Please note that the analysis presented in this thesis is pre-RC; for post-RC see BAD #1413.

A study of $B^+ \to \overline{K}^*(892)^0 K^+$ using the *BABAR* detector

James Patrick Burke

Department of Physics
H H Wills Physics Laboratory

April 5, 2007

A thesis submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy in the Faculty of Science

$\approx 40,000$ words

Abstract

A search for the decay of a charged B meson to a charged K meson and a neutral $K^*(892)$ meson is presented. The $K^*(892)$ resonance is detected through its decay to a charged pion and a charged kaon. The final state is therefore $K^{\pm}K^{\mp}\pi^{\pm}$. The analysis uses a data sample with an integrated luminosity of $210.6\,\mathrm{fb}^{-1}$, recorded by the *BABAR* detector at the PEP-II asymmetric B factory. This sample corresponds to 232.3 million charged B mesons. The branching ratio averaged over charge conjugate states, $\mathcal{B}$, is found to be

- $\mathcal{B}(B^+ \to \overline{K}^*(892)^0 K^+) = (0.6 \pm 0.3 \pm 0.2) \times 10^{-6}$.

The first uncertainty is statistical, the second is systematic. The significance of this measurement, 1.6σ, falls below the required significance for experimental evidence. As such an upper limit, at 90% confidence level (CL), is placed on the branching ratio:

- $\mathcal{B}(B^+ \to \overline{K}^*(892)^0 K^+) < 1.1 \times 10^{-6}$.

The mode $B^+ \to \overline{K}_0^*(1430)^0 K^+$ shares the same final state as the channel under investigation. The following 90% CL upper limit is placed on the branching ratio for this mode:

- $\mathcal{B}(B^+ \to \overline{K}_0^*(1430)^0 K^+) < 2.2 \times 10^{-6}$.

Acknowledgements

A gargantuan thanks to Tom and Jamie, who are responsible for the majority of knowledge that I've acquired in the field of experimental particle physics during the last four years. Their know-how and helpfulness have been thoroughly invaluable.

A huge thanks also to Fergus, for continuing to supervise me despite leaving Bristol mid-way through my PhD, to Nick, for the thesis-writing feedback and for helping whenever I asked, and to Nicole, for her assistance throughout my first year.

I'm especially grateful to the above people as they all moved on from Bristol at one time or another, but continued to support me, unobligated, and as enthusiastically as if they were still Bristolians.

I had a terrific time in the incredible state of California, and for this I'd like to thank `ajh`, `bhimji`, `boyd`, `carlosch`, `djbard`, `ejhill`, `forster`, `keford`, `mattb`, `mn98`, `mnaisbit`, `nbarlow`, `olaiya`, `paulj`, `smolloy`, `tlatham`, `wp99`, SSF *soccer*, Maggie, and Brits in years below and above – and indeed crazy Frenchmen.

In Bristol too I've made many good friends ... cheers to Adam, Pad, Steve (Bubble Chamber national champions '03!), Fraz, Metson, Anatoly, TPRFC and all those that I've played 8-a-side with. And of course a big cheers to my housewife Mooooooooooay.

I'd like to thank the University of Bristol particle physics group for giving me the opportunity and tools to carry out this PhD, PPARC for funding me, and SLAC for being excellent hosts for 15 months. I'm very grateful for the exceptionally hard work carried out by the PEP-II team and the *BABAR* collaboration, which has made

my analysis possible. In particular, I'd like to acknowledge the charmless 3-body AWG and the trigger group for their help and expertise.

Biggest thanks of all are left for me mam and dad, Colleen, and mama and grandad for their infinite support.

Declaration

I declare that the work in this thesis was carried out in accordance with the Regulations of the University of Bristol. No part of the thesis has been submitted for any other academic award at this or any other university. Any views expressed in the dissertation are those of the author.

The detector hardware, electronics and data acquisition system, as discussed in Chapters 2 and 3, and the offline reconstruction software described in Chapter 4, were developed by the *BABAR* collaboration and are the tools used for recording the data used in the analysis presented in this thesis. The author contributed to the running of the detector through the taking of general shifts and being the Commissioner of the Electromagnetic Calorimeter Trigger for 12 months. During this period the author developed online monitoring code. This is briefly described in Section 3.3.4. The author also spent seven months assisting with the production of Monte Carlo data.

The maximum likelihood analyses described in Chapters 5 and 6 were undertaken solely by the author. The event selection code used was developed by numerous *BABAR* collaborators, including, in part, the author. The fitting package used, `CharmlessFitter`, was developed by Thomas Latham. The author has added functionality to this package and implemented numerous improvements. The package builds upon the more general fitting code, `RooFit`, also developed by *BABAR* collaborators, and also to which the author has made a contribution.

SIGNED: ..... DATE: .....

Contents

List of Figures

List of Tables

Introduction

The Big Bang theory has been widely accepted as a description of the birth of the Universe. In this theory matter particles and antimatter antiparticles are created in equal amounts. When a particle meets its antiparticle the two annihilate to produce radiation. In today's Universe all of the antimatter has annihilated with matter[1]. However, rather than a Universe consisting purely of radiation, there is a surplus of matter. This can only be explained by a matter-antimatter asymmetry. The asymmetry is very small and subtle, but nevertheless responsible for the Universe as we know it and our very existence.

Turning our attention now from physics on the largest possible scale, the cosmos, to physics on the smallest scale. The Standard Model of particle physics (SM) is a relativistic quantum field theory that describes the fundamental particles that make up all matter (and antimatter) and the interactions that occur between these particles. The SM is a mathematical *gauge* theory, the conceptual results of which can be described as follows. All observed matter is hypothesised to consist of various combinations of 12 elementary, spin $\frac{1}{2}$ fermions (6 quarks and 6 leptons). Fundamental interactions are explained by the exchange of spin 1 bosons, whilst a further boson, with spin 0, accounts for particles' mass. The theory is incomplete as it

[1]Antimatter continues to be produced in small quantities in certain nuclear reactions and in physics experiments on Earth, but these quantities are negligible in comparison to the amount of matter present in the Universe.

does not incorporate the very weak (relatively speaking) gravitational interaction. Furthermore, the SM is unable to address numerous fundamental questions. For example, why are there three generations of fermions? Why do fermions have the masses they do and why are they so varied? Why does the Higgs bare mass in the SM Lagrangian have to be so unnaturally fine-tuned to one part in 10^{17}? These issues point toward more fundamental physics at higher energies (for further reading see [2]). Nonetheless, rigorous and precise experimental measurements demonstrate that the SM is an extremely successful model at currently probed energies.

CP violation – in which the rates of various processes differ if particles are exchanged with antiparticles and spatial directions are reversed – is one of three conditions identified by Sakharov [3] that must be satisfied in order for the Universe to evolve from a system in which the amount of matter and antimatter is equal to one in which it is not. (The other two conditions are baryon number violation and a withdrawal from thermal equilibrium for some period.) *CP* violation is accommodated in the SM, but the amount predicted by the SM mechanism is too small by several orders of magnitude to account for the observed imbalance between matter and antimatter in the Universe. The sector is therefore an interesting area to look for new physics beyond the SM.

To establish experimentally the effect that allowed matter to dominate is a central theme in high energy physics research, and one which the *BABAR* experiment [4] at the Stanford Linear Accelerator Center (SLAC) [5] in California addresses. The *BABAR* collaboration consists of approximately 600 physicists and engineers from around the world.

The analysis presented in this thesis sets out to determine the branching ratio of the decay $B^+ \to \overline{K}^*(892)^0 K^+$. The author conducted a *cut and count* analysis in 2004 in which an upper limit, at 90% CL, was placed on this branching ratio [6]:

$$\mathcal{B}_{Exp}(B^+ \to \overline{K}^*(892)^0 K^+) < 2.2 \times 10^{-6}$$

(reduced from 5.3×10^{-6} [7]). In 2005 Fleischer and Recksiegel placed a theoretical lower limit on this mode [8]:

$$\mathcal{B}_{Thry}(B^+ \to \overline{K}^*(892)^0 K^+) > \left(0.76^{+0.10}_{-0.12}\right) \times 10^{-6}.$$

This window and the use of a larger data sample and a more sensitive technique (maximum likelihood rather than cut and count) indicate that a measurement is within present experimental sensitivity at *BABAR*.

$B^+ \to \overline{K}^*(892)^0 K^+$ occurs at the quark level as a $b \to d$ transition. These transitions are suppressed, and there is currently little experimental information available on them. Studying these processes provides important tests of the quark flavour sector of the SM. This is discussed in [9]. The study presented in this thesis can also be used to help clarify issues concerning a possible difference between determinations of a particular *CP* violation parameter, namely $\sin 2\beta$ using the modes $B^0 \to J/\psi\, K^0_S$ and $B^0 \to \phi(1020) K^0_S$ [10]. Such a difference would require physics beyond the SM to explain it.

Chapter 1 describes the portions of the SM relevant to the *BABAR* experiment and to the decay mode under investigation. The physics requirements, design and performance of the *BABAR* detector and the PEP-II B factory are discussed in Chapter 2, whilst the online flow of data from the detector is described in Chapter 3. In Chapter 4, details are given for the reconstruction of the raw detector data into a form that can be used for physics analyses. Chapter 5 presents the analysis method, including event selection, treatment of background, the fit procedure and validation of this procedure. The results are presented in Chapter 6 and summarised in Chapter 7.

1

Theoretical overview

1.1 Introduction

In this chapter relevant aspects of the underlying theory are presented. The main goal of *BABAR* is to conduct high precision studies of *CP* violation in the B meson system. As such, the CKM sector of the SM – in which *CP* violation occurs – is examined in some detail. The role of the mode under investigation in this thesis is then discussed in relation to the *CP* violation parameter $\sin 2\beta$. The discussion then turns to the kinematics of B decays to 3-body final states. Finally, the effects of the strong interaction on the decay under study are very briefly discussed.

1.2 The CKM description of flavour changing processes

This section describes how *CP* violation arises in the SM. Mixing in neutral mesons is described as a precursor to discussing the three types of *CP* violation. Current knowledge of the CKM parameters, accumulated experimentally, is then presented. Lastly, an explanation of how the parameter sin2β is extracted using *BABAR*'s *golden channel* is given.

1.2.1 An introduction to *CP* violation

C is the charge conjugation operator; it is discrete and transforms a particle into its antiparticle by reversing the sign of its internal quantum numbers. **P** is the parity operator – also discrete – which inverts three-dimensional space: $\vec{x} \rightarrow -\vec{x}$. In contrast to the strong and electromagnetic interactions, the symmetries of both the **C** and **P** transformations are violated by the weak force. That is, only particles with left-handed chirality and antiparticles with right-handed chirality participate in weak interactions in the SM[1]. Parity violation was first proposed by Lee and Yang in 1956 [11], and experimentally verified (using β decay of cobalt-60) in 1957 by Wu *et al.* [12]. In the same year Ioffe *et al.* showed that **P** violation meant that **C** invariance must also be violated in weak decays [13]. At this stage, however, there was no evidence to suggest that the combined transformation of **C** and **P**, **CP**, was

[1]Chirality is a Lorentz invariant quantity, which for massless particles, is the same as helicity (the projection of a particle's spin onto the direction of its motion). The introduction of mass terms renders helicity non-absolute and makes possible weak interactions that otherwise would not be. For example – assuming massless neutrinos for the sake of this argument – if electrons were massless, the decay $\pi^- \rightarrow e^-\overline{\nu}_e$ would be forbidden. The electron is not massless and the process therefore occurs. However, the electron is much closer to being massless than the muon and as such $\pi^- \rightarrow e^-\overline{\nu}_e$ is heavily suppressed compared to $\pi^- \rightarrow \mu^-\overline{\nu}_\mu$ – despite the available phase spaces suggesting the opposite.

not conserved. In 1964 Christenson *et al.* presented experimental evidence for the violation of symmetry under the **CP** transformation – *CP* violation – in the K^0 meson system [14]. The K_L^0 particle was observed to decay to two pions, rather than three, at the 10^{-3} level.

A further exclusive property of the weak interaction is that it is able to change flavour. The mechanism that allows this in quarks, introduced by Cabibbo in 1963 [15], is *quark mixing*. Cabibbo proposed that the weak eigenstate d' is a linear superposition of the flavour eigenstates d and s. This can be extended as follows:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{1.1}$$

The matrix in Eq. (1.1) can be parameterised using the single Cabibbo angle, θ_C, measured experimentally to be 12.3° [16]. In 1973 Kobayashi and Maskawa proposed a third generation of quarks and leptons [17] with the 2×2 Cabibbo matrix becoming the 3×3 *CKM matrix*, $\mathbf{V}_{\mathrm{CKM}}$. Three real angles and one complex phase are now required to parameterise the matrix. It is the complex phase that is the only source of *CP* violation in the SM[2]. The first experimental evidence for the existence of third generation fermions followed shortly after (the b quark [19] and the τ lepton [20] were discovered in 1977).

1.2.2 The CKM matrix and the Unitarity Triangle

With a third generation of quarks Eq. (1.1) becomes

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \mathbf{V}_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.2}$$

[2]In fact, there are natural terms in the QCD Lagrangian that are also able to break CP symmetry. However, it turns out that the relevant phase is fine-tuned such that this contribution can effectively be ignored. This is known as the *Strong CP Problem* [18].

Each $\mathbf{V}_{\rm CKM}$ matrix element, V_{ij}, represents the weak coupling between quarks i and j. The standard parameterisation of the CKM matrix is shown in Eq. (1.3):

$$\mathbf{V}_{\rm CKM} = \begin{pmatrix} c_{12}\,c_{13} & s_{12}\,c_{13} & s_{13}\,e^{-i\delta} \\ -s_{12}\,c_{23} - c_{12}\,s_{23}\,s_{13}\,e^{i\delta} & c_{12}\,c_{23} - s_{12}\,s_{23}\,s_{13}\,e^{i\delta} & s_{23}\,c_{13} \\ s_{12}\,s_{23} - c_{12}\,c_{23}\,s_{13}\,e^{i\delta} & -c_{12}\,s_{23} - s_{12}\,c_{23}\,s_{13}\,e^{i\delta} & c_{23}\,c_{13} \end{pmatrix}, \quad (1.3)$$

where $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$ and δ represents the phase. θ_{ij} is the mixing angle between the ith and jth generations. Experimental results show a hierarchy in the magnitudes of the matrix elements: $|V_{ii}| \approx 1$, $|V_{12}| \approx |V_{21}| \approx \lambda$, $|V_{23}| \approx |V_{32}| \approx \lambda^2$ and $|V_{13}| \approx |V_{31}| \approx \lambda^3$, where $\lambda = s_{12} \approx 0.22$ is the sine of the Cabibbo angle. This leads to an alternative (approximate) parameterisation of the CKM matrix, developed by Wolfenstein [21]:

$$\mathbf{V}_{\rm CKM} \approx \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (1.4)$$

To conserve the number of quarks and be self-contained the CKM matrix must be unitary. Two relevant types of relation arise from $\mathbf{V}^\dagger_{\rm CKM}\mathbf{V}_{\rm CKM} = \mathbf{I} = \mathbf{V}_{\rm CKM}\mathbf{V}^\dagger_{\rm CKM}$:

$$\sum_j |V_{ij}|^2 = 1, \quad \text{and} \quad (1.5)$$

$$\sum_i V_{id}V^*_{is} = 0, \qquad \sum_i V_{is}V^*_{ib} = 0, \qquad \sum_i V_{id}V^*_{ib} = 0,$$

$$\sum_j V_{uj}V^*_{cj} = 0, \qquad \sum_j V_{cj}V^*_{tj} = 0, \qquad \sum_j V_{uj}V^*_{tj} = 0, \quad (1.6)$$

where $i = u, c, t$ and $j = d, s, b$. Eq. (1.5) describes *weak universality*, which implies that the sum of all the couplings of any up-type quark to the down-type quarks is generation independent. Of the relations of the type described by Eq. (1.6), the one of particular interest to us is

$$V_{ud}\,V^*_{ub} + V_{cd}\,V^*_{cb} + V_{td}\,V^*_{tb} = 0. \quad (1.7)$$

This can be depicted as a triangle – the *Unitarity Triangle* – in the complex plane, as illustrated in Figure 1.1. Here, a phase convention is chosen such that $V_{cd}\,V^*_{cb}$ is

real and the triangle has been scaled such that two of the three corners lie at $(0,0)$ and $(1,0)$, thus leaving only the apex to find, $(\overline{\rho},\overline{\eta})$.

$$\overline{\rho} = \rho\left(1-\lambda^2/2\right), \quad \overline{\eta} = \eta\left(1-\lambda^2/2\right), \tag{1.8}$$

where λ, ρ and η are quantities used in the Wolfenstein parameterisation of the CKM matrix (Eq. (1.4)). The internal angles of the triangle are given by

$$\alpha \equiv \arg\left[-\frac{V_{td}\,V_{tb}^*}{V_{ud}\,V_{ub}^*}\right], \quad \beta \equiv \arg\left[-\frac{V_{cd}\,V_{cb}^*}{V_{td}\,V_{tb}^*}\right], \quad \gamma \equiv \arg\left[-\frac{V_{ud}\,V_{ub}^*}{V_{cd}\,V_{cb}^*}\right]. \tag{1.9}$$

The five remaining relations of the type described by Eq. (1.6) can also be represented as triangles, but these triangles are much thinner (one side much shorter than the other two) and are therefore less useful for experimental probing.

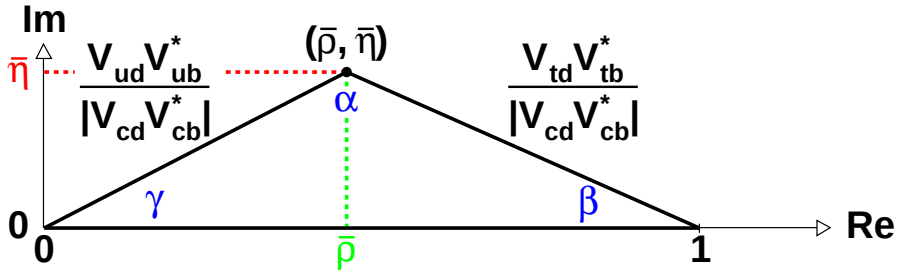

Figure 1.1: *The Unitarity Triangle (not drawn to scale).*

It is the aim of the *BABAR* collaboration to measure the sides and angles of the Unitarity Triangle, using as many independent decay modes as possible. The area of the Unitarity triangle is equal to $J/2$ where J is the Jarlskog invariant [22] and is a model independent measure of the amount of *CP* violation in the SM:

$$J = c_{12}c_{23}c_{13}^2 s_{12}s_{23}s_{13}\sin\delta \approx A^2\eta\lambda^6. \tag{1.10}$$

1.2.3 The charged current weak interaction Lagrangian

The charged current weak interaction is mediated by massive, charged gauge bosons $W^{\pm}$ and is described, in the weak basis, by the following Lagrangian:

$$\mathcal{L}_{\text{Int}}^{\text{CC}} = -\frac{g}{\sqrt{2}} \left(\overline{u}'_L, \overline{c}'_L, \overline{t}'_L\right) \gamma^\mu \begin{pmatrix} d'_L \\ s'_L \\ b'_L \end{pmatrix} W_\mu^\dagger + h.c., \tag{1.11}$$

where g is the weak coupling constant, γ^μ are the Dirac matrices, $W_\mu^\dagger$ are the weak gauge bosons and *h.c.* indicates the hermitian conjugate of the first term. Left-handed projections of the *weak* eigenstates of the quark fields are represented by the q'_L.

The quarks have mass in the SM due to their Yukawa couplings to the Higgs doublet. It is therefore useful to rewrite Eq. (1.11) in terms of left-handed projections of the *mass* eigenstates of the quark fields (q_L) by substituting in Eq. (1.2):

$$\mathcal{L}_{\text{Int}}^{\text{CC}} = -\frac{g}{\sqrt{2}} \left(\overline{u}_L, \overline{c}_L, \overline{t}_L\right) \gamma^\mu \mathbf{V}_{\text{CKM}} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} W_\mu^\dagger + h.c. \tag{1.12}$$

Let us now rewrite Eq. (1.12) in terms of the full quark fields, (u_i, d_j), on which the left-handed projection operator $\frac{1}{2}(1-\gamma_5)$ acts:

$$\mathcal{L}_{\text{Int}}^{\text{CC}} = -\frac{g}{2\sqrt{2}} (\overline{u}_i \gamma^\mu W_\mu^+ \left(1-\gamma_5\right) V_{ij} d_j + \overline{d}_j \gamma^\mu W_\mu^- \left(1-\gamma_5\right) V_{ij}^* u_i). \tag{1.13}$$

Applying the *CP* operator to the field terms in Eq. (1.13) transforms them as:

$$\overline{u}_i \gamma^\mu W_\mu^+ \left(1-\gamma_5\right) d_j \rightarrow \overline{d}_j \gamma^\mu W_\mu^- \left(1-\gamma_5\right) u_i. \tag{1.14}$$

It can be seen that the field terms are interchanged. The $\mathbf{V}_{\text{CKM}}$ couplings however (V_{ij} and V_{ij}^*) are unchanged, and since $V_{ij} \neq V_{ij}^*$ is possible with a complex element in $\mathbf{V}_{\text{CKM}}$, *CP* violation can occur in the SM.

1.2.4 Mixing in neutral mesons

The box diagrams shown in Figure 1.2 are an example of *mixing* in the SM, in which particle and antiparticle states can oscillate. For the majority of particle-antiparticle

systems this is not possible because various quantum numbers must be conserved regardless of whether the interaction is strong, electromagnetic or weak. However, for a small number of systems, including the neutral K, D and B systems, there are no quantum numbers to conserve when taking into account the weak interaction. When this is the case, the observed physical particles correspond not to the flavour eigenstates themselves ($|P^0\rangle$ and $\left|\overline{P}^0\right\rangle$, with definite quark content), but to linear combinations of them.

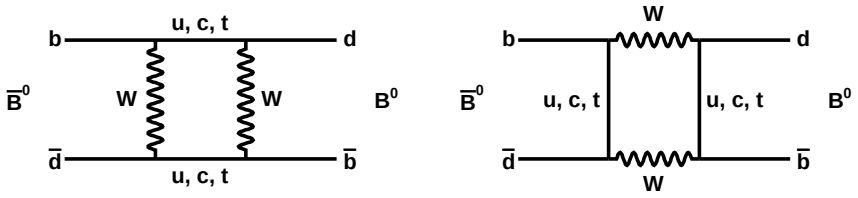

Figure 1.2: *Feynman diagrams representing the second order weak interactions that give rise to $B^0 - \overline{B}^0$ mixing. The top quark dominates these transitions due to its large mass and because $V_{tb} \approx 1$.*

The time evolution of an arbitrary state consisting of a linear superposition of the flavour eigenstates,

$$|P(t)\rangle = a(t)\left|P^0\right\rangle + b(t)\left|\overline{P}^0\right\rangle, \tag{1.15}$$

is governed by the time dependent Schrödinger equation (TDSE)

$$i\frac{\partial}{\partial t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \mathbf{H}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}\right)\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}. \tag{1.16}$$

$\mathbf{H}$ is the Hamiltonian matrix, whilst $\mathbf{M}$ and $\mathbf{\Gamma}$ are 2×2 Hermitian matrices that describe mixing and decay respectively.

Invariance under the **CPT** transformation (a prerequisite of quantum field theory and assumed to be true) – where $\mathbf{T}$ is the time reversal operator – requires that H_{11}

and H_{22} are equal. The off-diagonal elements, H_{12} and H_{21} are the amplitudes for mixing. If these elements are zero, there is no mixing.

1.2.4.1 Mixing without *CP* violation

For *CP* symmetry to hold, we require that $H_{12}^* = H_{21}$. In the absence of *CP* violation Eq. (1.16) thus becomes

$$i\frac{\partial}{\partial t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} A & B \\ B^* & A \end{pmatrix}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ \left(M_{12} - \frac{i}{2}\Gamma_{12}\right)^* & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}. \tag{1.17}$$

By diagonalising $\mathbf{H}$ we can move into the mass basis. $\mathbf{X}$ is defined to be the matrix whose columns contain the eigenvectors of $\mathbf{H}$ such that

$$\mathbf{H} = \mathbf{X}\begin{pmatrix} A + |B| & 0 \\ 0 & A - |B| \end{pmatrix}\mathbf{X}^{-1}, \tag{1.18}$$

where $A + |B|$ and $A - |B|$ are eigenvalues of $\mathbf{H}$. The resulting mass eigenstates are given by

$$|P_{1,2}\rangle = \frac{1}{\sqrt{2}}\left(\left|P^0\right\rangle \pm \left|\overline{P}^0\right\rangle\right), \tag{1.19}$$

The masses and widths of the mass eigenstates are given by the real and imaginary parts of the eigenvalues respectively:

$$M_{1,2} = \Re(A \pm |B|), \quad -\frac{\Gamma_{1,2}}{2} = \Im(A \pm |B|). \tag{1.20}$$

With exact *CP* invariance, the mass eigenstates are also *CP* eigenstates with eigenvalues ± 1:

$$\begin{aligned}
\mathbf{CP}\,|P_1\rangle &= \mathbf{CP}\left(\frac{1}{\sqrt{2}}\left(\left|P^0\right\rangle + \left|\overline{P}^0\right\rangle\right)\right) && (1.21)\\
&= \frac{1}{\sqrt{2}}\left(\mathbf{CP}\left|P^0\right\rangle + \mathbf{CP}\left|\overline{P}^0\right\rangle\right) && (1.22)\\
&= \frac{1}{\sqrt{2}}\left(e^{i\delta}\left|\overline{P}^0\right\rangle + e^{-i\delta}\left|P^0\right\rangle\right) && (1.23)\\
&= \frac{1}{\sqrt{2}}\left(\left|P^0\right\rangle + \left|\overline{P}^0\right\rangle\right), \text{ choosing the convention } \delta = 0 && (1.24)\\
&= (+1)\,|P_1\rangle, && (1.25)
\end{aligned}$$

$$\begin{aligned}
\mathbf{CP}\,|P_2\rangle &= \mathbf{CP}\left(\frac{1}{\sqrt{2}}\left(\left|P^0\right\rangle - \left|\overline{P}^0\right\rangle\right)\right) && (1.26)\\
&= \frac{1}{\sqrt{2}}\left(\mathbf{CP}\left|P^0\right\rangle - \mathbf{CP}\left|\overline{P}^0\right\rangle\right) && (1.27)\\
&= \frac{1}{\sqrt{2}}\left(e^{i\delta}\left|\overline{P}^0\right\rangle - e^{-i\delta}\left|P^0\right\rangle\right) && (1.28)\\
&= -\frac{1}{\sqrt{2}}\left(\left|P^0\right\rangle - \left|\overline{P}^0\right\rangle\right), \text{ choosing the convention } \delta = 0 && (1.29)\\
&= (-1)\,|P_2\rangle\,. && (1.30)
\end{aligned}$$

1.2.4.2 Mixing with *CP* violation

With *CP* invariance no longer exact, the off-diagonal elements of **H** are no longer required to be of equal magnitude, and **H** can be rewritten as

$$\mathbf{H} = \begin{pmatrix} A & B/r \\ rB^* & A \end{pmatrix}. \tag{1.31}$$

The mass eigenstates are now

$$\left|P'_{1,2}\right\rangle = \frac{1}{\sqrt{1+|r|^2}}\left(\left|P^0\right\rangle \pm r\left|\overline{P}^0\right\rangle\right). \tag{1.32}$$

The masses and widths of the two states come from the real and imaginary parts of $A \pm |B|$ as before. The mass eigenstates are no longer *CP* eigenstates.

1.2.4.3 $B^0 - \overline{B}^0$ mixing

The neutral B meson mass eigenstates, $|B_L\rangle$ and $|B_H\rangle$, are linear superpositions of the flavour eigenstates, $|B^0\rangle$ and $\left|\overline{B}^0\right\rangle$, as described by Eq. (1.33):

$$\begin{aligned}
|B_L\rangle &= p\left|B^0\right\rangle + q\left|\overline{B}^0\right\rangle,\\
|B_H\rangle &= p\left|B^0\right\rangle - q\left|\overline{B}^0\right\rangle, && (1.33)
\end{aligned}$$

where p and q are complex coefficients (to allow for a phase difference between the two states) that satisfy the normalisation condition

$$|p|^2 + |q|^2 = 1. \tag{1.34}$$

The mass, M, mass difference, Δm_B, and lifetime difference, $\Delta\Gamma_B$, are useful quantities to define:

$$\begin{aligned} M &= \frac{1}{2}\left(M_H + M_L\right), \\ \Delta m_B &= M_H - M_L, \\ \Delta\Gamma_B &= \Gamma_H - \Gamma_L; \end{aligned} \tag{1.35}$$

$M_{H,L}$ and $\Gamma_{H,L}$ are specific cases of the general quantities described by Eq. (1.20).

It has been shown experimentally [23] that

$$\Delta m_B \gg \Delta\Gamma_B \approx \mathcal{O}(1\%) \tag{1.36}$$

(this is in contrast to the K system in which $\Delta\Gamma_K \approx -2\Delta m_K$, with the mass eigenstates therefore dubbed short (S) and long (L) rather than light (L) and heavy (H)).

A general B state $|\psi(t)\rangle$ propagating through space is an admixture of the mass eigenstates, and evolves with time like:

$$|\psi(t)\rangle = a_L(t)\,|B_L\rangle + a_H(t)\,|B_H\rangle\,, \tag{1.37}$$

where time dependent amplitudes $a_L(t)$ and $a_H(t)$ are solutions of the TDSE and are given by:

$$\begin{aligned} a_L(t) &= a_L(0)e^{-iM_Lt}e^{-\frac{1}{2}\Gamma_Lt}, \\ a_H(t) &= a_H(0)e^{-iM_Ht}e^{-\frac{1}{2}\Gamma_Ht}. \end{aligned} \tag{1.38}$$

It can be seen from Eq.s (1.33), (1.37) and (1.38) that for a pure $|B^0\rangle$ state at time $t = 0$, we require

$$a_L(0) = a_H(0) = \frac{1}{2p}, \tag{1.39}$$

and for a pure $\left|\overline{B}^0\right\rangle$ state at time $t = 0$,

$$a_L(0) = -a_H(0) = \frac{1}{2q}. \tag{1.40}$$

Substituting Eq.s (1.33) and (1.38) into (1.37) and using Eq.s (1.35), (1.39) and (1.40), we are now able to show how a state that is initially pure $|B^0\rangle$ evolves to become pure $\left|\overline{B}^0\right\rangle$ (and vice versa) and oscillates between the two states (taking $\Gamma_L = \Gamma_H = \Gamma$ as implied by Eq. (1.36)):

$$\begin{aligned} \left|B^0(t)\right\rangle &= g_+(t)\left|B^0\right\rangle + \frac{q}{p} g_-(t)\left|\overline{B}^0\right\rangle, \\ \left|\overline{B}^0(t)\right\rangle &= \frac{p}{q} g_-(t)\left|B^0\right\rangle + g_+(t)\left|\overline{B}^0\right\rangle, \end{aligned} \tag{1.41}$$

with

$$\begin{aligned} g_+(t) &= e^{-iMt} e^{-\frac{1}{2}\Gamma t} \cos\left(\frac{\Delta m_B t}{2}\right), \\ g_-(t) &= e^{-iMt} e^{-\frac{1}{2}\Gamma t} \sin\left(\frac{\Delta m_B t}{2}\right). \end{aligned} \tag{1.42}$$

Thus the probability that a state that starts off as a pure $|B^0\rangle$ will decay as a $\left|\overline{B}^0\right\rangle$ oscillates sinusoidally with a frequency that depends on Δm_B.

1.2.4.4 Flavour tagging

The *BABAR* experiment produces $B\overline{B}$ pairs in a coherent $L = 1$ state. The two B mesons therefore have a common wavefunction that evolves in time such that there is always exactly one B^0 and one $\overline{B}^0$. At this stage however the flavour of each meson is indeterminate. This is an example of the Einstein-Podolski-Rosen effect [24]. At the instant one B decays, let us say as a B^0 and at time $t = t_{tag}$, the wavefunction collapses. Knowledge of one meson has an instantaneous influence on the other, namely that in our example its flavour must be $\overline{B}^0$ at time $t = t_{tag}$. This $\overline{B}^0$ will continue to evolve with time until it itself decays at time $t = t_{decay}$.

The flavour of the first B to decay can be *tagged* by studying its decay products. The flavour of the other B can then be established at time $t = t_{tag}$. The amount of time that it has to mix before decaying is simply given by $t_{decay} - t_{tag}$.

The method of flavour tagging has been used by *BABAR* collaborators to measure the $B^0 - \overline{B}^0$ oscillation frequency, Δm_B. See, for example, [25]. The world average

value is $\Delta m_B = 0.502 \pm 0.007\,\mathrm{ps}^{-1}$ [23].

1.2.5 Three types of *CP* violation

CP violation can manifest itself in the SM in three distinct ways. Each of these are now discussed.

1.2.5.1 *CP* violation in decay

CP violation in decay, also called *direct CP violation*, occurs when the amplitude for a particular decay is different to that of the *CP* conjugate decay.

The amplitude for the decay $B \to f$, where f is any final state, can be written

$$A_f = \langle f\,|\mathbf{H}|\,B\rangle = \sum_j A_j e^{i(\delta_j+\phi_j)}, \tag{1.43}$$

where A_j, δ_j, and ϕ_j are the amplitude, *strong phase* and *weak phase* respectively of a contributing process. Likewise, for the process $\overline{B} \to \bar{f}$, we can write

$$\bar{A}_{\bar{f}} = \left\langle \bar{f}\,|\mathbf{H}|\,\overline{B}\right\rangle = \sum_j A_j e^{i(\delta_j-\phi_j)}. \tag{1.44}$$

The weak phase comes from the electroweak interaction terms of the SM Lagrangian. It is this phase that is able to violate *CP* symmetry; it appears in A_f and $\bar{A}_{\bar{f}}$ with opposite signs. The strong phase is *CP* conserving and appears in A_f and $\bar{A}_{\bar{f}}$ with the same sign. For any particular term in A_j, both the weak and strong phases are convention dependent such that an overall phase rotation of the entire amplitude would have no overall observable impact. The relative phase differences however are phase convention independent and are therefore physically meaningful.

For *CP* violation we require that

$$A_f \neq \bar{A}_{\bar{f}}. \tag{1.45}$$

In order for this condition to be met there must be contributions from at least two processes with different strong and weak phases, as can be seen from Eq. (1.46)

where we have two contributing processes:

$$|A_f|^2 - |\bar{A}_{\bar{f}}|^2 = -2\sum_{ij} A_i A_j \sin(\phi_i - \phi_j)\sin(\delta_i - \delta_j). \tag{1.46}$$

It is helpful when attempting to measure this effect if the processes have amplitudes with similar magnitudes.

Figure 1.3 illustrates (as an example) different processes for a neutral B meson decaying to a charged kaon and a charged pion. The tree diagram is a purely weak interaction, whilst the higher order *penguin* diagram also involves gluon exchange. (As an aside it is worthwhile ensuring that the reader is acquainted with so-called penguin diagrams. A penguin diagram is defined to be one with an internal quark loop that radiates a gluon, photon or Z^0. They are suppressed with respect to the tree-level diagrams due to the additional vertices in the diagram. They can also be CKM suppressed or enhanced with respect to the tree depending on the CKM elements at the vertices). Direct *CP* violation can also occur where only penguin diagrams occur (for example $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$, the channel under investigation in this thesis, for which a tree level diagram does not exist). This is because the quark in the loop in the penguin diagram can be any one of three.

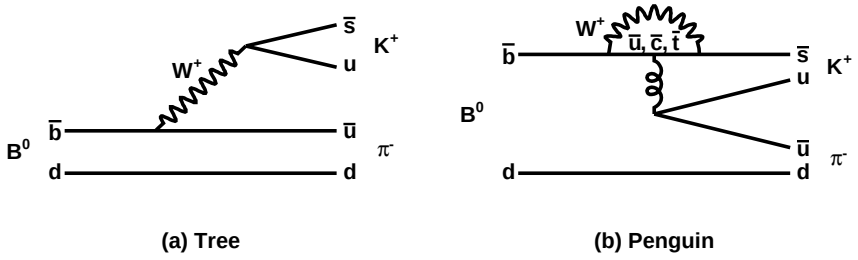

Figure 1.3: *Feynman diagrams for the decay $B^0 \to K^+\pi^-$.*

Direct *CP* violation can occur in both charged and neutral decays. It is the only type that can occur in charged decays, where it is most cleanly observed as it cannot be confused with other types of *CP* violation. An observed asymmetry $\mathcal{A}_{CP}^D$ can be

written using the decay rates (Γ) as:

$$\mathcal{A}_{CP}^{D} = \frac{\Gamma(\overline{B} \to \bar{f}) - \Gamma(B \to f)}{\Gamma(\overline{B} \to \bar{f}) + \Gamma(B \to f)} = \frac{\left|\bar{A}_{\bar{f}}/A_f\right|^2 - 1}{\left|\bar{A}_{\bar{f}}/A_f\right|^2 + 1}. \tag{1.47}$$

Direct *CP* violation has been observed in K decays [26,27] and, more recently, in B decays in which $\frac{\Gamma(B^0 \to K^-\pi^+) - \Gamma(B^0 \to K^+\pi^-)}{\Gamma(B^0 \to K^-\pi^+) + \Gamma(B^0 \to K^+\pi^-)}$ was measured to be -0.133 ± 0.031 by the *BABAR* collaboration [28].

1.2.5.2 *CP* violation in mixing

CP violation in mixing is observed as a difference in the rates for the $P^0 \to \overline{P}^0$ and $\overline{P}^0 \to P^0$ transitions. This difference arises due to interference between the box diagrams (for example in Figure 1.2) proceeding via the t quark and those proceeding via other flavours. It occurs when the mass eigenstates are not *CP* eigenstates:

$$\mathbf{CP}\left|P_{1,2}\right\rangle \neq \pm\left|P_{1,2}\right\rangle \tag{1.48}$$

(i.e. the mass eigenstates are described by Eq. (1.32) rather than Eq. (1.19)).

Following the steps described in Eq.s (1.21)-(1.30), it is easily shown for the B system – as described by Eq. (1.33) – that for $|B_{L,H}\rangle$ to be *CP* eigenstates and *CP* thus conserved, (q/p) must equal 1. However, for a physically meaningful result, we require a phase convention-independent approach (i.e. δ can take on any value, not necessarily zero):

$$\begin{aligned}
\mathbf{CP}\left|B_{L,H}\right\rangle &= \mathbf{CP}\left(p\left|B^0\right\rangle \pm q\left|\overline{B}^0\right\rangle\right) \\
&= \mathbf{CP}\left(|p|\,e^{i\arg(p)}\left|B^0\right\rangle \pm |q|\,e^{i\arg(q)}\left|\overline{B}^0\right\rangle\right) \\
&= |p|\,e^{i\arg(p)}\,e^{i\delta}\left|\overline{B}^0\right\rangle \pm |q|\,e^{i\arg(q)}\,e^{-i\delta}\left|B^0\right\rangle \\
&= \pm\left(|p|\left|\frac{q}{p}\right|\,e^{i(\arg(q)-\delta)}\left|B^0\right\rangle \pm |q|\left|\frac{p}{q}\right|\,e^{i(\arg(p)+\delta)}\left|\overline{B}^0\right\rangle\right) \\
&= \pm\left|B_{L,H}\right\rangle, \text{ if} \qquad (1.49) \\
\left|\frac{q}{p}\right|\,e^{i(\arg(q)-\delta)} &= e^{i\arg(p)}, \text{ and} \qquad (1.50) \\
\left|\frac{p}{q}\right|\,e^{i(\arg(p)+\delta)} &= e^{i\arg(q)}. \qquad (1.51)
\end{aligned}$$

It can be seen that there is always a value of δ that satisfies Eq. (1.49), provided $|q/p| = 1$.

For *CP* violation then we require that:

$$\left|\frac{q}{p}\right| \quad \left(= \left|\sqrt{\frac{\left\langle \overline{B}^0 \left|\mathbf{H}\right| B^0\right\rangle}{\left\langle B^0 \left|\mathbf{H}\right| \overline{B}^0\right\rangle}}\,\right|\right) \quad \neq 1. \tag{1.52}$$

CP violation in mixing is also known as *indirect CP violation*, and was the first type to be shown to exist experimentally using the K system. The level of indirect *CP* violation in the B system is expected to be small ($\mathcal{O}(10^{-4})$) [29]. This is because

$$\text{Eq. (1.36)} \Rightarrow |q/p| \approx 1. \tag{1.53}$$

1.2.5.3 *CP* violation in interference between decays with and without mixing

Interference between mixing and decay processes leads to a third type of *CP* violation in the SM. This is observed in the decays of neutral mesons to the same final state. For the discussion here we take the final state to be a *CP* eigenstate, i.e. $f = \bar{f} = f_{CP}$.

If $A_{f_{CP}}$ is the amplitude of $B^0 \to f_{CP}$ and $\bar{A}_{f_{CP}}$ is the amplitude of $\overline{B}^0 \to f_{CP}$ then we can define the following phase convention independent quantity $\lambda_{f_{CP}}$:

$$\lambda_{f_{CP}} = \frac{q}{p}\frac{\bar{A}_{f_{CP}}}{A_{f_{CP}}}. \tag{1.54}$$

CP violation occurs when $\lambda_{f_{CP}}$ deviates from unity. This can occur due to either direct ($|\frac{\bar{A}_{f_{CP}}}{A_{f_{CP}}}| \neq 1$) or indirect ($|\frac{q}{p}| \neq 1$) *CP* violation. However, it can also occur where there is no *CP* violation in mixing or decay. For this to happen we require that the imaginary part of $\lambda_{f_{CP}}$ take on a non-zero value:

$$\Im\left(\lambda_{f_{CP}}\right) \neq 0; \quad |\lambda_{f_{CP}}| = 1. \tag{1.55}$$

In the case of Eq. (1.55) being satisfied, $\lambda_{f_{CP}}$ is a pure phase that can be calculated without hadronic uncertainties.

Eq. (1.41) can be used to write the time dependent amplitudes for $B^0, \overline{B}^0 \to f_{CP}$:

$$\begin{aligned}
\left\langle f_{CP} \left| \mathbf{H} \right| B^0(t) \right\rangle &= A_{f_{CP}} \left(g_+(t) + \lambda_{f_{CP}} g_-(t) \right), \\
\left\langle f_{CP} \left| \mathbf{H} \right| \overline{B}^0(t) \right\rangle &= A_{f_{CP}} \frac{p}{q} \left(g_-(t) + \lambda_{f_{CP}} g_+(t) \right). \qquad (1.56)
\end{aligned}$$

The rates of the processes are obtained by taking the modulus squared of the amplitudes,

$$\begin{aligned}
\Gamma(t)(B^0 \to f_{CP}) &= \left| \left\langle f_{CP} \left| \mathbf{H} \right| B^0(t) \right\rangle \right|^2 \\
&= |A_{f_{CP}}|^2 e^{-\Gamma t} \Bigg(\frac{1 + |\lambda_{f_{CP}}|^2}{2} + \frac{1 - |\lambda_{f_{CP}}|^2}{2} \cos\left(\Delta m_B t\right) \\
&\quad - \Im\left(\lambda_{f_{CP}}\right) \sin\left(\Delta m_B t\right) \Bigg), \\
\Gamma(t)(\overline{B}^0 \to f_{CP}) &= \left| \left\langle f_{CP} \left| \mathbf{H} \right| \overline{B}^0(t) \right\rangle \right|^2 \\
&= |A_{f_{CP}}|^2 e^{-\Gamma t} \Bigg(\frac{1 + |\lambda_{f_{CP}}|^2}{2} - \frac{1 - |\lambda_{f_{CP}}|^2}{2} \cos\left(\Delta m_B t\right) \\
&\quad + \Im\left(\lambda_{f_{CP}}\right) \sin\left(\Delta m_B t\right) \Bigg), \qquad (1.57)
\end{aligned}$$

where the definitions of $g_\pm$ from Eq. (1.42) have been used and $|q/p| = 1$ (Eq. (1.53)) is taken to be true.

The time dependent asymmetry, $\mathcal{A}^I_{CP}(t)$, is constructed as the difference between the two rates in Eq. (1.57) divided by their sum:

$$\begin{aligned}
\mathcal{A}^I_{CP}(t) &= \frac{\Gamma(t)(\overline{B}^0 \to f_{CP}) - \Gamma(t)(B^0 \to f_{CP})}{\Gamma(t)(\overline{B}^0 \to f_{CP}) + \Gamma(t)(B^0 \to f_{CP})} \\
&= \frac{-\left(1 - |\lambda_{f_{CP}}|^2\right) \cos\left(\Delta m_B t\right) + 2\Im\left(\lambda_{f_{CP}}\right) \sin\left(\Delta m_B t\right)}{1 + |\lambda_{f_{CP}}|^2} \\
&= -C_{f_{CP}} \cos\left(\Delta m_B t\right) + S_{f_{CP}} \sin\left(\Delta m_B t\right), \qquad (1.58)
\end{aligned}$$

where

$$C_{f_{CP}} = \frac{1 - |\lambda_{f_{CP}}|^2}{1 + |\lambda_{f_{CP}}|^2} \quad \text{and} \quad S_{f_{CP}} = \frac{2\Im\left(\lambda_{f_{CP}}\right)}{1 + |\lambda_{f_{CP}}|^2}. \qquad (1.59)$$

With $|\lambda_{f_{CP}}| = 1$ it can be seen that Eq. (1.58) reduces to

$$\mathcal{A}^I_{CP}(t) = \Im\left(\lambda_{f_{CP}}\right) \sin\left(\Delta m_B t\right). \qquad (1.60)$$

For the majority of B decays, the level of direct and indirect *CP* violation is negligible. *CP* violation in interference between decays with and without mixing occurs at a much greater level and, with $|\lambda_{f_{CP}}| \approx 1$, is the easiest type to measure at *BABAR*.

1.2.6 Current knowledge of CKM parameters

The B meson system is an ideal environment in which to measure numerous $\mathbf{V}_{\mathrm{CKM}}$ parameters. Generally speaking, angles in the Unitarity Triangle can be measured using time dependent *CP* studies, whilst branching ratios are used to determine the lengths of the sides.

The Unitarity Triangle angle β has been measured to high precision by both the *BABAR* and Belle [30] collaborations in the form of $\sin 2\beta$ using the so-called golden channel $B^0 \to J/\psi\, K^0_S$ (see Section 1.2.7):

- *BABAR*: $\sin 2\beta = 0.72 \pm 0.05 \pm 0.02$ [31].
- Belle: $\sin 2\beta = 0.67 \pm 0.05 \pm 0.02$ [32].

These measurements rule out the possibility of the area of the Unitarity Triangle being zero, thus providing solid experimental evidence of *CP* violation in the SM.

To measure the angles α and γ at such a precision requires considerably larger datasets. α is best measured using interference of $b \to u$ decay amplitudes with $B^0 - \overline{B}^0$ mixing. However, $b \to u$ is CKM suppressed, and non-negligible penguin contributions (and therefore possible significant direct *CP* violation) further complicate matters. The measurement of γ is even more challenging as it suffers, depending on the method employed, from either experimental (e.g. suppressed rates) or theoretical (e.g. large hadronic uncertainties) difficulties. Methods using B_s mesons are promising but these shall be best utilised using LHC data, and, at the time of writing, the LHC [33] experiments have not yet commenced data-taking. A more detailed discussion of methods used to extract Unitarity Triangle angles can be found

in [34–36].

Let us now summarise the current experimental knowledge of the CKM sector using the four parameters λ, A, $\overline{\rho}$ and $\overline{\eta}$ as defined in the Wolfenstein parameterisation (Eq.s (1.4) and (1.8)). The parameter λ ($= |V_{us}|$) is known to be 0.22 from $K^+ \to \pi^0 \ell^+ \nu_\ell$ decays, with an accuracy of $\approx 2\%$ [37]. The parameter A ($= |V_{cd}|$) is also fairly well known from B decays to charm states. Its value is quoted as $(40.2^{+2.1}_{-1.8}) \times 10^{-3}$ [37]. The parameters $\overline{\rho}$ and $\overline{\eta}$ are less well known. It is useful to plot the various constraints on the $(\overline{\rho}, \overline{\eta})$ plane that are obtained from the measurements of numerous parameters including:

- Unitarity Triangle angles.
- B mixing parameters (Δm_s for $B^0_s - \overline{B}^0_s$ mixing, Δm_d for $B^0_d - \overline{B}^0_d$ mixing ($B^0_d \equiv B^0$, $\Delta m_d \equiv \Delta m_B$ – as defined in Eq. (1.35))).
- The neutral kaon mixing parameter ϵ_K (defined as $\frac{q}{p} = \frac{1-\epsilon_K}{1+\epsilon_K}$, where p and q are the values in the K system analogous to p and q in Eq. (1.33)).
- The two side lengths, which can be written as $R_u = (1 - \lambda^2/2)\,|V_{ub}/V_{cb}|/\lambda$ and $R_t = |V_{tb}/V_{cb}|/\lambda$. Inclusive semileptonic B decays to charmless states can be used to measure $|V_{ub}|$ [38], allowing R_u to be calculated to a precision of $\approx 20\%$. R_t is primarily constrained by mixing analyses since $\Delta m_d \propto |V_{td} V^*_{tb}|$; theoretical hadronic uncertainties lead to an uncertainty in the measurement of $|V_{td}|$ of the order of 20%.

The resulting plot is shown in Figure 1.4. It is generated using the `CKMfitter` package [37]. The non-shaded areas are excluded at 95% CL. For sin2β, 68% CLs are also shown.

CP violation has now been observed in numerous channels; the current values obtained experimentally (at 90% CLs) for the magnitudes of the $\mathbf{V}_{\rm CKM}$ elements

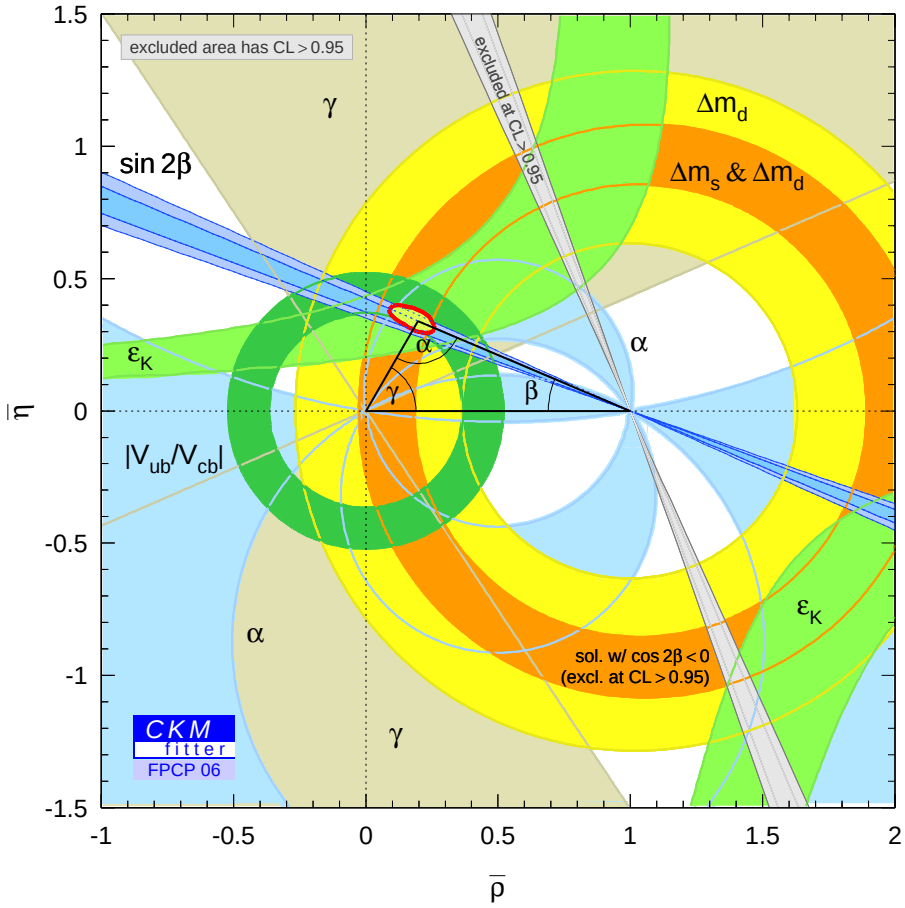

Figure 1.4: *Constraints on the CKM matrix depicted in the $(\overline{\rho}, \overline{\eta})$ plane. The apex of the Unitarity Triangle is constrained to the pale yellow area with the red outline.*

are [23]

$$|\mathbf{V}_{\mathrm{CKM}}| = \begin{pmatrix} 0.9739 - 0.9751 & 0.221 - 0.227 & 0.0029 - 0.0045 \\ 0.221 - 0.227 & 0.9730 - 0.9744 & 0.039 - 0.044 \\ 0.0048 - 0.014 & 0.037 - 0.043 & 0.9990 - 0.9992 \end{pmatrix}. \tag{1.61}$$

Although the constraints depicted in Figure 1.4 are impressive and represent millions of man hours, it is necessary to *overconstrain* the system and thereby be sensitive to deviations from the SM description of flavour changing processes. Any observed discrepancies between measured parameters using independent decay modes would provide a strong indication for physics beyond the SM. This is discussed in Section 1.3.

1.2.7 Measuring $\sin 2\beta$ using $B^0 \to J/\psi K^0_S$

$B^0 \to J/\psi K^0_S$ is dubbed the golden channel at *BABAR* and is the mode responsible for *BABAR*'s most notable result, namely the precision measurement of $\sin 2\beta$ (as presented in Section 1.2.6). Decays of B^0 and $\overline{B}^0$ mesons to the *CP* eigenstate $J/\psi K^0_S$ occur as $\bar{b} \to \bar{c}c\bar{s}$ transitions that can proceed via the Feynman diagrams shown in Figure 1.5. The amplitudes for the tree-level diagram and two of the penguin diagrams are of order λ^2 (this is the Wolfenstein parameter (Section 1.2.2) and is not to be confused with $\lambda_{f_{CP}}$ defined in Eq. (1.54)). The penguin with the u quark and the virtual D^0 exchange diagram involving long distance rescattering are suppressed by an additional factor of λ^2. The total amplitude of $B^0 \to J/\psi K^0_S$ is given by the sum of the individual amplitudes of each of the contributing diagrams and can be written as:

$$A_{f_{CP}} = A_{J/\psi K^0_S} = (V^*_{cb}V_{cs})\,T + (V^*_{tb}V_{ts})\,P_t + (V^*_{cb}V_{cs})\,P_c + (V^*_{ub}V_{us})\,P_u. \tag{1.62}$$

The virtual D^0 exchange diagram has been absorbed into the u quark penguin diagram since it contains the same $V^*_{ub}V_{us}$ term. Rearranging the relevant relation from Eq. (1.6) and substituting into Eq. (1.62) we can write

$$A_{f_{CP}} = A_{J/\psi K^0_S} = (V^*_{cb}V_{cs})\,(T + P_c - P_t) + (V^*_{ub}V_{us})\,(P_u - P_t)\,. \tag{1.63}$$

It follows that for $\overline{B}^0 \to J/\psi K^0_S$,

$$\bar{A}_{f_{CP}} = \eta_{f_{CP}}\bar{A}_{\bar{f}_{CP}} = -\bar{A}_{J/\psi K^0_S} = -\,(V_{cb}V^*_{cs})\,(T + P_c - P_t) - (V_{ub}V^*_{us})\,(P_u - P_t)\,, \tag{1.64}$$

where the *CP* eigenvalue $\eta_{J/\psi K^0_S} = -1$.

Working up to and including order λ^3 in the Wolfenstein parameterisation (in which V_{tb} and V_{ub} are complex and the remaining elements are real) we can write

$$\begin{aligned}
\frac{\bar{A}_{J/\psi K^0_S}}{A_{J/\psi K^0_S}} &= -\frac{(V_{cb}V^*_{cs})\,(T + P_c - P_t) + (V_{ub}V^*_{us})\,(P_u - P_t)}{(V^*_{cb}V_{cs})\,(T + P_c - P_t) + (V^*_{ub}V_{us})\,(P_u - P_t)} \\
&= -\frac{\left(\frac{V_{cb}V^*_{cs}}{V_{cb}V^*_{cs}}\right)A_1 + \left(\frac{V_{ub}V^*_{us}}{V_{cb}V^*_{cs}}\right)A_2}{\left(\frac{V^*_{cb}V_{cs}}{V_{cb}V^*_{cs}}\right)A_1 + \left(\frac{V^*_{ub}V_{us}}{V_{cb}V^*_{cs}}\right)A_2}
\end{aligned}$$

$$= \quad -\frac{A_1 + \kappa\, e^{-i\gamma}\, A_2}{A_1 + \kappa\, e^{+i\gamma}\, A_2}, \tag{1.65}$$

where $V_{ub} = |V_{ub}|e^{-i\gamma}$, $V_{us} = |V_{us}|$, $V_{cb} = |V_{cb}|$, $V_{cs} = |V_{cs}|$, $\kappa = |(V_{ub}V_{us})/(V_{cb}V_{cs})|$, $A_1 = T + P_c - P_t$ and $A_2 = P_u - P_t$; γ is the Unitarity Triangle angle defined in Eq. (1.9).

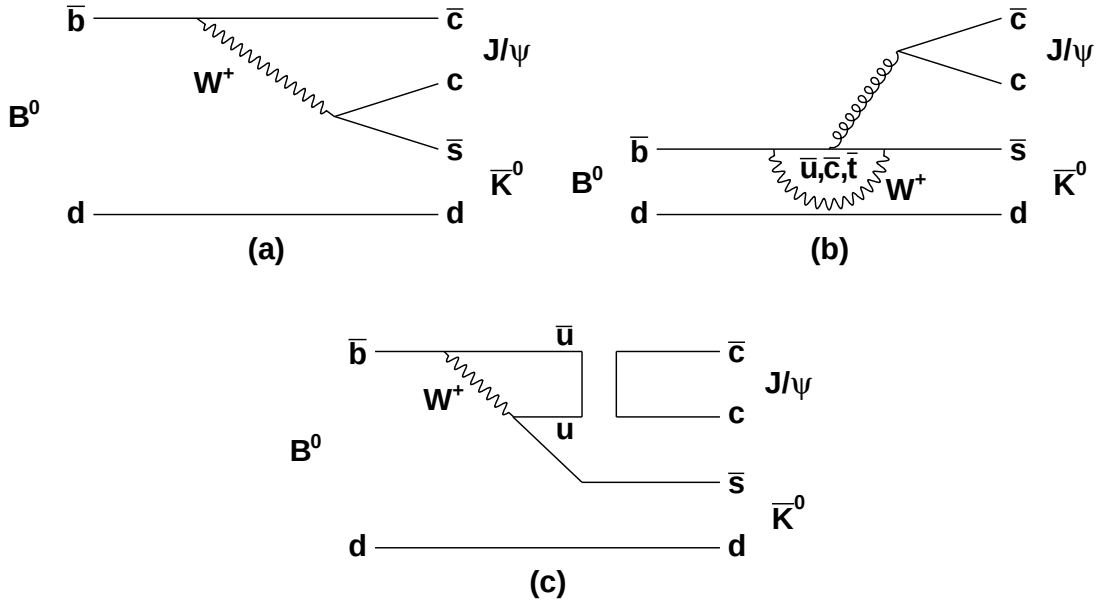

Figure 1.5: *(a) Tree-level diagram, (b) penguin diagrams, and (c) virtual D^0 exchange diagram for the $B^0 \to J/\psi\, K^0_S$ decay mode.*

For $B^0, \overline{B}^0 \to J/\psi\, K^0_S$, Eq. (1.54) can be written as

$$\lambda_{J/\psi\, K^0_S} = \left(\frac{q}{p}\right)_B \left(\frac{\bar{A}_{J/\psi\, K^0_S}}{A_{J/\psi\, K^0_S}}\right) \left(\frac{q}{p}\right)_K . \tag{1.66}$$

The first term represents $B^0 - \overline{B}^0$ mixing. Although the ratio $|q/p|$ is very close to unity (Eq. (1.53)) the relative phase from the box diagram (Figure 1.2 – in which we take the t quark to dominate) must be taken into account when quoting $\lambda_{J/\psi\, K^0_S}$:

$$\left(\frac{q}{p}\right)_B = \frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}} = \left|\frac{V_{tb}V_{td}}{V_{tb}V_{td}}\right| \frac{e^{-i\beta}}{e^{+i\beta}} = e^{-i2\beta} \tag{1.67}$$

where $V_{td} = |V_{td}|e^{-i\beta}$ and $V_{tb} = |V_{tb}|$ (still working up to and including order λ^3 in the Wolfenstein parameterisation). β is the Unitarity Triangle angle defined in

Eq. (1.9). It is also necessary to take into account mixing in the final state, hence the need for the additional multiplicative phase $\left(\frac{q}{p}\right)_K$ in Eq. (1.66). This phase, for $K^0 - \overline{K}^0$ mixing, is given by

$$\left(\frac{q}{p}\right)_K = \frac{V_{cd}^* V_{cs}}{V_{cd} V_{cs}^*} = 1 \text{ up to and including } \mathcal{O}\left(\lambda^3\right). \tag{1.68}$$

Substituting Eq.s (1.65), (1.67) and (1.68) into (1.66) gives

$$\lambda_{J/\psi K_S^0} = -e^{-i2\beta} \frac{A_1 + \kappa\, e^{-i\gamma}\, A_2}{A_1 + \kappa\, e^{+i\gamma}\, A_2}. \tag{1.69}$$

The tree diagram in Figure 1.5 (a) and the dominant penguin diagrams (those containing the c and t quarks) in Figure 1.5 (b) share the same weak phase to a very good approximation [39]. Direct *CP* violation in this mode is therefore expected to be negligible. $A_1 \gg A_2$ and $\kappa < 1$ and so, from Eq. (1.69),

$$\lambda_{J/\psi K_S^0} = -e^{-i2\beta} \tag{1.70}$$

is true to within 1% [23, 40].

Substituting Eq. (1.70) into Eq. (1.60) yields

$$\mathcal{A}_{CP}^I(t) = \sin\left(2\beta\right) \sin\left(\Delta m_B t\right). \tag{1.71}$$

This result, together with the relatively high branching ratio of $B^0 \to J/\psi\, K_S^0$ and its clean reconstruction, allows for a clean, unpolluted measurement of the angle β at *BABAR*.

1.3 Motivation for studying $B^+ \to \overline{K}^*(892)^0 K^+$

The channel $B^+ \to \overline{K}^*(892)^0 K^+$ can be used, together with other modes, to place a SM upper bound on a possible discrepancy between determinations of sin2β using different decay channels of neutral B mesons, namely $B^0 \to J/\psi\, K_S^0$ and $B^0 \to \phi(1020) K_S^0$. An observed discrepancy in excess of such a bound would be a strong indication of new physics beyond the SM.

1.3.1 Measuring $\sin 2\beta$ using $B^0 \to \phi(1020) K^0_S$

Decays of neutral B mesons to the *CP* eigenstate $\phi(1020)K^0_S$ occur as a $\bar{b} \to \bar{s}s\bar{s}$ transition that proceed via penguin diagrams or through a virtual kaon exchange diagram with long distance rescattering. These diagrams are shown in Figure 1.6. The amplitudes for the two dominant penguins are of order λ^2 whilst the third penguin (with the u quark) and the virtual kaon exchange diagram are suppressed by an additional factor of λ^2. Following the same recipe as that presented in Section 1.2.7 we can write an equation for $B^0 \to \phi(1020)K^0_S$ analogous to Eq. (1.69):

$$\lambda_{\phi(1020)K^0_S} = e^{-2i\beta} \frac{A_3 + \kappa\, e^{-i\gamma} A_4}{A_3 + \kappa\, e^{+i\gamma} A_4}, \tag{1.72}$$

where

$$\begin{aligned} A_3 &= p_c - p_t, \\ A_4 &= p_u - p_t \end{aligned} \tag{1.73}$$

(p_t, p_c and p_u are analogous to P_t, P_c and P_u in Section 1.2.7) and the total amplitude is given by

$$A_{\phi(1020)K^0_S} = V^*_{cb} V_{cs} A_3 + V^*_{ub} V_{us} A_4. \tag{1.74}$$

In contrast to $B^0 \to J/\psi\, K^0_S$ in which $A_1 \gg A_2$, for $B^0 \to \phi(1020)K^0_S$ it is expected that $A_3 \sim A_4$ [40]. This is because the leading terms begin at the one-loop order rather than at the tree-level order. Setting $A_3 = A_4$ in Eq. (1.72) results in an additional factor of $\mathcal{O}(\lambda^2) \sim 5\%$ in the determination of $\lambda_{\phi(1020)K^0_S}$, leading to so-called *SM pollution* in the $\sin 2\beta$ measurement using the $B^0 \to \phi(1020)K^0_S$ channel. The parameter $\xi_{f_{CP}}$ is often used to evaluate the level of SM pollution. For $B^0 \to \phi(1020)K^0_S$, this parameter is given by

$$\xi_{\phi(1020)K^0_S} = \frac{V^*_{ub} V_{us} A_4}{V^*_{cb} V_{cs} A_3}. \tag{1.75}$$

Eq. (1.74) can now be rewritten

$$A_{\phi(1020)K^0_S} = V^*_{cb} V_{cs} A_3 \left(1 + \xi_{\phi(1020)K^0_S}\right). \tag{1.76}$$

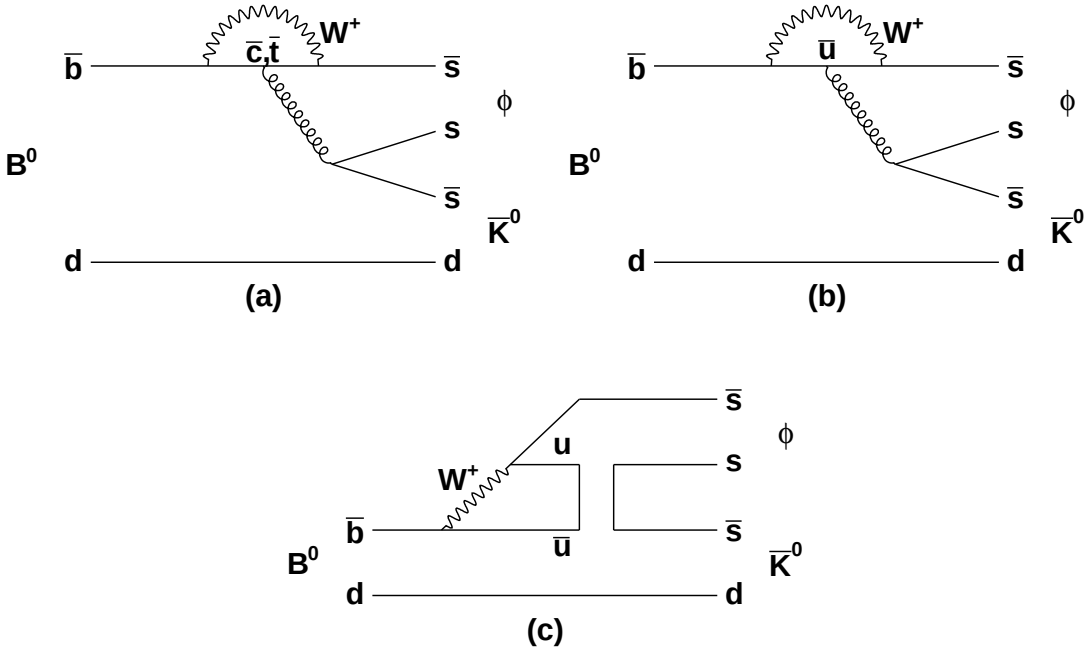

Figure 1.6: *Penguin diagrams (a) (dominant) (b) (suppressed) and (c) virtual kaon exchange diagram for the $B^0 \to \phi(1020)K^0_S$ decay mode.*

It is shown in [10] that, to first order in $\xi_{f_{CP}}$, the deviation from sin2β due to SM pollution is given by:

$$-\eta_{f_{CP}} S_{f_{CP}} - \sin 2\beta = 2\cos 2\beta \sin\gamma \cos\delta_{f_{CP}} |\xi_{f_{CP}}|. \tag{1.77}$$

We can therefore define the parameter $\Delta S_{\phi(1020)K^0_S}$ as the difference between sin2β as measured by the $J/\psi\, K^0_S$ mode (in which $\sin 2\beta = -\eta_{J/\psi\, K^0_S} S_{J/\psi\, K^0_S}$ is taken to be true) and that measured by the $\phi(1020)K^0_S$ mode:

$$\Delta S_{\phi(1020)K^0_S} = 2\cos 2\beta \sin\gamma \cos\delta_{\phi(1020)K^0_S} |\xi_{\phi(1020)K^0_S}|, \tag{1.78}$$

where $\delta_{\phi(1020)K^0_S} = \arg\left(\frac{A_4}{A_3}\right)$ is unknown, β is known to high precision (Section 1.2.6) and γ is taken to lie in the interval 38° to 79° at the 95% CL [37].

1.3.2 Using $B^+ \to \overline{K}^*(892)^0 K^+$ to bound $\Delta S_{\phi(1020)K^0_S}$

Grossman *et al.* introduce a method to bound $\Delta S_{\phi(1020)K^0_S}$ using SU(3) relations [10]. SU(3) flavour symmetry is used to relate the SM pollution terms for the channel[3] $B^+ \to \phi(1020)K^+$ to penguin dominated, strangeness conserving processes, such as those shown in Figure 1.7 for $B^+ \to \overline{K}^*(892)^0 K^+$.

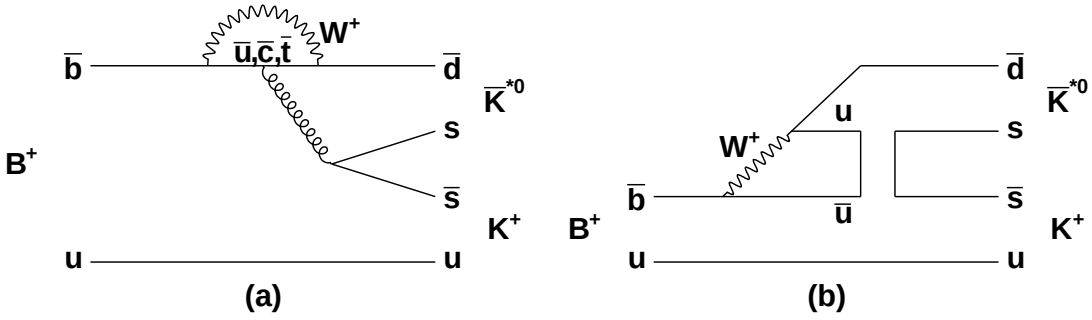

Figure 1.7: *Feynman diagrams for $B^+ \to \overline{K}^*(892)^0 K^+$. The penguin diagram is shown in (a); the long distance final state rescattering diagram is shown in (b).*

We can write general equations for total amplitudes that are dominated by penguin diagrams (analogous to the specific case of Eq. (1.74)) for final states, f and f', with and without non-zero strangeness respectively:

$$A_f = V^*_{cb}V_{cs}A^f_i + V^*_{ub}V_{us}A^f_j, \tag{1.79}$$

$$B_{f'} = V^*_{cb}V_{cd}B^{f'}_i + V^*_{ub}V_{ud}B^{f'}_j, \tag{1.80}$$

where the sub-i and sub-j terms come from the charm penguin minus top penguin and up penguin minus top penguin contributions respectively, as in Eq. (1.73). SU(3) relations provide an upper bound on $|V^*_{cb}V_{cd}A_i + V^*_{ub}V_{ud}A_j|$ in terms of the measured branching ratios (or upper limits on them) of zero strangeness final states [10]. We

[3]Feynman diagrams for this channel are as Figure 1.6 (b) and (c), but with the spectator d quarks replaced with u quarks.

can therefore bound the quantity

$$\hat{\xi_f} = \left| \frac{V_{us}}{V_{ud}} \times \frac{V_{cb}^* V_{cd} A_i + V_{ub}^* V_{ud} A_j}{V_{cb}^* V_{cs} A_i + V_{ub}^* V_{us} A_j} \right| = \left| \frac{\xi_f + \frac{V_{us} V_{cd}}{V_{ud} V_{cs}}}{1 + \xi_f} \right|, \tag{1.81}$$

where

$$\xi_f = \frac{V_{ub}^* V_{us} A_j}{V_{cb}^* V_{cs} A_i} \tag{1.82}$$

is the general case of Eq. (1.75). With $\lambda^2 \ll \hat{\xi_f} < 1$ (appropriate since we want to contain the possibility $|A_j/A_i| \gg 1$), we can take

$$|\xi_f| = \hat{\xi_f} \tag{1.83}$$

to be a good approximation [10].

SU(3) relations lead to the following relations between physical states [10]:

$$A_{i,j}^{\phi(1020)K^+} = B_{i,j}^{\phi(1020)\pi^+} + B_{i,j}^{\overline{K}^*(892)^0 K^+}. \tag{1.84}$$

Using Eq.s (1.79)-(1.81) and (1.83)-(1.84), and taking branching ratios ($\mathcal{B}$) to be the square of the absolute value of the total amplitude, we can write

$$|\xi_{\phi(1020)K^+}| < \frac{\lambda}{1 - \frac{\lambda^2}{2}} \left(\left(\frac{\mathcal{B}(B^+ \to \overline{K}^*(892)^0 K^+)}{\mathcal{B}(B^+ \to \phi(1020) K^+)} \right)^{\frac{1}{2}} + \left(\frac{\mathcal{B}(B^+ \to \phi(1020)\pi^+)}{\mathcal{B}(B^+ \to \phi(1020) K^+)} \right)^{\frac{1}{2}} \right), \tag{1.85}$$

where $\mathcal{B}(B^+ \to \overline{K}^*(892)^0 K^+)$ and $\mathcal{B}(B^+ \to \phi(1020)\pi^+)$ can be either measurements or upper limits. The average measurement for $\mathcal{B}(B^+ \to \phi(1020)K^+)$ using results presented by the *BABAR*, Belle, CLEO and CDF collaborations is $(8.30 \pm 0.65) \times 10^{-6}$ [41–44]. The current upper limit on $\mathcal{B}(B^+ \to \phi(1020)\pi^+)$ is 0.24×10^{-6} as determined by the *BABAR* collaboration at 90% CL [45].

Under the assumption that the up penguin minus top penguin contribution for $B^0 \to \phi(1020)K_S^0$, A_3, is not much larger than the corresponding contribution for $B^+ \to \phi(1020)K^+$, the bound for the charged mode also applies for the neutral mode. The resulting bound on $\xi_{\phi(1020)K_S^0}$ can then be used to bound $\Delta S_{\phi(1020)K_S^0}$ using Eq. (1.78).

For further details please see [10].

1.4 Kinematics of 3-body decays

Consider the decay of a spin 0 B meson, at rest and with mass m_B, to three particles. The three daughter particles have masses m_i, momenta $\vec{p}_i$, energies E_i and four momenta p_i, where $i = 1, 2, 3$ and $p_i^2 \equiv E_i^2 - |\vec{p}_i|^2 = m_i^2$. Defining $p_{ij} = p_i + p_j$, it follows that

$$m_{12}^2 + m_{23}^2 + m_{13}^2 = m_B^2 + m_1^2 + m_2^2 + m_3^2, \tag{1.86}$$

$$m_{ij}^2 = (p_B - p_k)^2 = m_B^2 + m_k^2 - 2m_B E_k, \tag{1.87}$$

with $(i, j, k) = (1, 2, 3)$, $(2, 3, 1)$ or $(1, 3, 2)$. The momenta of the three daughter particles lie in a plane in the B rest frame. Their energies provide information about their orientation with respect to each other.

It is shown in [46] that the decay rate for a scalar particle decaying to a 3-body final state depends not only on the matrix element of the decay $\mathcal{M}$ (where $i\mathcal{M}$ is the Lorentz invariant amplitude), but also on the kinematic constraints of such a decay in the form of available phase space:

$$d\Gamma = \frac{1}{(2\pi)^3} \frac{1}{32 m_B^3} |\mathcal{M}|^2 \, dm_{ij}^2 dm_{jk}^2. \tag{1.88}$$

The *Dalitz plot* [47] is defined as a scatter plot in the variables represented in Eq. (1.88), m_{ij}^2 versus m_{jk}^2. As such, it represents the kinematically allowed available phase space. The boundaries of the Dalitz plot for a given value of m_{ij}^2 occur when $\vec{p}_j$ is parallel or anti-parallel to $\vec{p}_k$, yielding [46]

$$\begin{aligned} (m_{jk}^2)_{min} &= (E_j + E_k)^2 - (p_j + p_k)^2, \\ (m_{jk}^2)_{max} &= (E_j + E_k)^2 - (p_j - p_k)^2. \end{aligned} \tag{1.89}$$

Decays that proceed only according to phase space are uniformly distributed over the Dalitz plot, whilst non-uniformity in the plot indicates a kinematics dependent matrix element.

1.4.1 Resonance variables

For $B^+ \to K^+K^-\pi^+$, it is most useful to construct the Dalitz plot in the variables $m^2_{K^+K^-}$ and $m^2_{K^-\pi^+}$. $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ then appears as a band around $m_{K^-\pi^+} = m_{K^*(892)^0}$ where $m_{K^*(892)^0}$ is the mass of the $K^*(892)^0$. This is demonstrated in Figure 1.8 in which it can be seen that a suitable window in $m_{K^-\pi^+}$ can be defined to study the $B^+ \to \overline{K}^*(892)^0K^+$ contribution. It can also be seen that for some resonant bands, there are non-uniform distributions in both dimensions of the Dalitz plot. This is the case for the $B^+ \to \overline{K}^*(892)^0K^+$ contribution in which events are more concentrated towards the minimum and maximum kinematically allowed values of $m_{K^+K^-}$ and less concentrated towards the centre.

The structure in m^2_{jk}, where the intermediate resonance appears as a band in m^2_{ij}, is determined by the *helicity angle* of the resonance. The helicity of particle i is given by

$$\lambda_i = \frac{\vec{p_i} \cdot \vec{s_i}}{|\vec{p_i}|}, \tag{1.90}$$

where $\vec{p_i}$ is the particle's momentum and $\vec{s_i}$ is its spin. The helicities of the B, K and π mesons in $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ are zero since they are spin 0 particles. By conservation of angular momentum the helicity of the intermediate resonance, $K^*(892)^0$, must therefore be zero. The angular distribution of the decay of the intermediate resonance (with spin s_R) into two scalar daughters is given by a matrix element with the value $|P_{s_R}(\cos\theta_H)|^2$. P_{s_R} is a Legendre polynomial of order s_R [48]; θ_H is the helicity angle of the resonance. The helicity angle is the angle between the momentum vector of the intermediate resonance in the B rest frame and the momentum vector of one of the resonance daughters in the resonance rest frame. The daughters of spin 0 scalar resonances are thus expected to be uniform in $\cos\theta_H$, whilst the daughters of spin 2 tensor resonances will be distributed according to $|3\cos^2\theta_H - 1|^2$. For spin 1 vector resonances, such as $K^*(892)^0$, the daughters will be distributed as $\cos^2\theta_H$.

The helicity angle of a resonance in m^2_{ij} can be written in terms of m^2_{jk} as

$$\cos\theta_H = \frac{(m^2_{jk})_{max} + (m^2_{jk})_{min} - 2m^2_{jk}}{(m^2_{jk})_{max} - (m^2_{jk})_{min}}. \tag{1.91}$$

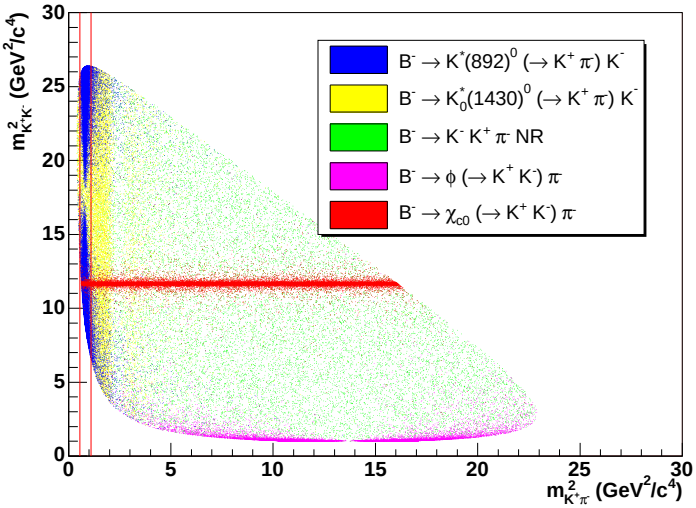

Figure 1.8: *Example $B^+ \to K^+K^-\pi^+$ Dalitz plot showing simulated events for the contributions listed in the legend. The solid, red vertical lines represent a suitable window in which to study $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$.*

1.4.2 Interference

Modes in the Dalitz plot that decay to the same final state interfere with each other quantum mechanically. For two modes with matrix elements $\mathcal{M}_1$ and $\mathcal{M}_2$, the $|\mathcal{M}|^2$ term in Eq. (1.88) can be written as

$$\begin{aligned} |\mathcal{M}|^2 &= \left|\mathcal{M}_1 + \mathcal{M}_2\, e^{i\delta}\right|^2 \\ &= |\mathcal{M}_1|^2 + |\mathcal{M}_2|^2 + 2\Re\left(\mathcal{M}_1\mathcal{M}_2^*\, e^{i\delta}\right). \end{aligned} \tag{1.92}$$

The effect of the interference terms is proportional to the area of overlap between resonances in the Dalitz plot.

The resonant matrix elements are the product of two parts, a dynamical amplitude and an angular amplitude. The angular amplitude is described by $P_{s_R}(\cos\theta_H)$ as defined in Section 1.4.1. The orthogonality property of Legendre polynomials,

$$\int_{-1}^{+1} P_m(\cos\theta_H)P_n(\cos\theta_H)\,d\cos\theta_H = \frac{2}{2n+1}\delta_{nm}, \tag{1.93}$$

ensures that the decay rate (Eq. (1.88)) is unaffected by interference between resonances with different spins when taken over the full $\cos\theta_H$ range[4].

1.4.3 Invariant mass lineshapes

The time dependent wave function for a particle with mass m_0 and decay rate Γ is given by

$$\psi(t) = \psi(0)e^{-t(im_0+\Gamma/2)}. \tag{1.94}$$

The amplitude of this state as a function of E is obtained by performing a Fourier transform:

$$A(E) = \int \psi(t)e^{iEt}\,dt = \frac{N}{(m_0 - E) - i\Gamma/2}, \tag{1.95}$$

where N is an arbitrary constant. The probability of finding the state with energy E is then

$$|A(E)|^2 = \frac{N^2}{(m_0 - E)^2 + \Gamma^2/4}. \tag{1.96}$$

This function is the *Breit-Wigner* curve, and is the expected shape for resonances with $\Gamma \ll E$. For broad resonances ($\Gamma \sim E$) the effect of changes in phase space must also be considered. The $K^*(892)^0$ is described by the Breit-Wigner lineshape with $m_0 = (896.1 \pm 0.3)$ MeV/c^2 and $\Gamma = (50.7 \pm 0.6)$ MeV/c^2.

For the $K_0^*(1430)^0$ resonance, the situation is more complicated. The observed lineshape is distorted with respect to the Breit-Wigner shape. The exact mechanism is unknown, but it is believed that the observed lineshape is a superposition of the

[4]This is true from a theoretical viewpoint. Experimental issues however may render this statement unusable, for example a reconstruction efficiency that is not constant as a function of $\cos\theta_H$.

$K_0^*(1430)^0$ and a further spin 0 contribution. A generally accepted description of this $K\pi$ *S-wave* is as yet undetermined. The *LASS* parameterisation [49] is used for the analysis presented in this thesis. This parameterisation has been successfully utilised in several recent *BABAR* analyses (e.g. [50,51]). It originates from the LASS experiment [52] in which $K\pi$ scattering in the reaction $K^-p \to K^-\pi^+n$ is studied. Dunwoodie [49] postulates a generalisation of the LASS parameterisation in order to extend it to the context of the E791 analysis of $D^+ \to K^-\pi^+\pi^+$ [52]. The parameterisation takes the form

$$\mathcal{M} = B\frac{m_{K\pi}}{q\cot(\delta_B+\phi_B)-iq} + Re^{i\phi_R}e^{2i(\delta_B+\phi_B)}\frac{m_0\Gamma_0\frac{m_0}{q_0}}{(m_0^2-m_{K\pi}^2)-im_0\Gamma_0\frac{q}{m_{K\pi}}\frac{m_0}{q_0}}, \quad (1.97)$$

where the first term describes a non-resonant contribution with an effective range (rather than being uniform over the available phase space) and the second term describes the resonant contribution. B, R, ϕ_B and ϕ_R are constants, whilst

$$q = \sqrt{\frac{\left(m_{K\pi}^2-(m_K+m_\pi)^2\right)\left(m_{K\pi}^2-(m_K-m_\pi)^2\right)}{4m_{K\pi}^2}}, \quad (1.98)$$

$$q_0 = q\left(m_{K\pi}=m_0\right), \quad (1.99)$$

$$\cot\delta_B = \frac{1}{aq}+\frac{rq}{2}, \quad (1.100)$$

where r is the effective range and a is the scattering length. m_0 and Γ_0 are the mass and width of the resonant contribution respectively.

Figure 1.9 illustrates the Breit-Wigner and LASS lineshapes. It can be seen that the lineshape has a large impact in the region of $m_{K\pi}$ in which $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ is studied.

1.5 The strong interaction

Although the work presented in this thesis focuses on the weak interaction, it is worthwhile pointing out that the channel under investigation is also subject to QCD effects since the initial and final states are hadronic. Radiative corrections, stemming from gluons with a range of momenta being emitted and absorbed, greatly

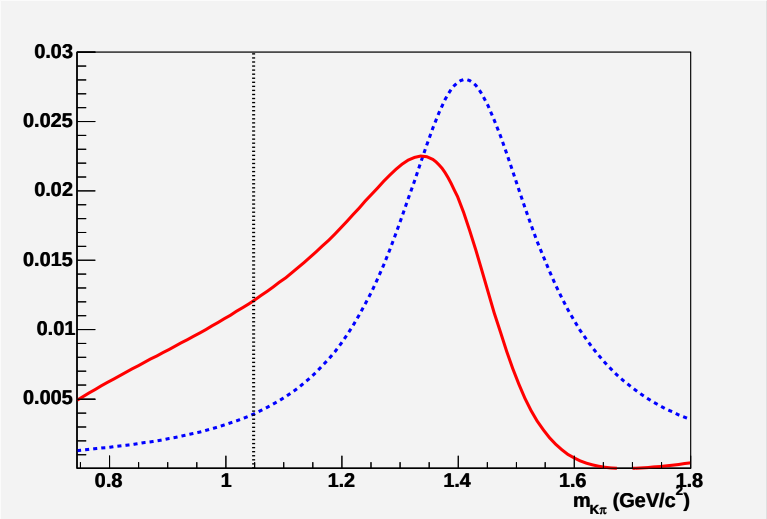

Figure 1.9: *Breit-Wigner (blue, dashed line) and LASS (red, solid line) line-shapes. The integral of each shape over the displayed range is equal. The $m_{K\pi}$ range to the left of the black, dotted line is used to study $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$.*

complicate the calculations of cross-sections and *CP* asymmetries for decays of this nature. Experimental results for such decays play an important role in the ongoing theoretical work[5] being undertaken aimed at estimating these QCD effects.

[5]See for example [53] and [54].

2

The *BABAR* detector

2.1 Introduction

The main motivation behind the construction of the *BABAR* experiment was to provide an environment ideal for carrying out extensive studies of time dependent *CP* violation in the decays of neutral B_d mesons. Further B physics goals are precision measurements of the rates of rare decays of both neutral and charged B mesons. Other objectives include being able to conduct detailed studies of τ, charm and two-photon physics. With these aims in mind, the PEP-II B factory and the *BABAR* detector were constructed. Construction was completed, and data acquisition began, in the summer of 1999. This chapter describes PEP-II and *BABAR*; the physics

requirements, design and performance are discussed.

2.2 The PEP-II accelerator

2.2.1 Overview

The PEP-II B factory is an asymmetric-energy electron positron collider operating at a centre of mass (CM) energy of 10.58 GeV. This energy corresponds to the $\Upsilon(4S)$ resonance, which decays almost exclusively to $B\overline{B}$ pairs ($\approx 50\%$ B^+B^-, $\approx 50\%$ $B^0\overline{B}^0$ [55]). The cross-sections for the production of fermion pairs at 10.58 GeV CM energy are shown in Table 2.1.

$e^+e^- \rightarrow$	***Cross-section*** (nb)	***Production Rate*** (Hz) ***At Design Luminosity*** ($3.0 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
$b\overline{b}$	1.05	3.2
$c\overline{c}$	1.30	3.9
$s\overline{s}$	0.35	1.1
$u\overline{u}$	1.39	4.2
$d\overline{d}$	0.35	1.1
$\tau^+\tau^-$	0.94	2.8
$\mu^+\mu^-$	1.16	3.5
e^+e^-	~ 40	~ 120

Table 2.1: *e^+e^- production cross-sections at CM energy 10.58 GeV [34]. For e^+e^- scattering, the cross-section given applies only within detector coverage.*

Use of an asymmetric collider (9.0 GeV e^-, 3.1 GeV e^+) results in a Lorentz boost to the $\Upsilon(4S)$ of $\beta_z\gamma = 0.56$ with respect to the laboratory frame. Because the $\Upsilon(4S)$ energy is just above the $B\overline{B}$ production threshold, the B mesons are produced almost

at rest in the CM frame, and therefore inherit the relativistic boost of the $\Upsilon(4S)$. This boost allows the distance between the vertices of the B^0 and the $\overline{B}^0$ decays to be large enough to measure at the required precision. This is essential for time dependent studies of *CP* violation. The fact that the $\Upsilon(4S)$ decays to a coherent $B\overline{B}$ pair means that flavour tagging can be used – as discussed in Section 1.2.4.4 – without the uncertainty of whether the tagged B mixed before assuming the role of tagger. Flavour tagging is also essential for time dependent studies.

PEP-II is a high luminosity machine with a design luminosity of $3.0 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Such a luminosity, together with very clean events (compared to a hadronic environment), facilitate the studies of many rare B meson decays.

PEP-II can also operate at an energy 40 MeV below the $\Upsilon(4S)$. This is below the threshold of $B\overline{B}$ production. Approximately 10% of data are taken in this *off-resonance* mode. These data are used to study continuum ($q\overline{q}$) background ($e^+e^- \to c\overline{c}, s\overline{s}, u\overline{u}, d\overline{d}$). Data taken at a CM energy of 10.58 GeV shall be referred to as *on-resonance* throughout this thesis.

A complete description of the PEP-II machine can be found in [56].

2.2.2 The injection system

Figure 2.1 shows the two mile long linear accelerator (linac), which acts as the source of electrons and positrons injected into PEP-II. Electron bunches, created by an electron gun at one end of the linac, are stored in the north damping ring, before being accelerated by the linac to the required 9.0 GeV and injected into PEP-II's High Energy Ring (HER) at the other end. Separate electron bunches are produced by the electron gun, which are accelerated to an energy of ~ 30 GeV and collided with a stationary Beryllium target. This produces positron bunches, which are stored in the south damping ring before being accelerated by the linac to 3.1 GeV and injected into PEP-II's Low Energy Ring (LER).

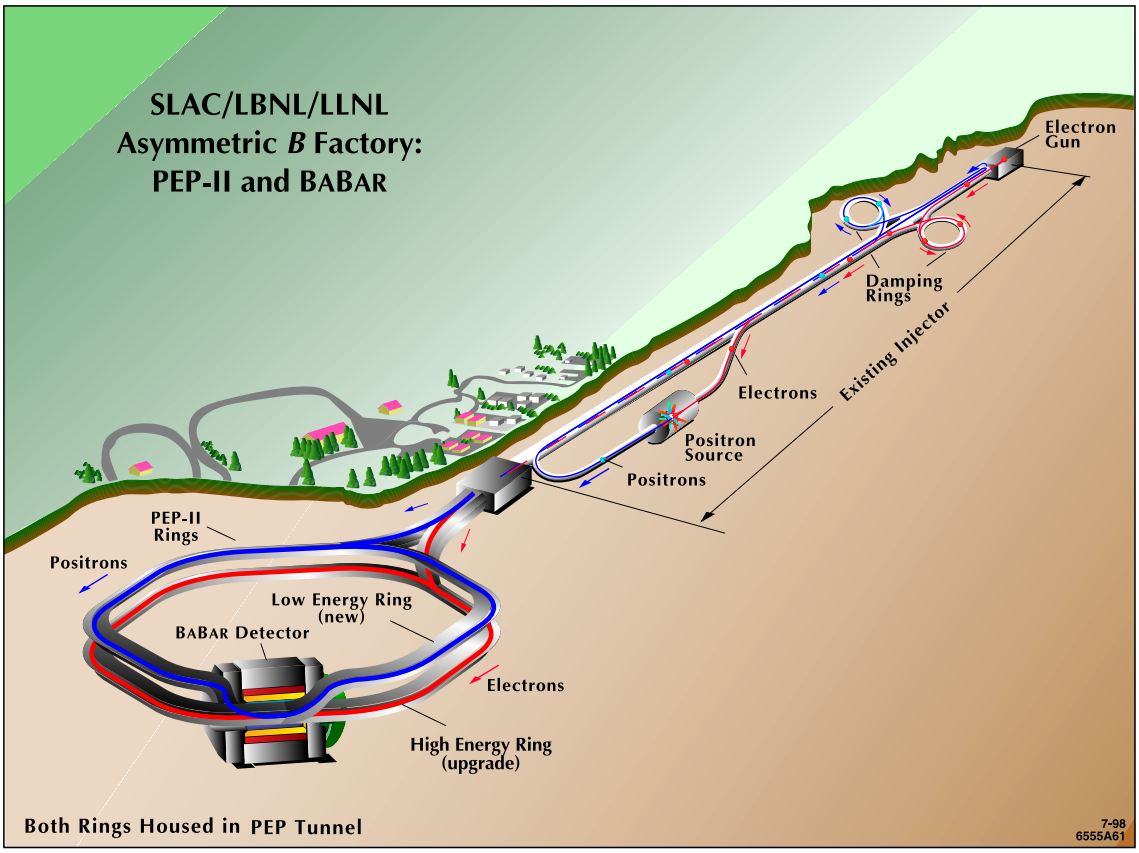

Figure 2.1: *The linac injection system and the PEP-II storage rings.*

Until recently, PEP-II operated by filling up the HER and LER to some maximum value, at which point injection stopped and data taking began. Data taking would continue until the instantaneous luminosity reached a lower limit. It then stopped whilst the injection system topped up the two beams. This cycle was repeated as illustrated by Figure 2.2. This method worked well because machine background is high during injection and for a short time after whilst the beams become stable. The high voltages of the various subdetectors could be ramped down during this period, and the detector was protected from high machine backgrounds, which bring about radiation damage. However, this method is not optimal when considering integrated luminosity delivered.

An alternative approach is to continuously inject into the two rings, whilst simulta-

neously taking data. This is known as *trickle injection*. Trickle injection increases the integrated luminosity, as illustrated by Figure 2.2, but is technically challenging (to both the accelerator and detector teams) and increases machine backgrounds. In November 2003, tests were carried out that successfully demonstrated that, for the LER, these technical challenges could be met, and the machine backgrounds could be managed. In March 2004, these tests were repeated, again successfully for the HER, and for both the LER and HER together. Default injection mode has been trickle injection since March 2004. Injection occurs at a very low rate ($\sim 10\,\mathrm{Hz}$). A short inhibit window is employed immediately after a bunch is injected – whilst background is high – that prevents events for that bunch being accepted by the Level-1 Trigger.

For full details of the trickle injection method, see [57, 58].

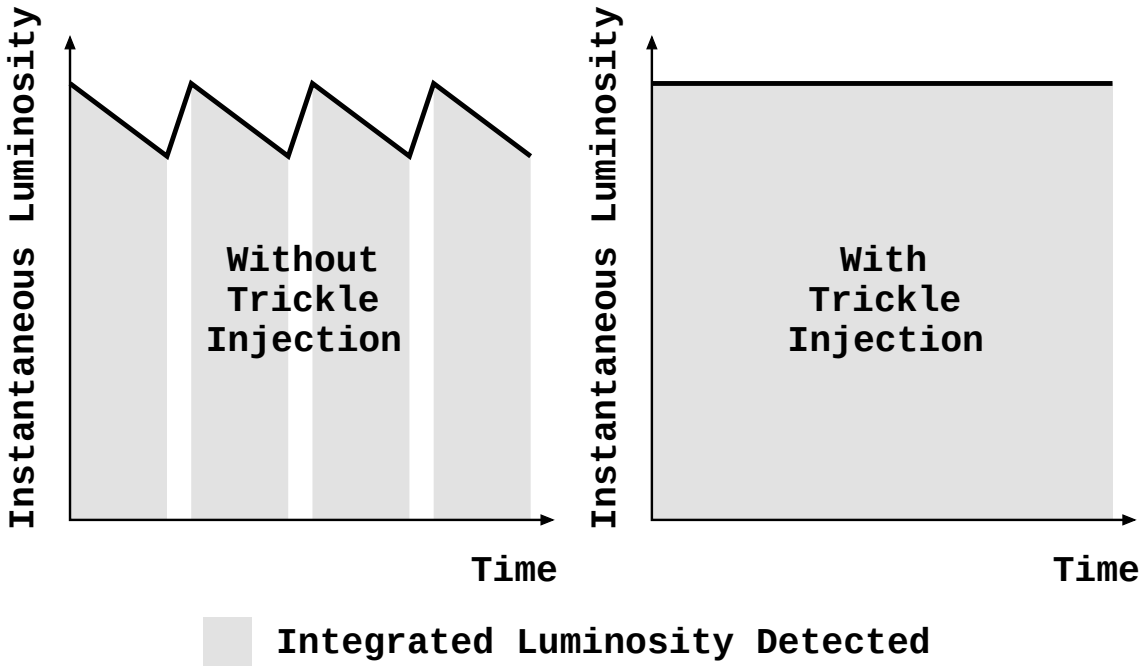

Figure 2.2: *Trickle injection enhances integrated luminosity.*

2.2.3 The interaction region

The high luminosities necessary to meet the requirements of the physics program demand that the distance between bunches in the HER and LER are small (of the order of the length of the *BABAR* detector). As such it is necessary to keep the two beams separated – until just before the interaction point (IP) – in order to prevent secondary collisions occurring, the products of which would be detected by *BABAR*.

To allow the HER and LER beams to meet head on, and to ensure that they do not meet until just before the IP, the beams are subjected to magnetic fields generated by strong dipole magnets, which are located inside the *BABAR* detector, very close to the IP.

Figure 2.3 illustrates how the incoming beams are diverted by the dipole magnets labelled B1. The same magnets are used to separate the beams after they are collided to prevent further, secondary collisions. Focusing of the beams is achieved using sets of quadrupole magnets. Q4 and Q5 (in Figure 2.3) focus the HER beam, Q2 focuses the LER beam. Q4, Q5 and Q2 are iron magnets located outside of the detector. Q1, a permanent magnet that partially enters the detector volume, is a final focus and affects both beams.

The presence of the dipole magnets in the *BABAR* detector, and the synchrotron radiation emitted by the beams as a result of being diverted by these magnets causes obvious problems. Background conditions are adversely effected, and detector components are prone to radiation damage. An alternative approach is employed by the Belle experiment [30], based in Japan, in which the beams are collided at a cross-over angle, rather than head on. This eliminates the need for the dipole magnets and the associated synchrotron radiation. However, during the construction of the *BABAR* experiment, this method was untried and risky. A tried and tested, and more conservative approach was decided upon.

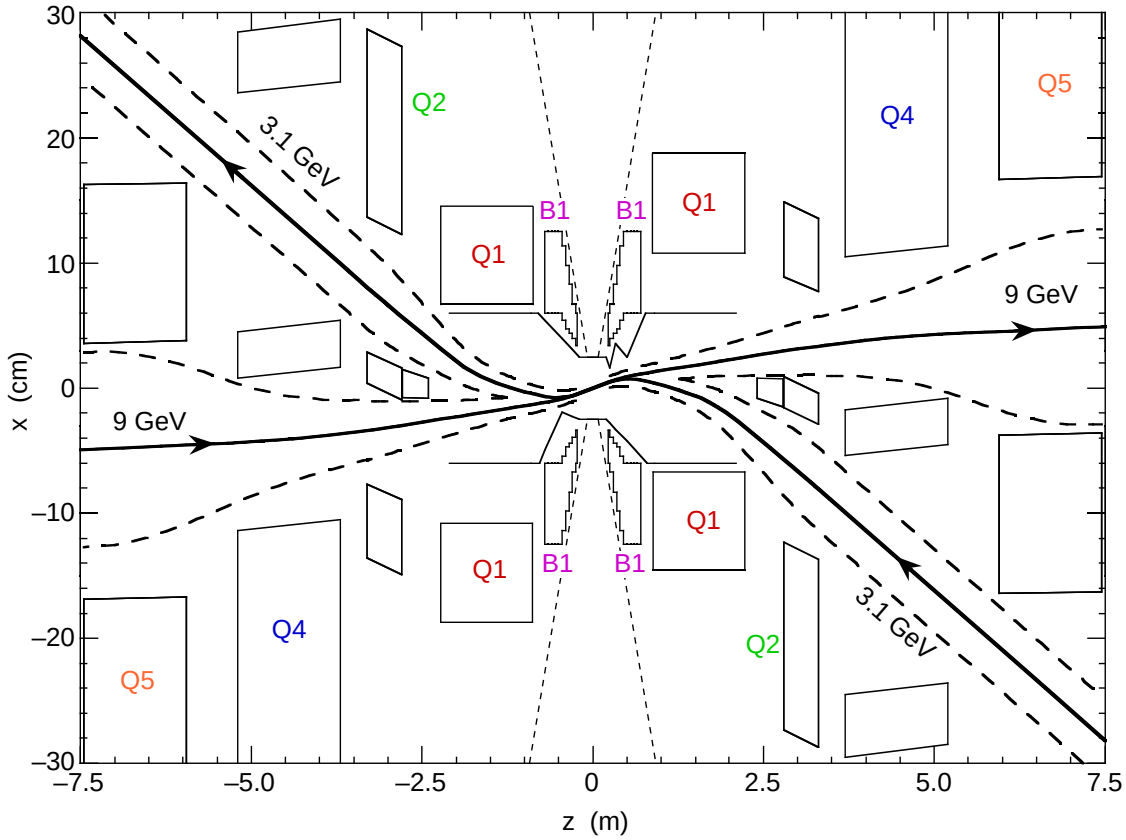

Figure 2.3: *The PEP-II interaction region, with exaggerated vertical scale.*

2.2.4 Machine backgrounds

Machine backgrounds increase occupancy in the *BABAR* detector, causing physics measurements to be degraded. Backgrounds also increase the trigger rate with the adverse effect that deadtime increases (more interesting physics events are lost). Furthermore, machine backgrounds contribute to radiation damage.

There are several sources of machine background. One source is synchrotron radiation, which is due to the complicated optics near the IP – as discussed in Section 2.2.3. The geometry of the interaction region was designed such that the majority of the synchrotron radiation passes through the detector with minimal interaction. Copper masks are also used in an attempt to prevent interaction with the beam pipe.

Gas molecules in the beam pipe bring about another source of machine background. Beam particles undergo bremsstrahlung and Coulomb interactions with these gas molecules, and also with the beam pipe resulting in electromagnetic showers. This background is minimised by keeping a very good vacuum in the beam pipe near the IP.

Beam-beam interactions are also a source of machine background. The electric field generated by one beam can perturb particles in the other beam, some of which may be knocked out of orbit. These particles are then likely to interact with gas molecules or the beam pipe.

A further source of background is due to bhabha scattering. Bhabha scattering is the scattering of an electron and a positron, and so is inevitable in an electron-positron collider. When an electron or positron hits material a short distance from the IP, an electromagnetic shower enters the detector.

Machine background due to bhabha scattering worsens, approximately proportionally, as luminosity is increased. The level of improvement in the luminosity of the PEP-II machine since data taking began (see Section 2.2.5) is reflected by the fact that radiative bhabha debris – minimal when *BABAR* first started running – is now the largest source of machine background. Beam-beam and beam-gas interactions are more dependent on the current carried by each beam rather than the luminosity (beam-beam interactions in a non-trivial way, beam-gas interactions approximately linearly). The overall machine background, therefore, deviates slightly from being linearly related to the luminosity.

2.2.5 Performance

To date PEP-II has performed exceptionally well. During the first year of operation, both instantaneous and integrated design luminosities were achieved, and the collider has gone from strength to strength since. Records for the integrated luminosities

delivered over eight hour, daily, weekly and monthly periods have been broken with impressive regularity. The most recent records are shown in Table 2.2. Recent trickle injection developments have seen a striking increase in performance. Figure 2.4 shows integrated luminosity delivered by month in the period since first collisions in May 1999 up to the end of July 2004.

Parameter	***Design***	***Best Achieved***	***Date Achieved***
HER Current (A)	0.75	1.68	October 3, 2005
LER Current (A)	2.14	2.66	October 3, 2005
Luminosity (10^{33} $cm^{-2}s^{-1}$)	3.000	9.378	October 3, 2005
Luminosity (pb^{-1}/8 hour shift)	—	246.3	May 21, 2004
Luminosity (pb^{-1}/day)	130.0	710.5	May 24, 2004
Luminosity (fb^{-1}/week)	—	4.464	July 25-31, 2004
Luminosity (fb^{-1}/month)	—	17.036	July 2004
Total Delivered Luminosity	$312\,fb^{-1}$		

Table 2.2: *PEP-II machine performance records, as of October 4, 2005. Total delivered luminosity is on- and off-resonance data.*

2.3 The *BABAR* detector

To maximise the physics potential of the asymmetric PEP-II machine, the *BABAR* detector is offset from the IP by 37 cm in the HER direction and is also asymmetric in design. This allows the maximum possible coverage of the $\Upsilon(4S)$ CM frame. *BABAR* is designed as a general purpose detector optimised for its primary physics goals, namely precision measurements of *CP* violation parameters and of beauty, tau and charm physics. To achieve these goals, the detector must satisfy the following:

- High reconstruction efficiency for both charged and neutral particles.

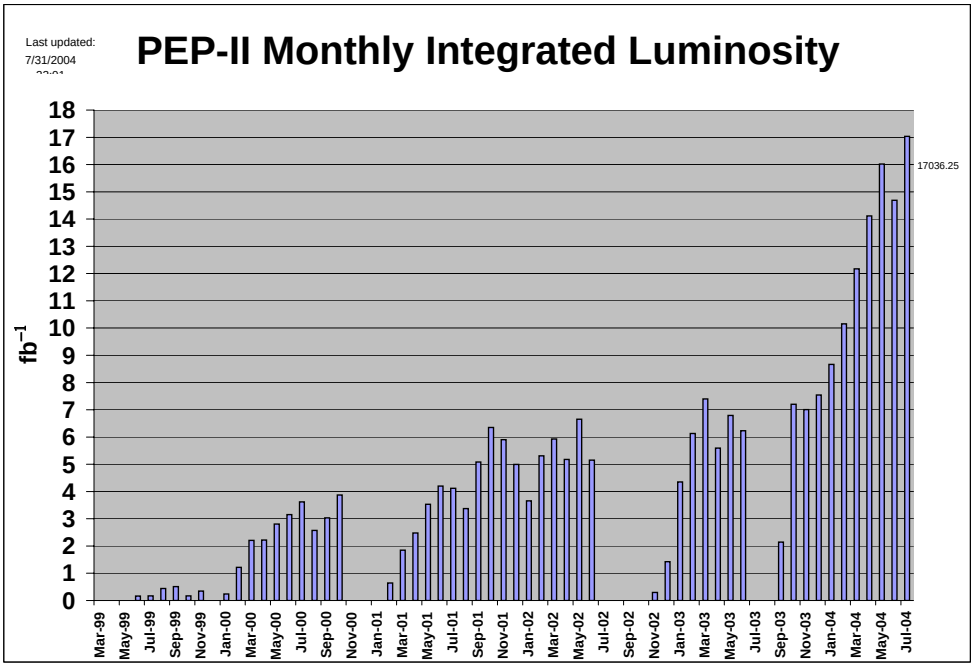

Figure 2.4: *PEP-II integrated luminosity per month.*

- Good position and momentum resolution for charged particles over the range $60\,\mathrm{MeV}/c - 4\,\mathrm{GeV}/c$ (the range of momenta of the decay products for the analysis described in this thesis are illustrated in Figure 2.6). In particular, excellent vertex resolution in the z direction is required to allow accurate measurements of the time difference between the neutral B mesons' decays. Excellent vertex resolution in the x and y directions is also required – to allow reconstruction of secondary charm and tau vertices.

- Good energy and angular resolutions for neutrals in the range $20\,\mathrm{MeV} - 4\,\mathrm{GeV}$ – especially important for π^0 and η detection.

- Excellent particle identification for e, μ, π, K and p over a wide range of

momenta. This is essential for flavour tagging of the non-*CP* B in *CP* violation measurements and for separating important final states such as $K^{\pm}\pi^{\mp}$ and $\pi^{\pm}\pi^{\mp}$.

- A reliable and efficient data acquisition system that can handle the huge volume of data produced at high luminosities.
- An efficient and flexible trigger that can reduce the event rate to a level that the data acquisition system can handle, without losing interesting physics events.
- The detector must be able to operate in the presence of the high machine backgrounds expected at high luminosities.

To achieve the criteria set out above, whilst also taking into account considerations including financial cost and reliability, the final design is one that consists of a system of five subdetectors and a superconducting solenoidal electromagnet (with a 1.5 T magnetic field). This is illustrated in Figure 2.5. Each of the subdetectors shall be discussed in more detail in the following sections. They are, starting with the innermost, the Silicon Vertex Tracker (SVT), the Drift Chamber (DCH), the Detector of Internally Reflected Cherenkov Radiation (DRC), the Electromagnetic Calorimeter (EMC) and the Instrumented Flux Return (IFR). These are labelled the Vertex Detector, the Tracking Chamber, the Cherenkov Detector, the Electron/Photon Detector and the Muon/Hadron Detector respectively in Figure 2.5. The first four of these systems are enclosed in the 1.5 T magnetic field of the superconducting magnetic coil, also marked in Figure 2.5.

A complete description of the *BABAR* detector can be found in [59].

2.3.1 The *BABAR* co-ordinate system

BABAR uses a right-handed co-ordinate system, with the origin at the IP. The z-axis corresponds to the principle axis of the Drift Chamber in the direction of the HER.

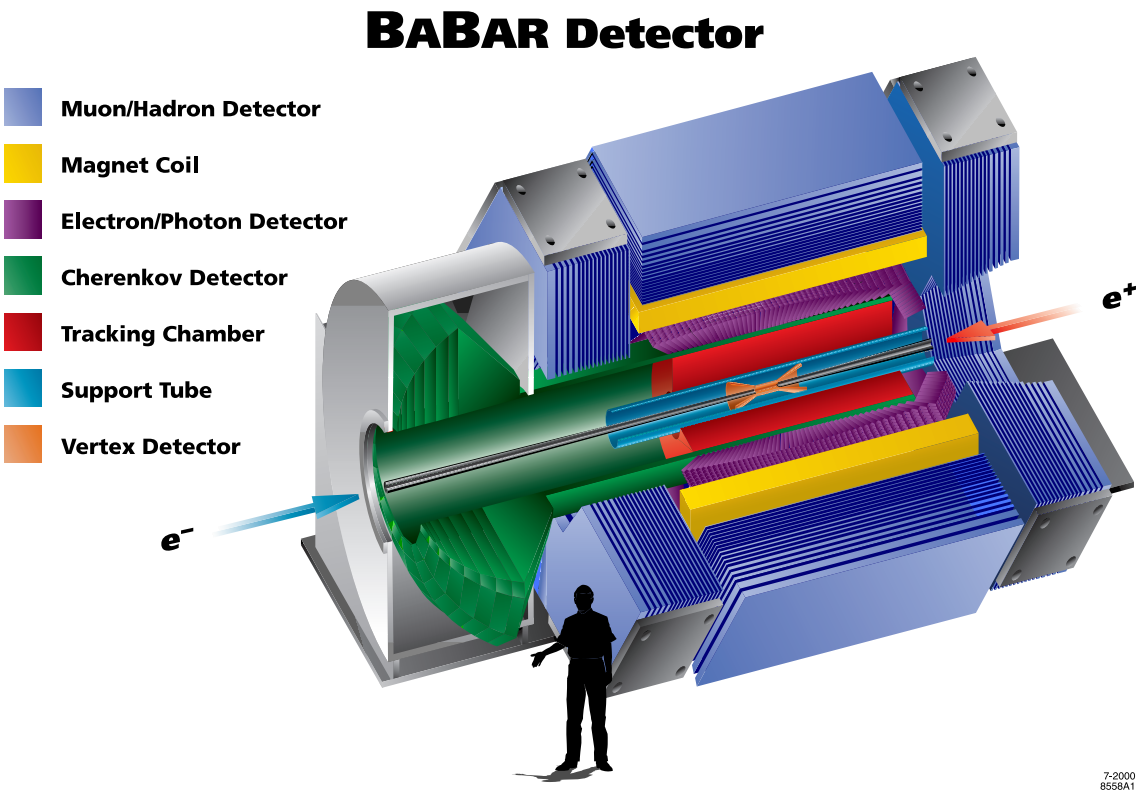

Figure 2.5: *The BABAR detector.*

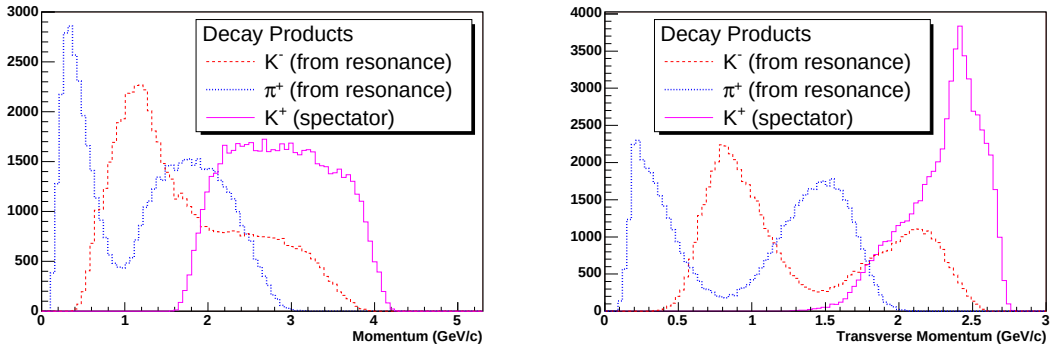

Figure 2.6: *Momentum, $|\vec{p}|$, and transverse momentum, p_T, distributions – in the laboratory frame – for reconstructed $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ Monte Carlo events.*

The x-axis points horizontally outwards from the centre of the PEP-II ring; the y-axis points vertically upwards. The polar (θ) and azimuthal (ϕ) angles are defined

as in the standard spherical co-ordinate system. This is illustrated in Figure 2.7.

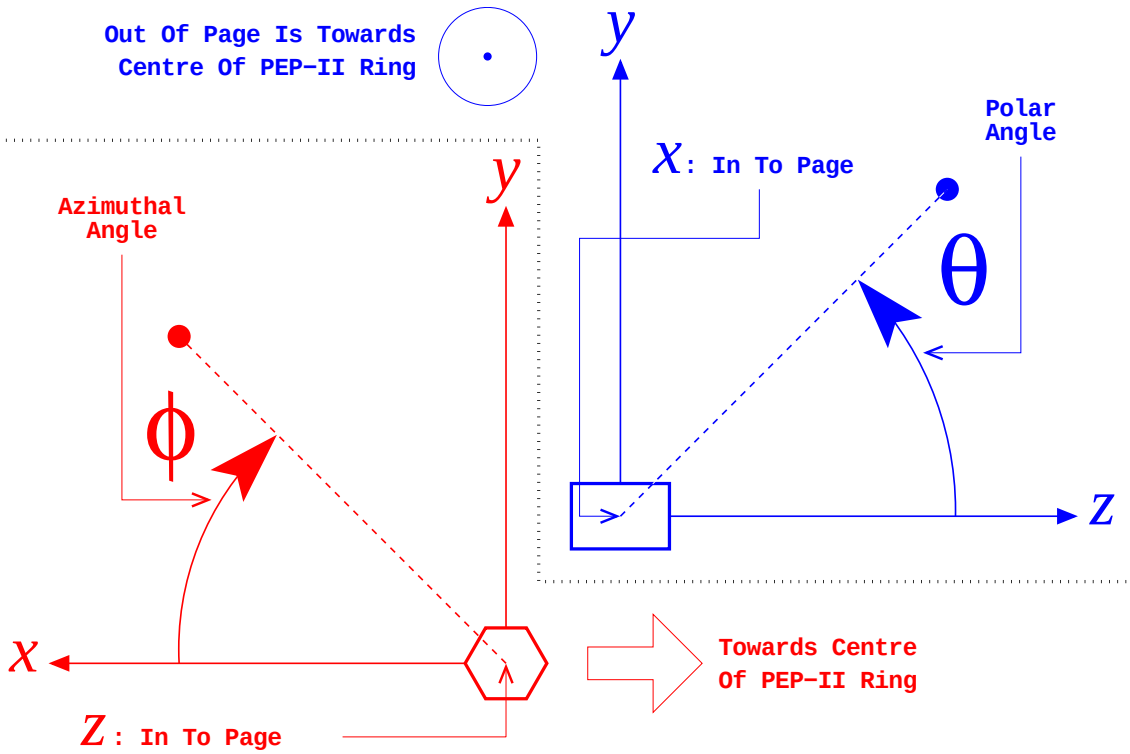

Figure 2.7: *The BABAR co-ordinate system.*

2.4 The Silicon Vertex Tracker (SVT)

2.4.1 SVT physics requirements

For all time dependent *CP* asymmetry measurements the difference between the vertices of the two B mesons must be measured to high precision. To satisfy this primary physics goal, the SVT is required to measure this difference with an accuracy of $\approx 135\,\mu$m. Monte Carlo simulation studies show that this corresponds to a single vertex resolution of $\approx 80\,\mu$m [60]. Other physics goals place further requirements on the SVT: a resolution of $100\,\mu$m in the x-y plane is required in order to make precise measurements of secondary decay vertices, such as D mesons and τ leptons.

The primary tracking system in the *BABAR* detector, the DCH, will not reliably detect tracks with low transverse momentum[1], p_T, since its coverage for such tracks is limited (indeed some low-p_T particles exit the *BABAR* detector without ever making it to the DCH). The SVT must therefore provide excellent efficiency for such tracks. These include slow pions from decays of D^* particles, which are very common B-decay products. Furthermore, the track angle measured by the SVT for high p_T tracks is used to achieve maximum resolution on measurements of the Cherenkov angle in the DRC. Finally, the SVT is used for particle identification, measuring the energy loss (dE/dx) of particles with momenta less than 700 MeV/c.

2.4.2 SVT design

The SVT is an asymmetric p-n junction semi-conductor detector. Semi-conductor detectors provide excellent resolution due to their fine granularity. Additionally, the energy required to produce an electron-*hole* pair in a semi-conductor is ~ 10 times less than the energy required to create a electron-*ion* pair in a gas-based detector, such as the DCH. A further benefit is their size. Semi-conductor detectors are very small and compact.

The principle of operation of a semi-conductor detector is as follows: if an ionising particle penetrates the detector it produces electron-hole pairs along its track, the number being proportional to the energy loss of the charged particle. An externally applied electric field collects the pairs; electrons drift towards the anode, holes to the cathode. The collected charge produces a current pulse on the electrode, whose integral equals the total charge generated by the incident particle.

In addition to the physics requirements described in Section 2.4.1, there are other factors that affect the design of the SVT. The design is influenced by the presence of magnets needed for the PEP-II machine close to the IP (discussed in Section 2.2.3). This places (manageable) constraints on the acceptance of the SVT, with 10% of

[1]Momentum in the x-y plane.

the solid angle in the CM frame not covered. The acceptance in the polar angle (θ) is 20.1° to 150.2°. The SVT design must also take into account the high levels of radiation that it will experience during its lifetime. The budget used in the design was 2 MRad and an instantaneous limit of up to 1 Rad/ms. The design must limit the amount of material through which the tracks pass in order to reduce multiple scattering and bremsstrahlung.

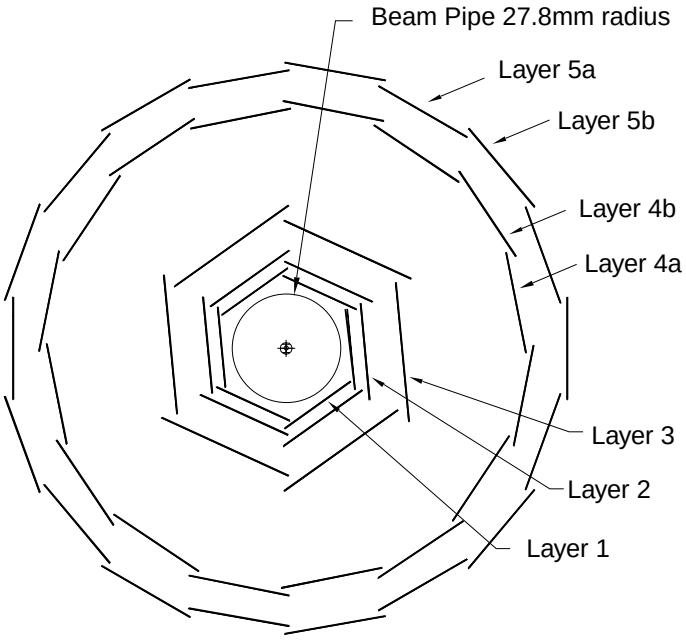

Figure 2.8: *End on view of the SVT showing the five layer structure.*

The SVT consists of five layers of silicon strip sensors, divided azimuthally into modules, as shown in Figure 2.8. Layers $1-3$ have 6 modules each, which are rotated by 5° in ϕ so that they overlap slightly, both to provide complete coverage and to aid with alignment. They have radii 32, 40 and 54 mm respectively. Layers 4 and 5 have 16 and 18 modules respectively. These are arch shaped in the longitudinal plane to increase angular coverage whilst minimising the amount of material that tracks pass through. This is illustrated in Figure 2.9. For arch shaped strips overlapping is achieved by splitting the modules between two sub-layers at slightly different radii (124 and 127 mm for layer 4, 140 and 144 mm for layer 5). The silicon strip sensors

are double-sided, the sensors on the two sides being orthogonal. That is, the inner side gives z measurements and the outer side provides ϕ measurements. Modules are split into *wafers* in the z direction, as it is impractical to produce silicon strips that are as long as the length of the SVT. For layer 1, there are 4 such wafers; for layer 2 there are also 4, 6 for layer 3, 7 for layer 4 and 8 for layer 5.

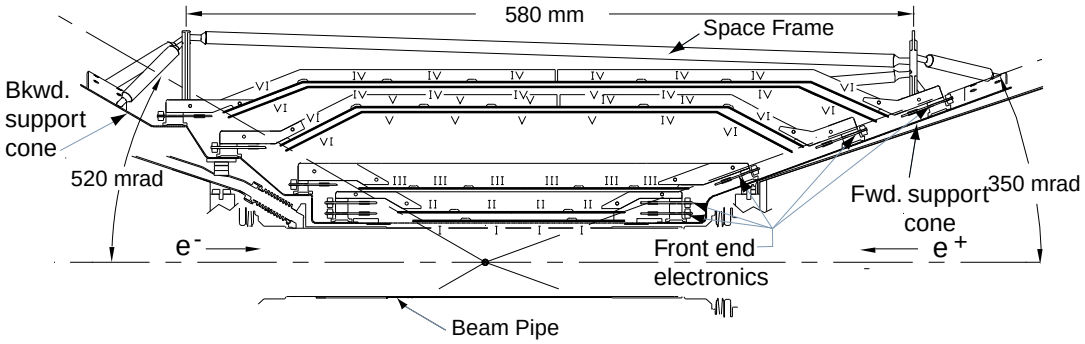

Figure 2.9: *Side on view of the SVT showing the five layer structure, and the arch design of the outer two layers.*

The main purpose of the inner two layers is to precisely measure vertex positions. These layers are also used for angular measurements (since multiple scattering is minimal). The outer layers (4 and 5) are mainly used for alignment with the DCH. Layer 3 provides momentum information used for low p_T tracks that do not make it to the DCH.

Each module is split into a front and backward section (in z), which are electrically isolated. Each half module is read out by electronics that are fixed to the module, but which are outside the coverage of the detector acceptance. These electronics use custom made radiation-hard chips that rely on a time-over-threshold (TOT) technique to determine the deposited charge in a strip. The signal is amplified and shaped before being compared with a threshold that depends on background conditions. The TOT method uses a scale that is approximately logarithmically related to the deposited charge. This allows a large range to be covered with a small number of bits.

To achieve optimal resolution for position and momentum measurements, precise knowledge of the alignment of the SVT is essential. This includes the alignment of the SVT modules with respect to each other (local alignment) and alignment of the whole SVT system with respect to the rest of the detector (global alignment).

2.4.3 SVT performance

The spatial resolution and efficiency of hits in SVT layers are calculated by comparing the hit position in the layer under study with the trajectory of the track as fitted using all other hits. The uncertainty on the track trajectory is subtracted from the width of the residual distribution to give the hit resolution. It is found, at all angles, to be better than 40 μm in each of the first three layers. This means that the B vertex resolution is well within design specifications. The SVT tracking efficiency as measured in data is 97%, excluding defective read out sections, which account for fewer than 5% of the total (a section is one side of a half module, there being 208 in total). Tracking efficiency calculations are described in Section 2.5.3.

The dE/dx resolution for minimum ionising particles is found to be 14%, which makes possible a 2σ separation of pions and kaons up to a momentum of 500 MeV/c and of kaons and protons up to 1 GeV/c.

The SVT is still within its radiation budget of 2 MRad. However, this budget did not take into account the impressive improvements in luminosity delivered by the PEP-II machine, and the corresponding increase in the radiation dose received by the SVT. Replacement modules for the SVT have been constructed, should they be required. However, replacing SVT modules is a particularly intrusive operation that will be avoided if possible.

2.5 The Drift Chamber (DCH)

2.5.1 DCH physics requirements

The main tracking system of the *BABAR* detector is the DCH. It is required to make precision measurements of particle momenta and angular positions for charged particles with momenta greater than 120 MeV/c and in the range $0.1 < p_T/(\,\mathrm{GeV}/c) < 5.0$.

To perform exclusive reconstruction of B and D decays, the momentum resolution (σ_{p_T}/p_T) must be less than 0.3% and the spatial hit resolution must be better than 140 μm. The DCH is also the main source of reconstruction information for K^0_S ($\rightarrow \pi^+\pi^-$) particles, which feature in many time dependent *CP* asymmetry studies (e.g. $B^0 \rightarrow J/\psi\, K^0_S$). This places the further requirement of measuring longitudinal position with a resolution better than 1 mm.

The DCH is also used for dE/dx information needed for particle identification – particularly at low momenta where the DRC is ineffective, as well as for areas that fall outside the DRC acceptance. This requires dE/dx measurements with a minimum resolution of around 7%.

The DCH must cover as large a solid angle as possible whilst keeping at a minimum the amount of material that particles must pass through. It must be able to operate under the harsh background conditions generated by the PEP-II machine.

Finally, the DCH passes information to the Level-1 Trigger every 269 ns.

2.5.2 DCH design

The DCH is a 2.8 m long cylindrical chamber with an inner radius of 23.6 cm and an outer radius of 80.9 cm. This chamber envelops a support tube, which contains the beam pipe and the SVT (see Figure 2.5). Its centre (in z) is offset from the IP in order to increase coverage in the forward direction, as illustrated in Figure 2.10.

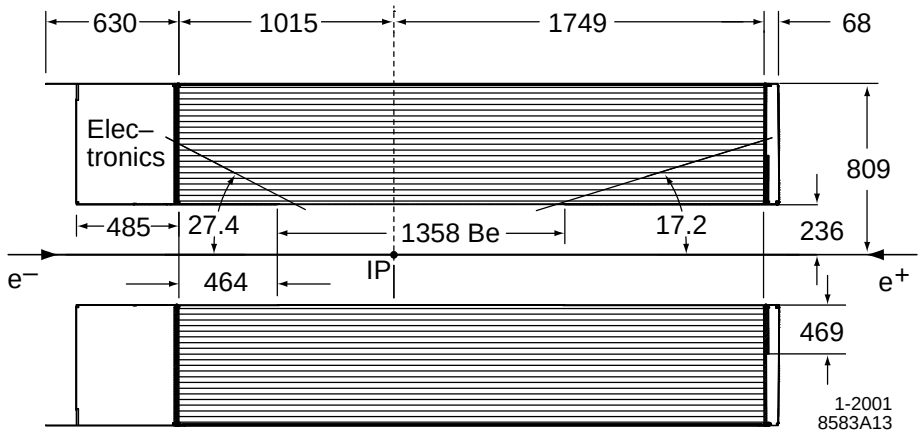

Figure 2.10: *Side on view of the DCH.*

The chamber is filled with a low mass gas mixture consisting of helium and isobutane in a ratio of 4:1. A small amount (0.3%) of water vapour is added to prolong the life of the chamber. This mixture provides good spatial and dE/dx resolution, a short drift time, and low multiple scattering. Field wires and sense wires extend between the two ends of the DCH. Together, these create hexagonal drift cells in the x-y plane, as illustrated in Figure 2.11.

There are 7,104 of these cells, each typically $1.2 \times 1.8\,\mathrm{cm}^2$ in size, arranged into 40 radial layers. The 40 layers are arranged as 10 4-layer *superlayers*. Each cell consists of a gold plated tungsten-rhenium sense wire $20\,\mu$m in diameter, surrounded by six gold plated aluminium field wires $120\,\mu$m or $80\,\mu$m in diameter, as shown in Figure 2.11. The sense wires are held at a high, positive voltage (design voltage: 1960 V); the field wires are grounded. This creates an electric field that has an almost circular symmetry near the sense wires, becoming more irregular near the field wires.

The wires in six out of the ten superlayers are orientated at a small angle to the z-axis. This allows positional information in the z direction to be extracted. The superlayers alternate between axial (A) and positive and negative stereo (U, V) from the innermost superlayer outwards according to the following pattern: AUVAU-

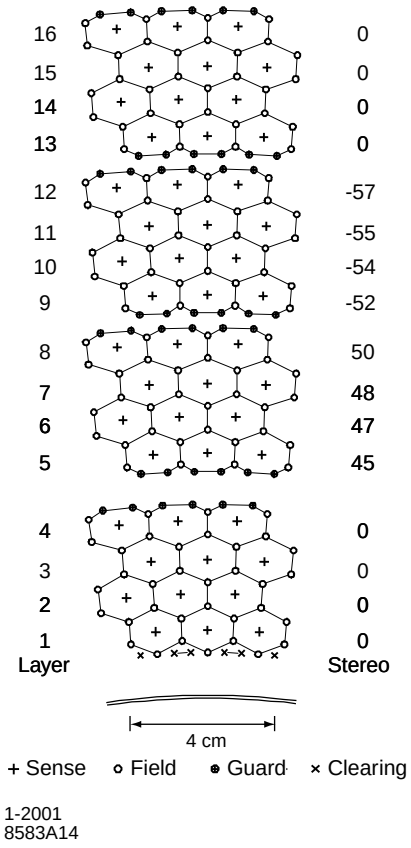

Figure 2.11: *DCH cell layout for the first four superlayers. The stereo angle of the layers in* mrad *is shown in the right hand column.*

VAUVA. The angle of each stereo layer increases from 45 mrad for the innermost stereo layer to 76 mrad for the outermost.

Apart from the IFR, the outermost component of the *BABAR* detector is a superconducting coil (see Figure 2.5) that generates a 1.5 T magnetic field in the z direction, $\vec{B}$. This field causes the curvature of the paths of charged particles in the x-y plane due to the force $q\vec{v} \times \vec{B}$, where q is the particle's charge and $\vec{v}$ is its velocity. This centripetal force can also be written as $(m\vec{v}^2/r)\hat{r}$ where m is the particle's mass and r is the radius of curvature; $\hat{r}$ represents the unit vector along r. Using momentum, $\vec{p} = m\vec{v}$, and rearranging slightly, this can be rewritten $p_T = qBr$, where now only magnitudes are shown, and momentum is the transverse momentum.

The main purpose of the DCH is to measure the curvature of the paths of charged tracks, which in turn provides a momentum measurement. This is achieved by fitting to a number of hits in the DCH, which combined make up a track. A hit occurs when a charged particle passing through a drift cell ionises the gas mixture. The liberated electrons are accelerated in the electric field towards the sense wire. These electrons further ionise the gas causing a charge avalanche (a gain of 5×10^4 for an operating voltage of 1960 V) that descends on the sense wire. The leading edge of the signal is used to determine an accurate drift time with a 1 ns resolution, and hence positional information. The total charge deposited is also determined for use in dE/dx measurements. The signals are digitised and corrected on a per channel basis for time offsets, pedestals and gain constants and read out by the DCH electronics.

2.5.3 DCH performance

High momentum two-prong events (e.g. di-muon events and bhabhas) are used to calibrate the drift time to track distance relation. The calibration is performed for each cell, the drift distance being estimated by calculating the distance of closest approach of the best fit to the track in question, where the fit is performed omitting the cell being calibrated. dE/dx measurements are also calibrated to remove biases from several sources including changes in gas pressure and temperature.

Track reconstruction efficiency, for both the SVT and the DCH, is calculated by reconstructing tracks in the SVT and DCH independently, and then extrapolating into the other subdetector. The efficiency is then, for the DCH, the number of tracks in both subdetectors as a fraction of the number of tracks found in the SVT.

Figure 2.12 shows this efficiency as a function of both transverse momentum and polar angle for both the design voltage of 1960 V and for 1900 V. A small section of the DCH was damaged during commissioning and so as a safety precaution an operating voltage of 1900 V was initially used. This has since been increased to

1930 V. The average tracking efficiency of the DCH at both design voltage and at 1930 V is $(96 \pm 1)\%$.

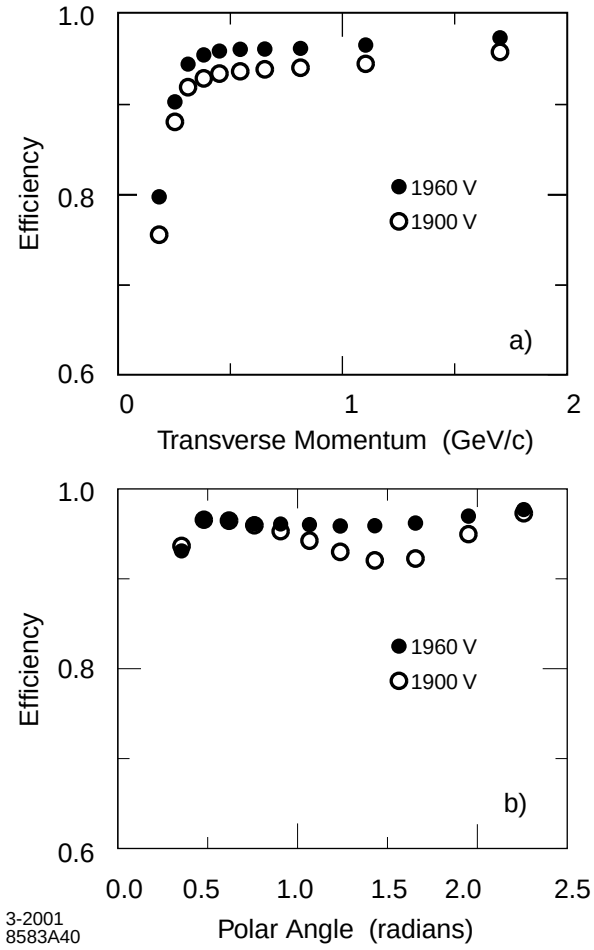

Figure 2.12: *Tracking efficiency for the DCH shown as a function of p_T (top) and of polar angle (bottom).*

Figure 2.13 illustrates the distribution of dE/dx measurements as a function of track momentum; the expected Bethe-Bloch [23] curve for the five stable charged particle types is superimposed. This shows that a good π/K separation can be achieved up to 0.6 GeV/c. The dE/dx resolution for e^+e^- events is 7.5%, which is near to the design value of 7.0%.

The transverse momentum resolution determined from cosmic ray muons is found

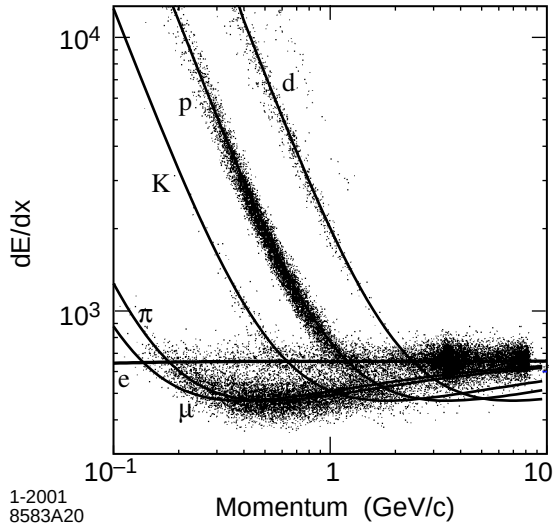

Figure 2.13: *dE/dx measurements in the DCH shown as a function of track momentum, with superimposed Bethe-Bloch predictions.*

to be well described by the function

$$\sigma_{p_T}/p_T = \frac{(0.13 \pm 0.01)\%}{\mathrm{GeV}/c} \cdot p_T + (0.45 \pm 0.03)\% \qquad (2.1)$$

where p_T is the transverse momentum in units of GeV/c. Furthermore this is in good agreement with Monte Carlo simulations and close to the design resolution.

The complete gas and wire system of the DCH is 0.28% of a radiation length for tracks with normal incidence. The DCH is well within its radiation budget of 20 kRad.

2.6 The Detector of Internally Reflected Cherenkov Radiation (DRC)

2.6.1 DRC physics requirements

The DRC is *BABAR*'s primary charged hadron particle identification system. It plays a vital role in distinguishing between pions and kaons. This is essential for B flavour

tagging in *CP* violation measurements, in which pions and kaons from the tagged event can have momenta up to $2\,\mathrm{GeV}/c$. π/K separation is also crucial for rare B decays of the type investigated in this thesis. For such decays, the final state pions and kaons have a wide range of momenta up to around $4.2\,\mathrm{GeV}/c$ (see Figure 2.6).

The DCH is only able to provide effective π/K separation up to $0.6\,\mathrm{GeV}/c$, as discussed in Section 2.5.3. The DRC is therefore required to provide acceptable π/K separation over the momentum range $0.6\,\mathrm{GeV}/c - 4.2\,\mathrm{GeV}/c$.

Further requirements placed on the DRC are that it must be small both in terms of radiation lengths – so as not to impact the resolution of the EMC – and physical size in the radial direction. The EMC is the most expensive component of the *BABAR* detector. As such, it is desirable for the EMC's inner (and therefore outer) radius to be as small as possible. This in turn is determined by the size of more internal subdetectors.

Finally, as with all *BABAR* subdetectors, in order to operate in the high luminosity environment of PEP-II the DRC must have a fast signal response time and be able to operate in high backgrounds.

Combining momentum measurements made by the DCH and velocity measurements made by the DRC allows determination of the charged particle's mass.

2.6.2 DRC design

Cherenkov radiation occurs when a charged particle travels through a medium faster than the speed of light in that medium. The angle at which the photon is emitted (the Cherenkov angle) depends on the particle's velocity. The DRC takes advantage of this phenomenon using a new and innovative design. This design relies on the principle that the magnitude of the Cherenkov angle is preserved when undergoing reflection from a plane surface.

Figure 2.14 illustrates a schematic of the DRC principle. 144 synthetic quartz bars

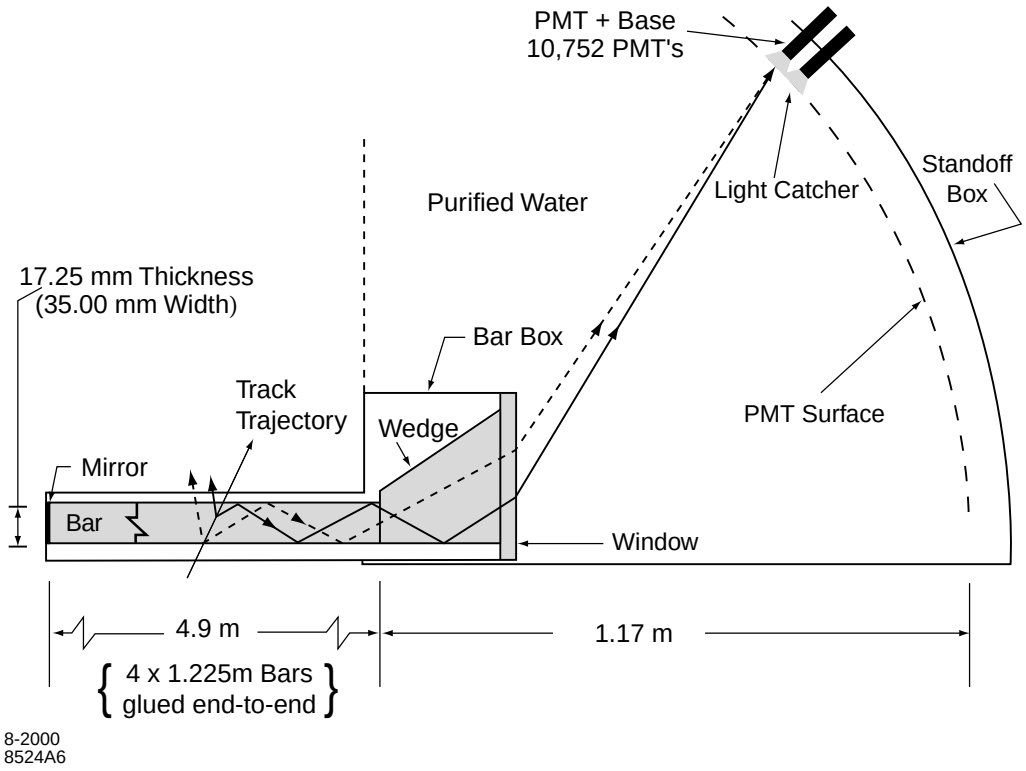

Figure 2.14: *Structure and concept of the DRC. A charged particle traversing the quartz bars with adequate velocity produces Cherenkov photons. Photons that are internally reflected within the quartz bars retain information about the initial Cherenkov angle. At the backward end the photons enter a standoff box where they are detected by an array of PMTs.*

are arranged into a 12-sided barrel. When a charged particle travels through a quartz bar, it emits Cherenkov radiation provided $\beta > 1/n$, where β is the speed of the charged particle as a fraction of the speed of light in a vacuum and $n = 1.473$ is the refractive index of the quartz. Photons are emitted at an angle $\theta_C = \cos^{-1}(\frac{1}{n\beta})$. Some of these photons are trapped by total internal reflection in the quartz bar. Those that travel forwards are reflected by a mirror at the front end, so that only the backward end of the DRC is instrumented. This allows room for the EMC endcap and reduces the background levels that the DRC instrumentation is subjected to.

On arriving at the backward end of the quartz bar, photons retain information about

their Cherenkov angle. Here, they enter the *standoff box*, a $6\,\mathrm{m}^3$ expansion region filled with purified water. Water is chosen as it has a refractive index close to that of quartz, minimising total internal reflection at the quartz-water junction. The standoff box is lined with an array of 10,752 Photo-Multiplier Tubes (PMTs), each 2.82 cm in diameter. Magnetic shielding is placed around the standoff box to reduce the effect of the magnetic field on the PMTs.

The measured Cherenkov angle has a certain number of ambiguities attached to it due to the numerous possible paths that could have been taken by a photon to give a signal in a particular PMT. These ambiguities are treated by reconstruction code. The arrival time of the signal is used to resolve many of these ambiguities, and also to suppress hits from beam background and other tracks in the event. The angle measured also depends on the angle at which the charged particle enters the DRC. SVT and DCH information is used to determine this angle.

The DRC contributes just 17% of a radiation length for tracks at normal incidence, and takes up only 8 cm of radial space. It is designed to have a coverage of 94% in the azimuthal angle, 83% in the polar angle.

2.6.3 DRC performance

The Cherenkov angle and time resolutions can be determined from di-muon events. The angular resolution of a single DRC photon is 10.2 mrad; the timing resolution is 1.7 ns. The Cherenkov angle resolution for a track is 2.5 mrad. This gives a π/K separation of just over 4σ at $3\,\mathrm{GeV}/c$. Figure 2.15 shows π/K separation as a function of momentum.

Figure 2.16 illustrates the discriminating power provided by the DRC. The peak in the $K\pi$ invariant mass spectrum corresponds to the decay $\overline{D}^0 \to K^+\pi^-$. It can be seen that using DRC information greatly reduces combinatorial background without having an adverse effect on the signal.

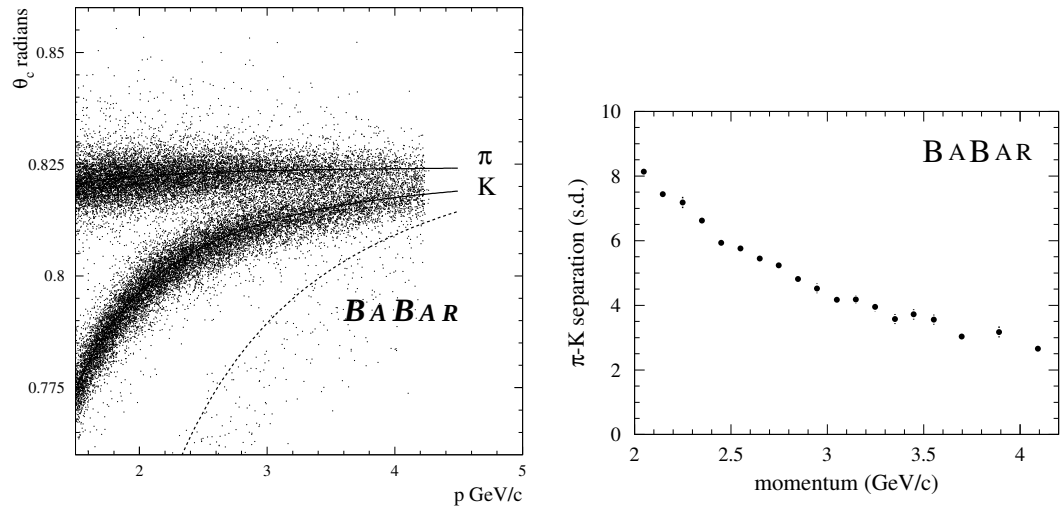

Figure 2.15: *Cherenkov angle and π/K separation in the DRC as functions of track momentum.*

2.7 The Electromagnetic Calorimeter (EMC)

2.7.1 EMC physics requirements

The EMC must provide efficient photon detection – with excellent energy and angular resolution – over a wide range of energies. For example, there are numerous important B decays that include neutral pions and η particles. These often decay into low energy photons (less than 200 MeV). On the other hand, photons produced in processes such as $e^+e^- \to e^+e^-\gamma$ and $e^+e^- \to \gamma\gamma$ – important for calibration and luminosity monitoring – are highly energetic (as high as 9 GeV in the laboratory frame). The EMC is also required to provide electron identification for B flavour tagging, and to make measurements of semi-leptonic decays. The EMC passes cluster and timing information to the Level-1 Trigger every 269 ns.

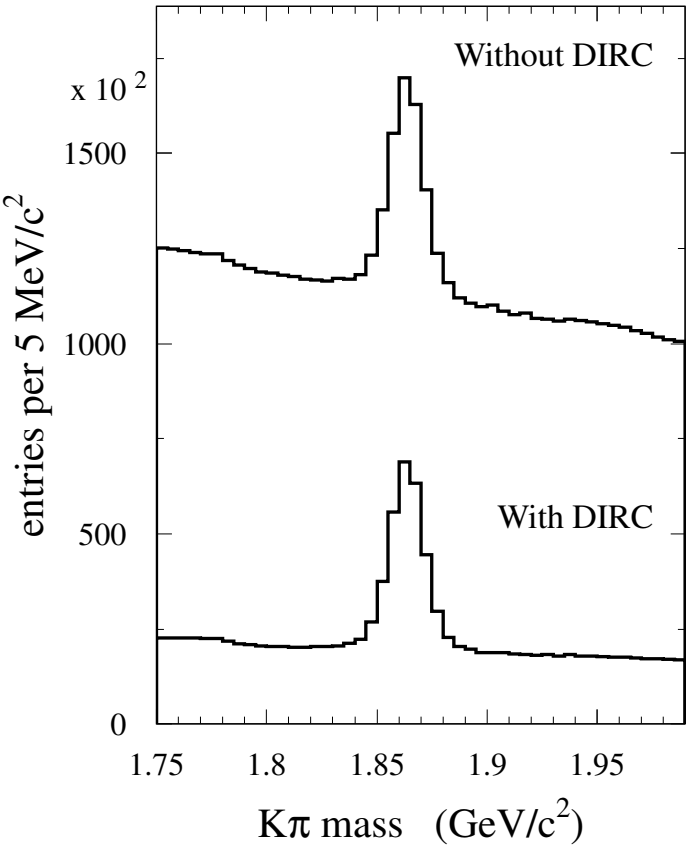

Figure 2.16: *Reconstructed $K\pi$ mass with and without the use of DRC (labelled DIRC in the plot) information for kaon identification. The peak corresponds to the decay $\overline{D}^0 \to K^+\pi^-$. The products of this decay can be identified using kinematics and so this process provides an excellent tool with which to test particle identification efficiencies and mis-identification rates.*

2.7.2 EMC design

The EMC consists of 6,580 thallium doped caesium iodide (CsI) crystals arranged in a barrel and a forward endcap. The crystals are arranged into 56 rings in θ: 8 rings in the endcap, 48 in the barrel. Each barrel ring contains 120 crystals in ϕ, as do the outer three rings of the endcap. The inner two endcap rings contain 80 crystals in ϕ, the remaining three consist of 100. In ϕ, all crystals are angled such that they point in the direction of the IP. In θ however, the crystals are aligned such that they point

in the general direction of the IP – but not in its *exact* direction, as this would allow particles to travel straight through the gaps between individual crystals. Crystals in the backward direction are 29.6 cm in length, corresponding to 16 radiation lengths. Those that are subjected to more energetic Lorentz boosted particles in the forward direction have a length of 32.4 cm (equivalent to 17.5 radiation lengths), the exception being the crystals in the innermost ring of the endcap, which are shorter by 1 radiation length due to space constraints. Inside this innermost ring (at smaller θ) lead shielding is installed to reduce the effect of machine background on the endcap crystals. The face size of the crystals is $\sim 5\,\mathrm{cm}^2$. Figure 2.17 illustrates the layout of the EMC.

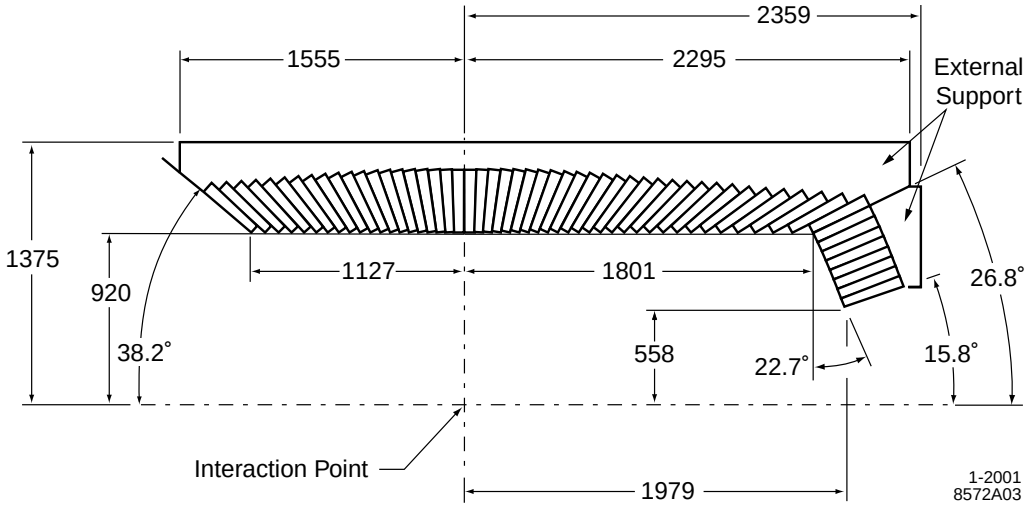

Figure 2.17: *Side on view of the EMC showing only the top half. The detector is rotationally symmetric about the z-axis.*

Particles entering the EMC are subjected to large electric fields due to the nuclei of the material from which the EMC is constructed. This causes incident particles to interact: electrons predominantly by bremsstrahlung, photons by pair production. Secondary electrons, positrons and photons also interact, as do further generations, until the energies of the electrons and positrons fall below a critical energy. At this point ionisation losses equal those from bremsstrahlung. The resulting electromagnetic shower is spread over numerous adjacent crystals, forming a *cluster*. In

addition to acting as an absorber, the EMC crystals also play the role of scintillator. Each crystal has two $1\,\mathrm{cm}^2$ silicon PIN diodes glued to the back face for photon detection.

Despite their financial cost, CsI crystals are used because of their high light yield, and their small Molière radius. These characteristics provide excellent energy and angular resolution respectively. The light yield is the number of photons emitted per unit of energy lost when an ionising particle passes through the medium. The Molière radius describes the transverse dimension of electromagnetic showers. Additionally, a small radiation length allows for the crystals to be small in physical size, desirable for such expensive material. The wavelength of the scintillated light makes it relatively easy to detect. Finally, and importantly, the signal response is fast. Table 2.3 shows some of the properties of CsI when used as a scintillator.

Parameter	***Value***	***Parameter***	***Value***
Radiation length	1.85 cm	Peak† λ	565 nm
Molière radius	3.8 cm	Signal decay time	940 ns
Light yield	$50,000\gamma/\,\mathrm{MeV}$	Radiation hardness	$10^3 - 10^4\,\mathrm{rad}$

†: The scintillator emits a spectrum of light, the peak of which lies at 565 nm.

Table 2.3: *CsI scintillator parameters.*

The EMC provides full coverage of the azimuthal angle and covers the polar angle range $15.8° - 141.8°$. This corresponds to a solid angle coverage of 90%.

2.7.3 EMC performance

The total energy in scintillation photons detected by the PIN diodes is not equal to the amount of energy deposited in the crystal. Each crystal must therefore be calibrated in order to determine the relationship between the physical energy deposited and the light yield detected. This is carried out by using photons or

electrons of known energies. A radioactive source (^{16}N) to which the crystals are exposed produces low energy (6.1 MeV) photons, whilst bhabha events – in which $e^{\pm}$ energies can be calculated from their polar angle – provide a tool for calibration at high energies ($\approx$ 7.5 GeV). In addition to individual crystal calibrations, it is also necessary to conduct cluster calibrations. This is because the energy deposited in one crystal is unlikely to account for the total energy of the incident particle. Corrections must be applied for crystal leakage, absorption in material in front of the EMC and between crystals, as well as energy in the crystals that is not associated with the incident particle. $\pi^0 \to \gamma\gamma$ processes are used for cluster calibrations.

The energy and angular resolutions of the EMC can be parameterised in the following way ("$\oplus$" represents adding the terms in quadrature):

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt[4]{E(\,\mathrm{GeV})}} \oplus b, \qquad \sigma_\theta = \sigma_\phi = \frac{c}{\sqrt{E(\,\mathrm{GeV})}} \oplus d \tag{2.2}$$

where, using the procedures described above, the parameters have been determined as $a = (2.3\pm0.4)\%$, $b = (1.9\pm0.1)\%$, $c = (3.9\pm0.1)$ mrad and $d = (0.00\pm0.04)$ mrad. This gives an angular resolution that is consistent with the design value, whilst a and b are slightly worse than design goals. The energy dependent terms arise mainly from fluctuations in photon statistics, and to a lesser extent from electronics noise. The constant term results from crystal leakage, absorption, non-uniformities and uncertainties in the calibration methods. Good angular resolution is important in, for example, the measurement of the mass of π^0 candidates for the decay $\pi^0 \to \gamma\gamma$.

For electron identification, EMC energy measurements are combined with DCH measurements of momentum. Electrons and photons typically deposit all of their energy in the EMC. For an electron, this energy should be approximately equal to its momentum, since its mass is very small in comparison. Muons on the other hand travel straight through the EMC depositing very little of their energy. Such particles are known as *minimum ionising particles* (MIPs). Hadrons such as kaons and pions can pass through the EMC as MIPs or can interact, depositing a fraction of their energy. Hadronic showers and electromagnetic showers have different longitudinal

and lateral shapes, allowing for further discrimination.

2.8 The Instrumented Flux Return (IFR)

2.8.1 IFR physics requirements

Muon identification is essential for precise *CP* measurements. Many of the most important *CP* modes contain a J/ψ, which is reconstructed in its decay to e^+e^- or $\mu^+\mu^-$. Muons are also used in the flavour tagging of the non-*CP* B meson through semi-leptonic decays. The IFR is required to identify muons, with good efficiency and high background rejection, with a minimum momentum of 1 GeV/c. It must also be able to detect neutral hadrons with high efficiency and good angular resolution. K_L^0 identification is especially important for the *CP* mode $B^0 \to J/\psi\, K_L^0$, which can be used to measure sin2β.

In addition to the physics requirements placed on the IFR, it also fulfills two important mechanical demands. It provides a support structure for the rest of the *BABAR* detector, and it acts as a flux return for the 1.5 T magnetic field.

The IFR passes hit information to the Level-1 Trigger every 269 ns.

2.8.2 IFR design

The IFR consists of alternating layers of iron and Resistive Plate Chambers (RPCs). The iron layers have two roles. They act as the flux return for the 1.5 T magnetic field. They also provide discrimination between muons, which, as MIPs, generally travel straight through the iron layers, and hadrons, which are more readily absorbed due to hadronic showers. The RPCs are the detecting components. Figure 2.18 shows the mechanical structure of the IFR. It consists of a hexagonal barrel and forward and backward endcaps. The barrel is segmented into 19 layers of RPCs with

the thickness of iron between each layer increasing from 2 cm for the innermost layer, to 10 cm for the outermost layer. This layout was determined by detailed Monte Carlo studies to optimise muon filtering and hadron absorption. The endcaps each have 18 layers of RPCs arranged in a similar way to the barrel. In addition to the barrel and endcaps, there are two cylindrical RPC layers between the EMC and the magnet. These are used to detect particles exiting the EMC and link tracks and/or EMC clusters to IFR clusters.

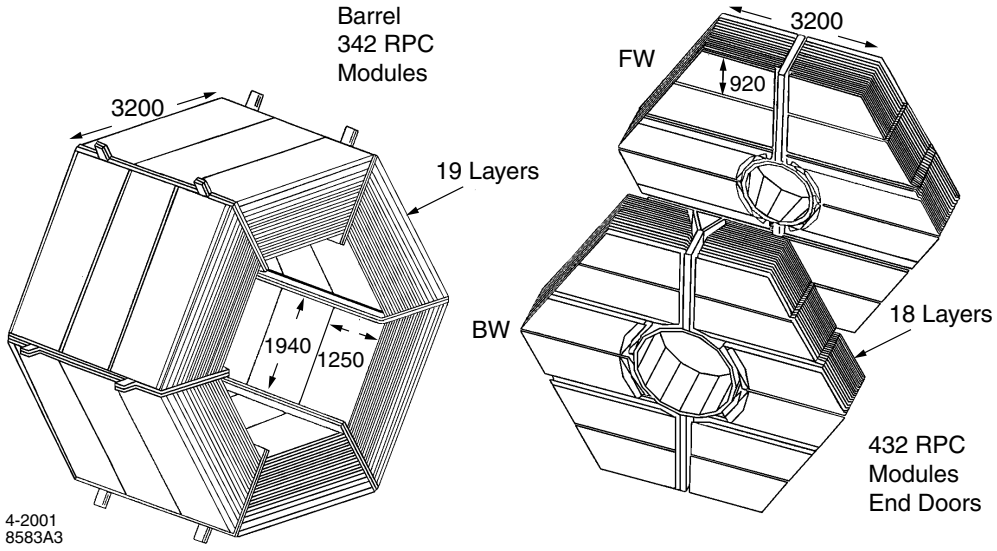

Figure 2.18: *IFR geometry, showing the hexagonal barrel (left) and forward (FW) and backward (BW) endcaps.*

A cross-section of an RPC is shown in Figure 2.19. The gas used is a non-flammable mixture of 57% argon, 39% freon-134a and 4% isobutane. A charged particle travelling through this mixture causes ionisation of the gas molecules, leaving a trail of ions and electrons. A high voltage is applied between the graphite coated bakelite electrodes. As in the gaseous mixture in the DCH, this causes an avalanche, in which ionised electrons cause further ionisation. In the DCH, the amount of ionised electrons collected at the sensor wires is proportional to the energy loss of the incident, ionising-causing particle. RPCs, on the other hand, operate in *streamer* mode. Here,

the avalanche effect becomes saturated causing vast ionisation that goes beyond the proportional mode.

The principle of the RPC is that it acts in *limited* streamer mode. That is, the gas is chosen so that the streamer is contained locally (it absorbs photons and electrons, preventing a lateral spread). High resistance electrodes are used for the applied voltage. The combined *quenching* effect of the gas and electrode material used prevents a complete breakdown of the medium where the resulting widespread plasma would prevent any further detection. The sudden movement of charged streamers onto the bakelite induces a charge on the capacitive read out strips. These are labelled "X Strips" and "Y Strips" in Figure 2.19, and are orthogonally arranged to give, together with the known radial position of the RPC, positional information in three dimensions.

The signals from the read out strips are passed to a front end readout card. Here, signals are compared to a threshold to determine whether the channel is active. Signals from active channels are digitised.

2.8.3 IFR performance

For data taken in 1999 and the first half of 2000, a mean muon efficiency of $\sim 90\%$ with a pion mis-identification rate of less than 8% was achieved for the momentum range $1.5 - 3.0\,\mathrm{GeV}/c$. These values were determined using clean control samples from the data. K_L^0 particles, and other neutral hadrons, are detected as IFR clusters with no associated charged track. The detection efficiency of these particles varies from 20% to 40% over the momentum range $1.0 - 4.0\,\mathrm{GeV}/c$.

Unfortunately, RPC efficiencies have deteriorated with time at an alarming and unforeseen rate. Reasons for this are not entirely clear. One likely contributor is the linseed oil with which the bakelite electrodes are treated to increase their longevity. It has been observed in some cases that the linseed oil has improperly

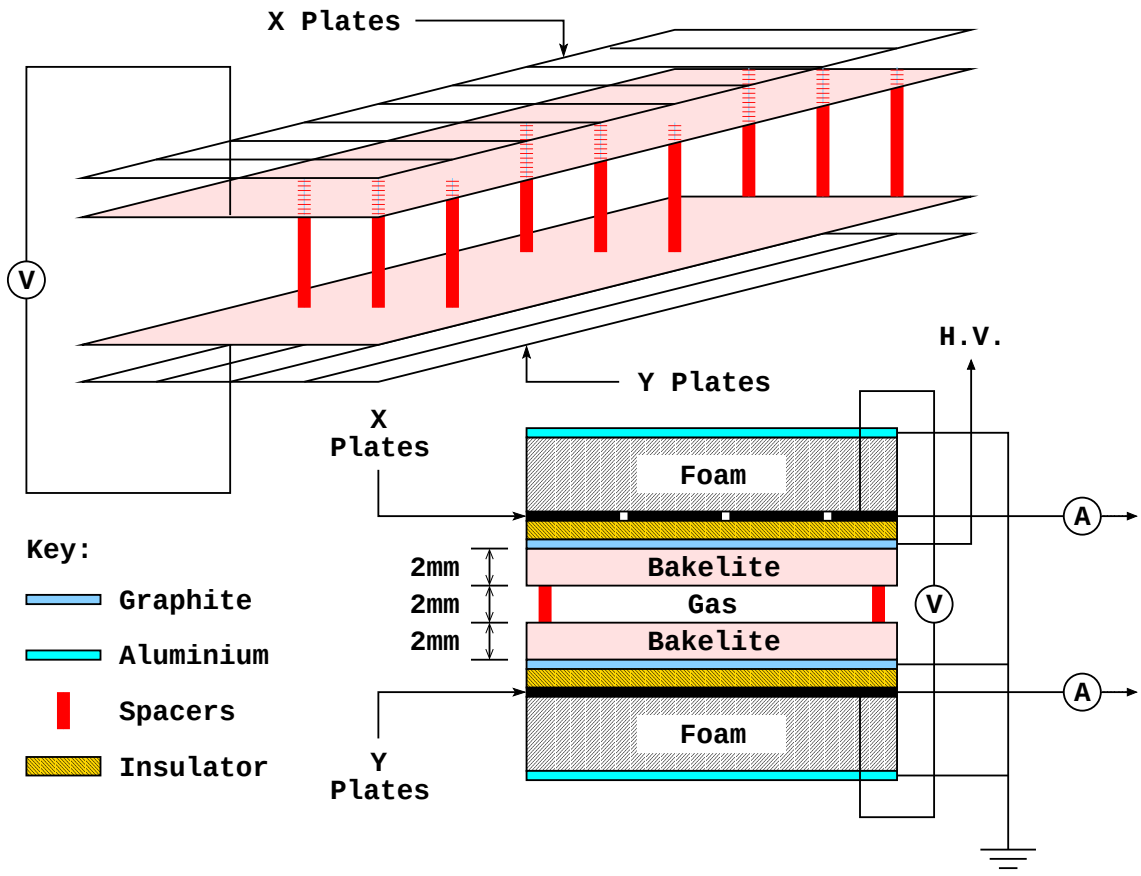

Figure 2.19: *RPC cross-section.*

cured, pooling into droplets and leaving other areas thinly covered. This can lead to field reduction to a level that is below streamer mode, or even outright shorting between plates. Despite numerous and lengthy efforts, remedies have not been found for this problem. As such, it was decided to replace all barrel RPCs with an alternative technology: Limited Streamer Tubes (LSTs). LSTs use a different gas, applied voltage and geometry, but the "limited streamer mode" physics is very similar. LSTs are described in more detail in [61]. In the summer shutdown of 2004, two of the six barrel sextants had their RPCs replaced by LSTs. In the summer shutdown of 2006, the remaining four barrel sextants had their RPCs replaced by LSTs. These sextants see significant improvements in detection efficiency.

3

The Trigger and data acquisition system

3.1 Introduction

The previous chapter describes the equipment used to produce B mesons and detect their decay products. In this chapter we discuss how the data representing particle interactions with the detector are recorded.

3.2 The Trigger (TRG)

3.2.1 Overview

The purpose of the *BABAR* trigger system is to accept interesting physics events whilst rejecting background events, thereby reducing the flow of data to a rate that is acceptable for processing and storage by the *BABAR* data acquisition system. The trigger efficiency[1] must be high, stable and well understood. *CP* violation studies require that the trigger efficiency for $B\overline{B}$ ($q\overline{q}$) events be greater than 99% (95%). Secondary physics goals require that the efficiency for $\tau^+\tau^-$ events be greater than 95%. Processes such as bhabha scattering – important for luminosity measurements – must have a trigger efficiency that is known to within 0.5%.

The trigger consists of a Level-1 hardware trigger (L1T), designed to retain nearly all physics events whilst rejecting background, and a Level-3 software trigger (L3T) that then selects the physics events of interest. The design accommodates the possibility of adding a Level-2 trigger, should it become necessary with increasing luminosities.

The L1T receives information from the detector every 269 ns – giving it an input rate of 3.72 MHz. There are approximately 30 bunch crossings in a typical 269 ns window, the majority of which do not result in an interesting physics event. The L1T is designed to accept events with an output rate of around ~ 2 kHz. The L3T reduces this further to around the ~ 200 Hz level. Compare this to the rates given in Table 2.1. The L1T rate has a direct impact on the level of deadtime in the data acquisition system. For an L1T rate less than approximately 3 kHz, deadtime rises steadily (approximately linearly) with L1T rate. For higher L1T rates, deadtime starts to rise significantly (up to 20% for high occupancy – i.e. 20% of interesting

[1]For the process $X \to Y$, the trigger efficiency is given by the number of $X \to Y$ events triggered divided by the product of the *collected* luminosity and the $X \to Y$ cross-section. The collected luminosity is given by the PEP-II luminosity multiplied by the difference between unity and the deadtime.

physics events are lost).

Further information regarding the *BABAR* trigger system can be found in [62].

3.2.2 Level-1 Trigger (L1T)

There is no event identification at Level-1; rather, trigger *lines* are defined to capture minimal event topologies.

The L1T consists of a DCH Trigger (DCT), an EMC Trigger (EMT), an IFR Trigger (IFT) and a Global Trigger (GLT), as illustrated in Figure 3.1. The DCT, EMT and IFT constantly receive data from their parent systems, which are used to produce trigger *primitives*. Combined, the trigger primitives provide a 269 ns snapshot summary of the detector in terms of position and energy or momentum. The primitives are then passed to the GLT where they are used to form trigger lines, which are indicators of certain topologies. If certain criteria are met, based on the trigger primitives, then a line is said to *fire*, or to be active. There are 24 L1T trigger lines in total, arranged to give priority to high multiplicity multi-hadronic events. If one or more line fires (this decision is made every 67 ns), at a time that corresponds to a bunch crossing, the trigger lines are passed from the GLT to the Fast Control and Timing System (FCT). Here, trigger lines can be masked and prescaled. Lines corresponding to processes that occur very frequently, and that are used more for calibration and efficiency studies rather than physics analyses (bhabhas, for example), are prescaled. If any lines remain active after prescaling/masking, a Level-1 Accept (L1A) is issued, causing all detector subsystems – including the trigger itself – to read out their event buffers to the data acquisition system. L1T lines and FCT prescales and masks are configurable on a run-by-run basis.

The IFT is used to collect events with muon particles, which are used for calibration and diagnostic purposes. It is not designed primarily to be efficient for B mesons. The IFT primitives simply correspond to single clusters or pairs of back to back

clusters with the further information of whether they are located in the barrel or endcap.

The *BABAR* L1T system is designed to be redundant. That is, the DCT and EMT each nearly satisfy the trigger requirements independently, each having an efficiency for $B\overline{B}$ events of over 99%. This allows their efficiencies to be studied in detail by using events triggered by the other system. Together they have an efficiency of over 99.9%.

Despite receiving signals (ticks) from the detector every 269 ns (Clock4), the L1T is able to pinpoint the start-time of a L1A to within 67 ns (output is Clock16). The better the time resolution of the L1A, the narrower the DAQ window, which means less data volume and deadtime associated with the DAQ data transfer for some of the subdetectors. For example, for the SVT, if there is a large time window before the real event, there is a large possibility of any background hits killing the interesting hit since single-hit electronics are employed. The DRC too has excellent time resolution and so a shorter time window eliminates much of the background photons detected by the PMTs making it easier to form ring segments out of the remaining hits and associate them with tracks in the DCH. The DCT and EMT receive Clock4 inputs from the DCH and EMC respectively. They then carry out algorithms that convert their Clock4 inputs into Clock8 (134 ns) outputs to be passed on to the GLT. The DCT achieves this by using information from previous ticks and look-up tables created using calibrations. The EMT uses zero-crossing interpolation. The GLT is then able to carry out a further (Clock8 $\rightarrow$ Clock16) conversion using the combined information from the EMT and the DCT.

3.2.3 The Drift Chamber Trigger (DCT)

The DCT consists of the Track Segment Finder (TSF), the Binary Link Tracker (BLT) and the z_0 p_T Discriminator (ZPD), which has recently replaced the p_T Discriminator (PTD). z_0 is the distance between the IP and the point on the z-axis to

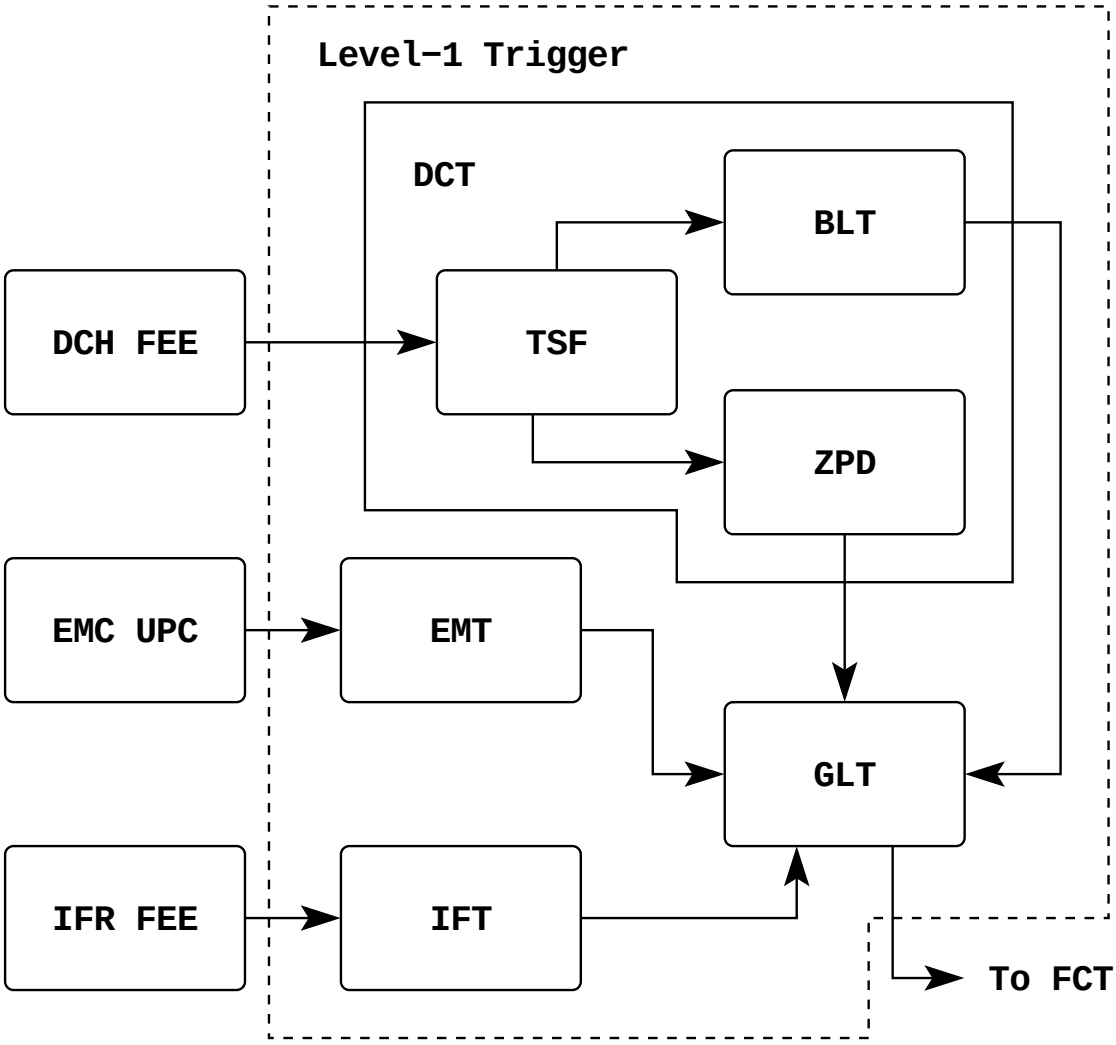

Figure 3.1: *The BABAR Level-1 Trigger.*

which the track can be extrapolated; p_T is the transverse momentum (the momentum in the x-y plane).

The DCH consists of 7,104 small hexagonal drift cells, organised into 40 layers, and grouped further into 10 radial superlayers (as discussed in Section 2.5.2). Each individual layer contains between 96 and 256 cells, which are grouped into 32 supercells per superlayer – 3-8 cells/supercell (the DCH has a 16-fold symmetry). The TSF looks for sets of adjacent DCH hits in a group of eight cells in a superlayer known as a *pivot group*. A 2-bit counter is associated with each cell in a pivot group. Depend-

ing on the path of the track, the resulting ions will take between one and four 269 ns ticks to reach each of the eight cells in the pivot group (assuming they are reached at all). This is reflected in the value of the two-bit counter. It is this time delay, or drift time, that the TSF uses in order to establish the accurate position (in ϕ) and time of the track segment. The 16-bit address (8 cells, each with a 2-bit counter) returned by a pivot group has corresponding spatial and temporal co-ordinates, which are obtained from a look-up table containing the 65,536 possible addresses for each pivot group. Each address also holds an uncertainty on the position and time of the track segment. The TSF receives and processes DCH Clock4 data continuously to determine which of the 32 supercells in each superlayer (320 supercells in total) contain track segments of interest. It reports the results to the BLT and ZPD in Clock8. The TSF consists of 24 TSF boards.

The BLT tries to link track segments to form tracks. If a track can be formed that reaches the outermost superlayer (A10) it is called a long track (type A), whilst one that reaches at least half way through the chamber (superlayer U5) is called a short track (type B). Output to the GLT is at Clock8 and is in the form of two 16-bit words corresponding to a type A track and a type B track. Each bit in a word represents the azimuthal supercell position of a linked track in either superlayer A10 (A tracks) or U5 (B tracks). The BLT consists of 1 BLT board.

When in operation, the PTD examined segments found in axial superlayers to determine whether they are consistent with being part of a track with p_T greater than a threshold value, typically 800 MeV/c. Many background events that passed during this PTD era were due to beam particle interactions with beamline components that are located 20 cm either side of the IP along the z-axis. This is illustrated in Figure 3.2. Information in z was not available using the 2D track reconstruction offered by the old-style DCT. These background events, therefore, could not be distinguished from those that derived from e^+e^- collisions.

With luminosities provided by PEP-II continuing to improve, it became necessary

for the L1T to accommodate the increasing data rates. The ZPD, which replaced the PTD permanently in the spring of 2005, having ran in parallel with it during testing in 2004, provides 3D tracking capabilities. The z_0 resolution of the ZPD is 4.1 cm. The majority of background tracks come from $|z_0| > 20$ cm. In addition to providing discrimination in z, the ZPD boasts more sophisticated algorithms than its comparatively crude predecessor. It performs algorithms that attempt to find and fit tracks originating from the IP that pass through *seed segments* (track segments located in axial superlayers A7 or A10). The ZPD consists of 8 ZPD boards. To accommodate the ZPD, TSF algorithms had to be modified. ZPD primitives use information on z_0, the error on z_0, the angle made by the track with the z-axis, λ, p_T and the number of segments for each track.

The primitive definitions for the DCT are listed in Table 3.1 along with those for the EMT.

For further information on the DCT, see [63] and [64].

3.2.4 The Electromagnetic Calorimeter Trigger (EMT)

Like the DCT, the EMT does not make a decision on whether to issue an L1A. It simply processes data from its parent subsystem, to be used by the GLT.

The EMC has 6,580 crystals in total, arranged in 56 rings in θ (8 endcap, 48 barrel), with between 80 and 120 crystals in ϕ in each ring (Section 2.7.2). The eight rings in the endcap are numbered $\theta_{crystal} = 1 \rightarrow 8$. The rings in the barrel are numbered $\theta_{crystal} = 9 \rightarrow 56$. For EMT purposes, the crystals are grouped into 280 *towers*. In the barrel, the towers all correspond to a 8×3 crystal array in $\theta \times \phi$. Towers in the barrel are numbered:

$$\begin{aligned} \theta_{tower} &= (\theta_{crystal} - 1)/8 \\ \phi_{tower} &= \phi_{crystal}/3 \end{aligned}$$

– using integer division, i.e. rounding down when the division is not exact. For the

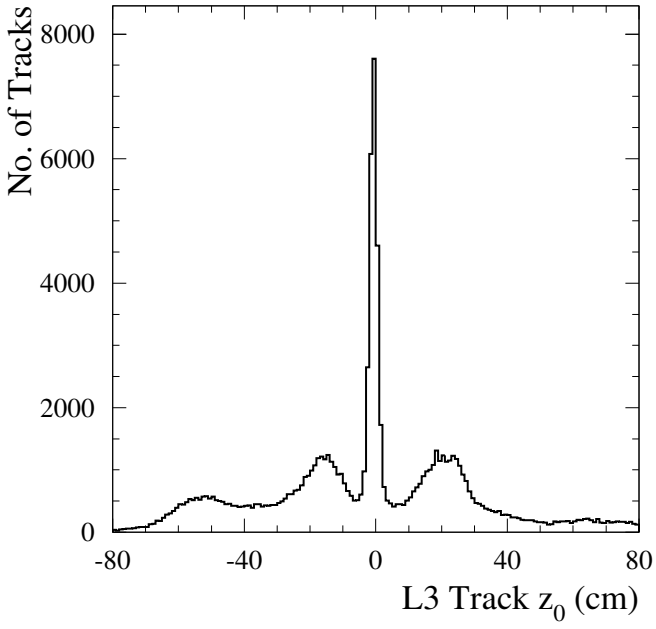

Figure 3.2: *Origin along the z-axis of DCH tracks that passed the old-style DCT (PTD era, pre-ZPD). The peak at $z_0 = 0$ is from e^+e^- collisions; the remaining tracks are primarily background tracks from beam particle collisions with beamline components.*

endcap, towers are numbered $\theta_{tower} = 0$ and:

$$\phi_{tower} = \begin{cases} \phi_{crystal}/2 & \theta = 1 \rightarrow 2 \\ 2\phi_{crystal}/5 & \theta = 3 \rightarrow 4 \\ 2(\phi_{crystal}+1)/5 & \theta = 5 \\ \phi_{crystal}/3 & \theta = 6 \rightarrow 8. \end{cases}$$

This gives us 7 towers in θ by 40 towers in ϕ.

The probability of multiple particles arriving close together in the same region of the calorimeter at the same time is low. It is also true that the energy deposited by a passing particle is not necessarily confined to one tower. The EMT uses these two facts to simplify its logic. Towers are clustered by summing the 40 towers in ϕ over the 7 towers in θ, yielding 40 ϕ *strips*. The EMT performs calculations using ten Trigger Processor Boards (TPBs). The TPBs receive energy samples from the 280

EMC towers. These are summed over the polar angle to form the 40 ϕ strips. The TPBs then perform algorithms to determine peak energy deposited and the time at which this peak energy occurs. Each TPB processes 4 ϕ strips independently, and reports the results directly to the GLT in the form of 4 (previously 5) trigger primitives: M, G, E (, X) and Y. These primitives indicate whether the peak energy exceeds a certain threshold, which corresponds to a physics process as described in Table 3.1. M, G and E have complete coverage in θ; X and Y provide discrimination in the polar angle. Specifically they single out MIP hits in the forward endcap ($\theta_{tower} = 0$) and high energy hits in the backward barrel ($\theta_{tower} = 5 \rightarrow 6$). Before information is shipped out to the GLT, neighbouring ϕ strips are OR'ed in pairs to give a 20-fold ϕ map for M, G and E (and previously X). Again, this is to ensure that the reconstructed energy accurately represents the physical deposit, since the deposit may occupy space in two ϕ strips. It also means that the amount of data to be shipped to and processed by the GLT is halved. For Y, groups of four ϕ strips are OR'ed to give a 10-fold map. The X primitive has recently been retired due to an upgrade of the DCT.

Further information regarding the EMT can be found in [65] and [66].

3.2.5 Level-3 Trigger (L3T)

The software based L3T uses all available information from the detector. It receives information every time an L1A is issued by the L1T. The Level-3 code runs on the Online Event Processing (OEP) computing farm, consisting of several hundred nodes. Events are distributed one per node at any given time.

The data received by the L3T contain superior positional information and increased energy and momentum resolution compared to the data received by the L1T. The L3T is therefore able to employ more sophisticated algorithms to event data. Timing information, as well as simple track finding and calorimeter clustering is used to accept or reject events. Those that are accepted are classified into various physics

Primitive	***Origin***	***Description*** (SL = superlayer)	ϕ ***Segmentation***	p_T, ***Energy Threshold***	z_0 ***Cut*** (cm)
B	DCT	Short track reaching SL U5	16	120 MeV/c	—
A	DCT	Long track reaching SL A10	16	180 MeV/c	—
Z	DCT	Track reaching SL A7	16	220 MeV/c	12
Z_t	DCT	Track reaching SL A7, tighter z_0 cut	8	220 MeV/c	10
Z′	DCT	High p_T track reaching SL A7	8	800 MeV/c	15
Z_k	DCT	Moderate p_T track reaching SL A7	4	350 MeV/c	10
M	EMT	All-θ MIP energy	20	120 MeV	—
G	EMT	All-θ intermediate energy	20	307 MeV	—
E	EMT	All-θ high energy	20	768 MeV	—
X (retired)	EMT	Forward endcap MIP energy	20	100 MeV	—
Y	EMT	Backward barrel high energy	10	922 MeV	—
U	IFT	Muon IFR sextant hit pattern	—	—	—

Table 3.1: *Definition of L1T primitives. The threshold values are configurable and those shown here are typical (preliminary) values for the BLT/EMT (ZPD). These primitives are used by the GLT to form trigger lines. An example of a trigger line is "*`3B&2A&2M`*" in which 3 or more B primitives plus 2 or more A primitives plus 2 or more M primitives are required for this line to be active. The U primitive is a 3-bit pattern allowing for seven combinations of muon topologies in the detector.*

categories as well as cosmic ray, bhabha and other types used for calibration and monitoring. The L3T forms trigger lines, which are logical combinations of the algorithm outputs, much as the GLT forms trigger lines from the Level-1 primitives. Level-3 output lines can be prescaled to reduce the rate at which certain processes are recorded, for example bhabha events, which are required for calibration and luminosity measurements, but not at the huge rate at which they occur.

To allow efficiencies to be calculated, some events that do not pass Level-3 are accepted regardless at a certain prescaled rate. These events are known as "Level-1

Pass-through" events.

All events that are accepted by Level-3 are logged to disk.

3.3 The data acquisition system (DAQ)

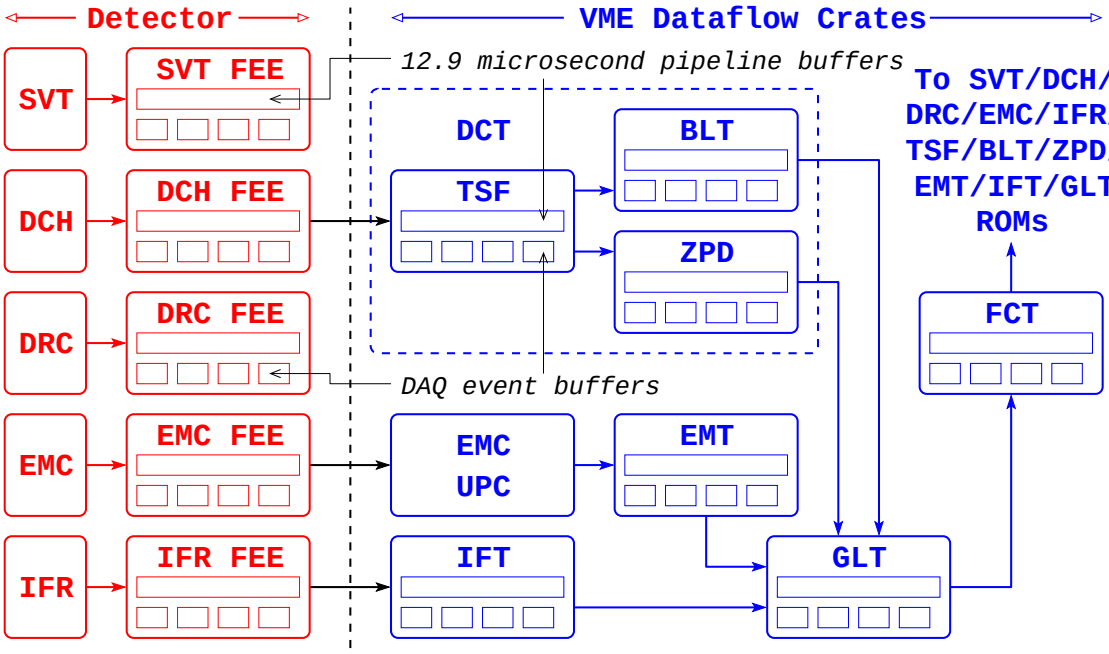

Figure 3.3: *The continuous flow of data from the BABAR detector.*

Figures 3.3 and 3.4 show schematics of the *BABAR* data acquisition system (DAQ).

The L1T operates in continuous sampling mode, generating trigger information at regular, fixed time intervals. The DCT and EMT receive data from the DCH and EMC respectively approximately $2\,\mu$s after the e^+e^- collision. The DCT and EMT event processing times are $4-5\,\mu$s, followed by another $\sim 3\,\mu$s of processing in the GLT.

First, this continuous flow of data from the *BABAR* detector, as illustrated by Figure 3.3, shall be examined. Second, as shown in Figure 3.4, the movement and processing of data when an L1A is issued shall be discussed. The time taken for an L1A to

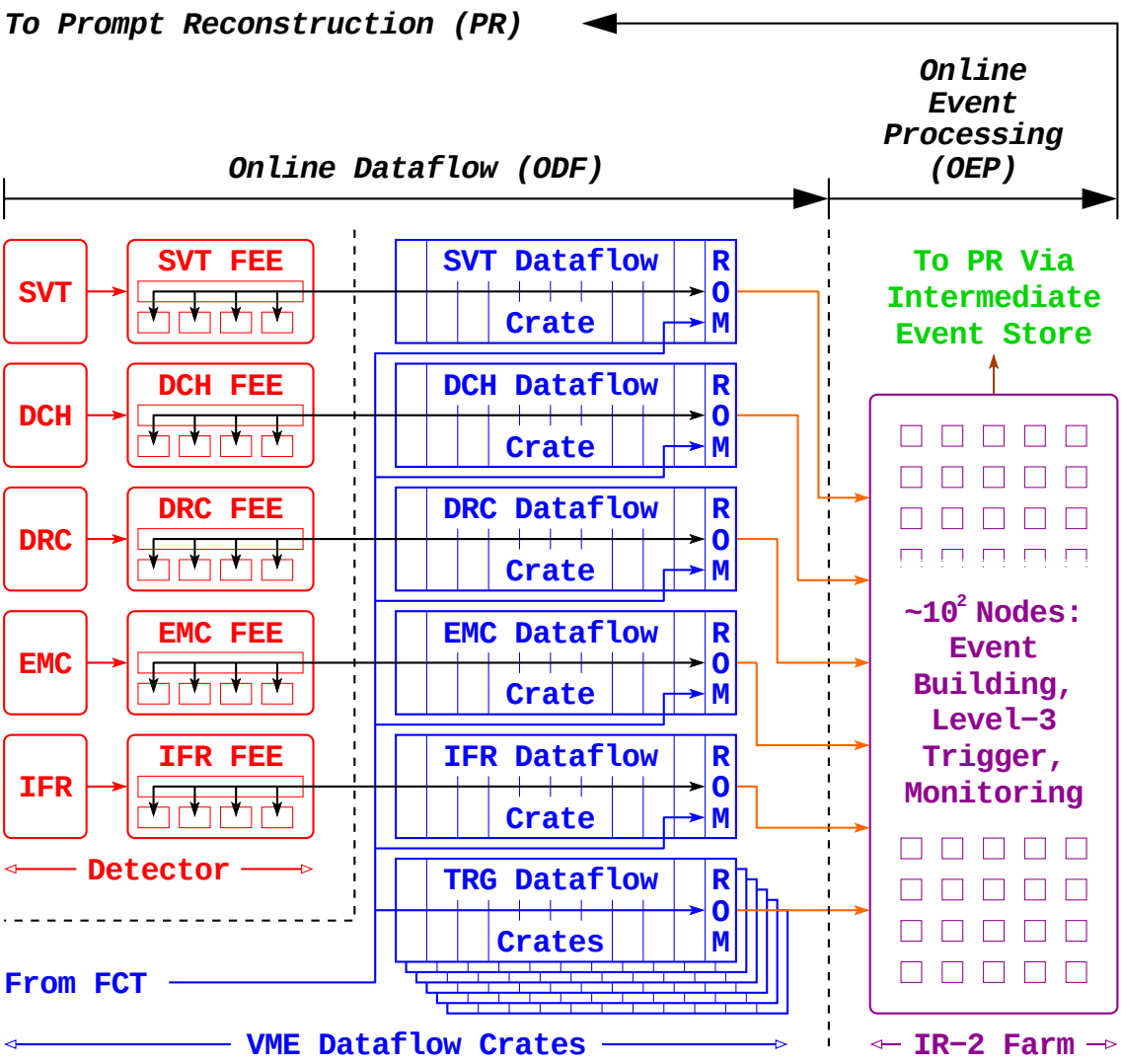

Figure 3.4: *The flow of data when an L1A is issued. Although not shown here, subsystem ROMs also talk to the conditions and configuration databases.*

propagate through the FCT and for subsystems to initiate event read out is $\sim 1\,\mu$s.

The steps described above – from the DCH and EMC shipping out their data, to an L1A being issued, to subsystems initiating event read out – are all accomplished within the $12.9\,\mu$s capacity of the subsystems' circular trigger latency buffers.

3.3.1 *BABAR* online: the continuous flow of data

Raw, analogue signals from each of the sub-detectors are processed by Front End Electronics (FEE), which lie within the detector. FEE processing is subsystem dependent, but in all cases the raw detector output is amplified and digitised. The digital signals are stored in 12.9 μs circular trigger latency buffers, located on the FEE.

Data from the DCH, EMC and IFR are shipped out of the detector, via optical fibre links, to be used by the *BABAR* L1T, as discussed in Section 3.2.2 and illustrated in Figure 3.3. These data are processed by the DCT, EMT (via EMC UPC boards) and IFT into trigger primitives, which are 269 ns snapshots of the detector, and are used by the GLT to form 24 trigger lines. These lines are passed to the FCT, which then uses the information to decide whether to issue an L1A.

Like the subdetectors' FEE, the subtrigger systems have 12.9 μs pipeline trigger latency buffers, which store the continuous flow of data that they receive.

3.3.2 *BABAR* online: the flow of data when an L1A is issued

Each subsystem is linked to a Read Out Module (ROM): the subdetectors by fibre optics between the FEE and the dataflow crates, the trigger subsystems by connections within the dataflow crates. Each of these ROMs contain *Triggered* Personality Cards (TPCs). This means that signals are only collected from the subsystem buffers on receipt of an L1A from the FCT. The only exception to this is the EMC. EMC ROMs contain *Untriggered* Personality Cards (UPCs). As such, signals are continuously received by the ROM (i.e. every 269 ns, rather than every time an L1A is issued), where they are processed and on receipt of an L1A are passed to a TPC ROM. As briefly mentioned in Section 3.3.1, these UPCs create the tower sums that are continuously sent to the EMT.

On receipt of an L1A, the subsystem ROMs collect data from the subsystem circular

trigger latency buffers. The L1T outputs at Clock16. That is, the exact time of an L1A is correct to within 67 ns. The amount of data extracted from the latency buffers is subsystem dependent. Naturally, the amount of data must correspond to an amount of time that is sufficient to be able to observe the whole event. A small amount of time before and after the event is also extracted to enable background studies and to allow for the uncertainty in the exact L1A start-time and other jitters. The data taken from the FEE is then subjected to algorithms, which perform *feature extraction* (FEX). Subsystem specific FEX attempts to isolate signals and suppress background and noise.

A master crate ensures that subsystems are synchronised with one another.

BABAR accumulates data on a run to run basis. Generally, the time taken for one run of data taking does not exceed one hour. On completion of one run, a new run is started. Runs are used to enhance the convenience of bookkeeping. Exact detector conditions, such as temperature, voltage, gas supply and humidity are known for each run. These conditions are extracted from the ROMs and stored in the conditions database (they can be modified using Online Detector Control, ODC). The output of each run is stored in an individual, offline file.

The ROMs, and other boards in the dataflow crates, are configurable for each run of data taking using Online Run Control (ORC). This is achieved by using the configuration database in which system-wide and subsystem specific configuration objects are stored. System-wide configurations include "physics", which is the normal mode of data taking for physics analyses, "cosmics" – in which the detector detects naturally occurring particles rather than particles created as a result of the PEP-II collider, and "calibration".

Both the conditions and configuration databases use Objectivity^{TM} [67] technology. They are used in Monte Carlo production to ensure that simulated events are generated using the same detector conditions and configurations as those for which real data were collected.

FEX processed data are passed to the Online Event Processing (OEP) platform (a computer farm consisting of several hundred nodes), where event building software is implemented before the L3T is employed. Online monitoring also takes place here. This is known as fast monitoring. *BABAR* collaborators continuously monitor the data being acquired in real time using the Java Analysis Studio (JAS) interface.

Events that pass the L3T are written to temporary files, which are then picked up by the Prompt Reconstruction (PR) farms.

Once offline, data are not only used for physics analyses, but also for thorough offline monitoring to ensure that the detector and DAQ are functioning correctly.

3.3.3 *BABAR* online: deadtime

Online Dataflow (ODF) and OEP, described in Section 3.3.2 (and illustrated in Figure 3.4), takes a finite amount of time from start to finish – despite the numerous and ingenious techniques aimed at minimising this time. Deadtime occurs when an L1A is issued, but the path is blocked by the movement of data triggered by a preceding L1A. The L1T fires at intervals of 67 ns (using 269 ns snapshots of the DCH and EMC). To counter deadtime each 12.9 μs circular trigger latency buffer (discussed in Section 3.3.1) is accompanied by four parallel DAQ event buffers. When an L1A is issued, but the ODF/OEP path is busy, the relevant data from the 12.9 μs buffer are stored in one of the four DAQ event buffers until the path is freed up, preventing it from being permanently overwritten when the 12.9 μs cycle has elapsed. Of course, deadtime still occurs when the four DAQ event buffers are full.

For normal beam backgrounds, *BABAR* deadtime generally occurs at the $2 - 5\%$ level.

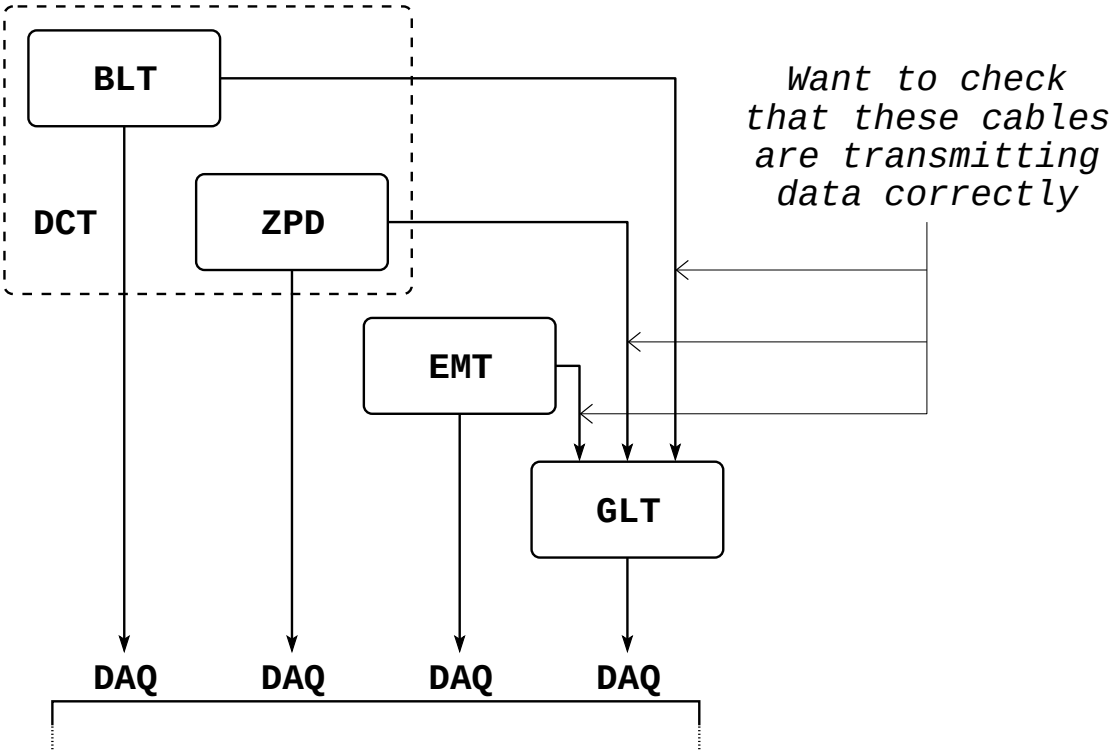

Figure 3.5: *The BLT, ZPD and EMT transmit information to the GLT. To check these transmissions, DAQ data from each of the subsystems are used.*

3.3.4 *BABAR* online: an example of online monitoring

Figure 3.5 illustrates that the BLT, ZPD and EMT pass data to the GLT, and that all 4 subsystems ship out data to DAQ on L1A. It is not possible to look directly at the data being transmitted between the EMT and the GLT, and the DCT (BLT and ZPD) and the GLT, whilst *BABAR* is online. To check whether the data received by the GLT are the same as that sent out by the EMT/DCT, DAQ data must be used. This involves bit manipulation software that formats the DAQ outputs before making tick by tick comparisons. A tick is the finite amount of time that elapses between successive, discrete, digitised signals read out by the detector subsystems. The period of one tick depends on the clock speed of the subsystem, and varies from 16.8 ns (Clock64) to 269 ns (Clock4).

Numerous other checks of cable transmission and of board algorithms are carried out throughout the DAQ system.

Physics events are also studied to ensure that what comes out of the detector is what one would expect.

4

BABAR software and data

4.1 Introduction

The previous chapter describes how the *BABAR* online system is deployed to select and record information from events produced by the PEP-II machine. This chapter describes how the data are further manipulated offline, facilitating physics analyses such as the one described in this thesis.

Large numbers of simulated *Monte Carlo* (MC) events are required for an interpretation of the data. Furthermore, for branching ratio analyses, the total number of B events must be established. These items are also discussed in this chapter.

Discriminating variables, devised to distinguish between signal and background

events, are also presented.

4.2 Reconstruction and data distribution

Data outputted by the *BABAR* online system are written to flat files called *extended tagged container* (XTC) files. These files are stored in a mass storage (HPSS) system shortly after production. XTC files contain raw detector hits for events that have passed the L3T. An average file contains $\mathcal{O}(10^5 - 10^6)$ events, corresponding to one online run. Figure 4.1 illustrates how XTC files are inputted into the *Prompt Reconstruction* (PR) system. PR consists of *Prompt Calibration* (PC) and *Event Reconstruction* (ER). PC is run over small subsets of the total number of events – enough to adequately perform data quality monitoring and calculate calibration constants – typically within a few hours of the event being logged to disk. ER is run over all events and is usually completed within a few days of the events being logged. Reconstruction routines are executed in which charged tracks and calorimeter clusters are reconstructed using the raw detector hits (as discussed in Sections 4.2.1 and 4.2.2). The PR code (`Elf`) runs on a large computing farm, on which numerous events can be processed in parallel.

PR output is in the form of *collections*, which are written to an *event store* as `ROOT` [68] files. Collections can be written to the event store in numerous formats from low level "raw" to high level "tag". Intermediate formats include "reco", "mini" and "micro". Data in the raw format provide raw hit information from the detector. Reco data contains track and cluster information. At this point there is no physics information about any possible particle candidates. This comes in the mini format where various hypotheses are applied to the reconstructed hits and hypothetical particle `C++` objects called `BtaCandidate`s are formed. The micro format is a higher level subset of the mini. Data in the tag format contains only tagbits in the form of booleans, integers and floating points. These tagbits are indications of various physics events.

The volume of data recorded by the *BABAR* online system is colossal (the event store database once boasting to be the largest in the world [69]). It is completely infeasible that a single analysis could make practical use of such an extensive dataset. Hence the need for a centrally managed processing system. The various levels of collection formats are used to accomplish acceptable compromises between processing time and level of detail.

A further technique to enhance the usability of datasets – and one that makes use of the collection formats discussed above – is *skimming*. Output from PR is for all events (collection name: `AllEvents`). The majority of analyses are interested only in a small fraction of `AllEvents`, and so processing all of them would be a waste of effort and resources. Skimming code, which makes use of tagbits and `BtaCandidate`s, applies various filters to `AllEvents` and outputs *skims*. Skimmed collections share the same formats as `AllEvents` collections. An example of a skim is `BCCC03a3body`, in which events that resemble inclusive B decays to final states with three "stable" charged particles have been selected. The filter algorithm that produces this skim is briefly discussed in Section 5.3.1.

Physics analysis code is based on a common *BABAR* framework that makes use of `BtaCandidate`s. Information from PR is combined not only to form particle candidates that are directly detected, but also, using vertexing, composite particle candidates that are inferred from their decay products. Particle identification (PID) selectors are employed at this stage. PID selectors are discussed in Section 4.2.3. Physics analysis code is generally run over skimmed collections in the micro format. This is the case for `NonCharm3BodyUser` [70], the package used in the analysis described in this thesis.

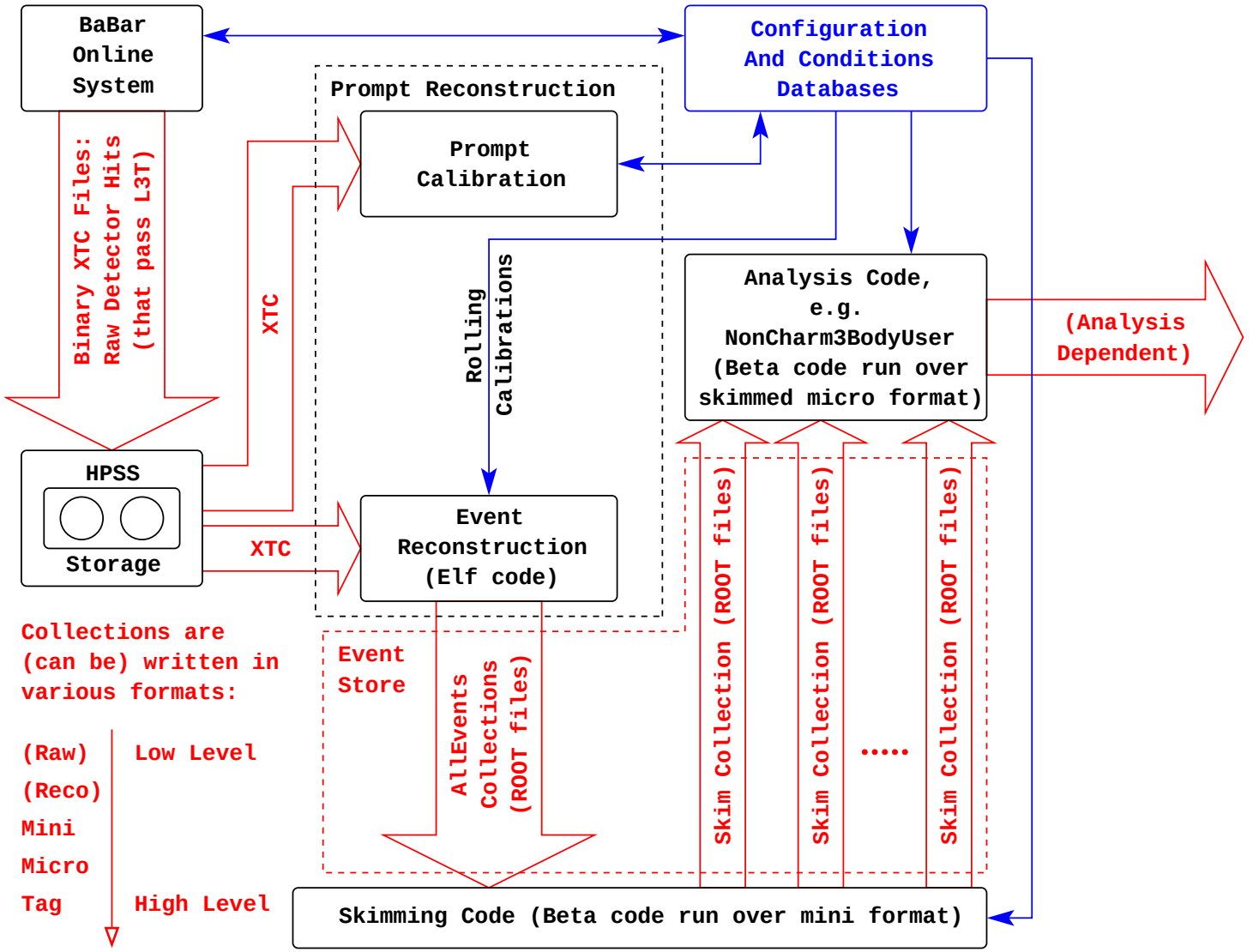

Figure 4.1: *The BABAR offline reconstruction and data distribution system.*

4.2.1 Track finding

Charged, helix-shaped tracks are described using five quantities:

- d_0 – the distance in the x-y plane from the origin to the point of closest approach (POCA) of the track trajectory (illustrated in Figure 4.2 (a)).
- ϕ_0 – the azimuthal angle as determined by the direction of the track at the POCA (illustrated in Figure 4.2 (a)).
- ω – equal to $1/r$, where r is the radius of curvature in the x-y plane (illustrated in Figure 4.2 (a)).
- z_0 – the distance in the z direction from the origin to the z co-ordinate of the POCA in the x-y plane (illustrated in Figure 4.2 (b)).
- $\tan\lambda$ – the tangent of the dip angle (equivalent to the cotangent of the polar angle, θ), which defines the slope of the track. If $\theta = 0$, the track is unaffected by the magnetic field in the z direction and continues in a straight line; if $\theta = 90°$, the particle's path is described by a flat circle (assuming no energy loss or Coulomb scattering). These are both special cases of a helix.

Each of these quantities can take on positive or negative values, as defined by the following equations:

$$\phi = \phi_0 + \omega\, L, \tag{4.1}$$

$$x = r\, \sin\phi - (r + d_0)\, \sin\phi_0, \tag{4.2}$$

$$y = -r\, \cos\phi + (r + d_0)\, \cos\phi_0, \tag{4.3}$$

$$z = z_0 + L\, \tan\lambda. \tag{4.4}$$

Eq. (4.1) gives the track angle, whilst Eq.s (4.2)-(4.4) give the track position; L is the path length along the track in the x-y plane. The sign of r depends on the direction in which the track proceeds in the x-y plane (clockwise or counter-clockwise). This in turn depends on the particle's charge.

For further details please see [71].

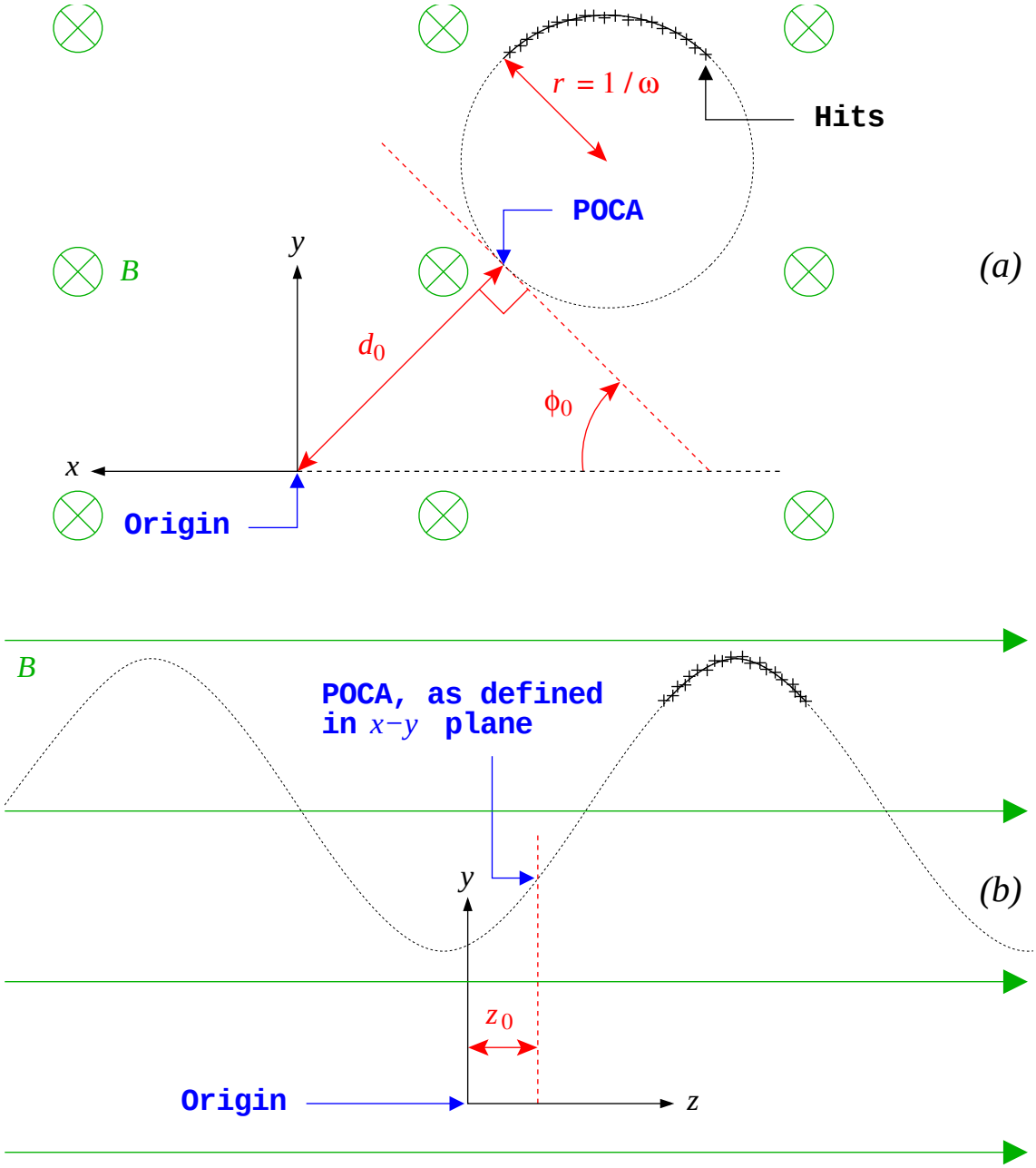

Figure 4.2: *Plot (a) illustrates the magnitudes of the track parameters d_0, ϕ_0 and ω; the x-y plane is shown with the positive z-axis and magnetic field perpendicular to and going into the page. Plot (b) illustrates the magnitude of the track parameter z_0; the positive x-axis is perpendicularly into the page.*

PR track reconstruction algorithms use, as their starting point, tracks found by the L3T. A Kalman fitting technique [72] is applied to the hits associated with these tracks. This technique accounts for detector material distribution and magnetic field variations. Additional DCH hits consistent with the tracks (but not used by the cruder, more time-conscious algorithms of the L3T) are added at this stage, and the Kalman fit is performed again. The resulting tracks are used to estimate a refined value of the collision time, t_0. Further track finding code is applied to remaining DCH hits to search for tracks that do not span the whole chamber, or that do not originate from the IP.

All tracks found in the DCH are extrapolated into the SVT and, where consistent, SVT hits are added in (again taking into account material and field variation in the detector). SVT hits that are not consistent with projected DCH tracks are subjected to a further track finding algorithm in order to identify SVT-only tracks (tracks that don't reach the DCH due to either low p_T or scattering by the support structure of the SVT).

The performance and efficiency of the track finding algorithms presented here are discussed in Section 2.5.3.

The tracks are placed into *lists* in the event store depending on the quality of the track. The track list used in the analysis described in this thesis is the `GoodTracksLoose` list. The requirements of this list are presented in Table 4.1.

p_T	$\geq 0.1\,\mathrm{GeV}/c$
$\lvert\vec{p}\rvert$	$\leq 10.0\,\mathrm{GeV}/c$
Number of hits in DCH	≥ 12
$\lvert d_0\rvert$	$< 1.5\,\mathrm{cm}$
$\lvert z_0\rvert$	$< 10.0\,\mathrm{cm}$

Table 4.1: *Requirements of the* `GoodTracksLoose` *list.* $|\vec{p}|$ *is the magnitude of the particle's momentum;* p_T *is its transverse momentum.*

4.2.2 EMC cluster finding

Clusters are identified by first looking for individual crystals that have detected an energy greater than 10 MeV. For crystals that pass this criterion, neighbouring crystals are examined. Those that detect an energy greater than 1 MeV, and those that neighbour a crystal with an energy greater than 3 MeV and already included in the cluster, are taken to be a component of the cluster and their neighbours in turn are examined. This process is iterated until all crystals outside of the cluster fail the energy threshold cut.

On establishing a cluster, a bump-finding algorithm is run over all contributing crystals. This algorithm searches for local maxima (in energy) within the cluster. This is necessary to determine whether a cluster is the result of one or more showers.

A further algorithm attempts to associate cluster bumps with reconstructed tracks. Where this is possible, the association is accounted for in subsequent reconstruction code. Where an association is not possible, the shower is assumed to derive from a neutral particle and is placed in a list analogous to the `GoodTracksLoose` list for charged tracks.

4.2.3 Particle identification (PID)

After track and cluster finding has been carried out, particle identification (PID) *selectors* are executed. PID selectors use information from a combination of *BABAR* subdetectors, potentially all of them. Each selector uses probability density functions (PDFs) to form a per-track likelihood for the particle type that it's designed to select. The five charged particle types that can give rise to tracks in the SVT and the DCH are, by increasing mass: electron, muon, pion, kaon and proton. PDFs and likelihoods are discussed in more detail in Section 5.6. The selectors operate in various modes that allow the user to strike the required balance between efficiency and purity.

Control samples are of paramount importance both in the design of PID selectors and in determining their performance. These samples are not selected using PID information. Instead, highly pure samples are selected using known physics processes and kinematics. For the selectors discussed here, the following channels are utilised:

- For pions: $K^0_S \to \pi^+\pi^-$ and $D^*(2010)^+ \to D^0(\to K^-\pi^+)\pi^+$. The lifetime of the K^0_S is such that there is an appreciable distance between the B vertex and the K^0_S vertex, making $K^0_S \to \pi^+\pi^-$ easy to identify.
- For kaons: $D^*(2010)^+ \to D^0(\to K^-\pi^+)\pi^+$, which can be identified using kinematical cuts only.
- For electrons: radiative bhabha events.

The analysis described in this thesis is concerned with identifying kaons and pions. To achieve this, a kaon selector `SMSKaonSelector` and an electron selector `LHSelector` are employed. Despite pions being responsible for the vast majority of tracks in multi-hadronic events, there is no specific pion selector. Instead, tracks that fail the kaon and electron selectors are generally assumed to be pions.

Tables 4.2 and 4.3 present an outline of the information used by the kaon and electron selectors.

For the analysis described in Chapter 5, `SMSKaonSelector` is employed in *Tight* mode: DRC information is used for tracks with momenta greater than 0.6 GeV/c; dE/dx information from the SVT and DCH is used for tracks with momenta less than 0.7 GeV/c. If no PID information exists for a track within the valid ranges (i.e. if the track has momentum greater than 0.7 GeV/c and lies outside the DRC acceptance), the candidate fails the selection. Likelihoods for kaon, pion and proton (l_K, l_π and l_p) are formed for each candidate. The candidate is accepted if

$$l_K > r_\pi l_\pi \qquad \text{and} \qquad l_K > l_p, \tag{4.5}$$

where

$$r_\pi = \begin{cases} 1 & \text{if momentum} < 0.5\,\text{GeV}/c, \\ 15 & \text{if } 0.5 < \text{momentum}/(\,\text{GeV}/c) < 0.7, \\ 1 & \text{if } 0.7 < \text{momentum}/(\,\text{GeV}/c) < 2.7, \\ 80 & \text{if momentum} > 2.7\,\text{GeV}/c. \end{cases} \tag{4.6}$$

These values of r_π represent the fact that π/K separation is superior for the momentum ranges $< 0.5\,\text{GeV}/c$ – using dE/dx in the SVT and DCH – and $0.7 - 2.7\,\text{GeV}/c$ – using the DRC. Between these two ranges, neither system is as effective. As momentum increases above $2.7\,\text{GeV}/c$, π/K separation deteriorates despite the angle resolution in the DRC improving. This is because the difference of the mean value of the expected angles for kaons and pions becomes smaller.

Likelihoods l_e, l_K, l_π and l_p are formed in a similar manner by `LHSelector`, where l_e is electron likelihood. The mode being used for the analysis described in this thesis is such that candidates pass the electron selector when

$$\frac{l_e}{l_e + 5.0 l_\pi + l_K + 0.1 l_p} > 0.95. \tag{4.7}$$

Figure 4.3 illustrates the efficiencies of the kaon and electron selectors for kaons and electrons respectively. The rates at which pions are mis-identified as kaons and electrons are also shown. Figure 4.4 demonstrates how PID eliminates vast amounts of combinatoric background, in which random tracks are combined to form signal candidates. The performance of π/K separation for the SVT, DCH and DRC are discussed in Sections 2.4.3, 2.5.3 and 2.6.3 respectively.

The PID selectors discussed here have been developed by and are maintained by the *BABAR* PID group. Further details can be found in [73, 74].

4.2.4 Vertexing

The B meson decay studied in this thesis is that with the 3-body final state $K^\pm K^\mp \pi^\pm$. Intermediate resonances (e.g. $K^*(892)^0$, $K_0^*(1430)^0$) are so short lived that their vertices are indistinguishable from that of the B. It follows that the three charged tracks

Information Used	***Sub-detector***	***Prob. Density Func.***		***Further Details***
		Shape	***Parameters From***	
dE/dx pull†	SVT, DCH	Gaussian	Control sample pull distribution	Expected value obtained from Bethe-Bloch functions (as illustrated in Figure 2.13)
Cherenkov angle (θ_C) pull	DRC	Gaussian	Control sample pull distribution	Expected value depends on particle type and momentum, $\cos\theta_C = (n\beta)^{-1}$ (Section 2.6.1)
Number of Cherenkov photons detected	DRC	Poissonian	Poissonian mean depends on particle type, momentum and polar angle (expected means as functions of these parameters stored in look-up table)	

Table 4.2: *Information used by the BABAR kaon selector,* `SMSKaonSelector`*. PDFs from these sources are assumed to be uncorrelated. The overall likelihood can therefore be formed by taking their product. Likelihoods for various particle hypotheses are formed. Tracks then pass or fail the selector depending on whether the ratios of likelihoods for various particle hypotheses pass cuts. The values of these cuts depend on the mode that the selector is running in:* NotAPion*,* VLoose*,* Loose*,* Tight *or* VeryTight*.*

†*: Pull is defined to be the difference between the measured value and the expected value, divided by the experimental error. Pulls are discussed in more detail in Section 5.10.1.*

should originate from the B decay vertex position, and their invariant mass should peak at the B mass.

Information Used	***Sub-detector***	***Prob. Density Func.***		***Further Details***
		Shape	***Parameters From***	
E/p	EMC	Gaussian with exponential tail	Control sample distribution	For electrons, $E/\|p\| \approx 1$; muons are MIPs, whilst hadrons generally deposit a fraction of their energy (Section 2.7.3)
EMC shower shape variables	EMC	Double Gaussian	Control sample distribution	For hadrons, correlations between variables need accounting for (can take as uncorrelated for electrons)
dE/dx	DCH	Gaussian	Control sample distribution	As kaon selector
Additional electron/hadron separation from θ_C for momentum $< 1.5\,\mathrm{GeV}/c$:				
Cherenkov angle, θ_C	DRC	Gaussian	Control sample distribution	For kaons and protons: as kaon selector
Cherenkov angle, θ_C	DRC	Double (Triple) Gaussian	Control sample distribution	For electrons (pions): to account for bremsstrahlung scattering (decay to muons/ electron emission in quartz)

Table 4.3: *Information used by the BABAR electron selector,* `LHSelector`*. Loose selections are first applied to separate muons. PDFs for the variables shown here are then created using control samples for the four remaining charged track types: electron, pion, kaon and proton.*

Geometric and kinematic fits are carried out to find the B decay vertex position. The geometric fit requires that the three tracks combined to form the B candidate originate from the same point in space; the kinematic fit requires that momentum

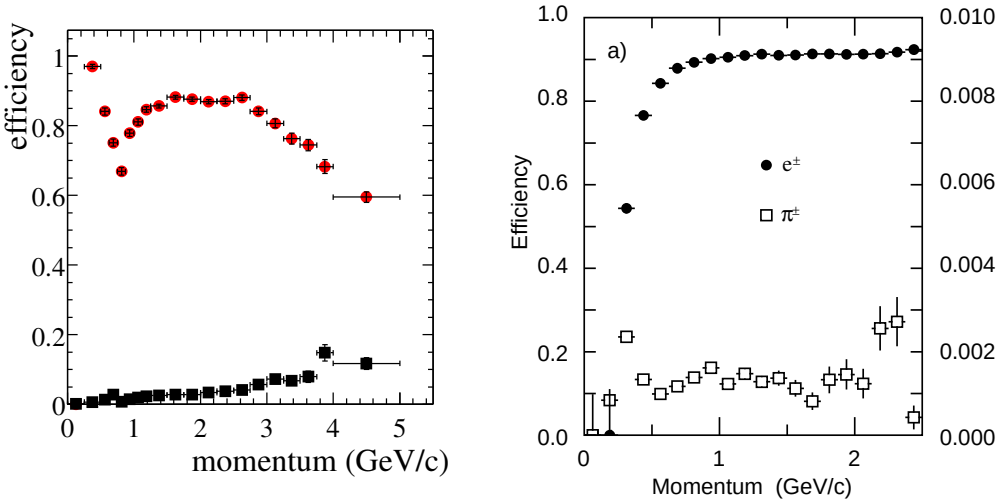

Figure 4.3: *Efficiencies of the kaon and electron selectors in* Tight *mode, as a function of momentum. Left plot:* `SMSKaonSelector` *efficiency on kaons (red circles) and pions (black squares). Right plot:* `LHSelector` *efficiency (left scale) on electrons and mis-identification probability (right scale) for pions.*

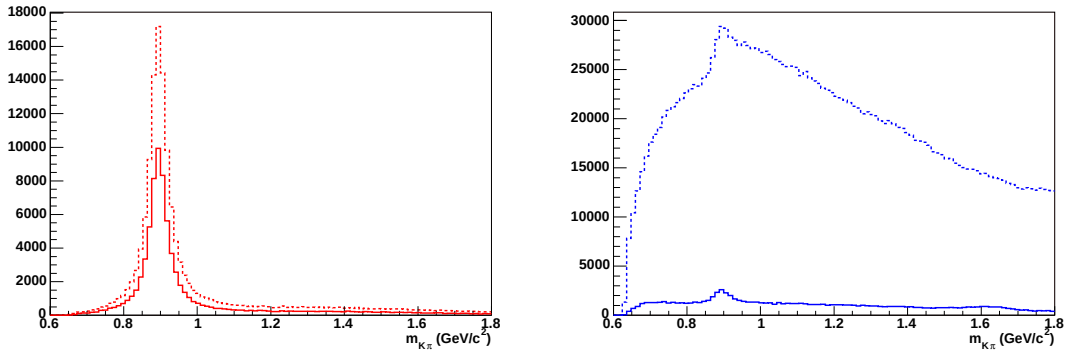

Figure 4.4: *Reconstructed* $K^*(892)^0$ *mass peak for MC data simulating the decay* $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ *(left) and for on-resonance data (right). Solid lines are for events that have been subjected to PID selection criteria; PID has not been applied to events represented by the dashed lines.*

is conserved at the vertex. `GeoKin`, the fitting routine, carries out an iterative procedure – aimed at minimising χ^2 – in which track positions and momenta are adjusted within their measured errors.

A second fit is performed, this time with the further constraint that the set of three tracks must have an invariant mass equal to that of the B mass. The advantage gained by applying this third constraint is that all reconstructed events lie within the kinematic boundaries as defined by the Dalitz plot. This avoids unwelcome complications in subsequent analyses – especially amplitude-level analyses. A further benefit, and one applicable to the intensity-level analysis described in this thesis, is the improvement of the mass resolution of intermediate resonances. An adverse effect is that background events are forced to appear more signal-like, making discrimination a little more difficult.

4.3 Monte Carlo (MC) simulation

Simulated MC data play a vital role in many areas of particle physics. Without them, interpretation of real data would not be possible. In the analysis described in this thesis, MC data are used to determine reconstruction and selection criteria efficiencies, to construct PDFs and to determine sources and amounts of background originating from B decays.

The first stage in producing MC data is event generation in which decays of B mesons and other particles and resonances are simulated. Each event is generated given an e^+e^- IP, and e^+ and e^- three momenta. These co-ordinate and three momentum values are randomly selected from distributions that reflect what occurs in the PEP-II machine. Output of this first stage is in the form of four-vectors $(E,\vec{p})$ and vertex positions. `EvtGen` [75] and `JETSET` [76] are the packages used for event generation.

The second step is detector simulation. The `GEANT4` [77] package is used to model the geometry of the *BABAR* detector and the interaction of its material with traversing particles. `GEANT4` output is in the form of *gHits*, each gHit representing an individual interaction of a detector subsystem with a passing particle.

The third stage simulates the response of the *BABAR* electronics to gHits. This includes the complete path of electronic signals from the FEE, through the dataflow crates, the trigger system, and – for events that are triggered – to DAQ.

The events generated by `EvtGen` and `JETSET`, and "detected" by `GEANT4` are interesting physics events. That is, events that involve the interaction of an electron from the HER beam and a positron from the LER beam (usually an e^+e^- annihilation event, but bhabhas are also useful). For the MC to accurately represent real data, it is necessary to account also for machine backgrounds (as discussed in Section 2.2.4). To do this, background events are overlaid with the simulated physics events. These background events are taken during acquisition of real data. The trigger system issues *cyclic* triggers at regular intervals (typically 1 Hz) at which point the digital signals stored in the subsystem buffers are read out to DAQ. It is improbable that a physics event is in progress during these randomly selected windows. As such the data accumulated using this approach provide a sample that accurately represents background conditions in the detector.

The simulated MC data are then subjected to the same reconstruction routines as the real data. The conditions and configuration databases, as discussed in Section 3.3, are used to ensure that MC provides as accurate a representation of real data as possible. The reconstructed MC collections are written to the event store as `ROOT` files. They share the same formats as those of the real data collections, the only difference being that MC contains *truth* information. Truth information allows the user to determine the particles – produced at the event generator level – responsible for each gHit. This is very useful for calculating mis-reconstruction efficiencies.

Control samples for MC and for real data are used to check whether MC accurately represents true data. For example, for $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$, the number of data events is inadequate for a comparison to be made with MC, so the control channel $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$ is used. $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$ is kinematically and topologically very similar to $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$, and the same

reconstruction tools can be employed for the two modes. However, the branching ratio of $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$ is large enough for a comparison to be made between MC and real data events. Another example of a control channel is $D^{*+} \to D^0(\to K^-\pi^+)\pi^+$, which is used to test both the performance of the kaon PID selector, and the agreement between MC and data. Where discrepancies are observed between MC and real data, control samples can also be used to apply corrections to the results obtained from MC.

4.4 B counting

All B meson branching ratios measured at *BABAR* must be normalised such that the B meson's total branching ratio is unity. To achieve this it is necessary to accurately determine the number of $B\overline{B}$ pairs ($N_{B\overline{B}}$) present in the dataset under investigation, and – for the analysis described in this thesis – the fraction that are charged (i.e. B^+B^-, rather than $B^0\overline{B}^0$). The B *counting* technique is employed to find $N_{B\overline{B}}$. This method allows a greater level of precision compared to using the cross-section of $e^+e^- \to \Upsilon(4S)$ and beam luminosity. The fractions of charged and neutral B mesons from $\Upsilon(4S)$ decays are determined by measuring the rates at which $B^+ \to J/\psi\, K^+$ and $B^0 \to J/\psi\, K^0_S$ occur [55].

B counting uses the number of multi-hadronic events[1], N^{On}_{MH} and N^{Off}_{MH}, recorded on- and off-resonance respectively (see Section 2.2.1). Since off-resonance data is acquired at energies beneath the $\Upsilon(4S)$ production threshold, and the branching ratio of $\Upsilon(4S) \to B\overline{B}$ is taken to be 100%, the weighted difference can be attributed entirely to $B\overline{B}$ production:

$$N_{B\overline{B}} = \frac{1}{\epsilon_{B\overline{B}}}\left(N^{On}_{MH} - N^{Off}_{MH}\kappa\frac{N^{On}_{\mu\mu}}{N^{Off}_{\mu\mu}}\right). \tag{4.8}$$

[1]Multi-hadronic events are those in which the e^+ and e^- annihilate, and from the resulting energy a quark pair is formed. Subsequent decays result in numerous hadrons being formed, which interact with the detector material.

$\epsilon_{B\overline{B}} = 0.96$ is the efficiency with which $B\overline{B}$ events pass the multi-hadronic selection cuts (determined from MC simulation); κ is a constant close to unity that accounts for the energy dependence of the continuum cross-section and selection efficiency. The ratio of the numbers of $\mu^+\mu^-$ pairs, $N_{\mu\mu}^{On}/N_{\mu\mu}^{Off}$, accounts for the different integrated luminosities of the on- and off-resonance samples.

The on-resonance data sample used in the analysis described in this thesis corresponds to

$$N_{B\overline{B}} = (231.58 \pm 0.08 \pm 2.55) \times 10^6. \tag{4.9}$$

The first error is statistical, the second is systematic. The main source of systematic uncertainty is the time variation in the number of multi-hadronic and muon pair events recorded off-resonance. The quantity $N_{MH}^{Off}/N_{\mu\mu}^{Off}$ is expected to be constant in time, but variations are observed for different off-resonance running periods. Another contribution is the uncertainty in the tracking efficiency, leading to variations in $\epsilon_{B\overline{B}}$. This is because the multi-hadronic filter uses inferior tracking (since it is applied before full reconstruction).

A further discussion on B counting, including error calculations, can be found in [78].

It is found that the ratio of charged and neutral B mesons from $\Upsilon(4S)$ decays, $R^{+/0} = \Gamma(\Upsilon(4S) \to B^+B^-)/\Gamma(\Upsilon(4S) \to B^0\overline{B}^0)$, is $1.006 \pm 0.036_{Stat} \pm 0.031_{Syst}$ [55]. Systematic uncertainties are those on the efficiencies (tracking, PID, K_S^0 reconstruction, selection criteria) and those attributed to the modelling and subtraction of background. The number of B^+B^- pairs present in the dataset under investigation, $N_{B^+B^-}$, is therefore given by:

$$\begin{aligned} N_{B^+B^-} &= \frac{R^{+/0}}{1 + R^{+/0}} N_{B\overline{B}} \\ &= (116.14 \pm 2.07 \pm 2.19) \times 10^6, \end{aligned} \tag{4.10}$$

and the number of charged B mesons, $N_{B^\pm}$, by:

$$\begin{aligned} N_{B^\pm} &= 2 \times N_{B^+B^-} \\ &= (232.28 \pm 4.14 \pm 4.39) \times 10^6. \end{aligned} \tag{4.11}$$

4.5 Discriminating variables

The analysis described in this thesis involves extracting a very small number of signal events ($10^0 - 10^2$) from vast backgrounds ($10^4 - 10^5$ events). In order to be able to achieve this, powerful discriminating variables are developed. These variables are used to *cut* – where signal and background distributions share different densities in different ranges, and to *fit* – when distributions differ in shape. The fitting techniques employed for the analysis presented in this thesis are discussed in Section 5.6.

4.5.1 Topological variables

The mass difference between the $\Upsilon(4S)$ and the $B\overline{B}$ pair is very small. As such the $B\overline{B}$ pair is produced almost at rest in the CM frame – resulting in a *spherical* distribution of its decay products, i.e. no preferred direction. This fact can be used to develop powerful discriminating variables against continuum events – by far the dominant source of background in the analysis described in this thesis. This is because, by contrast, continuum event decay products form highly collimated jets due to the large kinetic energy of the $q\overline{q}$ pair produced in the e^+e^- collision.

4.5.1.1 Thrust

The thrust axis of a collection of final state particles is the direction, in the CM frame, in which their combined longitudinal momentum is maximised. The thrust angle, θ_T, is defined here as the angle between the thrust axis of the reconstructed B candidate, and the thrust axis of particles constituting the rest of the event (ROE). The distribution of $\cos\theta_T$ is approximately flat for signal events, whilst peaking strongly at ± 1 for continuum events. This is illustrated in Figure 4.5.

4.5.1.2 The Fisher discriminant

The Fisher discriminant, $\mathcal{F}$, is a linear combination of discriminating variables, x_i, which individually may not provide a large amount of discrimination. Combined however, using optimised coefficients a_i, $\mathcal{F}$ provides appreciable levels of signal-background separation:

$$\mathcal{F} = \sum_i a_i x_i. \tag{4.12}$$

The Fisher discriminant used in the analysis described in this thesis uses the following five variables:

- The zero and second order Legendre polynomials:

$$L_0 = \sum_i^{N_{ROE}} p_i, \tag{4.13}$$

$$L_2 = \sum_i^{N_{ROE}} p_i \times \frac{1}{2}(3\cos^2\theta_i - 1). \tag{4.14}$$

  p_i and θ_i are the momentum and polar angle respectively of each track and neutral cluster in the ROE (of which there are N_{ROE} in total).

- θ_{Bmom}, the angle between the momentum of the reconstructed B candidate and the beam direction. For true B events, in which a spin 1 $\Upsilon(4S)$ decays to two spin 0 B mesons, the θ_{Bmom} distribution is proportional to $\sin^2\theta_{Bmom}$; it is approximately flat for continuum.

- θ_{Bthr}, the angle between the thrust axis of the reconstructed B candidate and the beam direction. For true B events the distribution should be uniform due to the isotropic nature of B decays, whilst for continuum it will show a distribution proportional to $1 + \cos^2\theta_{Bthr}$ consistent with spin $\frac{1}{2}$ particles coupling to a spin 1 boson. The quantities θ_{Bthr} and θ_{Bmom} rely on the principle of the conservation of angular momentum.

- `TFlv`, the output of the flavour tagging algorithm [79]. This algorithm is designed to look for processes specific to B decays. As such, many continuum

events will have failed or weak tags, whilst a larger proportion of correctly reconstructed B events will have strong tags.

Studies were carried out in which $\mathcal{F}$ was constructed using other sets of variables to determine whether $\mathcal{F}(L_0, L_2, \theta_{Bmom}, \theta_{Bthr}, \texttt{TFlv})$ was superior in terms of discriminatory power. For each $\mathcal{F}$, coefficients were calculated using signal-like and continuum-like data samples. Separation power was then tested using a second sample of each of these data types. For a given background rejection, $\mathcal{F}(L_0, L_2, \theta_{Bmom}, \theta_{Bthr}, \texttt{TFlv})$ was seen to give the greatest signal efficiency. This is illustrated in Figure 4.6 and is in agreement with other studies [80, 81]. Further details can be found in Appendix A. $\mathcal{F}(L_0, L_2, \theta_{Bmom}, \theta_{Bthr}, \texttt{TFlv})$ will simply be referred to as $\mathcal{F}$ throughout the remainder of this thesis.

The distribution of $\mathcal{F}$ is approximately Gaussian for both signal and continuum events – but with different means. This is illustrated in Figure 4.5.

For a further discussion of the Fisher discriminant method, see [82].

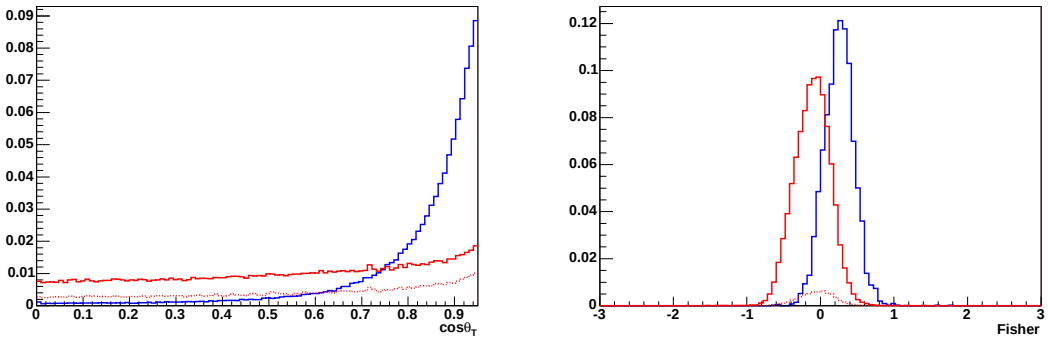

Figure 4.5: *$|\cos\theta_T|$ (left) and $\mathcal{F}$ (right) distributions for MC simulating the signal mode $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ (red, solid line) and off-resonance data consisting entirely of continuum background (blue, solid line). These distributions are normalised to unity. MC self-cross-feed distributions are also shown, normalised to the fraction of signal events that are self-cross-feed (red, dotted line) (self-cross-feed is discussed in Sections 5.5.2.4 and 5.7.1).*

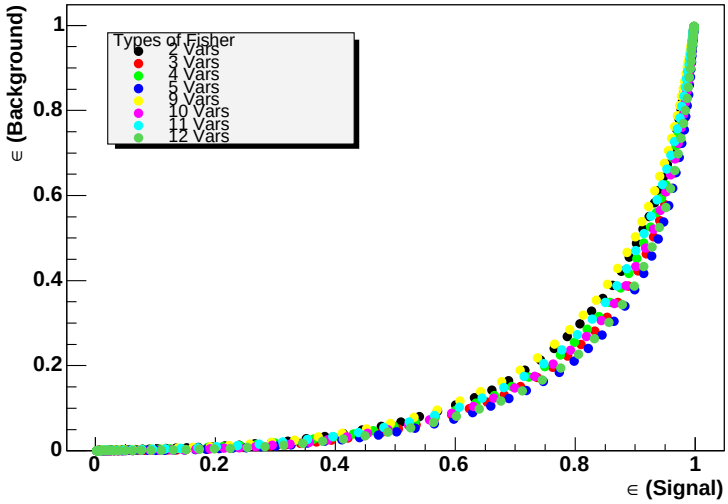

Figure 4.6: *Signal efficiency versus continuum background efficiency for various Fisher discriminants. It can be seen that the 5-variable fisher, $\mathcal{F}(L_0, L_2, \theta_{Bmom}, \theta_{Bthr}, \texttt{TFlv})$, provides the greatest discrimination (background efficiency is smallest for a given signal efficiency). A description of the variables used for each of these Fisher discriminants can be found in Appendix A.*

4.5.2 Resonance variables

The resonance variables $m^2_{K^-\pi^+}$, $m^2_{K^+K^-}$ and $\cos\theta_H$ are described in Section 1.4.1. For the remainder of this thesis, the invariant mass variables shall be written $m_{K\pi}$ and m_{KK} (+/−'s dropped). $m_{K\pi}$ and $\cos\theta_H$ distributions for reconstructed signal and continuum background events are shown in Figure 4.7.

4.5.3 Kinematic variables

When choosing suitable kinematic event variables, there are numerous issues to consider. These include:

- Optimising resolution.
- Avoiding correlations.

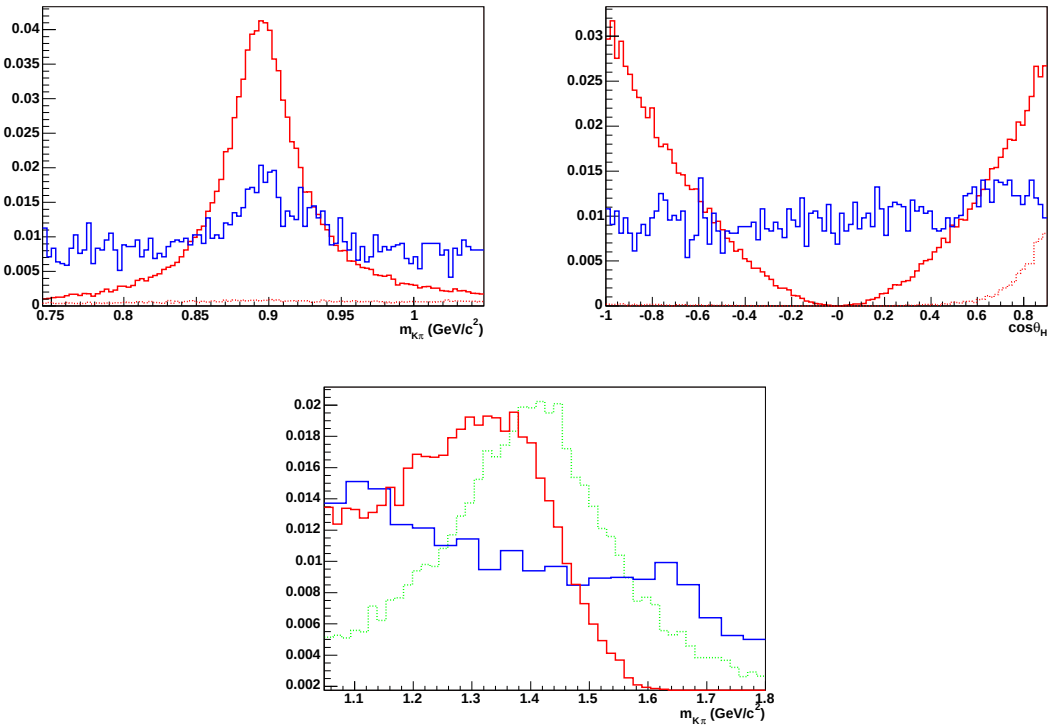

Figure 4.7: *$m_{K\pi}$ (top left) and $\cos\theta_H$ (top right) distributions for $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ MC (red, solid line) and off-resonance data (blue, solid line). The fraction of reconstructed $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ MC events that are self-cross-feed is depicted by the red, dotted line. The bottom plot shows $m_{K\pi}$ distributions for off-resonance data (blue, solid line) and for MC simulating the mode $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$. The red, solid line (green, dotted line) represents a MC sample in which $m_{K\pi}$ is modelled with a LASS (BW) lineshape.*

- Taking full advantage of the available information.
- The asymmetric nature of the PEP-II collider.
- Using variables that are free of non-essential information and which do not depend on parameters that may vary (for example the boost of the e^+e^- CM frame with respect to the laboratory frame), thus avoiding potential confusion and resolution degradation.

For pair production, $e^+e^- \rightarrow X\overline{X}$, the initial state has four momentum $(E_i, \vec{p}_i)$, where $E_i = E_{e^+} + E_{e^-}$ and $\vec{p}_i = \vec{p}_{e^+} + \vec{p}_{e^-}$. Let $\vec{p}_f$ be the resultant momentum for the group of final state particles originating from X, and E_X ($E_{\overline{X}}$) be X's ($\overline{X}$'s) energy; m_X ($m_{\overline{X}}$) is the rest mass of X ($\overline{X}$) and $m_X = m_{\overline{X}}$. All quantities defined above in this section are in the laboratory frame.

Conserving energy and momentum, and using the energy-momentum relation (with $c = 1$), we have

$$\begin{aligned} E_i &= E_X + E_{\overline{X}}, && (4.15)\\ \vec{p}_{\overline{X}} &= \vec{p}_i - \vec{p}_f, && (4.16)\\ E_X &= \sqrt{m_X^2 + \vec{p}_f^{\,2}}, && (4.17)\\ E_{\overline{X}} &= \sqrt{m_X^2 + (\vec{p}_i - \vec{p}_f)^2}. && (4.18) \end{aligned}$$

Substituting Eq. (4.18) into Eq. (4.15) and rearranging,

$$\begin{aligned} E_X &= \frac{E_i^2 - \vec{p}_i^{\,2} + 2\,(\vec{p}_i \cdot \vec{p}_f)}{2\,E_i} && (4.19)\\ &= \frac{s + 2\,(\vec{p}_i \cdot \vec{p}_f)}{2\,E_i}, && (4.20) \end{aligned}$$

where

$$s = E_i^2 - \vec{p}_i^{\,2}. \qquad (4.21)$$

For a symmetric collider, $\vec{p}_i = 0$, and Eq. (4.19) reduces to

$$\begin{aligned} E_X &= \frac{E_i}{2} \\ &= \frac{\sqrt{s}}{2}, && (4.22) \end{aligned}$$

i.e. the energy of one of the beams.

For an asymmetric collider, let us boost into the e^+e^- CM frame:

$$E_X \longrightarrow E_X' = \frac{1}{\sqrt{1 - \vec{\beta}^{\,2}}} \left(E_X - \vec{\beta} \cdot \vec{p}_f \right). \qquad (4.23)$$

$|\vec{\beta}|$ is the speed of the e^+e^- CM frame with respect to the laboratory frame as a fraction of the speed of light in a vacuum, *c*. $\vec{\beta}$ can be written as

$$\vec{\beta} = \frac{\vec{p}_i}{E_i}. \tag{4.24}$$

Substituting Eq. (4.24) into Eq. (4.23),

$$\begin{aligned} E_X' &= \frac{1}{\sqrt{1-(\vec{p}_i/E_i)^2}}\left(E_X - \frac{\vec{p}_i}{E_i}\cdot\vec{p}_f\right) \\ &= \frac{E_i}{\sqrt{E_i^2-\vec{p}_i^{\,2}}}\,\frac{E_i^2-\vec{p}_i^{\,2}+2\,(\vec{p}_i\cdot\vec{p}_f)}{2\,E_i} - \frac{E_i}{\sqrt{E_i^2-\vec{p}_i^{\,2}}}\,\frac{\vec{p}_i\cdot\vec{p}_f}{E_i} \\ &= \frac{E_i^2-\vec{p}_i^{\,2}+2\,(\vec{p}_i\cdot\vec{p}_f)}{2\sqrt{E_i^2-\vec{p}_i^{\,2}}} - \frac{2\,(\vec{p}_i\cdot\vec{p}_f)}{2\sqrt{E_i^2-\vec{p}_i^{\,2}}} \\ &= \frac{\sqrt{E_i^2-\vec{p}_i^{\,2}}}{2} \\ &= \frac{\sqrt{s}}{2}. \end{aligned} \tag{4.25}$$

This is the same result as that presented in Eq. (4.22), as expected.

The variable E_X describes the expected energy of the reconstructed X candidate, in the laboratory frame, using the beam parameters E_i and $\vec{p}_i$ and the three momenta of the detected decay products. It has been shown to be a valid quantity for asymmetric collisions (Eq. (4.25)). We are now able to use E_X to introduce two kinematic event variables for our $X = B$ system[2]: m_{ES} and ΔE. These two variables make maximum use of the available information whilst being largely uncorrelated.

ΔE is the frame-independent difference between the expected energy of the reconstructed B candidate, E_X, and the energy of the candidate as obtained using decay product masses as measured by the detector's PID systems in the laboratory frame, E_f:

$$\Delta E = E_f - E_X. \tag{4.26}$$

m_{ES} is the *beam-energy substituted mass* and is given by

$$m_{ES} = \sqrt{E_X^2 - \vec{p}_f^{\,2}}. \tag{4.27}$$

[2]For a collider operating just above the $B\overline{B}$ threshold, such as PEP-II, B's result only from pair production.

It can be seen that m_{ES} is evaluated using the reconstructed candidates' three (rather than four – as with ΔE) momenta, avoiding the need to assign mass hypotheses to the B daughters. Instead, the beam energy is used. This is allowed because, operating at the $\Upsilon(4S)$ energy, the four momenta of the B mesons are very well constrained by the beam energy. The approach is advantageous because the beam energy is known to a greater precision compared to the combined effect of the detector mass resolutions for each of the B decay products.

ΔE, on the other hand, does use reconstructed masses of the B candidate decay products (as such, its resolution is considerably larger, typically – for $B^+ \to K^+K^-\pi^+$ – $\sim 20\,\text{MeV}$ versus $\sim 2.5\,\text{MeV}$ for m_{ES}). This is useful when mass hypotheses aid discrimination (see Section 5.5.2.1).

For a correctly reconstructed B meson, m_{ES} is expected to peak at the B mass, $5.279\,\text{GeV}/c^2$ [23]. ΔE should peak at zero. For continuum background events, the distributions are slowly varying over the allowed kinematic ranges. Figure 4.8 illustrates the distributions in m_{ES} and ΔE of reconstructed events for signal and continuum background samples.

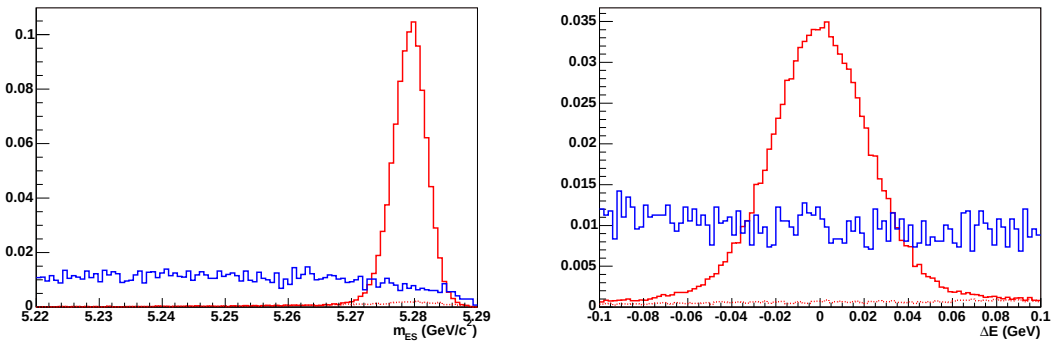

Figure 4.8: *m_{ES} (left) and ΔE (right) distributions for reconstructed signal and continuum background samples. The red, solid line represents $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ MC; the blue, solid line is for off-resonance data. The fraction of reconstructed $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ MC events that are self-cross-feed is depicted by the red, dotted line.*

For further information regarding m_{ES} and ΔE, please see [83–85].

5

Analysis method

5.1 Introduction

The analysis method described in this chapter is for the measurement of the branching ratio of the charmless, hadronic decay

$$B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+.$$

Charge conjugate states are assumed throughout. The final state is the 3-body state $K^+K^-\pi^+$.

In keeping with *BABAR* collaboration policy, the analysis described here is performed *blind* [86]. That is, the signal yield is not observed (i.e., *unblinded*) until all aspects

of the analysis are complete.

Statistics permitting, an amplitude-level analysis of the full $B^+ \to K^+K^-\pi^+$ Dalitz plot is preferred. For such an analysis, the relative magnitudes and phases of the contributing resonant and non-resonant components are extracted using the *maximum likelihood* (ML) technique. The amplitude-level approach thus makes possible measurements of interference between intermediate states, in addition to the measurements of rates. The ML method is discussed in Section 5.6.

However, despite PEP-II's impressive luminosities, the amount of on-resonance data acquired by the *BABAR* experiment at the time of writing are not sufficient to carry out an analysis at the amplitude level. Instead an intensity-level analysis is conducted. For such an analysis it is not possible to quantify the amount of interference between intermediate states (e.g. $\overline{K}^*(892)^0K^+$ and $\overline{K}^*_0(1430)^0K^+$). It must therefore be treated as a source of systematic error.

The analysis presented is a *quasi-two-body* analysis in that $\overline{K}^*(892)^0K^+$ is inferred from the structure observed in 3-body phase space. A ML fit is applied to extract a signal yield in a particular region of the Dalitz plot where the $K^*(892)^0$ is expected to dominate. This yield, N_S, is used to calculate the branching ratio, $\mathcal{B}$:

$$\mathcal{B} = \frac{N_S}{\epsilon N_{B^\pm}}. \tag{5.1}$$

ϵ is the efficiency of reconstruction and the selection criteria applied. It is estimated using MC data. $N_{B^\pm}$ is the number charged B mesons used, as discussed in Section 4.4.

The main issue faced in this analysis is due to the possible presence of the channel $B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$ for which the following are unknown:

- Branching ratio.
- $m_{K\pi}$ lineshape (as discussed in Section 1.4.3).
- Amount of interference with $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$.

Using the current 90% CL upper limit of 6.3×10^{-6} on $\mathcal{B}(B^+ \to \overline{K}_0^*(1430)^0 K^+, \overline{K}_0^*(1430)^0 \to K^-\pi^+)$ [23] leads to a fractional systematic uncertainty on the measurement of the rate of the 892 mode that is in excess of 100%. To decrease this uncertainty it is necessary to obtain a more sensitive measurement than the one quoted above for the rate at which $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ occurs.

The following strategy is therefore employed:

- Two independent fits are conducted in a region of the Dalitz plot where the $K^*(892)^0$ is expected to dominate. The two fit models are designed to extract:
  - $K^+K^-\pi^+$ final state events that occur via the $K^*(892)^0$ resonance.
  - All final state $K^+K^-\pi^+$ events.

  These fits will be referred to as the "*5-variable fit in the 892 window*" and the "*3-variable fit in the 892 window*" respectively throughout the remainder of this thesis.

- A fit is conducted in a region of the Dalitz plot where the $K_0^*(1430)^0$ is expected to dominate. This fit will be referred to as the "*3-variable fit in the 1430 window*" throughout the remainder of this thesis.

If final state $B^+ \to K^+K^-\pi^+$ events are present in the 892 window, which do not proceed via the $K^*(892)^0$ resonance, we would expect the signal yield extracted from the 3-variable fit in the 892 window to be larger than the signal yield obtained using the 5-variable fit. The most likely sources for such events are, as discussed above, $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$, and also non-resonant $B^+ \to K^+K^-\pi^+$. Upper limits on the numbers of $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ and non-resonant $B^+ \to K^+K^-\pi^+$ events present in the data sample can be estimated using the yield extracted from the 3-variable fit in the 1430 window[1]. If the 3-variable fits indicate

[1]Ideally, a *non-resonant window* – a region of the $B^+ \to K^+K^-\pi^+$ Dalitz plot free of any intermediate resonances – would be used to estimate the number of non-resonant $B^+ \to K^+K^-\pi^+$ events present in the dataset (see Figure 5.3). However, time did not permit such a study here.

that there are no $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ or non-resonant $B^+ \to K^+K^-\pi^+$ events present, the yield obtained for the 5-variable fit can be taken to be due only to $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ (and should be consistent with the yield from the 3-variable fit in the 892 window). Otherwise – since the 5-variable fit doesn't discriminate 100% between final states that do and do not occur via the $K^*(892)^0$ resonance – it is necessary to estimate the bias on the 5-variable fitted yield due to $B^+ \to K^+K^-\pi^+$ events that occur via an alternative – or no – intermediate state. Only then can a measurement of the $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ rate be made.

In the next section, an overview of the analysis method is given before describing the analysis in detail in the rest of the chapter. The results obtained are presented in Chapter 6.

5.2 Overview

The specifics of the independent 5- and 3-variable fits introduced in the previous section are presented in Table 5.1.

The general procedure used for each fit is now outlined:

- Events are reconstructed and skimmed as described in Section 4.2.
- Selection criteria are applied to greatly enhance the signal to continuum background ratio. This is achieved using the PID selectors and discriminating variables discussed in Sections 4.2.3 and 4.5 respectively.
- Sources of B related background are identified using MC.
- Signal and background events are modelled in the form of *Probability Density Functions* (PDFs).
- A fit model is built using these PDFs.

- This model is validated using *toy* and real MC. Real MC is MC of the type described in Section 4.3; toy MC is discussed in Section 5.10.1.
- The model is applied to the on-resonance data sample.

5.3 Event selection

The selection of the reconstructed events that are outputted by PR (as discussed in Section 4.2) occurs in three stages. These are now described.

5.3.1 Preselection

Preselection is where `AllEvents` are subjected to a filter algorithm, which outputs the skimmed collection `BCCC03a3body`. Each track in the `GoodTracksLoose` list (see Table 4.1) is kinematically fitted as a kaon or as a pion. For each triplet of tracks with overall charge ± 1 the algorithm forms all possible distinct combinations (eight in total, e.g. track 1 = K^+, track 2 = π^-, track 3 = π^+). If any of the combinations pass the criteria in Table 5.2, the event makes it into the `BCCC03a3body` skim.

5.3.2 Batch level pre-analysis

The `NonCharm3BodyUser` package, briefly discussed in Section 4.2, is used to further refine the `BCCC03a3body` collection. More demanding criteria are placed on the skimmed events, and variables specific to subsequent analyses are calculated using more general, fundamental quantities.

Vertexing of B candidates is carried out, with geometric and kinematic constraints applied (vertexing is discussed in Section 4.2.4). m_{ES} and ΔE are then recalculated using the track momenta obtained from the vertex fit. Next, a further vertex fit is performed, identical except that this time the B mass constraint is applied. PID

Fit	***Variables Used In Fit***	***Fit Components***	***Dalitz Plot Window***	***Selection Criteria***	***Signal Split?***
5-variable fit in the 892 window	m_{ES}, ΔE, $\mathcal{F}$, $m_{K\pi}$, $\cos\theta_H$	S, $q\overline{q}$, $(B\overline{B})_1$, $(B\overline{B})_2$, $(B\overline{B})_3$	*892 window* : $0.744 < m_{K\pi}/(\,\mathrm{GeV}/c^2) < 1.048$	Summarised in Column 1, Table 5.3	Yes, into S_{TRU} and S_{SXF}
3-variable fit in the 892 window	m_{ES}, ΔE, $\mathcal{F}$	S, $q\overline{q}$, $(B\overline{B})_1$, $(B\overline{B})_2$, $(B\overline{B})_3$	As above	Summarised in Column 2, Table 5.3	No
3-variable fit in the 1430 window	m_{ES}, ΔE, $\mathcal{F}$	S, $q\overline{q}$, $(B\overline{B})_2$, $(B\overline{B})_3$, $(B\overline{B})_4$	*1430 window* : $1.048 < m_{K\pi}/(\,\mathrm{GeV}/c^2) < 1.800$, $m_{KK} > 2.500\,\mathrm{GeV}/c^2$; veto: $3.284 < m_{KK}/(\,\mathrm{GeV}/c^2) < 3.618$	Summarised in Column 3, Table 5.3	No

Table 5.1: *Summary of fits performed and Dalitz plot windows used. The 892 window and the 1430 window are illustrated in Figure 5.3. The fit method is described in Section 5.6. Hypotheses for which fit components are used are:*
S: signal hypothesis.
$q\overline{q}$: continuum background hypothesis.
$(B\overline{B})_{1-4}$: B related background hypotheses.
Splitting the signal component into truth matched and self-cross-feed (SXF) sub-components is discussed in Section 5.7.1.

selectors are then run, and the topological and resonance discriminating variables (described in Sections 4.5.1 and 4.5.2 respectively) are calculated (using the B mass

Requirement		***Further Comments***
Total charge of B candidate	$= \pm 1$	Charged B meson decay
Total number of tracks in event	≥ 4	At least one track from ROE
Total energy of event	$< 20\,\text{GeV}$	Basic sanity check
$\lvert m_{ES} - \frac{\sqrt{s}}{2} \rvert$	$\leq 0.1\,\text{GeV}/c^2$	$\frac{\sqrt{s}}{2} = 5.29\,\text{GeV}/c^2$
$\lvert \Delta E \rvert$	$< 0.45\,\text{GeV}$	ΔE and m_{ES} are defined in Section 4.5.3

Table 5.2: *Preselection criteria.*

constrained vertex fit). Finally, a loose cut of $|\Delta E| < 0.35\,\text{GeV}$ is applied.

`NonCharm3BodyUser` processed events are outputted to `Root` [68] ntuples, and are general to any three-charged-track mode.

5.3.3 $B^+ \to K^+K^-\pi^+$ selection

Further selection (using the `CharmlessFitter` package) is performed on the three-charged-track combinations to single out the $K^+K^-\pi^+$ final state and further suppress backgrounds (in the order shown):

- The B candidate is required to have a valid kinematic vertex fit for a $KK\pi$ track hypothesis.
- Kaon candidate tracks must pass the kaon selector, `SMSKaonSelector`, in *Tight* mode; the pion candidate track must fail `SMSKaonSelector` in *Tight* mode.

- All tracks must fail the electron selector, `LHSelector`, in the mode defined by Eq. (4.7).
- Kaon candidate tracks must have opposite charges.
- $|\cos\theta_T| < 0.9$.
- $|\mathcal{F}| < 3.0$.
- Cuts are applied to the invariant masses, $m_{K\pi}$ and, for the 1430 window, m_{KK}. Dalitz plot windows are discussed in the next section.
- For the 5-variable fit in the 892 window, $\cos\theta_H < 0.9$. This cut is applied not to enhance the signal to continuum background ratio (in fact it is detrimental to this ratio), but to make possible a fit to the $\cos\theta_H$ signal distribution. Figure 5.1 illustrates that, for signal events, there is a sharp drop off as $\cos\theta_H$ approaches +1. This is because the momentum of the resonance's daughter pion is low at high $\cos\theta_H$, and difficult to reconstruct.

  When used as a fit variable, $\cos\theta_H$ provides excellent discrimination between states that occur via the spin 1 $K^*(892)^0$ resonance and states that do not.

  For constant reconstruction and selection criteria efficiency as a function of $\cos\theta_H$, a symmetric $\cos\theta_H$ cut ensures that interference between the intermediate resonances $K^*(892)^0$ and $K_0^*(1430)^0$ cancels. However, $|\cos\theta_H| < 0.9$ is not used despite the asymmetric cut introducing a source of systematic error due to the unknown amount of interference. This is because the analysis is statistics dominated with the extra source of systematic error deemed a worthwhile trade-off for the enhanced sensitivity. A more detailed discussion can be found in Section 4.6 of [87].
- $5.22\,\mathrm{GeV}/c^2 < m_{ES} < 5.29\,\mathrm{GeV}/c^2$.
- $|\Delta E| < 0.1\,\mathrm{GeV}$.

- For the 1430 window, the potential $B\overline{B}$ background $B^+ \to \chi_{c0}(\to K^+K^-)K^+$ is vetoed (see Section 5.4).

- Finally, the requirement of a single candidate per event is imposed. In the rare (less than $< 5\%$ for signal MC) case of multiple candidates existing, the single candidate is chosen arbitrarily to avoid introducing bias.

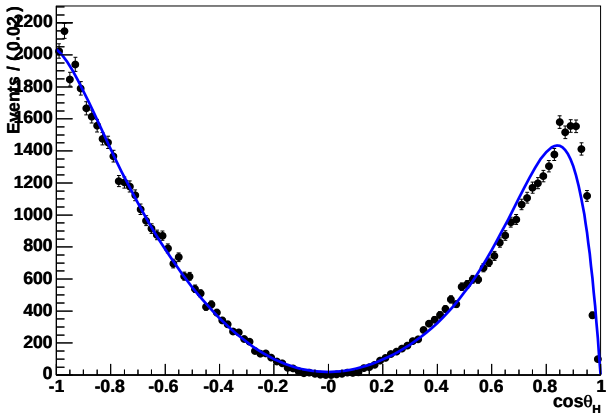

Figure 5.1: *$\cos\theta_H$ distribution for reconstructed $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ MC (black points). The blue, solid line represents one of many failed attempts to successfully fit to this distribution in the range $-1 \longrightarrow +1$. These unsuccessful fits are due to the sharp drop off toward $\cos\theta_H = 1$.*

The requirements listed here are summarised for the 5- and 3-variable fits in the 892 window, and for the 3-variable fit in the 1430 window, in Table 5.3. `CharmlessFitter` output is in the form of `Root` files with ntuples containing a smaller number of variables and fewer events relative to the input files (the `NonCharm3BodyUser` output).

The efficiencies of the selection criteria discussed here for relevant data samples are shown in Table 5.4. For the 5-variable fit in the 892 window, a breakdown of the efficiencies for individual cuts for $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ MC and for on- and off-resonance data are given in Tables 5.5-5.6.

At this point it is useful to introduce the following notation:

Selection ***(Cut)*** ***{Veto}***		***Column 1***	***Column 2***	***Column 3***
Reconstruction, preselection		✓	✓	✓
$KK\pi$ hypothesis		✓	✓	✓
PID – `SMSKaonSelector` *Tight*		✓	✓	✓
PID – `LHSelector` *Tight*		✓	✓	✓
Oppositely charged kaon cands.		✓	✓	✓
$\lvert\cos\theta_T\rvert$	(< 0.9)	✓	✓	✓
$\lvert\mathcal{F}\rvert$	(< 3.0)	✓	✓	✓
$m_{K\pi}/(\,\mathrm{GeV}/c^2)$	$(> 0.744, < 1.048)$	✓	✓	
$m_{K\pi}/(\,\mathrm{GeV}/c^2)$	$(> 1.048, < 1.800)$			✓
$m_{KK}/(\,\mathrm{GeV}/c^2)$	(> 2.500)			✓
$\cos\theta_H$	(< 0.9)	✓		
$m_{ES}/(\,\mathrm{GeV}/c^2)$	$(> 5.22, < 5.29)$	✓	✓	✓
$\lvert\Delta E\rvert/\,\mathrm{GeV}$	(< 0.1)	✓	✓	✓
$m_{KK}/(\,\mathrm{GeV}/c^2)$	$\{> 3.284, < 3.618\}$			✓
One candidate/event imposed		✓	✓	✓

Table 5.3: *Summary of selection criteria applied for the 5-variable fit in the 892 window (column 1), the 3-variable fit in the 892 window (column 2), and the 3-variable fit in the 1430 window (column 3).*

- RaSC—reconstruction and the selection criteria summarised in Table 5.3.
- Final on-resonance sample—on-resonance data that have been subjected to RaSC.

	Signal MC	***On-res. Data***	***Off-res. Data***
5-variable fit in the 892 window (signal MC simulates the mode $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$):			
MC Events Generated	$242,000$	—	—
Reconstructed/Preselected	$174,951$	2.487×10^7	2.487×10^6
Passed Selection Criteria	$64,391$	$38,690$	$4,077$
3-variable fit in the 892 window (signal MC simulates the mode $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$):			
MC Events Generated	$242,000$	—	—
Reconstructed/Preselected	$174,951$	2.487×10^7	2.487×10^6
Passed Selection Criteria	$68,747$	$40,156$	$4,239$
3-variable fit in the 1430 window (signal MC simulates the mode $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$):			
MC Events Generated	$117,000$ $^{\beta}$, $128,150$ $^{\lambda}$	—	—
Reconstructed/Preselected	$82,838$ $^{\beta}$, $91,620$ $^{\lambda}$	2.487×10^7	2.487×10^6
Passed Selection Criteria	$26,055$ $^{\beta}$, $24,143$ $^{\lambda}$	$46,657$	$4,792$

$^{\beta}$: Breit-Wigner $m_{K\pi}$ lineshape; $^{\lambda}$: LASS $m_{K\pi}$ lineshape.

Table 5.4: *Efficiencies of reconstruction and selection criteria. The number of reconstructed/preselected events is the number remaining after PR, the skimming algorithm and* `NonCharm3BodyUser` *code have been run (described in Sections 4.2, 5.3.1 and 5.3.2 respectively). The number of events to pass the selection criteria is the number remaining after* `CharmlessFitter` *code has been applied.* `CharmlessFitter` *implements the cuts described in Section 5.3.3, and summarised in Table 5.3.*

Selection ***(Cut)***	***Truth Match***	***SXF***	***MC Total***
Reconstruction, preselection	0.563	0.723	0.723
$KK\pi$ hypothesis	0.999	0.984	0.986
PID – `SMSKaonSelector` *Tight*	0.565	0.316	0.575
PID – `LHSelector` *Tight*	1.000	1.000	1.000
Oppositely charged kaon cands.	1.000	0.939	0.967
$\lvert\cos\theta_T\rvert$ (< 0.9)	0.926	0.909	0.931
$\lvert\mathcal{F}\rvert$ (< 3.0)	0.998	0.998	0.998
$m_{K\pi}/(\,\mathrm{GeV}/c^2)$ $(> 0.744, < 1.048)$	0.940	0.326	0.821
$\cos\theta_H$ (< 0.9)	0.945	0.811	0.946
$m_{ES}/(\,\mathrm{GeV}/c^2)$ $(> 5.22, < 5.29)$	0.999	0.932	0.996
$\lvert\Delta E\rvert/\,\mathrm{GeV}$ (< 0.1)	0.979	0.455	0.934
One candidate/event imposed at this stage ...			
Overall Efficiency	**0.2499 ± 0.0009**	**0.0162 ± 0.0003**	**0.2661 ± 0.0009**

Table 5.5: *Breakdown of reconstruction and selection criteria efficiencies for* $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ *signal MC. The cuts are hierarchical – performed in the order shown in the table. Each line refers to the efficiency of that cut after all previous cuts have been applied. The efficiencies shown here are calculated by counting the number of events that have at least one candidate, which passes the relevant cut. Self-cross-feed (SXF) is discussed in Section 5.5.2.4. This table is relevant to the 5-variable fit in the 892 window.*

Selection	***(Cut)***	***On-res. Data***	***Off-res. Data***
Reconstruction, preselection		$(2.487 \times 10^7$ events)	$(2.487 \times 10^6$ events)
$KK\pi$ hypothesis		0.889	0.890
PID – `SMSKaonSelector` *Tight*		0.087	0.087
PID – `LHSelector` *Tight*		0.960	0.963
Oppositely charged kaon cands.		0.834	0.837
$\lvert\cos\theta_T\rvert$	(< 0.9)	0.621	0.613
$\lvert\mathcal{F}\rvert$	(< 3.0)	1.000	1.000
$m_{K\pi}/(\,\mathrm{GeV}/c^2)$	$(> 0.744, < 1.048)$	0.210	0.222
$\cos\theta_H$	(< 0.9)	0.968	0.966
$m_{ES}/(\,\mathrm{GeV}/c^2)$	$(> 5.22, < 5.29)$	0.685	0.682
$\lvert\Delta E\rvert/\,\mathrm{GeV}$	(< 0.1)	0.291	0.292
One candidate/event imposed at this stage ...			
Efficiency, relative to preselection		$\mathbf{(1.556 \pm 0.008) \times 10^{-3}}$	$\mathbf{(1.639 \pm 0.026) \times 10^{-3}}$

Table 5.6: *Breakdown of selection criteria efficiencies for on- and off-resonance data. The cuts are hierarchical – performed in the order shown in the table. Each line refers to the efficiency of that cut after all previous cuts have been applied. The efficiencies shown here are calculated by counting the number of events that have at least one candidate, which passes the relevant cut. This table is relevant to the 5-variable fit in the 892 window.*

5.4 Dalitz plot windows

The regions of the $B^+ \to K^+K^-\pi^+$ Dalitz plot to be used are chosen such that:

- The region contains a large proportion of the decays of the dominant resonance of interest.
- Interference from neighbouring modes is minimised.
- Contamination by B related background is minimised (an in-depth discussion of B related background is presented in Section 5.5.2).

The channels shown in Table 5.7 are identified as modes that could cause contamination, if care is not taken in choosing suitable $m^2_{K\pi}$ vs. m^2_{KK} windows. Figure 5.2 illustrates where simulated events for these modes lie in the Dalitz plot. These events have been subjected to RaSC (except that no cuts are applied to $m_{K\pi}$, m_{KK} and $\cos\theta_H$).

Where a mode forms a clear band in the Dalitz plot, and when the number of events expected in the final on-resonance sample for this mode is not insignificant, a veto can be applied (should the band intersect a window to be used to study signal events). A Breit-Wigner is fitted to the invariant mass distribution ($m_{K\pi}$ or m_{KK}) from the relevant MC distribution. The veto is then centered at the fitted Breit-Wigner's mean, with a width of ten times the Breit-Wigner's width.

The dimensions of the Dalitz plot windows used are presented in Table 5.1. The windows are illustrated in Figure 5.3.

There are further sources of B related background for which the use of vetoes is not appropriate. These sources must be modelled in the fit. This is discussed in Section 5.5.2.

Channel, $B^+ \to$	***Number Of Events Expected In Final On-resonance Sample***
$J/\psi(\to l^+l^-)K^+$	$138.2 \pm 4.7 \pm 6.0$
$\chi_{c0}(\to K^+K^-)K^+$	$98.9 \pm 0.8 \pm 42.3$
$\chi_{c0}(\to \pi^+\pi^-)K^+$	$9.0 \pm 0.3 \pm 3.7$
$\psi(2S)(\to l^+l^-)K^+$	$11.6 \pm 0.8 \pm 1.0$
$\overline{D}^0(\to K^+K^-)\pi^+$	$1371.2 \pm 14.6 \pm 97.0$
$\overline{D}^0(\to K^+\pi^-)\pi^+$	$565.6 \pm 10.7 \pm 36.1$
$\overline{D}^0(\to K^+\pi^-\pi^0)\pi^+$	$837.8 \pm 22.4 \pm 71.6$
$\overline{D}^0(\to K^+\pi^-)K^+$	$51.6 \pm 0.7 \pm 8.5$
$\phi(1020)(\to K^+K^-)\pi^+$	$17.5 \pm 0.1 \pm 17.5$
$J/\psi(\to K^+K^-)K^+$	$6.4 \pm 0.1 \pm 0.9$
$\overline{D}^0(\to K^-\pi^+)K^+$	$3.9 \pm 0.1 \pm 0.7$

Table 5.7: *Background modes expected to be present in the $B^+ \to K^+K^-\pi^+$ Dalitz plot, and the number of events expected in the final on-resonance sample. This number is calculated using RaSC efficiencies obtained from exclusive MC, and from PDG [23] branching ratios. The first error is statistical (from the statistical errors on the MC efficiency and $N_{B^\pm}$); the second is systematic (from the uncertainty on the branching ratios, and the systematic uncertainty on $N_{B^\pm}$).*

5.5 Background

The two types of background to be accounted for in this analysis have been touched upon in previous chapters, namely continuum ($q\overline{q}$) background and B related ($B\overline{B}$) background. Here, they are discussed in a little more detail. Self-cross-feed (SXF) is also discussed.

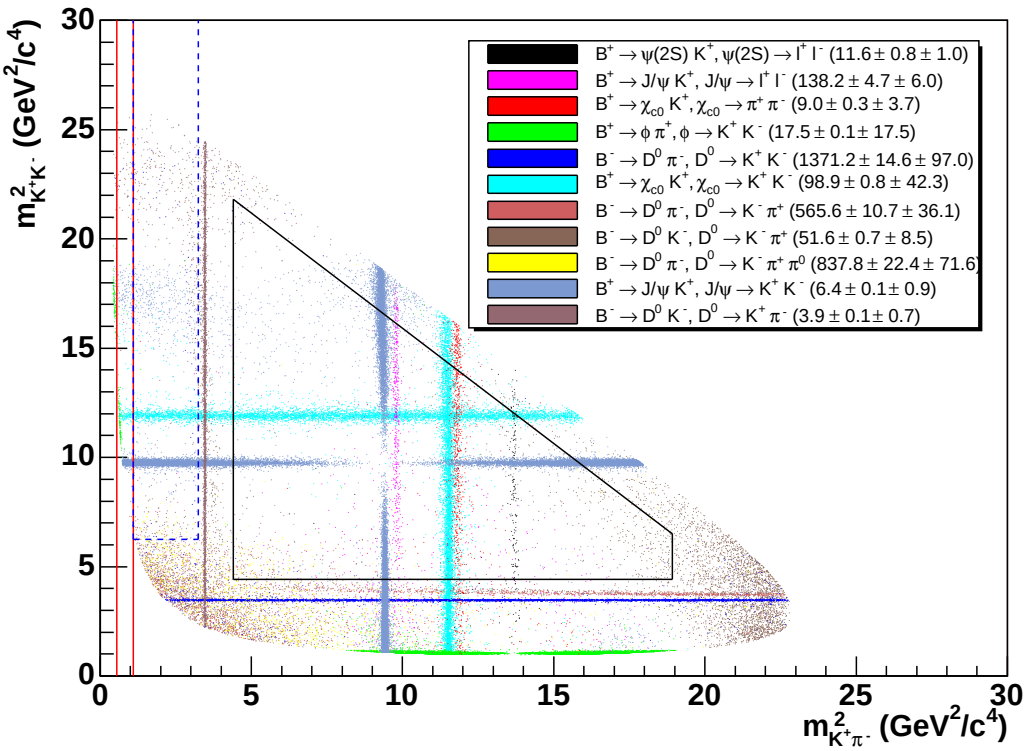

Figure 5.2: *Distributions of the events of the background modes listed in Table 5.7 in the $B^+ \to K^+K^-\pi^+$ Dalitz plot. The number of events plotted for each mode depends only on the number of MC events available and is irrelevant. It is the positions of the events shown and the number of events expected in the final on-resonance sample (shown in the legend and in Table 5.7) that must be considered when choosing appropriate regions of the Dalitz plot. The solid, red lines illustrate the window where $K^*(892)^0$ is expected to dominate; the blue, dashed lines depict the region expected to be dominated by $K_0^*(1430)^0$ (avoiding the bottom left corner of the Dalitz plot where B related background is more abundant). Additionally, vetoes will be applied where deemed necessary. (The black, solid lines represent an area suitable for measuring the non-resonant contribution after vetoes have been applied).*

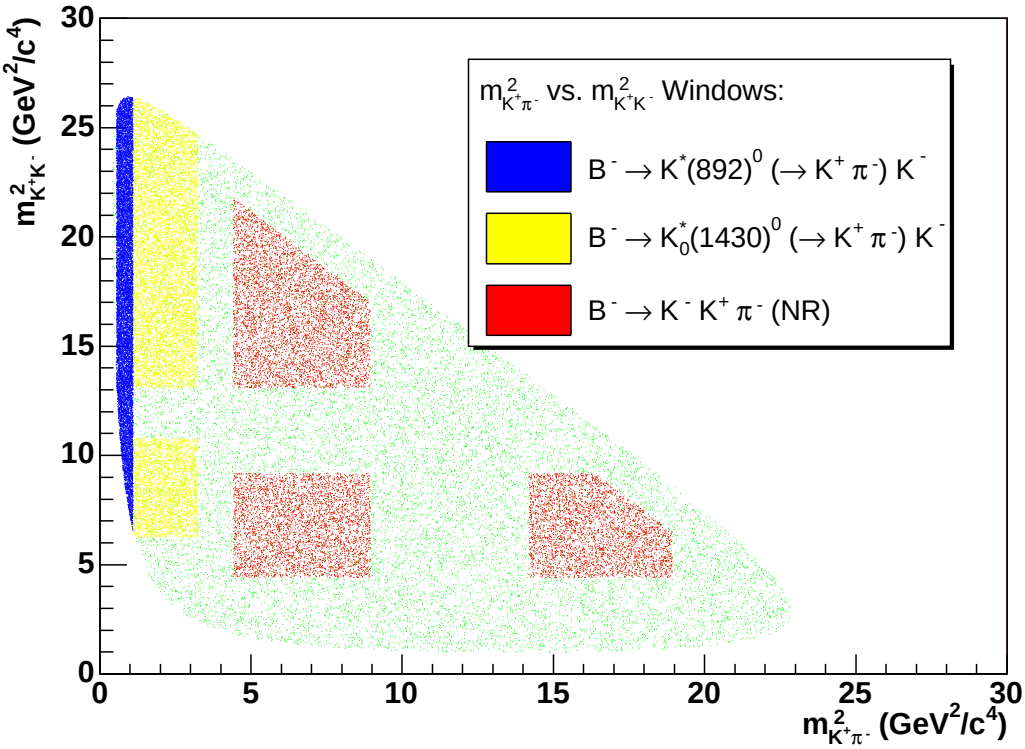

Figure 5.3: *Blue: the 892 window; yellow: the 1430 window. Also shown is the available Dalitz plot phase space (green) and an appropriate window to measure the non-resonant contribution (red).*

5.5.1 Continuum ($q\overline{q}$) background

$q\overline{q}$ background is the dominant source of background in this analysis. It is the result of incorrectly reconstructing random tracks that originate from di-jet quark sources ($e^+e^- \to c\overline{c}$, $s\overline{s}$, $u\overline{u}$, $d\overline{d}$) as a signal B meson. $e^+e^- \to \tau^+\tau^-$ also contributes. The selection criteria presented in Section 5.3.3 eliminate the vast majority of $q\overline{q}$ background. However, the events that remain greatly outnumber the signal events.

$q\overline{q}$ background is modelled using off-resonance data, which consist purely of continuum events. Modelling the background involves choosing PDFs that accurately represent the distributions of the fit variables, as shown in Sections 5.7-5.9. A *sideband*

in on-resonance data is also used as a cross-check. RaSC are applied to these data samples before modelling takes place. The exception is that for the on-resonance sideband, $0.10 < \Delta E/\,\mathrm{GeV} < 0.35$.

5.5.2 B related ($B\overline{B}$) background

$B\overline{B}$ background is the result of B decays to modes that are not the signal mode ($e^+e^- \rightarrow B\overline{B} \rightarrow X \neq \overline{K}^*(892)^0(\rightarrow K^-\pi^+)K^+$), but that are mis-reconstructed as such. It is potentially more dangerous, as it can be more difficult to distinguish from signal, and it must be ensured that no modes that are a potential source of $B\overline{B}$ background are ignored.

Sources of $B\overline{B}$ background fall into several categories. These are now discussed (Sections 5.5.2.1-5.5.2.4). A discussion of how $B\overline{B}$ background is identified and treated is then presented (Section 5.5.2.5).

5.5.2.1 Particle mis-identification

The inclusive (charmless) intensities for the channels $B^+ \rightarrow K^+K^-K^+$ and $B^+ \rightarrow K^+\pi^-\pi^+$ are $\approx 1-2$ orders of magnitude larger than that for $B^+ \rightarrow K^+K^-\pi^+$. Therefore, although the rate at which pions are incorrectly identified as kaons (and vice versa) is low, the number of $B\overline{B}$ background events occurring as a result of particle mis-identification is appreciable and cannot be ignored. Due to the tight PID cut applied to kaon track candidates, and the relatively looser PID cut applied to pion track candidates, the rate at which kaons are mis-identified as pions is approximately one order of magnitude larger than the rate at which pions are mis-identified as kaons. For the same reason, $B^+ \rightarrow \pi^+\pi^-\pi^+$ has very little impact. The ΔE variable assists in discriminating against $B\overline{B}$ background of this type. As discussed in Section 4.5.3, for correctly reconstructed signal events, ΔE is expected to peak at zero. For reconstructed candidates where particle mis-identification has

occurred, the peak is expected to be shifted from zero due to the difference between the mass of the actual particle and the hypothesised mass of the reconstructed candidate (see, for example, the second row plots in Figure 5.16). This is useful when the difference in masses between the reconstructed and actual particle types is appreciable, such as for pion versus kaon – and also for muon/electron versus kaon and electron versus pion.

5.5.2.2 Combinatorics

Combinatoric $B\overline{B}$ background occurs where a $B\overline{B}$ event has three random, unrelated tracks incorrectly reconstructed as a signal event. These events have discriminating variable distributions very similar to those of $q\overline{q}$ background. This is useful because although there are a considerable number of events of this type present, the fit will assign a background hypothesis to the majority and the impact on the signal yield extracted from the fit will be small.

5.5.2.3 4- and 2-body decays

B decays to final state 4- and 2-body modes are a source of $B\overline{B}$ background when a particle is lost in reconstruction (usually a low momentum neutral), or where a track from the other B in the event is incorrectly attributed to the reconstructed signal B. Decays of neutral B mesons, where combinations of a lost neutral daughter and the attribution of a track from the other B occur, are also a possible source of $B\overline{B}$ background.

5.5.2.4 Self-cross-feed (SXF)

SXF is where signal events have been mis-reconstructed by switching one or more tracks from the decay of the true signal B with tracks from the decay of the other B in the event. The amount of SXF present is estimated from MC studies. For the 5-

variable fit in the 892 window, of the 242,000 $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ signal MC events generated, a total of 64,391 pass RaSC of which 3,918 lack a MC truth match (see Section 4.3) and so are considered to be SXF. Despite being described in this section, SXF is taken to belong to the signal hypothesis in this analysis.

5.5.2.5 Identification and treatment of $B\overline{B}$ background

To identify sources of $B\overline{B}$ background RaSC are applied to MC samples containing generic B^+B^- and $B^0\overline{B}^0$ decays (535.9 and 541.3 million events respectively). MC truth information is then used to determine the channels that significantly contribute. These channels are treated individually using exclusive MC samples. World average branching ratios and the efficiencies of applying RaSC to the exclusive MC samples are used to calculate the number of events expected in the final on-resonance sample for each source, $n_{B\overline{B}\ Bkgd}$. If this number – plus its error, $\Delta n_{B\overline{B}\ Bkgd}$ – is in excess of 0.5 events, *toy tests* are conducted. These toy tests are described in Section 5.10.4. Their purpose is to determine whether the presence of the $B\overline{B}$ background source induces a bias on the signal yield extracted from the fit, and if so how large.

The events remaining in the generic MC samples – after the modes to be treated individually have been removed – provide ideal combinatoric samples. As with the exclusive samples, events from these samples are embedded into the toy tests described in Section 5.10.4 to establish the impact on the signal yield.

Whilst generic MC provides an excellent tool to identify sources of $B\overline{B}$ background, it does not necessarily provide 100% coverage of every possible mode. In addition to those identified using the generic samples, further channels are considered as being possible sources of $B\overline{B}$ background. All modes considered as possible sources of $B\overline{B}$ background for the *BABAR* $B^+ \to K^+K^-K^+$ [88] and $B^+ \to K^+\pi^-\pi^+$ [50] analyses are also considered for the analysis described here. Modes with higher K^* resonances (e.g. $K_0^*(1430)^0$) are investigated, as are charm and charmonium modes.

Charm and charmonium modes are a potential problem due to their relatively much larger branching ratios. These large branching ratios do mean, however, that such modes are very likely to be accurately represented in the generic samples.

Recent analyses of the $B^+ \to K^+K^-K^+$ Dalitz plot strongly suggest the presence of a resonant structure at $m_{KK} \approx 1500\,\mathrm{MeV}/c^2$ [42, 88]. Furthermore, the non-resonant $B^+ \to K^+K^-K^+$ contribution is modelled for these analyses as being non-uniform across the Dalitz plot. $B^+ \to K^+K^-K^+$ is a major source of $B\overline{B}$ background for the analysis described in this chapter. As such it is important that it is modelled correctly. For the *BABAR* non-resonant $B^+ \to K^+K^-K^+$ MC sample, the events are distributed evenly over the available 3-body phase space. There is no available MC for the aforementioned $\approx 1500\,\mathrm{MeV}/c^2$ structure. It has therefore been decided to generate MC that simulates $B^+ \to K^+K^-K^+$ events that occur via numerous resonances, plus the non-resonant contribution, with both relative magnitudes and phases modelled. The model used is Belle's published model B_0, solution 1 [42], minus the $B^+ \to \phi(1020)K^+$ contribution[2]. This model accounts for the $\approx 1500\,\mathrm{MeV}/c^2$ structure and for the non-uniform non-resonant distribution across the Dalitz plot. Further models are used as a cross-check, and to quantify the model dependent systematic error.

Table 5.8 (Table 5.9) lists the sources of $B\overline{B}$ background that contribute $n_{B\overline{B}\ Bkgd} + \Delta n_{B\overline{B}\ Bkgd} > 0.5$ events to the final on-resonance sample for the 5-(3-)variable fit in the 892 (1430) window. The modes listed in Table 5.8 are also the modes that contribute $n_{B\overline{B}\ Bkgd} + \Delta n_{B\overline{B}\ Bkgd} > 0.5$ events for the 3-variable fit in the 892 window, though values of $n_{B\overline{B}\ Bkgd}$ are slightly different.

[2]Ideally, MC simulating both the $B^+ \to K^+\pi^-\pi^+$ and $B^+ \to K^+K^-K^+$ Dalitz plots in their entirety is preferred, since interference between intermediate resonances is accounted for. However, the process used to generate the Dalitz plot modelled MC is too inefficient for the $B^+ \to K^+\pi^-\pi^+$ plot. Due to $\phi(1020)$'s narrow width, the process also becomes too inefficient for the $B^+ \to K^+K^-K^+$ plot when the $B^+ \to \phi(1020)K^+$ contribution is included. $B^+ \to \phi(1020)(\to K^+K^-)K^+$ is treated individually, as are the contributions that make up the $B^+ \to K^+\pi^-\pi^+$ plot. This is adequate in the context of the analysis described in this chapter.

Modes listed in Table 5.10 are also identified as possible sources of $B\overline{B}$ background for the 5-variable fit in the 892 window. These modes share the same final state as the mode under investigation. Interference is therefore an issue.

A full list of MC samples that are subjected to RaSC can be found in Appendix B. Also shown in Appendix B is a list of the most abundant modes present in the combinatoric generic samples, after the modes that are treated exclusively have been removed.

5.6 The maximum likelihood (ML) method

The analysis described in this chapter is a multivariate analysis. That is, numerous variables are used to form a model that discriminates between various hypotheses – *signal* and *background* being examples of such hypotheses. Fitting a model to data can be done in many ways, for example the χ^2 method. Perhaps one of the most elegant, however, is that of the *maximum likelihood* (ML) approach. The advantages of using the ML method include:

- Its robustness for low statistics fits.
- It can be performed on unbinned data eliminating the need to choose bin sizes, which can lead to bias and inaccuracy.

The ML method is discussed in detail in [91]. The basic principle is now described.

Consider numerous independent measurements of some experimental variable x. The probability of observing x in the infinitesimal interval $[x, x + dx]$ is given by $\mathcal{P}(x;\vec{\Theta})\,dx$, where $\mathcal{P}(x;\vec{\Theta})$ is the *Probability Density Function* (PDF) with parameters $\vec{\Theta}$. The PDF is normalised such that the probability of measuring a value of x that lies within the complete sample range $[x_{min}, x_{max}]$ is unity:

$$\int_{x_{min}}^{x_{max}} \mathcal{P}(x;\vec{\Theta})\,dx = 1. \tag{5.2}$$

$B\overline{B}$ ***Background Source***	***Number Of Events To Pass RaSC***	***# MC Events Run Over***	***Branching Ratio*** ($\times 10^{-6}$)	***# Events Expected In Final On-res. Sample***
Combinatoric charged generics	316	5×10^8	$(5.0 \pm 2.5) \times 10^5$	$68.3 \pm 3.8 \pm 34.2$
Combinatoric neutral generics	118	5×10^8	$(5.0 \pm 2.5) \times 10^5$	$25.2 \pm 2.3 \pm 12.6$
$B^+ \to K^+K^-K^+$ Dalitz plot model, minus the $B^+ \to \phi(1020)K^+$ contribution				
Belle model B_0, sol. 1, minus ϕK [42]	402	85,190	21.1 ± 2.0 [89]	$23.0 \pm 1.1 \pm 2.2$
(Belle model B_0, sol. 2, minus ϕK [42]	188	47,464	21.1 ± 2.0 [89]	$19.3 \pm 1.4 \pm 1.9$)
(*BABAR* solution A, minus ϕK [88]	306	49,568	21.1 ± 2.0 [89]	$30.1 \pm 1.7 \pm 2.9$)
(*BABAR* solution B, minus ϕK [88]	295	48,172	21.1 ± 2.0 [89]	$29.9 \pm 1.7 \pm 2.9$)
$B^+ \to \phi(1020)(\to K^+K^-)K^+$	3,266	163,000	4.4 ± 0.3 [23] [89]	$20.6 \pm 0.4 \pm 1.5$
$B^0 \to K^*(892)^0(\to K^+\pi^-)\overline{K}^*(892)^0(\to K^-\pi^+)$ (L)	740	122,000	< 9.8 [89]	$6.9 \pm 0.3 \pm 6.9$
$B^+ \to \eta'(958)(\to \gamma\rho(770)^0(\to \pi^+\pi^-))K^+$	175	145,000	20.5 ± 1.1 [23] [89]	$5.7 \pm 0.4 \pm 0.3$
$B^0 \to K^+K^-\pi^0$ (NR)	299	126,000	< 19.0 [23]	$5.2 \pm 0.3 \pm 5.2$
$B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$	742	206,000	5.1 ± 0.9 [89]	$4.3 \pm 0.2 \pm 0.7$
$B^0 \to K^+K^-K^0_S$ (NR)	178	220,000	12.4 ± 1.2 [89]	$2.3 \pm 0.2 \pm 0.2$
$B^+ \to f_0(980)(\to \pi^+\pi^-)K^+$	128	148,000	9.1 ± 1.1 [89]	$1.8 \pm 0.2 \pm 0.2$
$B^0 \to K^+\rho(770)^-(\to \pi^-\pi^0)$	398	647,000	9.9 ± 1.6 [89]	$1.4 \pm 0.1 \pm 0.2$
$B^+ \to K^*(892)^+(\to K^+\pi^0)\overline{K}^*(892)^0(\to K^-\pi^+)$ (L)	657	126,000	< 1.6 [89] [90]	$1.0 \pm 0.0 \pm 1.0$
$B^+ \to \overline{D}^0(\to K^+K^-)K^+$	638	250,000	1.4 ± 0.2 [23]	$0.9 \pm 0.0 \pm 0.1$
$B^0 \to K^+\pi^-\pi^0$ (NR)	236	2,307,000	35.6 ± 3.4 [89]	$0.8 \pm 0.1 \pm 0.1$
$B^+ \to K_1(1270)^0(\to \rho(770)^+(\to \pi^+\pi^0)K^-)K^+$	138	117,000	$< 5.0^\dagger$	$0.7 \pm 0.1 \pm 0.7$
$B^0 \to K_1(1270)^+(\to \rho(770)^0(\to \pi^+\pi^-)K^+)K^-$	125	117,000	$< 5.0^\dagger$	$0.6 \pm 0.1 \pm 0.6$
$B^0 \to K^*(892)^0(\to X)\gamma$	16	392,000	43.0 ± 4.0 [23]	$0.4 \pm 0.1 \pm 0.0$
$B^+ \to K^*(892)^+(\to K^+\pi^0)\overline{K}^*(892)^0(\to K^-\pi^+)$ (T)	28	126,000	< 15.8 [89]	$0.4 \pm 0.1 \pm 0.4$

L(T): Longitudinally (transversely) polarised; †: Conservative upper limit.

Table 5.8: *Potential sources of $B\overline{B}$ background for the 5-variable fit in the 892 window. MC efficiencies are calculated by applying RaSC to the relevant samples. For the number of events expected in the final on-resonance sample, the first error is statistical – from the statistical error on the MC efficiency and on $N_{B^\pm}$. The second error is systematic – from the systematic uncertainty on $N_{B^\pm}$ and on the branching ratio (BR). Where only an upper limit (UL) is available, $BR = UL/2 \pm UL/2$ is used.*

The ML method requires that N measurements of x have been made, $\{x_1, x_2, \ldots x_n\}$, and that the form of the PDF, $\mathcal{P}(x; \vec{\Theta})$, is known. It sets out to determine $\vec{\Theta}$. It

$B\overline{B}$ ***Background Source***	***Number Of Events To Pass RaSC***	***# MC Events Run Over***	***Branching Ratio*** ($\times 10^{-6}$)	***# Events Expected In Final On-res. Sample***
Combinatoric charged generics	2,234	5×10^8	$(5.0 \pm 2.5) \times 10^5$	$477.9 \pm 10.1 \pm 239.0$
Combinatoric neutral generics	710	5×10^8	$(5.0 \pm 2.5) \times 10^5$	$153.4 \pm 5.8 \pm 76.7$
$B^+ \to K^+K^-K^+$ Dalitz plot model, minus the $B^+ \to \phi(1020)K^+$ contribution				
Belle model B_0, sol. 1, minus ϕK [42]	1,561	85,190	21.1 ± 2.0 [89]	$89.4 \pm 2.2 \pm 8.6$
(Belle model B_0, sol. 2, minus ϕK [42]	879	47,464	21.1 ± 2.0 [89]	$90.4 \pm 3.0 \pm 8.7$)
(*BABAR* solution *A*, minus ϕK [88]	864	49,568	21.1 ± 2.0 [89]	$85.1 \pm 2.9 \pm 8.2$)
(*BABAR* solution *B*, minus ϕK [88]	834	48,172	21.1 ± 2.0 [89]	$84.5 \pm 2.9 \pm 8.1$)
$B^0 \to K^+K^-\pi^0$ (NR)	1,070	126,000	< 19.0 [23]	$18.7 \pm 0.6 \pm 18.7$
$B^0 \to K^*(892)^0(\to K^+\pi^-)\overline{K}^*(892)^0(\to K^-\pi^+)$ (L)	2,005	122,000	< 9.8 [89]	$18.6 \pm 0.4 \pm 18.6$
$B^+ \to f_0(980)(\to \pi^+\pi^-)K^+$	848	148,000	9.1 ± 1.1 [89]	$12.0 \pm 0.4 \pm 1.4$
$B^+ \to \overline{D}^0(\to K^+K^-)K^+$	8,608	250,000	1.4 ± 0.2 [23]	$11.5 \pm 0.1 \pm 1.9$
$B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$	64	302,000	189.3 ± 11.9 [23]	$9.3 \pm 1.2 \pm 0.6$
$B^+ \to \overline{D}^0(\to K^+K^-)\pi^+$ $\star$	42	20,000	19.4 ± 1.4 [23]	$9.4 \pm 1.5 \pm 0.7$
$B^0 \to K^+K^-K^0_S$ (NR)	687	220,000	12.4 ± 1.2 [89]	$8.9 \pm 0.3 \pm 0.8$
$B^+ \to \eta'(958)(\to \gamma\rho(770)^0(\to \pi^+\pi^-))K^+$	270	145,000	20.5 ± 1.1 [23] [89]	$8.8 \pm 0.5 \pm 0.5$
$B^0 \to K^+\rho(770)^-(\to \pi^-\pi^0)$	1,308	647,000	9.9 ± 1.6 [89]	$4.6 \pm 0.1 \pm 0.8$
$B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$	614	206,000	5.1 ± 0.9 [89]	$3.5 \pm 0.1 \pm 0.6$
$B^+ \to K^*(892)^+(\to K^+\pi^0)\overline{K}^*(892)^0(\to K^-\pi^+)$ (L)	1,988	126,000	< 1.6 [89] [90]	$2.9 \pm 0.1 \pm 2.9$
$B^0 \to K^+\pi^-\pi^0$ (NR)	747	2,307,000	35.6 ± 3.4 [89]	$2.7 \pm 0.1 \pm 0.3$
$B^+ \to \overline{D}^0(\to K^+\pi^-)K^+$	290	362,000	14.1 ± 2.3 [23]	$2.6 \pm 0.2 \pm 0.4$
$B^+ \to K^*(892)^+(\to K^+\pi^0)\overline{K}^*(892)^0(\to K^-\pi^+)$ (T)	138	126,000	< 15.8 [89]	$2.0 \pm 0.2 \pm 2.0$
$B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^*(892)^+(\to K^+\pi^0)$ (L)	222	201,000	3.4 ± 1.3 [23] [89]	$0.9 \pm 0.1 \pm 0.3$
$B^+ \to J/\psi(\to K^+K^-)K^+$	4,392	280,000	0.2 ± 0.0 [23]	$0.9 \pm 0.0 \pm 0.1$
$B^+ \to K^+\pi^+\pi^-$ (NR)	1,621	1,298,985	2.9 ± 1.1 [89]	$0.8 \pm 0.0 \pm 0.3$
$B^+ \to K_1(1270)^0(\to \rho(770)^+(\to \pi^+\pi^0)K^-)K^+$	160	117,000	$< 5.0^\dagger$	$0.8 \pm 0.1 \pm 0.8$
$B^0 \to K_1(1270)^+(\to \rho(770)^0(\to \pi^+\pi^-)K^+)K^-$	154	117,000	$< 5.0^\dagger$	$0.8 \pm 0.1 \pm 0.8$
$B^0 \to K^+\pi^-$	323	1,873,500	18.9 ± 0.7 [89]	$0.8 \pm 0.0 \pm 0.0$
$B^+ \to \phi(1020)(\to K^+K^-)K^+$	111	163,000	4.4 ± 0.3 [23] [89]	$0.7 \pm 0.1 \pm 0.1$

L(T): Longitudinally (transversely) polarised; †: Conservative upper limit.

$\star$: Final state $K^+K^-\pi^+$, but zero interference assumed here.

Table 5.9: *Potential sources of $B\overline{B}$ background for the 3-variable fit in the 1430 window.*

does so by choosing values of $\vec{\Theta}$ such as to maximise the probability of obtaining the observed measurements, according to $\mathcal{P}(x;\vec{\Theta})$. This probability is given by the

$B\overline{B}$ ***Background Source***	***Number Of Events To Pass RaSC***	***# MC Events Run Over***	***Branching Ratio*** ($\times 10^{-6}$)	***# Events Expected In Final On-res. Sample***
$B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$				
$m_{K\pi}$ modelled as LASS	9,624	128,150	$< 3.8^{\ddagger}$	$32.9 \pm 0.3 \pm 32.9$
$m_{K\pi}$ modelled as Breit-Wigner	3,771	117,000	$< 3.8^{\ddagger}$	$14.1 \pm 0.2 \pm 14.1$
$B^+ \to K^+K^-\pi^+$ (NR)	15,211	1,323,000	< 6.3 [23]	$8.4 \pm 0.1 \pm 8.4$

‡: Estimated using Eq. (C.1).

Table 5.10: *Final state $K^+K^-\pi^+$ channels that are a potential source of $B\overline{B}$ background for the 5-variable fit in the 892 window.*

$B\overline{B}$ ***Background Source***	***Number Of Events To Pass RaSC***	***# MC Events Run Over***	***Branching Ratio*** ($\times 10^{-7}$)	***# Events Expected In Final On-res. Sample***
$B^+ \to K^+K^-\pi^+$(NR)	55,203	1,323,000	< 63.0 [23]	$30.4 \pm 0.1 \pm 30.4$
$B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$	17,845	242,000	$< 14.5^{\S}$	$12.4 \pm 0.1 \pm 12.4$
$B^+ \to \overline{K}^*_2(1430)^0(\to K^-\pi^+)K^+$	19,428	97,000	$< 1.2^{\ddagger}$	$2.6 \pm 0.0 \pm 2.6$
$B^+ \to \overline{K}^*(1680)^0(\to K^-\pi^+)K^+$	19,422	117,000	$< 1.1^{\ddagger}$	$2.4 \pm 0.0 \pm 2.4$

§: From unpublished *cut and count* analysis [6].

‡: Estimated using Eq. (C.1).

Table 5.11: *Final state $K^+K^-\pi^+$ channels that are expected to be present in the 1430 window.*

likelihood function:

$$\mathcal{L}(\vec{\Theta}) = \mathcal{P}(x_1;\vec{\Theta})\,\mathcal{P}(x_2;\vec{\Theta})\ldots\mathcal{P}(x_n;\vec{\Theta})\,dx^N. \tag{5.3}$$

It is convenient to write $\mathcal{L}(\vec{\Theta})$ as

$$\mathcal{L}(\vec{\Theta}) = \prod_{i=1}^{N}\mathcal{P}(x_i;\vec{\Theta}), \tag{5.4}$$

where the proportionality constant dx^N has been omitted.

More generally, $\mathcal{P}$ can be a combination of a number of normalised PDFs, e.g. $\mathcal{Q}$, $\mathcal{R}$ and $\mathcal{S}$, which are functions of uncorrelated variables x, y and z, say. Then $\mathcal{P}$ becomes

$$\mathcal{P}(\vec{x};\vec{\Theta}) = \mathcal{P}(x,y,z;\vec{\Theta}) = \mathcal{Q}(x;\vec{\theta}_{\mathcal{Q}})\,\mathcal{R}(y;\vec{\theta}_{\mathcal{R}})\,\mathcal{S}(z;\vec{\theta}_{\mathcal{S}}), \tag{5.5}$$

where $\vec{\theta}_{\mathcal{Q},\mathcal{R},\mathcal{S}}$ are subsets of $\vec{\Theta}$. For a fit to M hypotheses the weighted sum is taken:

$$\mathcal{P}(\vec{x};\vec{\Theta}) = \sum_{j=1}^{M} f_j\,\mathcal{Q}_j(x;\vec{\theta}_{\mathcal{Q}})\,\mathcal{R}_j(y;\vec{\theta}_{\mathcal{R}})\,\mathcal{S}_j(z;\vec{\theta}_{\mathcal{S}}); \tag{5.6}$$

f_j is the fraction of events in the sample belonging to hypothesis j, and

$$\sum_{j=1}^{M} f_j = 1. \tag{5.7}$$

The likelihood function is now written

$$\mathcal{L}(\vec{\Theta}) = \prod_{i=1}^{N}\left(\sum_{j=1}^{M} f_j\,\mathcal{Q}_j(x_i;\vec{\theta}_{\mathcal{Q}})\,\mathcal{R}_j(y_i;\vec{\theta}_{\mathcal{R}})\,\mathcal{S}_j(z_i;\vec{\theta}_{\mathcal{S}})\right). \tag{5.8}$$

One of the most common data analysis scenarios (and the one faced in this chapter) is where *yields* are the required output of a ML fit (the number of events in each of the M hypotheses). The ML method can be used to achieve this, but Poissonian fluctuations in the number of events are not taken into account.

5.6.1 The extended maximum likelihood formalism

The *extended* likelihood function is given by

$$\mathcal{L}(\nu,\vec{\Theta}) = \frac{\nu^N}{N!}e^{-\nu}\prod_{i=1}^{N}\mathcal{P}(\vec{x}_i;\vec{\Theta}) = \frac{e^{-\nu}}{N!}\prod_{i=1}^{N}\nu\mathcal{P}(\vec{x}_i;\vec{\Theta}), \tag{5.9}$$

where

$$\mathcal{P}(\vec{x}_i;\vec{\Theta}) = \sum_{j=1}^{M}\left(\prod_{l=1}^{V}\mathcal{P}_j^l(x_i^l;\vec{\Theta})\right). \tag{5.10}$$

V is the number of discriminating variables used. Eq. (5.9) is the product of Eq. (5.4) (the normal likelihood function) and the total event yield. This yield is in the form of a pseudo-constant factor that represents a Poissonian distribution with mean ν. The sample size, N, is now defined to be part of the result of the experiment, as are the yields, n_j, for each of the hypotheses.

Assuming no functional relation between ν and $\vec{\Theta}$, it can easily be shown that the estimator

$$\hat{\nu} = N\left(=\sum_{j=1}^{M} n_j\right), \tag{5.11}$$

as one would expect, and that the function to maximise (dropping constant factors) is

$$\mathcal{L}(\vec{n},\vec{\Theta}) = \exp\left(-\sum_{k=1}^{M} n_k\right)\prod_{i=1}^{N}\left(\sum_{j=1}^{M} n_j\left(\prod_{l=1}^{V}\mathcal{P}_j^l(x_i^l;\vec{\Theta})\right)\right). \tag{5.12}$$

With $\mathcal{L}$ now a function of $\vec{n}$, in addition to $\vec{\Theta}$, and with components of $\vec{n}$ considered to be random variables, the additional source of statistical fluctuation that was not accounted for by the normal ML is now considered.

It is sometimes necessary (as with the analysis described in this chapter) to fix certain components of $\vec{\Theta}$ and $\vec{n}$. The values to which they are fixed are obtained from separate data samples, MC for example.

5.6.2 Fitting tools

The dedicated fitting package used to maximise user-defined likelihood functions for the analysis described in this chapter is `Minuit` [92,93]. `Minuit`'s `MIGRAD` routine is used to find the function maximum[3] and return the parameter values. `MIGRAD` also

[3] `Minuit` actually finds the minimum value of the negative $\log_e$ of the likelihood function. This is computationally more convenient since the product over N events becomes a summation.

makes a first attempt at calculating parameter errors. The `HESSE` and `MINOS` routines perform more precise calculations of parameter errors, as discussed in Section 6.2.1.

The modelling package `RooFit` [94] provides a high-level user interface to `Minuit` via the `Root` class `TMinuit`. Likelihood function normalisation is carried out automatically by `RooFit` before the function is passed to `Minuit`. Normalisation can be to unity for the normal ML or to the Poisonially modelled total number of events for the extended version. Numerous PDF shapes are available in the `RooFit` package. These can be combined by addition, multiplication and convolution to form the desired fit model. PDF shapes can also be added by the user. `RooFit` makes use of further `Root` classes for histogram plotting and data manipulation.

`CharmlessFitter`, in addition to being used to implement the selection cuts described in Section 5.3.3, is also used to provide a further level of simplification between the user and various common fitting tasks. It builds upon the `RooFit` framework and is the most heavily used package in the analysis described in this chapter.

5.7 Signal and $q\overline{q}$ bkgd. PDFs for the 5-var. fit in the 892 window

We first construct a fit model with two fit components (hypotheses) for which PDFs are constructed: signal, S, and continuum background, $q\overline{q}$. Toy tests – referred to in Section 5.5.2.5, and described in Section 5.10.4 – are then conducted to determine whether $B\overline{B}$ background PDFs are required.

5.7.1 Splitting the signal component

For the way in which we implement the ML method, it is assumed that components of $\vec{x}$ (see Eq. (5.12)) are uncorrelated. Correlations may bring about biases to the

fit results.

For the 5-variable fit in which the variable $\cos\theta_H$ is employed, the decision is taken to split S into truth matched and SXF sub-components, S_{TRU} and S_{SXF}. This approach is chosen because of a large correlation for un-split S between the discriminating variables m_{ES} and $\cos\theta_H$ in signal MC. This is illustrated in Figure 5.4. The correlation is brought about due to SXF events being concentrated towards $\cos\theta_H = 1$ (also shown in Figure 5.4), and having, on average, lower m_{ES} values than their truth matched counterparts.

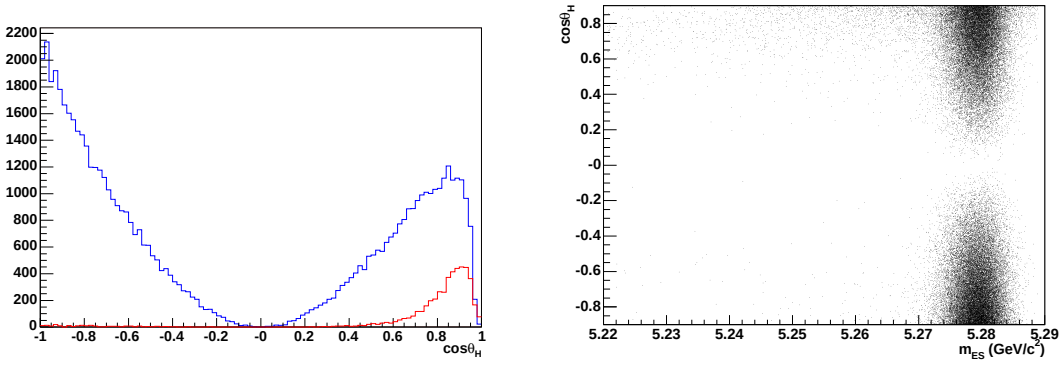

Figure 5.4: *Left:* $\cos\theta_H$ *distributions for* $B^+ \rightarrow \overline{K}^*(892)^0(\rightarrow K^-\pi^+)K^+$ *signal MC (with RaSC applied). The blue (red) histogram represents truth matched (SXF) events. It can be seen that SXF events tend towards high* $\cos\theta_H$*. Right: for an un-split signal component this brings about a large and undesired correlation between* $\cos\theta_H$ *and another fit variable,* m_{ES}*.*

Table 5.12 shows the correlations between the five fit variables for S_{TRU}, S_{SXF} and $q\overline{q}$. Correlations, $\rho_{x,y}$, are calculated using

$$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y}, \tag{5.13}$$

where x and y are fit variables, $cov(x,y)$ is their covariance, and σ_x and σ_y are their respective variances. The values obtained are shown to be at an acceptable level (toy tests, described in Section 5.10.3, are conducted to determine whether any biases

are present due to these small correlations, and if so how large). Correlation plots are presented in Appendix E.

	m_{ES}	ΔE	$\mathcal{F}$	$m_{K\pi}$	$\cos\theta_H$
m_{ES}	1	−0.010, 0.006	0.010, 0.014	0.000, −0.001	−0.008, −0.007
ΔE	−0.114, 0.024	1	−0.017, −0.015	0.000, −0.000	−0.025, −0.013
$\mathcal{F}$	0.003, −0.063	0.001, −0.041	1	0.007, −0.003	0.000, −0.024
$m_{K\pi}$	0.005, −0.054	0.004, 0.009	−0.005, 0.000	1	0.044, 0.042
$\cos\theta_H$	−0.012, 0.016	0.018, 0.104	−0.021, −0.044	0.004, 0.006	1

Table 5.12: *Correlations (calculated using Eq. (5.13)) between variables used in the 5-variable fit in the 892 window. The values in the bottom left section of the table are for signal MC. For each cell, the first value is for S_{TRU}, the second value is for S_{SXF}. The values in the top right section of the table are for $q\overline{q}$ background. For each cell, the first value is obtained using the off-resonance sample, the second value is obtained using the on-resonance sideband sample (these samples are discussed in the main text (Section 5.7.2)).*

The functional forms of $\mathcal{P}_j^l$ (from Eq. (5.12)), where $l = m_{ES}$, ΔE, $\mathcal{F}$, $m_{K\pi}$, $\cos\theta_H$ and $j = S$, $q\overline{q}$ are presented in the remainder of this section. Splitting the S component gives

$$\mathcal{P}_S = (1 - f_{SXF})\mathcal{P}_{S_{TRU}} + f_{SXF}\mathcal{P}_{S_{SXF}}, \tag{5.14}$$

where f_{SXF} is the SXF fraction.

For the 3-variable fits $\cos\theta_H$ is not used and as such the signal component is not split.

5.7.2 Data samples

The functional forms and parameters of $\mathcal{P}_{S_{i=TRU,SXF}}^{l=m_{ES},\Delta E,\mathcal{F},m_{K\pi},\cos\theta_H}$ are determined by performing unbinned, one-dimensional fits to a MC sample in which $B^+ \to \overline{K}^*(892)^0$ $(\to K^-\pi^+)K^+$ events are simulated. The fits are to 60,473 truth matched events and 3,918 SXF events – the number of events remaining after RaSC have been applied to an original sample containing 242,000 generated events ($f_{SXF} = 6.08\%$).

$q\overline{q}$ PDFs are constructed by performing fits to an off-resonance data sample that has been subjected to RaSC. This sample contains 4,077 events. The PDF functional forms are checked using an on-resonance sideband data sample, containing 30,781 events after RaSC.

In the fit to the final on-resonance sample, the functional forms of all PDFs are fixed. Ideally, all PDF parameters would be floated, as would f_{SXF}. However, due to the very small number of signal events expected, to ensure a stable fit it is necessary to fix PDF parameters for the signal component. f_{SXF} must also be fixed.

Fit parameter values, and their errors, are presented in Appendix D.

5.7.3 m_{ES} PDFs

The *Crystal Ball* function (CB) [95] – a Gaussian modified by a power law tail:

$$\mathrm{CB}(m_{ES};\mu,\sigma,\alpha,n) = \frac{1}{N}\times$$

$$\begin{cases} \exp\left(-\frac{1}{2}(m_{ES}-\mu)^2/\sigma^2\right) & , \text{ if } (m_{ES}-\mu)/\sigma \geq -|\alpha| \text{ and } \alpha \geq 0 \\ (n/|\alpha|)^n \exp(-|\alpha|^2/2) \times & \\ (-(m_{ES}-\mu)/\sigma + n/|\alpha| - |\alpha|)^{-n} & , \text{ if } (m_{ES}-\mu)/\sigma < -|\alpha| \text{ and } \alpha \geq 0 \\ \exp\left(-\frac{1}{2}(\mu-m_{ES})^2/\sigma^2\right) & , \text{ if } (\mu-m_{ES})/\sigma \geq -|\alpha| \text{ and } \alpha < 0 \\ (n/|\alpha|)^n \exp(-|\alpha|^2/2) \times & \\ (-(\mu-m_{ES})/\sigma + n/|\alpha| - |\alpha|)^{-n} & , \text{ if } (\mu-m_{ES})/\sigma < -|\alpha| \text{ and } \alpha < 0 \end{cases} \tag{5.15}$$

– is fitted to both the truth matched and SXF signal distributions. The *Argus* function (AF) [96],

$$\mathrm{AF}(m_{ES}; m_0, \xi) = \frac{1}{N} m_{ES} \sqrt{1-(m_{ES}/m_0)^2} \exp\left(\xi - \xi(m_{ES}/m_0)^2\right), \qquad (5.16)$$

is fitted to the $q\overline{q}$ background distribution. The Argus *endpoint*, m_0, is fixed at half the $\Upsilon(4S)$ mass (5.29 GeV/c^2). These fits are illustrated in Figure 5.5.

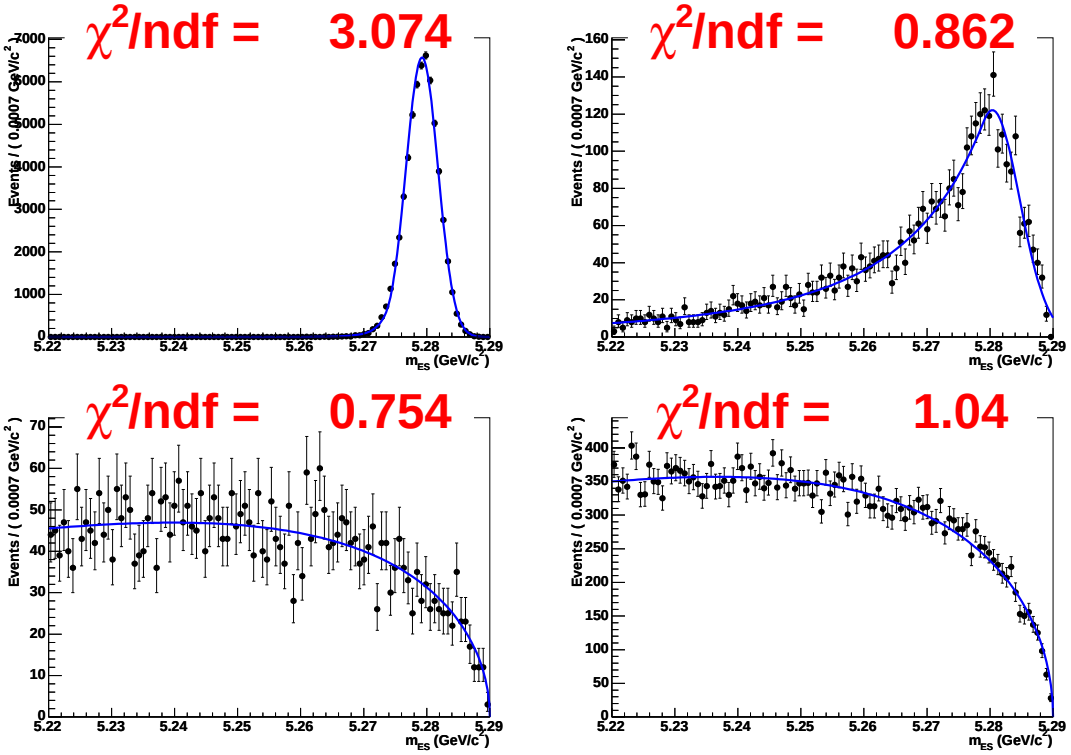

Figure 5.5: *m_{ES} distributions after the application of RaSC (black points with error bars) with fitted PDFs superimposed (blue, solid line). The PDF functional forms and the data samples used are described in the main text (Section 5.7). Signal PDF parameter values are given in Tables D.1 and D.2; $q\overline{q}$ background PDF parameter values are given in Table D.3. Top left: truth matched signal; top right: SXF; bottom: $q\overline{q}$ background (bottom left: using off-resonance sample; bottom right: using on-resonance sideband sample). These PDFs are used for the 5-variable fit in the 892 window.*

5.7.4 ΔE PDFs

Truth matched signal events are fitted with a double Gaussian (DG):

$$\begin{aligned} \mathrm{DG}(\Delta E; \mu_1, \sigma_1, \mu_2, \sigma_2, f_2) &= \frac{1}{N}\left((1-f_2)\,\exp\left(-\frac{(\Delta E-\mu_1)^2}{2\sigma_1^2}\right)+\right. \\ &\left. f_2\,\exp\left(-\frac{(\Delta E-\mu_2)^2}{2\sigma_2^2}\right)\right). \end{aligned} \tag{5.17}$$

SXF events are fitted with a first order polynomial, as are $q\overline{q}$ background events. These fits are illustrated in Figure 5.6.

5.7.5 $\mathcal{F}$ PDFs

Double Gaussians are fitted to signal (truth matched and SXF) and $q\overline{q}$ background distributions, as illustrated in Figure 5.7.

5.7.6 $m_{K\pi}$ PDFs

A Breit-Wigner is fitted to the truth matched signal distribution. SXF is fitted with a Gaussian (centred at approximately the $K^*(892)^0$ mass) plus a first order polynomial. A Breit-Wigner and a *Voigtian* (a Breit-Wigner convolved with a Gaussian) – both plus a first order polynomial – are also tried, but the Gaussian gives the best fit. The $K^*(892)^0$ is also present in continuum events. As such, a Breit-Wigner plus a first order polynomial is fitted to the $q\overline{q}$ background distribution. These fits are illustrated in Figure 5.8.

5.7.7 $\cos\theta_H$ PDFs

For constant RaSC efficiency as a function of $\cos\theta_H$, a second order polynomial is expected to describe the $\cos\theta_H$ distribution for truth matched signal events. However, this efficiency is not constant, as illustrated in Figure 5.9. Attempts are made

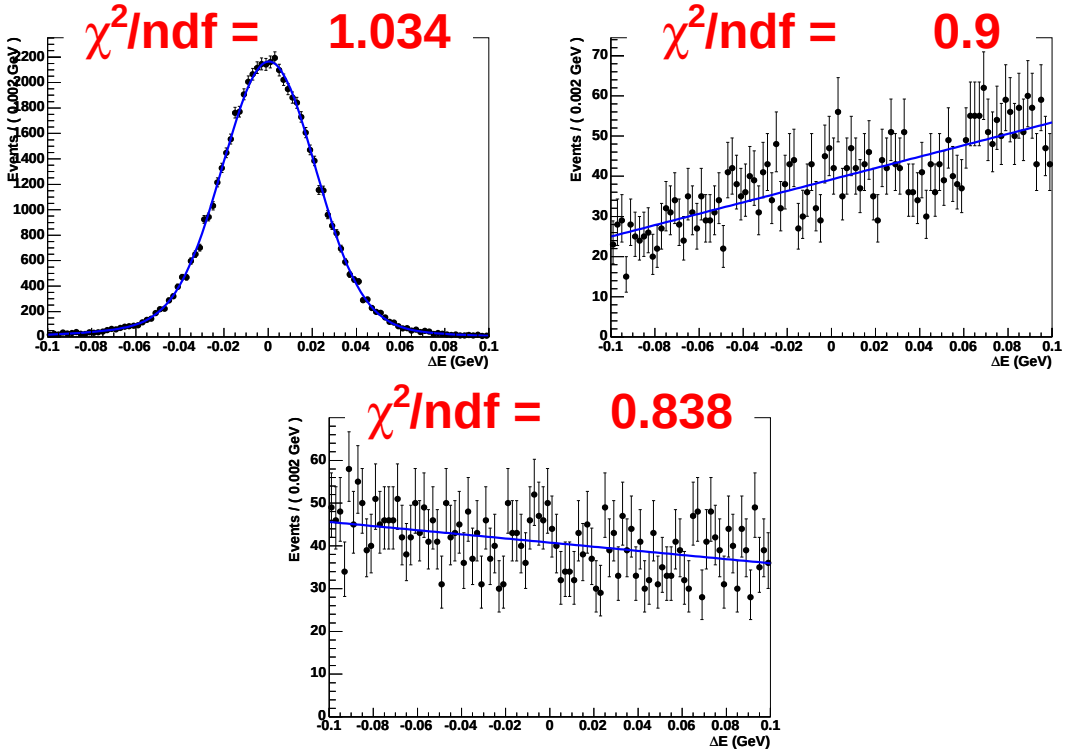

Figure 5.6: *ΔE distributions after the application of RaSC (black points with error bars) with fitted PDFs superimposed (blue, solid line). The PDF functional forms and the data samples used are described in the main text (Section 5.7). Signal PDF parameter values are given in Tables D.1 and D.2; $q\overline{q}$ background PDF parameter values are given in Table D.3. Top left: truth matched signal; top right: SXF; bottom: $q\overline{q}$ background (using off-resonance sample). These PDFs are used for the 5-variable fit in the 892 window.*

at fitting numerous PDF shapes to the truth matched signal distribution, the most successful by far being the function $e^{Poly8(\cos\theta_H)} + C$. $Poly8(\cos\theta_H)$ is an eighth order polynomial in $\cos\theta_H$ and C is a real scalar. $e^{Poly5(\cos\theta_H)} + C$ is fitted to the SXF distribution, where $Poly5(\cos\theta_H)$ is a fifth order polynomial. Continuum events are fitted with a second order polynomial. Fitted PDFs to $\cos\theta_H$ distributions are illustrated in Figure 5.10.

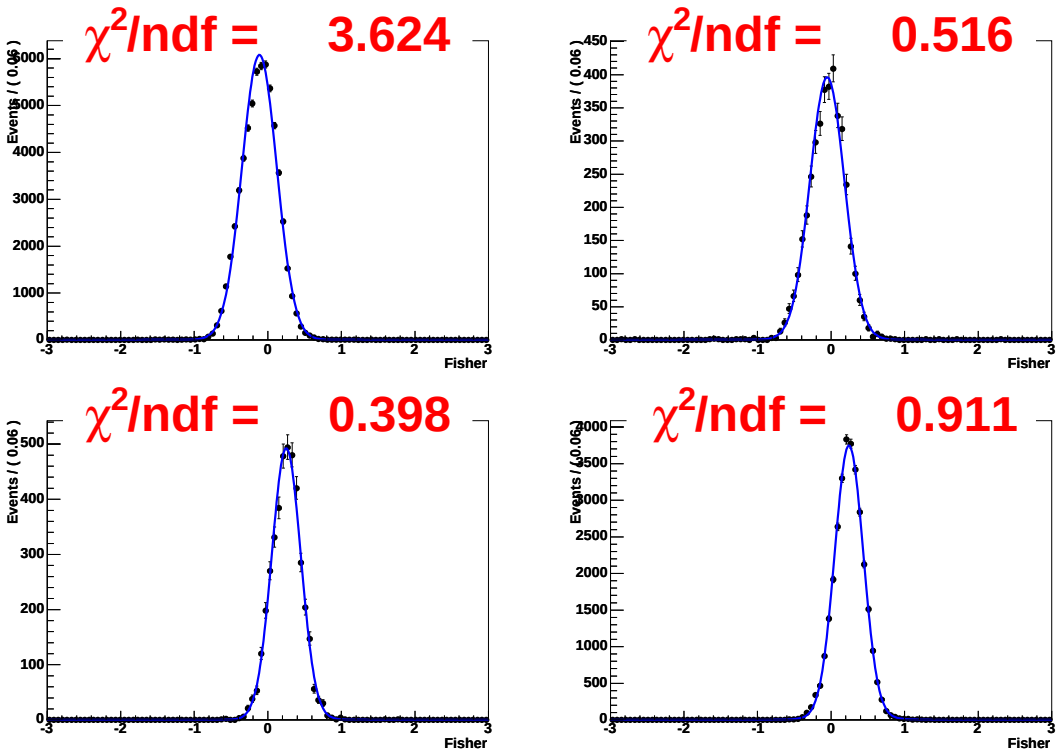

Figure 5.7: *$\mathcal{F}$ distributions after the application of RaSC (black points with error bars) with fitted PDFs superimposed (blue, solid line). The PDF functional forms and the data samples used are described in the main text (Section 5.7). Signal PDF parameter values are given in Tables D.1 and D.2; $q\overline{q}$ background PDF parameter values are given in Table D.3. Top left: truth matched signal; top right: SXF; bottom: $q\overline{q}$ background (bottom left: using off-resonance sample; bottom right: using on-resonance sideband sample). These PDFs are used for the 5-variable fit in the 892 window.*

5.8 Signal and $q\overline{q}$ bkgd. PDFs for the 3-var. fit in the 892 window

Signal PDFs are constructed by fitting to $68,747$ $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ signal MC events – the number remaining after RaSC have been applied to an original sample containing $242,000$ generated events. An off-resonance sample containing

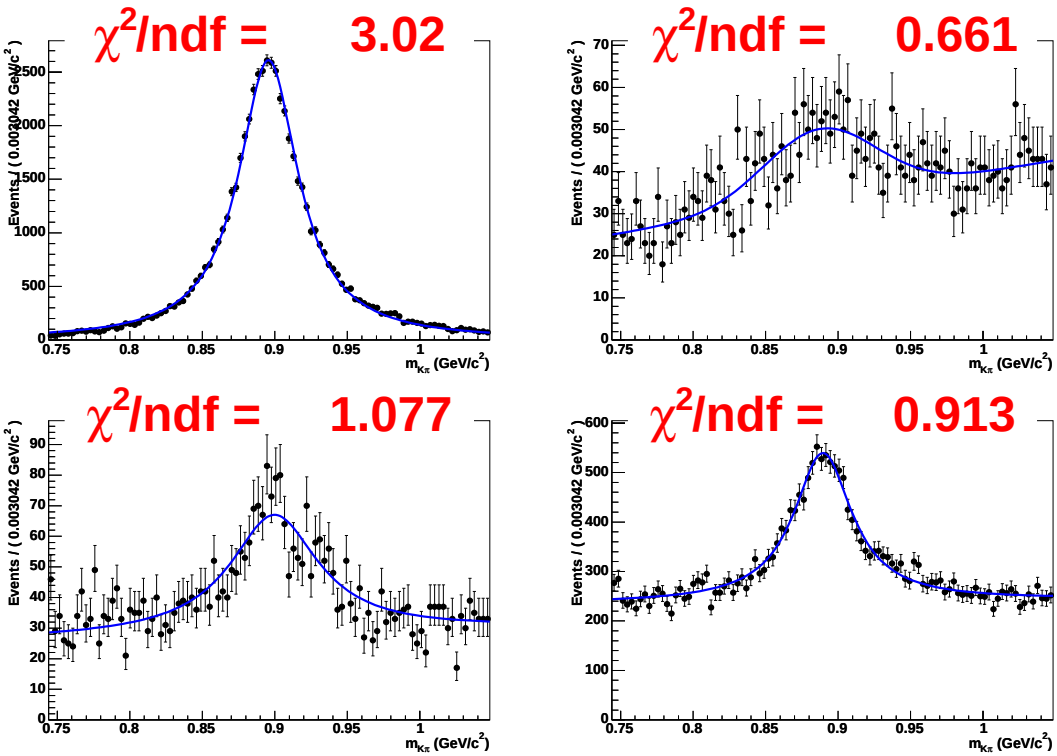

Figure 5.8: *$m_{K\pi}$ distributions after the application of RaSC (black points with error bars) with fitted PDFs superimposed (blue, solid line). The PDF functional forms and the data samples used are described in the main text (Section 5.7). Signal PDF parameter values are given in Tables D.1 and D.2; $q\overline{q}$ background PDF parameter values are given in Table D.3. Top left: truth matched signal; top right: SXF; bottom: $q\overline{q}$ background (bottom left: using off-resonance sample; bottom right: using on-resonance sideband sample). These PDFs are used for the 5-variable fit in the 892 window.*

4,239 events and an on-resonance sideband sample ($0.10 < \Delta E/\,\mathrm{GeV} < 0.35$) containing 31,912 events – both after the application of RaSC – are used to construct $q\overline{q}$ PDFs. Table 5.13 lists the functional forms of the PDFs. PDF parameters are given in Appendix D. Figure 5.11 illustrates fit variable distributions with PDFs overlaid. Table 5.14 shows correlation values between fit variables; correlation plots are presented in Appendix E.

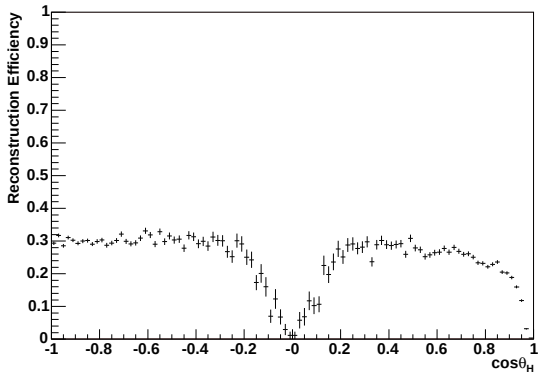

Figure 5.9: *Efficiency of RaSC, as a function of* $\cos\theta_H$.

	S	$q\overline{q}$
m_{ES}	Crystal Ball	Argus
ΔE	Double Gaussian	Linear
$\mathcal{F}$	Double Gaussian	Double Gaussian

Table 5.13: *PDF shapes for the 3-variable fit in the 892 window.*

	m_{ES}	ΔE	$\mathcal{F}$
m_{ES}	1	-0.013, 0.008	0.013, 0.012
ΔE	-0.055	1	-0.019, -0.017
$\mathcal{F}$	-0.036	-0.000	1

Table 5.14: *Correlations (calculated using Eq. (5.13)) between variables used in the 3-variable fit in the 892 window. The values in the bottom left section of the table are for the S component. The values in the top right section of the table are for the* $q\overline{q}$ *component (the first value is obtained using the off-resonance sample, the second using the on-resonance sideband sample).*

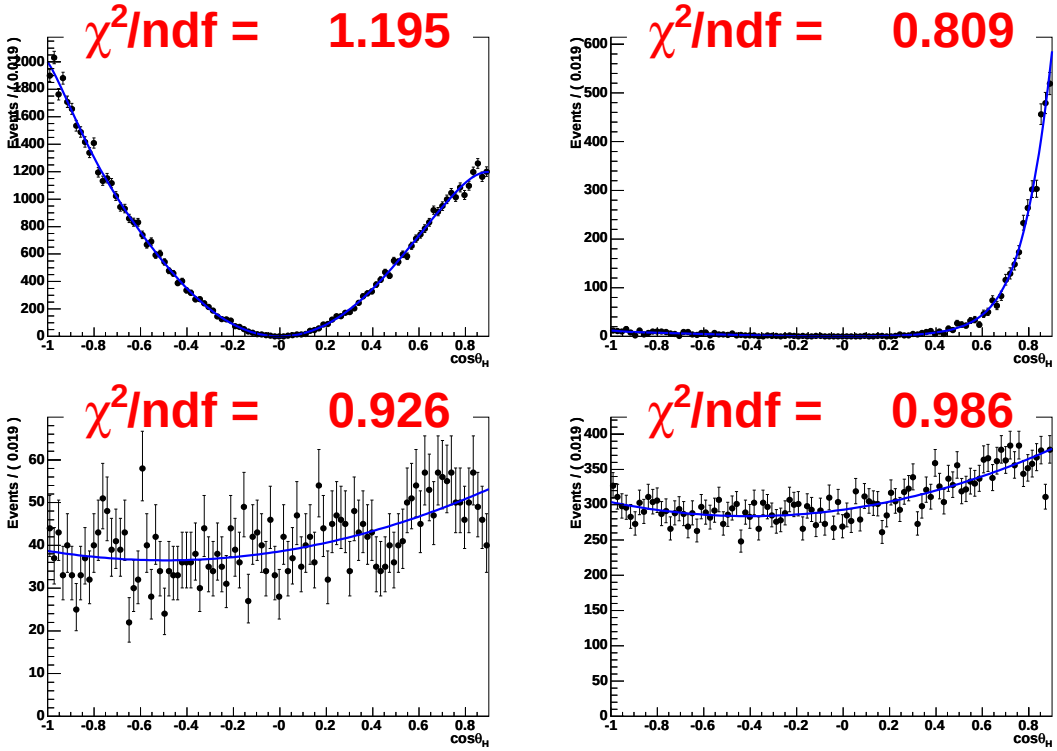

Figure 5.10: $\cos\theta_H$ *distributions after the application of RaSC (black points with error bars) with fitted PDFs superimposed (blue, solid line). The PDF functional forms and the data samples used are described in the main text (Section 5.7). Signal PDF parameter values are given in Tables D.1 and D.2;* $q\overline{q}$ *background PDF parameter values are given in Table D.3. Top left: truth matched signal; top right: SXF; bottom:* $q\overline{q}$ *background (bottom left: using off-resonance sample; bottom right: using on-resonance sideband sample). These PDFs are used for the 5-variable fit in the 892 window.*

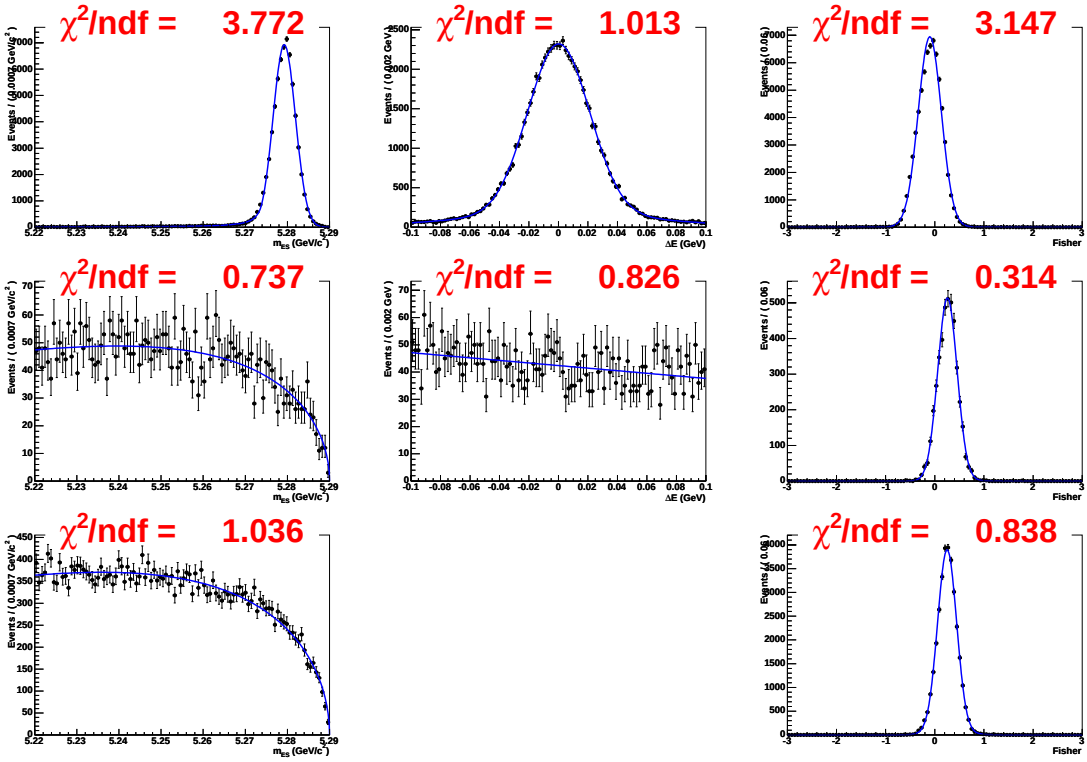

Figure 5.11: *Fit variable distributions and fitted PDFs for the 3-variable fit in the 892 window. Left column: m_{ES}; middle column: ΔE; right column: $\mathcal{F}$. Top row: S; second row: $q\overline{q}$ (off-resonance sample); bottom row: $q\overline{q}$ (on-resonance sideband sample). The functional forms of the PDFs are listed in Table 5.13, whilst the PDF parameter values are listed in Tables D.7 (signal) and D.8 ($q\overline{q}$ background). The data samples used are discussed in the main text (Section 5.8).*

5.9 Signal and $q\overline{q}$ bkgd. PDFs for the 3-var. fit in the 1430 window

Figure 5.12 compares fit variable distributions for two separate MC samples – both containing simulated $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ events that have passed RaSC. For one sample the $m_{K\pi}$ lineshape is modelled as a Breit-Wigner. For the other, this lineshape is modelled as the LASS shape (see Appendix F for LASS parameters). It can be seen that the shapes of the fit variable distributions are very similar, regardless of $m_{K\pi}$ lineshape. The sample in which $m_{K\pi}$ is modelled as the LASS shape is used to construct PDFs for the S component. This sample contains $26,055$ events.

$q\overline{q}$ PDFs are constructed by performing fits to an off-resonance sample, which contains $4,792$ events after the application of RaSC. An on-resonance sideband sample ($0.10 < \Delta E/\,\mathrm{GeV} < 0.35$) that has been subjected to RaSC is also used to ensure that the PDF functional forms are in agreement for on- and off-resonance data. This sample consists of $37,773$ events.

PDF functional forms and parameters are listed in Table 5.15 and Appendix D respectively. Figure 5.13 illustrates fit variable distributions with overlaid PDFs. Table 5.16 and Appendix E present correlation values and plots respectively.

	S	$q\overline{q}$
m_{ES}	Crystal Ball	Argus
ΔE	Double Gaussian	Linear
$\mathcal{F}$	Gaussian	Double Gaussian

Table 5.15: *PDF shapes for the 3-variable fit in the 1430 window.*

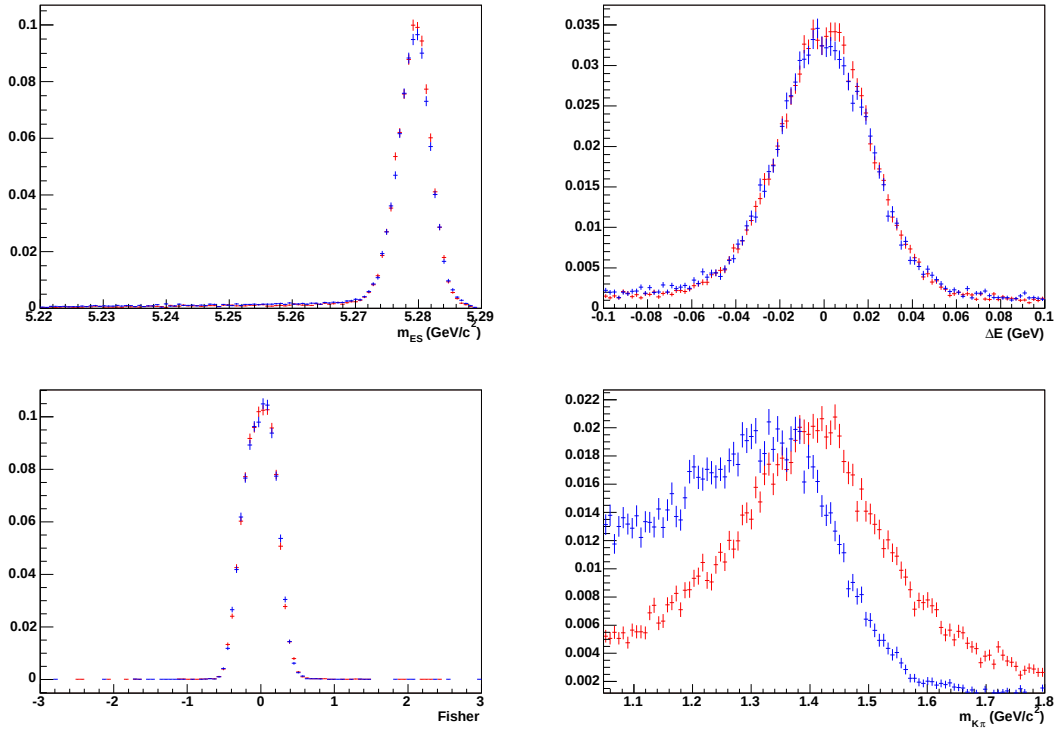

Figure 5.12: *m_{ES} (top left), ΔE (top right), $\mathcal{F}$ (bottom left) and $m_{K\pi}$ (bottom right) distributions for MC events simulating the decay $B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$, after the application of RaSC. The red (blue) points represent a sample in which the $m_{K\pi}$ lineshape is modelled as Breit-Wigner (LASS).*

5.10 Tests and validations for the 5-var. fit in the 892 window

The PDFs constructed in Section 5.7 are combined according to Eq. (5.12) to form a prototype fit model. Before applying the fit model to the final on-resonance sample, it must first be verified that it behaves as we would expect and does not return biased results. The effect of $B\overline{B}$ background on the behaviour of the fit model is also studied, with the possibility that further components shall be added to account for this type of background.

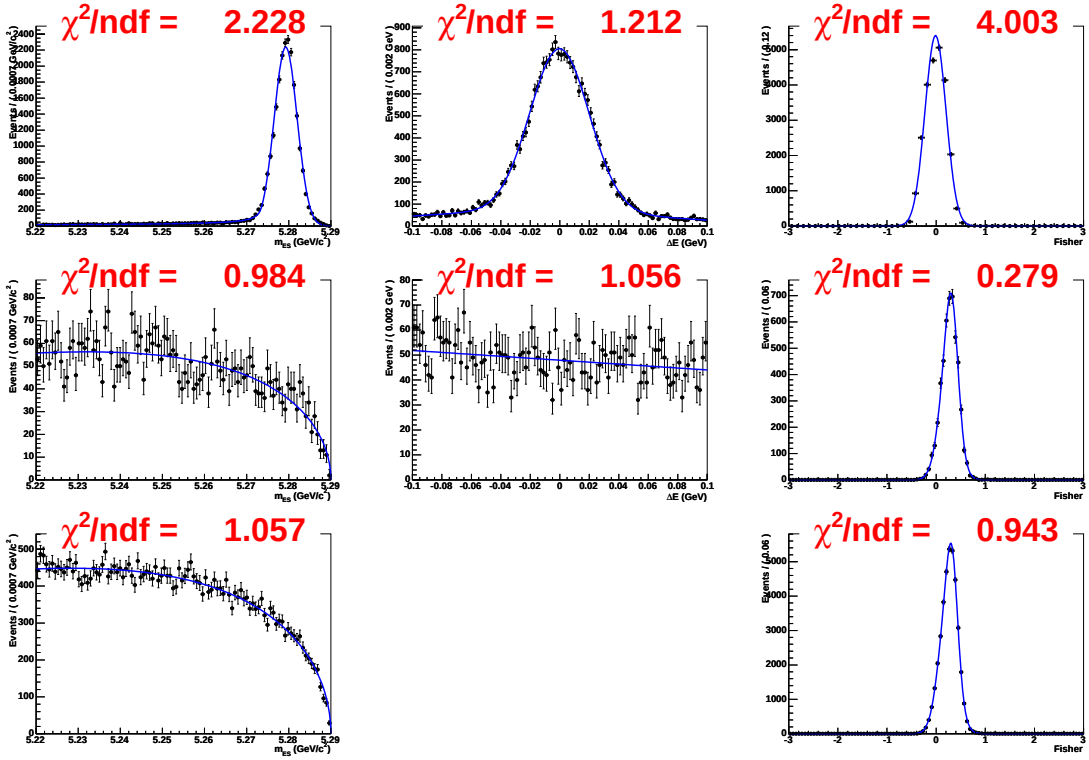

Figure 5.13: *Fit variable distributions and fitted PDFs for the 3-variable fit in the 1430 window. Left column: m_{ES}; middle column: ΔE; right column: $\mathcal{F}$. Top row: S; second row: $q\overline{q}$ (off-resonance sample); bottom row: $q\overline{q}$ (on-resonance sideband sample). The functional forms of the PDFs are listed in Table 5.15, whilst the PDF parameter values are listed in Tables D.10 (signal) and D.11 ($q\overline{q}$ background). The data samples used are discussed in the main text (Section 5.9).*

5.10.1 Toy MC and toy tests

In addition to being used for fitting, PDFs are also used to generate *toy MC*. The Von Neumann accept/reject algorithm [97] is one example of a technique used to generate toy MC events. N random numbers are thrown determining where in N-dimensional space the event is to be generated (for the 5-variable fit, $N = 5$). The probability that the event will be accepted at this point in variable space, $\vec{x}$, is

$$P_{Accept}(\vec{x}) = \frac{\mathcal{P}(\vec{x})}{\mathcal{P}_{Max}}. \tag{5.18}$$

	m_{ES}	ΔE	$\mathcal{F}$
m_{ES}	1	0.022, −0.004	−0.000, 0.017
ΔE	0.023, −0.007	1	−0.022, 0.001
$\mathcal{F}$	−0.054, −0.030	−0.007, −0.003	1

Table 5.16: *Correlations (calculated using Eq. (5.13)) between variables used in the 3-variable fit in the 1430 window. The values in the bottom left section of the table are for signal MC (the first value is obtained using MC in which the $m_{K\pi}$ lineshape is modelled as LASS, the second value is obtained using MC in which the $m_{K\pi}$ lineshape is modelled as Breit-Wigner). The values in the top right section of the table are for $q\overline{q}$ background (the first value is obtained using the off-resonance sample, the second using the on-resonance sideband sample).*

$\mathcal{P}(\vec{x})$ is the value of the PDF at position $\vec{x}$; $\mathcal{P}_{Max}$ is the maximum value of the PDF in the available variable space. A further random number is generated and if it is less than P_{Accept} then the event is accepted, otherwise it is rejected. The process is iterated until the requested number of events have been accepted. There are other techniques, the most efficient of which is chosen automatically within the `RooFit` framework. The user simply passes the fit model to `RooFit` (via `CharmlessFitter`). Sampling datasets from PDFs in this way allows us to generate similar but statistically independent samples on which to exercise the fit.

A poorly constructed fit model or a lack of statistics can bring about fit instabilities and biases. These must be removed, or at least, in the case of biases, accounted for. *Toy tests* are employed to determine whether biases or instabilities are present. A toy test consists of numerous experiments. One experiment is defined to be the process of acquiring an independent dataset and applying the multivariate fit to extract

floated PDF parameter values and, in the case of the extended ML, floated yields. Samples from one experiment to the next differ only due to statistical fluctuations. For each experiment i, a bias (b_i) and a pull (p_i) can be calculated for each floated parameter or yield:

$$b_i = f_i - t, \tag{5.19}$$

$$p_i = \frac{b_i}{(\sigma_f)_i}. \tag{5.20}$$

t is the true value, f_i is the value obtained from the fit and $(\sigma_f)_i$ is the error on f_i. For a stable fit model and a large enough number of experiments, the *residual distribution* and the *pull distribution* that result in histogramming b_i and p_i respectively are expected to be Gaussian. These Gaussians should be centred at zero for an unbiased fit. The overall bias, β, for a floated parameter or yield in the fit model is taken to be the mean of the Gaussian fitted to the residual distribution. The pull distribution also provides a tool for testing error coverage. For correct error coverage the Gaussian width of the pull distribution should be unity. Notable deviations from unity indicate error undercoverage (width > 1) or error overcoverage (width < 1).

5.10.2 Pure toy tests

Pure toy tests are used to check for fit instabilities and intrinsic biases. They are also used for fit error studies.

A single pure toy test consists of numerous experiments (typically 600 here[4]). For each of these experiments, signal and $q\overline{q}$ background datasets are sampled – as described in Section 5.10.1 – using the PDFs presented in Section 5.7. The number of events generated for each hypothesis is varied around a mean value according to the Poisson distribution (termed *Poissonian smearing*). This mean value (t in

[4]The number of experiments conducted in a toy test is limited by the amount of available processing (CPU) time.

Eq. (5.19)) is set to be equal to the number of events expected in the final on-resonance sample. The extended ML fit is then applied to extract the floated PDF parameters and the signal and $q\overline{q}$ background yields. On completion of the final experiment, a Gaussian is fitted to the pull distribution. For an unbiased fit model with correct error coverage, this Gaussian should be centred on zero with unit width.

Several toy tests are carried out, each time varying the Poissonian mean that determines the number of events generated and which reflects the number of events expected in the final on-resonance sample. Table 5.17 lists the results of these pure toy tests. No significant biases or error discrepancies are observed. Figure 5.14 illustrates as an example the residual and pull distributions obtained from one of these toy tests for the signal and $q\overline{q}$ background yields; Figure 5.15 shows the pull distributions of the floated PDF parameters.

5.10.3 Toy tests with embedded signal events

The toy tests described in Section 5.10.2 are repeated this time using fully simulated (rather than toy) MC events for the signal component. This is achieved by randomly sampling subsets from the available MC, which has been subjected to RaSC. Using fully simulated signal events provides a test of whether any subtle correlations between fit variables give rise to biases. Here, we are not concerned with validating the error coverage of the fit. Poissonian smearing is therefore turned off. For a large number of experiments per toy test this does not affect the value of the bias β, but it does enable this value to be established with greater precision. The results of toy tests with embedded MC signal events are listed in Table 5.18. No significant biases are observed.

Number Of Signal And $q\overline{q}$ Background Events Generated (Poissonially Smeared)	***Average Yield Extracted Over 600 Experiments (And Average Error)***		***Bias β; Mean, Width Of Gaussian Fitted To Pull Distribution***	
	Signal	***$q\overline{q}$ Bkgd.***	***Signal***	***$q\overline{q}$ Bkgd.***
122 signal, 38,568 $q\overline{q}$ bkgd.	122.40 ± 0.71 (17.78)	$38,565 \pm 8$ (197)	0.40 ± 0.71; -0.026 ± 0.040, 0.987 ± 0.029	-2.90 ± 7.63; -0.017 ± 0.039, 0.949 ± 0.027
61† signal, 38,629 $q\overline{q}$ bkgd.	61.27 ± 0.57 (14.13)	$38,632 \pm 8$ (197)	0.27 ± 0.57; -0.055 ± 0.041, 1.015 ± 0.029	3.39 ± 8.31; 0.015 ± 0.042, 1.033 ± 0.030
51 signal, 38,639 $q\overline{q}$ bkgd.	50.29 ± 0.55 (13.26)	$38,647 \pm 8$ (197)	-0.71 ± 0.55; -0.141 ± 0.043, 1.046 ± 0.030	8.25 ± 8.27; 0.039 ± 0.042, 1.029 ± 0.030
41 signal, 38,649 $q\overline{q}$ bkgd.	41.91 ± 0.52 (12.52)	$38,630 \pm 8$ (197)	0.91 ± 0.52; -0.025 ± 0.043, 1.046 ± 0.030	-19.08 ± 7.94; -0.099 ± 0.040, 0.988 ± 0.029
31 signal, 38,659 $q\overline{q}$ bkgd.	31.00 ± 0.49 (11.43)	$38,661 \pm 8$ (197)	-0.01 ± 0.49; -0.119 ± 0.044, 1.088 ± 0.031	2.20 ± 7.92; 0.009 ± 0.040, 0.986 ± 0.028
21 signal, 38,669 $q\overline{q}$ bkgd.	21.38 ± 0.43 (10.29)	$38,662 \pm 8$ (197)	0.38 ± 0.43; -0.093 ± 0.044, 1.069 ± 0.031	-7.19 ± 8.25; -0.039 ± 0.042, 1.027 ± 0.030
11 signal, 38,679 $q\overline{q}$ bkgd.	10.79 ± 0.42 (8.85)	$38,672 \pm 8$ (197)	-0.21 ± 0.42; -0.163 ± 0.048, 1.169 ± 0.034	-7.46 ± 7.99; -0.040 ± 0.041, 0.996 ± 0.029

†: 61 is the number of signal events expected in the final on-resonance sample for $\mathcal{B}(B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+) = 1.0 \times 10^{-6}$.

Table 5.17: *Pure toy tests for the 5-variable fit in the 892 window.*

5.10.4 Toy tests with embedded $B\overline{B}$ bkgd. events

Having established that there are no intrinsic biases present in the fit model and that there are no biases due to the small correlations between the S component fit variables, the next stage is to study the effect of the presence of $B\overline{B}$ background

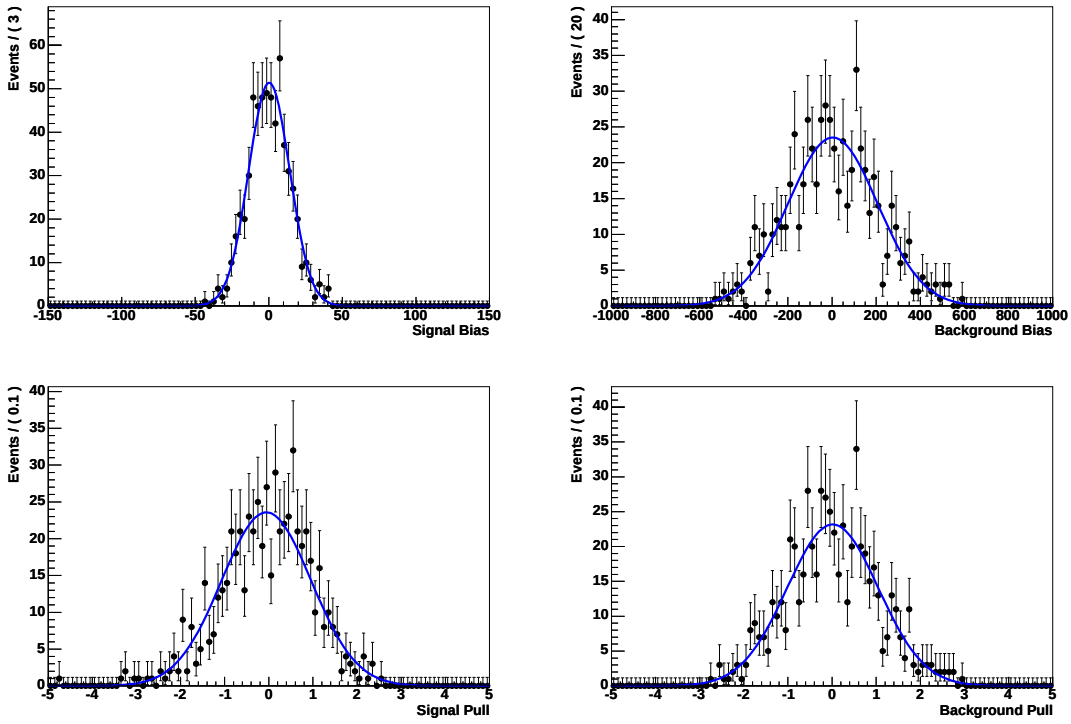

Figure 5.14: *A pure toy test consisting of 600 experiments is conducted. For each experiment, 61 signal and 38,629 $q\overline{q}$ background events (Poissonially smeared) are generated and subjected to the extended ML fit. Top row: residual distributions with fitted Gaussians for signal (left) and $q\overline{q}$ background (right) yields. Bottom row: pull distributions with fitted Gaussians for the same yields. The parameters of the fitted Gaussians are given in Table 5.17.*

events. For each of the modes listed in Table 5.8, a toy test is conducted in which $B\overline{B}$ backgrounds are embedded. For example, for the mode $B^+ \to \phi(1020)(\to K^+K^-)K^+$, $n_{B\overline{B}\ Bkgd} = 21$ events are embedded in addition to the S and $q\overline{q}$ events that are sampled as in Section 5.10.3 (recall from Section 5.5.2.5, that for a given $B\overline{B}$ background source, $n_{B\overline{B}\ Bkgd}$ is the number of expected events in the final on-resonance sample due to this source). The embedded $B\overline{B}$ background events are sampled from MC samples fully simulating the mode under study. These MC samples are first subjected to RaSC. At this stage the fit model consists only of an S component and a $q\overline{q}$ component. It follows then that the presence of the

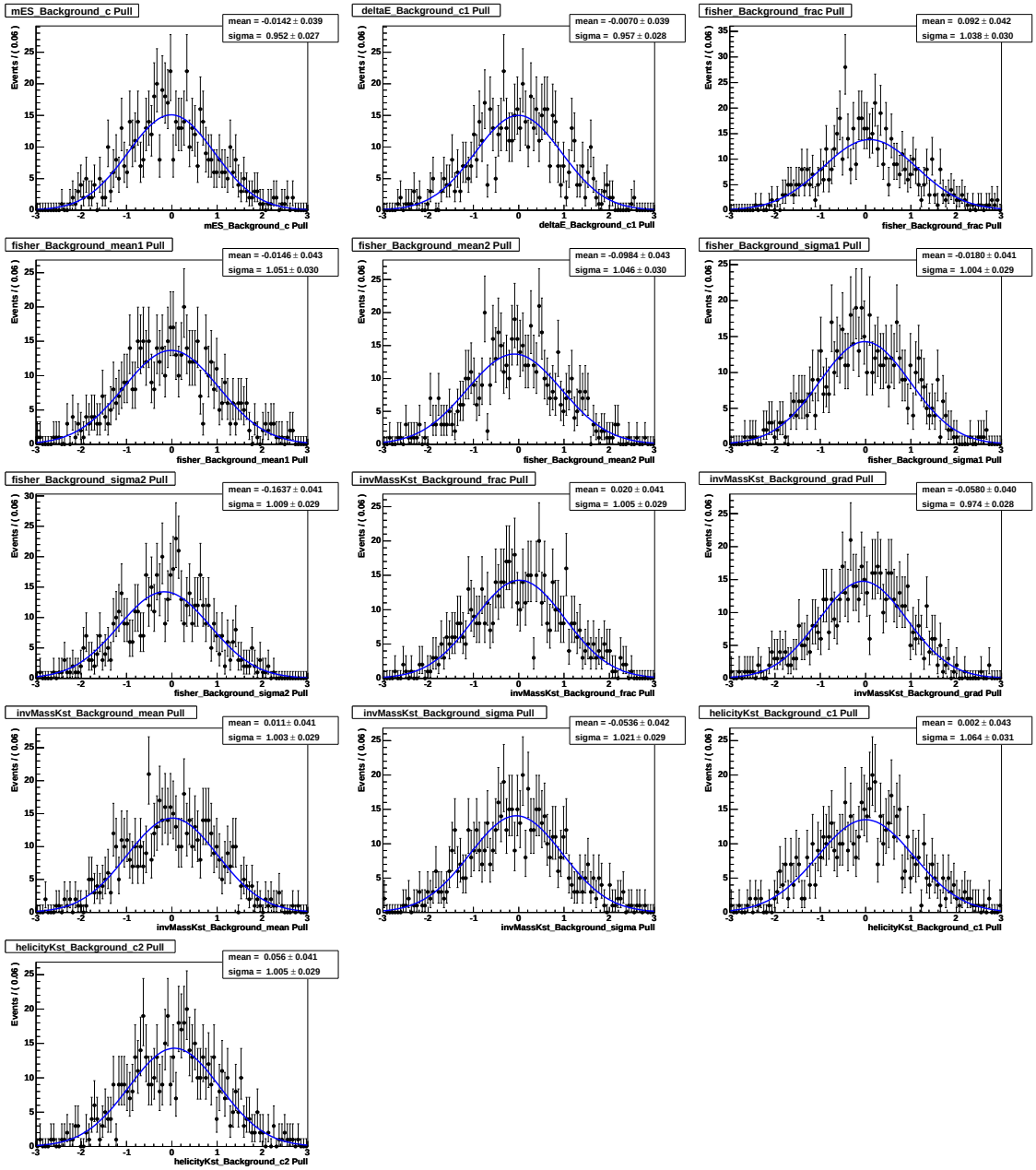

Figure 5.15: *PDF parameter pull distributions for the pure toy test described in Figure 5.14.*

$B\overline{B}$ background events will introduce a bias to one or both of the yields extracted from the fit. It is the bias on the signal yield that is of interest. Where a source of $B\overline{B}$ background induces a significant bias on the signal yield, it can be treated in one of two ways. The estimated bias can be subtracted from the signal yield,

Number Of MC Signal Events Sampled	***β On Signal Yield***	***Number Of $q\overline{q}$ Background Events Generated***	***β On $q\overline{q}$ Bkgd. Yield***
122	1.04 ± 0.55	38,568	-0.85 ± 0.55
61	0.94 ± 0.47	38,629	-0.76 ± 0.47
51	0.97 ± 0.44	38,639	-0.80 ± 0.44
41	0.37 ± 0.43	38,649	-0.21 ± 0.43
31	0.21 ± 0.41	38,659	-0.08 ± 0.41
21	0.66 ± 0.39	38,669	-0.57 ± 0.39
11	0.09 ± 0.39	38,679	-0.05 ± 0.39

Table 5.18: *Biases on the signal and $q\overline{q}$ background yields for the 5-variable fit in the 892 window. These values are obtained using toy tests with embedded MC signal events. 600 experiments are run per toy test.*

or a component can be added to the fit model to account for the $B\overline{B}$ background hypothesis.

Table 5.19 lists the biases obtained from running toy tests with embedded $B\overline{B}$ background events. Table 5.20 lists further biases, but here the results are hypothetical, since zero interference is assumed for modes with the same final state. For modes that cause a significant bias on the signal yield, in addition to the toy tests in which $n_{B\overline{B}\ Bkgd}$ $B\overline{B}$ background events are embedded, further toy tests are conducted in which $n_{B\overline{B}\ Bkgd} \pm \Delta n_{B\overline{B}\ Bkgd}$ events are embedded. This is helpful where it is necessary to apply systematic errors (i.e. when subtracting the bias due to a particular source from the signal yield). Note that for some of the modes listed in Table 5.19, the number of events sampled per experiment multiplied by the number of experiments per toy test is in excess of the number of MC events available (Table 5.8). Oversampling in this way means that the subset samples used for each experiment are pseudo-independent in that some events are used in more than one subset. The

level of oversampling required to noticeably skew results is avoided here with a large margin to spare.

The three sources of $B\overline{B}$ background that induce the largest biases on the signal yield are:

- $B^+ \to \phi(1020)(\to K^+K^-)K^+$.
- $B^+ \to K^+K^-K^+$ Dalitz plot model, minus the $B^+ \to \phi(1020)K^+$ contribution.
- $B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$.

The numbers of events estimated to be present in the final on-resonance sample for these sources are $n^1_{B\overline{B}\ Bkgd} = 20.6 \pm 0.4 \pm 1.5$, $n^2_{B\overline{B}\ Bkgd} = 23.0 \pm 1.1 \pm 2.2$ and $n^3_{B\overline{B}\ Bkgd} = 4.3 \pm 0.2 \pm 0.7$ respectively ($n^2_{B\overline{B}\ Bkgd}$ depends on the Dalitz plot model used, the value given here is using the Belle model B_0 (solution 1) [42], as discussed in Section 5.5.2.5). The fit model is modified by adding components to account for these three sources of $B\overline{B}$ background. This gives five components in total:

- Signal, S (split into S_{TRU} and S_{SXF}).
- $q\overline{q}$ background, $q\overline{q}$.
- $B\overline{B}$ background, $(B\overline{B})_1$ ($B^+ \to \phi(1020)(\to K^+K^-)K^+$).
- $B\overline{B}$ background, $(B\overline{B})_2$ ($B^+ \to K^+K^-K^+$ Dalitz plot model, minus the $B^+ \to \phi(1020)K^+$ contribution).
- $B\overline{B}$ background, $(B\overline{B})_3$ ($B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$).

Mode And Number Of Events Embedded		β ***(Signal)***	β ***($q\overline{q}$ Bkgd.)***	***Mode And Number Of Events Embedded***		β ***(Signal)***	β ***($q\overline{q}$ Bkgd.)***
Combinatoric charged generics	69	0.43 ± 0.47	68.56 ± 0.47	$B^0 \to K^*(892)^0(\to K^+\pi^-)\overline{K}^*(892)^0(\to K^-\pi^+)$ (L)	7	0.14 ± 0.48	6.86 ± 0.48
Combinatoric charged generics	107	0.82 ± 0.48	106.21 ± 0.48	$B^0 \to K^*(892)^0(\to K^+\pi^-)\overline{K}^*(892)^0(\to K^-\pi^+)$ (L)	14	0.45 ± 0.48	13.56 ± 0.47
Combinatoric charged generics	31	0.45 ± 0.45	30.57 ± 0.45	$B^0 \to K^*(892)^0(\to K^+\pi^-)\overline{K}^*(892)^0(\to K^-\pi^+)$ (L)	1	-0.04 ± 0.47	1.03 ± 0.47
Combinatoric neutral generics	26	0.75 ± 0.48	25.25 ± 0.48	$B^+ \to \eta'(958)(\to \gamma\rho(770)^0(\to \pi^+\pi^-))K^+$	6	0.57 ± 0.48	5.44 ± 0.48
Combinatoric neutral generics	41	1.01 ± 0.47	39.98 ± 0.47	$B^+ \to \eta'(958)(\to \gamma\rho(770)^0(\to \pi^+\pi^-))K^+$	7	0.84 ± 0.48	6.17 ± 0.48
Combinatoric neutral generics	11	0.52 ± 0.48	10.48 ± 0.48	$B^+ \to \eta'(958)(\to \gamma\rho(770)^0(\to \pi^+\pi^-))K^+$	5	0.41 ± 0.47	4.60 ± 0.47
$B^+ \to K^+K^-K^+$ Dalitz plot model, minus the $B^+ \to$				$B^0 \to K^+K^-\pi^0$ (NR)	6	0.26 ± 0.48	5.77 ± 0.48
$\phi(1020)K^+$ contribution				$B^0 \to K^+K^-\pi^0$ (NR)	11	0.71 ± 0.48	10.31 ± 0.48
Belle model B_0, sol. 1, minus ϕK	24	1.97 ± 0.47	22.04 ± 0.47	$B^0 \to K^+K^-\pi^0$ (NR)	1	-0.02 ± 0.47	1.04 ± 0.47
Belle model B_0, sol. 1, minus ϕK	27	2.54 ± 0.46	24.48 ± 0.46	$B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$	5	1.35 ± 0.47	3.64 ± 0.47
Belle model B_0, sol. 1, minus ϕK	20	1.92 ± 0.46	18.08 ± 0.46	$B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$	6	1.83 ± 0.49	4.18 ± 0.49
(Belle model B_0, sol. 2, minus ϕK	20	1.73 ± 0.46	18.27 ± 0.46)	$B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$	4	1.17 ± 0.45	2.83 ± 0.45
(Belle model B_0, sol. 2, minus ϕK	23	1.61 ± 0.46	21.40 ± 0.46)	$B^0 \to K^+K^-K^0_S$ (NR)	3	-0.08 ± 0.47	3.09 ± 0.47
(Belle model B_0, sol. 2, minus ϕK	17	1.26 ± 0.50	15.76 ± 0.50)	$B^+ \to f_0(980)(\to \pi^+\pi^-)K^+$	2	0.02 ± 0.48	1.97 ± 0.48
(*BABAR* solution A, minus ϕK	31	2.51 ± 0.47	28.49 ± 0.47)	$B^+ \to f_0(980)(\to \pi^+\pi^-)K^+$	3	0.13 ± 0.48	2.88 ± 0.48
(*BABAR* solution A, minus ϕK	35	2.86 ± 0.47	32.15 ± 0.47)	$B^0 \to K^+\rho(770)^-(\to \pi^-\pi^0)$	2	0.16 ± 0.47	1.83 ± 0.47
(*BABAR* solution A, minus ϕK	26	2.29 ± 0.46	23.71 ± 0.46)	$B^+ \to K^*(892)^+(\to K^+\pi^0)\overline{K}^*(892)^0(\to K^-\pi^+)$ (L)	2	0.04 ± 0.47	1.97 ± 0.47
(*BABAR* solution B, minus ϕK	30	2.03 ± 0.46	27.99 ± 0.46)	$B^+ \to \overline{D}^0(\to K^+K^-)K^+$	2	-0.08 ± 0.47	2.09 ± 0.47
(*BABAR* solution B, minus ϕK	35	2.43 ± 0.47	32.61 ± 0.47)	$B^0 \to K^+\pi^-\pi^0$ (NR)	1	-0.01 ± 0.46	1.04 ± 0.46
(*BABAR* solution B, minus ϕK	26	2.11 ± 0.46	23.91 ± 0.46)	$B^+ \to K_1(1270)^0(\to \rho(770)^+(\to \pi^+\pi^0)K^-)K^+$	2	-0.03 ± 0.47	2.02 ± 0.47
$B^+ \to \phi(1020)(\to K^+K^-)K^+$	21	3.41 ± 0.46	17.59 ± 0.46	$B^0 \to K_1(1270)^+(\to \rho(770)^0(\to \pi^+\pi^-)K^+)K^-$	2	-0.05 ± 0.47	2.05 ± 0.47
$B^+ \to \phi(1020)(\to K^+K^-)K^+$	23	3.93 ± 0.47	19.08 ± 0.48	$B^0 \to K^*(892)^0(\to X)\gamma$	1	0.11 ± 0.46	0.91 ± 0.46
$B^+ \to \phi(1020)(\to K^+K^-)K^+$	19	3.62 ± 0.49	15.39 ± 0.49	$B^+ \to K^*(892)^+(\to K^+\pi^0)\overline{K}^*(892)^0(\to K^-\pi^+)$ (T)	1	-0.05 ± 0.46	1.08 ± 0.46

Table 5.19: *Biases obtained from toy tests with embedded $B\overline{B}$ background events (5-variable fit in the 892 window). Fit model components are S and $q\overline{q}$; no $B\overline{B}$ components. 600 experiments are run per toy test. Poissonian smearing is turned off.*

Mode And Number Of Events Embedded		β ***(Signal)***	β ***($q\overline{q}$ Bkgd.)***
$B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$			
$m_{K\pi}$ modelled as LASS	33	14.13 ± 0.49	18.86 ± 0.50
$m_{K\pi}$ modelled as LASS	67	30.10 ± 0.55	36.89 ± 0.56
$m_{K\pi}$ modelled as LASS	1	0.35 ± 0.46	0.68 ± 0.46
$m_{K\pi}$ modelled as Breit-Wigner	15	5.71 ± 0.48	9.29 ± 0.48
$m_{K\pi}$ modelled as Breit-Wigner	29	10.85 ± 0.49	18.17 ± 0.49
$m_{K\pi}$ modelled as Breit-Wigner	1	0.30 ± 0.46	0.71 ± 0.46
$B^+ \to K^+K^-\pi^+$ (NR)	9	3.41 ± 0.50	5.62 ± 0.50
$B^+ \to K^+K^-\pi^+$ (NR)	17	6.68 ± 0.50	10.34 ± 0.50
$B^+ \to K^+K^-\pi^+$ (NR)	1	0.34 ± 0.46	0.69 ± 0.46

Table 5.20: *As Table 5.19, except the $B\overline{B}$ background modes share the same final state as signal events. Biases shown are hypothetical in that the assumption of zero interference is made.*

5.10.5 $B\overline{B}$ background PDFs

Table 5.21 lists the PDF functional forms used for the $(B\overline{B})_i$ $(i = 1, 2, 3)$ components. These PDFs are illustrated in Figure 5.16. They are generated by performing one-dimensional fits to the appropriate MC samples[5]. A `keys` PDF (used to fit to $\cos\theta_H$ for the $(B\overline{B})_2$ component) is a continuous, non-parametric PDF. See [98] for further details. PDF parameters are presented in Appendix D.

Correlation values are given in Table 5.22. These numbers are large. Correlation plots are presented in Appendix E. Viewing these plots, the correlations do not look at all severe – suggesting that the large values obtained using Eq. (5.13) are, in part, due to low statistics. One solution to a large correlation between two fit variables is to use a two-dimensional PDF, in which PDF parameters of one variable are functions of the other variable. However, two-dimensional PDFs are not employed here since to do so would be to model one correlation (between, for example, m_{ES} and ΔE) but then ignore another one (e.g. m_{ES} and $\cos\theta_H$). This is deemed arbitrary and without gain. One-dimensional PDFs are therefore used throughout. It is shown in the following section that the biases brought about by these correlations are small.

[5]The number of available MC events used to construct the PDFs for the $(B\overline{B})_2$ component is smaller than desired. However, the resulting large uncertainties on the PDF parameters are propagated as described in Section 6.4.1.2. Also, $N_{MC} \gg n^2_{B\overline{B}\ Bkgd}$, where N_{MC} is the number of available MC events and $n^2_{B\overline{B}\ Bkgd}$ is the number of actual events expected in the final on-resonance sample.

	$(B\overline{B})_1$	$(B\overline{B})_2$	$(B\overline{B})_3$
m_{ES}	Crystal Ball	Gaussian + Argus	Crystal Ball
ΔE	Gaussian + Linear	Gaussian + Linear	Gaussian + Linear
$\mathcal{F}$	Double Gaussian	Double Gaussian	Gaussian
$m_{K\pi}$	Voigtian + Exponential	Quadratic	Gaussian + Linear
$\cos\theta_H$	Exponential	`Keys`	Exponential

Table 5.21: *$B\overline{B}$ background PDF types used for the 5-variable fit in the 892 window.*

ΔE	$-0.248, -0.406, 0.210$			
$\mathcal{F}$	$-0.080, 0.007, -0.107$	$0.025, -0.017, -0.053$		
$m_{K\pi}$	$-0.386, 0.122, -0.069$	$0.207, -0.046, -0.003$	$0.069, 0.006, 0.011$	
$\cos\theta_H$	$-0.520, -0.448, 0.136$	$0.202, 0.318, -0.017$	$0.067, -0.053, -0.033$	$0.176, -0.043, -0.231$
	m_{ES}	ΔE	$\mathcal{F}$	$m_{K\pi}$

Table 5.22: *Correlations (calculated using Eq. (5.13)) between variables used in the 5-variable fit in the 892 window. For each cell, the first value is for the $(B\overline{B})_1$ component, the second is for the $(B\overline{B})_2$ component and the third is for the $(B\overline{B})_3$ component.*

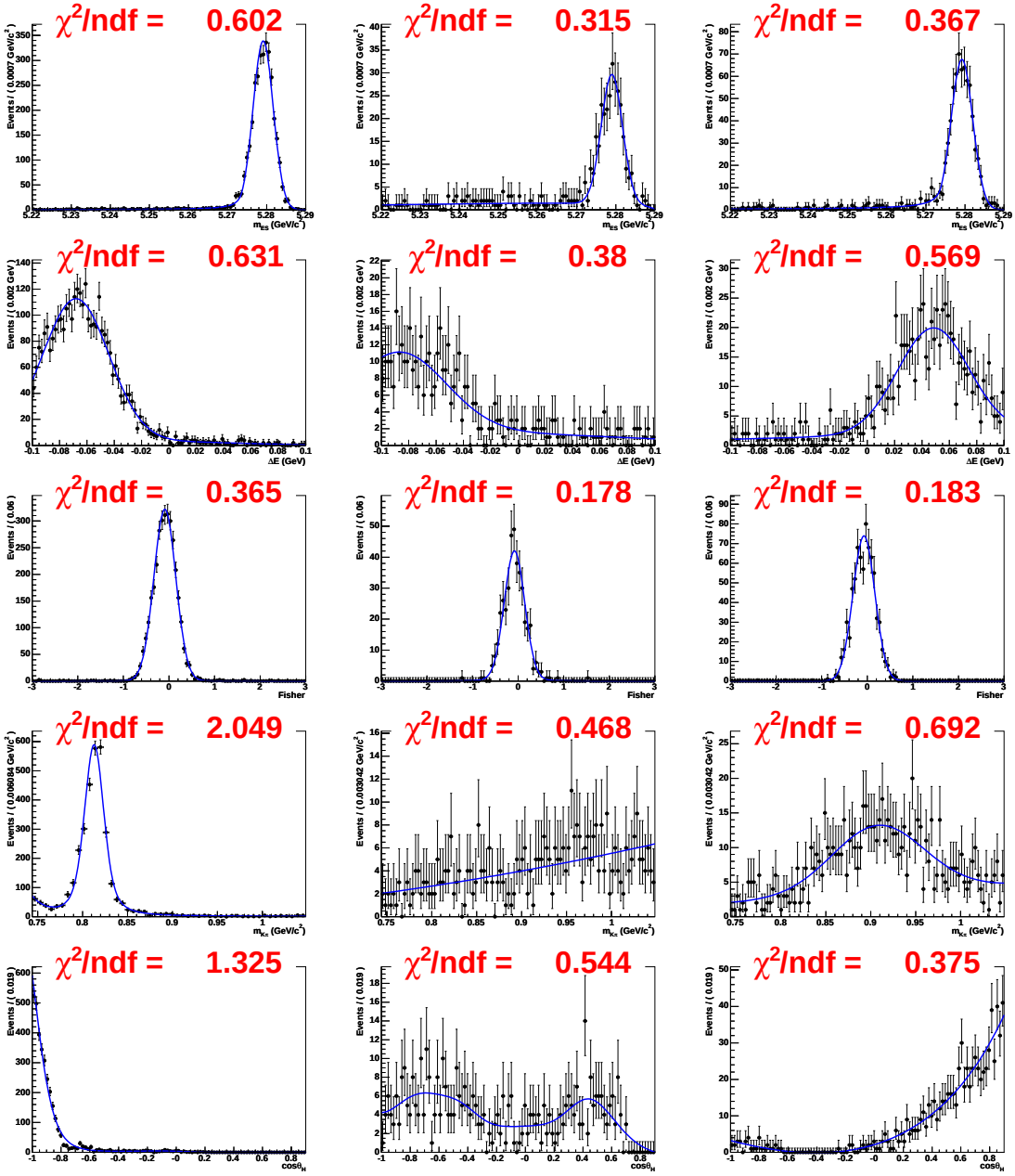

Figure 5.16: $(B\overline{B})_i$ $(i = 1, 2, 3)$ *PDFs (blue, solid lines) for the 5-variable fit in the 892 window, obtained by fitting to appropriate MC distributions (black points with error bars). PDF types are presented in Table 5.21; PDF parameter values are given in Tables D.4, D.5 and D.6 (for components* $(B\overline{B})_1$, $(B\overline{B})_2$ *and* $(B\overline{B})_3$, *respectively). Left column:* $(B\overline{B})_1$ *component; middle column:* $(B\overline{B})_2$ *component; right column:* $(B\overline{B})_3$ *component. Top row:* m_{ES}*; second row:* ΔE*; third row:* $\mathcal{F}$*; fourth row:* $m_{K\pi}$*; bottom row:* $\cos\theta_H$.

5.10.6 Toy tests with $B\overline{B}$ background components

The toy tests described in Sections 5.10.2 and 5.10.3 are repeated, this time with five components (S, $q\overline{q}$, $(B\overline{B})_1$, $(B\overline{B})_2$, $(B\overline{B})_3$) rather than two (S and $q\overline{q}$). Toy tests are conducted for:

- Floated $(B\overline{B})_i$ $(i = 1, 2, 3)$ yields.
- Fixed $(B\overline{B})_i$ $(i = 1, 2, 3)$ yields.

As with the signal component, and for the same reason, $(B\overline{B})_i$ $(i = 1, 2, 3)$ PDF parameters are fixed (to the values obtained in Section 5.10.5).

Tables 5.23 and 5.24 show, as an example, the yield biases obtained from two pure toy tests – one in which the $B\overline{B}$ background yields are floated, and one in which they are fixed. For fixed-yield $(B\overline{B})_i$ $(i = 1, 2, 3)$ components, no significant[6] biases are observed. For floated-yield $(B\overline{B})_i$ $(i = 1, 2, 3)$ components, the model is arguably a little less stable with a bias and error miscoverage present for the $(B\overline{B})_3$ component.

Tables 5.25 and 5.26 list the biases obtained for floated yields when running toy tests in which embedded S and $(B\overline{B})_i$ $(i = 1, 2, 3)$ events are fully simulated MC events. For Table 5.25, the $(B\overline{B})_i$ $(i = 1, 2, 3)$ yields are fixed. They are fixed to the central values of $n^1_{B\overline{B}\ Bkgd}$, $n^2_{B\overline{B}\ Bkgd}$ and $n^3_{B\overline{B}\ Bkgd}$ respectively (as defined in the final paragraph of Section 5.10.4). In fixing the yields of the $B\overline{B}$ background components, a systematic error is introduced due to the uncertainties on the values to which the yields are fixed. It is ensured that this error is not excessively large by embedding n_{Embd} $(B\overline{B})_i$ events (where $n_{Embd} \neq n^i_{B\overline{B}\ Bkgd}$), fixing the $(B\overline{B})_i$ yield to $n^i_{B\overline{B}\ Bkgd}$, and observing the resulting bias on the floated signal yield. To avoid this source of systematic error, the $(B\overline{B})_i$ $(i = 1, 2, 3)$ yields can be floated. This is the case for the embedded toy tests whose yield biases are listed in Table 5.26. The downside to floating these yields is that the fit model may become unstable due to the lack of statistics for these components.

[6] When compared with the statistical uncertainty on the yield and other systematic effects.

$B\overline{B}$ background events that originate from sources for which there is no component in the fit do not have a large impact on the signal yield extracted. However, the effect is not negligible. Non-$K^+K^-\pi^+$-final-state sources of $B\overline{B}$ background that are expected to contribute $n_{B\overline{B}\ Bkgd} + \Delta n_{B\overline{B}\ Bkgd} > 0.5$ events to the final on-resonance sample, excluding those sources for which a fit component is included, are now defined collectively as "$(B\overline{B})_{Other}$" (all sources listed in Table 5.8 excluding $B^+ \to \phi(1020)(\to K^+K^-)K^+$, $B^+ \to K^+K^-K^+$ Dalitz plot model (minus the $B^+ \to \phi(1020)K^+$ contribution) and $B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$). $n^{Other}_{B\overline{B}\ Bkgd}$ is defined to be the value obtained when rounding the central values of $n_{B\overline{B}\ Bkgd}$ for each of the $(B\overline{B})_{Other}$ sources to the nearest non-zero integer, and taking the sum. $n^{OtherMax}_{B\overline{B}\ Bkgd}$ is defined to be the value obtained when rounding the values of ($n_{B\overline{B}\ Bkgd}$+ $\Delta n_{B\overline{B}\ Bkgd}$) for each of the $(B\overline{B})_{Other}$ sources to the nearest non-zero integer, and taking the sum. To estimate the number of events that should be subtracted from the signal yield due to the presence of $(B\overline{B})_{Other}$ events, a toy test is carried out. The bias on the signal yield is estimated when we have $n^{Other}_{B\overline{B}\ Bkgd}$ $(B\overline{B})_{Other}$ real MC events embedded. To help evaluate the uncertainty on the subtraction, a further toy test is carried out in which $n^{OtherMax}_{B\overline{B}\ Bkgd}$ $(B\overline{B})_{Other}$ events are embedded. The second section of Table 5.25 and the last two rows of Table 5.26 show that the vast majority of $(B\overline{B})_{Other}$ events are assigned to the background hypotheses.

It can be seen, from Table 5.24 and Table 5.26, that the fit remains reasonably stable with floated-yield $(B\overline{B})_i$ $(i = 1, 2, 3)$ components. For embedded toy tests, there are small biases on the background hypotheses, but, with no $(B\overline{B})_{Other}$ events embedded, the bias on the signal yield is consistent with zero. However, it is decided that the more conservative approach of fixing the $(B\overline{B})_i$ $(i = 1, 2, 3)$ component yields shall be taken. This provides the most stable fit model, and the cost in fixing these yields – in terms of introducing an additional source of systematic error – is small.

	Gaussian Fitted To Pull Distribution		β
	Mean	***Width***	
S	-0.089 ± 0.060	1.119 ± 0.042	0.54 ± 0.65
$q\overline{q}$	-0.015 ± 0.053	0.990 ± 0.037	-2.42 ± 10.48
$(B\overline{B})_1$	-0.170 ± 0.056	1.045 ± 0.039	-0.33 ± 0.34
$(B\overline{B})_2$	-0.103 ± 0.061	1.133 ± 0.043	0.46 ± 1.21
$(B\overline{B})_3$	-0.335 ± 0.068	1.274 ± 0.049	-2.81 ± 0.84

Table 5.23: *Example of a pure toy test for the 5-variable fit in the 892 window, with floated-yield components S, $q\overline{q}$, $(B\overline{B})_1$, $(B\overline{B})_2$ and $(B\overline{B})_3$.*

	Gaussian Fitted To Pull Distribution		β
	Mean	***Width***	
S	-0.180 ± 0.050	1.117 ± 0.035	-0.66 ± 0.54
$q\overline{q}$	-0.026 ± 0.048	1.064 ± 0.034	-4.50 ± 9.36

Table 5.24: *Example of a pure toy test for the 5-variable fit in the 892 window, with floated-yield components S and $q\overline{q}$, and fixed-yield components $(B\overline{B})_1$, $(B\overline{B})_2$ and $(B\overline{B})_3$.*

$\#\ (B\overline{B})_1$ ***Embedded*** †	$\#\ (B\overline{B})_2$ ***Embedded*** †	$\#\ (B\overline{B})_3$ ***Embedded*** †	β ***(Signal)***	β ***($q\overline{q}$ Bkgd.)***
—	—	—	0.24 ± 0.42	-0.11 ± 0.41
$n^1_{B\overline{B}\ Bkgd}$	—	—	0.66 ± 0.45	-1.23 ± 0.46
$n^1_{B\overline{B}\ Bkgd} - \Delta n^1_{B\overline{B}\ Bkgd}$	—	—	0.45 ± 0.45	-1.96 ± 0.46
$n^1_{B\overline{B}\ Bkgd} + \Delta n^1_{B\overline{B}\ Bkgd}$	—	—	0.75 ± 0.44	-0.27 ± 0.45
—	$n^2_{B\overline{B}\ Bkgd}$ ♣	—	1.00 ± 0.46	-1.10 ± 0.46
—	$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♣	—	0.62 ± 0.47	-3.54 ± 0.47
—	$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♣	—	0.85 ± 0.45	1.81 ± 0.44
—	$n^2_{B\overline{B}\ Bkgd}$ ♢	—	0.52 ± 0.44	-4.16 ± 0.45
—	$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♢	—	-0.35 ± 0.46	-6.11 ± 0.46
—	$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♢	—	0.56 ± 0.47	-1.35 ± 0.47
—	$n^2_{B\overline{B}\ Bkgd}$ ♠	—	1.66 ± 0.46	4.19 ± 0.45
—	$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♠	—	1.09 ± 0.45	1.06 ± 0.45
—	$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♠	—	1.18 ± 0.47	8.35 ± 0.47
—	$n^2_{B\overline{B}\ Bkgd}$ ♡	—	1.46 ± 0.47	4.34 ± 0.47
—	$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♡	—	1.51 ± 0.44	0.68 ± 0.43
—	$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♡	—	1.41 ± 0.46	8.10 ± 0.47
—	—	$n^3_{B\overline{B}\ Bkgd}$	0.51 ± 0.44	-0.67 ± 0.44
—	—	$n^3_{B\overline{B}\ Bkgd} - \Delta n^3_{B\overline{B}\ Bkgd}$	-0.18 ± 0.45	-0.97 ± 0.45
—	—	$n^3_{B\overline{B}\ Bkgd} + \Delta n^3_{B\overline{B}\ Bkgd}$	0.68 ± 0.46	1.07 ± 0.46
$n^1_{B\overline{B}\ Bkgd}$	$n^2_{B\overline{B}\ Bkgd}$ ♣	$n^3_{B\overline{B}\ Bkgd}$	0.57 ± 0.46	-1.85 ± 0.46
$n^1_{B\overline{B}\ Bkgd} - \Delta n^1_{B\overline{B}\ Bkgd}$	$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♢	$n^3_{B\overline{B}\ Bkgd} - \Delta n^3_{B\overline{B}\ Bkgd}$	-0.20 ± 0.45	-9.04 ± 0.46
$n^1_{B\overline{B}\ Bkgd} + \Delta n^1_{B\overline{B}\ Bkgd}$	$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♠	$n^3_{B\overline{B}\ Bkgd} + \Delta n^3_{B\overline{B}\ Bkgd}$	1.18 ± 0.44	8.71 ± 0.46
$n^1_{B\overline{B}\ Bkgd}$	$n^2_{B\overline{B}\ Bkgd}$ ♣ $+ n^{Other}_{B\overline{B}\ Bkgd}\ (B\overline{B})_{Other}$ events	$n^3_{B\overline{B}\ Bkgd}$	3.25 ± 0.46	123.96 ± 0.46
$n^1_{B\overline{B}\ Bkgd}$	$n^2_{B\overline{B}\ Bkgd}$ ♣ $+ n^{OtherMax}_{B\overline{B}\ Bkgd}\ (B\overline{B})_{Other}$ events	$n^3_{B\overline{B}\ Bkgd}$	4.07 ± 0.51	192.67 ± 0.51
Extreme cases:				
$n^1_{B\overline{B}\ Bkgd}/2$	—	—	-0.26 ± 0.45	-5.34 ± 0.45
$2n^1_{B\overline{B}\ Bkgd}$	—	—	1.68 ± 0.44	7.54 ± 0.45
—	$n^2_{B\overline{B}\ Bkgd}/2$ ♣	—	-0.88 ± 0.44	-10.50 ± 0.44
—	$2n^2_{B\overline{B}\ Bkgd}$ ♣	—	1.97 ± 0.46	19.42 ± 0.47
—	—	$n^3_{B\overline{B}\ Bkgd}/2$	0.23 ± 0.45	-2.35 ± 0.45
—	—	$2n^3_{B\overline{B}\ Bkgd}$	1.16 ± 0.45	3.52 ± 0.45
$n^1_{B\overline{B}\ Bkgd}/2$	$n^2_{B\overline{B}\ Bkgd}/2$ ♣	$n^3_{B\overline{B}\ Bkgd}/2$	-1.32 ± 0.44	-17.88 ± 0.44
$2n^1_{B\overline{B}\ Bkgd}$	$2n^2_{B\overline{B}\ Bkgd}$ ♣	$2n^3_{B\overline{B}\ Bkgd}$	3.78 ± 0.48	30.39 ± 0.49

†: rounded to nearest integer

$n^1_{B\overline{B}\ Bkgd} = 20.6 \pm 0.4 \pm 1.5$, $n^3_{B\overline{B}\ Bkgd} = 4.3 \pm 0.2 \pm 0.7$

♣: Belle model B_0, solution 1 (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 23.0 \pm 1.1 \pm 2.2$; ♢: Belle model B_0, solution 2 (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 19.3 \pm 1.4 \pm 1.9$; ♠: *BABAR* solution A (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 30.1 \pm 1.7 \pm 2.9$; ♡: *BABAR* solution B (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 29.9 \pm 1.7 \pm 2.9$

Table 5.25: *Embedded toy tests for the 5-variable fit in the 892 window with fixed-yield $B\overline{B}$ background components. Where events are embedded for component $(B\overline{B})_i$ ($i = 1, 2, 3$), the yield for this component is fixed at the central value of $n^i_{B\overline{B}\ Bkgd}$ ($i = 1, 2, 3$).*

	β $((B\overline{B})_1)$	β $((B\overline{B})_2)$	β $((B\overline{B})_3)$	β (S)	β $(q\overline{q})$
—	—	—	—	0.72 ± 0.40	-0.60 ± 0.40
$21, 23, 4$	2.30 ± 0.24	4.21 ± 1.03	-0.36 ± 0.74	-0.27 ± 0.59	-5.79 ± 1.20
$21, 23, 4, n^{Other}_{B\overline{B}\ Bkgd}$	2.28 ± 0.25	11.34 ± 1.05	5.67 ± 0.73	-0.12 ± 0.57	110.12 ± 1.20
$21, 23, 4, n^{OtherMax}_{B\overline{B}\ Bkgd}$	2.65 ± 0.26	15.89 ± 1.10	6.67 ± 0.63	2.35 ± 0.54	171.68 ± 1.26

Table 5.26: *Embedded toy tests for the 5-variable fit in the 892 window with floated-yield $B\overline{B}$ background components. Values given in the left column are the number of $(B\overline{B})_1$, number of $(B\overline{B})_2$, number of $(B\overline{B})_3$ and number of $(B\overline{B})_{Other}$ events embedded respectively.*

5.10.7 Likelihood value comparisons

Two further tests are performed to ensure that our fit model is behaving as it should. These tests involve applying the fit model to the final on-resonance sample. The object is to test whether the likelihood values obtained from fitting to the data are sensible (i.e. in agreement with toy tests, in which the stability of the fit has been rigourously tested).

The values of the yields extracted from the fit to data are used at this point, but they are not observed and so the analysis remains blind. Toy MC events are generated using the PDFs presented in Sections 5.7 and 5.10.5. The number of events generated for the S and $q\overline{q}$ hypotheses is Poissonially smeared about mean values that are equal to the yield values obtained in the fit to the final on-resonance sample.

The first test is a pure toy test of the type described in Section 5.10.2. 500 experiments are performed. For each experiment, events are generated as described above (that is, using the yields obtained from the fit to the final on-resonance sample). The minimum value of the negative log likelihood function obtained in each of the 500 fits is plotted. The resulting distribution – as with the residual and pull distributions – is expected to be Gaussian. The value of the likelihood function obtained when applying the fit model to the final on-resonance sample is then compared with the values obtained for the 500 pure toy tests. Figure 5.17 demonstrates that these values are in excellent agreement. This is clearly an encouraging sign, but it should be noted that the success of this test is a necessary but insufficient condition. The same is true of the second test, which is now described.

Toy MC events are again generated in quantities determined by the yields obtained from the fit to the final on-resonance sample, but this time scaled to give a high-statistics toy sample (since this time we are generating one toy sample rather than 500). Here, rather than comparing the overall likelihood function values obtained from the extended fit, individual event likelihoods are compared. The distributions of the likelihood ratio "Likelihood (Signal) / Likelihood (Total)" for each event are

plotted for the final on-resonance sample and for the high-statistics toy MC sample. Likelihood (Signal) $= \mathcal{P}_S$, whilst Likelihood (Total) $= \mathcal{P}_S + \mathcal{P}_{q\overline{q}} + \mathcal{P}_{(B\overline{B})_1} + \mathcal{P}_{(B\overline{B})_2} + \mathcal{P}_{(B\overline{B})_3}$ (see Section 5.6). The two distributions should be very similar and this is shown to be the case in Figure 5.17.

For the first test, it is feasible that the positive result is a "lucky" coincidence, that event-by-event likelihoods – which individually would indicate discrepancies, "cancel out" to give what appears to be an encouraging result. The findings of the second test imply that this is not the case.

5.11 Tests and validations for the 3-var. fits

The procedure described in Section 5.10 is repeated for the 3-variable fits in the 892 and 1430 windows. For the sake of brevity, the results of the toy tests conducted are summarised in this section with the results in full being reserved for Appendix G.

For pure toy tests of the type described in Sections 5.10.2 and 5.10.6, biases are not found to be present and correct error coverage is observed. This is true for the 3-variable fits in both windows (892 and 1430).

Since the 3-variable fits are a measure of all $B^+ \to K^+K^-\pi^+$ final state modes, separate toy tests are conducted in which embedded signal events are sampled from datasets fully simulating decays for the three modes that are expected to give the largest contributions:

- $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$.
- $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ (Breit-Wigner $m_{K\pi}$ lineshape).
- $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ (LASS $m_{K\pi}$ lineshape).
- Non-resonant $B^+ \to K^+K^-\pi^+$.

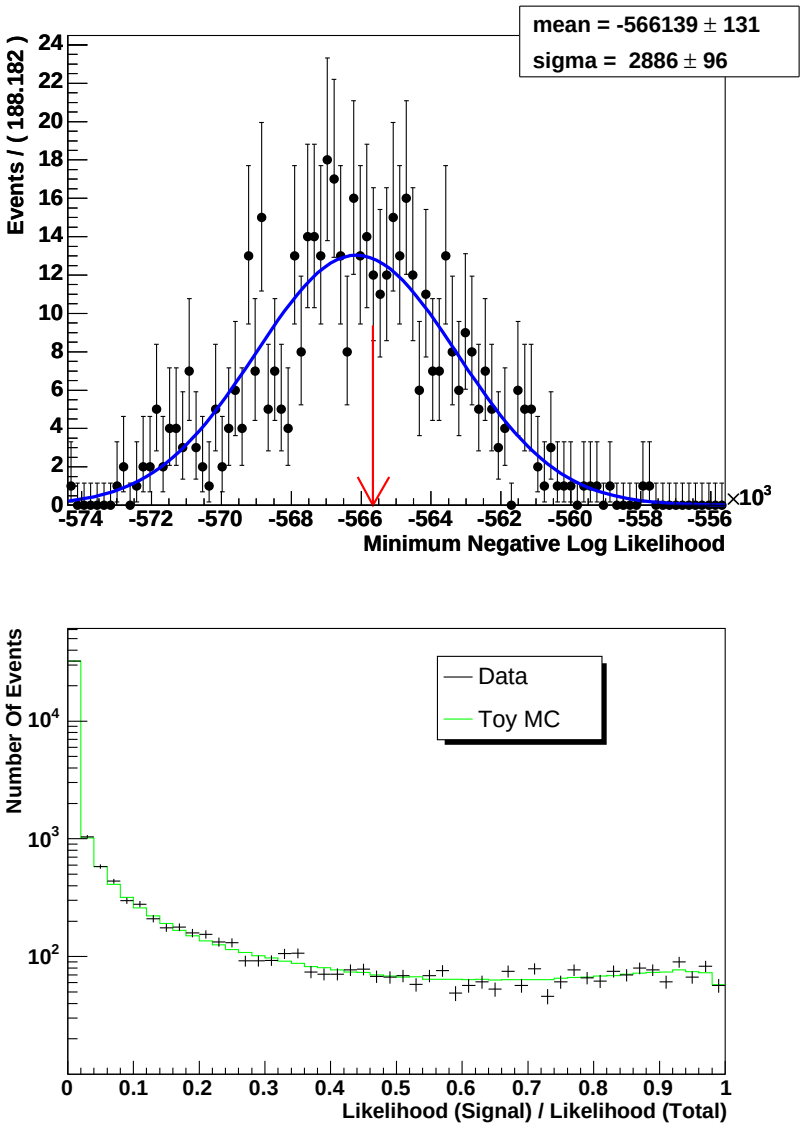

Figure 5.17: *Both plots are relevant to the 5-variable fit in the 892 window, with floated-yield components S and $q\overline{q}$, and fixed-yield components $(B\overline{B})_1$, $(B\overline{B})_2$ and $(B\overline{B})_3$. Top: the distribution of $(-\ln \mathcal{L}_{Max})$ values for 500 pure toy experiments, with a superimposed fitted Gaussian. The red arrow indicates the value of $(-\ln \mathcal{L}_{Max})$ obtained when applying the fit model to the final on-resonance sample. Bottom: event likelihood ratio distributions for toy MC and the final on-resonance sample.*

One mode is embedded per toy test. For each toy test, the fixed PDF functional forms and parameters for the signal component are those obtained in Sections 5.8 and 5.9 for the 892 and 1430 windows respectively. Table 5.27 (892 window) and Table 5.28 (1430 window) list the biases on the floated yields obtained from such toy tests. It can be seen, for the 3-variable fit in the 892 window, that the embedded signal events introduce biases. Although smaller than the average error on the yield, these biases are nevertheless significant. The same is true of the 1430 window. It is necessary to apply a systematic error to account for these observed biases (see Section 6.3.1).

The $B\overline{B}$ background modes that cause significant biases to the fitted signal yield for the 5-variable fit in the 892 window are, as expected, also the modes that give rise to significant biases for the 3-variable fit in the 892 window. For the 3-variable fit however, the biases are larger due to the decreased discriminatory power of 3 variables versus 5. As with the 5-variable fit, fit components shall be included for the following $B\overline{B}$ background modes:

- $B^+ \to \phi(1020)(\to K^+K^-)K^+$: component $(B\overline{B})_1$.
- $B^+ \to K^+K^-K^+$ Dalitz plot model, minus the $B^+ \to \phi(1020)K^+$ contribution: component $(B\overline{B})_2$.
- $B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$: component $(B\overline{B})_3$.

The PDF functional forms are the same as those used for the 5-variable fit (Table 5.21); PDF parameters are presented in Appendix D.

For the 1430 window, the following $B\overline{B}$ background sources are to be accounted for in the form of including components in the fit model:

- $B^+ \to K^+K^-K^+$ Dalitz plot model, minus the $B^+ \to \phi(1020)K^+$ contribution: component $(B\overline{B})_2$.
- $B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$: component $(B\overline{B})_3$.

- $B^+ \to f_0(980)(\to \pi^+\pi^-)K^+$: component $(B\overline{B})_4$.

$B^+ \to \phi(1020)(\to K^+K^-)K^+$ is not an issue in this window, but $B^+ \to f_0(980)(\to \pi^+\pi^-)K^+$ causes a significant bias on the signal yield and a component for this mode is therefore included. The functional forms of the PDFs used for the $(B\overline{B})_i$ ($i = 2, 3, 4$) components are listed in Table 5.29; PDF parameters are given in Appendix D.

As with the 5-variable fit, the yields of the $B\overline{B}$ background components are fixed, as are all PDF parameters for the signal and $B\overline{B}$ background components. This is to ensure a stable fit model. The m_{ES} Argus endpoint for the $q\overline{q}$ background is also fixed. All remaining parameters are floated when the fit model is applied to the final on-resonance sample.

Table 5.30 shows the estimated biases on the floated yields due to the presence of events originating from $B\overline{B}$ background sources for which components are not included in the fit model. For the 1430 window, these estimated biases do not account for the modes $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$ and $B^+ \to \overline{D}^0(\to K^+K^-)\pi^+$. For these modes, there are an inadequate number of MC events remaining after RaSC to be able to model them in the form of fit components, or to estimate the bias they induce on the signal yield using toy tests. They are therefore treated separately from the remaining $(B\overline{B})_{Other}$[7] sources. $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$ shall be taken to induce a bias on the signal yield of $(9.3+\sqrt{1.2^2+0.6^2})/2$ events. A 100% systematic uncertainty is applied to this value. The final state of $B^+ \to \overline{D}^0(\to K^+K^-)\pi^+$ is $K^+K^-\pi^+$. It is therefore expected to be signal-like in m_{ES}, ΔE and $\mathcal{F}$. As such the bias that it induces on the signal yield shall be taken to be 9.4 events, with an uncertainty of $\sqrt{1.5^2+0.7^2}$. The choice of these values becomes clear when considering the relevant entries in Table 5.9.

Two-dimensional fit variable scatter plots are presented in Appendix E, as are fit variable correlation values for the $B\overline{B}$ background components.

[7] $(B\overline{B})_{Other}$ is defined in Section 5.10.6.

The likelihood value comparisons that are described in Section 5.10.7 for the 5-variable fit are repeated for the 3-variable fits. The results are shown in Figure G.3. Toy MC is found to be in excellent agreement with the final on-resonance sample.

Number Of MC Signal Events Sampled	***Average Error On Signal Yield***	***β On Signal Yield***	***Number Of $q\overline{q}$ Background Events Generated***	***Average Error On $q\overline{q}$ Bkgd. Yield***	***β On $q\overline{q}$ Bkgd. Yield***
Embedded signal events sampled from $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ MC:					
200	25.23	6.62 ± 0.66	39,956	201	-6.45 ± 0.66
150	23.05	4.82 ± 0.63	40,006	201	-4.65 ± 0.63
100	20.56	4.65 ± 0.58	40,056	201	-4.49 ± 0.58
80	19.34	4.45 ± 0.55	40,076	201	-4.27 ± 0.55
60	17.96	3.63 ± 0.53	40,096	201	-3.46 ± 0.53
50	17.19	3.41 ± 0.50	40,106	201	-3.28 ± 0.50
40	16.23	2.44 ± 0.47	40,116	201	-2.30 ± 0.46
30	15.15	1.89 ± 0.48	40,126	201	-1.79 ± 0.48
20	14.11	1.76 ± 0.45	40,136	201	-1.65 ± 0.45
10	12.85	0.80 ± 0.44	40,146	201	-0.71 ± 0.44
Embedded signal events sampled from $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ MC (LASS $m_{K\pi}$ lineshape):					
150	22.60	-5.10 ± 0.62	40,006	201	5.28 ± 0.62
100	20.16	-2.57 ± 0.56	40,056	201	2.74 ± 0.56
80	18.98	-1.24 ± 0.53	40,076	201	1.42 ± 0.53
60	17.67	-0.21 ± 0.50	40,096	201	0.35 ± 0.50
50	16.90	0.01 ± 0.50	40,106	201	0.14 ± 0.49
40	15.99	-0.26 ± 0.48	40,116	201	0.39 ± 0.48
30	14.97	-0.24 ± 0.47	40,126	201	0.35 ± 0.47
20	13.91	-0.01 ± 0.44	40,136	201	0.12 ± 0.43
10	12.70	0.15 ± 0.43	40,146	201	-0.07 ± 0.43
Embedded signal events sampled from $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ MC (Breit-Wigner $m_{K\pi}$ lineshape):					
60	16.80	-10.99 ± 0.51	40,096	201	11.12 ± 0.51
50	16.11	-9.15 ± 0.48	40,106	201	9.30 ± 0.48
40	15.28	-7.33 ± 0.45	40,116	201	7.45 ± 0.45
30	14.39	-5.50 ± 0.45	40,126	201	5.60 ± 0.45
20	13.49	-3.49 ± 0.45	40,136	201	3.60 ± 0.44
10	12.49	-1.79 ± 0.43	40,146	201	1.88 ± 0.43
Embedded signal events sampled from non-resonant $B^+ \to K^+K^-\pi^+$ MC:					
60	17.26	-5.71 ± 0.52	40,096	201	5.87 ± 0.52
50	16.53	-4.42 ± 0.50	40,106	201	4.57 ± 0.50
40	15.68	-3.57 ± 0.46	40,116	201	3.70 ± 0.46
30	14.69	-2.62 ± 0.46	40,126	201	2.72 ± 0.46
20	13.73	-1.35 ± 0.44	40,136	201	1.45 ± 0.44
10	12.73	-0.69 ± 0.43	40,146	201	0.78 ± 0.43

Table 5.27: *Biases on the signal and $q\overline{q}$ background yields for the 3-variable fit in the 892 window. These values are obtained using toy tests with embedded MC signal events. 1,000 experiments are run per toy test. Poissonian smearing is turned off.*

Number Of MC Signal Events Sampled	***Average Error On Signal Yield***	***β On Signal Yield***	***Number Of $q\overline{q}$ Background Events Generated***	***Average Error On $q\overline{q}$ Bkgd. Yield***	***β On $q\overline{q}$ Bkgd. Yield***
Embedded signal events sampled from $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ MC (LASS $m_{K\pi}$ lineshape):					
200	30.71	11.88 ± 0.91	46,352	217	-11.76 ± 0.85
150	28.63	9.95 ± 0.83	46,402	217	-9.35 ± 0.79
100	26.14	5.99 ± 0.77	46,452	217	-6.81 ± 0.76
80	24.87	3.23 ± 0.76	46,472	217	-4.19 ± 0.76
60	23.58	2.46 ± 0.69	46,492	217	-3.39 ± 0.71
40	21.95	0.25 ± 0.68	46,512	217	-0.89 ± 0.68
30	21.10	-1.12 ± 0.64	46,522	217	0.11 ± 0.64
20	20.29	-1.44 ± 0.64	46,532	217	0.35 ± 0.64
10	18.81	-4.48 ± 0.67	46,542	217	3.37 ± 0.67
5	18.44	-5.30 ± 0.66	46,547	217	4.17 ± 0.66
Embedded signal events sampled from $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ MC (Breit-Wigner $m_{K\pi}$ lineshape):					
200	30.93	18.27 ± 1.10	46,352	217	-17.54 ± 0.96
150	28.78	12.46 ± 0.95	46,402	217	-12.83 ± 0.91
100	26.30	9.18 ± 0.83	46,452	217	-9.82 ± 0.83
80	25.01	5.87 ± 0.87	46,472	217	-6.16 ± 0.85
60	23.75	4.17 ± 0.76	46,492	217	-5.10 ± 0.78
40	22.12	1.76 ± 0.71	46,512	217	-2.62 ± 0.72
30	21.17	-0.67 ± 0.73	46,522	217	-0.45 ± 0.72
20	20.11	-1.19 ± 0.73	46,532	217	0.19 ± 0.74
10	18.97	-4.29 ± 0.74	46,542	217	3.14 ± 0.73
5	18.36	-4.79 ± 0.70	46,547	217	3.69 ± 0.71
Embedded signal events sampled from non-resonant $B^+ \to K^+K^-\pi^+$ MC:					
80	24.74	0.16 ± 0.85	46,472	217	-0.84 ± 0.83
60	23.50	-0.62 ± 0.77	46,492	217	-0.28 ± 0.77
40	21.88	-1.61 ± 0.68	46,512	217	0.71 ± 0.68
30	21.04	-2.31 ± 0.73	46,522	217	1.38 ± 0.73
20	20.11	-2.82 ± 0.73	46,532	217	1.76 ± 0.72
10	18.80	-4.25 ± 0.72	46,542	217	3.13 ± 0.73
5	18.51	-4.96 ± 0.71	46,547	217	3.88 ± 0.72

Table 5.28: *Biases on the signal and $q\overline{q}$ background yields for the 3-variable fit in the 1430 window. These values are obtained using toy tests with embedded MC signal events. For the first section, 1,000 experiments are run per toy test. For the remaining two sections, 800 experiments are run per toy test. Poissonian smearing is turned off.*

	$(B\overline{B})_2$	$(B\overline{B})_3$	$(B\overline{B})_4$
m_{ES}	Gaussian + Argus	Gaussian + Argus	Double Gaussian + Argus
ΔE	Gaussian + Linear	Gaussian + Linear	Gaussian + Linear
$\mathcal{F}$	Double Gaussian	Gaussian	Gaussian

Table 5.29: *$B\overline{B}$ background PDF types used for the 3-variable fit in the 1430 window.*

Number Of $(B\overline{B})_{Other}$ ***Events Embedded***	β ***(Signal)***	β ***($q\overline{q}$ Bkgd.)***
3-variable fit in the 892 window (fit model components: S, $q\overline{q}$, $(B\overline{B})_1$, $(B\overline{B})_2$, $(B\overline{B})_3$):		
$n^{Other}_{B\overline{B}\ Bkgd}$	4.83 ± 0.51	146.86 ± 0.51
$n^{OtherMax}_{B\overline{B}\ Bkgd}$	7.14 ± 0.52	209.33 ± 0.52
3-variable fit in the 1430 window (fit model components: S, $q\overline{q}$, $(B\overline{B})_2$, $(B\overline{B})_3$, $(B\overline{B})_4$):		
$n^{Other}_{B\overline{B}\ Bkgd}$	6.16 ± 0.71	713.63 ± 0.71
$n^{OtherMax}_{B\overline{B}\ Bkgd}$	10.20 ± 0.70	$1,081.80 \pm 0.70$

Table 5.30: *Toy tests are conducted to determine the biases brought about on the floated yields due to the presence of events originating from $B\overline{B}$ background sources for which there are no components in the fit model ("$(B\overline{B})_{Other}$", "$n^{Other}_{B\overline{B}\ Bkgd}$" and "$n^{OtherMax}_{B\overline{B}\ Bkgd}$" are defined in Section 5.10.6). Events due to the modes $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$ and $B^+ \to \overline{D}^0(\to K^+K^-)\pi^+$ are not embedded in the toy tests for the 1430 window, despite contributing $n_{B\overline{B}\ Bkgd}+\Delta n_{B\overline{B}\ Bkgd}$ > 0.5 events (see main text, Section 5.11). 1,000 experiments are run per toy test. Poissonian smearing is turned off.*

6

Analysis results

6.1 Introduction

This chapter presents the results of the analysis described in the previous chapter. The raw yields obtained from the fits to the final on-resonance samples are first presented. These yields are then amended to account for non-$K^+K^-\pi^+$-final-state $B\overline{B}$ background. Sources of systematic uncertainties are examined and their values determined from various studies. The statistical errors on the yields extracted from the fits are also discussed. Finally the branching ratio values are presented, and $\mathcal{B}(B^+ \to \overline{K}^*(892)^0 K^+)$ is used to calculate an improved bound on $\Delta S_{\phi(1020)K^0_S}$.

6.2 Yields

Table 6.1 presents the values obtained for the floated yields for each of the fits conducted. The values obtained for the floated PDF parameters are presented in Appendix D. Correlations between floated PDF parameters and yields are presented in Appendix H. The statistical significance is calculated using $\sqrt{-2\ln(\mathcal{L}_{Max}/\mathcal{L}_{Max}^{S=0})}$ where $\mathcal{L}_{Max}^{S=0}$ is the maximum value of the likelihood function obtained when carrying out the standard extended ML fit, with the exception that the signal yield is fixed to zero.

Fit	***# Events In Final Onres. Sample***	***Signal Yield***	***Stat. Sig. (σ)***	***$q\overline{q}$ Background Yield***	***Fixed Yields***
5-variable fit in the 892 window	$38,690$	$30.1^{+13.0}_{-11.8}$	3.09	$38,609 \pm 197$	$(B\overline{B})_1$: 20.6 $(B\overline{B})_2$: 23.0 $(B\overline{B})_3$: 4.3
3-variable fit in the 892 window	$40,156$	$45.9^{+18.8}_{-17.5}$	2.95	$40,058 \pm 201$	$(B\overline{B})_1$: 20.6 $(B\overline{B})_2$: 23.0 $(B\overline{B})_3$: 4.8
3-variable fit in the 1430 window	$46,657$	$34.7^{+23.8}_{-21.6}$	1.69	$46,503 \pm 217$	$(B\overline{B})_2$: 89.4 $(B\overline{B})_3$: 3.5 $(B\overline{B})_4$: 12.0

Table 6.1: *Yields obtained from the extended ML fits. Systematic errors are not shown, and corrections due to $B\overline{B}$ background have not yet been applied.*

Projection plots of the fit onto the fit variables are shown in Figures 6.1-6.3. For

illustrative purposes the signal component is enhanced. For each variable, only events which pass a cut on a signal likelihood to background likelihood ratio are plotted. The per event likelihoods used to calculate this ratio are functions of all fit variables except the projected variable that is plotted (to avoid bias). The blue, superimposed curves are a projection onto the variable of the fit model. Fit model components are also projected.

Figure 6.4 shows the contributions from each component for the toy sample created for the second test described in Section 5.10.7. It can be seen that signal events are concentrated at high values of the ratio "Likelihood (Signal) / Likelihood (Total)", as expected.

Scans of the negative log likelihood (NLL) values as a function of the number of fitted signal events are shown in Figure 6.5. Each curve is obtained by conducting 400 extended ML fits to the final on-resonance sample, with the signal yield fixed to values between 0 and 100. The fits are otherwise standard (i.e. as those conducted to give the results shown in Table 6.1). A polynomial is fitted to the signal yield versus minimum NLL points to give the curves.

6.2.1 Statistical errors

The errors associated with the PDF parameters and yields that are floated in the fits are now briefly discussed. There are two different ways that these errors can be calculated during the fitting procedure.

`Minuit`'s `HESSE` routine, introduced in Section 5.6.2, uses the covariance matrix of the floated parameters, $\mathbf{H}$, the elements of which are given by

$$H_{ij} = \frac{\partial^2(\ln \mathcal{L})}{\partial\theta_i \, \partial\theta_j}. \tag{6.1}$$

The vector of parameter errors is then given by

$$\vec{\sigma}_\theta = \mathbf{H}^{-1}\vec{\Theta}. \tag{6.2}$$

The error for a given parameter calculated by `Minuit`'s `MINOS` routine – also introduced in Section 5.6.2 – is defined as the change in the value of that parameter that causes the minimum value of the negative log likelihood function to increase by an amount ℓ. ℓ is set such that the probability of the true parameter value lying between the negative and positive errors is 68.3%. For example, for the floated signal yields, the `MINOS` algorithm follows the curves shown in Figure 6.5 from their minima, in both directions, until it reaches the two signal yield values that give minimum negative log likelihood function values that are larger than the original minima by ℓ.

Both `HESSE` and `MINOS` take into account the correlations between floated parameters. Where they differ is that `HESSE` simply calculates the covariance matrix (from which the errors are obtained) at $(-\ln \mathcal{L}_{Max})$ and assumes that the negative log likelihood curve is parabolic about the minimum. `MINOS` on the other hand iterates over the path of the negative log likelihood curve. For a fit model that is a linear function of all of its floated parameters, and with an infinite amount of data, `HESSE` and `MINOS` would yield identical values for parameter errors. In reality, the difference between the three error values (the two asymmetric `MINOS` errors and the symmetric `HESSE` error) are a reasonable measure of the non-linearity of the problem.

Where symmetric errors are presented in the results of fits in this thesis, the `HESSE` routine has been employed. This is usually the case, as `MINOS` errors are considerably more expensive – computationally – to calculate, and `HESSE` and `MINOS` errors are generally in good agreement. For the signal yields, however, `MINOS` errors are used as these values are used to calculate the main results of the analysis (see Table 6.1; for comparison, the `HESSE` errors are 30.1 ± 12.4, 45.9 ± 18.2 and 34.7 ± 22.6).

For further details, please refer to [92] and [93].

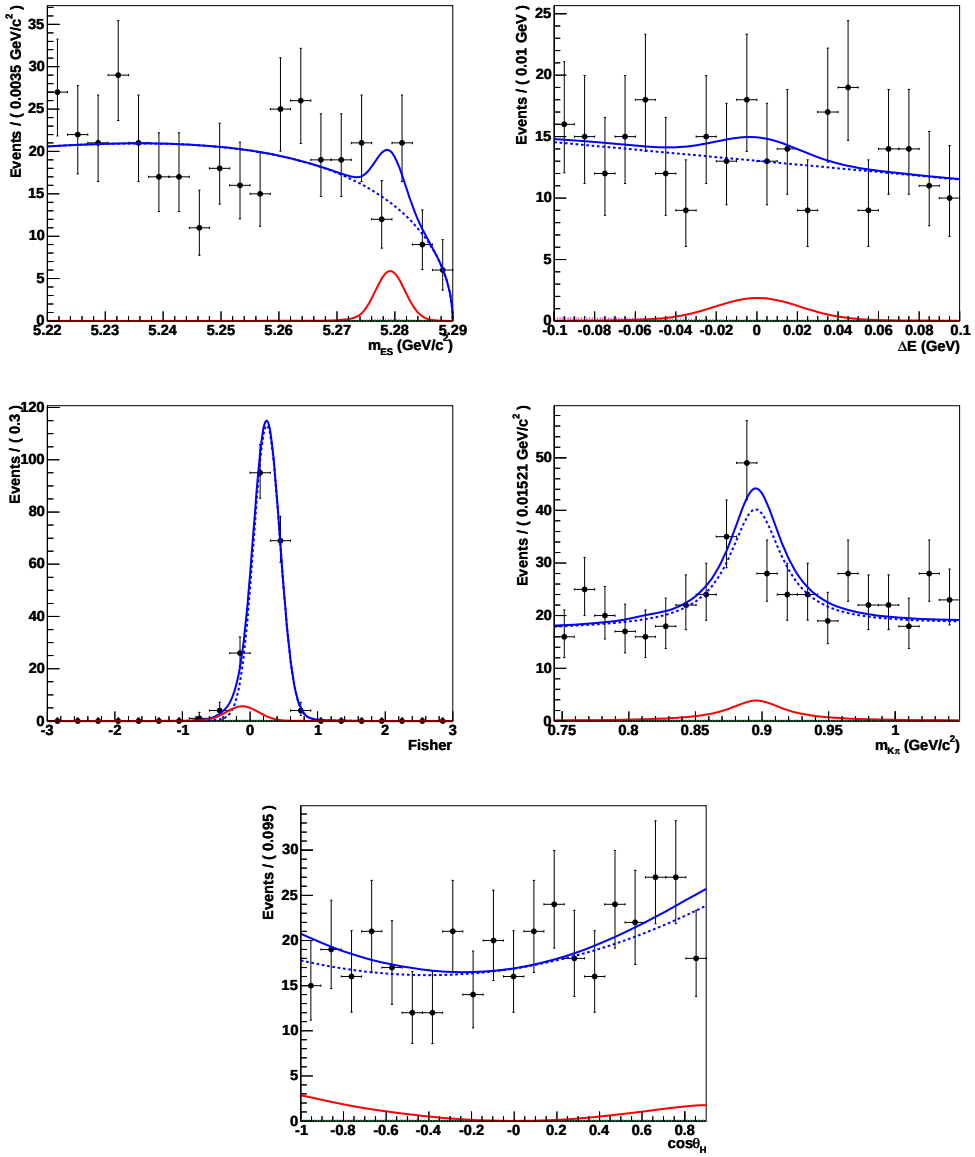

Figure 6.1: *Projection plots for the 5-variable fit in the 892 window. Fit variable distributions are shown (black points with error bars) for final on-resonance sample candidates that pass a likelihood ratio, which favours signal-like events. The solid, blue line represents the fit model. The solid, red line represents the signal component, whilst the dashed, blue line represents the $q\overline{q}$ background component. Also superimposed, though not clearly visible due to their low integrated values, are lines representing the $B\overline{B}$ background components.*

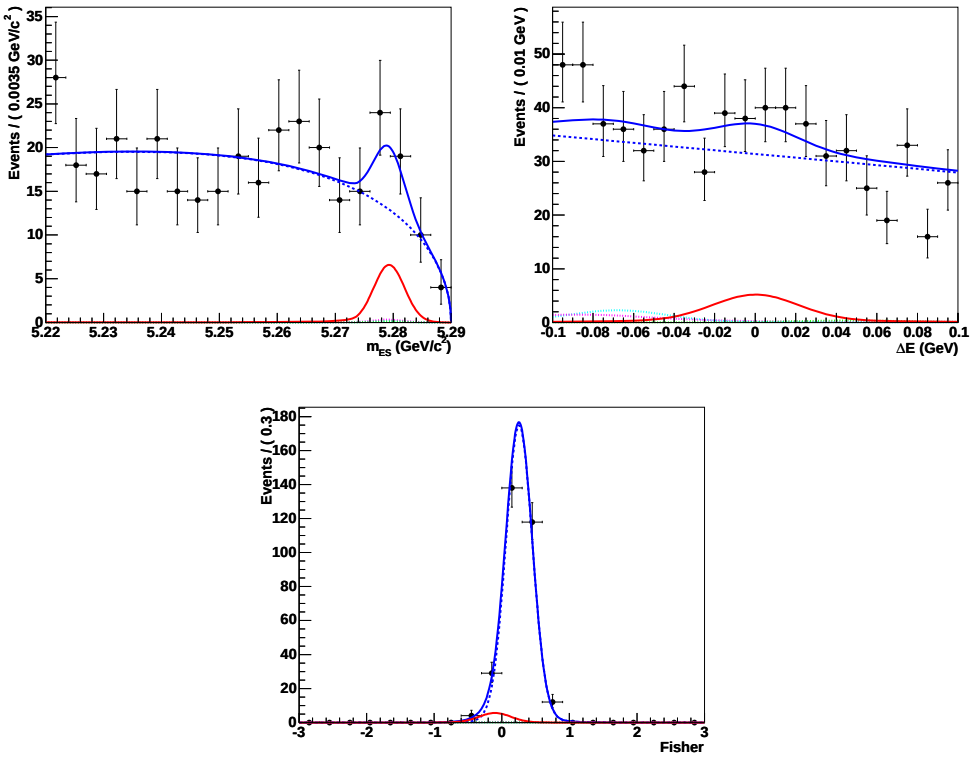

Figure 6.2: *Projection plots for the 3-variable fit in the 892 window. Marker/line colours/styles are as Figure 6.1, with* $(B\overline{B})_1$*: dotted, cyan line;* $(B\overline{B})_2$*: dotted, magenta line;* $(B\overline{B})_3$*: dotted, green line.*

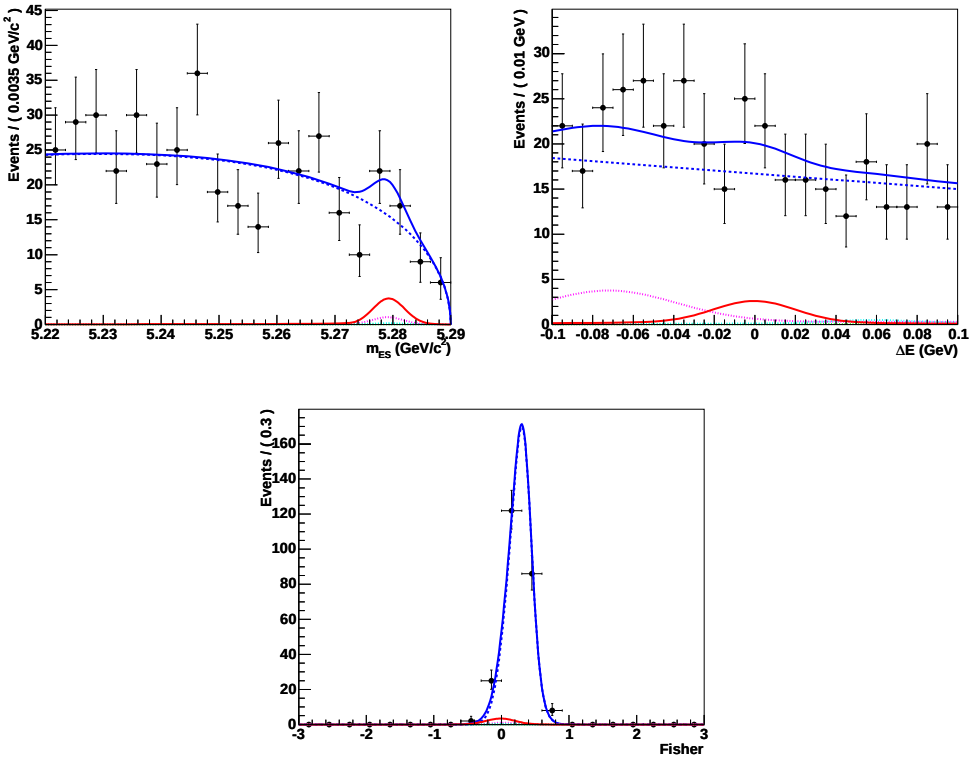

Figure 6.3: *Projection plots for the 3-variable fit in the 1430 window. Marker/line colours/styles are as Figure 6.1 and Figure 6.2 with the exception that the dotted, cyan line (barely visible) represents the* $(B\overline{B})_4$ *component.*

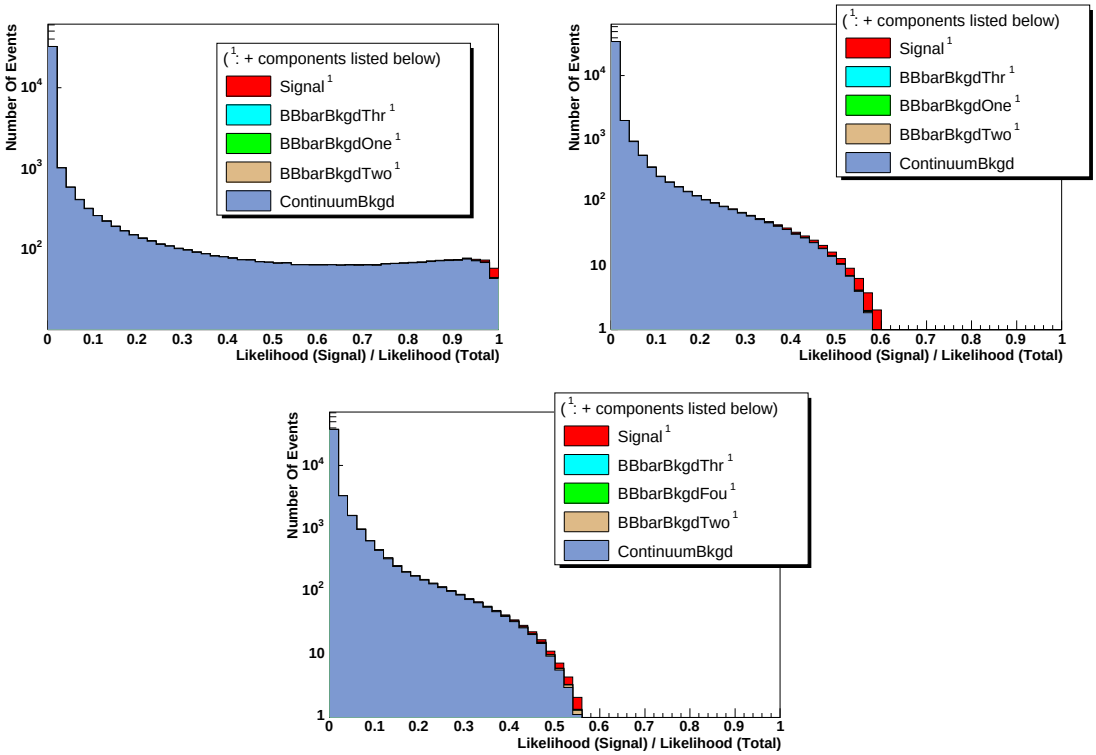

Figure 6.4: *Top left: 5-variable fit in the 892 window. The event likelihood ratio distribution for toy MC as shown in Figure 5.17 (bottom plot), this time split into components. The parameters and yields used to generate the toy MC are those obtained from applying the fit model to the final on-resonance sample. Top right and bottom: 3-variable fits in the 892 window and 1430 window respectively (see also Figure G.3).*

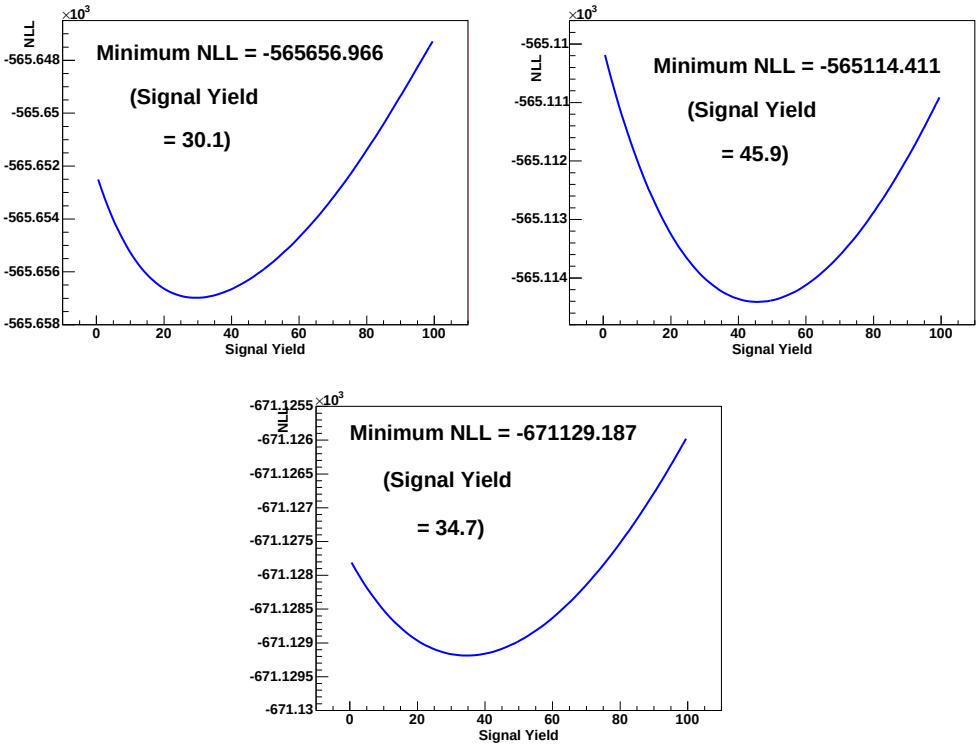

Figure 6.5: *Negative log likelihood scans. Top left: 5-variable fit in the 892 window; top right and bottom: 3-variable fits in the 892 window and 1430 window respectively.*

6.3 $B\overline{B}$ background subtraction

The discussion presented in this section (Section 6.3) applies to $B\overline{B}$ background modes that do not share the same final state as the signal mode. $B\overline{B}$ background with the final state $K^+K^-\pi^+$ is discussed in Section 6.6.

It is estimated that the signal yields presented in Table 6.1 are biased due to the presence of $B\overline{B}$ background events originating from sources for which components are not included in the fit model. The magnitudes of these biases are estimated using toy tests as described in Sections 5.10.6 and 5.11 (in which the numbers of events sampled/generated for each hypothesis correspond to the yields obtained/fixed in the fits to the final on-resonance samples). For the 1430 window, the modes $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$ and $B^+ \to \overline{D}^0(\to K^+K^-)\pi^+$ must also be accounted for as discussed in Section 5.11.

Table 6.2 lists the signal yields extracted from the extended ML fits. It also lists the number of these events that are estimated to be $B\overline{B}$ background, and the corrected signal yields. Excluding the modes $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$ and $B^+ \to \overline{D}^0(\to K^+K^-)\pi^+$ (the uncertainties for which are discussed in Section 5.11), it can be seen that a 100% systematic uncertainty has been applied to the number of estimated $B\overline{B}$ background events for each of the fits. This highly conservative error accounts for the following possible sources of uncertainty[1]:

- The number of $(B\overline{B})_{Other}$ MC events embedded in the toy tests described in Section 5.10.6 is not exactly equal to $n^{Other}_{B\overline{B}\ Bkgd}$, rather it is rounded to the nearest non-zero integer.
- Where the number of MC events remaining after the application of RaSC is small for a given $B\overline{B}$ background mode, oversampling occurs. MC events may be embedded more than once for a single toy test (but not more than once for a single experiment). Oversampling is kept to a minimum, and the effect, if

[1] $n_{B\overline{B}\ Bkgd}$ is defined in Section 5.5.2.5; $(B\overline{B})_{Other}$ and $n^{Other}_{B\overline{B}\ Bkgd}$ are defined in Section 5.10.6.

any, is expected to be very small to negligible.

- The uncertainties on the values of $n_{B\overline{B}\ Bkgd}$, as listed in Tables 5.8 and 5.9. For the toy tests in which $n_{B\overline{B}\ Bkgd} \pm \Delta n_{B\overline{B}\ Bkgd}$ are embedded rather than $n_{B\overline{B}\ Bkgd}$, the difference between the corresponding signal yield biases is well within the assigned 100% systematic.

- $B\overline{B}$ background modes that contribute less than $0.5 - \Delta n_{B\overline{B}\ Bkgd}$ events to the final on-resonance sample are not studied using toy tests. It is possible that, combined, some of these modes may induce a very small – but not negligible – bias on the signal yield.

6.3.1 Further corrections

For the 3-variable fits, biases were observed for toy tests conducted with embedded signal MC events, as described in Section 5.11. It can be seen from Tables 5.27 and 5.28 that for both the 892 and the 1430 windows the biases on the signal yield are both positive and negative. A correction will therefore not be applied. However, a systematic uncertainty of ± 5 events will be applied (to both windows) to account for these biases.

6.4 Systematic uncertainty estimation

In this section sources of systematic uncertainty are considered, and techniques for estimating their values are described.

6.4.1 Systematics on the yields

Systematic uncertainties on the signal yields are listed in Table 6.4 and detailed below.

Fit	***Initial Signal Yield***	$B\overline{B}$ ***Background***	***Amended Signal Yield***
5-variable fit in the 892 window	$30.1^{+13.0}_{-11.8}$	3.2 ± 3.2 $^{\triangle}$	$26.8^{+13.0}_{-11.8} \pm 3.2$
3-variable fit in the 892 window	$45.9^{+18.8}_{-17.5}$	4.8 ± 4.8 $^{\triangle}$	$41.1^{+18.8}_{-17.5} \pm 4.8$
3-variable fit in the 1430 window	$34.7^{+23.8}_{-21.6}$	(6.2 ± 6.2) $^{\triangle}$ $+(5.3 \pm 5.3)$ $^{\square}$ $+(9.4 \pm 1.6)$ $^{\diamond}$ $= 20.9 \pm 8.3$	$13.8^{+23.8}_{-21.6} \pm 8.3$

$^{\triangle}$: Source is $(B\overline{B})_{Other}$ – discussed in Sections 5.10.6 and 5.11. See Tables 5.25 and 5.30.

$^{\square}$: Source is $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$ – discussed in Section 5.11.

$^{\diamond}$: Source is $B^+ \to \overline{D}^0(\to K^+K^-)\pi^+$ – discussed in Section 5.11.

Table 6.2: *Corrections to signal yields due to non-$K^+K^-\pi^+$-final-state $B\overline{B}$ background. Only systematic errors due to these corrections are shown.*

6.4.1.1 $B\overline{B}$ background

The yields of the $B\overline{B}$ background components included in the fit models are fixed. There is an uncertainty on the values to which they are fixed (see Table 5.8). To estimate the resulting signal yield uncertainties, each $B\overline{B}$ background yield, one at a time, is shifted by plus and minus its error, and the otherwise standard fit is reperformed. The total uncertainties on the signal yields are estimated to be:

- 5-variable fit in the 892 window: $^{+0.4}_{-0.6}$ events.
- 3-variable fit in the 892 window: $^{+0.9}_{-1.5}$ events.
- 3-variable fit in the 1430 window: $^{+2.8}_{-1.5}$ events.

We arrive at these values by adding in quadrature the differences between the signal yield obtained in the standard fit and the signal yields obtained for each of the fits in which the $B\overline{B}$ background yields are varied.

The PDF parameters of the $B\overline{B}$ background components are also fixed. This is discussed in the next section. The systematic uncertainties brought about by sources of $B\overline{B}$ background for which components are not included in the fit are discussed in Section 6.3.

6.4.1.2 Fixed PDF parameters

Whilst the $q\overline{q}$ background PDF parameterisations are determined from the fit to data, the PDF shapes and parameter values for the signal and $B\overline{B}$ background contributions are determined from simulation. It is necessary to fix these parameter values in the fit to data to avoid an unstable fit in which too much information is requested from too little data. In doing so however, we introduce systematic uncertainty. To estimate this uncertainty each fixed PDF parameter in the fit model is varied by plus and minus its error (one standard deviation). Correlated parameters

in the PDF are adjusted accordingly. The otherwise standard fit is then repeated with these new PDF parameter values. The resulting shift in the signal yield is taken to be the systematic error. This is repeated for all fixed PDF parameters. The positive and negative shifts are added separately in quadrature to give the total systematic uncertainty.

Previous analyses indicate that there is a small discrepancy between the mean value of the ΔE distribution for data and for MC. The analysis of $B^+ \to K^+K^-K^+$ described in [88] – kinematically similar to the channel in this analysis, and with similar selection criteria – report a shift of 8.3 MeV. In addition to varying PDF parameters according to their errors, fits are also reperformed with parameter values reflecting this data-MC discrepancy.

The overall estimated uncertainties due to fixing the signal and $B\overline{B}$ background PDF parameters are:

- 5-variable fit in the 892 window: $^{+1.7}_{-0.6}$ events.
- 3-variable fit in the 892 window: $^{+1.6}_{-0.8}$ events.
- 3-variable fit in the 1430 window: $^{+2.6}_{-2.2}$ events.

6.4.1.3 Fixed SXF fraction

The SXF fraction, used in the 5-variable fit in the 892 window, is also fixed. As with the PDF parameters, its value is obtained from MC. To date, no studies have been performed that estimate the efficiency of the truth matching algorithm. A deviation from 100% will result in an overestimation of the SXF fraction. An underestimation could result if, for example, candidate multiplicity is underestimated by the MC. In such a case the likelihood of choosing the true B would be overestimated. We choose to vary the SXF fraction by a highly conservative $\pm 20\%$ (compare this with, for example, [81]). The shift in the fitted number of signal events is then found to be:

- 5-variable fit in the 892 window: ± 0.7 events.

6.4.2 Systematics on the reconstruction efficiency

The efficiency of RaSC on the number of signal events is obtained from MC (see Tables 5.4 and 5.5). MC does not perfectly model real data and so it is necessary to conduct studies to determine whether corrections need to be made to account for discrepancies. There are uncertainties associated with such corrections. There are also uncertainties where a correction is not deemed necessary as we cannot be 100% confident that a correction is not required.

6.4.2.1 Tracking and PID

A small difference in efficiency is observed when applying the track finding and PID algorithms – discussed in Sections 4.2.1 and 4.2.3 – to samples of MC and to samples of data. To correct for these differences control samples are studied and the efficiency of reconstruction is tabulated for both the data and MC samples in bins of momentum, polar and azimuthal angles and track multiplicity. When MC events are reconstructed these tables are retrieved from a database and used to correct the efficiency for the event. As such the efficiencies shown in Tables 5.4-5.5, 5.8-5.11, and those in Appendix B, have been subjected to tracking and PID corrections. Each correction has an associated systematic error. For tracking, the value of this error for $B^+ \to K^+K^-\pi^+$ is given as

- $\pm 2.4\%$,

which is a linear addition of 0.8% per track [99]. For PID we use the variables ϵ_{Data} and ϵ_{MC}, which are defined to be the efficiencies for candidates passing the PID selector for the two samples. If $\epsilon_{Data} < \epsilon_{MC}$, an MC track accepted before the correction has a probability $1 - \frac{\epsilon_{Data}}{\epsilon_{MC}}$ of being rejected. If $\epsilon_{Data} > \epsilon_{MC}$, a rejected

MC track has a probability $1 - \frac{1-\epsilon_{Data}}{1-\epsilon_{MC}}$ of being accepted. The systematic error for the PID correction is calculated from control sample studies and is found to be 1.4% per corrected track [100], which, for $B^+ \to K^+K^-\pi^+$, add linearly to give a total of

- $\pm 4.2\%$.

6.4.2.2 Selection criteria

The efficiencies of the cuts on our discriminating variables are also contenders for possible corrections. The cuts on the variables m_{ES}, ΔE and $\cos\theta_T$ for the *BABAR* analysis $B^+ \to K^+K^-K^+$ [88] are very similar to those in this analysis. In the $B^+ \to K^+K^-K^+$ analysis the authors perform a study using the calibration channel $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$. This mode is kinematically very similar to both $B^+ \to K^+K^-\pi^+$ and $B^+ \to K^+K^-K^+$, and the same tools that are used to study these modes can also be used to analyse $B^+ \to \overline{D}^0(\to K^+\pi^-)\pi^+$. Its branching ratio however is much larger at $(189 \pm 12) \times 10^{-6}$ [23]. This allows accurate measurements of the efficiencies of the applied cuts on signal events from both data and MC, and corrections can be obtained. The correction factors obtained in [88] are consistent with unity when taking into account errors (of less than 5%). In a previous (cut and count) analysis of $B^+ \to \overline{K}^*(892)^0K^+$, the author found that the correction factor on the efficiency of $\cos\theta_H$ was also consistent with unity within errors [6]. To estimate a correction factor for the cut on $m_{K\pi}$, Eq. (6.3) is used:

$$\text{Correction factor} = \frac{\int_{0.744\,\text{GeV}/c^2}^{1.048\,\text{GeV}/c^2} \left((\mu_{PDG} - m_{K\pi})^2 + \frac{\Gamma_{PDG}^2}{4}\right)^{-1} dm_{K\pi}}{\int_{0.744\,\text{GeV}/c^2}^{1.048\,\text{GeV}/c^2} \left((\mu_{PDF} - m_{K\pi})^2 + \frac{\Gamma_{PDF}^2}{4}\right)^{-1} dm_{K\pi}}, \qquad (6.3)$$

where μ_{PDG} and Γ_{PDG} are the world average values for the mass and width of the $K^*(892)^0$ [23], and μ_{PDF} and Γ_{PDF} are the values for the mass and width obtained from fitting a Breit-Wigner shape to $B^+ \to \overline{K}^*(892)^0K^+$ MC (as discussed in Section 5.7.6) – in which detector resolution is taken into account. A value of 0.998 ± 0.002 is obtained. The cut on $\mathcal{F}$ of $< |3.0|$ is very loose as can be seen in Figure 4.5. A correction is therefore not deemed necessary.

It is decided that it is not necessary to apply a correction to the efficiencies of the selection criteria to which our discriminating variables are subjected. Furthermore, we can be conservative when assigning a systematic error to cover the uncertainty of this non-action, since the errors on the fit yields are considerably larger (Table 6.1). A highly conservative[2] value of

- $\pm 8.8\%$

is chosen. Adding in quadrature with the uncertainties for tracking and PID gives a total systematic uncertainty on the efficiency (ϵ in Eq. (5.1)) of $\pm 10\%$.

6.4.3 Systematics on the number of B events

From Eq. (4.11) it can be seen that the systematic uncertainty on $N_{B^\pm}$ is

- $\pm 1.9\%$.

6.5 Upper limits on $\mathcal{B}(B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)\ K^+)$ and $\mathcal{B}(B^+ \to K^+K^-\pi^+$ non-resonant$)$

The signal yield for the 3-variable fit in the 1430 window can be used to place an upper limit on the branching ratio of the process $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$. In assigning an upper limit to this branching ratio, we assume that all events that constitute the signal yield are $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ events. We also assume that $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ does not interfere with other $B^+ \to K^+K^-\pi^+$ events. The latter assumption, although unlikely to be true, is an acceptable approximation for the situation described.

[2]The systematic uncertainties on the selection criteria efficiencies for the *BABAR* $B^+ \to K^+K^-K^+$ [88] and $B^+ \to K^+\pi^-\pi^+$ [50] analyses are $\approx 5\%$.

Since the fit does not use information on intermediate states, the signal yield can also be used to place an upper limit on the branching ratio of the non-resonant channel $B^+ \to K^+K^-\pi^+$. The same assumptions are made for non-resonant $B^+ \to K^+K^-\pi^+$ as those described above for $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$.

The method used to produce the NLL scans (Figure 6.5) described in Section 6.2 is again employed. This time the normalised maximum likelihood values (e^{-NLL+C}) are plotted as a function of the branching ratio (given by Eq. (5.1)), and the curves are broadened to account for systematic errors using the method described in Appendix G.1 of [101]. The resultant likelihood scans are presented in Figure 6.6.

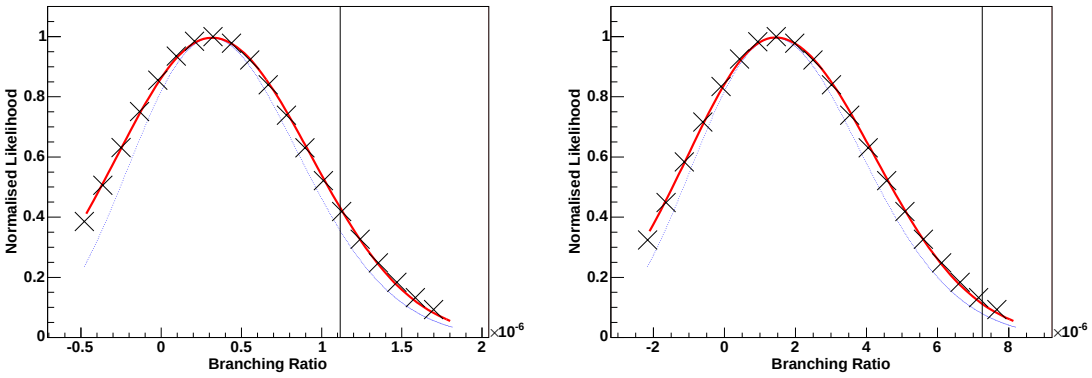

Figure 6.6: *For each plot, the black crosses represent 20 of 400 likelihood values obtained from varying the branching ratio. The red curve is a bifurcated Gaussian fitted to all 400 of these points. The black markers/red curve take into account systematic uncertainties (including the uncertainty on the $m_{K\pi}$ lineshape). The blue points are the 400 likelihood values before systematic uncertainties are considered. For the left (right) plot, the signal yield obtained for the 3-variable fit in the 1430 window is taken to be solely due to the process $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ (non-resonant $B^+ \to K^+K^-\pi^+$). Zero interference with other $B^+ \to K^+K^-\pi^+$ contributions is assumed. The vertical, black lines indicate 90% CL upper limits on $\mathcal{B}(B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+)$ and on $\mathcal{B}(B^+ \to K^+K^-\pi^+$ non-resonant$)$, obtained by integrating the red curves.*

The 90% CL upper limit for the branching ratio is defined as that point on the likelihood curve where the integral from zero to that point equals 90% of the total area under the curve between $\mathcal{B} = 0$ and $\mathcal{B} = \infty$. The 90% CL upper limits are found to be:

- $\mathcal{B}(B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+) < 1.2 \times 10^{-6}$.
- $\mathcal{B}(B^+ \to K^+K^-\pi^+ \text{ non-resonant}) < 7.3 \times 10^{-6}$.

The upper limit on $\mathcal{B}(B^+ \to K^+K^-\pi^+$ non-resonant) obtained using the above method offers no improvement on the current upper limit of 6.3×10^{-6} [89,102]. This is unsurprising. The technique used here is far from optimal for such a measurement.

6.6 $\mathcal{B}(B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+)$

Before converting the signal yield from the 5-variable fit in the 892 window into a branching ratio for $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$, $B\overline{B}$ background with the same final state, namely $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ and non-resonant $B^+ \to K^+K^-\pi^+$, must be taken into account.

Embedded toy tests are conducted using the same components as in the standard fit. The number of events embedded (generated for the $q\overline{q}$ component) is consistent with the yields obtained from the standard fit. Also added to the mixture are MC events simulating $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ and non-resonant $B^+ \to K^+K^-\pi^+$ in quantities determined by the branching ratio central values and upper limits obtained for the 3-variable fit in the 1430 window (see Section 6.5). The biases induced on the signal yield due to the presence of these same-final-state $B\overline{B}$ background events are shown in Table 6.3 and are used to make an appropriate subtraction from the signal yield obtained from the standard fit. This correction, and the associated systematic uncertainty is set at:

- $(-5.3^{+5.3}_{-6.6})$ events.

Mode And Number Of Events Embedded		***Bias*** β ***(Signal)***	***Bias*** β ***(*** $q\overline{q}$ ***Bkgd.)***
$B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$ $^\beta$	$2^\dagger$	0.71 ± 0.55	1.32 ± 0.56
$B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$ $^\beta$	$9^\ddagger$	3.11 ± 0.54	5.92 ± 0.51
$B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$ $^\lambda$	$5^\dagger$	1.96 ± 0.53	2.98 ± 0.52
$B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$ $^\lambda$	$19^\ddagger$	7.75 ± 0.55	11.22 ± 0.54
$B^+ \to K^+K^-\pi^+$ (non-resonant)	$9^\dagger$	3.33 ± 0.54	5.64 ± 0.54
$B^+ \to K^+K^-\pi^+$ (non-resonant)	$17^\ddagger$	6.48 ± 0.56	10.44 ± 0.53

Table 6.3: *Biases obtained from toy tests with embedded same-final-state* $B\overline{B}$ *background events for the 5-variable fit in the 892 window. Fit model components are* S, $q\overline{q}$ *and* $(B\overline{B})_i$ *(*$i = 1, 2, 3$*). 500 experiments are run per toy test. Poissonian smearing is turned off.*

$^\beta$*:* $m_{K\pi}$ *lineshape modelled as Breit-Wigner;* $^\lambda$*:* $m_{K\pi}$ *lineshape modelled as LASS.*

‡*: Branching ratio value used to calculate expected number of events is upper limit presented in Section 6.5;* †*: branching ratio value used to calculate expected number of events is central value obtained for 3-variable fit in 1430 window.*

It is also necessary to consider the possible effect of interference (discussed in Section 1.4.2). A study is performed in which toy MC events are generated using the `Laura++` Dalitz plot package [103]. Several hundred thousand datasets are generated. The total number of events generated is the same for each dataset, and is equal to the total number of observed signal events in our two windows (the 892 window and the 1430 window) from the final on-resonance samples. Events are generated only in the area of the Dalitz plot that corresponds to the 892 and 1430 windows. $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ events are generated according to the Breit-Wigner distribution (Eq. (1.96)), whilst $B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$ events are generated according to the LASS distribution (Eq. (1.97)). The relative magnitudes and

phases of the two amplitudes are varied randomly between datasets. If the number of events observed in the 892 and 1430 windows in the resultant dataset is not consistent with our signal yields from the final on-resonance samples, the dataset is discarded. This accounts for the majority of datasets generated. For those that remain we are interested in the values of the fraction

$$f_{892} = \frac{\int\int_{\mathrm{DP}} \left| c_{892}\, a_{892}\left(m^2_{K\pi}, m^2_{KK}\right)\right|^2 \quad d\left(m^2_{K\pi}\right)\, d\left(m^2_{KK}\right)}{\int\int_{\mathrm{DP}} \left| c_{892}\, a_{892}\left(m^2_{K\pi}, m^2_{KK}\right) + c_{1430}\, a_{1430}\left(m^2_{K\pi}, m^2_{KK}\right)\right|^2 \quad d\left(m^2_{K\pi}\right)\, d\left(m^2_{KK}\right)} \tag{6.4}$$

(plotted in Figure 6.7), which cause us to observe the yields that we do in our two windows. a_{892} and a_{1430} are the dynamical amplitudes for $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ and $B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$ respectively (as determined by the invariant mass lineshapes and $\cos\theta_H$ distributions), and c_{892} and c_{1430} are complex coefficients (the magnitudes and phases of which are varied randomly). "DP" indicates that we are integrating over the $B^+ \to K^+K^-\pi^+$ Dalitz plot.

The fractional systematic uncertainty due to possible interference is then taken to be the standard deviation of the random variable f_{892} divided by its mean value.

The procedure is repeated for the non-resonant contribution, and the uncertainty is found to be smaller than for the 1430 contribution. The uncertainty obtained from the 1430 study described above,

- $\pm 8.3\%$,

is therefore used as the systematic uncertainty for interference.

The signal yield to be used in the calculation of $\mathcal{B}(B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+)$ can now be quoted:

- $N_S = 21.6^{+13.0+6.7}_{-11.8-7.6}$.

This value is plugged into Eq. (5.1) together with

- $\epsilon = 0.266 \pm 0.001 \pm 0.027$, and

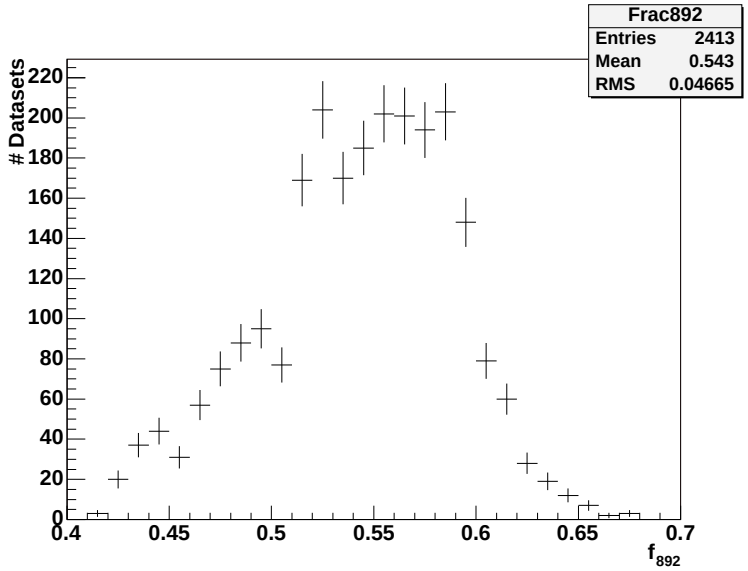

Figure 6.7: *Interference study: distribution of the variable f_{892} (Eq. (6.4)) for datasets in which the phases and magnitudes of the $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ and $B^+ \to \overline{K}^*_0(1430)^0(\to K^-\pi^+)K^+$ amplitudes are such that the number of events observed in the 892 and 1430 windows are equal to the signal yields obtained from the final on-resonance samples.*

- $N_{B^\pm} = (232.28 \pm 4.14 \pm 4.39) \times 10^6$,

to give a central value on the branching ratio of $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ of

- $\mathcal{B}(B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+) = (0.35^{+0.21+0.11}_{-0.19-0.13}) \times 10^{-6}$.

The first uncertainty is statistical whilst the second is systematic. The significance of this result is taken to be $\chi(\mathcal{B} = 0)$ where

$$\chi^2(\mathcal{B}) = \frac{\chi^2_{stat}(\mathcal{B})\,\chi^2_{syst}(\mathcal{B})}{\chi^2_{stat}(\mathcal{B}) + \chi^2_{syst}(\mathcal{B})}, \tag{6.5}$$

$$\chi^2_{stat}(\mathcal{B}) = -2\ln\left(\frac{\mathcal{L}_{Max}(\mathcal{B}_{Nom})}{\mathcal{L}_{Max}(\mathcal{B})}\right), \text{ and} \tag{6.6}$$

$$\chi^2_{syst}(\mathcal{B}) = \frac{\mathcal{B} - \mathcal{B}_{Nom}}{\sigma_{syst}}. \tag{6.7}$$

$\mathcal{B}_{Nom}$ is the branching ratio that gives the maximum value of $\mathcal{L}_{Max}$, i.e. the central value quoted above. σ_{syst} is the systematic uncertainty and $\mathcal{L}$, recall, is the likelihood

function defined in Eq. (5.12). See Section 6.1.3 of [104] for further information regarding Eq. (6.6). The significance obtained using Eq.s (6.5)-(6.7) is

- 1.61σ.

A likelihood curve analogous to those in Figure 6.6 is constructed for $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$, and is used to place an upper limit on $\mathcal{B}(B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+)$:

- $\mathcal{B}(B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+) < 0.7 \times 10^{-6}$.

This curve is shown in Figure 6.8.

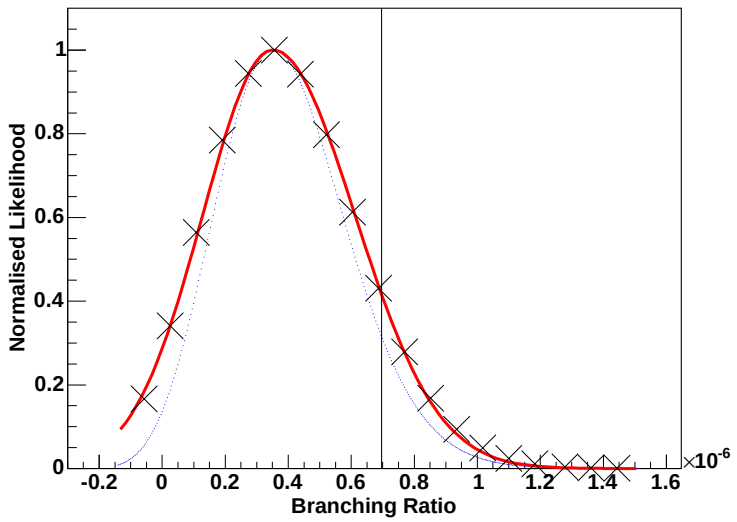

Figure 6.8: *Maximum likelihood values as a function of branching ratio for the 5-variable fit in the 892 window. The black crosses/red curve take into account systematic uncertainties whilst the blue points do not. The vertical, black line represents the 90% CL upper limit on* $\mathcal{B}(B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+)$.

Source	***Systematic Uncertainty On Signal Yield (Number Of Events)***		
	5-var. Fit In 892 Window	***3-var. Fit In 892 Window***	***3-var. Fit In 1430 Window***
Fit procedure (Section 6.3.1)	—	± 5.0	± 5.0
Fixed PDF parameters (Section 6.4.1.2)	$^{+1.7}_{-0.6}$	$^{+1.6}_{-0.8}$	$^{+2.6}_{-2.2}$
Fixed $B\overline{B}$ background yields (Section 6.4.1.1)	$^{+0.4}_{-0.6}$	$^{+0.9}_{-1.5}$	$^{+2.8}_{-1.5}$
Subtraction of $B\overline{B}$ background events for which components are not included in fit model[†] (Section 6.3)	± 3.2 $^{+5.3}_{-6.6}$ ‡	± 4.8	± 8.3
Fixed SXF fraction (Section 6.4.1.3)	± 0.7	—	—
Interference (Section 6.6)	± 1.8	—	—
Total	$\mathbf{^{+6.7}_{-7.6}}$	$\mathbf{^{+7.2}_{-7.1}}$	$\mathbf{^{+10.4}_{-10.0}}$

[†]: Final state $B\overline{B}$ background is not $K^+K^-\pi^+$, unless otherwise stated.

[‡]: Final state $B\overline{B}$ background is $K^+K^-\pi^+$ (Section 6.6).

Table 6.4: *Sources of systematic errors on the signal yields. The total systematic errors are obtained by adding the individual contributions in quadrature.*

6.7 $\Delta S_{\phi(1020)K^0_S}$ bound

Taking $\mathcal{B}(\overline{K}^*(892)^0 \to K^-\pi^+) = \frac{2}{3}$, and using the result presented in Section 6.5, we can quote the following:

$$\begin{aligned} \mathcal{B}(B^+ \to \overline{K}^*(892)^0 K^+) \text{ central value} &= (0.6 \pm 0.3 \pm 0.2) \times 10^{-6}. \\ \mathcal{B}(B^+ \to \overline{K}^*(892)^0 K^+) \text{ upper limit} &= 1.1 \times 10^{-6}. \end{aligned}$$

For the central value, the first uncertainty is statistical, the second is systematic. The upper limit is at the 90% CL.

This information can be used to evaluate an improved bound on $\Delta S_{\phi(1020)K^0_S}$ (as defined in Section 1.3). Using

- Eq. (1.85),

$$|\xi_{\phi(1020)K^+}| < \frac{\lambda}{1-\frac{\lambda^2}{2}} \left(\left(\frac{\mathcal{B}(B^+ \to \overline{K}^*(892)^0 K^+)}{\mathcal{B}(B^+ \to \phi(1020)K^+)} \right)^{\frac{1}{2}} + \left(\frac{\mathcal{B}(B^+ \to \phi(1020)\pi^+)}{\mathcal{B}(B^+ \to \phi(1020)K^+)} \right)^{\frac{1}{2}} \right),$$

- The assumption that the up penguin minus top penguin contribution for $B^0 \to \phi(1020)K^0_S$ is not much larger than the corresponding contribution for $B^+ \to \phi(1020)K^+$ (see Section IV.B of [10]),

- Eq. (1.78),

$$\Delta S_{\phi(1020)K^0_S} = 2\cos 2\beta \sin\gamma \cos\delta_{\phi(1020)K^0_S} |\xi_{\phi(1020)K^0_S}|,$$

- The technique described in Section VI of [1] (the branching ratio and phase values generated, and the resulting $\Delta S_{\phi(1020)K^0_S}$ values, are illustrated in Figure 6.9),

- The likelihood curve (analogous to Figure 6.8) obtained in the *BABAR* $B^+ \to \phi(1020)\pi^+$ analysis [45],

- $\mathcal{B}(B^+ \to \phi(1020)K^+) = (8.3 \pm 0.7) \times 10^{-6}$ [41–44] and

- $\mathcal{B}(B^+ \to \overline{K}^*(892)^0 K^+) = (0.6 \pm 0.4)^3 \times 10^{-6}$,

we are able to place the following SM bound (at 90% CL) on $\Delta S_{\phi(1020)K^0_S}$:

$$|\Delta S_{\phi(1020)K^0_S}| < 0.11.$$

The most recent sin2β measurements are, from *BABAR*,

- $\sin 2\beta\,(B^0 \to J/\psi\, K^0_S) = 0.697 \pm 0.041_{Stat}$ [105],
- $\sin 2\beta^{\rm eff}\,(B^0 \to \phi(1020) K^0_S) = 0.12 \pm 0.31_{Stat} \pm 0.10_{Syst}$ [106],

and from Belle,

- $\sin 2\beta\,(B^0 \to J/\psi\, K^0_S) = 0.643 \pm 0.038_{Stat}$[4] [107],
- $\sin 2\beta^{\rm eff}\,(B^0 \to \phi(1020) K^0_S) = 0.50 \pm 0.21_{Stat} \pm 0.06_{Syst}$ [107].

The $B^0 \to \phi(1020) K^0_S$ measurements are illustrated in Figure 6.10.

[3]The statistical and systematic uncertainties have been added in quadrature.

[4]The systematic errors presented in [105] and [107] are for several charmonium modes combined, rather than just for $B^0 \to J/\psi\, K^0_S$. The systematic uncertainties on $\sin 2\beta\,(B^0 \to J/\psi\, K^0_S)$ are smaller than the statistical uncertainties ($\sim 50\%$).

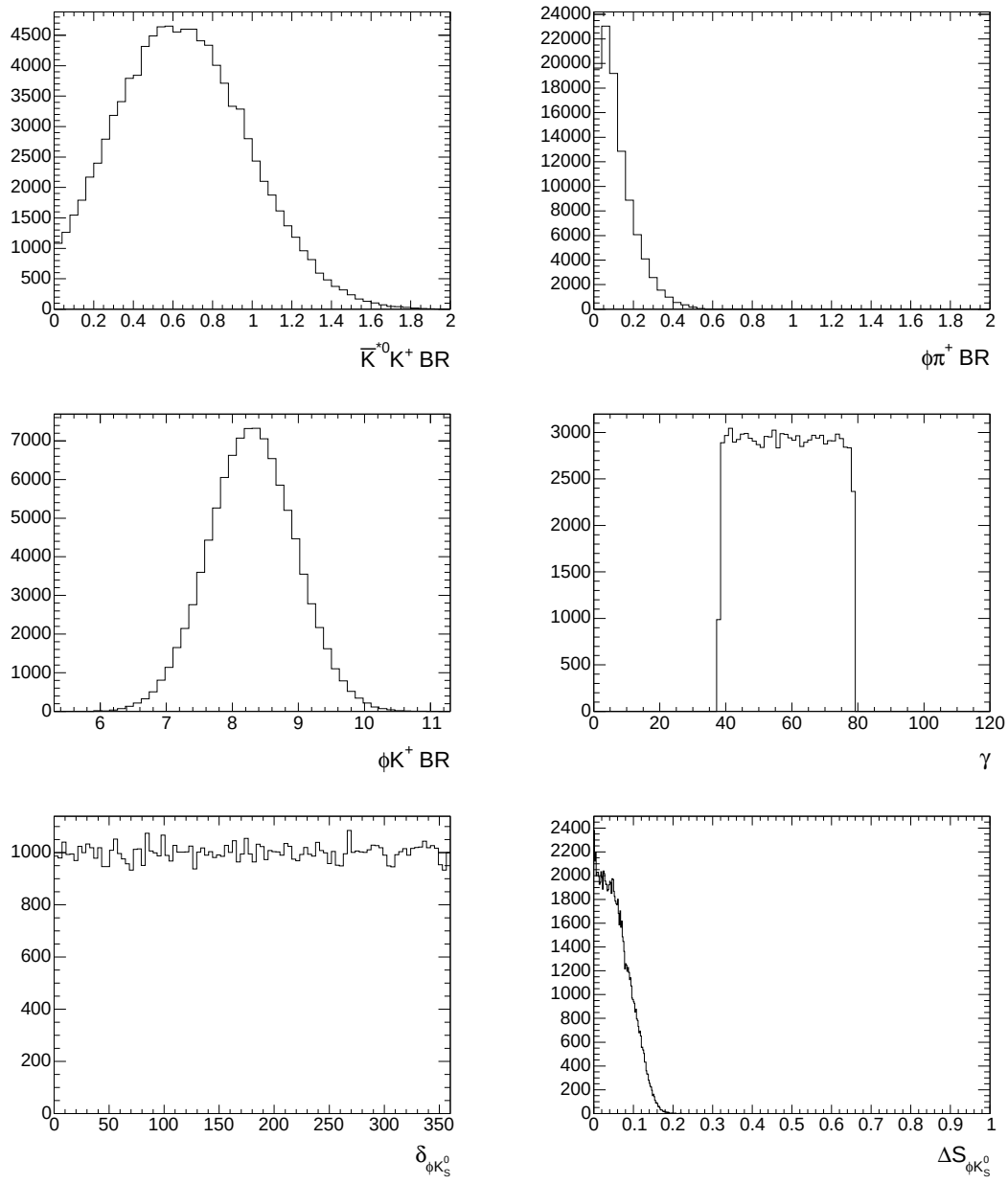

Figure 6.9: *Branching ratio and phase values generated, and resulting $\Delta S_{\phi(1020)K^0_S}$ values, using the technique described in Section VI of [1].*

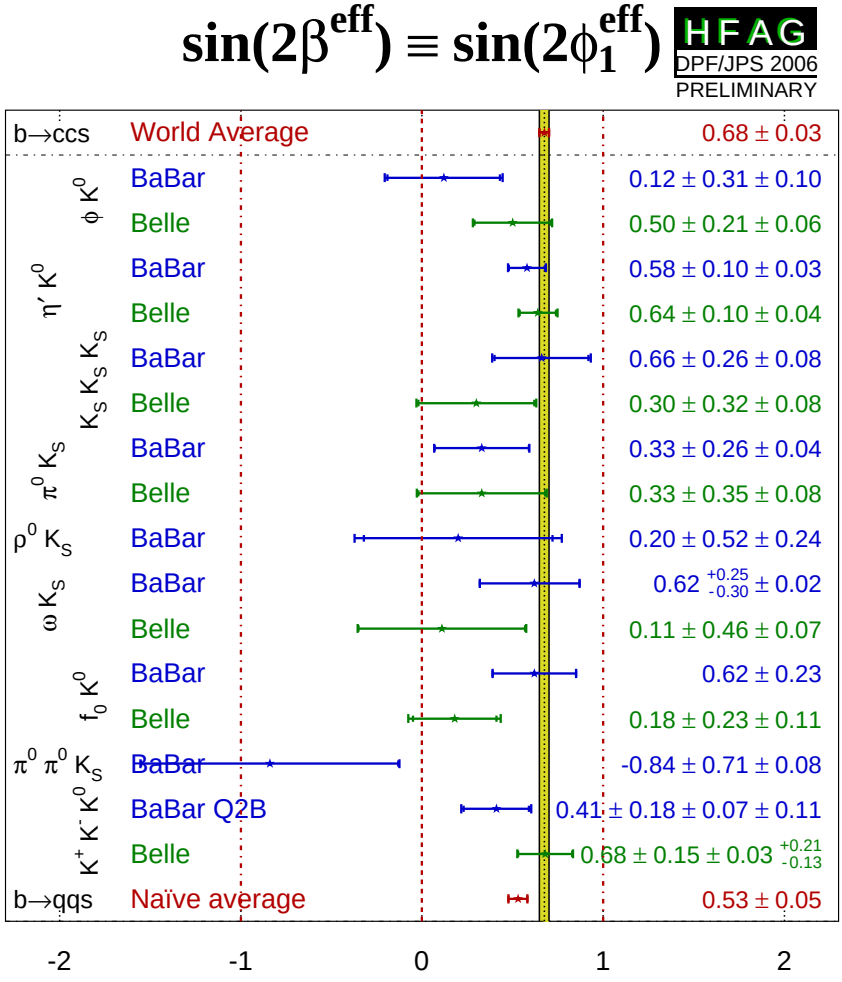

Figure 6.10: *Current status of* $\sin 2\beta^{\rm eff}$ *from penguin dominated modes. The entries relevant to this thesis are the first two entries (not including the* $\bar{b} \to \bar{c}c\bar{s}$ *world average). The* $\bar{b} \to \bar{c}c\bar{s}$ *world average is heavily influenced by* $B^0 \to J/\psi\, K^0_S$.

7

Conclusion

The optimal technique to measure the rates of B meson decays to the final state $K^+K^-\pi^+$ is an amplitude-level study of the full Dalitz plot. The results of the analysis described in this thesis indicate that we may have to wait for the next generation of B factories before such a study is feasible.

With the currently available data, which correspond to 232.3 million charged B mesons recorded by the *BABAR* detector, the most sensitive measurements possible are those that make use of an intensity-level unbinned extended maximum likelihood approach, as described in this thesis.

The 90% CL upper limit on the branching ratio of the channel $B^+ \to \overline{K}{}^*_0(1430)^0 K^+$ has been reduced here, significantly, from 10.2×10^{-6} [89, 102] to 2.2×10^{-6} (where

we use $\mathcal{B}(\overline{K}_0^*(1430)^0 \to K^-\pi^+) = \frac{2}{3} \times (0.93 \pm 0.10)$ [23]).

We are unable to comment on whether the difference between the fitted signal yields for the 5- and 3-variable fits in the 892 window is due to a statistical fluctuation or the presence of a further final state $K^+K^-\pi^+$ contribution (in addition to $\overline{K}^*(892)^0K^+$). The uncertainties are such that both scenarios are feasible.

The branching ratio of the decay $B^+ \to \overline{K}^*(892)^0K^+$ has been measured to be $(0.6\pm0.3_{Stat}\pm0.2_{Syst})\times10^{-6}$. The significance of this measurement, when taking into account systematic uncertainties, is 1.6σ. This falls short of the significance required to represent evidence for the process. As such, an upper limit of 1.1×10^{-6} (90% CL) is placed on the aforementioned branching ratio. This has been considerably reduced from 5.3×10^{-6} as presented by the CLEO collaboration in 2000 [7].

Our result can be used to evaluate an improved SM bound on $\Delta S_{\phi(1020)K_S^0}$, the deviation from sin2β of the time dependent *CP* asymmetry for the mode $B^0 \to \phi(1020)K_S^0$ (as discussed in Section 1.3). We find that $|\Delta S_{\phi(1020)K_S^0}| < 0.11$. This represents a significant reduction on the bound given in [1] of $|\Delta S_{\phi(1020)K_S^0}| < 0.42$, which is evaluated using neutral B decays[1]. A deviation exceeding this SM bound would provide a strong indication for physics beyond the SM.

[1]It should be noted that the assumption described in Section 6.7 (second bullet point) is not necessary when using neutral processes.

A

Fisher Discriminants

Eight Fisher discriminants are devised using the following discriminating variables:

- 2-variable Fisher, $\mathcal{F}(L_0, L_2)$.
- 3-variable Fisher, $\mathcal{F}(L_0, L_2, \texttt{TFlv})$.
- 4-variable Fisher, $\mathcal{F}(L_0, L_2, \theta_{Bmom}, \theta_{Bthr})$.
- 5-variable Fisher, $\mathcal{F}(L_0, L_2, \theta_{Bmom}, \theta_{Bthr}, \texttt{TFlv})$.
- 9-variable Fisher, $\mathcal{F}$(Nine CLEO cones).
- 10-variable Fisher, $\mathcal{F}$(Nine CLEO cones, $\texttt{TFlv}$).
- 11-variable Fisher, $\mathcal{F}$(Nine CLEO cones, $\theta_{Bmom}, \theta_{Bthr}$).

- 12-variable Fisher, $\mathcal{F}$(Nine CLEO cones, θ_{Bmom}, θ_{Bthr}, `TFlv`).

The variables L_0, L_2, θ_{Bmom}, θ_{Bthr} and `TFlv` are described in Section 4.5.1.2.

The *CLEO cone* variables [108] are constructed as follows: a 90° cone around the thrust axis of the reconstructed B is defined and split into nine independent concentric 10° cones. The energy flow of all charged tracks and neutral candidates in the ROE into these cones (in both the forward and backward hemispheres) is summed.

The nine CLEO cone variables are an alternative to L_0 and L_2 as a measure of the energy flow. They can be used instead of – rather than in addition to – L_0 and L_2.

L_0 and L_2 are shown to give superior separation compared to the CLEO cones. The 2-, 3-, 4- and 5-variable Fisher discriminants therefore give superior separation compared to their 9-, 10-, 11- and 12-variable Fisher discriminant counterparts, respectively.

For every added variable there is a certain amount of noise, which adversely affects the overall discrimination. It is only worth adding variables therefore when their discriminating power exceeds the effect of this noise. This is true for the variable `TFlv` and, used together, for the variables θ_{Bmom} and θ_{Bthr}.

Optimal Fisher coefficients (a_i in Eq. (4.12)) are presented in Tables A.1-A.3; signal versus $q\overline{q}$ background distributions are illustrated in Figure A.1 for the eight Fisher discriminant variables considered for the 5-variable fit in the 892 window.

	2-variable Fisher	*3-variable Fisher*	*4-variable Fisher*	*5-variable Fisher*	*9-variable Fisher*	*10-variable Fisher*	*11-variable Fisher*	*12-variable Fisher*
L_0	-0.049	-0.012	-0.017	-0.006	—	—	—	—
L_2	0.951	0.322	0.359	0.210	—	—	—	—
10° CLEO cone	—	—	—	—	0.206	0.143	0.149	0.113
20° CLEO cone	—	—	—	—	0.231	0.162	0.169	0.130
30° CLEO cone	—	—	—	—	0.116	0.083	0.085	0.067
40° CLEO cone	—	—	—	—	0.149	0.108	0.112	0.089
50° CLEO cone	—	—	—	—	0.028	0.023	0.023	0.020
60° CLEO cone	—	—	—	—	-0.024	-0.014	-0.014	-0.008
70° CLEO cone	—	—	—	—	-0.065	-0.042	-0.046	-0.033
80° CLEO cone	—	—	—	—	-0.074	-0.050	-0.054	-0.040
90° CLEO cone	—	—	—	—	-0.105	-0.069	-0.078	-0.056
θ_{Bmom}	—	—	0.326	0.183	—	—	0.142	0.104
θ_{Bthr}	—	—	0.298	0.168	—	—	0.127	0.093
`TFlv`	—	-0.668	—	-0.433	—	-0.307	—	-0.246

Table A.1: *Fisher coefficients for the 5-variable fit in the 892 window.*

	2-variable Fisher	*3-variable Fisher*	*4-variable Fisher*	*5-variable Fisher*	*9-variable Fisher*	*10-variable Fisher*	*11-variable Fisher*	*12-variable Fisher*
L_0	-0.044	-0.009	-0.015	-0.005	—	—	—	—
L_2	0.956	0.323	0.355	0.209	—	—	—	—
10° CLEO cone	—	—	—	—	0.210	0.145	0.151	0.115
20° CLEO cone	—	—	—	—	0.233	0.164	0.170	0.131
30° CLEO cone	—	—	—	—	0.117	0.084	0.086	0.068
40° CLEO cone	—	—	—	—	0.149	0.108	0.111	0.088
50° CLEO cone	—	—	—	—	0.029	0.023	0.023	0.019
60° CLEO cone	—	—	—	—	-0.021	-0.011	-0.012	-0.006
70° CLEO cone	—	—	—	—	-0.064	-0.042	-0.045	-0.032
80° CLEO cone	—	—	—	—	-0.075	-0.050	-0.054	-0.040
90° CLEO cone	—	—	—	—	-0.102	-0.068	-0.076	-0.055
θ_{Bmom}	—	—	0.332	0.186	—	—	0.146	0.106
θ_{Bthr}	—	—	0.298	0.171	—	—	0.127	0.095
`TFlv`	—	-0.668	—	-0.429	—	-0.306	—	-0.245

Table A.2: *Fisher coefficients for the 3-variable fit in the 892 window.*

	2-variable Fisher	***3-variable Fisher***	***4-variable Fisher***	***5-variable Fisher***	***9-variable Fisher***	***10-variable Fisher***	***11-variable Fisher***	***12-variable Fisher***
L_0	0.100	0.024	0.027	0.015	—	—	—	—
L_2	0.900	0.216	0.240	0.132	—	—	—	—
10° CLEO cone	—	—	—	—	0.243	0.159	0.161	0.121
20° CLEO cone	—	—	—	—	0.064	0.040	0.042	0.030
30° CLEO cone	—	—	—	—	0.240	0.155	0.161	0.119
40° CLEO cone	—	—	—	—	0.159	0.110	0.109	0.086
50° CLEO cone	—	—	—	—	0.053	0.038	0.039	0.032
60° CLEO cone	—	—	—	—	−0.000	−0.000	−0.000	−0.000
70° CLEO cone	—	—	—	—	−0.022	−0.014	−0.015	−0.010
80° CLEO cone	—	—	—	—	−0.119	−0.075	−0.082	−0.059
90° CLEO cone	—	—	—	—	−0.101	−0.065	−0.069	−0.050
θ_{Bmom}	—	—	0.392	0.212	—	—	0.163	0.119
θ_{Bthr}	—	—	0.341	0.179	—	—	0.158	0.113
`TFlv`	—	−0.760	—	−0.463	—	−0.344	—	−0.263

Table A.3: *Fisher coefficients for the 3-variable fit in the 1430 window.*

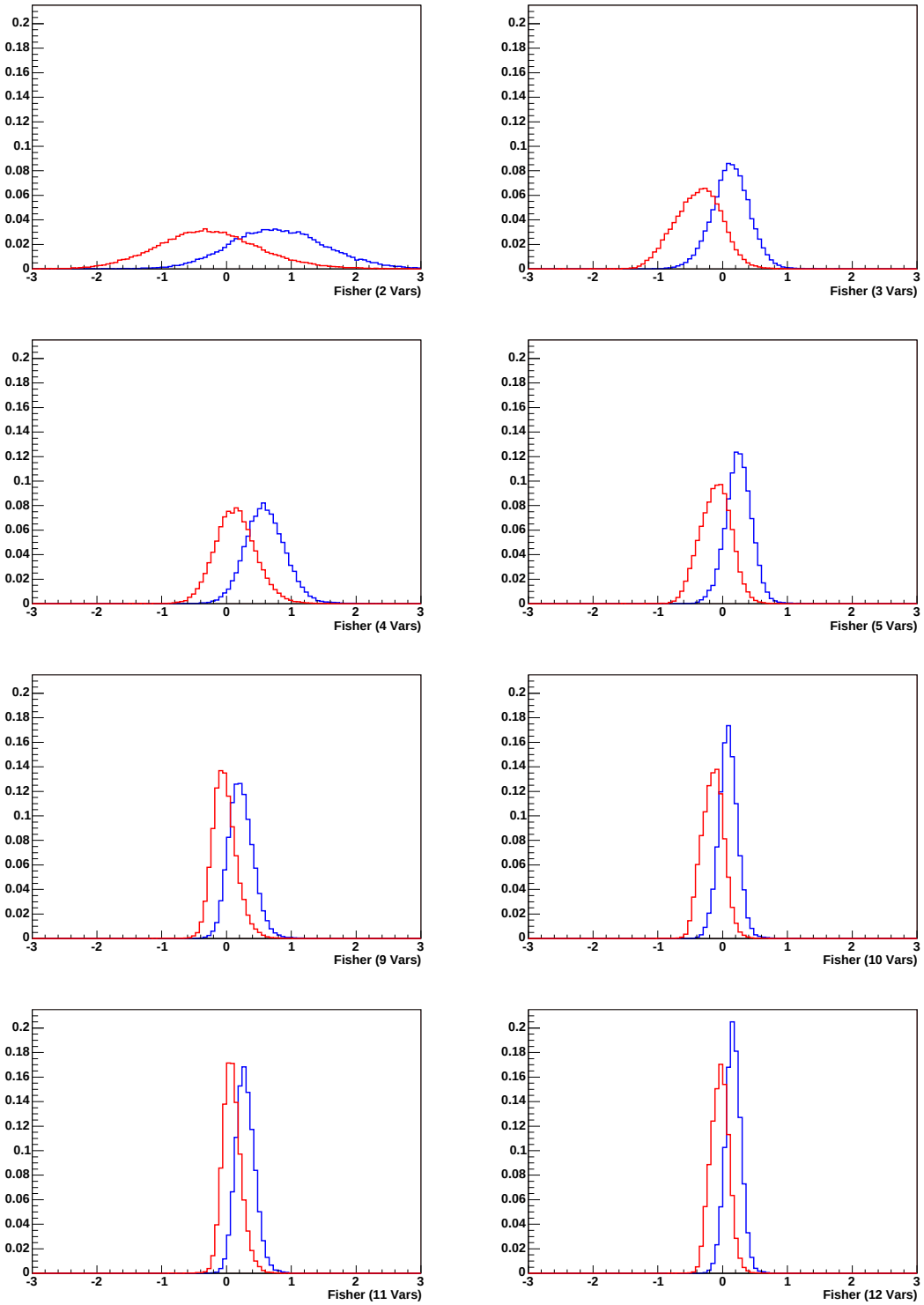

Figure A.1: *Signal (red) versus $q\overline{q}$ background (blue) distributions for eight Fisher discriminants.*

B

$B\overline{B}$ background MC studies

The following is for the 5-variable fit in the 892 window. The procedure is repeated for the 3-variable fits.

Tables B.1-B.2 and B.3-B.4 list all exclusive MC samples that are subjected to reconstruction and the selection criteria summarised in Table 5.3 (RaSC). Tables B.1-B.2 list modes that occur in the generic MC samples after the application of RaSC. Tables B.3-B.4 list further possible sources of $B\overline{B}$ background that are considered for the analysis.

Table B.5 lists the modes that are removed from the generic sample to be treated separately or simply ignored. Tables B.6-B.9 list the most abundant modes present in the resulting combinatoric *soups*, after the application of RaSC. Figure B.1 illus-

trates the fit variable distribution shapes for these combinatoric samples, compared to corresponding signal shapes.

Table B.1: *Modes that are a possible source of $B\overline{B}$ background, identified by running RaSC over generic MC samples. The number of events expected in the final on-resonance sample, for each mode, is estimated using the efficiency of applying RaSC to appropriate MC samples and world average branching ratio (BR) values. The first error is statistical – from the statistical uncertainty on the MC efficiency; the second is systematic – due to the error on the branching ratio. Statistical and systematic errors on $N_{B^\pm}$ are also accounted for (see Eq. (4.11)). Where only an upper limit (UL) is available, value used = UL/2± UL/2. Some of the modes shown here are treated exclusively, whilst others are left in the generic samples to be treated as a combinatoric soup. Continued in Table B.2 ...*

Exclusive MC Mode, $B^+ \to \{B^0 \to\}$	***Number Of Events To Pass RaSC***	***# MC Events Run Over***	***Branching Ratio*** ($\times 10^{-6}$)	***# Events Expected In Final On-res. Sample***
$\phi(\to K^+K^-)K^+$	3,266	163,000	4.4 ± 0.3 [23] [89]	$20.6 \pm 0.4 \pm 1.5$
$K^+K^-\pi^+$ (NR)	15,211	1,323,000	< 6.3 [23]	$8.4 \pm 0.1 \pm 8.4$
$K^+K^+K^-$ (NR)	3,755	1,314,000	< 23.1 [89]	$7.6 \pm 0.1 \pm 7.6$
$\overline{D}^0(\to K^+\pi^-)\rho^+(\to \pi^+\pi^0)$	16	415,000	509.2 ± 69.5 [23]	$4.5 \pm 1.1 \pm 0.6$
$\rho^0(\to \pi^+\pi^-)K^+$	742	206,000	5.1 ± 0.9 [89]	$4.3 \pm 0.2 \pm 0.7$
$\overline{D}^0(\to K^+\pi^-\pi^0)\rho^+(\to \pi^+\pi^0)$	3	294,000	$1,742.0 \pm 257.4$ [23]	$4.1 \pm 2.4 \pm 0.6$
$\overline{D}^0(\to K^+\pi^-)\mu^+\nu_\mu$	3	294,000	817.0 ± 85.8 [23]	$1.9 \pm 1.1 \pm 0.2$
$\overline{D}^0(\to K^+\pi^-\pi^0)\mu^+\nu_\mu$	1	292,000	$2,795.0 \pm 333.7$ [23]	$2.2 \pm 2.2 \pm 0.3$
$f_0(980)(\to \pi^+\pi^-)K^+$	128	148,000	9.1 ± 1.1 [89]	$1.8 \pm 0.2 \pm 0.2$
$\chi_{c0}(\to K^+K^-)K^+$	177	114,000	3.6 ± 1.5 [23]	$1.3 \pm 0.1 \pm 0.6$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-)\gamma)\mu^+\nu_\mu$	2	288,000	311.3 ± 40.4 [23]	$0.5 \pm 0.4 \pm 0.1$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-\pi^0)\gamma)\mu^+\nu_\mu$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	294,000	$1,064.9 \pm 150.8$ [23]	$0.04 \pm 0.04 \pm 0.01$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-)\pi^0)\mu^+\nu_\mu$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	294,000	505.7 ± 58.2 [23]	$0.020 \pm 0.020 \pm 0.002$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^0)\mu^+\nu_\mu$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	294,000	$1,730.1 \pm 221.9$ [23]	$0.07 \pm 0.07 \pm 0.01$
$\vdots$				

⋮				
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+K^-)\gamma)\pi^+$	15	147,000	6.8 ± 0.8 [23]	$0.16 \pm 0.04 \pm 0.02$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+K^-)\pi^0)\pi^+$	11	147,000	11.1 ± 1.2 [23]	$0.19 \pm 0.06 \pm 0.02$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-)\gamma)\pi^+$	8	192,500	66.6 ± 7.9 [23]	$0.6 \pm 0.2 \pm 0.1$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^0)\pi^+$	6	350,000	370.2 ± 43.1 [23]	$1.5 \pm 0.6 \pm 0.2$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-)\pi^0)\pi^+$	4	232,000	108.2 ± 11.0 [23]	$0.43 \pm 0.22 \pm 0.04$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-\pi^0)\gamma)\pi^+$	2	239,000	227.8 ± 29.8 [23]	$0.4 \pm 0.3 \pm 0.1$
$\eta'(\to \gamma\rho^0(\to \pi^+\pi^-))K^+$	175	145,000	20.5 ± 1.1 [23] [89]	$5.7 \pm 0.4 \pm 0.3$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-)\pi^0)e^+\nu_e$	2	294,000	505.7 ± 58.2 [23]	$0.8 \pm 0.6 \pm 0.1$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-)\gamma)e^+\nu_e$	1	289,000	311.3 ± 40.4 [23]	$0.25 \pm 0.25 \pm 0.03$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-\pi^0)\gamma)e^+\nu_e$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	283,000	$1,064.9 \pm 150.8$ [23]	$0.04 \pm 0.04 \pm 0.01$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^0)e^+\nu_e$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	292,000	$1,730.1 \pm 221.9$ [23]	$0.07 \pm 0.07 \pm 0.01$
$\overline{D}^0(\to K^+\pi^-)e^+\nu_e$	3	294,000	817.0 ± 85.8 [23]	$1.9 \pm 1.1 \pm 0.2$
$\overline{D}^0(\to K^+\pi^-\pi^0)e^+\nu_e$	1	294,000	$2,795.0 \pm 333.7$ [23]	$2.2 \pm 2.2 \pm 0.3$
$\overline{D}^0(\to K^-\pi^+)K^+$	61	10,000	0.05 ± 0.01 [23]	$0.07 \pm 0.01 \pm 0.01$
$\overline{D}^0(\to K^+\pi^-)K^+$	32	362,000	14.1 ± 2.3 [23]	$0.29 \pm 0.05 \pm 0.05$
$\overline{D}^0(\to K^+K^-)\pi^+$	5	20,000	19.4 ± 1.4 [23]	$1.1 \pm 0.5 \pm 0.1$
$\overline{D}^0(\to K^+\pi^-)\pi^+$	3	216,000	189.2 ± 11.9 [23]	$0.61 \pm 0.35 \pm 0.04$
$\overline{D}^0(\to K^+\pi^-\pi^0)\pi^+$	3	249,000	647.4 ± 54.8 [23]	$1.8 \pm 1.0 \pm 0.2$
$\chi_{c0}(\to \pi^+\pi^-)K^+$	2	111,000	4.4 ± 1.8 [23]	$0.02 \pm 0.01 \pm 0.01$
$J/\psi(\to e^+e^-, \mu^+\mu^-)K^+$	1	169,000	118.1 ± 4.9 [23]	$0.16 \pm 0.16 \pm 0.01$
$\overline{D}^0(\to K^+K^-)K^+$	638	250,000	1.4 ± 0.2 [23]	$0.83 \pm 0.03 \pm 0.12$
$\{K^+\rho^-(\to \pi^-\pi^0)\}$	398	647,000	9.9 ± 1.6 [89]	$1.4 \pm 0.1 \pm 0.2$
$\{D^{*-}(\to \overline{D}^0(\to X)\pi^-)\pi^+\}$	15	4,226,000	$1,868.5 \pm 142.8$ [23]	$1.5 \pm 0.4 \pm 0.1$
$\{D^{*-}(\to \overline{D}^0(\to X)\pi^-)\rho^+(\to \pi^+\pi^0)\}$	3	1,052,000	$4,603.6 \pm 610.2$ [23]	$3.0 \pm 1.8 \pm 0.4$
$\{D^-(\to X)\rho^+(\to \pi^+\pi^0)\}$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	112,000	$7,700.0 \pm 1,300.0$ [23]	$0.8 \pm 0.8 \pm 0.1$
$\{D^-(\to \pi^-\pi^0)K^+\}$	66	295,000	0.5 ± 0.2 [23]	$0.026 \pm 0.003 \pm 0.010$
$\{D^-(\to X)K^+\}$	7	246,000	200.0 ± 60.0 [23]	$1.3 \pm 0.5 \pm 0.4$
$\{K^+K^-\pi^0$ (NR)$\}$	299	126,000	< 19.0 [23]	$5.2 \pm 0.3 \pm 5.2$
$\overline{K}_2^*(1430)^0(\to K^-\pi^+)K^+$	1,045	97,000	$< 0.2^\ddagger$	$0.24 \pm 0.01 \pm 0.24$
$K^+\pi^+\pi^-$ (NR)	467	1,298,985	2.9 ± 1.1 [89]	$0.24 \pm 0.01 \pm 0.09$
$\phi(\to K^+K^-)\pi^+$	111	121,000	< 0.2 [23] [89]	$0.021 \pm 0.002 \pm 0.021$

‡: Estimated using Eq. (C.1).

Table B.2: *Continued from Table B.1.*

Table B.3: *Further modes that are considered as possible sources of $B\overline{B}$ background, including all modes taken into account for the BABAR analyses $B^+ \to K^+K^-K^+$ [88] and $B^+ \to K^+\pi^-\pi^+$ [50]. Continued in Table B.4 ...*

Exclusive MC Mode, $B^+ \to \{B^0 \to\}$	***Number Of Events To Pass RaSC***	***# MC Events Run Over***	***Branching Ratio*** $(\times 10^{-6})$	***# Events Expected In Final On-res. Sample***
$\overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ (LASS)	9,624	128,150	$< 3.8^{\ddagger}$	$32.9 \pm 0.3 \pm 32.9$
$\overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$ (BW)	3,771	117,000	$< 3.8^{\ddagger}$	$14.1 \pm 0.2 \pm 14.1$
$\{K^{*0}(\to K^+\pi^-)\overline{K}^{*0}(\to K^-\pi^+)$ (L)$\}$	740	122,000	< 9.8 [89]	$6.9 \pm 0.3 \pm 6.9$
$\{K^+K^-K_S^0$ (NR)$\}$	178	220,000	12.4 ± 1.2 [89]	$2.3 \pm 0.2 \pm 0.2$
$K^{*+}(\to K^+\pi^0)\overline{K}^{*0}(\to K^-\pi^+)$ (L)	657	126,000	< 1.6 [89] [90]	$0.97 \pm 0.04 \pm 0.97$
$\{K^+\pi^-\pi^0$ (NR)$\}$	236	2,307,000	35.6 ± 3.4 [89]	$0.8 \pm 0.1 \pm 0.1$
$\{D^{*-}(\to X)K^+\}$	4	247,000	200.0 ± 50.0 [23]	$0.8 \pm 0.4 \pm 0.2$
$K_1(1270)^0(\to \rho^+(\to \pi^+\pi^0)K^-)K^+$	138	117,000	$< 5.0^{\dagger}$	$0.7 \pm 0.1 \pm 0.7$
$\{K_1(1270)^-(\to \rho^0(\to \pi^+\pi^-)K^-)K^+\}$	125	117,000	$< 5.0^{\dagger}$	$0.6 \pm 0.1 \pm 0.6$
$\overline{D}^0(\to K^+\pi^-\pi^0)K^+$	6	134,000	48.1 ± 8.3 [23]	$0.5 \pm 0.2 \pm 0.1$
$\{\pi^+\pi^-\pi^0$ (NR)$\}$	10	2,130,000	< 720.0 [23]	$0.4 \pm 0.1 \pm 0.4$
$K^{*+}(\to K^+\pi^0)\overline{K}^{*0}(\to K^-\pi^+)$ (T)	28	126,000	< 15.8 [89]	$0.4 \pm 0.1 \pm 0.4$
$\{K^{*0}(\to X)\gamma\}$	16	392,000	43.0 ± 4.0 [23]	$0.41 \pm 0.10 \pm 0.04$
$\rho^0(\to \pi^+\pi^-)K^{*+}(\to K^+\pi^0)$ (L)	69	201,000	3.4 ± 1.3 [89] [90]	$0.27 \pm 0.03 \pm 0.10$
$\{K^{*+}(\to K^+\pi^0)\rho^-(\to \pi^-\pi^0)\}$	3	12,000	< 8.0 [89]	$0.2 \pm 0.1 \pm 0.2$
$\overline{K}^*(1680)^0(\to K^-\pi^+)K^+$	1,211	117,000	$< 0.2^{\ddagger}$	$0.24 \pm 0.01 \pm 0.24$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^0)K^+$	3	120,000	29.0 ± 8.4 [23]	$0.17 \pm 0.10 \pm 0.05$
$\{K^+\pi^-\}$	101	1,873,500	18.9 ± 0.7 [89]	$0.24 \pm 0.02 \pm 0.01$
$J/\psi(\to K^+K^-)K^+$	956	280,000	0.24 ± 0.03 [23]	$0.19 \pm 0.01 \pm 0.02$
$K_S^0(\to \pi^+\pi^-)K^+$	222	182,000	0.3 ± 0.3 [23]	$0.09 \pm 0.01 \pm 0.09$
$\{\rho^0(\to \pi^+\pi^-)K^{*0}(\to K^+\pi^-)$ (L)$\}$	96	202,000	< 1.8 [89]	$0.10 \pm 0.01 \pm 0.10$
$K^{*+}(\to K_S^0(\to \pi^+\pi^-)\pi^+)\overline{K}^{*0}(\to K^-\pi^+)$ (L)	7	122,000	< 10.9 [23] [89]	$0.07 \pm 0.03 \pm 0.07$
$K_0^*(1430)^0(\to K^+\pi^-)\pi^+$ (LASS)	3	116,497	23.7 ± 3.8 [23] [89]	$0.14 \pm 0.08 \pm 0.02$
$K_0^*(1430)^0(\to K^+\pi^-)\pi^+$ (BW)	4	177,000	23.7 ± 3.8 [23] [89]	$0.12 \pm 0.06 \pm 0.02$
$\{\overline{D}^0(\to K^+\pi^-)\rho^0(\to \pi^+\pi^-)\}$	3	70,000	11.0 ± 4.2 [23]	$0.11 \pm 0.06 \pm 0.04$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+K^-)\gamma)K^+$	180	265,000	0.5 ± 0.2 [23]	$0.08 \pm 0.01 \pm 0.03$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+K^-)\pi^0)K^+$	20	30,000	0.9 ± 0.2 [23]	$0.14 \pm 0.03 \pm 0.03$
$\vdots$				

⋮				
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-)\pi^0)K^+$	16	285,000	8.5 ± 2.4 [23]	$0.11 \pm 0.11 \pm 0.03$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-)\gamma)K^+$	14	259,500	5.2 ± 1.5 [23]	$0.06 \pm 0.06 \pm 0.02$
$\rho^+(\to \pi^+\pi^0)K^{*0}(\to K^+\pi^-)$ (L)	52	800,000	4.6 ± 0.9 [89] [90]	$0.07 \pm 0.07 \pm 0.01$
$K^0_S(\to \pi^+\pi^-)\pi^+$	9	219,000	8.3 ± 0.4 [23] [89]	$0.079 \pm 0.079 \pm 0.004$
$\{\pi^+\pi^-\}$	2	2,595,000	5.0 ± 0.4 [89]	$(8.9 \pm 8.9 \pm 0.7) \times 10^{-4}$
$\eta'(\to \gamma\rho^0(\to \pi^+\pi^-))\pi^+$	4	146,000	2.5 ± 0.6 [23] [89]	$0.016 \pm 0.016 \pm 0.004$
$\eta(\to \pi^+\pi^-\pi^0)K^{*+}(\to K^+\pi^0)$	4	121,000	1.8 ± 0.2 [23] [89]	$0.014 \pm 0.014 \pm 0.002$
$\{K^*_2(1430)^0(\to X)\gamma\}$	3	184,000	13.0 ± 5.0 [23]	$0.05 \pm 0.05 \pm 0.02$
$\{\overline{D}^{*0}(\to \overline{D}^0\pi^0, \overline{D}^0\gamma)\rho^0(\to \pi^+\pi^-),\ \overline{D}^0 \to K^+\pi^-\}$	1	140,000	< 19.8 [23]	$0.02 \pm 0.02 \pm 0.02$
$\rho^+(\to \pi^+\pi^0)K^{*0}(\to K^+\pi^-)$ (T)	2	180,000	2.4 ± 0.6 [89] [90]	$0.006 \pm 0.006 \pm 0.002$
$\rho^0(\to \pi^+\pi^-)K^{*+}(\to K^+\pi^0)$ (T)	4	178,000	0.1 ± 0.5 [89] [90]	$(0.5 \pm 0.5 \pm 2.0) \times 10^{-3}$
$\{\rho^0(\to \pi^+\pi^-)K^{*0}(\to K^+\pi^-)$ (T)$\}$	10	182,000	< 1.8 [89]	$0.01 \pm 0.01 \pm 0.01$
$\rho^+(\to \pi^+\pi^0)\rho^0(\to \pi^+\pi^-)$ (L)	2	352,000	25.6 ± 6.5 [89] [90]	$0.03 \pm 0.03 \pm 0.01$
$\{\rho^0(\to \pi^+\pi^-)\rho^0(\to \pi^+\pi^-)$ (L)$\}$	6	203,000	< 1.1 [89]	$0.004 \pm 0.004 \pm 0.004$
$\{\rho^+(\to \pi^+\pi^0)\rho^-(\to \pi^-\pi^0)$ (L)$\}$	29	4,383,000	25.4 ± 3.7 [89] [90]	$0.04 \pm 0.04 \pm 0.01$
$\{\overline{D}^0(\to K^+\pi^-)\pi^0(\to \gamma\gamma)\}$	1	154,000	10.1 ± 3.0 [23]	$0.015 \pm 0.015 \pm 0.004$
$K^{*+}(\to X)\gamma$	2	436,000	38.0 ± 5.0 [23]	$0.04 \pm 0.04 \pm 0.01$
$\pi^+\pi^+\pi^-$ (NR)	32	2,220,000	< 4.6 [89]	$0.01 \pm 0.01 \pm 0.01$
$K^{*0}(892)(\to K^+\pi^-)\pi^+$	5	185,000	7.6 ± 0.7 [89]	$0.048 \pm 0.048 \pm 0.004$
$\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-\gamma)\pi^0)K^+$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	42,000	17.8 ± 5.3 [23]	$0.005 \pm 0.005 \pm 0.001$
$\psi(2S)(\to e^+e^-, \mu^+\mu^-)K^+$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	50,100	10.1 ± 0.8 [23]	$(2.3 \pm 2.3 \pm 0.2) \times 10^{-3}$
$K^*(1410)^+(\to X)\gamma$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	290,000	$< 8,000.0^\dagger$	$0.16 \pm 0.16 \pm 0.16$
$\{K^*(1410)^0(\to X)\gamma\}$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	288,000	$< 8,000.0^\dagger$	$0.16 \pm 0.16 \pm 0.16$
$K^*(1680)^+(\to X)\gamma$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	290,000	$< 1,900.0$ [23]	$0.04 \pm 0.04 \pm 0.04$
$\{\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-)\gamma)\gamma\}$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	183,000	< 0.8 [23]	$(2.5 \pm 2.5 \pm 2.5) \times 10^{-5}$
$K^{*+}(\to K^0_S(\to \pi^+\pi^-)\pi^+)\overline{K}^{*0}(\to K^-\pi^+)$ (T)	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	121,000	< 10.9 [23] [89]	$0.001 \pm 0.001 \pm 0.001$
$\{K^{*0}(\to K^+\pi^-)\overline{K}^{*0}(\to K^-\pi^+)$ (T)$\}$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	121,000	< 9.8 [89]	$(4.7 \pm 4.7 \pm 4.7) \times 10^{-4}$
$\{\overline{D}^0(\to K^+\pi^-\pi^0)\pi^0(\to \gamma\gamma)\}$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	70,000	34.7 ± 10.5 [23]	$0.006 \pm 0.006 \pm 0.002$
$\{\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-)\pi^0/\gamma)\pi^0\}$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	140,000	10.1 ± 1.9 [23]	$(8.4 \pm 8.4 \pm 1.6) \times 10^{-4}$
$\{\overline{D}^{*0}(\to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^0/\gamma)\pi^0\}$	0 ($\Rightarrow$ 0.1 UL @ 90% CL)	140,000	34.7 ± 6.8 [23]	$0.003 \pm 0.003 \pm 0.001$

L(T): Longitudinally (transversely) polarised.

†: Conservative upper limit.

‡: Estimated using Eq. (C.1).

Table B.4: *Continued from Table B.3.*

$B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$	$B^+ \to \eta'(958)(\to \rho(770)^0(\to \pi^+\pi^-)\gamma)K^+$	$B^+ \to \overline{D}^0(\to K^+K^-)K^+$
$B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+\gamma)K^+$	$B^+ \to \overline{K}_2^*(1430)^0(\to K^-\pi^+)K^+$	$B^+ \to \overline{D}^0(\to \pi^+\pi^-)\pi^+$
$B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+\gamma$	$B^+ \to \overline{K}_2^*(1430)^0(\to K^-\pi^+)K^+\gamma$	$B^+ \to \phi(1020)(\to K^+K^-)\pi^+$
$B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+\gamma\gamma)K^+$	$B^+ \to K^+K^-\pi^+$ (NR)	$B^+ \to K_2^*(1430)^0(\to K^+\pi^-)\pi^+$
$B^+ \to \phi(1020)(\to K^+K^-)K^+$	$B^0 \to \rho^-(\to \pi^0\pi^-)K^+$	$B^0 \to D^-(\to \pi^0\pi^-)K^+$
$B^+ \to \phi(1020)(\to K^+K^-)K^+\gamma$	$B^0 \to \rho^-(\to \pi^0\pi^-)K^+\gamma$	$B^0 \to K^{*0}(\to K^+\pi^-)\overline{K}^{*0}(\to K^-\pi^+)$
$B^+ \to K^+K^+K^-$ (NR)	$B^+ \to K^+\pi^+\pi^-$ (NR)	$B^0 \to K_S^0K^+K^-$
$B^+ \to K^+K^+K^-\gamma$ (NR)	$B^+ \to \overline{D}^0(\to K^+K^-)\pi^+$	$B^+ \to \eta_c(1S)(\to K^+K^-)K^{+\square}$
$B^+ \to f_0(980)(\to K^+K^-)K^+$	$B^+ \to J/\psi(\to K^+K^-)K^+$	$B^+ \to \chi_{c1}(1P)(\to K^+K^-)K^{+\square}$
$B^+ \to f_0(980)(\to K^+K^-)K^+\gamma$	$B^0 \to \pi^0K^+K^-$	$B^+ \to a_0(980)^0(\to K^+K^-)K^{+\diamond}$
$B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$	$B^+ \to \chi_{c0}(1P)(\to K^+K^-)K^+$	$B^+ \to J/\psi(\to K^+K^-)\pi^{+\triangle}$
$B^+ \to \rho(770)^0(\to \pi^+\pi^-\gamma)K^+$	$B^+ \to f_0(980)(\to \pi^+\pi^-)K^+$	

$^\square$: Standard Model forbidden. $^\diamond$: Very small branching ratio – ignored by *BABAR* [88] and Belle [42] $B^+ \to K^+K^-K^+$ Dalitz plot teams. $^\triangle$: Tiny branching ratio ($\sim 10^{-9}$ [23]).

Table B.5: *Modes removed from the charged and neutral generic MC samples. These modes are treated separately, or, in the case of* $^\square$, $^\diamond$ *and* $^\triangle$, *ignored. The resulting generic MC samples are combinatoric, with no single mode dominating (see Tables B.6-B.9).*

Table B.6: *Modes remaining in the $B^+B^- \to X$ generic MC sample after the application of RaSC and the modes listed in Table B.5 have been removed. B daughters only are shown. Granddaughters and great granddaughters for the most popular modes are shown in Table B.8. In total, 316 events remain. Modes not shown contribute less than any of those shown. The purpose of listing these modes is to demonstrate that no single mode is responsible for making a major contribution. Rather, what remains is a combinatoric mixture of (mainly charm) B meson decays. Continued in Table B.7 . . .*

$B^+ \to$	***Number Of Events Present After RaSC Are Applied To Generic MC Sample Containing 535.9 Million $B^+B^- \to X$ Events*** †			***# Events Expected In Final On-res. Sample (231.8 Million $B\overline{B}$ Events) By Extrapolating Numbers From Generics Study*** ‡
	B^+	B^-	***Total***	
$\rho^+\overline{D}^0$	36	29	65	14.0 ± 1.7
$\overline{D}^{*0}\rho^+$	14	22	36	7.8 ± 1.3
$\overline{D}^0\mu^+\nu_\mu$	14	14	28	6.1 ± 1.1
$\overline{D}^{*0}\pi^+$	10	16	26	5.6 ± 1.1
$\overline{D}^{*0}\mu^+\nu_\mu$	14	10	24	5.2 ± 1.1
$\overline{D}^{*0}e^+\nu_e$	13	10	23	5.0 ± 1.0
$\overline{D}^0e^+\nu_e$	11	8	19	4.1 ± 0.9
$a_1^+\overline{D}^0$	11	8	19	4.1 ± 0.9
$\overline{D}^0K^+$	7	11	18	3.9 ± 0.9
$\overline{D}^0\pi^+$	6	7	13	2.8 ± 0.8
$D_s^{*+}\overline{D}^{*0}$	5	6	11	2.4 ± 0.7
$\overline{D}^{*0}a_1^+$	5	4	9	1.9 ± 0.6
$\overline{D}^{*0}D^{*0}K^+$	3	4	7	1.5 ± 0.6
$\vdots$				

⋮				
$\overline{D}^0 D_s^+$	5	1	6	1.3 ± 0.5
$\overline{D}^{*0}\overline{K}^0 K^+$	2	3	5	1.1 ± 0.5
$\overline{D}_2^{*0}\pi^+$	3	2	5	1.1 ± 0.5
$D^{*-}\pi^0\pi^+\pi^+$	3	2	5	1.1 ± 0.5
$\overline{D}^0\gamma e^+\nu_e$	2	3	5	1.1 ± 0.5
$\overline{D}_1^0\mu^+\nu_\mu$	3	2	5	1.1 ± 0.5
`string`	1	3	4	0.9 ± 0.4
$\overline{D}^{*0}\tau^+\nu_\tau$	1	3	4	0.9 ± 0.4
$D_s^{*+}\overline{D}^0$	2	2	4	0.9 ± 0.4
$\overline{D}^{*0}\gamma e^+\nu_e$	0	4	4	0.9 ± 0.4
$\overline{D}^0\tau^+\nu_\tau$	2	2	4	0.9 ± 0.4
$J/\psi\, K^+$	1	2	3	0.6 ± 0.4
$\overline{D}_2^{*0} D_s^{*+}$	2	1	3	0.6 ± 0.4
$\overline{D}^0\gamma\mu^+\nu_\mu$	3	0	3	0.6 ± 0.4
$D_s^+\overline{D}^0\pi^0\pi^0$	3	0	3	0.6 ± 0.4
$\overline{D}_1^{*0}\pi^+$	0	3	3	0.6 ± 0.4
$\overline{D}^0\rho^+\pi^-\pi^+$	0	3	3	0.6 ± 0.4

†: *It is not necessarily the case that all of the events shown here pass RaSC. It may be that it is the other B in the event that passes.*

‡: *Estimated from a somewhat crude extrapolation using the scale (231.8/535.9)/2. Uncertainty on the values of the branching ratios used in the generation of the generics are not accounted for in the errors shown.*

Table B.7: *Continued from Table B.6.*

$B^+ \to$	*# Generic MC Events Present After RaSC*			*Number Of Events Extrapolated In To Final On-res. Sample*
	B^+	B^-	***Total***	
$\boldsymbol{\rho^+\overline{D}^0}$:	**36**	**29**	**65**	$\mathbf{14.0 \pm 1.7}$
$\rho^+ \to \pi^+\pi^0$, $\overline{D}^0 \to K^+\pi^-$	17	12	29	6.3 ± 1.2
$\rho^+ \to \pi^+\pi^0$, $\overline{D}^0 \to \mu^-\overline{\nu}_\mu K^+$	6	7	13	2.8 ± 0.8
$\rho^+ \to \pi^+\pi^0$, $\overline{D}^0 \to e^-\overline{\nu}_e K^+$	3	5	8	1.7 ± 0.6
$\rho^+ \to \pi^+\pi^0$, $\overline{D}^0 \to K^+\pi^-\pi^0$	3	4	7	1.5 ± 0.6
$\rho^+ \to \pi^+\pi^0$, $\overline{D}^0 \to K^+K^-$	3	1	4	0.9 ± 0.4
$\boldsymbol{\overline{D}^{*0}\rho^+}$:	**14**	**22**	**36**	$\mathbf{7.8 \pm 1.3}$
$\overline{D}^{*0} \to \overline{D}^0(\to K^+\pi^-)\gamma$, $\rho^+ \to \pi^+\pi^0$	5	9	14	3.0 ± 0.8
$\overline{D}^{*0} \to \overline{D}^0(\to K^+\pi^-)\pi^0$, $\rho^+ \to \pi^+\pi^0$	3	5	8	1.7 ± 0.6
$\overline{D}^{*0} \to \overline{D}^0(\to \mu^-\overline{\nu}_\mu K^+)\pi^0$, $\rho^+ \to \pi^+\pi^0$	2	2	4	0.9 ± 0.4
$\overline{D}^{*0} \to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^0$, $\rho^+ \to \pi^+\pi^0$	1	2	3	0.6 ± 0.4
$\boldsymbol{\overline{D}^0\mu^+\nu_\mu}$:	**16**	**14**	**30**	$\mathbf{6.5 \pm 1.2}$
$\overline{D}^0 \to K^+\pi^-$	7	4	11	2.4 ± 0.7
$\overline{D}^0 \to K^+\pi^-\pi^0$	1	6	7	1.5 ± 0.6
$\overline{D}^0 \to \mu^-\overline{\nu}_\mu K^+$	1	3	4	0.9 ± 0.4
$\boldsymbol{\overline{D}^{*0}\mu^+\nu_\mu}$:	**16**	**10**	**26**	$\mathbf{5.6 \pm 1.1}$
$\overline{D}^{*0} \to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^0$	1	7	8	1.7 ± 0.6
$\overline{D}^{*0} \to \overline{D}^0(\to K^+\pi^-\pi^0)\gamma$	2	3	5	1.1 ± 0.5
$\overline{D}^{*0} \to \overline{D}^0(\to K^+\pi^-)\pi^0$	2	1	3	0.6 ± 0.4
$\overline{D}^{*0} \to \overline{D}^0(\to a_1^+K^-)\pi^0$	1	2	3	0.6 ± 0.4
$\boldsymbol{\overline{D}^{*0}\pi^+}$:	**10**	**16**	**26**	$\mathbf{5.6 \pm 1.1}$
$\overline{D}^{*0} \to \overline{D}^0(\to K^+\pi^-\pi^0)\gamma$	2	3	5	1.1 ± 0.5
$\overline{D}^{*0} \to \overline{D}^0(\to K^+\pi^-\pi^0)\pi^0$	1	2	3	0.6 ± 0.4
$\overline{D}^{*0} \to \overline{D}^0(\to e^-\overline{\nu}_e K^+)\pi^0$	3	0	3	0.6 ± 0.4
$\boldsymbol{\overline{D}^{*0}e^+\nu_e}$:	**13**	**10**	**23**	$\mathbf{5.0 \pm 1.0}$
$\overline{D}^{*0} \to \overline{D}^0(\to a_1^+K^-)\pi^0$	2	3	5	1.1 ± 0.5
$\overline{D}^{*0} \to \overline{D}^0(\to K^+\pi^-)\gamma$	0	3	3	0.6 ± 0.4
$\boldsymbol{\overline{D}^0e^+\nu_e}$:	**11**	**8**	**19**	$\mathbf{4.1 \pm 0.9}$
$\overline{D}^0 \to K^+\pi^-$	4	3	7	1.5 ± 0.6
$\overline{D}^0 \to \mu^-\overline{\nu}_\mu K^+$	3	1	4	0.9 ± 0.4
$\overline{D}^0 \to K^+\pi^-\pi^0$	2	1	3	0.6 ± 0.4
$\boldsymbol{a_1^+\overline{D}^0}$:	**11**	**8**	**19**	$\mathbf{4.1 \pm 0.9}$
$a_1^+ \to \rho^0(\to \pi^-\pi^+)\pi^+$, $\overline{D}^0 \to K^+\pi^-$	6	3	9	1.9 ± 0.6
$\boldsymbol{\overline{D}^0K^+}$:	**7**	**11**	**18**	$\mathbf{3.9 \pm 0.9}$

Table B.8: *Breakdown of the most populous modes shown in Tables B.6-B.7.*

$\overline{B}^0 \rightarrow$	*Number Of Events Present After RaSC Are Applied To Generic MC Sample Containing 541.3 Million $B^0\overline{B}^0 \rightarrow X$ Events* †			*# Events Expected In Final On-res. Sample (231.8 Million $B\overline{B}$ Events) By Extrapolating Numbers From Generics Study* ‡
	B^+	B^-	***Total***	
$D^{*+}\mu^-\overline{\nu}_\mu$	5	5	10	2.2 ± 0.7
$D^{*+}\pi^-$	4	3	7	1.5 ± 0.6
`string`	3	3	6	1.3 ± 0.5
$K^{*+}\pi^-$	1	5	6	1.3 ± 0.5
$\rho^- D^{*+}$	1	5	6	1.3 ± 0.5
$D^+\pi^-$	0	6	6	1.3 ± 0.5
$K_2^{*-}\pi^+$	5	0	5	1.1 ± 0.5
$\rho^+ D^-$	5	0	5	1.1 ± 0.5
$K^{*+}K^-$	0	5	5	1.1 ± 0.5
$K_2^{*+}\pi^-$	0	5	5	1.1 ± 0.5
$D^{*-}e^+\nu_e$	0	5	5	1.1 ± 0.5
$D^{*+}e^-\overline{\nu}_e$	4	0	4	0.9 ± 0.4
$D^{*+}\overline{D}^{*0}K^-$	1	3	4	0.9 ± 0.4
$D^{*-}a_1^+$	4	0	4	0.9 ± 0.4
$K^{*-}K^+$	3	0	3	0.6 ± 0.4
$D^+\mu^-\overline{\nu}_\mu$	3	0	3	0.6 ± 0.4
$D^{*+}\overline{D}^0K^-$	2	1	3	0.6 ± 0.4
$\overline{K}^{*0}K_S^0$	0	3	3	0.6 ± 0.4
$D^-\mu^+\nu_\mu$	0	3	3	0.6 ± 0.4
$D^{*-}\gamma\mu^+\nu_\mu$	2	1	3	0.6 ± 0.4

Table B.9: *As Tables B.6-B.7, except generated generic MC sample used is for 541.3 million $B^0\overline{B}^0 \rightarrow X$ events. For † and ‡, see foot of Table B.7.*

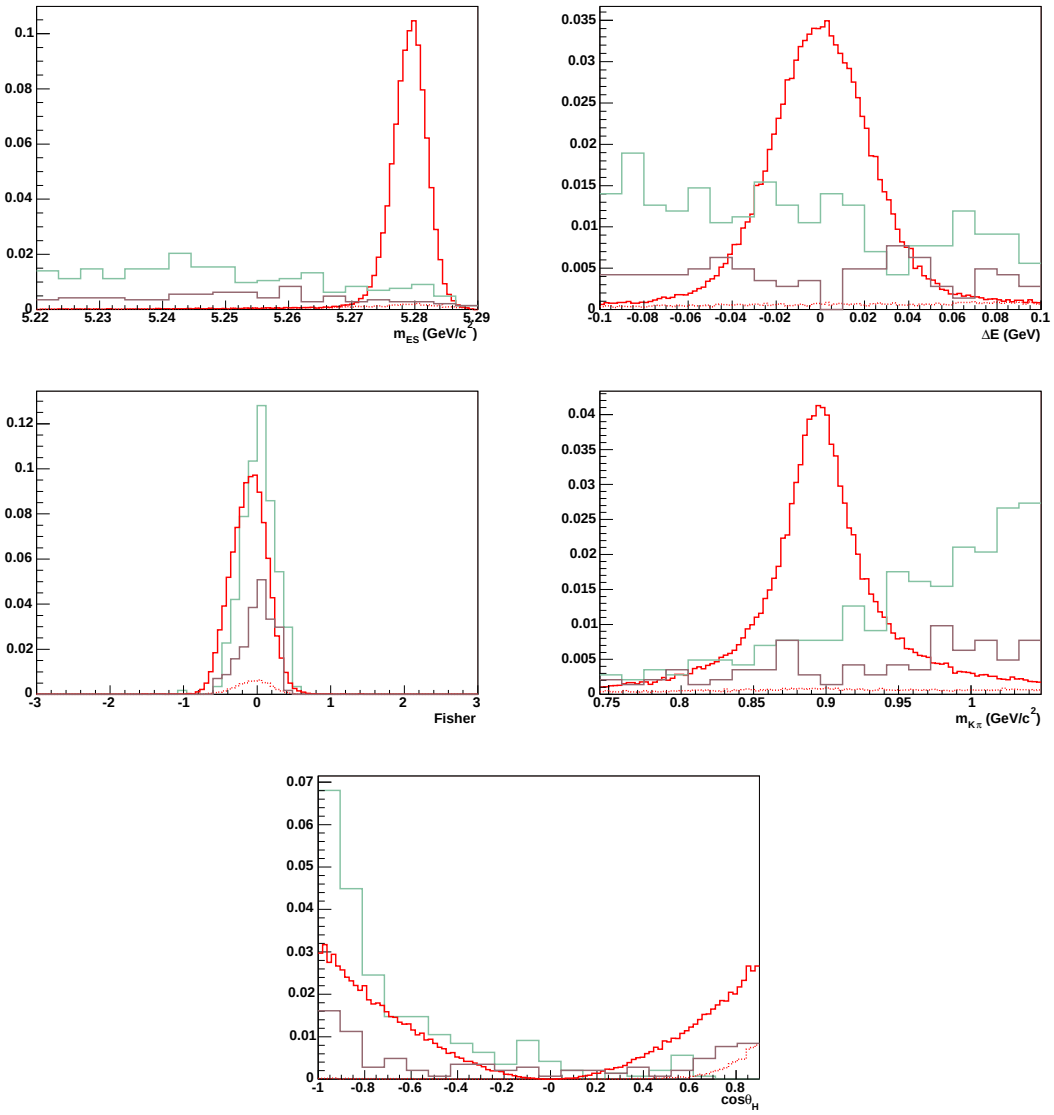

Figure B.1: *Distributions of fit variables for signal (red), charged combinatoric generic MC (green) and neutral combinatoric generic MC (purple), normalised w.r.t. each other to reflect the number of expected events in the final on-resonance sample (with* $\mathcal{B}(B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+) = 1.0 \times 10^{-6}$*). RaSC have been applied, and modes listed in Table B.5 have been removed from the generic samples.*

C

Branching ratio upper limit estimates for K^{**} modes

Let "K^{**}" represent any K^* resonance that is greater in mass than the $K^*(892)$ resonance (for example $K_0^*(1430)$, $K_2^*(1430)$, $K^*(1680)$, etc.).

Penguins are much more important in pseudoscalar-pseudoscalar decays ($\mathcal{B}(B \to \pi\pi)/\mathcal{B}(B \to \pi K) \sim 1/4$) than in vector-pseudoscalar ones ($\mathcal{B}(B \to \rho\pi)/\mathcal{B}(B \to \rho K) \sim 2$) – a result also found in QCD factorisation. Using

$$\mathcal{B}(B^+ \to \overline{K}^{**} K^+) < \frac{\mathcal{B}(B^+ \to K^{**}\pi^+)}{\mathcal{B}(B^+ \to K^*(892)^0\pi^+)} \mathcal{B}(B^+ \to \overline{K}^0 K^+) \qquad \text{(C.1)}$$

therefore, should be entirely conservative. We have as a safety range the unused expectation of small penguins in vector-pseudoscalar decays.

D

PDF parameters

PDF parameters are presented in Tables D.1-D.12. Mass units are GeV/c^2; energy units are GeV.

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Signal MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Signal MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
S_{TRU}	m_{ES}	μ	5.2792 ± 0.0000	Fixed	S_{TRU}	$m_{K\pi}$	μ	$(8.9558 \pm 0.0014) \times 10^{-1}$	Fixed
		σ	$(2.5344 \pm 0.0102) \times 10^{-3}$	Fixed			Γ	$(5.003 \pm 0.031) \times 10^{-2}$	Fixed
		α	1.8670 ± 0.0459	Fixed		$\cos\theta_H$	c_1	$(0.008 \pm 1.567) \times 10^{-2}$	Fixed
		n	5.9244 ± 0.5219	Fixed			c_2	3.602 ± 0.465	Fixed
	ΔE	μ_1	$(4.42 \pm 1.12) \times 10^{-4}$	Fixed			c_3	$(-0.299 \pm 1.554) \times 10^{-1}$	Fixed
		σ_1	$(2.083 \pm 0.016) \times 10^{-2}$	Fixed			c_4	-5.217 ± 0.964	Fixed
		μ_2	$(-6.913 \pm 1.095) \times 10^{-3}$	Fixed			c_5	$(0.672 \pm 4.194) \times 10^{-1}$	Fixed
		σ_2	$(4.703 \pm 0.197) \times 10^{-2}$	Fixed			c_6	4.926 ± 1.071	Fixed
		f_2	$1.0 - (0.8807 \pm 0.0111)$	Fixed			c_7	$(-2.769 \pm 3.211) \times 10^{-1}$	Fixed
	$\mathcal{F}$	μ_1	$(-1.147 \pm 0.010) \times 10^{-1}$	Fixed			c_8	-2.075 ± 0.513	Fixed
		σ_1	$(2.368 \pm 0.007) \times 10^{-1}$	Fixed			C	-2.716 ± 0.003	Fixed
		μ_2	$(-1.993 \pm 0.794) \times 10^{-1}$	Fixed					
		σ_2	1.346 ± 0.091	Fixed					
		f_2	$1.0 - (0.9937 \pm 0.0005)$	Fixed					

Table D.1: *PDF parameter values for the S_{TRU} component for the 5-variable fit in the 892 window. For further information, see Section 5.7.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Signal MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Signal MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
S_{SXF}	m_{ES}	μ	5.2804 ± 0.0004	Fixed	S_{SXF}	$m_{K\pi}$	μ	$(8.8745 \pm 0.0493) \times 10^{-1}$	Fixed
		σ	$(4.3385 \pm 0.2735) \times 10^{-3}$	Fixed			σ	$(3.868 \pm 0.579) \times 10^{-2}$	Fixed
		α	0.3262 ± 0.0432	Fixed			c_1	-3.196 ± 0.773	Fixed
		n	3.0615 ± 0.7204	Fixed			f_{Gaus}	0.1370 ± 0.0230	Fixed
	ΔE	c_1	3.622 ± 0.268	Fixed		$\cos\theta_H$	c_1	$(1.121 \pm 1.568) \times 10^{-1}$	Fixed
	$\mathcal{F}$	μ_1	$(-5.267 \pm 0.381) \times 10^{-2}$	Fixed			c_2	5.115 ± 0.623	Fixed
		σ_1	$(2.351 \pm 0.028) \times 10^{-1}$	Fixed			c_3	4.447 ± 0.502	Fixed
		μ_2	$(-3.914 \pm 2.652) \times 10^{-1}$	Fixed			c_4	-1.640 ± 0.506	Fixed
		σ_2	1.245 ± 0.262	Fixed			c_5	-2.451 ± 0.449	Fixed
		f_2	$1.0 - (0.9914 \pm 0.0026)$	Fixed			C	-2.718 ± 0.034	Fixed

f_{SXF}	0.0608 ± 0.0122	Fixed

Table D.2: *PDF parameter values for the S_{SXF} component for the 5-variable fit in the 892 window. For further information, see Section 5.7.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To $q\bar{q}$ Bkgd. Sample* †	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To $q\bar{q}$ Bkgd. Sample* †	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
$q\overline{q}$	m_{ES}	ξ	-25.962 ± 2.226	-24.028 ± 0.727	$q\overline{q}$	$m_{K\pi}$	μ	$(8.997 \pm 0.019) \times 10^{-1}$	$(8.951 \pm 0.006) \times 10^{-1}$
		m_0	Fixed, to 5.29	Fixed			Γ	$(5.241 \pm 0.677) \times 10^{-2}$	$(4.811 \pm 0.211) \times 10^{-2}$
	ΔE	c_1	-1.186 ± 0.267	-1.166 ± 0.088			c_1	$(-1.370 \pm 1.710) \times 10^{-1}$	$(2.247 \pm 1.000) \times 10^{-1}$
	$\mathcal{F}$	μ_1	$(2.506 \pm 0.032) \times 10^{-1}$	$(2.485 \pm 0.001) \times 10^{-1}$			f_{BW}	0.2503 ± 0.0243	0.2183 ± 0.0072
		σ_1	$(1.972 \pm 0.025) \times 10^{-1}$	$(1.952 \pm 0.001) \times 10^{-1}$		$\cos\theta_H$	c_1	$(2.194 \pm 0.329) \times 10^{-1}$	$(2.163 \pm 0.109) \times 10^{-1}$
		μ_2	$(4.022 \pm 1.453) \times 10^{-1}$	$(2.772 \pm 0.221) \times 10^{-1}$			c_2	$(2.226 \pm 0.666) \times 10^{-1}$	$(2.697 \pm 0.221) \times 10^{-1}$
		σ_2	$(5.690 \pm 1.249) \times 10^{-1}$	$(4.625 \pm 0.252) \times 10^{-1}$					
		f_2	$1.0 - (0.9926 \pm 0.0048)$	0.0267 ± 0.0051					
$q\overline{q}$	m_{ES}	ξ	-24.381 ± 0.811		$q\overline{q}$	$m_{K\pi}$	μ	$(8.895 \pm 0.007) \times 10^{-1}$	
		m_0	Fixed, to 5.29				Γ	$(4.906 \pm 0.245) \times 10^{-2}$	
	ΔE	c_1	—				c_1	$(1.235 \pm 0.969) \times 10^{-1}$	
	$\mathcal{F}$	μ_1	$(2.433 \pm 0.012) \times 10^{-1}$	As			f_{BW}	0.2218 ± 0.0083	As
		σ_1	$(1.938 \pm 0.012) \times 10^{-1}$	above		$\cos\theta_H$	c_1	$(1.547 \pm 0.117) \times 10^{-1}$	above
		μ_2	$(2.784 \pm 0.250) \times 10^{-1}$				c_2	$(1.913 \pm 0.238) \times 10^{-1}$	
		σ_2	$(4.438 \pm 0.270) \times 10^{-1}$						
		f_2	$1.0 - (0.9751 \pm 0.0056)$						

†: Top section: $q\overline{q}$ background sample used is off-resonance sample; bottom section: $q\overline{q}$ background sample used is on-resonance sideband sample.

Table D.3: *PDF parameter values for the $q\overline{q}$ component for the 5-variable fit in the 892 window. For further information, see Section 5.7.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Exclusive MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Exclusive MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
$(B\overline{B})_1$	m_{ES}	μ	5.2791 ± 0.0000	Fixed	$(B\overline{B})_1$	$m_{K\pi}$	μ	$(8.129 \pm 0.003) \times 10^{-1}$	Fixed
		σ	$(2.5774 \pm 0.0393) \times 10^{-3}$	Fixed			Γ	$(1.113 \pm 0.115) \times 10^{-2}$	Fixed
		α	2.1393 ± 0.0998	Fixed			σ	$(8.144 \pm 0.562) \times 10^{-3}$	Fixed
		n	1.1752 ± 0.1513	Fixed			c_1	$(6.597 \pm 1.649) \times 10^{1}$	Fixed
	ΔE	μ	$(-6.787 \pm 0.066) \times 10^{-2}$	Fixed			c_2	$(-8.603 \pm 1.221) \times 10^{1}$	Fixed
		σ	$(2.400 \pm 0.063) \times 10^{-2}$	Fixed			C	0.205 ± 2.898	Fixed
		c_1	-8.843 ± 0.820	Fixed			f_{Voig}	0.9344 ± 0.0127	Fixed
		f_{Gaus}	0.9005 ± 0.0135	Fixed		$\cos\theta_H$	c_1	-5.896 ± 0.819	Fixed
	$\mathcal{F}$	μ_1	$(-9.017 \pm 0.428) \times 10^{-2}$	Fixed			c_2	$(-2.595 \pm 0.240) \times 10^{1}$	Fixed
		σ_1	$(2.408 \pm 0.031) \times 10^{-1}$	Fixed			c_3	$(-2.144 \pm 0.294) \times 10^{1}$	Fixed
		μ_2	$(-0.654 \pm 2.968) \times 10^{-1}$	Fixed			c_4	$(3.243 \pm 0.334) \times 10^{1}$	Fixed
		σ_2	1.447 ± 0.368	Fixed			c_5	$(2.867 \pm 0.220) \times 10^{1}$	Fixed
		f_2	$1.0 - (0.9903 \pm 0.0028)$	Fixed			C	$(-0.826 \pm 4.355) \times 10^{-2}$	Fixed

Table D.4: *PDF parameter values for the $(B\overline{B})_1$ component for the 5-variable fit in the 892 window. For further information, see Section 5.10.5.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Exclusive MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Exclusive MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
$(B\overline{B})_2$	m_{ES}	μ	5.2790 ± 0.0000	Fixed	$(B\overline{B})_2$	$\mathcal{F}$	μ_1	$(-8.015 \pm 0.377) \times 10^{-2}$	Fixed
		σ	$(2.7098 \pm 0.0763) \times 10^{-3}$	Fixed			σ_1	$(2.243 \pm 0.099) \times 10^{-1}$	Fixed
		ξ	-48.508 ± 4.396	Fixed			μ_2	$(-0.801 \pm 2.042) \times 10^{-1}$	Fixed
		m_0	Fixed, to 5.29	Fixed			σ_2	$(7.473 \pm 2.300) \times 10^{-1}$	Fixed
		f_{Argus}	0.3146 ± 0.0272	Fixed			f_2	$1.0 - (0.9695 \pm 0.0212)$	Fixed
	ΔE	μ	$(-8.598 \pm 0.369) \times 10^{-2}$	Fixed		$m_{K\pi}$	c_1	$(-4.982 \pm 6.945) \times 10^{1}$	Fixed
		σ	$(3.363 \pm 0.300) \times 10^{-2}$	Fixed			c_2	$(0.973 \pm 1.216) \times 10^{2}$	Fixed
		c_1	-4.771 ± 0.800	Fixed		$\cos\theta_H$	`Keys`	—	Fixed
		f_{Gaus}	0.6252 ± 0.0415	Fixed			[98]		

Table D.5: *PDF parameter values for the $(B\overline{B})_2$ component for the 5-variable fit in the 892 window. For further information, see Section 5.10.5.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Exclusive MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Exclusive MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
$(B\overline{B})_3$	m_{ES}	μ	5.2792 ± 0.0001	Fixed	$(B\overline{B})_3$	$m_{K\pi}$	μ	$(9.082 \pm 0.049) \times 10^{-1}$	Fixed
		σ	$(2.7625 \pm 0.0926) \times 10^{-3}$	Fixed			σ	$(5.107 \pm 0.580) \times 10^{-2}$	Fixed
		α	1.8895 ± 0.1521	Fixed			c_1	-1.948 ± 0.458	Fixed
		n	$(9.312 \pm 1.921) \times 10^{-1}$	Fixed			f_{Gaus}	0.5532 ± 0.0749	Fixed
	ΔE	μ	$(4.822 \pm 0.172) \times 10^{-2}$	Fixed		$\cos\theta_H$	c_1	$(1.987 \pm 3.261) \times 10^{-1}$	Fixed
		σ	$(2.612 \pm 0.230) \times 10^{-2}$	Fixed			c_2	$(2.859 \pm 4.039) \times 10^{-1}$	Fixed
		c_1	3.768 ± 3.877	Fixed			c_3	$(6.911 \pm 4.480) \times 10^{-2}$	Fixed
		f_{Gaus}	0.7733 ± 0.0825	Fixed			C	-2.623 ± 0.163	Fixed
	$\mathcal{F}$	μ	$(-7.267 \pm 0.882) \times 10^{-2}$	Fixed					
		σ	$(2.402 \pm 0.062) \times 10^{-1}$	Fixed					

Table D.6: *PDF parameter values for the $(B\overline{B})_3$ component for the 5-variable fit in the 892 window. For further information, see Section 5.10.5.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Signal MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Signal MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
S	m_{ES}	μ	5.2793 ± 0.0000	Fixed	S	$\mathcal{F}$	μ_1	$(-1.015 \pm 0.009) \times 10^{-1}$	Fixed
		σ	$(2.6450 \pm 0.0092) \times 10^{-3}$	Fixed			σ_1	$(2.357 \pm 0.007) \times 10^{-1}$	Fixed
		α	2.0219 ± 0.0223	Fixed			μ_2	$(-1.738 \pm 0.708) \times 10^{-1}$	Fixed
		n	1.3348 ± 0.0370	Fixed			σ_2	1.303 ± 0.079	Fixed
	ΔE	μ_1	$(1.89 \pm 1.08) \times 10^{-4}$	Fixed			f_2	$1.0 - (0.9936 \pm 0.0005)$	Fixed
		σ_1	$(2.103 \pm 0.015) \times 10^{-2}$	Fixed					
		μ_2	$(-0.986 \pm 1.035) \times 10^{-3}$	Fixed					
		σ_2	$(6.849 \pm 0.344) \times 10^{-2}$	Fixed					
		f_2	$1.0 - (0.8284 \pm 0.0080)$	Fixed					

Table D.7: *PDF parameter values for the S component for the 3-variable fit in the 892 window. For further information, see Section 5.8.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To $q\overline{q}$ Bkgd. Sample* †	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To $q\overline{q}$ Bkgd. Sample* †	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
$q\overline{q}$	m_{ES}	ξ	-25.825 ± 2.183	-23.745 ± 0.717	$q\overline{q}$	$\mathcal{F}$	μ_1	$(2.495 \pm 0.031) \times 10^{-1}$	$(2.560 \pm 0.011) \times 10^{-1}$
		m_0	Fixed, to 5.29	Fixed			σ_1	$(1.976 \pm 0.025) \times 10^{-1}$	$(1.948 \pm 0.011) \times 10^{-1}$
	ΔE	c_1	-1.116 ± 0.262	-1.103 ± 0.086			μ_2	$(4.033 \pm 1.513) \times 10^{-1}$	$(2.857 \pm 0.212) \times 10^{-1}$
							σ_2	$(5.745 \pm 1.275) \times 10^{-1}$	$(4.569 \pm 0.241) \times 10^{-1}$
							f_2	$1.0 - (0.9931 \pm 0.0045)$	$1.0 - (0.9735 \pm 0.0049)$
$q\overline{q}$	m_{ES}	ξ	-24.100 ± 0.797		$q\overline{q}$	$\mathcal{F}$	μ_1	$(2.507 \pm 0.012) \times 10^{-1}$	
		m_0	Fixed, to 5.29				σ_1	$(1.934 \pm 0.012) \times 10^{-1}$	
	ΔE	c_1	—	As above			μ_2	$(2.955 \pm 0.253) \times 10^{-1}$	As above
							σ_2	$(4.444 \pm 0.264) \times 10^{-1}$	
							f_2	$1.0 - (0.9762 \pm 0.0052)$	

†: Top section: $q\overline{q}$ background sample used is off-resonance sample; bottom section: $q\overline{q}$ background sample used is on-resonance sideband sample.

Table D.8: *PDF parameter values for the $q\overline{q}$ component for the 3-variable fit in the 892 window. For further information, see Section 5.8.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Exclusive MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Exclusive MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
$(B\overline{B})_1$	m_{ES}	μ	5.2791 ± 0.0000	Fixed	$(B\overline{B})_2$	ΔE	μ	$(-8.646 \pm 0.364) \times 10^{-2}$	Fixed
		σ	$(2.5774 \pm 0.0393) \times 10^{-3}$	Fixed			σ	$(3.403 \pm 0.287) \times 10^{-2}$	Fixed
		α	2.1393 ± 0.0998	Fixed			c_1	-4.784 ± 0.777	Fixed
		n	1.1752 ± 0.1513	Fixed			f_{Gaus}	0.6285 ± 0.0416	Fixed
	ΔE	μ	$(-6.787 \pm 0.066) \times 10^{-2}$	Fixed		$\mathcal{F}$	μ_1	$(-5.619 \pm 0.369) \times 10^{-2}$	Fixed
		σ	$(2.400 \pm 0.063) \times 10^{-2}$	Fixed			σ_1	$(2.239 \pm 0.086) \times 10^{-1}$	Fixed
		c_1	-8.843 ± 0.820	Fixed			μ_2	$(3.877 \pm 2.031) \times 10^{-1}$	Fixed
		f_{Gaus}	0.9005 ± 0.0135	Fixed			σ_2	$(8.657 \pm 2.271) \times 10^{-1}$	Fixed
	$\mathcal{F}$	μ_1	$(-8.048 \pm 0.425) \times 10^{-2}$	Fixed			f_2	$1.0 - (0.9788 \pm 0.0108)$	Fixed
		σ_1	$(2.392 \pm 0.031) \times 10^{-1}$	Fixed	$(B\overline{B})_3$	m_{ES}	μ	5.2793 ± 0.0001	Fixed
		μ_2	$(-0.226 \pm 2.738) \times 10^{-1}$	Fixed			σ	$(2.8916 \pm 0.0873) \times 10^{-3}$	Fixed
		σ_2	1.383 ± 0.326	Fixed			α	1.9240 ± 0.1248	Fixed
		f_2	$1.0 - (0.9900 \pm 0.0028)$	Fixed			n	$(8.372 \pm 1.533) \times 10^{-1}$	Fixed
$(B\overline{B})_2$	m_{ES}	μ	5.2790 ± 0.0001	Fixed		ΔE	μ	$(4.816 \pm 0.163) \times 10^{-2}$	Fixed
		σ	$(2.7063 \pm 0.0749) \times 10^{-3}$	Fixed			σ	$(2.665 \pm 0.233) \times 10^{-2}$	Fixed
		ξ	-48.562 ± 4.332	Fixed			c_1	2.727 ± 3.956	Fixed
		m_0	Fixed, to 5.29	Fixed			f_{Gaus}	0.7604 ± 0.0814	Fixed
		f_{Argus}	0.3165 ± 0.0272	Fixed		$\mathcal{F}$	μ	$(-6.597 \pm 0.831) \times 10^{-2}$	Fixed
							σ	$(2.397 \pm 0.059) \times 10^{-1}$	Fixed

Table D.9: *PDF parameter values for the $B\overline{B}$ background components for the 3-variable fit in the 892 window. For further information, see Section 5.11.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Signal MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Signal MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
S	m_{ES}	μ	5.2793 ± 0.0000	Fixed	S	$\mathcal{F}$	μ	$(-1.188 \pm 0.138) \times 10^{-2}$	Fixed
		σ	$(2.7187 \pm 0.0157) \times 10^{-3}$	Fixed			σ	$(2.142 \pm 0.010) \times 10^{-1}$	Fixed
		α	2.1511 ± 0.0304	Fixed					
		n	$(5.998 \pm 0.304) \times 10^{-1}$	Fixed					
	ΔE	μ_1	$(-1.57 \pm 1.84) \times 10^{-4}$	Fixed					
		σ_1	$(2.016 \pm 0.024) \times 10^{-2}$	Fixed					
		μ_2	$(-2.051 \pm 0.456) \times 10^{-2}$	Fixed					
		σ_2	$(8.907 \pm 0.872) \times 10^{-2}$	Fixed					
		f_2	$1.0 - (0.7735 \pm 0.0110)$	Fixed					

Table D.10: *PDF parameter values for the S component for the 3-variable fit in the 1430 window. For further information, see Section 5.9.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To $q\overline{q}$ Bkgd. Sample* †	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To $q\overline{q}$ Bkgd. Sample* †	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
$q\overline{q}$	m_{ES}	ξ	-22.425 ± 2.057	-21.493 ± 0.668	$q\overline{q}$	$\mathcal{F}$	μ_1	$(2.451 \pm 0.082) \times 10^{-1}$	$(2.385 \pm 0.021) \times 10^{-1}$
		m_0	Fixed, to 5.29	Fixed			σ_1	$(1.873 \pm 0.058) \times 10^{-1}$	$(1.863 \pm 0.013) \times 10^{-1}$
	ΔE	c_1	$(-8.146 \pm 2.468) \times 10^{-1}$	-1.030 ± 0.080			μ_2	$(3.234 \pm 0.132) \times 10^{-1}$	$(3.355 \pm 0.039) \times 10^{-1}$
							σ_2	$(1.211 \pm 0.106) \times 10^{-1}$	$(1.114 \pm 0.029) \times 10^{-1}$
							f_2	$1.0 - (0.6437 \pm 0.0996)$	$1.0 - (0.7237 \pm 0.0194)$
$q\overline{q}$	m_{ES}	ξ	-21.017 ± 0.733		$q\overline{q}$	$\mathcal{F}$	μ_1	$(2.355 \pm 0.023) \times 10^{-1}$	
		m_0	Fixed, to 5.29				σ_1	$(1.857 \pm 0.014) \times 10^{-1}$	
	ΔE	c_1	—	As above			μ_2	$(3.300 \pm 0.043) \times 10^{-1}$	As above
							σ_2	$(1.067 \pm 0.036) \times 10^{-1}$	
							f_2	$1.0 - (0.7341 \pm 0.0221)$	

†: Top section: $q\overline{q}$ background sample used is off-resonance sample; bottom section: $q\overline{q}$ background sample used is on-resonance sideband sample.

Table D.11: *PDF parameter values for the $q\overline{q}$ component for the 3-variable fit in the 1430 window. For further information, see Section 5.9.*

Component	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Exclusive MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*	*Component*	*Variable*	*Parameter*	*Value Obtained From One-dimensional Fit To Exclusive MC Sample*	*Value Obtained From Multi-dimensional Fit To Final On-res. Sample*
$(B\overline{B})_2$	m_{ES}	μ	5.2791 ± 0.0001	Fixed	$(B\overline{B})_3$	ΔE	μ	$(6.311 \pm 1.129) \times 10^{-2}$	Fixed
		σ	$(2.5313 \pm 0.0768) \times 10^{-3}$	Fixed			σ	$(3.975 \pm 1.375) \times 10^{-2}$	Fixed
		ξ	-54.538 ± 8.516	Fixed			c_1	1.501 ± 3.906	Fixed
		m_0	Fixed, to 5.29	Fixed			f_{Gaus}	0.3248 ± 0.2052	Fixed
		f_{Argus}	0.3245 ± 0.0158	Fixed		$\mathcal{F}$	μ	$(5.268 \pm 0.879) \times 10^{-2}$	Fixed
	ΔE	μ	$(-7.223 \pm 0.236) \times 10^{-2}$	Fixed			σ	$(2.179 \pm 0.062) \times 10^{-1}$	Fixed
		σ	$(3.344 \pm 0.248) \times 10^{-2}$	Fixed	$(B\overline{B})_4$	m_{ES}	μ	5.2793 ± 0.0001	Fixed
		c_1	-0.078 ± 4.247	Fixed			σ_1	$(2.8534 \pm 0.1744) \times 10^{-3}$	Fixed
		f_{Gaus}	0.7983 ± 0.0567	Fixed			σ_2	$(8.819 \pm 5.592) \times 10^{-4}$	Fixed
	$\mathcal{F}$	μ_1	$(1.950 \pm 0.507) \times 10^{-2}$	Fixed			f_2	0.0586 ± 0.0620	Fixed
		σ_1	$(1.974 \pm 0.037) \times 10^{-1}$	Fixed			ξ	-84.353 ± 12.976	Fixed
		μ_2	1.012 ± 0.521	Fixed			m_0	Fixed, to 5.29	Fixed
		σ_2	$(3.474 \pm 3.095) \times 10^{-1}$	Fixed			f_{Argus}	0.3141 ± 0.0245	Fixed
		f_2	$1.0 - (0.9981 \pm 0.0020)$	Fixed		ΔE	μ	$(5.776 \pm 0.192) \times 10^{-2}$	Fixed
$(B\overline{B})_3$	m_{ES}	μ	5.2795 ± 0.0004	Fixed			σ	$(1.902 \pm 0.281) \times 10^{-2}$	Fixed
		σ	$(2.9029 \pm 0.5180) \times 10^{-3}$	Fixed			c_1	7.434 ± 1.017	Fixed
		ξ	-104.560 ± 9.657	Fixed			f_{Gaus}	0.3997 ± 0.0656	Fixed
		m_0	Fixed, to 5.29	Fixed		$\mathcal{F}$	μ	$(2.575 \pm 0.764) \times 10^{-2}$	Fixed
		f_{Argus}	0.7772 ± 0.0408	Fixed			σ	$(2.225 \pm 0.054) \times 10^{-1}$	Fixed

Table D.12: *PDF parameter values for the $B\overline{B}$ background components for the 3-variable fit in the 1430 window. For further information, see Section 5.11.*

E

Fit variable correlations

Correlation values between fit variables for the $B\overline{B}$ background components in the 3-variable fits are presented in Tables E.1 and E.2 (for the 892 and 1430 windows respectively). All other relevant correlation values are given in the main text.

Figures E.1-E.17 illustrate correlations between fit variables in the form of two-dimensional scatter plots.

ΔE	$-0.248, -0.413,$ 0.220	
$\mathcal{F}$	$-0.081, 0.007,$ -0.096	$0.026, -0.018,$ -0.022
	m_{ES}	ΔE

Table E.1: *Correlations (calculated using Eq. (5.13)) between variables used in the 3-variable fit in the 892 window. For each cell, the first value is for the $(B\overline{B})_1$ component, the second is for the $(B\overline{B})_2$ component and the third is for the $(B\overline{B})_3$ component.*

ΔE	$-0.399, 0.156,$ 0.329	
$\mathcal{F}$	$-0.014, -0.124,$ -0.123	$0.025, -0.050,$ -0.081
	m_{ES}	ΔE

Table E.2: *Correlations (calculated using Eq. (5.13)) between variables used in the 3-variable fit in the 1430 window. For each cell, the first value is for the $(B\overline{B})_2$ component, the second is for the $(B\overline{B})_3$ component and the third is for the $(B\overline{B})_4$ component.*

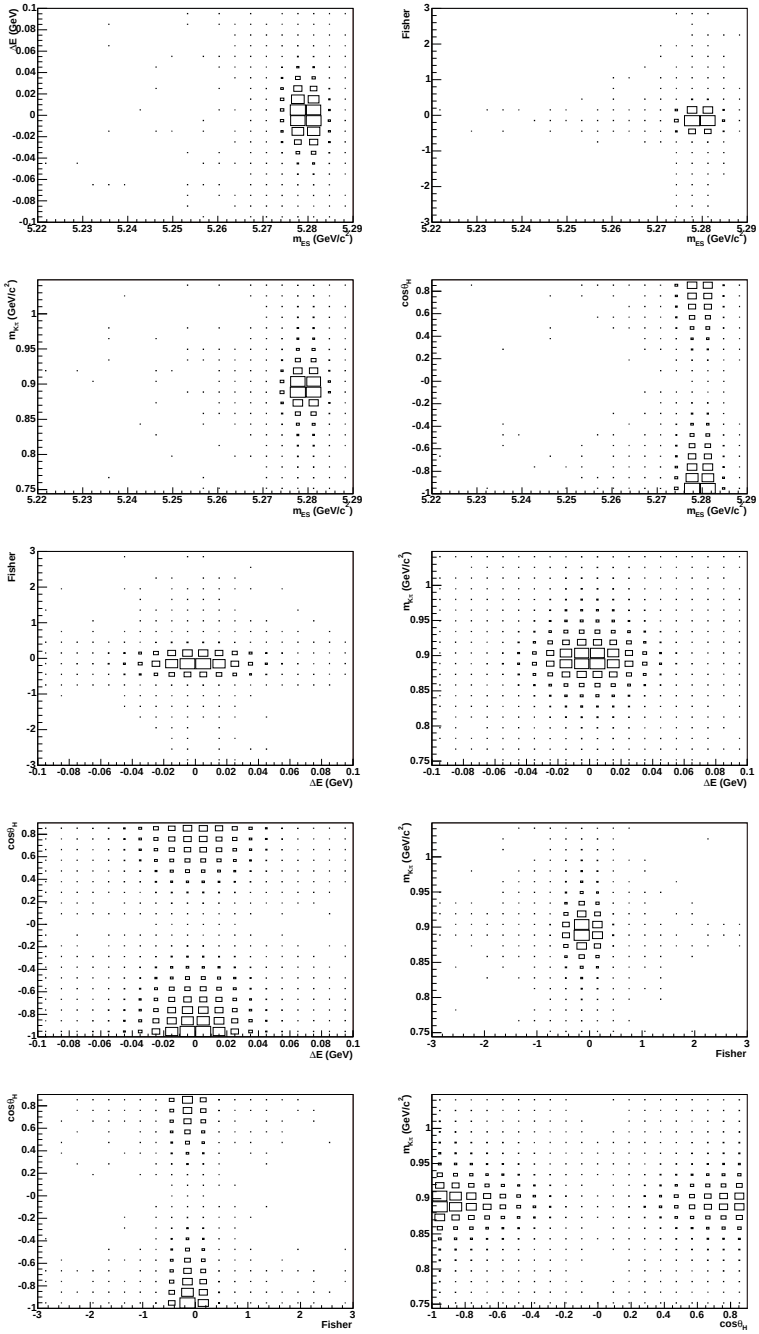

Figure E.1: *Scatter plots illustrating correlations between fit variables for the S_{TRU} component for the 5-variable fit in the 892 window. The sample used is the MC sample discussed in Section 5.7.2. Correlation values, calculated using Eq. (5.13), are presented in Table 5.12 (Section 5.7.1).*

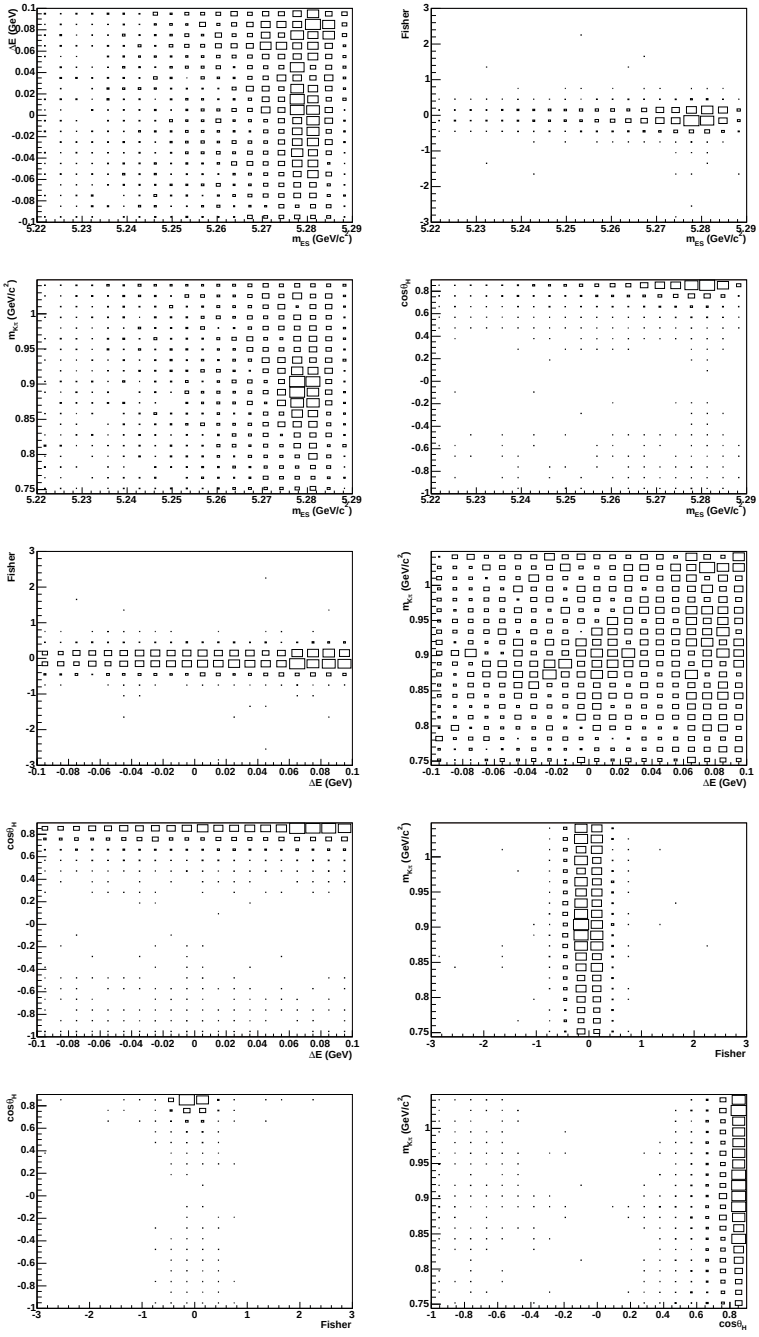

Figure E.2: *Scatter plots illustrating correlations between fit variables for the S_{SXF} component for the 5-variable fit in the 892 window. The sample used is the MC sample discussed in Section 5.7.2. Correlation values, calculated using Eq. (5.13), are presented in Table 5.12 (Section 5.7.1).*

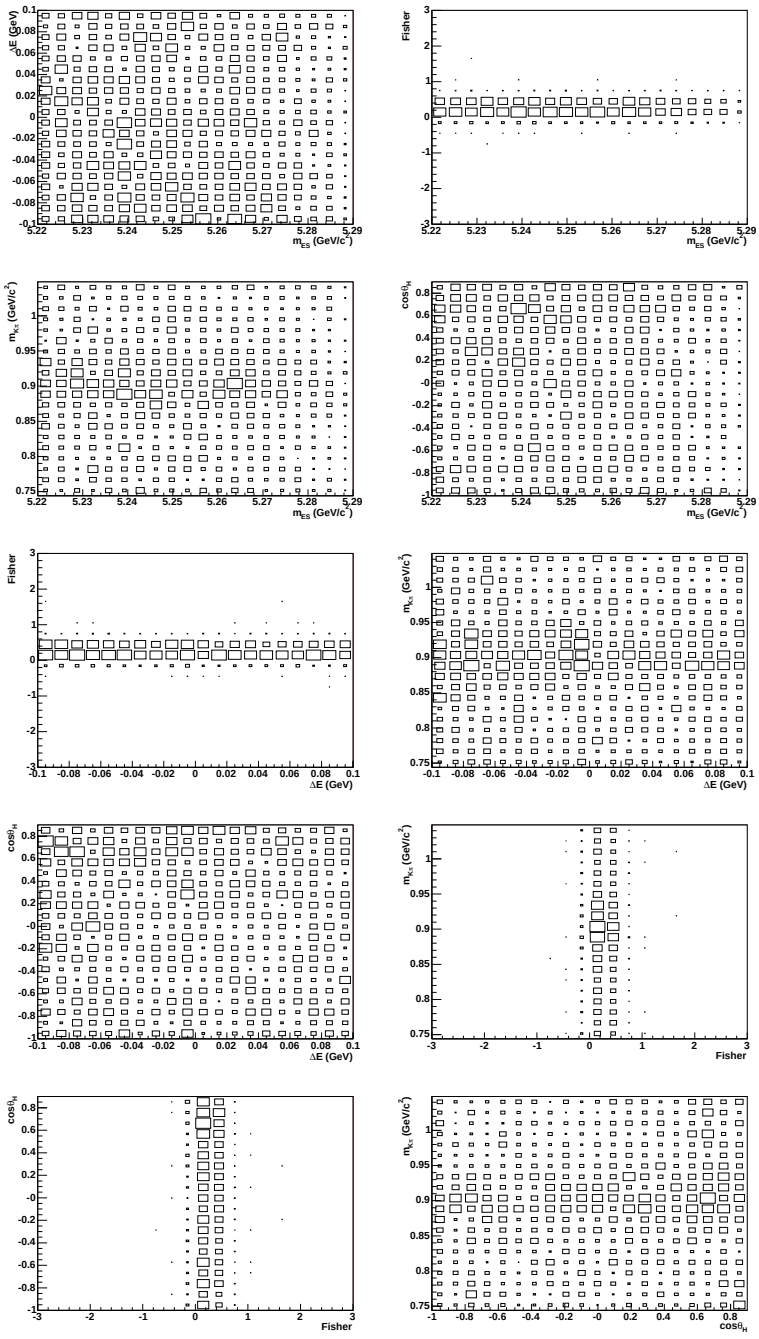

Figure E.3: *Scatter plots illustrating correlations between fit variables for the $q\overline{q}$ component for the 5-variable fit in the 892 window. The sample used is the off-resonance sample discussed in Section 5.7.2. Correlation values, calculated using Eq. (5.13), are presented in Table 5.12 (Section 5.7.1).*

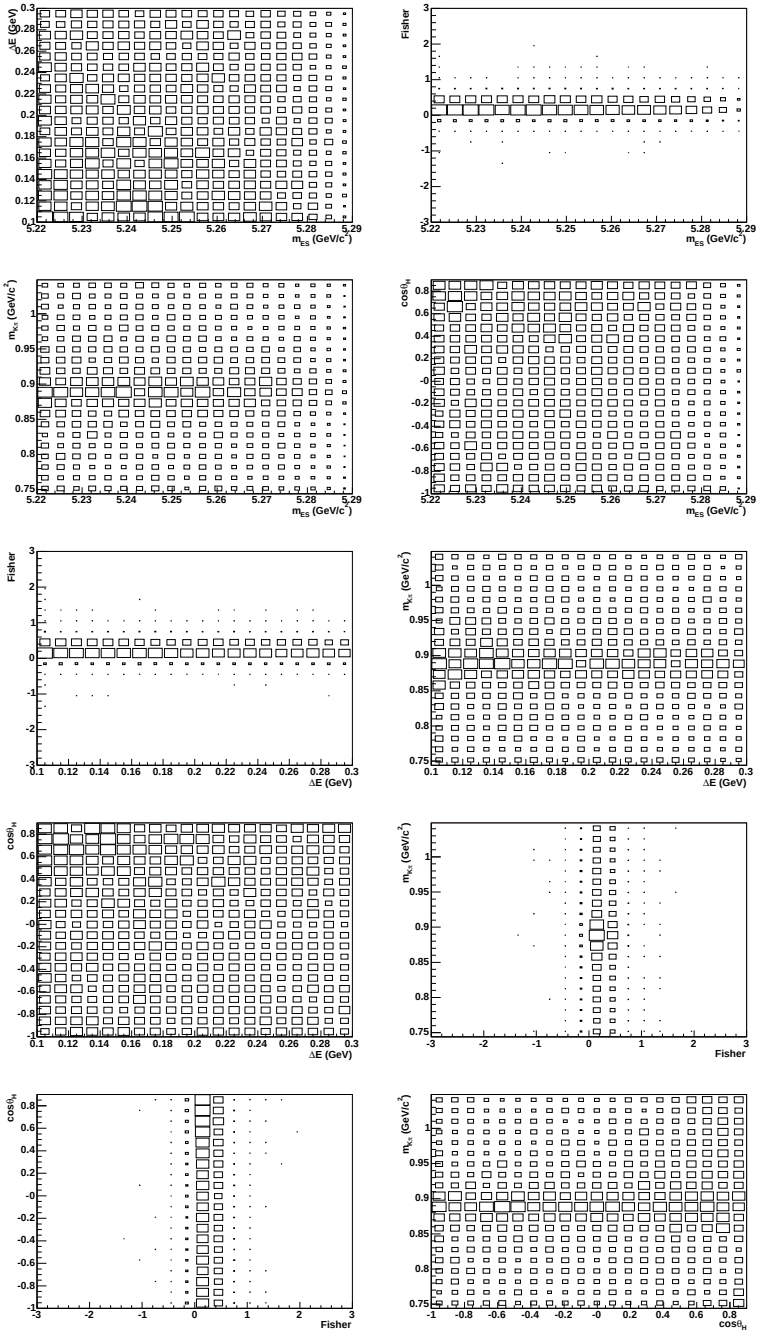

Figure E.4: *Scatter plots illustrating correlations between fit variables for the $q\bar{q}$ component for the 5-variable fit in the 892 window. The sample used is the on-resonance sideband sample discussed in Section 5.7.2. Correlation values, calculated using Eq. (5.13), are presented in Table 5.12 (Section 5.7.1).*

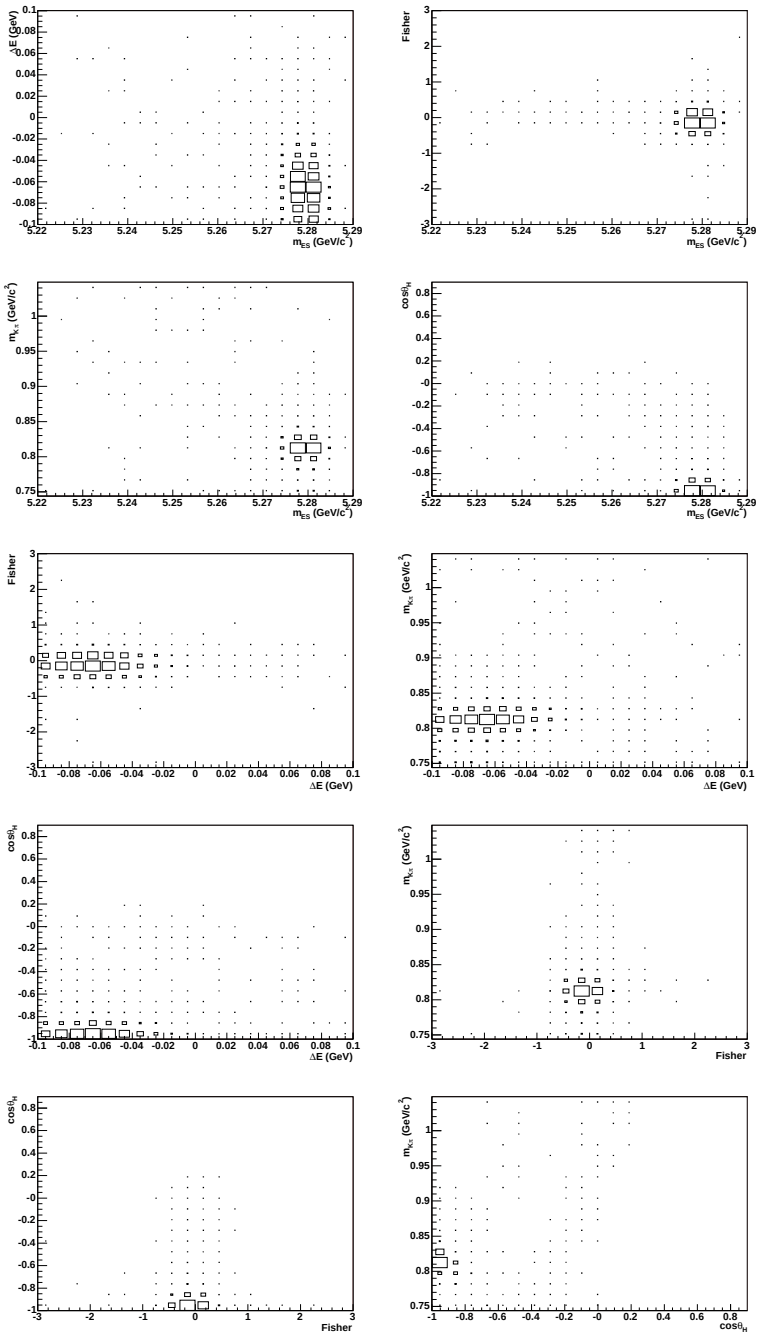

Figure E.5: *Scatter plots illustrating correlations between fit variables for the* $(B\overline{B})_1$ *component for the 5-variable fit in the 892 window. The sample used is a MC sample in which the decay* $B^+ \to \phi(1020)(\to K^+K^-)K^+$ *is simulated. Correlation values, calculated using Eq. (5.13), are presented in Table 5.22 (Section 5.10.5).*

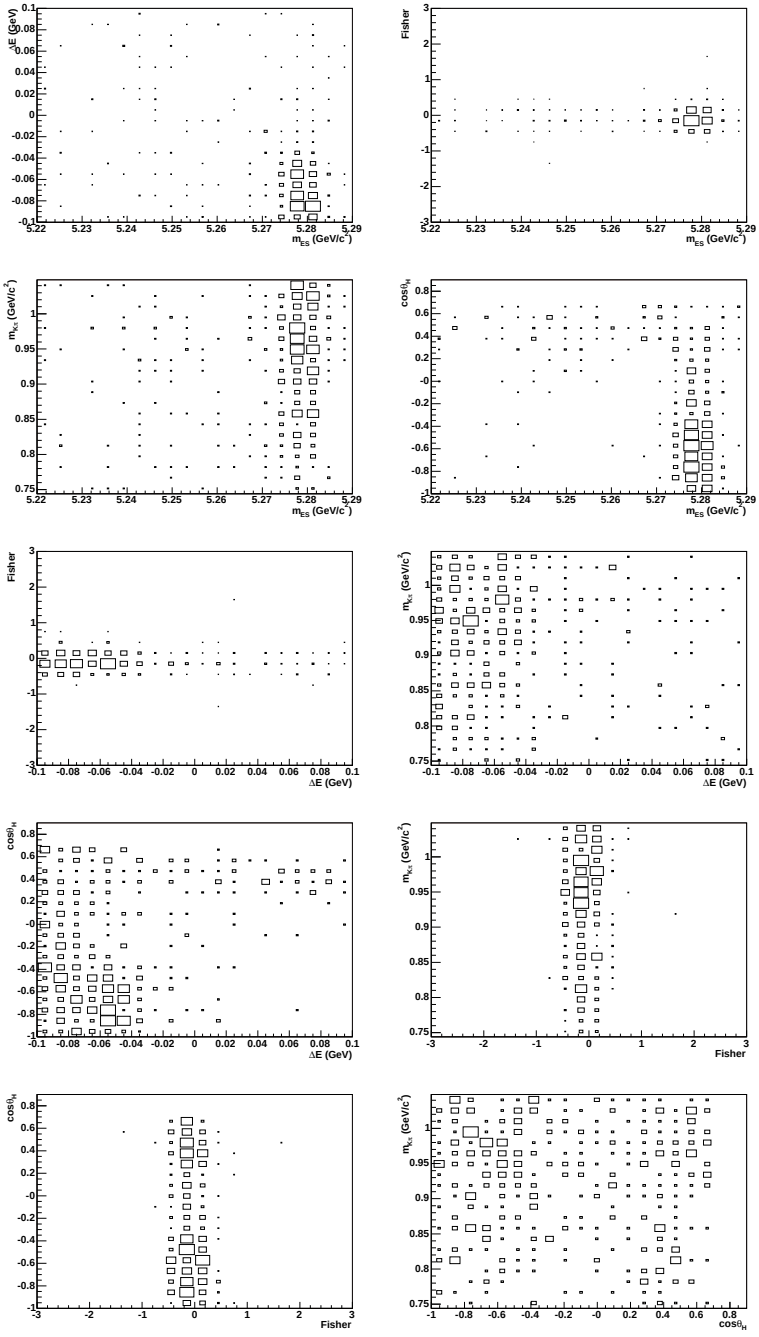

Figure E.6: *Scatter plots illustrating correlations between fit variables for the* $(B\overline{B})_2$ *component for the 5-variable fit in the 892 window. The sample used is a MC sample in which the decay* $B^+ \to K^+K^-K^+$ *(Belle Dalitz plot model* B_0*, solution 1 [42], minus the* $B^+ \to \phi(1020)K^+$ *contribution) is simulated. Correlation values, calculated using Eq. (5.13), are presented in Table 5.22 (Section 5.10.5).*

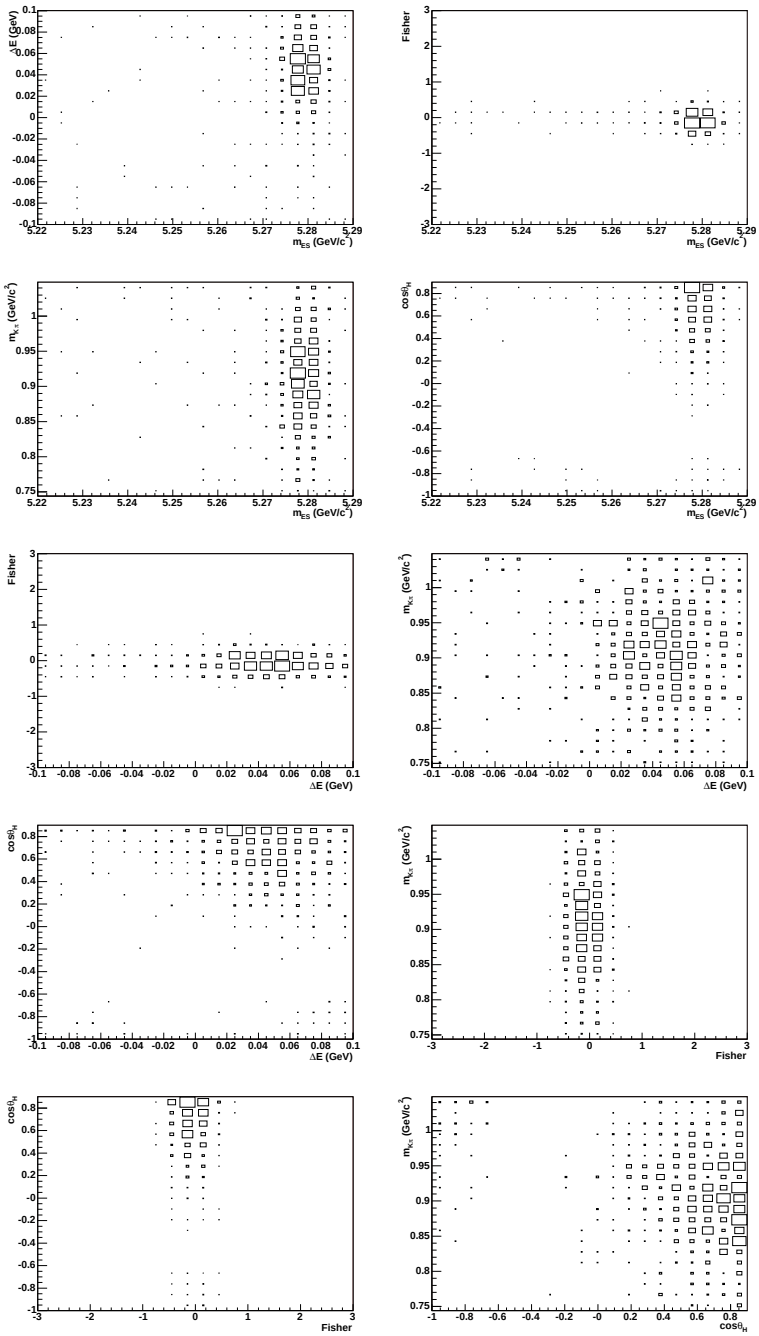

Figure E.7: *Scatter plots illustrating correlations between fit variables for the* $(B\overline{B})_3$ *component for the 5-variable fit in the 892 window. The sample used is a MC sample in which the decay* $B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$ *is simulated. Correlation values, calculated using Eq. (5.13), are presented in Table 5.22 (Section 5.10.5).*

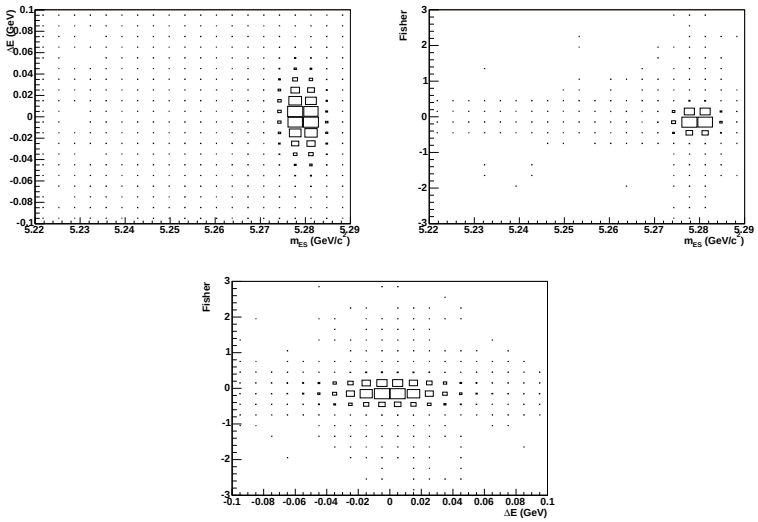

Figure E.8: *Scatter plots illustrating correlations between fit variables for the S component for the 3-variable fit in the 892 window. The sample used is the MC sample discussed in Section 5.8. Correlation values, calculated using Eq. (5.13), are presented in Table 5.14 (Section 5.8).*

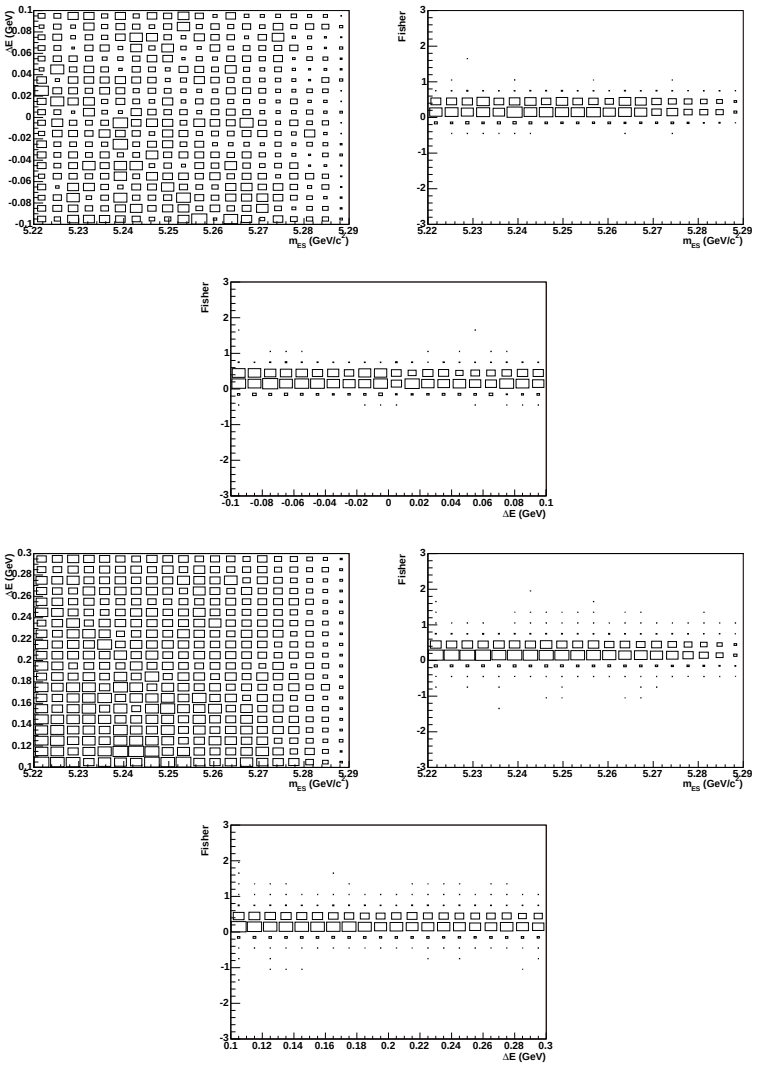

Figure E.9: *Scatter plots illustrating correlations between fit variables for the $q\overline{q}$ component for the 3-variable fit in the 892 window. Top three plots: the sample used is the off-resonance sample discussed in Section 5.8; bottom three plots: the sample used is the on-resonance sideband sample discussed in Section 5.8; Correlation values, calculated using Eq. (5.13), are presented in Table 5.14 (Section 5.8).*

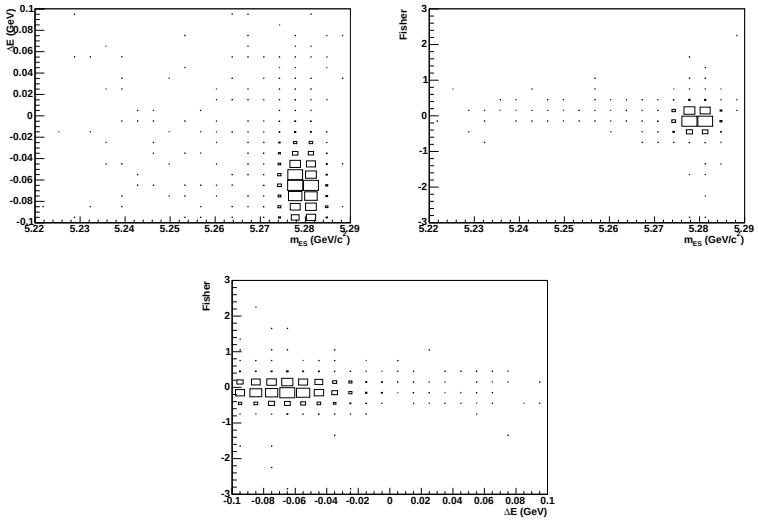

Figure E.10: *Scatter plots illustrating correlations between fit variables for the $(B\overline{B})_1$ component for the 3-variable fit in the 892 window. The sample used is a MC sample in which the decay $B^+ \to \phi(1020)(\to K^+K^-)K^+$ is simulated. Correlation values, calculated using Eq. (5.13), are presented in Table E.1.*

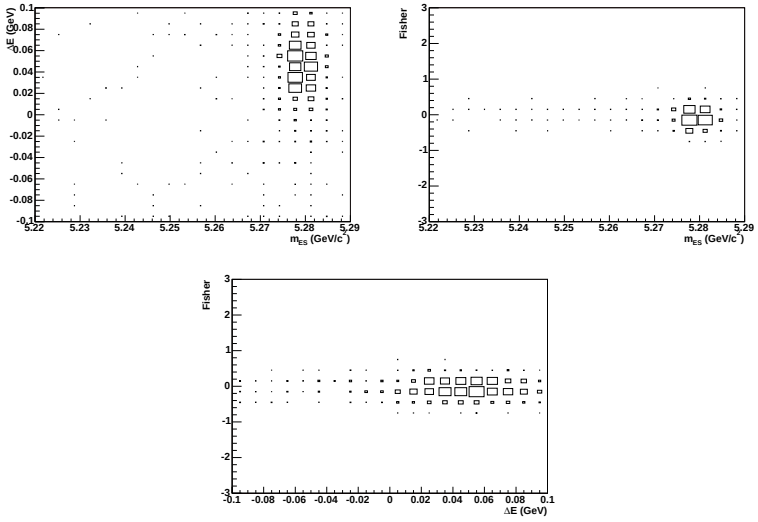

Figure E.11: *Scatter plots illustrating correlations between fit variables for the $(B\overline{B})_3$ component for the 3-variable fit in the 892 window. The sample used is a MC sample in which the decay $B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$ is simulated. Correlation values, calculated using Eq. (5.13), are presented in Table E.1.*

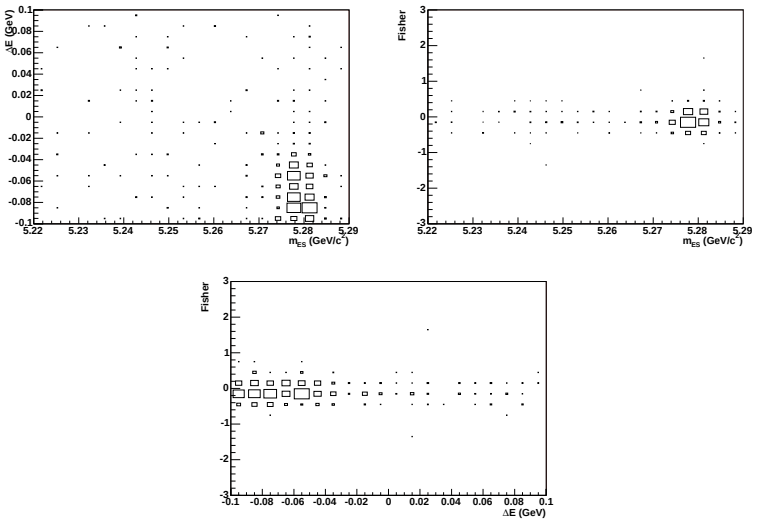

Figure E.12: *Scatter plots illustrating correlations between fit variables for the* $(B\overline{B})_2$ *component for the 3-variable fit in the 892 window. The sample used is a MC sample in which the decay* $B^+ \to K^+K^-K^+$ *(Belle Dalitz plot model* B_0*, solution 1 [42], minus the* $B^+ \to \phi(1020)K^+$ *contribution) is simulated. Correlation values, calculated using Eq. (5.13), are presented in Table E.1.*

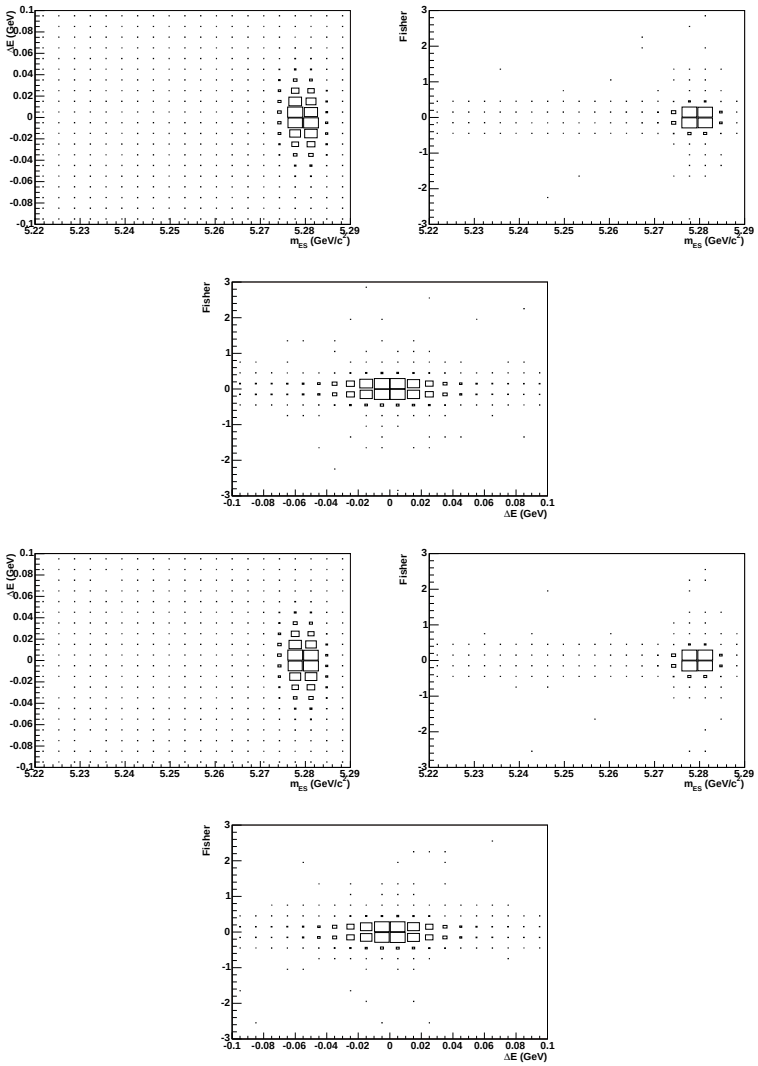

Figure E.13: *Scatter plots illustrating correlations between fit variables for the S component for the 3-variable fit in the 1430 window. The samples used are the MC samples discussed in Section 5.9. For the top three plots, the $m_{K\pi}$ lineshape is modelled as LASS; for the bottom three plots, the $m_{K\pi}$ lineshape is modelled as Breit-Wigner. Correlation values, calculated using Eq. (5.13), are presented in Table 5.16 (Section 5.9).*

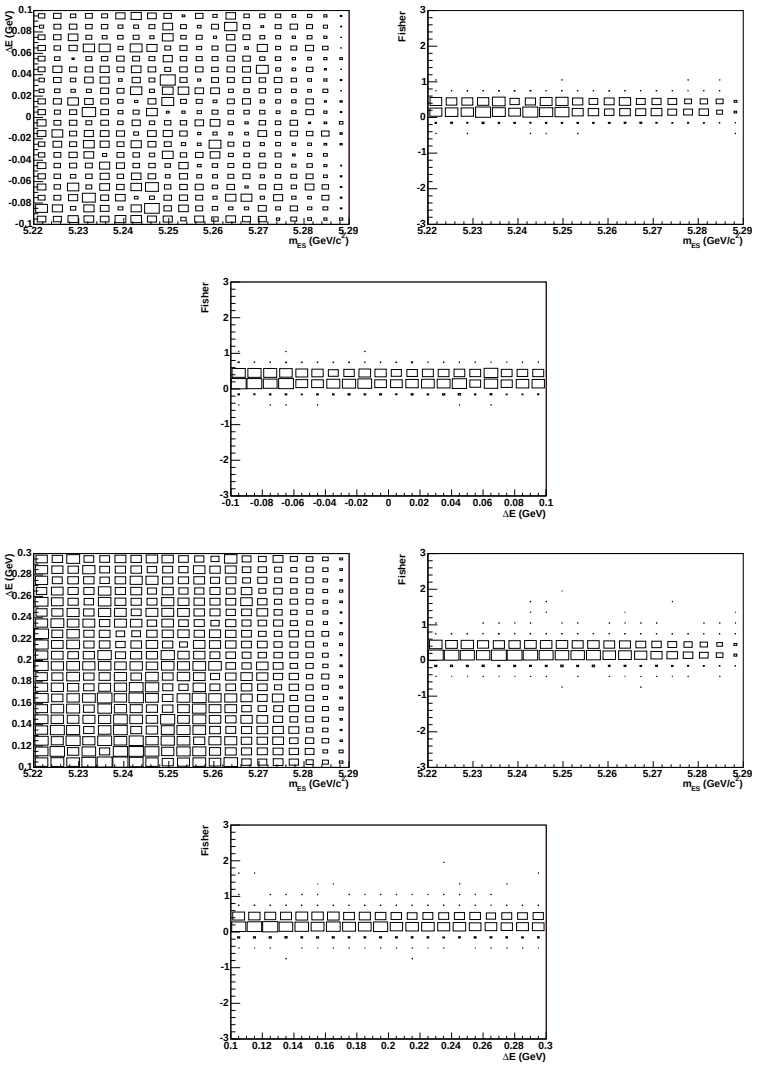

Figure E.14: *Scatter plots illustrating correlations between fit variables for the $q\overline{q}$ component for the 3-variable fit in the 1430 window. Top three plots: the sample used is the off-resonance sample discussed in Section 5.9; bottom three plots: the sample used is the on-resonance sideband sample discussed in Section 5.9; Correlation values, calculated using Eq. (5.13), are presented in Table 5.16 (Section 5.9).*

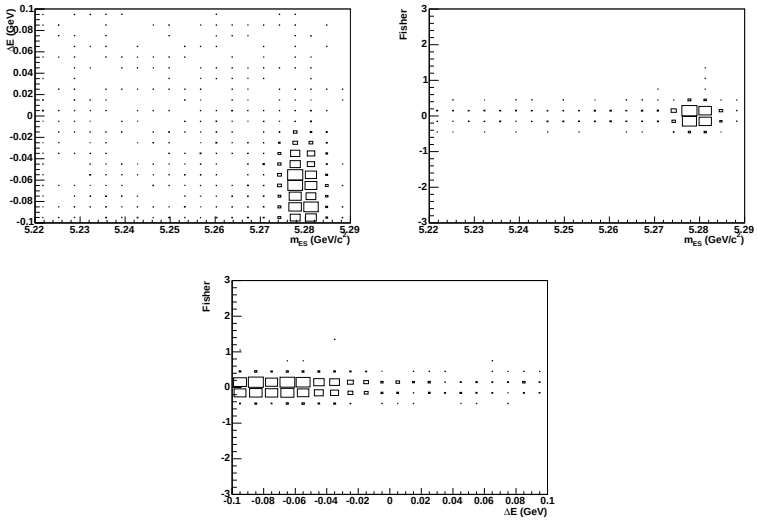

Figure E.15: *Scatter plots illustrating correlations between fit variables for the* $(B\overline{B})_2$ *component for the 3-variable fit in the 1430 window. The sample used is a MC sample in which the decay* $B^+ \to K^+K^-K^+$ *(Belle Dalitz plot model* B_0*, solution 1 [42], minus the* $B^+ \to \phi(1020)K^+$ *contribution) is simulated. Correlation values, calculated using Eq. (5.13), are presented in Table E.2.*

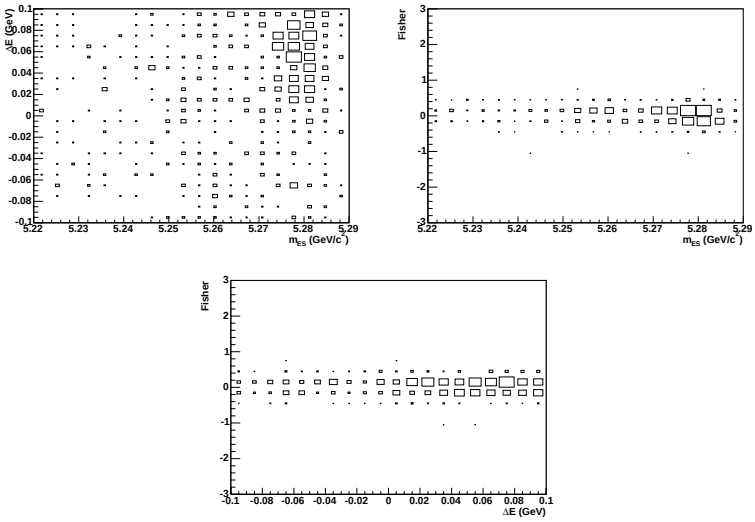

Figure E.16: *Scatter plots illustrating correlations between fit variables for the $(B\overline{B})_3$ component for the 3-var. fit in the 1430 window. The sample used is a MC sample in which the decay $B^+ \to \rho(770)^0(\to \pi^+\pi^-)K^+$ is simulated. Correlation values, calculated using Eq. (5.13), are presented in Table E.2.*

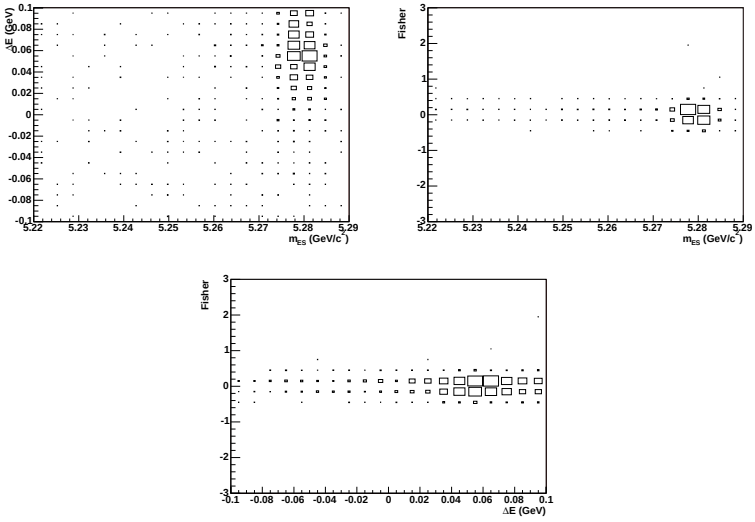

Figure E.17: *Scatter plots illustrating correlations between fit variables for the $(B\overline{B})_4$ component for the 3-var. fit in the 1430 window. The sample used is a MC sample in which the decay $B^+ \to f_0(980)(\to \pi^+\pi^-)K^+$ is simulated. Correlation values, calculated using Eq. (5.13), are presented in Table E.2.*

F

LASS parameters

Where the lineshape of $m_{K\pi}$ is modelled as the LASS shape for the decay $B^+ \to \overline{K}_0^*(1430)^0(\to K^-\pi^+)K^+$, the parameters listed in Table F.1 are used. These values are taken from [49].

LASS Parameter	***Value***
m_0	$(1.415 \pm 0.003)\,\mathrm{GeV}/c^2$
Γ_0	$(0.300 \pm 0.006)\,\mathrm{GeV}/c^2$
a	$(2.07 \pm 0.10)(\,\mathrm{GeV}/c)^{-1}$
r	$(3.32 \pm 0.34)(\,\mathrm{GeV}/c)^{-1}$
B	1.0
ϕ_B	0.0
R	1.0
ϕ_R	0.0

Table F.1: *37 pts LASS parameters [49]. See Eq.s (1.97) and (1.100) and Section 1.4.3 for a definition of these variables.*

G

Tests and validations for the 3-var. fits

The toy tests, fit model validations and studies of the effects of $B\overline{B}$ background on the fit model, which are described for the 5-variable fit in Section 5.10, are repeated for the 3-variable fits. The findings are summarised in Section 5.11. For completeness, the tables and figures not shown in Section 5.11 are presented here.

Tables G.1 and G.2 list the results of pure toy tests (of the type described in Section 5.10.2) conducted for the 892 and 1430 windows respectively.

Figures G.1 and G.2 illustrate the $B\overline{B}$ background PDFs used for the 892 and 1430 windows respectively.

Tables G.3 and G.4 show the results of toy tests in which signal and $B\overline{B}$ background events are embedded real[1] MC events (that have passed RaSC), and in which the $B\overline{B}$ background yields are fixed.

Figure G.3 shows the plots obtained from the likelihood value comparison tests (see final paragraph, Section 5.11). The absence of events with a likelihood ratio above ~ 0.6 (top right and bottom right plots, Figure G.3) can be explained as follows. Events with a high signal likelihood also have high $(B\overline{B})_i$ likelihoods ($i = 1, 2, 3, 4$), since their PDFs in m_{ES} and $\mathcal{F}$ are very similar, with only ΔE providing any significant discrimination. A large amount of discrimination between the signal and $B\overline{B}$ background components is lost when the variables $m_{K\pi}$ and $\cos\theta_H$ are removed from the fit model. Binning events by $\mathcal{P}_S/(\mathcal{P}_S + \mathcal{P}_{q\overline{q}})$, bins are filled up to $\mathcal{P}_S/(\mathcal{P}_S + \mathcal{P}_{q\overline{q}}) = 1.0$. However, it is highly probable – due to the aforementioned lack of discrimination – that events with large $\mathcal{P}_S$ also have large $\mathcal{P}_{(B\overline{B})_i}$ ($i = 1, 2, 3, 4$). For a given event therefore, it is most likely that $\mathcal{P}_S/(\mathcal{P}_S + \mathcal{P}_{q\overline{q}} + \mathcal{P}_{(B\overline{B})_1} + \mathcal{P}_{(B\overline{B})_2} + \mathcal{P}_{(B\overline{B})_3}) < \mathcal{P}_S/(\mathcal{P}_S + \mathcal{P}_{q\overline{q}})$.

[1]By "real" MC, we mean MC that is generated as described in Section 4.3, rather than toy MC as discussed in Section 5.10.1.

Number Of Signal And $q\overline{q}$ Background Events Generated (Poissonially Smeared)	***Average Yield Extracted Over 1,000 † Experiments (And Average Error)***		***Bias β; Mean, Width Of Gaussian Fitted To Pull Distribution***	
	Signal	***$q\overline{q}$ Bkgd.***	***Signal***	***$q\overline{q}$ Bkgd.***
200 signal, 39,956 $q\overline{q}$ bkgd.	200.34 ± 0.78 (25.18)	$39,959 \pm 6$ (201)	0.33 ± 0.80; -0.023 ± 0.031, 0.982 ± 0.022	2.68 ± 6.32; 0.011 ± 0.031, 0.995 ± 0.022
150 signal, 40,006 $q\overline{q}$ bkgd.	151.30 ± 0.70 (23.08)	$40,008 \pm 6$ (201)	1.30 ± 0.70; 0.015 ± 0.031, 0.967 ± 0.022	1.69 ± 6.34; 0.006 ± 0.032, 0.998 ± 0.022
100 signal, 40,056 $q\overline{q}$ bkgd.	100.89 ± 0.63 (20.52)	$40,058 \pm 6$ (201)	0.90 ± 0.63; -0.009 ± 0.031, 0.982 ± 0.022	2.07 ± 6.34; 0.008 ± 0.032, 0.998 ± 0.022
80 signal, 40,076 $q\overline{q}$ bkgd.	79.28 ± 0.62 (19.21)	$40,077 \pm 6$ (201)	-0.72 ± 0.62; -0.107 ± 0.033, 1.046 ± 0.023	1.23 ± 6.15; 0.004 ± 0.031, 0.969 ± 0.022
60 signal, 40,096 $q\overline{q}$ bkgd.	60.18 ± 0.56 (17.84)	$40,104 \pm 6$ (201)	0.19 ± 0.56; -0.071 ± 0.033, 1.035 ± 0.023	8.05 ± 6.29; 0.038 ± 0.031, 0.990 ± 0.022
50 signal, 40,106 $q\overline{q}$ bkgd.	50.39 ± 0.55 (17.06)	$40,112 \pm 6$ (201)	0.39 ± 0.55; -0.067 ± 0.033, 1.047 ± 0.023	5.61 ± 6.36; 0.026 ± 0.032, 1.001 ± 0.022
40 signal, 40,116 $q\overline{q}$ bkgd.	39.97 ± 0.53 (16.10)	$40,114 \pm 6$ (201)	-0.04 ± 0.53; -0.092 ± 0.034, 1.069 ± 0.024	-2.29 ± 6.32; -0.014 ± 0.031, 0.995 ± 0.022
30 signal, 40,126 $q\overline{q}$ bkgd.	30.54 ± 0.49 (15.14)	$40,136 \pm 6$ (201)	0.54 ± 0.49; -0.074 ± 0.034, 1.069 ± 0.024	9.59 ± 6.47; 0.045 ± 0.032, 1.018 ± 0.023
20 signal, 40,136 $q\overline{q}$ bkgd.	19.29 ± 0.45 (13.93)	$40,137 \pm 6$ (201)	-0.72 ± 0.45; -0.168 ± 0.034, 1.079 ± 0.024	0.66 ± 6.39; 0.001 ± 0.032, 1.006 ± 0.022
10 signal, 40,146 $q\overline{q}$ bkgd.	9.98 ± 0.45 (12.81)	$40,144 \pm 6$ (201)	-0.02 ± 0.45; -0.111 ± 0.036, 1.123 ± 0.025	-1.99 ± 6.45; -0.012 ± 0.032, 1.016 ± 0.023
0 signal (not Poissonially smeared), 40,156 $q\overline{q}$ bkgd. (300 experiments, not 1,000)	-1.82 ± 0.82 (11.20)	$40,147 \pm 10$ (201)	-1.82 ± 0.82; -0.262 ± 0.072, 1.239 ± 0.051	-8.78 ± 10.73; -0.046 ± 0.053, 0.924 ± 0.038

†: Unless stated otherwise.

Table G.1: *Pure toy tests for the 3-variable fit in the 892 window.*

Number Of Signal And $q\overline{q}$ Background Events Generated (Poissonially Smeared)	***Average Yield Extracted Over 1,000 Experiments (And Average Error)***		***Bias β; Mean, Width Of Gaussian Fitted To Pull Distribution***	
	Signal	***$q\overline{q}$ Bkgd.***	***Signal***	***$q\overline{q}$ Bkgd.***
200 signal, 46,352 $q\overline{q}$ bkgd.	200.59 ± 1.05 (30.46)	$46,348 \pm 7$ (217)	0.57 ± 1.05; -0.013 ± 0.032, 0.998 ± 0.022	-4.28 ± 6.62; -0.022 ± 0.030, 0.964 ± 0.022
150 signal, 46,402 $q\overline{q}$ bkgd.	150.13 ± 0.93 (28.35)	$46,398 \pm 7$ (217)	0.13 ± 0.93; -0.030 ± 0.031, 0.991 ± 0.022	-3.84 ± 6.61; -0.020 ± 0.030, 0.962 ± 0.022
100 signal, 46,452 $q\overline{q}$ bkgd.	99.24 ± 0.85 (25.80)	$46,449 \pm 7$ (217)	-0.76 ± 0.85; -0.084 ± 0.032, 1.024 ± 0.023	-2.72 ± 6.59; -0.015 ± 0.030, 0.960 ± 0.021
80 signal, 46,472 $q\overline{q}$ bkgd.	80.04 ± 0.82 (24.66)	$46,461 \pm 7$ (217)	0.04 ± 0.82; -0.052 ± 0.034, 1.061 ± 0.024	-10.53 ± 6.89; -0.051 ± 0.032, 1.005 ± 0.022
60 signal, 46,492 $q\overline{q}$ bkgd.	60.61 ± 0.76 (23.35)	$46,493 \pm 7$ (217)	0.60 ± 0.76; -0.057 ± 0.033, 1.041 ± 0.023	0.62 ± 6.91; 0.001 ± 0.032, 1.007 ± 0.023
40 signal, 46,512 $q\overline{q}$ bkgd.	40.19 ± 0.73 (21.78)	$46,496 \pm 7$ (217)	0.18 ± 0.73; -0.078 ± 0.034, 1.088 ± 0.024	-16.01 ± 7.07; -0.076 ± 0.033, 1.032 ± 0.023
30 signal, 46,522 $q\overline{q}$ bkgd.	29.11 ± 0.68 (20.80)	$46,514 \pm 7$ (217)	-0.89 ± 0.68; -0.143 ± 0.034, 1.064 ± 0.024	-7.65 ± 7.04; -0.038 ± 0.032, 1.028 ± 0.023
20 signal, 46,532 $q\overline{q}$ bkgd.	20.55 ± 0.65 (19.89)	$46,526 \pm 7$ (217)	0.54 ± 0.65; -0.078 ± 0.034, 1.073 ± 0.024	-5.97 ± 7.00; -0.030 ± 0.032, 1.022 ± 0.023
10 signal, 46,542 $q\overline{q}$ bkgd.	8.98 ± 0.61 (18.79)	$46,527 \pm 7$ (217)	-1.02 ± 0.61; -0.139 ± 0.033, 1.048 ± 0.023	-15.07 ± 7.09; -0.072 ± 0.033, 1.036 ± 0.023
5 signal, 46,547 $q\overline{q}$ bkgd.	2.79 ± 0.65 (18.48)	$46,542 \pm 7$ (217)	-2.20 ± 0.65; -0.228 ± 0.037, 1.177 ± 0.026	-4.68 ± 6.83; -0.024 ± 0.031, 0.996 ± 0.022
0 signal (not Poissonially smeared), 46,552 $q\overline{q}$ bkgd.	-1.98 ± 0.65 (17.78)	$46,558 \pm 7$ (217)	-1.98 ± 0.65; -0.207 ± 0.038, 1.185 ± 0.027	5.67 ± 6.75; 0.024 ± 0.031, 0.984 ± 0.022

Table G.2: *Pure toy tests for the 3-variable fit in the 1430 window.*

# $(B\overline{B})_1$ ***Embedded*** †	# $(B\overline{B})_2$ ***Embedded*** †	# $(B\overline{B})_3$ ***Embedded*** †	β ***(Signal)***	β ***($q\overline{q}$ Bkgd.)***
$n^1_{B\overline{B}\ Bkgd}$	—	—	-0.00 ± 0.49	21.47 ± 0.49
$n^1_{B\overline{B}\ Bkgd} - \Delta n^1_{B\overline{B}\ Bkgd}$	—	—	0.78 ± 0.51	18.79 ± 0.51
$n^1_{B\overline{B}\ Bkgd} + \Delta n^1_{B\overline{B}\ Bkgd}$	—	—	1.08 ± 0.51	22.04 ± 0.51
—	$n^2_{B\overline{B}\ Bkgd}$ ♣	—	0.05 ± 0.51	20.85 ± 0.51
—	$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♣	—	-0.01 ± 0.49	17.99 ± 0.49
—	$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♣	—	0.82 ± 0.48	22.90 ± 0.48
—	$n^2_{B\overline{B}\ Bkgd}$ ♢	—	-0.64 ± 0.50	17.77 ± 0.50
—	$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♢	—	-1.10 ± 0.48	15.30 ± 0.48
—	$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♢	—	-0.04 ± 0.49	20.08 ± 0.49
—	$n^2_{B\overline{B}\ Bkgd}$ ♠	—	1.93 ± 0.51	25.28 ± 0.51
—	$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♠	—	1.17 ± 0.48	22.21 ± 0.48
—	$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♠	—	2.65 ± 0.50	28.31 ± 0.50
—	$n^2_{B\overline{B}\ Bkgd}$ ♡	—	1.85 ± 0.51	25.29 ± 0.51
—	$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♡	—	1.43 ± 0.49	21.93 ± 0.49
—	$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♡	—	3.07 ± 0.49	27.82 ± 0.49
—	—	$n^3_{B\overline{B}\ Bkgd}$	1.17 ± 0.49	19.16 ± 0.49
—	—	$n^3_{B\overline{B}\ Bkgd} - \Delta n^3_{B\overline{B}\ Bkgd}$	0.81 ± 0.48	18.56 ± 0.48
—	—	$n^3_{B\overline{B}\ Bkgd} + \Delta n^3_{B\overline{B}\ Bkgd}$	1.06 ± 0.50	21.22 ± 0.50
$n^1_{B\overline{B}\ Bkgd}$	$n^2_{B\overline{B}\ Bkgd}$ ♣	$n^3_{B\overline{B}\ Bkgd}$	-0.05 ± 0.53	20.25 ± 0.54
$n^1_{B\overline{B}\ Bkgd} - \Delta n^1_{B\overline{B}\ Bkgd}$	$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♢	$n^3_{B\overline{B}\ Bkgd} - \Delta n^3_{B\overline{B}\ Bkgd}$	-1.53 ± 0.50	12.63 ± 0.51
$n^1_{B\overline{B}\ Bkgd} + \Delta n^1_{B\overline{B}\ Bkgd}$	$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♠	$n^3_{B\overline{B}\ Bkgd} + \Delta n^3_{B\overline{B}\ Bkgd}$	2.59 ± 0.52	30.51 ± 0.52
$n^1_{B\overline{B}\ Bkgd}$	$n^2_{B\overline{B}\ Bkgd}$ ♣ $+ n^{Other}_{B\overline{B}\ Bkgd}$ $(B\overline{B})_{Other}$ events	$n^3_{B\overline{B}\ Bkgd}$	4.83 ± 0.51	146.86 ± 0.51
$n^1_{B\overline{B}\ Bkgd}$	$n^2_{B\overline{B}\ Bkgd}$ ♣ $+ n^{OtherMax}_{B\overline{B}\ Bkgd}$ $(B\overline{B})_{Other}$ events	$n^3_{B\overline{B}\ Bkgd}$	7.14 ± 0.52	209.33 ± 0.52
Extreme cases:				
$n^1_{B\overline{B}\ Bkgd}/2$	—	—	-0.88 ± 0.48	12.29 ± 0.48
$2n^1_{B\overline{B}\ Bkgd}$	—	—	4.40 ± 0.51	36.02 ± 0.51
—	$n^2_{B\overline{B}\ Bkgd}/2$ ♣	—	-1.17 ± 0.49	10.54 ± 0.49
—	$2n^2_{B\overline{B}\ Bkgd}$ ♣	—	3.22 ± 0.50	39.80 ± 0.51
—	—	$n^3_{B\overline{B}\ Bkgd}/2$	0.20 ± 0.49	18.16 ± 0.49
—	—	$2n^3_{B\overline{B}\ Bkgd}$	2.57 ± 0.49	22.71 ± 0.49
$n^1_{B\overline{B}\ Bkgd}/2$	$n^2_{B\overline{B}\ Bkgd}/2$ ♣	$n^3_{B\overline{B}\ Bkgd}/2$	-3.93 ± 0.50	1.39 ± 0.50
$2n^1_{B\overline{B}\ Bkgd}$	$2n^2_{B\overline{B}\ Bkgd}$ ♣	$2n^3_{B\overline{B}\ Bkgd}$	7.44 ± 0.53	57.34 ± 0.55

†: rounded to nearest integer

$n^1_{B\overline{B}\ Bkgd} = 20.6 \pm 0.4 \pm 1.5$, $n^3_{B\overline{B}\ Bkgd} = 4.8 \pm 0.2 \pm 0.8$

♣: Belle model B_0, solution 1 (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 23.0 \pm 1.1 \pm 2.2$; ♢: Belle model B_0, solution 2 (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 19.3 \pm 1.4 \pm 1.9$; ♠: *BABAR* solution A (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 30.1 \pm 1.7 \pm 2.9$; ♡: *BABAR* solution B (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 29.9 \pm 1.7 \pm 2.9$

Table G.3: *Embedded toy tests for the 3-variable fit in the 892 window with fixed-yield $B\overline{B}$ background components. Where events are embedded for component $(B\overline{B})_i$ $(i = 1, 2, 3)$, the yield for this component is fixed at the central value of $n^i_{B\overline{B}\ Bkgd}$ $(i = 1, 2, 3)$.*

# $(B\overline{B})_2$ ***Embedded*** †	# $(B\overline{B})_3$ ***Embedded*** †	# $(B\overline{B})_4$ ***Embedded*** †	β ***(Signal)***	β ***($q\overline{q}$ Bkgd.)***
$n^2_{B\overline{B}\ Bkgd}$ ♣	—	—	0.07 ± 0.78	-0.80 ± 0.78
$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♣	—	—	-1.55 ± 0.79	-7.45 ± 0.78
$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♣	—	—	0.60 ± 0.77	7.92 ± 0.76
$n^2_{B\overline{B}\ Bkgd}$ ♢	—	—	0.62 ± 0.76	-1.02 ± 0.75
$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♢	—	—	-1.34 ± 0.77	-7.27 ± 0.77
$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♢	—	—	1.74 ± 0.79	6.95 ± 0.79
$n^2_{B\overline{B}\ Bkgd}$ ♠	—	—	0.49 ± 0.79	-5.19 ± 0.77
$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♠	—	—	-1.05 ± 0.78	-11.89 ± 0.78
$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♠	—	—	2.56 ± 0.79	1.11 ± 0.78
$n^2_{B\overline{B}\ Bkgd}$ ♡	—	—	0.05 ± 0.77	-5.60 ± 0.77
$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♡	—	—	-0.57 ± 0.74	-13.15 ± 0.74
$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♡	—	—	1.09 ± 0.80	2.60 ± 0.79
—	$n^3_{B\overline{B}\ Bkgd}$	—	3.38 ± 0.75	-1.90 ± 0.75
—	$n^3_{B\overline{B}\ Bkgd} - \Delta n^3_{B\overline{B}\ Bkgd}$	—	2.93 ± 0.76	-3.53 ± 0.76
—	$n^3_{B\overline{B}\ Bkgd} + \Delta n^3_{B\overline{B}\ Bkgd}$	—	3.51 ± 0.76	-1.07 ± 0.76
—	—	$n^4_{B\overline{B}\ Bkgd}$	2.84 ± 0.78	-1.87 ± 0.77
—	—	$n^4_{B\overline{B}\ Bkgd} - \Delta n^4_{B\overline{B}\ Bkgd}$	2.55 ± 0.73	-3.54 ± 0.72
—	—	$n^4_{B\overline{B}\ Bkgd} + \Delta n^4_{B\overline{B}\ Bkgd}$	3.41 ± 0.76	-0.51 ± 0.75
$n^2_{B\overline{B}\ Bkgd}$ ♣	$n^3_{B\overline{B}\ Bkgd}$	$n^4_{B\overline{B}\ Bkgd}$	-0.64 ± 0.79	0.56 ± 0.78
$n^2_{B\overline{B}\ Bkgd} - \Delta n^2_{B\overline{B}\ Bkgd}$ ♢	$n^3_{B\overline{B}\ Bkgd} - \Delta n^3_{B\overline{B}\ Bkgd}$	$n^4_{B\overline{B}\ Bkgd} - \Delta n^4_{B\overline{B}\ Bkgd}$	-1.49 ± 0.76	-15.69 ± 0.75
$n^2_{B\overline{B}\ Bkgd} + \Delta n^2_{B\overline{B}\ Bkgd}$ ♠	$n^3_{B\overline{B}\ Bkgd} + \Delta n^3_{B\overline{B}\ Bkgd}$	$n^4_{B\overline{B}\ Bkgd} + \Delta n^4_{B\overline{B}\ Bkgd}$	2.15 ± 0.77	10.22 ± 0.79
$n^2_{B\overline{B}\ Bkgd}$ ♣	$n^3_{B\overline{B}\ Bkgd}$ + $n^{Other}_{B\overline{B}\ Bkgd}$ $(B\overline{B})_{Other}$ events	$n^4_{B\overline{B}\ Bkgd}$	6.35 ± 0.71	712.63 ± 0.71
$n^2_{B\overline{B}\ Bkgd}$ ♣	$n^3_{B\overline{B}\ Bkgd}$ + $n^{OtherMax}_{B\overline{B}\ Bkgd}$ $(B\overline{B})_{Other}$ events	$n^4_{B\overline{B}\ Bkgd}$	10.40 ± 0.70	$1,080.80 \pm 0.70$
Extreme cases:				
$n^2_{B\overline{B}\ Bkgd}/2$ ♣	—	—	-5.55 ± 0.74	-35.61 ± 0.73
$2n^2_{B\overline{B}\ Bkgd}$ ♣	—	—	11.91 ± 0.82	70.36 ± 0.81
—	$n^3_{B\overline{B}\ Bkgd}/2$	—	2.93 ± 0.76	-3.53 ± 0.76
—	$2n^3_{B\overline{B}\ Bkgd}$	—	3.56 ± 0.77	0.92 ± 0.76
—	—	$n^4_{B\overline{B}\ Bkgd}/2$	1.41 ± 0.75	-6.32 ± 0.74
—	—	$2n^4_{B\overline{B}\ Bkgd}$	6.05 ± 0.78	6.77 ± 0.77
$n^2_{B\overline{B}\ Bkgd}/2$ ♣	$n^3_{B\overline{B}\ Bkgd}/2$	$n^4_{B\overline{B}\ Bkgd}/2$	-6.37 ± 0.75	-42.02 ± 0.74
$2n^2_{B\overline{B}\ Bkgd}$ ♣	$2n^3_{B\overline{B}\ Bkgd}$	$2n^4_{B\overline{B}\ Bkgd}$	16.47 ± 0.90	81.18 ± 0.86

†: rounded to nearest integer

♣: Belle model B_0, solution 1 (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 89.4 \pm 2.2 \pm 8.6$; ♢: Belle model B_0, solution 2 (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 90.4 \pm 3.0 \pm 8.7$; ♠: *BABAR* solution A (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 85.1 \pm 2.9 \pm 8.2$; ♡: *BABAR* solution B (minus ϕK) — $n^2_{B\overline{B}\ Bkgd} = 84.5 \pm 2.9 \pm 8.1$;

$n^3_{B\overline{B}\ Bkgd} = 3.5 \pm 0.1 \pm 0.6$, $n^4_{B\overline{B}\ Bkgd} = 12.0 \pm 0.4 \pm 1.4$

Table G.4: *Embedded toy tests for the 3-variable fit in the 1430 window with fixed-yield $B\overline{B}$ background components. Where events are embedded for component $(B\overline{B})_i$ ($i = 2, 3, 4$), the yield for this component is fixed at the central value of $n^i_{B\overline{B}\ Bkgd}$ ($i = 2, 3, 4$).*

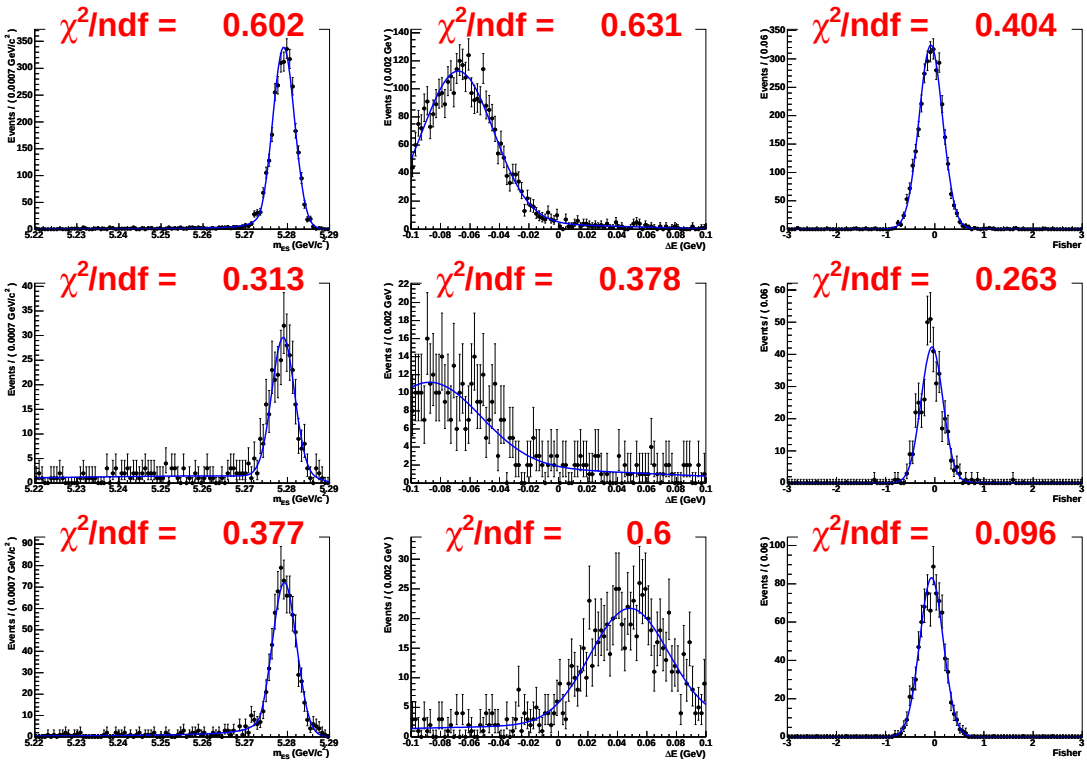

Figure G.1: *Fit variable distributions and fitted PDFs for the 3-variable fit in the 892 window. Left column: m_{ES}; middle column: ΔE; right column: $\mathcal{F}$. Top row: $(B\overline{B})_1$; second row: $(B\overline{B})_2$; bottom row: $(B\overline{B})_3$. The functional forms of the PDFs are listed in Table 5.21. PDF parameters are given in Table D.9.*

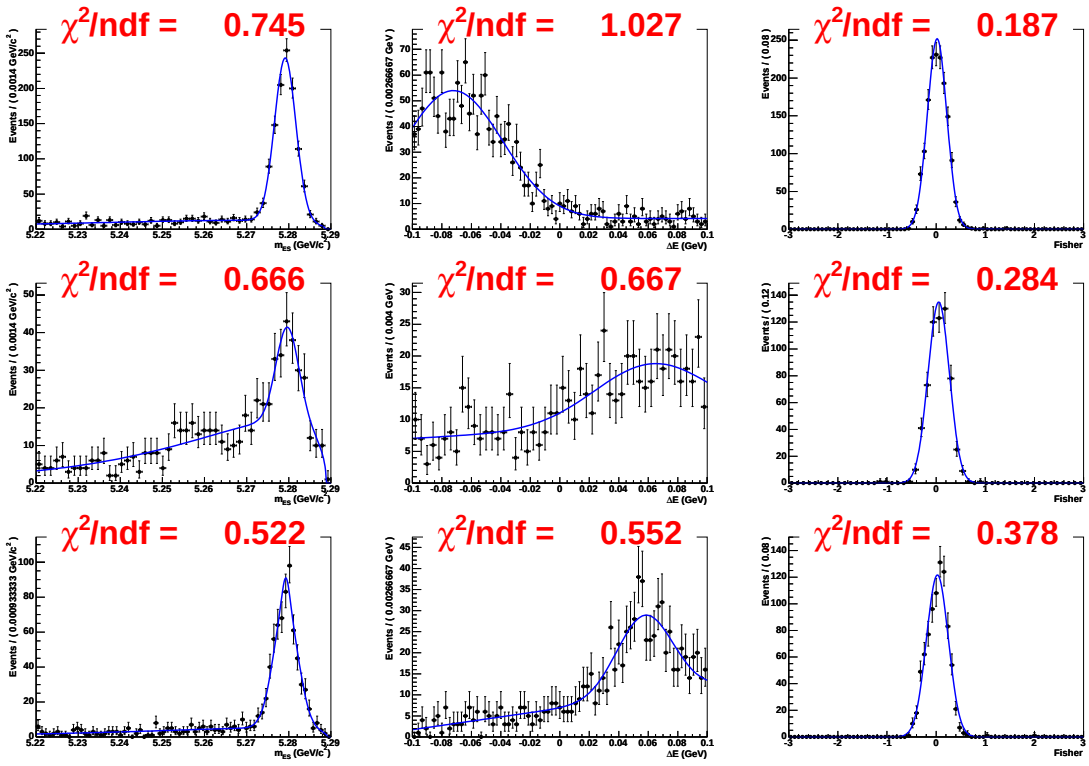

Figure G.2: *Fit variable distributions and fitted PDFs for the 3-variable fit in the 1430 window. Left column: m_{ES}; middle column: ΔE; right column: $\mathcal{F}$. Top row: $(B\overline{B})_2$; second row: $(B\overline{B})_3$; bottom row: $(B\overline{B})_4$. The functional forms of the PDFs are listed in Table 5.29. PDF parameters are given in Table D.12.*

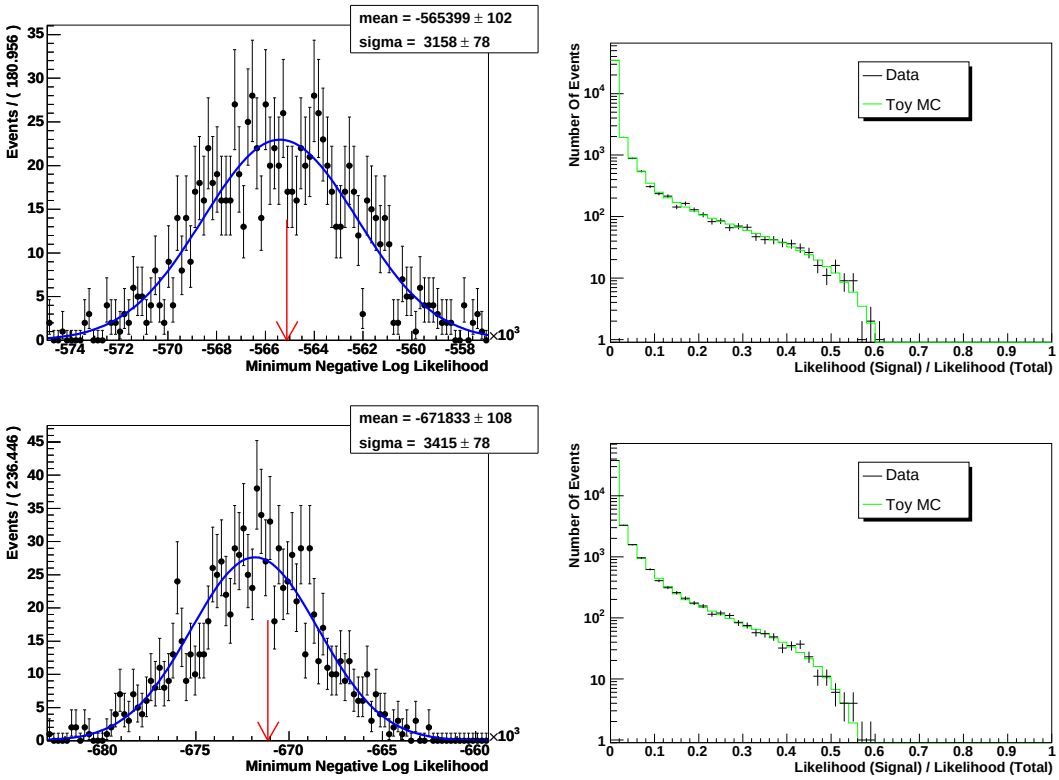

Figure G.3: *Top row: 3-variable fit in the 892 window; bottom row: 3-variable fit in the 1430 window. Left: the distribution of* $(-\ln \mathcal{L}_{Max})$ *values for 1,000 pure toy experiments, with a superimposed fitted Gaussian. The red arrow indicates the value of* $(-\ln \mathcal{L}_{Max})$ *obtained when applying the fit model to the final on-resonance sample; right: event likelihood ratio distributions for toy MC and the final on-resonance sample.*

H

Floated parameter correlations

Tables H.1-H.3 present the correlation values between the floated PDF parameters and yields for the extended ML fits to the final on-resonance samples.

	$n_{q\bar{q}}$	n_S	$c^{(q\bar{q})}_{1(\Delta E)}$	$f^{(q\bar{q})}_{2(\mathcal{F})}$	$\mu^{(q\bar{q})}_{1(\mathcal{F})}$	$\mu^{(q\bar{q})}_{2(\mathcal{F})}$	$\sigma^{(q\bar{q})}_{1(\mathcal{F})}$	$\sigma^{(q\bar{q})}_{2(\mathcal{F})}$	$c^{(q\bar{q})}_{1(\cos\theta_H)}$	$c^{(q\bar{q})}_{2(\cos\theta_H)}$	$f^{(q\bar{q})}_{BW(m_{K\pi})}$	$c^{(q\bar{q})}_{1(m_{K\pi})}$	$\mu^{(q\bar{q})}_{(m_{K\pi})}$	$\Gamma^{(q\bar{q})}_{(m_{K\pi})}$
n_S	−0.087													
$c^{(q\bar{q})}_{1(\Delta E)}$	−0.001	−0.000												
$f^{(q\bar{q})}_{2(\mathcal{F})}$	−0.042	−0.080	−0.000											
$\mu^{(q\bar{q})}_{1(\mathcal{F})}$	0.048	0.123	−0.001	−0.291										
$\mu^{(q\bar{q})}_{2(\mathcal{F})}$	−0.001	0.045	0.001	0.071	−0.260									
$\sigma^{(q\bar{q})}_{1(\mathcal{F})}$	0.034	0.060	0.000	−0.855	0.227	−0.048								
$\sigma^{(q\bar{q})}_{2(\mathcal{F})}$	0.057	0.015	−0.002	−0.686	0.191	−0.058	0.556							
$c^{(q\bar{q})}_{1(\cos\theta_H)}$	−0.012	0.012	−0.002	0.012	−0.021	0.004	−0.006	−0.016						
$c^{(q\bar{q})}_{2(\cos\theta_H)}$	0.004	−0.053	−0.003	0.011	−0.009	−0.011	−0.017	−0.008	0.195					
$f^{(q\bar{q})}_{BW(m_{K\pi})}$	0.020	−0.055	0.000	−0.011	0.004	−0.013	−0.002	0.010	0.003	0.008				
$c^{(q\bar{q})}_{1(m_{K\pi})}$	0.020	−0.034	−0.000	−0.103	0.095	0.018	0.093	0.073	−0.026	−0.023	−0.026			
$\mu^{(q\bar{q})}_{(m_{K\pi})}$	−0.005	0.017	−0.001	0.021	−0.035	0.000	−0.018	−0.025	0.006	0.001	0.103	−0.231		
$\Gamma^{(q\bar{q})}_{(m_{K\pi})}$	0.009	−0.038	0.000	−0.019	0.015	−0.007	0.008	0.019	0.002	−0.000	0.759	−0.028	0.127	
$\xi^{(q\bar{q})}_{(m_{ES})}$	−0.011	0.078	0.003	−0.001	0.004	0.006	0.003	−0.002	0.026	0.012	−0.004	−0.007	0.000	0.012

Table H.1: *Correlation matrix for fitted PDF parameters and yields for the 5-variable extended ML fit to the final on-resonance sample in the 892 window.*

n_S	-0.076							
$c_{1(\Delta E)}^{(q\bar{q})}$	-0.001	-0.001						
$f_{2(\mathcal{F})}^{(q\bar{q})}$	-0.004	0.037	0.001					
$\mu_{1(\mathcal{F})}^{(q\bar{q})}$	-0.009	0.084	0.001	-0.064				
$\mu_{2(\mathcal{F})}^{(q\bar{q})}$	-0.006	0.061	0.001	0.403	-0.229			
$\sigma_{1(\mathcal{F})}^{(q\bar{q})}$	0.005	-0.044	-0.000	0.720	-0.063	0.303		
$\sigma_{2(\mathcal{F})}^{(q\bar{q})}$	-0.001	0.015	0.001	0.874	-0.042	0.332	0.600	
$\xi_{(m_{ES})}^{(q\bar{q})}$	-0.011	0.109	0.001	0.005	0.012	0.008	-0.007	0.002
	$n_{q\bar{q}}$	n_S	$c_{1(\Delta E)}^{(q\bar{q})}$	$f_{2(\mathcal{F})}^{(q\bar{q})}$	$\mu_{1(\mathcal{F})}^{(q\bar{q})}$	$\mu_{2(\mathcal{F})}^{(q\bar{q})}$	$\sigma_{1(\mathcal{F})}^{(q\bar{q})}$	$\sigma_{2(\mathcal{F})}^{(q\bar{q})}$

Table H.2: *Correlation matrix for fitted PDF parameters and yields for the 3-variable extended ML fit to the final on-resonance sample in the 892 window.*

n_S	-0.093							
$c_{1(\Delta E)}^{(q\bar{q})}$	-0.002	0.003						
$f_{2(\mathcal{F})}^{(q\bar{q})}$	-0.023	0.126	0.000					
$\mu_{1(\mathcal{F})}^{(q\bar{q})}$	-0.022	0.160	0.002	0.760				
$\mu_{2(\mathcal{F})}^{(q\bar{q})}$	-0.016	0.088	0.000	0.712	0.250			
$\sigma_{1(\mathcal{F})}^{(q\bar{q})}$	0.021	-0.157	-0.002	-0.726	-0.397	-0.750		
$\sigma_{2(\mathcal{F})}^{(q\bar{q})}$	0.024	-0.110	0.000	-0.857	-0.621	-0.676	0.558	
$\xi_{(m_{ES})}^{(q\bar{q})}$	-0.011	0.129	0.002	0.029	0.029	0.020	-0.025	-0.030
	$n_{q\bar{q}}$	n_S	$c_{1(\Delta E)}^{(q\bar{q})}$	$f_{2(\mathcal{F})}^{(q\bar{q})}$	$\mu_{1(\mathcal{F})}^{(q\bar{q})}$	$\mu_{2(\mathcal{F})}^{(q\bar{q})}$	$\sigma_{1(\mathcal{F})}^{(q\bar{q})}$	$\sigma_{2(\mathcal{F})}^{(q\bar{q})}$

Table H.3: *Correlation matrix for fitted PDF parameters and yields for the 3-variable extended ML fit to the final on-resonance sample in the 1430 window.*

Glossary

AF	Argus function
AWG	Analysis working group
BLT	Binary Link Tracker
BR	Branching ratio
BW	Breit-Wigner
CB	Crystal Ball
CKM	Cabibbo Kobayashi Maskawa
CL	Confidence level
CM	Centre of mass
CP	Charge Parity
CPT	Charge Parity Time
DAQ	Data acquisition system
DCH	Drift Chamber
DCT	Drift Chamber Trigger
DG	Double Gaussian
DRC	Detector of Internally Reflected Cherenkov Radiation
EMC	Electromagnetic Calorimeter
EMT	Electromagnetic Calorimeter Trigger
ER	Event Reconstruction
FCT	Fast Control and Timing System

FEC	Front End readout Card
FEE	Front End Electronics
FEX	Feature extraction
GLT	Global Trigger
HER	High Energy Ring
IFR	Instrumented Flux Return
IFT	Instrumented Flux Return Trigger
IP	Interaction point
JAS	Java Analysis Studio
L1A	Level-1 Accept
L1T	Level-1 Trigger
L3T	Level-3 Trigger
LER	Low Energy Ring
LHC	Large Hadron Collider
LST	Limited Streamer Tube
MC	Monte Carlo
MIP	Minimum ionising particle
ML	Maximum likelihood
NLL	Negative log likelihood
NR	Non-resonant
ODC	Online Detector Control
ODF	Online Dataflow
OEP	Online Event Processing
ORC	Online Run Control
PC	Prompt Calibration
PDF	Probability Density Function
PDG	Particle Data Group
PEP-II	Positron Electron Project II
PID	Particle identification

PIN	Positive Intrinsic Negative (semiconductor)
PMT	Photo-Multiplier Tube
POCA	Point of closest approach
PPARC	Particle Physics and Astronomy Research Council
PR	Prompt Reconstruction
PTD	p_T Discriminator
QCD	Quantum chromodynamics
RaSC	Reconstruction and the selection criteria summarised in Table 5.3
ROE	Rest of the event
ROM	Read Out Module
RPC	Resistive Plate Chamber
SLAC	Stanford Linear Accelerator Center
SM	Standard Model of particle physics
SVT	Silicon Vertex Tracker
SXF	Self-cross-feed
TDSE	Time dependent Schrödinger equation
TOT	Time-over-threshold
TPB	Trigger Processor Board
TPC	Triggered Personality Card
TRG	Trigger
TSF	Track Segment Finder
UL	Upper limit
UPC	Untriggered Personality Card
XTC	Extended tagged container (binary files)
ZPD	z_0 p_T Discriminator

References

[1] B. Aubert et al. Search for the decay of a B^0 or $\overline{B}^0$ meson to $\overline{K}^{*0}K^0$ or $K^{*0}\overline{K}^0$. *Phys. Rev.*, D74:072008, 2006.

[2] G. Burdman. New solutions to the hierarchy problem. 2007. hep-ph/0703194.

[3] A. D. Sakharov. Violation of *CP* invariance, *C* asymmetry, and baryon asymmetry of the Universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, 1967.

[4] http://www-public.slac.stanford.edu/babar/.

[5] http://www.slac.stanford.edu/.

[6] J. P. Burke. Search for $B^+ \to \overline{K}^{*0}K^+$, $\overline{K}^{*0} \to K^-\pi^+$. *BABAR* Analysis Document #986.

[7] C. P. Jessop et al. Study of charmless hadronic B meson decays to pseudoscalar vector final states. *Phys. Rev. Lett.*, 85:2881–2885, 2000.

[8] R. Fleischer and S. Recksiegel. $b \to d$ penguins: *CP* violation, general lower bounds on the branching ratios and standard model tests. *PoS*, HEP2005:255, 2006.

[9] R. Fleischer and S. Recksiegel. Model-independent lower bounds for $b \to d$ penguin processes. *Phys. Rev.*, D71:051501, 2005.

[10] Y. Grossman et al. SU(3) relations and the *CP* asymmetries in B decays to $\eta' K^0_S$, ϕK^0_S and $K^+K^-K^0_S$. *Phys. Rev.*, D68:015004, 2003.

[11] T. D. Lee and C. N. Yang. Question of parity conservation in weak interactions. *Phys. Rev.*, 104:254–258, 1956.

[12] C. S. Wu et al. Experimental test of parity conservation in β decay. *Phys. Rev.*, 105:1413–1414, 1957.

[13] B. L. Ioffe et al. *Sov. Phys. JETP*, 5:327, 1957.

[14] J. H. Christenson et al. Evidence for the 2π decay of the K^0_2 meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[15] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–532, 1963.

[16] K. Hagiwara et al. Review of particle physics. *Phys. Rev.*, D66:010001, 2002.

[17] M. Kobayashi and T. Maskawa. *CP* violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[18] M. Neubert. B physics and *CP* violation. *Int. J. Mod. Phys.*, A11:4173–4240, 1996.

[19] L. M. Lederman. The discovery of the upsilon, bottom quark, and B mesons. Given at 3rd International Symposium on the History of Particle Physics: The Rise of the Standard Model, Stanford, CA, 24-27 Jun 1992.

[20] M. L. Perl. Review of heavy lepton production in e^+e^- annihilation. Invited talk presented at International Symposium on Lepton and Photon Interactions at High Energies, Hamburg, Aug 25-31, 1977.

[21] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa matrix. *Phys. Rev. Lett.*, 51:1945, 1983.

[22] C. Jarlskog. Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal *CP* violation. *Phys. Rev. Lett.*, 55:1039, 1985.

[23] S. Eidelman et al. Review of particle physics. *Phys. Lett.*, B592:1, 2004.

[24] H. J. Lipkin. EPR in B physics and elsewhere. Presented at 20th Anniversary Symposium: Twenty Beautiful Years of Bottom Physics, Chicago, IL, 29 Jun - 2 Jul 1997.

[25] C. Voena. B_d mixing measurements with the *BABAR* detector. 2002. hep-ex/0211045.

[26] V. Fanti et al. A new measurement of direct *CP* violation in two pion decays of the neutral kaon. *Phys. Lett.*, B465:335–348, 1999.

[27] A. Alavi-Harati et al. Observation of direct *CP* violation in $K_{S,L} \to \pi\pi$ decays. *Phys. Rev. Lett.*, 83:22, 1999.

[28] B. Aubert et al. Observation of direct *CP* violation in $B^0 \to K^+\pi^-$ decays. *Phys. Rev. Lett.*, 93:131801, 2004.

[29] W. L. Lambert et al. Eq. (21), LHCb's potential to measure flavour-specific *CP*-asymmetry in semileptonic and hadronic B_s^0 decays. LHCb Note, http://www.slac.stanford.edu/∼phjpb/index/work/papers/asl.pdf.

[30] http://belle.kek.jp/.

[31] B. Aubert et al. Improved measurement of *CP* asymmetries in $B^0 \to c\overline{c}K^{(*)0}$ decays. 2006. hep-ex/0607107.

[32] K. Abe et al. Time-dependent *CP* asymmetries in $b \to s\overline{q}q$ transitions and $\sin(2\phi_1)$ in $B^0 \to J/\psi K^0$ decays with 386 million $B\overline{B}$ pairs. 2005. hep-ex/0507037.

[33] http://cern.ch/lhcb.

[34] P. F. Harrison and H. R. Quinn. The *BABAR* physics book: Physics at an asymmetric B factory. Papers from Workshop on Physics at an Asymmetric B Factory (*BABAR* Collaboration Meeting), Rome, Italy, 11-14 Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and Pasadena, CA, 22-24 Sep 1997.

[35] M. Beneke et al. Angles of the Unitarity Triangle: Measurements and perspectives. Invited talk at 2nd Workshop on the CKM Unitarity Triangle, Durham, England, 5-9 Apr 2003.

[36] A. Jawahery. CKM unitarity angles α (ϕ_2) and γ (ϕ_3). *Int. J. Mod. Phys.*, A19:975–990, 2004.

[37] J. Charles et al. *CP* violation and the CKM matrix: Assessing the impact of the asymmetric B factories. *Eur. Phys. J.*, C41:1–131, 2005.

[38] A. Bornheim et al. Improved measurement of $|V_{ub}|$ with inclusive semileptonic B decays. *Phys. Rev. Lett.*, 88:231803, 2002.

[39] Y. Grossman and M. P. Worah. *CP* asymmetries in B decays with new physics in decay amplitudes. *Phys. Lett.*, B395:241–249, 1997.

[40] Y. Grossman et al. *CP* asymmetry in $B_d \to \phi K_S^0$: Standard model pollution. *Phys. Rev.*, D58:057504, 1998.

[41] B. Aubert et al. Dalitz plot analysis of the decay $B^\pm \to K^\pm K^\pm K^\mp$. *Phys. Rev.*, D74:032003, 2006.

[42] A. Garmash et al. Dalitz analysis of the three-body charmless decays $B^+ \to K^+\pi^+\pi^-$ and $B^+ \to K^+K^+K^-$. *Phys. Rev.*, D71:092003, 2005.

[43] R. A. Briere et al. Observation of $B \to \phi K$ and $B \to \phi K^*$. *Phys. Rev. Lett.*, 86:3718–3721, 2001.

[44] D. Acosta et al. First evidence for $B_s^0 \to \phi\phi$ decay and measurements of branching ratio and A_{CP} for $B^+ \to \phi K^+$. *Phys. Rev. Lett.*, 95:031801, 2005.

[45] B. Aubert et al. Search for $B^+ \to \phi\pi^+$ and $B^0 \to \phi\pi^0$ decays. *Phys. Rev.*, D74:011102, 2006.

[46] W. M. Yao et al. Review of particle physics. *J. Phys.*, G33:1–1232, 2006.

[47] R. H. Dalitz. *Phil. Mag.*, 44, 1953.

[48] H. E. Haber. Spin formalism and applications to new physics searches. 1994.

[49] W. M. Dunwoodie. Fits to $K\pi$ I=$\frac{1}{2}$ S-wave amplitude and phase data. http://www.slac.stanford.edu/∼wmd/kpi_swave/kpi_swave_fit.note.

[50] B. Aubert et al. Dalitz-plot analysis of the decays $B^\pm \to K^\pm\pi^\mp\pi^\pm$. *Phys. Rev.*, D72:072003, 2005.

[51] Blanc F. et al. Exclusive charmless quasi-two-body B decays to $\eta K^*(892)$, $\eta K^*(\mathrm{S}-\mathrm{wave})$, and $\eta K_2^*(1430)$. *BABAR* Analysis Document #1341, Version 14, Jul 2006.

[52] D. Aston et al. A study of $K^-\pi^+$ scattering in the reaction $K^-p \to K^-\pi^+n$ at 11 GeV/c^2. *Nucl. Phys.*, B296:493, 1988.

[53] G. Buchalla et al. Weak decays beyond leading logarithms. *Rev. Mod. Phys.*, 68:1125–1144, 1996.

[54] M. Beneke and M. Neubert. QCD factorization for $B \to PP$ and $B \to PV$ decays. *Nucl. Phys.*, B675:333–415, 2003.

[55] B. Aubert et al. Measurement of the b^+/b^0 production ratio from the $\upsilon(4s)$ meson using $b^+ \to j/\psi k^+$ and $b^0 \to j/\psi k_s^0$ decays. *Phys. Rev.*, D69:071101, 2004.

[56] W. Kozanecki. The PEP-II B-factory: Status and prospects. *Nucl. Instrum. Meth.*, A446:59–64, 2000.

[57] J. Seeman et al. Results and plans of the PEP-II B-factory. Presented at 9th European Particle Accelerator Conference (EPAC 2004), Lucerne, Switzerland, 5-9 Jul 2004.

[58] W. Kozanecki et al. Trickle-charge: A new operational mode for PEP-II. Presented at the 9th European Particle Accelerator Conference (EPAC 2004), Lucerne, Switzerland, 5-9 Jul 2004.

[59] B. Aubert et al. The *BABAR* detector. *Nucl. Instrum. Meth.*, A479:1–116, 2002.

[60] A. Snyder. Effect of vertex cuts on *CP* reach. Sep 1994. SLAC-BABAR-NOTE-177.

[61] A. J. S. Smith et al. A barrel IFR instrumented with limited streamer tubes. http://www.slac.stanford.edu/BFROOT/www/Detector/LST/documentation /LSTprop-A4-May30-ajss.pdf, May 2003.

[62] R. Bartoldus et al. Trigger and filter documentation for run 1. *BABAR* Analysis Document #194, Apr 2002.

[63] F. A. Kirsten et al. *BABAR* level 1 drift chamber and global trigger implementation. http://www.slac.stanford.edu/BFROOT/www/Detector/Trigger/ Kirsten42.ps, Sep 1998.

[64] S. J. Bailey et al. Rapid 3d track reconstruction with the *BABAR* trigger upgrade. *IEEE Trans. Nucl. Sci.*, 51:2352–2355, 2004.

[65] J. C. Andress et al. The electromagnetic calorimeter level 1 trigger hardware design. *BABAR* Note #519, http://www.slac.stanford.edu/BFROOT/www/ Detector/Trigger/emt/babarnotes/BaBarNote519/BaBar519.ps.gz, May 2000.

[66] P. D. Dauncey et al. Design and performance of the level 1 calorimeter trigger for the *BABAR* detector. *IEEE Trans. Nucl. Sci.*, 48:541–546, 2001.

[67] The Objectivity^{TM} database. http://www.objectivity.com.

[68] ROOT, an object oriented data analysis framework. http://root.cern.ch/.

[69] Stanford researchers may have largest database. http://archives.cnn.com/2002/TECH/industry/04/18/stanford.database.idg/.

[70] http://www.slac.stanford.edu/BFROOT/www/Physics/Analysis/AWG/chrmls_hadronic/ThreeBody/Software/NonCharm3BodyUser_cm2.html.

[71] http://www.slac.stanford.edu/~grl/TrackParProposal.ps.

[72] P. Billoir. Track fitting with multiple scattering: A new method. *Nucl. Instr. Meth.*, A225:352, 1984.

[73] S. Spanier and G. Mancinelli. Kaon selection at *BABAR*. *BABAR* Analysis Document #116.

[74] T. Brandt. Likelihood-based electron identification. *BABAR* Analysis Document #396.

[75] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[76] T. Sjöstrand. High-energy-physics event generation with PYTHIA 5.7 and JETSET 7.4. *Comput. Phys. Commun.*, 82:74–90, 1994.

[77] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[78] C. Hearty. Measurement of the number of $\Upsilon(4S)$ mesons produced in run 1 (B counting). *BABAR* Analysis Document #134.

[79] T. Lange et al. Tag04: an improved tagging algorithm for *CP* violation studies. *BABAR* Analysis Document #729, Version 2.

[80] A. Hoecker et al. Measurement of *CP*-violating asymmetries in $B^0 \to \pi^+\pi^-\pi^0$. *BABAR* Analysis Document #637.

[81] D. A. Bowerman et al. Measurement of the branching fraction and charge asymmetry for the decay $B^\pm \to K^{*\pm}\pi^0$. *BABAR* Analysis Document #814.

[82] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals Eugen.*, 7:179–188, 1936.

[83] W. M. Dunwoodie. Simple observations on $B\overline{B}$ kinematics. http://www.slac.stanford.edu/∼wmd/bbkinematics/kinematics.note.

[84] J. Smith et al. Recommendation for exclusive B reconstruction analysis variables. *BABAR* Note #134.

[85] W. T. Ford. Choice of kinematic variables in B meson reconstruction—take 3. *BABAR* Analysis Document #53.

[86] P. F. Harrison. Blind analysis. *J. Phys.*, G28:2679–2692, 2002.

[87] J. P. Burke. Search for $B^+ \to \overline{K}^*(892)^0(\to K^-\pi^+)K^+$ using the extended maximum likelihood method. *BABAR* Analysis Document #1413, Version 3, http://www.slac.stanford.edu/∼phjpb/index/work/papers/BAD1413V3.ps.

[88] M. Barrett and A. J. Hart. Analysis of the $B^\pm \to K^+K^-K^\pm$ Dalitz plot. *BABAR* Analysis Document #865, Version 6.

[89] Heavy flavor averaging group – B^+ branching fractions, 2005. http://www.slac.stanford.edu/xorg/hfag/rare/leppho05/charmless/OUTPUT/TABLES/charmless.pdf.

[90] Heavy flavor averaging group – B^+ polarization measurements, 2005. http://www.slac.stanford.edu/xorg/hfag/rare/leppho05/polar/OUTPUT/TABLES/polar.pdf.

[91] G. Cowan. *Statistical data analysis.* Oxford University Press Inc., 1998.

[92] F. James and M. Roos. 'MINUIT' a system for function minimization and analysis of the parameter errors and correlations. *Comput. Phys. Commun.*, 10:343–367, 1975.

[93] F. James. MINUIT, function minimization and error analysis, reference manual. http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html.

[94] Kirkby D. and Verkerke W. RooFit users manual v2.07. http://roofit.sourceforge.net/docs/index.html.

[95] T. Skwarnicki. A study of the radiative cascade transitions between the upsilon-prime and upsilon resonances. DESY F31-86-02.

[96] H. Albrecht et al. Exclusive hadronic decays of B mesons. *Z. Phys.*, C48:543–552, 1990.

[97] I. M. Sobol. *The Monte Carlo Method*. University Chicago Press, Chicago, 1974.

[98] Kyle S. Cranmer. Kernel estimation in high-energy physics. *Comput. Phys. Commun.*, 136:198–207, 2001.

[99] The *BABAR* Tracking Efficiency Task Force. http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEfficTaskForce/TrackingTaskForce-2005.html.

[100] The *BABAR* PID Group. http://www.slac.stanford.edu/BFROOT/www/Physics/Tools/Pid/pid.html.

[101] S. J. Bailey et al. Exclusive charmless two-body B decays to η and η' mesons. *BABAR* Analysis Document #521, Version 26, http://www.slac.stanford.edu/~phjpb/index/work/papers/BAD521V26AppG.ps.

[102] J. J. Back. Studies of $B^+ \to h^+h^-h^+(h = \pi/K)$ decays on the full summer 2002 dataset. *BABAR* Analysis Document #505, Version 8.

[103] J. J. Back et al. LAURA^{++} – likelihood analysis unbinned reconstruction of amplitudes. *BABAR* Analysis Document #806, Version 1.

[104] R. J. Barlow et al. Recommended statistical procedures for *BABAR*. *BABAR* Analysis Document #318, Version 2, http://www.slac.stanford.edu/∼phjpb/index/work/papers/BAD318V2.ps.

[105] The *BABAR* Collaboration. Improved measurement of *CP* asymmetries in $B^0 \to (c\overline{c})K^{*0}$ decays. 2006. hep-ex/0607107.

[106] The *BABAR* Collaboration. Measurement of *CP*-violating asymmetries in the $B^0 \to K^+K^-K^0$ Dalitz plot. 2006. hep-ex/0607112.

[107] The Belle Collaboration. Observation of time-dependent *CP* violation in $B^0 \to \eta' K^0$ decays and improved measurements of *CP* asymmetries in $B^0 \to \phi K^0$, $K^0_S K^0_S K^0_S$ and $B^0 \to J/\psi\, K^0$ decays. 2006. hep-ex/0608039.

[108] D. M. Asner et al. Search for exclusive charmless hadronic B decays. *Phys. Rev.*, D53:1039–1050, 1996.