Preliminary Studies Concerning $\Delta \Gamma_{B_s}$ Measurements in Proton Antiproton Collisions at $\sqrt{s} = 2.0$ TeV

Matthew S. Martin, Particle Physics, Jesus College Oxford

Michaelmas Term 2001
Acknowledgements

First and foremost I acknowledge the LORD my God. Without you I would be lost. Secondly I want to thank my family, near and far, who have borne with my lunatic ways. Thankyou especially Elisabeth. You will be my wife in ten short days. How much have you have had to put up with already! Next I owe a huge debt of gratitude to Todd my supervisor. Not only have you provided invaluable scientific insight and direction, but you have patiently borne my tiresome, and at times down right rude antics. It is a strange phenomena that you are almost invariably right! Thankyou Dad for loving me enough to point out where you thought I was going astray. Thankyou Mum for tirelessly encouraging, and sending Pizza into the Oxford Physics Dept! Thankyou Phil for being there when I needed you. Thankyou Randy for being so patient with me. I could not wish for a more understanding friend and guide. Thankyou Will for bearing with the failings of a demanding room-mate. I could always count on you. Thankyou Jeff for keeping me up to the mark. Your standards of excellence always inspire me to do the job properly. Thankyou Glenn and Rich for your tireless patience with my rather unreasonable computing demands! Thankyou Farukh for mucking in and helping when I needed you to. Thankyou Louis for always being interested in what I was doing. Thankyou Jonas for your uncompromising, but invaluable criticism. Thankyou Dave for so many little pieces of advice and help. You always managed to find time. Thankyou Armin for being so generous with your time near the start. I was very green! Thankyou Giulia, Tracey and Thilo for all your support. I’m sure I was unbearable at times. Thankyou Paul, Ankush and Dan for an office that not only buzzed with ideas, but that didn’t take itself too seriously. Thankyou Izi for remaining sane whilst having to share the caverness luxury that is the CDF trailers. Thankyou Bill, Steve, Helen, Joe and Dillard for putting up with a young whipper-snapper who is quite new to this teaching lark!
ABSTRACT

Studies pertaining to a measurement of $\Delta \Gamma_{B_s}$ at CDF Run II are presented. The emphasis of the work is on the systematic effects of the impact parameter cut in the level 2 trigger on the observed $B_s$ lifetime distribution. The mode chosen to investigate this effect is:

$$B_s \rightarrow D_s^- \pi^+ \text{ where } D_s^- \rightarrow \phi \pi^- \text{ and } \phi \rightarrow K^+ K^-$$

A monte-carlo study of the $B_s$ lifetime from a subset of the two track multi-body trigger is made with comparable statistics to run IIa. The precision is:

$$\delta \lambda = (+0.062)(-0.058) \text{ (stat)} \, ps \pm 0.023 \, ps \text{ (syst1)} + 0.059 \, ps \text{ (syst2)}$$

where $\lambda$ is the average $B_s$ lifetime.
# Contents

1 **Theoretical Introduction** ................................................. 1
   1.1 The Standard Model .................................................. 1
      1.1.1 Quark Mixing .................................................. 2
   1.1.2 The CKM Matrix .................................................. 4
   1.2 $\Delta \Gamma_{B_s}$ Phenomenology ..................................... 5
      1.2.1 General Formalism ............................................. 5
      1.2.2 Operator Product Expansion .................................. 9
      1.2.3 Towards a low energy effective theory ..................... 10
      1.2.4 Wilson Coefficients ......................................... 11
      1.2.5 Calculating the Wilson Coefficients ....................... 13
      1.2.6 The renormalisation group ................................... 13
      1.2.7 Summary ....................................................... 14
      1.2.8 Long Range Hadronic effects ................................ 15
      1.2.9 Bag Parameters and the Breakdown of the Vacuum
            Saturation Approximation ........................................ 16
      1.2.10 Pulling it all together ..................................... 16
      1.2.11 $\Gamma_{12}$ .................................................... 18
   1.3 Current Results ...................................................... 19
   1.4 Epilogue ............................................................. 20

2 **The CDF Detector** ....................................................... 21
   2.1 Introduction .......................................................... 21
   2.2 Tevatron ............................................................. 21
   2.3 The CDF Run II Detector ............................................ 22
      2.3.1 Tracking Systems ............................................. 22
      2.3.2 COT ............................................................. 26
      2.3.3 SVX II ........................................................ 26
      2.3.4 ISL ............................................................. 29
3 Experimental Techniques for measuring $\Delta \Gamma_{B_s}$

3.1 Sculpting the Observed $B_s$ lifetime distribution .......... 46
3.2 Two different ways to measure $\Delta \Gamma_{B_s}$ ............... 48
3.3 The Two Sample Method .................................... 50
   3.3.1 $B_s \rightarrow D_s^\pi$ and $B_s \rightarrow D_s^\pi \pi \pi$ .... 51
   3.3.2 $B_s \rightarrow J/\psi \phi$ .................................. 53
   3.3.3 $B_s \rightarrow D_s^+D_s^-$ ................................. 55
3.4 The specifics of a lifetime measurement ....................... 55
   3.4.1 Candidate selection .................................... 56
   3.4.2 Lifetime reconstruction ................................. 62
   3.4.3 Likelihood fit to an exponential ....................... 64
3.5 Other potentially useful modes ............................... 65
   3.5.1 Mixed CP Modes ....................................... 65
   3.5.2 CP Separable Channels ................................. 65

4 Sample Purity .................................................. 67
   4.0.3 Strategy ................................................ 68
   4.0.4 Cuts ................................................... 71
   4.0.5 Background Discrimination ............................. 71
   4.0.6 Effect of Correlations ................................. 72
   4.0.7 Optimal Cut values .................................... 72
4.1 Signal to Background Calculation .............................. 74
   4.1.1 Limitations of the prediction .......................... 74
   4.1.2 The Calculation ........................................ 75
   4.1.3 $B_s$ invariant mass cut factor ......................... 81
   4.1.4 Extra cut factor ...................................... 85
   4.1.5 Main cut factors for signal and background .......... 86
   4.1.6 Signal normalisation : B ............................. 87
   4.1.7 Final value and error .................................. 88
5 Level 2 Trigger Efficiency Studies .......................... 90
  5.1 Level 2 Trigger Simulation ................................. 91
      5.1.1 General Strategy ................................... 91
      5.1.2 Generator Level Trigger Simulation: GenTrig .......... 91
      5.1.3 Generation p_t Cut .................................. 92
  5.2 Aims of the Study ......................................... 95
      5.2.1 Trigger Configurations ............................... 97
      5.2.2 Pure Signal Trigger Configurations .................. 97
      5.2.3 Pure Background Trigger Configurations .............. 98
      5.2.4 Mixed Trigger Configurations ......................... 99
  5.3 Global Results ........................................... 99
      5.3.1 Global Efficiency ..................................... 99
      5.3.2 Efficiency as a function of B_s proper time .......... 100
  5.4 Different Configuration Efficiencies ...................... 105
      5.4.1 Pure Signal Trigger Configurations ................... 105
      5.4.2 Mixed Signal Trigger Configurations .................. 109
      5.4.3 Pure background configurations ....................... 109
  5.5 L2 Efficiency Summary .................................... 110
      5.5.1 (f_B,f_B) Sub-sample ................................ 112
      5.5.2 (sig,sig) Sub-sample ................................ 112
  5.6 Parameterising the Efficiency Curve ....................... 113
      5.6.1 The dominance of π_b ................................ 113
      5.6.2 The effect of a d_0 cut on the \tau_p proper time probability distribution ......................... 116
      5.6.3 \nu Probability Distribution .......................... 119
      5.6.4 Proper time Efficiency functional form ............... 119

6 Likelihood Fit to the B_s lifetime ........................ 123
  6.1 Unbinned Log Likelihood Hypothesis Testing ............... 123
      6.1.1 The Probability Function ............................ 124
      6.1.2 Parameter Fitting .................................... 125
      6.1.3 Minimising the Negative Log Likelihood ............... 125
      6.1.4 Testing the Goodness of Fit .......................... 125
  6.2 The Proper Time Likelihood Function ....................... 126
      6.2.1 Signal Proper Time Probability Function .............. 127
      6.2.2 \sigma(t) dependence of the efficiency curve ........... 128
      6.2.3 Background Proper Time Probability Function .......... 134
      6.2.4 Signal Mass Probability Function ..................... 138
6.2.5 Background Mass Probability Function .......................... 138
6.3 Results ........................................................................ 138
6.4 Goodness of Fit .......................................................... 140

7 Systematic Uncertainties ................................................. 143
  7.1 Uncertainty in the Efficiency Function ............................ 143
  7.2 Effect of the cut on Proper Time Uncertainty .................. 145
  7.3 Fitted Range of Proper Time ........................................ 146
  7.4 Effect of the Size of \( \sigma(t) \) bins ................................. 146
  7.5 Bias due to assuming a lifetime central value .................. 148
  7.6 \( B_s \) meson \( p_t \) spectrum .......................................... 148
  7.7 Summary .................................................................. 149

8 Conclusions ..................................................................... 151
  8.1 Discussion of Results .................................................. 151
  8.2 Epilogue ..................................................................... 153

A Glossary of Definitions .................................................. 154
  A.1 CDF Coordinate Definition .......................................... 154
  A.2 \( L_{xy} \) and \( d_0 \) ........................................................ 156
  A.3 \( \eta \) definition ............................................................ 156

B Theoretical Appendix ...................................................... 158
  B.1 Connecting the Weak and Strong Eigenvalues ................. 158
    B.1.1 Starting Point ...................................................... 158
    B.1.2 Prescription ......................................................... 159
    B.1.3 Working it through ............................................... 159

C \( p_t \) pull discussion ..................................................... 164
# List of Figures

1.1 A box diagram for $B_s$ mixing ........................................ 7
1.2 A one loop correction to $c \to s u d$ .................................. 12
1.3 The lowest order contribution to $\Gamma_{12}$ .......................... 19

2.1 An Elevation view of the CDF detector ............................... 23
2.2 The tracking systems geometry ........................................ 24
2.3 A COT endplate ......................................................... 27
2.4 A Run2 data $J/\Psi$ mass peak ...................................... 28
2.5 An end view of the SVXII detector .................................. 30
2.6 An end view of a VRB .................................................. 30
2.7 A view of 2 HDI’s, or 1 ladder ....................................... 30
2.8 A close up end view of ISL forward region ........................... 32
2.9 A schematic view of one half of the ISL .............................. 32
2.10 An end view of Layer 00 .............................................. 33
2.11 The effect of Layer 00 .................................................. 34
2.12 A Run2 data SVT impact parameter resolution ..................... 37
2.13 A forward decay cut example ........................................ 41
2.14 Another forward decay cut example ................................ 42

3.1 The $B_s$ proper time before Level 1 .................................. 47
3.2 The $B_s$ proper time after Level 2 ................................... 48
3.3 The relative orbital angular momentum in $B_s \to J/\psi \phi$ ........ 54
3.4 A graphical description of a pointing constraint .................... 59
3.5 A sample Vertex Probability distribution (signal) ................ 60
3.6 A sample Vertex Probability distribution (generic B) ............ 60
3.7 The helicity angle definition ......................................... 61
3.8 The helicity angular distribution .................................... 62
3.9 Why the $p_t$ unit vector is in the lifetime definition .......... 63
### LIST OF FIGURES

4.1 A Feynman diagram for the decay: $B_s \rightarrow D_s^- \pi^+$ .......................... 68
4.2 A Feynman diagram for the decay: $D_s^- \rightarrow \phi \pi^-$ .......................... 68
4.3 A Feynman diagram for the decay: $D_s^- \rightarrow \phi \pi^-$ .......................... 69
4.4 The $p_t$ separation power ................................................. 73
4.5 The $\phi$ mass separation power ........................................ 74
4.6 The $D_s$ mass separation power ........................................ 75
4.7 The helicity separation power ........................................... 76
4.8 The kaon $p_t$ correlation ............................................... 77
4.9 The $B_s$ mass with $p_t$ cuts ........................................... 82
4.10 The $B_s$ mass with $p_t$ and $\phi$ mass cuts ....................... 82

5.1 A $p_t$ pull distribution ............................................. 93
5.2 The pythia b-Quark $p_t$ distribution ................................. 94
5.3 The level 2 efficiency vs b-quark $p_t$ ............................... 96
5.4 The pythia $B_s$ proper time distribution .......................... 101
5.5 The $B_s$ proper time distribution after Level 2 ................ 102
5.6 The level 2 Efficiency vs $B_s$ meson proper time ............. 103
5.7 The level 2 Efficiency with no upper impact parameter cut ... 104
5.8 The (signal,signal) contributions to level 2 trigger rate ....... 107
5.9 The angular correlation of daughter products ..................... 108
5.10 The (background,background) contributions to level 2 ....... 111
5.11 The $B_s$ pion $d_0$ ..................................................... 114
5.12 The $D_s$ pion $d_0$ ..................................................... 115
5.13 The effect of impact parameter cut on proper time .......... 117
5.14 The $v$ distribution from Generator level ....................... 120
5.15 The $v$ distribution, with a $p_t$ cut ............................... 120
5.16 The $v$ distribution with a fit ..................................... 121
5.17 A fit to the efficiency function ................................. 122

6.1 A likelihood gaussian hypothesis ................................. 124
6.2 The effect of different proper time uncertainty distributions 129
6.3 The proper time uncertainty distributions ......................... 130
6.4 The proper time uncertainty distribution ratios .................. 131
6.5 A linear plot of proper time uncertainty ratio .................... 132
6.6 The $t$ distributions for first two $\sigma(t)$ bins ................ 135
6.7 The $t$ distribution for last $\sigma(t)$ bin. Also all 3 bins. ....... 136
6.8 The efficiency function fit pulls .................................. 137
6.9 The $B_s$ mass distribution ..................................... 139
<table>
<thead>
<tr>
<th>Number</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.10</td>
<td>The shape of the likelihood minimum</td>
<td>141</td>
</tr>
<tr>
<td>6.11</td>
<td>A toy monte-carlo likelihood function values</td>
<td>142</td>
</tr>
<tr>
<td>7.1</td>
<td>The lifetime spread due to efficiency function uncertainty</td>
<td>144</td>
</tr>
<tr>
<td>7.2</td>
<td>The lifetime central values for different fit ranges</td>
<td>147</td>
</tr>
<tr>
<td>8.1</td>
<td>The fits to the efficiency function for the 3 $\sigma(t)$ bins</td>
<td>153</td>
</tr>
<tr>
<td>A.1</td>
<td>Beam axis and definition of $\theta$</td>
<td>154</td>
</tr>
<tr>
<td>A.2</td>
<td>Definition of $\phi_0$ (high $p_t$ track)</td>
<td>155</td>
</tr>
<tr>
<td>A.3</td>
<td>Diagram defining $L_{xy}$ and $d_0$</td>
<td>157</td>
</tr>
</tbody>
</table>
# List of Tables

1.1 Current Lattice Parameter Calculations .......................... 18
2.1 A comparison of some projected modes of running ............... 22
2.2 The coarse Geometrical parameters of Tracking Subsystems ....... 25
2.3 The SVX II Strip Parameters ..................................... 29
2.4 The physical ISL Sensor Parameters ............................... 31
2.5 The average B-Physics lifetimes .................................. 38
2.6 The level 1 Two Track Trigger definition ........................... 39
2.7 The level 2 Two Track trigger cuts ................................. 43
2.8 The $B^0 \rightarrow \pi \pi$ Two Track Trigger ............................... 44
2.9 The $B_s$ multi - body Two Track Trigger ............................ 44
2.10 The current Level 3 Operating parameters .......................... 45
3.1 The sample sizes for mixed CP $B_s$ decay channels ................ 52
3.2 The expected sample sizes for $B_s \rightarrow D_s^+ D_s^-$ ................. 55
3.3 The basic kinematic cuts for $B_s \rightarrow D_s \pi$ reconstruction ........... 57
4.1 The optimum cut values .............................................. 78
4.2 The effect of signal cuts ............................................. 79
4.3 The $B_s$ mass function parameters .................................. 83
4.4 The $B_s$ mass rejection factor ...................................... 84
4.5 The $B_s$ mass rejection factor with revised uncertainty ............. 85
4.6 The extra cuts rejection factor ..................................... 86
4.7 The main background rejection factor ................................ 86
4.8 The main signal rejection factor .................................... 87
4.9 The $B_s$ production fraction and mode branching ratios ........... 87
4.10 The signal to Background Prediction with error ....................... 88
4.11 The signal to background with no proper time uncertainty cut ...... 89
\textit{LIST OF TABLES} \\

5.1 The (signal,signal) contribution to level 2 \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 106
5.2 The (signal,$f_B$) contributions to level 2 \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 109
5.3 The (bgd,bgd) contributions to level 2 \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 110
6.1 The signal to background 1:2 Likelihood fit results. \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 140
7.1 A summary of the studied systematic effects. \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 149
Chapter 1

Theoretical Introduction

In this chapter the present studies for measurements of the $B_s$ width difference ($\Delta \Gamma_{B_s}$) are theoretically motivated. The first half of the chapter constitutes an overview of the underlying Standard Model (SM) theory. The second half concentrates on the calculation of $\Delta \Gamma_{B_s}$.

The chapter begins with a description of the SM itself. The place of quarks within the theory is then discussed, concentrating on the freedom for the different generations to mix.

Then in the second half of the chapter, the theoretical techniques necessary for a calculation of the width difference are surveyed.

1.1 The Standard Model

The SM is the best theory that exists for describing three out of the four fundamental interactions present in the universe. In a remarkably elegant fashion, it unifies the Electromagnetic and Weak forces into the Electroweak force. It provides a description of the Strong force, and crucially for high energy experimentalists it allows precision calculations to be made of a wide variety of physical observables. The only force which is not described is Gravity.

To date, the only observable field predicted by the SM which has not been directly observed is the scalar Higgs field. This is a critical element of the theory, which would not survive in its absence.

While the theory has immense calculational power, it has several undesirable aspects, including the 19 (25 with $m_\nu \neq 0$) free parameters and the
fine tuning of the gauge boson masses[1]. Some of the main features of the SM are now discussed.

The SM Lagrangian (see below) is invariant under transformations of the gauge group:

\[ SU(3) \times SU(2) \times U(1) \]  \hspace{1cm} (1.1)

The three components of this overall group are associated (albeit in an entangled fashion) with the three forces in the following way:

\[ SU(3) \text{ assigned: } SU(3)_C \]  \hspace{1cm} (1.2)

where \( C \) is a three fold degree of freedom color. The gluons and quarks which carry color mediate and couple to the strong interaction respectively.

\[ SU(2) \text{ assigned: } SU(2)_L \]  \hspace{1cm} (1.3)

where \( L \) stands for Left, and indicates that, unmixed with \( U(1)_Y \), \( SU(2)_L \) only couples to left handed fermion doublets. All the left handed fermions (quarks and leptons) couple to this interaction. The three fold degree of freedom, Weak-Isospin is carried by the left handed fermions and the unphysical triplet of vector bosons \( (W^-_1, W^0, W^+_1) \) which mediate \( SU(2)_L \)

\[ U(1) \text{ assigned: } U(1)_Y \]  \hspace{1cm} (1.4)

where \( Y \) stands for Hypercharge and is the degree of freedom carried by all the fermions except \( \nu_R \) which hasn’t been observed and since it is not expected to couple to anything is omitted from the theory. \( SU(2)_L \) and \( U(1)_Y \) mix to give rise to the physically observed weak and electromagnetic interactions.

### 1.1.1 Quark Mixing

Several terms in the SM Lagrangian will now be examined to indicate how the CKM matrix arises. This will then lead to a discussion of the width difference in the \( B_s \) meson system.

The Higgs field gives mass to the fermions and bosons in the standard model. The \( u \) and \( d \) quark mass terms may be written:

\[ \mathcal{L}_S^{\text{fer}} = -\frac{g}{\sqrt{2}M} \bar{\psi}_L m_u K u_R + \frac{g}{\sqrt{2}M} \bar{\psi}_L m_d K^C d_R + h.c. \]  \hspace{1cm} (1.5)
CHAPTER 1. THEORETICAL INTRODUCTION

where $m_u$ and $m_d$ are the masses of up type and down type fermions respectively, $M$ is the bare W-boson mass, and $K$ is the scalar higgs doublet. Confining ourselves simply to the quarks, this expression can be generalised:

$$
\mathcal{L}_{S}^{\text{fer}} = -\frac{g}{\sqrt{2}M} (\bar{\psi}_L)_{\alpha} (M^U)_{\alpha \beta} K (u_R)_{\beta} + \frac{g}{\sqrt{2}M} (\bar{\psi}_L)_{\alpha} (M^D)_{\alpha \beta} K^C (d_R)_{\beta} + h.c. \quad (1.6)
$$

where $(u_R)_{\beta}$ and $(d_R)_{\beta}$ are columns containing right handed up and down type quark fields respectively:

$$
(u_R)_{\beta} = \begin{pmatrix} u \\ c \\ t \\ \end{pmatrix}_R \quad (d_R)_{\beta} = \begin{pmatrix} d \\ s \\ b \end{pmatrix}_R \quad (1.7)
$$

and $\bar{\psi}_L$ are a string of left handed Dirac-conjugated quark fields:

$$
\bar{\psi}_L = (\bar{u}, \bar{c}, \bar{t}, \bar{s}, \bar{b})_L = (\bar{U}^t, \bar{D}^t)_L \quad (1.8)
$$

The mass matrices $M^U$ and $M^D$ are $3 \times 3$ complex matrices. The primes on the U’s and D’s indicate that the U and D fields are eigenstates of the weak interaction.

When $K$ is expanded we obtain the generalized mass term:

$$
\mathcal{L}_{S}^{\text{fer}} = -\frac{g}{\sqrt{2}M} \bar{U}_L M^U U^t_R - \frac{g}{\sqrt{2}M} \bar{D}_L M^D D^t_R + h.c. \quad (1.9)
$$

Now the mass matrices $M^U$ and $M^D$ are not diagonal, which means we do not observe these eigenstates directly. Rather, we observe the combinations of these eigenstates which form the basis in which the mass matrices are diagonal. Thus we write:

$$
M^U = U^t_L m_u U_R, \quad M^D = D^t_L m_d D_R \quad (1.10)
$$

where the matrices $U_{R,L}$ and $D_{R,L}$ are chosen so that $m_u, m_d$ are diagonal, yielding the observable masses. These mass eigenstates are represented with unprimed quantities so that the mass term of the scalar part of the Lagrangian becomes:

$$
\mathcal{L}_{S}^{\text{fer}} = -\frac{g}{\sqrt{2}M} \bar{U} m_u U - \frac{g}{\sqrt{2}M} \bar{D} m_d D \quad (1.11)
$$
where $m_u$ and $m_d$ are now diagonal mass matrices in 3-dimensional $U$ and $D$ space.

The weak interaction eigenstates and the mass eigenstates are then related as:

$$U_L = U' U_L \quad U_R = U'_ R \quad D_L = D' D_L \quad D_R = D' D_R$$  \tag{1.12}

It is this process of diagonalising the fermion mass matrices which mathematically introduces what will become the CKM matrix. To see this, consider the fermion-boson interaction Lagrangian:

$$\mathcal{L}^\text{fer, int}_V = \Sigma_f [ig \sin \theta Q_f A_\mu \overline{f} \gamma_\mu f + i \frac{g}{2 \cos \theta} Z_\mu \overline{f} \gamma_\mu (v_f + a_f \gamma_5) f]$$

$$+ i \frac{g}{2 \sqrt{2}} W^{+}_\mu \overline{U} \gamma_\mu (1 + \gamma_5) C D + i \frac{g}{2 \sqrt{2}} W^-_\mu \overline{D} \gamma_\mu (1 + \gamma_5) C^T U$$  \tag{1.13}

Here, the weak interaction eigenstates ($U', D'$) have been expanded using equation 1.12, and the transformation matrices ($U, D$) have been collected together into the matrix:

$$C = U_L D^\dagger_L$$  \tag{1.14}

This is the CKM matrix which describes the coupling between the different generations via the W boson field\(^1\).

### 1.1.2 The CKM Matrix

The CKM matrix $C$ introduced in the last section is a complex $3 \times 3$ unitary matrix. There are therefore 9 independent parameters. However there are 6 possible fermionic fields involved in the the charged weak processes leading to the absorption of 5 of these parameters into redefinitions of the fields. The remaining 4 independent parameters are chosen as 3 parameters to describe the magnitudes of the elements and 1 to describe a complex phase. It is this phase which allows CP-Violation in the SM.

One popular parameterisation of the CKM matrix is due to Wolfenstein [3] and uses parameters $A$ ($\sim 1$), a complex number $(\rho + i \eta)$ and a small number $\lambda$ ($\sim 0.22$):

\(^1\)See [2] for more details of this algebra.
\[ V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \simeq \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \] (1.15)

Current experimental values of the magnitudes of the CKM parameters\[^4\] at the 90\% confidence level are\[^4\]:

\[
\begin{pmatrix}
0.9742 \text{ to } 0.9757 & 0.219 \text{ to } 0.226 & 0.002 \text{ to } 0.005 \\
0.219 \text{ to } 0.225 & 0.9734 \text{ to } 0.9749 & 0.037 \text{ to } 0.043 \\
0.004 \text{ to } 0.014 & 0.035 \text{ to } 0.043 & 0.9990 \text{ to } 0.9993 
\end{pmatrix}
\] (1.16)

It may be seen that the diagonal elements of this matrix (which represent the couplings within a given generation) are close to unity (ie 1 minus the square of the small number \(\lambda\)). The off diagonal elements however are much smaller, the extreme off diagonals (representing the couplings: \(u \rightarrow b\) and \(t \rightarrow d\)) being suppressed by the cube of a small number.

It is this inter-generational coupling which mediates the physical processes responsible for the width difference.

### 1.2 \(\Delta \Gamma_{B_s}\) Phenomenology

The remainder of this chapter is taken up with a discussion of the width difference in the \(B_s\) system, \(\Delta \Gamma_{B_s}\).

First, a general framework for the physical observables is introduced, and then in the subsequent sections the various theoretical techniques necessary for the calculation of \(\Delta \Gamma_{B_s}\) are described.

#### 1.2.1 General Formalism

There are three different sets of eigenstates which can be used to characterise the \(B_s\) sector. These are:

- \((B_H, B_L)\) : Mass Eigenstates
- \((B, \bar{B})\) : Strong interaction Eigenstates
- \((B_1, B_2)\) : CP Eigenstates
CHAPTER 1. THEORETICAL INTRODUCTION

The mass eigenstates are the states that have the physically observable quantities of mass and lifetime. Since these states decay via the weak interaction, the mass eigenstates are often called the weak interaction eigenstates. \(\Delta \Gamma_B\) is the width difference between these eigenstates.

The Strong interaction eigenstates are the states in which the particles are created. The name derives from the interaction which creates them (in this case in a collision between a \(p\) and a \(\overline{p}\)). Since a given meson is purely one of these states at creation time, it is the time evolution of the mass eigenstates from this initial condition which must be calculated.

The CP Eigenstates are defined as the eigenstates of the combined operations\(^2\) of first P and then C.

The time evolution of the mass eigenstates of a B meson is governed by:

\[
\begin{align*}
|B_H(t)\rangle &= \exp(-iM_H t) \exp(-\Gamma_H t/2) \ |B_H\rangle \quad (1.17) \\
|B_L(t)\rangle &= \exp(-iM_L t) \exp(-\Gamma_L t/2) \ |B_L\rangle \quad (1.18)
\end{align*}
\]

where \(H\) stands for the heavier of the two eigenstates and \(L\) stands for the lighter.

In contrast to this, the time evolution of the strong interaction eigenstates (|\(B\rangle\) and |\(\overline{B}\rangle\)) is much more complicated. An arbitrary linear combination of the strong interaction eigenstates:

\[
a|B\rangle + b|\overline{B}\rangle
\]

obeys a Schrodinger-like equation\(^3\):

\[
\frac{i}{\hbar} \frac{d}{dt} \begin{pmatrix} a \\ b \end{pmatrix} = \left( M - \frac{i}{2} \Gamma \right) \begin{pmatrix} a \\ b \end{pmatrix} \quad (1.20)
\]

where \(M\) and \(\Gamma\) are \(2 \times 2\) Hermitian matrices:

\[
M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad \Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix} \quad (1.21)
\]

The off diagonal elements of \(M\) and \(\Gamma\) are the dispersive and absorptive parts respectively\(^6\)\(^5\) of the transition amplitudes from \(|B^0\rangle \rightarrow |\overline{B}^0\rangle\) and visa versa.

\(^2\)P is the parity operation \((\vec{r} \rightarrow -\vec{r})\), and C is the charge conjugation operation (particles \(\rightarrow\) antiparticles)

\(^3\)The derivation of this equation is given in O.Nachtmann’s book\(^5\) and is not trivial.
In the SM, the off diagonal elements in M arise because of second order charged weak interactions described by box diagrams like the one shown in figure 1.1.

The matrix element for this diagram is roughly proportional to the masses of the two internal quark lines, so that even though these internal quark lines can be any of $u, c, t$ the diagram where both are $t$ dominates. Note that for the $M$ term of equation 1.20 the intermediate states are allowed to be virtual (hence the top quark is allowed in the loop). However when the second $\Gamma$ term is considered, it is a fundamental feature of equation 1.20 that the intermediate states have to be real\(^4\), and so only kinematically allowed states will be allowed in the intermediate state. It is this distinction that the terms absorptive and dispersive refer to.

The weak interaction eigenstates which decay, $B_L$ and $B_H$, are given in terms of strong interaction eigenstates:

$$
|B_L\rangle = p|B\rangle + q|\bar{B}\rangle \\
|B_H\rangle = p|B\rangle - q|\bar{B}\rangle
$$

(1.22)

where $p$ and $q$ are complex coefficients which obey the normalisation condition:

$$
|q|^2 + |p|^2 = 1
$$

(1.23)

The mass difference $\Delta m_B$ and width difference $\Delta \Gamma_B$ are defined:

\(^4\)This arises because of the $\delta$ function in Nachtmann I.32[5]. The term the $\delta$ function multiplies becomes the $\Gamma$ term in equation 1.20
\[ \Delta m_B \equiv M_H - M_L, \quad \Delta \Gamma_B \equiv \Gamma_H - \Gamma_L \]  
(1.24)

where \( \Delta m_B \) is consequently positive by definition. The matrices defined
in equation 1.21 are not diagonal which means \(|B\rangle\) and \(|\overline{B}\rangle\) are not the
eigenvalues of equation 1.20.

When equations 1.17, 1.18 and 1.22 are substituted into 1.20 the following
three equations are obtained:

\[ (\Delta m_B)^2 - \frac{1}{4}(\Delta \Gamma_B)^2 = 4(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2) \]  
(1.25)

\[ \Delta m_B \Delta \Gamma_B = 4\Re(M_{12}\Gamma_{12}^*) \]  
(1.26)

\[ \frac{q}{p} = \frac{\Delta m_B - \frac{i}{2}\Delta \Gamma_B}{2(M_{12} - \frac{i}{2}\Gamma_{12})} = -\frac{2(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}{\Delta m_B - \frac{i}{2}\Delta \Gamma_B} \]  
(1.27)

This algebra is considered in detail in appendix B.1. As is shown in
that section, when the approximations \( \Delta m_B \gg \Delta \Gamma_B, \ |M_{12}| \gg |\Gamma_{12}| \) are
applied, these three equations simplify into:

\[ \Delta m_B = 2|M_{12}|, \quad \Delta \Gamma_B = 2\Re(M_{12}\Gamma_{12}^*)/|M_{12}| \]  
(1.28)

\[ \frac{q}{p} = -\frac{|M_{12}|}{M_{12}} \]  
(1.29)

Equation 1.29 suggests that the approximation \( |\frac{q}{p}| \simeq 1 \) is valid. This is
usually the case and has two physical consequences. Firstly, it means that
the CP eigenstates reduce to the mass eigenstates.\(^5\) This result is used many
times in the current work. Secondly, this approximation is the same as saying
that CP violation in mixing can be neglected.

In order to proceed with a calculation of these quantities it is necessary
to introduce several theoretical techniques. The first of these is the Operator
Product Expansion (OPE).

\(^5\)Under this approximation, \(|B_L\rangle\) corresponds to CP-Even, and \(|B_H\rangle\) corresponds to
CP-Odd. This comes from inspection of equation 1.22
1.2.2 Operator Product Expansion

Motivation

At first glance it may seem possible to calculate \( M_{12} \) by just calculating the box diagram of figure 1.1. However, this diagram treats the \( b \) and \( s \) quarks as if they were free, whereas the reality is that they are confined by the strong interaction within the \( B_s \) meson.

It is therefore desirable to factor the calculation into three different parts, each pertaining to a given distance \( (x) \) scale:

- \( x \sim \frac{1}{c M_W} \) the distance scales of the virtual \( t \) quarks and \( W \) mesons. This is the scale where the interaction appears as in figure 1.1.

- \( x \sim \frac{1}{c m_b} \) the distance scale where the fundamental charged weak interaction appears as a point force, but where perturbation theory is still valid. This is where gluon loop corrections to the lowest order weak interaction will be included.

- \( x \sim \frac{1}{c \Lambda_{QCD}} \) the distance scale where non-perturbative QCD (lattice field theory) has to be employed. This corresponds to the long distance meson part of the calculation.

This factorization can be achieved with a theoretical tool called operator product expansion[8].

If the different distance scales cannot be disentangled, the ability to calculate the box diagram of figure 1.1 is useless. The box diagram only describes the very short range \( W \)-boson interaction between the quarks. But there are initial and final meson states in addition to radiative QCD corrections. These will be considered in the following sections, but without the OPE, it is not possible to consider these contributions separately. This important technique will now be outlined.

Approximating the Propagator

The essence of the mathematical manipulation involved is an expansion of the propagator(s) in the interaction amplitude(s) to be calculated in terms of a taylor expansion. Consider the example of a weak decay of a \( b \)-quark. The \( W \) boson is the propagator in question and the amplitude is proportional to:
This propagator may be expanded in a Taylor series:

\[
\frac{(ig_\gamma)^2/4}{p^2 - M_W^2} = \sum_{n=0}^{\infty} \frac{g_\gamma^2}{8M_W^2} \left[ 1 + \frac{p^2}{M_W^2} + \frac{p^4}{M_W^4} + \cdots \right]
\]  \hspace{1cm} (1.31)

which is only valid when \( p^2 \ll M_W^2 \). Since for B decays the external kinematics requires \( p^2 \leq m_b^2 \) this condition is well satisfied.

Thus the momentum dependent propagator of equation 1.30 is identical to the infinite series of local\(^6\) operators:

\[
\frac{G_F}{\sqrt{2}} \left( 1 + \frac{(i\partial)^2}{M_W^2} + \frac{(i\partial)^4}{M_W^4} + \cdots \right)
\]  \hspace{1cm} (1.32)

where the multiplicative constants have been reinterpreted using the Fermi decay constant. Thus the first term represents the lowest order charged weak interaction, and the higher order terms are local operators of higher mass dimension (2, 4 etc).

1.2.3 Towards a low energy effective theory

Given that the series may be truncated at the first term as mentioned, the matrix element may be written in its approximated form and a low energy hamiltonian may be deduced. This is what is meant by writing down a low energy effective theory. The propagator is expanded, the series truncated, and the hamiltonian which would have given the resulting expression is thus deduced.

In order to gain an intuitive understanding of what has just been performed consider the physical picture which we are left with in light of the above procedure. Including the rest of the amplitude, in the absence of QCD effects, the low energy effective theory may be written as:

\[
\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} V_{cs}^* V_{ud} (\bar{c}c)_{V-A} (\bar{u}d)_{V-A}
\]  \hspace{1cm} (1.33)

\(^6\)local because of the derivatives on top of the fractions in contrast to the non local operator in equation 1.30
where the eff subscript on the hamiltonian stands for effective, and \((\bar{s}c)_{V-A}\), \((\bar{u}d)_{V-A}\) are V-A currents.

This has the same propogator structure as the old Fermi theory of beta-decay. There, as here, the physical picture lying behind the mathematics is that of a point interaction.

### 1.2.4 Wilson Coefficients

The low energy effective hamiltonian introduced in section 1.2.3 resulted from a taylor expansion of the propogator of equation 1.30. So far only the first distance scale of \(x \sim \frac{\hbar}{c M_W}\) has been described. The hamiltonian of equation 1.33 describes the interaction between the W-bosons and the quarks (the shortest distance scale) but does not make any statement about the radiative QCD corrections which must be considered at the typical energy scales of a B-meson (ie \(\alpha_s\) is stronger as lower energies are considered). Considering these QCD corrections will constitute incorporating the second distance scale of \(x \sim \frac{\hbar}{c m_q}\). Wilson coefficients are the way of parameterising the magnitude of these QCD corrections.

Consider the concrete example of the quark-level transition \(c \rightarrow s u d\). In the absence of QCD effects, the low energy effective theory may be written:

\[
H_{\text{eff}} = C \frac{G_F}{\sqrt{2}} V_{cs}^* V_{ud}(\bar{s}c)_{V-A}(\bar{u}d)_{V-A}
\]  

(1.34)

which is identical to equation 1.33 except for the multiplicative constant \(C\). This \(C\) is a wilson coefficient. It is the weighting factor in front of a given term in the low energy effective hamiltonian. In the absence of QCD, \(C = 1\), however, when we consider radiative gluonic corrections to the tree level process, the situation becomes more complicated and \(C \neq 1\).

Consider the following diagram:

Such radiative corrections to the underlying weak interaction process must be taken into account. This can be accomplished within normal perturbation theory crucially, because the strong interaction is asymptotically free. In other words, the energy range of the virtual gluon inside the loop is high enough so that \(\alpha(s)\) is small.

The only alteration to the low energy hamiltonian of equation 1.34 is the addition of a term with a different color structure to the pure weak interaction expression. This is because it is now necessary to consider the
strong interaction (gluons) which have a more complicated colour structure. The effective hamiltonian becomes:

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} V_{cs}^* V_{ud} (C_1 Q_1 + C_2 Q_2)$$  \hspace{1cm} (1.35)$$

where

$$Q_1 = (\bar{s}_i c_j)_{V-A} (\bar{u}_j d_i)_{V-A}$$  \hspace{1cm} (1.36)$$

and

$$Q_2 = (\bar{s}_i c_i)_{V-A} (\bar{u}_j d_j)_{V-A}$$  \hspace{1cm} (1.37)$$

The indices $i, j$ are color indices and each run over the three colors. The color structure of the two operators in equations 1.36 and 1.37 is different. The different indices $(i,j$ and $j,i)$ in equation 1.36 indicate color mixing within the operators. In contrast the similar indices $(i, i$ and $j, j)$ in equation 1.37 indicate a simpler color structure associated with the pure weak interaction[9].

So the second term in equation 1.35 is the one which corresponds to equation 1.34. However, the wilson coefficient ($C_2$) is different now and needs to be calculated by matching the effective theory onto the full theory. If the results from the low energy effective theory do not match the full theory when calculated to the appropriate order then the formulation of the low energy theory is wrong.
1.2.5 Calculating the Wilson Coefficients

The Wilson Coefficients are calculated by matching the low energy theory onto the full theory. In this process, the amplitude in the full theory (with the propagator unexpanded) has to be the same as the amplitude in the low energy effective theory. The crucial simplification which is possible here though is it doesn't matter what the initial and final states are. The whole challenge of this calculation is that the physical initial and final states are mesons, and therefore incalculable in normal perturbation theory ($\alpha(s) > 1$ for the meson states). However, simply for the purposes of calculating the wilson coefficients, which after all represent the short range part of the calculation, it does not matter what the initial and final states are. Therefore, free quark initial and final states can be used, which just amounts to a straightforward calculation of the relevant feynmann diagrams. When this procedure is applied, the wilson coefficients are calculated:

$$C_1 = -3\frac{\alpha_s}{4\pi} \ln \left( \frac{M_W^2}{\mu^2} \right) \quad C_2 = 1 + \frac{3}{N} \frac{\alpha_s}{4\pi} \ln \left( \frac{M_W^2}{\mu^2} \right)$$ (1.38)

where $N$ is the number of colours. In calculating the amplitude of the full theory to the appropriate order ($\alpha_s$) it is necessary to introduce a cut-off in the integration over some of the loops. This is to keep the expression from becoming uncontrollably divergent. This cut-off is referred to as the renormalisation scale, and is the symbol $\mu$ in equation 1.38[10].

There is one final complication before the correct wilson coefficients are obtained. The coefficients in equation 1.38 have been calculated in normal perturbation theory, with a normalization scale that makes logarithms like $\ln \left( \frac{M_W^2}{\mu^2} \right)$ small (eg $\mu \sim M_W$). However, the energy scale in the problem at hand is the b-meson regime of a few GeV where the logarithms that appear in equation 1.38 are not small. In order to extrapolate the wilson coefficients to this energy regime, a theoretical technique called the renormalisation group is employed.

1.2.6 The renormalisation group

Despite the name, the renormalisation group doesn't have anything to do directly with group theory. Rather, it is a group of transformations which describe how renormalised quantities change with the renormalisation scale $\mu$. 
Mathematically these transformations are calculated by solving differential
equations.

Once the right renormalisation group equation has been applied, the wil-
son coefficients can be written as:

\[
C_\pm(\mu) = \left( \frac{1}{1 + \beta_0 \frac{\alpha_s(\mu)}{4\pi} \ln \frac{M_W^2}{\mu^2}} \right)^{\gamma_\pm(0)}
\]  

(1.39)

where \(\alpha_s(\mu)\) is the strong coupling at the renormalisation scale \(\mu\). The
\(\pm\) subscripts on \(C\) indicate a change of basis \(C_\pm = C_2 \pm C_1\). \(\beta_0\) is a factor
dependent on the number of flavours and number of colours, and so is
\(\gamma_\pm(0)\)[11].

It is a feature of equation 1.39 that it can be expanded in terms of an
infinite series of logarithms of the form \(\ln \frac{M_W^2}{\mu^2}\). This stands in contrast to
equation 1.38 which only contains one such term. Recall that equation 1.38
gives the wilson coefficients before the renormalisation group equations have
been used to transform to the correct energy scale.

Thus, another way of looking at the application of the renormalisation
group equations is that these leading log terms will have been summed to all
orders in perturbation theory. At \(\mu \sim M_W\) these logs were not significant,
but at \(\mu \sim m_b\) they would cause the series to diverge. So they need to be
summed to all orders of the series.

1.2.7 Summary

This concludes the short distance part of the entire calculation. It is worth
reiterating what has been introduced thus far:

- With the application of the operator product expansion, the full ma-
  trix element for a low energy B-physics process can be split up into
  short distance and long distance contributions. This factorisation is
evident in equation 1.40. That is to say the multiplicative terms in
the expression which account for the different distant scales are now
separate. This stands in contrast to the situation where the amplitude
expression is written in terms of the propagator in equation 1.30. Here
the propagator (which stands for the effect of the W-bosons) sits be-
between (and therefore acts upon) the initial and final meson states. The distance scales are entangled. This is the importance of the OPE.

- The short distance contributions are fully contained within *wilson coefficients* which have to be calculated by matching the low energy effective theory onto the full theory.

- When QCD short range effects are taken into account, the calculation of these coefficients is complicated by the fact that the energy range of the gluons in the loops is down at about 1 GeV.

- This inadequacy is overcome by the application of the differential equations of the renormalisation group which extrapolate down to the right energy scale.

The amplitude for the entire physical process is now:

$$ A = -i \frac{G_F}{\sqrt{2}} V_{cs} V_{ud} (C_1 \langle Q_1 \rangle + C_2 \langle Q_2 \rangle) \quad (1.40) $$

where $\langle Q_1 \rangle$ and $\langle Q_2 \rangle$ are the long range hadronic effects calculated between the initial and final meson states.

The wilson coefficients, $C_1, C_2$ are now fully determined. There are of course uncertainties which can be beaten down by considering higher order corrections, but the short distance part of the calculation is over. All that remains is to evaluate $\langle Q_1 \rangle$ and $\langle Q_2 \rangle$. It is this topic which takes up the next section.

### 1.2.8 Long Range Hadronic effects

Now the last distance scale ($x \sim \frac{\hbar c}{\Lambda_{QCD}}$) in the problem has been reached. Mathematically, the difficulty with calculating the hadronic matrix elements mentioned at the end of section 1.2.7 has to do with the break down of perturbation theory. The strong interaction coupling constant $\alpha_s$ is almost 1 for the long distance part of the amplitude because the energies involved are so low. This means that the perturbation series does not converge and other non-perturbative techniques must be employed. Currently, the most powerful technique is lattice gauge theory, where space time is approximated as a discrete lattice of points and the integrals are evaluated by *brute force* [12].
1.2.9 Bag Parameters and the Breakdown of the Vacuum Saturation Approximation

When lattice field theory is initially applied to calculate the hadronic matrix element it is necessary to apply a further approximation called the **Vacuum Saturation Approximation**. In the case of $B_s$ mixing the following hadronic matrix element needs to be calculated:

$$\langle \bar{B}^0 | \slashed{D} \gamma^5 d \bar{B} | B^0 \rangle$$

(1.41)

Recall that the propagator has been removed by the application of the operator product expansion. The result is the matrix element of a **four quark operator** (ie equation 1.41). In vacuum saturation a complete set of states is inserted in between the two currents in equation 1.41 (in between the $d\bar{b}$ pair). Then the assumption is made that the sum is dominated by the vacuum. In this case, the currents have effectively been decoupled from each other. Each current interacts with the vacuum separately. The uncertainty that this adds into the result for the hadronic matrix element is parameterised by the **Bag parameter** $B_B$. In the limit that the vacuum saturation approximation holds exactly, $B_B = 1$. This approximation sounds very unsatisfactory since it would mean that the mesons are non-interacting. However, calculations on the lattice indicate that a value of $B_B = 1$ is favoured.

1.2.10 Pulling it all together

Recall that the expression for $\Delta \Gamma_B$ is:

$$\Delta \Gamma_B = \frac{2\Re(M_{12} \Gamma_{12}^*)}{|M_{12}|}$$

(1.42)

It is now possible to give an expression for the factor $|M_{12}|$ in this expression.

Due to the operator product expansion already discussed the off diagonal mass matrix element may be written:

$$M_{12} = \frac{\langle B^0 | H^{\Delta B=-2} | \bar{B}^0 \rangle}{2m_B}$$

(1.43)

where
\[ H^{\Delta B} = \frac{G_F^2}{16\pi^2} (V_{tb} V_{ts}^*)^2 C^{\Delta B} (m_t, M_W, \mu) Q(\mu) + \text{h.c.} \] (1.44)

Where \( Q \) is the long range hadronic portion and \( C \) is the wilson coefficient portion.

The result of applying lattice gauge theory to the hadronic matrix element under the approximations already discussed is:

\[ \langle B^0 | Q(\mu) | B^0 \rangle = \frac{8}{3} B_B(\mu) f_B^2 m_B^2 = \frac{8}{3} \tilde{B}_B(\mu) f_B^2 m_B^2 \] (1.45)

Where \( B_B(\mu) \) is the bag parameter introduced in connection to the vacuum saturation approximation, \( f_B \) is known as a leptonic decay constant, defined as \( \langle 0 | A_\mu | B(p) \rangle = f_B p_\mu \), and \( m_B \) is the mass of the B meson.

The wilson coefficient, \( C \) is calculated to be:

\[ C^{\Delta B} (m_t, M_W, \mu) = M_W^2 S(x) \eta_B b_B(\mu) \] (1.46)

where \( x = f \left( \frac{m_t}{m_W} \right) \) and \( S(x) \) is calculated from the box diagram in figure 1.1 to be:

\[ S(x) = x \left[ \frac{1}{4} + \frac{9}{4} \frac{1}{1-x} - \frac{3}{2} \frac{1}{(1-x)^2} \right] - \frac{3}{2} \frac{x}{1-x} \left[ \frac{x}{1-x} \right] ^3 \ln(x) \] (1.47)

The parameters \( \eta_B \) and \( b_B(\mu) \) are:

\[ \eta_B = |\alpha_s(M_W)|^{6/23} \quad b_B(\mu) = [\alpha_s(\mu)]^{-6/23} \] (1.48)

The reason why the quantity \( \tilde{B}_B \) in equation 1.45 was introduced was to illustrate that the \( \mu \) dependence (the renormalisation scale) drops out in the product of the long and short range parts of the calculation.

Putting all of this together, and making the expression explicitly for the \( B_s \) as opposed to the \( B_d \) we obtain:

\[ |M_{12}| = \frac{G_F^2}{12\pi^2} \eta_B m_B \tilde{B}_B f_{B_s}^2 M_W^2 S(x) |V_{tb} V_{ts}^*|^2 \] (1.49)

The current values for the lattice parameters \( f_{B_s}, f_{B_d}, \tilde{B}_{B_s} \) and \( \tilde{B}_{B_d} \) are given in table 1.1. Note that distinction between \( B_s \) and \( B_d \) is made in these values since an analogous expression to 1.49 exists for \( B_d \).
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{B_s}$</td>
<td>210 MeV</td>
<td>30 MeV</td>
</tr>
<tr>
<td>$f_{B_d}$</td>
<td>245 MeV</td>
<td>30 MeV</td>
</tr>
<tr>
<td>$\hat{B}_{B_d}$</td>
<td>1.4</td>
<td>0.1</td>
</tr>
<tr>
<td>$\hat{B}_{B_s}$</td>
<td>1.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 1.1: Current calculations for the lattice parameters $f_{B_s}$, $f_{B_d}$, $\hat{B}_{B_d}$ and $\hat{B}_{B_s}$[13].

1.2.11 Γ_{12}

Recall the equation for $\Delta \Gamma_B$:

$$\Delta \Gamma_B = \frac{2\Re(M_{12} \Gamma_{12}^*)}{|M_{12}|}$$ (1.50)

The only factor that needs to be calculated is now $\Gamma_{12}$.$^7$

The feynmann diagram which contributes to $\Gamma_{12}$ at lowest order is shown in figure 1.3. It is a fundamental feature of equation 1.20 that the $M$ term allows for off-shell intermediate states (which is why the top quark is allowed in the mixing diagram figure 1.1), while the $i\Gamma$ term requires the intermediate states to be real[5]. Therefore the top quark cannot appear in the intermediate state of figure 1.3. The charm quark has the largest contribution. This form of interaction term is known as a $bilocal$ term because it has two local interactions. These interactions have been contracted to vertices in figure 1.3, since they are short range and can be absorbed into wilson coefficients in a manner similar to the way described in section 1.2.4.

When this diagram is calculated and summed over real intermediate states the following expression is obtained:

\footnote{\textsuperscript{7}note that the phase of $M_{12}$ and $\Gamma_{12}$ must also be known. Beneke \textit{et al} give a comprehensive treatment of the $\Delta \Gamma_B$, calculation in [\textsuperscript{14} section 2.}
\[ \Gamma_{12} = -\frac{G_F^2}{24\pi m_{B_s}} (V_{cb} V_{cs}^*)^2 m_b^2 \sqrt{1 - 4z} \left[ (1 - z) K_1 + \frac{1}{2} (1 - 4z) K_2 \right] \]

\[ \langle B_s^0|Q|B_s^0\rangle + (1 + 2z)(K_1 - K_2)\langle B_s^0|Q_S|B_s^0\rangle \]  \hspace{1cm} (1.51)

where \( z = \frac{m_s^2}{m_b^2} \) and

\[ K_1 = 3C_1^2 + 2C_1 C_2 \quad K_2 = C_2^2 \]  \hspace{1cm} (1.52)

where the C’s are wilson coefficients as before.

The only difference is that since we have another operator \( Q_S \) another

\textit{bag parameter} \( B_S(\mu) \) is required.

### 1.3 Current Results

A current theoretical prediction for the SM value of \( \frac{\Delta \Gamma_{B_s}}{\Gamma_{B_s}} \) is\cite{48}:

\[ \frac{\Delta \Gamma_{B_s}}{\Gamma_{B_s}} = 0.12 \pm 0.06 \]  \hspace{1cm} (1.53)

The only current experimental prediction is\cite{49}:

\[ \frac{\Delta \Gamma_{B_s}}{\Gamma_{B_s}} < 0.69 \text{ at the 95\% C.L.} \]  \hspace{1cm} (1.54)
A precise measurement of \( \frac{\Delta \Gamma_{B_s}}{\Gamma_{B_s}} \) would therefore be a good test of the SM. However, the compelling reason to measure \( \frac{\Delta \Gamma_{B_s}}{\Gamma_{B_s}} \) is the new physics potential. It is a fairly robust prediction[50] that physics beyond the standard model acts to reduce \( \frac{\Delta \Gamma_{B_s}}{\Gamma_{B_s}} \). In particular, a precise\(^8\) constraint to zero would be a powerful indicator of new physics.

1.4 Epilogue

Thus most of the important theoretical techniques employed in a calculation of \( \Delta \Gamma_{B_s} \) have been surveyed. Much of the rest of the current work is taken up with studies directly pertaining to the measurements of this quantity. The following chapters will first introduce how one might go about experimentally measuring it. Then the studies themselves will be presented.

\( ^8 \delta(\frac{\Delta \Gamma_{B_s}}{\Gamma_{B_s}}) \approx 0.03 \)
Chapter 2

The CDF Detector

The CDF Run II detector is a general purpose particle physics experiment located on the Tevatron ring at Fermilab Illinois. The following sections describe the experiment. First the accelerator is very briefly described. The remainder of the chapter provides an overview of the Run II CDF detector.

2.1 Introduction

The CDF detector is a general purpose High Energy Physics (HEP) experiment located at the Fermi National Accelerator Laboratory (FNAL). It is designed to detect particles resultant from \( p\bar{p} \) collisions. These collisions are provided by the Tevatron collider. This is a multi-stage accelerator which is described in section 2.2.

During the period from 1987 to 1995 CDF (and D0) took data which amongst other things led to the discovery of the top quark [15][16], and the \( B_s \) meson [17]. After 1995 there was a shut down while an extensive upgrade was designed and implemented.

2.2 Tevatron

The Tevatron accelerator is the highest energy stage in a sequence of accelerators at Fermilab. The final \( p\bar{p} \) centre of mass energy is designed to be almost 2 \( TeV \) as compared to 1.8 \( TeV \) in Run I. There are several scenarios for luminosity in Run II. These are characterised by different bunch spacing,
number of bunches, and beam crossing angle. A comparison of these features is given in table 2.1 (note: The exact values of some parameters such as luminosity will not be clear until the run begins).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Lower $\mathcal{L}$</th>
<th>Higher $\mathcal{L}$</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antiproton Bunches</td>
<td>36</td>
<td>121</td>
<td>$\mu$rad</td>
</tr>
<tr>
<td>Proton Bunches</td>
<td>36</td>
<td>140</td>
<td>$cm^{-2} sec^{-1}$</td>
</tr>
<tr>
<td>Crossing Angle</td>
<td>0</td>
<td>136</td>
<td>$n$sec</td>
</tr>
<tr>
<td>Design Luminosity</td>
<td>$0.9 \times 10^{32}$</td>
<td>$1.6 \times 10^{32}$</td>
<td></td>
</tr>
<tr>
<td>Bunch Spacing</td>
<td>396</td>
<td>132</td>
<td></td>
</tr>
<tr>
<td>Ave Interactions/crossing</td>
<td>2.3</td>
<td>1.3</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Comparison of some projected modes of running

### 2.3 The CDF Run II Detector

This section will not discuss all detector elements with equal weight. Instead, the elements which are relevant to the studies in hand will be emphasised. The important elements are thus the tracking systems and the trigger. Figure 2.1 shows an elevation view of the CDF detector. See section A.1 for a definition of the CDF coordinate system.

#### 2.3.1 Tracking Systems

The tracking systems sub divide into the main tracking volume (a large open cell drift chamber) and the silicon tracking systems (closer to the beam pipe). The main tracking volume is called the Central Outer Tracker (COT) [18]. Table 2.2 describes the coarse geometric characteristics of these systems, and figure 2.2 shows most of this information graphically. There are three distinct silicon sub-systems. The main sub-system is the Silicon Vertex Detector (SVX II) [20]. There are two extensions to this. First the Intermediate Silicon Layers (ISL) [19] outside SVX II, and second the innermost layer of silicon, Layer Zero-Zero (L00) inside SVX II [21][22].
Figure 2.1: An Elevation view of the CDF detector
Figure 2.2: Coarse Tracking systems geometry. (Layer 00 not shown.) The large parallel lines taking up the upper part of the figure are represent the COT, the irregularly spaced bold lines below this represent the ISL, and the smaller closely spaced lines at the bottom left represent the SVX.
<table>
<thead>
<tr>
<th>Sub-system</th>
<th>$R(cm)$</th>
<th>$\Delta Z(cm)$</th>
<th>$\eta$ max</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>COT</td>
<td>$40 \to 137$</td>
<td>320</td>
<td>1</td>
<td>R range from innermost to outermost wafer</td>
</tr>
<tr>
<td>SVX II</td>
<td>$2.1 \to 12.9$</td>
<td>87</td>
<td>2</td>
<td>$Z$ range describes extent of space frame</td>
</tr>
<tr>
<td>ISL</td>
<td>$22.6 \to 29.0$</td>
<td>175</td>
<td>1.9</td>
<td>R refers to radius of the centres of sensors</td>
</tr>
<tr>
<td>L00</td>
<td>$1.35, 1.62$</td>
<td>96</td>
<td>4</td>
<td>$Z$ refers to the length of SVX inner screen</td>
</tr>
</tbody>
</table>

Table 2.2: Coarse Geometrical parameters of Tracking Subsystems.
CHAPTER 2. THE CDF DETECTOR

Unless otherwise specified, for table 2.2, the R and Z range refer to the mechanical extent of the detector, while the η range refers to a track which intersects the entire sub-system in question (see section A.3 for the definition of η).

2.3.2 COT

Figure 2.3 shows $\frac{1}{6}$ of an end plate of the COT. This gives an idea of the geometry of the detector. The main design constraint which the COT had to fit was that the drift time be less than 132μs due to the decreased bunch spacing. Between the wires there is a 50 : 35 : 15 Ar – Et – CF₄ gas mix. This has a drift velocity of $\sim 100\mu m/\mu s$ which allows a drift distance of $\sim 1 cm$. There are 3 main elements to the main volume of the detector apart from the gas. There are tungsten sense wires, tungsten shaper wires, and gold on Mylar field panels. The sense wires are part of the circuit which actually registers the potential caused by the passage of ions resultant from the passage of a high energy particle. The shaper wires shape the potential field, and the field panel forms the cathode of the detection circuit. The electric field is $\sim 2.5 \text{ kV/cm}$.

The single hit resolution is expected to be $\sim 120 \mu m$, and the best case $p_t$ resolution is expected to be $\frac{\sigma_{(p_{t})}}{p_{t}} \approx 0.0009$ where $p_{t}$ is in $\text{GeV}$. Figure 2.4 shows an early Run II $J/\Psi$ mass peak from the $\mu^+\mu^-$ decay channel.

2.3.3 SVX II

SVX II is a fine resolution strip silicon detector. The cylindrical geometry divides into three barrels, where each barrel further splits into 12 wedges. One wedge of a given barrel consists of 5 double sided layers. Each layer consists of 4 silicon sensors arranged as two pairs. The sensors in each pair are wire bonded together so that both ends of a given barrel are read out. The most coarse read-out entity is the VRB (VME Readout Buffer) which corresponds to two wedges on one side of a barrel. There are thus $2 \times 6 = 12$ VRB’s per barrel, 36 in total. This information is represented graphically in figures 2.5, 2.6, and 2.7. Figure 2.5 shows an end view of SVX II, the bulkheads and hybrids are clearly visible. Figure 2.6 is a blowup of $\frac{1}{6}$ of SVX II, end-on, showing the section of a barrel which is read out by one VRB. Figure 2.7 shows what one layer from a wedge looks like from the top, in schematic form. The SVX 3 readout chip positions are shown here.
Figure 2.3: A view of $\frac{1}{6}$ of an end plate of the COT. The beam axis is out of the page.
Figure 2.4: An early Run II $J/\psi$ mass peak from data taken in August 2000
There are 3 different sorts of strips in SVX II: \( r - \phi \), small angle stereo, and 90° stereo. Table 2.3 summarises how these are allocated across the layers.

One important feature of the SVX II detector is that a significant proportion of the readout electronics is actually mounted on the detector. At the time the detector was designed, it was judged that this was the best way of ensuring an acceptable level of loss of information. The portcards [23] are mounted outside the SVX II detector, but the hybrids (structures the readout chips are mounted on) are mounted on the layers of silicon, and are therefore much closer to the interaction region. The effect of these hybrids is discussed in section 2.3.5.

<table>
<thead>
<tr>
<th>Property</th>
<th>Layer 1</th>
<th>Layer 2</th>
<th>Layer 3</th>
<th>Layer 4</th>
<th>Layer 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>nos ( \phi ) strips</td>
<td>256</td>
<td>384</td>
<td>640</td>
<td>768</td>
<td>896</td>
</tr>
<tr>
<td>nos Z strips</td>
<td>256</td>
<td>576</td>
<td>640</td>
<td>512</td>
<td>896</td>
</tr>
<tr>
<td>stereo angle</td>
<td>90°</td>
<td>90°</td>
<td>+1.2°</td>
<td>90°</td>
<td>-1.2°</td>
</tr>
<tr>
<td>( \phi ) strip pitch (( \mu m ))</td>
<td>60</td>
<td>62</td>
<td>60</td>
<td>60</td>
<td>65</td>
</tr>
<tr>
<td>Z strip pitch (( \mu m ))</td>
<td>141</td>
<td>125.5</td>
<td>60</td>
<td>141</td>
<td>65</td>
</tr>
</tbody>
</table>

Table 2.3: SVX II Strip Parameters

Impact parameter resolutions for central high momentum tracks are \( \sigma_\phi < 30\mu m \) and \( \sigma_z < 60\mu m \). At normal incidence, the position resolution is described by \( \frac{\text{strip pitch}}{\sqrt{12}} \), whereas for particles with oblique incident angles the resolution is expected to vary between 12 and 25\( \mu m \).

### 2.3.4 ISL

The ISL has a different configuration in the central low \( \eta \) region than in the forward region. The central barrel of the ISL consists of ladders that are placed alternately at radii of 22.6 or 23.1 cm. The staggering arrangement can be seen in figure 2.8. This layer is extended in a more forward position by a set of ladders at 19.7 and 20.2 cm. The outermost layer of the ISL, or the forward barrel consists of ladders mounted at 28.6 and 29.0 cm. This three-fold arrangement of the ISL can be seen schematically in figure 2.9.
Figure 2.5: End view of the SVXII detector

Figure 2.6: End view of a VRB

Figure 2.7: 2 HDI’s, or 1 ladder
Each ISL sensor is the same, having $r - \phi$ strips on one side and small-angle stereo on the other. The sensor parameters are summarised in table 2.4.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>nos $\phi$ strips</td>
<td>1024</td>
</tr>
<tr>
<td>nos stereo strips</td>
<td>768</td>
</tr>
<tr>
<td>stereo angle</td>
<td>1.2°</td>
</tr>
<tr>
<td>$\phi$ strip pitch ($\mu m$)</td>
<td>55</td>
</tr>
<tr>
<td>stereo strip pitch ($\mu m$)</td>
<td>73</td>
</tr>
<tr>
<td>$\phi$ position resolution ($\mu m$)</td>
<td>$&lt; 16$</td>
</tr>
<tr>
<td>stereo position resolution ($\mu m$)</td>
<td>$&lt; 23$</td>
</tr>
</tbody>
</table>

Table 2.4: Physical ISL Sensor Parameters

2.3.5 Layer 00

Layer 00 is the innermost layer of silicon. The mechanical support and cooling apparatus is mounted directly onto the beampipe. The layers of silicon are single sided (axial layers) at radii of 1.35 cm and 1.62 cm. This arrangement can be seen in figure 2.10.

Figure 2.11 shows the importance of L00 (see [22] for this plot). The presence of the SVX 3 chips right on the hybrids in SVX II seriously compromises the impact parameter resolution in these localised areas. The figure shows impact parameter resolutions calculated for stiff tracks, perpendicular to the beam line. The effect of L00 may be seen by the lower graph. This radical improvement in the resolution is due to the chips being mounted at the end of the hybrids, rather than in the middle. The positions of the readout electronics may be seen underneath the graphs. The small green rectangles indicate the positions of the SVX 3 chips for the inner SVX layer. These line up with the degradation regions for the impact parameter resolution. The end position of the L00 readout electronics is shown by the blue rectangles. It may readily be seen that such an improvement is vital for B-physics. It is worth mentioning that L00 was motivated by radiation considerations also. SVX II is not expected to survive in its entirety for the whole of Run II. The inner layers particularly are vulnerable to radiation damage because of the
Figure 2.8: Close up end view of ISL forward region

Figure 2.9: Schematic view of one half of the ISL
higher fluxes. L00 was constructed using LHC detector technology which is necessarily a more rad-hard design than CDF.

2.4 Trigger

In the following section, the CDF trigger is described. The emphasis is again placed on aspects of the trigger which are relevant for the studies in hand. Thus, the Two-Track hadronic trigger [24] is described in some detail.

The interaction rate in Run II will be $> 10$ MHz. The rate to tape will be a fraction of this ($\sim 80$ $Hz$) [25]. In order to cope with the vast rejection factor that this necessitates, the CDF Run II trigger is a three level architecture. Each level provides a reduction sufficient to allow for processing in the next level with minimal dead time. Level 1 is purely hardware and has an output rate to Level 2 of about 40 $KHz$. Level 2 consists of programmable alpha processors and a dedicated impact parameter trigger. Its output rate to Level 3 is expected to be between about 500 $Hz \to 800$ $Hz$. Level 3 is purely software.

2.4.1 Level 1

Level 1 [26] consists of 3 sub-systems which synchronously feed the global Level 1 decision unit. These three are:

- L1CAL: Calorimeter based objects.
Figure 2.11: The effect of Layer 00 to improve the Impact Parameter resolution due to the hybrids
• MUON, PRIM-L1MUON: Muon objects.

• XFT, XTRP-L1 TRACK: Track based objects.

XFT stands for Xtremely Fast Tracker, and XTRP stands for Extrapolation Unit. The XTRP measures $p_t$ and $\phi$ of tracks.

Since the muon and calorimeter streams require the presence of a track pointing at their respective outer detector elements, the tracks found by the XFT must be sent to the other streams in addition to the global Level 1 decision unit.

The XFT and the XTRP are the most important elements of Level 1 for the purposes of B-physics. The minimum design requirements for the XFT are given by the performance of the old run I counterpart:

• Track finding efficiency > 96% when the single-hit efficiency of the COT > 92%

• Momentum resolution reported to the trigger: $\delta(\frac{1}{p_t}) < 0.02 GeV^{-1}$

• Resolution on $\phi_0$ better than 6 mrad (see appendix A.1).

• Fake-track rejection at least twice as good as the CFT (Run I Central tracking chamber).

• Efficient down to $p_t \sim 1.5 GeV/c$

The XFT takes a subsection of COT data and applies a two stage algorithm to locate high $p_t$ tracks. For each one of four layers of the COT (the axial layers), the first stage of the XFT algorithm applies a mask of possible hit configurations for $p_t \geq 1.5 GeV$. The second stage of the algorithm then examines 1.25° angular bins to put the high $p_t$ hits in these layers together to form high $p_t$ track candidates. This trigger forms the basis for all of the hadronic B-Physics triggers at CDF Run2.

The XTRP receives tracks from the XFT and distributes them to the L1 and L2 trigger sub-systems. For the purposes of B-physics, L1 TRACK is the most important part of L1 which the XTRP serves. If a track is above the $p_t$ threshold, the XTRP feeds it to L1 TRACK. If 6 or more tracks are found, an automatic Level 1 accept is generated. If 2, 3, 4 or 5 tracks are found, the $P_T$ and $\phi$ information is used in conjunction with look-up tables to evaluate other Level 1 triggers. This is the hardware which forms the Two-Track trigger (see section 2.4.3) at Level One.
2.4.2 Level 2

The Level 2 sub-systems relevant to the studies in hand are the SVT (Secondary Vertex Trigger) [27] and the alpha processors [28]. The alpha-processors are reprogrammable and are used more generally than simply for track triggers, but they provide valuable capability to impose kinematic cuts necessary for the Two-Track-Trigger.

The SVT uses the SVX II detector (NOT the ISL and L00) to find high impact parameter tracks. Impact parameter is defined in section A.2. At Level 2 it is a 2 dimensional quantity (the projection of its 3 dimensional definition onto the transverse plane).

The SVT operates on a subset of the data from the SVX and the COT. It takes the high $pt$ track candidates from Level 1 and forms superstrips which are used in conjunction with SVX II data in the pattern recognition stage. The digitised data from the SVX II is fed into the Hit Finder which clusters the strips into hits. The combination of hits and superstrips are then fed in parallel to a number of Associative Memory chips which compare the data to a list of previously computed legitimate combinations. Finally, these legitimate combinations (roads) are sent to the Track Fitter Farm where they are fed into a linear approximation fit to yield 2D tracks with impact parameter information. Figure 2.12 shows the impact parameter distribution from an early Run 2 data set. The data is from a single wedge of the SVX. Internal detector misalignments are not taken into account, resulting in a worse impact parameter resolution than is ultimately expected [29].

For rough comparison with the SVT resolution the $c\tau$ for some of the important mesons is given in Table 2.5. For the comparison recall that:

- A typical boost for a B meson would be $\gamma = 1.5$.
- The SVT measures impact parameter which is always smaller than flight distance.

---

1The Associative memory system comprises an Associative Memory Sequencer (AMS) which feeds the Associative Memory (AM) boards. The boards each have 128 AM chips. Each chip stores 128 allowed combinations.

2Note that the $D$ lifetimes are smaller than the B lifetimes. This means that the prompt charm component of the SVT bandwidth is smaller than the B component.
Figure 2.12: An early Run 2 SVT Impact Parameter resolution plot. The resolution here of 48\(\mu\)m is larger than the final SVT resolution because of the large uncertainty in beam spot position (30\(\mu\)m)[29]
Table 2.5: *Some of the important $c\tau$ (average lifetimes) in CDF Run II B-Physics* [4]. This is to be roughly compared to the predicted SVT resolution of 35$ \mu$m [27]. A typical B boost is $\gamma = 1.5$. Recall that impact parameter is mathematically less than flight distance.

<table>
<thead>
<tr>
<th>Meson</th>
<th>$c\tau \ (\mu m)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0_s$</td>
<td>$448 \pm 19$</td>
</tr>
<tr>
<td>$D^+_s$</td>
<td>$149 \pm 3$</td>
</tr>
<tr>
<td>$B^0_d$</td>
<td>$464 \pm 10$</td>
</tr>
<tr>
<td>$D^+$</td>
<td>$315 \pm 4$</td>
</tr>
<tr>
<td>$D^0$</td>
<td>$124 \pm 1$</td>
</tr>
</tbody>
</table>

2.4.3 Two Track Trigger

The Two Track trigger is composed of a recipe of cuts applied at Level 1 and Level 2 on combinations of 2 tracks. The most important cuts are applied using the information from the SVT described in section 2.4.2. Using the impact parameter measurement provided by the SVT, it is possible to reduce the trigger cross-section out of Level 2 to a sufficient level without losing all the signal. The reason for this is because the signal this trigger is designed to accept is primarily heavy flavour which has a characteristically long lifetime, and therefore has daughter particles which can leave high impact parameter tracks.

The definition and utility of the Two-Track trigger is discussed in [24]. A synopsis of this discussion now follows.

**Level 1 Definition**

At Level 1, there is a $p_t$ cut on each track in the 2 track pair, and on the scalar sum. There is also a cut on the opening angle of the two tracks (in the transverse plane). This angle is defined as the angle measured at Superlayer 6 in the COT (see section 2.3.2).

There are three operating scenarios for the Tevatron in Run II[30], and the Level 1 definition of the Two-Track trigger is different for each. These different configurations are defined in table 2.6.

In addition to the cuts defined in table 2.6 the requirement is imposed that
if more than 6 tracks pass the individual cuts the event automatically passes Level One (see section 2.4.1). This is to constrain the potential number of combinations that need to be considered.

**Level 2 Definition**

At Level 2 the SVT forms the most important part of the trigger. The information in the axial layers of SVX II is analysed in the manner discussed in section 2.4.2 to yield a list of tracks where three quantities are measured:

- $p_t$
- $\phi_0$ (as measured at the beamline)
- $d_0$

These quantities are discussed in sections A.1 and A.2. It should be noted that all three of these quantities are defined in the transverse plane here. There is no other way to define $p_t$ and $\phi_0$, but $d_0$ has a 3-dimensional definition. Here it is transverse.

The list of tracks from the SVT is sent to the Level 2 Alpha processors where the first thing that is done is that the Level 1 decision is verified. Then four separate classes of cuts are made:

1. A range cut on $\phi_0$.
2. A range cut on $d_0$ defined in the transverse plane.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$T_{bunch}$ (ns)</th>
<th>$\mathcal{L}$ ($\times 10^{32}cm^{-2}s^{-1}$)</th>
<th>$\langle N_{\overline{p}p} \rangle$</th>
<th>$p_{t_{\text{min}}}$ (GeV/c)</th>
<th>$\langle \Sigma p_t \rangle_{\text{min}}$ (GeV/c)</th>
<th>$\delta \phi_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>396</td>
<td>0.7</td>
<td>2</td>
<td>2.00</td>
<td>5.5</td>
<td>135°</td>
</tr>
<tr>
<td>B</td>
<td>132</td>
<td>2.0</td>
<td>2</td>
<td>2.25</td>
<td>6.0</td>
<td>135°</td>
</tr>
<tr>
<td>C</td>
<td>396</td>
<td>1.7</td>
<td>5</td>
<td>2.50</td>
<td>6.5</td>
<td>135°</td>
</tr>
</tbody>
</table>

Table 2.6: *Level One Two Track Trigger definition.* The columns are from left to right: The luminosity scenario label; The bunch crossing time; The luminosity; the average number of $\overline{p}p$ interactions per crossing; 1 track $p_t$ cut, 2 track combined $p_t$ cut; $\phi$ measured at COT super-layer 6 cut.
3. A range cut on the combined transverse $d_0$ of the track pair.

4. A Positive decay length cut.

Cuts 1 and 2 are self explanatory. The combined $d_0$ required for the 3rd cut is calculated by adding the vector transverse momenta of the two tracks in the pair (equation 2.2). This is combined with the intersection point of the pair to yield a pseudo-track which represents the transverse path of a potential particle which decayed into the two particles yielding the two track pair.

The quantity required for the 4th cut is found by calculating the sign of:

$$\overrightarrow{p_T} \cdot \overrightarrow{X_v}$$

(2.1)

where $\overrightarrow{p_T}$ and $\overrightarrow{X_v}$ are defined:

$$\overrightarrow{p_T} = p_t(track1) + p_t(track2)$$

(2.2)

$$\overrightarrow{X_v} = intersectionpoint - beamline$$

(2.3)

Where equation 2.3 defines the vector in the transverse plane pointing from the beamline to the intersection point.

The utility of this cut is that for generic QCD background the quantity defined in equation 2.1 will be measured positive and negative with equal probability. This is because the intersection point calculated for generic QCD background\(^3\) is solely due to the measurement error on the tracks. QCD background by its very nature emanates from the beam-line. However, for the heavy flavour component of interest, the intersection points calculated will represent true flight distance. Consider the diagrams in figures 2.13 and 2.14 for a graphical description of this cut.

In the figure 2.13 two possible cases for $\overrightarrow{p_T}$ and $\overrightarrow{X_v}$ are considered. Firstly, for a true heavy flavour case ($\overrightarrow{p_{T1}}$ and $\overrightarrow{X_{v1}}$) the hypothesised particle is decaying in front of the primary vertex. The combined momentum vector of 2 of its daughters is likely to be somewhat correlated to the decay vector. This stands in contrast to the generic QCD case where $\overrightarrow{p_{T2}}$ and $\overrightarrow{X_{v2}}$ are not at all correlated (any non-zero $|\overrightarrow{X_{v2}}|$ arises from the resolution of the detector). Therefore the dot-products calculated (see figure 2.14) will strongly

\(^3(u,d,s)\)
Figure 2.13: Example $X_v$ and $P_T$ for true heavy flavour and QCD background
\[ P_{T1} \cdot \vec{X}_{v1} \]

\[ \theta_1 < 90^\circ \Rightarrow P_{T1} \cdot \vec{X}_{v1} > 0 \]

\[ P_{T2} \cdot \vec{X}_{v2} \]

\[ \theta_2 > 90^\circ \Rightarrow P_{T2} \cdot \vec{X}_{v2} < 0 \]

Figure 2.14: Dot product signs for example heavy flavour and QCD background cases. The top diagram represents a real heavy flavour decay where the daughter particle direction is correlated to the parent direction. The bottom diagram represents a generic QCD prompt daughter particle where the supposed parent direction is really just due to detector smearing, and so uncorrelated to the daughter direction.
peak positive in the heavy flavour case, but uniformly positive and negative in the QCD background case.

The Level 2 Two-Track trigger is split into two physics cases:

1. $B^0 \rightarrow \pi^+\pi^-$

2. $B_s$ multi-body

The second of these two cases is the one very much in view for the studies in hand.

Table 2.7 describes the Level 2 trigger cuts for these two cases.

<table>
<thead>
<tr>
<th>$B_s$ multi-body</th>
<th>$B^0 \rightarrow \pi^+\pi^-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$120 \mu m &lt;</td>
<td>d_0</td>
</tr>
<tr>
<td>$2^\circ &lt; \Delta \phi_0 &lt; 90^\circ$</td>
<td>$20^\circ &lt; \Delta \phi_0 &lt; 135^\circ$</td>
</tr>
<tr>
<td>$\vec{p}_f \cdot \vec{X}_v &gt; 0$</td>
<td>$\vec{p}_f \cdot \vec{X}_v &gt; 0$</td>
</tr>
<tr>
<td>-</td>
<td>$d_B &lt; 140 \mu m$</td>
</tr>
</tbody>
</table>

Table 2.7: Level 2 Two Track trigger cuts

In the first case, the signal sought has a large opening angle and this is reflected in a looser $\phi_0$ range. The $d_0$ range can correspondingly be loosened. $d_B$ is the impact parameter of the pseudo track described earlier on in this section. The cut on this quantity is $< d_0 (cut)$ because the original $B^0$ comes from the primary vertex.\footnote{A $d_B$ from two signal tracks should be zero within experimental resolution}

For the second case, one important signal sought sometimes has a small opening angle (the $\phi$ decay has a very small Q-value). The reason there is a lower cut on the angle is because this reduces the trigger rate by a factor of 2. This is attributed to a high fake rate in dense hit environments [24]. The resulting inefficiency is not serious since the trigger decision doesn’t usually depend on the kaons from the $\phi$. This is because the kaons usually have lower momenta than the pions from the $B_s$ and $D_s^-$. The upper angle cut in the $B_s$ multi-body case is less than the $B^0 \rightarrow \pi^+\pi^-$ case because there is not as much phase space for the $B$-daughters to decay into in general for these sorts of decays. The non-trivial boost of the $B_s$ therefore means the lab-frame

\footnote{$B_s \rightarrow D_s^- \pi^+$ where $D_s^- \rightarrow \phi\pi^-$ and where $\phi \rightarrow K^+K^-$}
opening angles of the B-daughters won’t be as large as the $B^0 \rightarrow \pi^+\pi^-$ case.

Finally in this section a summary is given (tables 2.8 and 2.9) of the Two-Track trigger definitions for the two physics cases, $B^0 \rightarrow \pi^+\pi^-$ and $B_s$ multi-body for the scenario A (table 2.6) of Tevatron running.

<table>
<thead>
<tr>
<th>Trigger level</th>
<th>Cuts on each track</th>
<th>Cut on the pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>$pt &gt; 2 GeV$</td>
<td>$\Sigma pt &gt; 5.5 GeV, \Delta \phi &lt; 135^\circ$</td>
</tr>
<tr>
<td>L2</td>
<td>$100 \mu m &lt; d_0 &lt; 1 mm$</td>
<td>$20^\circ &lt; \Delta \phi &lt; 135^\circ$ &lt;br&gt;$p_T \cdot \vec{X}_v &gt; 0$ &lt;br&gt;$d_B &lt; 140 \mu m$</td>
</tr>
</tbody>
</table>

Table 2.8: $B^0 \rightarrow \pi\pi$ Two Track Trigger ($d_0$ is impact parameter)

<table>
<thead>
<tr>
<th>Trigger level</th>
<th>Cuts on each track</th>
<th>Cut on the pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>$pt &gt; 2 GeV$</td>
<td>$\Sigma pt &gt; 5.5 GeV, \Delta \phi &lt; 135^\circ$</td>
</tr>
<tr>
<td>L2</td>
<td>$120 \mu m &lt; d_0 &lt; 1 mm$</td>
<td>$2^\circ &lt; \Delta \phi &lt; 90^\circ$</td>
</tr>
</tbody>
</table>

Table 2.9: $B_s$ multi-body Two Track Trigger ($d_0$ is impact parameter)

### 2.4.4 Level 3

The final level of the trigger will have various roles in Run II. Its initial and probably most important role will be to clean up the results from Level 2.

Level 3 consists of a farm of Linux PCs and a very fast ATM\textsuperscript{6} switch [31]. The current operating parameters for Level 3 are given in table 2.10.

These operating parameters mean that there is sufficient time at Level 3 ($\sim 1$ CPU second) to do sophisticated offline-like reconstruction. So the full SVX II detector including the stereo layers will be used to track yielding 3-dimensional track measurements.

The overall rejection factor of Level 2 combined with the clean-up effect of verifying this information at Level 3 will be high. To begin with it is envisaged that this rejection will be sufficient so that the entire bandwidth

\textsuperscript{6}Asynchronous Transfer Mode
<table>
<thead>
<tr>
<th>L3 Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Nodes</td>
<td>130</td>
</tr>
<tr>
<td>Specification of Nodes</td>
<td>Pentium 3 Dual 550 → 800 MHz</td>
</tr>
<tr>
<td>Input rate from Level 2</td>
<td>300 → 1000Hz</td>
</tr>
<tr>
<td>Event size</td>
<td>\sim 250Kb</td>
</tr>
<tr>
<td>Output rate to tape</td>
<td>\sim 30 → 75Hz</td>
</tr>
</tbody>
</table>

Table 2.10: Current Level 3 Operating parameters

out of Level 3 can be written to tape. However, as the run progresses, and
the instantaneous luminosity goes up, it may well become necessary to begin
making more sophisticated cuts at Level 3 in order to bring the bandwidth
down to an acceptable level.

2.4.5 Conclusion

The CDF detector has now been surveyed, with a particular emphasis on
the elements important for a measurement of $\Delta \Gamma_{B_s}$. The next chapter gives
an introduction to how one might go about measuring $\Delta \Gamma_{B_s}$ in CDF Run
II. The subsequent chapters then detail studies which have been carried out
which are relevant to such a measurement.
Chapter 3

Experimental Techniques for measuring $\Delta \Gamma_{B_s}$

In the following chapter, the experimental techniques that will be employed to measure $\Delta \Gamma_{B_s}$ in CDF Run II will be discussed. The primary focus of the present studies is introduced, and then the wider experimental context of these studies is discussed. Finally, the specific techniques used in the present studies are described.

This work focuses on the effect of the level 2 trigger on the observed $B_s$ lifetime distribution measured in the mode $B_s \rightarrow D_s \pi$. The context of this analysis is the measurement of the quantity $\Delta \Gamma_{B_s}$ introduced in chapter 1. The experimental techniques used are those pertaining to a lifetime measurement.

3.1 Sculpting the Observed $B_s$ lifetime distribution

The core of a $\Delta \Gamma_{B_s}$ measurement is a lifetime analysis. Many of the samples to be used for this will come in on the two track hadronic trigger and therefore are subject to the level 2 impact parameter cut. Recall from chapter 2 that for the $B_s$ multi-body stream of the trigger, this cut is $(120 \mu m < d_0 < 1000 \mu m)$. This cut severely affects the observed $B_s$ lifetime distribution. Figure 3.1 shows the generator level proper time before level 1. Figure 3.2 shows the generator level proper time after level 2. Before level 1, the distribution
is almost precisely an exponential with the same lifetime as the input MC value of 1.551 ps\textsuperscript{1}. However, after level 2, the true $B_s$ decay exponential has been seriously distorted. The central value of the exponential fit even over a partial proper time range is completely different to the monte-carlo input value (1.149 ± 0.009 ps compared to 1.551 ps).\textsuperscript{2} Also, the $\chi^2$ of the fit is bad. For 50 d.o.f it is 76 so that only 1% of similar data sets would have a worse fit.

\textsuperscript{1}The small systematic difference from 1.551 ps is due to assuming a $B_s$ mass of 5.3693 GeV in the calculation of the generator level proper time. The correct mass was 5.38 GeV. This introduces a small but systematically negative effect on the lifetime. This mistake was not made for the reconstructed proper time.

\textsuperscript{2}Note the same small systematic mistake due to the mass assumption is present for this distribution, but is tiny compared to the systematic shift of the central value due to the trigger effects (1%).

Figure 3.1: The generator level $B_s$ proper time before Level 1. The fitted value for the lifetime is consistent with the input MC lifetime value of 1.551 ps (with a small systematic shift due to assuming slightly the wrong mass).
Figure 3.2: The generator level $B_s$ proper time after Level 2. Not only does one exponential not fit over the entire range, but it fits badly, and to the wrong value. The $\chi^2$ is 76 for 50 d.o.f, meaning 1% of such plots have a worse $\chi^2$.

3.2 Two different ways to measure $\Delta \Gamma_{B_s}$

As presented in section 1.2.1 $\Delta \Gamma_{B_s}$ is defined:

$$\Delta \Gamma_{B_s} = \Gamma_H - \Gamma_L$$  \hspace{1cm} (3.1)

where $\Gamma_H$, $\Gamma_L$ are the widths of the heavy and light Mass eigenstates of the $B_s$ meson.

It is a subtlety in the most general framework that these mass eigenstates are not the same as the CP eigen-states. However, to a rather good approximation they are. The validity of this approximation is discussed further in section 3.3.1. From here on in this chapter, the mass and CP eigenstates will be used interchangeably, except in section 3.3.1 where the particular case of $B_s \rightarrow D_s \pi$ is discussed. $^3$

A measurement of $\Delta \Gamma_{B_s}$ thus consists of measuring the difference in lifetime between the two mass eigen-states, which are to a very good approximation the CP eigen-states.

A useful quantity to define is the fractional lifetime difference, $\frac{\Delta \Gamma_{B_s}}{\Gamma_{\text{ave}}}$:

$^3$Recall from the text after equation 1.29 that this comes from the approximation $|\xi| \approx 1$
\[
\frac{\Delta \Gamma_{B_s}}{\Gamma_{\text{ave}}} = \frac{\Gamma_H - \Gamma_L}{\Gamma_{\text{ave}}}, \quad \Gamma_{\text{ave}} = \frac{\Gamma_H + \Gamma_L}{2}
\] (3.2)

The most obvious way of measuring this lifetime difference would be to find a mode which contains both CP eigenstates of the $B_s$ in approximately equal proportion (e.g., $B_s \rightarrow D_s \pi$), and directly fit for two different exponential decays. If the lifetime difference between the CP eigenstates were large then with enough events in a data sample one might be able to successfully fit for two separate exponentials. However, there are two big problems with this.

Firstly, even if $\frac{\Delta \Gamma_{B_s}}{\Gamma_{\text{ave}}}$ were large enough ($\sim 0.5$), such a two exponential fit is still very unstable and the number of events required to make it stable would be more than CDF Run II will see in the first few years\(^4\). In fact during the course of attempting to parameterise the background of the $B_s \rightarrow D_s \pi$ sample\(^5\) a similar two exponential fit was attempted.\(^6\)

Secondly, given the instability of such a two exponential fit when $\frac{\Delta \Gamma_{B_s}}{\Gamma_{\text{ave}}}$ is small, it is a huge problem for the two exponential method that the theoretical prediction for $\frac{\Delta \Gamma_{B_s}}{\Gamma_{\text{ave}}}$ is $\lesssim 0.15$. This prediction is fairly robust since independent of the model used, it is possible to show that the maximum value of $\frac{\Delta \Gamma_{B_s}}{\Gamma_{\text{ave}}}$ cannot be larger than a number similar to this [14].

The experimental solution to the quandary outlined above is to find some way of isolating one or both of the two CP eigenstates. One way to do this is to find a mode which is purely one CP eigenstate (e.g., $B_s \rightarrow D_s^+ D_s^-$). Another way is to find a variable which discriminates between the two different eigenstates in a CP mixed mode. An example of this latter approach would be the mode $B_s \rightarrow J/\psi \phi$. Here angular momentum considerations imply that the angular distribution of the decay products is sensitive to the CP of the mode.

In both of the cases of $B_s \rightarrow D_s^+ D_s^-$ and $B_s \rightarrow J/\psi \phi$, only one CP eigenstate is isolated. This is evident in the case of $B_s \rightarrow D_s^+ D_s^-$ since it

\(^4\)There is an alternative analysis which uses a fit to one exponential, but it only yields information about $(\Delta \Gamma)^2$\(^{[32]}\)

\(^5\)see section 6.2.3

\(^6\)For example, some simulation was fitted with a 2 exponential function with $\lambda_1 = 1.24 \text{ps}$, $\lambda_2 = 0.64 \text{ps}$ and a 1:1 relative abundance. A further toy set of simulation was generated with these parameters, but they were not recovered in the subsequent fit. This is indicative of an unstable functional form. Some other way of separating the exponentials is needed.
CHAPTER 3. EXPERIMENTAL TECHNIQUES FOR MEASURING $\Delta \Gamma_{B_S}$

is purely CP even (see 3.3.3). The mode $B_s \rightarrow J/\psi \phi$ has the freedom to have either CP eigenvalue (see 3.3.2), but a Run I measurement at CDF [33] has shown that this mode is largely CP even.

Therefore, if either of these modes are employed for a measurement of $\frac{\Delta \Gamma_{B_s}}{\Gamma_{ave}}$, a second mode will be required to yield the information about the other CP eigenstate. For example, given that $B_s \rightarrow D_s \pi$ has each CP eigen-state in equal abundance, in conjunction with a measurement of $\Gamma_L$, $\frac{\Delta \Gamma_{B_s}}{\Gamma_{ave}}$ would follow as:

$$\frac{\Delta \Gamma_{B_s}}{\Gamma_{ave}} = \frac{\Gamma_H - \Gamma_L}{\Gamma_{ave}} = \frac{\Gamma_H + \Gamma_L - 2\Gamma_L}{\Gamma_{ave}}$$

(3.3)

$$\therefore \frac{\Delta \Gamma_{B_s}}{\Gamma_{ave}} = 2 \left( 1 - \frac{\Gamma_L}{\Gamma_{ave}} \right)$$

and in the approximation:

$$\Gamma_{B_s \rightarrow D_s \pi} = \frac{\Gamma_H + \Gamma_L}{2} = \Gamma_{ave}$$

(3.4)

it follows that:

$$\frac{\Delta \Gamma_{B_s}}{\Gamma_{ave}} = 2 \left( 1 - \frac{\Gamma_L}{\Gamma_{B_s \rightarrow D_s \pi}} \right)$$

(3.5)

(Note: the approximation in equation 3.4 is good in the absence of direct CP Violation in this mode, which is expected to be very small).

Using one sample for isolation of one CP eigen-state, and combining this measurement with another measurement in a mixed CP mode will be called the Two Sample Method in these studies.

### 3.3 The Two Sample Method

There are four main modes which will be used early in Run II in conjunction with the two sample method just described in order to measure $\frac{\Delta \Gamma_{B_s}}{\Gamma_{ave}}$. These are:

1. $B_s \rightarrow D_s \pi$
2. $B_s \rightarrow D_s \pi \pi \pi$
3. $B_s \rightarrow J/\psi \phi$

4. $B_s \rightarrow D_s^+ D_s^-$

1 and 2 are used for the $\Gamma_{\text{ave}}$ measurement, while 3 and 4 are used to isolate one CP eigen-state.

While these four modes are the main modes of interest, there are also other modes which may prove useful (eg $B_s \rightarrow D_s^{*+} D_s^{*-}$). These modes will be discussed briefly in section 3.5.

### 3.3.1 $B_s \rightarrow D_s \pi$ and $B_s \rightarrow D_s \pi \pi$

The mode $B_s \rightarrow D_s \pi$ and its counterpart $B_s \rightarrow D_s \pi \pi$ are the most interesting modes for the $\Gamma_{\text{ave}}$ measurement. The other modes discussed in section 3.5.1 are less interesting because they are subject to larger systematic uncertainties. In the studies in the following chapters, only $B_s \rightarrow D_s \pi$ is considered. One reason for this is that the mode $B_s \rightarrow D_s \pi \pi$ has some disadvantages\(^7\) when compared to $B_s \rightarrow D_s \pi$.

**CP composition**

Since these modes are flavour specific it is always possible to tell from which of the strong interaction eigen-states each one arises\(^8\). For the case of $B_s \rightarrow D_s \pi$:

$$B_s \rightarrow D_s^- \pi^+$$

$$\bar{B}_s \rightarrow D_s^+ \pi^-$$

Note that the charge of the decay mode indicates the flavour of the $B$-meson at decay. This can be different to the flavour at creation because $B_s$ mixing is a large effect.

From section 1.2.1 recall that:

$$|B_L\rangle = p|B\rangle + q|\bar{B}\rangle$$

$$|B_H\rangle = p|B\rangle - q|\bar{B}\rangle$$

\(^7\)Primarily the 3 $\pi$ mesons have less phase space than the single $\pi$ in $B_s \rightarrow D_s \pi$.

\(^8\)It is possible to tell from the charge of the $D_s$ which flavour of $B_s$ ($b$ or $\bar{b}$) the $D_s$ comes from.
CHAPTER 3. EXPERIMENTAL TECHNIQUES FOR MEASURING $\Delta \Gamma_{B_s}$

This means that:

$$|B\rangle = \frac{1}{2p}(|B_L\rangle + |B_H\rangle)$$

$$|\overline{B}\rangle = \frac{1}{2q}(|B_L\rangle - |B_H\rangle)$$

Therefore, it is expected that any mode which is constrained to come from a $B_s$ or a $\overline{B_s}$ (eg by its charge) will be an equal admixture of $B_H$ and $B_L$.

However, this expectation breaks down in the case of direct CP-Violation. Now this sort of CP-Violation requires two amplitudes leading to the same final state ($D_s\pi$ in this case) to interfere. Therefore in the case of $B_s \rightarrow D_s\pi$ where the amplitude is completely dominated by the tree-level contribution, direct CP-Violation is expected to be very small.

It will therefore be assumed in the following studies that $B_s \rightarrow D_s\pi$ has exactly 50% $B_L$, 50% $B_H$ content.

Sample Sizes

The expected sample sizes (reconstructible events) for $B_s \rightarrow D_s\pi$ and $B_s \rightarrow D_s\pi\pi\pi$ are given in [34], and are summarised in table 3.2.

<table>
<thead>
<tr>
<th>$B_s$ decay</th>
<th>$D_s$ decay</th>
<th>$N_{Run, H}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_s \rightarrow D_s\pi$</td>
<td>$D_s \rightarrow \phi\pi$</td>
<td>16100</td>
</tr>
<tr>
<td>$B_s \rightarrow D_s\pi$</td>
<td>$D_s \rightarrow K^* K$</td>
<td>15300</td>
</tr>
<tr>
<td>$B_s \rightarrow D_s\pi$</td>
<td>$D_s \rightarrow \pi\pi\pi$</td>
<td>5500</td>
</tr>
<tr>
<td>$B_s \rightarrow D_s\pi$</td>
<td>$D_s \rightarrow$ all above</td>
<td>36900</td>
</tr>
<tr>
<td>$B_s \rightarrow D_s\pi\pi\pi$</td>
<td>$D_s \rightarrow \phi\pi$</td>
<td>15100</td>
</tr>
<tr>
<td>$B_s \rightarrow D_s\pi\pi\pi$</td>
<td>$D_s \rightarrow K^* K$</td>
<td>17000</td>
</tr>
<tr>
<td>$B_s \rightarrow D_s\pi\pi\pi$</td>
<td>$D_s \rightarrow \pi\pi\pi$</td>
<td>6200</td>
</tr>
<tr>
<td>$B_s \rightarrow D_s\pi\pi\pi$</td>
<td>$D_s \rightarrow$ all above</td>
<td>38300</td>
</tr>
</tbody>
</table>

Table 3.1: Expected sample sizes for the most important Mixed CP $B_s$ decay channels ($2 fb^{-1}$)[34] (Note: $B_s$ and $\overline{B_s}$ yields are taken into account)
The numbers in the table above are for $2fb^{-1}$ of running which corresponds to the expected integrated luminosity for the first year of running. The numbers summarise the expected sample size for all the statistically significant $D_s$ decay modes. Further the factor of two implicit in the fact that for every $B_s$ decay there is a corresponding $\overline{B}_s$ decay is taken into account.\footnote{The events are also required to have the signal tracks inside the fiducial volume.} It will be noted that some of the branching ratios necessary for predictions of the sort given in table 3.2 have not been measured (for example, $B_s \rightarrow D_s \pi$ and $B_s \rightarrow D_s \pi \pi \pi$ have not been observed). Where this is the case, the branching ratios are inferred from their $B_d$ counterparts ($B_d \rightarrow D \pi$ and $B_d \rightarrow D \pi \pi \pi$ in this case.)

**Trigger Strategy**

Since this mode contains no leptons, the only trigger it can come in on in CDF Run II is the two-track hadronic trigger (section 2.4.3). As has been noted already, the most important new aspect of this trigger is the impact parameter cut at Level 2. However, this impact parameter cut seriously affects the lifetime distribution observed offline. It will become clear in the following chapters that the systematic effect on the lifetime distribution of the impact parameter cut at Level 2 is the main emphasis of the present studies\footnote{See section 3.1 for a preview of the discussion.}.

**3.3.2 $B_s \rightarrow J/\psi \phi$**

This mode will be one of the first employed to isolate the CP even eigen-state of the $B_s$.

**CP Composition**

This mode stands in contrast to the flavour specific modes discussed in section 3.3.1. It is not required to originate purely from one of $B_s$ or $\overline{B}_s$. Therefore it could in principle be a CP eigenstate. However, this does not take into account the effect of relative orbital angular momentum between the $J/\psi$ and the $\phi$. The initial state $B_s$ is a Pseudo-scalar (spin 0, negative parity). However the final state consists of two Vector particles (spin 1). Consider the diagram in figure 3.3.
CHAPTER 3. EXPERIMENTAL TECHNIQUES FOR MEASURING $\Delta \Gamma_{B_s}$

![Diagram showing relative orbital angular momentum in $B_s \rightarrow J/\psi \phi$](image)

Figure 3.3: Relative orbital angular momentum in $B_s \rightarrow J/\psi \phi$ (between the $J/\psi$ and the $\phi$

The two possibilities for two Spin 1 particles plus an orbital angular momentum to add up to zero total angular momentum are shown. Since $L$ can be either 0 or 1, the parity of the state can be odd or even. Hence it can be either odd or even CP. A Run 1 measurement has been carried out to measure the fraction of CP odd in this mode. The result was found [33] to be $\frac{L_{\text{CP}}}{L_{\text{tot}}} = 0.229 \pm 0.188(\text{stat}) \pm 0.038(\text{syst})$. This angular momentum structure allows the CP eigenstates to be separated. By forming the appropriate angular distributions of the decay products, it is possible to do the isolation necessary to measure the lifetime of the CP eigenstates.\(^{11}\) As the Run I measurement indicates [33], it is likely that one CP eigenstate dominates $B_s \rightarrow J/\psi \phi$. This means one of the weak interaction eigenstate lifetimes will be measured with more accuracy than the other. This mode is therefore more useful in conjunction with another mode (ie in the 2 sample method) than on its own.

**Sample Size**

The expected yield of this mode is approximately 4000 events\(^{35}\). A possible $J/\psi \rightarrow e^+e^-$ trigger could add another 3000 events, while reconstructing $\psi(2S)$ through the decay modes $\psi(2S) \rightarrow \mu^+\mu^-$ and $\psi(2S) \rightarrow J/\psi \pi^+\pi^-$ could add another 600 events. These estimates are based on Run I yields\(^{36}\).

\(^{11}\)Note that this CP eigenstate separation is not possible on an event by event basis, since the information about the CP eigenstates is only accessible from an angular distribution. The separation is therefore statistical by nature.
CHAPTER 3. EXPERIMENTAL TECHNIQUES FOR MEASURING $\Delta \Gamma_{B_s}$

Trigger Strategy

Because of the $J/\psi$, this mode does not primarily come in on the 2 track trigger. Instead the bulk of the rate for this mode will come from the dimuon trigger, with some events coming from the supplementary triggers mentioned in the last section. Consequently the lifetime distribution of the $B_s$ is largely free from trigger effects.

3.3.3 $B_s \rightarrow D_s^+ D_s^-$

This mode is a very clean $\frac{\Delta \Gamma_{B_s}}{\Gamma_{B_s}^{ave}}$ mode, since it is entirely CP even$^{12}$. Like $B_s \rightarrow D_s \pi$ it is Pseudoscalar $\rightarrow$ 2 Pseudoscalars. But unlike $B_s \rightarrow D_s \pi$, it has definite CP because one daughter particle is the charge conjugate of the other (the C operation).$^{13}$ The only disadvantage with this mode is that the number of events obtained will be severely limited. The expected yields [34] are given in table 3.2.

<table>
<thead>
<tr>
<th>$B_s$ decay</th>
<th>$D_s$ decay</th>
<th>$N_{RunII}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_s \rightarrow D_s^+ D_s^-$</td>
<td>$D_s^+ \rightarrow \phi \pi^-$</td>
<td>370</td>
</tr>
<tr>
<td>$B_s \rightarrow D_s^+ D_s^-$</td>
<td>$D_s^- \rightarrow K^{*0} K^-$</td>
<td>430</td>
</tr>
<tr>
<td>$B_s \rightarrow D_s^+ D_s^-$</td>
<td>$D_s^- \rightarrow \pi^- \pi^- \pi^+$</td>
<td>70</td>
</tr>
<tr>
<td>$B_s \rightarrow D_s^+ D_s^-$</td>
<td>$D_s^- \rightarrow$ all above</td>
<td>870</td>
</tr>
</tbody>
</table>

Table 3.2: Expected sample sizes for $B_s \rightarrow D_s^+ D_s^-$ ($2fb^{-1}$)[34]

3.4 The specifics of a lifetime measurement

The following sections give an overview of the experimental techniques employed in measuring the lifetime distribution of a given particle. The main steps are:

1. Candidate selection

$^{12}$As with $B_s \rightarrow D_s \pi$ this only holds in the absence of direct CP Violation

$^{13}$Angular momentum considerations also influence the CP.
2. Lifetime reconstruction

3. Likelihood fit to an exponential

In the following, the discussion will be confined to the modes of the $B_s$ which come in on the two-track hadronic trigger (e.g. $B_s \rightarrow D_s \pi$ or $B_s \rightarrow D_s^+ D_s^-$).

The particular emphasis will be on $B_s \rightarrow D_s \pi$ since this is the mode which has been reconstructed in the present studies.

### 3.4.1 Candidate selection

It is expected that the bulk of the bandwidth resultant from the SVT (see section 2.4.2) will actually be B-mesons. The selection of candidates thus comes down to selecting one B-meson decay mode over another one.

The tracking at CDF is the most powerful detector element for detecting the decay products of these B-mesons. Therefore, in a given attempt to reconstruct a B decay mode, a set of tracks is decided on as *well measured*, and then each track in this set is assumed to be a candidate in the decay in question. There is an obvious decision that can be taken regarding the charge of the particle, but the mass for example is hypothesised, not measured. Consider the example of $B_s \rightarrow D_s \pi$. One especially clean version of this decay mode is:

$$B_s \rightarrow D_s^- \pi^+ \text{ where } D_s^- \rightarrow \phi \pi^- \text{ and } \phi \rightarrow K^+ K^-$$  \hspace{1cm} (3.9)

There are four final state particles in this case, two pions and two kaons. A procedure for reconstruction the full $B_s$ decay is then\textsuperscript{14}:

- Loop over all negative tracks assuming the kaon mass.
- Within this loop, iterate over all positive tracks assuming the kaon mass.
- Form the invariant mass of the $\phi$.
- Within the $K^+$ loop, iterate over all other (i.e. not counting the $K^-$ track) negative tracks assuming the $\pi$ mass.

\textsuperscript{14} Note that this procedure is not strictly valid for $\overline{B_s}$, which would require reversal of the charges on the pions
• Form the invariant mass of the $D_s^-$

• Within the $\pi^-$ loop, iterate over all other (ie not counting the $K^+$ track) positive tracks assuming the $\pi$ mass.

• Form the invariant mass of the $B_s$.

The obvious cuts to make in the above procedure are on the invariant mass of the various intermediate particles. At this early stage it is wise to make these fairly wide (about 200 MeV), since there are more cuts to come.

### Basic Kinematic Cuts

After this most basic procedure has been done, the subsequent procedure consists of making more kinematic cuts on the various particles. There are a very large number of possibilities here. Consider the cuts listed in table 3.3.

<table>
<thead>
<tr>
<th>Particle</th>
<th>$p_t$</th>
<th>$d_0$</th>
<th>$L_{xy}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi$ from $B_s$</td>
<td>X</td>
<td>X</td>
<td>N/A</td>
</tr>
<tr>
<td>$\pi$ from $D_s$</td>
<td>X</td>
<td>X</td>
<td>N/A</td>
</tr>
<tr>
<td>$K^+$ from $\phi$</td>
<td>X</td>
<td>X</td>
<td>N/A</td>
</tr>
<tr>
<td>$K^-$ from $\phi$</td>
<td>X</td>
<td>X</td>
<td>N/A</td>
</tr>
<tr>
<td>$B_s$</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>$D_s$</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>$\phi$</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 3.3: List of basic kinematic cuts for $B_s \rightarrow D_s\pi$ reconstruction. $X$ indicates the cut is possible for this particle, N/A indicates that it is not applicable (see appendix A.2 for definition of $d_0, L_{xy}$).

However, it is important not to go overboard here. There are a number of cuts in the list in table 3.3 which distort the lifetime distribution. For example, one of the main systematic effects in measuring the $B_s$ lifetime in the mode $B_s \rightarrow D_s\pi$ is the effect of the $d_0$ cut in the level 2 trigger. As has already been noted, this forms the emphasis of the present studies. It would be wrong therefore, having gone to all the trouble of removing the effects of the trigger $d_0$ cut to then introduce another 4 cuts on $d_0$ in the
reconstruction phase! A similar observation can be made for the $L_{xy}$ cuts. In fact, for $B_s \rightarrow D_s \pi$, a sufficient signal to background ratio was obtained without needing to cut on any of the basic kinematic quantities in table 3.3 except for the $p_t$ of the daughter particles$^{15}$. However, in addition to the basic kinematic quantities there are certain more complicated physics cuts that are extremely powerful.

**Vertex Cut**

The most obvious more complicated physics cut to make for a B decay is a vertex cut. The secondary decay point of the hypothesised B meson reconstructed (this was only possible with the advent of vertex detectors), and the probability that the daughter tracks do in fact pass through a common vertex is formed. This probability is then used as the cut variable. At CDF, there are several software options for calculating whether tracks are likely to pass through common vertices, but probably the most useful one for a lifetime analysis is a piece of software called CTVMFT [37]. The input data to CTVMFT are the track parameters and associated errors$^{16}$. A combination of physical constraints is applied to improve the initial track parameters. The most useful physical constraint to apply as already mentioned is requiring some of the particles to pass through a common 3D vertex. However there are other physical constraints that one can require in the fit. In the present studies, the only one of these extra constraints used was a pointing constraint. This is described in figure 3.4. The way CTVMFT works is that it changes the input track parameters according to their errors, and picks the best set consonant with the requirements with which it has been initialised. The statistical method is a $\chi^2$ fit.

It is important to realise that for a given topology there could be more than one secondary vertex (there is always the primary vertex). In the case of $B_s \rightarrow D_s \pi$ for example, the $D_s$ flies some appreciable distance from the $B_s$ decay vertex before it decays$^{17}$. In the configuration used for $B_s \rightarrow D_s \pi$ in the present studies there were three vertices, the primary, the $B_s$ decay vertex and the $D_s$ decay vertex.

---

$^{15}$There will also be mass cuts on the $B_s$, $\phi$ and $D_s$.

$^{16}$See section A.1

$^{17}$c$\tau = 149 \mu m$, and a typical boost is $\gamma \sim 1.5$. $D_s$ flight distances from decay are thus $\sim 200 \mu m$
CHAPTER 3. EXPERIMENTAL TECHNIQUES FOR MEASURING $\Delta \Gamma_{B_s}$

Direction of the Hypothesised $B_s$
Formed from $D_s$ and $\pi$ paths
Constrained to point to the Primary

Hypothesised $\pi$ path
Hypothesised $D_s$ path

Primary Vertex

Figure 3.4: Graphical description of a pointing constraint for the vertex algorithm CTVMFT. Here the track parameters of the $D_s$ daughter tracks and the $B_s$ $\pi$ are varied to make the hypothesised $B_s$ trajectory point back to the Primary

The probability distribution\textsuperscript{18} that CTVMFT yields has a shape as shown in figure 3.5 for $B_s \rightarrow D_s \pi$ signal monte carlo. In this case, a majority of the events contain a reconstructible set of vertices which fit into the CTVMFT configuration assumed. Figure 3.6 shows a corresponding plot of the vertex probability for generic B events where slightly less events have a reconstructible vertex which fits the assumptions given to CTVMFT. This may be seen from the fact that the spike at zero probability is proportionally smaller for the signal monte carlo than for the generic B monte carlo. The flat distribution which takes up most of the ordinate axis in figures 3.5 and 3.6 is due to track combinations which did vertex. It makes sense therefore to cut on the vertex probability at a value which is quite small\textsuperscript{19}.

Furthermore, given the improved track parameters from the vertex fit, it is natural at this point to apply invariant mass cuts where possible. All three of the $B_s$, $D_s$ and $\phi$ have sufficiently narrow masses to make these cuts useful\textsuperscript{20}.

\textsuperscript{18}This is the calculated probability that the input track combination conformed to the vertex fit constraints
\textsuperscript{19}0.01 is usually chosen (the width of the 1st bin in figures 3.5 and 3.6).
\textsuperscript{20}Note, in reconstructing the mode, loose versions of these mass cuts are made on pre-vertex-fit track parameters to throw out combinatorics. Optimised cuts come later.
Figure 3.5: Sample Vertex Probability distribution from $B_s \rightarrow D_s \pi$ signal monte-carlo. The $K^+$, $K^-$ and $\pi^-$ have been constrained to pass through a common vertex (the $D_s^{-}$ vertex). The resulting track and the $\pi^+$ have been constrained to pass through a separate common vertex (the $B_s$ vertex), and the resulting $B_s$ trajectory has been constrained to point at the primary.

Figure 3.6: Sample Vertex Probability distribution from generic $B$ monte-carlo. The vertex algorithm constraints are identical to the upper figure.
Helicity Cut

There are a class of cuts which can be made for certain modes which rely on more complicated physics quantities than the vertex probability and mass cuts. An example of such a cut for $B_s \rightarrow D_s \pi$ is one made on the angular distribution of one of the Kaons relative to the $D_s$ direction in the rest frame of the $\phi$ meson.\textsuperscript{21} This quantity is illustrated in figure 3.7.

![Diagram](image)

Figure 3.7: The angle $\theta$ is measured in the rest frame of the $\phi$ relative to the direction of the $D_s$ in this frame.

The initial total angular momentum before the decay of the $D_s$ is $J = 0$. Since the $\phi$ has spin $S = 1$ this means that the $\phi$ and $\pi$ must have a relative orbital angular momentum of $L = 1$. Furthermore, because this is a 2 body decay of the $D_s$, the total relative orbital angular momentum can have no component along the direction of the decay in the $D_s$ rest frame. This means that the $\phi$ is polarised perpendicular to the flight axis in the $D_s$ rest frame. Its helicity ($h = \frac{\vec{S} \cdot \vec{L}}{|\vec{L}|}$) is therefore zero. Helicity is a Lorenz invariant and reduces to $S_z$ (the projection of spin onto the preferred axis) in the rest frame of the $\phi$. Therefore, in the decay of the $\phi$, $J = 1$, and $J_z = 0$. These quantities are conserved, and so the angular distribution of the decay amplitude is given

\textsuperscript{21}Recall the decay chain: $B_s \rightarrow D_s \pi$ where $D_s \rightarrow \phi \pi$ and $\phi \rightarrow K^+ K^-$
Figure 3.8: The angular distribution of the $K^-$ relative to the $D_s$ direction in the $\phi$ rest frame. The $\frac{d\sigma}{d\cos \theta} \propto \cos^2 \theta$ structure can clearly be seen, arising from the polarisation of the $\phi$. The $\chi^2$ given indicates 23.7% of such fits have a worse $\chi^2$.

by the d-function $d_{J^P,J_z}^I = d_{00}^1 = \cos(\theta)$. The angular dependence of the differential cross-section is thus $\frac{d\sigma}{d\cos \theta} \propto \cos^2 \theta$. A histogram versus $\cos \theta$ for the signal can be seen in figure 3.8.

### 3.4.2 Lifetime reconstruction

The next stage in measuring the lifetime distribution is to reconstruct the proper lifetime of the particle in question. In order for this to be done, the vertex of the particle must be reconstructed, but this has already been done in the course of candidate selection (section 3.4.1). Reconstructing the proper lifetime thus consists of taking the results of the vertex fit and forming the
quantity:

\[
\tau_{\text{proper}} = \frac{(L_{xy} \cdot \vec{p}_t) M_B}{|p_t|}
\]  

(3.10)

where all three quantities are those of the B-meson, and are calculated from the fit. The quantity \( L_{xy} \) is defined in section A.2. The reason why 2 dimensional quantities are used instead of 3 dimensional quantities is because the resolution on the 2D flight distance is much better than its 3D counterpart. This is true even in CDF Run II (see section 2.3.3), and especially at the start of the run, where the first efforts will go into trying to understand 3D silicon tracking.

The reason the dot product with the unit vector \( \vec{p}_t \) is used is because the \( p_t \) direction is much better measured than the \( L_{xy} \) direction. Consider the diagram in figure 3.9.

![Diagram](image)

Figure 3.9: Reason for including \( p_t \) unit vector in lifetime definition. The direction of the \( p_t \) is much better measured than the direction of \( L_{xy} \).

Diagram A shows a likely signal event where the \( p_t \) of the B meson is headed in the same direction as the 2D decay vector \( L_{xy} \). In this case the dot product is positive, leading to a positive proper lifetime of the correct magnitude. In contrast, diagram B shows a possible background where the \( L_{xy} \) is fake, and is in fact due to the measurement resolution on the vertex. In this case, the direction of \( L_{xy} \) is entirely uncorrelated to the direction of \( p_t \), and the situation in diagram B can easily arise. In this case the dot product is negative, leading to a negative proper lifetime. The dot product in the
proper time definition therefore separates out poorly measured $L_{xy}$ events.\footnote{Note that events where the $L_{xy}$ is poorly measured, but not badly enough to make the dot product negative will also be separated. The dot product in this case will weight the lifetime towards zero, and these events will form a short lived positive lifetime background.}

As an aside, it is the entry of $p_t$ into the proper time calculation which makes the lifetime resolution worse for $B_s \rightarrow D_s l \nu X$. Here, the vertex is reconstructible (the $D_s$ and the $l$ meet at the $B_s$ vertex). However the measured $p_t$ is less than the actual $p_t$ by an uncertain amount because it is impossible to reconstruct the $\nu$.

### 3.4.3 Likelihood fit to an exponential

The final stage in the measurement of the proper time distribution of a given particle is a likelihood fit to an exponential distribution. Usually this fit is unbinned (see section 6.1) since there are usually zero bins which would need to be handled carefully in a binned fit. Also, an unbinned fit avoids any dependence of the result on bin size.

It is at this stage that the analysis runs straight into the systematic effect which forms the emphasis of the current study: the effect of the level 2 impact parameter cut. The proper time distribution obtained from signal monte-carlo without taking the trigger into account bears little resemblance to an exponential (figure 3.2). A level 2 efficiency as a function of proper time must be calculated and applied if the original exponential is to be regained. These effects will be studied in detail in chapters 4, 5 and 6. For now, the general character of the fit will be described.

The likelihood function must take account of the signal and the background in a given sample. It is also highly advantageous if there is a way to put a quantity that discriminates between the signal and background into the fit. Such a quantity is the invariant mass distribution of the $B$ in question. The background for this quantity is effectively flat over the small invariant mass window resultant from the candidate selection, but the signal is normally distributed according to a gaussian due to the measurement resolution of the detector. The general form of one term of the likelihood function is thus:

$$
\mathcal{L}_i = F_i^{\text{sig}} \times (M_i^{\text{sig}} T_i^{\text{sig}}) + (1 - F_i^{\text{sig}}) \times (M_i^{\text{bck}} T_i^{\text{bck}}) \quad (3.11)
$$

The quantities sub-scripted with $i$ indicate quantities which change in
each minimisation stage in the fit (parameters in the fit). The fraction of signal in the sample, $F^{\text{sig}}$, is fitted, and the fraction of background is constrained by this. The factor $T^{\text{sig}}$ contains the efficiency function mentioned above, which will usually be fixed in the fit according to some prior measurement (either from monte-carlo, or possibly from data from a different trigger). The quantity $T^{\text{bck}}$ is a parameterised form of the background. The exact form of this parameterisation is discussed more for the case of $B_s \rightarrow D_s \pi$ in section 6.2.3. The quantities $M^{\text{sig}}$, $M^{\text{bck}}$ are the signal and background mass shapes which are gaussian and flat respectively. It is this mass dependence which allows the simultaneous fit for the background and signal.

### 3.5 Other potentially useful modes

The modes discussed in section 3.3 do not represent the only prospects for measuring $\Delta \Gamma_{B_s}$ at CDF Run II. For the sake of completeness the following sections provide an overview of the other possible candidate modes. Because of the volume of statistics needed, these modes will not be investigated early in Run II, but depending on the final integrated luminosity they may yet prove to be useful. The paper [36] which has been used for some of the numbers above was used as the primary source for the discussion below.

#### 3.5.1 Mixed CP Modes

For the Mixed CP modes (analogous to $B_s \rightarrow D_s \pi$) $D_s^- l^+ \nu X$ will be useful. This mode is subject to kinematic uncertainty due to the $\nu$. It is possible to apply a Monte-Carlo based correction to this, but the lifetime resolution is better for the fully reconstructed modes. The expected sample size in Run II for this set of modes is about 18000 [36]. The mode $B_s \rightarrow K^- \pi^+$ could also be used for the $\Delta \Gamma_{\text{ave}}$ measurement. The main problem here is overlap with $\pi \pi$, $\pi K$, and $KK$ decay modes of the $B_d$ and $B_s$. If it could be separated from these other modes, the yield could be as high as 4000 events [36].

#### 3.5.2 CP Separable Channels

The mode $B_s \rightarrow K^+ K^-$, like $B_s \rightarrow D_s^+ D_s^-$ is a (Pseudoscalar $\rightarrow$ 2 Pseudoscalars) decay. This means it is expected to be purely CP even for the same reason as $B_s \rightarrow D_s^+ D_s^-$. This would make it a good candidate
for isolating one of the CP eigenstates if it weren’t for the overlap problems
with $\pi\pi$ and $K\pi$ modes of the $B_d$ and $B_s$. The expected yield here is of the
order of 10000 events. A further complication with this mode is that CP
Violation is expected to be significant, which would therefore complicate the
interpretation of a lifetime measurement as a $\Delta \Gamma_{B_s}$ measurement.

$B_s \rightarrow D_s^+ D_s^-$ is also a potential mode for isolating one of the CP
eigenstates. However, it suffers from the same problems as $B_s \rightarrow J/\psi \phi$.
Since it is a (Pseudoscalar $\rightarrow 2$ Vectors) decay, there can be an odd or even
parity factor due to the relative orbital angular momentum of the two $D_s$
particles. An added complication in this mode is the momentum uncertainty
introduced by the two photons or $\pi^0$ mesons. In addition there are overlap
problems with $B_s \rightarrow D_s^* D_s$. The expected sample size in this mode for
2 $fb^{-1}$ is about 2000 events.

Finally in this section, it is worth mentioning in passing a number of
other modes which may yet prove useful in the distant future for a $\frac{\Delta \Gamma_{B_s}}{T_{ave}}$
measurement. $B_s \rightarrow J/\psi K^0_S$ would be primarily CP odd, but since it was
not seen in Run I, very few ($\sim 100$) will be observed in Run II. $B_s \rightarrow \pi\pi$
would be CP even, but again has a very small branching ratio. $B_s \rightarrow \overline{D^0} K^0$
could be used for a CP eigenstate isolation in CP eigen-state final states (eg
$\overline{D^0} \rightarrow K^0_S \rho^0$ and $K^0$ decaying as $K^0_S$). Finally, the decay modes $B_s \rightarrow
J/\psi \pi^0$, $B_s \rightarrow J/\psi \eta$, $B_s \rightarrow J/\psi \eta'$ could be used where
$\pi^0 \rightarrow \gamma\gamma$, $\eta \rightarrow \gamma\gamma$, $\eta' \rightarrow \gamma\gamma$, or $\eta' \rightarrow \rho^0 \gamma$. However, this would require fairly
low energy photon detection, which has been done at CDF (by detecting
$\gamma \rightarrow e^+e^-$) but is experimentally challenging.

\begin{footnote}
23 $D_s^* \rightarrow D_s\gamma = 94.2 \pm 2.5\%$ and $D_s^* \rightarrow D_s\pi_0 = 5.8 \pm 2.5\%$
\end{footnote}
Chapter 4
Sample Purity

The following chapter builds on the introduction to the techniques for measuring $\Delta \Gamma_{B_s}$ given in the previous chapter. The process of candidate selection undertaken for the present studies is described in detail.

The sample which is gathered by the trigger discussed in section 2.4.3 will consist of both signal and background events. It is not possible to tell on an event by event basis whether an event is signal or background. Both will be present in the sample from which the lifetime is extracted. The likelihood fit will include a fit to the $B_s$ mass distribution which should select out the signal (the background should be nearly flat in this variable\(^1\)). However the ability of the fit to lock onto the signal will depend on the ratio:

\[
\frac{S}{\sqrt{(S + B)}} \tag{4.1}
\]

where $S$ is the number of signal events in the sample, and $B$ is the number of background events. The quantity $\sqrt{(S + B)}$ is the 1σ poisson fluctuation of the sample, and so the above ratio is the size of the sample compared to this fluctuation. Maximising this ratio will yield the best result.

In order to do this, a set of kinematic cuts are applied to the sample which preferentially select out the signal over the background. The first part of the chapter describes this process. The second part of the chapter describes an

\(^1\)The main background will come from other B-decays because of the trigger. However, it is expected to be nearly flat in the mass distribution because the $B_s$ peak will be well separated from the $B_d$ peak. Rather, the background is expected to be mainly combinatoric in nature. Reflections due to assuming the wrong mass for a particle are expected to be small.
estimate of the signal to background ratio which would be obtained if the cuts defined were used on a real data sample.

The signal mode chosen for the present study is: $B_s \rightarrow D_s^- \pi^+$ where $D_s^- \rightarrow \phi \pi^-$ and $\phi \rightarrow K^+ K^-$. Feynman diagrams for these decays can be seen in figures 4.1, 4.2 and 4.3.

### 4.0.3 Strategy

**Samples**

In order to make this study, two samples were used.

- Pure signal sample of $B_s \rightarrow D_s^- \pi^+$ where $D_s^- \rightarrow \phi \pi^-$ and $\phi \rightarrow K^+ K^-$ (7659 events).
- Generic background comprised entirely of $b\bar{b}$ pairs. The decays of this sample are performed by the CLEO QQ Monte Carlo [38] (429843 events).
Both of these samples had previously been filtered by the Level 2 trigger simulation described in section 5.1 in order to describe as accurately as possible what will be observed in the data prior to offline cuts. (For the moment, it is assumed that no significant kinematic cuts will be made at Level 3, although this could change when the luminosity configuration of the experiment is upgraded).

The rationale for using a generic $b\bar{b}$ background sample is the expectation that the output of Level 2 will be almost entirely comprised of heavy flavour events. There is an assumption here that Level 3 will be able to clean up the output of Level 2, and in particular verify the impact parameter measurement. Given this assumption, the vast majority of the data which gets written to tape from the Two-Track trigger will be heavy flavour. It is expected that a small fraction of the SVT rate will be due to prompt charm decays. However the kinematic cuts used to eliminate $b$ background will be even more effective for charm. Since this is expected to be a small component of the background to begin with its effect here is ignored.

**Optimisation Procedure**

Given these two samples, a set of offline cuts are then chosen. The starting value for the cuts is chosen so as to let all the signal through. Then each cut is in turn optimised. The optimisation procedure consists of the following sequence of tasks:

1. Run the reconstruction code on the signal sample with the input cut values.
2. Run the reconstruction code on the background sample with the input
cut values.

3. Form histograms of the cut variable to be optimised for the signal and the background.

4. Calculate signal to background distribution for the cut \(^2\).

5. Read off the optimised value for the cut.

6. Form a new set of input cut values where the optimal value of the cut just optimised is used as its input value in the next iteration.

This sequence is then repeated for all the cuts to be optimised. It is necessary to point out four things about this procedure.

Firstly, in principle, the final values of the cuts obtained could depend on the order in which the optimisation is carried out.\(^3\) This is investigated in section 4.0.6.

Secondly, the definitions of signal and background for the purpose of optimisation are:

**signal**: Number of signal events which appear in the cut region for the right reasons. For example, if a signal event passes the cut, but the track(s) which actually pass the cut were not signal, then it is discounted.

**background**: Number of candidates in the cut region. So in principle there could be more than one candidate per event. This occurs often for the early cuts, like \(p_t\), but much more rarely for the latter cuts like \(D_s^\pm\) invariant mass.

Thirdly, the relative sizes of the signal and background samples don’t really matter for the purpose of optimisation. At this stage the signal to background ratio simply needs to be maximised, not measured\(^4\).

Finally, only a part of the \(B_s \rightarrow D_s \pi\) rate is used in the optimisation. This part is defined by requiring that the pions from the \(B_s\) and \(D_s\) pass the

---

\(^2\)Each point in the signal to background distribution is calculated by adding up the amount of signal and background respectively above a given bin in the individual distributions.

\(^3\)For example two cuts may act on the same set of events. If they are optimised sequentially as in the above procedure, the optimal values will depend on the order.

\(^4\)Note that the *signal* in the denominator of equation 4.1 is actually present in the right amount in the background sample.
two track trigger as a pair. See sections 5.2.2, 5.4.1 and 5.6 for the reasons this is done.

4.0.4 Cuts

Section 3.4.1 describes the sorts of kinematic cuts to be applied. These are the cuts which were optimised to maximise the signal to background ratio:

- $p_t$ of the $K^-$
- $p_t$ of the $K^+$
- $p_t$ of the $\pi^-$ from the $D_s^-$
- $p_t$ of the $\pi^+$ from the $B_s$
- Invariant mass of $\phi$
- Invariant mass of the $D_s^-$
- Helicity angle between the $K^l's$ and $D_s^-$ direction.

These are the cuts which were applied, but not optimised to maximise the signal to background ratio:

- Invariant mass of the $B_s$
- Probability of the vertex fit.

The $B_s$ invariant mass cut is not optimised because the $B_s$ mass is used to discriminate between signal and background in the likelihood fit (see section 6.2). The probability from the vertex fit is not usually optimised.

4.0.5 Background Discrimination

The utility of the various cuts can be seen graphically by plotting the signal and background on the same axes so that the separation power of the cut can be seen.

In these plots, the signal is scaled so that it appears to be the same height as the background. This is simply for display purposes. Also, the calculated signal to background ratio is scaled. The feature to look for in all the separation power plots is the relative shapes of the distributions. If all the
signal is concentrated in one portion of the plot, whereas all the background is concentrated elsewhere, this means the separation power of the cut is high.

Figure 4.4 shows the separation power of the \( p_t \) cuts for the 4 signal tracks. Figure 4.5 shows the separation power of the invariant mass cut on the \( \phi \), and figure 4.6 shows the separation power of the invariant mass cut on the \( D_s \). Finally figure 4.7 shows the separation power of the helicity cut.

In each plot, all the cuts except the one plotted have been applied. So for example in figure 4.7 the optimal \( p_t \) invariant mass cuts have been applied, but the helicity cut has not. This has the advantage of showing the effect each cut over and above the others. If each cut variable was plotted without any of the other cuts applied, overlap between the different cuts could be hidden.

The signal cuts off very sharply at 2\( GeV \) in the two pion plots. This is because for each event in the sample both the pions formed part of the trigger decision (ie \( p_t > 2\text{GeV} \)). The same is not true for the kaons.

### 4.0.6 Effect of Correlations

The final values for the cuts obtained from the procedure outlined in section 4.0.3 could be dependent on the order of optimisation. This is because two cut variables could be correlated. This effect was investigated and was found to have a minimal effect on the optimal cut values quoted in table 5.1. In fact, the worst correlation occurs between the \( p_t \) of the two kaons.\(^5\) This correlation is shown in figure 4.8. It is due to the fact that the \( \phi \) mother particle has very little phase space to decay into. The kaons are therefore highly collimated. However, even this correlation did not change the optimal value of either kaon cut by more than 100 MeV.

### 4.0.7 Optimal Cut values

Table 4.1 summarises the optimum values for the various cuts. Note that the \( p_t \) cuts on the kaons are below the trigger threshold for the two track trigger because typically they are not trigger tracks.

\(^5\)Note that there is a weak dependence of the invariant masses of the \( \phi \) and \( D_s \) on the \( p_t \) of the daughter tracks, but it is not large enough to influence the optimal values of the cuts.
Figure 4.4: Separation power plots for $K^-, K^+, \pi^-$ and $\pi^+ p_t$. Background in blue, signal in green. $\frac{S^2}{(S+B)}$ in black. The top left plot shows the $K^- p_t$, the top right plot shows the $K^+ p_t$, the bottom left plot shows the $\pi^- p_t$ (from the $D_s^-$), and the bottom right plot shows the $\pi^+ p_t$ (from the $B_s$).
Figure 4.5: Separation power of the $\phi$ invariant mass cut. The left plot shows how the signal (green) and background (blue) differ. The right hand plot shows how the quantity $\frac{S^2}{S+B}$ varies with different cut windows. Note that the separation power is plotted against cut window size in contrast to the $\phi$ mass plot.

4.1 Signal to Background Calculation

Armed with the optimal cut values obtained the preceding section it is now possible to make a prediction for the approximate signal to background ratio which might be expected in an analysis of the data.

4.1.1 Limitations of the prediction

It cannot be emphasised enough that these studies do not take into account many detector effects. This means that the actual signal to background ratio observed is likely to be worse. The main source of contaminant in the background which is not taken into account in a generator level study are secondary tracks due to the presence of the detector\(^6\).

Another inadequacy of the signal to background estimate is that it does not take into account the effect of prompt charm mesons. However, there are two reasons to suspect that the amount of background due to this source

\(^6\)For example protons knocked out of the beryllium beampipe.
will be significantly lower than the generic B meson case treated in these studies. Firstly, the D mesons are shorter lived than the B’s. The mean decay distances of the $D_0$ and $D_+$ are 120 $\mu m$ and 300 $\mu m$ respectively. This is to be compared with 450 $\mu m$ as the mean decay distance of the $B_s$. The lower cut of 120$\mu m$ (100 $\mu m$ for the $B_d \rightarrow \pi^+\pi^-$ stream) in the trigger will therefore favour the $B_s$. Secondly, the invariant mass of the D mesons is a great deal less than the $B_s$ mass (1.8$GeV$ compared to 5.38$GeV$). Therefore the number of combinations of tracks in a prompt charm event which can have an invariant mass which adds up to the $B_s$ mass is much less than a generic B event$^7$.

4.1.2 The Calculation

Normally, a signal to background calculation would be carried out by fitting the data to a distribution known to distinguish between signal and background (such as the invariant mass of the $B_s$ in this case). The relative

---

$^7$See discussion in section 4.0.3
Figure 4.7: Separation power of the kaon helicity. $|\cos(\theta)|$ is plotted. Signal is in green, background is in blue, and $\frac{S}{\sqrt{B}}$ is in black.
Figure 4.8: The 2D distribution of the kaon $p_t$. A significant correlation is visible due to the fact that the $\phi$ has very little phase space to decay into. Therefore the kaons are fairly tightly collimated.
<table>
<thead>
<tr>
<th>Cut</th>
<th>Optimum Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^- p_t$</td>
<td>0.6 GeV</td>
</tr>
<tr>
<td>$K^+ p_t$</td>
<td>0.6 GeV</td>
</tr>
<tr>
<td>$\pi$ (from $D_s$) $p_t$</td>
<td>1.9 GeV</td>
</tr>
<tr>
<td>$\pi$ (from $B_s$) $p_t$</td>
<td>3.0 GeV</td>
</tr>
<tr>
<td>$\phi$ invt mass window</td>
<td>0.0072 GeV</td>
</tr>
<tr>
<td>$D_s$ invt mass window</td>
<td>0.01 GeV</td>
</tr>
<tr>
<td>$\text{Abs} (\cos(\theta))$ (helicity)</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Table 4.1: Optimum Cut values for the optimisation of $D_s$ reconstruction. These are the values used in the analysis.

fractions could then be inferred directly from this fit (for a given range of mass).

However, in the present studies there were two samples used. The signal sample, after level 2 is only about 10 times smaller than the background sample. This contrasts to the situation in the data where the background sample after level 2, but before reconstruction cuts, will be many orders of magnitude larger than the signal. This is the point of the reconstruction cuts.

Therefore, a relative normalisation must be taken into account when calculating the signal to background ratio expected. This normalisation will parameterise the prior knowledge about how much more background is expected. The ingredients of this factor will thus be the various branching ratios of the decay $B_s \rightarrow D_s \pi$, and the $B_s$ production fraction itself.

Consider the relationship between the numbers of events prior to all cuts (trigger and reconstruction), and the final numbers of events after all cuts:

$$N^S_I = F_S N^S_F$$  \hspace{1cm} (4.2)
$$N^B_I = F_B N^B_F$$  \hspace{1cm} (4.3)

where $N^S_I$ is the number of events in the signal monte-carlo sample before the level 2 trigger, and $N^F_I$ is the number of events in the signal monte-carlo sample after all cuts. Similar quantities are defined for the background.

Consider further the factor $B$ defined:

$$N_B = BN_S$$  \hspace{1cm} (4.4)
where \( N_B \) is the number of generic B events required to have \( N_S \) signal events inside the generic B sample. This is the normalisation factor mentioned above.

The signal to background ratio (R) is then given by:

\[
R = \frac{N_S}{N_B} = \frac{F_B}{B \times F_S}
\]  \hspace{1cm} (4.5)

It then remains to calculate \( F_B \) and \( F_S \). This is straightforward for \( F_S \), since the reconstruction cuts by design don’t kill this sample. In addition to the cuts defined in section 4.0.7, there are two more cuts on the reconstructed proper time and the reconstructed proper time uncertainty. These cuts are necessary because the proper time efficiency curve is only measured over a finite range in these variables\(^8\). The range of reconstructed proper time in the fit was \( 0.1 \text{ ps} < t < 8 \text{ ps} \). The cut on the uncertainty in reconstructed proper time was \( 0.075 \text{ ps} < \sigma(t) < 0.117 \text{ ps} \). Table 4.2 summarises the signal efficiency for the sample used in the current calculation.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Value</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Size</td>
<td>(1.785 \times 10^7)</td>
<td></td>
</tr>
<tr>
<td>After Main Cuts</td>
<td>1394</td>
<td>37</td>
</tr>
<tr>
<td>After all Cuts</td>
<td>1070</td>
<td>33</td>
</tr>
<tr>
<td>Overall Efficiency</td>
<td>(5.99 \times 10^{-5})</td>
<td>(1.83 \times 10^{-6})</td>
</tr>
<tr>
<td>Rejection Factor</td>
<td>(1.67 \times 10^4)</td>
<td>510</td>
</tr>
</tbody>
</table>

Table 4.2: The effect of the cuts, and the corresponding efficiency and rejection factor. The Main Cuts in the 2nd row refer to those defined in table 5.1. The 3rd row includes the effects of the cuts on reconstructed proper time and its uncertainty. The rejection factor in the last row is the reciprocal of the overall efficiency.

In contrast to \( F_S \) the \( F_B \) calculation is much tougher. There is about 1 background event left after the main reconstruction cuts have been applied, and this is not enough to base a prediction of the signal to background ratio.

\(^8\)The trigger is only efficient on the signal above \( t \sim 0.1 \text{ps} \).
The strategy which was decided upon to overcome this problem was to open up one of the cleanest cuts (the $B_s$ invariant mass) to a point where a number $\sim 100$ of background events survived the cuts. The statistical fluctuation as a percentage of the total is far less on this number of events than on $\sim 1$ event. Then, using the knowledge of the shape of the background distribution in this variable, a factor to take account of the opening up of the cut variable was deduced.

Since the reconstructed proper time and its uncertainty are cut in the signal, these same cuts must also be applied in the background. Further, recall that the signal sample is a subcomponent of the total level 2 bandwidth$^9$. In the real data, this pseudo-cut will be imposed by requiring that the b-pion, and d-pion candidates appear as a pair in the level 2 trigger decision. This pseudo-cut has to be applied to the background.

The order of the cuts in the background is thus:

1. All optimised and quality cuts, without the tight $B_s$ mass cut.
2. Reconstructed proper time and uncertainty cut.
3. b-pion, d-pion trigger decision cut.
4. Tightened $B_s$ mass cut.

It is safe to leave the $B_s$ mass cut until last, and calculate the rejection factor from the shape of the larger sample because it is largely independent of all the other cuts$^{10}$.

The second and third cuts enumerated above will be quoted as one factor so that the total background rejection factor is given by:

$$F_B = F_B(m_B) \times F_B(extra) \times F_B(main) \quad (4.6)$$

where $F_B(m_B)$ is the $B_s$ invariant mass factor to be examined in the next section, $F_B(extra)$ is the factor due to the proper time, proper time uncertainty, and trigger decision cuts, and $F_B(main)$ is the factor due to all of the other reconstruction cuts and the trigger.

$^9$See section 4.0.3 and section 5.2.2.
$^{10}$Except the proper time uncertainty cut
4.1.3 $B_s$ invariant mass cut factor

Various ranges of the reconstructed $B_s$ mass were investigated to evaluate how many background combinations survived. The range of 2.98 GeV $\rightarrow$ 6.58 GeV was decided on in the end since this yielded an acceptable number of background events.$^{11}$

The background $B_s$ mass shape distribution was deduced by applying all the $p_t$ cuts, but leaving all the other cuts wide.$^{12}$ This is because although the invariant mass is not strongly correlated to the $p_t$ of the daughter tracks, over such a large range of $B_s$ invariant mass, there is a significant effect on the calculated cut factor. Figure 4.9 shows the shape of the $B_s$ invariant mass distribution for the background with the $p_t$ cuts applied, but the invariant mass cuts left open.$^{13}$

Figure 4.10 shows the same thing as figure 4.9, except the optimised $\phi$ invariant mass cut has been applied. The fits shown in each figure are of the form:

$$F(m) = A + Be^{-\frac{(m-2.8)}{\alpha}}$$  \hspace{1cm} (4.7)

The parameters for the fits are summarised in table 4.3.

As can be seen from the table, the parameters are basically consistent with each other, showing that the shape has not changed when the $\phi$ invariant mass cut has been introduced. This indicates that the invariant mass cuts are independent to each other.

The ratio of $\frac{A}{B}$ is expected to be the same for each fit if the shape is staying the same since a common factor is simply an overall normalisation. The bad $\chi^2$ per d.o.f of the upper plot is not a problem since it just reflects the very high statistics of the bins being fit. We simply require a parameterisation to give a functional form which will then be used on the $\sim$ 100 events to calculate a cut factor. In other words, the distribution of $B_s$ invariant mass after all the other cuts will be assumed to have the same shape as the distributions in figures 4.9 and 4.10.

A very similar unchanged shape is obtained when the optimal $D_s$ invariant mass cut is applied, and the $\phi$ invariant mass cut is left loose.

$^{11}$133 events was enough to investigate the effect of the other cuts on the background.
$^{12}$In principle, the $B_s$ mass could be correlated to some of the other cuts. This will be investigated at the end of this section.
$^{13}$The rise at large invt mass is due to the $p_t$ cuts
Figure 4.9: Reconstructed background $B_s$ Invariant mass distribution, with $p_t$ cuts, but loose $\phi$ and $D_s$ invariant mass cuts. The rise at large mass is due to the $p_t$ cuts.

Figure 4.10: Reconstructed background $B_s$ Invariant mass distribution, with $p_t$ cuts, and optimised $\phi$ Invt mass cut, but loose $D_s$ invariant mass cut. The rise at large mass is due to the $p_t$ cuts.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>With $p_t$ cuts (figure 4.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A$</td>
<td>991</td>
<td>8.8</td>
</tr>
<tr>
<td>$B$</td>
<td>28700</td>
<td>159</td>
</tr>
<tr>
<td>$B_A$</td>
<td>29.0</td>
<td>0.93</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.49</td>
<td>0.002</td>
</tr>
<tr>
<td>With $p_t$ and $m_\phi$ cuts (figure 4.10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$A$</td>
<td>498</td>
<td>2.0</td>
</tr>
<tr>
<td>$B$</td>
<td>1420</td>
<td>34</td>
</tr>
<tr>
<td>$B_A$</td>
<td>28.53</td>
<td>4.09</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.52</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Table 4.3: Parameters in equation 4.7 from the fits in figures 4.9 and 4.10

The value of the $B_s$ mass cut used in the final fit is not optimised in the same way as the invariant mass cuts on the $\phi$ and $D_s$. The final $B_s$ mass cut window is 5.34 GeV $\rightarrow$ 5.42 GeV. The factor is then simply the ratio of two definite integrals of equation 4.7:

$$F_B(m_B) = \frac{BW^{-1} \times \int_{m_h'}^{m_h} A + Be^{-\frac{(m-m_\phi-2.8)}{\chi}} \, dm}{BW^{-1} \times \int_{m_l}^{m_l} A + Be^{-\frac{(m-m_\phi-2.8)}{\chi}} \, dm} \quad (4.8)$$

where the upper integral is between the wide $B_s$ invariant mass limits ($m_h' = 6.58, m_h = 2.98$), and the other is between the final cut limits ($m_h = 5.42, m_l = 5.34$). The factor BW is the bin width which comes into any integral of a function fitted to a histogram. The uncertainty of this rejection factor is estimated by calculating the effective binomial fluctuation on the number of events in the narrow mass range in figure 4.9.

This procedure yields a value for the $F_B(m_B)$ factor defined in equation 4.6 which is given in table 4.4.

However, this way of calculating the $B_s$ rejection factor does not take into account the effect of all the cuts. In particular, if one cut overlaps with the $B_s$ mass cut the rejection factor of table 4.4 will be over estimated.

---

14To enable the likelihood fit to separate based on the mass, it has to be left a bit wider. See sections 4.1.7 and 6.2.
Table 4.4: The value for $F_B(m_B)$ defined in equation 4.6 with its uncertainty

<table>
<thead>
<tr>
<th>$F_B(m_B)$</th>
<th>Percentage Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>148</td>
<td>3.3%</td>
</tr>
</tbody>
</table>

The $p_t$ cuts do not affect this result because it is calculated after they have been applied. The invariant mass cuts on the $\phi$ and the $D_s$ have been shown to decouple from the $B_s$ mass. Furthermore, it is likely that the $B_s$ mass cut will be independent of the helicity cut. There is no reason for the angular distribution of the kaons relative to the $D_s$ in the $\phi$ rest frame to be correlated to the invariant mass of the $B_s$.

However, this still leaves several cuts:

- Vertex Probability cut
- $(\pi_b,\pi_d)$ trigger decision cut
- Proper Time cut
- Proper Time Uncertainty cut

The effect of the vertex probability cut was estimated by plotting the $B_s$ mass shape after both $p_t$ and the vertex probability cuts. The mass shape in figures 4.9 and 4.10 was again recovered indicating that there is no significant overlap.\textsuperscript{15}

The effect of the $(\pi_b,\pi_d)$ trigger decision cut and the proper time cut were investigated by imposing a representative $B_s$ mass cut on the event sample.\textsuperscript{16} A cut window of $3.5 \, GeV < m_B < 5.5 \, GeV$ was chosen because this resulted in a rejection factor of 2. It was observed that this rejection factor did not change significantly when the $(\pi_b,\pi_d)$ trigger decision cut and proper time cut were applied separately.\textsuperscript{17}

\textsuperscript{15}The rejection factor calculated using the corresponding plot was statistically consistent $(F_B(m_B) = 129 \pm 26)$ to the one in table 4.4.

\textsuperscript{16}If the full $B_s$ mass cut were imposed, there would be $< 1$ event expected.

\textsuperscript{17}Without the trigger decision cut: $F_B(m_B) = 2.08 \pm 0.18$, with the trigger decision cut: $F_B(m_B) = 1.73 \pm 0.35$. Without the proper time cut: $F_B(m_B) = 2.08 \pm 0.18$, with the proper time cut: $F_B(m_B) = 1.98 \pm 0.23$
CHAPTER 4. SAMPLE PURITY

However, the proper time uncertainty cut seems to have a significant overlap with the $B_s$ mass. Before the proper time uncertainty cut, the representative $B_s$ mass cut gave $F_B(m_B) = 2.08 \pm 0.18$, while afterwards, it yielded $F_B(m_B) = 1.29 \pm 0.17$. Unfortunately, this observation cannot be used to adjust the central value of the final $F_B(m_B)$ cut because it was calculated with a representative $F_B(m_B)$ cut. Therefore, this effect is added in as an uncertainty on the final value of $F_B(m_B)$.

This leads to a revised final value for $F_B(m_B)$ given in table 4.5.

<table>
<thead>
<tr>
<th>$F_B(m_B)$</th>
<th>Percentage Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>148</td>
<td>61%</td>
</tr>
</tbody>
</table>

Table 4.5: Revised value for $F_B(m_B)$ defined in equation 4.6 with its uncertainty

4.1.4 Extra cut factor

The factor $F_B(extra)$ defined in equation 4.6 is calculated by applying the three cuts listed below to the sample after the main reconstruction cuts but with the wide $B_s$ mass cut.

Reconstructed proper time cut

Reconstructed proper time uncertainty cut

Requirement that the B-pion and D-pion be present in the trigger

The rejection factors for these cuts are summarised in table 4.6.

---

18 The ratio of the representative $F_B(m_B)$ before and after the proper time uncertainty cut has been imposed is the necessary multiplication: $\frac{2.08}{1.29} = 1.61$ leading to a 61\% uncertainty.

19 It would be advantageous to apply the proper time uncertainty before the $B_s$ mass shape is calculated to take account of the overlap between these two cuts. However, this reduces all subsequent sample sizes by an order of magnitude and is therefore unfeasible (the effect of proper time and trigger decision would have to be estimated with a sample of 13 events).

20 see section 4.1.2
<table>
<thead>
<tr>
<th>Quantity</th>
<th>Value</th>
<th>Percentage Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size after main cuts</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>Size after $t$ and $\sigma(t)$ cuts</td>
<td>$12 \pm 3.3$</td>
<td>28%</td>
</tr>
<tr>
<td>Size after trigger requirement</td>
<td>$2 \pm 1.4$</td>
<td>70%</td>
</tr>
<tr>
<td>Combined rejection factor ($F_B(extra)$)</td>
<td>$65.17 \pm 45.61$</td>
<td>70%</td>
</tr>
</tbody>
</table>

Table 4.6: The values for $F_B(extra)$ defined in equation 4.6 with errors

4.1.5 Main cut factors for signal and background

The final part of $F_B$ is $F_B(main)$ as defined in equation 4.6, and is straightforward to calculate. It is the ratio of the $\sim 100$ events left after the loose version of the cuts and the number of events before level 2. However, there is one more wrinkle. The number of events after level 2 but before the reconstruction cuts, is known for the background; but there is no easy way of calculating the number of events before level 2. However, it is easy to run a relatively small number of events through the level 2 trigger simulation to calculate the efficiency. Therefore, $F_B(main)$ is itself a product:

$$F_B(main) = F(L2) \times F(reco)$$  (4.9)

where $F(L2)$ is the level 2 rejection factor for the background, and $F(reco)$ is the factor between level 2 and the reconstruction cuts.

These parameters with their uncertainties are given in table 4.7.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Abs Uncertainty</th>
<th>Fractional Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F(L2)$</td>
<td>555</td>
<td>6.19</td>
<td>1.1%</td>
</tr>
<tr>
<td>$F(reco)$</td>
<td>3231</td>
<td>280</td>
<td>8.7%</td>
</tr>
</tbody>
</table>

Table 4.7: The values for $F_B(main)$ components defined in equation 4.6 with uncertainties.
CHAPTER 4. SAMPLE PURITY

The signal factor is simply:

\[
F_S = \frac{\text{events prior to } L_2}{\text{events after all cuts}}
\] (4.10)

Its value and error are given in table 4.9, along with the final value of \( F_B \) with its error.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Abs Error</th>
<th>Fractional Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F_S )</td>
<td>( 1.67 \times 10^4 )</td>
<td>510</td>
<td>3%</td>
</tr>
<tr>
<td>( F_B )</td>
<td>( 1.73 \times 10^{10} )</td>
<td>( 1.60 \times 10^{10} )</td>
<td>93%</td>
</tr>
</tbody>
</table>

Table 4.8: The values for \( F_S \) defined in equation 4.6 with errors

4.1.6 Signal normalisation : B

The quantity defined in equation equation 4.4 is given in table 4.9 with its error. Here, the central value for the branching ratio \( BR(B_s \rightarrow D_s \pi) \) has been assumed to be the same as its \( B_d \) analogue, \( B_d \rightarrow D \pi \) [4].21 The \( B_s \) production fraction is taken from the LEP measurement [4].

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Abs Error</th>
<th>Fractional Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B_s ) prod</td>
<td>10.7%</td>
<td>1.4%</td>
<td>13%</td>
</tr>
<tr>
<td>( BR(B_s \rightarrow D_s \pi) )</td>
<td>( 3.0 \times 10^{-3} )</td>
<td>( 0.4 \times 10^{-3} )</td>
<td>13%</td>
</tr>
<tr>
<td>( BR(D_s \rightarrow \phi \pi) )</td>
<td>3.6%</td>
<td>0.9%</td>
<td>25%</td>
</tr>
<tr>
<td>( BR(\phi \rightarrow K + K^-) )</td>
<td>49.2%</td>
<td>0.007%</td>
<td>0.01%</td>
</tr>
<tr>
<td>B</td>
<td>( 1.76 \times 10^{5} )</td>
<td>( 5.5 \times 10^{4} )</td>
<td>31%</td>
</tr>
</tbody>
</table>

Table 4.9: Branching Ratios and \( B_s \) production fraction giving \( R \) in equation 4.4

\[21\] By SU(3) flavour symmetry
4.1.7 Final value and error

When all the factors in equation 4.5 are combined a value for R is obtained that is shown in table 4.10.

The definition of R is repeated for clarity:

\[ R = \frac{N_S}{N_B} = \frac{F_B}{B \times F_S} \]  

(4.11)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Abs Uncertainty</th>
<th>Fractional Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>5.88</td>
<td>5.76</td>
<td>98%</td>
</tr>
</tbody>
</table>

Table 4.10: Signal to Background Prediction with error

This value unfortunately is not very illuminating. A signal to background of \( R = 0 \) is distinctly allowed. The reasons for this are the two errors:

- The binomial fluctuation on the pre \( B_s \) mass cut background sample: 70%
- The uncertainty in \( B_s \) mass cut rejection factor due to the overlap with the proper time uncertainty: 61%

However, if the cut on the proper time uncertainty were not made then these two large sources of uncertainty would be greatly reduced. The background sample before the \( B_s \) mass cut would be about 10 times bigger, meaning a much smaller fractional uncertainty on the background rejection factor. Further, there would be no overlap with the proper time uncertainty to estimate, so the error on the \( B_s \) mass rejection would reduce to the one given in table 4.4.

Leaving out the proper time uncertainty cut corresponds to understanding the efficiency functions of chapters 5 and 6 extremely well. That is beyond the scope of this analysis. However, from the shape of the initial and final proper time uncertainty distribution, it seems that cutting on this quantity improves the signal to background ratio.\(^{22}\) Table 4.11 therefore shows the

\(^{22}\) See the plot in figure 6.3. Note that this does not account for probable overlap with other cuts.
signal to background ratio for the case of no proper time uncertainty cut.\footnote{\(F_S\) changes from 1.49 \(\times\) \(10^4\) \(\rightarrow\) 1.28 \(\times\) \(10^4\), and \(F_B(\text{extra})\) changes from 65.17 \(\rightarrow\) 5.32. The relative errors on \(F_S\) and \(F_B(\text{extra})\) are 2.7\% and 18\% respectively.}

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Abs Uncertainty</th>
<th>Fractional Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>0.63</td>
<td>0.23</td>
<td>37%</td>
</tr>
</tbody>
</table>

Table 4.11: Projected Signal to Background Prediction with error for the case of no proper time uncertainty cut.

The dominant uncertainty in this result is the branching ratio \(BR(D_s \rightarrow \phi \pi)\). However, this result suggests a reasonably favourable signal to background ratio. Furthermore, as noted above, the proper time uncertainty cut is likely to improve the signal to background ratio. Therefore, it is conservative to use a signal to background ratio in the analysis of chapter 6 which is similar to the value shown in table 4.11. The value \(R = 0.5\) was chosen.

Finally, it is worth pointing out that it would be possible to cut much tighter on the \(B_s\) mass. The reason this is not done is to provide a way for the fit to discriminate between signal and background.\footnote{The \(B_s\) mass distribution is very different for the signal and background} However, this technique is not the only way to fit for the lifetime, particularly if there are enough events in the sample.\footnote{A lifetime analysis using a very small sample requires the simultaneous fit.\cite{40}} For example if a sideband subtraction method is employed, the optimum \(B_s\) mass window is about 4 times smaller than in the present analysis. This would improve the signal to background ratio at the expense of assuming the background behaviour under the peak matched the sidebands.

\footnote{In sideband subtraction, the shape of the background in the lifetime is deduced from the sample defined by 2 windows each side of the signal mass peak.}
Chapter 5

Level 2 Trigger Efficiency Studies

One of the most important aspects of any analysis at a hadron collider is the effect of the trigger on the physics samples. Not only does it determine the final sample sizes, but it significantly influences the measured quantities in the samples. This is particularly true at CDF Run II, since the trigger has features which are entirely new. For example, in any B-Meson life-time measurement, such as the ones inherent in a $\Delta \Gamma$ analysis, the least well understood effect of the trigger is the impact parameter cut at Level 2. This is because the impact parameter of a B-daughter is strongly correlated to the B-lifetime.\footnote{See appendix A.2 for a definition of impact parameter.}

The following sections detail a study made of the effect of the Level 2 Two-Track trigger (see section 2.4.3) for the mode of the $B_s$ meson, $B_s \rightarrow D_s^- \pi^+$ where $D_s^- \rightarrow \phi \pi^-$ and $\phi \rightarrow K^+ K^-$.\footnote{The conjugate mode $B_s^* \rightarrow D_s^{+} \pi^-$ where $D_s^+ \rightarrow \phi \pi^+$ is also understood to be under consideration unless otherwise stated.}

A study of Level 2 efficiency will indicate how the $B_s$ lifetime is to be measured in this mode. The techniques developed in the present chapter will be employed in chapter 6 in a monte carlo analysis of the $B_s$ lifetime. The Level 2 Efficiency will also enter the signal to background estimate given at the end of chapter 4. Expected sample sizes for CDF Run II in the mode mentioned above are given in \cite{34}.
5.1 Level 2 Trigger Simulation

5.1.1 General Strategy

For the purposes of this study it was decided to use a generator level simulation of the Two-Track Trigger. The primary disadvantage of the hit based simulation of the level 2 trigger is that it relies upon the full GEANT [41] simulation of the silicon tracking volume. At the time this study was carried out, this simulation was still under construction, and was yet to be thoroughly tested. In addition to this, even if the results of the simulation could be trusted, previous experience with an earlier draft had indicated a prohibitive processing time. Even at its quickest, the old simulation had taken of the order of 20 seconds per event to run. Since the Level 2 efficiencies are of the order of 1%, obtaining a comparable sample size to the generator level technique employed would have taken too much time.\footnote{The generator level simulation is several orders of magnitude faster, but it still took many days to collect enough events. Therefore, using the old simulation would have taken well over a year of computer time.} There was a plan to combine these techniques (where the generator level version would be run first in loose mode and then the hit based simulation would be run subsequently), but this has not happened largely due to the lack of a tested simulation.

In addition to the utility of these studies in providing actual predictions of the effects mentioned above, they will also prove very useful as a cross-check when similar studies are performed with the full simulation.

5.1.2 Generator Level Trigger Simulation: GenTrig

Work began on constructing a simulation of the Two-Track Trigger (Level One and Level Two) early in 2000. This work began as a collaboration between the Oxford and MIT groups, but the later development was largely due to MIT [42]. The end product of these efforts is a software package called GenTrig.

The input to GenTrig is the Pythia LUND monte carlo [43]. For all B-Physics studies, Pythia is run with the \textit{MSEL} switch set to 5 [44] which ensures that for every $p\overline{p}$ interaction, there is a $b\overline{b}$ pair produced. The version of pythia used is 6.129.
The beamspot in pythia is a point, whereas the beamspot at CDF is a long thin tube. The standard deviations of collisions from the point (0, 0, 0) are approximately \( \sigma_x = 25\mu m \), \( \sigma_y = 25\mu m \), \( \sigma_z = 30cm \). Therefore a further module, GenPrimVert, must be included in the analysis path to smear the primary vertex with these widths.

In order to be as realistic as possible, the measurements that the trigger makes of kinematic quantities are smeared according to gaussians of the appropriate widths. The three quantities that the trigger simulation smears at Level 2 are \( d_0 \) (the impact parameter), \( \phi_0 \) (\( \phi \) at distance of closest approach to the primary), and \( p_t \). Of these, \( d_0 \) and \( \phi_0 \) are smeared according to a gaussian of uniform width (30\( \mu m \) for \( d_0 \) and \( 0.001 \text{ radians} \) for \( \phi_0 \)). The \( p_t \) is smeared according to the relationship:

\[
p_t(smear) = \frac{1}{p_t + R\sigma_C} \tag{5.1}
\]

where \( \sigma_C \) is the resolution of the measurement of curvature, which is assumed to be a gaussian.

A cross check to make sure that this smearing is being done properly is to plot pull distributions:

\[
pull(x) = \frac{x_{smear} - x_{true}}{x_{\sigma}} \tag{5.2}
\]

An example of these pull distributions are given below in figure 5.1\(^4\).

### 5.1.3 Generation \( p_t \) Cut

Even though every event generated has a \( \bar{b}b \) pair, the vast majority will not get passed the Two-Track Trigger. For this reason, it makes good sense to generate only in the region of phase space where the trigger will pass. The most effective switch to utilise in pythia therefore is the minimum \( p_t \) of the b-quarks generated. This is accessible from the Run \( \Pi \text{ pythia module via the } \text{common menu, and the parameter } \text{CKIN3. The plot in figure 5.2 shows a typical } b \text{-quark } p_t \text{ distribution which one obtains using this switch (note the bin errors would be too small to see on this plot).}

\(^4\)It will be noted that there is a slight bias towards positive values in the \( p_t \) pull distribution. This effect is discussed in appendix C
Figure 5.1: Example of a $p_t$ pull distribution from GenTrig. The slight bias towards positive values is discussed in appendix C.
Figure 5.2: The $b$-Quark $p_t$ distribution out of pythia with a soft rest frame $p_t$ cut on the individual $b$-quarks. See section 5.1.3.
Though one might expect a sharp cut-off in the $p_t$ distribution at 2.0GeV with CKIN3 = 2.0 GeV, this is evidently not the case in figure 5.2. The reason for the significant number of events with b-quark $p_t$ less than 2.0GeV is that the cut on $p_t$ is made in the rest frame of the $b\bar{b}$ pair. This is not the same as the lab frame, and there is a corresponding smearing of the threshold.

In order to argue that this generation cut does not significantly affect the results of the proposed efficiency study, it must be demonstrated that a negligibly small fraction of the Level 2 accepts occur in the region where the simulated $p_t$ distribution departs from a physical $p_t$ distribution (before the turn-over\textsuperscript{5} occurs in figure 5.2). It may be seen that this turn-over is beginning to occur at $p_t \simeq 5.0 GeV$ in figure 5.2. Figure 5.3 shows the fraction of total level two accepts occurring below this threshold is negligibly small.

### 5.2 Aims of the Study

The $B_s/\overline{B_s}$ decay channel\textsuperscript{6} which is largely in view in the current study is:

$$B_s \rightarrow D^- \pi^+ \text{ where } D^- \rightarrow \phi \pi^- \text{ and where } \phi \rightarrow K^+ K^-$$
$$\overline{B_s} \rightarrow D^+ \pi^- \text{ where } D^+ \rightarrow \phi \pi^+ \text{ and where } \phi \rightarrow K^+ K^- \quad (5.3)$$

The primary aim of this study is to examine the Level 2 Two-Track trigger efficiency for this mode. The global Level 2 efficiency is an interesting result of the study because it allows us to estimate the sample size in this mode that we might expect. However, this number has already been estimated [34]. What has not been studied is how this efficiency varies with the distributions one would wish to measure in an analysis of this mode. For example in a $\Delta \Gamma$ measurement at CDF Run II, a measurement of the life-time of the $B_s$ is required. In Run I, since there was no impact parameter cut at Level 2, the effect of the trigger on the lifetime distribution was not marked [45].

\textsuperscript{5}Although some turn over is expected due to the boost of the b-quarks, it is safe to assume that the unphysical behaviour due to the $p_t$ cut under discussion is confined to the region of the turn-over.

\textsuperscript{6}The second mode ($\overline{B_s}$) is equivalent for most purposes to the first, and is understood to be under consideration whenever the $B_s$ mode is mentioned (unless explicitly stated).
Figure 5.3: The Level 2 Two-Track Trigger efficiency vs b-quark $p_t$. It may be seen that a negligible fraction of the level 2 trigger rate comes from b-quarks with $p_t$ below about 5 GeV.
However, with the presence of the SVT in Level 2, a study such as the present one is very important to understand.

5.2.1 Trigger Configurations

The visible products of the mode defined in equation 5.3 have different kinematic characteristics. A given Two-Track trigger decision will contain a combination of these tracks and background. Therefore the efficiency as a function of proper-lifetime of the $B_s$ for the different configurations of trigger accept is expected to be different. Estimating the size and importance of these differences is an aim of this study.

This is important because it is extremely difficult to accurately model the relative abundances of the different configurations of trigger accept. This is especially true for some of the trigger configurations which contain background.

5.2.2 Pure Signal Trigger Configurations

Consider the mode defined in equation 5.3. There are a maximum of 4 visible tracks for such a decay of the $B_s$. In principle any combination of these tracks can account for a trigger accept. A meaningful way to categorise the possible combinations is given below.

The two kaons are kinematically symmetric except that they have opposite charge. There are small charge dependent detector effects at CDF, but since this is a generator level study they are ignored. Therefore, it is useful to analyse the mode in terms of:

$\pi_b$ : the $\pi$ from the $B_s$

$\pi_d$ : the $\pi$ from the $D_s$

$K$ : either K from the $\phi$

The meaningful categorisation of pure signal pairs which can account for a trigger accept is therefore:

1. $(\pi_b, \pi_d)$

2. $(\pi_b, K)$
3. $(\pi_d, K)$

4. $(K, K)$

where for example $(\pi_b, \pi_d)$ means that both pions from the $B_s$ and $D_s$ were part of the same pair in the two-track trigger.

### 5.2.3 Pure Background Trigger Configurations

Apart from the direct daughters of the $B_s$ decay in question, there are in general two other sorts of track in a signal event:

- **$B$-daughters** from any other $B$ meson in the event.
- **Fragmentation tracks** from the underlying event.

Usually there will only be one other $B$ meson in the event since $b$-quarks are produced in pairs. However, when the instantaneous luminosity reaches its maximum there will be an average of $> 1$ interactions in a given beam-crossing\(^7\). This leads to a small number of events where there are more than 2 $b$-quarks in the event. This effect is small because the $b$-quark production cross-section is very small compared to the total elastic cross-section.

The main kinematic difference between the $B$-daughter background and the fragmentation background is the separation from the primary vertex. The $B$-meson in general is long-lived. Since there is an impact parameter cut in the Level 2 Two-Track trigger, it is expected that the main background in the trigger will be from $B$-daughters (somebody else’s signal).

However, sometimes a fragmentation track will form part of a trigger accept. This will either be because the particle was scattered, or because it really did live long enough to make it past the impact parameter cut. Detector effects cannot be properly estimated in a generator level study, but the contribution of non-$B$ particles which are truly long lived can be estimated\(^8\).

Categorising the two different sorts of background as $t_B$ (tracks from $B$-daughters) and $t_F$ (tracks from underlying event fragmentation), there are the following pure background trigger configurations:

---

\(^7\)See section 2.2

\(^8\)Note that measurement resolution is a detector effect, and is estimated in this study
5.2.4 Mixed Trigger Configurations

As one might expect, the contribution to the total Level 2 Two-Track trigger efficiency in signal events from pure background trigger configurations is fairly small. However there is one more class of trigger configuration to be categorised: a decision where the pair is a mixture of signal and background. The possible combinations are:

1. \((\pi_0, t_B)\)
2. \((\pi_d, t_B)\)
3. \((K, t_B)\)
4. \((\pi_0, t_F)\)
5. \((\pi_d, t_F)\)
6. \((K, t_F)\)

Recall that the two-track trigger requires an impact parameter cut on both tracks in any given pair. Prompt fragmentation tracks will hardly ever pass this cut. Therefore, the contribution from possibilities \((1, 2, 3)\) could be marginally significant whereas the contribution from the possibilities \((4, 5, 6)\) is expected to be very small.

5.3 Global Results

5.3.1 Global Efficiency

The basic result of an efficiency study is the number of signal events that made it through the set of cuts in question. In this case the total fraction of pythia events containing a decay defined by equation 5.3 which made it past the GenTrig Two-Track trigger simulation was \(1.05 \times 10^{-3} \pm 3.62 \times 10^{-6}\).
The pythia sample was generated with the minimum $p_t$ parameter set at 2 GeV. The effect of the non-physical nature of the pythia $p_t$ spectrum is discussed in section 5.1.3.

5.3.2 Efficiency as a function of $B_s$ proper time

Figure 5.4 shows the events generated by pythia for the mode in question as a function of the $B_s$ meson proper time. This was calculated by taking the unsmeared generator level quantities and boosting into the $B_s$ rest frame. The numbers on the plot are from the fit shown. This should be compared to the monte carlo input of 1.551 ps from the QQ-Decay Monte Carlo (the piece of code which actually decays the B-mesons [38]).

Figure 5.5 shows the subsection of these events which passed the Level 2 trigger simulation, while figure 5.6 shows the ratio of these two histograms, which is the Level 2 efficiency as a function of $B_s$ proper time.

It will be noted that the efficiency in figure 5.6 drops off at larger values of the proper time. This is the effect of the upper impact parameter cut in Level 2. This can be seen by comparing figure 5.6 to figure 5.7 which is an efficiency plot generated without the upper impact parameter cut.

The drop-off in efficiency at large proper time will influence a $B_s$ measurement in this mode. The shape of the exponential fall off will be changed by this efficiency curve, and therefore this curve must be modeled accurately if the lifetime is to be measured at all. The curve generated by the current study could be used as a starting point for taking this shaping into account.

---

9 Note: The sample used for this plot and most other plots in this chapter is a representative sample, but not identical to the sample used to calculate the efficiency curve in chapter 6.

10 The small systematic difference from 1.551 ps is due to assuming a $B_s$ mass of 5.3693 GeV in the calculation of the generator level proper time. The correct mass was 5.38 GeV. This introduces a small but systematically negative effect on the lifetime. This mistake was not made for the reconstructed proper time.

11 Note: A much smaller special sample was generated for this plot.

12 It will become clear in chapter 6 that in addition to the lifetime efficiency shape, the lifetime error efficiency shape is also important.
Figure 5.4: The pythia $B_s$ proper time distribution. This distribution is an exponential with the almost the monte-carlo simulated lifetime. See section 5.3.2. The small systematic shift from the correct simulated lifetime is due to an incorrect assumption of the mass in the proper time calculation.
Figure 5.5: The $B_s$ proper time distribution after Level 2. One exponential does not fit over the whole range. The fit shown has a $\chi^2$ of 76 with 50 d.o.f meaning that only 1% of such fits would have a worse $\chi^2$. Furthermore, the value of $(1.149 \pm 0.009)$ ps is totally different to the MC input value of 1.55 ps.
Figure 5.6: The Level 2 Efficiency as a function of $B_s$ meson proper time. In Run I this function would have been flat for much of the range plotted here. However, mainly because of the impact parameter cuts in the trigger it is anything but flat.
Figure 5.7: The Level 2 Efficiency as a function of $B_s$ meson proper time with no upper impact parameter cut. This plot is to be compared with figure 5.6 where the efficiency dies off at higher proper time. Here, the efficiency function for a smaller sample without the upper impact parameter cut can be seen. There is no discernible turn off at high proper time.
5.4 Different Configuration Efficiencies

The following section subdivides the sample that passes the Level 2 trigger according to the trigger configurations introduced in sections 5.2.2 to 5.2.4. For each section of configurations, the fraction of the total Level 2 trigger rate is given, and the Level 2 efficiencies versus proper time are given.

5.4.1 Pure Signal Trigger Configurations

Table 5.1 summarises the contributions of the various pure (signal,signal) configurations to the total Level 2 trigger rate. The entries with the word only in the description refer to trigger decisions which contained only this configuration and no other. Conversely, the entries with the word present in the description refer to the trigger decisions where a pair of this particular configuration was present, but other pairs could also be present. See the aforementioned sections for the definitions of the configurations mentioned in the table below.

The first 4 rows of the table provide a coarse description of how the configurations breakdown. The (sig, sig) only row refers to decisions where there were one or more (signal, signal) combinations in the trigger decision, but no other combinations. The NOT (sig, sig) row refers to decisions which do not contain a pure (sig, sig) combination. The mixed row refers to decisions which contain a mixture of these two sorts of track pairs. The Dimuon row refers to a small subsection of the two track trigger simulation. It is not relevant to the current studies, and therefore only appears in the Level 2 total. It does not appear anywhere else[39].

The first thing to note is that trigger decisions where the only pairs to fire the trigger are pure (signal,signal) account for 85.86% of the total trigger rate. These are in principle the cleanest sorts of events to understand, since modeling the Level 2 proper time efficiency depends on our understanding of how a $B_s$ meson decays into these particular modes. This stands in contrast to the more complicated configurations to come in later sections. The contribution from these more complicated configurations is given by the NOT (sig, sig) row. Its worth noting also that these two sorts of trigger decision largely decouple, as shown by the small percentage in the Mixed row. In principle these 1.5% of events are also relatively straight-forward to include in the efficiency modeling since they contain the simple (sig, sig) combinations.

It is also worth noting that the relative proportions of the various con-
Table 5.1: Table 5.1: Contributions of the various pure \((\text{signal,signal})\) configurations to the Level 2 trigger rate. See section 5.2.1 for the definitions of the trigger configurations. Note that only the first 4 rows of the table add up to the total level 2 trigger rate. The quantities in the second section overlap, while those in the third section are exclusive, but not exhaustive.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Number in L2</th>
<th>Fraction of Level 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\text{sig, sig})) only</td>
<td>73532</td>
<td>85.86 ± 0.32</td>
</tr>
<tr>
<td>NOT ((\text{sig, sig}))</td>
<td>9783</td>
<td>11.42 ± 0.12</td>
</tr>
<tr>
<td>Mixed</td>
<td>1276</td>
<td>1.49 ± 0.04</td>
</tr>
<tr>
<td>Dimuon</td>
<td>1055</td>
<td>1.23 ± 0.04</td>
</tr>
<tr>
<td>((\pi_b, \pi_d)) present</td>
<td>35837</td>
<td>41.84 ± 0.22</td>
</tr>
<tr>
<td>((\pi_b, K)) present</td>
<td>24429</td>
<td>28.52 ± 0.18</td>
</tr>
<tr>
<td>((\pi_d, K)) present</td>
<td>15132</td>
<td>17.67 ± 0.14</td>
</tr>
<tr>
<td>((K, K)) present</td>
<td>13205</td>
<td>15.42 ± 0.13</td>
</tr>
<tr>
<td>((\pi_b, \pi_d)) only</td>
<td>31826</td>
<td>37.13 ± 0.21</td>
</tr>
<tr>
<td>((\pi_b, K)) only</td>
<td>16624</td>
<td>19.41 ± 0.15</td>
</tr>
<tr>
<td>((\pi_d, K)) only</td>
<td>8769</td>
<td>10.24 ± 0.11</td>
</tr>
<tr>
<td>((K, K)) only</td>
<td>6386</td>
<td>7.46 ± 0.09</td>
</tr>
<tr>
<td>Level 2 Total</td>
<td>85646</td>
<td></td>
</tr>
</tbody>
</table>

configurations reflect what we know about the kinematics of the various tracks. For example, the kaons are much softer (lower \(p_t\)) than the \(\pi\) from the \(B_s\). Thus only 7.5% of the trigger rate comes purely from \((K, K)\), whereas 37% comes purely from \((\pi_b, \pi_d)\) (the \(\pi\) from the \(D_s\) is also harder than the kaons).

The plots in figure 5.8 show the event contributions of the configurations where a particular \((\text{sig, sig})\) pair was present. The total Level 2 event distribution is overlayed on each plot for comparison.

The proper time behaviour of the different distributions in figure 5.8 can be understood by consideration of the kinematics of the different configurations. For example, the \((\pi_b, \pi_d)\) configuration has a very low efficiency for low proper time. This is simply the effect of the lower impact parameter cut. The \(\pi_b\) comes from the \(B_s\) meson decay vertex, and for a very low \(B_s\) proper time, the impact parameter is guaranteed to be small, failing the cut. The same behaviour can be seen in the \((\pi_b, K)\) configuration. This stands in contrast to the behaviour of the other two configurations. The effect is
Figure 5.8: The contributions of $(sig,sig)$ present to the Level 2 trigger rate. In the configuration terminology defined in section 5.2.2: Top Left: Config 1, Top right: Config 3, Bottom left: Config 2, Bottom right: Config 4. The x-axis in each plot is the generator level $B_\perp$ proper time.
Figure 5.9: The relative angle in radians between the $\phi$ and the $D_s$. Since the $\phi$ has a bigger mass than the $\pi$ from the $D_s$, its direction is more correlated to the $D_s$ direction. In the same way, the $D_s$ is correlated to the $B_s$ and therefore when the $B_s$ is short lived, the Kaons from the $\phi$ often have short impact parameter (see section 5.4.1)

most marked in the $(\pi_d, K)$ plot, where the distribution effectively plateaus near zero $B_s$ meson proper time. This is because these two tracks come from the tertiary $D_s$ vertex. Even when the $B_s$ has zero proper time, the $D_s$ can still fly a significant distance ($ct = 140 \mu m$)\cite{4}. The drop in efficiency in the $(K, K)$ plot is due to the correlation between the kaon direction and $B_s$ flight direction. The kaon and $\phi$ direction are correlated because the kaons have very little phase space. The $\phi$ and $D_s$ directions are correlated because the $\phi$ mass is very much more than the $\pi_d$ mass (see figure 5.9). The $D_s$ and $B_s$ directions are correlated because the $D_s$ mass is very much more than the $\pi_b$ mass.
5.4.2 Mixed Signal Trigger Configurations

The table below gives the contributions of the configurations where one trigger track was signal \(^{13}\) and the other one was another B-daughter from the same event.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Number in L2</th>
<th>Fraction of Level 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\pi_b, f_B)) only</td>
<td>257</td>
<td>0.30 ± 0.02</td>
</tr>
<tr>
<td>((\pi_d, f_B)) only</td>
<td>66</td>
<td>0.08 ± 0.01</td>
</tr>
<tr>
<td>((K, f_B)) only</td>
<td>85</td>
<td>0.10 ± 0.01</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>408</td>
<td><strong>0.48 ± 0.02</strong></td>
</tr>
<tr>
<td><strong>Level 2 Total</strong></td>
<td><strong>85646</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.2: Table 5.2: Contributions of the various mixed \((\text{signal}, f_B)\) configurations to the Level 2 trigger rate.

It may be seen that the contribution to the total Level 2 trigger rate of these mixed configurations is very small and so these can be thrown out of the total sample without significantly hurting the sample size. This had better be the case, because modeling these events will be a very tough problem. Getting the right efficiency distribution would entail a model of B-fragmentation which was physically accurate.

5.4.3 Pure background configurations

The only configurations left to be considered are the ones where both trigger tracks are background (section 5.2.3). The table below summarises the contributions to the total Level 2 trigger rate of these configurations.

The first row of table 5.3 refers to the total number of trigger decisions where every pair was some combination of background, but no signal was in the trigger (where signal refers to the mode in question, *not* to other B-decays). This accounts for 10.9% of the total trigger rate. Of this, 10.8% is in fact due to configurations where the only pairs present are ones where both tracks come from another B-decay in the event. This is entirely expected because this is someone else’s signal. The significant thing to point out is

\(^{13}\)One of \((\pi_b, \pi_d, K)\)
<table>
<thead>
<tr>
<th>Configuration</th>
<th>Number in L2</th>
<th>Fraction of Level 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(bkgrd, bkgrd) any</td>
<td>9297</td>
<td>10.9 ± 0.11</td>
</tr>
<tr>
<td>(f_B, f_B) only</td>
<td>9259</td>
<td>10.81 ± 0.11</td>
</tr>
<tr>
<td>(f_B, f_X) only</td>
<td>16</td>
<td>0.02 ± 0.01</td>
</tr>
<tr>
<td>(f_X, f_X) only</td>
<td>22</td>
<td>0.03 ± 0.01</td>
</tr>
<tr>
<td>Level 2 Total</td>
<td>85646</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.3: Table 5.3 : Contributions of the various background (bkgrd, bkgrd) configurations to the Level 2 trigger rate. The last row shows the total Level 2 trigger rate for comparison.

that the contribution from decisions where only pairs involving a fragmentation track (f_X) appear in the decision only account for 0.03% of the total trigger rate. In other words this means that it doesn’t matter that detailed information is not available about the fragmentation of the underlying event. The high impact parameter requirements of the trigger cut out virtually all fragmentation tracks.

Figure 5.10 shows the (f_B, f_B) configuration with the total Level 2 shown for comparison (versus proper time).

It may be seen that the lifetime distribution for this type of configuration rises at low proper time. This is because the tracks in these configurations are almost entirely uncorrelated with the B_s proper time because they are from the other B meson in the event. The proper time of one B-meson in an event is uncorrelated to the proper time of the other one. Therefore, the shape of the (f_B, f_B) curve in figure 5.10 should be similar to the exponential decay of the B_s meson.

## 5.5 L2 Efficiency Summary

The bottom line of this study really is that the Level 2 trigger efficiency for the mode under study (B_s → D_s π where D_s → φπ and φ → K K) is a complicated quantity. When one examines how the efficiencies vary for the various different configurations of the trigger tracks, it becomes clear that the sample must be broken down into these various sub divisions. The two main subsamples are the sample where the L2 trigger contains (sig, sig) and the one where the trigger contains (f_B, f_B).
Figure 5.10: Contribution of $f_B, f_B$ only configuration to L2 trigger rate. The x axis is the generator level $B_s$ proper time.
5.5.1 \((f_B, f_B)\) Sub-sample

There are two main disadvantages to this sub-sample. Firstly only a small proportion of the total L2 trigger rate will contain usable events of this type. Secondly it is hard to model accurately in monte carlo.

The way to identify the \((f_B, f_B)\)-only subsample would be to require that none of the \(B_s\) candidate tracks fired the L2 trigger. However, although table 5.3 states that 10.81% of the total signal trigger rate is \((f_B, f_B)\) only, not all of this will be usable. This is because not all of the \(B_s\) tracks will be inside the acceptance.

In addition to this problem, in order to make an efficiency curve calculation, an accurate model of B-fragmentation would be required. For example, the angle between the two B-mesons will influence the efficiency curve, and this is very sensitive to the B-fragmentation model. The same is not true of the \((\text{sig}, \text{sig})\) sub-sample, because the opposite side B-meson does not significantly influence any measured quantity.

However, despite the difficulties outlined above, the \((f_B, f_B)\) sub-sample does have one redeeming feature in that its \(B_s\) lifetime efficiency curve is expected to have a less pronounced shape than the \((\text{sig}, \text{sig})\) sub-sample discussed in the next section (see section 5.4.3). In fact, when a plain exponential is fitted to the lower points in figure 5.10 a lifetime of \(1.48 \pm 0.02\) ps is obtained, which is not far off the input value of 1.55 ps.

5.5.2 \((\text{sig}, \text{sig})\) Sub-sample

This sub-sample accounts for the vast bulk of the usable rate of the level 2 trigger. Of this, the \((\pi_b, \pi_d)\) sample is the largest sub-component (41.84% of the total level 2 trigger rate). This is one of the reasons, the \((\pi_b, \pi_d)\) sub-component was chosen for the analysis of chapter 6.

It will be advantageous to analyse the different sub-components of the \((\text{sig}, \text{sig})\) sub-sample separately. The reason is that each component has a different efficiency shape. If all sub-components were considered together, an error in the relative abundance of each component would lead to a wrong efficiency shape\(^{14}\). However, if each sub-component is analysed separately the relative normalisation between sub-components will not matter.

In the present studies, the level 2 efficiency curve was taken into account by developing a physically motivated parameterisation, and introducing the

\(^{14}\) Each sub-component has a different efficiency shape.
CHAPTER 5. LEVEL 2 TRIGGER EFFICIENCY STUDIES

resulting functional form into a log-likelihood fit. The fit is described in chapter 6. The following sections describe the physical motivation behind the parameterisation adopted for the level 2 efficiency shape\footnote{It will become clear in chapter 6 that the lifetime error efficiency function is also important. A discussion of this is postponed until then}.

5.6 Parameterising the Efficiency Curve

It was decided to limit the focus of the attention to the special case trigger configuration $(\pi_b, \pi_d)$. It is the largest single contribution to the signal rate. Conceptually it is also one of the simpler configurations.

The following section details a derivation of an approximate functional form for the measured Level 2 $B_s$ proper time efficiency function for this configuration.

5.6.1 The dominance of $\pi_b$

It was realised early on that the dominant effect of the level 2 trigger on the $B_s$ proper time would be the impact parameter cuts. Consider the diagram in figure 5.11. A simple schematic of the first stage of the $B_s$ decay is shown. The angle $\theta_L$ describes the direction of the $\pi_b$ relative to the initial $B_s$ direction in the lab. The relationship between $d_0, \tau_p$ and $\theta_L$ is then:

$$d_0 = (\gamma c \tau_p)\sin(\theta_L)$$  \hspace{1cm} (5.4)

where $\gamma$ is the boost of the $B_s$. Thus it is clear that the $\tau_p$ directly influences the impact parameter of the $\pi_b$. Therefore a cut on this impact parameter will directly affect the observed $\tau_p$ spectrum. The situation for $\pi_d$ is somewhat more complicated. Consider the diagram in figure 5.12.

The color coding in figure 5.12 is as follows:

- Black represents physically present quantities (eg directions of $\pi_b$ and $\pi_d$, and the decay distances of the $B_s$ and $D_s$).

- Red represents the construction lines relevant to impact parameters ($d_0$).

- Blue represents the construction lines relevant to the decay angles.
Figure 5.11: Diagram relating $d_0$ of $\pi_b$ to $\tau_p$ and $\theta_L$. All quantities are in the transverse plane.
Figure 5.12: Diagram relating $d_0$ of $\pi_d$ to $\gamma_{\pi}^{B_s}, \gamma_{\pi}^{D_s}$ and the various decay angles $(\theta, \rho, \psi)$. $\theta$ is between $\pi_d$ and $D_s$, $\psi$ is between $B_s$ and $D_s$, and $\rho$ is between $B_s$ and $\pi_d$. All quantities are in the transverse plane.
• Green represents arrows which have no geometrical significance.

Careful study of figure 5.12 indicates that the impact parameter of the \( \pi_d \) is given by the expression:

\[
d_0 = (\gamma_{B_s} c \tau_{D_s}^{B_s}) \sin(\theta - \psi) + (\gamma_{D_s} c \tau_{D_s}^{D_s}) \sin(\theta)
\]  

(5.5)

where \( \theta \) describes the decay direction of the \( \pi_d \) relative to the \( D_s \) direction, and \( \psi \) describes the decay direction of the \( D_s \) relative to the direction of the \( B_s \) (in figure 5.12, \( \rho = \theta - \psi \)). Here, the relationship between \( d_0 \) for the \( \pi_d \) and \( \tau_{D_s}^{B_s} \) is much weaker than in equation 5.4. The angles \( \theta \) and \( \psi \) are independent. The angle \( \psi \) is correlated but not identical to the decay direction of the \( \pi_b \).

Therefore, it is entirely expected that the effect of the impact parameter of the \( \pi_b \) will be much more marked in the \( B_s \) proper time efficiency distribution. In fact, this effect can be clearly seen in figure 5.8, and was remarked on in that section. The configurations which contain the \( \pi_b \) have a much more marked turn-off near zero \( B_s \) proper-time.

In fact the form of the expression for the \( \pi_d \) impact parameter suggests that there could be cancellation between the two terms, which is indeed the case. This means that even for large values of \( \tau_{D_s}^{B_s} \) the \( \pi_d \) impact parameter can still fall inside the upper impact parameter cut.

The result of the above arguments is that the effect of the \( \pi_b \) impact parameter dominates the reconstruction efficiency curve for the \((\pi_b, \pi_d)\) configuration. In fact the effect of the upper impact parameter cut for the \( \pi_d \) will be able to be ignored (see section 5.6.4).

### 5.6.2 The effect of a \( d_0 \) cut on the \( \tau_p \) proper time probability distribution

Consider the diagram in figure 5.13.

The curve is the simple exponential that represents what the \( B_s \) proper time probability density function is known to be before it gets influenced by the level 2 trigger. Qualitatively, it is expected that this probability density will be damped down more at small proper time, than intermediate proper time (the effect of the lower impact parameter cut). This is what the group of arrows on the left of figure 5.13 is meant to represent. The upper impact parameter cut will produce a similar effect, as shown by the group of arrows on the right.
Figure 5.13: Qualitative description of the effect of the impact parameter cut at level 2 on the observed proper time probability density distribution. The area $dS$ is reduced by the effects of the trigger cuts. The arrows indicate this effect on other parts of the exponential. $d\tau_p$ is a proper time infinitesimal.
CHAPTER 5. LEVEL 2 TRIGGER EFFICIENCY STUDIES

This observation will now be made more quantitative. Consider the equation:

\[ d_0 = (\gamma \cos \theta_L) \tau_p \]  \hspace{1cm} (5.6)

This is a slight rearrangement of equation 5.4. Now define:

\[ v = \gamma \cos \theta_L \]  \hspace{1cm} (5.7)

So that:

\[ d_0 = v \tau_p \]  \hspace{1cm} (5.8)

It is clear that the quantity \( v \) is entirely uncorrelated to \( \tau_p \). The proper time at which the \( B_s \) decays is not sensitive to the direction of its decay products, or to the boost of the \( B_s \). Thus, in the absence of the level 2 trigger, all phase-space allowed \( v \) values are permitted. In other words, defining the probability density distribution for \( v \) as \( P(v) \), the integral over all permitted \( v \) values is:

\[ \int P(v) dv = 1 \]  \hspace{1cm} (5.9)

In terms of figure 5.13, it is possible to imagine taking the small proper time element \( d\tau_p \), and integrating over all possible \( v \) in order to obtain the area element \( dS \). In the absence of level 2 \( dS \) would be as it appears in figure 5.13, that is to say it lines up with the unaffected proper time distribution.

However, for any given proper time, \( \tau_p \) the range of allowed \( v \) (neglecting detector smearing) is bounded by \( v_{\text{low}} \) and \( v_{\text{high}} \):

\[ v_{\text{low}} = \frac{d_0(\text{low})}{\tau_p} \]  \hspace{1cm} (5.10)

\[ v_{\text{high}} = \frac{d_0(\text{high})}{\tau_p} \]  \hspace{1cm} (5.11)

where \( d_0(\text{low}) \) and \( d_0(\text{high}) \) are the lower and upper level 2 impact parameter cuts respectively.

The integral:

\[ \int_{v_{\text{low}}}^{v_{\text{high}}} P(v) dv < 1 \]  \hspace{1cm} (5.12)
and when the new area infinitesimal $dS'$ is calculated for a given proper time, it will no longer line up with the curve in figure 5.13, but will lie lower.

If $P(v)$ is properly normalised, this integral of $P(v)$ over the allowed range of $v$ as a function of proper time will yield the proper time efficiency function shape (the resultant function would have to be normalised again).

All that is necessary therefore, is to obtain the functional form of $P(v)$. This problem is addressed in the next section.

### 5.6.3 $v$ Probability Distribution

This probability distribution function is best parameterised by looking at results from monte-carlo. Consider the plots in figures 5.14 and 5.15. This represents a very quick look at the $v$ distribution, with and without one of the other cuts that might influence it (the $p_t$ cut on the $\pi_b$). The basic behaviour of the quantity is unaffected by the 2 GeV cut on the $\pi_b$ $p_t$. The plot in figure 5.16 shows a fit to the second plot in figure 5.16. The intention is not to fit this function exactly (as the chisquare per d.o.f of ~ 3 indicates), but rather to model its overall behaviour.

Therefore, the important point to note from these plots is that the probability distribution for $v$ will be adequately\(^{16}\) modeled by a functional form:

$$P(v) = \frac{A}{(x-B)^2 + 1} \quad (5.13)$$

The great advantage of this functional form, is that it integrates simply to an $arctan$ function.

### 5.6.4 Proper time Efficiency functional form

The loose parameterisation of the $v$ probability distribution means that it is now possible to do the integration in equation 5.12:

$$\int_{v_{low}}^{v_{high}} P(v)dv = \int_{v_{low}}^{v_{high}} \frac{A}{B(v-C)^2 + 1}dv \quad (5.14)$$

When this integration is done, and the functional form of $v \propto \frac{1}{\tau_p}$ is substituted into the limits, the following general functional form for the proper time efficiency is obtained:

\(^{16}\)The turn on, turn off, and central peak of the distribution are all reproduced. An exact correspondance is not necessary.
Figure 5.14: The $\nu$ distribution from Generator level, with no level 2 cuts. Note the small high $\nu$ tail.

Figure 5.15: The $\nu$ distribution from Generator level, with a 2 GeV cut on the $\pi_b$ (see section 5.6.3) Note that the tail from the upper plot is much reduced, but the overall shape is very similar.
Figure 5.16: The $\nu$ distribution with a fit to $A \frac{1}{B(\nu-C)^{2+1}}$ (see section 5.6.3)

$$
\epsilon(\tau_p) = N[\arctan\left(\frac{D_h}{\tau_p} - C_1\right) - \arctan\left(\frac{D_l}{\tau_p} - C_2\right)]
$$

(5.15)

where $D_h$ is a function of the upper impact parameter cut, and $D_l$ is a function of the lower impact parameter cut.

It would be unwise to try and push the preceding arguments too far (eg try to calculate exact values for $D_h$ and $D_l$). The aim of arguing in broad terms about the proper time efficiency function is to obtain a functional form, but not necessarily the parameters in the function. The acid test then is whether this functional form describes the efficiency calculated from monte-carlo.

Since there are two impact parameters in the two-track trigger the efficiency function will be a product of two terms of the form of equation 5.16. As was argued in section 5.6.1 for the ($\pi_h, \pi_d$) configuration, the effect of the $\pi_h$ impact parameter cut dominates the effect of the $\pi_d$ in the efficiency function. Because of the cancellation that can occur in equation 5.5 the effect of the $\pi_d$ upper impact parameter cut can be ignored. This corresponds to an efficiency function of the form:
Example of an Efficiency Function Fit

\[ \chi^2 \text{ per d.o.f.} = 0.82 \]

Figure 5.17: An example of a fit to the efficiency function derived in this chapter. See section 6.2.1 to see this fit in its context.

\[ \epsilon(\tau_p) = N[\arctan\left(\frac{D_h}{\tau_p} - C_1\right) - \arctan\left(\frac{D_l}{\tau_p} - C_2\right)] \times [K - \arctan\left(\frac{D'_l}{\tau_p} - C'_2\right)] \] (5.16)

where the upper impact parameter dependence has been reduced to a constant in the second half of the expression. The first term describes the effect of the impact parameter cut on the \( \pi_b \) and the second term describes the effect of the impact parameter cut on the \( \pi_d \).

As an example of how this expression fits to an efficiency histogram consider figure 5.17. The \( \chi^2 \) per d.o.f is \( \chi^2 \) per d.o.f = 0.82. A more extended discussion of how this function was used in the analysis at hand is given in section 6.2.1.\(^{17}\)

\(^{17}\) The interplay between the two turn-on terms can be seen in the figure in the range between 1 and 2 ps.
Chapter 6

Likelihood Fit to the $B_s$ lifetime

In the first part of the following chapter the Unbinned Log Likelihood method for hypothesis testing is described. In the second part, the specific form of the likelihood function used in the present studies is given. Finally, the results of the likelihood fit in the present studies are given.

6.1 Unbinned Log Likelihood Hypothesis Testing

Consider an experiment which yields a set of data:

$$x_i \; i = 1, 2...N$$

where $N$ is the number of data points.

It is hypothesised that this data is described by a function $F(x)$. This hypothesis may be tested by forming a likelihood function which describes the probability for a given $x_i$, that the function $F(x)$ does indeed describe this data point. The probability that the whole data set fits to this distribution is then:

$$P_{tot} = \prod_{i=1}^{N} P(x_i)$$

It will be noticed that there is no need to divide the data points into bins using this method. The probability for each data point is calculated independently. This is why the method is called unbinned.
Figure 6.1: The probability function for a gaussian hypothesis. The gaussian is both the hypothesised function and the probability function.

### 6.1.1 The Probability Function

The best function to use as the probability function turns out to be the hypothesised function itself, in this case $F(x)$. Consider the simple example of a gaussian (figure 6.1). The two blue arrows in the figure represent two data points. If the data points follow the gaussian hypothesised, they are expected to populate the region of the x axis under the gaussian (eg the arrow on the right). However, if the hypothesis is a bad one (for example if the gaussian has far too small a width) then the data points may not populate only the x axis under the gaussian (eg the arrow on the left). The probability function evaluated at the two arrows reflects this discrimination in the case that it is the gaussian. In this case, it will be very small for the left hand arrow, but much larger for the right hand one.
6.1.2 Parameter Fitting

The process of finding the best hypothesis function $F(x)$ consists of maximising the probability defined in equation 6.2. Consider the case where the function $F(x)$ is dependent on a parameter $p$ so that:

$$F(x) \rightarrow F(x, p)$$

(6.3)

The probability in equation 6.2 may be maximised by recalculating the total probability for different $p$ values. This is how the best parameter values for the fit are decided. There is the possibility of course that the functional form doesn’t fit very well for any value of the parameter $p$. This will be discussed more in section 6.1.4.

It should be emphasised that the probability function must be correctly normalised to within a global constant factor. This is necessary because if the normalisation depends on the parameters of the function, it must vary as the parameter is varied. For example, in the case of a gaussian, the normalisation depends on the width of the gaussian.

6.1.3 Minimising the Negative Log Likelihood

It is inconvenient to calculate the product in equation 6.2, so this is transformed to a sum by taking the logarithm. This makes no difference to the position of the maximum sought.

Further, it is convenient to minimise rather than maximise the likelihood quantity, and so the final likelihood quantity is defined as:

$$-1 \times \sum_{i=1}^{N} \log(P(x_i))$$

(6.4)

This quantity is then minimised using the CERNLIB package MINUIT[46].

6.1.4 Testing the Goodness of Fit.

The following method is the one presented in Dr L Lyons book [47].

The simplest way, if not the quickest, of finding out how well the function fits to the data is to carry out the fit many times with toy monte-carlo data sets.
The minimisation will yield a final best probability for the fit, and it is this probability which must be judged.

Therefore, many data sets are generated according to the fitted distribution, and the negative log likelihood minimisation is performed for each. Since these data sets are known to follow the fitted distribution, the distribution of minimised negative log likelihoods will provide an idea of how the actual fit should have turned out. If the minimum negative log likelihood from the original fit falls reasonably within this generated distribution of likelihoods, then it can be inferred that the fit is good. However, if it is much larger than the mean of the generated distribution, then the chances are that this functional form does not fit well to the data.

6.2 The Proper Time Likelihood Function

Recall the likelihood function defined in equation 3.11:

\[ \mathcal{L} = F^{\text{sig}}(M^{\text{sig}}T^{\text{sig}}) + (1 - F^{\text{sig}})(M^{\text{bck}}T^{\text{bck}}) \]  \hspace{1cm} (6.5)

All the quantities in equation 6.5 differ for each stage of minimisation. Note that \( M^{\text{sig}}, T^{\text{sig}}, M^{\text{bck}}, T^{\text{bck}} \) are evaluated for each data point in a given stage of minimisation whereas \( F^{\text{sig}} \) is not.

The expression in equation 6.5 subdivides as follows:

- Signal
  - Mass Hypothesis Probability : \( M^{\text{sig}} \)
  - Proper Time Hypothesis Probability : \( T^{\text{sig}} \)

- Background
  - Mass Hypothesis Probability : \( M^{\text{bck}} \)
  - Proper Time Hypothesis Probability : \( T^{\text{bck}} \)

\( F^{\text{sig}} \) parameterises the abundance of the signal within the sample, so that the abundance of the background is given by \( (1 - F^{\text{sig}}) \).

The reason the mass distribution appears in what is supposed to be a lifetime fit is because it provides an excellent way of discriminating between signal and background. Consider a data point which would give similar values
for the signal and background proper time probability hypotheses. The mass probability hypothesis will then provide needed extra discriminatory power to make the two probabilities different. This is especially important in an analysis where the background proper time distribution is somewhat similar to the signal proper time distribution. In the current studies there are regions of proper time where the two proper time distributions follow each other. This is mainly at larger proper time ($\gtrsim 1\text{ps}$).

For each data point, there are four stochastic variables:

$$(\tau_i, \sigma_{\tau_i}, m_i, \sigma_{m_i})$$

where $\tau_i$ is the proper time, $\sigma_{\tau_i}$ is its error, $m_i$ is the reconstructed $B_s$ mass, and $\sigma_{m_i}$ is its error.\(^1\)

Each of the terms in equation 6.5 is dependent on some combination of these stochastic variables, parameters in the fit (e.g. the $B_s$ lifetime), and constants calculated prior to the fit. The functional form of the component terms is now given.

### 6.2.1 Signal Proper Time Probability Function

The theoretical distribution for the proper time is an exponential. However, there are two physical effects to this distribution which must be parameterised in order to describe what is actually observed.

- The resolution of the detector.
- The efficiency of the level 2 trigger.

The effect of detector resolution is taken into account by convoluting the exponential with a gaussian, whose width is described on an event by event basis by $\sigma_{\tau_i}$. The effect of the level 2 trigger is taken into account by multiplying the convoluted exponential by the efficiency function calculated in chapter 5. The efficiency function in the equations below includes both the effect of level 2 and the reconstruction cuts. The effect of level 2 is dominant, but it is conservative to calculate the histogram after reconstruction to take into account any small extra efficiency effects. This yields the following expression:

\(^1\)The subscript $i$ runs over the data points.
\[ T_i^{\text{sig}}(t_i, \sigma(t_i), \lambda_{B_*}) = \epsilon(t_i) \times \frac{1}{\sqrt{2\pi} \sigma(t_i) \lambda_{B_*}} \int_0^\infty \exp\left(-\frac{(t - t_i)^2}{2\sigma(t_i)^2}\right) \exp\left(-\frac{t}{\lambda_{B_*}}\right) dt \]

When the integration is done the following expression is obtained:

\[ T_i^{\text{sig}}(t_i, \sigma(t_i), \lambda_{B_*}) = \epsilon(t_i) \times \frac{1}{2\lambda_{B_*}} \exp\left(\frac{-\sigma(t_i)^2}{2\lambda_{B_*}^2} - \frac{t_i}{\lambda_{B_*}}\right) \left[1 - \text{erf}\left(\frac{\sigma(t_i)}{\sqrt{2}\lambda_{B_*}} - \frac{t_i}{\sqrt{2}\sigma(t_i)}\right)\right] \]

Normalisation of this quantity is done numerically.

### 6.2.2 \( \sigma(t) \) dependence of the efficiency curve

Consider the simple case where every event generated gives rise to a reconstructed proper time with the same uncertainty. The sub-sample after all cuts would have the same uncertainty. Therefore an efficiency function calculated as:

\[ \epsilon(t) = \frac{N_{\text{post-cuts}}(t)}{N_{\text{pre-cuts}}(t)} \]  

would correctly normalise the distribution after all cuts back to the distribution before all cuts.

Now consider the situation shown in figure 6.2.

The two curves shown are exponential-gaussian convolutions. Each convolution contains an exponential with the same width (1.55 ps). But whereas the dotted blue convolution contains a gaussian with a width of 0.1 ps, the red convolution contains a gaussian with a width of 0.33 ps. The shapes of the convolutions are distinctly different\(^2\).

The distribution of errors before level 1 is observed to be significantly different to after (see figure 6.5). The dotted blue convolution thus represents what the post-cuts distribution should look like once it has been corrected for the efficiency factors under discussion. The red convolution represents the initial distribution.

The efficiency function defined in equation 6.9 is calculated by dividing the mutated blue convolution (as it appears after cuts) by the red convolution. Therefore, the efficiency function will normalise to the red shape, but

\(^2\)The area under each convolution curve is 1.0
Figure 6.2: The two convolutions (Exponential $\otimes$ Gaussian) represent smeared lifetime exponentials. Both have exponential width of 1.55 ps. The red solid line has a gaussian width of 0.33 ps, and the blue dotted line has a gaussian width of 0.1 ps. This illustrates how a difference in lifetime error distribution before (red) level 1 and after (blue) level 2 means that a single efficiency curve cannot be used to correct for the trigger.
Figure 6.3: The proper time uncertainty distributions before and after the trigger. The distributions peak in a slightly different place. Both fall very steeply after about 0.1 ps, but the post level 2 distribution falls much more steeply than the pre level 1 distribution.

with the blue error distribution. In other words, the function in the likelihood fit will have the shape of the red distribution, but the gaussian width of the blue. Therefore it will yield the wrong exponential width.

The cleanest way to correct for this problem is to calculate the efficiency function as a function of $\sigma(t)$. In other words, the efficiency function would become 2 dimensional, depending both on $t$ and its uncertainty. However, this is impossible in practice because it would require impossibly large statistics.

The proper time uncertainty distributions before and after the trigger can be seen in figure 6.3. Each curve rises very sharply, and falls very sharply, but the post level 2 distribution falls much quicker than the pre level 1 distribution. While there are significant numbers of events after 0.15 for the post level 2 distribution, the vast bulk is concentrated below about 0.15. For this reason it is sufficient to choose a small number of bins in the range below
Figure 6.4: The ratio of the two plots in figure 6.3: The proper time uncertainty distribution after Level 2 divided by the proper time uncertainty before Level 1. The rise and fall in ratio is extremely sharp, and the fall off is almost exponential. The ratio remains non zero until about 1.0 ps.
Figure 6.5: The ratio of the proper time uncertainty distribution after reconstruction cuts and the proper time uncertainty distribution before Level 1. The structure is very similar to figure 6.4. A narrower range on a linear scale is shown here. The range for the cut on proper time uncertainty (0.075 $\rightarrow$ 0.117) is such that it encompasses most of the high ratio region where the trigger is most efficient.
0.15 and calculate a t-dependent efficiency function for each. An obvious extension of the current studies is to calculate finer bins. Figures 6.4 shows the ratio of proper time uncertainty distributions before level 1 and after level 2. Figure 6.5 shows the ratio of proper time uncertainty distributions before level 1 and after reconstruction. All of these plots are signal monte-carlo.

The exact \( \sigma(t) \) range chosen was 0.075 \( \to \) 0.117. This was split into three equal width bins. The resulting efficiency histograms can be seen in figures 6.6 and 6.7. The shape of the efficiency histograms becomes flatter with increasing \( \sigma(t) \). There is no discernible plateau in the plot of the first bin (0.075 \( \to \) 0.089), there is the beginnings of one in the plot of the second bin (0.089 \( \to \) 0.103), and there is quite a pronounced plateau in the plot of the third bin (0.103 \( \to \) 0.117). This is probably because the events with larger lifetime uncertainty are less sensitive to the upper impact parameter cut. Larger lifetime uncertainty usually arises because the tracks constrained to a vertex are fairly parallel. This situation arises when the tracks are high momentum and often point back towards the primary. Their impact parameters are on average less, and therefore even large proper time events can make it passed the upper impact parameter cut.

For each histogram in figures 6.6 and 6.7 a fit to the function derived in chapter 5 is shown. The function for the first two histograms is the product of two terms of the form of equation 5.16. Each term represents one impact parameter. Because the \( \pi_d \) is less sensitive to the upper impact parameter cut, this half of that term is a constant. The resulting functional form for the first two bins is then:

\[
\epsilon(\tau_p) = N[\text{atan}(\frac{D_1}{\tau_p} - C_1) - \text{atan}(\frac{D_2}{\tau_p} - C_2)] \times [K - \text{atan}(\frac{D_3}{\tau_p} - C_3)]
\]  

(6.10)

The third bin did not fit to this function probably because of its lower statistics, and so the simpler function:

\[
\epsilon(\tau_p) = N[\text{atan}(\frac{D_1}{\tau_p} - C_1) - \text{atan}(\frac{D_2}{\tau_p} - C_2)]
\]  

(6.11)

was used.

Note also that the rising behaviour at low proper time is very well modeled by the function. The third bin fit was to the form of the efficiency

\footnote{The possible bias introduced by only fitting for part of the error distribution is discussed in section 7.2}
function which only has two arctan functions. This has a fairly simple rising behaviour.

Consider the form of the function in equation 6.10. The two turn-on arctans are the 2nd and 3rd in equation 6.10, while the turn-off arctan is the 1st. This gives significant insight into the influence of the various impact parameter cuts in this mode.

The last plot in figure 6.7 shows all three curves together. This clearly shows that they have different shapes. This is the clearest evidence that the efficiency function needs to be binned in $\sigma(t)$.

The plots in figure 6.8 show the quantity:

$$\text{Pull} = \frac{\text{curve} - \text{histogram}}{\text{error}}$$

(6.12)

for the centres of the non-zero bins. The functions appear to fit the rising part of the curve better than the rest.

### 6.2.3 Background Proper Time Probability Function

The background proper time probability function was estimated by looking at the background distribution before some of the cuts had been applied. This was necessary because of lack of statistics. It was observed that the sample of 133 events discussed in section 4.1.3 fit reasonably well to a decaying exponential with width $1.34 \pm 0.15\,ps$. Without more background statistics, this is the best that can be done. It is not expected that this decaying exponential model will be too wrong because most of the important cuts have been applied by the time the 133 events is obtained. The only cuts left are the $B_s$ mass cut, the $t$ and $\sigma(t)$ cuts, and the requirement that the $(\pi_b, \pi_d)$ candidate pair be in the trigger.

The background proper time probability function $T_i^{bck}$ is thus given by:

$$T_i^{bck} = \frac{1}{\lambda_{bck}} \exp\left( -\frac{t_i}{\lambda_{bck}} \right)$$

(6.13)

where $\lambda_{bck}$ is left as a parameter in the fit.
Figure 6.6: Reconstructed proper time distributions for the first two of three proper time uncertainty bins. The first bin is: \((\sigma_t = 0.075 \rightarrow 0.089)\). The second is \((\sigma_t = 0.089 \rightarrow 0.103)\). A fit to the efficiency functions derived in chapter 5, is shown for each distribution. For bin 1, \(\chi^2 = 0.82\) per d.o.f. corresponds to the statement that 63\% of such fits have a worse \(\chi^2\). Similarly for bin 2, 32\% of such fits have a worse \(\chi^2\).
Figure 6.7: The proper time distribution for the 3rd proper time uncertainty bin ($\sigma_t = 0.103 \rightarrow 0.117$). The slightly poor $\chi^2$ per d.o.f is partly due to the one point at about 2 ps which is probably statistical fluctuation. The fit would probably be better if there were more statistics and a more complicated function could be used. This $\chi^2$ corresponds to the statement that 2% of such fits have a worse $\chi^2$. The lower plot shows the functions fit to all three of the bins on the same scale. Each bin higher in proper time uncertainty has a more pronounced plateau.
Figure 6.8: The quantity $\text{Pull} = \frac{\text{curve-histogram}}{\text{error}}$ is plotted for the three proper time uncertainty bins. Zero bins are suppressed. It seems that the functions fit the early behaviour of the distributions better than the later behaviour. Most of the points fall within the $(1 \rightarrow -1)$ range expected.
6.2.4 Signal Mass Probability Function

The signal mass probability function is given by the expression:

\[
M^s_{i} = \frac{1}{\sqrt{2\pi}\sigma_{m_{i}}} \exp\left(-\frac{(m_{i} - M_{B_{s}})^2}{2\sigma_{m_{i}}^2}\right)
\]  \hspace{1cm} (6.14)

where the width of the gaussian enters on an event by event basis. Now there is a complication with the trigger simulation sample here. It was expected that all the mass errors would cluster around a similar value, however figure 6.9 shows the signal mass distribution with a fit to 2 gaussians. The widths of these two gaussians from the fit are $10 \pm 0.7 \text{MeV}$ and $3.3 \pm 0.3 \text{MeV}$. This combination of gaussians does not affect the final fit in any way since the width of the probability gaussian enters on an event by event basis, and the errors do reflect the mass distributions. The reason for this 2 gaussian structure is that the mass error distribution is a reasonably continuous spectrum. A distribution of this sort will have a good fit to 2 gaussians of different widths.

6.2.5 Background Mass Probability Function

This is assumed to be flat over the small range of candidate $B_{s}$ masses allowed\(^4\), and so the probability function is given as:

\[
M^{bck}_{i} = \frac{1}{m_{\text{high}} - m_{\text{low}}}
\]

\hspace{1cm} (6.15)

where $m_{\text{high}}$ and $m_{\text{low}}$ are the highest and lowest mass in the sample, thus taking into account the necessary normalisation.

6.3 Results

The final fit was carried out with the s:b ratio of 0.5. This means that the fraction of the signal should be 33\%. The results for the fit are shown in table 6.1\(^5\).

\(^4\)5.34 GeV \rightarrow 5.42 \text{ GeV}

\(^5\)Note that the Minos errors are asymmetric errors calculated without assuming a parametric shape to the likelihood function in the region of the minimum. Rather, the likelihood function is sampled in the space around the minimum to calculate the uncertainty.
Figure 6.9: Signal $B_s$ candidate invariant mass distribution. The fit is to two gaussians with widths $10 \pm 0.7\,\text{MeV}$ and $3.3 \pm 0.3\,\text{MeV}$. This combination of gaussians will not affect the final fit since the width enters on an event by event basis.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Central Fit Value</th>
<th>Minos Neg Uncertainty</th>
<th>Minos Pos Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{B_s}(ps)$</td>
<td>1.460</td>
<td>-0.058</td>
<td>0.062</td>
</tr>
<tr>
<td>$\lambda_{bck}(ps)$</td>
<td>1.318</td>
<td>-0.020</td>
<td>0.020</td>
</tr>
<tr>
<td>$F^{sig}(Fraction)$</td>
<td>0.35</td>
<td>0.008</td>
<td>0.008</td>
</tr>
<tr>
<td>$m_{B_s}(GeV)$</td>
<td>5.38</td>
<td>-0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 6.1: Signal to background 1:2 Likelihood fit results.

Finally, figure 6.10 shows the shape of the likelihood function in the neighbourhood of the minimum for the average proper time parameter.

### 6.4 Goodness of Fit

The goodness of fit was evaluated by applying the method discussed in section 6.1.4. A toy Monte-Carlo of 30 experiments was done. The resulting spread in the likelihood function values can be seen in figure 6.11. The real likelihood function value is shown by the arrow and is well inside the range of generated values. The fraction of the toy experiments which had a worse fit than the full Monte-Carlo is approximately 70%.
Figure 6.10: Shape of the Likelihood function in the neighbourhood of the proper time minimum. The statistical error is about 0.06 ps.
Figure 6.11: The spread in Likelihood Function values for 30 toy monte-carlo data sets with the fitted parameters as inputs. The real likelihood function value (the arrow) is well within the above range. This plot implies that approximately 70% of the toy data sets fit worse than the full monte carlo set.
Chapter 7

Systematic Uncertainties

The following chapter gives an account of the important systematic effects in the present analysis. Each effect is explained and estimated where possible.

7.1 Uncertainty in the Efficiency Function

The efficiency functions derived in chapter 5 are calculated by fitting to efficiency histograms. The errors on the fitted function parameters are not a very good way of estimating how the uncertainty in the function will influence the fit. This is because the different parameters are very correlated in the fit.

Instead, a toy monte-carlo of efficiency histograms is generated. This is done by using the central value and binomial uncertainty of each bin of the original histograms to define a set of binomial distributions for the new ones. A binomially distributed random deviate is then generated for each bin of the new histograms. The efficiency functions are then fitted to these dummy histograms to evaluate the effect of not knowing the efficiency histograms exactly. In this way, the correlations between the efficiency function parameters are correctly taken into account.

The toy efficiency functions resultant from this procedure are then used to fit the data sample. A spread of lifetime central values is obtained, and a systematic uncertainty may be inferred. This spread is shown for a toy monte carlo of 500 experiments in figure 7.1.

A systematic effect of \((2.07 \pm 0.07) \times 10^{-2}\) ps is therefore inferred. This is probably one of the bigger systematic effects and can be reduced by cal-
Figure 7.1: The spread in the central value of the fitted lifetime for different generated efficiency histograms. This shows the systematic effect of not knowing the efficiency functions precisely.
culating the efficiency functions with more statistics.

7.2 Effect of the cut on Proper Time Uncertainty

The cut on the $B_s$ proper time uncertainty will almost certainly improve the signal to background ratio, but it may also introduce a bias into the measurement of the lifetime. It is not inconceivable that a $B_s$ that flies further would have a different uncertainty distribution from one that does not fly as far. Unfortunately, it is very difficult to investigate the events with very large proper time uncertainty because there aren’t many of them. This means that an extremely large amount of monte-carlo statistics would have to be collected to measure efficiency functions in this region. Furthermore the events outside the proper time uncertainty cut are spread thinly. The three bins used in the analysis in chapter 6 span a range of 0.042 $ps$ and contain 76% of the events over the entire proper time uncertainty range. The remaining 24% of events are spread over a region approximately 24 times as wide. This means that in order to obtain comparable statistics in the sparse regions of uncertainty space, while keeping the narrow bin size, about 60 efficiency functions would have to be generated, requiring about 180 times as much statistics as was available in the present analysis.$^1$

The best that could be done to investigate this potential source for systematic error was to lump all the events above an uncertainty of 0.117$ps$ into one bin and calculate one single 4th efficiency function. A 0th efficiency function was also calculated for the $\sim 4\%$ of events with an uncertainty below 0.075$ps$. Efficiency functions were calculated for these bins, and the entire proper time uncertainty range was fitted. The result was $\lambda = 1.519 \pm 0.037\ ps$. This is significantly higher than the central value obtained with just the central three proper time uncertainty bins. Furthermore, when the 4th bin was used by itself, a value of $\lambda = 1.634 \pm 0.072\ ps$ was obtained.

The problem with employing this technique however is that it is very sensitive to the large uncertainty bin width of the 4th bin. It was consistently observed that when the entire uncertainty range was lumped together into one bin, and a single efficiency function was calculated that the central value was systematically too high. For example, for one monte-carlo set, a value of

$^1180 = \frac{76}{24} \times 60$
\[ \lambda = 1.747 \pm 0.052 \] was obtained when lumping all proper time uncertainties together in one bin.

Therefore, it is entirely possible that some of the tendency for the 4th bin to pull the central value high could be due to the width of the 4th bin \((0.117 \rightarrow 1.0 \text{ ps})\). However, it is also possible that some of the effect could be a real bias due only to not fitting using the entire uncertainty distribution.

The best that can be done with the current statistics is therefore to quote the difference between the central values obtained using the 5 bins and using only the narrow central 3 as a systematic. This difference is: \[ \delta(\lambda) = 1.519 - 1.460 = 0.059 \text{ ps}. \]

### 7.3 Fitted Range of Proper Time

The central value of the lifetime is observed to depend slightly on the start and end lifetime values in the fit. This effect can be seen for two different sets of efficiency functions in figure 7.2. Three efficiency functions in the \( \sigma(t) \) range were used for the upper plot. Two functions were used for the lower plot. The standard deviation of central values is approximately the same for each \((6.95 \times 10^{-3} \text{ and } 5.93 \times 10^{-3})\). The arithmetic mean is thus \( 6.44 \times 10^{-3} \). This is an extremely small effect. Furthermore, it is difficult to tell whether it is systematic or statistical in nature. It will therefore be omitted from the quoted systematic error. For better statistics, this effect may become more important.

### 7.4 Effect of the Size of \( \sigma(t) \) bins

The plot in figure 7.2 can also be used to estimate the effect of not binning \( \sigma(t) \) finely enough. There is a systematic trend between the 3 and 2 \( \sigma(t) \) bin cases in the figure. This could be due to two things. Firstly, because the efficiency histograms are binned differently in the two fits to the efficiency functions, the statistics are distributed differently. This was the effect estimated in section 7.1. However, this systematic trend could also be due to the different \( \sigma(t) \) bin widths. The average difference between the central values for the 2 and 3 bin efficiency function case is \( 0.012 \text{ ps} \). It is therefore prudent to assign this as the systematic uncertainty due to the \( \sigma(t) \) bin width.
Figure 7.2: The spread in lifetime central values of the lifetime for different fit ranges is shown. The bars show the parabolic errors for each data point. The red line is for reference, and is at 1.46 ps on each plot. The x-axis is the end point of lifetime in the fit. The start point is always 0.1 ps.
7.5 Bias due to assuming a lifetime central value

In a real analysis, the true value of the lifetime will not be known. However, in order to calculate the efficiency curves, a value must be assumed. It is expected that in an efficiency expression:

\[ \epsilon = \frac{N_{\text{fit}}(1.55 \text{ ps assumed})}{N_{\text{true}}(1.55 \text{ ps assumed})} \]  

(7.1)

that the any dependence on the assumed value for the proper time will cancel. In other words, it is expected that the probability of seeing an event of given reconstructed proper time does not depend on the average proper time of the distribution of events in the whole detector. There are very good reasons to believe this in the absence of some of the other systematic effects in this analysis. For example, the finite \( \sigma(t) \) bin width discussed in the previous section could lead to a dependence of the calculated efficiency function on the assumed average lifetime.

At the very least, this effect should be investigated. Unfortunately, that requires many different data sets to be generated, and is well outside the scope of the current studies. All that was done was to generate a separate data set with average proper time 1.7ps and fit using the 1.55ps efficiency functions. A value of 1.61 \( \pm 0.045 \) ps was obtained. When systematic effects are taken into account this is a little under 2\( \sigma \) away from the 1.7ps central value. This is not far enough away from 1.7ps to indicate a definite effect. All it does show is that at least to first order, the assumed value for the average proper time does cancel from the efficiency function.

If this does prove to be a problem in a higher statistics analysis, the effect of assuming the lifetime can investigated and reduced by iteratively following the fitted lifetime value with the assumed value in the efficiency function generation.

7.6 \( B_s \) meson \( p_t \) spectrum

Finally in this chapter, the possible inadequacy of the simulation is discussed. The primary shortcoming of the simulation is the momentum dependence of the production cross-section of the \( B_s \) mesons. This is related to the momentum production cross-section of the b-quarks via the Peterson fragmentation
function. It is possible to measure this production cross-section down to about 2 GeV at the absolute minimum\textsuperscript{2}. To be conservative it is assumed that it is only possible to measure the $B_s$ meson production spectrum down to 3 GeV.

The possible effect of not knowing this spectrum below 3 GeV was investigated by weighting the events with $B_s$ $p_t$ in the range 0 $\rightarrow$ 3GeV by a factor:

$$y = 10 - 9x$$ (7.2)

where $y$ is the weight factor at a $p_t$ of $x$. This weights a 3GeV $B_s$ with 1.0, and a 1GeV $B_s$ with 9.0.

It is already known from section 5.1.3 and especially figure 5.3 that level 2 is very inefficient at accepting low $p_t$ B-quarks. This will propagate through indirectly to the $p_t$ of the B-Mesons. Therefore, weighting low $p_t$ B-Mesons down would only serve to make a small effect less. Rather, they need to be weighted up in order to see if they influence the efficiency functions.

However, absolutely no discernible effect was observed in the parent histograms to the efficiencies. This effect is not judged to be important.

### 7.7 Summary

Table 7.1 summarises the systematic effects discussed above with their estimated values.

<table>
<thead>
<tr>
<th>Systematic Uncertainty</th>
<th>Estimated Effect (ps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency Function</td>
<td>0.02</td>
</tr>
<tr>
<td>$\sigma(t)$ cut bias</td>
<td>0.059</td>
</tr>
<tr>
<td>Range of Proper Time</td>
<td>N/A</td>
</tr>
<tr>
<td>Size of $\sigma(t)$ bins</td>
<td>0.012</td>
</tr>
<tr>
<td>Assumption of Central Value</td>
<td>N/A</td>
</tr>
<tr>
<td>$B_s$ meson production spectrum</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 7.1: Summary of the studied systematic effects.

\textsuperscript{2}At 2 GeV, it is kinematically possible to reconstruct 4500 MeV tracks at CDF
The systematic uncertainty due to the lack of knowledge of the efficiency functions (0.02 ps) is essentially another statistical uncertainty. With enough statistics, this source of uncertainty would effectively go away. This source of uncertainty is symmetric.

The uncertainty due to the finite proper time uncertainty bin size (0.012 ps) may not be symmetric. But there simply weren’t enough statistics to tell, so it will be lumped together with the 0.02 ps.

Finally, the uncertainty due to not including the whole proper time uncertainty distribution (0.059 ps) is likely to be asymmetric. It is not possible to subtract off this effect with the current statistics because of the possible overlap of the effect of the 4th bin width.

\footnote{The question as to whether one 	extit{trusts} the efficiency functions is a separate issue.}
Chapter 8

Conclusions

8.1 Discussion of Results

The result of the present analysis performed on the $B_s$ lifetime as measured in the $B_s \rightarrow D_s \pi$ where $D_s \rightarrow \phi \pi$, and $\phi \rightarrow K^+ K^-$. Is:

$$\lambda = 1.460^{+0.062}_{-0.058} \pm 0.023 \text{ ps (syst1)} + 0.059 \text{ ps (syst2)} \quad (8.1)$$

Here (syst1) is the combined effect of the lack of accuracy of the efficiency functions, and the possible effect of the proper time uncertainty bin width. The second systematic (syst2) due to not using the whole proper time uncertainty distribution is quoted as positive, but could contain an appreciable contribution from the large width of the 4th proper time uncertainty bin.

If the effect of (syst2) is added to the central value, and the statistical (syst1) are added in quadrature, the central value lies 0.5 standard deviations away from the input monte-carlo value of 1.551ps. However, as has already been stated, there is an uncertain bias as part of (syst2) due to the 4th bin. Therefore, the measurement is probably not quite as close as this to the input monte-carlo value.

If the total statistical-like error is combined as\(^1\):

$$\sigma(\lambda)_{TOT} = \sqrt{(0.06^2 + 0.023^2)} = 0.064 \quad (8.2)$$

Then expressed as a fraction of 1.55 ps, this is $\frac{\sigma(\lambda)}{\lambda} = 0.042$.

\(^1\) (syst2) will not be taken into account in the following discussion since it is asymmetric, and could in principle be subtracted off with sufficient knowledge of the proper time uncertainty distribution, and associated proper time efficiency functions.
CHAPTER 8. CONCLUSIONS

The standard model predicts that $\Delta \Gamma$ should be $\Delta \Gamma = 0.12 \pm 0.06[48]$. Given this central value, and assuming that the lifetime measured in the other sample$^2$ could be measured with the same accuracy as the present analysis, this would enable approximately a $2\sigma$ statement to be made.$^3$ New physics would be seen in a smaller value of $\Delta \Gamma[50]$.

The signal sample size used for this analysis compares reasonably well with the expectation for Run 2a. The number of reconstructible events expected in the channel $B_s \rightarrow D_s \pi$, where $D_s \rightarrow \phi \pi$, and $\phi \rightarrow K^+K^-$ is $16100[34]$. If the reconstruction efficiency from the current analysis is applied to this prediction, 1352 events would be seen. This compares to a sample size used in this analysis of 3293 events.

However, this assumes that only the trigger configuration ($\pi_b, \pi_d$) is used. If table 5.1 is used to extrapolate to using all of the exclusive (sig,sig) configurations, the number of expected events rises to 2353. In this case, the 0.064 $ps$ error from this analysis scales to 0.076 $ps$ which would be a $\frac{\sigma(\Lambda)}{\Lambda} = 0.05$ relative measurement of the $B_s$ lifetime.

These predictions assume the analysis is confined to the mode of interest in these studies. If other $B_s$ modes and $D_s$ modes are considered, the expected number of events after optimised cuts rises to 15040. This would result in a $\frac{\sigma(\Lambda)}{\Lambda} = 0.02$ relative measurement of the $B_s$ lifetime.$^4$ Furthermore, after the 2 $fb^{-1}$ of integrated luminosity expected in Run 2a, there are a further 15 $fb^{-1}$ expected in Run 2b.$^5$ Therefore, all statistics will improve by about a factor of 7.

It becomes clear then that the limiting factor of a measurement of $\Delta \Gamma$ will probably not be the mixed CP sample. Rather it will be the CP eigenstate sample. The best example is probably $B_s \rightarrow J/\psi \phi$. Here the contribution to the uncertainty on $\Delta \Gamma$ depends on the CP content of the mode. The Run 2a ($2 fb^{-1}$) prediction here is an uncertainty on $\frac{\Delta \Gamma}{\Gamma}$ of 0.05$[51]$. The best case for new physics in Run 2a would therefore be a measurement of $\frac{\Delta \Gamma}{\Gamma} = 0 \pm 0.05$, which would be a 2 sigma statement. A higher central value would weaken this statement.

$^2$ie the other sample in the two sample method discussed in chapter 3. $B_s \rightarrow J/\Psi \phi$ is a good example.

$^3$ie assuming $\Delta \Gamma$ was measured to be 0.1, the statement that $\Delta \Gamma$ would be non-zero would be at the 2 sigma level.

$^4$Note that scaling from the error of the current analysis assumes a signal to background ratio of 0.5

$^5$Expected to begin in approximately 2004.
**Chapter 8. Conclusions**

![Figure 8.1: The fits to the efficiency function from chapter 6. The three different efficiency functions are fits to the three different proper time uncertainty bins of the analysis.](image)

**8.2 Epilogue**

Consider the plot in figure 8.1. One of the most significant discoveries of the present work was the form and dependence of the efficiency functions shown. In particular it is very surprising that one has to split the data sample into proper time uncertainty bins, as the figure makes clear.\(^6\)

The present studies only represent a pioneering effort at this analysis. Further work should concentrate on calculating and understanding these efficiency functions more precisely. An extremely large amount of monte-carlo statistics will be necessary in order to investigate these systematic effects on the \(B_s\) lifetime to the point where they are smaller than the statistical uncertainty expected in Run 2a.

---

\(^6\)Each function has a different shape.
Appendix A

Glossary of Definitions

A.1 CDF Coordinate Definition

The CDF detector follows a cylindrical geometry. The Beam axis forms the z-axis and the transverse plane cutting a cross-section through the detector defines the xy plane.

However, the paths of charged particles in the detector do not follow a cylindrical geometry. Rather, since the central tracking volume sits in a magnetic field, they follow helical trajectories originating from a point very close to the z-axis. The beam-spot is very narrow in x and y. In Run I the rms beam radius at the collision point was $\sigma = 0.006 \text{ mm}$.[53] For this reason it makes sense to define an angle $\theta$ which is the angle between the z-axis and the direction of a charged particle at its distance of closest approach to the beam axis (z-axis). See figure A.1 for a graphical definition.

![Beam Axis Diagram](image)

Figure A.1: Beam axis and definition of $\theta$

154
A helical track is then defined by the following 5 parameters:

1. $\cot \theta$ : cotangent of the polar angle at closest approach to the beam axis (figure A.1).

2. $C$ : half curvature. This is a signed quantity with the same sign as the particle. It is related to $p_t$ by $C \propto \frac{1}{p_t}$

3. $z_0$ : $z$ position at the point of minimum approach to the beam axis.

4. $d_0$ : signed impact parameter: distance between helix and origin at the point of minimum approach to the beam axis (see figure A.3).

5. $\phi_0$ : Direction in the $xy$ plane of the track at the point of minimum approach to the beam axis (see figure A.2).

![Figure A.2: Definition of $\phi_0$ (high $p_t$ track)](image-url)
A.2 $L_{xy}$ and $d_0$

The variables $L_{xy}$ and $d_0$ occur very frequently in this work. In words they are defined as follows:

$L_{xy}$ : The 2D distance from the decay point of a particle to the primary vertex.

$d_0$ : The distance from the primary vertex to the line from the decay point of a particle, back along the direction of flight, where $d_0$ is orthogonal to this line.

Consider the diagram in figure A.3. Note that $d_0$ usually has a 3D interpretation, but since the level 2 trigger measures only transverse quantities, in the current work it will always be 2D.

A.3 $\eta$ definition

The pseudo-rapidity $\eta$ is defined:

$$\eta = -\ln(\tan \frac{\theta}{2}) \quad (A.1)$$

where $\theta$ is defined in section A.1.

Since $\theta$ is defined in the range $0 \rightarrow \pi$, $\eta$ is defined in the range:

$\theta \rightarrow 0 : \eta \rightarrow -\infty$ (since $\ln x \rightarrow -\infty$ as $x \rightarrow 0$).

$\theta \rightarrow \pi : \eta \rightarrow +\infty$ (since $\tan x \rightarrow +\infty$ as $x \rightarrow \frac{\pi}{2}$).

This quantity approximates to rapidity in the high $\gamma$ limit (as the name might suggest), where true rapidity is defined:

$$y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) \quad (A.2)$$
Figure A.3: Diagram defining $L_{xy}$ and $d_0$ (red). Note that all quantities are projected onto the xy plane. $d_0$ is the perpendicular distance of closest approach of a track to the primary vertex (transverse here). $L_{xy}$ is the 2D distance from the secondary decay point to the primary vertex.
Appendix B

Theoretical Appendix

B.1 Connecting the Weak and Strong Eigenvalues

In section 1.2.1 equation 1.20 describes the time evolution of the strong interaction B-meson eigenstates. A calculation of the magnitude of $\Delta \Gamma_{B_s}$ and $B_s$ mixing requires us to connect the Weak interaction mass and lifetime eigenvalues to their strong interaction counterparts algebraically. The following section outlines the algebra of this procedure.

B.1.1 Starting Point

The equation governing the time evolution of the strong interaction eigenstates is:

$$i \frac{d}{dt} \begin{pmatrix} |B\rangle \\ \end{pmatrix} = (M - \frac{i}{2} \Gamma) \begin{pmatrix} |B\rangle \\ \end{pmatrix}$$  \hspace{1cm} (B.1)

where $M$ and $\Gamma$ are $2 \times 2$ Hermitian matrices:

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad \Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix}$$  \hspace{1cm} (B.2)

The weak interaction eigenstates are linked to the strong interaction states via the equations:
\[ |B_L\rangle = p|B\rangle + q|\bar{B}\rangle \] (B.3)
\[ |B_H\rangle = p|B\rangle - q|\bar{B}\rangle \] (B.4)

Their time evolution is given by:

\[ |B_H(t)\rangle = \exp(-i M_H t - \Gamma_H t/2) |B_H\rangle \] (B.5)
\[ |B_L(t)\rangle = \exp(-i M_L t - \Gamma_L t/2) |B_L\rangle \] (B.6)

This is all the information we need to calculate the weak interaction mass and lifetime eigenvalues in terms of the strong interaction eigenvalues.

**B.1.2 Prescription**

The recipe to connect the eigenvalues (which will be worked out in detail in the next section) is then:

1. Invert equations B.3 and B.4 to get expressions for \( |B\rangle \) and \( |\bar{B}\rangle \) in terms of \( |B_L\rangle \) and \( |B_H\rangle \).

2. Substitute these into equation B.1:
   - using equations B.5 and B.6 to calculate the effect of the \( \frac{d}{dt} \)
   - and expanding \( M \) and \( \Gamma \) on the r.h.s using equation B.2 (simplified using the Hermitian Property).

3. Gather coefficients of \( |B_H\rangle \) and \( |B_L\rangle \).

4. Eliminate the on diagonal elements of \( M \) and \( \Gamma \) from the resulting equations.

**B.1.3 Working it through**

Inverting B.3 and B.4 yields:

\[ |B\rangle = \frac{1}{2p}(|B_L\rangle + |B_H\rangle) \] (B.7)
\[ |\bar{B}\rangle = \frac{1}{2q}(|B_L\rangle - |B_H\rangle) \] (B.8)
In anticipation of the \( \frac{d}{dt} \) in B.1 equations B.5 and B.5 yield:

\[
\frac{d}{dt}|B_H(t)| = -(iM_H + \Gamma_H/2) |B_H(t)| \tag{B.9}
\]

\[
\frac{d}{dt}|B_L(t)| = -(iM_L + \Gamma_L/2) |B_L(t)| \tag{B.10}
\]

Substituting B.7 and B.8 into B.1 and using B.9,B.9 yields:

\[
i \left( \frac{1}{2p}[ -(iM_L + \Gamma_L/2)|B_L\rangle + -(iM_H + \Gamma_H/2)|B_H\rangle \right)
- \frac{1}{2q} \left( -(iM_L + \Gamma_L/2)|B_L\rangle - -(iM_H + \Gamma_H/2)|B_H\rangle \right)
= (M - \frac{i}{2}\Gamma) \left( \frac{1}{2p}(|B_L\rangle + |B_H\rangle)
- \frac{1}{2q}(|B_L\rangle - |B_H\rangle) \right) \tag{B.11}
\]

Next the hermitian property of B.2 is used so that \( M = M^\dagger \) and \( \Gamma = \Gamma^\dagger \):

\[
\begin{pmatrix}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{pmatrix}
= \begin{pmatrix}
M^*_{11} & M^*_{12} \\
M^*_{21} & M^*_{22}
\end{pmatrix}
= \begin{pmatrix}
\Gamma_{11} & \Gamma_{12} \\
\Gamma_{21} & \Gamma_{22}
\end{pmatrix}
\tag{B.12}
\]

so that:

\[
M_{21} = M^*_{12} \quad \Gamma_{21} = \Gamma^*_{12} \tag{B.13}
\]

The right hand side of B.11 becomes:

\[
\begin{pmatrix}
M_{11} \frac{1}{2}\Gamma_{11} & M^*_{12} \frac{1}{2}\Gamma^*_{12} \\
M_{12} \frac{1}{2}\Gamma_{12} & M^*_{22} \frac{1}{2}\Gamma^*_{22}
\end{pmatrix}
\left( \frac{1}{2p}(|B_L\rangle + |B_H\rangle) - \frac{1}{2q}(|B_L\rangle - |B_H\rangle) \right) \tag{B.14}
\]

Equation B.14 expands out so that the fully expanded version of B.11 is:

\[
i \left( \frac{1}{2p}[ -(iM_L + \Gamma_L/2)|B_L\rangle + -(iM_H + \Gamma_H/2)|B_H\rangle \right)
- \frac{1}{2q} \left( -(iM_L + \Gamma_L/2)|B_L\rangle - -(iM_H + \Gamma_H/2)|B_H\rangle \right)
= \left( (M_{11} \frac{1}{2}\Gamma_{11}) \frac{1}{2p}(|B_L\rangle + |B_H\rangle) + (M^*_{12} \frac{1}{2}\Gamma^*_{12}) \frac{1}{2q}(|B_L\rangle - |B_H\rangle) \right)
+ \left( (M_{12} \frac{1}{2}\Gamma_{12}) \frac{1}{2p}(|B_L\rangle + |B_H\rangle) + (M^*_{22} \frac{1}{2}\Gamma^*_{22}) \frac{1}{2q}(|B_L\rangle - |B_H\rangle) \right) \tag{B.15}
\]

Gathering the coefficients of \( |B_H\rangle \) and \( |B_L\rangle \) yields the following 4 equations:
\[
\begin{align*}
-\frac{i}{2p}(iM_L + \frac{\Gamma_L}{2}) &= (M_{11} - \frac{i}{2}\Gamma_{11})\frac{1}{2p} + (M_{12}^{*} - \frac{i}{2}\Gamma_{12})\frac{1}{2q} & (B.16) \\
-\frac{i}{2p}(iM_H + \frac{\Gamma_H}{2}) &= (M_{11} - \frac{i}{2}\Gamma_{11})\frac{1}{2p} - (M_{12}^{*} - \frac{i}{2}\Gamma_{12})\frac{1}{2q} & (B.17) \\
-\frac{i}{2q}(iM_L + \frac{\Gamma_L}{2}) &= (M_{12} - \frac{i}{2}\Gamma_{12})\frac{1}{2p} + (M_{22} - \frac{i}{2}\Gamma_{22})\frac{1}{2q} & (B.18) \\
\frac{i}{2q}(iM_H + \frac{\Gamma_H}{2}) &= (M_{12} - \frac{i}{2}\Gamma_{12})\frac{1}{2p} - (M_{22} - \frac{i}{2}\Gamma_{22})\frac{1}{2q} & (B.19)
\end{align*}
\]

Equations B.16 and B.17 can be combined because the first term on the right hand side of both is the same. Equations B.18 and B.19 can be combined similarly (the 2nd term on the right hand side is the same). This substitution yields 2 equations:

\[
\begin{align*}
-\frac{i}{2p}(iM_L + \frac{\Gamma_L}{2}) - (M_{12}^{*} - \frac{i}{2}\Gamma_{12})\frac{1}{2q} &= -\frac{i}{2p}(iM_H + \frac{\Gamma_H}{2}) + (M_{12}^{*} - \frac{i}{2}\Gamma_{12})\frac{1}{2q} & (B.20) \\
-\frac{i}{2q}(iM_L + \frac{\Gamma_L}{2}) - (M_{12} - \frac{i}{2}\Gamma_{12})\frac{1}{2p} &= -\frac{i}{2q}(iM_H + \frac{\Gamma_H}{2}) + (M_{12} - \frac{i}{2}\Gamma_{12})\frac{1}{2p} & (B.21)
\end{align*}
\]

Multiplying B.20 by q and rearranging yields:

\[
-\frac{i}{2p}q[(iM_L + \frac{\Gamma_L}{2}) - (iM_H + \frac{\Gamma_H}{2})] = (M_{12}^{*} - \frac{i}{2}\Gamma_{12})
\]  

(B.22)

Which can be rearranged to yield:

\[
\frac{1}{2p}q[(M_L - M_H) - \frac{i}{2}(\Gamma_L - \Gamma_H)] = (M_{12}^{*} - \frac{i}{2}\Gamma_{12})
\]  

(B.23)

And since \(\Delta m_B = M_H - M_L\), \(\Delta \Gamma_B = \Gamma_H - \Gamma_L\) B.23 can be rearranged to yield:

\[
\frac{q}{p} = \frac{-2(M_{12}^{*} - \frac{i}{2}\Gamma_{12})}{\Delta m_B - \frac{i}{2}\Delta \Gamma_B}
\]  

(B.24)

Equation B.21 can be rearranged similarly to yield:

\[
\frac{q}{p} = \frac{\Delta m_B - \frac{i}{2}\Delta \Gamma_B}{2(M_{12} - \frac{i}{2}\Gamma_{12})}
\]  

(B.25)
These two equations are thus the same as 1.27. Equating B.24 and B.25 yields:

\[ \frac{2(M_{12}^* \pm i \Gamma_{12}^*)}{\Delta m_B - \frac{i}{2} \Delta \Gamma_B} = \frac{\Delta m_B - \frac{i}{2} \Delta \Gamma_B}{2(M_{12}^* \pm i \Gamma_{12}^*)} \]  
(B.26)

Rearranging gives:

\[ (\Delta m_B)^2 - \frac{1}{4}(\Delta \Gamma_B)^2 - i \Delta \Gamma_B \Delta m_B = 4|M_{12}|^2 - |\Gamma_{12}|^2 - 2i\Gamma_{12}^* M_{12} - 2i\Gamma_{12} M_{12}^* \]  
(B.27)

Equating real and imaginary parts yields:

\[ (\Delta m_B)^2 - \frac{1}{4}(\Delta \Gamma_B)^2 = 4|M_{12}|^2 - |\Gamma_{12}|^2 \]  
(B.28)

\[ \Delta \Gamma_B \Delta m_B = 2(\Gamma_{12}^* M_{12} + \Gamma_{12} M_{12}^*) \]  
(B.29)

which rearrange to:

\[ (\Delta m_B)^2 - \frac{1}{4}(\Delta \Gamma_B)^2 = 4(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2) \]  
(B.30)

\[ \Delta m_B \Delta \Gamma_B = 4 \Re(\Gamma_{12}^* M_{12}) \]  
(B.31)

which are the same as equations 1.25 and 1.26.

Finally in this section, the effect of the approximations, \( \Delta m_B \gg \Delta \Gamma_B \) and \( |M_{12}| \gg |\Gamma_{12}| \) is considered. The first approximation is true model independently. The second follows from the fact that the size of the diagrams that contribute to the absorptive part of the amplitude scale as the mass of the exchanged particle (see diagram 1.1). In this case the mass of the t-quark enters the absorptive amplitude, where as only the masses of the u and c quarks enter the amplitude for the dispersive amplitude.

The effect of the first of these approximations on equation B.26 are to change it to:

\[ \frac{q}{p} = \frac{2(M_{12}^* \pm i \Gamma_{12}^*)}{\Delta m_B} = \frac{\Delta m_B}{2(M_{12}^* \pm i \Gamma_{12}^*)} \]  
(B.32)

Thus:
\[ \Delta m_B^2 = 4(|M_{12}|^2 - \frac{1}{4} |\Gamma_{12}|^2 - \frac{i}{2} M_{12}\Gamma_{12}^* - \frac{i}{2} \Gamma_{12} M_{12}^*) \] (B.33)

Equating real parts and applying the second approximation (\(|M_{12}| \gg |\Gamma_{12}|\)) yields:

\[ \Delta m_B^2 = 4|M_{12}|^2 \] (B.34)

So that:

\[ \Delta m_B = 2|M_{12}| \] (B.35)

which is the same as the first equation of 1.28.

Substituting B.35 into B.31 and B.32 yields:

\[ \Delta \Gamma_B = 2\Ree(M_{12}\Gamma_{12}^*)/|M_{12}|, \quad \frac{q}{p} = -\frac{|M_{12}|}{M_{12}} \] (B.36)

which are the same as equation 1.29 and the second equation in 1.28.
Appendix C

\( p_t \) pull discussion

This section refers to the pull distribution in figure 5.1. The \( p_t \) is not actually the quantity which is gaussian distributed. Rather it is the curvature. The smearing is then most easily understood as:

\[
C_s = C + R\sigma_C
\]

(C.1)

where \( C_s \) is the smeared curvature, \( C \) is the unsmeared curvature, \( R \) is a unit gaussian deviate\(^1\), and \( \sigma_C \) is the curvature smearing width. When this equation is inverted, and expressed in terms of \( p_t \) the following equation is obtained:

\[
R = \left( \frac{p_t - p_{ts}}{\sigma_C} \right) \frac{1}{p_t p_{ts}}
\]

(C.2)

where \( p_{ts} \) is the smeared \( p_t \) and \( p_t \) is the unsmeared \( p_t \).

The slight systematic shift toward positive values in the pull distribution of figure 5.1 arises because in the code the approximation \( p_{ts} \sim p_t \) was made in the second fraction so that:

\[
R \sim \left( \frac{p_t - p_{ts}}{\sigma_C} \right) \frac{1}{p_t^2}
\]

(C.3)

A small amount of manipulation reveals that the approximate pull is related to the exact pull by:

\[
R(appx) = R + R^2 \left( \frac{\sigma_C}{C_S} \right)
\]

(C.4)

The discrepancy is therefore positive definite, but very small, since \( \sigma_C \ll C_S \)

---

\(^1\)Random number gaussian distributed about zero
Bibliography


165


[28] CDF Collaboration The CDF II Detector Technical Design Report Level Two Trigger Decision Hardware, Chapter R.Blair et al. (http://www-cdf.fnal.gov/upgrades/tdr/tdr_12_trig.html) Section 12.4.4


[33] S. Pappas, Proc. of the APS Meeting of the Division of Particles and Fields, UCLA, Los Angeles, January 5-9, 1999
[34] I. Furic, Ch. Paus Event Yield Estimates for Hadronic $B_s$ decays in Run II CDF Note 5349 (http://wwwcdf.fnal.gov/cdfnotes/cdf5349_BsHadrYields.ps)


[39] Paus et.al Generator Level Dimuon Simulation CDF/DOC/ BOTTOM/CDFR/5686

[40] PhD Dissertation F.Azfar 1995 Measurement of The Lifetime of the $B_s$ Meson Using The Exclusive Decay Mode $B_s \rightarrow J/\psi$

[41] GEANT Detector simulation tool (version 3) M.Goossens, J.Apostolakis (CERN).


[45] Heiss et.al A new Measurement of Exclusive B Lifetimes using Run 1 Data (Section 4.3) CDF Note 5331 (CDF/ANAL/ BOTTOM/CDFR/5331)

[46] James et.al Minuit, Function Minimization and Error Analysis CERN Program Library Long Writeup D506

[48] *B Physics at the Tevatron Run II and Beyond* FERMILAB-Pub-01/197 Page 358. equation 8.86

[49] Study of $B_s^0 - \bar{B}_s^0$ oscillations and $B_s^0$ lifetimes using hadronic decays of $B_s^0$ mesons Delphi Collaboration CERN-EP/2000-087

[50] *B Physics at the Tevatron Run II and Beyond* FERMILAB-Pub-01/197 Ch 8 (page 363)

[51] *B Physics at the Tevatron Run II and Beyond* FERMILAB-Pub-01/197 Ch 8 (page 415, section 8.7.1.1)

[52] CDF Collaboration Run II offline repository. Access restricted to collaboration members usually. Contact: kreymer@fnal.gov.

[53] Beams Division FNAL Design Report Tevatron 1 Project Section 10-2.


