methods* **A315**, 45 (1992);  
L. Antonazzi *et al.* (E 771 Collaboration), *Nucl. Instr. Methods* **A315**, 92 (1992);  
C. Bacci *et al.* (RD 5 Collaboration), *Nucl. Instr. Methods* **A315**, 102 (1992).
- [Bal94] R. Baldini-Ferroli, R. de Sangro, A. Palano, A. Zallo, “ $\pi/\mu$  Discrimination with the Instrumented Flux Return,” *BABAR Note* # 198 (1994).
- [Calc94] A. Calcaterra, *BABAR Note* # 144 (1994).
- [Cosmi] F. D’Aquino *et al.* (MINI Collaboration), *Nucl. Instr. Methods* **A324**, 330 (1993).  
M. Ambrosio *et al.* (Cover Plastex Collaboration), *Nucl. Instr. Methods* **A344**, 350 (1994).
- [Jaf94] D.E. Jaffe *et al.*, “Treatment of Weighted Events in a Likelihood Analysis of  $B_s$  oscillations or  $CP$  Violations,” *BABAR Note* # 132 (1994).
- [Lis94] L. Lista and S. Mele, “Study of Iron Segmentation Optimization of the I.F.R. Detector for  $\pi\mu$  separation,” *BABAR Note* # 194 (1994).
- [LHC94] ATLAS Collaboration, *Nucl. Instr. Methods* **A340**, 466 (1994);  
CMS Collaboration, *Nucl. Instr. Methods* **A334**, 98 (1994).
- [L3RPC] A. Aloisio *et al.*, Proceedings of the Sixth Pisa Meeting on Advanced Detectors, Isola d’Elba, Italy (1994); submitted to *Nucl. Instr. Methods*.
- [Plant] General Tecnica, Colli (FR) Italy.
- [Pia94] M.G. Pia, “ $B$  Tagging with Muons,” *BABAR Note* # 192 (1994).
- [PDG94] L. Montanet *et al.* (Particle Data Group), *Phys. Rev.* **D50** (1994).
- [San81] R. Santonico and R. Cardarelli, *Nucl. Instr. Methods* **A187**, 377 (1981).
- [San88] R. Santonico and R. Cardarelli, *Nucl. Instr. Methods* **A263**, 20 (1988).
- [Say68] G.A. Sayer and E.F. Beall, *Phys. Rev.* **169**, 1045 (1968).
- [Wri94] D.M. Wright, “BEGET: The  $B$ -Factory Event Generator Version 21,” *BABAR Note* # 149 (1994).

- 
- [Wri94b] D.M. Wright, “ $K_L^0$  identification in  $B^0 \rightarrow J/\psi K_L^0$  Decays,” *BABAR Note # 201* (1994).



---

---

# Magnet Coil and Flux Return

---

## 9.1 Physics Requirements and Performance Goals

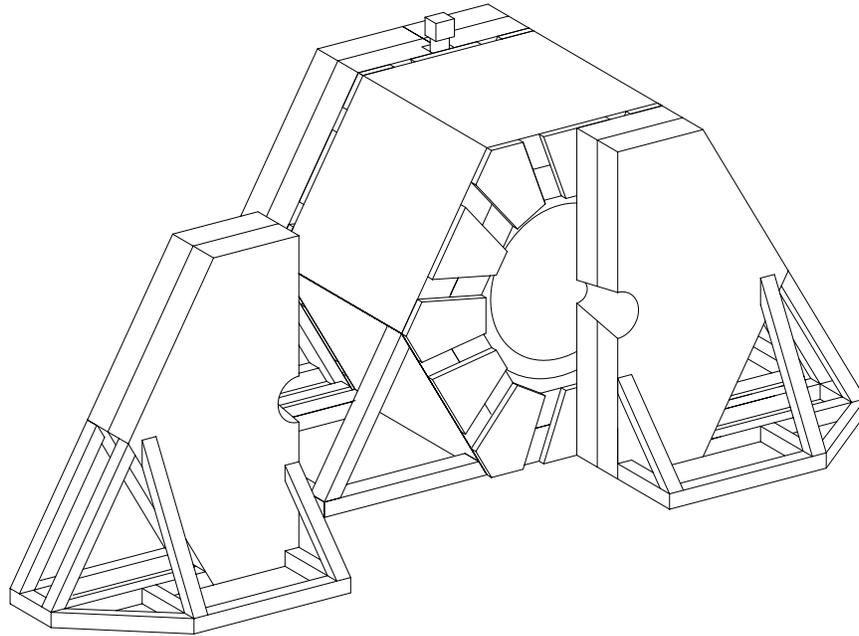
---

The *BABAR* magnet is a thin, 1.5 T superconducting solenoid within a hexagonal flux return, as shown in Figure 9-1. Detector performance criteria and geometry considerations drive the design of the solenoid and the flux return. The magnitude and uniformity specifications for the magnetic field are derived from drift chamber track finding and momentum resolution requirements. Studies of  $B^0 \rightarrow \pi^+\pi^-$  suggest that a magnetic field of 1.5 T is necessary to achieve a mass resolution of 21 MeV/ $c^2$ . The combined thickness of the vertex detector, drift chamber, particle identification system, electromagnetic calorimeter, and appropriate clearances set the solenoid inner diameter. Solenoid length is also determined by the length of the nested subsystems. The solenoid thickness limits the momentum threshold for detecting muons and the efficiency of  $K_L^0$  detection within the instrumented flux return.

The segmented geometry of the flux return allows tracking of muons and provides for detection of  $K_L^0$ s with adequate angular resolution. The total thickness of the steel layers in the barrel and end door is determined both by the minimum steel required to avoid magnetic saturation and by the need for sufficient thickness to ensure that most of the pions interact in the steel. The minimal steel thickness to prevent pion punch-through is 55 cm ( $\sim 3.6 \lambda_{int}$ ). Plate segmentation and thicknesses are specified both for efficient identification of  $K_L^0$ s and for distinguishing muons from pions based on range measurements. For more information on the meson detection system refer to Chapter 8.

The overall thickness of the flux return is the sum of both the steel thickness and the number and thickness of the RPC layers. Cost is also a factor in determining the number of RPC layers. Separation and movement of the end doors are constrained by beam line components and by the need to provide ready access to detector subsystems.

The physics performance and operational requirements for the solenoid and flux return (Table 9-1) are similar to those of many operating detector magnets (Table 9-2).



**Figure 9-1.** *Geometry of the solenoid within the flux return.*

## 9.2 Overview

---

The design of the superconducting solenoid for the *BABAR* detector is conservative and within the state of the art [Des85, And82, Coils] for detector magnets. It is based on the experience gained over the past 15 years with thin superconducting solenoids. Although specifically tailored to meet the requirements of *BABAR* (Table 9-1), this design is similar to many operating detector magnets. A common feature of all these magnets is the use of aluminum-stabilized conductors that are indirectly cooled by liquid helium pipes connected to an aluminum alloy support structure. This technique was developed for CELLO, the first thin solenoid, and has been improved in subsequent designs. Table 9-2 shows the main characteristics of some of these solenoids compared to the *BABAR* design. All of these designs used a Rutherford-type cable made of NbTi superconductor encased in an aluminum stabilizer that allows for adequate quench protection.

The *BABAR* detector schedule identifies the magnet as a critical procurement item. The three-and-one-half-year-long critical path is formed by: solenoid design and procurement; assembly with the flux return; verification testing and mapping; and detector subsystem installation and commissioning. While these task durations may be shortened, such reductions expose the project to higher budget and schedule risks. The solenoid design and fabrication duration

<b>Solenoid Requirements</b>	
Central induction	1.5 T
Field uniformity in the tracking region	$\pm 2\%$
Nuclear interaction length	$0.25\text{--}0.4 \lambda_{int}$
Cryostat inner radius	1400 mm
Cryostat outer radius	1730 mm
Minimize thermal cycling	✓
Comply with ES&H requirements	✓

<b>Flux Return Requirements</b>	
Provide an external flux path for a 1.5 T field	✓
Provide 3 cm spacing between the steel plates for IFR instrumentation	✓
Provide the gravitational and seismic load path for the barrel detector components to the concrete foundation	✓
Fit in IR-2 (3.5 m radial distance from beam axis to the concrete floor)	✓
Movable end doors to allow access inside the barrel	✓
Comply with ES&H requirements	✓

**Table 9-1.** *Physics performance and operational requirements.*

of 24–26 months requires the contract to be awarded in the fall of 1995 to meet the overall detector schedule.

The magnet cryostat will be designed, fabricated, and inspected according to the intent of the ASME Boiler and Pressure-Vessel Code, Section VIII, Division 2 [ASME94], but will not be code-stamped. The magnet will be subject to seismic design requirements described in the SLAC Seismic Design Manual for mechanical systems [SDM91]. The magnet design will also follow the requirements outlined in the Safety Analysis Document (SAD), which will address ES&H issues. For steel structures, the allowable design stresses follow the standard guidelines as specified in the AISC Manual of Steel Construction, 9th edition. Bolted connections and fasteners will conform to their recommended torques and allowable stresses depending on

	CDF	ZEUS	CLEO-II	ALEPH	<i>BABAR</i>
Location	FNAL	DESY	Cornell	CERN	SLAC
Manufacturer	Hitachi	Ansaldo	Oxford	Saclay	?
Year Completed	1984	1988	1987	1986	1997
Central Field (T)	1.5	1.8	1.5	1.5	1.5
Inner Bore (m)	2.86	1.85	2.88	4.96	2.80
Length (m)	5	2.5	3.48	7	3.46
Stored Energy (MJ)	30	12.5	25	137	25
Current (A)	5000	5000	3300	5000	7110
Total Weight (t)	11	2.5	7.0	60	6.5
Radiation Length	0.85	0.9	n/a	1.6	1.4 max
Conductor Dimensions (mm)	3.89×20	4.3×15, 5.56×15	5×16	3.6×35	3.2×30
Current Density (A/mm <sup>2</sup> )	64	78	42	40	74

**Table 9-2.** Comparison of solenoids similar to *BABAR*.

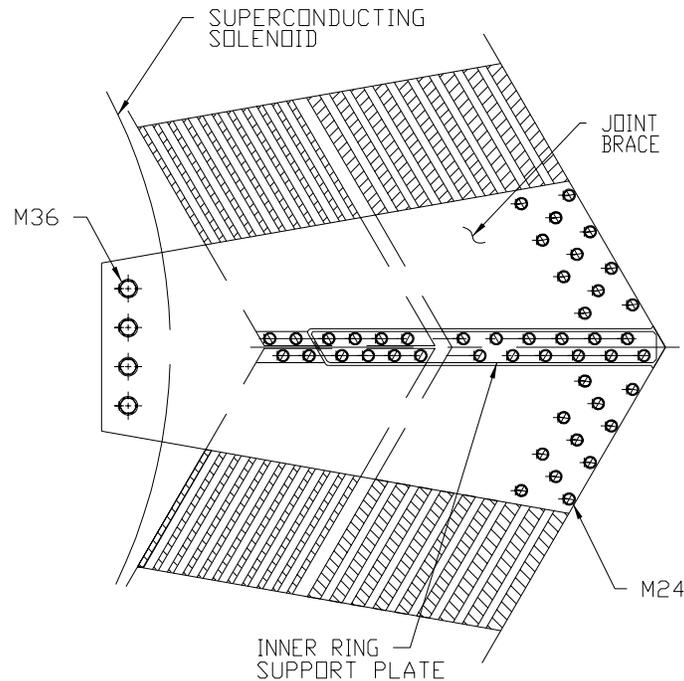
the connection. The flux return is fabricated from ASTM A36 structural steel plates or a material with similar mechanical and magnetic properties.

### 9.2.1 Description of Key Interfaces

**Superconducting Solenoid and Flux Return.** The radial distance between the outer diameter of the solenoid and the inner surface of the barrel flux return is 50 mm. The solenoid weight and magnetic forces are transmitted to the inner and outer hexagonal rings of the flux return as shown in Figure 9-2. This attachment, located at the vertical mid-plane of the detector, also provides the load path of the inner detector components to the barrel flux return.

The backward end doors provide a chase for the cryostat chimney. The chase is 400 mm wide and extends 400 mm into the backward end doors.

**Barrel and End Door.** Both ends of the barrel flux return have a 60% solid steel contact area at the interface with the end doors. This area is composed of the 150 mm-thick inner ring support plates, 150 mm-thick joint braces, and 150 mm-thick steel gap filler plates. The remaining 40% open area on the barrel ends is reserved for cabling and utilities from the inner detector components. The end doors are attached to the barrel with tie plates that are bolted to the end door structure and to the barrel.



**Figure 9-2.** *Superconducting solenoid support bracket attached to the mid-plane of the flux return.*

**Particle Identification System.** A vertical slot between the backward end doors permits the support structure for the DIRC to penetrate into the detector. This structure also supports the backward beam magnets Q2, Q4, Q5, and the backward flux return field shaping plug located physically inside the DIRC. The final design details of the DIRC and the mounting of the backward beam magnets are not yet fully resolved.

**Forward Q2 Beam Magnet Shielding.** The forward beam magnet Q2 is physically located within the forward end doors. A specially designed, three-piece, conical magnetic shield plug is mounted to the end doors to isolate Q2 from the detector magnet. The shielding plugs are split along their vertical centerline, and each half is attached to a half-round mounting flange that is bolted to the face of each forward end door.

**Inner RPC Detector and Solenoid.** There is an RPC detector located between the calorimeter and the solenoid. This RPC detector attaches directly to the inner diameter of the solenoid cryostat with a 20 mm clearance gap between the RPC detector and the calorimeter.

**Movable End Door Skids and the Beam Line.** The end doors are mounted on skids equipped with rollers so that they can be moved away from the barrel for maintenance access. The end door skids move on tracks installed in the floor of IR-2. The end doors clear the beam line magnets, vacuum pumps, magnet stands, and other beam line equipment during door opening.

**External Platforms, Stairways, and Walkways.** The external platforms necessary to install and service electronic racks and cryogenic equipment are supported from the flux return. The requirements of these components have not yet been determined.

## 9.3 Summary of Projected Magnet Performance

---

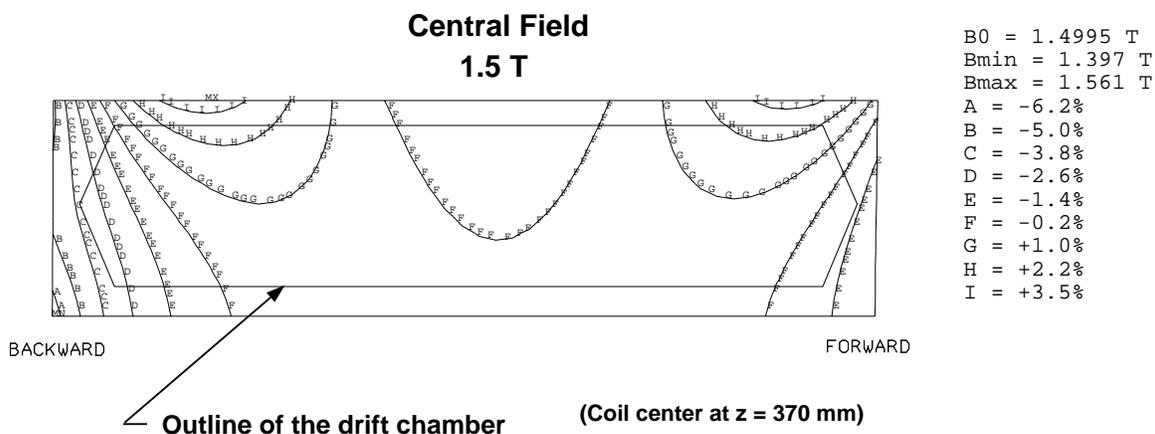
### 9.3.1 Central Field Magnitude and Coil Performance

The magnetic field of 1.5 T is obtained by energizing the solenoid with a constant current of 7110 A. The conductor is operated at 45% of the critical current, with a peak field in the conductor of 2.5 T. This gives a large safety margin.

Magnetic uniformity is achieved by doubling the current density in regions at both ends of the solenoid. This is done by adding more aluminum stabilizer to the central region conductor, which reduces the current density there. Figure 9-3 shows the field uniformity in the central region. The areas in which the field nonuniformity is greater than 2% are small and are located in regions in which they do not affect the performance of the drift chamber. In addition, once the solenoid parameters are optimized, the corners of the drift chamber should also be within  $\pm 2\%$  of 1.5 T.

The radial pressure on the conductor during operation is 1.5 MPa in the high current-density regions and 0.78 MPa in the central region of the conductor. An aluminum support cylinder surrounds the coiled conductor to react against these radial pressures and keep the conductor from yielding.

The integrated axial force on the winding is 3.5 MN. The conductor winding and support cylinder are mechanically coupled by an epoxy bond. This epoxy bond allows some of the axial load to be transmitted in shear to the outer aluminum cylinder, which keeps the conductor from yielding. There is an axial 18 kN de-centering force applied to the conductor winding due to an asymmetry in the iron, mainly due to the differences in the forward and backward Q2 shielding.



**Figure 9-3.** Field uniformity inside the drift chamber. The central field is within  $\pm 2\%$  of 1.5 T. Once the solenoid parameters are optimized, the corners of the drift chamber should also be within  $\pm 2\%$  of 1.5 T. (The BABAR coordinate system is defined in Section 14.2.)

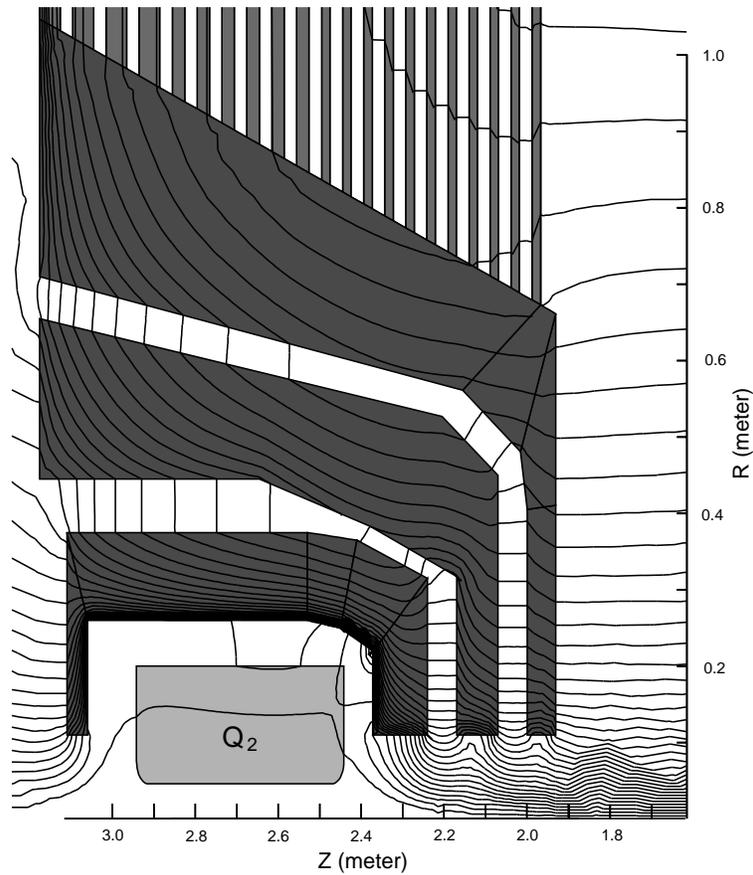
### 9.3.2 Shielding of Forward Q2

The high luminosity of PEP-II requires that Q2, a non-superconducting septum quadrupole magnet, be placed close to the interaction region. Consequently, Q2 is situated within the forward end of the instrumented end door flux return (Figure 9-15). Q2 is subject to induced multipole moments resulting from the magnetic field in its vicinity, the octopole moment being the major one. The luminosity is critically dependent upon the Q2 field quality. Hence, it is necessary to provide adequate shielding of the BABAR central field to ensure the quality of the quadrupole field in Q2.

The present Q2 shield design is shown in Figure 9-4, where the three high-permeability shields (dark gray) surrounding Q2 (light gray) are visible, along with the logarithms of the magnetic equipotentials. The present design appears to shield Q2 from the detector magnet but does not provide a safety margin, should actual parameters, *e.g.*, steel permeability, differ from those used in the magnetic modeling programs. Work, including a full three-dimensional analysis, is continuing to improve the shield design.

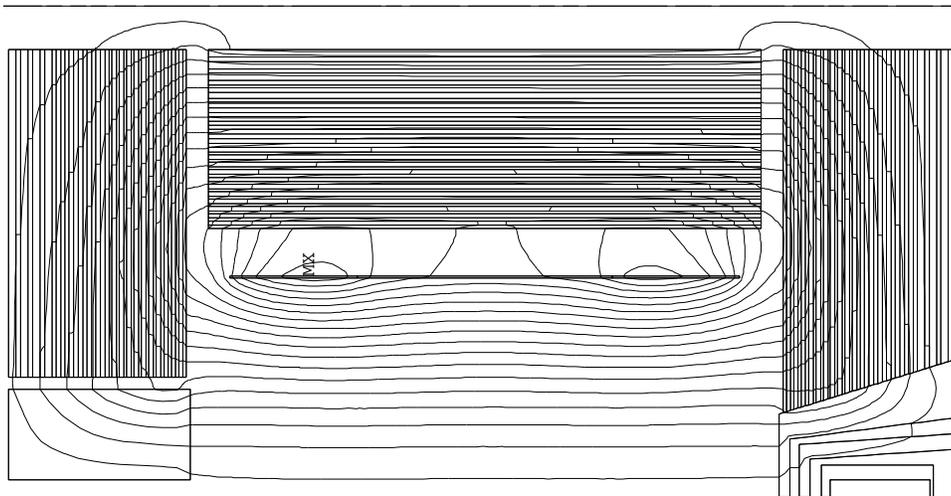
### 9.3.3 Flux Return

The flux return assembly provides an external flux path for the magnetic field of the superconducting solenoid. Figure 9-5 shows the flux lines from the magnetic analysis. There are large body forces in the first few plates as a result of the magnetic field. Figure 9-6 shows

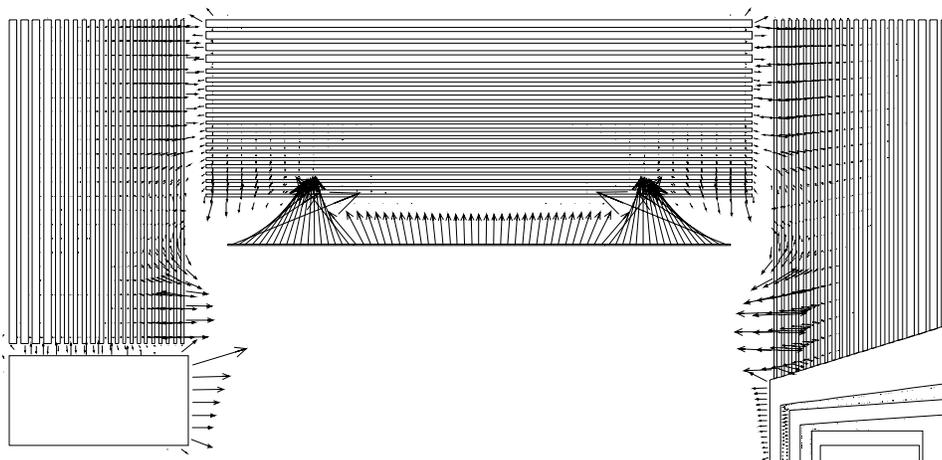


**Figure 9-4.** *Logarithms of magnetic equipotentials in the region of the Q2 septum quadrupole.*

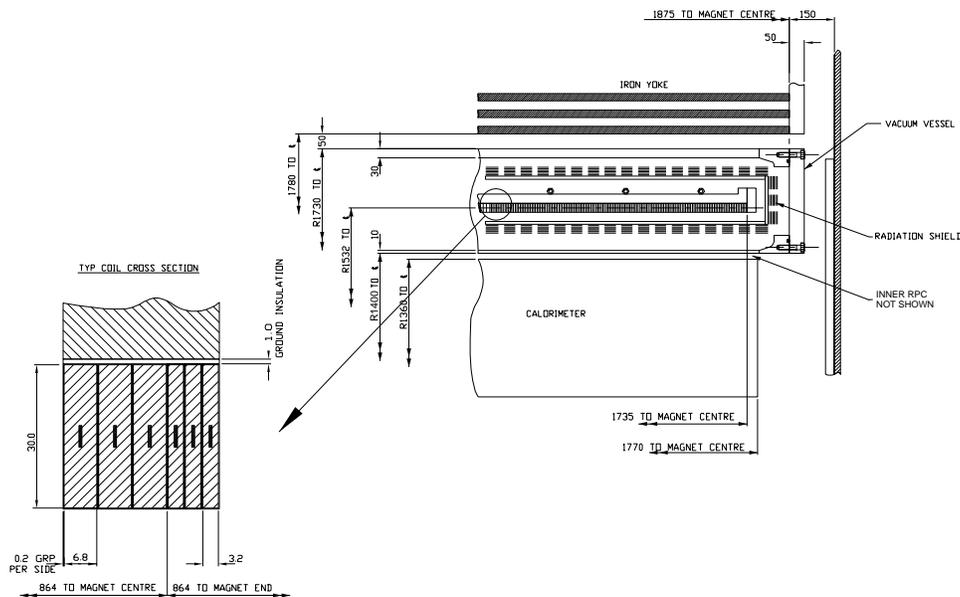
the force vectors in the barrel and end doors of the flux return. Preliminary results show that stiffeners are needed in the end door plates to resist these axial forces. The present design has two stiffeners in each end door. As the end door design is refined, the locations and number of the plate stiffeners may change to keep the deflections and stresses in the plates within acceptable levels.



**Figure 9-5.** *Vector potential lines in the full detector region. The potential lines are horizontal in the region of the drift chamber, representing good field uniformity.*



**Figure 9-6.** *Forces on the flux return plates, solenoid, and end plugs as a result of the magnetic field. The forces are the highest in the first few plates of the flux return.*



**Figure 9-7.** Overall view of the solenoid showing the cryostat, both conductor cross sections, and radiation shielding.

## 9.4 Superconducting Solenoid

### 9.4.1 Magnetic Design

This section describes the main features of the superconducting solenoid. For a more detailed description of the solenoid, refer to Reference [BF95]. A cross section of the solenoid is shown in Figure 9-7, and parameters are given in Table 9-3.

The magnetic analysis is based on a two-dimensional axially symmetric model. This model includes the solenoid, flux return plates, forward Q2 shield, backward shield, and the 150 mm gap between barrel and end doors.

The backward shield is designed to accommodate the DIRC. Its main functions are to improve the field uniformity in the backward region of the drift chamber and to balance the magnetic force on the solenoid due to the forward Q2 shield. A detailed design of this shield is underway. The iron properties used for computation (ANSYS code [ANS95]—two-dimensional magnetic element) are those of hot-rolled carbon steel.

The magnet design provides a magnetic field of 1.5 T with a uniformity of  $\pm 2\%$  in the tracking region. This is obtained by grading the current density of the solenoid in three

Parameter	Value
Central Induction	1.5 T
Conductor Peak Field	2.5 T
Uniformity in the Tracking Region	$\pm 2\%$
Winding Length	3456 mm
Winding Mean Radius	1532 mm
Amp Turns	$5.1192 \times 10^6$
Operating Current	7110 A
Inductance	0.985 H
Stored Energy	25 MJ
Total Length of Conductor	7000 m

**Table 9-3.** Overall coil parameters.

regions connected in series. The central region is 1728 mm in length with 240 turns. Two end regions are 864 mm in length with 240 turns each. The current density in the end regions is twice that of the central part. A better field uniformity may be obtained by reducing the axial length of the two end regions and increasing the current to generate the same field, but this would cause a reduction in stability against thermal disturbance. For the initial design, the maximum allowed current density in the conductor has been limited to the maximum currently attainable for magnets of this kind, *i.e.*,  $\sim 80 \text{ A/mm}^2$  (ZEUS magnet). Thus, a cross section of  $\sim 90 \text{ mm}^2$  for the smaller conductor corresponds to a maximum current of  $\sim 7000 \text{ A}$ .

Figure 9-5 shows the graph of the field lines over the full detector region. Figure 9-3 shows the field uniformity in the central region defined by  $r < 800 \text{ mm}$  and  $-1170 < z < 1910 \text{ mm}$  with respect to the IP. The magnetic field is essentially symmetric. A field uniformity of  $\pm 2\%$  is obtained. Field uniformity is required up to  $z = 1670 \text{ mm}$  in the forward region, and the present design provides a uniform field up to  $z = 1910 \text{ mm}$ , providing a factor of safety. Further adjustment of the backward shield geometry may improve field symmetry, which would improve field uniformity in the backward region.

## 9.4.2 Cold Mass Design

### Aluminum-Stabilized Conductor

The conductor is composed of a superconducting Rutherford cable embedded in a very pure aluminum matrix by a coextrusion process that ensures a good bond between aluminum and superconductor. Table 9-4 shows the main parameters of the conductor.

The operating current for this conductor is 45% of the critical current at the peak field, giving a large safety margin. In the case of local heating up to 5.2 K, there is still a significant margin on the critical current ( $I = 0.6I_c$ ). At 2.5 T, the conductor-critical temperature is  $T_c = 8.23$  K, and the current sharing temperature is 6.5 K. A simple method to evaluate the stability of the winding consists of considering the enthalpy margin per unit length between the operating and the sharing temperature. This stability parameter for the *BABAR* solenoid is 0.5 Joule/m, which is the same value obtained for the ALEPH magnet.

The cross section of the conductor is  $3.2 \times 30$  mm<sup>2</sup> for the higher current density regions and  $6.8 \times 30$  mm<sup>2</sup> for the central region. The coil winding can be made using six 1200 m lengths of conductor, requiring five electrical joints. Each joint, made by either welding of the aluminum matrix or soft soldering, has a resistance less than  $5 \times 10^{-10}$   $\Omega$  after the electro-deposition of copper, limiting the power dissipation to a few milliwatts.

### Winding Support

The winding will be supported by an external aluminum alloy cylinder similar to other existing detector magnets. The winding support is designed for all aspects of force containment, *i.e.*, its weight and the radial and axial magnetic forces. Figure 9-6 shows these magnetic forces on the solenoid and the flux return.

The highest radial pressure, 1.5 MPa, is generated in the high current density regions at the ends of the coils. A pressure of 0.78 MPa is generated in the central region. An aluminum alloy (5083) support cylinder surrounds the coiled conductor to react against these radial pressures and prevent coil movement. The radial pressure is applied through the coil winding to the outer support cylinder. Assuming that the pure aluminum coil winding can be loaded to a stress of 20 MPa, the stress in the outer support cylinder can be estimated, using a thickness of 30 mm, to be 65 MPa. This stress of 65 MPa is well below the 170 MPa yield stress of the 5083 aluminum alloy. This simple analysis suggests that after the first charge the pure aluminum stabilizer will never be stressed beyond the elastic limit. This will help prevent premature quenching during coil energizing. A more complete analysis including both the Rutherford cable and the fiberglass epoxy insulation is not expected to significantly change the calculated stresses.

Parameter	Value
Conductor Type	NbTi, Pure Al-stabilized, Co-extruded
Aluminum RRR	> 500
Conductor Unit Length	1.2 km
Number of Lengths	6
Dimensions: Bare	$3.2 \text{ and } 6.8 \times 30.0 \text{ mm}^2$
Insulated	$3.6 \text{ and } 7.2 \times 30.4 \text{ mm}^2$
Superconducting Cable	Rutherford
Dimensions	$9 \times 1.23 \text{ mm}^2$
Strands Diameter	0.84 mm
Number of Strands	20
Cu/Sc	1.8
Filament Diameter	$20 \mu\text{m}$
$I_c$ ( $B = 2.5 \text{ T}$ , $T = 4.5 \text{ K}$ )	> 16 kA
Insulation Type	Fiberglass Tape
Insulation Thickness	0.4 mm

**Table 9-4.** Conductor parameters.

An integrated compressive axial force of 3.5 MN is induced in the winding. The distribution of the axial force within the coil is complex. The end regions, with higher current density, compress the central part with 5.4 MN. The central part is axially stressed outward by a force of 1.9 MN. For preliminary calculations of the axial stress, the maximum force (6 MN) was considered. This would lead to an axial stress of 18 MPa on the pure aluminum, with only the winding supporting the axial forces. However, if the axial force is transmitted to the outer cylinder, the stress is lowered by a factor of two, with the pure aluminum working well below its elastic limit. In this case, the shear stress between the winding and outer supporting cylinder is less than 2 MPa. This low value of shear stress will allow the winding and support cylinder to be mechanically coupled through an epoxy impregnation without applying any axial prestress to the winding (as was done for the ZEUS magnet). Epoxy impregnation can support a shear stress higher than 20 MPa, providing a high safety margin. This leads to a simplification and cost saving in the winding fabrication.

The current design causes axial de-centering forces on the coil due to the iron asymmetry and a residual force of 18 kN is applied to the winding. A more careful design of the backward shield can help reduce the amount of this residual axial force by a factor of two or three.

Offset forces have been calculated as follows. An axial displacement of the solenoid of 10 mm causes an axial force of 98 kN in the direction of the displacement. A radial misalignment

of 10 mm gives rise to a force of 89 kN. These values will be taken into consideration in designing the support system and should not present any significant problems.

Table 9-5 shows the main features of the cold mass. The values are given at a temperature of 4.5 K. The dimensions at room temperature are higher by a factor of approximately 1.004.

### Electrical Insulation

Electrical insulation is an important aspect of solenoid design and manufacture. Two categories of insulation are required: ground plane insulation between the coil and support cylinder, and turn-to-turn insulation.

- The ground plane insulation must operate at relatively high voltages during quench conditions and will be subjected to strict QA controls. The design of the quench protection systems is based on a maximum voltage to ground of 250 V. The ground plane insulation will be made by a 1 mm layer of fiberglass epoxy laminate that is bonded to the support cylinder before winding. The insulation will be fully tested at 2 kV before winding.
- The conductor will be insulated with a double wrap of  $\sim 0.1$  mm glass tape during winding to give an insulation thickness of 0.2 mm. The resulting turn-to-turn insulation thickness will be 0.4 mm and will be fully impregnated in the bonding process. Electrical tests will be carried out during winding to detect any failure of insulation. The tests will include regular/continuous testing for turn-to-turn and turn-to-ground insulation.

### 9.4.3 Quench Protection and Stability

#### Protection Concept

The solenoid will be protected by an external dump resistor which will determine the current decay under quench conditions and allow extraction of  $\sim 75\%$  of the stored magnetic energy. The quench protection concept is shown in Figure 9-8, and quench parameters are given in Table 9-6. The protection concept is based on two main criteria.

- A voltage limit of 500 V across the solenoid applies during fast discharge. Center-tapping of the fast dump resistor to ground will limit the voltage to ground to 250 V. The center-tapped resistor will also allow the measurement of ground leakage currents as a safety and diagnostic tool.

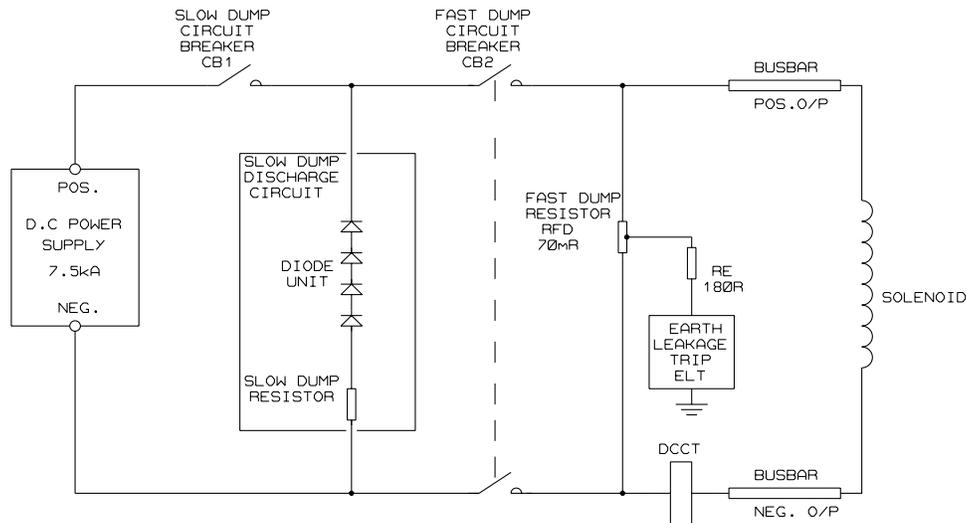
Parameter	Value
Winding:	
ID	3033.8 mm
OD	3095.2 mm
Length	3456.0 mm
Weight	2.4 tonne
Supporting Cylinder:	
Material	Al alloy 5083
ID	3456.40 mm
OD	3516.40 mm
Length	3506.00 mm
Weight	2.5 tonne
Ground Insulation:	
Material	Fiberglass epoxy
Thickness	1.0 mm
Total Solenoid Weight:	4.9 tonne
Nuclear Interaction Length: (Assuming Aluminum)	
Maximum	$0.19 \lambda_{int}$
Minimum	$0.15 \lambda_{int}$

**Table 9-5.** Cold mass (4.5 K) parameters.

- An upper temperature limit of 100 K applies during quench conditions. This limit will give very good safety margins against peak temperature rise and thermally induced stresses at quench.

### Quench Analysis

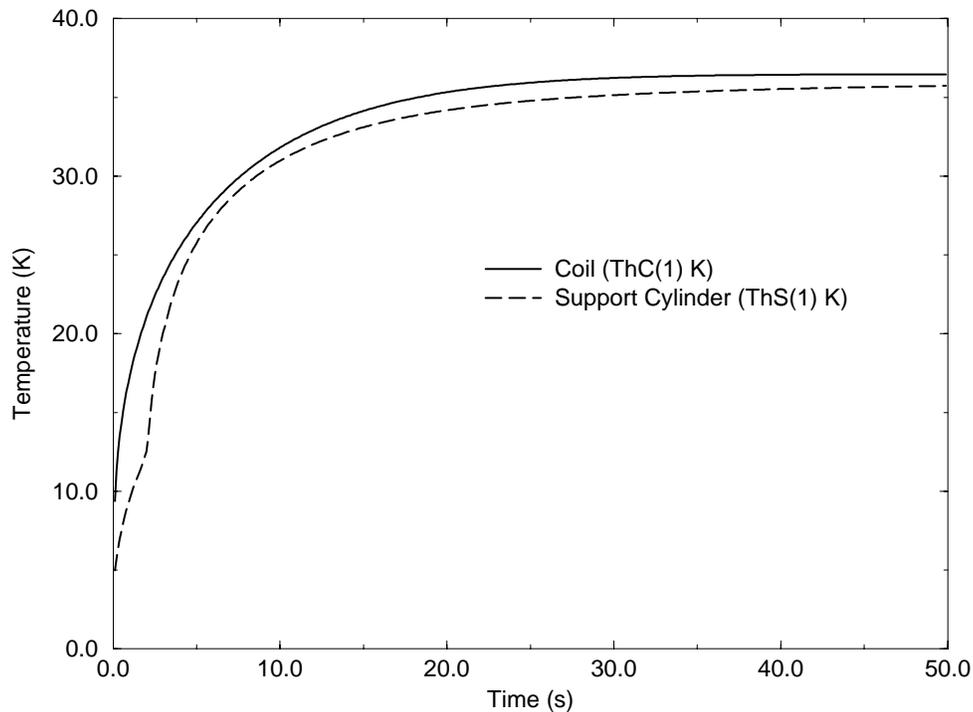
A preliminary quench analysis of the *BABAR* solenoid has been made using a code developed for the DELPHI solenoid design. The code models the thermal and inductive behavior of the solenoid in order to take into account quench-back effects and heat transfer to the support cylinder. This analysis shows that quench-back is predicted about two seconds after opening



**Figure 9-8.** Solenoid power and quench protection concept.

Parameter	Value
Operating Current	7.11 kA
Stored Energy	25 MJ
Inductance	1.2 H
Quench Voltage	500 V
Protection Resistor	0.070 $\Omega$
Time Constant	17.1 s
Adiabatic Peak Temperature	100 K
Overall Current Density: Conductor 1	74 A/mm <sup>2</sup>
Conductor 2	35 A/mm <sup>2</sup>
Aluminum Stabilizer RRR Zero Field	500

**Table 9-6.** Quench parameters.



**Figure 9-9.** *Temperature variation during quench. The temperature rise in the coil and support cylinder during a quench should not exceed 40 K.*

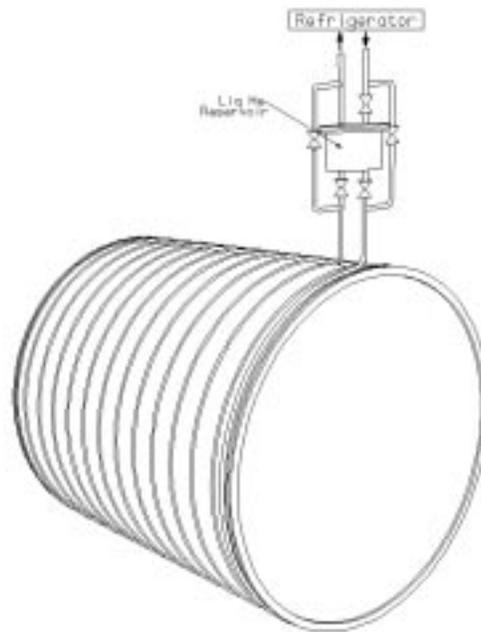
the protection circuit breakers. Figure 9-9 shows that the temperature rise in the coil and support cylinder during a quench should not exceed 40 K.

### Stability

The *BABAR* solenoid coil will be indirectly cooled using the technology established for existing detector magnets such as DELPHI and ALEPH. The reliable operation of those magnets has demonstrated that safe stability margins can be achieved using high-purity, aluminum-clad superconductors in a fully bonded, indirectly cooled coil structure.

Conductor stability has been estimated using an analysis code in order to establish the minimum quench energy (MQE) for transient heat pulses. The computed MQE = 1.4J. The computed minimum quench length (MQZ) is 0.6 m.

These margins are considered to be safe for the *BABAR* solenoid due to its low-stress design. The stability margin will be optimized during the full design study.



**Figure 9-10.** Cold mass cooling circuit. The cryogenic supply chimney passes through a cut-out in the backward end of the barrel flux return.

#### 9.4.4 Cold Mass Cooling

**Cooldown.** Cold mass cooldown is accomplished by circulating cold helium gas either directly from the refrigerator or from a storage dewar with gas mixing. A preliminary cold mass cooldown analysis has been performed. A cooldown mass flow rate of  $\sim 15$  g/s will lead to a cooldown time of five days. The maximum temperature difference across the cold mass is limited to 40 K in order to minimize thermal stress during the cooldown from 300 to 100 K.

**Operating Conditions.** Under operating conditions, the cold mass is cooled by circulating two-phase helium in circuits mounted on the coil support cylinder. The thermo-syphon technique will be used to drive the cooling circuit. This technology is established and yields the simplest operational mode. The thermo-syphon cooling circuit is designed for high flow rates to ensure the correct quality factor for the helium. The conceptual layout of the cold mass cooling circuit is shown in Figure 9-10. The circuit is fed through a manifold at the bottom of the support cylinder. The cooling circuits are welded to the support cylinder surface with a spacing of  $\sim 0.3$  m to limit the temperature rise to less than 0.1 K. The cooling pipes terminate in an upper manifold. The circuit will be designed to provide operation during quench conditions.

**Heat Loads.** The estimated static heat loads for the solenoid are given in Table 9-7. Eddy current heating in the support cylinder will cause additional heat loads during charging of the solenoid. However, for a solenoid charging time of 30 minutes, the estimated transient power is  $\sim 2$  W, which is small compared to static heat loads.

### 9.4.5 Cryostat Design

**Vacuum Vessel.** The cryostat consists of an annular vacuum vessel equipped with radiation shields and superinsulation (Figure 9-7). The vacuum vessel is designed to satisfy a number of basic criteria:

1. Support vacuum loads in accordance with recognized pressure vessel codes;
2. Carry the cold mass and radiation shield weight through the insulating supports;
3. Operate with deflections of less than 2 mm under all loads when mounted in the flux return barrel;
4. Carry the loading of the inner detectors; and
5. Operate under defined seismic loadings.

The vacuum vessel is designed as two concentric cylinders with thick annular end plates, all of aluminum alloy 5083; its basic parameters are given in Table 9-8. A preliminary finite element (FE) structural analysis of the vessel has confirmed that design criteria (1)–(4) can be met with reasonable safety factors. Maximum vessel deflections are less than 2 mm, and stress levels are generally lower than 40 MPa with all loads applied. Deflections are minimized when the vessel is supported on the horizontal centerline and detector loads are also applied at that point. Performance of the vessel under seismic loadings (5) is still under consideration.

**Thermal Shielding.** The cryostat is equipped with radiation shields, which operate at 40–80 K, and superinsulation. The shields are cooled by helium gas supplied directly from the refrigerator. About 30 layers of superinsulation separate the vacuum vessel walls from the radiation shields. Another five layers will be installed between the shields.

**Services.** Cryogenic supplies and current supplies are connected from a services turret to the cryostat through the service chimney in the backward end door. Current leads and local control valves are mounted in the services turret. Cryogenic relief valves are also mounted in the service turret for quench and refrigeration failure conditions.

Magnet Heat Loads at 4.2K			
Item	Parameter	Load (watts)	Liquifaction (l/h)
Cold Mass	7000 kg		
Total Surface Area	100 m <sup>2</sup>		
Radiation Heat Flux (Design)	0.4 W/m <sup>2</sup>		
Radiation Heat Load (Design)		40 W	
Conduction Heat Load		10 W	
Transient Heat Load (30 min)		2 W	
Total 4.5 K		52 W	73 l/h

Magnet Shield Heat Loads at 80K			
Item	Parameter	Load (watts)	Liquifaction (l/h)
Shield Mass	1000 kg		
Total Surface Area	100 m <sup>2</sup>		
Radiation Heat Flux (Design)	3 W/m <sup>2</sup>		
Radiation Heat Load (Design)		300 W	
Conduction Heat Load		50 W	
Total 80 K		350 W	

Current Leads			
2 leads × 7 kA	0.72 g/s	16 W	22 l/h

Cryoplant Heat Loads at 4.2K			
Item	Parameter	Load (watts)	Liquifaction (l/h)
Valve Box and Valves		10 W	15 l/h
Transfer Lines (Liquefier-dewar & Return)	4 m (×2)	6 W	8.5 l/h
Transfer Line (Dewar-valve Box)	4 m	3 W	4 l/h
Coaxial Transfer Line (Valve Box-magnet)	60 m	3 W	4 l/h
4000L Dewar		6 W	8.5 l/h
Total 4.2K		28 W	40 l/h

**Table 9-7.** Cryogenic heat loads.

Envelope Dimensions	
Inner Radius	1400 mm
Outer Radius	1730 mm
Length	3850 mm
Materials	AL5083

Design Loads	
Vessel Weight	6 tonne
Cold Mass	6 tonne
Calorimeter	50 tonne

Seismic Load Factors (Max)	
Horizontal	1.2 g
Vertical	2.0 g

**Table 9-8.** *Vacuum vessel parameters.*

### 9.4.6 Coil Assembly and Transportation

The coil will be assembled inside the cryostat at the manufacturer's plant. Electrical and cryogen connections will be made at the chimney so that the coil can be tested before shipping.

A complete cooldown will be carried out from room temperature to the operating temperature of 4.5 K. The cooldown will allow checking of cooldown time, temperature control, heat loads, and full operation of sensors. A magnetic test will also be performed at low field (30% of the operating current) to check superconductor operation, the joint resistance, and the additional losses due to the energization.

Before delivering the magnet, but after the tests at the factory, the end flanges will be dismantled to allow a hard connection of the cold mass to the cryostat walls. Depending on the transport facilities, the chimney may also be dismantled. In this case, the electrical and cryogen connections also must be dismantled and protected against breakage.

## 9.5 Cryogenic Supply System and Instrumentation

---

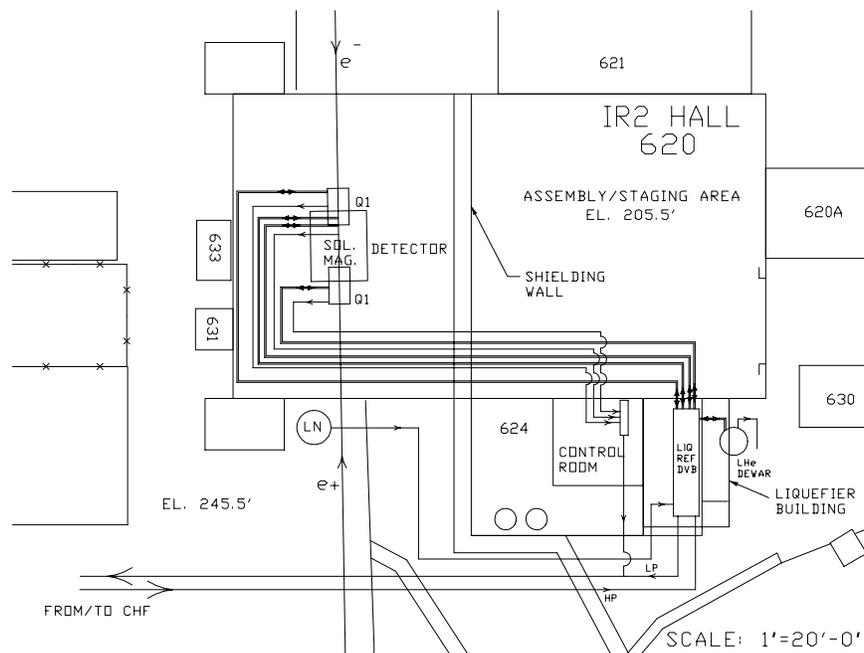
Operation of the superconducting solenoid requires both liquid helium and cold helium gas (20 K to 100 K) for cooldown and refrigeration of the thermal shields. Similar systems have been used successfully throughout the HEP community. A summary of the cryogenic loads is given in Table 9-7.

The helium plant, consisting of a helium liquefier/refrigerator, a 4000  $\ell$  supply dewar, and a distribution valve box (DVB), is located adjacent to the experimental hall, approximately 60 m from the magnet, as shown in Figure 9-11. It is possible that the DVB will be incorporated within the helium liquefier cold box, depending upon the helium liquefier selected. This liquid helium plant meets all of SLAC's requirements and is sized conservatively at 150–200  $\ell$ /hr. It will supply LHe to the *BABAR* superconducting solenoid, the two future superconducting beam line focusing magnets (Q1), and an auxiliary dewar/trailer for all other SLAC experiments. The detector solenoid is expected to consume less than 100  $\ell$ /hr of LHe.

Liquid nitrogen is required for the helium liquefier, for the initial stages of coil cooldown, and miscellaneous uses in IR-2. It will be supplied from a 20,000  $\ell$  tank located on the apron above the experimental hall. This tank, which is an existing SLAC unit, will be refurbished to serve all cryogenic system requirements. A second, similar vessel is also available if a reserve LN2 supply is required. Vacuum-insulated transfer lines connect the LN2 tank to the helium plant.

The 4000  $\ell$  liquid helium dewar is a refurbished SLAC unit fed directly from the liquefier. This volume provides approximately 30 hours of autonomous operation of the solenoid and quadrupole magnet systems in the event of a minor liquefier/compressor malfunction. LHe from the dewar is supplied at  $\sim 1.25$  bar, via proportional control valves in the DVB, to the appropriate magnet/auxiliary dewar system as required. These electro-pneumatic operated valves are actuated by process controllers and superconducting LHe level gauges.

Liquid helium and cold helium gas are supplied to the detector solenoid and Q1 superconducting magnets in flexible, vacuum-insulated, low-loss transfer lines. These transfer lines provide the cold gas return path to the liquefier/refrigerator. This type of transfer line has been used successfully by SLD. The transfer lines are designed to be compatible with the SLD lines so that the latter can serve as spares and/or Q1 transfer lines. LHe and cold shield gas for the solenoid are routed to the magnet service chimney at the detector backward end door penetration (north end). Magnet cooldown to  $\sim 20$  K is achieved with cold helium gas via the DVB. Warm helium gas from magnet current leads is routed in uninsulated piping to compressor suction.



**Figure 9-11.** Layout of the cryogenic system.

The solenoid is equipped with a full set of instrumentation sensors for monitoring, control, and diagnostic purposes. Instrumentation includes temperature sensors for the cold mass, shield cryogen flow monitoring, and strain gauges in the coil support cylinder. Voltage taps monitor the electrical resistance of the conductor joints and quench detection. The quench detection systems are hard-wired to interlocks. The solenoid instrumentation and controls are integrated with the *BABAR* experiment and refrigeration controls.

The liquid helium plant, which is fully automatic, is furnished with a process control system and all requisite logic and software necessary for all operational modes. Control and monitoring of the cryogenic plant and the magnet coil, together with remote control and monitoring of the compressor room, is carried out from a control room adjacent to the plant room and IR-2. Main operating parameters are interfaced with the *BABAR* data acquisition and monitoring systems.

## 9.6 Flux Return

---

### 9.6.1 Overview

The flux return assembly provides the external flux path for the magnetic field from the superconducting solenoid. The steel plate segmentation enhances low momentum pion/muon separation and  $K_L^0$  detection efficiency. The flux return also provides the gravitational and seismic load path for the barrel detector components to the concrete foundation.

The flux return assembly consists of a barrel, four external support legs, two sets of end doors, roller mechanisms, and vertical adjustment systems. The design of the flux return and its components reflects the limitations of the IR-2 experimental hall.

End door components must be movable once the detector is assembled to allow maintenance access to inner detector subsystems. The design of the flux return provides a means for proper alignment of all flux return components with respect to the detector magnet and the beam line. This alignment must be reestablished after maintenance is performed.

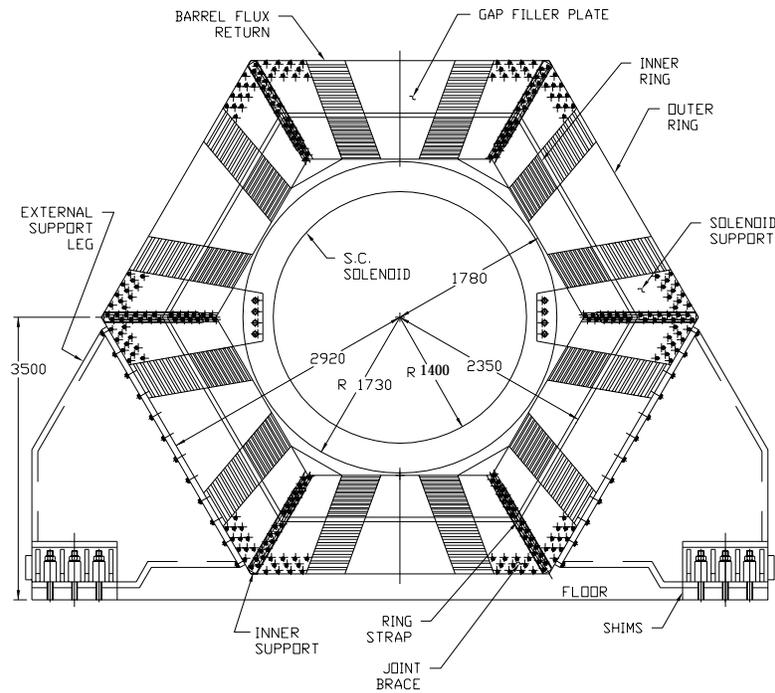
### Analysis of the Flux Return

The flux return analysis will investigate the overall structural integrity, including that of the supporting structures. This includes looking at all the components and connections involved in the operation, assembly, and seismic loadings of the structure. Operating loads for the barrel include both the gravitational and the magnetic forces. Assembly loads include any additional fixture weights that need to be attached during assembly. Seismic loadings assume a site-specific peak ground acceleration of 0.6 g. This seismic design standard exceeds that of the minimum base shear requirements as formulated by the Uniform Building Code.

A finite element model program, ANSYS [ANS95], revision 5.1, is currently being used to study the stresses and deflections of the barrel and its supporting structures under the different loading conditions. It is also being used to model the structure's response to oscillatory ground motion. The results from these studies will be used to verify that the design requirements are met.

### 9.6.2 Barrel Flux Return Description

The barrel flux return assembly is shown in Figure 9-12. The barrel flux return assembly extends 4050 mm in  $z$ , and the center is positioned 370 mm from the interaction point in

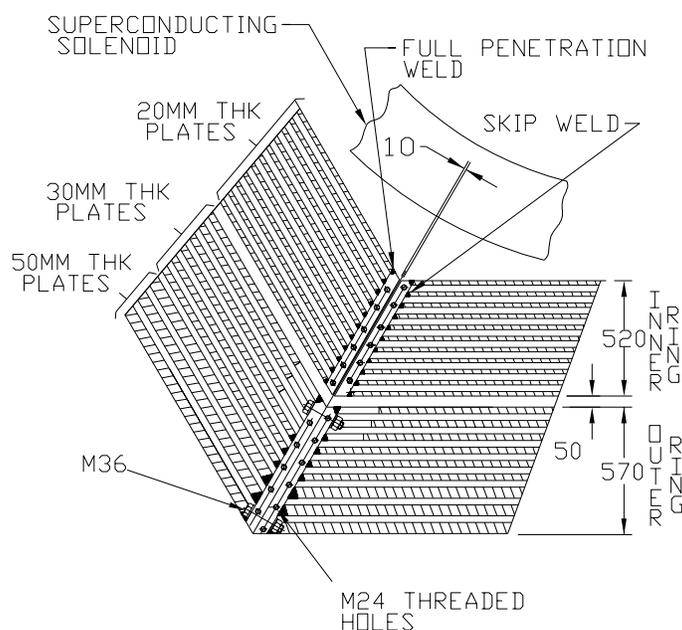


**Figure 9-12.** Barrel flux return assembly showing the support of the solenoid, gap filler plates, and the external support legs.

the positive  $z$  direction. The barrel flux return extends radially from the detector axis from 1780 mm to 2920 mm and consists of inner and outer hexagonally-shaped concentric steel rings. The radial gap between the inner and outer rings is 50 mm. The rings consist of blocks with multiple steel plates oriented parallel to the axis of the solenoid. Two side plates and the inner and outer parallel plates of each block are welded into a rigid box using full penetration welds along their entire length in the  $z$  direction (parallel to the beam line). The remaining parallel plates are held in place using skip welds at their connections with the side plates. The use of double block construction for each of the hexagonal segments of the barrel flux return provides a significant safety margin with respect to the 45-tonne rated capacity of the overhead crane in IR-2. The blocks are designed to provide continuous muon detection in azimuth. The weight of the barrel flux return is 312 tonne excluding the external support legs.

### Outer Ring Description

The outer ring is the primary structural support for the *BABAR* detector. This ring is composed of six blocks. Each block consists of six 30 mm-thick and three 50 mm-thick parallel

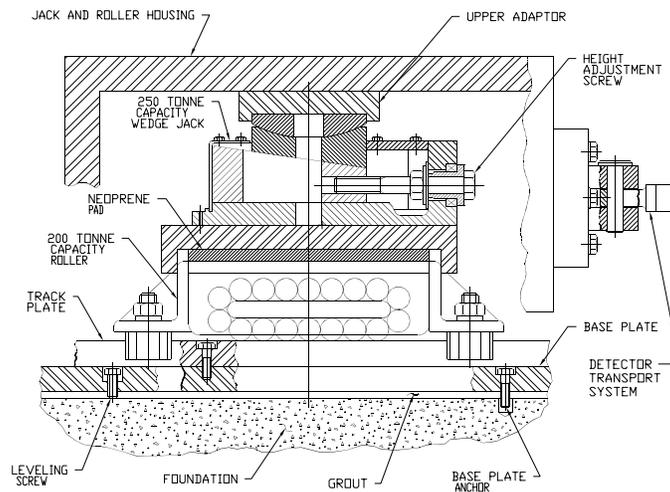


**Figure 9-13.** *Block-to-block interface in the flux return.*

structural plates oriented so that the thicker plates are positioned towards the outermost portion of the barrel hexagon. The plates are positioned in each block to provide 30 mm of clear space between each plate and the next for the installation of Resistive Plate Chambers (RPCs). The inner surface of the outer block is located 2350 mm radially from the interaction point. The 30 mm spacing between plates is to be maintained during block fabrication. Therefore, the tolerance build-up due to variations in the plate thickness and plate flatness extends outward in the radial direction. The spacing of the parallel plates is held fixed by 60 mm-thick side plates located at the block-to-block interface, as shown in Figure 9-13. Two of the blocks are provided with threaded holes in their outer plates for mounting of the external support legs. Each block is 3750 mm in length, 570 mm in nominal height, 2714 mm wide at the inner trapezoid base, and 3372 mm wide at the nominal outer trapezoid surface. Each outer block weighs approximately 31.2 tonne.

### Inner Ring Description

The inner ring is also composed of six blocks. Each block consists of eleven 20 mm-thick parallel structural plates positioned to provide 30 mm of clear space between plates for the installation of RPCs. The inner surface of the inner block is located 1780 mm radially from the interaction point. Like the outer plates, the 30 mm spacing is held fixed, and the plate manufacturing tolerances are allowed to accumulate in the outward direction. A gap



**Figure 9-14.** A 200 tonne capacity roller and a 250 tonne capacity jack are located at the end of each support leg to support the barrel.

of 20 mm is provided between the inner and outer hexagonal rings for this purpose. The spacing of the parallel plates is held fixed by 50 mm-thick side plates located at the block-to-block interface as shown in Figure 9-13. The side plates are provided with threaded holes to mount a 150 mm-thick plate that supports the inner blocks from the outer ring. The inner blocks are 5 mm shorter in  $z$  length than the outer blocks and will have to be shimmed during assembly for a secure fit. There is 10 mm of clearance provided at the block-to-block connections for manufacturing and assembly tolerances. Each block is 3745 mm in length, 520 mm in nominal height, 2050 mm wide at the inner trapezoid base, and 2650 mm wide at the nominal outer trapezoid surface. Each block weighs approximately 16.4 tonne.

### External Support Legs

Two external support leg assemblies provide the gravitational and seismic load path from the barrel flux return to the concrete foundation. Each assembly consists of two legs positioned 3200 mm apart in  $z$ . The legs are fabricated from 50 mm-thick structural steel plates. Included at each support leg is a 200 tonne capacity roller and a 250 tonne capacity jack as illustrated in Figure 9-14. A preformed fabric pad is positioned between the jack and the roller to add compliance to the system. The horizontal spacing of the jack and roller assemblies is nominally 8000 mm. Each support leg is approximately 3750 mm in length, 2400 mm in width, and 3300 mm in height. The weight of each support leg assembly is approximately 20 tonne.

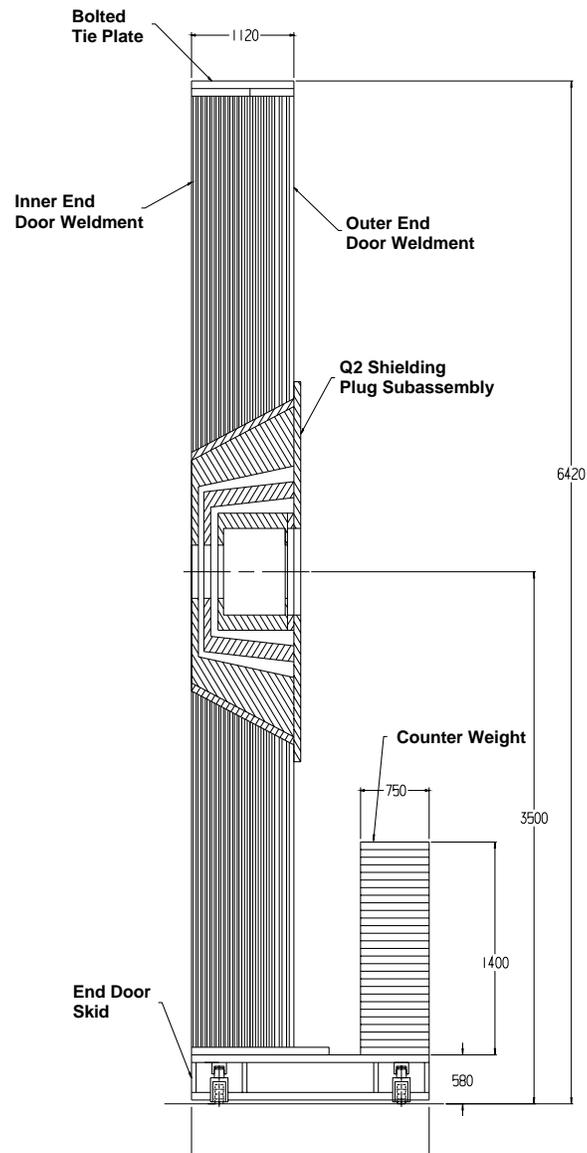
### 9.6.3 End Door Description

The forward and backward end doors are an array of steel plates that form a regular hexagon approximately 5840 mm across the flats and 1120 mm thick. The array consists of eleven 20 mm-thick plates, six 30 mm-thick plates, and three 50 mm-thick plates. There is a nominal gap width of 30 mm between plates in the array for locating the RPCs. The plates for both the forward and backward end doors weigh approximately 255 tonne. This weight does not include plates and bars that tie the plates together for structural reasons, the support members that attach the end doors to a movable skid, or the skid itself (Figures 8-23 and 9-15). Both the forward and backward end doors are split along the vertical centerline of the detector forming right and left end doors.

Each end door is mounted on a skid that permits it to be raised onto high load capacity rollers by hydraulic jacks and to be moved on tracks located in the IR-2 hall. This provides a means to move the end doors into proper alignment with the barrel prior to being bolted in place, and to be moved away from the barrel for maintenance access to the detector. The center of gravity of the end door plates is high compared to the depth of the base. The skid provides additional stability during the horizontal accelerations experienced during moving or seismic events.

During operation, the magnet exerts an inward axial force that causes a significant bending moment on the end door plates. The forward end doors also support the weight of the Q2 shielding plug and carry the axial magnetic load induced in the plug. To resist gravitational and seismic loads, and to limit the plate stresses and deflections, all the end door plates are joined together to form a single structural member.

The design of the end doors permits each door to be assembled and disassembled inside the IR-2 hall with the existing facility crane. Since the capacity of the IR-2 crane is 45 tonne, each separable part, together with all rigging and lifting fixtures needed, cannot exceed this limit. Each end door is therefore composed of two weldments that are fastened together at installation. The inner weldment consists of eleven 20 mm plates and two 30 mm plates, with each plate welded to a channel-shaped frame that extends along the top and bottom of the hexagon shape and along the boundary between the right and left doors. Two horizontal stiffeners are positioned between the 20 mm plates. Additional stiffeners are also required to stiffen the weldment and help maintain the necessary gaps for the RPCs. The outer weldment consists of four 30 mm plates and three 50 mm plates welded together with similar stiffening members. The detailed analysis of the response of the plates to the magnetic force distribution is not yet available, nor has a detailed seismic analysis been done for the end door plates. Several design options are being studied that can provide the necessary strength and stiffness.



**Figure 9-15.** Side view of the forward end door showing the Q2 shielding, counter weights, and support rollers.

The inner and outer weldments are joined together at installation by bolting each weldment to the top of the skid along the bottom of the hex, and by bolting tie plates around the remaining perimeter of the weldments, except where the shielding plugs are located. The outside tie plates are bolted in place after the RPCs are installed, and provisions for cabling are provided in these outside plates. These tie plates are also used to attach the end doors to the barrel of the flux return.

### End Door Skids

Each end door is mounted on a skid that is equipped with four 70 tonne capacity rollers, and four 45 tonne hydraulic jacks, one in each corner, which allow each of the doors to be moved relative to the barrel for maintenance access. The rollers ride on hardened steel tracks that are permanently located in the floor of IR-2. Because the bases of the end doors are narrow compared to their high centers of gravity, 30 tonne of existing scrap steel is bolted to each skid to lower the center of gravity and move it toward the middle of the skid.

The end doors are bolted either to the barrel of the flux return when in the operating position or to seismic restraint brackets when in the parked position. While the doors are being moved, the 30 tonne counter weight provides lateral stability for a horizontal acceleration of 0.3 g.

### Forward Q2 Shielding Plug Support

The preferred option for mounting the Q2 shielding plug is to fabricate and assemble the shield in two sections split along the vertical axis. These assemblies are mounted to large half-round flanges that are bolted to the back 50 mm plates of the forward end doors (Figure 9-15). This simplifies access to the front portion of the detector when the end doors are moved. It is likely that the halves of the end door will be tied together by two additional half-round flanges oriented at  $90^\circ$  to the shielding plug supports to stiffen the assembly. This option requires that each half of the shield be installed in the end doors at final assembly since each half of the shield assembly weighs approximately 20 tonne. Special lifting and assembly fixtures are required to accomplish this task.

The alignment of the shield and the Q2 magnet is done by adjusting the position of each half-round flange on the end door and installing dowel pins to maintain the alignment. Sufficient clearance between the shield and the forward Q2 magnet is provided to permit the shield to move vertically with the stroke of the jack as the end door is raised to be rolled out of the way for detector access.

### 9.6.4 Options and Detailed Design Issues

A detailed stress and deflection analysis is proceeding for the finalized overall envelope dimensions. A detailed magnetic field and force analysis of the end doors ensures that they will have adequate strength and stiffness to meet all the requirements imposed on them.

The tolerance on plate flatness must be defined together with the envelope dimensions of the RPCs. Standard mill tolerances for plate flatness do not meet our requirements for the end door plates to permit reliable RPC installation; these tolerances exceed 15 mm, half the nominal gap width. This issue, together with weld distortion in the plates during fabrication, must be resolved with both the plate supplier and the barrel and end door fabricator. Other manufacturing tolerances must be established as the system interface dimensions are finalized, and the effects of these tolerances on the physics performance of the detector are reviewed.

### 9.6.5 Procurement, Fabrication, Assembly, and Schedule

The barrel and the end doors will be built by the same fabricator as part of the same procurement contract. This will eliminate some duplication in the review of vendor qualifications, quality control plans and actions, and many contract administration issues. In addition, control of other characteristics of the plate material, such as the chemistry of the plates as it affects weldability and machinability, the mechanical properties of the plate, *etc.*, may be more easily tracked by having one set of acceptance inspection criteria from one supplier for all plates.

The barrel and end doors will be fully assembled and inspected at the fabricator's shop. In this way, any problems that arise can be solved before final assembly in the IR-2 hall. This will also permit a thorough review of the assembly procedure by an experienced fabricator and ensure that the necessary lifting and assembly fixtures are functional and are provided with the barrel and end doors. The details of the fabrication and assembly plan will be developed by the fabricator subject to the review of the responsible design engineer in the *BABAR* collaboration.

## References

---

- [And82] D. Andrews *et al.*, “A Superconducting Solenoid for Colliding Beam Experiments,” *Ad. Cryogenic Eng.* **27** (1982).
- [ANS95] ANSYS, INC, “General Finite Element Code,” Rev. 5.1 (1995).
- [ASME94] ASME, “1992 ASME Boiler and Pressure Vessel Code” (1992).
- [BF95] E. Baynham and P. Fabbriatore, “Superconducting Solenoid for the *BABAR* Detector,” *BABAR Note* in preparation.
- [Coils] A. Bonito Oliva *et al.*, “ZEUS Magnets Construction Status Report,” 11th Int. Conference on Magnet Technology, MT-11, 229 (1989).  
P. Clee *et al.*, “Towards the Realization of Two 1.2 T Superconducting Solenoids for Particle Physics Experiments,” 11th Int. Conference on Magnet Technology MT-11, 206 (1989).  
H. Desportes *et al.*, “Construction and Test of the CELLO Thin-Wall Solenoid,” *Ad. Cryogenics Eng.* **25**, 175 (1980).  
H. Desportes *et al.*, “General Design and Conductor Study for the ALEPH Superconducting Solenoid,” *J. Phys. (Paris)* **C1-341**, S1-T45 (1984).  
Y. Doi *et al.*, “A 3T Superconducting Magnet for the AMY Detector,” *Nucl. Instr. Methods* **A274**, 95 (1989).  
M.A. Green *et al.*, “Construction and Testing of the Two-Meter-Diameter TPC Thin Superconducting Solenoid,” *Ad. Cryogenics Eng.* **25**, 194 (1980).  
M.A. Green *et al.*, “A Large Superconducting Thin Solenoid for the STAR Experiment at RHIC,” *IEEE Trans. App. Superconductivity*, 104 (1993).  
H. Hirabayashi, “Detector Magnet Technology for High Energy Accelerators,” ICEC-11, 115 (1986).  
H. Minemura *et al.*, “Fabrication of a 3 m-dia.  $\times$ 5 m Superconducting Solenoid for the Fermilab Collider Detector Facility,” *J. Phys. (Paris)* **C1-333**, S1-T45, (1984).  
C.M. Monroe *et al.*, “The CLEO-II Magnet—Design, Manufacture, and Tests,” ICEC-12, 773 (1988).  
M. Wake *et al.*, “Excitation of a Superconducting Large Thin Solenoid Magnet,” MAG-23, 1236 (1987).  
F. Wittgenstein *et al.*, “Construction of the L3 Magnet,” 11th International Conference on Magnet Technology, MT-11, 131, (1989).  
A. Yamamoto *et al.*, “Thin Superconducting Solenoid Wound with the Internal Winding Method for Colliding Beam Experiments,” *J. Phys. (Paris)* **C1-337**, S1-T45 (1984).

- 
- [Des85] H. Desportes, "Recent Progress in the Design and Construction of Beam and Detector Magnets," in Proceedings of the Int. Symposium on Magnet Technology (Zurich, Switzerland), MT-9, 149 (1985).
- [SDM91] Earthquake Safety Committee, "Seismic Design Manual - Mechanical," Rev. 1 Stanford Linear Accelerator Center, Stanford, California (1991).



---

## 10.1 Overview

### 10.1.1 Introduction

**T**he electronics system encompasses all front-end, trigger, and data acquisition electronics. The environment of the high-luminosity PEP-II machine poses certain unique challenges to this system. PEP-II beam crossings occur at a rate of 238 MHz, which, in terms of the response time of the electronics, is essentially continuous, unlike previous  $e^+e^-$  colliders. PEP-II may have severe backgrounds, producing high occupancies. These high occupancies present a significant burden on the transport and recording of data. The rate of all physics processes to be recorded at the design luminosity of  $3 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$  is expected to be 100 Hz. The electronics system will also need to accommodate the rate and background implications of a luminosity upgrade to  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . These factors have resulted in the adoption of an innovative system architecture that departs significantly from previous implementations at conventional  $e^+e^-$  experiments. In this architecture, the electronics systems are capable of collecting data from the detector in realtime, and processing and transporting the data in parallel. The data are extensively buffered to avoid deadtime losses while making trigger decisions. In addition, the detector electronics, trigger logic, and data acquisition system are capable of simultaneous trigger processing of multiple events.

As described in Chapter 12, the backgrounds expected in PEP-II include particles lost near the IP, beam-gas interactions, and cosmic rays. The dominant source of occupancy is predicted to be the lost particle background, which depends on the beam intensity and the quality of the vacuum in the machine in the region from 3 to 50 m from the IP. At the design pressure of 1 nTorr, simulations predict raw hit rates of several tens of megahertz per layer in the silicon vertex detector, 100 kHz per layer in the drift chamber, and 15 kHz per crystal in the calorimeter. The first device to become unusable as background increases is the silicon vertex detector. This will happen at approximately ten times the nominal rate. Both long-term (ten-year) radiation damage and very high occupancies in the inner layers provide limits to useful running conditions. The electronics system is fully operational under background conditions up to ten times the nominal rate. At higher background conditions,

System	Number of Channels
Vertex Detector	150,000
Muon System (IFR)	42,000
DIRC	13,400
Calorimeter	7,000
Drift Chamber	7,000
Aerogel (ATC)	144

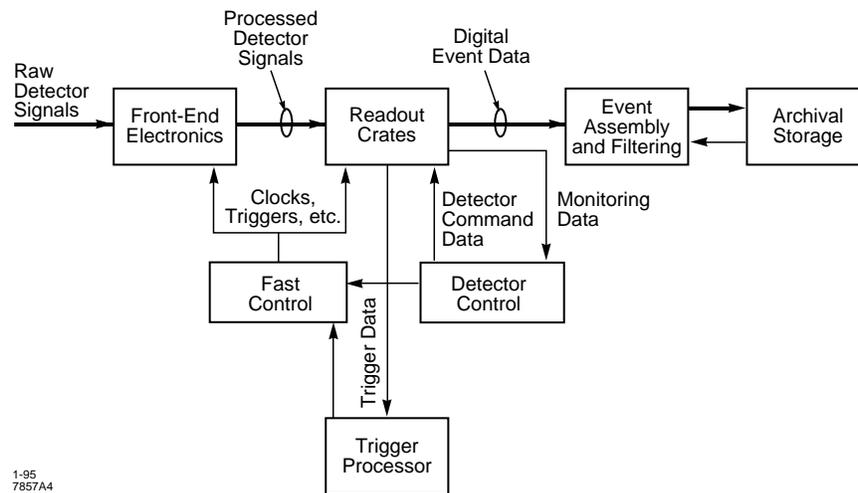
**Table 10-1.** Channel count for each detector system.

the system degrades gracefully, albeit with increasing deadtime, allowing the background and detector performance to be studied.

The principal features of the electronics system are: a common architecture for all detector electronic front-ends; common trigger, timing, and control interfaces; a common data acquisition interface; and a common online software environment from the lowest level within the front-end signal processors to the highest level user interfaces and event processing farms. By generalizing the architecture for all systems down to the point of contact with signal digitization, the overall maintenance, software, and training needed to operate the system are minimized, and engineering can properly be focused on detector-specific requirements.

The *BABAR* detector has approximately 220,000 signal channels (summarized in Table 10-1) distributed among six types of detectors, delivering signals with shaping times varying from 1 ns to 3  $\mu$ s. Since the accelerator operates essentially continuously, the fast trigger will be unable to identify the specific beam crossing associated with an event. In fact, the event time will be uncertain to hundreds of nanoseconds. Consequently, a requirement of the electronics systems is that they record the raw data in real time and extract the data within the time uncertainty window. Providing sufficient bandwidth to transmit all of the raw event data within these time windows would be overwhelming, so as much signal processing of the data as possible is performed prior to transmission. This requires storage of the data while filtering is being performed and while the pertinent data (features) are being extracted.

All detector systems have a common interface for event readout, fast control, trigger, and detector control. Yet, within this commonality, all detector front-ends have unique signal processing requirements. While it is not possible to devise a uniform interface to the trigger for each detector system or to provide a uniform implementation of the storage of data awaiting a Level 1 decision, the same architecture of data storage and control has been applied to the signal processing chains for all detector systems. A block diagram of the major components of the electronics is shown in Figure 10-1. The following subsections give overviews of the front-end electronics, trigger, and data acquisition systems.

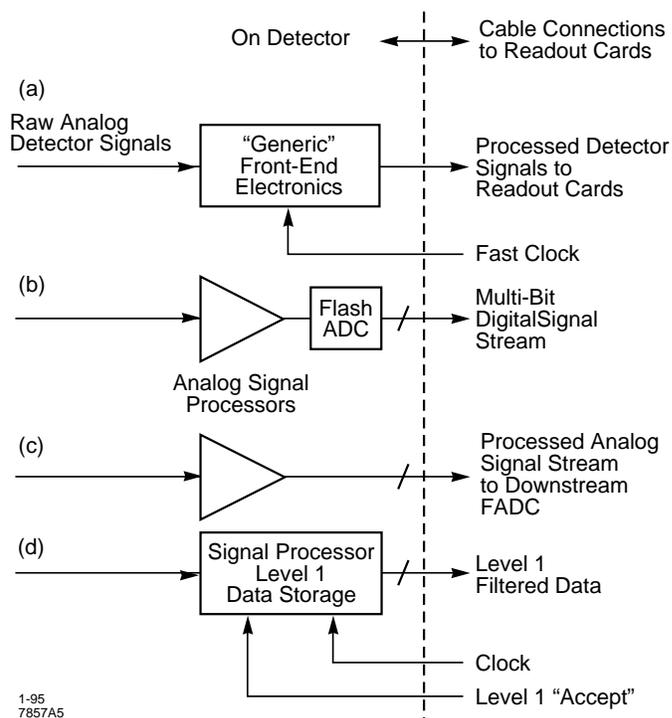


**Figure 10-1.** Block diagram of the electronics system showing the principal communication paths. The dark arrows show the event data path.

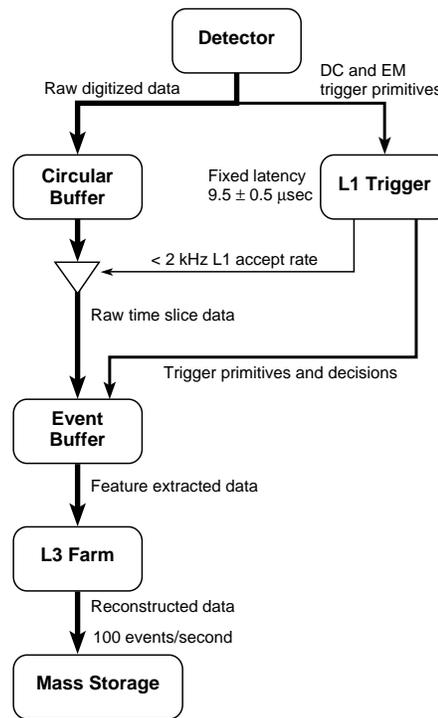
### 10.1.2 Front-End

In the *BABAR* nomenclature, the front-end is that portion of the electronics from the signal source on the detector through the digitizer. For those systems which have their digitizers mounted on the detector, the front-end also includes any additional detector-mounted parts. As shown in Figure 10-2, the front-ends are customized to fit the needs of the particular detector system. Figure 10-2(a) shows a generic front-end. The raw detector signals are appropriately processed and transmitted to the readout modules, which are housed in VME crates in the electronics house. The electronics house (in contrast to the detector) is accessible when the machine is running. Figure 10-2(b) depicts a front-end for the calorimeter. Since the event time is not known at this point, the digitization is done periodically, at a rate optimized for each system. Figure 10-2(c) is the front-end for the drift chamber. Here the digitization is done on the readout module. Figure 10-2(d) shows the front-end for the vertex detector, the instrumented flux return and the DIRC.

The processed detector signals are transmitted to the off-detector readout crates. For systems in which the first stage of buffering is on the readout modules, processed signals constitute a continuous data flow; selection of data by the triggers is accomplished on the readout modules. Systems which have buffers in the detector-mounted electronics send data selected by a Level 1 strobe only upon the request of the readout module. Examples of the first case are the drift chamber and the calorimeter. Examples of the second case are the vertex detector, the instrumented flux return and the DIRC.



**Figure 10-2.** Examples of front-end electronics for various detector technologies. Variations (a) through (d) are explained in the text.



**Figure 10-3.** Trigger structure and data acquisition data flow, showing levels and buffers. See text for a discussion of Level 2 options.

### 10.1.3 Trigger

#### Definition of Trigger Levels and their Functional Requirements

The structure of the trigger and data flow are shown in Figure 10-3. The trigger:

- provides a front-end strobe (Level 1 strobe) after a latency between 9 and 10  $\mu\text{s}$  as required by the finite depths of the front-end buffers (Chapter 4);
- initiates, in the readout modules, the extraction of hit times and magnitudes from the sample streams (feature extraction);
- initiates the collection of the data associated with an event in a single processor (event assembly); and

- selects a sample of events to be written to mass storage at a rate of no more than 100 Hz, as required by the online computing system.

This architecture is the simplest robust solution that performs the trigger functions, given the rate considerations.

The only practical method for delivering a Level 1 strobe in less than  $10\ \mu\text{s}$  is to use a hardware processor operating on a reduced representation of a subset of the data (trigger primitives). The most powerful method for reducing the trigger accept rate to 100 Hz with no loss to the physics program is to use the vertex detector to select events with multiple tracks originating from a narrowly defined primary vertex near the  $e^+e^-$  luminous region. Simulation studies show that the required vertex detector information cannot be made available to the Level 1 hardware trigger logic because the bandwidth required to supply sufficient position granularity to suppress background is comparable to the raw-data bandwidth required *after* the Level 1 strobe [Lev94c]. Because the vertex information cannot be included, the Level 1 accept rate will be considerably higher than the eventual rate of writing events to archival storage. Event assembly could proceed comfortably at a rate of 2.0 kHz, based on experience with other experiments [Boz93, Dou94, Buo, Cul91, Pat94]. Once the events have been assembled, there will be ample information with which to filter out the background events. Therefore, a two-level trigger containing a flexible Level 1 filter with a maximum output rate less than the maximum event assembly rate is necessary and sufficient.

To support the system architecture, dedicated data paths provide trigger data to the first level trigger hardware (Level 1). The maximum average acceptance rate for this trigger is required to be less than 2.0 kHz (Table 10-11 in Section 10.9). A final software filter (Level 3) running in commercial processors analyzes the event data from all detector systems and reduces the rate at which events are written to archival storage to less than 100 Hz. The trigger protocol is expected to be highly efficient for a variety of physics modes in addition to  $B$  physics events.

Each trigger level uses a unique set of data and algorithms, summarized in Table 10-2, that limits the accepted events. Successive trigger levels use increasing amounts of time to perform more intricate analyses of the data. Not all detectors contribute trigger data; those that do, preprocess it, *e.g.*, extract track segments from a collection of raw drift chamber hits. The results of this preprocessing are referred to generically as primitives. The primitives used by the Level 1 trigger are track segments from the drift chamber and energy sums from fixed sets of crystals in the calorimeter. The Level 1 trigger rate is dominated by the lost particle background with current estimates showing 2.0 kHz at ten times the nominal beam-pipe pressure. If needed, additional trigger criteria can reduce this rate with some small loss of non- $B$ -physics efficiency.

Trigger	New Information	Algorithms
Level 1	DC Segments, EM Energy Sums	Track Counting, Energy Cuts, $p_t$ Cuts, Matching
Level 2 or 3	Vertex System Occupancies	Primary Vertex Cuts
Level 3	Full Event Data	Final Event Selection

**Table 10-2.** *Trigger inputs and algorithms.*

Additional robustness is designed into the system via hooks for an additional Level 2 filter in case the actual Level 1 accept rate is higher than the maximum Level 3 event build rate, or if the feature extraction function requires more computing power and/or bandwidth than is currently estimated. The Level 2 trigger adds requirements on the track point-of-origin to reject non-annihilation backgrounds. It might also simplify the Level 1 trigger by doing the track transverse momentum cuts. Inputs to this trigger level would include all inputs to Level 1 and any additional signals needed as the algorithms are developed.

All trigger levels have the flexibility to examine event topology, count tracks, and match charged tracks to energy clusters. All levels have some capability in setting adjustable thresholds in energy or transverse momentum. This high degree of flexibility is required to reject known backgrounds under a range of conditions while maintaining very high efficiency for physics events.

#### 10.1.4 Data Acquisition

The readout modules contain large memories that preserve the streams of digitized data. For systems with detector-mounted buffers, these memories store events after the Level 1 accept and are used to hold data that is ready for readout. The remaining systems also use this memory to buffer data during the Level 1 trigger latency.

Within the data acquisition system, there are four principal communication paths:

- An event data path by which all event data pass from the system electronics to the event collection system;
- A high speed clock, trigger, and control pathway for synchronizing the flow of data and controlling reset, initialization, and throttling of data;
- A relatively slow detector control interface between the system electronics and the online system; and

Detector System	Bytes per Event	Number of Sources	Bytes per Source per Event	Source Data Rates (MBytes/s)
Vertex	8400	5	1680	3.4
Drift Chamber	3600	18	216	0.4
DIRC	2700	1	2700	3.5
Aerogel	300	1	300	0.3
Calorimeter	7700	18	275	0.6
IFR	800	1	800	1.6
Trigger	800	3	400	0.8
Totals	24,300	47		

**Table 10-3.** Summary of event data contributions from each detector type at ten times nominal occupancies and data rates per data source for each detector system for a 2.0 kHz trigger rate.

- A trigger data path that conducts detector data to the trigger processor.

The detector control interface is responsible for initializing the system and for monitoring and debugging. Overall management, error logging, and databases are part of the online system described in Chapter 11.

The data acquisition interfaces are envisioned to be very similar across the systems in the following respects. The crates for each detector system will have

- readout modules,
- a fast control distribution module,
- a readout controller, and
- an event-data port card unit (data source).

For each event, the event-assembly network moves the approximately 25 kbytes of event data, summarized in Table 10-3, from the data sources to the Level 3 processor farm. The event-assembly network is implemented with commercial network technology. At a Level 1 trigger rate of 2.0 kHz, this technology is expected to be able to transport the approximately 50 Mbytes/s of data expected at 10 times the nominal design point. This type of technology is easily upgradable and can accommodate the needs of luminosity, trigger processing, and bandwidth upgrades.

## 10.2 Silicon Vertex Detector

---

This section describes the storage and transport of data from the silicon vertex detector transition card into the data acquisition system. The readout chip electronics and the transition card are described in Chapter 4.

### 10.2.1 Requirements

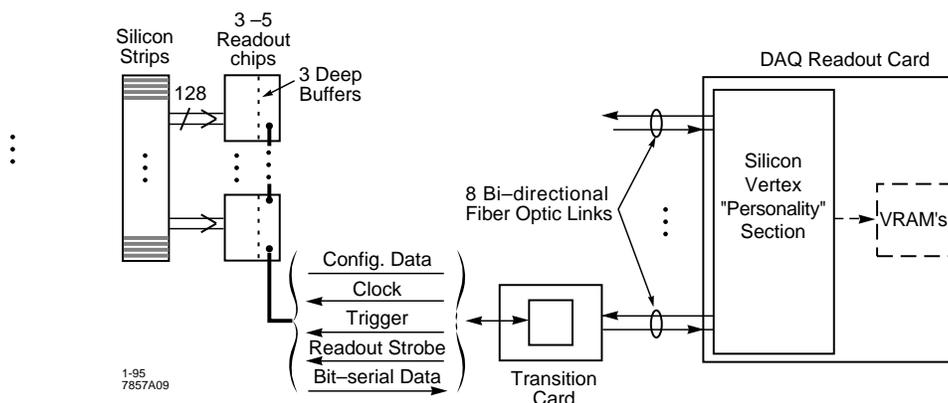
In the vertex detector readout, the data are digitized and sparsified at the detector front-end. Unlike the data from the drift chamber and calorimeter, the silicon vertex detector data must be buffered before being moved from the detector. This minimizes the deadtime that can arise from multiple triggers occurring on a timescale which is short compared with the readout time, or from events that generate large numbers of hits.

Because of the large number of signal channels in the silicon vertex detector (approximately 150,000), the electronics for this detector are required to produce sparsified data streams. For each hit that occurs within a specified time window (*e.g.*,  $\pm 0.5 \mu\text{s}$ ) of a Level 1 trigger, the system must generate three numbers: the number of the channel (strip) that was hit; a rough measure of the time-over-threshold of the pulse produced by the hit; and the time at which the pulse first crossed the threshold.

Data fragments consisting of the above three numbers are generated by the silicon vertex readout chips and placed into the chip's three-deep buffers. Subsequently, the data in these buffers are gathered event by event, and merged into larger data fragments, each covering a large number of hits associated with a particular trigger. Further numbers on particular performance parameters like dynamic range, noise levels, and speeds are given in Chapter 4.

### 10.2.2 Overview

A block diagram depicting the organization of the vertex detector electronics is shown in Figure 10-4. Each chip processes the signals from 128 strips as follows. A time-over-threshold signal is created, which is a logarithmic function of the pulse height. For each clock period, depending on whether the signal is above or below threshold, a binary digit with a value of **1** or **0**, respectively, is input to a latency buffer, awaiting Level 1 decisions. These bits are preserved for the latency period of the trigger, then erased. When a trigger is received, the time history of the stored bits is examined, and the number of **1**s within a certain period before and after the trigger is determined. This number represents a measure of the pulse height. In addition, the time with respect to the trigger of the start of the stored **1**s is



**Figure 10-4.** Block diagram of the vertex detector electronics system.

found. This represents a measure of the time of arrival of the pulse from the strip. These two numbers are then stored in the three-deep (FIFO) buffer for that channel. If there are no 1s in the latency buffer close to the trigger, then a null-data indicator is stored in the three-deep buffer.

During the readout of an event, the data in the last levels of the buffers of all channels containing valid data are read out in bit-serial mode, and channel serial numbers are added. In this operation, several readout chips are connected in a daisy chain as shown in Figure 10-4. The data travel a short distance to a transition card, which launches the bit-serial stream on an optical fiber link destined for a readout module (Section 10.10.4).

The optical fiber link is bidirectional. In addition to the bit-serial data stream traveling toward the readout card, a set of signals is carried in the opposite direction by a second fiber. These signals include: a system clock, Level 1 triggers, event data readout strobes and controls, and several bits of configuration data. The configuration data can control certain functions of the readout chips such as putting a chip into calibration mode.

A single bundle of data lines plus control lines is assumed for each readout section. A readout section is either the  $z$  or  $\phi$  side of a detector. To provide redundancy in case of a single point failure, pairs of readout sections ( $z$  and  $\phi$ ) on the same detector are grouped together such that the data and control lines of one side can service chips on both sides, if necessary. There are 208 readout sections servicing the silicon vertex detector.

The readout chips are daisy-chained on a serial communications pathway operating at 80 MHz. The data are passed through the chips rather than on a common bus. The advantage of this is that a chip can only break the line in one place as opposed to taking down an entire bus. The chips can pass their respective data in either direction on the pathway, and both ends of the pathway are accessible to external electronics.

The data transmission control is implemented as a redundant set of readout strobes. A pulse on the readout strobe gives the first chip in the daisy chain priority to begin sending its data. The packet of serialized data is always prefaced with a header and terminated by a trailer. When this data has been transmitted, the token giving priority to the data pathway is passed to the next chip in the daisy chain, and the process is repeated.

### 10.2.3 Data Format

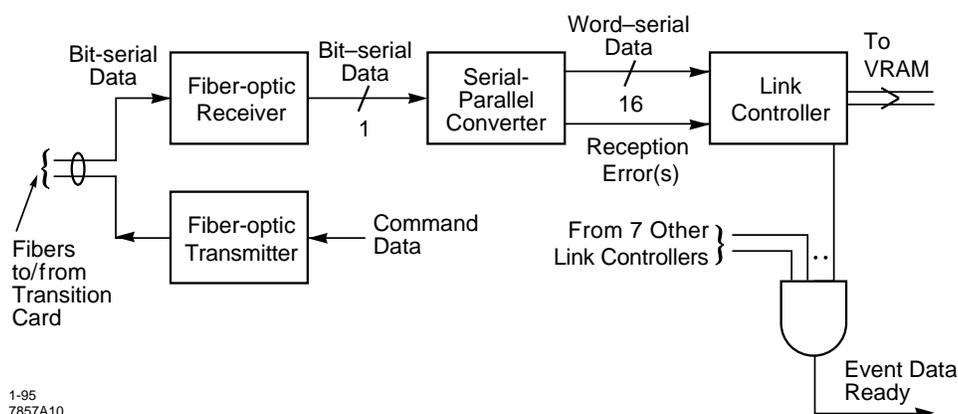
The format of the bit-serial data travelling between the transition card and the readout card contains three fields: a header, a body, and a trailer. The header has a unique code that enables the personality card to locate unambiguously the start of an event data message. The body contains the information from a number of hits. The number of hits will, of course, vary from event to event. The information from each hit is contained in a subfield containing the three numbers—strip identification, pulse height, and time of arrival. These subfields are of constant length. The end of the subfields is announced by a unique trailer bit pattern.

### 10.2.4 Description of the Readout Module

As shown in Figure 10-5, the signals on eight optical fiber pairs are received by the personality module, which is physically part of the readout card. A generic readout card is shown in Figure 10-30 in Section 10.10.4. Here, the optical signals on the fiber are converted to bit-serial digital logic signals by the fiber optic receiver. Next, the serial-parallel converter converts the bit-serial data to word-serial data in which the words are 16 bits wide. The words are then passed to the link controller.

The link controller on the vertex personality module must identify the start of data from the header code and the end of data from the trailer code. Since each chip in the chain has included a trailer after its data, the data acquisition system can count how many chips sent data in response to a single read strobe. The system knows from a configuration database how many chips should report data from a given readout section, so it can recognize the end of the data for an event.

Relying on the format described above, the link controller in the personality module is able to extract the hit data from the incoming data stream and prepare it for forwarding to the VRAM on the DAQ readout card. When it senses the last trailer field, the link controller raises a “Done” flag. When all such flags are raised, an event-data-ready signal is strobed to signal the onboard CPU that a new event has arrived and to enable the issuance of the next readout strobe.



**Figure 10-5.** Block diagram of the vertex detector specific section of the readout module.

The readout strobe is not buffered on the chips. The DAQ readout card must never send a readout strobe until all of the data corresponding to the previous readout strobe have been read out.

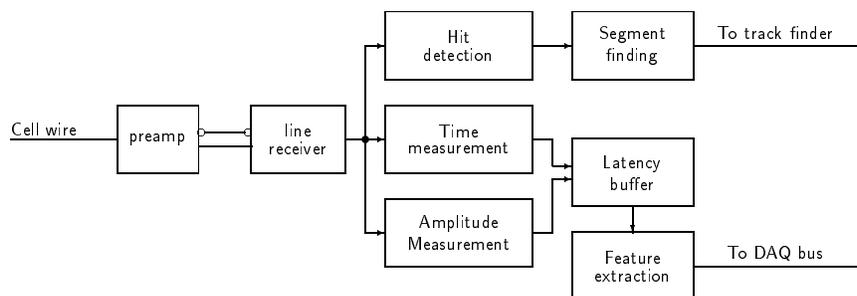
The serial data stream enters the VRAM buffers synchronously with the 80 MHz serial transmission clock (5 MHz after 16 bit deserialization). The VRAM buffers are large enough to store approximately 500 events.

Each of the readout modules also has a processor for signal processing and message handling. Upon receipt of a Level 1 trigger, the processor copies all data from the relevant time slot of the VRAM to conventional RAM. The processor extracts the physical quantities (times and amplitudes) from the information stored and passes the result on for event assembly. Additional buffers are required to hold the post-processed events until they can be read out.

## 10.3 Drift Chamber

### 10.3.1 Requirements

The readout system for the drift chamber should not introduce any significant degradation of the performance of the chamber. The timing resolution should be sufficient to determine the distance of closest approach of a track to each wire with a precision better than the inherent chamber resolution. The charge deposition should be measured with an accuracy compatible with the statistical fluctuations in the formation of the primary clusters. Cell hit information should be made available for the segment finder logic with an efficiency approaching 100%



**Figure 10-6.** Functional diagram of the drift chamber readout system.

and with minimal contributions from electronic noise or neighboring channels. A functional diagram of the readout chain is shown in Figure 10-6.

The position measurement should have a precision of  $140\ \mu\text{m}$ . With an average drift velocity of about  $25\ \mu\text{m}/\text{ns}$ , this distance translates into a time precision of approximately 5 ns. A timing resolution of about 2 ns would then degrade the measurement by 10%.

The energy loss in the gas is usually measured by integrating the charge collected on a sense wire. The fluctuation in the number of primary clusters is about 7%, so the charge measurement requires a precision of a few percent.

### 10.3.2 Preamplifiers

The preamplifier is of the transimpedance type, giving an output voltage proportional to the instantaneous current flowing into the sense wire. The current shape from the drift chamber wire has a rapid rise when the charge multiplication occurs, and then falls inversely with time. For a shaping time of 6 ns and a gas gain of  $2 \times 10^4$ , the current pulse from one produced electron is about 60 nA, with a dynamic range of about 250. Thus, a noise figure of about 1000 electrons is required. The current cuts off when the last positive ion reaches a field wire  $100\ \mu\text{s}$  later. In the inner wires of the chamber, where the count rates might reach 80 kHz under  $\times 10$  background conditions, there would be pileup on the tail of the previous event. The pileup would have little influence on the  $dE/dx$  information quality, but would cause some inefficiency in the discriminator. Therefore, the preamplifier includes a pole zero cancellation circuit with a time constant on the order of the drift period (300–500 ns). The preamplifier risetime will be optimized to the digitizing scheme selected.

The use of an unsaturated gas in the drift chamber requires careful control of the parameters that determine the gas density, *e.g.*, the gas temperature. The power consumption of the preamplifiers which are all connected to one end of the drift chamber is of concern

in this regard. The preamplifier will be built as a bipolar semicustom ASIC, which allows a reduction in power of a factor of 10 compared to previous designs. A goal of 4 mW/channel is envisioned. The chip will have six channels, and several chips can be assembled on a preamplifier card to match the geometry of the drift chamber. Each preamplifier chip is individually protected by a polyfuse. In the event of a chip failure which shorts the power input at the chip, the polyfuse will open and allow the other preamplifier chips on the board to remain powered. The fuse can recover from a temporary short or a surge by cycling the power off and on. It should be noted that a loss of six cells on a layer does not prevent the segment finder from working.

### 10.3.3 Digitizer Options

Two classes of implementation schemes have been studied for the digitizer: those in which the time and amplitude measurements are done separately, and those in which the two measurements are combined in a single device. In the latter class, the flash analog-to-digital converter (FADC) is the natural candidate. It provides both time and amplitude measurement capability, does not require external gating, can digitize continuously with no deadtime, and is naturally compatible with a clocked dual-ported latency buffer. When operated synchronously at a submultiple of the machine frequency, the FADC provides an automatic rejection of the machine-related background noise. An FADC readout system can be assembled from off the shelf components from multiple manufacturers. An FADC-based readout system is also a formidable debugging tool—it is essentially a 7000-channel digital storage oscilloscope.

The second class of implementation schemes uses a specialized device to measure the time. For deadtimeless operation, a flash time to digital converter (FTDC), or time memory cell, will fulfill the requirements. The amplitude measurement is still achieved with an FADC but no time information is expected to be extracted from the data; thus the timing objectives can be achieved at a substantially lower clock rate. It is slightly more complex, not immune to machine related noise, and requires a low-time-walk discriminator to trigger the FTDC. In the all-FADC approach, a discriminator is also present to generate the hit information for the segment finder logic, but then the timing requirements are very modest.

An interesting variation of the FTDC/FADC approach is to build many such channels on a single ASIC, including a shallow latency buffer. The advantages are lower power dissipation, lower cost per channel for the components, and a considerable saving of space required on the readout boards. A larger number of channels can then be packed on every readout board, thus reducing the number of readout crates.

### 10.3.4 Implementing the All-FADC Scheme

An FADC readout system has been studied on the test drift chamber together with various algorithms to extract the time information. The sampling frequency for the tests was 250 MHz, close to the 238 MHz machine frequency. The performance at lower frequencies (250/2 or 250/3 MHz) could be deduced from the same measurements. Since commercial FADCs have a very short time aperture, it is sufficient to use one sample out of every two or three to simulate the operation at reduced clock rate. At 83 MHz(250/3), the position resolution achieved with the FADC readout channel was 135–142  $\mu\text{m}$ . That corresponds to a degradation of 20% compared to the 115  $\mu\text{m}$  achieved with the TDC readout. At 125 MHz, the degradation reduces to 10%. It should be noted that the above figures are somewhat unfair to the FADC method. The time-to-distance function is obtained by an iterative process that uses only the TDC time measurement information. The same function is applied to the FADC time measurement, though it has been optimized for the TDC system. Expected improvements in the time extraction algorithms may lead to the firm conclusion that one-third of the machine frequency (79 MHz) is a satisfactory sampling rate. But at this time, the all-FADC system operating at 119 MHz (one-half machine frequency) remains our baseline design. No technical uncertainty is involved in building such a system, and it has been demonstrated to achieve acceptable performance with a drift chamber very similar to the final one. FADCs which operate at 119 MHz are expensive and power hungry. The use of pairs of slower FADCs with interleaved sampling is being pursued. The uncertainties of this approach are discussed in Section 10.3.6.

#### Time Measurement

One of the difficulties in achieving a very good timing resolution with an FADC in the *BABAR* drift chamber originates from the low-pressure, low-Z gas. This choice minimizes the multiple scattering and yields very well-defined track trajectories. However, it also reduces the density of primary electron clusters along the track to about one cluster per millimeter. For tracks that directly cross a wire, this represents a fluctuation in drift time of about 15 ns for the arrival of the first cluster. Fortunately, this time difference reduces rapidly as the distance of the track from the wire increases and soon becomes a marginal effect. However, the coarse cluster granularity has a serious effect on the pulse shape of the signal extracted from the wire. It causes a succession of current pulses that can cover the full drift period in the case of an anode-crossing track. As a result, the pulse shapes are very different from event to event, making the time extraction a difficult task.

Various de-convolution methods have been tried. They succeed in separating the individual clusters when they are sufficiently far apart in time, but at the leading edge, where the geometrical effect makes the clusters concentrate for tracks away from the sense wire, the

situation is very difficult. So far, the so-called TDC-mimic algorithm, which is a linear extrapolation of the first two bins to the baseline, gives the best results. Other algorithms are being studied. The timing performance might also benefit from tuning the preamplifier shaping. New, faster preamplifiers with adjustable shaping are now being installed on the test chamber to improve these measurements.

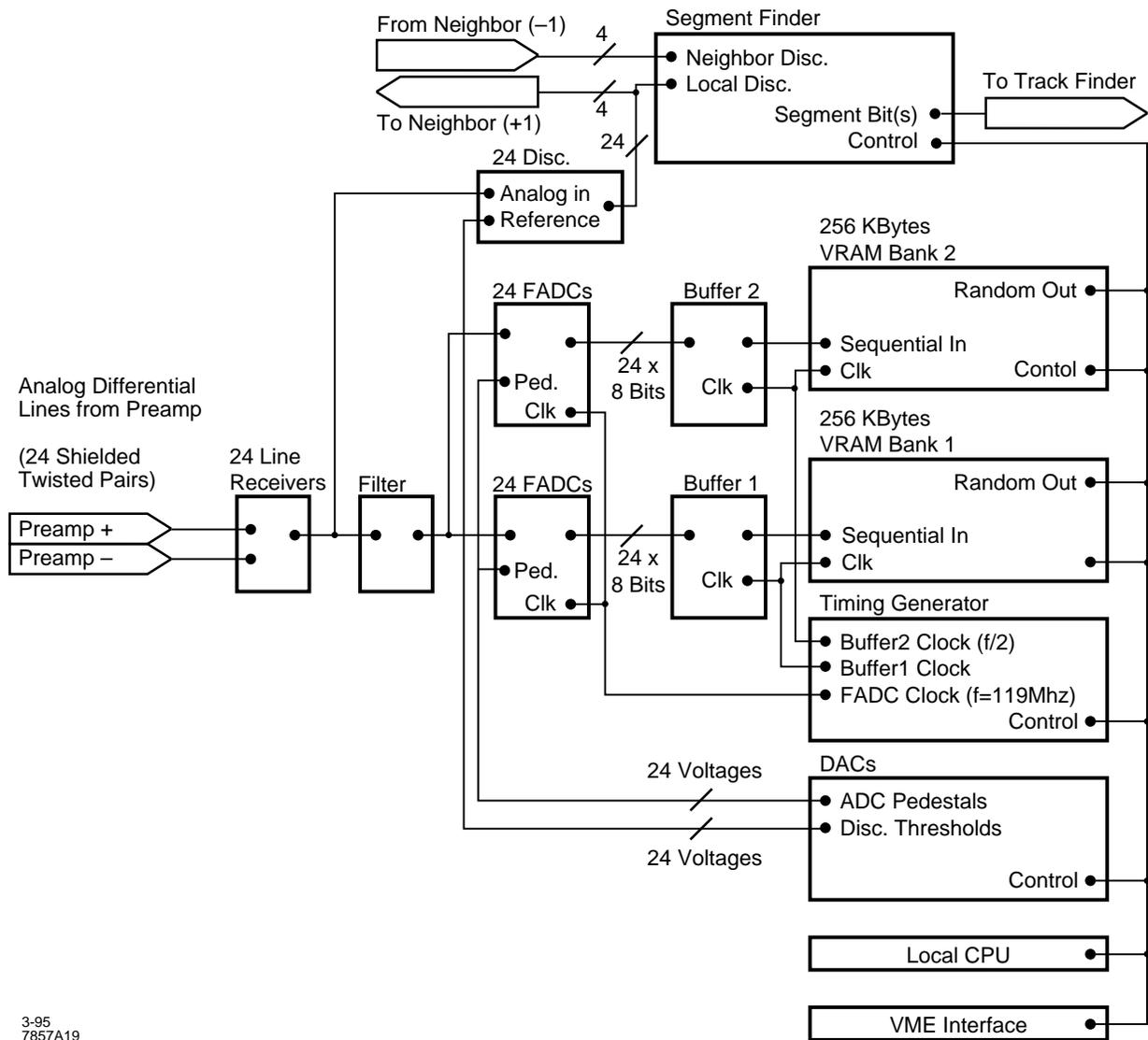
For the extrapolation algorithm to work accurately with small signals, it is necessary that the smallest signals, corresponding to single clusters, have an amplitude on the order of 10 ADC counts. This partially determines the dynamic range of the FADC. To prevent the digitizer from overflowing with large pulses, other multiplying factors have to be taken into account: 2 for track angle, 2.5 for highly ionizing particles, and 4 or 5 for cluster pileup, for a total factor of 250, or 8 bits. An effective dynamic range of 9 bits would be more comfortable.

### Charge Collection

The evaluation of the primary charges deposited in the volume of a cell is made by summing the FADC samples over the drift interval. The statistical fluctuations in the formation of the primary clusters already introduce an average variation of 7% along the full length of the track. Thus, a measurement of the integrated charge with a precision of a few percent will not degrade the  $dE/dx$  evaluation, and a precision of 6 bits is sufficient for the charge measurement. If this requirement is combined with that of time evaluation, the FADCs should have an effective dynamic range of 8 to 9 bits and an accuracy of 6 bits. It would then be satisfactory in principle to use a 6-bit digitizer, and obtain the 8- to 9-bit dynamic range with a nonlinear or a multislope conversion.

### Readout Module

The readout module is composed of two distinct sections: a front-end specific to the drift chamber and a common data acquisition section. The common section contains the latency buffers, some local RAM, the processor that will extract the physical features from the raw data, and an interface to the VME crate bus. The front-end section carries the FADC chips, the signal conditioning electronics, the discriminators that provide the wire hit information, and the segment finder logic, as shown in Figure 10-7. Each differential signal line from the preamplifier arrives at the readout board at a line receiver featuring a large common mode rejection. The output of the line receiver feeds a discriminator and a driver amplifier for the FADC chip input. This driver also includes a low pass filter to optimize the signal shape for the time extraction. An offset voltage is supplied to the driver from a 6-bit CMOS DAC, in order to adjust the FADC pedestal.



3-95  
7857A19

Figure 10-7. Drift chamber module block diagram.

The FADC chips are 8-bit devices operated at a clock frequency of 119 MHz, which is almost twice the maximum speed at which the VRAM buffer memories operate. Thus, it is necessary to use two interleaved VRAM units for each FADC, interfaced through two buffer latches. The chamber geometry will determine how many FADC/discriminator channels are put onto each board. For the baseline design, the chamber is organized into sets of four layers each, called superlayers, which are themselves organized into groups of 24 cells called supercells. Since the data acquisition board supplies the information to the segment finder, and the segment finder needs the information from one supercell, the number of channels on the board will be equal to, or a multiple of, the number of cells in a supercell. The segment finder logic also requires that some overlap be provided between adjacent supercells. An interconnect is thus required between neighboring readout modules, forming a daisy-chained closed loop. In the baseline configuration, three signals would be carried from one module to the next.

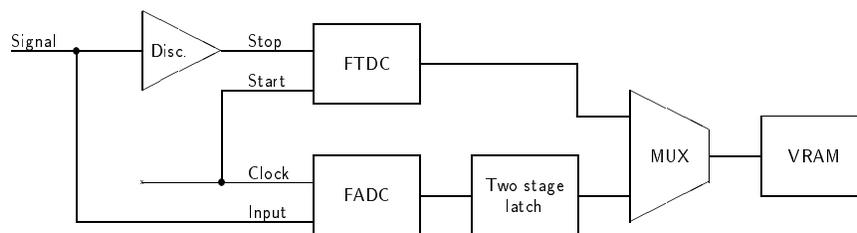
### 10.3.5 Calibration

The gain calibration is accomplished by injecting pulses at the preamplifier input. To limit the number of pulser cables going to the preamplifiers, a compromise is made between the flexibility in the trigger patterns and the size of the system. Each channel of a preamplifier integrated circuit can be pulsed independently. It is therefore possible to detect channels shorted together and to measure crosstalk. Each preamp cable has two or four pulser lines, for a total of about 1400 lines in the latter case. The pulser signals originate from the readout module. A lesser number of pulser lines is required if the pulses are distributed in a ring geometry. One pulser line is then connected to one of the four channels of each preamplifier chip belonging to that layer. This requires 320 pulser lines. Every line serves at least 20 channels, but this granularity is still small enough to generate useful patterns to exercise the trigger logic and the data acquisition system, in addition to the basic gain calibration function. The pulser system only calibrates the electronic gain. It must be combined with the information extracted from the data itself, such as the histograms of the integrated charge for each cell, in order to account for the wire gain.

### 10.3.6 Research and Development

#### Operation at a Lower Sampling Frequency

The research program concerning the FADC time extraction algorithms will continue, since even a slight improvement would justify reducing the sampling rate from 119 MHz to 79 MHz. At this lower rate, a CMOS FADC may be used, reducing cost by a factor of 5, power



**Figure 10-8.** *Flash TDC/FADC channel principle.*

consumption by a factor of 4, and space on the the readout module by a factor of 10. Current candidate CMOS FADCs have only 6 bits and thus would have to be operated in a nonlinear mode. This has been done for other technologies but needs to be studied for CMOS devices. A test board will be constructed to evaluate this possibility and measure the maximum practical dynamic range expansion of the 6-bit CMOS FADC that is available commercially. The LBL group also has designed a 6-bit 100 MHz CMOS FADC prototype chip. It will be available for testing shortly. An 8-bit device could in principle be built with the same technology, with a somewhat lower maximum sampling frequency yet to be specified. It is also possible to use two slower FADC chips and operate them with split phase clocks to increase the sampling rate. Two CMOS units could be operated at 59.5 MHz to achieve an effective 119 MHz sampling rate. However, this requires a good matching of the transfer functions of the two FADC chips and might require a more expensive, more linear version of the chip. A test board will be made with CMOS devices to check the dynamic range compression method and the split phase technique.

### Time Measurement with an FTDC

An alternate approach, aimed at lowering the sampling frequency, is to complement the FADC readout with an FTDC time measurement. The FTDC measures the time with very good resolution, and the FADC measures the charge only, while still providing a redundant lower precision time measurement. This approach is interesting if the sampling frequency is brought down to that of the VRAM buffers (40–60 MHz). Only one VRAM chip would then be required per channel. This is quite feasible since the FADC data would not be the primary source of time information.

As shown in Figure 10-8, the FTDC produces an encoded measurement of the elapsed time between a pulse of the sampling clock and the instant the discriminator fires. There is only one byte of TDC information for each hit, and it can be inserted in front of the FADC byte stream. The two latch stages following the FADC provide the required digital delay for that operation. This approach requires only one VRAM per channel instead of two, but

the FTDC is an extra device. The FTDC function could also be achieved coarsely with a 6-bit CMOS FADC sampling a voltage ramp synchronized with the sampling frequency. In either case, the cost of this approach would be equivalent to that of the 79 MHz pure FADC operation and be subject to the same uncertainties relative to the dynamic range compression feasibility. However, since the FADC would no longer measure the time, the dynamic range requirement could probably be relaxed by a factor of 2.

### Readout ASIC

An integrated drift chamber readout circuit could reduce the cost and complexity of the readout system. This readout IC operates like that for the silicon vertex detector, *i.e.*, multiple channels of electronics are read out through a single circuit only after a Level 1 accept. This architecture reduces the processing time on a DAQ card since only channels with hits are transferred from the readout IC. The architecture also minimizes the number of readout (VRAM) channels required on the DAQ cards and reduces the total number of readout crates required. Eight channels of drift chamber readout electronics are expected to fit on one integrated circuit.

The proposed drift chamber readout IC is composed of four main components: a time-to-digital converter (TDC) marks the time of the threshold crossing to an accuracy of less than 1 ns; a 6-bit FADC samples the drift chamber output pulse at 100 MHz, a dual-port DRAM stores the digitized drift chamber output pulse for the entire trigger latency of 10  $\mu$ s; and a SRAM secondary buffer stores up to 3  $\mu$ s-long sets of drift chamber hits.

The TDC is crucial in determining the exact time of the threshold crossing. The TDC operates by dividing the input clock period into 32 bins. The values of the bins are stored during each threshold crossing, and these values determine the portion of the clock period in which the threshold crossing occurred. The operation of the core of the TDC, the delay elements and the phase detector, has already been tested with a previous IC. An updated TDC currently in fabrication will test a new phase detector and decoding scheme.

An FADC built with auto-zeroed comparators has been designed. Simulations show that 6-bit operation at frequencies up to 100 MHz is feasible. The comparison scheme itself is good up to 10 bits. A 6-bit test circuit has been submitted for fabrication. Verification of operating parameters must be tested.

The DRAM and SRAM cores are designed and in fabrication. Some simple tests will be undertaken to determine if the memories operate correctly. These memories are built with memory generators. This allows adjustable size memories to be implemented for each design. Some slight modifications might be needed once the drift chamber IC channel pitch is determined. There should be no effect on the memory operation.

Once the above blocks have been tested, they can be connected at the board level to test the system functionality. A complete drift chamber channel can be designed with these blocks at the IC level concurrently with the system test. A four- or eight-channel prototype chip must be built to study power, noise, and crosstalk issues. The die size of an eight-channel chip is expected to be 16 mm<sup>2</sup>.

## 10.4 DIRC

---

### 10.4.1 Requirements

In the DIRC, the Cherenkov angle is determined from the position of hit photomultipliers among the array of 13,400 tubes situated on the standoff region backplane. The pulses are the result of the impacts of single photons on the photocathodes. The times of arrival of the photons from a given event are scattered within a 50 ns time window.

The electronics requirements are: to record single photoelectron hits quickly enough to avoid being swamped by the accidental rates, to provide 1 ns time resolution, to discriminate against background, to implement modest amplitude measurement (6 bits) in order to monitor the gain of the tubes and maintain the threshold settings, and to provide adequate bandwidth to safely transfer the physics data to the data acquisition system. Special calibration runs are needed to determine pedestals and gains.

### Possible Readout Implementations

Several options have been considered, and their cost and performance evaluated: (1) full sampling, (2) full sampling with zero-suppression, (3) zero-suppression with local buffering with the Level 1 strobe starting the transmission of the data that arrived during the trigger latency, (4) the same as (3) but with only the data which arrived during the trigger timing window being transmitted. The above architectures can be implemented many ways. They differ in location of the components (on the detector or on the readout modules), the cable plant, number of readout modules and crates, amount of VRAM, and bandwidth of the transmission medium between the detector and the readout modules. A comparison of architectures is shown in Table 10-4.

At this time, options (3) and (4) are preferred because the number of cables is kept to a minimum, the amount of data to be sent to the data acquisition system is sufficiently reduced so that the whole DIRC readout resides in one or two crates, and custom integrated circuits can be used. The price is modest, and the custom ICs are not too complex.

Option	Number of Cables	On Standoff	Mbytes/s	PMs per VRAM Chip	VRAM Chips	Readout Modules	Readout Crates
1	13.4k	No	215k	2	6700	210	14
2	—	Yes	6k	27	500	65	4
3	—	Yes	100	1.34k	10	3	1
4	—	Yes	10	13.4k	1	1	1

**Table 10-4.** Comparison of four DIRC readout options.

Choosing between options (3) and (4) requires additional detailed study to judge whether the gain in bandwidth justifies the extra complexity of the digital front-end logic, since these added features can be easily implemented on the readout module.

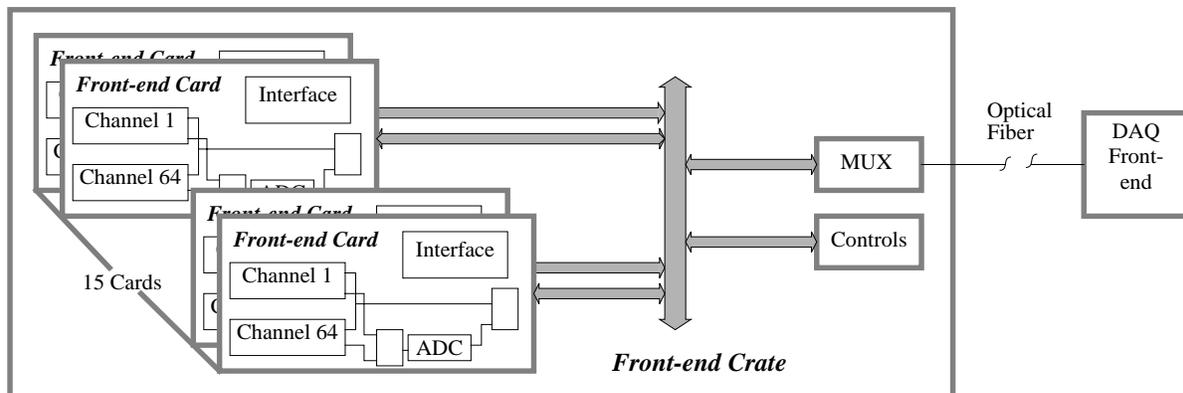
Although the proposed architectures make the use of custom chips attractive, it has been decided to present, as the baseline implementation, a system in which discrete components are used wherever possible. It is believed, however, that an optimal system will be built mostly out of custom or semicustom integrated circuits for lower cost, improved performance, and more reliability. Such an implementation is explored in Section 10.4.3. The baseline solution implements option (4) as far as the functionalities are concerned, but the bandwidth characteristics are derived assuming option (3). This allows option (3) to be a fall-back solution for option (4).

## 10.4.2 Baseline Design

### Overview

The plan is to install the front-end electronics as close as possible to the tubes. A data driven scheme has been devised in which digitization occurs only when the pulse is above a threshold. The time of each hit is measured to 1 ns with higher precision left as an option. Amplitude measurements with 6-bit precision are foreseen for calibration events only. Although the pulse height carries no physical information on a single hit, it would be useful to record it on an event-by-event basis to sort out pileups and provide assistance with debugging. This possibility is further explored in Section 10.4.3. An option is to add a high threshold discriminator on each channel to flag hits with anomalously high pulse heights.

Digitized data are stored for the duration of the Level 1 accept latency ( $10 \mu\text{s}$ ) and are read out upon receipt of a Level 1 strobe. An intermediate stage between the front-end and the readout module is needed to synchronize the data flow and multiplex it to 15 optical links. One readout crate houses the entire DIRC readout.



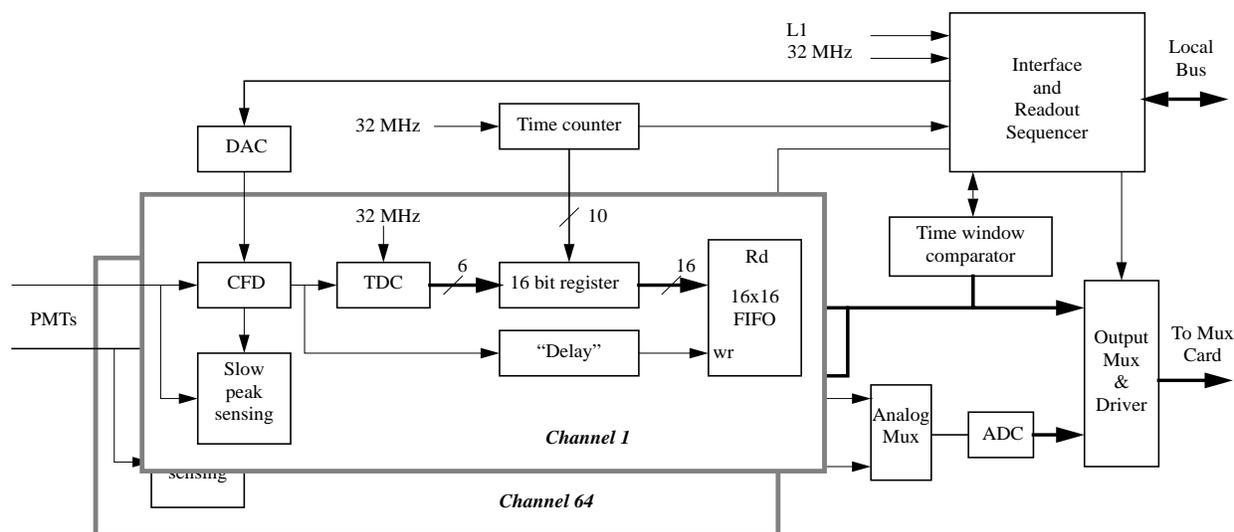
**Figure 10-9.** *Layout of the DIRC front-end electronics.*

The layout (Figure 10-9) calls for circuit boards located in crates situated near the tubes. Each card would carry the front-end electronics for 64 tubes. The idea of integrating the front-end electronics on boards shared with the high voltage system, attached to the back of the tubes, has been abandoned for the present because of the complexity of the interconnections and because the front-end electronics would get buried within the magnetic shield of the standoff box.

Short cables connect the bases to the circuit boards. A block diagram of the front-end is given in Figure 10-10. The main features of each channel are: timing—a constant fraction discriminator (CFD), TDC, and FIFO; the readout logic; and the slow analog calibration chain. A preamplifier with a gain of four and a bandwidth of 1 GHz may be needed at the input, depending on the choice of photomultiplier tube. Most of these features can be implemented with off-the-shelf components. A more detailed description of the items is given below.

## Timing Chain

After the CFD, the time is measured by a double-tiered TDC. A coarse range is obtained by parallel loading 10 bits of a time register from the time counter, which records one pulse out of 16 from the storage ring rf (1 tick every 33.6 ns). Time measurement over the total Level 1 latency requires 14 bits. Sixteen bits have been allowed. The vernier range is obtained by using a high-precision 6-bit TDC. Early versions of this device have been built for the DELPHI experiment [Gen93]. The output of the 16-bit register is connected to a FIFO deep enough to keep 16 hits. A time window comparator and interface, and readout



**Figure 10-10.** Block diagram of the DIRC front-end electronics.

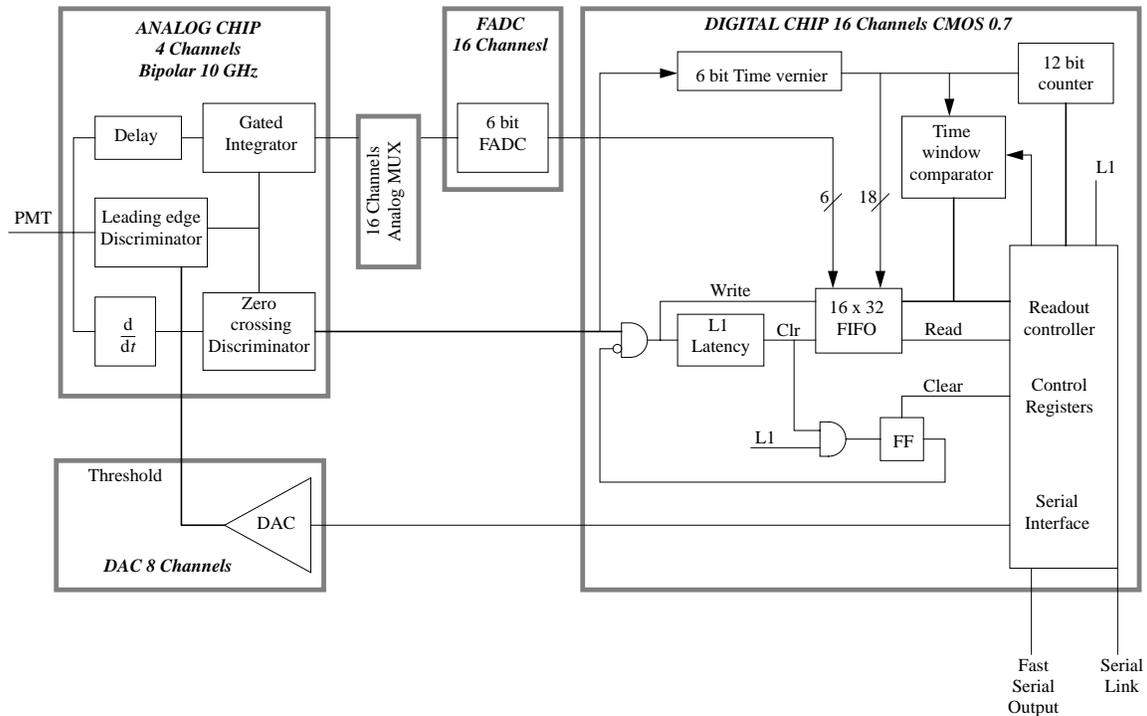
sequencer enable the transmission of the correct data. The trigger time is compared to the hit times stored in the FIFO, and those which are in time are transmitted. The TDC and the FIFOs will be implemented inside a 64-input standard cell CMOS digital chip. The interface and readout sequencer adds the address of the photomultiplier within the card to the data word, which becomes a three-byte string. The output MUX/driver sends the data to the multiplexer boards.

### Analog Chain

The pulse height will be measured during calibration. This requires an analog chain incorporating slow, charge-sensitive circuitry on each tube, which feeds a 6-bit ADC through a 16-to-1 analog multiplexer.

### Multiplexing

From each front-end circuit board, the three bytes of data are sent at 400 kbytes/s to an intermediate crate. Assuming 20 Mbytes/s for the bandwidth of the readout crate backplane bus and a safety factor of about 4 to absorb the fluctuations in event size and trigger timing leads to a layout with 15 circuit boards per front-end crate. To minimize the cable plant, it is advantageous to add a multiplexer board in the sixteenth slot of such a crate. Each multiplexer board will concentrate the data of its front-end crate onto a 12 Mbytes/s optical



**Figure 10-11.** Block diagram of the DIRC front-end electronics showing the suggested implementation with ICs.

fiber serial link. This board will also add a 4-bit board address to make the data word 4 bytes. There are 16 front-end crates, each with its own readout fiber.

### 10.4.3 Research and Development

Most of the front-end described above is assembled from commercial components. It is well understood and buildable and has been used primarily for budgeting purposes. Only the TDC and the FIFOs are ICs. It is believed that a somewhat improved architecture can be built by using ICs throughout (Figure 10-11). Although such a design entails more development work, it should cost less and lead to a more reliable system.

In this scheme, the basic granularity is 16, although a better value may emerge from simulations. The philosophy of the timing measurement in this implementation is exactly the same as in the baseline design. The difference is that the amplitude will be measured

during data taking. There is one FADC (6-bit, 100 MHz) serving 16 photomultipliers. If the rate is too high, not all tubes from a given readout section of 16 in an event will have their amplitudes measured.

The most elegant layout calls for circuit boards, attached behind the tube bases, which carry the front-end electronics for 16 tubes. There is enough room on these boards to consider sharing them with the high voltage system, should we choose to do so. Short cables connect the bases to the circuit boards. On each PC board, there are four analog chips, one or two digital chips, an analog multiplexer, a 6-bit 100 MHz FADC, two eight-fold DACs, and the output circuitry to send the data to the readout modules via digital multiplexers. More detailed descriptions of these are given below.

### Analog Chip

An analog chip services four photomultipliers. For each tube, there is a discriminator and a shaping circuit to feed a charge-sensitive ADC. From the analog chip emerge the four discriminated output pulses, which are fed to the digital chip, and the four analog outputs of the shaper, which are directed to a 16-input/1-output analog multiplexer. In addition to the four signal inputs, the analog chip receives four threshold inputs from DACs. The choice among bipolar, CMOS, and BiCMOS processes is being studied.

### Digital Chip

Each circuit board has one digital chip which accepts inputs from 16 tubes. The time measurement is identical to that of the baseline design. The interface and readout sequencer and the time window comparator could be made as a second separate chip or remain discrete components. These decisions will be made as the design matures. The digital chip receives 16 signals from the analog chip, the master clock, the trigger signal, and a calibration line used to measure the pedestals. The calibration proceeds as follows. A pulser controlled by the data acquisition system generates a signal which enables the discriminator threshold to be bypassed. The signal is also delayed to provide a pedestal trigger with the proper latency. For calibration or noisy channel suppression, individual tubes can be masked using a slow serial command line on the circuit board. The outputs of the digital chip are a 5-byte data bus daisy-chained across six circuit boards towards the fiber driver, and a 4-bit address bus to command the analog multiplexer. Standard cell CMOS is the technology chosen for the digital chip.

## Front-End Layout

Nearly one thousand circuit boards will be used, each having four analog chips, a digital chip, an analog multiplexer, an FADC, and two eight-fold DACs. On top of the input and output signals already mentioned, cabling will carry the low power distribution lines. Slow control will have to be implemented locally. These features have yet to be designed.

## Data Transmission to the Readout Crates

The data rate from all the digital chips is 100 Mbytes/s. Assuming the front-end readout modules accept data at 12 MHz, the flow must be shared among ten streams, each receiving data from 84 front-end sections. This corresponds to ten multiplexers with 84 inputs and one output, where a single address byte, that of the readout section, will be added to the data. Optical fibers similar to those considered for the vertex detector are appropriate. Alternatively, two 1 Gbytes/s optical fibers can be used.

### 10.4.4 High Voltage System

If a traditional resistor-divider photomultiplier base is chosen for the DIRC photodetector, a photomultiplier with a maximum operating voltage of 2 kV will draw a current of approximately 0.3 mA or less. In the proposed system, the high voltage will be distributed over long cables to a HV motherboard mounted on the photomultiplier support structure. Each motherboard services 16 bases, for a total current load of approximately 5 mA per HV channel. The motherboards will be connected to commercial multichannel high-voltage power supplies.

The high-voltage power would be distributed via standard RG-59 HV coaxial cable, utilizing a new connector scheme developed by LeCroy in collaboration with the AMP corporation. This connector is an eight-fold coaxial block connector, fully rated and UL listed for 5000 V. This results in significant savings over traditional SHV coaxial connectors. Connections to the motherboards would be made using standard SHV connectors.

### 10.4.5 Low Voltage Power Supplies and Control Systems

The on-detector front-end electronics has been estimated to dissipate a total of about 20 kW. The standard crates in the baseline design provide the needed low voltage and are water cooled. Control systems have yet to be designed.

---

## 10.5 Aerogel Threshold Cherenkov Counter (ATC)

---

### 10.5.1 Requirements and Overview

The signal processing provided by the off-detector electronics should be capable of fully exploiting the intrinsic performance of the ATC. For good efficiency, the threshold should be set between 1 and 1.5 photoelectrons and the timing resolution should be less than a few tens of nanoseconds in order to reduce the background from off-beam particles showering into the detector. An approach based on discriminators and TDCs, or one based on flash ADCs, can be used to achieve this goal. In the framework of the *BABAR* pipelined DAQ system, the most natural choice is the second of these options. This also provides some additional advantages:

- greater flexibility in setting different thresholds depending on the expected signal amplitude, which is a function of the particle momentum;
- shorter possible deadtime in high background conditions; and
- much more powerful diagnostic capability.

In the baseline design of the ATC detector, fine-mesh photomultiplier tubes (FM PMT) are used as photon detectors. After a possible front-end amplification and shaping, these can have an average pulse height and a shaping time similar to those expected from the drift chamber preamplifiers. Therefore, it may be convenient to use the drift chamber readout modules. The dynamic range of the expected ATC may be easily accommodated in the 8-bit effective range of the drift chamber readout system.

### 10.5.2 Front-End Electronics

With the Hamamatsu fine mesh PMT available today, the gain in a magnetic field of 1.5 T is about  $10^5$  and a low noise current preamplifier is required. To obtain a signal suitable for a commercial ADC, a transresistance of  $60 \text{ mV}/\mu\text{A}$  is sufficient. This would give a signal of  $\sim 50 \text{ mV}$  for a single photoelectron pulse at  $10^5$  PMT gain. This value is considered to be the maximum requirement, since the PMT response still can be optimized with respect to the orientation in the magnetic field and improvements are expected in the performance of the next generation of PMTs. Using the specifications of a typical amplifier like the LeCroy TRA1000, one obtains an estimate of the noise of 0.5 fC FWHM, which is equivalent to 0.03 photoelectrons. The preamplifier is to be mounted directly on the PC board that supports the PMT and can be connected to the off-detector electronics via twisted-pair shielded cables.

### 10.5.3 Readout

The details of the drift chamber readout board appear in Section 10.3. We plan to make use of the processing power available on the board to perform the feature extraction after the arrival of the first level trigger. Once converted into amplitude and time information, the amount of data to be shipped to the event builder would be negligible, on the order of a fraction of a kilobyte. Provisions will be made to optionally send the sampled waveforms for diagnostic and calibration purposes.

### High Voltage

The fine mesh PMT will be operated at 2.5 kV, with a steady current of approximately 2 mA. Existing commercial modular high voltage power supply systems are well-suited to the ATC HV power requirements. The ATC requires 144 HV channels. In the proposed system, the high voltage will be distributed over 12 cables to the PC boards where they will be split into two 6-pin connectors, each carrying the HV power needed by a PC board.

## 10.6 Calorimeter

---

The electronics for the CsI(Tl) calorimeter must read out the 6780 crystals in the calorimeter and provide information to the lowest level trigger as well as be able to respond to a trigger, and fetch and process the data corresponding to that trigger. This section describes the requirements of the calorimeter electronics and a proposed architecture which meets those requirements [Hal94].

### 10.6.1 Requirements

The following items are the primary requirements [Req94] for the calorimeter electronics:

**The calorimeter data acquisition system.** The system should not make the resolution for any photon with  $E_\gamma \geq 20$  MeV significantly worse than that due to the intrinsic properties of the detector alone. Specifically, the data acquisition contribution to the error should be less than half that of the CsI.

**Dynamic Range.** The calorimeter should have a dynamic range capable of reading out showers in beam events with energies from 20 MeV to 13 GeV with a resolution of 100–200 keV at 20 MeV and also provide adequate resolution for reading out low-energy gamma rays from radioactive sources for calibration. The dynamic range required for beam events is 17 bits. Allowing for a source calibration requires 18 bits of dynamic range, which is the design goal.

**Incoherent Noise.** The rms incoherent noise for a single crystal should not exceed an equivalent noise charge (ENC) of 350 electrons. This is the best that can be done at acceptable cost. This number combined with the number of photoelectrons collected for energy deposited (Chapter 7) determines the electronics contribution to the shower energy resolution.

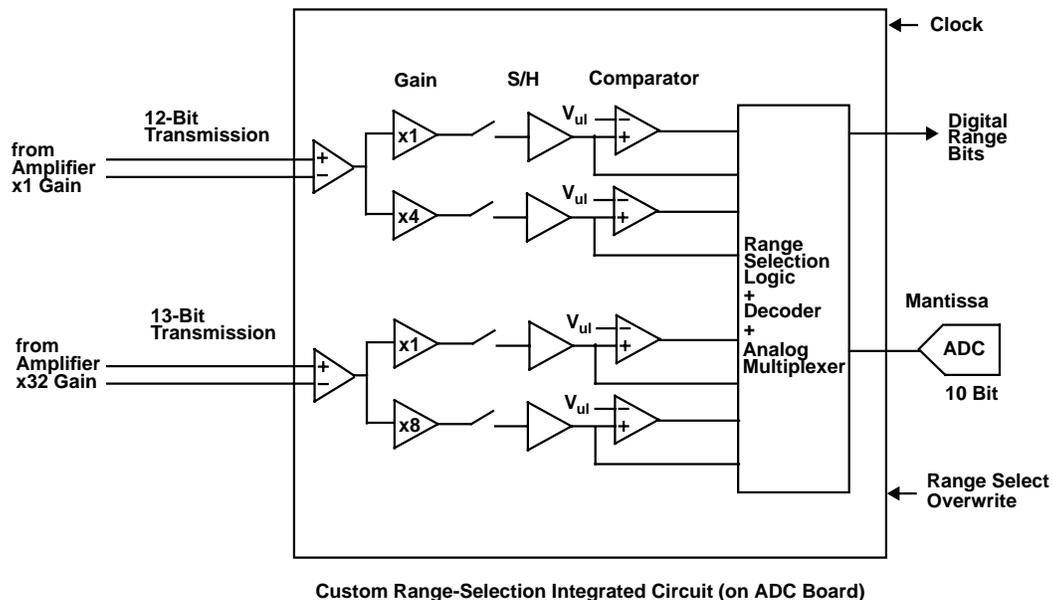
**Digitization Contribution.** The contribution to the resolution from the digitization should be less than half the CsI(Tl) resolution. This requirement can be met with a 10-bit ADC.

**Reliability.** The electronics components which are inaccessible must be provided in two independent sets. Except for the crystals in the endcaps, access to components mounted on the crystals will be impossible. Additional components are located inside the structure on the end of the detector, but will be accessible by opening the detector endcaps.

**Response to Background.** The electronics should be designed to provide optimal performance with the expected levels of background. At high background rates, soft photons provide a source of noise which must be dealt with in the calorimeter (Chapter 12). This requirement makes demands on the shaping time of the electronics and the signal processing that must be performed on the output of the calorimeter.

### 10.6.2 Overview

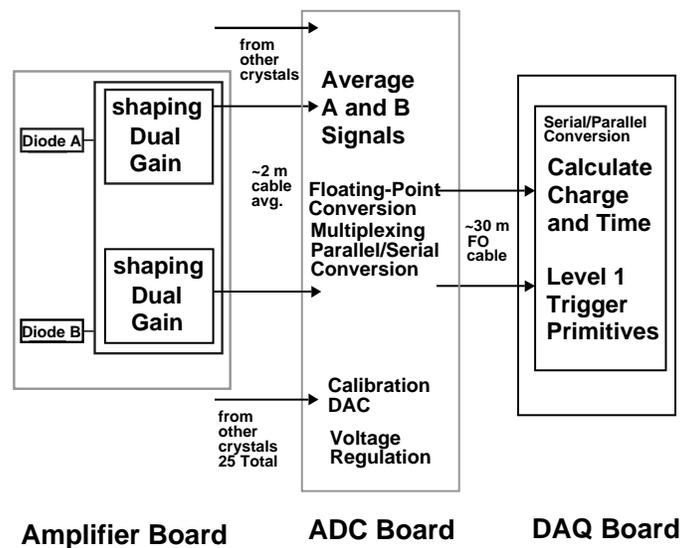
The dynamic range requirement of the calorimeter combined with the large potential for beam-related rf pickup at PEP-II make digitizing the signal close to the CsI(Tl) crystals advantageous. The proposed design minimizes the cable length and dynamic range of analog signals in the system, and places the digitization just outside the rf-shield of the calorimeter, allowing access to the ADCs.



**Figure 10-12.** Block diagram of gain stages and range encoding.

To minimize the distance the 18-bit dynamic range analog signal travels, an amplifier with two gain stages is mounted within a shielded assembly on the rear of each crystal. By choosing amplifiers with gains of unity and 32, only a 13-bit dynamic range is required on the analog cables which go from the crystals to the digitizing board. At the digitizing board, a custom IC receives the two analog signals and provides two additional gain ranges. The outputs of this device are an analog signal, which is sent to an ADC at one of four possible overall amplifications, and two bits which specify which of the four gain ranges is being used. The additional gain stages allow an 18-bit dynamic range floating-point number to be constructed from a 10-bit digitization of the analog signal (the mantissa) and the two bits specifying the range (the exponent). Figure 10-12 shows a block diagram of the layout of the gain stages and range encoding.

Because it will not be possible to know which samples contain useful information until a trigger decision is made, the digitized data flow down two paths. The first path places the digitized floating-point data in a VRAM buffer where they wait for a trigger decision. When a trigger is received, the data acquisition system must use the trigger time to select the correct range of stored data to analyze and process and then transfer these data to the next level. In addition, the data from the calorimeter are used in the trigger decision, so digital sums of energies must be formed and sent continuously to the central calorimeter trigger logic. The trigger sums are more easily performed on the digital stream as the multiple gains stages make analog sums more complicated.



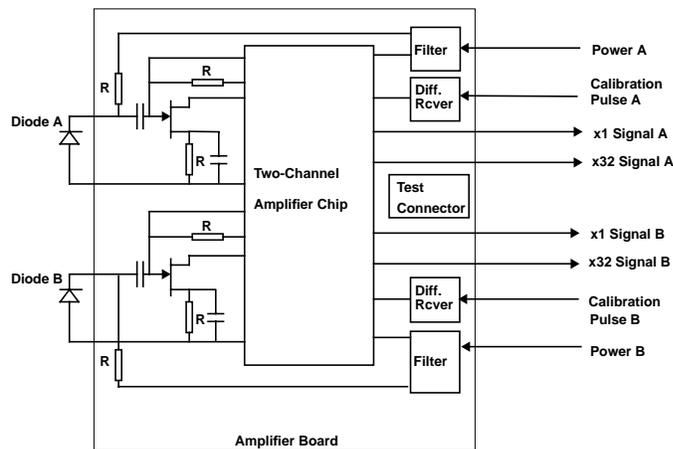
**Figure 10-13.** Calorimeter system diagram showing preamp, ADC, and data acquisition (DAQ) functions.

Figure 10-13 shows a block diagram of the components of the proposed readout system. The system has three main board types.

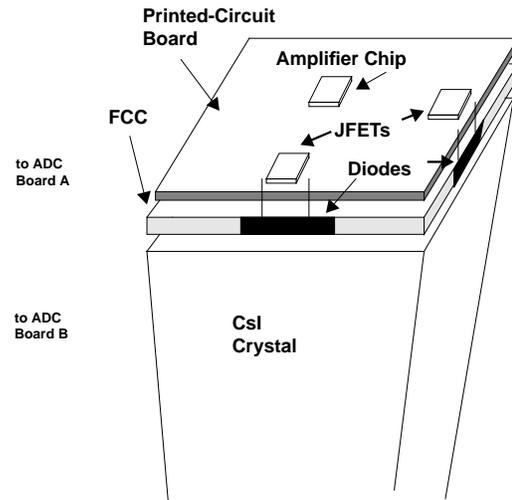
**Preamp boards.** One of these is mounted on the back of each CsI(Tl) crystal. They are not accessible after the calorimeter is constructed, so they must be reliable and contain a minimum of components. Each board provides an independent channel for each of the two photodiodes on the crystal (outputs A and B). Since it contains the first amplification stage, it must be well shielded. In addition, the power consumption should be low, as the temperature of the crystals needs to be regulated.

**ADC boards.** These are located at the end of the detector, just outside the rf-shield of the calorimeter. They should be accessible on the timescale of one day. Each board receives the signals from the A and B diodes on each of the approximately 25 crystals corresponding to the trigger towers. It contains the custom range encoding chip which chooses the gain then outputs two bits specifying the gain used. The board also includes the ADC chips, and its output is a serial digital stream.

**Readout modules.** These boards each receive the serial digital data from one ADC board, then de-serialize the data and store them in a VRAM buffer for later analysis. They also form trigger sums and ship the trigger data to the central calorimeter trigger logic. It is



**Figure 10-14.** Block diagram of preamplifier.



**Figure 10-15.** Calorimeter preamp/crystal shield assembly.

envisaged that there will be processing power on these boards to apply the gain and pedestal corrections to each channel and to extract the total charge and time information for each crystal when a trigger has been received.

Each of these components is discussed in more detail below.

### 10.6.3 The Preamplifier Card

Figure 10-14 shows a block diagram of the preamplifier board. It connects directly to the photodiodes on the CsI(Tl) crystals and provides, on output twisted pair cables, signals with  $\times 1$  and  $\times 32$  gain for both the A and B photodiodes. The main components are a custom two-channel amplifier chip (or perhaps two one-channel chips), and an external JFET. The shaping time for a CsI(Tl) pulse is optimally near  $3 \mu\text{s}$ ; however, the presence of background in the form of a large number of low-energy photons may mean that shorter shaping times are needed to give better background rejection. Shaping times less than  $\sim 1 \mu\text{s}$  mean a significant loss of signal from the crystal. In the following, a shaping time of between 1 and  $3 \mu\text{s}$  is assumed.

Providing most of the circuitry for the preamplifier on a single custom integrated circuit minimizes possible breakdown from faulty connections between components. The preamp IC contains completely separate paths for the power, signal, and calibration of both diodes. The only element in common between the two pathways is the substrate. The preamp chip

provides charge-to-voltage conversion, shaping, and dual-gain differential driver stages at  $\times 1$  and  $\times 32$  gain.

The 18-bit dynamic range input to the preamp board requires that the diode and preamp be contained within a single shielded assembly. One possible layout of the crystal/diode/preamp assembly is shown in Figure 10-15. In this design, the preamp board is mounted on a stand attached to the crystal assembly. The diodes are then connected directly to the JFET on the preamp board, minimizing the distance over which the 18-bit analog signal passes. This crystal/diode/preamp package is enclosed in a metal shield, which can also be used for cooling. The heat output of the preamplifier board is estimated to be between 50 and 200 mW per channel.

Each preamp board requires many cables in order to provide independent power, calibration, and signal paths for both photodiodes. A reasonable estimate of the cable plant is:

- Power:  $V_{dd}$ , bias for the photodiode, ground,  $V_{FET}$ ;
- Signal: twisted pair for  $\times 1$  and  $\times 32$  signals; and
- Calibration: differential calibration pulse.

Thus, there are ten wires for each diode. Separate cables are provided for the A and B diodes. These cables need to be approximately 2 m in length. Filtering for the power lines and the receivers for the calibration pulses are provided on the preamp board.

### 10.6.4 Digitizing Board

The digitizing board collects the analog signals from groups of preamp boards, digitizes the data, and ships the digital data in serial form to the readout boards. Approximately 280 of these are needed to read out the A and B diodes from all the calorimeter crystals. Figure 10-16 shows a block diagram of the ADC board. A working assumption is that each board serves both diodes from 25 preamplifier boards. This board must provide the following basic functions:

- For each input channel, it must provide a floating-point number covering an 18-bit dynamic range. Our solution is to have a 10-bit digital mantissa and 2-bit exponent. This requires a 10-bit ADC.
- It must sample with a period which is short compared to the shaping time after digital filtering. This is needed to reduce background and improve the feature extraction. A reasonable maximum rate at which the ADC board will be required to digitize a single channel is 4 MHz (five samples over a peak for a  $1.25 \mu\text{s}$  shaping time).

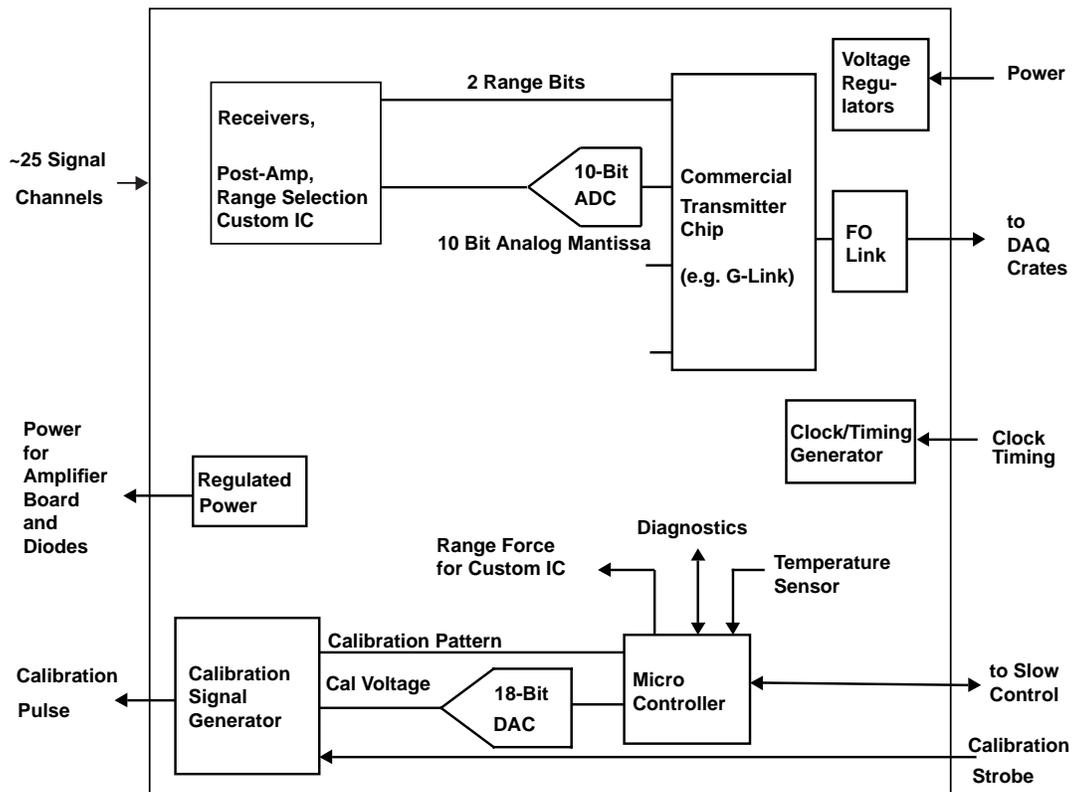


Figure 10-16. Block diagram of ADC board functions.

- It must provide the calibration signals to the preamp boards. In addition, the board should be able to control the range encoding chip and perform diagnostics and monitoring on the preamp and ADC boards.
- It must provide the power to the preamp boards.

One of the key components of the ADC board is the custom analog range encoding (CARE) chip which takes the four analog input signals from the A and B diodes ( $\times 1$  and  $\times 32$  for each of A and B), averages the A and B signals on the appropriate range, and then passes these through two more gain stages. The outputs of these four gain stages are sampled at the digitization sampling frequency and are compared to threshold levels. The chip places on its output the analog signal associated with the first gain stage not above threshold (for example 90% of the full range for that gain stage) and two bits signifying which of the four gain stages were used. The CARE chip can select one of three modes for operation:

- use input from only the A diode;

Energy Range (MeV)	Bin Size (MeV)	RMS Resolution (MeV)
0 to 46	0.050	0.029
46 to 366	0.40	0.229
366 to 2925	3.17	1.83
2925 to 13000	12.69	7.33

**Table 10-5.** *Calorimeter bin size and resolution as a function of energy.*

- use input from only the B diode; or
- average the inputs from the A and B diodes.

Normally, the signals from the diodes are averaged. In the case of the failure of one diode, the CARE chip can be switched to use only the input from the working diode, and the failed diode can be disconnected.

Table 10-5 summarizes an example configuration and shows how the resolution and bin size vary as functions of energy and range selection in the CARE chip. The numbers assume that the overall gain ranges are 1 ( $1 \times 1$ ), 8 ( $1 \times 8$ ), 32 ( $32 \times 1$ ), and 256 ( $32 \times 8$ ), and that a 10-bit ADC is used to digitize the data. The resolution is taken as the bin size divided by  $\sqrt{12}$  with an additional factor of 2 to account for losses in clustering. The results are summarized in Figure 10-17 which shows the resolution from the digitization compared with the specified resolution for the CsI(Tl) crystals.

Since the signal is presented to the ADC chip as a DC level, the ADC chips should be able to deliver the specified number of bits of accuracy. We are exploring whether it is more cost-effective to assign a single ADC to each channel or to multiplex several channels into fewer, higher-speed ADCs.

Assuming the highest rate of sampling (4 MHz) and multiplexing, the 12 bits from each of 25 channels into a single serial output requires a capability of 1.25 Gbits/s (1.5 Gbits/s with protocol overhead). For this rate, the best technology is optical fiber.

The ADC board also performs electronic calibrations and monitors the ADC board and the preamp. A small microcontroller is placed on the board, with a low speed serial link to the outside world. This controller can read the monitoring inputs and control the data acquisition and calibration lines.

The location of the preamps and ADC boards is shown in Figure 10-18. Each group of 25 preamp cards has a cable plant of 500 wires, many of which are power and ground, and the rest of which carry signals to the ADC board serving this group of cards. A possible solution to the problem of handling this number of cables is to have a fanout board which connects

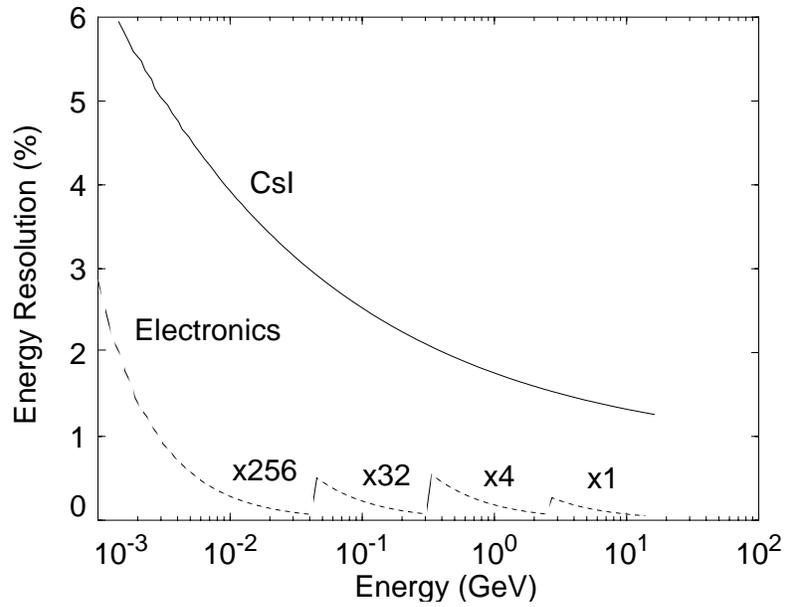


Figure 10-17. Digitization resolution vs. CsI Resolution.

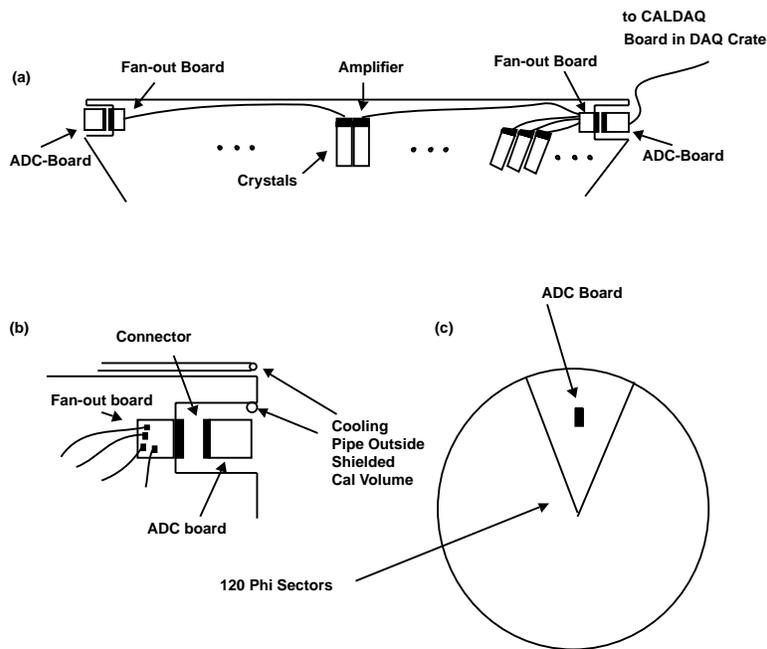


Figure 10-18. ADC board located at the end of the barrel.

**Figure 10-19.** *Block diagram of readout module functions.*

the 500 cables from the preamps to a bus connector into which the ADC board then plugs. Power lines are shared across this bus, while the signals are fed to the ADC cards. This allows ADC boards to be easily replaced during shutdowns. The inputs to the ADC board from the data acquisition system are: a sampling clock, a serial output link, a slow control link, and a serial output optical fiber.

The ADC boards are placed where they can be water cooled so that the heat generated on the boards does not get transmitted to the crystals. The expected heat generation on each ADC board is 5 W, leading to a total heat load of 1.5 kW.

### **10.6.5 Readout Module**

The readout modules for the calorimeter are located in the electronics house. They receive signals from the 280 ADC boards, provide trigger information to the Level 1 trigger, and send processed signal information to the higher level data acquisition system. This section discusses only the features unique to the calorimeter readout module. The features common

to the whole data acquisition system are discussed in Section 10.10. Figure 10-19 shows a block diagram of the readout module functions. The basic functions of these modules are:

- to receive serial data from the ADC boards;
- to de-serialize the data and place it in the VRAM;
- to contain the calibration look-up table for the trigger; and
- to apply calibration corrections and to extract amplitude and time information from the waveform samples.

For de-serialization of the input data, a clock is encoded on the optical fiber signal. This guarantees that the data flowing into the VRAM are synchronized with the digitization clock, which is synchronized with the overall system clock. Some information needs to be provided to verify the channel encoding. The simplest implementation would be a single bit to flag the first channel in a stream of 25. At the cost of extra bandwidth, 5 bits could be used to send individual channel numbers with the data. Data flow into the VRAM at 1.25 Gbits/s and can be stored in the VRAM for up to 6 ms before they must be moved to more permanent memory.

For trigger purposes, the readout modules provide information about the amount of energy deposited in the  $5 \times 5$  arrays of crystals which are seen by each board. This part of the module is discussed in Section 10.9.

A sophisticated algorithm to calculate the total charge and pulse height is needed to minimize the effect of the soft photon background on the resolution. Algorithms are being studied using models of the signal plus expected background. The floating-point conversion and calibration correction could be done continuously using a SRAM look-up table, but this is likely to be costly compared to doing this on demand using an onboard processor. Thus, in response to a trigger or a request containing better time information, the readout module processor executes the following steps to find the times and magnitudes of the showers: calibration constants are applied to the raw data, the charge and time of the pulse in each crystal are found, and if the magnitude is sufficiently large, this information is sent to the next level data acquisition system. Provision is also made to occasionally send the raw data to monitor the onboard algorithms.

The maximum data rate, allowing for online calculation of the time and pulse height in each crystal at a 2 kHz trigger rate, is estimated by allowing for 20 bits of charge information and 12 bits of time information. This requires 0.2 Mbytes/s to send data for all crystals on an ADC board, 3.2 Mbytes/s per readout crate, and 54 Mbytes/s for the entire calorimeter.

The design of the readout module is shared, as much as possible, with the other systems.

### 10.6.6 Electronic Calibration

For each channel, calibration corrections are applied on the readout module for both overall gain (Chapter 7) and the response of the analog and digitizing electronics. The latter information is provided by an electronic calibration subsystem.

In order to derive the calibration constants, the following tasks need to be performed:

- Injection of charge on the preamp: This is accomplished by sending a signal derived from an 18-bit digital-to-analog converter (DAC) on the ADC board through a capacitor onto the input of the preamp.
- Range forcing on the CARE chip: In order to understand the transition region from one range to another, the capability to force the CARE chip to remain in one gain range is essential. In this way, the operation of the automatic range selection can be verified.
- Independent control of the phase of the calibration signal with respect to the clock.
- Calibration of the CARE chip: it may be desirable to input a test signal to the CARE chip directly, and thus separate effects in the preamp from those in the CARE chip.

In addition to the calibration of the electronics, calibration with sources requires the readout of the calorimeter in an untriggered mode. This requires dedicated running using only the calorimeter and specialized software in the online processors to search for peaks in individual crystals and to create clusters. Since this calibration is performed without the beam, peak finding without a trigger should be feasible, as the design of the preamp ensures reasonably large signal/noise for source events. Both the electronic calibration runs and source calibration runs (Chapter 7) require special processing modes on the readout modules, during which the calorimeter must be in local mode.

### 10.6.7 Monitoring and Control

Most of the monitoring tasks are performed by the onboard microcontrollers on the ADC boards. These are interconnected with a slow low-cost link managed by an overall detector-control processor. The main variables which are monitored by the system are:

- Temperature: A tolerance of  $\pm 1^\circ\text{C}$  at the preamp boards and  $\pm 2^\circ\text{C}$  at the ADC board. A sample of preamp boards and all ADC boards are monitored.

- Bias Voltage Current: The bias voltage on and current from each photodiode are monitored. There are also provisions for measuring the total current.
- Supply Voltage: The low voltage supplies are likely to be shared across groups of ADC boards. These are monitored and safety interlocks applied to the power supplies.

Support software interprets the slow control data and provides diagnostics which allow problems to be isolated without opening the detector.

### 10.6.8 Research and Development

A preliminary version of the preamp IC has been constructed in a  $1.2\ \mu\text{m}$  BiCMOS technology. This chip uses a previous design [Bag94] modified to suit the needs of the *BABAR* calorimeter. The shaping time for this preliminary device can be modified using external components. This chip will be used in conjunction with various possible JFET candidates on the prototype calorimeter in order to evaluate the design, choice of shaping time, and performance of the preamp.

The R&D program for the ADC and readout modules centers around the following: the design and testing of the CARE chip; evaluation of candidate commercial ADCs; development of the electronic calibration system; development and testing of an appropriate VRAM control system; testing proposed trigger schemes; and benchmarking signal processing algorithms and evaluating the processor power required. A first version of the CARE chip has been built in CMOS and tested at SLAC, and a BiCMOS version of the chip capable of higher sampling rates is now in prototype fabrication. The remaining R&D issues will be the subject of parallel efforts to verify the concept and resolution obtained with the proposed methods. The new CARE chip and as many of the other components as possible will be included in the first tests of the calorimeter prototype with beam. Ongoing beam tests will study in detail the proposed production versions of all boards and modules and address system concerns which may arise only in the context of a realistic detector.

## 10.7 Muon System Electronics

---

### 10.7.1 Electronics Requirements

The functional requirements of the Instrumented Flux Return (IFR) electronics are to label hit strips and to associate a hit with a triggered event.

The resistive plate chamber (RPC) single counting rate is on the order of  $400 \text{ Hz/m}^2$  [Amb94], which is comparable to the cosmic ray rate at sea level. Assuming the probability of recording noise hits is on the order of  $10^{-4}$ , the number of accidental hits is below 20 per event. The physics rate also is quite low, with an event generating at most a few tens of hits. Thus the IFR is a very low occupancy detector.

The readout logic will utilize the Level 1 strobe which carries a fixed latency of  $10 \mu\text{s}$  with a maximum permitted jitter of  $\pm 0.5 \mu\text{s}$ . The average trigger rate is less than 2 kHz, and the maximum rate the readout must tolerate is 10 kHz. The minimum time between successive triggers is  $1.5 \mu\text{s}$ , and this is the time allotted to empty the front-end electronics. This small time in which to read the 42,000 channels is the main constraint on the design.

The baseline architecture presented here represents a working model which takes into account functional requirements and assumptions, and copes with constraints. The main goals of this design is to reduce the electronics cost by minimizing the number of modules and connections needed for the whole device, and to assure completely safe operation. This is accomplished by reducing the overall complexity, and by using well-defined existing technologies.

The first module (the Front-End Card, or FEC) will be located close to the detector in the iron gaps; the second module will be on the outside of the detector itself and accessible without opening the endcaps, while the last one (readout module) will be on the electronics house.

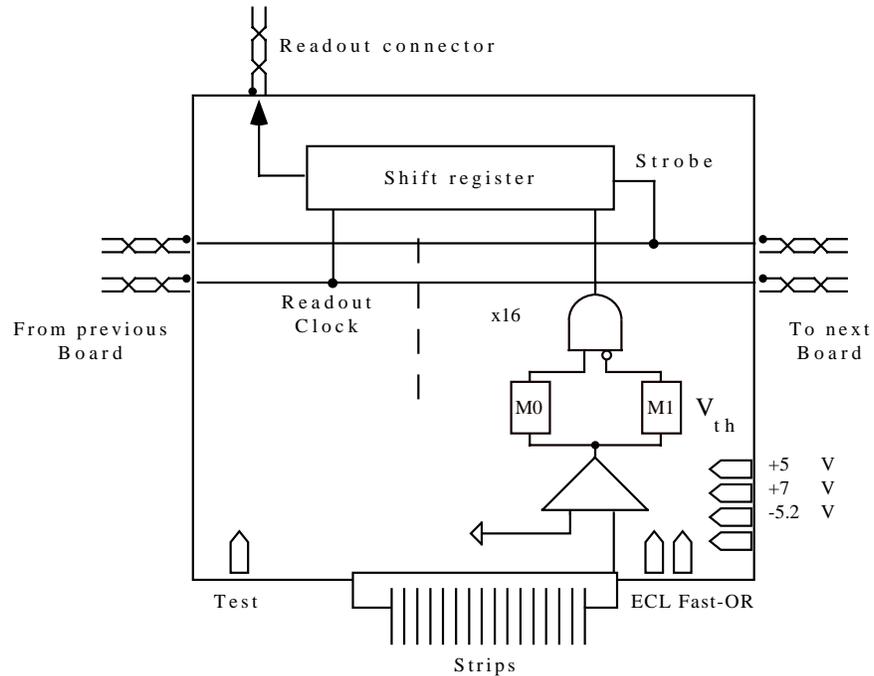
### 10.7.2 General Architecture

Because the RPC is practically a noiseless detector in which the rate is dominated by cosmic rays, immediate rejection of unwanted data is not as important as very rapid data collection. Buffering during the trigger latency is unnecessary because the probability of having two hits on the same strip during this time is negligible. Strip signals will be stored during the latency of the trigger decision by simply stretching them and using a fast readout card to prepare the front-end within the prescribed  $1.5 \mu\text{s}$ . The IFR electronics perform as a data-driven system.

Generally, a serial readout of front-end cards is suitable when a large number of strips must be handled and the acquisition speed is not a critical parameter. However, in this case, the time allowed to empty the front-end cards is relatively short, and we plan to use a mixed serial/parallel readout design.

Conceptually, the electronics system for the IFR is a sequence of blocks performing the following functions:

- record detected signals during Level 1 strobe latency;



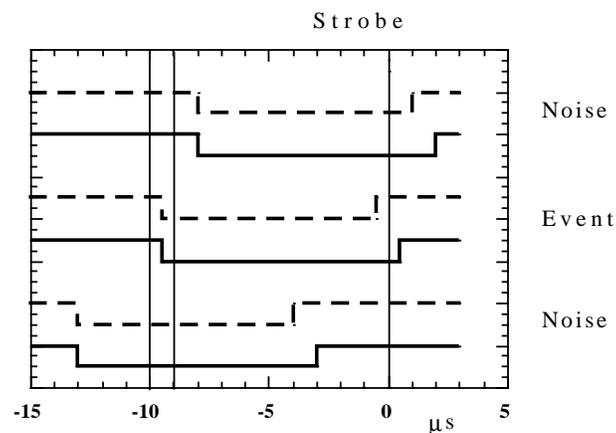
**Figure 10-20.** Schematic of the Front-End Card (FEC).

- unload the data from the front-end within the allowed time ( $1.5 \mu\text{s}$ ), skipping empty cards;
- encode the addresses of hit strips; and
- read hit data, assemble them into a full data stream, and transfer them to a storage medium.

### 10.7.3 Implementation

#### Front-End Card (FEC)

Because of the inaccessibility of the front-end cards, they are designed with a minimum number of components. A schematic is shown in Figure 10-20. Because RPCs produce pulses consistently near 200 mV in pulse height, the discriminator stage of the FEC is not critical. A single transistor will suffice to bring signal level to the stage of driving TTL chips [Alo94c].



**Figure 10-21.** *Double one-shot logic. Only strips fired in the  $1\ \mu\text{s}$  window will be acquired.*

Each FEC connects directly to the ends of 16 adjacent readout strips. The output is obtained from a logical AND of two one-shots (M0 and M1), which are  $9\ \mu\text{s}$  and  $10\ \mu\text{s}$  wide, respectively. Stretching the strip signals acts both as buffering for trigger latency and for noise reduction by identifying the strip fired within the  $1\ \mu\text{s}$  time window (Figure 10-21). The board also provides a Fast-OR signal of the 16 strips, which can be used for time measurements, monitoring, diagnostics, and time calibration.

They are connected in a 16-card chain. Each chain is read serially as a 16-bit-wide shift register, as shown in Figure 10-22. The FECs will be installed inside the IFR gaps, very close to the RPC strips.

### FIFO Buffering, Zero Suppression, and Digital Encoding (FIB)

After the Level 1 strobe is received, input data, consisting of one bit per strip, are latched into the FEC register. The chain is then read by the FIB, which is shown in Figure 10-23. This module buffers the data from the shift registers on 16 FECs and stores them in parallel in the data FIFO. The FIB module handles four different chains (1024 strips in total) and is located in a VME crate on the outside of the detector.

Each trigger has a time stamp which is read into the FIFO as a header to the data stream referring to that particular trigger. The readout chains are clocked at 12.5 MHz so that the FECs are ready for another trigger within  $1.5\ \mu\text{s}$ .

Independently, FIFO data are pushed into the zero suppressor and digital encoding logic, which encodes strip data into hit data. This is implemented with Xilinx LCAs and is a

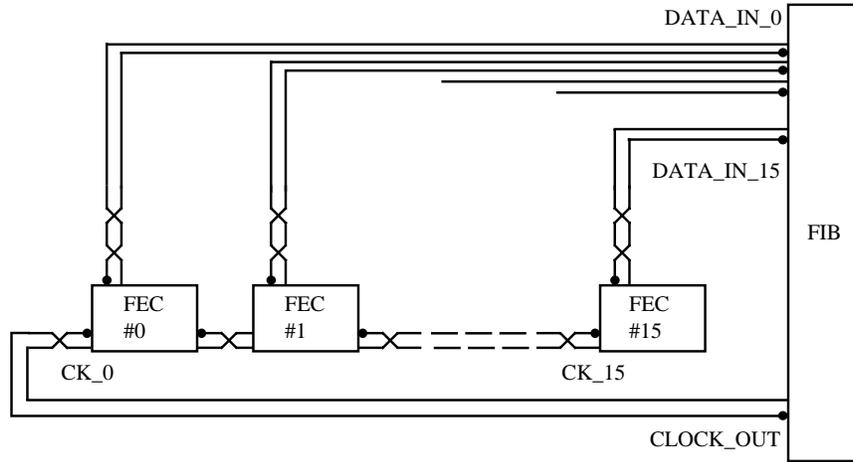


Figure 10-22. Schematic of the parallel readout mechanism for each chain.

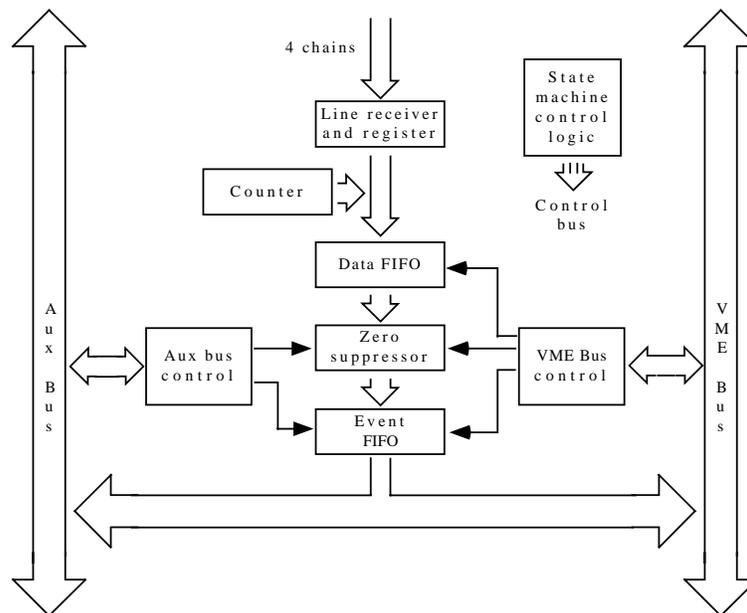


Figure 10-23. A simplified scheme of the FIFO board.

copy of that used in the L3 experiment at CERN [Alo94a, Alo94b, Alo94c, Alo94d, Alo94e]. Another Xilinx LCA handles VME protocol, while the state machine for the arbitration is implemented in a MACH gate array.

### Readout Integration

The FIB boards are arranged in seven VME crates, three for the barrel and two for each endcap. Each crate includes a data collector module (DCM) acting as interface between the encoding module and the readout module. The DCM asks for a specified trigger number from all FIBs plugged into the same crate. It uses an auxiliary bus connection and a token-in/token-out protocol. Data flow from an FIB until no more data are available on that FIB for that trigger number. The collected data are then ready to be transferred to the readout module. This architecture is deadtimeless even under the worst background conditions.

#### 10.7.4 Time Measurements

Time measurements can be achieved by using TDC channels with a granularity as coarse as a single layer. The measurements will be performed via common stop, VME-based, TDC modules located in the crates on the outside of the detector. The stop signal is supplied with a (synchronous) Level 1 strobe. The TDCs have a range covering a  $1.5 \mu\text{s}$  time window to ensure that the entire time jitter for the trigger is contained. Provisions have to be made for the  $10 \mu\text{s}$  strobe latency. When the Level 1 strobe is asserted, the TDCs are stopped and read into a FIFO buffer, associated to the right Level 1 strobe reference number, and sent to the data acquisition system. The required TDC resolution is  $0.5 \text{ ns}$ .

#### 10.7.5 Time Calibration

Since strip signals coming from different FECs have different path lengths, their arrival time will be spread out. A time-zero measurement is thus needed in order to calibrate out different cable lengths and electronics time response.

The time calibration is performed with an external pulser system. A synchronized test pulse is injected directly in the input stage of all the FEC channels. The difference in time between the various TDC outputs provides the factor needed to calibrate the different signal delays.

### 10.7.6 Monitoring

The complete hardware paths can be monitored by sampling unprocessed data (spying) during data taking. Furthermore, fast-OR signals directly collected by scalars provide front-end card rates for monitoring.

### 10.7.7 Diagnostics

The whole electronics paths will be checked in order to identify possible failures. A test input provided on the front-end board is used to inject a test pulse on all 16 input channels. This is the same pulse signal that is used in time calibration.

## 10.8 Trigger Requirements and Background Rates

---

### 10.8.1 Introduction

PEP-II may give rise to rather severe backgrounds, producing up to several kilohertz of interactions with two or more charged tracks reaching the calorimeter. This rate is to be contrasted with the desired logging rate of less than 100 Hz. The trigger and data acquisition subsystems are designed to record data at no more than the latter rate under a variety of background conditions and to accommodate luminosity upgrades up to  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ .

The purpose of the trigger is to reject backgrounds while selecting a wide variety of physics processes. This filtering function is performed through an increasingly sophisticated sequence of algorithms applying mostly topological cuts on increasingly refined data. Processes selected with high efficiency include final states of  $e^+e^-$  interactions ( $\Upsilon(4S)$ ,  $\Upsilon(5S)$ ,  $c\bar{c}$ , and  $\tau^+\tau^-$ ) as well as final states of  $\gamma\gamma$  and  $\gamma\gamma^*$  interactions. In addition, final states such as  $e^+e^-$ ,  $e^+e^-\gamma$ , and  $\mu^+\mu^-$  are selected for use with luminosity measurements, monitoring, and calibration. Finally, for diagnostic purposes, cosmic rays, beam-induced backgrounds, and random time slices are selected at a low rate.

In addition to providing high efficiency for a large number of interesting physics processes and rejecting most background events, the trigger is designed to allow precise measurements of its efficiency. The most powerful tool for this is the use of independent selection criteria. Orthogonal selection methods are used as much as possible for those filters operating on trigger primitives, *i.e.*, on quickly derived reduced representations of the data.

Physics Data Samples	Requirements
$B\bar{B}$ from the $\Upsilon(4S)$ : $CP$ Channels	High Efficiency
$B\bar{B}$ from the $\Upsilon(4S)$ : Non- $CP$ Channels	Precise Efficiency, High Efficiency
$B$ Physics at the $\Upsilon(5S)$	High Efficiency
Charm Physics	Precise Efficiency, High Efficiency
Continuum $q\bar{q}$	Precise Efficiency (Background Subtraction)
$\tau$ Physics: Rare Decays & Asymmetries	High Efficiency
$\tau$ Physics: Branching Ratios	Precise Efficiency
$\gamma\gamma$	Precise Efficiency

Other Data Samples	Purpose, Comments
Bhabhas: Two Detected Prongs	Luminosity, Calibration, High Efficiency
Bhabhas: One Detected Prong	Calibration, Beam Monitor
$\mu^+\mu^-$	Luminosity, Calibration
Cosmic Rays	Calibration
Beam Backgrounds	Beam Monitor, Trigger Studies
Random Events	Trigger Studies

**Table 10-6.** *Principal trigger requirements for each type of data sample.*

This section begins with a detailed discussion of the implications of  $BABAR$ 's physics goals for the trigger. It continues with a description of trigger algorithms, and concludes with the results of simulations of the proposed trigger algorithms, including both the efficiencies for some benchmark processes and the expected trigger rates.

## 10.8.2 Requirements

The trigger needs to be efficient, selective, flexible, measurable, and accurate. More details are given in the *BABAR Electronics Requirements* [Req94]. The relative importance of these attributes varies for different data sets. For asymmetry measurements and searches, it is more important to have a high efficiency than a well-known efficiency. For branching ratio and cross-section measurements, precisely measured efficiencies are crucial. The principal trigger requirements for each data set are listed in Table 10-6. In addition to the key physics events, a number of other events are recorded for uses such as luminosity determination, detector calibration, beam monitoring, and trigger performance studies. These impose additional requirements on the trigger and are listed at the bottom of the table.

Because of their relatively large multiplicities and visible energies, it is straightforward to achieve high trigger efficiencies for  $\Upsilon$  and  $c\bar{c}$  events. The challenge for the trigger designer lies in detecting low multiplicity  $\tau^+\tau^-$  and  $\gamma\gamma$  events while keeping the rate of accepted background events within the data acquisition limit, so as not to incur downtime. To ensure high efficiencies in offline analysis and large background rejection power, trigger algorithms need to have sharp efficiency turn-on curves. For example, the efficiency as a function of cluster energy should approximate a unit step function as closely as possible to allow a highly selective trigger.

Because PEP-II machine conditions will vary over the course of the experiment, the trigger needs to be flexible. Under temporarily severe background conditions, saturation of the bandwidth should result in downtime in such a way that a diminished number of whole events are collected rather than a larger set of partial events. When backgrounds are severe but data are useful for physics, prescaling factors can be applied to selection streams to ensure that the available bandwidth is filled with interesting events. This can be done on a relatively short timescale. It will be straightforward to change the global logic combinations of the selection streams in each of the trigger levels, and it will also be possible to vary selection parameters. Actual change of trigger selection algorithms is foreseen to occur rarely.

There are several ways to facilitate precise offline measurements of the trigger efficiency. First, the geometry of the detector and the readout electronics for each system contributing to the trigger are designed with trigger considerations in mind. Second, the trigger and data acquisition system is designed to:

- use orthogonal selection criteria;
- record all trigger primitives and decisions for accepted events;
- accept prescaled numbers of events with looser criteria;
- accept a prescaled number of rejected events;
- have a sharp efficiency turn-on as a function of the selection variables; and
- be simple to model with simulations.

Finally, the online system will be designed so that the recorded trigger primitives and decisions are readily available for offline analysis.

Use of the trigger results for precision physics measurements depends crucially on the trigger system working as a whole. Like the other systems that deliver data to the data acquisition, the trigger system should be designed with sufficient monitoring, calibration, and diagnostic testing capabilities so that trigger decisions and trigger data are known to be correct.

### 10.8.3 Backgrounds and Trigger Rates

The trigger architecture and selection algorithms designed to meet the requirements are based on simulation studies of the physics, backgrounds, and the *BABAR* detector.

The total trigger output rate is the sum of the rates for the physics and diagnostic processes listed in Table 10-6 and the rates for the background processes described in Chapter 12.

The main background processes that produce charged tracks traversing the drift chamber and calorimeter are:

- physics backgrounds from asymmetric Bhabhas with only one detected charged track and two-photon events below  $\gamma\gamma$  energies of interest;
- beam-induced backgrounds from hadronic interactions of lost beam particles, electromagnetic interactions of lost beam particles, and beam-gas collisions at the interaction point; and
- cosmic rays.

The main sources of occupancy in the vertex detector, drift chamber, and calorimeter are:

- particles from small-angle radiative Bhabhas, and
- beam-induced background from lost beam particles.

The rate of charged particles and the occupancy of physics backgrounds scale with the luminosity, whereas those of beam-induced backgrounds scale with the beam current and the pressure of the residual gas inside the PEP-II beam pipe.

Physics backgrounds are evaluated at the design luminosity, and the trigger and data acquisition system must be upgradable for higher luminosities. Because of their variable nature and because of uncertainties in the predictions, beam-induced backgrounds pose more stringent design requirements. The trigger and data acquisition system is required not to produce substantial impact on the physics program as long as data are useful. Based on simulation of the occupancies in the vertex detector as functions of lost-particle rate, pattern recognition becomes seriously compromised at ten times the nominal background rate. The simulation is described in Chapter 12.

## 10.8.4 Definitions of Trigger Filter Concepts

### Event Selection

Event signatures are used to separate signal from background. Combinations of the following global event properties are used in the trigger filters:

- charged track multiplicity;
- calorimeter cluster multiplicity;
- event topology; and
- primary vertex position.

These selection criteria have associated thresholds for the following parameters: charged-track transverse momentum ( $p_t$ ), energy of calorimeter clusters ( $E_{\text{clus}}$ ), solid angle separation ( $\phi$  and  $\theta$ ), track match quality (for electrons and muons) and primary vertex impact parameter ( $\rho_1$  and  $z_1$ , respectively). The trigger definition can contain selection criteria that differ only by the values of thresholds, such as the  $p_t$  or  $E_{\text{clus}}$  threshold values. A small fraction of random beam crossings and events with failed triggers are selected for diagnostic purposes.

### Selection Filters

Simulation studies both of processes of interest in the areas of B,  $\tau$ , and  $\gamma\gamma$  physics [Kra94b, Kra94a, Lev94a, Lev94b, Lev94c, Lev94d, Sny94], and of background events in the *BABAR* detector, have been used to guide the choice of trigger algorithms.

Table 10-7 lists “trigger objects” which are potentially definable at Level 1 and which have been determined to be useful in constructing global trigger definitions. Vertex detector information is available at higher levels. For a given trigger level, the global selection filter is a logical OR of a number of specific selection filters, each of which is the result of a boolean operation involving any of the ingredients available at that level. The types of selection filters may be arranged in the following hierarchy:

1. Multiplicity of drift chamber charged particles with implicit  $p_t$  and  $\cos\theta$  cuts and explicit  $\phi$ -separation cuts;
2. Multiplicity of calorimeter clusters with explicit  $E_{\text{clus}}$  and implicit  $\cos\theta$  cuts, and explicit  $\theta$ - and  $\phi$ -separation cuts;

Tracks in the Drift Chamber	
A tracks	Tracks that reach the outer layers of the drift chamber. For $B = 1.5$ T, this implicitly leads to approximately $p_t > 0.210$ GeV/ $c$ . These tracks are contained in the solid angle of approximately $-0.81 < \cos \theta < 0.90$ .
A' tracks	A tracks that satisfy an explicit cut on the reconstructed $p_t$ .
B tracks	Tracks that reach the middle layers of the drift chamber. This implicitly leads to approximately $p_t > 0.130$ GeV/ $c$ . The solid angle acceptance for these tracks that can exit the middle of the drift chamber end cone is substantially larger than for A tracks, $-0.92 < \cos \theta < 0.96$ .
Clusters in the Calorimeter	
M clusters	Clusters that have reconstructed energy deposits consistent with minimum ionizing particles (MIP), or larger, in the calorimeter barrel or endcap. The required value is near 150 MeV. A minimum ionizing $\mu^-$ deposits $181_{-8}^{+20}$ MeV, averaged over the solid angle of the barrel calorimeter, where the limits indicate the variation in energy deposit due to trajectories and varying crystal lengths [Kra94a]. For such a muon to reach the calorimeter barrel, it is implicitly required to have approximately $p_t > 0.250$ GeV/ $c$ . The solid angle acceptance is approximately $-0.78 < \cos \theta < 0.96$ . The cluster energy threshold is adjustable.
G clusters	Clusters with energies above those of M clusters; <i>i.e.</i> , $E_{\text{clus}} > 0.200$ GeV. The cluster energy threshold is adjustable.
E clusters	Clusters with energies large enough to be produced by $e^+e^-\gamma$ or $\gamma\gamma^*$ , <i>i.e.</i> , approximately $E_{\text{clus}} > 1.0$ GeV. The cluster energy threshold is adjustable.
Matches Between the Drift Chamber and Calorimeter	
A-M matches	A tracks correlated in $\phi$ with barrel M clusters.
A'-M matches	A' tracks correlated in $\phi$ with barrel M clusters.
B-M matches	B tracks correlated in $\phi$ with barrel or endcap M clusters.

**Table 10-7.** *Trigger objects.*

3. Special triggers for calibration, monitoring, and diagnostics;
4. Multiplicity of charged particles correlated between the drift chamber and calorimeter  $M$  clusters, with explicit matching in  $\phi$ ;
5. Multiplicity of drift chamber charged particles with implicit  $\cos \theta$  cuts, and explicit  $p_t$  and  $\phi$ -separation cuts; and
6. Multiplicity of charged particles correlated between the vertex detector and drift chamber with explicit matching in  $\phi$  or explicit cuts on  $\rho$  and  $z$ .

The list continues with more refined selection cuts applied to the full data set available in Level 3.

At each hierarchical step, there may be a number of selection filters (trigger lines), differing in whether certain cuts are applied or in the values of parameters used to make cuts. Let the isolated symbol for an object represent the number of such objects, *e.g.*,  $A$  is the number of  $A$  tracks. Assume that counting and geometry cuts are such that no physical object will fill a double role in the trigger. Then, using the notation of Table 10-7, category (2) might include:  $M \geq 2$  with minimal separation cuts; one or more versions of  $G \geq 1 \cdot M \geq 1$  with minimal separation cuts; or  $G \geq 1 \cdot M \geq 1$  with appreciable  $\phi$  separation. Each of the trigger lines will be subject to individual prescaling, which may have to be used to reduce rates for the more open definitions, while collecting all events satisfying more stringent definitions. The global selection OR at each level is formed after any such prescaling. Some specific filters will be evaluated for their effects on signal and background events in the following subsection.

### 10.8.5 Performance of Some Simple Triggers

#### Benchmark Physics Processes

In order to evaluate the efficiency of selection algorithms, a set of physics processes has been identified to serve as trigger benchmarks. As explained in Chapter 3, the *BABAR* physics program focuses on studies using  $\Upsilon(4S)$ ,  $\Upsilon(5S)$ ,  $q\bar{q}$ ,  $\tau^+\tau^-$ ,  $\gamma\gamma$ , and  $\gamma\gamma^*$  events. While the primary mission of the *BABAR* experiment is the study of  $CP$  violation in  $B$  decays, the trigger is designed to accept a large number of processes that may be important. The trigger benchmarks represent processes that are difficult to distinguish from background. Other similar processes have, or are expected to have, higher efficiencies. Here we use Monte Carlo information to study the effectiveness of some simple triggers on these benchmarks.

For  $B^0\bar{B}^0$   $CP$  studies, the most difficult tag to trigger on is  $\bar{B}^0 \rightarrow \mu X$  because of its low charged-track multiplicity and low cluster energy compared to the electron or kaon tags.

The low-multiplicity decay  $B^0 \rightarrow \pi^+\pi^-$  was chosen as a  $CP$  study decay mode to test the efficiency of our simple particle counting trigger algorithms. The high-multiplicity decay  $B^0 \rightarrow D^+D^-$  was chosen to test algorithms that require fast tracks.

For  $\tau^+\tau^-$  studies, events in which both taus decay to one charged track are the most difficult to select. To measure the muon branching ratio, events with  $\tau^- \rightarrow \mu^- \bar{\nu} \nu$  accompanied by  $\tau^+ \rightarrow e^+ \nu \bar{\nu}$  are used. This is a worst-case decay for the purpose of the benchmark table, as the measurement is not expected to be statistics limited. The other  $\tau^+\tau^-$  benchmark is an untagged event for which each tau decays to one charged pion, used to measure spin correlations for  $V-A$  tests.

The  $\gamma\gamma$  benchmarks are straightforward: two-pion events that are either both charged or both neutral. For two-charged-hadron final states, one can normalize the hadronic cross section to the better known muon-pair cross section. The  $\gamma\gamma^*$  benchmarks with two kaons and one pion in the final state, produced via an unspecified  $1^{++}$  resonance, represent searches for exotic mesons.

### Benchmark Efficiencies and Background Rates

An inclusive global trigger definition that efficiently selects each of the processes in Table 10-6 while rejecting some backgrounds requires a minimum of two charged tracks in the drift chamber, two M clusters, or one E cluster in the backward direction [Bol91]. The orthogonality of the two first criteria makes it straightforward to measure the efficiency of the filter.

The following discussion focuses on defining a flexible open trigger with high efficiency. Based on the first two criteria listed below, it is able to handle a large range of potential beam-induced backgrounds.

There are three versions of the two-track trigger to consider:  $A \geq 2$ ,  $A \geq 1 \cdot B \geq 1$ , and  $B \geq 2$ . The preferred definition is  $A \geq 1 \cdot B \geq 1$  (the so-called one and one-half track trigger) because it has larger solid angle acceptance than  $A \geq 2$  and accepts fewer looping tracks than  $B \geq 2$ . Simulations show that  $A \geq 2$  is only  $\sim 80\%$  as efficient as  $A \geq 1 \cdot B \geq 1$  for inclusive tau decays [Bol91, Kra94b]. One soft charged particle can loop back, resulting in two detected B tracks. This problem is less severe with the  $A \geq 1 \cdot B \geq 1$  definition.

Efficiencies and output rates for the two-particle trigger were simulated for two global-trigger input lines,  $D_2$  and  $C_2$ . The default two-track trigger  $D_2$  uses the  $A \geq 1 \cdot B \geq 1$  definition and is simulated. The default two-cluster trigger  $C_2$  requires  $M \geq 2$ , both with  $E_{\text{clus}} > 0.150$  GeV, where the chosen value for the cluster energy threshold is a compromise between efficiency and background rate. To suppress pile-up of soft photons due to electromagnetic interactions

Process	D <sub>2</sub>	C <sub>2</sub>	C <sub>2</sub> + D <sub>2</sub>
$B^0 \rightarrow \pi^+\pi^-, \bar{B}^0 \rightarrow \mu X$	1.00	1.00	1.00
$B^0 \rightarrow D^+D^-, \bar{B}^0 \rightarrow X$	1.00	1.00	1.00
$\tau^+ \rightarrow e^+\nu\bar{\nu}, \tau^- \rightarrow \mu^-\bar{\nu}\nu$	0.72	0.73	0.79
$\tau^+ \rightarrow \pi^+\bar{\nu}, \tau^- \rightarrow \pi^-\nu$	0.85	0.74	0.89
$\gamma\gamma \rightarrow \pi^+\pi^-, W_{\gamma\gamma} > 2 \text{ GeV}/c^2$	0.63	0.63	0.63
$\gamma\gamma \rightarrow \pi^0\pi^0, W_{\gamma\gamma} > 2 \text{ GeV}/c^2$		0.86	0.86
$\gamma\gamma^* \rightarrow X(1420) \rightarrow KK\pi$	0.48	0.62	0.62
$\gamma\gamma^* \rightarrow \eta(550) \rightarrow \pi^+\pi^-\pi^0$	0.07	0.21	0.21

**Table 10-8.** Open trigger efficiencies for benchmark physics processes. The efficiency denominator includes all events generated over the full  $4\pi$  solid angle. The simulation used 1000 events.  $W_{\gamma\gamma}$  is the mass of the  $\gamma\gamma$  system.

of lost beam particles, each calorimeter channel is required to have an energy deposit above a threshold of 10 MeV. An open global trigger is defined as  $C_2 + D_2$ .

Calculated efficiencies for benchmark physics processes using the defined trigger lines are listed in Table 10-8. These efficiencies are found to be very high for each of the benchmark processes. Estimated open-trigger output rates due to the processes listed in Table 10-6 and in Section 10.8.3 are shown in Table 10-9. The simulation is described in Section 10.9.5.

Using the open trigger, the maximum event-building rate goal is not satisfied at  $\times 10$ . Thus, the open trigger would need to be prescaled and a tighter criterion would need be applied when the total output rate exceeds 2.0 kHz. The best solution to this problem—*i.e.*, with the least impact on two-prong physics processes—is to require an  $A'$  track instead of an  $A$  track and a  $G$  cluster instead of one of the  $M$  clusters. These are the  $D'_2$  and  $C'_2$  trigger lines. To suppress lost-particle electromagnetic backgrounds, the  $C'_2$  line further required the  $G$  cluster to be back-to-back with an  $M$  cluster. The efficiency and rate for the two trigger lines are given in Tables 10-10 and 10-11, respectively.

## 10.9 Level 1 Trigger

### 10.9.1 Overview

The first level trigger architecture provides the Level 1 strobe for the front-end buffers and reduces the event output rate for feature extraction and event building. The trigger achieves

Process	D <sub>2</sub>	C <sub>2</sub>	C <sub>2</sub> + D <sub>2</sub>
Physics:			
$\Upsilon(4S)$	12 Hz	12 Hz	12 Hz
$q\bar{q}$	8 Hz	8 Hz	8 Hz
$\mu^+\mu^-$	2 Hz	2 Hz	2 Hz
$\tau^+\tau^-$	2 Hz	2 Hz	2 Hz
Bhabha	100 Hz	100 Hz	100 Hz
$\gamma\gamma$ signal	3 Hz	3 Hz	3 Hz
$\gamma\gamma$ background	50 Hz	50 Hz	50 Hz
Backgrounds:			
Hadronic nominal	300 Hz	140 Hz	380 Hz
Electromag. nominal	< 60 Hz	230 Hz	230 Hz
Beam-gas 1 nTorr	< 4 Hz	< 2 Hz	< 6 Hz
Hadronic $\times 10$	3000 Hz	1400 Hz	3800 Hz
Electromag. $\times 10$	450 Hz	3800 Hz	4000 Hz
Beam-gas 10 nTorr	< 40 Hz	< 20 Hz	< 60 Hz
Cosmic rays	80 Hz	410 Hz	440 Hz
Total nominal	620 Hz	970 Hz	1200 Hz
Total $\times 10$	3700 Hz	5800 Hz	8500 Hz

**Table 10-9.** *Open trigger rates. The background rates are subject to large uncertainties. The hadronic and electromagnetic backgrounds are due to lost beam particles. The rates of the first group of processes are dependent on the luminosity. The rates of the second and third groups are dependent on the product of the beam current and the residual pressure in the beam pipe. The cosmic rate is fixed.*

these goals through the use of a fast hardware trigger operating on primitives, provided via point-to-point links by the drift chamber and calorimeter. The drift chamber and calorimeter triggers feed information to the global trigger, which issues the Level 1 strobe.

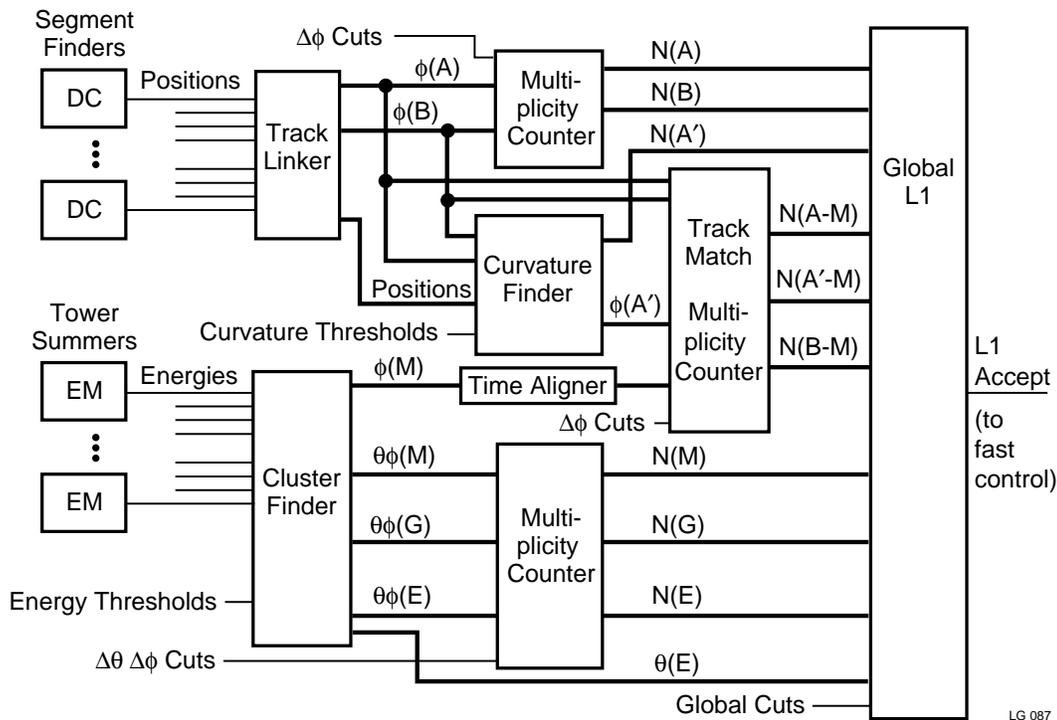
The structure of the Level 1 trigger is shown in Figure 10-24. Data from the drift chamber and calorimeter are converted into trigger primitives by local logic on the readout modules. There are three steps of trigger logic between the formation of primitives and the global Level 1 decision that produces the Level 1 accept signal: particle finders that identify drift chamber tracks and calorimeter clusters, particle counters that identify the number of distinct particles, and a track matcher that associates drift chamber tracks with calorimeter clusters.

Process	D' <sub>2</sub>	C' <sub>2</sub>	C' <sub>2</sub> + D' <sub>2</sub>
$B^0 \rightarrow \pi^+\pi^-, \bar{B}^0 \rightarrow \mu X$	0.99	0.75	0.99
$B^0 \rightarrow D^+D^-, \bar{B}^0 \rightarrow X$	0.97	0.80	0.99
$\tau^+ \rightarrow e^+\nu\bar{\nu}, \tau^- \rightarrow \mu^-\bar{\nu}\nu$	0.65	0.43	0.68
$\tau^+ \rightarrow \pi^+\bar{\nu}, \tau^- \rightarrow \pi^-\nu$	0.78	0.36	0.80
$\gamma\gamma \rightarrow \pi^+\pi^-, W_{\gamma\gamma} > 2 \text{ GeV}/c^2$	0.62	0.51	0.63
$\gamma\gamma \rightarrow \pi^0\pi^0, W_{\gamma\gamma} > 2 \text{ GeV}/c^2$		0.76	0.76
$\gamma\gamma^* \rightarrow X(1420) \rightarrow KK\pi$	0.18	0.29	0.31
$\gamma\gamma^* \rightarrow \eta(550) \rightarrow \pi^+\pi^-\pi^0$	0.05	0.06	0.06

**Table 10-10.** Safe trigger efficiencies for benchmark physics processes. The efficiency denominator includes all events generated over the full  $4\pi$  solid angle. The thresholds were  $p_t^A > 0.600 \text{ GeV}/c$  for D'<sub>2</sub> and  $E_{\text{clus}}^1 > 0.600 \text{ GeV}$  for C'<sub>2</sub>. Clusters were required to be separated by at least  $150^\circ$  in  $\phi$ .  $W_{\gamma\gamma}$  is the mass of the  $\gamma\gamma$  system.

Process	D' <sub>2</sub>	C' <sub>2</sub>	C' <sub>2</sub> + D' <sub>2</sub>
Physics Signal	47 Hz	124 Hz	127 Hz
$\gamma\gamma$ Background	50 Hz	20 Hz	50 Hz
Hadronic nominal	90 Hz	10 Hz	100 Hz
Electromag. nominal	< 60 Hz	< 60 Hz	< 60 Hz
Beam-gas 1 nTorr	0 Hz	< 1 Hz	< 1 Hz
Hadronic $\times 10$	900 Hz	120 Hz	980 Hz
Electromag. $\times 10$	< 220 Hz	220 Hz	220 Hz
Beam-gas 10 nTorr	< 2 Hz	< 10 Hz	< 12 Hz
Cosmic Rays	50 Hz	2 Hz	50 Hz
Total nominal	300 Hz	220 Hz	390 Hz
Total $\times 10$	1300 Hz	500 Hz	1400 Hz

**Table 10-11.** Safe trigger rates. The background rates are subject to large uncertainties. The D'<sub>2</sub> line required  $p_t^A > 0.600 \text{ GeV}/c$  and the C'<sub>2</sub> line required  $E_{\text{clus}}^1 > 0.600 \text{ GeV}$ . Clusters were required to be separated by at least  $150^\circ$  in  $\phi$ . The total output rate is the sum of the safe trigger rate, the prescaled open trigger rate, 10 Hz of prescaled asymmetric Bhabha triggers, and 5% of special triggers.

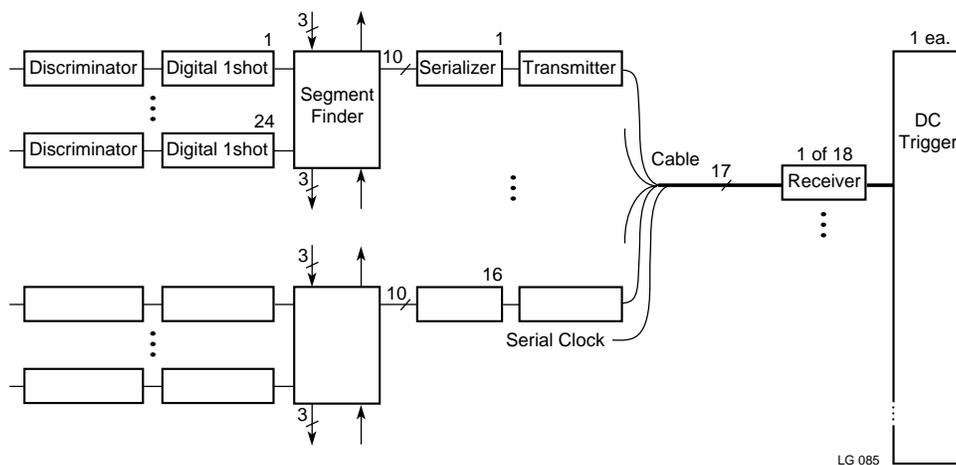


**Figure 10-24.** Level 1 trigger structure, data flow, and decision logic.

The different timing structure of the drift chamber and calorimeter primitives and the delay associated with the inputs to the track matcher require one stage of time alignment.

A number of variable selection cuts are available to these trigger logic steps. They could be tuned during initial running but should be stable during subsequent running so as not to complicate offline analysis. The global decision parameters are also flexible in the same way, to accommodate changing data-taking conditions. During more rapidly changing beam conditions, the Level 1 accept rate is controlled by prescaling certain components of the selected events. Information from the Level 1 global trigger is used for detailed accounting of prescaling factors and deadtimes.

In addition, there are important interfaces of the Level 1 trigger to the data acquisition, fast control, and online systems. These interfaces are discussed after a detailed description of the trigger decision implementation.



**Figure 10-25.** Level 1 drift chamber trigger electronics for primitives.

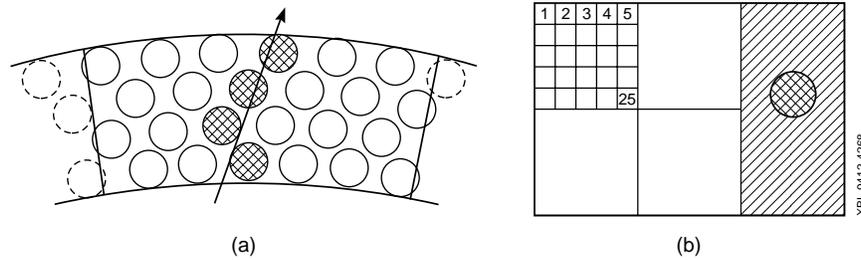
## 10.9.2 Drift Chamber Trigger

The drift chamber trigger performs three functions: local primitive formation, transmission and gathering of primitives, and global track finding. The structure of the drift chamber trigger is shown in Figure 10-24 with details relating to the primitives shown in Figure 10-25. The trigger algorithms consist of two track finders. The first identifies local track segments using information from readout cards that receive signals from 20 neighboring wire-chamber cells in the same superlayer. The second links these segments together to form track candidates. This section describes the algorithms and a preliminary electronics design for the baseline axial-stereo drift chamber [Kra94b].

### Primitive Finders: Local Segment Finding

The segment finder algorithm is closely tied to the electronics implementation [Bol91, SLA93]. As is shown in Figure 10-7, trigger primitives are formed on the readout modules, which receive analog signals from 24 channels comprising a  $6 \times 4$  supercell. To accommodate the maximum drift time, these signals are discriminated and converted to 600 ns-long signals by digital one-shots. Such signals from the 24 channels on the card are input, together with four signals from the neighboring readout modules, to a field-programmable gate array that finds track segments at a rate of one decision per 250 ns. Note that drift time information is not used.

Possible track-segment patterns consist of a hit in each of the four layers, with each hit being an immediate neighbor of the hit in the layer inside. An example of such a pattern is shown in



**Figure 10-26.** Level 1 trigger primitives for the (a) drift chamber and (b) calorimeter. A drift chamber track produces patterns like the one shown and uses majority logic to allow three of four layers hit. For the calorimeter, energies from 25 blocks are summed locally to form towers. The cluster sums shown for the example energy deposit are formed from the primitives in a central location.

Figure 10-26(a). The elimination of sideways tracks provides background rejection without significant loss in physics efficiency. Majority logic is implemented to guard the algorithm against inefficiencies. Patterns containing three of the four hits required for a pattern also produce valid segments. The segment finder generates 12 output bits, representing  $\phi$  position in the  $6 \times 4$  supercell. The local angle of the segment is not reported.

### Transmission Links: Segment Gathering

Transmission of drift chamber primitives is via a dedicated point-to-point serial link, as outlined in one implementation of the calorimeter trigger [Wun94]. Using a serial link reduces the number of wires needed and makes it possible to gather the signals in one location.

### Particle Finders: Global Track Linking

A simple and efficient track linking algorithm is the Binary Link Track finder (BLT), developed for the CLEO II trigger [CLE91] by K. Kinoshita [Kin89]. The method is to start at the second superlayer and move radially outward. At each superlayer, a segment is scored only if it is hit and any of its three immediate neighbors in the previous superlayer have been hit. Here, hit refers to a segment-finder hit. At the last superlayer, denoted a counting layer, any surviving scored segments represent tracks that traveled continuously from the first superlayer to the counting layer. The requirement of three neighbors imposes loose curvature cuts on tracks, while being forgiving to superlayer stereo stagger effects.

For *BABAR*, majority logic is incorporated by allowing combinations of four out of five superlayers to contribute segments. The B-track linker uses all five permutations of four

or five of the the inner five superlayers, while the A-track linker uses all 36 permutations of four or five of the inner superlayers together with four or five of the outer superlayers.

The BLT algorithm was chosen because it prevents double counting and because it avoids searching for tracks in “roads.” In the current drift chamber geometry, the latter method would require a sector with  $\phi = 74^\circ$  (one-fifth of  $2\pi$ ) in order to just barely contain a track with  $p_t = 180 \text{ MeV}/c$ . To contain any track, one would need to search in twice that region and then account for overlaps. Furthermore, unlike most road-based algorithms, the BLT is insensitive to the sign of the electric charge of the track. Because it can be implemented in combinational logic hardware, the BLT is also fast and inexpensive.

### Particle Counters: Global Track Multiplicity Finding

The BLT outputs are used to count distinct tracks using simple multiplicity logic. The output of the BLT is two digital strings with  $\phi$  positions of segments in the fifth and tenth superlayers, representing tracks passing through these points. The logic checks to see if there are at least two B tracks separated by an adjustable number (taken to be two) of supercells in the fifth superlayer (for a  $\phi$  isolation of at least  $22^\circ$ ). For events with two or more A tracks,  $\phi$  isolation is checked in superlayer ten. The multiplicities of distinct A tracks and B tracks are reported as 3-bit numbers to the global trigger.

### Transverse Momentum Determination: Curvature Finding

Using  $\phi$  positions of the axial layers of the drift chamber, it is straightforward to design a Level 1 drift chamber trigger that reports multiplicities above crude  $p_t$  thresholds [CLE91]. The thresholds can be used to suppress the most severe trigger backgrounds from lost-beam-particle interactions, as shown in Section 10.8.

### Drift Chamber Trigger for the All-Stereo Option

An all-stereo chamber with fully interleaved stereo layers is also being considered. This would require elaboration of the tracking trigger. Logical superlayers would have to be formed from alternating layers. A possible choice is six three-layer superlayers in each stereo view. The definition of logical supercells then leads to more than 300 unique segment-finding algorithms with increased interconnectivity of the readout modules.

The apparent curvature of tracks in one view varies with the track’s dip angle. This would be largely corrected by summing the curvatures of tracks found in each view. For events with small multiplicities, all pairs would be summed in parallel. Events with large multiplicities would be accepted without the  $p_t$  requirement.

### 10.9.3 Calorimeter Trigger

A calorimeter trigger based on towers (groups of contiguous crystals) will be more efficient than a trigger which uses only the energy deposits in individual crystals [Kra94a] because the center crystal of a shower will contain at most about 80% of the shower energy [SLA93]. Besides the improvement in efficiency, the better estimate of the total energy deposited in a tower compared with a crystal means that any energy threshold cut applied in the trigger is more precise, which is not only better for background rejection, but is also important for understanding and modeling the trigger response accurately.

The algorithm for the calorimeter Level 1 trigger [Kra94a] is summarized in this section. It is based on fixed towers of mostly  $5 \times 5$  crystals and performs three functions: local tower energy summing to form primitives, transmission of the primitives and cluster reconstruction, and central particle finding and counting. The formation of primitives occurs on the calorimeter readout modules (Section 10.6). The primitives are then transmitted to a central trigger crate where cluster reconstruction is carried out. Finally, the results of the clustering are compared and physical particles are found.

#### Primitive Finders: Local Tower Summing

Since the Level 1 calorimeter trigger is based on towers, the most straightforward arrangement is to organize the electronics by tower at the board level. Each tower, consisting of the signals of a  $5 \times 5$  group of crystals [Bol91, Kra94a], will be read out by a single calorimeter readout module (Section 10.6), and the calorimeter trigger primitives will be constructed from tower sums formed on the board. With the numbers of crystals listed in Chapter 7 and a  $5 \times 5$  tower structure, there will be 24 towers in  $\phi$  and ten towers in  $\theta$  along the barrel, with the towers at the backward edge of the barrel containing only  $4 \times 5$  crystals. For the endcap, there is a 20-fold  $\phi$  symmetry in the three superlayers, with each  $\phi$  sector containing 45 crystals. Two towers will be formed from each  $\phi$  sector, with five unused channels shared between each pair of boards. This gives a total of  $24 \times 10 + 20 \times 2 = 280$  towers. The trigger sampling rate is fixed to be the same as the data digitization rate (4 MHz).

For each time sample, the 25 channels of data are calibrated and converted to a cruder energy scale of 8 or 16 bits using a look-up table in SRAM. This table allows a large degree of flexibility; an individual channel threshold cut can be trivially applied to remove low-energy backgrounds. The 25 outputs are summed using a programmable logic device (PLD) to give a total tower energy for each sample, again in 8 or 16 bits, and the sums are transmitted to a central trigger crate. These sums form the calorimeter trigger primitives, although it is possible that, to obtain a sufficiently accurate time and energy for the tower, further processing might be performed on the readout modules rather than in the trigger crate. However, assuming the raw sums are sent, the data rate is  $4 \text{ MHz} \times 8\text{--}16 \text{ bits}$  or  $4\text{--}8 \text{ Mbytes/s}$ .

### Transmission Links: Tower Gathering

The transmission of the calorimeter primitives is done via a dedicated point-to-point serial link. The 4–8 Mbytes/s rate is reasonably low and is straightforward for a TAXI or similar link [Wun94]. The primitives are collected together either on several cards in one trigger crate, or possibly all on a single card.

### Particle Finders: Global Cluster Formation

Due to energy sharing between towers, the towers are considered in neighboring pairs, called dominos [Ell92], as indicated in Figure 10-26(b). All possible dominos are formed, and the total energy of each domino cluster is used for the trigger. There will be three separate energy thresholds for clusters in the trigger: an M cluster (above a low energy consistent with a minimum ionizing particle, *e.g.*, 150 MeV), a G cluster (higher than a MIP, *e.g.*, 600 MeV), and an E cluster (consistent with a Bhabha electron). These correspond to the trigger objects in Table 10-7. The multiplicities of each object required for a trigger are programmable in the central trigger logic. Due to the differing crystal lengths in different regions of the calorimeter (Table 7-3), the MIP signal size will depend on  $\theta$ , and different energy thresholds can be used in different towers if necessary.

### Particle Counters: Global Cluster Multiplicity Finding

To make isolation requirements on the clusters, the cluster pairs are grouped together in superclusters of between 9 and 12 towers each. For example, with eight superclusters in  $\phi$  and three in  $\theta$  in the barrel, and eight in  $\phi$  and one in  $\theta$  in the endcap, there will be a total of 32 superclusters. By counting the number of clusters above threshold in each supercluster, and comparing which superclusters have been hit, simple isolation requirements such as two non-neighboring superclusters can be easily set.

One subtlety is that low-angle Bhabhas with both the electron and positron hitting the calorimeter have a cross section on the order of 30 nb [Eis90], giving a rate of  $\sim 100$  Hz at design luminosity. These are prescaled.

## 10.9.4 Global Trigger

The global Level 1 trigger consists of two pieces: a track matcher and a global decider.

## Track Matcher

The track matcher uses  $\phi$  maps from the drift chamber and calorimeter triggers to find the match multiplicity. There are two types of matches: a track to barrel M cluster, and B track to barrel or endcap M cluster. The track matcher takes the different  $\phi$  granularities in the drift chamber (42 bins) and calorimeter (24 bins) into account.

For the drift chamber and calorimeter, the multiplicity logic defines distinct particles using variable minimum  $\Delta\phi$  cuts. The multiplicity of each type of match is reported as 3-bit numbers to the global decider.

The track matcher reduces the orthogonality of the two independent Level 1 triggers and is intended for use only during very severe beam backgrounds.

## Global Decider

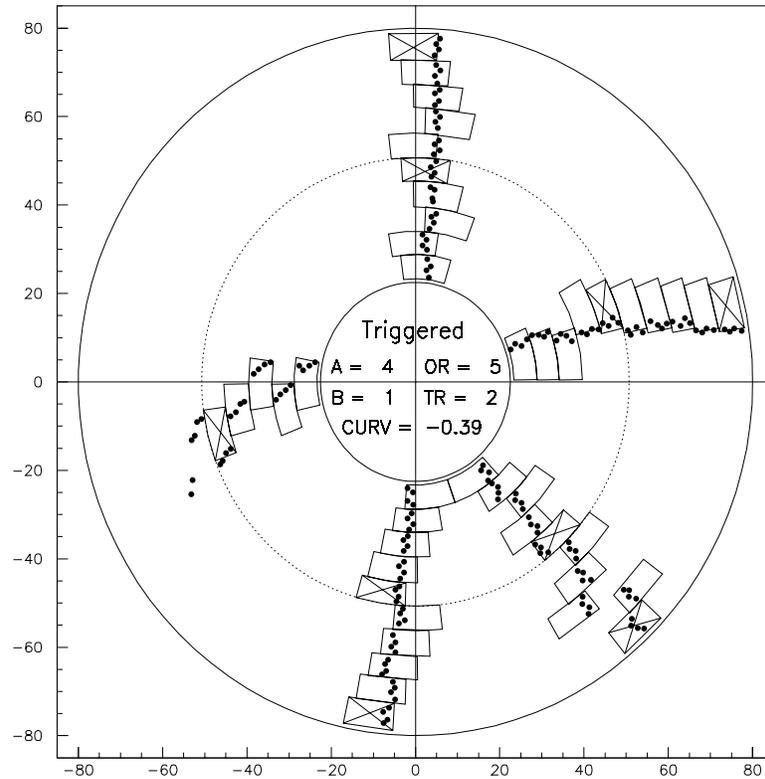
The global decider uses the multiplicity lines from the drift chamber, calorimeter triggers, and the track matcher to issue the Level 1 strobe. It also has access to the calorimeter  $\theta$ -slice number for single-prong Bhabha events that need to be prescaled as a function of  $\theta$ . For a diagram of the the trigger input lines, see Figure 10-24. A completely flexible trigger is implemented through the use of a RAM-based look-up table with 24 input bits (each of which can be masked off) and one output bit. This implementation is upgradable.

### 10.9.5 Simulation

The Level 1 trigger was simulated using the GEANT model of the *BABAR* detector, BBSIM. For the drift chamber trigger, the scored space-point hits in each layer of the GEANT model were converted to wire hits that were fed to segment finders. The BLT algorithm [Kin89] was used to find tracks, and global trigger criteria were constructed based on counting distinct A tracks and B tracks [Kra94b]. A display of an example  $B^0$  event is shown in Figure 10-27. A curvature finder was simulated using axial segment positions and patterns as inputs.

For the calorimeter trigger, the scored energy deposits in each CsI(Tl) block were summed into fixed  $5 \times 5$  towers that were paired to form two-tower (50-block) clusters. Isolation criteria in  $\theta$  and  $\phi$  were used to count clusters [Kra94a].

The implicit transverse-momentum and solid-angle trigger acceptances for single muon tracks, identified as the A-track, B-track, and M-cluster objects of Table 10-7, are shown in Figure 10-28. Note that the inefficiency due to tracks hitting adjacent corners of four towers,



**Figure 10-27.** Example of trigger response to a simulated  $B^0$  event. The dots represent wires that were hit. Boundaries of segment-finder supercells that were hit are drawn. Crosses mark  $\phi$  locations of A tracks and B tracks in counting layers. The event satisfied the open trigger requirements.

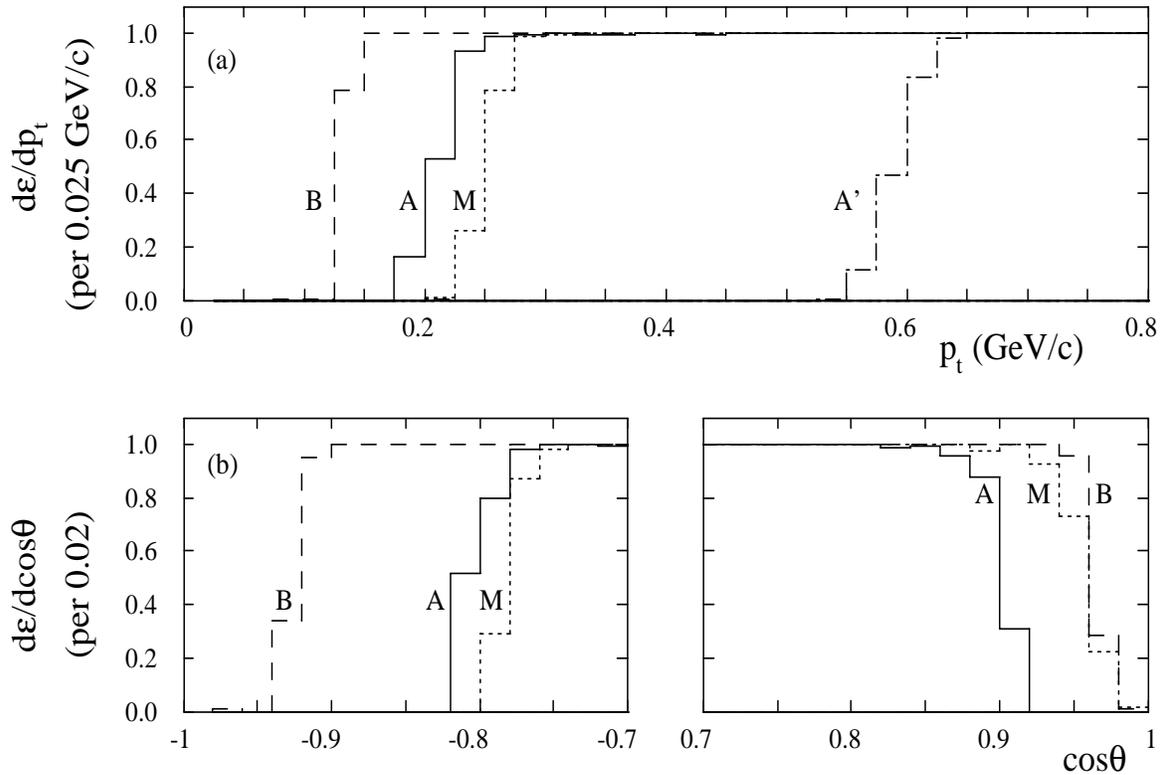
thus depositing only half their energy in one domino, is negligible. Event efficiencies and rates using this simulation are given in Section 10.8.5.

## 10.9.6 Trigger System

### Timing Considerations

The Level 1 trigger has a latency of  $9.5 \mu\text{s}$  and a jitter window of  $\pm 0.5 \mu\text{s}$ . These constraints are dictated by the characteristics of the front-end systems.

The latency time was set to a number that would not change and that would provide the maximum time to perform the Level 1 trigger while not degrading the physics of the vertex



**Figure 10-28.** Differential efficiencies for single muons to be identified as trigger objects A, A', B or M (see Table 10-7), as a function of (a)  $p_t$  with  $|\cos\theta| < 0.7$  and (b)  $\cos\theta$  with  $p_t > 0.4$  GeV/c.

detector. At full luminosity and 10 times background, there is little impact on the physics by the latency. At full luminosity and 10 times background, the time uncertainty will result in 30% occupancy in some sections of the vertex detector due to lost beam particles.

The pipeline length is composed of several elements, listed in Table 10-12. The trigger contains circuitry to align the track linker output to the late-arriving cluster finder output before these reach either the track matcher or global decider. The estimated minimum pipeline length of  $6\ \mu\text{s}$  is well within the fixed latency of  $9.5\ \mu\text{s}$ . Part of the safety margin is expected to be consumed as detailed designs are developed for the trigger. To make up any difference or account for potentially shorter calorimeter shaping times during high-background running, the output of the global decider passes through an adjustable-delay unit before producing the Level 1 strobe.

The jitter is set by the characteristics of the drift chamber and calorimeter. Calorimeter signals will have a  $1\text{--}2\ \mu\text{s}$  peaking time. These signals are sampled every  $0.25\ \mu\text{s}$  to form the trigger primitives. Simple peak finding is used in the Level 1 trigger. This results in

Path	Latency
Detector Signal Development	
Drift chamber Drift Time (600 ns)	
Calorimeter Peaking Time (4000 ns)	4000 ns
Transmission to readout modules	100 ns
Logic to Form Local Trigger Primitives	
Drift chamber Segment Finder (250 ns)	
Calorimeter Tower Summer (250 ns)	250 ns
Transmission to Particle Finders	
Drift chamber Track Linker (300 ns)	
Calorimeter Cluster Finder (250 ns to 450 ns)	450 ns
Logic to Form Trigger Decisions	500 ns
Transmission to Fast Control Distribution	100 ns
Transmission to Readout Modules	100 ns
Transmission to Vertex Detector Front-End Cards	100 ns
Total Minimum Pipeline Length	5600 ns

**Table 10-12.** *Level 1 trigger pipeline length.*

finding the peak to within one time bucket or a resulting jitter of  $\pm 0.25 \mu\text{s}$ . Drift chamber signals are sampled at a faster rate. They are extended by digital one-shots of 600 ns (the maximum drift time). Decisions are made every 250 ns by the segment finders. The resultant jitter window is 850 ns. A trigger jitter shorter than  $\pm 0.5 \mu\text{s}$  would reduce the storage and bandwidth requirements on the front-end systems, which could be significant for the vertex detector.

Globally, these triggers from the drift chamber and calorimeter are processed to form a global trigger by the track matcher and global decision maker every  $0.5 \mu\text{s}$ . This specification is set to be consistent with the Level 1 trigger jitter requirement of  $\pm 0.5 \mu\text{s}$ .

## System Functions

The Level 1 trigger system interacts with almost all other components of the *BABAR* electronics and online computing systems.

Global triggers will have prescaling of the selected trigger types. This prescaling is included to allow for optimization of the desired rate for physics data and for monitoring functions such as Bhabhas.

It is necessary to have timing and control interfaces to the fast control system. The trigger delivers the Level 1 accept signal to the fast control system. Fast control produces the Level 1 strobe for the front-end systems. The trigger is a consumer of fast control signals like all other front-end systems.

Monitoring of deadtime and efficiency must be part of the functionality of the electronics system. Deadtime monitoring is accomplished in the fast control system by measuring the physical time the trigger is not accepting new triggers and monitoring of scaled-down triggers such as Bhabhas. The efficiency of the trigger will be determined by the online system that is monitoring the primitives and resulting triggers. The scheme to perform this functionality must be included in the online design.

Readout by the data acquisition system will include all trigger primitives, drift chamber and calorimeter trigger output, and global decisions. This is needed to monitor the efficiency and operation of the trigger for both online and offline analysis.

The trigger will be partitionable such that it can run at the crate level, trigger system level, and *BABAR* system level. This functionality is included to allow for installation, debugging, and system testing. As part of the testing, the system will have the capability of injecting a calibration signal into the front-end electronics that allows for the determination of trigger primitive thresholds. Provisions will also be made for memories that can be downloaded with patterns which can be played through the trigger. These functions will facilitate the characterization of the trigger and help determine whether it is operating as desired.

## 10.10 Data Acquisition

---

### 10.10.1 Introduction

The key function of the *BABAR* data acquisition system is to collect the event data selected by the Level 1 trigger from the front-end and trigger systems, and transfer these data into a processor farm for packaging and filtering. Both hardware and software are required to concatenate the event fragments from all the data sources into complete events. Hardware and software protocols are needed to provide error-free transmission of the data, detect invalid operating conditions, prevent buffer overflows, and throttle the trigger rate. The data acquisition system also provides calibration, monitoring, and debugging facilities. The specialized equipment needed to calibrate the system electronics, including fast strobes, software processes, and hardware verification is a natural extension of data acquisition activities. Monitoring of front-end and trigger components, providing instantaneous status

information, localization of errors, and error recovery tasks such as the flushing of queues are a joint responsibility of the data acquisition system and the online computing system.

The data acquisition system interfaces to all detector front-end systems for event data collection, to the trigger systems for both data collection (of trigger information) and event data flow control, and to the online system for setup and runtime control as well as monitoring of the data acquisition system.

Data are transferred downstream within the data acquisition system in response to the identification of events by the trigger system. Under normal conditions, the Level 1 trigger rate is expected to be less than 1.5 kHz with the open trigger shown in Table 10-9. The largest component in this rate is cosmic ray events. Under severe machine background conditions, the Level 1 trigger rate can be kept below 2.0 kHz using the safe trigger shown in Table 10-11. In subsequent trigger levels, the drift chamber information and the silicon vertex detector information are used to restrict the tracks to those that originate near the interaction point. The final Level 3 trigger is expected to send approximately 100 events per second to be logged to tape. Contributions to this data rate are produced in many crates and flow through the event-assembly network over many individual data paths. Thus the total bandwidth can be easily achieved using any one of a variety of modern transport mechanisms. The *BABAR* detector trigger and data acquisition system is fully pipelined in both the trigger and the data flow paths. As a result, the system is nearly deadtime free at background levels up to ten times the rate given by machine simulations. Commercial equipment is used whenever suitable modules are available. The architecture is such that the capacity of the system is easily upgradable with the addition of data paths and processors.

## 10.10.2 Requirements

The requirements for the *BABAR* data acquisition system fall into four broad categories. First are quantitative requirements such as those on the bandwidths of data paths and the number of data sources to be accommodated. Next are architectural requirements that allow front-end and trigger systems to be operated as a complete system, or allow portions of each to be operated as separate partitions. Then there are requirements for initializing, controlling, and monitoring the system. Last are quality control requirements for items such as reliability, error rates, and maintainability. For a more complete discussion, see the *BABAR* electronics requirements document [Req94].

### Quantitative Requirements

The data acquisition system must satisfy the requirements for event rates, bandwidths, latencies, and deadtimes. The fundamental requirement of the data acquisition system

Detector System	Number of Readout modules	Number of Crates	Number of Data Sources
Vertex Detector	26	2	5
Drift Chamber	285	18	18
DIRC	3	1	1
Aerogel (ATC)	8	1	1
Calorimeter	280	18	18
Muon System (IFR)	7	2	1
Trigger	3	3	3
Total	612	45	47

**Table 10-13.** *Numbers of readout modules, crates, and data sources for each detector system.*

Maximum Trigger Rate	2 kHz
Max. Level 1 Latency	10 $\mu$ s
Max. Average Event Size	25 kbytes
Max. Event Data Rate	50 Mbytes/s
Max. Deadtime	10%
Number Data Sources	47
Number Level 3 Processors in Farm	10

**Table 10-14.** *Requirements of the data acquisition system.*

is that it be capable of collecting all event data for those events selected by the trigger, while imposing less than a specified amount of deadtime. Table 10-13 shows a list of all the detector systems that generate event data. The middle column of the table gives the number of readout crates required to house the modules associated with each detector system; the number of data sources for each detector system is shown in the last column. Most crates contain only one data source; the crates for the vertex detector each contain two or three data sources.

Together, all sources are expected to generate approximately 25 kbytes of data per event, with a combined event data rate of 50 Mbytes/s. Other requirements are derived from the expected physics and trigger rates, and the deadtime requirement of less than 10% at design luminosity. A list of important requirements is given in Table 10-14.

### Architectural Requirements

The system must scale modularly to accommodate upgrades, since it is inevitable that the data acquisition requirements will change during the life of the *BABAR* detector. A key element in this scalability is the design of the event-assembly network.

Furthermore, the data acquisition system must be partitionable. Partitioning means logically dividing the electronics of the detector into two or more independent subsystems, each with independent control. This allows multiple detector systems to perform calibration, diagnostic testing, and data acquisition tasks simultaneously and without interference. With this design, it will be possible to assign individual crates or entire subsystems to a partition under computer control without the need for recabling.

### Control Services Requirements

The data acquisition system must also provide a set of services to control and utilize the calibration and monitoring facilities built into the various front-end and trigger subsystems. It must recognize and handle error conditions—*i.e.*, errors in the transmission of data or malfunctions of components of the system. It must record non-event data streams such as magnet currents, high voltages, and accelerator parameters. It must initialize, download, and control the various processors and modules distributed throughout the system. Finally, it must deliver event data to consumer processes reliably.

### Quality Control Requirements

The data acquisition system must ensure the integrity of the event data stream. It must transport data from the front-end and trigger systems without errors in either data content or the identification of origin.

Care is being taken to ensure a high level of reliability. Proven industry standards and well-supported commercial software and hardware will be used wherever possible. Diagnostic features such as error detection and correction schemes will be implemented. To minimize downtime, the data acquisition system will be able to detect operational errors or malfunctioning modules and correct or report these conditions. Most system hardware will be implemented with modular components to permit the rapid replacement of malfunctioning units.

### 10.10.3 Architectural Overview

A block diagram showing the main components and interfaces of the data acquisition system is shown in Figure 10-29. In this figure, the data flow from left to right. The main components are the readout crates (VME), the event data flow control system (fast control), the event-assembly network, and the Level 3 event processing farm. The main interfaces shown are the front-end interface, the Level 1 trigger interface, and the slow control interface. Within each readout crate are the components of the various detector front-ends, trigger interfaces, and data acquisition system. These components collect analog data from the detector front-ends, process them in response to a trigger signal, and deliver digitized data to the downstream parts of the data acquisition system. The main components of the data acquisition system are described in the following sections.

#### Crate-Based Electronics Overview

Within the readout crates, there are four principal components or modules:

- readout modules that interface to the detector-mounted electronics and, for the drift chamber and calorimeter, provide specific information to the prompt trigger;
- fast control distribution modules that interface the crate to the fast control system;
- a readout controller that connects to the detector control system, provides monitoring and diagnostics, reads the event data fragments in the crate, and controls the flow of data through the buffers; and
- an event-data port card (data source) that interfaces the crate to the event-assembly network.

In each crate, the readout module connects the detector-mounted electronics to the detector-specific trigger, fast control system, and readout controller. Since most of these interfaces are identical for all detector systems, it is useful to look at the common features. Emphasizing the commonality of the systems creates a unified data acquisition structure. Although there are unique interfaces to the on-detector electronics and to the trigger, there is a common interface to the fast control system and another to the readout controller through the crate bus.

#### Event-Data Control Overview

The flow of the event data through the data acquisition system is controlled by the event-data control system. Its primary function is to keep data flowing efficiently from the

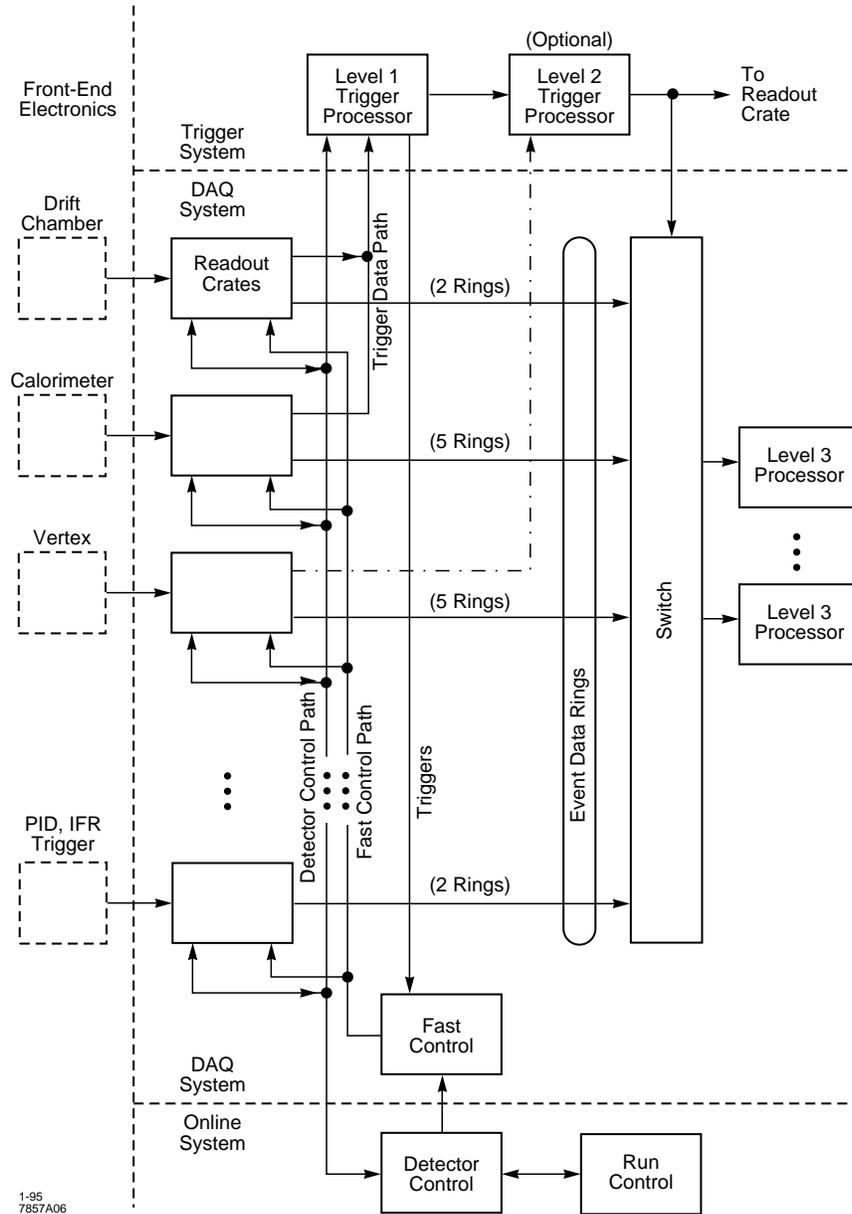


Figure 10-29. Block diagram of the Data Acquisition system.

front-end electronics, through the readout and trigger-system crates, through the event-assembly network, and to the Level 3 processor farm. The data acquisition system has three separate mechanisms for providing the event-data control: high speed control and trigger throttling (the fast control system); VME-based data collection and flow control algorithms; and network control and resource allocation.

The high-speed flow control uses a “backpressure” model to guarantee the orderly delivery of data. The mechanism works as follows. Ten  $\mu\text{s}$  after a particular beam interaction occurs, the Level 1 trigger declares a Level 1 accept. The fast control system fans out this Level 1 accept signal to all crates currently accessible to the main trigger partition. The signal is received synchronously at these readout crates by the fast control distribution modules. If this module is configured to respond to signals from the main partition, then the Level 1 accept is propagated through the crate’s fast control bus to each readout module. When input queues on the readout modules become full, a busy signal is transmitted back through the fast-control system to the trigger.

When a readout controller processor has completed processing an event fragment, the data will be located in a region of memory accessible by VME. The processor will also place the location of this event fragment and the length of the fragment in a separate queue, also accessible by VME. If either the data queue or the pointer queue becomes full the processor stops and the VRAM fills, resulting in a throttle condition. The locations of the heads and tails of both queues are accessible to both VME and the processor. The read-modify-write protocol of both the VME and processor ensure that the queues are properly maintained.

Each readout controller within a partition is assigned a unique position in the readout sequence. When the readout controller in the first crate in the sequence finds that it has a block of events ready and receives an event-request token, it transmits the event-data fragment to a designated Level 3 processor. When the transmission is complete, the token is passed to the next readout controller in the sequence. This process continues until the last crate in the sequence is reached. At that point, the token is sent back to the Level 3 processor. When the complete event has been successfully read, a message is broadcast to all readout crates signaling the successful completion of the event assembly. A suitable algorithm will distribute events among the Level 3 processors.

### Event Assembly Overview

The task of the event-assembly network is to send the event fragments from the 47 event-data port cards to the designated Level 3 processor.

Some of the requirements of the event-assembly network are listed in Table 10-14; these numbers will change as the design progresses. For example, the number of destination processors will depend on the computational speed of the processors, refined estimates of

the computational capability required to process a single event, and refined estimates of the event rates generated by the trigger processing.

The event-assembly network will be based on a commercial network similar to those presently being used or under development for the computer and telecommunications industries. Such a network, combined with the use of standard communication protocols, has many advantages. The hardware is readily available from multiple vendors. It is reliable and has a good price/performance ratio, especially when development costs are considered. The communication protocols are already implemented and include extensive error checking, flow control, and collision management. The data are usually self-directing and do not require any central intelligence to control the data flow. The amount of programming labor that is saved by employing this solution instead of a home-grown network is enormous.

The benefits of an interconnect based on commercial networks are not free. The key disadvantages are constraints imposed on the architecture by network hardware and protocols. The advantages, however, include the possibility of implementing flexible load-balancing algorithms in software without changes to the hardware configuration and the capability of scaling the system incrementally.

The flow control and error recovery procedures of the network protocol can be used to implement a simple data-flow architecture. A distributed flow control system allows high link utilization without special hardware to implement control functions. Readout synchronization in the trigger system and front-end electronics is required to achieve this benefit. A Level 1 trigger inhibit must be asserted if any readout crate lacks sufficient buffering to read out its portion of an event. This prevents partial readout of an event.

The network protocols and interfaces have been designed for efficient transmission of large messages and modest transaction rates. Thus, a message size of about 4 kbytes or more is required for efficient use of the network. This will require fragments containing data from several events to be sent in each message originating at a data source in a readout crate. This has the advantage of decoupling the real-time event rate from the rest of the data acquisition system. Sufficient memory must be provided at appropriate points in the system to buffer the large messages.

## Online Interface Overview

Software development for the readout controllers and readout module processors requires a commercial development environment and a commercial real-time operating system. The *BABAR* computing group is considering the VxWorks real-time operating system and the EPICS control system software tool kit for use in the control systems. Broad use of these tools throughout the computing, online, data acquisition, and trigger systems permits shared expertise and seamless integration of experiment control and data acquisition. The online

interface provides the physical means whereby operators and applications software have the ability to control and monitor the many processors in the front-end and data acquisition systems.

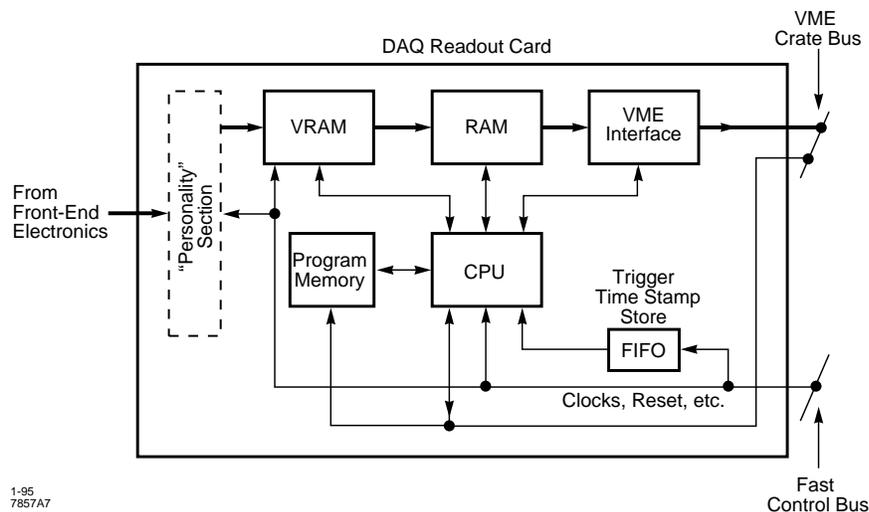
#### 10.10.4 Readout Crates

The readout crates provide housing for several types of modules required for the electronics system. These include the readout modules, the fast control distribution module, the readout controller, and the event-data port card module. The readout crates will be standard VME crates. All readout data, messages, and monitoring are to be transferred between modules on the VME backplane bus. Lines on the VME P2 bus are to be used to transmit high-speed control, trigger, and timing signals. Each crate will have spaces for 21 modules. These will be able to accommodate the readout controller, the event-data port card, the fast control distribution module, a temporary bus monitor module (for debugging), one space reserved for future use, and the 16 readout modules.

##### Readout Modules

All detector systems are able to present a common hardware and software interface to the central electronics system because of the design of the readout modules. As shown in Figure 10-30, each card contains a standard set of functions, which are shown as solid blocks. The “personality” section, shown in the figure as a dashed block, is designed to accommodate each detector system’s special characteristics. It is expected to make provision for bi-directional data flow between a readout module and its associated front-end electronics. This will permit, for example, dummy events or configuration data to be downloaded into the front-end electronics for tests of the overall system.

The readout module receives triggers, clocks, fast resets, *etc.*, via the crate’s fast control bus. Triggers are time stamped and buffered in the FIFO before being presented to the card’s CPU. This feature buffers the CPU from the rate fluctuations. The FIFO *almost full* condition results in trigger throttling; receipt of a trigger in the *full* condition generates an error signal. The interface between the personality section and the readout module will convey all necessary signals to permit any special processing needs of the detector systems. These may include analog signal processing, fast digitization, and data buffering. A bit-parallel digital data bus will convey data over the interface. A serial clock signal strobes the data into or out of the VRAM depending on the state of a direction signal. The clock line stops whenever there are no incoming data. The numbers of DAQ readout cards required by the various detector systems are given in Table 10-13.



**Figure 10-30.** Block diagram of a DAQ readout card.

To achieve event synchronization across the interface, all data streams and event serial numbers are initialized by a system-wide reset signal. The read strobe then ensures that the multiple data streams remain synchronized. The personality module sends an error signal when it detects transmission errors relative to the incoming data stream.

### Fast Control Distribution Module

The fast control system provides all of the high speed timing signals required for the selection, synchronization, and identification of data. These functions include fast reset, initialization, synchronization, clocking, calibration, triggering, and busy indications. The system also conveys those signals needed for rapidly enabling and disabling the triggers as required by conditions of filling and emptying of data buffers. A crate fast control bus will augment the VME backplane, and many of these signals will be transmitted over the unassigned pins of this bus.

The fast control distribution module is a VME card that receives the high-speed signals from the fast control system and repeats the signals on the crate fast control bus. This card is able either to drive directly the signals received from the clock and control interface or to generate these signals under program control through VME. This latter capability is needed to perform stand-alone operation and calibration of systems created by forming partitions. For example, the timing card can be told to issue a trigger or a calibration strobe at a specific moment, or select a particular partition with which to communicate. In addition, this card performs various control functions within the crate, including receiving a trigger throttle

signal and an error signal from each readout card. These signals are then made available to the central fast control system, to be used in the trigger throttling algorithm.

### Readout Controller

The readout controller is the only VME master contained within each readout crate and so is responsible for many activities, the most important of which is the event readout. In addition, the processor on this board is responsible for downloading software to all processors within the crate, responding to crate interrupts, performing detector control, monitoring, calibration, debugging, and operating in stand-alone mode. Since EPICS is likely to be used in the detector control subsystem, this processor must operate under VxWorks. While minimal memory requirements are needed to support the small event sizes of the *BABAR* data stream, excellent processor throughput will be required to assist in data formatting, queue management, and interrupt servicing.

The processor on the readout controller receives a trigger as an interrupt. After the processor services the interrupt, it executes user-defined codes to allocate buffers and read in the data. It builds a list of commands to transfer these data and then performs them as a sequence of high-speed input-output operations. When the readout controller determines that a complete event has been processed within the crate, it waits for a signal (token) before requesting transmission of the event fragment to the event-assembly network. Due to occasional long latencies in the readout, there may be several events ready for readout.

### Event-Data Port Card

The event-data port card (data source) in each crate interfaces to the transport medium for all event data streams and is responsible for transmitting the crate's event fragments via this medium to the event-assembly network. Since none of the expected data streams exceeds 5 Mbytes/s, many data transmission technologies and logical arrangements are possible candidates for this task. *BABAR* uses switched FDDI.

The data sources associated with the various detector systems output data at widely different rates. This makes it quite reasonable to arrange the data sources into a topology of several rings, thereby reducing the number of virtual sources seen by the event-assembly network. This arrangement is discussed in detail in Section 10.10.6.

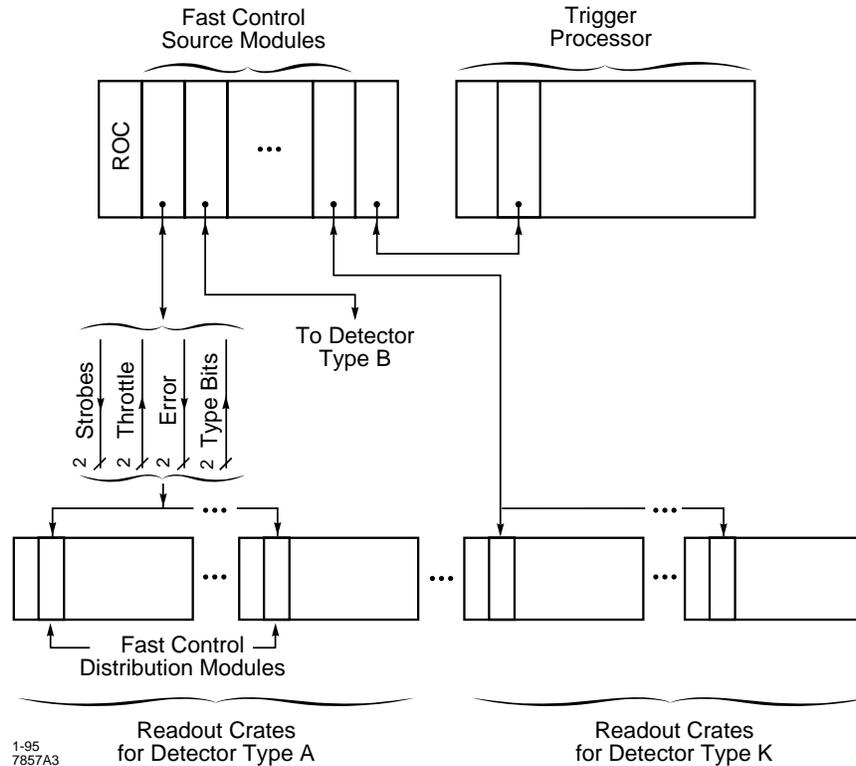


Figure 10-31. Partitioning logistics.

### 10.10.5 Event-Data Flow Control

#### Partitioning

It is important to be able to partition the electronics systems for the various detectors into independently operating systems for commissioning, calibration, and repairs. It is also useful to be able to run the detector in normal data-taking mode even though some parts are unavailable for various reasons, including being independently operated. Counting the trigger, *BABAR* contains seven independent systems. It will be possible to create at least seven partitions, allowing the systems to be initially commissioned independently. Figure 10-31 illustrates the partitioning scheme.

The lowest level of granularity in partitioning is the crate. A crate is either offline or online in a partition. A crate in a partition is under the direct control of the data acquisition system. An offline crate can still communicate with the online software for purposes of reading or

setting voltages, running dedicated software diagnostics, *etc.*, but it is not included in the normal data acquisition operation.

The partition containing the trigger system is referred to as the main data-taking partition. From the trigger system, triggers are distributed to other systems via the fast control distribution system. We do not envision the trigger system itself as partitionable, although it may be possible to take some of its crates offline for limited repairs. A crate which is offline, or which is not in the main partition, may not be generating valid trigger primitives and will be ignored by the trigger system. This will be set up as part of the partition initialization process.

The crates in a partition other than the main partition still communicate using the fast control distribution system and the data-transfer protocols. A system within that partition may be responsible for generating strobe signals for distribution in lieu of triggers, for example, to control calibration and diagnostic activities. It is also possible to define a partition that does not contain a module to drive the fast-control system; this partition is then fully functional from a data-transfer perspective but will require that all processing be initiated by software messages instead of by hardware strobes.

Crates can be read out via the main data transfer path or at a much slower rate via the controls local-area network. Depending on the final architecture for the main path, it may be possible to divide it logically among partitions or to share it on a first-come first-served basis. Either of these options will provide acceptable functionality for commissioning and calibrating the detector.

### Fast Control System

Clocks, strobes, type-bits, throttle signals, and error indications must be distributed among the data acquisition crates. These must be distributed in a flexible way to allow systems to use them for their own commissioning needs within special purpose partitions.

The free-running nature and small physical size of the *BABAR* electronics system makes timing and control distribution relatively easy. We plan to have a few central modules (fast control source modules) generating the required timing and control signals. The timing signals will be derived from the PEP-II rf clock and will maintain a stable phase relationship at the nanosecond level on the module output.

The set of strobe signals is used to carry time-specific information such as the Level 1 trigger and calibration strobes. Type-bits are used to indicate special purpose strobes, for example, an occasional diagnostic trigger which requires custom processing by the data acquisition system. Throttle signals are used to hold off further strobes when a system or crate is unable to process further requests for a short time, or they can be used to indicate that processing

of a given strobe is complete. Error indications are used to inform the online system that some condition has been detected that requires higher level intervention.

The main partition will have type-bits, to classify events as, for example, normal, Bhabha, uncompact, and snapshot. The last two are typical of several types used for occasional checks of the data acquisition logic. They will carry the Level 1 accept on a strobe and may use a second strobe to inform data acquisition modules in advance of a diagnostic trigger. Strobes in the main partition, including any that need interlocking to the throttle signal and special intertrigger time requirements. The fast control system monitors trigger deadtime.

During system-specific calibration, one might use several strobes to command the crates to generate local timing signals, and then trigger data readout on the result. It is the responsibility of the particular system to generate the required strobe signals, with interlocking to the throttle control as appropriate. If a system has only one set of electronics for generating these signals, it will not be possible to split the system into two fully functional partitions.

### **Fast Control Distribution Module**

We currently envision eight strobes, eight busy signals, eight error signals, and eight sets of type-bits. This is enough to allow all seven major systems of the *BABAR* detector to run in independent partitions when required.

To distribute these signals, each data acquisition crate will contain a fast control distribution module. It receives the entire set of signals and provides a standard interface to the crate. It can be programmed via the VME interface to route particular strobes and/or type-bits to either front panel connections or a set of reserved backplane signals, or both. Additionally, VME interrupts can be generated from strobes. Throttle and error signals can be driven from front panel connectors or from backplane signals. All signals can also be interrogated via VME read operations for diagnostic purposes. VME-readable scalars and LEDs are provided for most functions to help in commissioning and debugging.

### **Fast Control Source Module**

The fast control source module is used to drive the distribution system. It is configured via VME to drive strobes and type-bits from dedicated backplane signals or front panel connections, and to break out throttle and error signals onto the backplane, front panel, or into VME interrupts as required. Typically, the error signal will interrupt the local readout controller, which will poll the appropriate modules to find the cause of the problem. All signals can also be interrogated via VME read operations for diagnostic purposes. One or more of these fast control source modules is then used by the trigger or calibration control electronics in each system to distribute the required strobes, and other signals to the other

crates in the system for operation in a partition. It may be possible to combine both the fast control source module and fast control distribution module onto a single VME card to reduce expense.

Using this system, the online software can flexibly allocate signals to particular partitions by controlling the settings of the control module(s). This, in combination with informing the trigger system of which crates to monitor and setting up the data-transfer network appropriately, allows each partition to function independently.

### Throttling and Buffer Management

The basic hardware throttle for the *BABAR* data acquisition system is provided by the throttle lines of the fast-control system. Asserting the relevant line for a partition inhibits its strobes. In the data-taking partition, this means that distribution of the trigger is inhibited, causing deadtime. To allow an accurate determination of deadtime, there will be pairs of scalers counting system clock pulses—one gated by throttle and the other not. One pair of scalers will be assigned to each possible partition. In addition, the trigger will count Level 1 strobes internally, even when the data acquisition system is inhibited. All these scalers will be accessible as part of the normal event stream.

There are two sets of buffers that the fast control system must protect from overflow. The first of these are contained in the readout chips of the vertex detector electronics. These contain on-chip back-end buffers which are of fixed length and three events deep. They buffer against the variations in transmission time along the fiber to the readout modules. The fast-flow-control system must promptly assert throttle to prevent any overwriting of data if all three on-chip buffers on any readout chip are occupied.

Due to the limited space available, this portion of the fast control system will reside in the link controller at the data acquisition end of the line. The link controller will keep track of the chip buffer occupancy by monitoring a local model of the chip buffer states.

The other buffers controlled by the fast control system are the VRAM buffers on the readout modules. These are operated as circular buffers, in some cases (such as the drift chamber subsystem) with addresses linear in time, and in other cases, (*e.g.*, the vertex detector system) with addresses linear in event number. In the event-linear case, the throttle will be applied if there is too little space to accommodate two maximum-size events.

There is a third source of fast control signals. When a Level 1 strobe is presented to the readout module, the address in the VRAM of the corresponding data needs to be noted for later digestion by the readout module's CPU. This work queue will be maintained in a hardware FIFO. This FIFO will also be able to generate throttle if it is at risk of filling completely.

The throttle signals from each readout module will be presented individually to the fast control distribution module in its crate. This will allow disabling of these signals from broken hardware.

### Event Data Flow Control

The fast flow control system is the only form of event data flow control that directly throttles the trigger. All later stages of control operate by a backpressure mechanism.

In the backpressure scheme, any stage in the event flow—starting with a processor in the Level 3 farm—which fills up simply refuses to release resources used by its upstream neighbors. This eventually causes the upstream neighbor to fill. The process cascades until the upstream neighbor is the memory on the readout module. When that memory fills, throttle is asserted and deadtime is introduced.

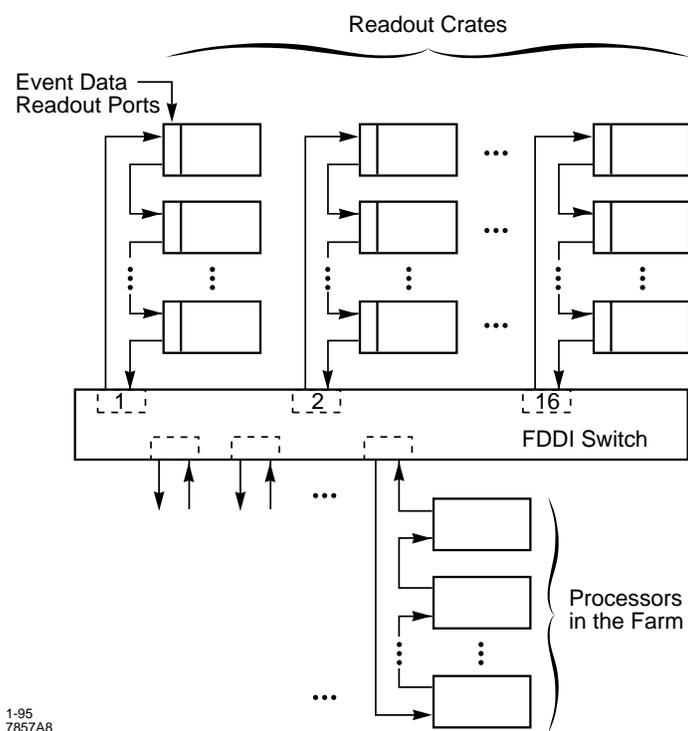
These algorithms become more complex in the case in which events are bundled into super events to amortize the cost of interrupts. Note that bundling events can also reduce the interrupt load within the crates; the fast control system could be programmed to provide interrupts to the readout controller only after a designated number of events is available in the readout modules.

### 10.10.6 Event Assembly

After the delivery of the Level 1 accept, the readout modules and readout controllers proceed autonomously to the point at which all useful event data from a crate are collected into that crate's readout controller. At this point, the data represent the crate's event fragment. The event assembly process collects all the crate-level fragments for a particular event into a single Level 3 node for further filtering, reconstruction, and storage. The baseline technology choice is switched FDDI, a commercially available high-speed network.

Such an assembly network must provide enough throughput to handle the worst-case average event size at the Level 1 accept rate with a reasonable latency. Instantaneous rates above the network throughput will have to be buffered before the network or deadtime will result. Since chance fluctuations may fill any buffer, the goal is to provide enough buffering to keep this deadtime below the few percent level at 10 times nominal background.

The pre-network buffers imply a latency for assembling events that is fluctuating and potentially large compared to the event rate. Since it is easy to provide this buffering on the readout controllers, and there are no hard real-time restrictions after the data have been



**Figure 10-32.** *Crate readout ring topology.*

rescued from the trigger-latency buffers, this latency is not important. The fluctuations introduced by the Level 3 nodes are likely to be much greater.

In a switched network with multiple destinations, only the switch needs to see the full data bandwidth. Typically, switches are designed with very high bandwidth backplanes. For example, one currently available FDDI switch has a capacity of over 3 Gbits/s (about 400 Mbytes/s). Real utilization approaching 80% of this figure has been observed at LBL, whereas the *BABAR* maximum demand is only 50 Mbytes/s. Hence, the switch itself should not be a limit.

After the switch, the next potential limit is the capacity of an optical fiber. At an FDDI data transfer rate of 100 Mbits/s (12.5 Mbytes/s), it is clear that a single fiber cannot cope with the load. On the other hand, assigning one fiber per data source would result in a large number of input ports on the event-assembly network, with many of the fibers significantly under-utilized in terms of data rate. A reasonable alternative is then to assign several data sources to each fiber data channel, with the number of sources per channel to be determined by the data rates of the sources.

System	# of Sources	# of Sources/Ring	# of Rings	Data Rate/Ring
Calorimeter	18	4	5	3.2 Mbytes/s
Drift Chamber	18	9	2	3.6 Mbytes/s
DIRC & ATC	2	2	1	3.8 Mbytes/s
Vertex	5	1	5	3.3 Mbytes/s
IFR & Trigger	4	4	1	3.2 Mbytes/s
Totals	47		14	

**Table 10-15.** Readout ring topology and bandwidth at 2.0 kHz.

Figure 10-32 shows a ring topology for combining several data sources or sinks onto a single fiber channel. Here, it is assumed that the event-assembly network is a 32-port device. Sixteen of the ports are assigned as input ports, with each port connected to a ring connection of data sources. The remaining ports are used as output ports to the processors in the Level 3 farm, which are also connected in rings. None of the expected maximum data rates per input ring exceed 25% of the 12.5 Mbytes/s bandwidth of switched FDDI.

These rings could be upgraded with a future higher performance technology at modest cost in order to take advantage of higher throughput. Technologies such as FCS (fiber channel standard), or SCI, can be used in a similar ring topology and provide significantly greater throughput. Technologies such as HIPPI and ATM do not have ring topologies, but suitable substitutes exist with similar performance improvements.

The sequencing of data onto the input rings is controlled by token passing. To gather event fragments, an input port of the event-assembly network will launch a token. The first data source on the ring will receive the token and will transmit its event fragment(s) via the fiber. When it is finished, or if it has no fragments to send, it immediately forward the token to the next source.

Tokens can be used to convey instructions to the data sources, *e.g.*, the Level 3 farm processor to which the current event data fragments are to be addressed.

The numbers used to assign data sources to particular rings are given in Table 10-3, where the last column shows the data rates for data sources assigned to each detector type. Table 10-15 shows how the 47 sources can be connected into 14 rings, with the rings having roughly equivalent data rates easily within the capacity of FDDI fibers.

On the output side of the switch, there needs to be at least six rings to keep the load within bounds. The number of farm nodes per output ring is essentially unlimited; *BABAR* will use whatever is needed to supply sufficient computing power.

The next potential bottleneck is the data rate capacity of the event-data port cards. These are the cards that interface a data source or a farm processor node to an FDDI optical fiber. Of course, every event-data port card will put bits on the fiber at the full 100 Mbits/s for the duration of its packet. The first question is how quickly it can or needs to supply packets so that the occupancy of the fiber is as high as possible. The second question is how much local processing power in the event-data port card is required to support a TCP/IP protocol at high rates. There are reports of event-data port card/processor combinations which can achieve full fiber loads even while running TCP/IP over the fiber, but these figures are likely to be reached only under ideal circumstances. Fortunately, doubling the number of nodes per ring directly cuts the event-data port card load in half, so there is a simple cure within this architecture. Also, this is a rapidly developing technology. *BABAR* will monitor developments and take advantage of performance improvements and cost reductions.

Note that, for the above minimal output configuration, the traffic of events out of Level 3, which are recorded on tape, represents less than 7 Mbits/s (800 kbytes/s). This is a sufficiently small load that it will not be discussed further.

The data rate per data source listed in Table 10-3 is the rate each crate on the input rings must handle. For some cases, such as the drift chamber, the load per event-data port card is quite modest. On these rings, the efficiency with which the FDDI protocol will multiplex the packets is of primary importance. On other rings, such as the trigger and vertex detector, the load per event-data port card is a significant fraction of the fiber bandwidth. Again, published reports indicate that such a throughput is attainable, but much depends on the event-data port card/processor combination; the *BABAR* readout controllers need to do more than simply feed the FDDI event-data port cards. Measurements on prototype configurations and modeling of more complete configurations will be done to decide if this is a serious problem.

Because the maximum throughput of the above configuration is determined by the input fibers, there is a simple chain of consequences. First, the input fibers must be used as efficiently as possible. Second, to use the fiber efficiently, each event-data port card should be ready at any time to send useful data. Third, for data to have a useful data density, the ratio of data to overhead should be high.

The second consequence suggests maximizing concurrence in the data gathering process both within events and across events. In particular, one would like each crate to send its portion of a given event to the destination farm node without waiting for other crates. In the FDDI protocol, if two nodes try to send to the same destination, one of them will be held off until the other finishes. Within a given input ring, this is not necessarily bad. If all crates tried to send at once, they would indeed serialize, but the fiber would be fully occupied. Unfortunately, however, the entire drift chamber fiber, for example, would have to be idle while the vertex detector crates on another ring are transmitting data. The solution is to have one ring's crates transmit a different event to a different destination while another

ring's crates are active. Hence, multiple events will be gathered simultaneously. Other data transmission technologies may not have this problem.

Because of contention for the output fiber bandwidth, some form of traffic shaping or control of fragment assembly order will improve event gathering throughput. An example of such an algorithm has been presented in Reference [Dou94]. In this scheme, the data acquisition crates are logically arranged in a ring. A farm node volunteers to handle an upcoming event by sending a token to the first crate in the ring. When the specified event arrives, the first crate sends its data to the farm node and forwards the token to the next crate in the ring. This crate sends its data and passes on the token. When the token leaves the last crate, the event is completely assembled. At no time is more than one crate talking to the destination node. Since each available farm node has registered a node with the first crate, that crate can begin delivering data for the next event before the first token has fully traversed the ring. Since the destinations of data for different tokens are different nodes, there is no contention for output bandwidth. Many variations on this idea are possible, and some may have better performance. For example, in the version outlined here, tokens do not overtake each other. This means that several events can be stalled waiting for the delivery of data from a particularly full crate. We intend to do simulation studies to evaluate the impact of such effects and ways around them.

Step three in the consequence chain is relevant since some crates, such as the drift chamber crates, have a low data rate. There is a risk that the useful data on the fiber will be swamped with headers and protocol messages. Fortunately, it is a simple matter to cluster multiple events together into bundles which are delivered as a unit into a single destination farm node. This amortizes message overhead over many events. The cost is more complex software in the farm nodes, which must deal with noncontiguous event fragments. Since the events are in random-access memory, a properly designed algorithm should minimize the performance loss. Any high rejection filters that run in Level 3 will be designed to run before the events are unscrambled. The cost of memory is such that bundles of 64 or even more events are practical.

### 10.10.7 Event Distribution

Once events have been fully filtered and reconstructed in the Level 3 farm nodes, they must be collected and stored on permanent media. The *BABAR* baseline design allows for both multiple partitions running simultaneously and multiple data streams within a single partition. This leads to several issues regarding event distribution.

Because crates are not partitionable, event distribution in a running partition involves little more than being sure that the participant at each stage—readout controller, Level 2, Level 3, and event consumer—knows where to forward its events. A resource-manager database

will be used to control configuration of partitions to prevent stages from being assigned to more than one partition. The resource-manager database will also be a source of current-configuration information for the control-source and distribution modules to ensure sensible configurations of strobes. This database can be queried by the event-recorder consumer at the beginning of a run to record the detailed configuration of the partition for use in later analysis.

Level 3 is made up of identical autonomous processors, which could easily be distributed across partitions. While *BABAR* will retain this ability, it is not likely to be used often; the full Level 3 farm will generally be assigned to the main partition. Partitions running in local mode will typically contain a few crates directly feeding a conventional workstation running an event consumer for monitoring and analysis. Such a workstation will run a single-node version of the Level 3 software to gather data from the partition's crates and supply events to the event consumer as described below. These event consumers could be located at the institutions responsible for the systems.

Partitions will be defined and configured through an interactive interface to the resource-manager database, with some common groupings of partitionable components such as the whole of Level 3 being predefined. Running partitions may not be reconfigured. Resolution of contention for resources, such as Level 3 nodes, will be achieved by discussion among the people on shift.

### Farm Node Allocation

Closely related to the problem of assigning fibers on the output side of the switch is the problem of assigning Level 3 processors to events. In CDF, for example, this decision is made by a central buffer manager which keeps track, via message exchange, of the availability of each farm node. The nodes are then queued for reassignment using a FIFO queue discipline. The disadvantage of this scheme is the high message load placed on the single buffer manager. An intrinsically distributed mechanism is needed to control the message load on any particular node.

The algorithm of Reference [Dou94], described above, may apply. This node allocation algorithm is presented as a near-round-robin queuing discipline for the farm nodes under the assumption that most of the time most of the nodes claim an event as the token passes it. Unfortunately, this will not be the case in general. First, note that the token must continue to circulate at all times. Under a lightly loaded system, most nodes will have already claimed an event which has not yet occurred and will therefore have no resources available; hence they will all pass the token. Under a heavy load, most nodes will be busy with their events and will again pass the token. In both cases, there is no speed control on the rotation of the token, and the token-related messages will increase rapidly in number. By recasting

this allocation problem in slightly different terms, one can choose from a rich literature of distributed algorithms [Lam74, Ric81, Dol82, Pet82].

### 10.10.8 Data Integrity

The data acquisition scheme is quite complex, with data fragments being bundled and sent through a network, and the bundles of fragments assembled into events. To be sure that events are being assembled correctly, each fragment will need an event label and content descriptor. As the events are assembled, these will need to be checked to ensure that all pieces of the event are completely and correctly included. A global process will need to track all triggers and verify that they are either rejected for cause or archived.

### 10.10.9 Research and Development

#### System Modeling

Simulations are being used to refine our knowledge of the data acquisition requirements. The trigger data-flows have been studied using our most detailed simulations of the expected backgrounds. The buffering required in the vertex detector [Lev94b] and drift chamber front-ends has also been examined using these background simulations.

The task for the immediate future is to simulate the data acquisition system as a whole to investigate signaling and protocol overheads, and to ensure that the bandwidth and latency requirements are understood. We intend to do this using commercially available simulation packages and to build up a family of detailed simulations of the various parts of the *BABAR* detector and its response to the expected backgrounds. The detailed simulations of the trigger and individual systems will also continue to be improved as more detail about their functions becomes available.

#### Alternative Network Technologies

Asynchronous Transfer Mode (ATM) networks will allow for data signaling rates per fiber of 155 Mbits/s (120 Mbits/s usable) in the near future and have the potential for 620 Mbits/s (480 Mbits/s usable). There is very widespread industrial support for ATM, which should ensure the availability of components at reasonable cost. There are some possible problems with use of ATM for data acquisition. The Fiber Channel runs at data signalling rates from 132 Mbits/s up to 1 Gbits/s. It has fairly widespread support in industry, especially in connection with high speed peripherals. At present, there are only a few manufacturers of

Fiber Channel switches, and cost per node is relatively high. Scalable Coherent Interface (SCI) networks operate at a data signaling rate of up to 1 Gbits/s. Industry support for SCI is modest.

## 10.11 Level 2 Trigger

---

The baseline design omits a Level 2 trigger stage and builds all events accepted by Level 1. Since having a Level 2 may prove to be desirable and even necessary, we describe this option here. The Level 2 process for an event begins with the Level 1 accept. It receives input data from the Level 1 trigger and some of the readout modules. A positive decision results in an event being built. Level 2 provides an opportunity for intermediate event filters which reduce the number of events built, and thus diminishes the burden on the readout module processors, the event assembly network, and the Level 3 processors. If all envisioned cuts are made, the rate of event building could drop from the 2.0 kHz of the Level 1 accept to 200 Hz.

### 10.11.1 Filter

The primary candidate for a Level 2 filter is a cut on the longitudinal component of the distance of closest approach of tracks to the interaction point (a  $z$  cut), for which the primary source of information for determining this quantity is the vertex detector. The volume of data needed for a vertex detector based  $z$  cut makes gathering the required information impractical before the Level 1 strobe.

A Level 2 trigger might also implement the cut on track transverse momenta (a  $p_t$  cut), thus simplifying Level 1. At a minimum, the  $p_t$  cut requires the Level 1 trigger information from the drift chamber. It could be accomplished in hardware or software, although for some of the proposed drift chamber geometries, a hardware implementation of the  $p_t$  cut may be complex and require more latency than the 10  $\mu$ s allowed for the Level 1 accept. A software implementation could benefit from additional information from the drift chamber.

Simulations of background processes indicate that appropriate cuts in  $z$  and  $p_t$  can reduce the trigger rate to near the logging rate with no substantial reduction in efficiencies, even for low-multiplicity tau decays [Lev94d].

## 10.11.2 Implementation

### Logical

The Level 2 design should have minimum impact on the design of the rest of the system and should require no substantial hardware developments. As an example of a possible implementation of Level 2, a network with a low-overhead fast protocol can be used to send one message per Level 1 accept to the crates serving the vertex detector, the drift chamber, and the Level 1 trigger. These crates return information to a Level 2 processor, which makes a decision before feature extraction takes place.

Using multiple processors for Level 2 decouples throughput from latency. The Level 1 accept rate (between 1 and 10 kHz), combined with a reasonable limit on the number of processors devoted to the Level 2 decision making (such as 10), implies an average latency of no more than 1–10 ms. This is the same order of magnitude as the length of buffers implemented with VRAM. If the Level 2 decision can be made in time to copy the relevant data from the VRAM, movement of the data upon a Level 1 accept can be avoided, thus reducing the load on the readout module processors.

Once a Level 2 accept has been issued, event building proceeds as described in Section 10.10, but at a much reduced rate. Events could be bundled if desired. The leading engineering constraints on implementations of Level 2 include:

- the rate of Level 1 triggers;
- the network latencies for message passing;
- the network bandwidths for data transport;
- the processing overhead incurred by the network protocol;
- the number of processors sharing the loads of filtering and event building;
- the execution times of their tasks; and
- the worst-case latencies for response to asynchronous messages (typically much different for workstations and embedded processors).

### Hardware

An example, based on two recent products, illustrates how a Level 2 trigger could be realized. Commercial reflective memory modules with interrupt capability have appeared in the PCI

Mezzanine Card (PMC) format. Reflective memories modules are connected together by a dedicated network. A message-passing mechanism, transparent to the user, keeps the contents of appropriately mapped locations in all modules the same. From the processors' point of view (PMC), they are accessing the same memory. VME PowerPC processors with PCI mezzanine slots are now on the market. This example uses several of these processors, each with a PMC reflective memory node. These appear as the readout controllers in the 20 drift chamber crates, the four vertex detector crates, and the Level 1 trigger crate, as ten Level 2 processors, and as the global controller. All reflective memory nodes form one logical network.

In response to the Level 1 accept signal, the global controller moves the Level 1 trigger primitives and time stamp into the next available buffer, then sends a message over the reflective memory network to the next available Level 2 processor, pointing to this buffer. The Level 2 processor then sends a request message to all the readout controllers on the reflective memory network for the Level 2 trigger symbols for the time interval specified by the stamp. Each readout controller then manages this request in its crate and responds, also over reflective memory. When all readout controllers have responded, the Level 2 processor completes the  $p_t$  and  $z$  cuts on Level 1 events.

Level 2 notifies the global controller of its decision and makes itself available for another event. For accepted events, the global controller notifies all readout modules, via their readout controllers, to save all data from the event, and then manages the remainder of the event processing in the same manner as in the baseline design.

### 10.11.3 Conclusion

A data acquisition system with a Level 2 filter would have greater architectural and data flow complexity than the baseline option. This may be offset by having a simpler Level 1 trigger and a reduced load on the event-assembly network and the processors. Furthermore, it may be required by the running conditions. We will continue to investigate its design, including prototyping the communication protocols and simulating the overall system, and we will keep its requirements in mind as we construct the event-assembly network. Simultaneously, we will track developments in networking technology, identify promising options, and select the most cost-effective low-risk implementation.

---

## 10.12 Global Support Electronics

---

### 10.12.1 Detector Monitoring and Control

#### Purpose

A primary tool for assuring the quality of data collected by *BABAR* is the careful control and monitoring of the detector hardware. This includes such traditional items as setting and monitoring power supplies and monitoring environmental factors such as temperature and humidity. It also includes monitoring to ensure the safety of both personnel and equipment, and communicating with the run control system to prevent it from taking actions which are not allowed by the current hardware status.

#### Design Strategy

The detector control includes areas of responsibility belonging both to the individual detector systems and to the central data acquisition group. The central group will provide a standard framework consisting of a set of displays, a set of procedures for setting and reading values from the hardware devices, and a messaging system for communication. The detector systems have the responsibility to identify their needs early in the design cycle and to provide the necessary interfaces to the standard hardware and software. The design process will use an object-oriented approach to abstract those parts of the implementation that are common to many systems, and from this, the central data acquisition group will provide a common framework. An important point is that the system designers think about these functions from the start so that the elements of the common framework can be identified as early as possible.

Among the quantities that will be routinely monitored for all detector systems are:

- all power supply voltages at the supply and as close to the point of use as practical;
- power supply currents where appropriate;
- power supply and crate temperatures;
- temperatures at sensitive points in the detector and in the electronics house;
- all gas pressures; and
- many additional items particular to individual detector systems.

The distribution of which functions are provided by software and which are provided by hardware will depend largely on the time-critical nature of that function. This will emerge as the design progresses.

### Implementation

The detector control system will be implemented using EPICS/VxWorks. This modern, modular approach frees the designer to choose the hardware technology which is most appropriate to the problem and most cost effective. For example, devices which communicate via GPIB, VME, RS232, or other means can all be accommodated easily by using the appropriate software driver. The effects of changing a hardware module are thus highly localized.

To eliminate interference with data gathering, the communication path for detector control will be kept separate from that used to collect events. An Ethernet-based system using the TCP/IP protocol will be used for the backbone. Processors on this backbone in turn communicate with hardware modules via the appropriate bus, *e.g.*, GPIB, VME, or RS232. These backbone processors will run VxWorks and are seen by EPICS as I/O Controllers (IOCs).

A number of software tools will be needed. These are discussed more fully in Chapter 11. Some of the tools needed for this task are:

- A set of standard displays such as strip chart plots, histograms, and status displays. These allow the operator to keep track of the hardware and are provided by the Display Manager in EPICS.
- A messaging system for communicating with the operator and the run control system. Messages will be classified according to hazard level as alarms which need immediate action, warnings, or information. Communication with the run control system will allow it to take appropriate automatic action for alarms and warnings and to allow it to check the status of appropriate hardware before allowing a change of the state of the experiment; *e.g.*, the high voltages must be on in order for data to be taken. The run control system is discussed more fully in Chapter 11.
- An interface to the online database, which is needed in order to store target values, tolerances, and measured values.

### 10.12.2 Data Monitoring Support

Monitoring is the continuous watching of the incoming data and the hardware to spot problems as soon as possible. It is a front-line tool to ensure data quality. Examples include synchronization checks, comparing histograms of raw data to expected distributions, and monitoring event sizes. Monitoring of the detector control hardware is discussed in Section 11.

Monitoring of the physics event stream is primarily a software task and is discussed in Chapter 11. An electronics implication is that data acquisition hardware may need to supply a buffer in which sampled events are stored in such a way that they do not interfere with the primary data gathering path. Individual monitoring processes could then access this data for analysis.

### 10.12.3 Diagnostic Support

#### Purpose

Diagnostic actions are taken to isolate a problem once it has been detected by monitoring or to perform routine checks to verify that the hardware is functioning properly. They involve operations that are not normally part of data taking. Examples include executing scope loops and measuring response to known stimuli such as test pulses.

#### Design

The choice of diagnostic operations for a particular detector system is best left to the designers of the system, but it is important that they address the issue early in the design process. Global support from the electronics and online groups will be in the form of a framework upon which the system designers can build. This framework will come from a global, object-oriented design of the needs of the systems and will consider hardware and software together. The design methodology will be iterative to allow information learned in late stages to be fed back into earlier stages. In many cases, the choice of using hardware or software for a given function is primarily a matter of the response time needed. By deferring this decision, we can optimize before making costly decisions.

## Implementation

Diagnostic runs can take place in either global or local mode. In global mode, there is central management for run control, trigger, data gathering, and timing. There can be only one such run in progress at a time. Global mode is used when the nature of the problem is such that two or more systems must run together. An example of a global diagnostic run is a pattern load of data into the system buffers to verify that the data are being read out correctly. In local mode, the system is in a partition that has its own local run control, local readout strobe, local data gathering, and local timing signals. This will be implemented by providing a local timing controller and a local readout controller for each partition. In global running, these units act under the control of a central partition. In local running, the timing controller can either pass on signals from the central partition, *e.g.*, clocks and synchronization pulses, or use locally generated ones. Similarly, the local readout controller can either pass the data on to the central partition or handle it locally. Many partitions can be running independently and concurrently, and within a given partition, it may be possible to have multiple concurrent activities. The local partitions must have access to the central database.

### 10.12.4 Calibration Support

#### Purpose

Calibration involves determining the parameters needed to turn digitized data like ADC and TDC counts into physical data such as energies and coordinates. This involves applying known stimuli to the analog portion of the data acquisition chain and observing the response through the digitization process.

#### Design and Implementation

In terms of data acquisition properties, calibration runs are very similar to diagnostic runs. As such, the general framework for detector diagnostics described above applies. The choices of how to calibrate and how often to do it are primarily up to each detector system and are discussed in their respective sections of this chapter.

Some general principles of calibration runs are: they should use the same analog and digitization electronics as the physics data; they should strive to measure one number instead of piecing together multiple measurements; and the absolute calibration may in many cases have to come from the physics data, *e.g.*, getting the absolute energy scale for the calorimeter from the Bhabha events.

Calibration runs can be in either global or local mode. An example of a global calibration run is using cosmic rays to do tracking alignments and to measure the  $dE/dx$  values from muons. An example of a local calibration run would be using calibrated electronics to inject known signals into the detector data acquisition system.

A special case of calibration involves automatic steps that need to be taken at specific times; *e.g.*, pedestal measurements which are taken at the start of each run. These processes will be fully automated.

---

## References

---

- [Amb94] M. Ambrosio *et al.*, “Performances of RPC Detectors with Tracking & Timing Electronics,” *Nucl. Instr. Methods* **A344**, 350 (1994).
- [Alo94a] A. Aloisio *et al.*, “High Density Zero Suppressor Card for the Muon Trigger with RPC for the L3 Experiment,” *IEEE-NS* **41**, 225 (1994).
- [Alo94b] A. Aloisio *et al.*, “The Suppressor Card for the Muon Trigger with RPC for the L3 Experiment,” INFN-TC-94-11 (1994).
- [Alo94c] A. Aloisio *et al.*, “The RPC Trigger System for the L3 Forward-Backward Muon Detector,” in Proceedings of the CHEP94 Conference, San Francisco, California, LBL-35822 (1994).
- [Alo94d] A. Aloisio *et al.*, “The L3 Forward-Backward Muon RPC Trigger System,” submitted to *Nucl. Instr. Methods* (1994).
- [Alo94e] A. Aloisio *et al.*, “Performances of the RPC Trigger System of the L3 Forward-Backward Muon Spectrometer,” submitted to *Nucl. Phys. B* (1994).
- [Bab94] *Letter of Intent*, SLAC-443 (1994).
- [Bag94] R. Baggs *et al.*, *Nucl. Instr. Methods* **A344**, 547-557 (1994).
- [Bau92] D. Bauer, “The Impact of Two-photon Physics on a *B* Factory Detector,” *BABAR Note # 77* (1994).
- [Bol91] T. Bolton *et al.*, “Report of the Trigger and Data Acquisition Group,” in Proceedings of the Workshop on Physics and Detector Issues for a High-luminosity Asymmetric *B* Factory at SLAC, SLAC-373, 597 (1991).
- [Boz93] W. Bozzoli *et al.*, “Data Transfer and Distribution at 70 Mbytes/s,” CERN/ECP 93-7, (1993).
- [Buo] S. Buono *et al.*, “Studies on RIO/HiPPI Based Event Building,” RD13 Note N.70 (1993).
- [CLE91] C. Bebek *et al.* (CLEO Collaboration), “CLEO-II Trigger System,” *Nucl. Instr. Methods* **A302**, 261-276 (1991).
- [Cul91] D. Cullen-Vidal *et al.*, “D0 Level-2/Data Acquisition; the New Generation,” in Proceedings of the 1991 International Conference on Computing in High-Energy Physics, 659 (1992).

- [Dol82] D. Dolev, M. Klawe, and M. Rodeh, "An  $O(n \log n)$  Unidirectional Distributed Algorithm for Extrema Finding in a Circle," *J. Algorithms* **3**, 245–260 (1982).
- [Dou94] D. Doughty Jr., D. Game, L. Mitchell, G. Heyes, and W.A. Watson III, "Event Building Using an ATM Switching Network in the CLAS Detector at CEBAF," Presented at the International Data Acquisition Conference on Event Building and Event Data Readout, Batavia, Illinois (1994).
- [Eis90] A. Eisner, "Bhabhas at an Asymmetric  $B$  Factory," *BABAR Note # 32* (1990).
- [Ell92] N. Ellis *et al.*, "A Calorimeter-Based Level-One Electromagnetic Cluster Trigger for LHC," in Proceedings of the 1992 International Conference on Computing in High-Energy Physics, edited by C. Verkerk and W. Wojcik (Annecy, France, 1992), 210, (1992).
- [Gen93] P. Bailly and J.-F. Genat, "A 100 Picosecond Resolution, 6 Microsecond Full Scale Multihit Time Encoder in CMOS Technology," in Proceedings of the Third Int. Conf. on Electronics for Future Colliders, Chestnut Ridge, p. 57 (1993). See also: J.-F. Genat, "High Resolution Digital Time to Digital Converters," in Proceedings of the First Int. Conf. on Electronics for Future Colliders, Chestnut Ridge (1991).
- [Hal94] G. Haller, D. Freytag, and J. Hoefflich, "Proposal for an Electronics System for the *BABAR* CsI Calorimeter," *BABAR Note # 184* (1994).
- [Kin89] K. Kinoshita, "A Fast Hardware Trigger for Charged Particles in Multilayer Tracking Devices," *Nucl. Instr. Methods* **A276**, 242 (1989).
- [Kra94a] F. Kral, "GEANT Simulation Results for a Fast Calorimeter Trigger," *BABAR Note # 133* (1994).
- [Kra94b] F. Kral and D. Worledge, "Initial Design and GEANT Simulation of a Fast Drift Chamber Trigger," *BABAR Note # 162* (1994).
- [Kra94c] F. Kral, "Simulation of Trigger Rates due to Beam-Gas Backgrounds," *BABAR Note # 165* (1994).
- [Kra94d] F. Kral, "Trigger Rates due to Cosmic Ray Muons," *BABAR Note # 164* (1994).
- [Lam74] L. Lamport, "A New Solution of Dijkstra's Concurrent Programming Problem," *Comm. ACM* **17**(8), 453 (1974).
- [Lam78] L. Lamport, "Time, Clocks, and the Ordering of Events in a Distributed System," *Comm. ACM* **21**(7), 558 (1978).

- [Lev94a] M. Levi and F. Kral, “Simulated Background Rates in Calorimeter Trigger Towers,” *BABAR Note # 134* (1994).
- [Lev94b] M. Levi, “Impact of Backgrounds on Silicon Vertex Detector Architecture and Detector Trigger,” *BABAR Note # 136* (1994).
- [Lev94c] M. Levi, “Obituary for the Prompt Silicon Trigger,” Netnews Item *slac.b-factory.vertex #54*, (1994).
- [Lev94d] M. Levi, “Efficacy of a Level 2 Trigger Algorithm,” Netnews Item *slac.b-factory.vertex #55*, (1994).
- [Lev94e] M. Levi, “Impact of the Electroproduction of Hadrons on the Trigger,” *BABAR Note # 191* (1994).
- [Pat94] J. Patrick *et al.*, “The CDF Ultranet-Based Data Acquisition System,” in Proceedings of the 1994 Int. Conference on Computing in High-Energy Physics, in preparation.
- [Pet82] G.L. Peterson, “An  $\mathcal{O}(n \log n)$  Unidirectional Algorithm for the Circular Extrema Problem,” *ACM Trans. on Programming, Languages and Systems* **4(4)**, 758 (1982).
- [Req94] “*BABAR* Electronics Requirements,” in preparation.
- [Ric81] G. Richart, and A.K. Agrawala, “An Optimal Algorithm for Mutual Exclusion in Computer Networks,” *Comm. ACM* **24(1)**, 9, (1981).
- [SLA93] “Status Report on the Design of a Detector for the Study of *CP* Violation at PEP-II at SLAC,” SLAC-419 (1993).
- [Sny92] A. Snyder, “Backgrounds in the *BABAR* Drift Chamber,” *BABAR Note # 88* (1992).
- [Sny94] A. Snyder, “Electro-Production Triggers in *BABAR*,” *BABAR Note # 180* (1994).
- [Wun94] S. Wunduke, M. Levi, and F. Kral, “Architecture for Readout Electronics,” *BABAR Note # 138* (1994).



---

---

# Computing

There are a number of computing challenges for the proposed *BABAR* detector. The data rates are high for an  $e^+e^-$  collider, and there is a large quantity of useful data which must be stored and analyzed. Convenient interactive access must be available for the efficient extraction of physics results. In addition, the wide geographical distribution of the collaboration implies significant issues for remote access to the data and code, collaboration communication, and for remote computing.

The *BABAR* Computing System, as defined here, includes several aspects: (1) the online data acquisition, control, and monitoring hardware and software; (2) the offline computing hardware—CPUs (“farm” and desktop), storage, X servers, and networking—whether local to SLAC or at remote locations; (3) the software environment in which the computing work is done; and (4) the code itself used by the collaboration. All of these are important items for *BABAR* computing, although parts of these, including most of (2), are properly considered institutional infrastructure.

The next section reviews the computing requirements, and Section 11.2 gives an overview of the proposed computing system. Following these are more detailed discussions of the computing model, software environment, online system, reconstruction and analysis framework, computing support plan, integration issues, system responsibilities and management, and cost and schedule.

## 11.1 Requirements

---

In this section, the basic computing requirements for *BABAR* are summarized. These are divided into operational requirements and technical requirements. Operational requirements are broad capabilities that must be present in order for the system to respond adequately to the environment within which it must operate. Technical requirements are specific capabilities that the system must exhibit in order to meet the operational requirements, for example, the CPU power and bandwidth. Estimates are given for the early years of the experiment. It is recognized that the requirements may increase as the experiment progresses; hence, a general further requirement exists that the system be scalable to meet future needs.

### 11.1.1 Operational Requirements

The *BABAR* Computing System must support several activities which compete for available resources. These include program development by a distributed developer community, data acquisition, control and monitoring, bulk data reduction and Monte Carlo simulations, and interactive physics analysis.

As the collaboration is worldwide, a working scheme which in no way reduces the ability of the remote collaborators to contribute to the experiment must be implemented. The remote computing issues cover the way in which the collaborators exchange information (at the moment WWW, ftp, netnews, and mailing lists), how software development is integrated, and how to implement updates of software production releases, data analysis, Monte Carlo production, database access, and software licensing. Access to the data will be provided at regional centers, which will keep some terabytes of data.

The computing system must provide a real-time response adequate to perform its many roles. This impacts several areas. The most obvious is that of data acquisition. To benefit fully from the available luminosity at PEP-II, it is important to reduce as much as possible the effective deadtime for the online system. The time for routine operations that can interfere with data taking, such as run start and close-down or system reboot, should be small. Finally, the latency between the acquisition of the data and its analysis for physics results should be short compared to the duration of an accelerator scheduling cycle.

The computing system should display a uniform face to the users that hides any intrinsic heterogeneity and allows them to be productive at their home institutions, at regional centers, or at SLAC. It must be robust, both in the sense of exhibiting a high degree of reliability and high availability, and also in allowing short recovery time from a failure. This in turn implies a high degree of maintainability. Given the extended time over which the experiment will be active and the probability of unforeseen demands and changes in the requirements, the computing system must also be highly flexible and extensible, this being achieved with a minimum of effort and in time intervals dictated by the accelerator scheduling cycle. Finally, the computer system must balance ease of use with security so that developers can be productive, and also ensure that high quality, well-understood, and reproducible physics results can be achieved.

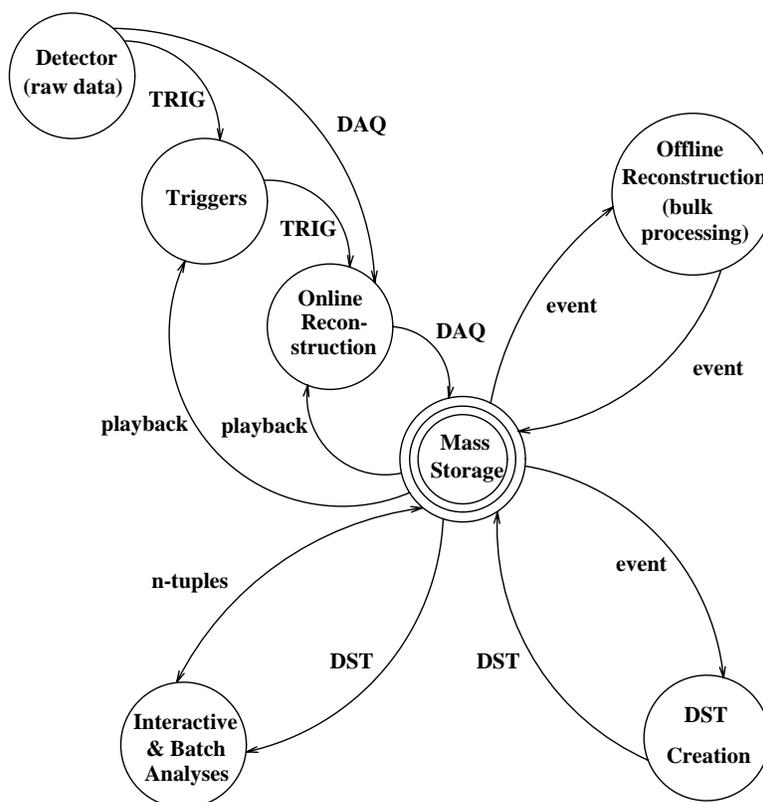


Figure 11-1. Schematic of the data flow for the BABAR experiment.

## 11.1.2 Technical Requirements

### CPU Requirements

There are several substantial computing tasks that require significant CPU resources: the online system, including the Level 3 trigger and data acquisition, event sampling, detector monitoring, calibration, and control; offline reconstruction; Monte Carlo simulation; creation of DSTs from the reconstructed data, involving a reduction in event size and/or a reduction in the number of events; analysis of large DST datasets; and interactive analysis tasks. These tasks are represented in a data-flow diagram in Figure 11-1. The CPU power<sup>1</sup> required to handle this data flow is summarized in Table 11-1.

<sup>1</sup>The familiar term MIPS is used to describe CPU processing power. In terms of the current industry standard specifications, one MIPS represents a power corresponding to approximately one SPECint92 (integer performance) and one SPECfp92 (floating point performance).

Online (Level 3)	MIPS	3000
Reconstruction	MIPS·s/event	25
	events/year	$10^9$
	passes/year	2
	MIPS	5000
Monte Carlo	MIPS·s/event	500
	events/year	$10^8$
	MIPS	5000
DST Creation	MIPS	500
DST Analysis	MIPS	2000
Interactive Analysis	MIPS	5000

**Table 11-1.** *Estimated CPU power requirements.*

The online CPU requirement is estimated assuming a 2000 events/s input rate to the Level 3 farm. This rate is higher than the average anticipated rate, but is used to ensure that the system is able to keep up with the data flow under extreme conditions. This is quickly reduced to the order of 200 events/s using  $p_t$  and  $z$  information from the trigger. More complete event information is then used to reduce the rate to tape to 100 events/s. A full reconstruction will be done on approximately 30 events/s (of which  $\sim 10$  events/s are hadronic via single photon) at  $\mathcal{L} = 3 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ .

Based in part on the CLEO-II experience [CLE94], similar figures are derived for the CPU requirements relevant to event reconstruction and Monte Carlo simulation (5000 MIPS each). The CPU requirement for Monte Carlo simulation, as for the data acquisition and reconstruction, is based on a peak rate requirement, namely the ability to keep up with the peak hadronic event data rate. This corresponds to a rate of  $10^8$  Monte Carlo events per  $10^7$  seconds.

DST creation refers to a major pass through the data to create a selected dataset (in which the output event size may or may not be reduced. The generic term DST is used here to include differing definitions of DST, mini-DST, *etc.*). The CPU power required is computed according to:

$$\text{MIPS} = 0.25(\text{MIPS}\cdot\text{s}/\text{event}) \times 10^9(\text{events}) \times 2(\text{users})/10^6(\text{s}),$$

which assumes that it is acceptable to have this task take less than a month, hence the  $10^6$  second time period; the two users allow for up to two such major passes through the data to be simultaneous. This estimate, of course, is a particular, and rather conventional, model for how an analysis progresses into selected subsets of the data. Alternatives will be investigated, such as using a direct access database technology [Bad90].

DST analysis is a substantial processing of a selected DST dataset for physics analysis. The formula used to estimate the CPU power in that case is

$$\text{MIPS} = 1(\text{MIPS}\cdot\text{s}/\text{event}) \times 10^8(\text{events}) \times 20(\text{users})/10^6(\text{s}).$$

This assumes 20 such analyses occur once per year, each requiring one month of CPU time to process the full dataset of interest. This estimate is in addition to the resources required to design and debug such analyses.

Interactive analysis involves debugging, exploratory analysis, graphics, and program development. One hundred active users, each employing a 50 MIPS workstation, are anticipated.

In summary, if the numbers above are summed, the total CPU requirement is 20,000 MIPS. At least 10,000 MIPS have to be located at SLAC for data acquisition, reconstruction, and other tasks. However, the total installed MIPS may be lower depending on the model (Section 11.2.2); the requirements are based on keeping up with the data acquisition, effectively realtime, at the peak luminosity rate, rather than at the rate averaged over a year.

## Throughput

Transformation of the raw data into the final reconstructed event information, and subsequent physics analysis, requires several data transfer steps (Figure 11-1). The most significant steps (in terms of required bandwidth) include:

- Transfer of data into the Level 3 online farm;
- Transfer of data from the Level 3 farm to mass storage (tape);
- Reading and writing the data in the bulk reconstruction;
- Writing of Monte Carlo events to tape;
- Reading through the data to create DSTs;
- Passes through the DSTs for physics analysis; and
- Interactive analysis on small datasets.

The estimated requirements are summarized in Table 11-2. The upper-limit requirement of 2000 events/s is assumed going into the Level 3 data acquisition farm, and the event size is estimated to be 25 kbytes [Por93], yielding a 50 Mbyte/s rate. The event size is assumed to double to 50 kbytes once the reconstruction information is added, but the rate decreases to 100 events/s. Hence, the reconstruction data rate is 7.5 Mbytes/s, 2.5 Mbytes/s input plus 5 Mbytes/s output. The Monte Carlo event size is also estimated at 50 kbytes, which, at the simulation rate of 10 events/s, implies a data rate of 0.5 Mbyte/s.

Task	Input (events/s)	Output (events/s)	Bandwidth (Mbytes/s)
DAQ to Level 3		2000	50
Level 3 to Tape		100	2.5
Reconstruction	100	100	7.5
Monte Carlo	10	10	0.5
DST Creation			$50(1 + f)$
DST Analysis			5/task
Interactive Analysis			0.2/user
Wide Area 80× Sessions Data Transfer			> 0.5 0.5

**Table 11-2.** *Estimated bandwidth requirements. The  $f$  factor is explained in the text.*

DST creation and analysis represent a large network load. Substantial DST creation may occur simultaneously with the reconstruction via multiple output streams. However, additional passes through the data should be anticipated here. In DST creation,  $f$  is the product of the sample fraction times the average sampled event size divided by 50 kbytes. If  $f$  is small, and two such tasks are running, 100 Mbytes/s is required. The task of DST creation will typically be heavily I/O bound. DST analysis, with 20 analysis tasks simultaneously running, requires 100 Mbytes/s.

The interactive analysis number is estimated assuming a 2 kbytes/event reduced DST, and a throughput of 1000 events in ten seconds, appropriate for small datasets which are assumed to be analyzed at a desktop workstation. There may be an occasional need for interactive analyses on larger DSTs, with larger bandwidth requirements. It is anticipated that such analyses would take place on the same machines as the DST analysis tasks, supported with higher bandwidth.

Based on this analysis, aggregate network capacity for offline tasks in excess of 200 Mbytes/s is required, dominated by the reconstruction and the multiple reads through the data required for DST creation and analysis. However, the DST creation task, which is a potentially large burden specifically on tape access, may be greatly mitigated by having this process proceed largely via multiple output streams during reconstruction.

The wide-area network will be used for code development, interactive analysis, user communications, and transfer of small datasets (less than 10 Gbytes). A rule of thumb is that an X Window session requires 50 kbits/s. Thus, supporting an estimated 80 sessions nominally requires 0.5 Mbyte/s, except that the scaling is sublinear with number of sessions over a

Tape Storage:	Tbytes/yr
Raw Data	25
Reconstruction Output	50
Monte Carlo	5
DSTs	5

Disk Storage:	Tbytes/yr
DSTs	2
Databases	0.03

**Table 11-3.** *Estimated storage requirements.*

network. To transfer a small dataset in an acceptable time (a few hours) requires a similar bandwidth.

## Storage

Table 11-3 summarizes the estimated storage requirements for the *BABAR* experiment. Approximately 100 Tbytes/yr of tape storage is required. If necessary, the 25 Tbytes of raw data could be moved offline after reconstruction. The 5 Tbytes/yr for DST storage is estimated by assuming 10 DSTs produced per year, with  $10^7$  events in each (sampling fraction of 0.01) and 50 kbytes/event.

Current experience at SLAC is that a (disk) staging space of 2% of the active tape storage is required. Alternatively, 2 Tbytes may be sufficient for 40% of the DSTs produced to be on disk. However, it is probably appropriate to regard this space as staging space, with the details of whether a particular job's data are already staged to disk transparent to the user. The database space requirement is implementation dependent but should be small compared with the event data requirement.

Disk space is also required for group code (and documentation, *etc.*) and for user space. The group code space will be relatively small, probably in the tens of gigabytes. The user space required will also be comparatively minor but cannot be neglected because of the large number of users. From developing experience, it appears that *at least* 100 Mbytes/user of disk space will be needed. If there are 300 users, this means at least 30 Gbytes.

## 11.2 Overview

---

This section gives a summary of the basic design choices which have been made and an overall description of the planned *BABAR* computing system.

### 11.2.1 Chosen Technologies

Choices for key computing technologies depend largely on capability and economics. Additional technology choices are then driven by those initial selections. Finally, some technologies are selected reflecting a judgment of how best to solve a particular problem. Given that the computing industry has become highly competitive and that technologies will evolve quickly on the timescale of the *BABAR* experiment, it is important to identify available and flexible solutions whenever possible. Thus, some of the initial selections may be considered working decisions; the initial design is expected to evolve between this design and the final implementation and final uses (a period of approximately 15 years).

The magnitude of the computing need overwhelms the traditional mainframe approach and dictates the use of a distributed, multiprocessor environment. In the following discussion, it is assumed that all computing, both online and offline, will share a common environment. It is recognized that to accommodate the diversity within a widespread collaboration and changes occurring within the computing marketplace, special care must be taken to plan for multiple platform support. This is both a hardware and a software issue; thus, generic tools which are platform independent have been and continue to be sought. At present, this implies the use of workstations utilizing inexpensive (RISC) technology and operating system software within the UNIX family. While it is too early to decide upon a specific networking medium, the industry-standard Transmission Control Protocol/Internet Protocol (TCP/IP) is expected to survive and will, therefore, play a major role both in local-area and wide-area network communications. The use of an object-oriented (OO) approach will be encouraged for all software development within the collaboration. The primary programming language will be C++. Fortran 90 will be an acceptable alternative. Windowing technology will be the X11 client-server system with the Motif library and window manager. Printing will support both text and the PostScript page description language. The management of program code releases will be, in part, based upon the Concurrent Versions System (CVS). World Wide Web (WWW) will serve as the primary mechanism for information access. The online system will be built around the Experimental Physics and Industrial Control System (EPICS) (Section 11.5.5). This, in turn, relies upon the VxWorks real-time executive; although it is

expected that this requirement will be relaxed within a few years as EPICS becomes POSIX<sup>2</sup> compliant.

## 11.2.2 Overall Description

### Computing Model

The *BABAR* computing model is based on systems of UNIX processors. Detailed choices will be deferred until need actually arises in order to take advantage of the rapid developments in computing technology. With the technology of today, a satisfactory solution can be built with a combination of Ethernet and FDDI local networking with nonblocking Ethernet-FDDI hubs and FDDI switches, SCSI-2 disk farms, and helical-scan tape technology in existing tape silos. A small number of remote computing centers, based on similar technology choices, are anticipated, with wide-area communications achieved by a combination of Internet connections and 3490 silo cartridge or other tape (*e.g.*, 8 mm) transfers.

### Software Environment

A consistent, robust, and easy-to-maintain software environment will be created by resorting to common facilities for writing, managing, and distributing code relevant to all the computing tasks. Recent advances in software engineering, such as the object-oriented methodology and programming languages, will be exploited. An approach has been chosen which includes the rapid creation of guidelines and standards documents, code templates, and sample programs, together with the establishment of a code development environment that supports the needs of the software developers while also ensuring that the end users' need for stability is also addressed.

### Online System

The online system has not only to support the steady state operation of the detector for physics data acquisition, but also has to deal with the conflicting demands of detector commissioning, calibration, and diagnostics. A system that accommodates these demands is described in terms of multiple viewpoints, from the most abstract which deals with the need to control multiple logical experiments or "partitions" during commissioning, to the most concrete which deals with the detailed control protocols among modules within the data acquisition system.

---

<sup>2</sup>IEEE Portable Operating System Interface for Computing Environments.

## Reconstruction and Analysis Framework

A common framework for applications is being developed that will allow physicist-developed code to be used for online triggering and monitoring, for bulk reconstruction and Monte Carlo simulations, and for physics analysis. This framework will shield the application programs from the different operating conditions in these areas and allow single applications to operate in all areas, accepting event data from real-time event servers in the online area, or from disk or tape files offline, and supporting output of analyzed data to a variety of output destinations including the real-time event server and disk or tape files. This framework must furthermore provide for both interactive and batch styles of operation and accommodate the conflicting needs of experts and novices. Both text-based and graphics-based user interfaces are under development in order to accommodate these needs, addressing also the needs for bulk job submission and bookkeeping.

## Computing Support

Computing support refers to both internal (within the collaboration) and external (university and laboratory support organizations) activities. Internally, this support will be coordinated directly by the computing system group. Designated members (Code Coordinators) of the collaboration will be assigned to look after the various software packages, to manage the data production and distribution tasks, to maintain the different databases, and to provide system management support for the online computers. The collaboration will see that appropriate documentation is generated, training needs are identified and covered, and that specific tools and utilities for the experiment are written. These support roles are expected to consume a significant amount of manpower within the collaboration.

External support will be provided by support staff at the collaborating institutions and is expected to include a wide variety of services. The overall computing system must be designed with the active participation and cooperation of these groups. Hardware elements must be planned, purchased, installed, operated, and maintained. This includes networking, mass storage, and CPUs. These heterogeneous and distributed computer systems must be managed with respect to user accounts, disk and other resources, backups, system tuning, *etc.* Wide-area networking operations will require continual monitoring with occasional attention. Software must be licensed, installed, and maintained. The details of data storage and access must be designed and implemented. A system for efficiently utilizing distributed computing cycles must be installed and maintained. Finally, the support staff must spend some fraction of their time in system analysis, technology tracking, and strategic planning.

---

## 11.3 Computing Model

---

A comprehensive computing model must satisfy the functional requirements outlined in Section 11.1. In addition, such a model must address practical considerations, for example: the “abilities” (usability, flexibility, scalability, reliability, maintainability, and affordability); smooth integration of all hardware and software components; ability to evolve in response to commercial developments; and compatibility of SLAC on-site computing with that of off-site collaborators. In the sections that follow, an architecture which addresses these requirements, a model to describe its overall operation, and a cost model based upon present and near-future technologies are presented.

### 11.3.1 Architectural Model

The overall computing system architecture includes a number of basic hardware and software components. Necessary hardware components include the elements of processing power (*e.g.*, computers), networking, and mass storage. The main software component is the operating system. A schematic model containing these components is given in Figure 11-2.

#### CPU Complex

The *BABAR* Collaboration has concluded that, currently, the most cost-effective way to satisfy the basic CPU requirements is through the use of multiple microprocessor-based computers. Microprocessors utilizing a RISC architecture, as found in UNIX workstations, represent the optimum in today’s market. A continued high rate of growth in computing power of this kind is expected for at least the next five years. Alternative computing architectures are not excluded in this model, and the possibility of a transition sometime during the period of this experiment has been considered.

A commitment to clusters of distributed workstations for an online system, reconstruction, Monte Carlo simulation, and interactive data analysis, along with general software development and debugging, requires a sophisticated and mature, multitasking, networked operating system. This requirement is met by using UNIX, offered by all major workstation vendors. Such a solution also tends to capitalize on relatively high-volume, inexpensive, components (*e.g.*, video monitors, SCSI disks, and RAM), thus offering a degree of vendor independence both for hardware and software.

Commercial development of workstation-based computing is proceeding in several directions. For example, workstation clusters currently tend to use single-processor implementations.

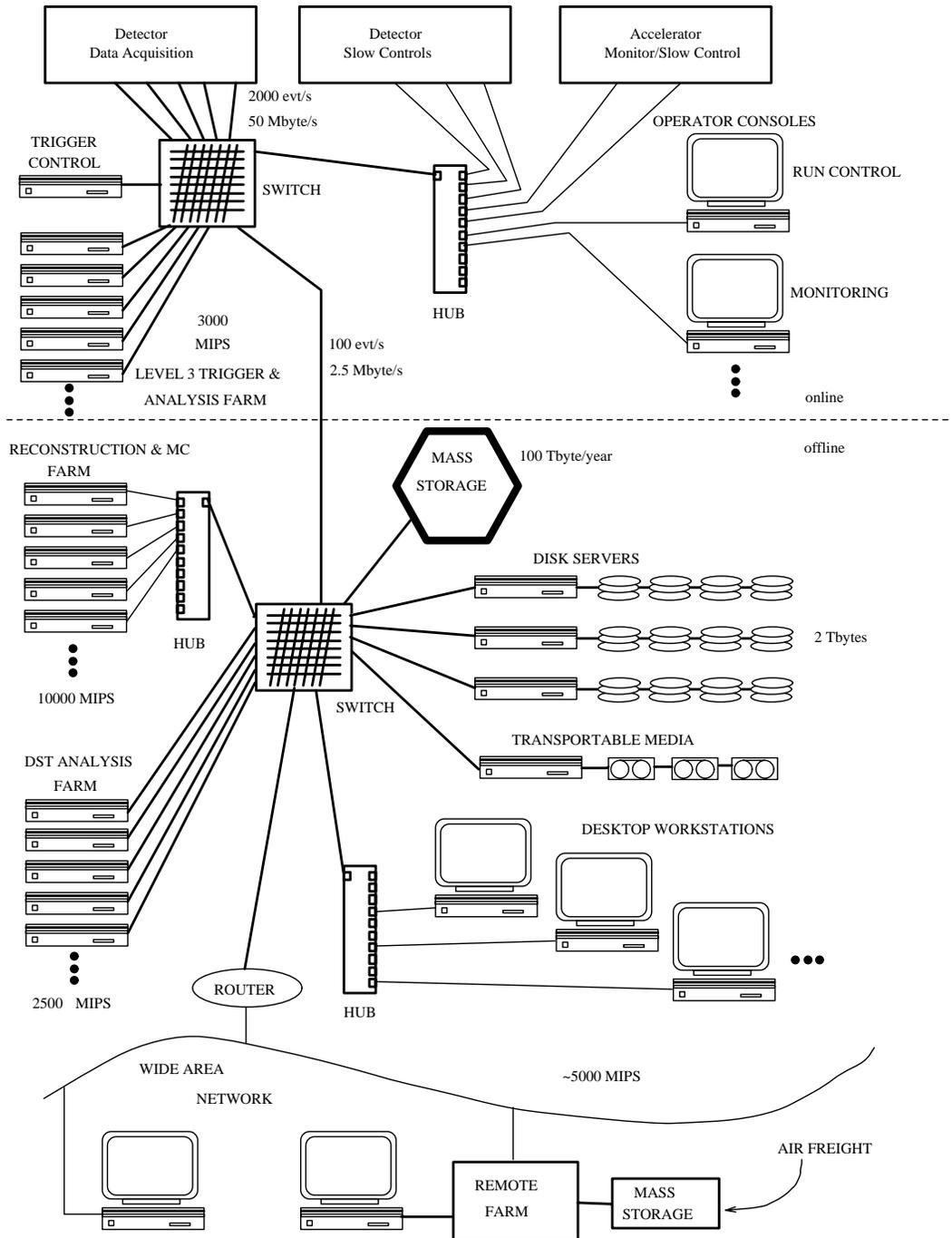


Figure 11-2. Conceptual computing architecture.

There is a trend towards multiprocessor cluster implementations that improves processing performance by virtue of a high-speed backplane interconnecting the individual CPUs. Current examples of this architecture include the SGI Power Challenge, the IBM SP-2, and the Convex Exemplar. Such multiprocessor implementations could be used to optimize the design, but only if the costs become competitive.

The UNIX operating system is also undergoing a series of changes. The direction of change appears to be toward standardization of both the basic command set and system libraries. By carefully adhering to such standards (*e.g.*, POSIX), one can attempt to minimize the work required to make a transition to another computer environment should that become necessary. Given the rapid pace of developments within the computing industry, the effort to abide by such standards within the collaboration is deemed essential.

### Networking

There are two levels at which networking is crucial to this experiment: local networking at SLAC, because of the large number of machines and high data rates involved both online and offline; and wide area networking, due to the large, widespread international collaboration. Local area networking refers to the links between the various components of the experiment, including those to desktop machines and the central computing complex. Wide area networking refers to those links between the various collaborating institutions and to SLAC. While the SLAC LAN is completely under lab control, the WAN is not. *BaBar* may hope to influence various funding agencies to accommodate WAN needs but cannot expect to sponsor significant upgrades due to the high costs, as shall be discussed later.

A networking system consists of various components and subsystems, including: workstation and other device interfaces; switches and hubs; various diagnostic devices; and the cable plant. The basic LAN architecture consists of a switched network of point-to-point links. In some cases, the I/O demands on a particular system may allow point-to-point links to be replaced by a short series of daisy-chained devices, or for a relatively expensive technology to be replaced by a cheaper (and lower performance) technology. The latter concept is illustrated in Figure 11-2 in the distinction between the Reconstruction & MC farm, characterized by a relatively high CPU:I/O ratio, and the DST Analysis farm, in which condensed data is read at a high rate with relatively little processing per event. The networking protocol will be TCP/IP, an industry standard available on all UNIX machines.

### Mass Storage

A large amount of reliable, inexpensive storage with fast, automated access from the UNIX environment is needed. Mass storage technologies, like those in other areas of computing,

are rapidly developing within the industry. Current mass storage candidates are typically expensive and unique to a particular vendor. While it seems unlikely that on *BABAR*'s timescale a *de facto* industry standard will emerge, several near-future products are expected to address these needs satisfactorily. Unlike the computing and network environments, however, the mass storage system is likely to be rather specialized.

The transfer of data between SLAC and remote collaborator sites may be complicated if the supported media at the various institutions are incompatible. For transfers to such institutions, an affordable and transportable data medium will be essential. Currently, 8 mm Exabyte tape cartridges are supported at SLAC and are in widespread use, although alternative formats (*e.g.*, 4 mm and DLT) are gaining in popularity. In general, capacity and performance are increasing while the cost per stored byte is decreasing. Hence, the decision as to which technology to use will be postponed until it is needed.

### Prototype Compute Farm

SLAC has been pursuing R&D towards this model in the form of a prototype compute farm, shown in Figure 11-3, designed to exercise these ideas. Two batch environments are available for testing: LoadLeveler from IBM and Load Sharing Facility from Platform Computing. SLAC Computing Services (SCS) is working with CERN on the latest data staging software being developed as part of CERN's SHIFT. It is anticipated that in the next several years, there will be investigations into commercial products which support the IEEE Mass Storage Model to provide data staging and hierarchical storage management.

## 11.3.2 Operational Model

### Hierarchy

A hierarchy of sites exists in the computing for the *BABAR* experiment. The data will be recorded at SLAC, giving SLAC a unique position. Demands of detector debugging and monitoring, physics analysis by collaborators resident at SLAC, and simplifications in data handling will mean that much, if not all, of the bulk data processing will be performed at SLAC. Next, there exists, within the collaboration several sites with the potential to provide computing facilities (data storage, computing capacity, and support) similar to those at SLAC. These regional centers will serve as analysis centers, providing convenient data access to institutes with poor network connections to SLAC and to those which cannot provide the levels of support required to maintain copies of the data locally. These centers could also be used for data processing, DST production, *etc.*, as performed at SLAC. Finally, there are the computing facilities in all of the other collaborating institutions. This final grouping

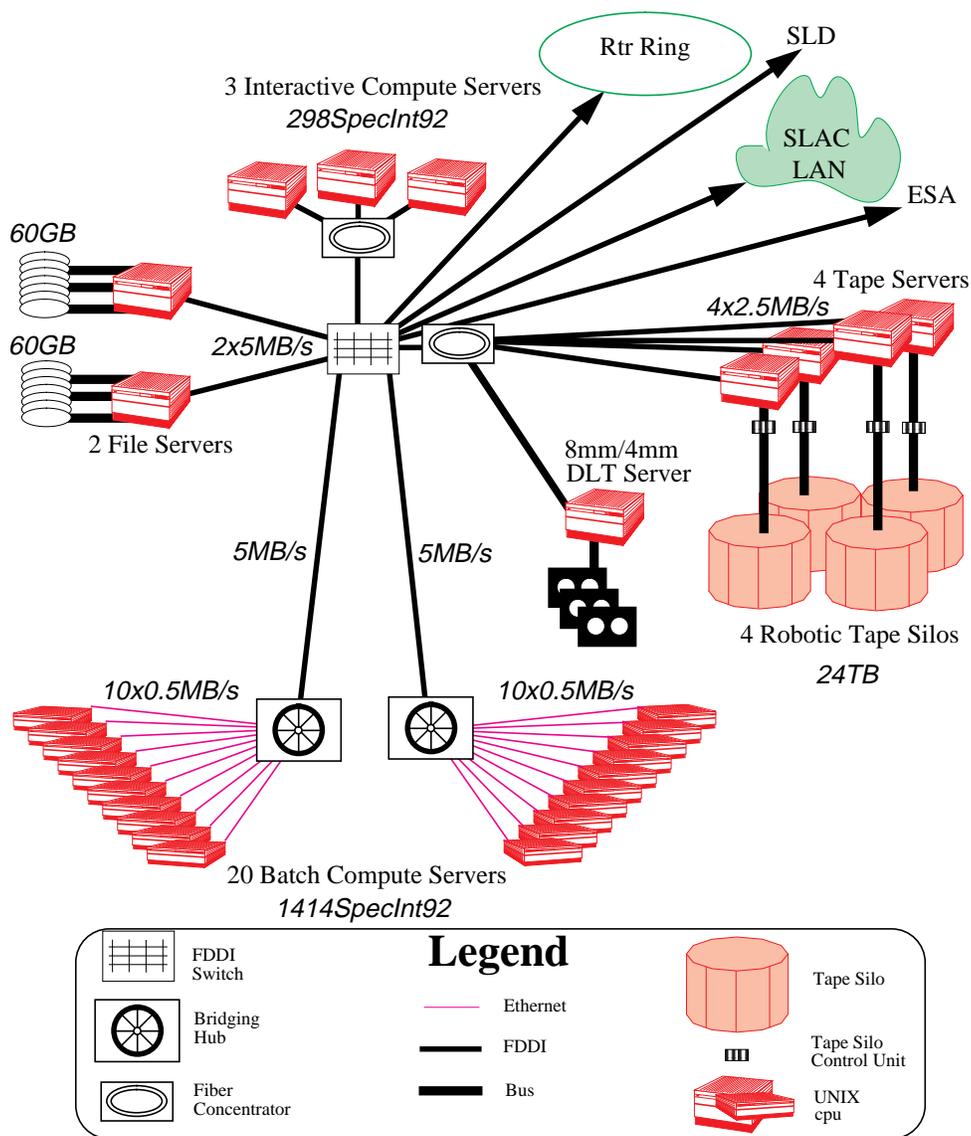


Figure 11-3. Prototype UNIX compute farm at SLAC.

of home institutions obviously includes a wide range of capabilities. They are characterized here, however, by the fact that they will not be used to perform bulk data processing, and are unlikely to want copies of significant fractions of the raw or processed data. It is envisaged that data analysis in the home institutions will be based (in terms of access to bulk data) around the data store at SLAC or a regional center, whichever is most convenient. Home institutions may want copies of DSTs. These again could come from SLAC or a regional center. It is quite possible that some of these institutions will contribute to Monte Carlo generation.

### SLAC Site

It is essential to have only one master copy of production software and databases. The SLAC site is an obvious choice for these, especially during the installation and early running of the experiment when updates and changes may be required frequently and quickly. This model of organization has already been implemented by *BABAR* for software development. The use of CVS/rCVS places central control of the master copies of all code at SLAC while allowing users at remote sites access to this code with a mechanism for including changes back into the master. The user's view of this process is essentially the same whether they are based at SLAC or at a collaborating institution.

The benefits of performing bulk processing at the site of the experiment have been proven many times. Most large experiments now do this task routinely (ALEPH, DELPHI, OPAL, L3, H1, ZEUS, CDF, D0). Though other models do exist (E791), the quasi-online processing possible with local facilities is an efficient use of resources and has significant advantages in producing results more quickly. Clearly, the ability to perform bulk data processing is required at SLAC. A batch system is needed to facilitate this processing.

The presence of the raw and processed data, as well as subsystem experts, at SLAC makes the SLAC site an obvious place for data analysis. To make the most effective use of this requires that an efficient analysis framework be maintained at SLAC, in particular, one which allows users not actually based at SLAC simple and efficient access. This system should not require that remote users log on to the SLAC computers. Remote users should have effective access to SLAC's batch system allowing creation of small DSTs or n-tuples which can easily be transferred back to the home institution for interactive analysis. Current WAN technology allows n-tuples of many tens of megabytes to be efficiently transferred in this way. For larger datasets, high-density tape technology provides a method of transport. The ability to ship data in this way requires additional copying facilities at SLAC.

## Regional Centers

Support for regional centers has significant manpower and cost implications both at the regional centers and at SLAC. However, there are distinct advantages to having access to a local data store. The regional centers will serve as data store, processing, and analysis centers. A *complete* data sample and support structure could be maintained at these sites. Regional centers will help affiliated remote users by providing easy access to the data and enabling them to build up the strength of a community.

The concentration of computing and data handling capabilities found at regional computing centers also makes them suitable for large scale Monte Carlo productions. Experience has shown that this is an effective use of these centers and an efficient use of collaboration resources. This also fits well with the regional center's role in general computing support for smaller collaborating institutions.

Regional centers already act as concentrations of support effort in *BABAR*. For example, rather than many institutions independently attempting to keep all software and libraries up to date over poor network connections, this effort is concentrated in a few places. Other institutions then make use of this local infrastructure directly via NFS/AFS and file transfer using the much better network connectivity that they have to the regional centers.

There will be varying forms of regional centers defined by the actual level of service and support they provide. This will be influenced by many factors. International networking and the costs of data transport may change significantly over the next five to ten years. The regional centers could play a role analogous to SLAC or any other home institution, or, more likely, somewhere in between.

Two sites have already expressed strong interest in the Regional Center model; CCIN2P3 in France and DRAL in the UK. Possible further sites in Europe or North America are not excluded, but the cost and effort associated with the copying of large quantities of data are likely to be such that the number of such centers will always be limited. Both CCIN2P3 and DRAL already provide large scale computing facilities for HEP experiments (including large UNIX workstation farms) on a similar scale to SLAC and are equipped with large tape robot systems compatible with the one at SLAC. Both the UK and France already make use of good national wide area networks, with *BABAR* software being maintained centrally and accessed by collaborating institutions via the network/NFS.

## Home Institutions

Physicists in the home institutions should be able to contribute to code development remotely. The mechanisms for this are already in place. The long term success requires that the collaboration computing model remains broad based in terms of supported hardware,

application of coding standards, and support for remote access and transport of small datasets.

It is anticipated that the analysis work of remote users will be centered at SLAC or a regional center in terms of access to the bulk data or full DST. Physicists will then transfer n-tuples or small DSTs to their own institutions for intensive interactive work.

Many home institutions have significant computing capability, and we may wish to make use of this for Monte Carlo generation, for example. The ability to generate Monte Carlo or reprocess small data samples locally, will almost certainly be required by many physicists.

### Code, Data, and Database Access

The access to code and data is covered elsewhere in this chapter. Providing convenient access to code for remote sites also has repercussions, however. For example, facilitating the running of *BABAR* code on essentially any home institution (UNIX) computer entails the risk of potential problems with compatibility of results from different machines, architectures, or flavors of operating system. This problem could become particularly important when considering reprocessing the data at one or more regional centers or home institutions. The collaboration is aware of these problems and aims to organize the production and analysis efforts so as to minimize their effects.

Implicit in having good access to the software and data samples is the requirement of access to the collaboration databases. Any institution processing or reprocessing data, generating Monte Carlo, monitoring detector performance, *etc.*, will require access to one or more databases. This access could be achieved by real-time queries to a master database at SLAC, over the WAN. This seems impractical today, especially when good real-time performance is required, for example, during a batch job. Most likely, copies of all or part of the database must be made available to collaborating institutions. These copies will almost certainly require frequent, automatic updates. The ability to do this will be important when choosing a suitable database technology for the experiment.

### License Issues

It is likely that at least some of the software used by the *BABAR* collaboration will be licensed commercially. Given that some of this will be duplicated in collaborating institutions, the collaboration should try to enter into blanket license agreements, covering SLAC and remote sites, in order to reduce the total cost of these licenses to the collaboration. For example, the possibility of purchasing AFS client licenses which are extensions of the Stanford AFS server license, as CERN has done with theirs, will be investigated.

### 11.3.3 Baseline Cost Model

It is too early to purchase the bulk of computing equipment needed for *BABAR* for many reasons: it is not needed immediately; decisions about inter-institution compatibility have not yet been made; technology is changing rapidly; and costs are decreasing. However, it is both useful (as a reality check) and necessary (to provide cost estimates for the funding agencies) to detail the design of a computing system using current and near-future technologies. In this section, such a system following the basic architectural plan set forth earlier (Figure 11-2) is described.

#### Networking

The baseline design specifies components available today (December 1994). The expectation is that the actual implementation will evolve with experience and product availability as items are required. The baseline design of the network should allow relatively easy migration to emerging technologies as they become cost effective. Based on the data flow, the various network links in the schematic diagram are addressed below.

#### Online Networking

The online networking is used for data acquisition, trigger, and online monitor/control. It is based on point-to-point FDDI links connecting the Level 1 single board computers to a cross-bar switch (Digital's GIGAswitch [GIG93]) and finally to the Level 3 CPU farm. A single GIGAswitch will provide 16 connections to the FDDI rings containing the front-end data acquisition crates; direct connections to 14 Level 3 farm CPUs; a link to a local Ethernet hub; and a single long-haul link from the IR hall to the computer center's mass storage silo. All other online networking for program development and monitor/control will use lower cost Ethernet links.

#### Local Area Networking

The design for local area networking is based on switched 10 Mbits/s Ethernet and 100 Mbits/s FDDI technologies. In particular, Digital's GIGAswitch and Alantec's Powerhub 3000 Ethernet switch [HUB94] are used to interconnect workstation farms, mass storage servers, and desktop machines as described below. The link between online and offline Gigaswitches will be an FDDI link running over existing multimode fiber between IR-2 and the computer center, a distance of about 5000 ft.

The need to follow activities in ATM (Asynchronous Transfer Mode) and the 100 Mbits/s Ethernet switching and interface markets as they evolve, is recognized. To gain practical experience and skills, it will be necessary to make serious on-site pilot studies of ATM and 100 Mbits/s Ethernet in 1995 or early 1996.

## Wide Area Networking

All of the computing tasks (other than data acquisition) can occur at the sites of remote collaborators. The degree to which this may be done is regulated by the computing investment at the remote sites and by the development of the WAN connectivity between SLAC and the collaborating institutions in the following years.

Today, SLAC has two T1 (1.544 Mbits/s) connections to ESnet, one to Caltech, the second to LBL. In addition, there is a 10 Mbits/s Ethernet microwave link to Stanford University that is used to access Stanford and to provide backup access to the rest of the world if the ESnet links are down. Wide area networking support to first order is outsourced to ESnet and Stanford University.

Data rates and response times between SLAC and other sites depend on the distance, the number of network hops, the speed of the links between hops on the route between SLAC and the other site, implementation details such as window and packet sizes, the error rates, and how busy the link is. Typical values for US sites vary from 90 kbytes/s and 12 ms ping response time, to 12 kbytes/s and 120 ms. For European nodes, the numbers are 10 kbytes/s and 200–360 ms.

As a minimum, all collaborators will need reasonable WAN connectivity to SLAC for the purposes of transferring code and performing interactive work using the X Window system. For X sessions, one study [Aba94] shows that the number of concurrent X sessions between adjacent T1 ESnet nodes is on the order of 24. Other studies [NCD89] indicate that the load imposed by an X terminal is approximately 50 kbits/s.

In spring of 1995, SLAC plans to upgrade its ESnet links to T3 links (45 Mbits/s). Recent measurements made by ESnet [Bos94] for one hop on unloaded T3 links between LLNL, LANL, and PPNL gave TCP data rates of between 0.58 and 1.72 Mbytes/s for optimized settings of window, packet, and internal write buffer sizes. Depending on budgets and demand, further upgrades of ESnet links to OC3 speeds (155 Mbits/s) may be initiated in 1995 for one or two sites, followed by other major sites a year later. The next step after OC3 is OC12 at 622 Mbits/s. Technically, upgrading to such links should be feasible in 1996/1997.

International links are more expensive than mainland US links. Today's cost for the ESnet T1 (1.54 Mbits/s) link from Princeton to Germany is about \$12K/month for the US half

circuit. This is one-fifth of the total cost of the link. One might expect a T3 link to be four to six times as expensive [Bos94].

Connectivity between the wide area network and the on-site network will be provided through a firewall FDDI router with connections to the SLAC FDDI backbone and the FDDI ring to which the off-site routers will be attached.

### Compute Farms and Other Processors

**Online System.** Software development, run control, and monitoring tasks will use general purpose workstations interconnected with each other and with the computer center as described above. The Level 3 farm, which functions as a part of the data trigger, online reconstruction, event tagging, and event monitoring, will also use general purpose workstations.

**Reconstruction and Monte Carlo Farm.** Event reconstruction is done in the offline reconstruction and Monte Carlo farm. Data from tape are first staged to a series of disk servers via FDDI connections through the offline switch. Each reconstruction farm machine is directly connected to a network hub, which in turn connects to the disk server via the switch. Both reconstruction and Monte Carlo tasks tend to be CPU bound rather than I/O bound, thus allowing the use of slower and less-expensive Ethernet network links to these machines.

The data aggregate rates for this farm are 8 Mbytes/s (Table 11-2). Assuming the nodes in the farm are rated at an average of 200 MIPS each, there will be about 50 nodes to provide the 10,000 MIPS required (Table 11-1). Using existing technology, the networking requirements can be met by connecting the nodes to two Alantec Powerhub 3000s. Each of the Powerhubs will support 12 switched full speed (10 Mbits/s) Ethernet connections and an FDDI connection. Each Ethernet connection will have two to three nodes attached via Unshielded Twisted Pair (UTP) connectors, for up to 36 nodes per Powerhub. Each of the two FDDI connections (one for each Powerhub) will attach to an FDDI port on a GIGAswitch.

**The DST Analysis Farm.** Creation and analysis of DSTs occur in a second farm of machines with direct FDDI connections to the offline switch. These jobs are I/O intensive, thus requiring the higher performance and more costly FDDI network links. Data flow is similar to that for the reconstruction and Monte Carlo, from the mass storage to disk servers and finally to/from the DST farm.

The aggregate network capacity required for the DST analysis farm, as estimated in the preceding requirements section, is 200 Mbytes/s. The compute requirements (Table 11-1) of 2500 MIPS could be met with a farm of thirteen 200 MIPS processors. To handle the bandwidth requirement of 200 Mbytes/s, however, about 20 nodes will be needed, assuming each node can support a maximum of 11 Mbytes/s FDDI transfer rate. In this model, each processor node will have its own dedicated FDDI port on an offline GIGAswitch. Measurements must be made to see what FDDI transfer rates can be supported by these nodes while still allowing them to deliver the requisite compute cycles for DST creation and analysis. One estimate for an ATM switching network [Dou94] indicates that approximately 8 MIPS of CPU power are required to deliver each megabyte per second of data. If this turns out to be a limitation, then an alternative is to double the number of farm nodes (hence, halving the required bandwidth for each node) and for pairs of nodes to share a single GIGAswitch port.

**Desktop Workstations.** Interactive analysis involves the manipulation (and sometimes creation) of  $n$ -tuple datasets, and other event data in the highly interactive manner of PAW or HippoDraw. This is also an I/O intensive task, but in general, it involves data sets of modest size. This type of exploratory analysis occurs on desktop workstations. The data flow is identical to that of the farm machines.

The desktop workstations are expected to be mainly located in the Central Lab, the Central Lab Annex, the Warehouse, IR-2, and the new  $B$  Factory building. The data rates for individual workstations are 0.2 Mbyte/s (Table 11-2). This is well within the capability of the 10 Mbits/s Ethernet interfaces of today's workstations. To ensure that multiple workstations do not require more aggregate bandwidth than can be sustained on a single 10 Mbits/s Ethernet segment, each segment will support only a few workstations. As more bandwidth is required for a particular desktop, the number of desktops on a segment can be reduced to one. The individual segments will be connected to an Ethernet switching hub located in the same building as the desktop, each with up to 36 nodes attached. In buildings with large ( $> 30$ ) concentrations of workstations, multiple switching hubs will be deployed, each hub being located close to a workgroup, and in turn, connected via a multimode fiber to a FDDI port on the offline GIGAswitch.

## Mass Storage

**Tape Mass Storage System.** A relatively small number of candidate mass storage technologies available today might serve  $BABAR$ 's needs. One seems especially well suited for use at the SLAC site, where four STK robotic silos with 3490 tape technology are currently in use. Each tape cartridge in this system holds 800 Mbytes (before hardware compression), while each tape controller (connected to four tape drives) has a maximum throughput of

about 2.5 Mbytes/s. The total storage capacity is thus 24 Tbytes. An attractive upgrade option involves migrating to STK's Redwood helical scan technology. The expectation is that this will initially provide 25 Gbytes of data per cartridge, about 15 Mbytes/s maximum throughput per tape drive/controller pair for a total storage capacity of 600 Tbytes. Cartridge capacity is expected to quickly rise to 40 Gbytes and eventually to 80 Gbytes. Such a system satisfies the *BABAR* storage requirements while reusing a significant amount of existing equipment. This solution may also be attractive to remote regional computing centers within the collaboration where there is already a large investment in STK silo equipment.

**Mass Storage Server.** The data rate from the online to the mass storage is 2.5 Mbytes/s. In addition, there will be 7.5 Mbytes/s from the reconstruction and 0.5 Mbyte/s from the Monte Carlo generation to tape (Table 11-2). To this must be added the requirements for staging analyzed data from tape to disk. This requirement is assumed to add an extra 5 Mbytes/s which would be sufficient to allow the 2 Tbytes in the disk farm to be staged from mass storage twice per week. The aggregate data transfer rates for the mass storage are thus about 15.5 Mbytes/s. To this must be added the requirements for the creation of major DSTs, which is foreseen to take place mainly during the reconstruction. To the extent that additional major passes must be accomplished in short time periods, the model may require scaling to additional bandwidth.

The baseline design calls for the mass storage to be served by four nodes. Each of these nodes will have its own dedicated FDDI connection to the offline GIGAswitch. The transfer rates achievable on the FDDI are on the order of 11 Mbytes/s [Che93]. Initially, the performance is expected to be limited by the workstation servers and the interfaces.

**Disk Mass Storage System.** The disk servers have bandwidth requirements similar to the DST analysis farm since that is where most of the data are needed. Using the GIGAswitch model, a need of about 20 disk server nodes each with a dedicated FDDI connection to the GIGAswitch is again anticipated. Today's model of GIGAswitch can only accommodate 32 FDDI connections, so at least two GIGAswitches will be needed to support the DST analysis farm and the disk servers. This will require segmenting the DST analysis farm and disk servers, each segment having connections to the reconstruction farm, the mass storage, the desktop workstations, the transportable media, and the external router. The connections between the two segments will be made by one or two FDDI links. This segmentation will need further evaluation to ensure a reasonable balance can be achieved. A new model of GIGAswitch is expected to have 52 ports with 155 Mbits/s ATM, and will be evaluated as part of the on-going design study.

**Transportable Media.** Standard 8 mm helical scan magnetic tapes will serve as the basis for data transfer to institutions without STK silo compatible drives. The data rate requirements for the transportable media are expected to be modest and easily handled by a single FDDI connection from the server to a GIGAswitch. If more than one server is required, then the servers will share a single FDDI port on the GIGAswitch.

The decision to use 8 mm tape in the baseline design stems from its widespread use within the physics community and current support at SLAC. Other technologies, such as 4 mm or DLT, are also possible for comparable costs. However, at most, one medium besides the silo medium would be supported by the collaboration.

## 11.4 Software Environment

---

The basic requirements for the software infrastructure are that it provide an environment in which developers are highly productive, provide the necessary stability to ensure that high-quality, reproducible physics results may be achieved, and also be flexible so that changes in both the detector hardware and physics goals may be accommodated quickly and efficiently. Applying the principles of commonality across all areas of software development, from the real-time systems within the data acquisition and trigger systems through the online and monitoring systems, bulk data reduction and Monte Carlo generation, and finally physics analysis, can bring about important benefits. However, application of these principles must not be at the expense of productivity and must take into account that there are different demands in the various application areas.

### 11.4.1 Software Methodologies and Languages

#### Types of Developers

Several types of software development exist within the collaboration, each type having different requirements in its support levels. Individual physicists developing their own analysis programs should not be overburdened with rules and regulations, whereas software developed in the real-time environment of the data acquisition and trigger systems must be subject to rigorous quality tests to ensure that physics data are not lost. Intermediate between these extremes is the software for the bulk reduction of the raw data and the online and monitoring software that ensures that the data from the experiment are of high quality. Within each of these levels, an appropriate level of quality control will be applied as described later.

### Software Methodology

Our overall approach to software development is based upon the object-oriented (OO) methodology. This methodology allows for a direct, intuitive association of physical entities (*e.g.*, detectors and particles) and abstract entities (*e.g.*, detector status and run control) with software objects. Commonalities between objects are also identified and exploited via the concepts of inheritance and overloading. Mapping of software objects into lines of code is most naturally done with an object-oriented programming language such as C++. The debugging and maintenance phases of development are well supported by the modularity imposed by the use of software objects. Many aspects of the object-oriented methodology can be automated through the use of computer aided software engineering (CASE) tools.

Several particular OO methodologies and notations have been investigated, including Booch, Schlaer-Mellor, OMT, and Objecteering [OOM94]. Two CASE tools, one supporting the Booch methodology, the other supporting the Objecteering methodology, are under detailed consideration. Each provides automatic C++ code generation as well as some reverse engineering capabilities for existing C++ code.

### Programming Languages

After much consideration and experimentation, it has been decided to base most of the software development on the C++ programming language while continuing to support the Fortran 90 language. Most of the real-time and online systems, together with the core of the offline system, will thus be implemented in C++, while adequate functionality will still be provided for the Fortran 90 user. Such an approach is believed to produce the most robust applications, taking advantage both of the power of the OO approach and C++ and of the large existing Fortran code base and many man-years of experience with Fortran. This approach implies a mixed language environment and therefore carries the burden of additional support, but it is believed to be the most realistic approach, given the schedule demands and learning curve of the new OO paradigm.

### Cross-Language Issues

Efficient access to software modules and data from both of the supported programming languages is obviously of crucial importance. Significant R&D has been performed in understanding how to accomplish this. Of central importance is the fact that the client programmer's view of the data format is established not by the data itself, which might be highly compressed in order to minimize the disk or tape space requirements, but by the interface that the corresponding object or module provides. Transforming the data between its compressed, internal form, to that seen by the user programmer, can be achieved in an

efficient manner for both C++ and Fortran 90. Further R&D is underway to see whether both representations can be derived from a common source. The Interface Definition Language (IDL), a component of the Common Object Request Broker Architecture (CORBA) [OMG91], allows this and has the additional advantage that it supports distributed object applications. An incomplete prototype IDL compiler that generates both C++ and Fortran 90 stubs and skeletons has been demonstrated and shows some promise, but more work in a real application environment will be necessary in order to decide whether to continue this project. This approach seems to be well matched to fairly coarse-grained clients and servers, such as the packages described in the code management section.

## 11.4.2 Development Environment

### Approach

In recommending that most program development be performed using the OO paradigm and implemented using C++, it is recognized that ease of use will have to be maintained, and therefore, support tools will be necessary in order to minimize the learning curve. The approach is aimed at establishing an environment in which newcomers can quickly be productive. Tools that are already in place, or shortly will be established, include code guidelines, code templates, sample programs, and tutorials.

### Code Guidelines, Templates, and Sample Programs

Code guidelines based on previously published documents [Ell92] gather rules and recommendations as well as lists of common pitfalls and suggestions for avoiding them. Code templates exist for both supported programming languages.

Some introductory tutorials have already been developed within the collaboration, but these will be extended to cover a wider range of topics including the selected OO methodology and notation, the CASE Tool, and C++ language tutorials for more proficient programmers. Other vendor tutorials will be combined with those written within the collaboration.

### Quality Assurance

Our software quality assurance (QA) strategy is based on a stable software release mechanism with responsibility for quality assurance being distributed among several software coordinators, as detailed in Section 11.4.4.

One tool in the overall QA strategy is code reviews, in which one model under consideration is that of a mentor who can informally review new code and advise the developer of improvements, rather than a formal review process with its large manpower overheads. A bug tracking system is being evaluated that should ensure that bugs are given sensible priorities and fixed in a timely manner.

A crucial component of software QA is adequate documentation, both at the level of the user and at the level of the maintainer. There is a commitment to the creation of documentation as an integral part of the development process, not as an afterthought. Word processors have been used to create documents (the Code Guidelines being an example of this) that can be viewed both in printable form (*i.e.*, PostScript documents) and also as browseable WWW documents for online use.

### Compilers, Debuggers, Application Builders, and CASE Tools

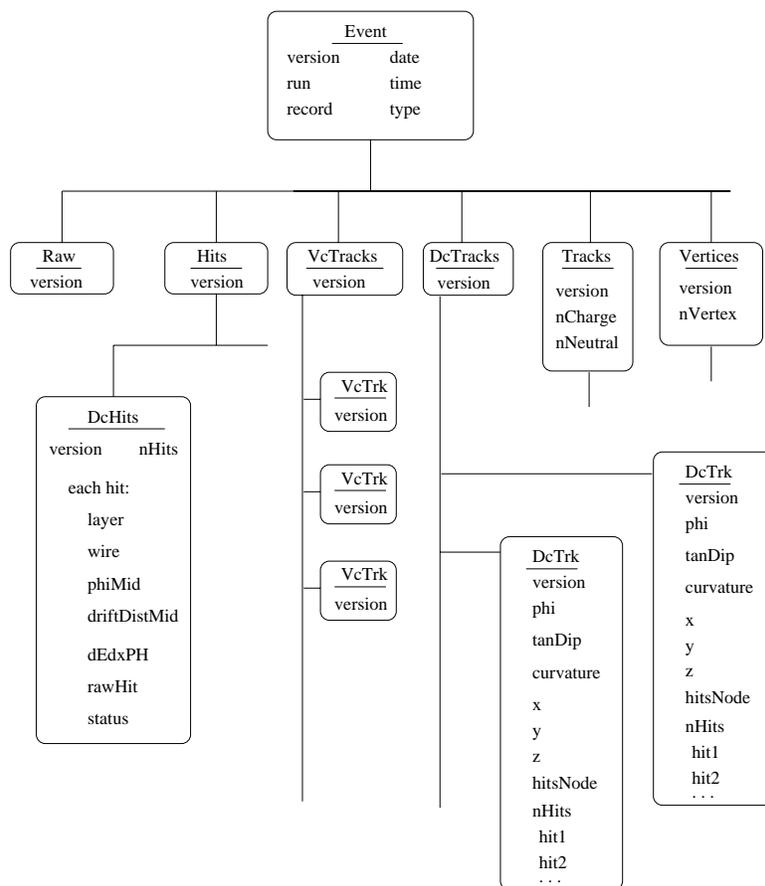
It is unfortunate that the C++ language has only a draft standard and that efficient Fortran 90 compilers are only just becoming widely available. The GNU C++ compiler has been identified as providing a suitable baseline in functionality, while excluding any extensions that are not part of the draft C++ language standard. Compilers from other vendors that conform to this minimum functionality are available on all hardware platforms being considered for use within *BABAR*. The situation for Fortran 90 is that one compiler (NAG) is available on many platforms and adheres strictly to the standards, and native compilers from the major vendors are either already released or shortly will be.

As mentioned in a previous section, some CASE tools are being evaluated, though it is realized that their use will probably be limited to a core set of developers. However, the more widespread use of the methodology and notation for which the CASE tools provide support is being promoted.

Application builders exist, although their use within HEP is not widespread. Several are being evaluated, some of which are commercial and therefore raise licensing issues. A prototype application framework that will support both batch and interactive environments is under development.

### 11.4.3 Data Model

The use of an object-oriented methodology for the software development environment leads naturally to the description of the data model for the experiment in terms of objects. This bypasses many of the solutions that have arisen throughout the FORTRAN era such as BOS and ZEBRA and integrates access to the data naturally into the programming language.



**Figure 11-4.** Portion of a FARFALLA event tree for the BABAR experiment.

A uniform data model will be applied throughout the software. Thus, the same model will apply to the access to raw, reconstructed, Monte Carlo, or calibration data, and to geometrical constants. This model will also allow for a smooth schema evolution. Advantages of this approach are that it makes possible access to the new generation of database management systems that manage objects, and that the implementation of the data storage is decoupled from its access.

An investigation of various approaches to the data model for BABAR has been carried out, including a number of products based on C and C++ [Sax94]. The FARFALLA [Wal94] package is the most promising existing object-oriented package on which to base the BABAR data model. A prototype effort using FARFALLA is underway.

The FARFALLA package provides a means to store an experiment's data on disk or tape, with convenient access in the user's program. It is based on the idea of nodes in a tree structure.

Some of the nodes in the *BABAR* prototype event tree are illustrated in Figure 11-4. As *FARFALLA* is written in C++, it provides a match with object-oriented design in C++ for other software. The access to the data from C++ programs is straightforward. The data itself is encapsulated (hidden), and basic access is provided via inline accessor member functions. This decouples the means of storage, which can be optimized for space, from the user interface, which can be optimized for convenience. Additional capabilities are provided by more complicated member functions. Because the detector simulation currently employs *GEANT*, written in Fortran, and some prototype reconstruction efforts are also being undertaken in Fortran, one of the objectives of the data model prototype is to provide access to this data. It has already been demonstrated that this can work, and current concern is in making this access in Fortran and C++ look similar.

### 11.4.4 Code Management and Distribution

#### Introduction

The code management environment must satisfy conflicting demands from the code developers, who want to be able to make changes rapidly with a minimum of controls and overheads, and the end users who wish to see a stable and bug-free code base. The situation is further complicated by the need to accommodate distributed program development throughout the collaboration and to make the resulting software rapidly available to the distributed end users. Finally, various machine architectures and dialects of UNIX must be accommodated.

Implementing a system that resolves these conflicting demands involves not only robust computer tools, but also policies on their use, and perhaps most important, acceptance by the developer and user communities that some controls need to be applied in order to ensure that the experiment produces reliable, reproducible, and defensible physics results.

#### Policy

The adopted policy is based on the concept of a *code package*. A package is a well-defined piece of code that has an internal coherency and well-defined interfaces to other packages. Examples of packages are track finding, vertex fitting, histogramming, *etc.* Many of these packages will come from within the *BABAR* collaboration, but others will come from outside. Individual applications are also treated as packages in this model. Each package is assigned a Package Coordinator, a responsibility that might be shared among several individuals. The coordinator is responsible for the development of stable snapshots or *versions* of the package, ensuring that a minimum set of test procedures has been performed. Different versions for each package might be generated quite rapidly during periods of intense program

development, but the development pace is dictated by the developers themselves, not outside constraints.

*BABAR* software releases are created at intervals dictated by stability and scheduling constraints. The release mechanism is controlled by a Code Coordinator. A release is formed from the most recent stable versions of every package, where stable is defined by the package coordinators. Testing focuses on interdependencies between the packages on the assumption that each package has already received some internal testing. As bugs are detected and corrected, the corrections are propagated forward to the most recent development version in a manner that does not interfere with the ongoing development. Software tools aid in the coalescing process.

Quality assurance is therefore a distributed activity, with all developers bearing some responsibility, although the main burden lies on the overall Code Coordinator and the Package Coordinators.

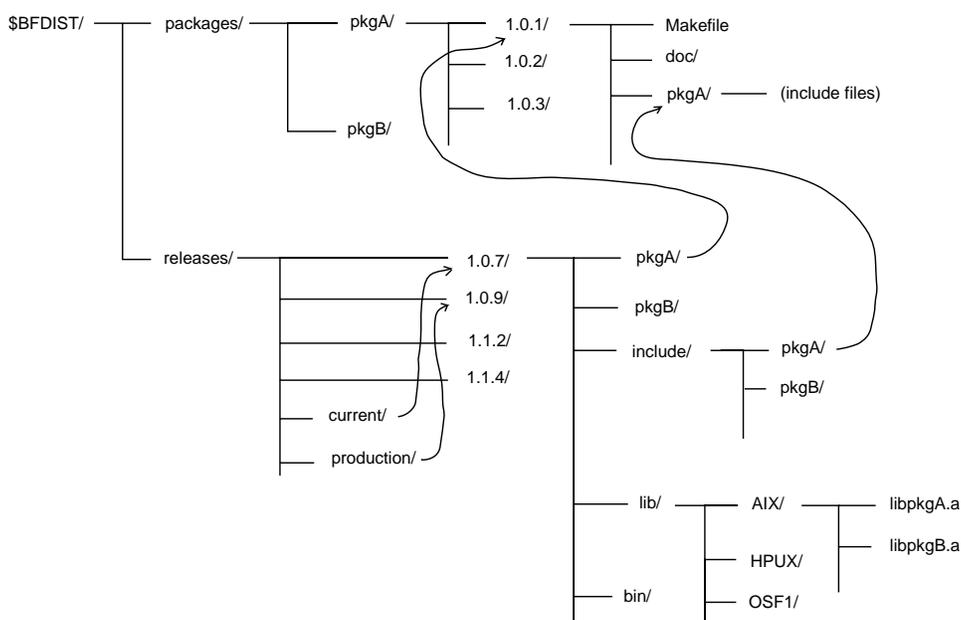
### File Organization

A simplified file organization is shown in Figure 11-5. The directory tree is split into package and release subtrees. All software packages reside under the package tree, with the primary code repository and subtrees corresponding to individual versions. Each package has directories for source code, libraries, executables, documentation, *etc.*, for each version. Under the release subtree are stored all the complete releases, although space limitations will probably cause only the more recent releases to be maintained at any one site.

### Tools

The primary code management tool is the CVS package, which allows multiple developers efficient access to a code repository while supporting tracking of code changes and their authors. The utility rCVS, developed at SLAC, extends the use of CVS to distributed developers by maintaining proxy code repositories at remote sites, via which changes are automatically propagated to and from the primary repository.

One weakness of CVS and the UNIX file system is the lack of Access Control Lists (ACLs), by which individual developers may be granted access to a subset of the overall code repository. Since it is expected that most of the collaboration will act as code developers for at least one package, accidental corruption of the repository is of some concern, and ACLs could play an important role in this. Advanced distributed file systems such as AFS and DFS support ACLs and would address these concerns, while simultaneously making the code repository available throughout the collaboration without the need for rCVS. More work remains to be



**Figure 11-5.** *Simplified file organization, showing the release and package subtree structure.*

done on understanding the performance issues of these file systems, and R&D is underway at SLAC with a small user community on AFS.

The various dialects of UNIX each have their own syntax for the make utility that builds the code dependencies and automates the building of application programs from the various packages or the libraries that might be part of a package. Two alternatives are being investigated in order to standardize on a version of make: `imake` (part of the X11 distribution) and GNU make.

Finally, a tool that automatically creates standardized version and release numbers and performs the functions for creation of a new release would greatly reduce the burden on the code coordinator. The Production Code Manager (PCM) prototype under development at SLAC is designed to perform these functions and is under evaluation.

## Code Packages

As has been mentioned previously, a package is a well-defined body of code. Any one package is likely to be referenced by other packages. One area that needs more work is ensuring that circular dependencies do not develop between packages, since this greatly impacts the ease

with which applications can be built. This implies a hierarchy in which a package can reference classes or functions from packages lower in the hierarchy, but not from its peers.

### Software Release Mechanisms

A release of the *BABAR* software will consist of a self-consistent and tested set of all software packages. It will be entirely self contained, applications having been built for all supported hardware platforms, thus relieving remote sites of the burden of building applications from source code. This implies a certain minimal level of consistency in operating system versions across the collaboration with only a small set of operating system versions being supported for each hardware platform. Remote sites can either have direct access to a release using a distributed file system such as AFS, or the directory tree corresponding to each release can be made available in compact (`tar`) form suitable for shipping across a network or via tape. The goal is to have reasonably stable releases (public releases) available at fairly frequent intervals (a few weeks to months) with more stable production releases having longer lifetimes for bulk production requirements.

An area that needs further work is that of integrating the bug reporting and tracking from remote sites in such a manner that fixes or patches are made available in a timely and coherent manner across the complete collaboration.

### 11.4.5 Databases

The *BABAR* project faces a formidable set of information management requirements. The varieties of data that must be managed range from conventional institutional information to highly technical engineering and physics data. The computing model in which it must be managed is distributed, heterogeneous, and geographically dispersed. Modern database management systems (DBMS) will be essential to satisfy these requirements. A partial list of databases required for *BABAR* includes the following:

- Detector information: geometry constants; calibration and alignment constants; hot-and dead-channel maps; monitor logs.
- Acquisition system information: constants for front-end electronics.
- Run information: run start and stop times; trigger and readout configurations; event and record counts; tapes written; online logbook.
- Offline information: location and status of datasets; status of jobs for each run (real data and Monte Carlo).

- Tape management system.
- Document catalog: *BABAR* notes, manuals, papers, drawings.
- Engineering data: component and subsystem test and calibration records.
- Personnel information: collaboration directory, distribution lists, author list, shift list, ES&H training, human resource estimates.
- Management information: budgets, schedules, equipment inventories, software license tracking.

The possibility of storing information about interesting physics events in a database will also be investigated.

Two different database technologies are relevant to *BABAR* requirements. Relational database management systems (RDBMS) are the current choice for managing institutional data in most large organizations, including other current HEP experiments. This is a mature technology with several major vendors and many high-quality third-party products on the market, a well-established theoretical foundation, and an expressive and standardized query language (SQL) to provide portability. On the other hand, RDBMS have not been integrated smoothly into programming languages, and their simple data model (rows and columns) is poorly matched to the complex data structures needed in a HEP environment and supported by an object-oriented language such as C++.

Object-oriented database management systems (OODBMS) provide an increasingly popular alternative to RDBMS. These systems generally map database elements directly to software objects. Most of the complexities of dealing with the database management system are hidden from the programmer. Unfortunately, these products are less mature and lack both a firm theoretical foundation and (as a consequence) satisfactory *ad hoc* query tools. Moreover, though several candidate interoperability standards exist, there is not yet strong, cross-vendor support within the industry for any of these specifications.

Oracle is already in place at SLAC and is in use by both *BABAR* and PEP-II for a variety of different types of institutional and engineering data. Most such data will likely remain in Oracle at least until OODBMS with equally expressive query languages become available. *BABAR* will investigate the use of OODBMS for those databases accessed heavily from physics programs, *e.g.*, calibration and alignment constants, channel maps, *etc.*

Some types of data, such as the run database, are likely to be frequently accessed both from within the physics code and from external query tools. The choice of technology for such databases will be deferred until more experience with OODBMS is gained and a better idea of the access requirements of the physics code is obtained. It is likely that the physics code

will require access to at least some relational databases, however, so *BABAR* will also select, or if necessary, develop a class library for accessing a generic RDBMS.

Large scale databases represent potentially major investments, both for licensing and administration. It is, therefore, unlikely that all collaborators will have local access to all database software. Various models can be imagined whereby data is shared with other collaborating institutions. For cases in which database software must be present, site licensing will be investigated, and a mechanism for periodically replicating and distributing certain databases will be developed. Where licensing is not essential or is impractical due to expense, services will be provided to access data over the network either via *BABAR*-developed client tools or via low-cost, third-party “data-dipper” applications.

### 11.4.6 Graphics

Computer graphics serve an important role in various stages of an experiment. During detector design, they aid the physics simulation tasks and CAD. In the course of code development, graphical detector representations and interactive event displays are necessary for pattern recognition and event reconstruction studies. For physics analysis of data, it is essential to have tools for event viewing, parameter representation, multidimensional statistical analysis, and histogramming. Graphics displays with graphic editors are useful for producing high-quality hard copies for presentation and publication. Also, graphical user interfaces (GUIs) are important for human interfaces in a variety of contexts, such as basic computer interaction (the desktop), software engineering tools, job submission, and detector control and monitoring. In this section, the available computer graphics and GUI options are summarized.

#### Graphics Hardware

Compared to the industry standards, the hardware requirements for graphics in high-energy experiments are moderate. The major factors are compute performance, variety of color display, screen resolution, and the ability to produce quality hard copies. For screen display, there exists a range of workstations with sufficient performance for most applications. Their price/performance ratio is improving with time. For hard copies, color printers and color copiers are becoming inexpensive. These trends are expected to continue.

### Software Options

For histogramming and plotting curves, there are several packages already available such as Handypak and Topdraw, along with the CERN graphics packages HBOOK, HPLOT, PAW++, *etc.* For event and detector displays, an assortment of approaches ranging from simple two-dimensional X11 or PostScript cartoons to complex three-dimensional structures can be useful. Most three-dimensional packages have high manpower overhead, and options should be examined carefully. One could also adopt a package like that of ALEPH's V-plot for representing three-dimensional data of events [Dre91].

In the UNIX environment, there are a number of commercially supported three-dimensional graphics packages available such as Graphics Kernel System (GKS), Programmer's Hierarchical Interactive Graphics System (PHIGS and PHIGS PLUS), Silicon Graphics Library (OpenGL), and Renderman. While X11 and PHIGS can be obtained free of charge from the Open Software Foundation, other packages must be purchased from commercial vendors. PEX (PHIGS Extensions to X) also supports three-dimensional graphics within the X-Window environment. Some of these packages have ISO and/or ANSI standards defined. They also provide device-independent interfaces. In the near future, a choice among these packages will be made based on relative performance, ease of use, and portability.

There are also a number of visualization packages available on the market such as AVS, apE, and PV. With these packages, the user can develop graphical displays interactively using the mouse and keyboard at a workstation.

### Graphical User Interfaces

A graphical user interface (GUI) is characterized by pointing, clicking, and dragging with a mouse; pull-down menus; dynamically resizable windows; and other features. GUIs offer the user a high degree of interaction and assistance in navigating through unfamiliar programs. Each component of the program can be associated with graphical icons illustrating the intuitive connection with its associated function. A properly designed GUI can dramatically aid the training to operate a program, reduce errors, and increase operator efficiency.

There can be a considerable cost associated with these user interfaces in terms of programming development. In fact, there are reports that 80–90% of a program's development may be required to design and produce a good GUI in the absence of a good interface builder. The availability of commercial packages called interface builders significantly eases the work load in developing GUIs. Products producing X11-based GUIs are numerous, including XVT-Power++, TAE+, UIM/X, Garnet, and TeleUSE.

## 11.5 Online System

---

### 11.5.1 Introduction

The online system provides for the operation of the *BABAR* detector while it is preparing for or engaged in data taking. The system's primary responsibility is for collecting and storing the experiment data. It will also monitor and maintain the integrity and quality of that data. The online system must maintain a high level of availability during collider operation. This system can be organized into several main areas as shown in Figure 11-6 and listed below:

**Experiment Control.** Controls and monitors the environment of the detector.

**Run Control.** Organizes, controls, and monitors the detector components, hardware, and software for data collection.

**Data Flow Control.** Controls and monitors (in real-time) the passage of data from the detector subsystem back-end electronics to the final data repository and other data consumers.

**Trigger Control.** Controls and monitors the decision making processes that determine whether data will be recorded.

**Clock and Control.** Provides a variety of precise timing signals and strobes to the subsystems.

The online system must interact with and link together the various components of the *BABAR* detector as illustrated in Figure 11-7. Some of the systems illustrated, such as the PEP-II control system, require only a communication protocol specification from the online system since they are external to the detector system's responsibilities. Other subsystems, such as the detector front-end electronics, require detailed specification so a complete range of operations is allowed. The online system will share much of its architecture and software support with the offline system.

The online system will use an object-oriented design methodology for many of its components. This approach stresses the reuse of computer code, logical constructs, and ideas.

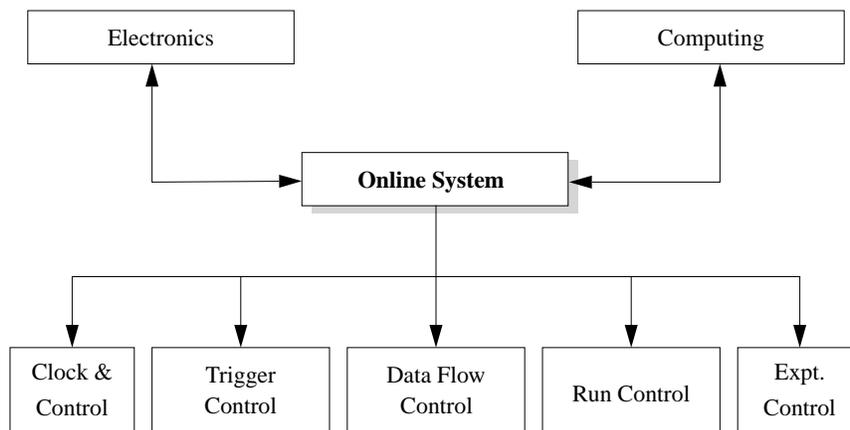


Figure 11-6. Schematic organization of the online system.

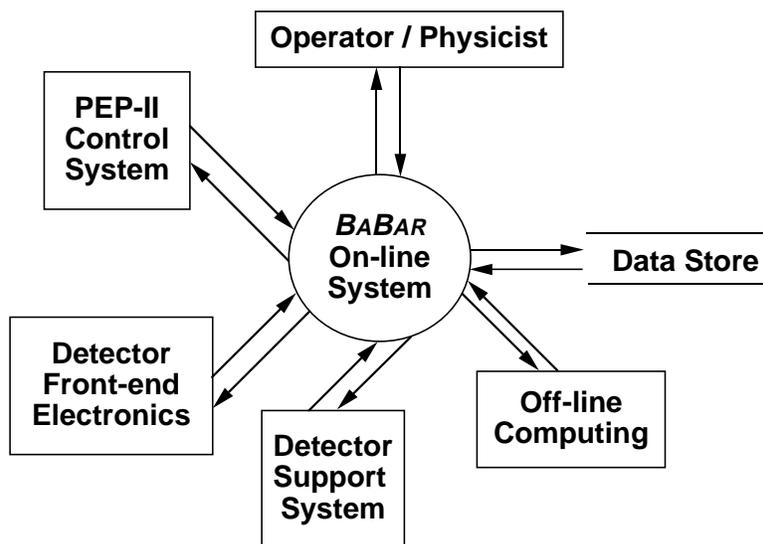


Figure 11-7. Context diagram of the online system.

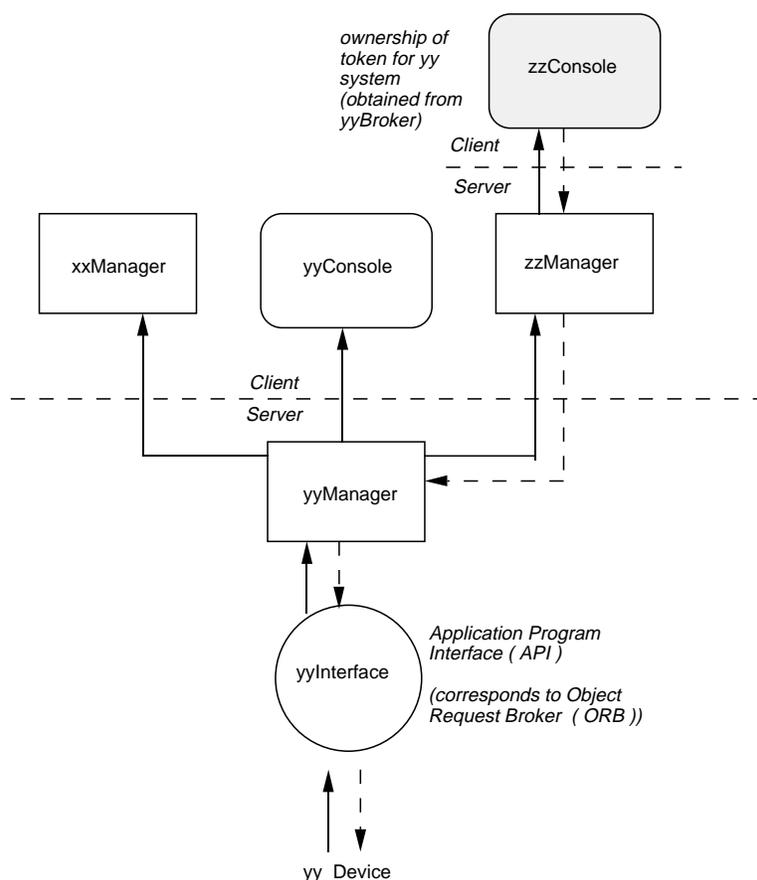
Within this methodology, the system will develop iteratively with each major release usable as a test system, but without complete functionality. Later versions will build on the existing system by filling in missing functionality. This model furnishes concrete deliverables for test stands, early beam tests, the high-energy ring commissioning, and the colliding beam commissioning. It also provides for adaptive improvement of and evolution in the system.

The description of the online system that follows presents different views of the system. The logical view focuses on the functional description of the system. The subsystem view details the interface to the detector subsystems and their common logical and physical constructs. The operator view describes the system's appearance to and allowed interactions with the physicist/operator. The implementation view describes the working decisions and specific implementations of the system. This approach aids in coping with some of the system's intrinsic complexity by focusing on only one aspect of it at a time.

### 11.5.2 Logical View of the Online System

The online system will use a client-server, distributed-object model for its architecture in which high-level control of the system is arbitrated by a series of console, manager, and interface processes (Figure 11-8). A console process such as a GUI program or a script always acts in the client role and allows the user to interact with and exert direct external control over the online system. A manager process provides resource allocation and control for the system's consoles, device interfaces, and other managers, and is composed of objects using Common Object Request Broker Architecture (CORBA) [OMG91] techniques to implement client-server communications. A device interface layer with well-defined functionality controls the system's low-level distributed devices, encapsulates the associated electronics, databases, and processes, and shields the rest of the online system from these complexities. In this architecture, multiple clients may need the services of a specific manager, but only one client at a time is allowed control. Token passing controlled by specialized managers called brokers will be used in this role.

**Experiment Control.** Experimental Control processes ensure correct functioning and efficient performance of the *BABAR* detector by continually monitoring its components. This involves periodic readback and analysis of the detector electronics and diagnostic procedures on the detector components. Experiment Control is not responsible for the data flow, trigger, or clock and control areas, which have their own control processes. Since the functionality of this system is common to most other high-energy physics experiments, the EPICS tool kit will form the basis of the Experiment Control framework. EPICS is freely available software developed originally at LANL and later at ANL and CEBAF. It was originally designed as an accelerator control system, although its current use includes several *experiment control*



**Figure 11-8.** Interaction of multiple consoles/managers, logical layers, and interfaces.

systems. Because EPICS is intrinsically a procedurally designed and implemented tool kit, a manager and interface software layer will be built above it.

**Run Control.** Run Control will guide the operator along the proper operational sequence by which the detector is readied for data collection, calibration procedures, or diagnostic procedures. Crucial hardware and software components will be automatically tested. During data taking, Run Control will continually monitor the detector and experiment data to detect failures and malfunctions. If an error is detected, an expert system will be used to help diagnose and fix the problem. This system is expected to evolve so that it can solve some routine problems on its own without involving the operator.

Run Control organizes and controls the detector for data taking using a configuration manager and a partition manager. It also has an expert system manager which aids the

operator in diagnosing and fixing problems. Configuration management provides the mapping among the physical hardware/software components and the appropriate control and monitor managers. All persistent attributes of these components will be stored in centralized object-oriented databases. For example, the device interface will build its EPICS databases from the central database according to the hardware needs. In particular, each component will be described by a state transition diagram. Partition management provides dynamic reorganization of the system's components into independently operating logical subsets. For example, the main partition could include the trigger and other system-wide resources, while several subsystem-level and individual readout crate partitions also exist. Configuration of a particular partition requires the instantiation of the objects representing the various hardware and software components on a workstation and a selective loading of data files and code to the associated embedded CPUs.

**Trigger, Data Flow and Clock and Control.** Trigger, data flow and clock and control are described in the Chapter 10. From the perspective of the online system, the trigger controls the data flow in the experiment. Trigger signals in the form of accepts and rejects, busy and ready, success and failure, are distributed as strobes via Clock and Control.

### 11.5.3 Subsystem View

Subsystems have several interfaces with the online system. Front-end electronics provide input to the trigger and data flow processes as described in Chapter 10. General purpose hardware and software perform routine monitoring and control functions. A set of special purpose hardware and software performs diagnostic testing for each subsystem. All devices are associated with a standard readback identification and status either directly or through the controlling processor. The identification is used for inventory, maintenance of repair records, and configuration details.

Most of the controlled and monitored components, including some purely logical components, can be described by a state transition diagram which has a direct EPICS implementation. The status is a state-dependent qualifier that describes the device's readback in terms of the device's expected behavior in that state.

The device manager drives the device interface through its transitions, enforcing an internal logical consistency and conformity with external constraints. For example, the magnet power supply in the state OFF is expected to have a current output of zero, within the tolerances of the readback. If the current monitored is not zero, then the returned status would be OVER CURRENT and color-coded red to signify a possible malfunction. The set of status values will depend on the allowed states of the device.

Data retrieved from the monitoring processes will be archived and error and message logs maintained. Status and state changes will be monitored by Run Control to maintain the integrity of the experiment data. During normal data taking, special readout options will be available. Nonsparsified data from a subsystem can be collected and monitored. This might check the validity of deleting below-threshold data from the normal data stream. A snapshot mode can monitor data from the VRAMs. This mode might check feature extraction algorithms by allowing comparison with full wave forms. An analysis mode can save all data buffers associated with a given trigger, so that the history of an event in the data acquisition can be fully reconstructed.

The process of controlling a device involves changing its state or the values of its functional parameters. A manager acquires a control token for the device and then takes the desired action. During the transition, a graphical representation of the action and its progress will be continually updated and displayed. For complex collection objects, the manager will enforce any necessary synchronization and present the most relevant progress display. Any error messages generated are sent to the operator and Run Control and are archived.

Most detector components will have associated calibration and diagnostic procedures to ensure their correct performance during operation. As required by partitioning, all subsystems must be capable of independent and concurrent calibration. In addition, individual readout controllers must be allowed to run in diagnostic mode while the others are participating in normal data-taking operations or calibration. As a result, a controller or set of data-flow strobe signals may either be replicated, or be made part of a diagnostics partition to allow independent operation.

A parallel load test will also be provided, using the controller to issue data pulses and triggers, which exercises the entire normal data-taking stream of the main partition.

#### 11.5.4 Operator View

The online system presents an operator of the detector with a standardized, hierarchically arranged set of control, diagnostic, and monitoring panels in which the most important parameters and controllers are also those most prominently displayed. Control panels allow an operator to perform and monitor normal operations on a particular subsystem. Diagnostic panels allow an operator to perform normal operations as well as the capability to test and exercise individual subsystems. Monitor panels only display information about a subsystem and do not allow an operator to perform any control functions. Detailed information and control of each element in a subsystem will also be available, but only when requested explicitly by the operator. The system will rely on the standard graphical user interface libraries and tools of the X-Window/Motif system.

For example, the calorimeter subsystem might have a main control panel that includes the status of the bias voltage power supplies with an overall gain knob, the most recently performed diagnostics, a summary of the average hit rates in the crystals, the status of the calorimeter's front-end electronics and DAQ modules and a set of histograms reflecting the overall performance and recent history of the calorimeter's functioning. The bias voltage power supply status would, for example, allow access to each of the individual power supply crates, which would in turn allow access to individual bias voltage channels through a standard point-and-click interface.

### 11.5.5 Implementation View

The ingredients for implementation of a successful online system must not only meet the technical and operational requirements, but must, in addition, contribute to a coherent overall architecture. Options available today are considerably different from those available only a few years ago. Rapid advances in electronics and computer technologies are likely to continue changing the landscape over the design and implementation period for the *BABAR* detector systems. Thus, these ingredients must exhibit sufficiently general properties so that they and other ingredients may be changed relatively freely and without requiring large architectural changes to the system. To this end, a preliminary selection of online components has been identified.

Industry standard VME will provide the backbone of the online system. VME is a mature technology in terms of reliability and availability of products. Its backplane performance is satisfactory for control applications and many data applications. Enhancements to a VME system are possible via VME-64 and/or VXI. Various auxiliary buses (*e.g.*, CAMAC) and cable interconnects (*e.g.*, MXI) can be used in conjunction with VME.

Single board computers residing in VME crates will assume the role of intelligent crate controllers. The need to provide portable online code implies the use of coding standards to avoid system-specific extensions, and the use of a real-time kernel to prevent code from being dependent upon the hardware specifics of a particular single board computer. The VxWorks kernel satisfies these requirements with excellent real-time performance, ANSI C/C++ language support, and compliance with the POSIX standards.

A framework within which to build an online system, perform overall control functions, and collect and analyze monitor data will be based on the EPICS package. EPICS is based upon an architecture of UNIX workstations and single-board computers running VxWorks connected by a network running the TCP/IP protocol. In its basic form, EPICS includes a communication system, a database, periodic monitors, a sequencer, and various device drivers for specific VME data acquisition cards. The toolkit also includes a number of applications for handling alarms, archiving data, editing the database, building new applications,

*etc.* Various third party software packages (*e.g.*, Mathematica and MatLab) interface to EPICS.

Another aspect of software development concerns the quality of design and maintainability. Consistent with overall computing guidelines, the online system will be based upon the object-oriented philosophy, utilizing the C and C++ programming languages. The ANSI C language will be used for compatibility with the EPICS package, while C++ will be used for higher level software components. CASE tools will be employed where applicable to assure sound software design.

Various networking technologies are candidates for the online system. The TCP/IP protocol is necessary for compatibility with the EPICS system and the rest of the SLAC site. Currently, both Ethernet and FDDI offer acceptable solutions for control and slow monitoring applications. It is expected that emerging technologies such as ATM (Asynchronous Transfer Mode), FCS (Fiber Channel Standard), and SCI (Scalable Coherent Interface) may become realistic candidates in the timescale of *BABAR*; such a solution might address both data acquisition and slow control needs. Bandwidth and latency are the primary issues facing the network system; these are discussed further in Chapter 10.

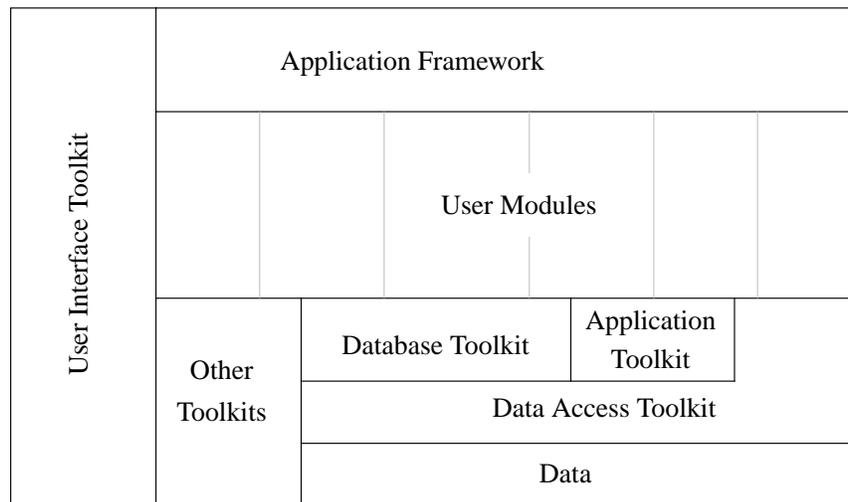
Generally, solutions for the online system will be chosen so as to be compatible with or identical to those for the offline computing system whenever reasonable. Mass storage, for example, will need to be equally available to both; thus a common network solution is desirable. Real-time event reconstruction will require executing the full offline code in the online environment, thus a standard workstation environment is needed.

## 11.6 Reconstruction and Analysis Framework

---

### 11.6.1 Scope

A common framework for application programs has to encompass several areas of deployment. For the data acquisition, trigger, and online systems, it must provide support for the monitoring processes and for the trigger algorithms within the online trigger farm. Within the offline environment, it must provide support for Monte Carlo simulations, the bulk processing and reduction of data, and physics analysis. This implies that the framework must provide support both for input of data from a variety of sources including disk and tape files and the online event server and for output to a similar set of permanent and transient destinations. Furthermore, it must provide an interface that is suitable for use in both interactive and batch environments, and ideally support both expert and novice users.



**Figure 11-9.** *Software framework.*

The question arises whether a framework is just a class library (using object-oriented terminology) or whether it encompasses greater functionality. The key distinction between a framework and an arbitrary collection of classes, however closely related those classes might be functionally, is that a framework describes not only the objects, but also their interactions with each other in a domain-specific manner. An application may therefore consist of several different frameworks, one providing access to the event data, another providing histogramming support, all under the overall control of the application framework. A framework that is available purely as a class library is called a toolkit.

### 11.6.2 Conceptual Overview

Our concept of an application framework is illustrated in Figure 11-9. User-written code, in the form of modules, can be incorporated with other modules under the overall control of the application framework, but with access to lower level frameworks and toolkits. Modules can be associated into paths, multiple paths being supported, where each path could correspond to a particular physics process, for example. The framework allows the user to select an input module and will then execute the modules in the various paths in the appropriate sequence. Filter modules may terminate or redirect processing of events in a particular path if the data do not meet the specified selection criteria. Events that reach the end of each path may be directed to a set of output streams, allowing for simultaneous stripping of selected data samples in the same pass through the data.

### 11.6.3 Toolkits

Several toolkits have been identified, including the application, data access, database, user interface, and visualization toolkits. The application toolkit is part of the framework itself and provides mechanisms whereby user code can find out attributes of the environment such as the run number and the presence of other toolkits. The data access toolkit provides access to data within an event from both C++ and Fortran, and the database toolkit provides access to the calibration, geometry, and other databases that specify the detector environment. The user interface toolkit allows users to prompt for operator input and to display quantities of interest, and the visualization toolkit provides a method whereby event data may be displayed graphically in a variety of modes. Thus, a track may be displayed in a Cartesian coordinate system, or as residuals from the fitted space points.

### 11.6.4 Framework

A framework that addresses the requirements is under development. It is written in C++ and uses the Tool Command Language (Tcl) [Ous94] as its primary user interface. It thus provides a flexible text-based interface that supports scripts and is therefore suitable for the expert user and for a batch mode of operation. It is intended that a graphical front-end to this framework based on the Tk package be created. Expect [Lib91] is a suitable candidate for this. The main focus of the development for this framework prototype is the data model prototype discussed elsewhere. In parallel with this development, several third-party application frameworks are being evaluated, including both public domain and commercial products. One problem appears to be their use in a batch environment, but they offer sophisticated user interfaces when used interactively.

### 11.6.5 Application Builder

The overall concept of the framework is that an application can be built that contains a large set of modules, only a few of which may be selected if so desired at run time. Thus, the same application may be used for reconstruction, diagnostics, or analysis depending on which modules are enabled. This model depends on there being dictionaries of standard modules that may be combined with user-written modules. It is possible that shared libraries may also be used to this end. Some of the application frameworks under investigation support the graphical creation of applications, and this capability has been shown to greatly improve the user friendliness of such systems.

### 11.6.6 Standard Applications

Several standard applications will be developed. One of these will be the stand-alone event display, although it is expected that an event display module for incorporation into other applications will also be available. Another application will support the dumping of event data in a variety of formats as a diagnostic aid. The applications running in the online trigger farm and for the main offline data processing may themselves be considered to be standard applications, and therefore available, in order that tests of the trigger efficiency and other systematic studies may be performed.

### 11.6.7 Job Submission

Much of the more time-consuming data processing will take place in a batch environment, and the application framework should interact seamlessly with the supported batch systems. Basing the framework prototype on Tcl was motivated by its availability as a scripting language within the UNIX shell environment as well as its use as a command parsing toolkit within applications. The generation of job submission scripts, spanning multiple jobs as well as single job submission, should therefore be a well-supported operation, capable of good integration into the overall batch environment.

### 11.6.8 Bulk Production

The submission of jobs for the bulk reduction of the raw data and generation of DSTs requires a level of support much greater than for casual job submission. Databases will need to track which data have reached which stages of the reduction, and which jobs failed and need further investigation. It might be possible to provide some tools to aid in quality checks over and above scanning of many histograms. Rapid feedback and high quality assurance are crucial to ensure that problems are detected quickly and do not affect the overall physics results.

---

## 11.7 Computing Support

---

### 11.7.1 Introduction

Support of the *BABAR* computing environment will be split essentially two ways. The collaboration will provide support for experiment-specific tools and the online system, as well as for offline analysis software. A substantial part of this effort will be handled using computing resources at remote institutions.

Infrastructure support including computing hardware, networking, and system software will be provided by the local computing center and/or departmental staff. SLAC Computing Services, for example, provides this support at the SLAC site.

### 11.7.2 Collaboration Support

The collaboration must design, maintain, and document the online and offline software framework, data model, analysis environment, and user interfaces to the various tasks. The personnel model is similar to recent typical approaches in HEP, in which physicists write most of the analysis and reconstruction code, including many of the peripheral tools and scripts, and software engineers address the other aspects where necessary, using appropriate commercial design tools. Examples include database technology and the design of the software framework. This model has the benefit that the physicists will have close contact with the physics aspects of the code, while the computer science and systems aspects are handled by people with the requisite expertise.

It must be stressed that the software for the *BABAR* experiment represents a very large system, and there are several major tasks that require special attention. The support for these activities is expected to consume a significant amount of manpower within the collaboration. Most of this manpower will consist of physicists and students, and each collaborating institution is expected to contribute a significant portion of its manpower to software development. The use of software engineers will have to be limited to those tasks for which they are essential, in order to minimize costs. Five full-time-equivalent (FTE) engineers are thought to be required. The following is a list of some of the major support tasks, most of which include some amount of infrastructure development, database technology, and coordination.

**Project Computing Engineer.** One of the FTEs mentioned is at a senior level to serve as the Project Engineer, responsible for the entire computing system.

**UNIX System Management.** There will be a UNIX system manager, responsible for the smooth running of the computing hardware, particularly online.

**Code Management.** Code coordinators will be appointed for all software packages, such as the online system, event reconstruction, and Monte Carlo. Their responsibilities include supervision, quality control, and management of production releases.

**Data Processing.** In the course of the experiment, a very large amount of data will be collected. The production of DSTs is therefore a major undertaking that needs careful management by data aides under the supervision of a physicist.

**Data Distribution.** Large amounts of data in DST form will need to be distributed to and received from various remote sites as discussed earlier. Data aides monitored by a physicist will perform this task.

**Database Management.** The *BABAR* experiment will have a large number and variety of databases. Database management is not a trivial task, and this will require at least one dedicated expert.

**Documentation and Training.** Documentation methods will be developed and enforced. Various topics may require training, such as C++ and databases. Identification of these topics and organizing training sessions is necessary.

**Tools and Utilities.** The experiment will require a great variety of tools and utilities. Some will be designed and written. All must be subsequently maintained.

### 11.7.3 Infrastructure Support

Infrastructure activities are expected to take place within all collaborating institutions, and these activities must be coordinated between institutions and with the collaboration. This coordination will be facilitated through a regional representative committee within the collaboration and through regular meetings between the collaboration and local support organizations. An example of the latter is already in place at SLAC where the majority of computing resources will reside. It is expected that similar arrangements will also be put into place at other regional centers.

An active discussion with SCS has been underway to identify the support that they will provide to the *BABAR* collaboration. A detailed proposal defining possible infrastructure support is given in Reference [SCS94] and is summarized below.

SCS will support the computing hardware, including CPUs, networking, and storage. This includes much of the offline environment and the network components at the detector. It also includes certain aspects of the Level 3 farm itself, which is envisioned to consist of general-purpose workstations. Hardware support includes planning, purchasing, installation, operation, tuning, and maintenance. Software aspects supported by SCS include system management of the UNIX workstations and the X/Motif graphical environment. In addition, a number of general programming and database tools will be managed by SCS, including compilers, editors, graphics, and Oracle. Software support also includes management of user accounts, disk and other resources, backups, a batch system, *etc.*

## 11.8 Integration Issues

---

The online computing system will be an important tool in the commissioning of the *BABAR* detector, and even of the PEP-II accelerator. The high-energy ring is scheduled to begin commissioning approximately a year prior to the completion of the low-energy ring. Background conditions will need to be studied during this period. The interaction region commissioning will involve simple detectors in advance of *BABAR* which will require some data acquisition. The essential components must, therefore, be ready relatively early, and this also has implications for the schedule for the control room and electronics areas.

The Level 3 data acquisition farm will include peripherals and CPUs, as well as a data link to the SLAC computer center. Operator interfaces will be discussed below. There already exists a fiber optic cable between the IR hall and the computer center that meets the requirements. The CPUs and peripherals should be installed in a clean, air conditioned environment with only occasional access necessary. The space for approximately four 19-inch racks is required—two for the CPU farm itself, and two for the peripherals.

Another crucial aspect for the effective running of an experiment is placement and number of computers and displays for control, monitoring, and debugging. Approximately ten monitors will be required in the control room to permit simultaneous development and control activities, and to dedicate screens to displays from the accelerator, environmental monitoring, event display, and run status. Several additional displays will be required in the electronics area to permit effective debugging and maintenance of the apparatus.

An external system that must be linked to experiment control is the PEP-II accelerator. The PEP-II Control System will be an extension of the SLC Control System. However, new developments such as longitudinal feedback and rf control will be based on EPICS, with

the IOCs (Input/Output Controllers) integrated and supported by PEP-II Controls. At this time, the role of the PEP-II Controls VAX is seen as limited to that of channel access client; data from the IOCs of the *BABAR* detector are thus directly available to PEP-II Controls. Conversely, information stored in the VAX database would be sent via the network to the *BABAR* Online System on a periodic basis (on a likely timescale on the order of minutes). This information would include beam position monitor data, beam energy, lifetime, current and emittance data, local pressure gauge readings from the ring vacuum system, *etc.* This data will be entered into the *BABAR* PEP-II database and be available for history plots and other display options.

Circumstances that may require a more immediate response than is possible with the network link to an EPICS controller are envisioned. One requirement is to assist in the process of beam tuning, both to reduce backgrounds and to increase luminosity. In response to a request, the online system will provide fast digital signals from detector chambers for noise metering, as well as signals from the detector luminosity monitor.

For proper control of both the detector and the accelerator, the state and status of many components must be monitored and the information shared between the online and the PEP-II Control Systems. Control over energizing the *BABAR* solenoid, over allowing access to the interaction region, and over allowing injection of beam into the ring, will be allocated on the basis of permissives exchanged between *BABAR* and PEP-II Controls. Information about the presence of colliding beams, the active logging of data, the value of the solenoidal field, *etc.*, will be made available to both control rooms. During injection, radiation damage to detector components will be minimized by reducing the injection rate if a threshold is exceeded in a PEP-II ionization detector at the IR.

## 11.9 System Responsibilities and Management

---

The overall responsibility for the *BABAR* computing system rests with two system co-managers and the computing project engineer. There will be designated people responsible for the various elements of the system, including, for example, code coordinators for the various packages, an online coordinator, *etc.*

Broad policy, communication, and planning issues are addressed by three committees:

- The detector subsystem (vertex chamber, drift chamber, *etc.*) reconstruction, simulation, calibration, and diagnostic code will largely be written by people closely involved in the design and construction of the subsystem. A committee of representatives from each of the subsystems will be responsible for ensuring a uniform approach within the context of the computing system and to allow communication of requirements.

- To ensure close and effective integration of computing efforts throughout the collaboration, a Regional Computer Representative Committee has been formed. This group brings together the Computing System Managers and Engineers plus a representative of each socio-geographical region in the collaboration. The group provides a forum for raising and solving problems of particular concern to collaborating institutions. It also provides a channel of communication from the Computing System to the widely spread computing effort of the collaboration.
- Finally, a planning committee exists with membership from both SCS and *BABAR*. This group is considering the support issues and resource requirements and is an important input to SCS for their planning and budget requests, and for coordinating which requirements will be the responsibility of the collaboration and which of the computing center.

## References

---

- [Aba94] F. Abar and W. Lidinsky, “The Effects of X-Window HEP Graphics Applications on ESnet,” presented at CHEP94, San Francisco, California (1994).
- [Ast92] D. Aston, A. Boyarski, A. Breakstone, D. Cords, T. Glanzman, F. Harris, T. Hung, P. Kunz, and F. Porter, “Computing for a *B* Factory,” *BABAR Note # 82*, (1992).
- [BaB94] The *BABAR* Collaboration, “Letter of Intent for the Study of CP Violation and Heavy Flavor Physics at PEP-II,” SLAC-443 (1994).
- [Bad90] A. Baden and R. Grossman, “A Model for Computing at the SSC,” Superconducting Super Collider Laboratory Technical Report No. 288 (1990)  
C. T. Day *et al.*, “Database Computing in HEP—Progress Report,” in *Proceedings of Computing in High-Energy Physics* (1992)  
C. T. Day *et al.*, “The PASS Project Architectural Model,” in *Proceedings of Computing in High-Energy Physics* (1994).
- [Boe94] C. Boeheim, K. Gounder, and R. Melen, “Computing Model Options for *B* Factory,” *BABAR Note # 154* (1994).
- [Bos94] Talk presented by Rebecca Bostwick at the ESnet Site Coordinating Committee at CEBAF (1994).
- [Boy90] A. Boyarski, T. Glanzman, F. A. Harris, and F. C. Porter, “Report of the Computing Group,” in *Proceedings of the Workshop on Physics and Detector Issues for a High-Luminosity Asymmetric *B* Factory at SLAC*, ed. by D. Hitlin, SLAC-373, LBL-30097, CALT-68-1697, 575 (1990).
- [Dou94] D. Doughty Jr., D. Game, L. Mitchell, G. Heyes, and W.A. Watson III, “Event Building Using an ATM Switching Network in the CLAS Detector at CEBAF,” presented at the International Data Acquisition Conference on Event Building and Event Data Readout, Batavia, Illinois (1994).
- [Che93] H. Chen, J. Hutchins, and J. Brandt, “Evaluation of DEC’s Gigaswitch for Distributed Parallel Computing,” SAND93-8013B, Sandia National Laboratories, California, (1993).
- [CLE94] Information concerning CLEO-II computing was kindly provided by M. Ogg, J. Urheim, and A. Weinstein.
- [Dre91] H. Drevermann *et al.*, in “Computing in High-Energy Physics ’91,” ed. by Y. Watase and F. Abe, Universal Academy Press, Tokyo, Japan (1991).
- [Ell92] Ellemtel Telecommunication Systems Laboratories, “Programming in C++: Rules and Recommendations,” Sweden (1992).

- [EPI92] “Experimental Physics and Industrial Control System (EPICS),” Argonne National Laboratory and Los Alamos National Laboratory (1992).
- [GIG93] Digital Equipment Corporation, “GIGAswitch Systems Manager’s Guide,” Part No. EK-GGMGA-MG.A01 (1993).
- [Gou94] K. Gounder, S. Bracker, K. Hendrix, and D. Summers, “A Simple Multiprocessor Management System for Event-Parallel Computing,” *BABAR Note # 128*, (1994).
- [HUB94] Alantec Corporation, “PowerHub Reference Manual,” vol. 2.2, Rev. C, Issue 1, Part No. 400-0336-000 (1994).
- [Lib91] “Expect: Scripts for Controlling Interactive Programs,” Don Libes, Computing Systems, Vol. 4, No. 2, University of California Press Journals (1991).
- [NCD89] “Network Loading and the NCD Network Display Station,” Note 3, available from NCD, (1989).  
F. Abar, C. O’Reilly, and E. Wicklund, “Network Load of X Terminals at CDF,” FNAL Report, (1992).  
R. Cottrell, “X Terminal White Paper,” SLAC (1993).
- [OMG91] Object Management Group, The Common Object Request Broker: Architecture and Specification, Revision 1.1, OMG TC Document 91.12.1 (1991).
- [OOM94] G. Booch, “Object Oriented Analysis and Design with Applications,” 2nd ed. (Benjamin Cummings, 1994) ISBN 0-8053-5340-2.  
D. Coleman, “Object Oriented Development, The Fusion Method” (1994)  
P. Desfray, “Object Oriented Modeling,” (Addison Wesley 1994).
- [Ous94] J.K. Ousterhout, “Tcl and the Tk Toolkit” (Addison-Wesley, 1994).
- [Por93] F. Porter, “BABAR Event Size Estimate,” *BABAR Note # 108* (1993).
- [Sax94] S. Saxena and T. Glanzman, *BABAR Note* (in preparation).
- [SCS94] “SCS Support of the B Factory,” *BABAR Note # 145* (1994).
- [Wal94] C. Walter and B. Nolty, “The FARFALLA User Guide, FARFALLA; V. 1.5” (1994).



---

---

## Interaction Region and Backgrounds

The PEP-II requirements of high beam currents, asymmetric energies, and head-on collisions conspire to make machine-induced backgrounds a significant challenge in the design of PEP-II and the *BABAR* detector. To achieve head-on collisions with asymmetric energies and 1.26 m bunch spacing, beam line elements must be positioned very close to the interaction point. The detector must be well protected from excessive occupancies and radiation damage, while maintaining a large solid-angle acceptance for  $\Upsilon(4S)$  decays. The expected sources of detector backgrounds are:

- synchrotron radiation photons produced in the machine magnetic elements;
- lost beam particles (bremsstrahlung and Coulomb scattering off residual gas molecules);
- luminosity-related backgrounds such as elastic and radiative Bhabha scattering; and
- other sources such as inelastic beam-gas interactions, cosmics, muons from collimators, etc.

These background sources can give rise to primary particles that either enter the detector directly or generate secondary debris that ultimately reaches the detector.

This chapter describes the various background sources and the results of simulations aimed at estimating the rates and impacts on the various detector components. Section 12.1 describes the interaction region design and the effect of background considerations on that design. Also in this section is a discussion of PEP-II operating modes, including issues relating to injection. Sections 12.2 and 12.3 describe the tools used to study the above backgrounds and estimates rates for each. Section 12.4 describes in detail the effect of the backgrounds on each of the major detector components in *BABAR*.

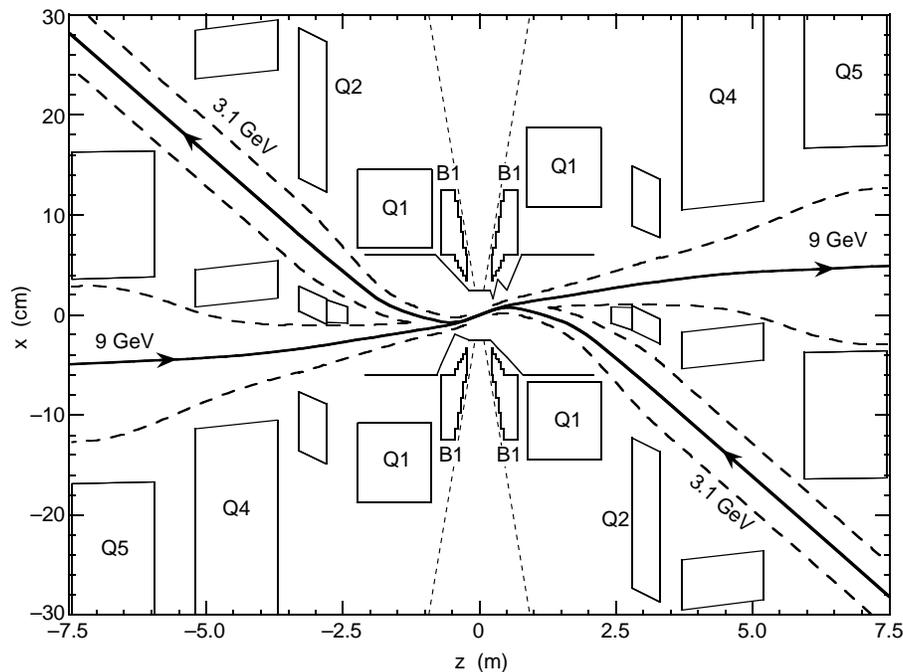
Parameter	HER	LER	Units
Center-of-Mass Energy	10.580		GeV
Peak Luminosity	$3.0 \times 10^{33}$		$\text{cm}^{-2}\text{s}^{-1}$
Luminosity Lifetime	1.55		hr
Bunch Spacing	1.26		m
Circumference	2219.3		m
Beam Energy	9.000	3.109	GeV
Bunch Length	1.0	1.0	cm
Number of Bunches	1658	1658	
Particles per Bunch	$2.73 \times 10^{10}$	$5.91 \times 10^{10}$	
Beam Current	0.986	2.140	Amps
Transverse Damping Time	36.8	54.0	msec
Horizontal Emittance	48.2	64.3	nm-rad
Horizontal IP Beta	50.0	37.5	cm
Horizontal IP Spot Size	155	155	$\mu\text{m}$
Horizontal Tune Shift	0.03	0.03	
Vertical Emittance	1.93	2.57	nm-rad
Vertical IP Beta	2.00	1.50	cm
Vertical IP Spot Size	6.2	6.2	$\mu\text{m}$
Vertical Tune Shift	0.03	0.03	

**Table 12-1.** *PEP-II parameters.*

## 12.1 PEP-II Design

### 12.1.1 Parameters

The PEP-II project is an  $e^+e^-$  colliding beam storage ring complex designed to produce a luminosity of at least  $3 \times 10^{33} \text{cm}^{-2}\text{s}^{-1}$  at  $E_{cm} = 10.58 \text{GeV}$ , equivalent to the mass of the  $\Upsilon(4S)$  resonance, with unequal energy beams of 3.1 and 9.0 GeV. The parameters of the accelerator are listed in Table 12-1. Both the high luminosity and the energy asymmetry require significant advances in storage ring design and construction. A full description of the PEP-II project may be found in the PEP-II Conceptual Design Report [PEP93].



**Figure 12-1.** Plan view of the interaction region. The low-energy beam enters from the lower right and exits from the upper left. The high-energy beam enters from the left and exits on the right. The vertical scale is highly exaggerated. The dashed lines represent the beam stay-clear envelopes, and the 300 mr detector acceptance cutoff.

### 12.1.2 Interaction Region Components

The high luminosity of PEP-II is achieved through high beam currents and strong focusing. The high current must be divided into a large number of low-charge bunches to avoid the beam-beam tune-shift limit, and the beams must collide only at the IP to avoid additional tune shift. PEP-II collides the beams head on but separates the beams horizontally before the next collision at 62 cm from the interaction point (IP), utilizing the energy asymmetry and separation dipoles (B1) within 20 cm of the IP. Strong focusing requires quadrupoles located close to the IP. The final quadrupole Q1 is common to both beams and starts at 90 cm from the IP, partially penetrating the central detector. The low-energy beam (LEB) is off-axis in Q1 to maximize the beam separation. The next quadrupole Q2 focuses only the LEB, with the high-energy beam (HEB) passing through a field-free region. The following pair of quadrupoles, Q4 and Q5, focus only the HEB, with the LEB passing through a field-free region. The layout of the PEP-II interaction region is shown in Figure 12-1. It is described in more detail in the PEP-II Conceptual Design Report [PEP93, pages 55–58 and 66–74].

The B1 and Q1 magnets are inside the 1.5T detector solenoid field. Conventional iron magnets do not function in this environment, so the only possibilities are superconducting and permanent magnets. There is not enough room between the beam pipe and the 300 mr detector acceptance envelope at the location of B1 to allow a cryostat, so the only option is a permanent magnet. In the PEP-II baseline design, Q1 is also a permanent magnet. (A superconducting Q1 design is presently under consideration. It would not change the IR optics in a fundamental way, nor would it change the luminosity, but it would allow greater operational flexibility at a significantly higher cost.)

B1 and Q1 are constructed of rings of SmCo permanent magnet material. Each ring is constructed of blocks with their magnetic moments normal to the beam direction and varying in azimuth around the beam in the manner described by Halbach [Hal81]. Permanent quadrupole magnets very similar to Q1 have been used at the Cornell Electron Storage Ring for many years [Her87a, Her87b].

The beam pipe at the interaction point is a double-wall structure, with an inner tube 50 mm in diameter and 800  $\mu\text{m}$  thick, separated by 2 mm from an outer tube of 400  $\mu\text{m}$  thickness, with helium gas flowing between the tubes for cooling. It is made of beryllium to minimize multiple scattering but is coated inside with 10 microns of gold to attenuate synchrotron radiation, totaling 0.6% of a radiation length. (The background studies have primarily assumed 25  $\mu\text{m}$  of copper rather than 10  $\mu\text{m}$  of gold.)

There are water-cooled masks inside the B1 magnet to prevent synchrotron radiation from striking the beam pipe. There are additional masks to protect the septum region of the Q2 quads from synchrotron radiation. The IP beam pipe and synchrotron radiation masks are described in the PEP-II Conceptual Design Report [PEP93, pages 357–371].

The two Q1 magnets, two B1 magnets, IP beam pipe, and vertex detector are assembled into a single rigid support barrel of 43 cm outside diameter, shown in Figure 12-2. The designs for B1, Q1, Q2, and the support tube are described in Reference [PEP93, pages 276–301].

The support barrel is assembled and internally aligned outside the detector and installed by sliding it through the drift chamber. The two end sections of the support barrel are made from stainless steel and carry the Q1 magnet assemblies. The middle barrel section is a carbon-fiber composite equivalent to  $0.005X_0$  in thickness. The total assembled weight is approximately 2500 kg, concentrated in the end sections.

### 12.1.3 Background Implications

The bending near the IP required to separate the beams generates a large synchrotron radiation flux that is not present in more conventional  $e^+e^-$  colliders. The bending also sweeps off-energy beam particles into the detector. There are also many more bunch crossings



### 12.1.4 Operating Modes

Most background calculations are based on stable colliding beams. The dominant beam loss mechanism is then beam-gas scattering distributed around the rings, although beam-beam bremsstrahlung at the IP can become comparable at high luminosity. The high dispersion in the bending arcs prevents damaging products of distant beam-gas bremsstrahlung from reaching the detector. Beam-gas bremsstrahlung in the upstream part of the IR straight section is stopped by collimators, but bremsstrahlung close to the detector produces irreducible backgrounds. Beam-gas coulomb scattering far from the detector is also absorbed by collimators and in general is less of a problem than bremsstrahlung. Most synchrotron radiation is directed away from the detector, and the remaining radiation is largely absorbed by masks. There is an irreducible background from scattering off the tips of the masks.

On the order of once per hour, it will be necessary to inject into PEP-II. The injection process and components are described in Reference [PEP93, pages 483–560]. Normal injection will be in top-off mode, *i.e.*, charge will be injected without dumping the stored beam. This mode and re-injecting after a total beam loss are projected to require on the order of six minutes to complete. The beams will be injected at collision energy into collision optics, with no necessity to ramp or squeeze the beams. The beam emittance provided by the injector (the Stanford Linear Collider with some modifications) will be smaller than the equilibrium emittance of PEP-II. Therefore, injection will be possible with all collimators at the settings assumed for background calculations.

A beam particle damps from its injected orbit to the stored orbit in a small fraction of a second, compared to the expected lifetime of many thousands of seconds. Thus, any radiation dose from the capture process is expected to be much less important than radiation from the normal stored condition. There is no known reason for the injection process not to be highly efficient, or for injection losses to produce larger detector radiation doses than losses from the stored beam. Since injection will take place when a few tens of percent of the stored beam has been lost, the dose from injection will not dominate unless the injection inefficiency is worse than tens of percent. There is a sufficient safety margin in the radiation damage thresholds that, if injection resulted in a dose comparable to normal running, this would be acceptable.

Signals from the detector or local radiation monitors will be used as inputs to the PEP-II beam-dumping system. If a rapid increase in radiation is observed, the stored beam will be dumped into a distant absorber. Such systems are in use already by the LEP detectors. The injection process will also be interrupted if radiation is observed to be abnormally high. It is assumed that particularly sensitive detector components like the drift chamber will be ramped down during injection to minimize radiation damage, and injection will be dependent on these components being ready.

---

## 12.2 Tools

---

### 12.2.1 QSRAD

The program QSRAD was built to study synchrotron radiation generated by the beam as it traveled through the last two focusing magnets prior to the collision point in PEP-I. The program traces weighted rays from a Gaussian beam profile through the specified magnetic optics and produces a geometric fan of synchrotron radiation with equal power density and constant critical energy for each magnetic element. These fans are then traced, and a tally is made of the fraction of each fan that strikes defined apertures. Apertures are specified in planes transverse to the beam. An aperture can be a circle, an ellipse, a horizontal slit, or a vertical slit. The distribution of photon critical energies is accumulated for each aperture and converted to a photon energy distribution. In addition, the beam rays that strike each aperture are recorded, and a table of  $x, y$  locations is produced, indicating the origin in the beam profile for those particles.

An enhanced version of QSRAD (called SYNC\_BKG) is used to calculate *B* Factory synchrotron radiation rates. The following improvements to SYNC\_BKG have been implemented: a non-Gaussian beam tail distribution can be added to the beam profile. The beam tail distribution is modeled as a lower flatter Gaussian beam profile; tilted and offset quadrupoles are allowed; horizontal and vertical bend magnets are recognized as magnetic elements; and individual magnetic elements can be subdivided to more accurately simulate long quadrupoles.

The information from SYNC\_BKG on photon rates striking various surfaces near the IP is used in a second stage to study backscattering and tip scattering from the mask surfaces. EGS4 [Nel85] (which includes Rayleigh scattering and K-shell fluorescence) is used to calculate the number and energy of photons that scatter from a mask tip, penetrate the detector beam pipe, and are absorbed in various layers of the silicon vertex detector.

### 12.2.2 Decay TURTLE

Decay TURTLE [Car91] is a program for tracking particles and decay secondaries through beam line components. A version of this package has been modified to do bremsstrahlung and Coulomb scattering. A scattering point is picked randomly along the beam line, a beam particle transported to that point, and the scattered secondary propagated down the remaining beam line until it hits an aperture or the end of the beam line. For bremsstrahlung scattering the photon is also propagated. Scoring planes are used to simulate collimators, masks, and other apertures. Selected particles, particularly ones hitting an aperture near

the IP, can be written to a disk file for later input to the GEANT or OBJEGS detector simulations described below.

In the studies presented here, the PEP-II machine lattice is simulated from the middle of the preceding arc to several meters downstream of the IP ( $\sim 185$  m). Placement of collimators in the interaction region reduces the rate of upstream scatters reaching the IP. A graded aperture is employed in this region of the machine lattice ( $\pm 60$  m from the IP), *i.e.*, progressively larger apertures are used as one approaches the IP. Multiturn effects have not yet been studied. Optimization and design of upstream collimation is in progress subject to space constraints and the avoidance of significant synchrotron power dissipation on the collimator.

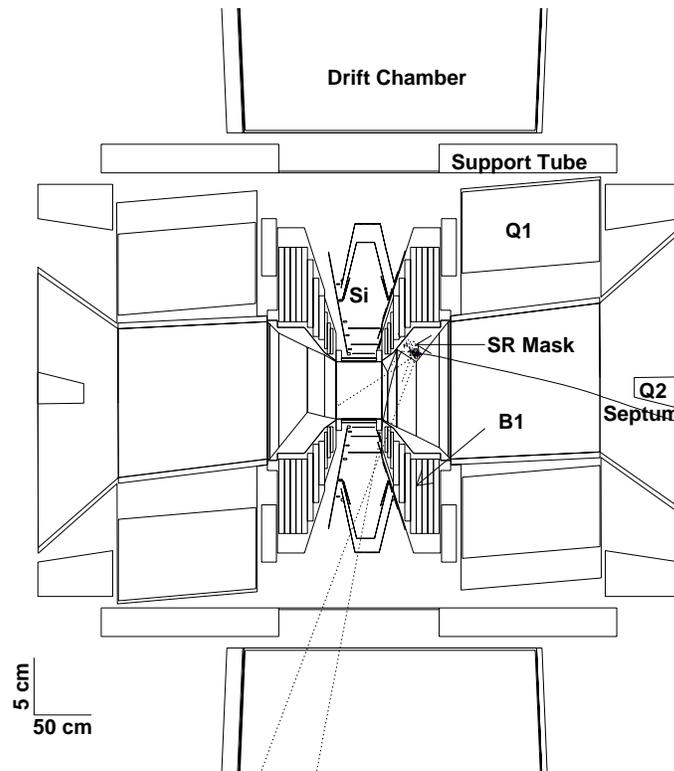
### 12.2.3 OBJEGS

OBJEGS [Hea91] interfaces EGS4 [Nel85] to simple cylindrical geometries. The user specifies beam line and detector components in terms of the inner and outer radius, lower and upper longitudinal limits, and material type. Files of lost particles from TURTLE or other sources of background such as radiative Bhabhas can be input from a file one particle at a time. Integration over time must be done as a separate step operating on the output of OBJEGS. This can be as simple as adding up total energy deposits to obtain the radiation dose or as complex as combining hits produced by multiple rays to make events for a trigger simulation.

The restriction to cylindrical geometries in OBJEGS makes it impossible to describe the *BABAR* detector perfectly, especially as the detector design has grown more sophisticated and realistic with features such as conical drift chamber end plates and variable crystal lengths. The biggest problem for OBJEGS geometries is the elliptical shape of the synchrotron radiation mask; to approximate this, two versions of OBJEGS geometries are used—one that reproduces the longitudinal profile of the low-energy beam synchrotron mask, and one that reproduces the high-energy synchrotron mask. The geometry corresponding to the side of the detector which is struck by a particle is used for the simulation of the shower produced by that particle. Due to these complications, *BABAR* has been moving toward a GEANT-based simulation (Section 12.2.4) as its primary detector simulation tool. A number of the studies in this document were performed using OBJEGS, but will be redone with GEANT as time permits.

### 12.2.4 GEANT

An extensive GEANT [GEA94] simulation of the *BABAR* detector, called BBSIM, is currently the primary tool for detector performance studies. As part of this implementation, there is



**Figure 12-3.** *GEANT simulation of beam line components. The interaction of a LEB lost beam particle is also shown (with a high shower cut-off energy of 10 MeV).*

a rather detailed simulation of the beam line components. Figure 12-3 shows an expanded view of this central part of the *BABAR* detector with a lost beam particle generated with the TURTLE package described above.

The simulation includes an approximation to the synchrotron radiation masks, which have an elliptical aperture, and the B1 and Q1 beam line magnet elements. The dipole and quadrupole fields of the B1 and Q1 (both SmCo permanent magnets), respectively, are also included. At the ends of the beam pipe are the Q2 septum masks that shield the Q2 coil.

Files of lost particles or radiative Bhabhas can be input to the simulation, either one particle scatter at a time or integrated over some detector time window and superimposed on a physics event.

### 12.2.5 GELHAD

GELHAD [Sny94] is a program for simulating the electro-production of hadrons via  $\gamma N$  interactions. The interactions of electrons with nuclei can be added using a virtual photon approximation; however, hadronic interactions in electromagnetic showers are dominated by real  $\gamma N$  interactions.

GELHAD is based on subroutines from the hadronic shower simulation GHEISHA [Fes85]. Two basic models are available: (1) the photon is absorbed by one nucleon that is then forced to interact with the residual nucleus; and (2) the photon is replaced by a pion of the same energy that interacts with the nucleus. In both models, energy conservation and charge conservation are respected, but in general, momentum conservation is violated. The default used in most of the studies for this document uses model (1) below  $E_\gamma = 150$  MeV and model (2) above.

GELHAD has been implemented both in OBJEGS and GEANT. Trigger studies with the OBJEGS version use the supplementary program TRACK [Sny94] to simulate the behavior of hadrons in OBJEGS geometries. GEANT-based studies have, of course, the full power of GEANT and the detailed BBSIM geometry available.

## 12.3 Background Sources

---

### 12.3.1 Synchrotron Radiation

There are several sources of synchrotron radiation backgrounds:

- direct synchrotron radiation;
- photons that scatter off a mask tip;
- synchrotron radiation from elements far upstream of the interaction point; and
- sources of backscattered photons from downstream surfaces.

Separating the unequal-energy beams by the use of bending magnets and offset quadrupoles generates several fans of synchrotron radiation. The geometry of the interaction region optics, however, is designed to minimize the amount of synchrotron radiation that strikes nearby surfaces. In particular, the S-bend geometry of the beam lines allows most of the synchrotron radiation generated by magnetic elements upstream of the interaction region

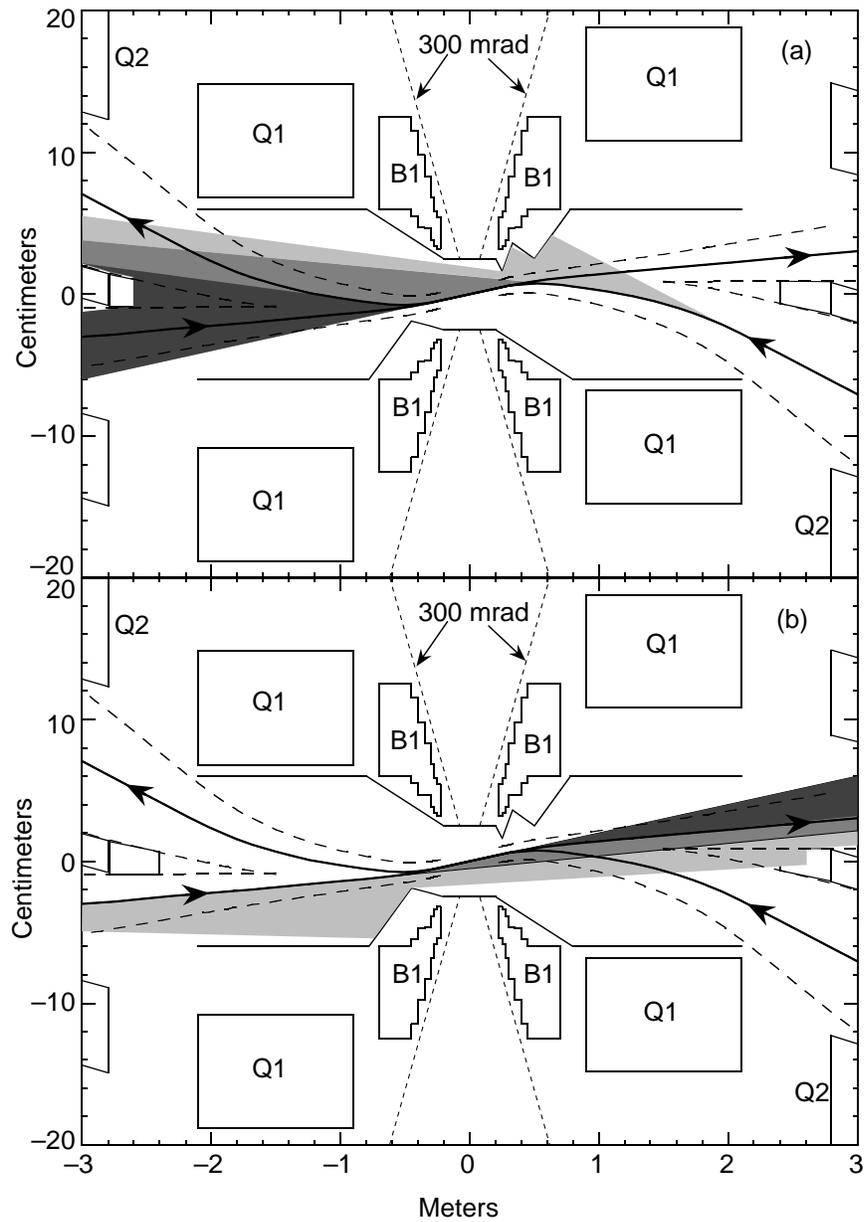
to pass through the detector region without hitting local surfaces. Primary masks near the collision point are used to prevent direct radiation from hitting the detector beam pipe, while at the same time keeping the number of photons that strike the mask tips to an acceptable level. The synchrotron radiation fans from both beams are shown in Figure 12-4.

The LEB generates synchrotron radiation fans as it passes through the Q1 and B1 magnets on its way to the collision point. The LEB mask is designed to prevent the synchrotron radiation (either fan or quadrupole) generated by the upstream magnets from directly striking the detector beam pipe. The surfaces of the LEB mask are sloped such that scattered photons cannot travel directly to the detector beam pipe. The LEB mask absorbs about 3.6 kW of synchrotron radiation power. Fans generated by the LEB in the two B1 magnets and in the downstream Q1 magnet pass through the interaction region without striking nearby surfaces. The first surface that intercepts the B1 fans is the septum mask in front of the Q2 septum quadrupole, located 2.8 m from the IP. This septum mask absorbs 3.5 kW of power.

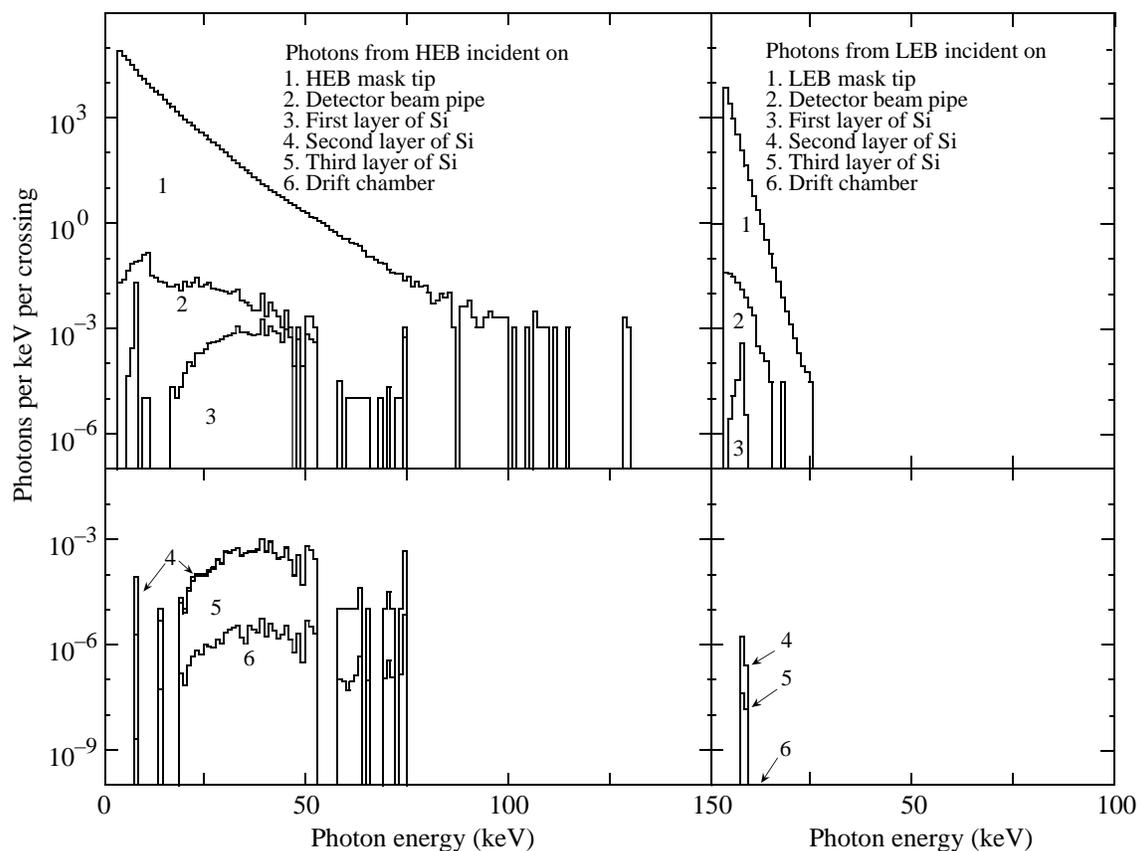
Synchrotron radiation fans generated by the HEB in upstream bend magnets and offset quadrupole magnets deposit nearly 1 kW of power on the HEB mask. These upstream HEB elements are positioned to ensure that radiation generated by beam particles even  $10\sigma$  off-axis in these elements does not strike the LEB mask. The fans of radiation generated by the HEB as it passes through the two B1 magnets do not strike any surfaces in the detector region and are absorbed in a dump about 17 m downstream of the detector.

The geometry of the masking near the IP is such that no synchrotron radiation can hit the IP beam pipe directly, nor can it scatter off the face of a mask onto the IP beam pipe. Scattering from the tips of the LEB and HEB masks is the dominant source of photons striking the detector beam pipe. This process has been simulated in great detail. The individual photons that strike the LEB and HEB mask tips are generated from an energy spectrum given by a program that traces the photons from sources to surfaces of interest. The photons that end up striking the detector beam pipe are followed through the detector components by the EGS Monte Carlo program [Nel85]. Further details of this simulation package can be found in Reference [PEP93, pages 108–112]. The results of the simulations are the photon spectra shown in Figure 12-5; backgrounds resulting from synchrotron radiation are shown in Table 12-2.

Other sources of synchrotron radiation, such as magnets farther upstream, and backscatter from the Q2 septum masks, produce negligible contributions to the detector background compared to primary mask tip scattering. While a high photon flux strikes the dump for HEB radiation from the B1 magnets, the small solid angle of the detector beam pipe as seen from this source, the photon's small angle of incidence on the beam pipe, and the fact that most of the beam pipe is shielded by the LEB mask, together reduce the background from the 17 m dump to a low level.



**Figure 12-4.** Synchrotron radiation fans from the low-energy (a) and high-energy (b) beams. The density of shading gives an indication of the relative photon intensity from the various radiation fans.



**Figure 12-5.** Photon spectra from synchrotron radiation. Left plot is for high-energy beam; right plot is for low-energy beam.

The effect of magnet misalignments on detector backgrounds has also been evaluated. Of the 24 different cases studied, two cases, a  $\pm 5$  mm displacement in  $x$  of Q1, Q4, and Q5, produced a threefold increase in the background rate. The rest of the misalignment checks produced small ( $< 50\%$ ) increases in backgrounds, with some configurations producing rates that are actually below the nominal background rate. None of these misalignment checks resulted in synchrotron radiation photons directly striking the detector beam pipe.

### 12.3.2 Lost Particles

Bremsstrahlung and Coulomb scattering of beam particles from residual gas molecules in the beam pipe can result in high-energy beam particles and photons striking masks and the beam pipe near the IP. The resulting electromagnetic showers can cause excessive detector

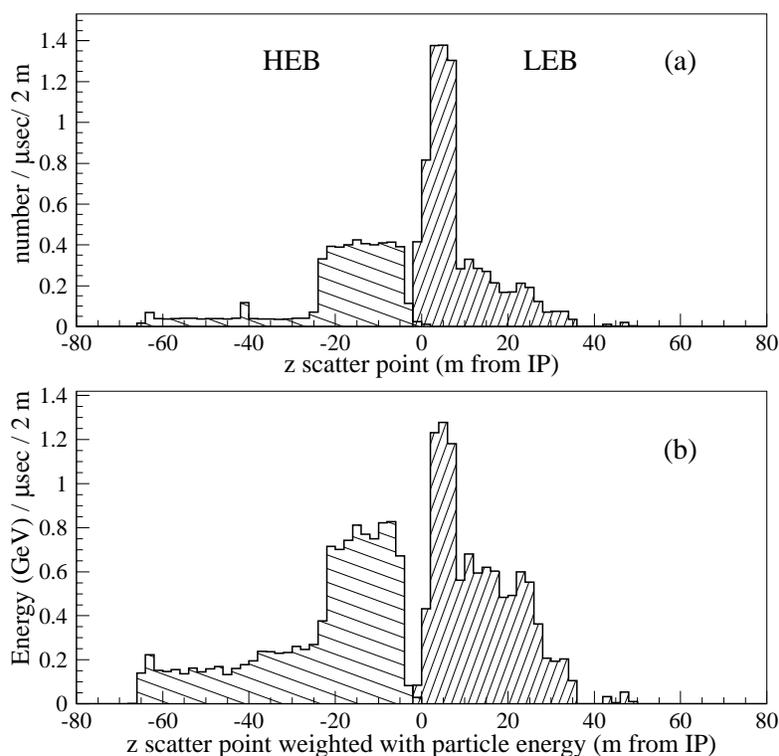
	Incident on Be Pipe	Absorbed in First Si Layer	Absorbed in Second Si Layer	Absorbed in Third Si Layer	Incident on Drift Chamber
$4 < E_\gamma < 100$ keV					
# of Photons	1.24	0.028	$6.9 \times 10^{-4}$	$5.6 \times 10^{-4}$	$2.4 \times 10^{-5}$
Energy (keV)	17.6	0.43	$2.2 \times 10^{-2}$	$2.0 \times 10^{-2}$	$3.2 \times 10^{-3}$
$4 < E_\gamma < 20$ keV					
# of Photons	0.96	0.022	$3.8 \times 10^{-5}$	$8.6 \times 10^{-6}$	$2.1 \times 10^{-7}$
Energy (keV)	9.34	0.17	$9.2 \times 10^{-4}$	$1.3 \times 10^{-4}$	$3.7 \times 10^{-6}$
Detector Limits					
# of Photons	—	2.3	$> 2.3$	$> 2.3$	125
Energy (keV)	—	95	$> 95$	$> 95$	5000

**Table 12-2.** Synchrotron radiation background simulation results. The numbers are for each crossing. Multiply by 238 to get photons per  $\mu$ s. In the simulation, the beam pipe consists of  $25 \mu$ m of Cu and 1 mm of Be.

occupancy and/or radiation damage. Hadronic showers from photoproduction can also occur in these showers and are discussed in Section 12.3.3.

Section 12.2.2 describes Decay TURTLE, the tool used to establish the rate of particles incident on the masks and beam pipe near the IP. The rate is reduced significantly with upstream collimation. Recent studies assume a uniform beam line pressure of 1 nTorr ( $N_2$  equivalent) to determine regions of sensitivity to pressure. Additional pumping in these regions can then be employed to further reduce this background. Figure 12-6(a) shows the location of the beam particle scatter for particles that end up striking near the IP. Rates are quoted per  $\mu$ s, a typical integration time for detector elements. Since higher energy particles deposit more energy in the detector, Figure 12-6(b) shows the same distribution weighted with the energy of the scattered particle.

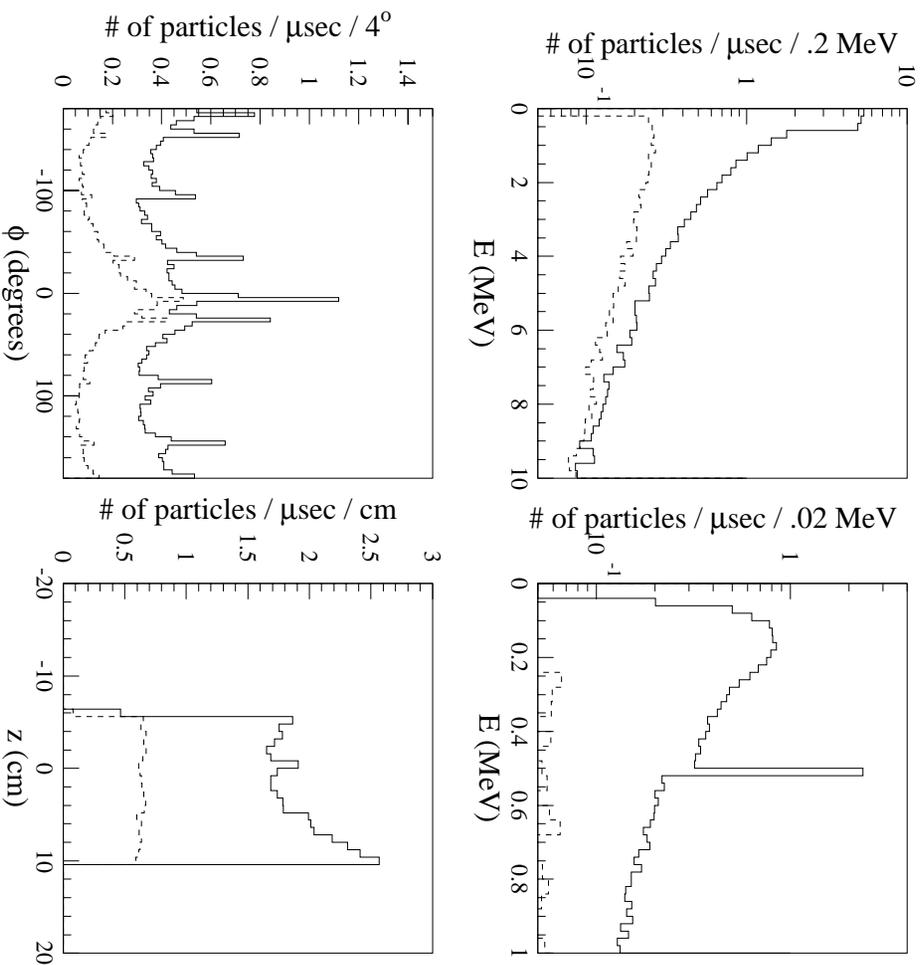
Studies are in progress to determine a realistic estimate of the IR vacuum pressure profile. Preliminary results suggest that, on average, the assumed uniform 1 nTorr distribution is a reasonable baseline for these studies. It appears possible to reach pressures below 1 nTorr (0.2–0.5 nTorr) in the regions from 3 to 40 meters from the IP in both the incoming HEB and LEB beam lines. Due to space constraints at the IP that preclude placement of pumps very close to the IP and the high heat load on the synchrotron radiation masks, the pressure within three meters of the IP will likely exceed 1 nTorr. Scatters very close to the IP ( $\pm 1$  m) will tend to make it out of the detector region before striking an aperture. Average background rates for these preliminary pressure profiles tend to be comparable to the rates found with the uniform 1 nTorr assumption, so this is taken to be the *nominal* level. Assuming this baseline, the HEB deposits 12.9 GeV per  $\mu$ s from bremsstrahlung scattering and 1.0 GeV



**Figure 12-6.** Distribution of the scatter point of beam particles that hit near the IP (a) unweighted and (b) weighted with the energy of the particle. A uniform 1 nTorr vacuum pressure is assumed.

from Coulomb scattering within  $\pm 1.5$  m of the IP. The LEB contribution is 9.2 GeV from bremsstrahlung scattering and 1.4 GeV from Coulomb scattering.

GEANT is used in the detector simulation package BBSIM to evaluate the impact of this energy deposition near the IP. Secondaries from electromagnetic showers in the beam line components are followed into the detector elements. The showers are cut off for electron and photon kinetic energies below 50 keV. Figure 12-7 shows the resulting energy, azimuthal, and  $z$  distributions for secondary electrons and photons incident on Layer 1 of the silicon vertex detector. The peak at 0.5 GeV in the energy distribution is due to  $e^+e^-$  annihilation. The strong peaks in the  $\phi$  distribution at  $0^\circ$  and  $180^\circ$  are from bremsstrahlung scattering in which the B1 dipoles bend the energy-degraded beam particles horizontally into the beam pipe. The other peaks in this distribution are from double counting the flux in regions of overlapping silicon detectors. Figure 12-8 shows the same distributions for Layer 1 of the central drift chamber. Since few photons interact in the drift chamber, the flux into the particle identification system and CsI calorimeter is nearly identical to the flux entering

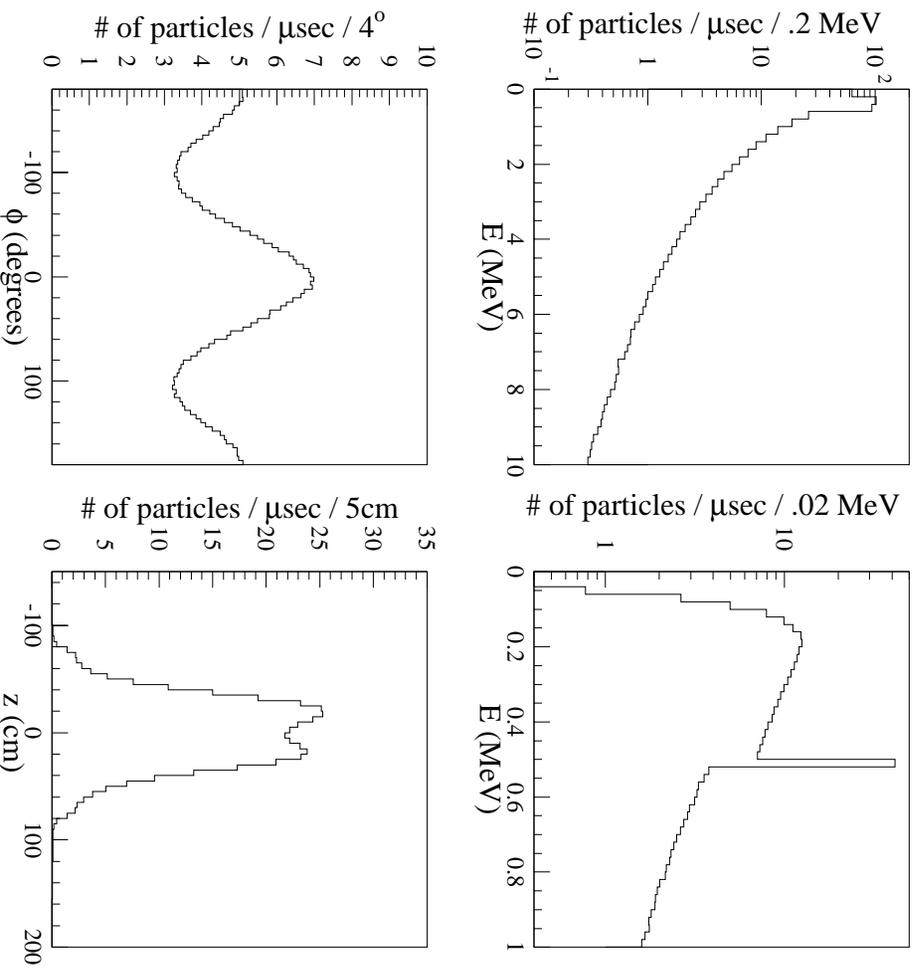


**Figure 12-7.** Energy, azimuthal, and  $z$  distributions of photons (solid) and electrons (dashed) incident on Layer 1 of the silicon detector for lost beam particle backgrounds. Rates correspond to the number of incident particles per  $\mu\text{s}$ .

the drift chamber. Figure 12-9 shows the distribution for the front face of the calorimeter, including the polar angle distribution of photons and energy flow.

### 12.3.3 Hadrons

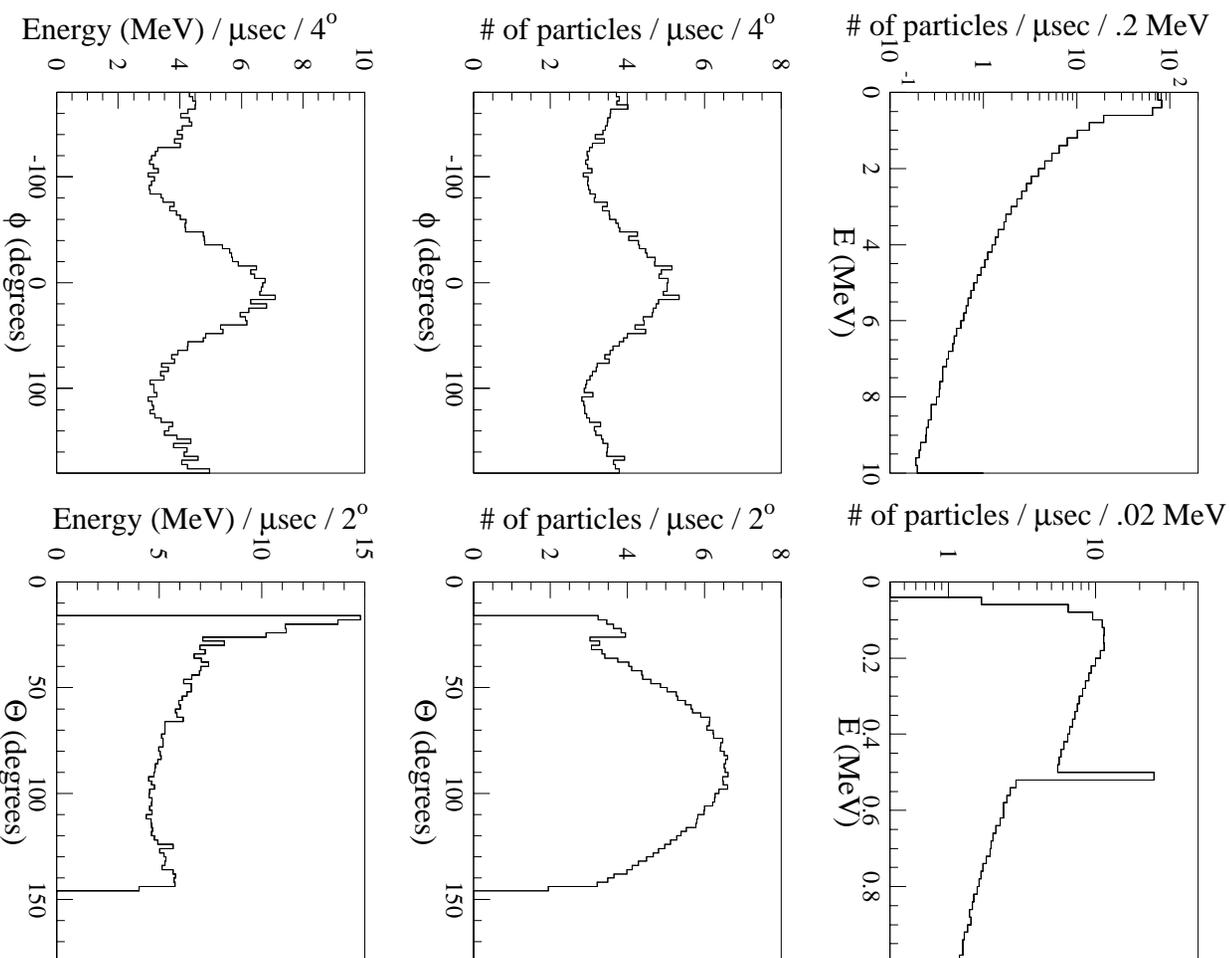
Estimates of trigger rates from photoproduced hadrons are given in Chapter 10. In this section, we seek only to check (using a simplified trigger simulation) whether or not the triggers are produced by photon interactions that could be expected to be well-modeled by GELHAD.



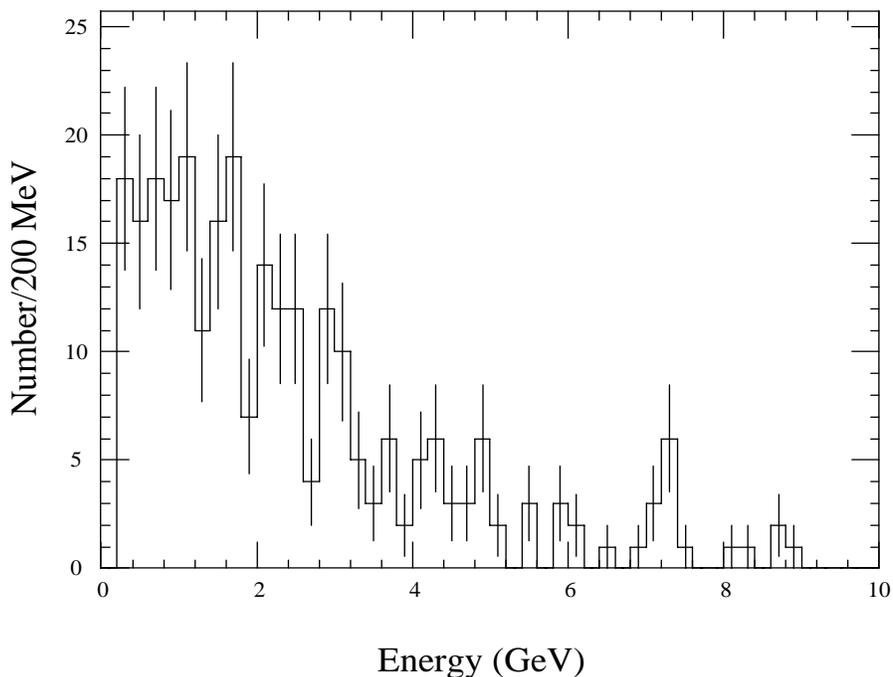
**Figure 12-8.** *Energy, azimuthal, and  $z$  distributions of photons incident on Layer 1 of the central drift chamber for lost beam particle backgrounds. The electron rates are negligible. Rates correspond to the number of incident particles per  $\mu\text{s}$ .*

Figure 12-10 shows the distribution of the energy of photons in electromagnetic showers that interact hadronically; only those photons that produced triggers are modeled. The GEANT version of GELHAD has been used to make this distribution. The trigger in this case is a one and one-half track trigger in which at least one track reached the outermost layer of the drift chamber and the other track reached at least the middle layer of the chamber. Note that the spectrum is rather hard, with an average energy of 2400 MeV. This gives us some confidence that our model, which at least conserves charge and energy, may do a reasonable job in the relevant region.

The process of comparing GELHAD with data from the PEGASUS [Deg94] test run using the TPC detector at PEP has just begun and will be vigorously pursued in the next few



**Figure 12-9.** Energy, azimuthal, and polar angle distributions of photons incident on the front face of the calorimeter. Also included are the angular distributions of energy flow. Rates correspond to number of incident particles and energy per  $\mu\text{s}$ .



**Figure 12-10.** *The energy distribution of photons in electromagnetic showers which interacted hadronically, causing a charged track trigger.*

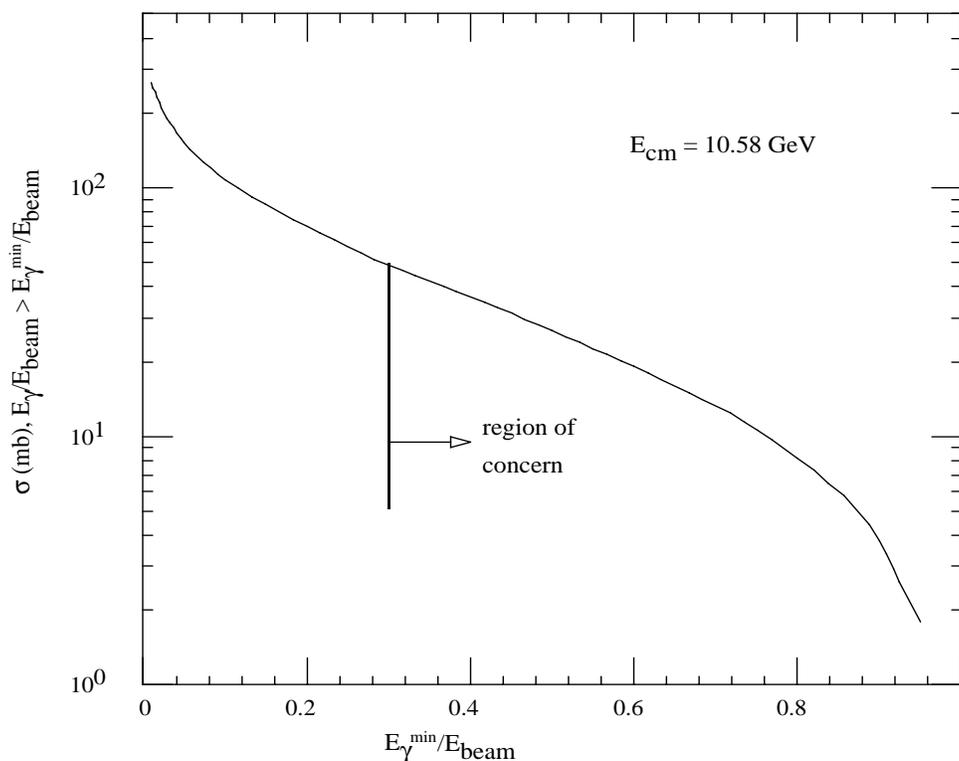
months. We are also actively investigating alternative models as a check of the current results.

At this stage in its development, it is difficult to evaluate the errors on GELHAD results. Using the model in which the photon is absorbed by a nucleus (model 1) at all energies instead of the model in which the photon is replaced by a pion (model 2) above 150 MeV results in a factor of 3 drop in the drift chamber trigger rate, while the calorimeter trigger rate remains virtually unchanged. The drop in the drift chamber trigger rate can be traced to a substantially lower pion multiplicity in model 1 compared to model 2. Only a program of comparison with data can establish the uncertainties and lead to an improved model.

### 12.3.4 Luminosity Backgrounds

#### Radiative Bhabhas

The radiative Bhabha process,  $e^+e^- \rightarrow e^+e^-\gamma$ , is a potentially large background because of its relatively large cross section. Outgoing degraded particles from the IP in each beam are

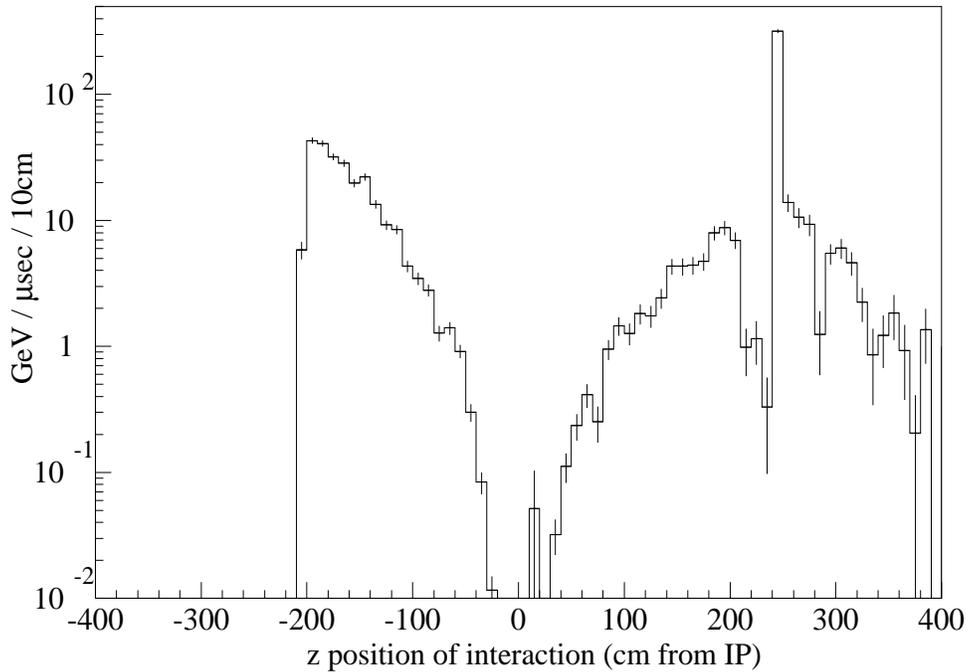


**Figure 12-11.** Radiative Bhabha integral cross section versus minimum fractional photon energy.

swept into the beam pipe by B1 and Q1. Figure 12-11 shows the radiative Bhabha cross section versus the minimum energy of the radiated photon that caused the degraded particle to hit the beam pipe, normalized to the beam energy.

To study the contribution of radiative Bhabhas to occupancy and radiation damage, we have used a Monte Carlo event generator called BBBREM, developed for LEP by Kleiss and Burkhardt [Kle94], and modified for an asymmetric collider. Individual events generated at the IP by BBBREM were input to OBJEGS and tracked until they hit beam line components and showered. Figure 12-12 shows the energy deposited per  $\mu\text{sec}$  in the beam pipe and other components (the peak at 240 cm is from radiative Bhabhas bent by B1 into the Q2 septum).

For particles that hit the beam pipe between the IP and the end of Q1 (*i.e.*, those less than 2.3 GeV), the rate from the LEB is about 5 times higher than that from the HEB due to the difference in cross section. However, in the HEB, there is an additional source from degraded electrons in the energy range 5.0–6.5 GeV that impinges on the Q2 septum at 2.4 m from the IP.



**Figure 12-12.** *Energy deposition from radiative Bhabhas versus distance from interaction point (in cm).*

Within  $\pm 2.5$  m of the IP, the power deposited by radiative Bhabhas is approximately two orders of magnitude more than the power from lost beam particles. Fortunately, the higher energy particles are hitting the beam pipe farther from the IP, and the resultant electromagnetic showers are mostly directed away from the vertex detector and central drift chamber.

To determine the effect of radiative Bhabhas on the detector, they were included in GEANT with lost particles from beam-gas bremsstrahlung and Coulomb scattering. With the addition of high- $Z$  shielding around Q1 and the magnetic shielding around Q2, the rate of photons into the detector was found to be small compared to the lost particle contribution.

### Elastic Bhabha Scattering

The cross section for elastic Bhabha scattering with angles large enough to cause the scattered particles to hit the beam pipe within Q1 is about four orders of magnitude smaller than the corresponding radiative Bhabha cross section. In terms of power lost within  $\pm 2.5$  m of the IP, elastic Bhabhas are at least two orders of magnitude lower than both radiative Bhabhas and lost particles. Therefore, we consider background from this process negligible.

## Two Photon Reactions

We neglect this process, in which the  $e^+$  and  $e^-$  both radiate in the same interaction, since its cross section is some 1000 times smaller than the radiative Bhabha process [Bai81], and the presence of two degraded beam particles arising from the same beam crossing does not give rise to any special background condition.

## Landau-Lifshitz Pairs

In this process, an  $e^+e^-$  pair is produced at the IP, and a low-energy particle with fairly large transverse momentum may hit the Be beam pipe [Che91]. The magnetic shielding from the 1.5 T solenoid greatly reduces this background contribution to the point where it is negligible.

## 12.3.5 Other Sources

### Beam-Gas Interactions at the IP

Electrons and positrons in the cores of the beams scatter off molecules inside the beam pipe near the interaction point, producing hadrons. This process does not include the more severe backgrounds due to lost-particle hadronic interactions. Rates of triggers were estimated using the electro-production code EPC [Lig88] interfaced to the detector GEANT simulation BBSIM.

The EPC generator program models the double-differential nucleon production cross section due to electrons scattering off thin targets of nuclei. The accuracy of the program is about 20% [Lig88]. Only single-particle cross sections are produced by the program. The EPC program was used to make tables of double-differential cross sections for each type of final-state hadron for beam electrons and positrons incident on a thin target of nitrogen molecules.

The single-particle final states  $p$ ,  $n$ ,  $\pi^0$ ,  $\pi^+$ , and  $\pi^-$ , were generated and tracked by GEANT through the detector. The simulation was lacking in that two-particle states were not generated. The main contribution to true two-track final states is expected to be from  $\Delta \rightarrow p\pi$ , which is the sole source of pions in the EPC program. To obtain a high upper limit of the two-track drift-chamber Level 1 trigger rate due to  $\Delta$  decay, it was assumed that all protons that give rise to a track were accompanied by an A-track pion. (An A track is defined as a particle track which passes through all layers of the drift chamber, while a B track passes through only the inner half of the chamber.) That is, the upper limit to

the two-track trigger rate was assumed to be the proton-induced B-track rate (including A tracks). A similar upper limit was estimated for the calorimetric trigger.

The estimated trigger rates due to beam-gas interactions near the interaction point [Kra94b] are given in Chapter 10. The upper bound for the total Level 1 rate was found to be comparable to the physics rate for a vacuum of 10 nTorr and thus is tolerable, with a large safety margin. The trigger rate due to beam-gas events would be significantly lowered by a trigger level that selects the track  $z$  position to be near the interaction point, such as a trigger using the vertex detector. The accuracy of the beam-gas trigger-rate calculation could be improved through the use of a program that provides final-state particle correlations.

### Cosmic Rays

Cosmic rays are caused by high-energy particles interacting with the atmosphere, producing pions and kaons that decay into muons, and producing showers. At ground level, the majority of penetrating particles are muons.

Cosmic rays are important to *BABAR* in two ways: they cause about 500 Hz of Level 1 trigger rate when using a trigger line containing only a loose calorimeter definition, and they can be used to calibrate elements of the detector.

The cosmic-ray muon generator HEMICOSM [Bri94] was developed for study of trigger rates and for future study of the detailed flux incident on various detector components. The generator HEMICOSM was interfaced to BBSIM and used to study detector response. The program uses an accurate theoretical flux prediction that varies as a function of both momentum  $p$  and zenith angle [Dar84]. The ratio of  $\mu^+$  to  $\mu^-$  was taken to be 1.25 [All84, PDG92].

There are reasonably accurate measurements of the flux incident on a unit horizontal area at sea level for hard muons, where a hard muon is defined to be one that penetrates 15 cm of lead, corresponding to a momentum slightly over 0.3 GeV/ $c$ . The measured value for the hard component of the flux,  $0.013 \text{ cm}^{-2} \text{ sec}^{-1}$  [PDG92], was used to calculate the incident muon rate.

The calculated trigger rates due to cosmic-ray muons [Kra94a] are given in Chapter 10. The rates for the theoretical momentum spectrum were found to agree to within 5% of those from a measured momentum spectrum [All84].

### Muons from Collimators

There is a potential background from muons produced by lost beam particles on either side of the detector near the IP. For a single beam lifetime of 1.5 hours, the HEB(LEB) beam

loss rate is  $3.5 \times 10^3/\mu\text{s}$  ( $7.7 \times 10^3/\mu\text{s}$ ). Assuming this loss is due to beam-gas scattering and is uniform around the ring, the fraction lost by each beam within the 60 m straight section on either side of the IP is  $60/2200 = 0.027$ . For muon momenta above  $1 \text{ GeV}/c$ , the number of muons/electron [Tsa74] from full energy lost particles on copper in the HEB(LEB) is  $5.0 \times 10^{-6}$  ( $2.9 \times 10^{-7}$ ), giving muon rates of  $4.8 \times 10^{-4} \mu\text{s}$  and  $0.6 \times 10^{-4} \mu\text{s}$  from the HEB and LEB, respectively. Making the conservative assumption that all the produced muons reach the detector, the combined rate is about 600 Hz.

### Tunnel Shine

The detectors at PEP-I had shielding walls, ranging from one-fourth of an inch Pb to eight inches of steel rods, separating the IR hall from the tunnels on either side, to protect against various radiations present in the tunnels. No doubt *BABAR* will require similar protection. The forward end, facing the proton alcove, is fairly well-protected by the steel plates of the magnet endcap. The DIRC water tank on the other end may be more sensitive to radiation.

## 12.4 Background Rates and Detector Responses

---

This section describes the effects of machine-induced backgrounds on the various components of the *BABAR* detector, specifically the vertex detector, drift chamber, particle identification system, calorimeter, and muon detector. As shown below, all detector components can function with virtually no degradation in performance at the nominal background levels for ten years or more. It is desirable to maintain reasonably good performance and lifetime at background levels above this nominal level for a number of reasons. First, there are inherent uncertainties in the above predictions of background levels, both in the assumptions that go into the generation of the source and in simulating the detector response. Second, the accelerator running conditions may vary significantly, especially in the early stages when the machine is being tuned to the optimal operating point.

For most of the significant sources described above, there are not large uncertainties in the generated rates. The dominant contributor to occupancy and radiation damage effects is from the electromagnetic showers produced by lost beam particles hitting near the IP (Section 12.3.2). This source also produces significant contributions to the background trigger rate through the photoproduction of hadrons. As discussed in the section above on lost particles, a major uncertainty in this process is lack of knowledge of the final vacuum pressure distribution in the interaction region, though preliminary designs suggest that the uniform 1 nTorr is probably a safe assumption. The photoproduction model and its resulting impact on trigger rates remain uncertain, and work is in progress to check them against data and other models.

A major concern relating to machine operation is the radiation dose received during injection. High voltage will be lowered during injection for some detector components to protect them from harm. Components such as silicon detectors, readout electronics, and calorimeter crystals will still be sensitive to this added dose from injection. This is currently being studied in simulations. Experience at other  $e^+e^-$  machines suggests that the integrated radiation dose received from injection can be as large as that received from normal running conditions (*i.e.*, the total dose should be assumed to be at least twice the nominal level).

One purpose of this section is to describe the ability of each component to handle backgrounds above the nominal value. As the background increases, the effects vary from reduced efficiency or high occupancy in a small region of the detector, with little impact on overall detector performance, to actual device failures that occur at very high background levels and result in downtime to replace components.

### 12.4.1 Vertex Detector

#### Vertex Detector Overview

The silicon vertex detector is described in detail in Chapter 4. The essential details of the geometry and other features relevant to the background analysis will be summarized here.

The vertex detector has five layers of sensors surrounding the beam pipe. The three inner layers are a traditional barrel-style geometry with six detector modules per layer, each covering  $60^\circ$  in azimuth. The minimum radii of the layers are located at 3.2, 4.0, and 5.2 cm, respectively. The outer two layers are arch shaped, with tilted detectors in the forward and backward directions to minimize the total area of silicon and to avoid small track crossing angles. There are 16 detector modules in azimuth in Layer 4, and 18 in Layer 5. The barrel sections of Layers 4 and 5 are located at minimum radii of 12 and 14 cm, respectively.

Double-sided silicon strip detectors are employed with  $\phi$  strips running parallel to the beam line and  $z$  strips oriented at  $90^\circ$  to the beam line. The vertex detector covers the solid angle over the region  $17.2^\circ < \theta < 150^\circ$ . The readout electronics are located outside of the active tracking volume in a 1 cm-thick space surrounding the B1 magnets.

#### Vertex Detector Backgrounds

The primary source of backgrounds in the vertex detector is lost beam particle interactions (Section 12.3.2). Synchrotron radiation is produced copiously near the IP but is carefully shielded so that very little reaches the vertex detector.

Beam-gas scattering produces off-energy electrons as well as photons. These lost particles can impact the vertex detector, where they have several effects. Radiation damage from the ionization they produce can degrade the performance of both the silicon sensors and the front-end readout electronics. The additional occupancy can result in inefficiency for hits from physics-related tracks and pose bandwidth problems for the DAQ system. Offline, background hits can confuse the pattern recognition algorithms. These effects have all been studied, and the results are summarized here.

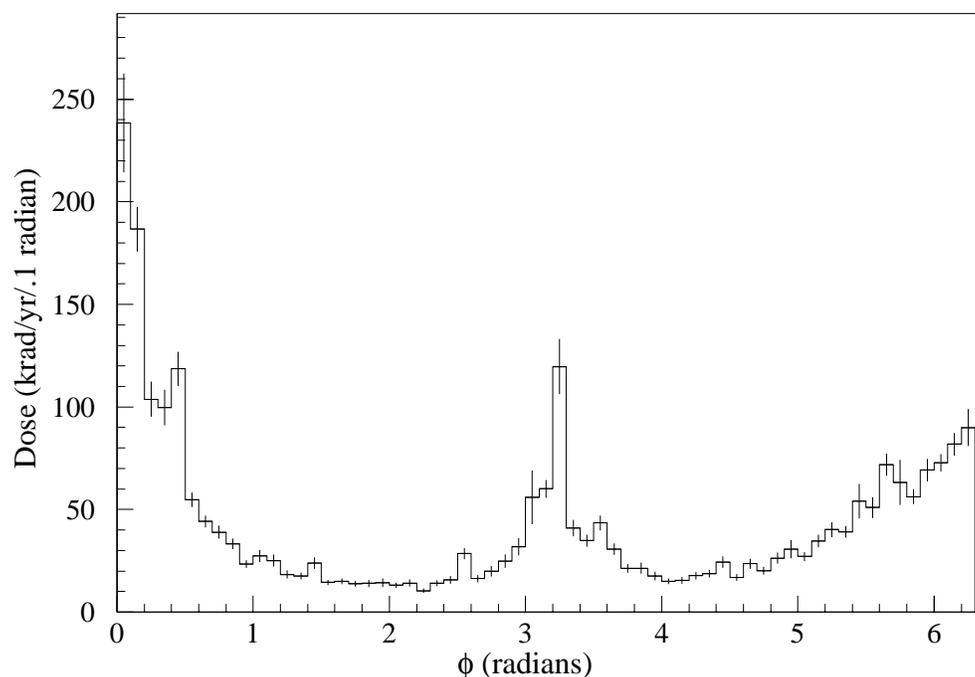
## Radiation Damage

The BBSIM detector simulation package is used to determine the flux of ionizing radiation through the silicon from lost beam particle showers. In units of krad per operational year, *i.e.*,  $10^7$  s, Layer 1 sees an average of about 33 krad/yr, Layer 2 sees 19 krad/yr and Layer 3 sees about 6 krad/yr [Cou95]. The dose is not uniformly distributed, as shown in Figure 12-13. There are sharp peaks in the bend planes of the machine at  $\phi = 0, \pi$ . This results in a highly nonuniform dose per stave in the inner layers. The worst case is in Layer 1, Stave 1, where the average yearly dose is 82 krad/yr, and the peak dose is 240 krad/yr over a region covering approximately  $6^\circ$  in azimuth.

This radiation will result in increased leakage current. Experimental measurements of the effects of ionizing radiation on silicon detectors have shown an increase in the leakage current of approximately  $0.7 \text{ nA/cm}^2/\text{krad}$  for detectors similar in design to those planned for *BABAR* [Joh09]. The vertex detector should withstand 10 times the yearly dose in order to have a reasonable safety factor over the estimated two-year period between access for repairs. There may be additional radiation damage caused during periods of non-optimal machine conditions, such as during injection or machine tuning; such effects have been known to increase the integrated radiation dose by a factor of 2 over the nominal calculated level. By designing the vertex detector to withstand 10 times the yearly dose, we account for this additional dose, integrated over a period of two years, with some additional margin for safety.

For Layer 1, Stave 1, 10 times the average yearly dose results in an estimated increase in leakage current of approximately 26 nA per strip in both  $\phi$  and  $z$ . The corresponding noise is  $166 e^-$  in equivalent noise charge (ENC), which must be added in quadrature to the ENC of the input amplifier; the latter ranges from about  $400 e^-$  for the  $z$  strips to a little over  $600 e^-$  for the  $\phi$  strips. The resulting increase in noise is less than 8% in both cases. In the worst azimuthal region, the increase is about 72 nA per strip resulting in an ENC of  $260 e^-$ , but this affects only the  $\phi$  strips. Adding this noise in quadrature with the noise of the input device results in an increase of about 9% in the total ENC.

The front-end readout ICs are located just outside of the active tracking volume and are mounted on hybrids supported from cones which surround the B1 magnets located on either



**Figure 12-13.** Dose as a function of  $\phi$  in vertex detector Layer 1, measured in  $\text{krad}/10^7 \text{ sec}/0.1 \text{ radian}$ .

side of the IP. Flexible circuits connect the silicon detector strips to the readout ICs, so they are not constrained to be at the exact radius of the detectors. In the inner layers, the ICs are located at a slightly larger radius than the detectors they service. The readout ICs for Layer 1 are located at approximately the same radius as the detectors for Layer 2, so we can use the Layer 2 radiation levels for the Layer 1 readout ICs. The worst case in Layer 2 is Stave 1, which receives an average dose of 47 krad/yr with a peak value of 110 krad/yr.

The readout ICs must tolerate 10 times the maximum expected dose; this requirement implies the use of radiation-hard circuits. For the planned CMOS technology, a radiation-hard process is available from several vendors. For the baseline design, the Honeywell process has been chosen because it is offered with a  $0.8 \mu\text{m}$  feature size with three metal layers; these options are desirable to optimize the IC layout and minimize the IC dimensions. The noise characteristics of the Honeywell process have been studied before and after irradiation. A slight increase in noise relative to the non-radiation-hard HP CMOS process is observed before irradiation, on the order of 20% at 3 MHz, which is the frequency of interest for the risetimes used in the inner layers. After irradiation to 500 krad, the noise increases by 15% at 3 MHz; after 1.5 Mrad the increase is 30%.

Stave	Layer 1		Layer 2	
	$z$	$\phi$	$z$	$\phi$
1	3.1%	1.4%	2.1%	1.3%
2	1.1%	0.6%	0.8%	0.5%
3	0.4%	0.3%	0.2%	0.3%
4	1.3%	0.6%	0.8%	0.5%
5	0.6%	0.4%	0.4%	0.3%
6	1.1%	0.6%	0.7%	0.5%

**Table 12-3.** Occupancy in Layers 1 and 2 at nominal background.

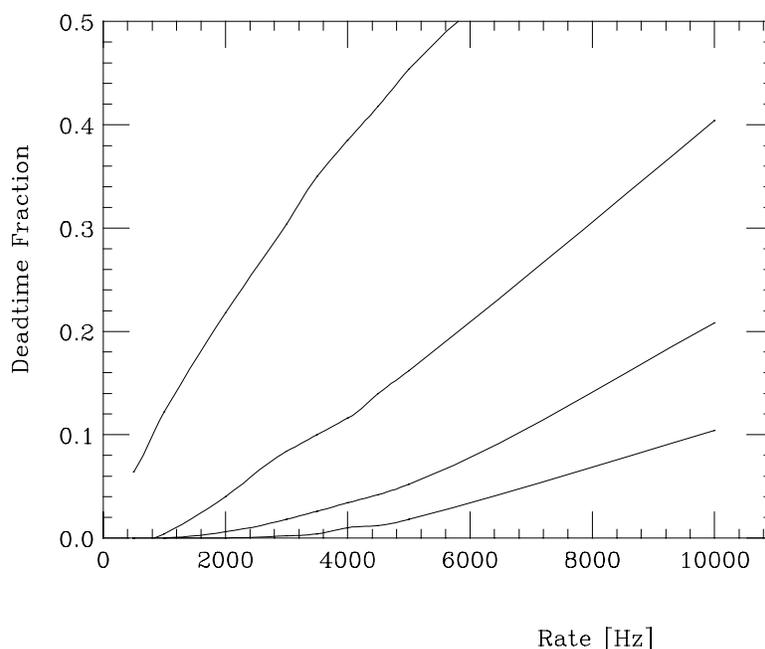
### Background Occupancies

The percent occupancy in a one  $\mu\text{s}$  window versus stave number for Layers 1 and 2 is shown in Table 12-3 [Cou95]. The distribution is peaked in a manner similar to the radiation damage, with significantly higher occupancy in Staves 1 and 4.

The background occupancy is fairly high, and this has some important consequences for the readout architecture. One notable consequence is that a data-push readout is not feasible due to the large data transmission bandwidth that would be required; this in turn implies a fast Level 1 trigger to enable the vertex detector readout [Lev94a]. The Level 1 trigger is described in Chapter 10; it will have a latency of no more than 10  $\mu\text{s}$  and an uncertainty of no more than 1  $\mu\text{s}$ .

The vertex detector readout system must be able to cope with up to 10 times the nominal background occupancy. The safety factor of 10 is taken into account for occasional increases in occupancy that may occur during non-optimal periods of running. This requirement is met by providing sufficient bandwidth in the data transmission system, together with enough buffers in the readout section of the front-end chip to even out the fluctuations in rate.

The worst case is again Layer 1, Stave 1, where the occupancy at 10 times nominal background is 30% for the  $z$  strips. Each stave will have four data links, two on each end, to independently transmit data from the  $\phi$  and  $z$  strips, and each data link is expected to have a transmission rate of at least 60 Mbits/s. At the maximum Level 1 trigger rate of 2 kHz, assuming 16 bits/hit plus a per-chip overhead of 16 bits, the total bandwidth out of the Layer 1, Stave 1  $z$  strips is about 6 Mbits/s per half detector module. The deadtime fraction for 10 times nominal background, under these assumed conditions, is plotted versus trigger rate in Figure 12-14 for one, two, three, and four back-end buffers [Lev94b]. For three buffers at the expected maximum Level 1 trigger rate of 2 kHz, the deadtime is about 1% and exhibits linear behavior.



**Figure 12-14.** *Deadtime fraction for 10 times nominal background vs. trigger rate for one, two, three, and four back-end buffers.*

### Pattern Recognition

The offline pattern recognition capabilities of the vertex detector have been studied using a detailed GEANT simulation, together with a Kalman filter track finding algorithm. The study assumed that tracks which generated at least ten hits in the drift chamber would be found by the drift chamber and extrapolated into the vertex detector. Tracks with a  $p_t$  below 90 MeV/ $c$  which fail this criteria are found using the vertex detector for stand-alone pattern recognition. The overall pattern recognition efficiency is found to be very high, approaching 100%, for generic  $B$  decays with nominal and with 10-times-nominal backgrounds superimposed (see Chapter 4). The fake rate, defined as the number of found tracks per event which have fewer than half of their hits from a real track, is seen to go up as background levels increase. Most fake tracks have a  $p_t$  below 90 MeV/ $c$  and are due to random combinations of background hits in the vertex detector alone. However, the five layers of the vertex detector provide a powerful pattern recognition capability, and the fake rate can be kept to an acceptable level (fewer than 0.05 fake tracks/event) even at 10 times the nominal background.

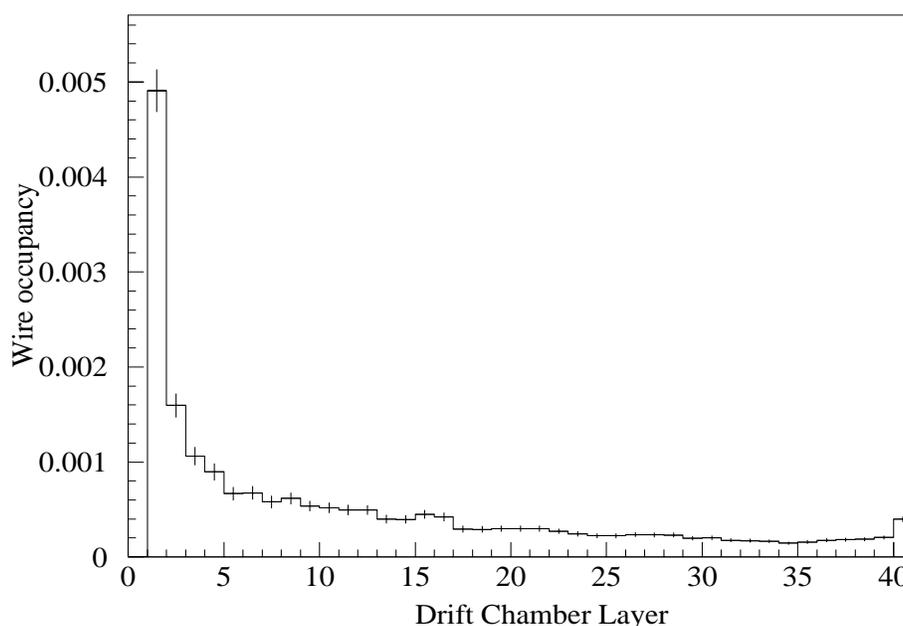
### 12.4.2 Drift Chamber

Wire aging and chamber occupancy due to beam-related backgrounds are two issues of concern for the operation of wire chambers. Aging refers to changes in gain or increases in noise due to material deposited on the wires from the gas-avalanche process. Chamber occupancy is the fraction of cells that contain random hits during the event resolving time, *i.e.*, the full drift time. If the occupancy exceeds a few percent, pattern recognition becomes difficult.

The design of the *BABAR* drift chamber has been guided by the desire to minimize the amount of material in the chamber, thus permitting the best possible measurement of the relatively low momentum tracks necessary for the *CP* asymmetry studies. The drift chamber is 280 cm in length and occupies the radial space between 22.5 cm and 80 cm, providing a polar angle acceptance down to 300 mr. It is bounded by the support tube at its inner radius and the particle identification device at its outer radius. The chamber contains approximately 40 layers of wires arranged in a small-cell geometry, with a total of about 7000 sense wires and 45,000 field wires. The sense wires are 20  $\mu\text{m}$ -diameter gold-plated tungsten and the field wires are 55  $\mu\text{m}$ -diameter gold-plated aluminum. The chamber is filled with a helium-based gas that, along with the choice of aluminum wire, represents a minimal number of radiation lengths. The gas alone has a radiation length of  $\sim 800$  m. The 40 layers of individual sense and field wires correspond to 0.02% and 0.08% of a radiation length, respectively, for normal incidence.

The chamber endplates and the inner and outer cylinder walls are also being designed using low-mass materials. The endplates are foreseen to be made of carbon fiber, corresponding to less than  $0.03X_0$ . The inner cylinder wall's function is to provide a gas seal and will be made of carbon fiber or beryllium, representing  $\lesssim 0.002X_0$ , while the outer wall, which supports the wire load, will be of aluminum or carbon fiber, giving  $\sim 0.03X_0$ .

Beam-related backgrounds which occur during beam injection are not a problem for the drift chamber, for which the voltage can and will be turned off. The expected backgrounds in the drift chamber have been calculated for a 1 nTorr beam vacuum. The drift chamber occupancies for a resolving time of 1  $\mu\text{s}$ , which corresponds to the maximal drift time, are shown for each layer in Figure 12-15. The distribution is peaked at the innermost layer, where occupancies of 0.5% are expected. For the outer layers, the occupancies are  $\sim 0.03\%$ . These results, based on a chamber with a radiation length of 800 m, do not include conversions in or scattering by the wires. A simple method of estimating an upper limit on the effect of the wires on the occupancies would be to smear the wire material over the chamber volume, resulting in a radiation length of 400 m, which would roughly double the occupancies. More realistically, the effect of the wires would increase the occupancies in the outer layers, while leaving the results for the innermost layers unchanged.



**Figure 12-15.** *Expected background occupancies for each drift chamber layer due to converted photons from showers produced by lost particle interactions.*

The occupancies have been calculated without including the effects of signal crosstalk, expected to be  $\sim 1\text{--}3\%$ . Since most background events would produce low-momentum tracks which spiral in  $z$  within a few cells (depending upon the wire orientations), producing very large signal pulses, the number of neighboring cells which also contain pulses due to electronic crosstalk could be sizable.

In general, chamber occupancies of a few percent are tolerable but render the problem of pattern recognition and segment-finding more difficult. Occupancies which approach or exceed  $\sim 10\%$  make track finding almost impossible.

For a uniform distribution along the wires, the energy deposition on a wire corresponds to an integrated charge of  $0.001\text{ C/cm}/10^7\text{ sec}$  for the innermost drift chamber layer. The charge deposition in the drift chamber is not uniform, showing factors of 2 variation in  $\phi$  and  $z$ . Taking these factors into account, the expected integrated charge for some regions of the inner layers could be as high as  $0.005\text{ C/cm}/10^7\text{ s}$  for the 1 nTorr beam vacuum.

Significant amounts of charge accumulated on the sense wires can lead over time to deposits on the wires, resulting in electrical discharges and current draw. A conservative limit on the integrated charge on a wire which can be safely tolerated is  $0.1\text{ C/cm}$ . This limit can be extended depending upon the ionizing gas mixture. For example, small amounts of

	Bars	Tank
# of Charged Particles	16	0.7
# of Hits	32	8.0

**Table 12-4.** Mean number of charged particles crossing the DIRC per  $\mu\text{s}$  due to beam lost secondaries.

$\text{H}_2\text{O}$  [Kad91] and  $\text{CF}_4$  [Ope91] in combination with ionizing gases have been shown to extend this limit to  $\gtrsim 1.0 \text{ C/cm}$ .

### 12.4.3 Particle ID

The particle identification system described in Chapter 6 consists of a ring-imaging Cherenkov system (DIRC) covering the barrel region and an aerogel threshold counter in the forward direction. The DIRC consists of quartz bars read out with phototubes mounted on a water tank in the backward ( $-z$ ) direction.

The background level in the aerogel detector has not been studied in detail. The principal source of background is likely to be the flux of low-energy photons from lost particle showers. Backgrounds from conversions of these photons in the aerogel itself are expected to be small due to its very low density. The contribution from the support structure and photodetectors are currently under study.

Several background sources might affect the DIRC's performance: event related backgrounds, synchrotron radiation photons, beam-gas interactions, cosmic rays, phototube dark noise, etc. The program BBSIM has been used to simulate the detector response. Each random background source has been integrated over an assumed 50 ns readout time.

**Lost Particle Backgrounds.** The background induced in the DIRC from lost beam particles has been estimated as described in Section 12.3.2 with the assumption of a uniform beam line pressure of 1 nTorr. The mean number of charged particles crossing the quartz bars and the water tank every  $\mu\text{s}$  and the number of hits produced at the nominal background rate are listed in Table 12-4. These numbers are to be compared to an average of 345 hits for a typical physics event in a 50 ns window. The deduced mean rates expected per PMT and per readout time at 10 times nominal background are around 20 kHz for a uniform beam line pressure of 1 nTorr.

**Radiative Bhabha.** The radiative Bhabha process,  $e^+e^- \rightarrow e^+e^-\gamma$  is a potentially important source of background. A BBSIM simulation of this process predicts that about 25 charged particles per  $\mu\text{s}$  from secondary interactions crossing the DIRC, yielding an average of 75 hits, 40% due to Cherenkov emission in the bar, 60% in the tank. The background from radiative Bhabhas is on the same order of magnitude as the nominal background from lost particle interactions. The showers from radiative Bhabhas tend to originate from around Q1 and Q2, suggesting the possibility of additional shielding to reduce this background.

Other external backgrounds such as synchrotron radiation and cosmic rays have very small or negligible impact on the DIRC response (see Chapter 6).

#### 12.4.4 Calorimeter

The main background in the CsI calorimeter is low-energy photons from lost beam particle showers. Energy and angular distributions for these photons are shown in Figure 12-9 for the nominal background level. About 330 photons per  $\mu\text{s}$ , with a total energy of 0.39 GeV, are incident on the calorimeter. The median photon energy is  $\sim 500$  keV, and about one photon per  $\mu\text{s}$  is above 20 MeV. Major background-related issues for the calorimeter are radiation damage, energy resolution degradation, creation of extra showers in events, and trigger rates.

The radiation dose was estimated by assuming that the photons of Figure 12-9 are absorbed in the first 2 cm ( $9 \text{ g/cm}^2$ ) of the calorimeter. At nominal background, the radiation dose for  $10^7$  s/yr running is about 45 rad/yr at  $\theta = 90^\circ$  and 90 rad/yr at  $\theta = 300$  mr. Larger doses are expected from injection and machine studies, based on the experience of existing storage ring operation. Scaling typical CESR operations to the PEP-II design beam currents gives a radiation dose of about 1.5 krad/yr at a radius of 45 cm, and 0.5 krad/yr at 100 cm [Blu86]. Both are well within the tolerable dose rate of 10 krad/yr.

The effect of backgrounds on photon energy resolution was studied using lost beam particles generated by TURTLE as input to BBSIM, where they were showered in the same events as the photons of interest. Backgrounds corresponding to nominal and 10 times nominal levels in  $1 \mu\text{s}$  windows were generated together with 100 MeV photons in the barrel and forward endcap. Comparison of the resolutions with the zero background case, for which  $\sigma = \text{FWHM}/2.36 = 1.8$  MeV, shows contributions, in quadrature, of 0.55 MeV from nominal background and 1.05 MeV from 10 times nominal background.

A small fraction of the photons from beam backgrounds has sufficient energy to be detected as extra photons in physics events. Efficient  $B$ -meson reconstruction requires measurement of photon energies as low as 20 MeV. The rate of photons with energies above 20 MeV from lost beam particles is 0.7 per  $\mu\text{s}$  in the barrel and 0.4 per  $\mu\text{s}$  in the forward endcap at nominal background. Timing information from the shaped calorimeter signals may be

used to discriminate between photons which belong to events and out-of-time background photons.

Some beam background particles reaching the calorimeter have sufficient energy to contribute to the trigger rate. Work is in progress to estimate this rate reliably and to improve the design of detector shielding if necessary. Section 10.8 contains further discussion of this issue.

### 12.4.5 Muons

The Instrumented Flux Return (IFR) is well shielded from potential background sources by the calorimeter, magnet, and flux return steel. Only the endcap chambers are close to the beam line, and these are protected by flux return and magnetic shielding for the beam line quadrupoles. Backgrounds from the previously discussed sources and from the irreducible cosmic ray flux are expected to be small. The outermost layer of the IFR may be sensitive to soft electromagnetic backgrounds in the detector hall originating from the PEP-II arcs. Shielding walls in the tunnel mouths will be necessary to isolate the detector from the arc radiation sources.

## 12.5 Physics Impact

---

The impact of backgrounds on the various detector components ultimately must be evaluated in terms of their effect on the overall physics goals of *BABAR*. Two requirements can be compromised by high background levels: trigger selection and data acquisition for events of interest; and offline reconstruction of the data.

As described above, a principal source of background trigger rates is the photoproduction of hadrons in the showers initiated by lost particles, contributing to both a charged-track trigger in the central drift chamber and to a calorimeter energy trigger. The rate is acceptable at nominal background levels but begins to saturate the Level 1 trigger rate at around 10 times nominal. Chapter 10 describes in detail the design of the trigger and steps taken to deal with the possibility of a high trigger rate from this source. The conclusion of this discussion is that some flexibility is required in the trigger to allow the energy and momentum thresholds to be raised if the rate from this source becomes too high. The loss of trigger rate for some events of interest from the higher thresholds is also discussed in this chapter.

The effect of backgrounds on event reconstruction within a particular detector component has been studied in several cases, including, for example, track reconstruction in the silicon vertex detector and the degradation of calorimeter resolution at above nominal background levels. These are examples of two major effects of backgrounds on offline event reconstruction:

the inability to accurately reconstruct events due to inefficiency and/or the confusion caused by the high occupancy, and the degradation in resolution from the overlap of hits from background and the physics event. The vertex detector shows a clear ability to maintain good reconstruction efficiency at 10 or more times nominal levels, due largely to its fine sampling and the redundancy of a five-layer detector. Similar studies with the drift chamber are in progress. Again the large number of layers (40) and low occupancy are expected to offer robust track reconstruction at above-nominal background levels. At 10 times nominal background the flux of photons noticeably degrades the calorimeter resolution only for photon energies below 50–70 MeV. This effect will be seen as a small worsening of the reconstructed mass resolution for generic  $B$  decays with a  $\pi^0$  in the final state but will not effect modes with higher momentum  $\pi^0$ s (*e.g.*,  $B \rightarrow \pi^0\pi^0$ ).

A real assessment of the effect of backgrounds on overall event reconstruction awaits the development of the reconstruction algorithms, a major effort that is in progress but not complete.

## 12.6 Summary

---

Extensive work has been done by the  $BABAR$  collaboration to develop the tools to predict the background rates in PEP-II and their effects on the various detector components. Work continues on refining these estimates and checking the results with data where available. There is also interaction with the PEP-II designers regarding aspects that affect background rates, *e.g.*, the degree of vacuum pumping in the interaction region. The principal aim in these studies is to ensure that the  $BABAR$  detector physics goals are not compromised by machine-induced backgrounds. The detector design performance can be degraded due to high occupancy causing inefficiencies, deadtime, or pile-up effects, and radiation damage, causing impaired performance or frequent downtime to replace components. Some safety margin (generally taken to be a factor of 10) is required to account for uncertainties in the calculations, the assumed PEP-II running conditions, and in the case of radiation damage, the dose accumulated during injection and machine studies. As explained in Section 12.1.4, the design of the injection system for PEP-II suggests that injection will be cleaner than for other machines. Nevertheless, the working assumption in the  $BABAR$  detector design is that the dose from injection will equal that of normal running (in the case of the calorimeter, a direct scaling of CLEO results is used).

Table 12-5 summarizes the contribution to occupancy from the dominant lost particle background for the more sensitive elements of the  $BABAR$  detector. Careful masking design and shielding around the beam line components has reduced the contribution from synchrotron radiation and radiative Bhabhas to a low level. Also in Table 12-5 is an estimate of the limit on how high a value could be tolerated without compromising the performance goals of

	Silicon Layer 1		Drift Chamber		Particle ID (DIRC)	Calorimeter Photons
	Average	Stave 1	Average	Layer 1		
Occupancy	1.3%	3.0%	0.05%	0.5%	0.015%	1.1/ $\mu$ s
Limit	20%	20%	10%	10%	2%	10/ $\mu$ s
Safety	15	6	200	20	130	9

**Table 12-5.** Occupancy in various detector components of BABAR. Calorimeter entries are for photon energies above 20 MeV.

	Silicon Layer 1		Drift Chamber		Calorimeter
	Average	Stave 1	Average	Layer 1	
Dose/ $10^7$ sec	33 krad	82 krad	0.0004 C/cm	0.0045 C/cm	1.5 krad
Limit	1500 krad	1500 krad	0.1 C/cm	0.1 C/cm	100 krad
Safety	45	18	250	22	67

**Table 12-6.** Radiation dose in various detector components of BABAR.

BABAR; the bottom row gives the resulting safety margin. The silicon detector has the lowest safety factors. The 20% limit on occupancy in this detector is set by the desire to maintain 90% or better hit detection efficiency. The worst stave of Layer 1 of the vertex detector has a safety margin below 10, but a larger loss in efficiency in this limited region of a single layer is not considered a serious compromise of its overall performance.

Table 12-6 gives the radiation doses in appropriate units for the detector. The first row gives the yearly dose, and the second row gives an estimate of the limit at which the performance begins to significantly degrade. The safety margin then determines the lifetime of the detector at nominal background levels. Included in this safety margin, though, must be some accounting for uncertainties in the calculated background levels and, particularly for the silicon and the calorimeter, the additional doses likely to be delivered to these systems during injection. The drift chamber will be ramped down in voltage during injection. The BABAR detector is expected to be operating for 10–15 years, but it is reasonable to assume that limited parts of the detector (such as sections of Layer 1 of the vertex detector or the forward calorimeter) could be replaced at intervals of three to five years. As noted in Section 12.4.1, small  $\phi$  regions of Layer 1 of the silicon vertex detector will see higher doses (240 krad). These regions will have degraded signal-to-noise due to increased leakage current, but the small extent of the damage will not seriously affect the overall performance. The conservative approach of scaling the CLEO results on dose from injection is used for the calorimeter dose.

## References

---

- [All84] O.C. Allkofer and P.K.F. Grieder, “Cosmic Rays on Earth,” Physik Daten 25–1, Karlsruhe, Germany, 1984.
- [Blu86] E. Blucher *et al.*, *Nucl. Instr. Methods* **A249**, 201 (1986).
- [Bai81] V.N. Baier, V.S. Fadin, V.A. Khoze, and E.A. Kuraev, “Inelastic Processes in Quantum Electrodynamics at High Energies,” *Phys. Repts.* **78**, 293 (1981).
- [Bri94] M. Bringle, “Cosmic Ray Muon Monte Carlo Generator using the Hemisphere Method,” *BABAR Note # 163* (1994).
- [Car91] D.C. Carey, K.L. Brown, and F.C. Iselin, “Decay TURTLE (Trace Unlimited Rays Through Lumped Elements): A Computer Program for Simulating Charged Particle Beam Transport Systems, Including Decay Calculations,” SLAC–246 (1980).
- [Che91] P. Chen, T. Tauchi, and D.V. Schroeder, “Pair Creation at Large Inherent Angles,” in *Linear Collider IR and Final Focus Introduction*, edited by J. Irwin and D. Burke, SLAC–PUB–5652 (1991); also in Proceedings of the 1990 DPF Summer Study (Snowmass, Colorado), pp. 751–756 (1990).
- [Cou95] D. Coupal, “Lost Particle Backgrounds in *BABAR*,” *BABAR Note # 210* (1995).
- [Dar84] A. Dar, “Cosmic Ray Muons at Ground Level and Deep Underground,” TECHNION–PHYS–84–41, Israel Institute of Technology, Haifa (1984).
- [Deg94] P.V. Degtyarenko *et al.*, *Phys. Rev.* **C50**, 541 (1994).
- [Fes85] H.C. Fesefeldt, “The Simulation of Hadronic Showers, Physics, and Applications,” Technical Report PITHA 85–02 (1985).
- [GEA94] GEANT Detector Description Tool, version 3.21, CERN Program Library W5103, CERN (1994).
- [Hal81] K. Halbach, “Design of Permanent Multipole Magnets with Oriented Rare Earth Cobalt Material,” *Nucl. Instr. Methods* **169**, 1 (1980).
- [Her87a] S.W. Herb, “Construction of Large Permanent Magnet Quadrupoles,” CLNS–87/61 (1987).
- [Her87b] S.W. Herb and J. Kirchgessner, “Operation of CESR with Permanent Magnet Interaction Region Quadrupoles,” CLNS–87/64 (1987).

- [Hea91] C. Hearty, “OBJEGS Users’ Manual,” *BABAR Note # 73*, (1991).
- [Joh09] R. Johnson, “Radiation Damage Measurements on ALEPH Double-Sided Silicon Strip Test Structures,” ALEPH Note 94-009 (1994).
- [Kad91] J. Kadyk, *Nucl. Instr. Methods* **A436**, (1991).  
J. Va’vra, SLAC-PUB-5207 (1990).  
M. Danilov *et al.*, *Nucl. Instr. Methods* **A274**, (1989).
- [Kle94] R. Kleis and H. Burkhardt, NIKHEF-H/94-01 (1994).
- [Kra94a] F. Kral, “Trigger Rates due to Cosmic Ray Muons,” *BABAR Note # 164* (1994).
- [Kra94b] F. Kral, “Simulation of Trigger Rates due to Beam-Gas Backgrounds,” *BABAR Note # 165* (1994).
- [Lev94a] M. Levi, “Impact of Backgrounds on the Silicon Vertex Detector Architecture and Detector Trigger,” *BABAR Note # 136* (1994).
- [Lev94b] M. Levi, “A Simulation of the Vertex Detector Readout Buffers Using the Machine Induced Occupancy,” *BABAR Note # 193* (1994).
- [Lig88] J.W. Lightbody, Jr. and J.S. O’Connell, “Modeling Single Arm Electron Scattering and Nucleon Production from Nuclei by GeV Electrons,” *Computers in Physics*, May–June 1988, 57–64.
- [Nel85] W. Nelson, H. Hirayama, and D. Rogers, “The EGS4 Code System,” SLAC-PUB-265 (1985).
- [Ope91] R. Openshaw *et al.*, *Nucl. Instr. Methods* **A**, 298 (1991).
- [PDG92] Review of Particle Properties, Particle Data Group, *Phys. Rev.* **D45** (1992).
- [PEP93] “PEP-II, An Asymmetric *B* Factory: Conceptual Design Report,” LBL-PUB-5379, SLAC-PUB-418, CALT-68-1869, UCRL-ID-114055, UC-IIRPA-93-01 (1993).
- [Sny94] A. Snyder, “Electroproduction Triggers in *BABAR*,” *BABAR Note # 180* (1994).
- [Tsa74] Y.S. Tsai, *Rev. Mod. Phys.* **46**, 815 (1974).

## 13.1 Introduction

---

This chapter provides a description of identified and potential hazards, their relative significance, and the proposed design controls needed to eliminate or reduce the associated risk to acceptable levels.

This chapter does not assess the risk for those identified hazards in terms of severity or probability. This will be accomplished in the Safety Assessment Document (SAD), which will further identify, evaluate, track, and resolve safety issues. The *BABAR* Detector will also be reviewed by SLAC safety committees such as the Hazardous Experimental Equipment Committee (HEEC), the Radiation Safety Committee, the Pressure Safety Committee, and the Earthquake Safety Committee.

The purpose of performing this analysis and the committee reviews is to make sure that safety is designed into systems, subsystems, equipment, facilities, and their interfaces, consistent with the objective of the detector. The policy of management is to design for minimum risk.

## 13.2 Detector Safety Overview

---

*BABAR* will be housed in the existing PEP Research Hall at Interaction Region 2 (IR-2). The detector weighs approximately 800 metric tonne. The floor of IR-2 is constructed of two-foot-thick reinforced concrete. Estimated dimensions of *BABAR* are 6.7 m in height, 9.5 m in width, and 8.3 m in length from the outer edge of the forward door to the outer edge of the DIRC water tank. The design of the Research Hall incorporates a concrete shield wall, which will contain penetrations for cableways and cryogenics. The fast electronics will be located in a counting house just outside the wall. The existing control room area in IR-2 will be utilized. The cryogenic plant will be located on an existing pad outside the southern wall of IR-2, and cryogenic storage tanks will be located on another pad on the hill at the western end of the building.

*BABAR* is similar to detectors operating at SLAC and other physics laboratories, therefore the hazards contained in *BABAR* are similar to previous examples.

Hazards contained in *BABAR* which are currently under assessment are:

- Gases—flammable, asphyxiant, toxic—possibly utilized in the drift chamber and RPCs;
- Electrical hazards in the various high voltage systems (the drift chamber, the RPCs, and the DIRC photomultiplier tubes), the electronics, the power supplies, the high current and stored energy associated with the magnet, and support equipment in the counting house and control room area;
- Radiation due to operation of the machine and possibly from a calibration system for the calorimeter;
- Cryogenics for the superconducting coil and possibly the Q1 magnets;
- Magnetic forces during operation of the magnet;
- Pressure/vacuum in the beam pipe, the drift chamber, the cryostat, and the associated supply systems; and
- Toxic materials such as beryllium, thallium-iodide, and cesium-iodide.

Generic hazards currently under assessment are

- Earthquake,
- Radiation,
- Fire,
- Occupational (Oxygen Deficiency and Confined Space), and
- Environmental.

### 13.3 Beam Pipe and Support Barrel

---

The Q1 magnets, the B1 magnets, the IP beam pipe, and the vertex detector are assembled into a single rigid support barrel. The beam pipe at the interaction point is a double-wall structure made of beryllium. The inner tube is 50 mm in diameter with a separation of

2 mm from the outer tube. Helium gas flows between the tubes to remove the beam-induced heating.

The health hazard of the beryllium is minimal in the form used. The particulate created as a result of machining, grinding, corrosion, *etc.*, is toxic if ingested. A coating will be applied which should control corrosion, and all machining will take place at an outside vendor under appropriate safety conditions. If the beryllium needs to be worked at SLAC, special procedures to control the particulate will be developed.

The thin wall beryllium tubes are easily damaged, and special handling requirements to reduce this potential will be developed.

## 13.4 Vertex Detector

---

The vertex detector consists of double-sided silicon microstrip detectors and readout hybrid circuits assembled into mechanical modules. These modules are glued to lightweight support structures. At each end of the module, beryllium oxide or aluminum blocks provide the required mechanical precision and act as a heat sink for the electronics.

The health hazard for beryllium is minimal in the form used (Section 13.3). The operating voltage is less than 50 V, with currents less than 100 mA. Dry nitrogen or air will be used to control the atmosphere. If nitrogen is used, the hazard of asphyxiation may exist depending on the quantity. Water may be used for cooling. This detector system appears to have minimal overall risk for causing injury to personnel or damage to equipment.

## 13.5 Drift Chamber

---

The drift chamber consists of a carbon-fiber shell, a helium-based gas which may be flammable, and a large number of wires with high voltage. Significant effort is under way to find an acceptable gas mixture that is nonflammable. The drift chamber may operate at constant gas pressure resulting in a pressure differential (as much as 30–40 mbar) between the chamber and the atmosphere. On both ends of the chamber, rf-shields will contain an inert gas (probably N<sub>2</sub>) that is circulated and will provide humidity control. The gas will be monitored for helium content, which will allow detection of leaks from the chamber.

The drift chamber gas is an asphyxiant and may be flammable; it is probably not toxic. The storage, transport, mixing, use, and disposal must be adequately controlled. The following are the proposed controls:

- Development of a safety system to sense leaks and/or oxygen deficiency, and to automatically bring the system to a safe state. This will require the shut-off of the flammable gas at its source, the interruption of electrical energy to potential ignition sources, the purging of the system with inert gas, the operation of a ventilation system to dilute any leaking gas, and the sounding of appropriate alarms to warn personnel;
- Isolation of the bulk storage and the mixing operations in separate areas outside the experimental hall;
- Utilization of flow restrictors at the storage area;
- Transportation and distribution of the gas in metal pipes to prevent leaks;
- Comparison of chamber supply and return flow rates to aid in the detection of leaks; and
- Utilization of an appropriate fire suppression system.

Control of the high voltage will be accomplished by conventional means. Design controls will include short-circuit protection; isolation by insulating, enclosing, or covering; and grounding. Equipment will be designed to meet appropriate codes and standards. Procedural controls will include utilization of properly trained personnel and adherence to established safe work practices and procedures.

Pressure is a concern in terms of damage to the chamber. The design must have features that will prevent overpressurization which could result in subsequent damage to the chamber and possibly other systems.

Accessibility to the chamber for assembly, maintenance or repairs (broken wire removal), may pose a threat to personnel due to the confined space and to the presence of hazardous gas. Reviews of the various scenarios to assure safe access will be accomplished in the SAD.

## 13.6 Particle Identification

---

### 13.6.1 DIRC

The DIRC system will consist of quartz radiator bars, a support structure, a standoff box containing a reservoir filled with water, and conventional photomultiplier tubes.

The hazards associated with the DIRC include high voltage at the photomultiplier tubes, water leakage from the standoff box reservoir, and the possible degradation of the tubes due to long-term exposure to helium. Control of high voltage is discussed in Section 13.5.

A water leak could damage electronics, short circuit electrical equipment, and cause corrosion in other detector systems. The presence of water near the calorimeter may degrade the calorimeter crystals. The interface between the standoff reservoir and the window frame/flange must provide a positive water seal. Flexing of the door due to the magnetic and seismic loads will need to be considered. In addition, features to control leaking water should be fairly simple to incorporate due to the location of the reservoir outside the backward door.

### 13.6.2 Aerogel

The second part of the particle identification system is constructed of silica aerogel. Silica aerogel is a glass-like structure which consists of amorphous  $\text{SiO}_2$ .

When aerogel is unprotected, it is very fragile and pieces easily break off at edges and corners. Cracks may develop inside during handling due to the low tensile strength. Aerogel also is damaged by most gases and liquids, including water. For these reasons, the aerogel blocks will be wrapped for protection. Aluminum, Teflon, and paper are under consideration. Photomultiplier tubes will be utilized and the issues addressed in the previous section pertaining to high voltage and He gas are applicable.

In the form used in the detector, aerogel appears to present minimal risk. The very fine dust of  $\text{SiO}_2$  generated as a result of machining requires the use of respiratory protection. No such machining of aerogel is planned at SLAC.

The presence of helium near the forward aerogel PID photomultiplier tubes could degrade the performance of the tubes. The drift chamber, the cryostat, and various cooling systems will use helium. It may be appropriate to direct the air/ $\text{N}_2$  flow of the ventilation system so that any possible helium present will be drawn away from the sensitive components.

Aerogel is an excellent insulator and the heat created by the tubes will need to be removed. Nitrogen gas is being considered for cooling.

## 13.7 Electromagnetic Calorimeter

---

The calorimeter consists of cesium-iodide crystals wrapped in PTFE Teflon or Tyvek, a support structure, photodiodes, and associated power supplies and regulators. The photodiodes will operate at approximately 75 V. Cooling will be required and freon is currently under consideration.

The principal hazard is the thallium iodide contained in the cesium iodide. The concentration of thallium iodide is approximately one part per thousand. The federal OSHA permissible

exposure level (PEL) is  $0.1 \text{ mg/m}^3$  for airborne thallium particles [CFR93]. The thallium iodide is contained in the crystal structure of the cesium iodide. Machining operations and handling of the uncoated crystal will require special precautions.

The possible presence of a mutagen in the fluorescent flux concentrator will also be evaluated. This is a laser dye that will be dispersed in solid plastic.

The effect of moisture on the crystals is detrimental, therefore the calorimeter will be contained in a closed, shielded environment, where the humidity will be strictly controlled to a level of approximately 3%.

One of the calibration systems under consideration contains a radiation hazard. This system would utilize polonium beryllium (Section 7.5.3) as a source of fast neutrons to activate freon. The freon would then be piped across the face of the crystals. The neutron source and the pipe carrying the pressurized, activated freon liquid through the detector would be located inside the IR concrete curtain wall. An alternative calibration scheme entails placing low-level sources on each crystal. Both schemes will be evaluated for radiation hazard in the SAD.

## 13.8 Muon and Neutral Hadron Detector

---

The system under consideration utilizes Resistive Plate Chambers (RPCs). These chambers consist of large-area parallel-plate electrodes held apart at a distance of 2 mm with spacer buttons and filled with a gas mixture. Approximately  $3000 \text{ m}^2$  will be covered by these RPCs.

The hazards associated with the chambers are related to electrical safety, flammability of material used in chamber construction, and the possible use of a flammable gas.

Electrical hazards are associated with the use of high voltage in the operation of the chambers. High voltage will be controlled through the use of appropriate cables and connectors, short circuit protection, isolation by insulating or enclosing, and grounding. High voltage supplies will not deliver currents in excess of 1 mA and will be interlocked with hazardous gas and fire alarms.

The RPCs are constructed of materials (bakelite) which are combustible only when exposed to an intense flame or arc. As the chambers are confined to narrow gaps between steel plates, there is a very limited supply of oxygen available and any combustion initiated at an exposed edge will be self-extinguishing. Smoke and heat-sensitive alarms will be located at strategic locations around the detector.

Proposed controls for the use of a flammable gas mixture are listed in Section 13.5. Due to the large number of RPCs and the large surface area covered, attention must be paid to the issue of leaks. Appropriate pressure testing before and after installation will be necessary.

Because the use of a nonflammable gas mixture is highly desirable, active R&D in this direction is currently being performed.

## 13.9 Magnet Coil and Flux Return

---

The magnet system consists of the superconducting solenoid, the flux return, and supporting systems including cryogenics. The superconducting solenoid consists of the coil in a cryostat; the vacuum system to provide thermal insulation; the power supply; the quench detection and protection system; and the control system. The flux return is composed of a barrel and endcaps, support/translation/alignment systems for each movable section and restraints to provide structural stability. The endcap doors are designed to split on a vertical line for access.

The principal hazards associated with the magnet are: cryogenic, electrical, magnetic, and the control of energy during a quench.

Liquid helium (LHe) will be utilized to cool the solenoid. Issues of concern are leakage, pressurization/vacuum, rupture, oxygen deficiency, and cryogenic burns. The helium plant will be located outside the hall. It will contain a 4000  $\ell$  dewar. The total volume of LHe in the experimental hall including the Q1 magnets and the solenoid is expected to be less than 250  $\ell$ .

Leakage of the LHe will be controlled in the distribution system by the use of coaxial lines with pressure switches interlocked to the supply valves. If vacuum in the line is lost (any leak will degrade the vacuum shield), the supply of LHe will be automatically interrupted at the source. Overpressurization of the aluminum restraining hoop cooling tubes and the cryostat will be controlled by pressure release valves and burst discs. The worst case release of LHe would require a catastrophic rupture of the dewar (such as a fork lift tine puncture). Barriers will be strategically positioned to prevent this kind of accident. If the dewar were opened in this way, the LHe would immediately evaporate and dissipate. Asphyxiation will be a hazard in enclosed areas. One liter of LHe vaporizes to approximately 750  $\ell$  of gaseous He at 20°C and atmospheric pressure.

The solenoid current will be 7000–8000 A. Appropriate protection for the power supplies, cabling, and connectors will be required.

Whenever the magnet is energized, it poses a hazard due to the damage that may be caused by objects, such as tools, being sucked into it. Procedural controls will be established to control the use of equipment during this time period.

A quench detection and energy extraction system will protect the coil in the event of a quench. The system to detect a quench will include the sensing of a voltage imbalance or an increase of pressure. The energy extraction system will utilize external dump resistors. Peak voltage across the magnet terminals during a quench will be less than 500 V.

## 13.10 Generic Hazards

---

The following is a brief discussion of the more general hazards:

**Earthquake.** The earthquake design requirements for the detector structure are defined in the SLAC Seismic Design Manual (this is more stringent than the Uniform Building Code). The Earthquake Safety Committee will review and approve the design based on compliance with the Seismic Design Manual [SDM91].

**Radiation.** Established radiation protections systems at SLAC include the Beam Containment System (BCS), the Personnel Protection System (PPS), the Beam Shut-Off Ion Chambers (BSOIC), shielding systems, and a radiological training program. For *BABAR*, the design of the BCS and shielding, implementation of the PPS, placement of the BSOICs, and development of operational requirements and procedures will be reviewed by the Radiation Safety Committee.

**Fire.** An evaluation of the fire prevention and fire protection systems will be performed. Life safety provisions will be provided for the facility in accordance with the NFPA Life Safety Code. Noncombustible materials will be utilized where possible.

**Occupational.** The SLAC safety standards for Environmental Safety and Health are contained in the SLAC ES&H Manual. During assembly, commissioning, operation, maintenance, and repair, the safety rules and regulations contained in this manual are applicable to SLAC personnel, collaboration personnel, subcontractors, and users.

Hazards to personnel during maintenance and repair of components inside the detector are under evaluation. The forward and backward end doors are designed in two halves, which allows retraction on rails for access to the aerogel, forward end of the drift chamber, and

calorimeter. Access to the backward components of the drift chamber will be possible through the DIRC strong tube. A rail system allows the endcap iron plug to be retracted back to a point where the crane could be used to set it aside. The rail system may then be used to support a cart to carry personnel. A full-size mock-up will be used to develop and simulate emergency evacuation procedures. The hazard of entrapment and the potential for hazardous atmospheres, may classify the inside of the detector as a permit-required confined space.

The hazards associated with elevated work will be considered in the design. Platforms, railings, barriers, stairways, and ladders will be designed in accordance with the Life Safety Code.

**Environmental.** Consideration of environmental issues such as the generation of hazardous waste, the uses of freon, the release of gas mixtures, and environmental reporting requirements will be addressed.

## References

---

- [CFR93] Code of Federal Regulations, Title 29, 1910.1000 (1993).
- [SDM91] Earthquake Safety Committee, “Seismic Design Manual—Mechanical R1 Draft 2,” (1991).

---

---

## Facilities, Assembly, Access and Integration

This chapter addresses the plan for on-site assembly of the detector subsystems and their installation in the experimental hall. It includes consideration of facilities required to assemble the detector, and presents the philosophy [BAR93] for integration of facilities and technical components in a smoothly flowing, continuous process of assembly and installation.

### 14.1 Facilities

---

Although many of the *BABAR* detector components will be fabricated elsewhere, the final assembly and testing procedures will be done at SLAC in the IR-2 hall where the experiment will be operated. This facility is large enough to accommodate, during the operation phase, the detector on the beam line, the radiation shield wall, and detector services on the apron (east side of the facility). The apron provides adequate working and storage space for the assembly of the detector before installation of the detector on the beam line.

Figure 14-1 shows the general configuration of the hall and indicates the designated north direction. There is a 5.96 m-wide by 8.53 m-high roll-up door at the east end of the building. All components associated with the erection and operation of this detector are designed to fit through this door. This facility is equipped with an overhead bridge crane that will be used during installation and maintenance. The rated capacity of this crane is 45 tonne with an auxiliary lift rated at 9 tonne. The crane has a maximum hook height from the concrete foundation of approximately 10.5 m and 11.1 m for the 45 tonne and 9 tonne lifts, respectively.

The building is divided into two separate areas by a concrete shielding wall. The apron in the east area will be used for the erection of the detector and initial testing. The west area will house the detector during operations. After the detector has been commissioned and experiments have begun, the apron will house the electrical, cryogenic, and data analysis support systems as well as maintenance and upgrade facilities. Additional offices for support personnel will be located on the second floor of the building.

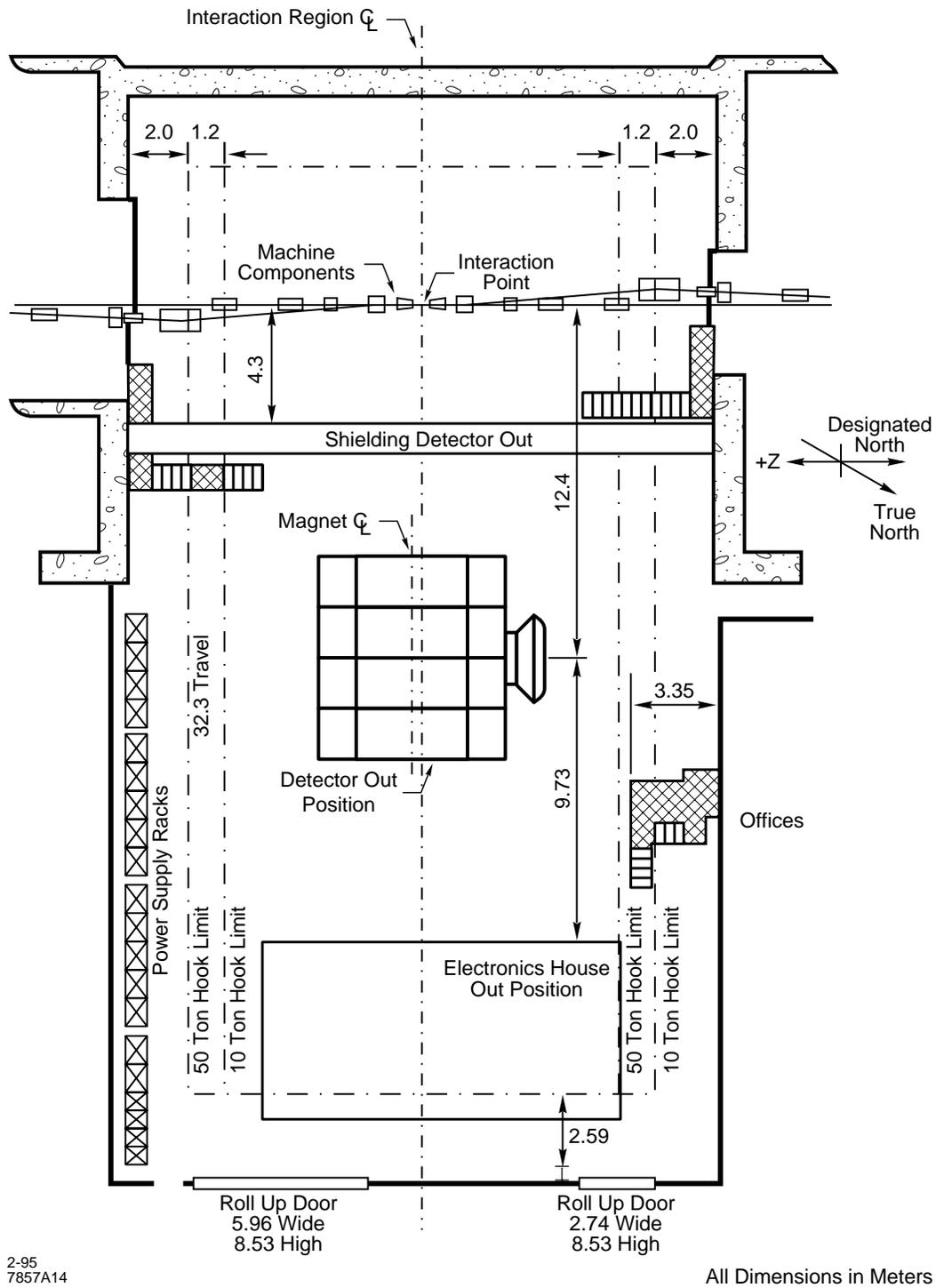


Figure 14-1. General plan of interaction hall with detector in assembly position.

---

## 14.2 Detector Coordinate System

---

The *BABAR* detector's right-handed coordinate system has its origin 370 mm axially from the center of the magnet. When installed on the beam line, the origin of the detector coordinate system will be coincident with the interaction point. The positive  $z$  axis is coincident with the detector axis and points toward the forward end of the detector. The positive  $y$  axis points up. The positive  $x$  axis points east. During assembly, the forward end of the detector is oriented toward the south end of IR-2. After installation on the beam line, the positive  $z$  direction will be set at an angle of 20 mrad with respect to the direction of the high-energy beam.

---

## 14.3 Structural Support of Systems

---

Because SLAC is located in a seismically active area, special care must be taken to design equipment to withstand seismic activity. Meeting this demand can be greatly eased by the use of earthquake dampers. These dampers can be expected to reduce the peak acceleration in the detector with respect to the ground by a factor of 3 or more. Such dampers were used in the support of buildings in Kobe, Japan, which suffered much less damage in the 1995 earthquake than neighboring buildings supported in a conventional manner. Dampers suited to the detector are available and were used as part of the SLD calorimeter assembly fixture.

The barrel steel rests on four feet which in turn rest on earthquake dampers resting on the floor. These feet have mechanical jacks to adjust the height of the detector. Both the forward and backward end doors are supported in the same manner as the barrel steel. The exception is that during transport of the detector, the doors are rigidly attached to the barrel steel.

The coil in its cryostat is supported inside of the barrel steel. These supports are strong enough to withstand the magnetic de-centering forces on the coil and the seismic forces, which are substantially larger than the expected unbalanced magnetic forces. The calorimeter is suspended from the cryostat. The DIRC is supported by the barrel steel so that the doors at each end can be opened to service the calorimeter electronics.

It is important to protect the relative alignment of the drift chamber and the vertex detector against temperature variations in the hall. The vertex detector is supported by the B1 magnets which are, in turn, mounted inside the support tube which holds the machine components B1 and Q1. The support tube is attached to the DIRC strong support tube in the backward direction and the barrel steel in the forward direction. In order to preserve the relative alignment of the vertex detector and the drift chamber, the drift chamber is

supported at its inner radius by the support tube. Large seismically induced relative motion of both Q2/Q4/Q5 pedestals with respect to the support tube must be minimized. The earthquake motion problem is exacerbated by the use of the earthquake dampers beneath the detector: the detector will probably move less than the pedestals.

## 14.4 Installation Overview

---

A preliminary installation plan, with schedule, has been formulated for the assembly of the *BABAR* detector in the IR-2 hall. The detector installation requires approximately 14 months. An additional month is needed for installation of the beam line components associated with the interaction region.

The installation critical path is driven by the configuration of the detector systems. The barrel assembly proceeds inward radially. The flux return is assembled. Then the solenoid is installed within that, and the barrel calorimeter is inserted into the solenoid. The DIRC is next installed, and the drift chamber inserted into the DIRC. The aerogel/forward calorimeter assembly is installed. The support tube is inserted, and the doors, containing their RPCs, are closed before a cosmic ray test. Each system's electronics is installed simultaneously with its mechanical components.

While the detector is being assembled and tested in the east end of the IR-2 hall, a concrete shielding wall is placed between the detector and the interaction point so that studies of the machine components may occur during detector assembly. The eastern edge of this wall is located approximately 4.5 m east of the interaction point.

After testing of the assembled detector is complete, the concrete shielding wall is disassembled, and a transport mechanism is used to move the detector to its final position in the PEP-II colliding beams. Figure 14-2 shows the detector in its operational location.

## 14.5 Detector Component Installations

---

### 14.5.1 IR-2 Detector Hall Preparation

Base plates of 50 mm nominal thickness are provided on the floor of the experimental hall. The plates are leveled using hex bolts and secured to the concrete foundation with adhesive anchors. The tops of the plates are positioned 60 mm above the floor of the experimental hall. After erection and leveling of the base plates, the space beneath the plates is grouted solid with a nonshrinking, cementitious, flowable grout. The purpose of the grout is to provide

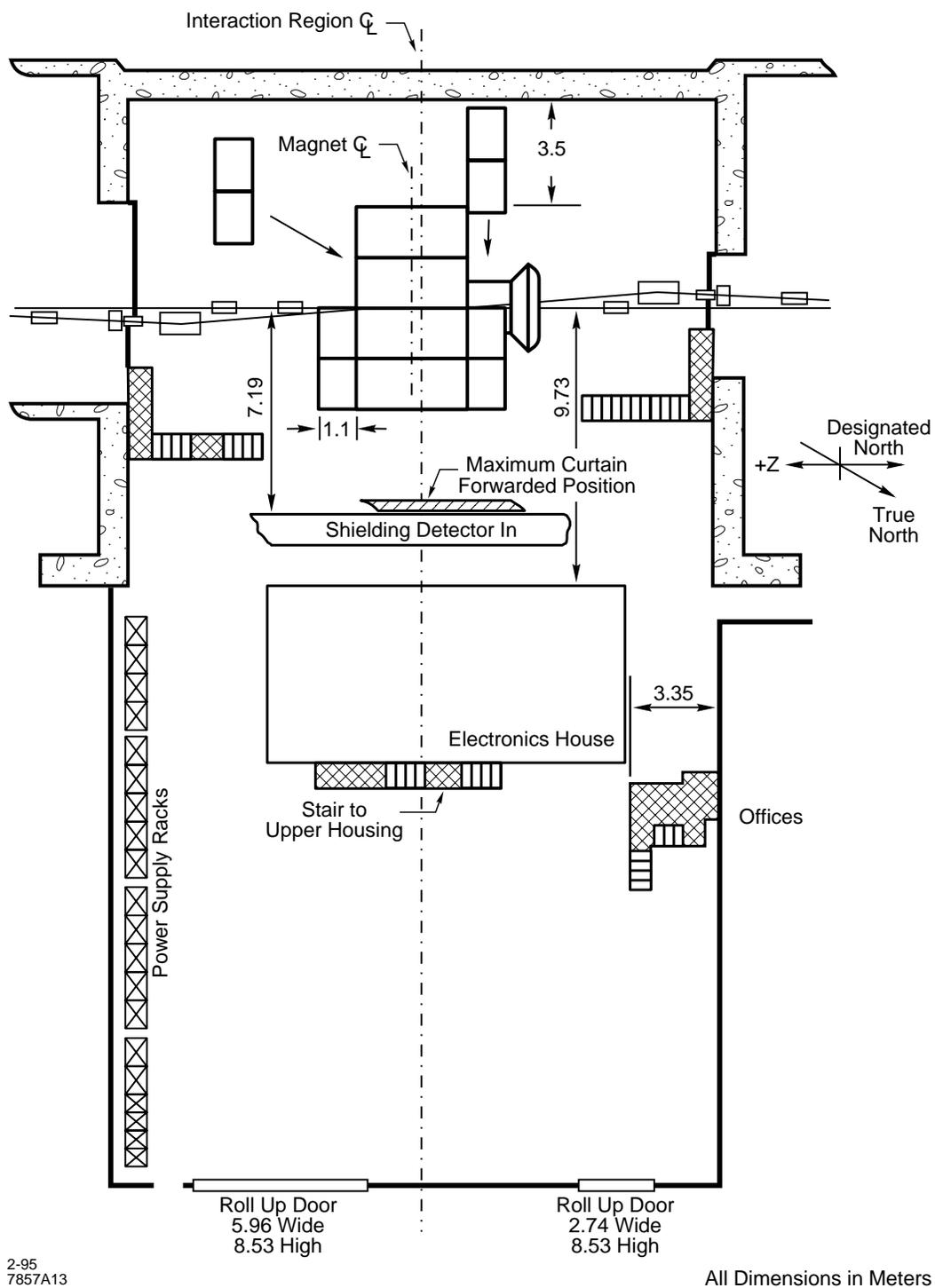


Figure 14-2. General plan of interaction hall with detector on beam line.

the base plates with a level supporting surface and to provide uniform bearing between the plate and the concrete foundation. The grout is placed as a continuous flow without air pockets and allowed to cure 28 days prior to commencement of the erection of the barrel flux return [KRE93].

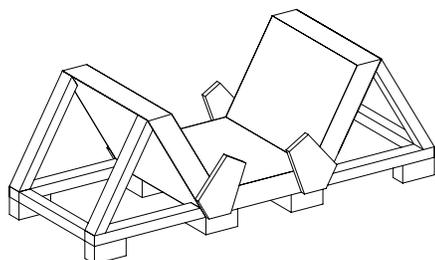
### 14.5.2 Coil and IFR Installation

The assembly of the barrel flux return assembly is begun at a point approximately 12.4 m east and 0.5 m north of the interaction point. The barrel axis is parallel to the bridge of the overhead crane. During assembly, it will be necessary to move the detector barrel toward the south end of the building. The magnet barrel is a self-supporting structure and houses its own transport mechanism. The transport mechanism will be tested before installation commences. Figure 14-1 shows the location of the detector in IR-2 during the assembly process.

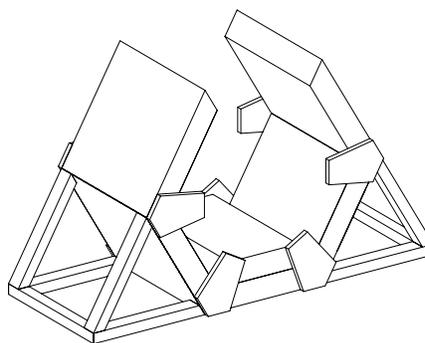
The barrel flux return is composed of six inner blocks and six outer blocks. The barrel RPCs are installed in these blocks while the blocks are horizontal. The muon system is serviced by a variety of signal cables, high voltage cables, and gas piping. All of the connections to the individual detectors are made during the assembly process. During the assembly process, the 45 tonne capacity overhead crane is used to lift the blocks into place. Figure 14-3 shows the assembly procedure.

The electronics platforms and stairways will be partially preassembled into large components and installed on the outer surfaces of the barrel flux return at this time. The manifold connections for the barrel RPCs are now made. These connections link the barrel detectors to the hall utilities [TRI92]. These manifold connections are located on the upper service platform. The manifold gas utility feeds run along the west wall of IR-2 and connect with the detector at both ends of the hall. The signal cables and high voltage cables are routed to crates located externally to the detector on the electronics platforms. The detailed routing of services from the RPC detectors to the crates and manifolds is still being investigated. All connections are inspected and tested to ensure proper operation of the system.

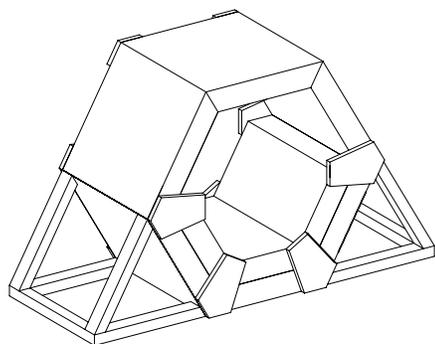
The solenoid junction box and cryo-chimney are mounted to the cryostat prior to its installation. The cryostat is installed into the barrel flux return from the backward direction as shown in Figure 14-4, using a beam and carriage system that will transport the coil's gravitational load along the axis of the beam. To install the coil into the magnet steel, a structural steel beam, referred to as the transfer beam, is inserted through the bore of the coil from the backward side. There are three tripod supports for this beam. One is located at the north end of the assembly area and a second at the south end. The third tripod support, the floater, is positioned on either the north or south side of the detector, depending on the device to be installed. For insertion of the coil, the floater is positioned on the north



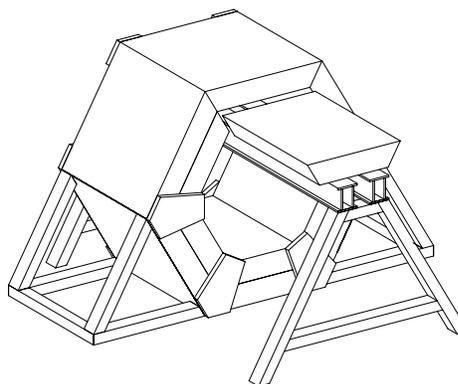
**Step 1**  
Position the bottom outer block on spacers. Attach the two diagonal outer blocks with their support legs.



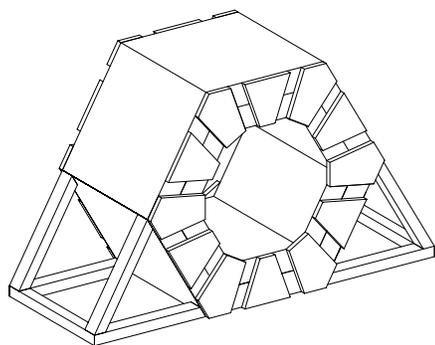
**Step 2**  
Remove the spacers. Install the bottom inner block. Attach the upper diagonal outer blocks.



**Step 3**  
Install the lower diagonal inner blocks. Attach the top outer block with the far corner braces.

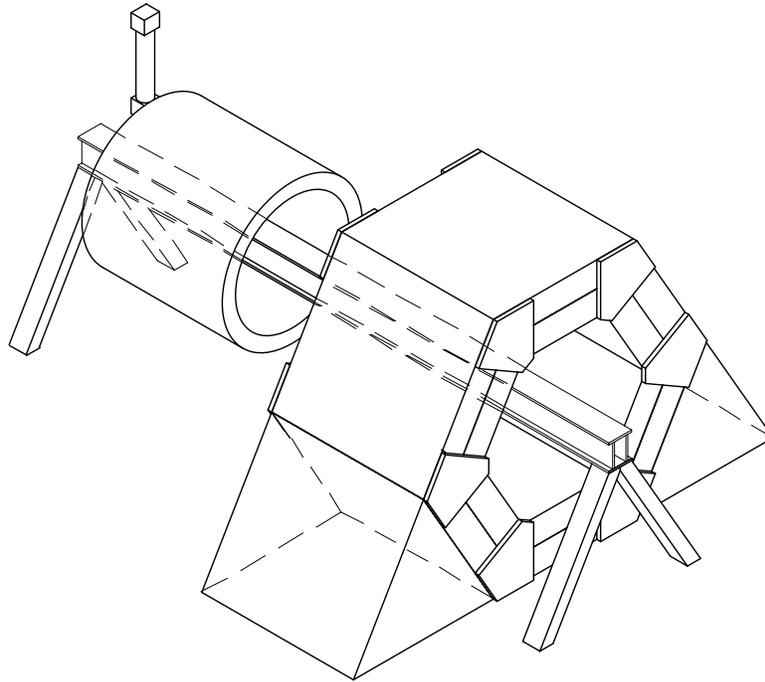


**Step 4**  
Install the upper diagonal inner blocks. Erect a raised platform to install the top inner block.



**Step 5**  
Roll the top inner block in and attach the corner braces. Remove the platform. Add the gap spacers.

**Figure 14-3.** *Magnet barrel assembly.*



**Figure 14-4.** *Installation of the Solenoid using a beam and carriage system.*

side. The transfer beam is temporarily supported on the two tripod supports on the north side and the strongback carrying the barrel calorimeter is positioned on the forward side of the barrel flux return. The transfer beam is then threaded through the coil and secured to the tripod located at the south end of the assembly area. The tripods are provided with  $x$  and  $y$  translation capability to enable accurate alignment of the coil relative to magnet steel. The coil will be cooled and the cryogenic safety systems tested. The cryogenic services are supplied via flexible transfer lines from the helium liquefier/refrigerator which is located outside the IR-2 experimental hall, near its southeast corner. The flexible transfer lines are routed along the building periphery from the refrigerator to the west wall. These lines pass through the wall to the cryo-chimney. This chimney is accessible from the upper platform. The location of the magnet dump resistor and the magnet power supply has not been decided. It is preferred that the power supply be installed on the apron outside the radiation wall to permit access for maintenance.

The DIRC support structure is mated to the backward end of the barrel flux return assembly using the overhead crane. This assembly consists of the DIRC strong support tube and the structural bracing that connects the assembly to the barrel steel. The structure is properly aligned and bolted into place.

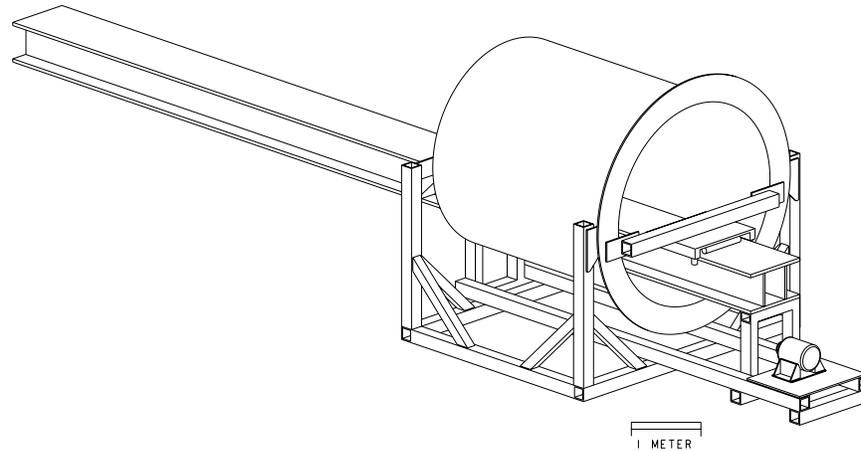
Each endcap flux return is assembled as two doors. Each end door has a transport mechanism built into its base. Using the overhead crane, an assembly of 20 vertical plates, containing RPCs which were installed along with cables and services while the assemblies were horizontal, is mounted on the base along with a 30 tonne counterweight. The counterweight helps minimize the potential for overturning during movement of the doors. The forward end doors also contain the plug assembly that is used to divert the magnetic flux away from the beam line components. The doors are attached to the barrel for magnet testing. The backward steel plug is temporarily inserted into the backward end doors and removed after field mapping is completed. Services for the door RPCs are connected.

The coil is powered at low current and the electrical safety systems tested. While the current is ramped up, the cooling system and decentering forces are monitored. It may be necessary to de-energize the magnet and adjust the position of the solenoid at this time to minimize the forces on the coil due to nonuniformities in the flux return. After initial testing is complete, the doors are opened, and the field mapping hardware is installed. NMR probes are installed to monitor the field during field mapping. These probes will be removed after field mapping, and reinstalled after drift chamber installation. After the field is mapped, the doors are removed and stored against the wall of the IR-2 experimental hall, and the field mapping hardware is removed. The location at which the end doors are to be stored has not yet been decided.

### 14.5.3 Barrel Calorimeter Installation

After magnetic mapping is complete, the installation of the barrel calorimeter commences. The fully aligned, cabled, and tested barrel calorimeter arrives at IR-2 from the calorimeter assembly building, B109. A system test is performed to ensure that no damage occurred during transport to the hall. A final check of the module alignment may also be done at this stage if necessary.

The installation of the calorimeter into the coil bore is very similar to the installation of the coil into the magnet steel. Once the transfer beam has been threaded through the coil a structured beam is bridged between two pins on each side of the calorimeter and tied to the end-flange pick up points. To maintain internal alignment of modules the pickup points and support points must be the same. Figure 14-5 shows the installation method and the transfer of loads. The flange on the forward end of the calorimeter support structure is bolted directly to the end flange of the solenoid cryostat. Eccentric pins on the end flange of the cryostat allow the horizontal and vertical positioning of the calorimeter. The backward end of the calorimeter support structure is secured to the cryostat using adjustable radial support arms. Shimming may be required between the flanges of the calorimeter and the cryostat.



**Figure 14-5.** *Sequence of barrel calorimeter installation into the coil.*

The distribution system for the freon cooling for the barrel calorimeter is installed at this time. It consists of multiple manifolds which are routed to the calorimeter along the cableways between the barrel flux return and the forward and backward doors. The fiber-optic cables and the dry nitrogen supply and return lines are also installed in this space.

#### 14.5.4 DIRC Tube Assembly Installation

The central support tube containing the DIRC quartz sectors and the standoff box assembly are transported to IR-2 while the barrel calorimeter is being installed.

The transfer beam is removed from the forward end of the detector barrel, which is then moved 0.87m to the south. No additional movement of the barrel will be necessary prior to installation into the beam line. The floating tripod support is then moved to the forward end of the barrel.

Using the overhead crane, the DIRC tube assembly is positioned at the backward end of the barrel. The transfer beam is then threaded through the DIRC tube assembly and the bore of the barrel calorimeter, and secured to the tripod support stands. The gravitational load of the DIRC tube assembly is transferred to the transfer beam and the DIRC tube assembly is rolled into the bore of the barrel calorimeter. The DIRC is aligned and secured to the strong support tube of the DIRC support structure. The central support tube is then suspended at the forward end from the barrel calorimeter end flange using temporary supports. The transfer beam and the tripod support stands are removed from the barrel and moved to storage.

Using the overhead crane, the DIRC standoff region is now mated to the backward flange of the DIRC strong support tube and securely fastened. A structural steel pedestal is then inserted and installed on the lower surface of the strong support tube. The pedestal is provided with a set of rails that extend beyond the backward end of the DIRC standoff box.

### 14.5.5 Drift Chamber Installation

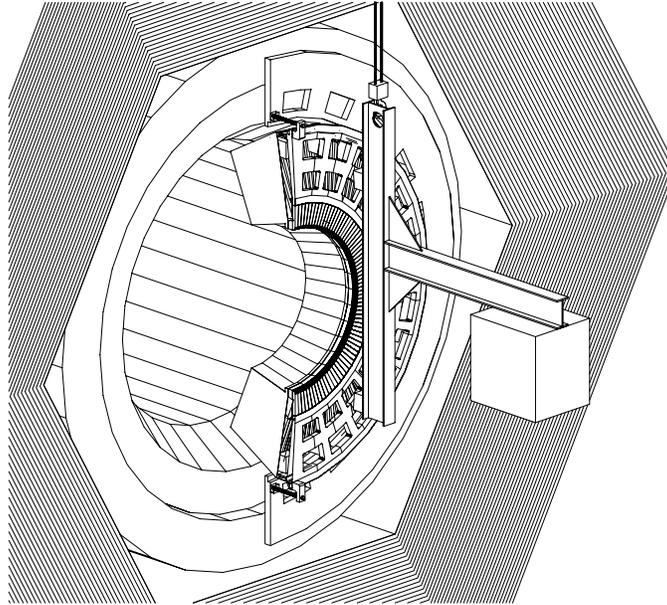
While the DIRC is being installed, the drift chamber and its assembly fixture are mated to an insertion frame. The base of the insertion frame is provided with rollers and also with vertical height adjustment capability. The drift chamber temporary backward support frame is inserted through the DIRC assembly from the backward end and secured. The drift chamber is then inserted into the bore of the DIRC central support tube from the forward direction. The drift chamber is properly positioned and temporarily cantilevered from the its backward support frame. The drift chamber insertion frame is then removed and taken to storage. The temporary supports that hold the DIRC central support tube are then removed from the forward end of the barrel.

The signal cables, high voltage cables, and gas lines exit the drift chamber from the backward end where they are routed outward radially to the inner surface of the DIRC strong support tube. They are then routed in  $z$  along this surface until they exit the detector at the backward end of the DIRC standoff box. The detailed routing of these cables and gas lines is under evaluation.

### 14.5.6 Aerogel/Forward Calorimeter Installation

The assembly frame, carrying the loaded and tested endcap segment, will be transported to the experimental area where the endcap segment is transferred to the installation jig. The aerogel segment is attached to the front of the calorimeter. The half-endcap is brought into its mating position with the barrel, defined by conical pins locating into holes drilled into the barrel end flange, and fixed in place, as shown in Figure 14-6. At this stage in the installation, there is 5–10 mm between the adjoining barrel and endcap conical surfaces. The endcap is aligned in the vertical plane to ensure uniform clearance, and final fixing holes drilled. The endcap is now in its *park* position.

The second endcap is similarly installed in its park position and the two segments bolted together to form the complete endcap, making a conic section. Provision is made for the endcap to be wound into the barrel, using adjustable bolts at three points on the barrel flange, and contact switches used to enable precision mating. Removal of the endcap involves winding it back to its park position, unbolting the two segments, attaching each in turn to



**Figure 14-6.** Schematic view of the installation of an endcap segment including mating to the barrel forward flange. (The aerogel is not shown.)

the installation frame and removing them one after the other. This procedure decouples the endcap from the barrel and ensures that the alignment of the latter is not disturbed.

The distribution system for the freon cooling for the forward calorimeter is installed at this time. These cooling tubes are routed along the cableways between the barrel flux return and the forward end door. The high voltage and signal cables for the aerogel are routed out and around the forward calorimeter. They also exit the detector radially through the space provided between the barrel assembly and the forward end door.

### 14.5.7 Vertex Assembly Installation

The vertex assembly that contains the beam pipe, vertex detector and the two B1 and two Q1 beam line magnets is installed from the forward direction. This is done by cantilevering the assembly from a temporary external support fixture and carefully inserting the assembly through the bore of the drift chamber using the overhead crane. When the vertex assembly is properly positioned, the permanent drift chamber supports are installed on both ends. The vertical pole structures that support the vertex assembly from the barrel flux return are installed in the backward and forward ends of the detector. The temporary supports on the drift chamber are then removed.

### 14.5.8 Backward End Plug Installation

Using the overhead crane, the backward end plug is mated with the DIRC rail system. The backward end plug is rolled into position and securely fastened to the DIRC strong support tube.

### 14.5.9 Electronics House

The initial processing of data occurs in the detector-mounted electronics. This equipment is not accessible while beams are circulating. The closest accessible electronics are located in the electronics house. The quantity and layout of the electronics located there are being evaluated. The electronics house is mounted to a structural steel frame. The frame is provided with heavy-duty rollers similar to those used for the detector barrel. Prior to movement of the detector into the beam line, the frame of the electronics house is mechanically connected to the barrel support structure so that they move as a single mechanical unit. The concrete radiation shielding wall is re-erected between the detector and the electronics house after the detector is moved onto the beam line.

### 14.5.10 Detector Transport System

The detector transport system, which is similar in concept to those used on previous detectors, is used to move the detector in the IR-2 hall. It consists of a structural steel frame with two hydraulic cylinders. The frame is bolted to the floor base plates, and the cylinder rams are pinned to the detector assembly. Pressurizing the hydraulic cylinders extends the rams, thus pushing the detector. When the ram reaches its maximum travel, the transport system is removed and the rams retracted. The system is then reinstalled closer to the detector and the process is repeated until the detector is located in the desired position.

### 14.5.11 Final Position of Detector in Beam Line

After the detector has been moved to the interaction point, it is rotated 20 mrad about its vertical axis using the detector transport system. The support legs of the detector rest on seismic isolation pads; the pads will absorb energy caused by ground motion during an earthquake. Jacks provided in the barrel assembly allow the vertex detector to be aligned vertically with the interaction point. The detector legs are shimmed appropriately at the seismic isolation pads and secured to the concrete foundation.

### 14.5.12 Service Space

Services for the various systems enter the detector via three paths. The first path is radially out between the barrel steel and the door. A gap of 100 mm is provided between the coil cryostat and the door for calorimeter services. The blocks between the barrel steel and the door occupy about 40% of the azimuth. These are placed at 60° intervals at the corners of the hexagon. Beyond the coil, a gap of 150 mm is provided between the barrel steel and the door. This allows space for the IFR services. All services passing this way will be in closed conduit to assure that they fit in the space available and to protect them when the door is open and when the door is being closed.

The services for the drift chamber exit the detector at the backward end along the inside of the DIRC tube. They proceed through the DIRC and around the water tank.

Routing of vertex detector services has not been resolved.

## 14.6 Detector Maintenance Access

---

Ready access to detector components eases the reliability requirements on those components, thus reducing cost without compromising performance. The ability to make accesses of short duration when needed to ensure data quality reduces loss of beam time. Parasitic use of accelerator downtime allows repair of less pressing detector problems if access times are commensurate with typical planned and random accelerator outages [BAR93]. The design timescale for access for minor repairs is a few hours to a day. Annual accelerator shut-downs will be used for more extensive repairs and for upgrades. Access to the electronics house is not limited. Access to electronics, power supplies, and services located inside the radiation wall but outside of the detector is possible by turning off the beams. All DIRC readout components can be repaired in this way. The timescale for such a controlled access is less than an hour.

Repairs that require access to the inside of the detector will be discussed in the balance of this section. Before the doors are opened, the solenoid is de-energized. However, the solenoid will remain cold. Only when specifically required will the solenoid be warmed to room temperature.

### Forward Detector Internal Access

Two schemes for access into the detector at the forward end are under consideration. In the first, the doors are moved outward at an angle of 30° relative to the beam line. This

option provides 4.6 m clearance between the barrel and the forward doors. In the second, the doors are moved outward at an angle of  $85^\circ$  relative to the beam line. This scheme provides minimal but acceptable access to the RPCs and minimizes potential conflicts with machine components Q2, Q4 and Q5. In both schemes, the barrel and endcap calorimeter electronics are readily accessible. The aerogel inner ring of readout devices can also be repaired. The time needed to open the doors is approximately one hour. In order to access the end of the drift chamber or the outer ring of aerogel readout, it will be necessary to remove one or both endcap segments. This requires use of special fixtures during a longer access.

### Backward Detector Internal Access

Inner detector access in the backward region is complicated by the extension of the DIRC assembly through the backward end doors. The barrel RPC detectors and the electronics on the backward end of the barrel calorimeter are accessed by opening the backward flux return end doors. These doors open perpendicular to the magnet axis. Access to the vertical DIRC support plate and the cryo-chimney is also available. The time needed to open the backward doors is approximately one hour.

The backward end of the drift chamber and some of the vertex detector electronics can be accessed by removing the backward end plug. This requires installation of a special jig and rails. This operation will take about one shift. The backward end plug is rolled back until it is positioned beyond the DIRC standoff box. It is then removed using the overhead crane. Once the plug is removed, it is possible to enter the interior of the DIRC along the beam line to access the face of the drift chamber. Quarters will be tight, since the inside diameter of the DIRC is only about 1.6 m. Special jigs will provide a temporary floor for access.

Repair of drift chamber wires requires that the forward doors be opened, the forward calorimeter and aerogel segments be removed, and the backward end plug removed.

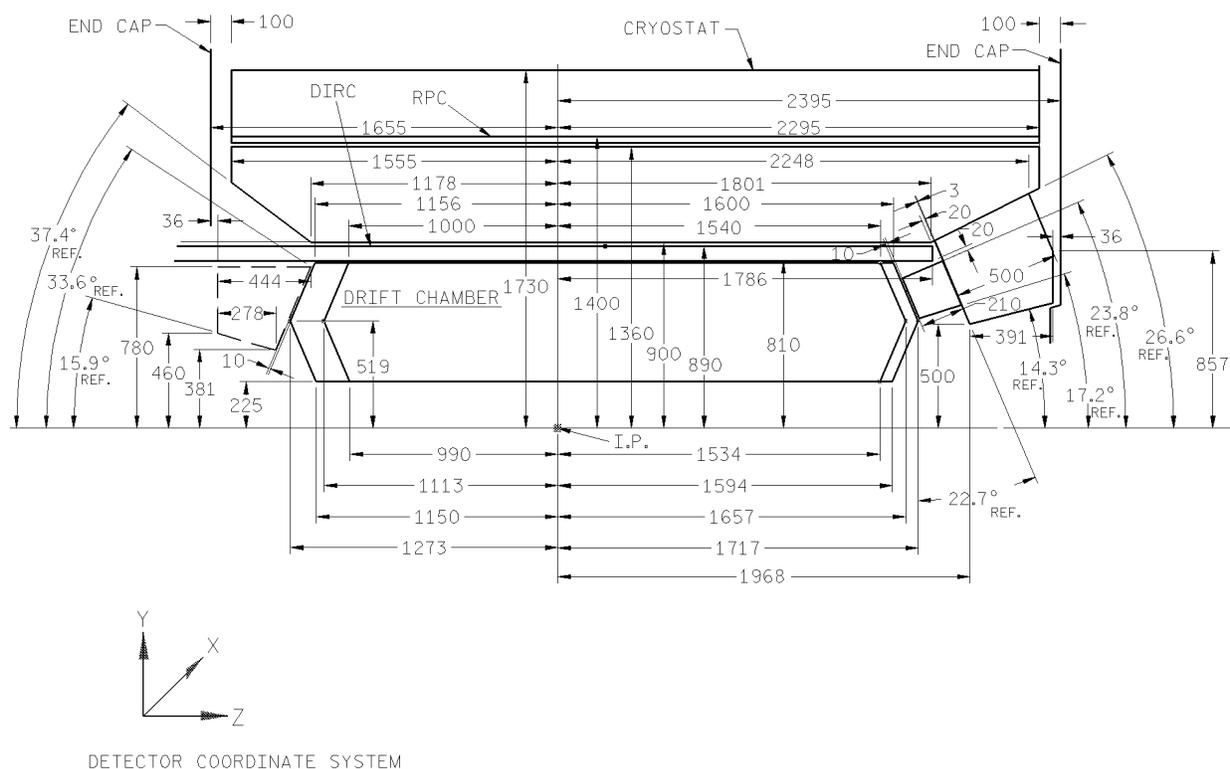
### Vertex Detector Access

Access to the vertex detector will require removing the support tube from *BABAR*. This requires dismantling numerous beam line components and will entail substantial downtime.

## 14.7 Detector Integration

---

The Detector Integration Group is responsible for ensuring integration of the detector subsystems and of the detector with the collider and the experimental facility. The detector



**Figure 14-7.** Overall detector dimensions(mm).

integration group maintains control of the physical interfaces including envelope dimensions, mechanical attachments, obscurations, cable and utility routing, and alignment [KIR93]. Figure 14-7 shows a cross section of the central section of the *BABAR* detector indicating the subsystem envelopes.

The detector integration group has the responsibility for installation of the detector elements in the experimental hall as well as for preparation of the experimental hall prior to commencement of the detector assembly process.

### 14.7.1 Assembly Clearances

In order to avoid damage to the detector systems during the assembly process, an inter-system clearance of 20 mm is generally required. The 20 mm criterion is based on experience

with SLD. The 20 mm clearance is intended to provide for inaccuracies in fabrication of the large parts, for tolerance build-up, and for assembly without extremely elaborate guidance fixtures. In addition, the clearance must be adequate to accommodate any relative motion of systems during an earthquake. As a typical example, the clearance between the cryostat and the inner flat RPC is 20 mm.

An exception to the 20 mm rule has been made for the clearance between the drift chamber and the DIRC central support tube because it is expected that both systems can be made with close circularity and straightness tolerances and that a simple but effective installation fixture can guide the drift chamber inside the DIRC for assembly. Another exception is the gap between the aerogel and the forward calorimeter, where only 3 mm are provided because the aerogel will be mounted directly on the calorimeter. Only 10 mm is provided between the aerogel and the drift chamber because these objects are not large and because there is good visibility available during the assembly process. The clearance between the forward calorimeter and the endcap steel is 36 mm because the steel is not expected to be very flat and is expected to deflect by about 8 mm due to magnetic forces.

### 14.7.2 Good Neighbor Policy

Each system must obey a Good Neighbor Policy, *i.e.*, the system must not pass its problems to its neighbor without the neighbor's agreement. For example, each system must dispose of any heat it generates without simply passing it on to its neighbor. This is particularly important in cases in which temperature stability is required for a system for mechanical alignment or accuracy of response. Each system must return its temperature to ambient by whatever means is required, and do so within its own geometrical envelope.

Heat is not the only operating problem. Electrical and mechanical noise must also be considered. Each system must take care to prevent or contain such interference with its neighbors.

It is expected that each system will take care to minimize the risk of leakage of fluids or gases which can harm other systems. For example, a leakage of helium could damage the phototubes of the DIRC, and water could damage CsI crystals.

### 14.7.3 Common Services

Use of common fixtures in the assembly process will be optimized. For example, the fixture used to install the cryostat may be suitable for installing the barrel calorimeter.

Easily erected, modular scaffolds will be used during the assembly phase and during the operation stage. Such scaffolds were employed to great advantage in SLD.

A centralized system will provide HVAC (heating, ventilation, and air conditioning), clean AC power, emergency power, uninterruptible power, and oxygen deficiency and fire alarms.

---

## References

---

- [BAR93] B. Barish *et al.*, “GEM Technical Design Report,” GEM-TN-93-262 (1993).
- [KIR93] T. Kirk *et al.*, “Solenoidal Detector Collaboration—Project Management Plan,” SGT-000006 (1993).
- [KRE93] H.J. Krebs *et al.*, “Solenoidal Detector Collaboration—Preliminary Design Report—Muon Barrel Toroid Support System,” SDT-000168 (1993).
- [TRI92] G. Trilling *et al.*, “Solenoidal Detector Collaboration—Technical Design Report,” SDC-92-201 (1992).



---

---

# Collaboration Issues and Project Management

The *BABAR* Collaboration had its inaugural meeting at SLAC in December of 1993, soon after the approval of the PEP-II accelerator project in October. An Interim Steering Committee was appointed to organize and direct the Collaboration until a permanent organizational structure could be set up.

The Steering Committee, in consultation with the collaboration membership and laboratory management, drafted a governance document for the Collaboration. The governance document was approved by the newly formed Collaboration Council in May of 1994. Following the procedures described in the governance document, the officers and management of the collaboration were selected. These structures are now fully in place.

In this chapter we summarize the collaboration organization as detailed in that document [BAB94] and describe the organization of the project management aspects of the construction and operation of *BABAR*. Figure 15-1 shows a schematic organization chart of the *BABAR* Collaboration.

## 15.1 Membership

---

Ph.D. physicists, engineers, and Ph.D. thesis students who contribute significantly to the *BABAR* detector, as well as those who contribute significantly to the accelerator and plan to participate in the physics program, are eligible to be members of the collaboration. Through the submission of the Technical Design Report, membership has been open to all individuals who meet these criteria, and their institutions. Henceforth, new groups may apply for membership in the collaboration by submitting an application to the Spokesperson, which will be voted on by the Collaboration Council.

The Technical Design Report has been signed by 77 institutions from 10 countries.

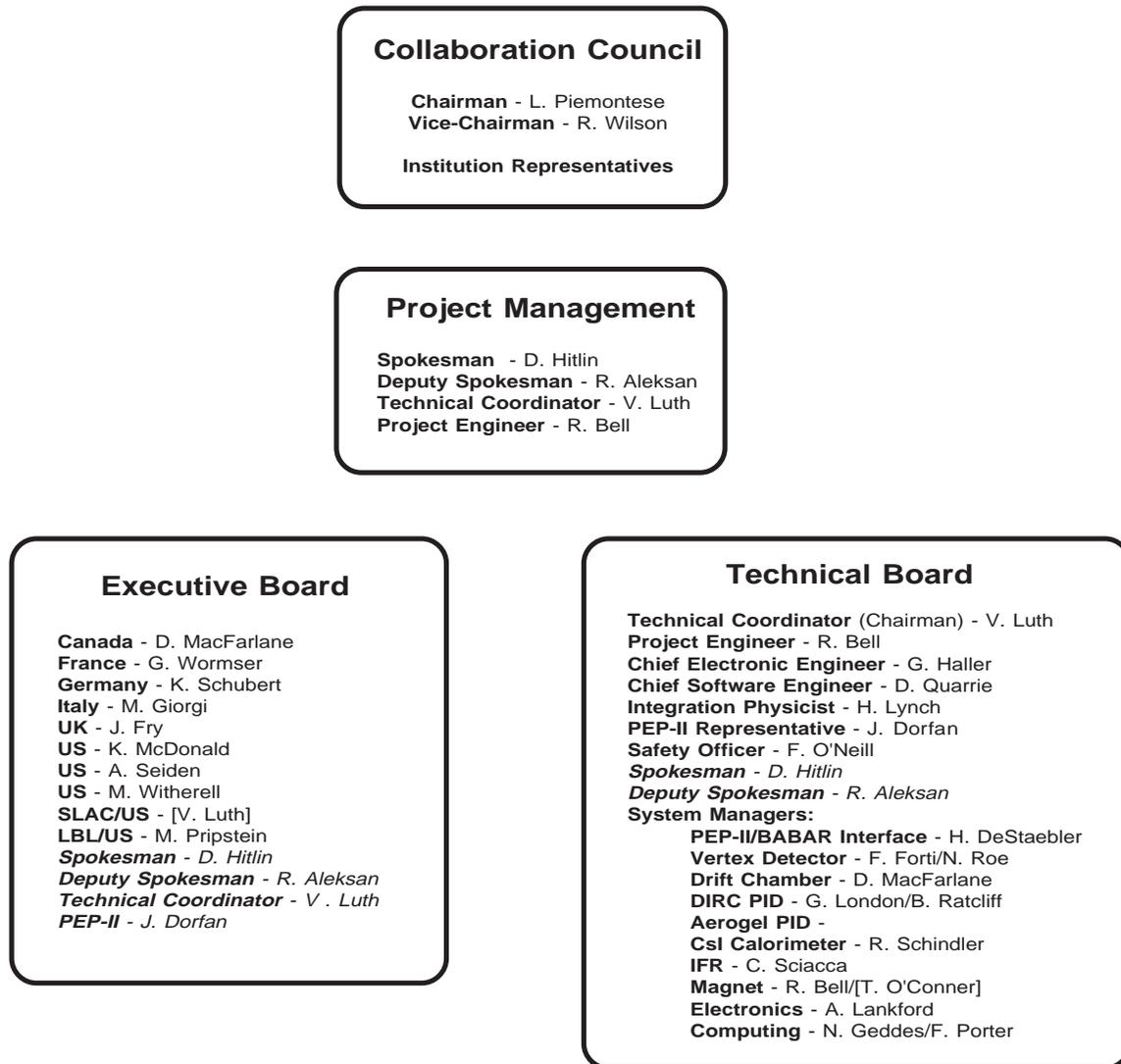


Figure 15-1. Organization of the BABAR Collaboration.

---

## 15.2 Collaboration Council

---

Institutions with three or more collaborating members who are Ph.D. physicists are represented directly on the Collaboration Council. Members of the collaboration from institutions with fewer than three Ph.D. physicists may affiliate with another institution for the purpose of representation on the Collaboration Council. Large institutions have one vote for every ten collaboration members. The Collaboration Council has an elected Chairperson and Vice-Chairperson.

The Collaboration Council deals with issues related to the overall framework of the collaboration. It is responsible for membership policy, publication policy, and the selection of speakers for conferences.

The Council also appoints a Nominating Committee to nominate the Spokesperson every three years, in consultation with SLAC management and with the Collaboration at large. This nominee must be ratified by a two-thirds majority of the Collaboration Council. The Council may vote to remove the Executive Board, also by a two-thirds majority.

---

## 15.3 Spokesperson

---

The Spokesperson is the scientific representative of the collaboration, and is responsible for all scientific, technical, organizational, and financial affairs. On financial matters, the Spokesperson's authority is consistent with the requirements of the various funding agencies. The Spokesperson consults with the Technical Board on technical matters. In addition, the Spokesperson is responsible for keeping SLAC management and the Collaboration Council informed about collaboration affairs.

The Spokesperson, after broad consultation with the collaboration and SLAC management, nominates a Deputy Spokesperson, who must be ratified by the Collaboration Council. The term of appointment for both offices is three years, renewable.

---

## 15.4 Executive Board

---

The Executive Board advises the Spokesperson on all scientific, financial, and organizational matters. It consists of members distinguished by their scientific judgment, technical expertise, and commitment to the experiment. The Executive Board may remove the Spokesperson with a two-thirds majority vote. The membership will reflect the national composition of the

collaboration: initially, the Executive Board consists of one representative each from Canada, France, Germany, Italy, and the U.K., and five representatives from the US. This composition may change as the collaboration grows. The Spokesperson, Deputy Spokesperson, Technical Coordinator, and a PEP-II representative serve as non-voting *ex-officio* members.

The Executive Board meets eight times per year. Several of these meetings are held jointly with the Technical Board.

## 15.5 Technical Board

---

The Technical Board advises the Spokesperson on technical and financial matters. It consists of the Technical Coordinator as chairperson, the Project Engineer, a PEP-II representative, the Collaboration Safety Officer, the Spokesperson *ex-officio*, and the managers of the detector subsystems. In cases in which there are co-system managers, one of the co-managers sits on the Technical Board. The Spokesperson nominates the subsystem managers in consultation with the Executive Board and may nominate additional members of the Technical Board.

The Technical Board meets monthly to review the progress of the project. Several of these meetings are held jointly with the Executive Board.

## 15.6 Finance Review Committee

---

The Finance Review Committee monitors the financial aspects of the experiment as set forth in the management plan for the detector and agreed upon between SLAC and the collaboration. This will be detailed in Memoranda of Understanding between SLAC and each participating institution. The committee will be chaired by the SLAC Director of Research, and will include representatives of each of the funding agencies. The Spokesperson will normally also attend meetings of the committee.

## 15.7 Communications

---

Collaboration Meetings are held four times per year, typically in March, June, September and December. At least one meeting each year will be held outside the United States.

The international nature of the *BABAR* collaboration requires the use of up-to-date communications techniques to facilitate communication. Frequent use of video and telephone conferences promotes communication with the collaboration.

The *BABAR* World Wide Web Home Page, located at

*<http://www.slac.stanford.edu/BFROOT/>*

and reachable from the SLAC home page and from the home pages of many collaborating institutions, is a major center of communication. Weekly event calendars, netnews groups, and *BABAR* Notes are posted there, as are announcements of special interest.

## 15.8 Construction Responsibilities

---

The design of the *BABAR* detector described herein has been arrived at through an extended process of optimizing the design of individual subsystems and the detector as a whole against the required physics performance, the interests and technical capabilities of the collaborating institutions, and available financial resources.

In cases in which there were competing technologies, the process of arriving at a single subsystem design has been managed by a series of *ad hoc* task forces charged with evaluating alternatives and recommending choices. In making these technology choices, the task forces have considered physics performance, cost, maturity of technology, schedule implications, and the integration of a given system into the overall detector context. Task forces have been employed in arriving at final designs for the vertex detector, particle identification, electromagnetic calorimeter and instrumented flux return systems. The reports of these task forces are available as *BABAR* notes.

Responsibilities for design and construction of the various detector subsystems have been assigned through the traditional process of matching interests, capabilities, and resources. Final responsibilities will be detailed in the Memoranda of Understanding. The following tables provide a brief synopsis of the expressed interest of many of the participating institutions in detector construction responsibilities. In the case of computing, the institutions listed are those that have been involved in the ongoing effort to plan the computing infrastructure. It is expected that responsibilities for the generation of simulation, reconstruction, and analysis code will be distributed more broadly throughout the Collaboration.

It has been agreed by the Collaboration that certain detector items will be financed through a Common Fund. These include, but are not limited to, the magnet and its supporting hardware, certain utilities, online computing hardware, and certain aspects of the software development project. Details are under discussion.

Subsystem	Institutions
Vertex Detector	Italy: Milano, Pavia, Pisa, Torino, Trieste US: UC Santa Barbara, UC Santa Cruz, LBL, Stanford
Main Tracking Chamber	Canada: UBC, Carleton, CRPP, McGill, Montreal, TRIUMF, Victoria, York US: Colorado, Colorado State, MIT, SLAC
DIRC PID	France: Ecole Polytechnique, Orsay, Paris 6/7, Saclay US: UC Santa Barbara, Cincinnati, Colorado State, LBL, Rutgers, SLAC
Aerogel PID	France: LAPP Annecy Italy: Ferrara, Milano, Padova, Roma Russia: BINP US: UCLA, Caltech, Maryland
CsI Calorimeter	China: Beijing, BGRI, SIC Germany: Dresden Russia: BINP UK: Bristol, Brunel, Edinburgh, Lancaster, Liverpool, Manchester, ICSTM, QMW, RHBNC, RAL US: UCIIRPA, UC Irvine, Caltech, Iowa, LLNL, UMass Amherst, Mississippi, Mt. Holyoke, Notre Dame, SLAC
Instrumented Flux Return	China: Beijing Italy: Bari, Frascati, Genova, Napoli US: LLNL, Vanderbilt, Wisconsin

**Table 15-1.** *Institutions interested in construction of BABAR detector systems.*

Subsystem	Institutions
Magnet	China: Beijing Italy: Genova Russia: Dubna, BINP UK: RAL US: LLNL, ORNL/Y12, SLAC
Electronics	Canada: Montreal France: Ecole Polytechnique, Orsay, Paris 6/7 Germany: Dresden Italy: Genova, Napoli, Milano, Pavia, Pisa, Torino Taiwan: Academia Sinica UK: Bristol, Edinburgh, Lancaster, ICSTM, QMW, RHBNC, RAL US: UCIIRPA, UC Irvine, UC Santa Cruz, Caltech, Colorado, Iowa, Iowa State, LBL, Penn, SLAC
Computing	Canada: McGill France: LAPP, Orsay, Saclay Germany: Dresden Italy: Padova UK: Manchester, RAL, US: Caltech, UC Davis, UC Irvine, UC Santa Cruz, LBL, LLNL, Mississippi, Pennsylvania, SLAC, Prairie View

**Table 15-2.** *Institutions interested in construction of BABAR detector systems (continued).*

## References

---

[BAB94] PEP-II Detector Collaboration Governance (1994).

---

---

## Cost and Schedule

---

### 16.1 Introduction

---

Since the *Letter of Intent*, the design of the *BABAR* detector has evolved as a result of a continuous performance analysis and cost optimization. These studies involved both individual systems as well as the detector as a whole. Much scrutiny was given to the requirements for acceptance, resolution, and granularity. Furthermore, the designs were updated following revisions of the fabrication, assembly and installation methods, and reassessments of the need for redundancy and reliability.

In cases where there were competing technologies, *ad-hoc* task forces were formed to evaluate the options taking into account the projected performance with regard to the *CP* measurements, the technical feasibility based on prototype measurements and comparisons with systems of similar design, the integration into the overall detector, and the projected cost and schedule implications. In this way the Collaboration selected the CsI calorimeter, the resistive plate chambers for the IFR, and the DIRC and aerogel for particle identification in the barrel section and forward sections, respectively.

The *BABAR* detector design is based on well established technologies and takes advantage of state-of-the art electronics, data acquisition and computing techniques. Furthermore, the design of each subsystem builds on extensive R&D work that was started, in some instances, well before the Collaboration was founded in December of 1993. The current focus is on detailed design, prototyping and testing; in some cases preproduction activities have been initiated. One of the highest priorities is the building of engineering teams to develop detailed designs and prepare the procurement of long lead-time items.

---

### 16.2 Project Cost

---

A cost estimate has been developed for the detector described in this Technical Design Report. This estimate includes all activities and materials associated with the project, except for R&D and the design, construction and tests of prototypes that are currently

underway and are funded from other sources. The current estimate is mostly based on grounds-up analyses, though in some areas engineering estimates remain to be replaced. The estimates are based on specific designs and where applicable, comparable work performed by the same group. In many cases, in particular for large items, multiple vendor quotes have been obtained and samples have been purchased and tested.

All costs are stated in 1995 US\$. The estimates cover the design, engineering, procurement and fabrication, the assembly, tests and installation. The duration of the project is assumed to be 42 months, beginning in April 1995 and ending in September 1998.

All activities necessary to complete the project have been organized into a detailed Work Breakdown Structure (WBS) which at present contains more than 8000 elements. Each WBS element is identified by a number and a short title. A description field is used for detailed definitions, key parameters, components, as well as information on the basis of estimate. Documentation related to the basis of estimate, vendor quotes, *etc.* will be collected separately in workbooks. The WBS is available in a uniform database format, and it is network accessible. At Level 2, the project is partitioned into nine systems, each headed by a system manager who is responsible for the design, construction and installation, as well as the cost and schedule management.

Because of the international nature of the collaboration, local conventions for cost accounting are being used. This means:

- The engineering and labor are recorded for each task in terms of the time required.
- All materials, fabrication and procurement costs are recorded on the basis of local rates.
- For work performed at US or Canadian institutions, local rates for engineering and labor are applied; the assigned rates are not burdened with operating overhead charges.
- For activities that are supported by non-US/non-Canadian institutions, expenses for engineering and labor are not recorded, they are included in operating budgets of those institutions.
- For items funded from US sources, contingency is added to the base cost to cover additional costs, above and beyond the base, that are necessary to ensure the completion of the task.
- For items funded from non-US sources, no contingency is added, since the base cost includes a certain allowance for uncertainties in the cost projections. Significant cost overruns can be dealt with by additional grants.
- In all cases support from operating or infrastructure funds are recognized and not included in the estimate. Also, work performed by physicists is not charged to the project.

	US Cost in US Accounting (1995 k\$)									
	Description	Funding Agency	EDI&A (mm)	EDI&A (\$)	Labor (mm)	Labor (\$)	M&S	Base Cost	Contingency	Total
1.1	Vertex	US DoE	112	821	114	527	733	2,081	25%	2,599
1.2	Drift Chamber	US DoE	3	24	16	103	255	382	22%	467
1.3.1	DIRC	US DoE	122	1,049	183	897	1,862	3,808	26%	4,799
1.3.2	Aerogel	US DoE	1	7	12	63	198	268	25%	334
1.4	Calorimeter	US DoE	136	1,219	436	1,507	11,257	13,983	16%	16,258
1.5	IFR	US DoE	7	58	63	219	196	473	16%	551
1.7	Electronics									
1.7.1	Vertex	US DoE	59	506	0	0	132	638	27%	813
1.7.2	Drift Chamber	US DoE	84	296	2	6	493	795	32%	1,051
1.7.3	DIRC	US DoE	79	529	9	56	341	926	20%	1,111
1.7.4	Calorimeter									
1.7.5	IFR									
1.7.6	Aerogel	US DoE	6	37	0	2	8	47	32%	62
1.7.7	Trigger Level I	US DoE	44	409	0		154	563	28%	718
1.7.8	Trigger Level II/DAQ	US DoE	60	523	0		187	710	32%	938
1.7.9	Controls	US DoE	5	44	0		47	91	37%	125
1.7.D	System Engineering	US DoE	71	533	0			533	42%	757
1.8	Computing	US DoE	156	1,289	137	446	243	1,978	31%	2,600
1.9	Management/Integration									
1.9.1	Tech. Coordination	US DoE	504	3,097	126	529	1,027	4,653	12%	5,227
1.9.2	QA/ES&H Oversight	US DoE	84	696			10	706	8%	764
SUM			1,533	11,137	1,098	4,355	17,143	32,635	20%	39,174

**Table 16-1.** *Estimated Detector Cost in FY1995 US dollars, for items supported by US Institutions. Not included in these tables are items that might be financed through the Common Fund. These include the magnet and some of the infrastructure, some of the expenses for computing equipment, the electronics common to all systems, and expenses for detector installation and project coordination.*

Following a procedure developed at other HEP laboratories, the contingency was derived for each WBS elements from a weighted sum of four factors that assess the potential risk or uncertainty in the design, engineering solution, cost and schedule. This method introduces a certain objectivity into the assessment, however, it can produce distortions. These have been corrected.

In Tables 16-1 and 16-2, the cost estimate is summarized at Level 2/3 of the WBS. The data are “rolled-up” from much more detailed information. Table 16-1 refers to the cost of items that are to be funded by US institutions, thus EDI&A and labor costs are fully accounted. Table 16-2 refers to items that are to be funded by non-US institutions. Not included in either table are items that could potentially be financed through a Common Fund. While there is consensus that such a fund should be established, the size of the fund,

Non-US Cost in Local Accounting (1995 k\$)										
	Description	Funding Agency	EDI&A (mm)	EDI&A (\$)	Labor (mm)	Labor (\$)	M&S	Base Cost	Contingency	Total
1.1	Vertex	INFN Italy	60		162		2,864	2,864	0%	2,864
1.2	Drift Chamber	NSERC Canada	95	316	149	602	888	1,806	0%	1,806
1.3.1	DIRC	France	111		106		2,641	2,641	0%	2,641
1.3.2a	Aerogel	France	12		12		173	173	0%	173
1.3.2b		INFN Italy	0		4		146	146	0%	146
1.4a	Calorimeter	BMFT Germany	0		89	286	2,768	3,054	0%	3,054
1.4b		PPARC U.K.	72		151		2,268	2,268	0%	2,268
1.5	IFR	INFN Italy	8		21		1,060	1,060	0%	1,060
1.7	Electronics									
1.7.1	Vertex	INFN Italy	0		0		120	120	0%	120
1.7.2	Drift Chamber	NSERC Canada	55	364	0		1,945	2,309	0%	2,309
1.7.3	DIRC	France	84				1,734	1,734	0%	1,734
1.7.4	Calorimeter	PPARC UK	185		18		1700	1,700	0%	1,700
1.7.5	IFR	INFN Italy	19		29		409	409	0%	409
1.7.6a	Aerogel	INFN Italy	6		0		126	126	0%	126
1.7.6b		France	2		0		33	33	0%	33
1.7.7	Trigger Level I	PPARC U.K.	23		0		78	78	0%	78
1.8	Computing	France	53		0		0	0	0%	0
SUM			785	680	741	888	18,953	20,521	0%	20,521

**Table 16-2.** *Estimated Detector Cost in FY1995 US dollars, for items supported by non-US Institutions. Not included in these tables are items that might be financed through the Common Fund.*

the obligations of various national funding agencies to this fund, and the list of items that should be supported from this fund remain to be negotiated.

The total cost of materials and services (in Europe referred to as investment costs) to be financed from the US and non-US resources, including the Common Fund, amounts to 45 million US\$.

## 16.3 Schedule

Having a detector ready to confront  $CP$  violation physics of PEP-II on a time scale commensurate with the PEP II completion date of October 1998 is a formidable task.

Detailed schedule information will be incorporated into the Integrated Project Schedule. The summary level schedule presented in Table 16-3 serves mainly to identify interactions

between systems and critical path items. This schedule is technology-limited and allows installation and commissioning of *BABAR* at a date compatible with the PEP-II schedule.

## 16.4 Detector Systems

---

A brief discussion of the cost and schedule projections for the individual detector systems follows.

### 16.4.1 Vertex Detector

The five-layer double-sided silicon strip vertex detector chosen for *BABAR* is an extrapolation of existing designs. The emphasis is on precision in the  $z$  direction, basic to the main physics thrust of the experiment. Reduction of material to decrease multiple scattering is a primary design requirement. The performance of  $300\ \mu\text{m}$  double-sided silicon strip detectors has been satisfactorily tested. For the read-out electronics, radiation-hard fabrication techniques developed for SSC and LHC are more than sufficient for the PEP-II environment.

Members of the vertex detector group have been active in the design and construction of similar devices for CLEO-II, SDC, ALEPH, D0, CDF and Mark II. There is substantial expertise within the group in detector and readout chip design, in trigger design and in mechanical structures. The cost estimate is based on this recent experience and actual vendor quotes. Full size prototype detectors are on order. The design and prototyping of the VLSI read-out chip is underway, it is on the critical path for the assembly and tests of the prototype detector modules and the beam tests scheduled for later this year. The assembly of the detector modules is a delicate, time consuming task that will be performed in parallel in the US and Italy.

### 16.4.2 Tracking Chamber

Several technical innovations in the main tracking chamber are motivated by the requirement of minimal multiple Coulomb scattering. Low-mass gases have been developed within the group over the past several years. Several chambers of comparable design and dimensions have been constructed for HEP experiments. Low-mass field wires can be made of aluminum, although thin gold plating of the Al wires remains to be developed.

Construction of the main tracking chamber on the needed timescale requires the early resolution of mechanical engineering issues involving endplate and feedthrough design, and choice

of field wire material. The electronics and data acquisition are reasonably straightforward. It is planned to use a full-length prototype test chamber as a test bench for chamber operation and monitoring, electronics and data acquisition.

The stringing of the wires is the most time consuming task and it is judged to be a critical path item. The task will be performed at a laboratory where chambers of similar dimensions have been assembled recently.

### 16.4.3 Particle Identification

While the DIRC is a novel concept for a Cherenkov detector, its construction is conceptually simple. The production and assembly of the quartz Cherenkov radiators present a relatively well-characterized problem that should have no schedule impact. Tests with small samples of quartz bars have produced excellent results. A test with a full length prototype detector is in preparation.

The readout system uses large numbers of conventional photomultiplier tubes; system engineering is thus the principal concern. A first-article procurement of 500 phototubes is the basis of the cost estimate, a firm offer for the whole photomultiplier purchase is in hand; it has a significant savings for large quantities. Concerns about the detailed design of the water tank and the modifications to the flux return have been addressed. Access to inner detector components has become easier with the elimination of the backward calorimeter endcap.

The aerogel threshold counter extends the coverage for particle identification in the forward direction to 300 mr. This system is by far the smallest in *BABAR*, but requires sizable support for R&D, engineering and photon detectors that can operate in the 1.5 T magnetic field. The present choice, fine-mesh photomultipliers, will become commercially available in the near future. The cost estimate is based on a preliminary vendor's quote. Several other experiments are in need of such tubes, and it is expected that the price will drop over the next few months.

In the past, the production of high quality aerogel has been a serious concern. Facilities exist for the production of the required quality and modest quantities. Prototype pieces have been tested in a beam and have shown very satisfactory light yield and absorption length.

The mechanical mounting structure, and the calibration and readout systems are relatively simple and of limited scale, requiring a modest engineering effort.

#### 16.4.4 CsI(Tl) Calorimeter

Large CsI calorimeters of similar design and comparable volume have been built in the past and have proven to work well at  $e^+e^-$  storage rings. Nevertheless, the production of 6,780 crystals of high quality is the principal challenge in this system, it is also the principal cost driver. The procurement of the crystals is clearly on the critical path, and there remain concerns about the projected rate of delivery of the crystals.

The *BABAR* calorimeter is divided into a barrel and a forward endcap section, with a total volume of 6 m<sup>3</sup>. The procurement plan assumes that a production rate of 250 crystals per month can be established by the end of 1996 and can be sustained for more than a year. The crystal procurement relies on three vendors in three different countries, each with prior experience in growing non-alkaline halide crystals and with interest in increasing the capacity of their production facilities. Extensive contacts with all three crystal manufacturers have been established and it is expected that new, more advanced growing techniques can be applied (similar to those being developed at SLAC) which will produce excellent quality at significantly lower prices than in the past. The choice of a single supplier for the raw material will assure uniform quality and will bring the benefit of volume and delivery discounts. It also allows for a beneficial arrangement for the recycling of crystal material.

The engineering and fabrication of the low mass structure that will support the crystals is the second largest item. The proposed concept is rather simple. The design has been tested using finite element analysis and involves the use of Kevlar-epoxy composites, while allowing fabrication by low-technology and relatively inexpensive methods. The present endcap design is based on more conventional support structure, but an alternate design, similar to the barrel is under study.

Based on experience in other experiments the readout system based on photodiodes and wavelength-shifters can be expected to be reliable and perform as designed. Prototypes for the silicon PIN photodiodes are presently being tested. The price for the 13,560 photodiodes is being negotiated. Since the production capacity is more than adequate, the delivery schedule is not expected to be a problem.

The modular design of mechanical support structure is simple and inexpensive and it provides great flexibility in scheduling the assembly and installation. The schedule allows for about one year of final assembly and tests. The production and assembly of the wavelength shifter/photodiode/preamplifier readout will be widely distributed and will be performed in parallel to the crystal production.

### 16.4.5 Flux Return Instrumentation

The construction of large area chambers for muon detection is a task of a scale that has been achieved by nearly every major  $4\pi$  detector built in the last decade. The design and operation of large numbers of resistive plate chambers (RPC) is well established. Facilities with adequate capacity for the fabrication of the RPCs exist, as do experienced teams of scientists at collaborating institutions where these chambers will be tested.

The schedule foresees the installation of the chambers prior to the assembly of the steel structure, and thus places constraints on the schedule for the chamber fabrication, tests, and final assembly at SLAC.

### 16.4.6 Magnet

The specified superconducting solenoid is similar in dimensions and specifications to several existing magnets. The preliminary design is being prepared by a collaboration of scientists and experienced engineers in the UK and Italy and will be completed in April 1995. It forms the basis for the cost estimate and will also be used to initiate the procurement process.

The segmented flux return is also similar in design to other detectors. It must be completely specified in the next few months, so procurement can be initiated in FY1995. The superconducting solenoid and the steel structure forming the flux return are both on the critical path since they must be completely assembled and the magnetic field must be measured prior to the installation of all other detector systems.

### 16.4.7 Electronics

Great emphasis has been placed on the adoption of a common architecture and the use of common solutions wherever possible. This approach is not only attractive from a design and operational point of view, it is also most cost effective, since it minimizes duplication of effort and allows for economies of scale.

The electronics system described in the *TDR* serves as a model based on present day technology and experience. The WBS divides the electronics system into three principal segments:

- the detector specific electronics, *i.e.*, front-end electronics,

- the trigger and data acquisition combined with electronics that are shared among detector systems, and
- the system engineering and the infrastructure.

For most detector systems, the design and prototyping is a very significant expense. The costs for the front-end electronics for the drift chamber, calorimeter and DIRC are largely driven by the channel count. For the drift chamber and calorimeter, the requirement for analog information and the extraction of trigger information adds significant costs. The fabrication costs include expenses for QC and tests. The front-end electronics for the vertex detector are very closely integrated with the detector design and assembly and they are therefore included in the detector costs.

The scale of the trigger and data acquisition system is conventional. Technical innovations relate mostly to the asynchronous nature of the system and its fully pipelined architecture. Trigger, data, acquisition and other electronics contribute about 20% to the total cost, shared about equally between engineering and M&S.

The schedule for the detector mounted electronics is given by the assembly and installation schedule for a given detector system, *i.e.*, the vertex detector, the drift chamber, PID, calorimeter, and the IFR. This means that there are 12 to 18 months for design and prototyping, and another 12 to 18 months for fabrication and testing. For all these systems it is crucial that prototype front-end circuits be available for the beam tests later this year.

The bulk of the off-detector electronics will not be needed until the assembly and installation is completed in early 1998, thus allowing a total of three years for design, prototyping and fabrication. However, it is planned to have a production model of the data acquisition system available for testing more than a year before the completion date, so that design problems that might not be apparent on a single circuit test can be detected and corrected prior to the large production. This would allow for complete system tests, *i.e.*, from a prototype detector to the DAQ for a substantial number of channels.

### 16.4.8 Computing

The quantity of data that will need to be transferred from the *BABAR* detector to mass storage and then analyzed pose significant challenges for the computing and network systems. For costing purposes the system is divided into two major segments:

- the online infrastructure, consisting of Level 3 data acquisition, control and monitoring hardware, as well as the networking, and

- the software development, including both online-specific items such as VxWorks, EPICS, expert system tools; and general items like IDL, methodology software, OODBMS, GUI builder, documentation and development aids.

Not costed are items that are conventionally part of laboratory or university infrastructure (*e.g.*, the large clusters of workstations that perform the offline analysis and simulation tasks), and activities that are generally supported from operating funds (*e.g.*, maintenance of hardware and software, software engineering that is available on a continuing basis throughout construction and operations, and data-aide activities). Also not included in the computing system cost estimate are computers for administration and workstations for engineering and design (covered in WBS 1.9), and workstations for software development (covered in WBS 1.7), as well as embedded CPUs in the DAQ/trigger systems (covered in WBS 1.7).

The largest cost items for the computing system are the required software engineering and the hardware for the Level 3 compute farm. The engineering effort is derived from three sources:

- professional software engineers, at a level of approximately three to four FTEs for the duration of the construction (not including the project computing engineer, covered in WBS 1.9);
- undergraduate students, who are frequently recruited from computer science departments. Past experience has proven that many of the programming tasks (such as development of specific computing utilities) can be performed by this kind of low-cost part-time student labor; and
- physicists, who will form a large part of the team for developing application code, but are not included in the construction cost.

For the purpose of this estimate, it assumed that the Level 3 farm will consist of a networked system of UNIX workstations. The principle data flow will be via FDDI connections, with a DEC Gigaswitch to provide routing.

Future cost savings that can be expected from advances in technology over the next few years are presently being examined. While there is incentive to delay the acquisition of the computing hardware in order to take advantage of rapid developments and lower costs, a large portion of the computing system needs to be in operation well before the detector is ready to take data. Prototype framework, data model, and code management capabilities need to be in place as soon as possible, because so many other development activities depend on them. Hence, these items represent relatively early expenditures in the *BABAR* project.

The software for detector simulation, data reconstruction and analysis will be developed and used by a large number of members of the collaboration, involving very few computer

specialists. Because of the geographic diversity of the collaboration, coordination is a complex task and organizational and technical systems are being developed to facilitate communication and remote operation.

### 16.4.9 Management, Installation and Integration

This WBS item includes the support of the *BABAR* management and detector integration, the QA and ES&H management as well as the detector installation.

Substantial emphasis is being placed on system integration from the start of the project. This involves coordination of the service and utility requirements of the various subsystems, as well as detailed planning for the assembly of the detector and its commissioning. This effort will be coordinated by an experienced staff of engineers and physicists.

The overall plan, the cost estimate and schedule projections are based on recent experience with SLD at SLAC. In particular, the estimate for the installation relies on the availability of a sizable crew of laborers and electricians to prepare the hall, assemble the electronics building, install utilities, and install the detector components. Special jigs and tooling for the installation of the individual detector systems are considered the responsibility of that system. The same is true for any detector specific utilities like cables, cooling and gas supplies.

The development of the detailed installation procedure has resulted in a top down schedule, determining the date at which each of the detector systems will have to be delivered to the experimental hall, together with any special tooling and jigs required for installation.

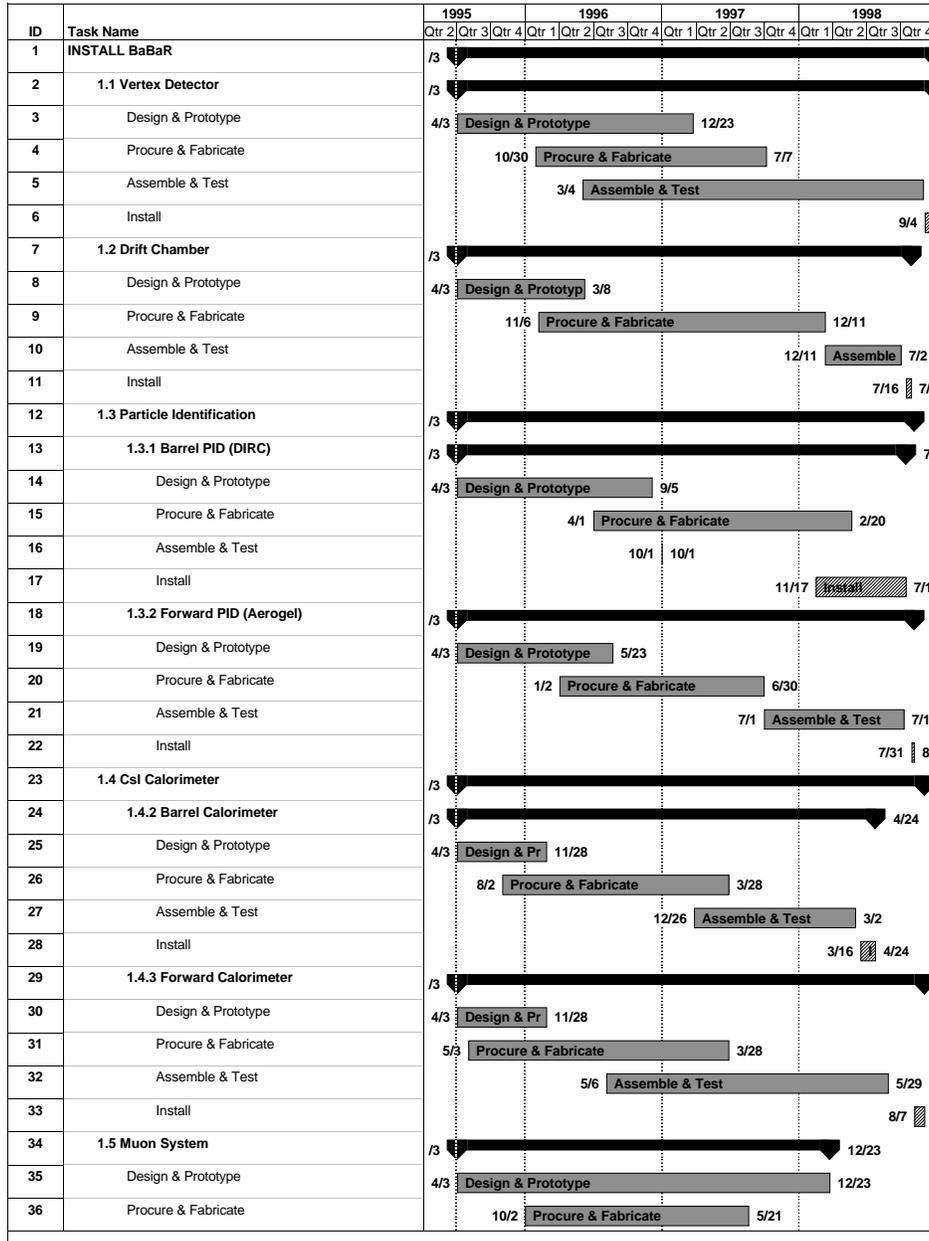


Table 16-3. Projected project schedule, aimed at completion in October 1998, part 1.

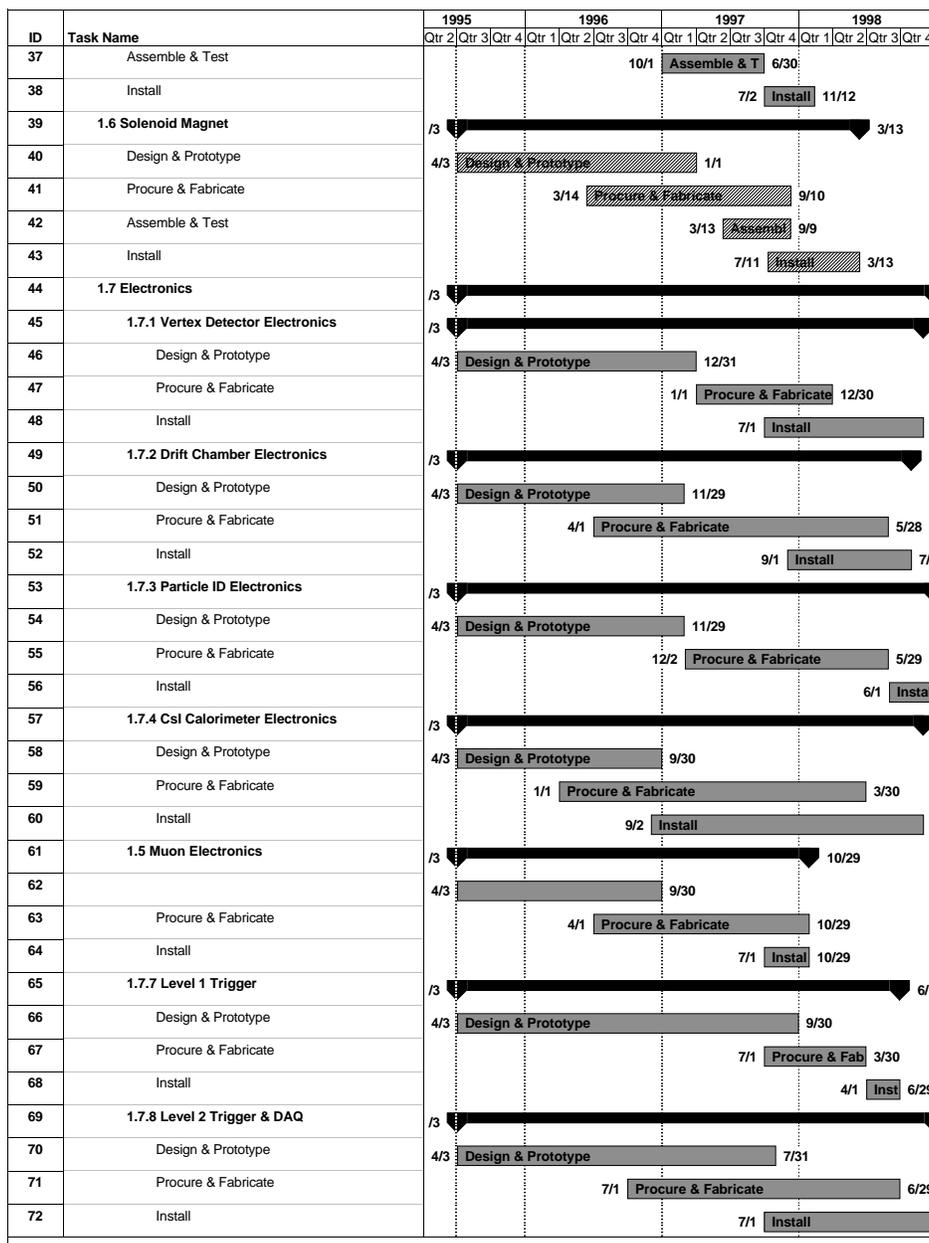


Table 16-4. Projected project schedule, part 2.

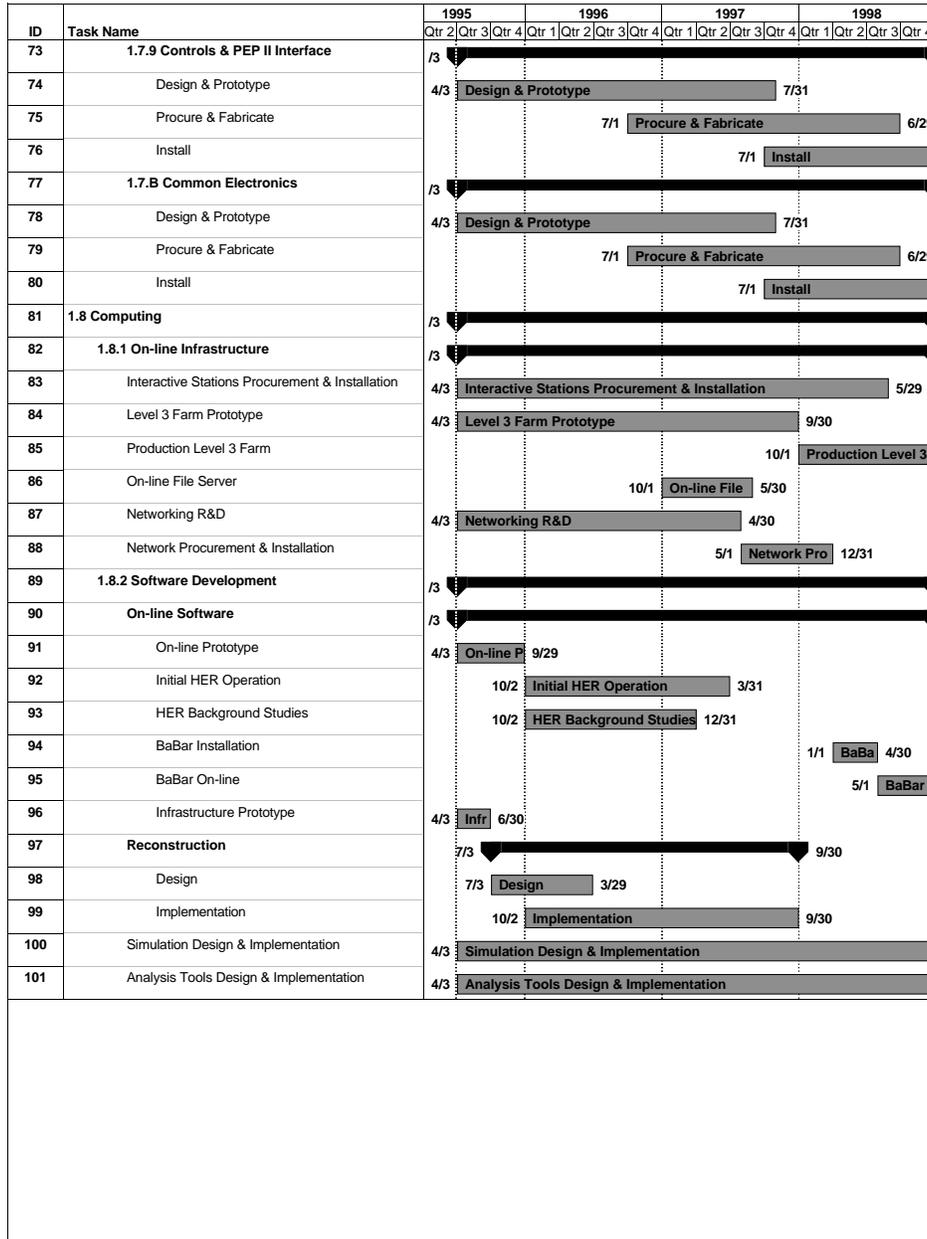


Table 16-5. Projected project schedule, part 3.