Measurement of the B_s^0 lifetime in $B_s^0 \to K^+K^-$ decays

Nicola Pounder
Wolfson College, Oxford

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Michaelmas Term, 2008

Abstract

A method is presented to simultaneously separate the contributions to a sample of $B^0_{(s)} \to h^+h'^-$ decays, where $h = \pi$ or K, and measure the B meson lifetimes in the sample while correcting for the bias in the lifetime distributions due to the hadronic trigger at the CDF experiment.

Using 1 fb^{-1}of data collected at CDF the B^0 lifetime is measured as

$$\tau_{B^0} = 1.558^{+0.050}_{-0.047\,\mathrm{stat}} \pm 0.028_{\mathrm{syst}}\ \mathrm{ps},$$

in agreement with the world average measurement. The B^0_s lifetime in the $B^0_s \to K^+K^-$ decay is measured as

$$\tau_{B^0_s \to K^+K^-} = 1.51^{+0.13}_{-0.11\,\mathrm{stat}} \pm 0.04_{\mathrm{syst}}\ \mathrm{ps}.$$

No difference is observed between this lifetime and other measurements of the average B^0_s lifetime or the lifetime of the light B^0_s mass eigenstate determined from $B^0_s \to J/\psi\phi$ decays.

With the assumptions that $B^0_s \to K^+K^-$ is 100% CP-even and that $\tau_{B^0_s} = \tau_{B^0}$ the width difference in the B^0_s system is determined as

$$\frac{\Delta\Gamma^{\mathrm{CP}}}{\Gamma} = 0.03^{+0.17}_{-0.15\,\mathrm{stat}} \pm 0.05_{\mathrm{syst}}$$

using the current world average B^0 lifetime. This is consistent with zero and with the current world average measurement.

Dedicated to Diane and Terry Pounder.
Thank you Mum and Dad.

Acknowledgements

Many people have given invaluable help, advice and support to me as I pursued the research for this D.Phil. I am grateful to you all.

I would like to thank Todd Huffman for his supervision and enthusiasm.

To Sneha Malde, you made a world of difference!

Thank you to Jonas Rademacker who's clever ideas began this work and who's support and guidance helped complete it.

I had so many useful discussions with the Oxford-Pittsburgh B-group, thank you Sneha Malde, Jonas Rademacker, Todd Huffman, Farrukh Azfar, Louis Lyons, Joe Boudreau and Aziz Rahaman.

I am extremely grateful to Michael Morrello for all his guidance in perfoming the statistical separation of $B^0_{(s)} \to h^+h'^-$.

Many friends have supported me in the last few years but I must particularly mention Matthew Leslie, Laura Somerville, Sam Harper, Ian McArthur, Cigdem Isserver, Louis Lyons and Sinead Farrington. Thank you.

Thank you to the members of Fermilab, particularly the CDF collaboration and Fermiposse, to the members of the University of Oxford Particle Physics department and my patient housemates.

I gratefully acknowledge funding form the Science and Technology Facilities Council (formerly the Particle Physics and Astronomy Research Council).

I am also grateful for additional funding from the University of Pittsburgh.

Contents

List of Figures

List of Tables

Introduction

This thesis is concerned with measuring the lifetime of the B_s^0 meson in the decay $B_s^0 \to K^+K^-$. The $B_s^0 \to K^+K^-$ decay can be described by Standard Model processes but may also be sensitive to physics beyond the Standard Model. Chapter 1 discusses the theoretical motivation for this analysis. The B_s^0 mesons used for this analysis have been produced at the Tevatron accelerator at the Fermi National Accelerator Laboratory and detected by the Collider Detector at Fermilab (CDF). These are described in chapter 2. The reconstruction and selection of the data from CDF for this analysis, as well as simulated data used for testing the analysis methods, are discussed in chapter 3.

Measuring the $B_s^0 \to K^+K^-$ lifetime in this data sample is complicated by two factors. Firstly, there are a number of very similar decays, such as $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$, which can not be separated from $B_s^0 \to K^+K^-$ decays on an event by event basis; however, they may be statistically disentangled and this is the subject of chapter 4. Secondly, the methods used to collect the data bias the distribution of lifetimes in the data sample. Chapter 5 describes a method to remove this bias. The methods of chapters 4 and 5 are combined in chapter 6.

Chapter 7 studies the performance of the methods used in this analysis and assesses the systematic uncertainties on the measurements. Finally the results of this analysis are presented and discussed in chapter 8.

Chapter 1

B meson lifetimes

1.1 Introduction

This thesis is concerned with the lifetime of the B_s^0 meson in its decay to K^+K^-. The physics describing this process is contained within the Standard Model of particle physics (SM) [1] which describes the properties of elementary particles and the interactions between them. The particles of the Standard Model can be divided in to fermions (fractional spin particles) and bosons (integer spin particles). The fundamental fermions consist of six quarks, six leptons and their anti-particles. The fundamental fermions of the Standard Model and some of their properties are summarised in table 1.1.

The Standard Model is a quantum field theory based on gauge invariance with the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. It describes the strong, weak and electromagnetic interactions. $SU(3)_C$ describes the strong force which is mediated by eight, massless, spin-1 gluons that interact with the colour charge, C, carried by all quarks and gluons. The electromagnetic and weak interactions are unified in the Standard Model as a single electroweak interaction, described by $SU(2)_L \times U(1)_Y$. Spontaneous symmetry breaking leads to the electromagnetic force, mediated by the

Quarks				
Generation	Particle	Symbol	Charge (e)	Mass
I	up	u	$+2/3$	1.5 to 3.3 MeV/c^2
	down	d	$-1/3$	3.5 to 6.0 MeV/c^2
II	charm	c	$+2/3$	$1.27^{+0.07}_{-0.11}$ GeV/c^2
	strange	s	$-1/3$	104^{+26}_{-34} MeV/c^2
III	top	t	$+2/3$	171.2 ± 2.1 GeV/c^2
	bottom	b	$-1/3$	$4.20^{+0.17}_{-0.07}$ GeV/c^2
Leptons				
Generation	Particle	Symbol	Charge (e)	Mass
I	electron	e	-1	0.51 MeV/c^2
	electron neutrino	ν_e	0	
II	muon	μ	-1	105.66 MeV/c^2
	muon neutrino	ν_μ	0	
III	tau	τ	-1	1776.84 ± 0.17 MeV/c^2
	tau neutrino	ν_τ	0	

Table 1.1: Properties of the observed fundamental fermions [2]. The observation of neutrino oscillations [3] indicates that the squared mass difference between the neutrino species is non zero so at least two of the neutrino types have mass; however, the mass of each type is yet to be determined.

massless photon, and the weak force, mediated by the massive, charged $W^{\pm}$ and neutral Z^0 bosons [4, 5].

The predictions of the Standard Model have been experimentally tested to remarkable accuracy. Much of current experimental research is looking for hints of physics that can not be described by the Standard Model and require a new theory. This is frequently referred to as new physics (NP).

The remainder of this chapter will concentrate on the physical processes behind B meson lifetimes, particularly in the decay $B_s^0 \to K^+K^-$.

1.2 The CKM description of quark transitions

B mesons are produced through quark hadronisation as flavour specific eigenstates: $B^0 = |\overline{b}d\rangle$; $B^+ = |\overline{b}u\rangle$; $B_s^0 = |\overline{b}s\rangle$; $B_c^+ = |\overline{b}c\rangle$ and the corresponding anti-particles. The lightest B meson states decay only via the weak interaction with the heavy quark

decaying as $\mathrm{b} \to \mathrm{W}q$ where $q = \mathrm{u}, \mathrm{c}$. The Lagrangian for the charged current weak interaction is

$$\mathcal{L}_{CC} = -\frac{g_2}{\sqrt{2}} \left(\overline{\mathrm{u}}_L, \quad \overline{\mathrm{c}}_L, \quad \overline{\mathrm{t}}_L \right) \gamma^\mu V_{\mathrm{CKM}} \begin{pmatrix} \mathrm{d}_L \\ \mathrm{s}_L \\ \mathrm{b}_L \end{pmatrix} W_\mu^\pm \tag{1.1}$$

where the Cabibbo-Kobayashi-Maskawa matrix (V_{CKM}) transforms between the mass and weak eigenstates [6, 7]:

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

The elements of V_{CKM} may be complex and so it contains 2×3^2 real parameters, four of which are independent. One common parameterisation is the perturbative parameterisaton proposed by Wolfenstein [8]:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3 (\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3 (1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.3}$$

The parameters are, from experiment, $\lambda \approx 0.2$, $A \approx 0.8$, $\rho \approx 0.1$ and $\eta \approx 0.3$ [2]. This parameterisation highlights the relative strengths of the quark couplings with the diagonal elements of order 1, which represent transitions within quark generations, transitions between neighbouring generations are of order λ or λ^2 and transitions across 2 generations are of order λ^3. It is the presence of a complex phase in the quark transitions that allows CP violation in the Standard Model.

1.3 Lifetimes of B mesons

A b quark decays to either a c or u quark via a virtual W boson. The lifetime of a free b quark for a hadronic W decay to $q_1\overline{q}_2$ is

$$\Gamma_{q_1\bar{q}_2}\,(\mathrm{b} \rightarrow q) = \frac{3G_F^2 m_b^5}{192\pi^3} 3|V_{q_1 q_2}|^2 |V_{\mathrm{qb}}|^2 F(\epsilon_q) \tag{1.4}$$

where m_b is the b quark mass, G_F the Fermi coupling constant and $F(\epsilon^2)$ is a phase space factor. The b quark lifetime is directly connected to the elements V_{qb} of the V_{CKM} matrix. To first order the lifetimes of B mesons can be determined according the spectator model. In this model the decay of the B meson is dominated by the b quark with the lighter quark, q, merely a spectator. In the spectator model the lifetime of the B mesons is determined by equation 1.4 and all B mesons should have the same lifetime. In fact the observed hierarchy of B meson lifetimes is

$$\tau_{\mathrm{B_c^+}} < \tau_{\mathrm{B_s^0}} \approx \tau_{\mathrm{B^0}} < \tau_{\mathrm{B^+}}. \tag{1.5}$$

The presence of a hierarchy of lifetimes indicates that the lighter quark in the B meson does influence the decay.

The differences in the B meson lifetimes can be understood using Heavy Quark Expansion Theory (HQE) (see for example [9, 10, 11]) which is based in Quantum Chromodynamics (QCD). In HQE the decay rate is expressed in an expansion series of $1/m_b$. The leading order terms are equal to the spectator model. Terms of order $1/m_b^2$ are higher order corrections associated only with the b quark and so do not lead to a difference in B meson lifetimes. Terms of order $1/m_b^3$ describe effects between the constituent quarks. In general, measurements of B meson lifetimes test HQE.

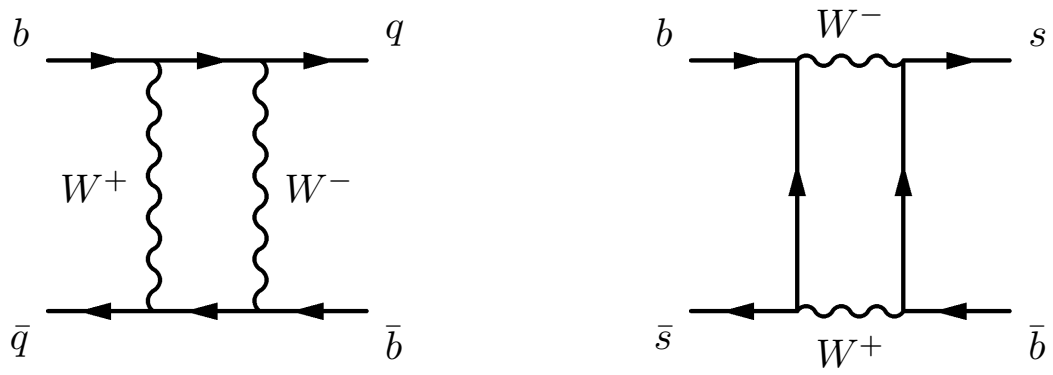


Figure 1.1: Lowest order Feynman diagrams that allow mixing in the B-meson system.

1.4 Mixing in neutral B mesons

The $B^0_{(s)}$ and $\overline{B^0_{(s)}}$ are produced as strong eigenstates. If $V_{\rm CKM}$ were diagonal these states would be stable; however, the small off diagonal terms allow transitions between $B^0_{(s)}$ and $\overline{B^0_{(s)}}$. Mixing between the two states allows the originally produced B meson to evolve into a quantum superposition of both states:

$$a(t)|\mathrm{B}\rangle + b(t)|\overline{\mathrm{B}}\rangle \tag{1.6}$$

where $a(t)^2 + b(t)^2 = 1$. Figure 1.1 shows the lowest order Feynman diagrams that allow mixing. The time evolution of the state is governed by the effective Schrödinger equation

$$i\frac{\partial}{\partial t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M^*_{12} - \frac{i}{2}\Gamma^*_{12} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix} \times \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \tag{1.7}$$

where the mass matrix $\mathbf{M}$ and the decay matrix $\mathbf{\Gamma}$ are time independent Hermitian matrices. Charge parity time (CPT) invariance requires that $M_{11} = M_{22} \equiv M$ and $\Gamma_{11} = \Gamma_{22} \equiv \Gamma$. The elements of $\mathbf{M}$ and $\mathbf{\Gamma}$ can be calculated from the amplitudes in the box diagrams, figure 1.1. The eigenstates of the Hamiltonian are the mass eigenstates and can be expressed as a linear combination of the flavour eigenstates:

$$|\mathrm{B_{L,H}}\rangle = p|\mathrm{B}\rangle \pm q|\overline{\mathrm{B}}\rangle \tag{1.8}$$

where the subscripts L and H indicate the light and heavy mass states respectively and $|p|^2 + |q|^2 = 1$. The corresponding eigenvalues are

$$\lambda_{\mathrm{L,H}} = \left(M - \frac{i}{2}\Gamma\right) \pm \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) \tag{1.9}$$

where

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{1.10}$$

The time evolution of the mass eigenstates is governed by the eigenvalues

$$|\mathrm{B_{L,H}}(t)\rangle = e^{-i\lambda_{\mathrm{L,H}}t}|\mathrm{B_{L,H}}\rangle \tag{1.11}$$

where $|\mathrm{B_{L,H}}\rangle = |\mathrm{B_{L,H}}(t=0)\rangle$. Defining $M_{\mathrm{L,H}} = \Re(\lambda_{\mathrm{L,H}})$ and $\Gamma_{\mathrm{L,H}} = -2\Im(\lambda_{\mathrm{L,H}})$ this becomes

$$|\mathrm{B_{L,H}}(t)\rangle = e^{-\left(iM_{\mathrm{L,H}} + \frac{\Gamma_{\mathrm{L,H}}}{2}\right)t}|\mathrm{B_{L,H}}\rangle. \tag{1.12}$$

This can be translated into the time evolution of the flavour eigenstates as

$$|\mathrm{B}(t)\rangle = g_+(t)|\mathrm{B}\rangle + \frac{q}{p}g_-(t)|\overline{\mathrm{B}}\rangle \tag{1.13}$$

$$|\overline{\mathrm{B}}(t)\rangle = \frac{q}{p}g_-(t)|\mathrm{B}\rangle + g_+(t)|\overline{\mathrm{B}}\rangle \tag{1.14}$$

where $g_\pm(t) = \frac{1}{2}(e^{-i\lambda_L t} \pm e^{-i\lambda_H t})$. The time dependent probability that an initial state decays as a particular flavour specific state can be calculated as

$$\left|\langle \mathrm{B}|\overline{\mathrm{B}}(t)\rangle\right|^2 = \left|\frac{p}{q}\right|^2 |g_-(t)|^2, \tag{1.15}$$

$$\left|\langle \overline{\mathrm{B}}|\mathrm{B}(t)\rangle\right|^2 = \left|\frac{q}{p}\right|^2 |g_-(t)|^2, \tag{1.16}$$

$$|\langle \mathrm{B}|\mathrm{B}(t)\rangle|^2 = |g_+(t)|^2, \tag{1.17}$$

$$\left|\langle \overline{\mathrm{B}}|\overline{\mathrm{B}}(t)\rangle\right|^2 = |g_+(t)|^2 \tag{1.18}$$

where

$$|g_+(t)|^2 = \frac{1}{2}\left(\cosh\left(\frac{\Delta\Gamma t}{2}\right) \pm \cos(\Delta m t)\right) e^{-\Gamma t}. \tag{1.19}$$

The mass and width difference have been defined as

$$\Delta m \equiv M_{\mathrm{H}} - M_{\mathrm{L}} \quad \text{and} \quad \Delta\Gamma \equiv \Gamma_{\mathrm{L}} - \Gamma_{\mathrm{H}}. \tag{1.20}$$

In addition the average mass and width are commonly defined as

$$m \equiv \frac{M_{\mathrm{H}} + M_{\mathrm{L}}}{2} \quad \text{and} \quad \Gamma \equiv \frac{\Gamma_{\mathrm{L}} + \Gamma_{\mathrm{H}}}{2}. \tag{1.21}$$

In this convention Δm is positive by definition. The sign of $\Delta\Gamma$ needs to be determined experimentally but the Standard Model predicts that it should be positive. Equations 1.13, 1.14 and 1.19 describe the phenomena of flavour oscillations where Δm describes the frequency of the oscillations and $\Delta\Gamma$ is the modulation term.

It is possible to express Δm and $\Delta\Gamma$ in terms of three physical quantities of neutral B mixing; that is $|M_{12}|$, $|\Gamma_{12}|$ and the relative phase between M_{12} and Γ_{12}:

$$\phi = \arg\left(-\frac{M_{12}}{\Gamma_{12}}\right). \tag{1.22}$$

A simple, approximate solution can be found by making the assumptions

$$|\Gamma_{12}| \ll |M_{12}| \quad \text{and} \quad \Delta\Gamma \ll \Delta m \tag{1.23}$$

which hold empirically for both the B^0 and $\mathrm{B}^0_{\mathrm{s}}$ systems. Expanding in Γ_{12}/M_{12} provides

$$\Delta m \approx 2\,|M_{12}|\,, \tag{1.24}$$

$$\Delta\Gamma \approx 2\,|\Gamma_{12}\cos\phi| \tag{1.25}$$

$$\frac{q}{p} \approx -e^{-i\phi_M}\left(1 - \frac{a}{2}\right) \tag{1.26}$$

where parameter a is defined as

$$a = \Im\frac{\Gamma_{12}}{M_{12}} = \left|\frac{\Gamma_{12}}{M_{12}}\right| \sin\phi \tag{1.27}$$

and ϕ_M is the phase of M_{12},

$$M_{12} = |M_{12}|\, e^{i\phi_M}. \tag{1.28}$$

1.5 $\Delta\Gamma_s$ in the Standard Model

Consider equations 1.8 and 1.10; if $|q/p|^2 = 1$ then the CP eigenstates are equivalent to the mass eigenstates. For the B_s^0 system:

$$|B_L\rangle = |B^{\text{CP-even}}\rangle = \frac{1}{\sqrt{2}}|B_s^0\rangle - \frac{1}{\sqrt{2}}|\overline{B_s^0}\rangle \tag{1.29}$$

$$|B_H\rangle = |B^{\text{CP-odd}}\rangle = \frac{1}{\sqrt{2}}|B_s^0\rangle + \frac{1}{\sqrt{2}}|\overline{B_s^0}\rangle. \tag{1.30}$$

CP-violation in B mixing is allowed if $|q/p|^2$ deviates from 1. This will only occur if $\Gamma_{12} \neq 0$, $M_{12} \neq 0$ and if the CP violating phase ϕ is different from 0 or π. In the Standard Model M_{12} arises from the virtual top quark contribution in the mixing diagram (figure 1.1). Γ_{12} is dominated by $b \to c\bar{c}s$ decays, transitions which do not involve a CP violating phase. Only contributions to Γ_{12} of the suppressed decays $b \to c\bar{u}s$, $u\bar{c}s$ and $u\bar{u}s$ can introduce a CP violating phase in the Standard Model and so CP violation in mixing is negligible in the B_s^0 system and can be assumed to be zero.

U. Nierste and A. Lenz have explored what physical processes can yield measurements of the B_s^0 width difference ($\Delta\Gamma_s$) in Standard Model processes. They have also explored how experimental measurements of these processes might be sensitive to physics beyond the Standard Model [12]. Their work forms the basis of the discussion which follows and provides the theoretical motivation for this analysis.

1.5.1 $b \to c\bar{c}s$ decay in the Standard Model

The time evolution of a $B_s^0/\overline{B_s^0}$ decay to a final state f can be described by

$$\Gamma(f,t) \propto \left|\langle f \mid B^0_{sL}\rangle\right|^2 e^{-\Gamma_L t} + \left|\langle f \mid B^0_{sH}\rangle\right|^2 e^{-\Gamma_H t}. \tag{1.31}$$

The Standard Model $b \to c\bar{c}s$ decay amplitude has a negligible CP violating phase. For the decay to a CP-even eigenstate such as $f = (J/\psi\phi)_{L=0}$ (where the orbital angular momentum between the final state particles, $L = 0$)

$$\langle (J/\psi\phi)_{L=0} \mid B^0_{sH}\rangle = \langle (J/\psi\phi)_{L=0} \mid B_s^{\text{CP-even}}\rangle = 0 \tag{1.32}$$

and the lifetime measured using $B_s^0(\overline{B_s^0}) \to (J/\psi\phi)_{L=0}$ determines Γ_L.

For a flavour specific decay, such as $B_s^0 \to D_s^{\pm}\pi^{\mp}$ or semi-leptonic B_s^0 decays, the decays $B_s^0 \to \overline{f}$ and $\overline{B_s^0} \to f$ are forbidden so $|\langle f \mid B^0_{sL}\rangle| = |\langle f \mid B^0_{sH}\rangle|$ and

$$\Gamma(f,t) \propto e^{-\Gamma_L t} + e^{-\Gamma_H t}. \tag{1.33}$$

The lifetime measured in a flavour specific mode provides a weighted average of Γ_L and Γ_H.

The lifetime measured using $B_s^0(\overline{B_s^0}) \to (J/\psi\phi)_{L=0}$ can be combined with the lifetime measured using flavour specific B_s^0 decays to determine $\Delta\Gamma_s$. An alternative approach is to use the B^0 rather than flavour specific B_s^0 lifetime. Theory predicts [13] that

$$\Gamma_{B_s^0} \equiv \frac{\Gamma_L + \Gamma_H}{2} = \Gamma_{B^0} + \mathcal{O}(1\%) \tag{1.34}$$

and $\Delta\Gamma_s$ can be calculated as

$$\Delta\Gamma_s = 2(\Gamma_L - \Gamma_{B_s^0}). \tag{1.35}$$

Measuring the B_s^0 lifetime in $B_s^0(\overline{B_s^0}) \to (J/\psi\phi)_{L=0}$ is not straightforward as the $L = 0$ contribution has to be disentangled from $L = 1$ (CP-odd) and $L = 2$ (CP-even) contributions to $B_s^0 \to J/\psi\phi$. This can be done using an angular analysis which can also be used to determine $\Delta\Gamma_s = \Gamma_L - \Gamma_H$ [14, 15]. The latest CDF measurement in this mode obtained [16]

$$\Delta\Gamma_s = 0.076^{+0.059}_{-0.063} \pm 0.006 \text{ ps}^{-1} \tag{1.36}$$

$$\tau_s = 1.52 \pm 0.04 \pm 0.02 \text{ ps}^{-1} \tag{1.37}$$

where in each case the first uncertainty is statistical and the second systematic.

In addition to measuring $\Delta\Gamma_s$ the angular analysis of $B_s^0 \to J/\psi\phi$ can simultaneously determine the angle β_s, which in the Standard Model is

$$\beta_s^{SM} = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) \tag{1.38}$$

and $\beta_s^{SM} \approx 0.02$ [17]. β_s may be modified by the presence of physics beyond the Standard Model. A 1.8σ deviation from the Standard Model prediction of $\Delta\Gamma_s$ and β_s has recently been observed at CDF [18].

1.6 $\Delta\Gamma_s$ beyond Standard Model

The B_s^0 width difference is, from equation 1.25, $\Delta\Gamma_s \approx 2\,|\Gamma_{12}\cos\phi|$. Γ_{12} derives from tree level processes and is difficult to change significantly in models of new physics so it can be assumed that $\Gamma_{12} = \Gamma_{12,SM}$, that is, Γ_{12} in the Standard Model. New physics models can, therefore, only affect $\Delta\Gamma_s$ through $\cos\phi$. ϕ can be written as the sum of Standard Model and new physics contributions so $\cos\phi = \cos(\phi^{SM} + \phi^{NP})$. The Standard Model mixing CP-violating phase, $\phi^{SM} = 0$ and so

$$\Delta\Gamma_s = \Delta\Gamma_{SM}\cos\phi^{NP}. \tag{1.39}$$

1.6.1 $b \to c\bar{c}s$ decay beyond the Standard Model

If $\phi^{NP} \neq 0$ then B^0_{sL} and B^0_{sH} are no longer CP-eigenstates. This means that both mass eigenstates can decay into $(J/\psi\phi)_{L=0}$. Consider the time evolution of a decay $B^0_s(\overline{B^0_s}) \to f$ described in equation 1.31. For a $b \to c\bar{c}s$ decay into a CP-even state, such as $(J/\psi\phi)_{L=0}$,

$$\left|\langle f_{CP+} \mid B^0_{sL}\rangle\right|^2 = \frac{1+\cos\phi^{NP}}{2}\left|\langle f_{CP+} \mid B_s^{\text{CP-even}}\rangle\right|^2 \tag{1.40}$$

and

$$\left|\langle f_{CP+} \mid B^0_{sH}\rangle\right|^2 = \frac{1-\cos\phi^{NP}}{2}\left|\langle f_{CP+} \mid B_s^{\text{CP-even}}\rangle\right|^2. \tag{1.41}$$

In the presence of a new mixing phase, ϕ^{NP}, equation 1.31 becomes

$$\Gamma(f,t) \propto \frac{1+\cos\phi^{NP}}{2}e^{-\Gamma_L t} + \frac{1-\cos\phi^{NP}}{2}e^{-\Gamma_H t} \tag{1.42}$$

and the lifetime measured in $B^0_s \to (J/\psi\phi)_{L=0}$ is no longer Γ_L.

1.6.2 $b \to s\bar{s}s$ decay beyond the Standard Model

Unlike $b \to c\bar{c}s$, which is a tree-level process, $b \to s\bar{s}s$ can only occur via penguin processes (see, for example, the lowest order Feynman diagrams in figures 1.2 and 1.3). In the Standard Model there is no direct CP-violating phase in the $b \to s\bar{s}s$ transition; however, models of new physics can introduce a CP phase, σ^{NP}, through the addition of new particles which may participate in the loops of the penguin diagrams.

For a $b \to s\bar{s}s$ decay into a CP-even eigenstate f_{CP+}, such as $f_{CP+} = (\phi\phi)_{L=0}$,

$$\langle f_{CP+} \mid B^0_s\rangle \propto e^{i\sigma^{NP}} \quad \text{and} \quad \langle f_{CP+} \mid \overline{B^0_s}\rangle \propto -e^{-i\sigma^{NP}}. \tag{1.43}$$

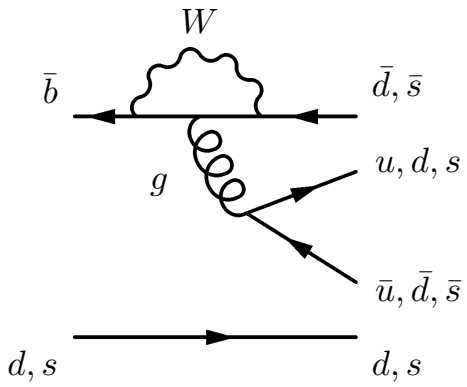


Figure 1.2: Strong penguin diagram for $B_s^0 \to \phi\phi$ and $B_{(s)}^0 \to h^+h'^-$ decays.

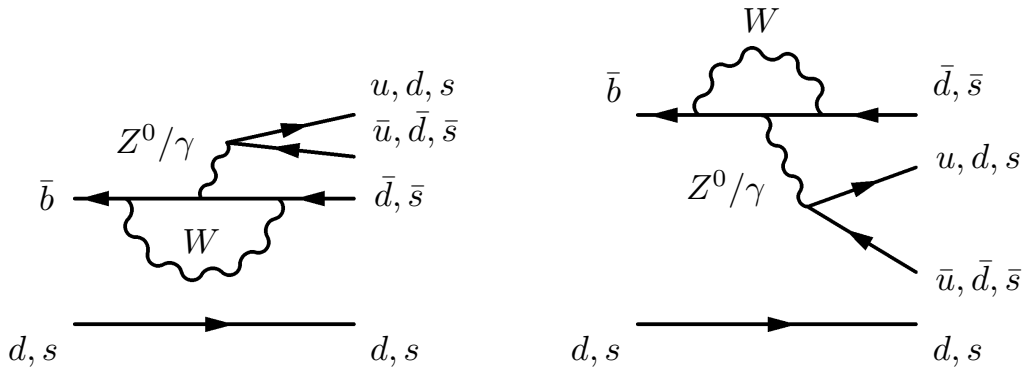


Figure 1.3: Weak penguin diagrams for the $B_s^0 \to \phi\phi$ and $B_{(s)}^0 \to h^+h'^-$ decays.

The coefficients in the time evolution of a decay $B_s^0(\overline{B_s^0}) \to f_{CP+}$ are, in this case,

$$\left|\langle f_{CP+} \mid B^0_{s\,L}\rangle\right|^2 = \frac{1+\cos(\phi^{NP}+2\sigma^{NP})}{2}\left|\langle f_{CP+} \mid B_s^{\text{CP-even}}\rangle\right|^2 \tag{1.44}$$

and

$$\left|\langle f_{CP+} \mid B^0_{s\,H}\rangle\right|^2 = \frac{1-\cos(\phi^{NP}+2\sigma^{NP})}{2}\left|\langle f_{CP+} \mid B_s^{\text{CP-even}}\rangle\right|^2. \tag{1.45}$$

For $b \to s\bar{s}s$ decays, equation 1.31 becomes

$$\Gamma(f,t) \propto \frac{1+\cos(\phi^{NP}+2\sigma^{NP})}{2}e^{-\Gamma_L t} + \frac{1-\cos(\phi^{NP}+2\sigma^{NP})}{2}e^{-\Gamma_H t}. \tag{1.46}$$

In the Standard Model there are no phases ϕ or σ in the $b \to s\bar{s}s$ decay and the lifetime measured in the decay $B_s^0 \to (\phi\phi)_{L=0}$ determines Γ_L. In this case the lifetimes measured in $B_s^0 \to (J/\psi\phi)_{L=0}$ and $B_s^0 \to (\phi\phi)_{L=0}$ will be the same.

Compare equation 1.42 for $b \to c\bar{c}s$ decays to equation 1.46 for $b \to s\bar{s}s$. If the lifetime measured in $B_s^0 \to (\phi\phi)_{L=0}$ is longer than $B_s^0 \to (J/\psi\phi)_{L=0}$ then this implies new physics in the decay amplitude $b \to s\bar{s}s$ with $\sigma^{NP} \neq 0$. The mixing phase may be either $\phi^{NP} = 0$ or $\phi^{NP} \neq 0$. If the lifetime measured in $B_s^0 \to (\phi\phi)_{L=0}$ is shorter than $B_s^0 \to (J/\psi\phi)_{L=0}$ then this implies new physics in both the decay amplitude $b \to s\bar{s}s$ and the mixing CP phase with $\sigma^{NP} \neq 0$ and $\phi^{NP} \neq 0$.

The difficulty with studying $B_s^0 \to (\phi\phi)_{L=0}$ experimentally is that, as with the $B_s^0 \to (J/\psi\phi)_{L=0}$ decay, it needs to be disentangled from $L = 1$ and $L = 2$ $B_s^0 \to (\phi\phi)_L$ decays. The branching ratio for $B_s^0 \to \phi\phi$ is $(1.4 \pm 0.8) \times 10^{-5}$ [2] and low statistics are problematic for angular analysis.

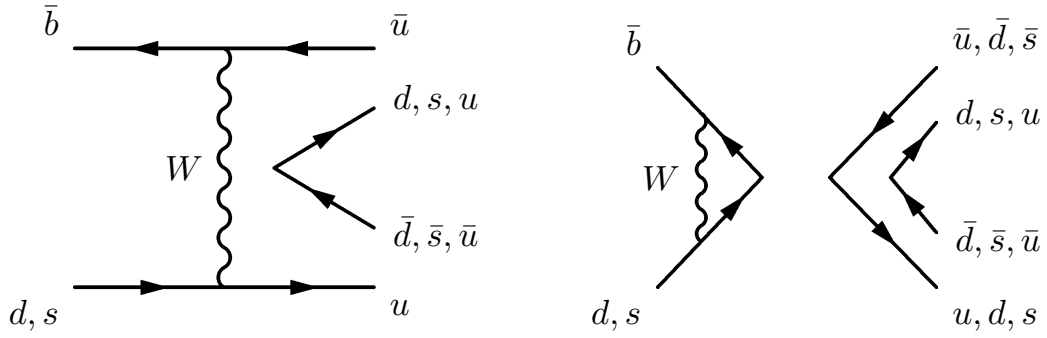


Figure 1.4: W-exchange and W-annihilation diagrams for the $B^0_{(s)} \to h^+h'^-$ decays. The unattached vertices involve the exchange of a neutral boson.

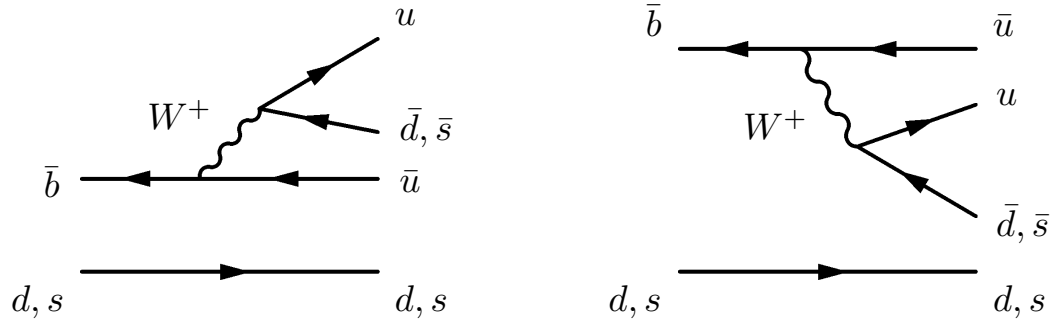


Figure 1.5: Colour allowed (left) and colour suppressed (right) tree-level diagrams for the $B^0_{(s)} \to h^+h'^-$ decays.

1.6.3 $b \to s\bar{u}u$ decay beyond the Standard Model

The decay $B^0_s \to K^+K^-$ also has a small branching ratio, $(3.3 \pm 0.9) \times 10^{-5}$ [2], however there is only one final state, $L = 0$. $B^0_s \to K^+K^-$ is facilitated by the decay $b \to s\bar{u}u$ and is dominated by penguin decays such as those in figures 1.2, 1.3 and 1.4. As with $B^0_s \to (\phi\phi)_{L=0}$, models of new physics can introduce a mixing phase ϕ^{NP} and, into the $b \to s\bar{u}u$ decay amplitude, a CP phase σ^{NP}. Unlike $b \to s\bar{s}u$ there is a small tree amplitude present (figure 1.5) which already includes a small, Standard Model, CP phase in the decay amplitude: σ^{SM}. The total direct CP phase is, therefore, $\sigma = \sigma^{SM} + \sigma^{NP}$.

The time evolution of a decay $B^0_s(\overline{B^0_s}) \to f$ for a $b \to s\bar{u}u$ decay into a CP-even

state (such as $B_s^0 \to K^+K^-$) is described by

$$\Gamma(f,t) \propto \frac{1+\cos(\phi^{NP}+2\sigma)}{2}e^{-\Gamma_L t} + \frac{1-\cos(\phi^{NP}+2\sigma)}{2}e^{-\Gamma_H t}. \quad (1.47)$$

As with $B_s^0 \to (\phi\phi)_{L=0}$, if the lifetime measured in $B_s^0 \to K^+K^-$ is longer than the lifetime measured in $B_s^0 \to (J/\psi\phi)_{L=0}$ then $\sigma \neq 0$ and the mixing phase may be $\phi^{NP} = 0$ or $\phi^{NP} \neq 0$. If the lifetime measured in $B_s^0 \to K^+K^-$ is shorter than the lifetime measured in $B_s^0 \to (J/\psi\phi)_{L=0}$ then $\sigma \neq 0$ and $\phi^{NP} \neq 0$. As has already been stated there is a small Standard Model phase σ^{SM} for $B_s^0 \to K^+K^-$ decays so the Standard Model predicts that $\sigma = \sigma^{SM} \neq 0$. $B_s^0 \to K^+K^-$ is predicted to be 95% CP-even [19].

Assuming $B_s^0 \to K^+K^-$ to be 100% CP-even, a value of $\Delta\Gamma_s$ could be calculated using $\Gamma_{B_s^0 \to K^+K^-} = \Gamma_s + |\Delta\Gamma_s|/2$. A value calculated in this way should not be combined with measurements in other modes as $B_s^0 \to K^+K^-$ would be sensitive to a new physics phase in the $b \to s\bar{u}u$ decay amplitude [12] while other decay modes are not.

1.7 Summary

This analysis is concerned with measuring the lifetime of the B_s^0 meson in $B_s^0 \to K^+K^-$ decays. In the Standard Model $B_s^0 \to K^+K^-$ is predicted to be 95% CP-even and the lifetime measured in a decay to a CP-even eigenstate gives the lifetime of the light mass eigenstate, $\tau_L = 1/\Gamma_L$. Comparison of τ_L to the B^0 lifetime allows $\Delta\Gamma$ to be calculated as, in the Standard Model, $\tau_{B^0} \approx \tau_{B_s^0}$.

The lifetime measured in $B_s^0 \to K^+K^-$ is particularly interesting because the $B_s^0 \to K^+K^-$ decay is dominated by penguin processes. This means it is sensitive to new physics processes that may effect the loops in the penguin diagrams; this shows up as an additional CP phase in the $b \to s\bar{u}u$ decay amplitude. It is because of this sensitivity that $\Delta\Gamma$ measured using $B_s^0 \to K^+K^-$, assuming only Standard Model processes, should

not be combined with measurements using other modes in a world average. Comparing the lifetime measured in $B_s^0 \to K^+K^-$, which is sensitive to both new mixing and direct CP phases, to the lifetime measured in $B_s^0 \to (J/\psi\phi)_{L=0}$, which is only sensitive to a new mixing CP phase, can probe new physics in the $b \to s$ decay amplitude.

Chapter 2

The CDF detector at the Fermilab Tevatron accelerator

2.1 Introduction

B_s^0 mesons can be produced in the collisions of protons and antiprotons. Beams of protons and antiprotons are produced, accelerated and collided at the Tevatron particle accelerator (described in section 2.2) at the Fermi National Accelerator Laboratory (FNAL) in Illinois, U.S.A. The products of the proton-antiproton collisions are detected by the Collider Detector at Fermilab (CDF), described in section 2.3. CDF is a multipurpose detector with many components including tracking detectors (section 2.4), calorimeters (section 2.6) and muon chambers (section 2.7). The rate at which collision events are produced at CDF, by the Tevatron, is orders of magnitude higher than the rate at which the information about them can be stored, so a trigger is employed to select select collision events which may contain interesting physics (such as $B_s^0 \to K^+K^-$ decays). The trigger and in particular the trigger path which selects $B_s^0 \to K^+K^-$ and other $B_{(s)}^0 \to h^+h'^-$ type events is discussed in section 2.8.

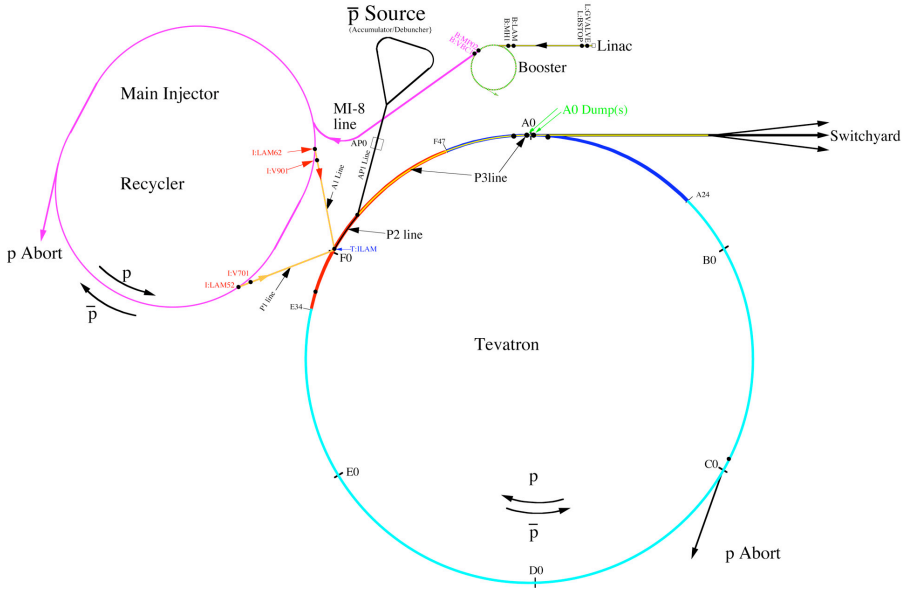


Figure 2.1: The Fermilab accelerator complex (reproduced from [20]).

2.2 The Tevatron

The Tevatron is a particle accelerator colliding protons and antiprotons at a centre of mass energy of 1.96 TeV. These energies are obtained by a series of accelerator systems which are shown in figure 2.1. A description of the accelerator complex can be found in [20].

The accelerator systems are made up of the proton source (the Pre-accelerator, Linac and Booster), the antiproton source (Debuncher and Accumulator), the Main Injector and the Tevatron. The Pre-accelerator is a Cockcroft-Walton accelerator which accelerates H^- ions to an energy of 750 keV. The H- ions are then sent to the Linear Accelerator (Linac) which further accelerates the H^- ions to 400 MeV. From here the H^- ions can be sent to the Booster, a 75 m radius synchrotron. At this stage the H^- ions are stripped of their electrons, by passing them through a carbon foil, leaving protons which are accelerated to 8 GeV in the Booster ready to be sent to the Main Injector.

The Main Injector is a synchrotron with 3 km circumference and several modes of operation. The Main Injector can accelerate the 8 GeV protons from the Booster

to either 120 GeV or 150 GeV depending on their final destination: the antiproton source or the Tevatron. The Main Injector also accepts 8 GeV antiprotons from the antiproton source which it accelerates to 150 GeV ready to be sent to the Tevatron.

The antiproton source consists of the Debuncher and the Accumulator. 120 GeV protons from the Main Injector are directed to strike a nickel target producing a spray of secondary particles. Antiprotons with a momentum of around 8 GeV are selected from these and sent to the Debuncher, a synchrotron, where they are stochastically cooled [21]. The cooled, 8 GeV antiprotons are then sent to the Accumulator (the second synchrotron in the antiproton source, sharing the same tunnel). Here the 8 GeV antiprotons are stored and cooled while more antiprotons are being produced. An electron cooling system has been included in the recycler since August 2005 [22]. When enough antiprotons have been produced they can be sent to the Main Injector to be accelerated further to 150 GeV.

The Tevatron, a synchrotron with a circumference of just over 6 km, is the final stage in the Accelerator chain. The Tevatron accelerates 150 GeV protons and antiprotons from the Main Injector to their final energy of 980 GeV. When loading the Tevatron 7 bunches of protons are injected into the Main Injector from the booster, These bunches are accelerated up to 150 GeV and coalesced into one bunch which is transferred to the Tevatron. This is repeated until there are 36 proton bunches orbiting the Tevatron. Then the antiprotons may be loaded. Antiprotons are sent to the Main Injector from the antiproton source in sets of 4 times 7 bunches which are accelerated to 150 GeV and coalesced into 4 bunches then sent on to the Tevatron. This is repeated 9 times so there are 36 antiproton bunches in the Tevatron. Once all of the proton and antiproton bunches are loaded the two, counter-rotating, beams are both accelerated up to 980 GeV. The interval between proton and antiproton bunches is 396 ns.

Once the protons and antiprotons have been accelerated to 980 GeV they are focused by quadrupole magnets and brought together to collide at the two interaction

regions, CDF and DØ. The two beams can be kept in the Tevatron for several hours until the luminosity drops below a useful value (or some problem causes the beam to be lost). The average store length is about 14 hours.

Figure 2.2 shows the integrated luminosity delivered by the Tevatron over time as well as the peak luminosity in each store. During the period from February 2002 to March 2006 approximately 1.2 fb^{-1} were recorded to tape at CDF. Data quality requirements reduce the sample to the 1 fb^{-1} used in this analysis[1]. To date, over 4 fb^{-1} has so far been delivered by the Tevatron with nearly 3.5 fb^{-1} written to tape at CDF. The Tevatron is planned to operate until 2010.

2.3 The CDF II detector

The CDF II detector is a general purpose detector for the study of interactions from $p\bar{p}$ collisions at an energy of 1.96 TeV at the Fermilab Tevatron collider. A cross-section of the detector is shown in figure 2.3. CDF II consists of tracking elements, which are located within a 1.4 T solenoidal magnet, and calorimeters and muon chambers which are located outside of the solenoid. A detailed description of the CDF II detector can be found in the Technical Design Report [25]. Here only a brief description of the detector components is given with more emphasis on those most important to this analysis, namely the tracking systems and trigger.

2.3.1 The co-ordinate system

CDF uses a right-handed Cartesian co-ordinate system with the origin at the centre of the detector. The direction of the proton beam defines the z-axis while the x-axis

[1]Good performance from the tracking detectors (see section 2.4) is required for this analysis. Data is retained only if the silicon systems and drift chamber are both operating well. In 2004 the drift chamber was found to be ageing unexpectedly quickly [24] and, in order to preserve it, was confined to operate below the standard necessary for this analysis. The ageing was reversed at the end of 2004 and the drift chamber returned to full operation.

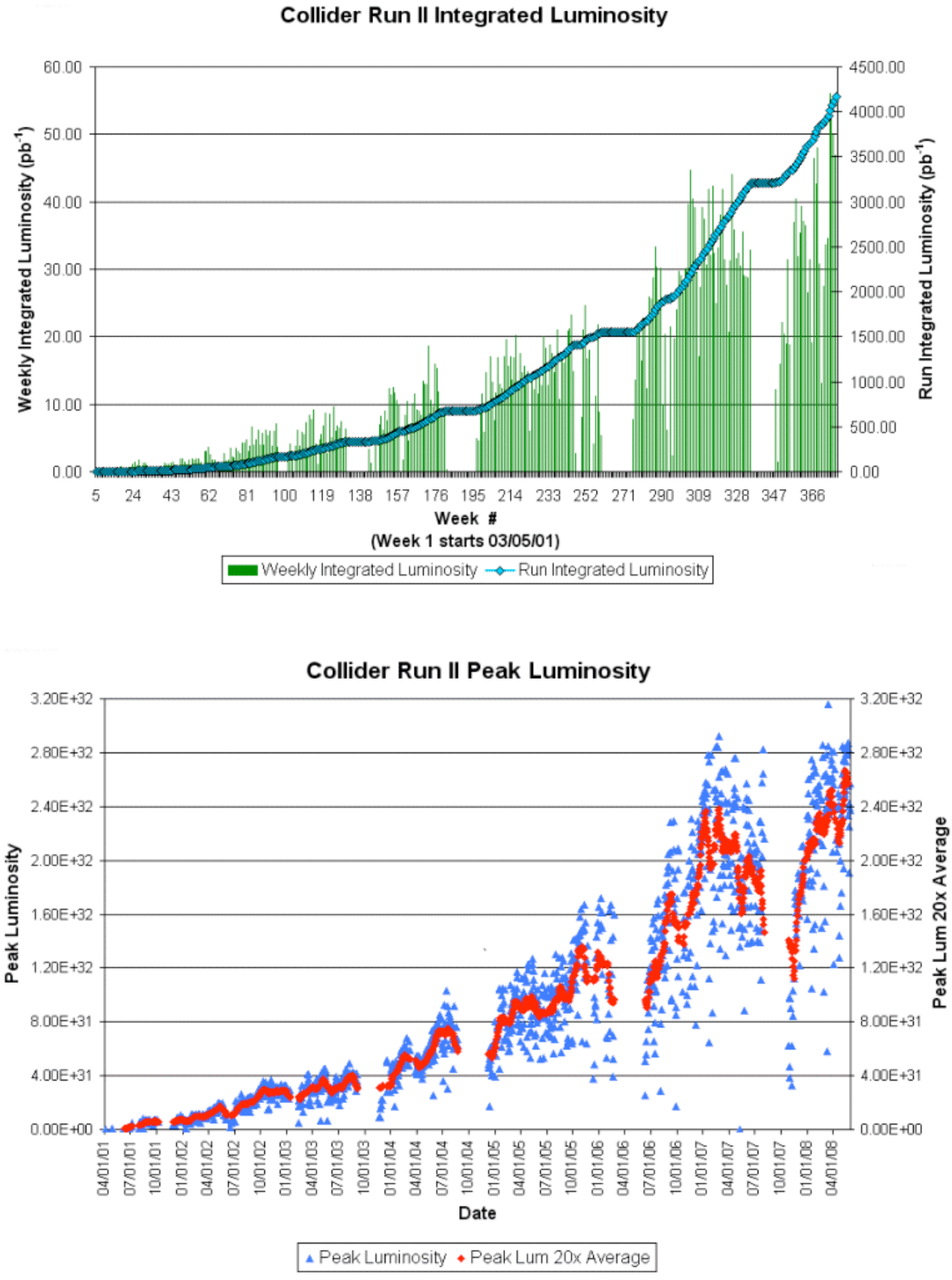


Figure 2.2: The integrated (top) and peak (bottom) luminosity delivered by the Tevatron over time. The bottom figure also shows the peak luminosity (times 20) averaged over 20 bins. Taken from [23].

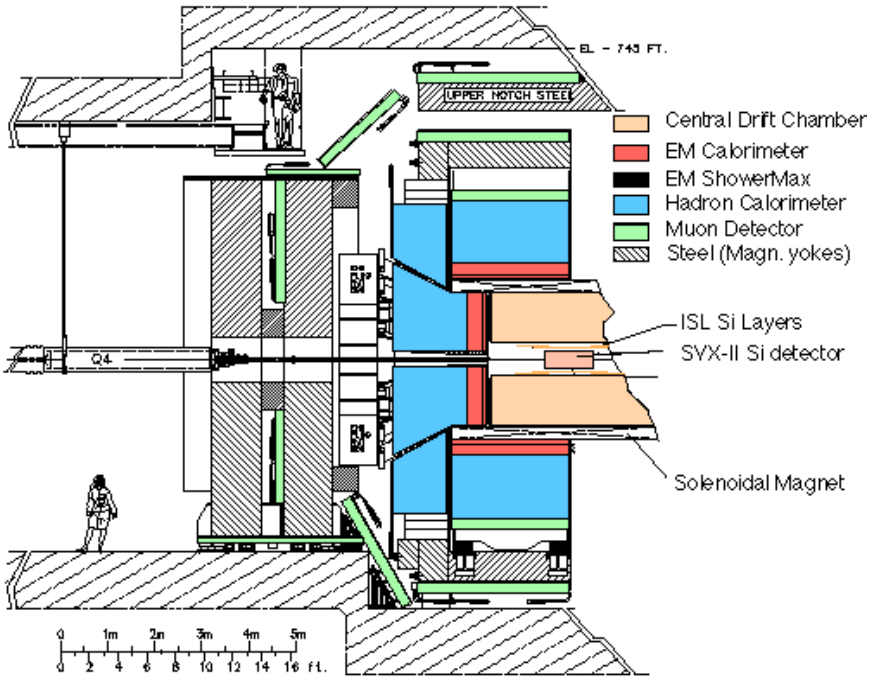


Figure 2.3: Cross-section of one half of the CDF detector [26].

points away from the centre of the Tevatron. The geometry of the detector lends itself to description using a cylindrical co-ordinate system (r, ϕ, z). Reference is often made of the polar angle, θ, and pseudorapidity, η, which is defined as

$$\eta \equiv -\log\left(\tan\frac{\theta}{2}\right). \tag{2.1}$$

The transverse plane is defined as the plane perpendicular to the z direction.

2.4 Tracking systems

The tracking systems at CDF provide precision 3-dimensional measurements of the tracks of charged particles as well as precision measurements of the displacement of tracks from the interaction point. The tracking system, shown in figure 2.4, consists of four sub-systems: an open cell drift chamber outside three silicon microstrip vertex detector systems. Surrounding the four sub-detectors is a 5 m long super-conducting

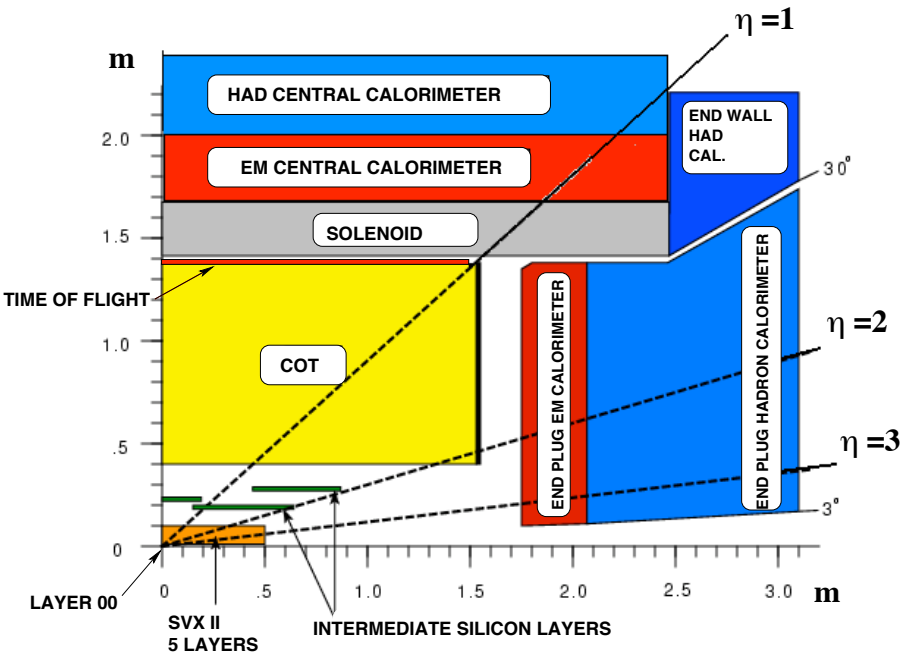


Figure 2.4: Elevation of one quadrant of the CDF tracking volume. Part of the calorimeter systems are also shown (the hadronic central calorimeter extends beyond the range of this diagram). The muon chambers are not shown. (Image adapted from [27].)

solenoid which provides a 1.4 Tesla magnetic field within the tracking volume.

The drift chamber (the Central Outer Tracker, COT) measures the trajectory of charged particles giving measurements of the curvature, azimuth and pseudo-rapidity of a particle's track in the central region. The silicon microvertex detectors extend the tracking capabilities towards the beam pipe and allow identification of track vertices displaced from the interaction point. The three silicon systems are Layer 00 (L00), the Silicon VerteX detector (SVX) and the Intermediate Silicon Layers (ISL).

Charged particles are detected through the ionisation of either the silicon in the microstrip detectors or of the gas in the drift chamber allowing their trajectory to be observed. The curvature of the tracks caused by the presence of the magnetic field is proportional to the transverse momentum of the particles, p_{T}:

$$p_{\mathrm{T}} = 0.3BR_c \text{ (GeV/c)}, \tag{2.2}$$

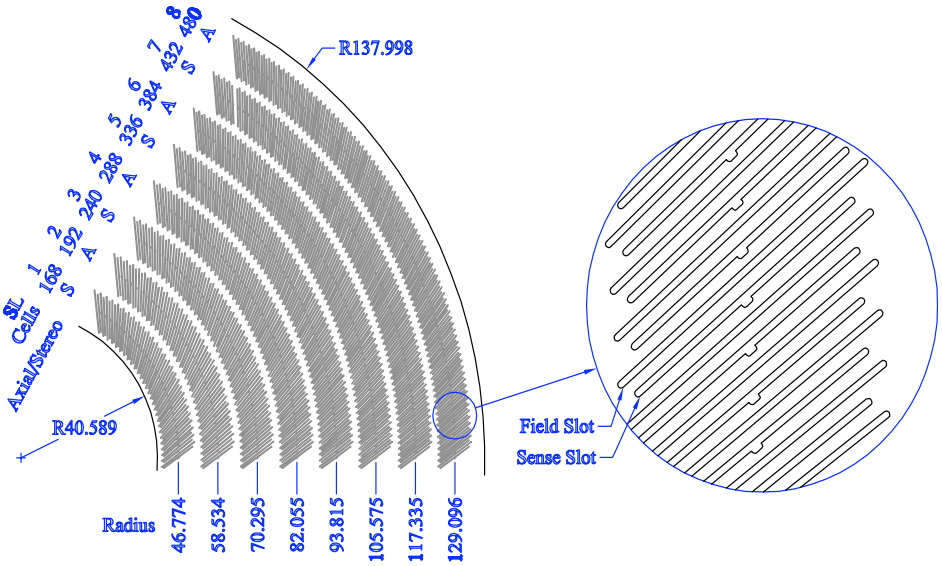


Figure 2.5: A 1/6 section of the COT end plate showing the 8 super-layers and cell layout. For each super-layer (SL) is given the number of cells, whether the wires are axial or stereo and the average radius (in cm). The enlargement shows where the wire planes (sense) and field sheet (field) are installed. (Reproduced from [27].)

where B is the magnetic field strength in Tesla and R_c the radius of curvature in meters. The precision of the p_{T} measurement from the tracking systems decreases as the p_{T} increases because R_c becomes harder to measure.

In the region $|\eta| \leq 1.0$ the combined tracking system provides 7 position measurements in the silicon detectors (1 in r-ϕ and 6 in r-ϕ-z) and 96 measurements in the COT (48 in r-ϕ and 48 in r-ϕ-z) throughout a radius $1.6 \leq r \leq 132$ cm. In the forward and background regions ($1 \leq |\eta| \leq 2$) The silicon systems provide 8 measurements (1 in r-ϕ and 7 in r-ϕ-z) across a radius of $1.6 \leq r \leq 29$ cm with additional measurements from part of the depth of the COT.

2.4.1 The Central Outer Tracker (COT)

The COT is a 3 m long open cell drift chamber consisting of 96 radial layers of sense wires [27]. Its active region covers a radius $40 \leq r \leq 132$ cm and it can provide a measurement in each layer for $|\eta| \leq 1.0$. The COT can be broken down into eight super-layers each of which can provide position measurements at 12 radii. There are

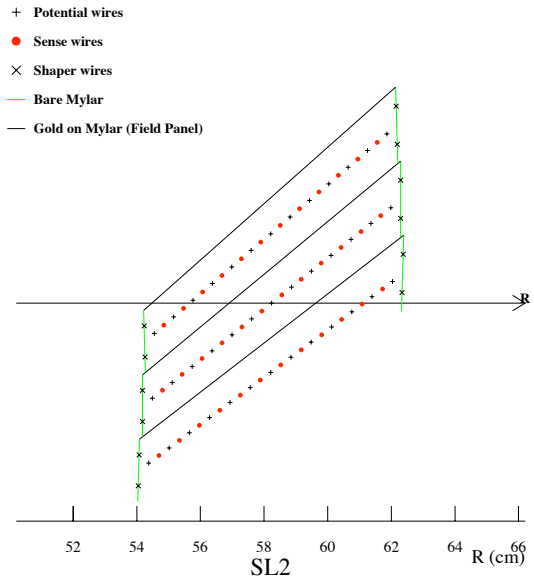


Figure 2.6: An axial cross-section of three cells in super-layer 2 showing the positions of the sense and potential wires and the field panels and shaper panels (bare mylar with shaper wires) [26]. The radial direction is marked by the arrow, **R**.

four axial super-layers alternated radially with four stereo super-layers. In the axial layers the wires are parallel with the beam line providing measurements in the r-ϕ plane. The wires in the stereo layers are angled alternately at $+3^\circ$ and -3° from the axial layers. Combined readout from the axial and stereo layers provides measurements in r-z of a particle moving through the chamber. The super-layer structure can be seen in figure 2.5 which shows the end plate for a 1/6 section of the COT. Each super-layer is divided in ϕ into open drift cells which each contain 25 wires: 13 potential wires alternating with 12 sense wires. The cell is closed in the azimuthal direction by field panels and they are closed in the radial direction both mechanically and electrostaticly by shaper panels. Figure 2.6 is a schematic of three cells from super layer 2. Dividing the chamber into cells means a broken wire would be contained within a single cell. The cells also limit the drift distance to 0.9 cm. The tilted geometry of the cells assists the drift velocity calibrations since every high p_{T}, radial track is sampled over the full range of drift distances within each super layer. The tilted geometry also removes the inherent left-right ambiguity since the fake track will be rotated by a large azimuthal angle and will not be found by pattern recognition. The tracking volume is

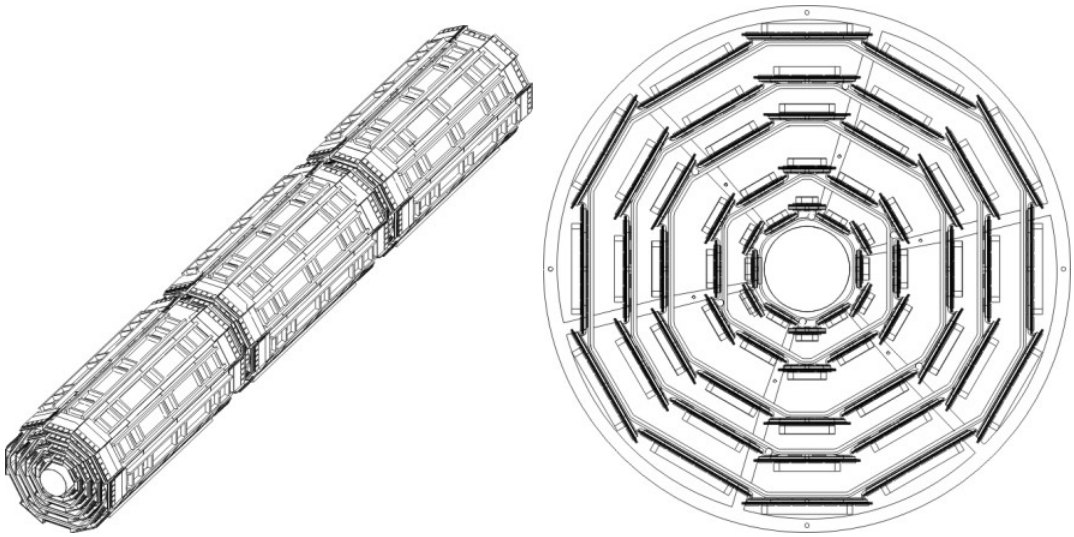


Figure 2.7: The SVX detector [26]. Left: showing the 3 barrel structure. Right: viewed in the r-ϕ plane. Each barrel is made up of five layers of silicon vertex microstrip detectors arranged in 12 ladders in ϕ.

filled with a gas which is a mixture of 50% argon, 50% ethane with a small amount of alcohol (1.7%) added. The COT provides a p_{T} dependent momentum resolution of $\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} \sim 0.15\%$ (GeV/c).

2.4.2 Particle identification using the Central Outer Tracker

The COT is optimised for tracking; however, the specific ionisation ($\frac{\mathrm{d}E}{\mathrm{d}x}$) of charged particles in the COT may be used for particle identification. The COT samples the amount of ionisation produced by a charged particle by measuring the time a pulse induced on a sense is over threshold. An 80% truncated mean of the $\frac{\mathrm{d}E}{\mathrm{d}x}$ samplings over the length of the particle track is used to estimate the average ionisation. The COT provides a 1.4σ separation between kaons and pions for particles with $2 < p_{\mathrm{T}} < 10$ GeV.

2.4.3 The Silicon VerteX detector (SVX)

The silicon vertex detector is made up of three cylindrical barrels with a combined length of 96 cm and covers the region $|\eta| < 2$ [28]. Each barrel consists of five layers of double-sided silicon microstrip detectors arranged in 12 wedges in ϕ as shown in figure 2.7. The radial position of each layer alternates from wedge to wedge within each layer. The layers are located radially at $r = 2.45(3.0)$, $4.1(4.5)$, $6.5(7.0)$, $8.2(8.7)$, $10.1(10.6)$ cm where the first number is the radius of the inner wedges and the bracketed number the radius of the outer wedges in each layer. One side of the microstrip detectors is axial, parallel to the z-axis, providing measurements of the r-ϕ position of charged particles. The other side is stereo. Layers 0, 1 and 3 are perpendicular to the z-axis providing measurements of position in z (90° r-z measurements). Layers 2 and 4 are at -1.2° and $+1.2^\circ$ to the z axis respectively, providing small angle stereo r-z measurements. These allow matching between the stereo measurements in layers 0, 1 and 3 and the axial measurements. The axial strips have implant widths of 14 to 15 μm and are spaced in r-ϕ by 60 to 65 μm depending on the layer. The stereo layers have, from the innermost layer out, a read-out pitch of 141, 125.5, 60, 141 and 65 μm and have implant widths of 20 μm for the 90° strips and 15 μm for the small angle stereo strips. Due to the high occupancy in the SVX particle tracks are reconstructed in the SVX using tracks measured in the COT as a starting point. The intrinsic impact parameter resolution of the SVX (as a function of p_{T}) without L00 is $\sigma_{d_0} = 9 \oplus \frac{50}{p_{\mathrm{T}}}$ μm.

2.4.4 Layer 00 (L00)

Layer 00 is an 80 cm long layer of light weight, radiation hard, single-sided silicon wafers mechanically attached to the beam pipe at a radius of 1.35 cm (see figure 2.8) [29]. L00 adds an additional position measurement to the silicon systems closer to the beam pipe than the SVX and recovers the degradation in resolution of vertex reconstruction due to multiple scattering on the SVX readout electronics and cooling system installed

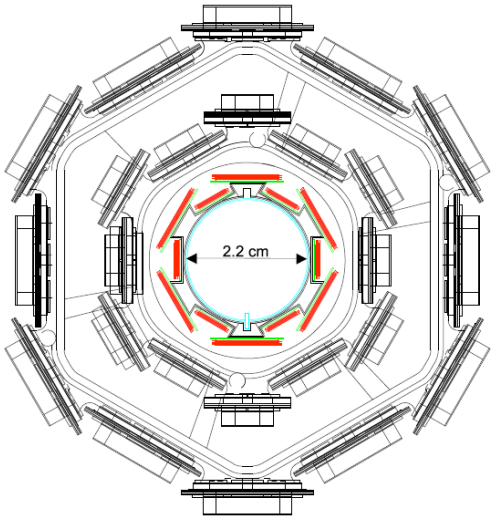


Figure 2.8: A schematic of L00 in the r-ϕ plane. L00 is shown in green and red along with the first two layers of the SVX. (Reproduced from [28].)

within the tracking volume. It will also extend the lifetime of the silicon systems as the SVX degrades due to integrated radiation damage. The SVX and L00 provide a combined intrinsic impact parameter resolution of $6 \oplus \frac{25}{p_{\mathrm{T}}}$ μm.

2.4.5 The Intermediate Silicon Layers (ISL)

The ISL consists of one central layer at a radius of 22 cm covering $|\eta| \leq 1.0$ and 2 forward layers at radii of 20 and 28 cm covering $1.0 \leq |\eta| \leq 2.0$ [30]. Each layer consists of double-sided, small angle stereo microstrip detectors. Both the axial and 1.2° stereo sides have a fixed strip pitch of 112 μm. Figure 2.9 shows schematics of the ISL with the rest of the silicon systems in the r-ϕ and r-z planes. The ISL is divided in ϕ into 12 wedges matching the SVX. The ISL provides intermediate measurements between the SVX and the COT and supplements the COT particularly for $1.0 \leq |\eta| \leq 2.0$ where the full depth of the COT is no longer available.

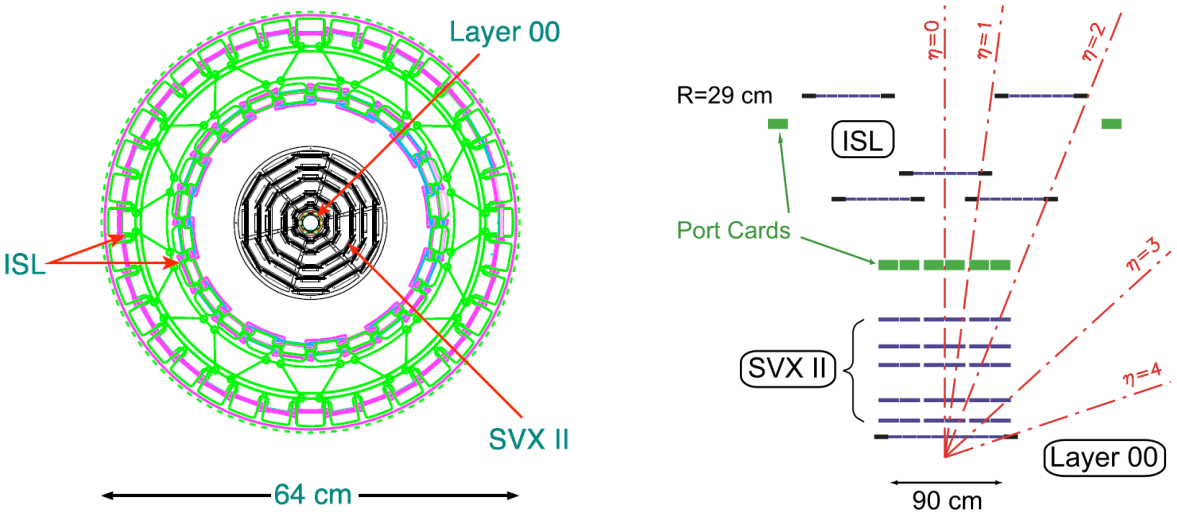


Figure 2.9: A schematic of all the silicon systems: left, in the r-ϕ plane; right, a side view (note the scale of the z-axis has been greatly compressed). L00, the SVX and the ISL are labelled. (Reproduced from [28].)

2.5 The Time Of Flight detector (TOF)

The Time Of Flight detector (TOF) indicated in figure 2.4 was designed specifically to provide particle identification for low momentum particles [31]. The mass of a charged particle is given by

$$m = \frac{p}{c}\sqrt{\frac{c^2t^2}{L^2} - 1} \tag{2.3}$$

where the particle momentum (p) and the distance travelled by the particle from the interaction point to the TOF (L) are measured precisely by the tracking systems. The time, t, taken for the particle to travel the path length, L, is determined by measuring the difference between the arrival time at the TOF and the bunch crossing time. The TOF consists of a cylindrical array of 216 scintillating bars orientated parallel to the beam line between the COT and the cryostat of the solenoid at an average radius of 140 cm. The polar coverage of the TOF is $|\eta| < 1$. As a charged particle passes through a scintillating bar the pulse of light produced is detected at both ends of the bar by photo-multipliers. This allows estimation of the z position of the particle hit in the bar. The interaction time is measured to an accuracy of 30 ps by combining measurements from all the photo-multipliers. The resolution of the particle travel time

measured by the TOF is 110 ps providing 2σ separation of pions and kaons up to $p_{\mathrm{T}} = 1.6$ GeV.

2.6 Calorimetry

Surrounding the solenoid and in the forward and background regions are electromagnetic and hadronic calorimeters. The calorimeters, consisting of alternating layers of passive absorber and plastic scintillator, use a shower sampling technique to determine the energy of incident particles. Particles with p_{T} greater than about 350 MeV/c reach the calorimeters and interact with material in the absorbers losing energy and producing daughter particles which in turn interact with the absorber material producing a cascade of particles. The shower of particles produced leads to a signal detected in the scintillators roughly proportional to the number of particles in the shower. Most particles (though not muons) lose all of their energy within the calorimeters and the sum of the signals sampled in the scintillators is proportional to the energy of the incident particles. Combined, the calorimeters provide a polar coverage of $|\eta| \leq 3.6$.

The calorimeters are segmented in solid angle around the nominal interaction point. This segmentation is organised into 'towers'. Each tower is of a truncated pyramid shape with the imaginary vertex pointing towards the nominal interaction point. The base is a rectangular cell in η-ϕ space. The Calorimeter is segmented radially into an inner, electromagnetic calorimeter and an outer, hadronic calorimeter. The electromagnetic calorimeters consist of layers of lead absorber and polystyrene scintillator. To maintain a constant thickness in radiation length some of the lead layers in the central region are replaced with increasing amounts of acrylic as a function of η. The hadronic calorimeters consist of layers of steel absorber with acrylic scintillator. Light from the scintillators is collected by acrylic wavelength shifters and guided to photomultiplier tubes. Electrons and photons can be distinguished from hadrons as they lose most of their energy in the electromagnetic calorimeters while hadrons lose most

of their energy in the hadronic calorimeters. There are several calorimeter subsystems (which can be seen in figure 2.4): the electromagnetic and hadronic central calorimeters; the end wall hadron calorimeter [32, 33] and the end plug electromagnetic and hadronic calorimeters [34] (one of each positioned in both positive and negative η). They each work in similar ways. The electromagnetic calorimeters provides a relative energy resolution, as a function of energy, of $\sigma_E/E = 13.5\%/\sqrt{E\sin(\theta)} \oplus 2\%$ for the central calorimeter and $\sigma_E/E = 16\%/\sqrt{E\sin(\theta)} \oplus 1\%$ for the plug calorimeters. The hadronic calorimeters provide a relative energy resolution, as a function of energy, of $\sigma_E/E = 50\%/\sqrt{E\sin(\theta)} \oplus 3\%$ for the central, $\sigma_E/E = 75\%/\sqrt{E\sin(\theta)} \oplus 4\%$ for the end wall and $\sigma_E/E = 74\%/\sqrt{E\sin(\theta)} \oplus 4\%$ for the plug calorimeters.

2.7 Muon detectors

The muon systems are the final layer of the CDF detector. Charged particles other than muons are stopped by a combination of the tracking systems, magnet return yoke, calorimeters and additional shielding. Charged particles reaching the muon systems have a high probability to be muons. The muon detectors are made up of a combination of four independent sub-systems whose positions are indicated in figure 2.4. Each are combinations of drift tubes, scintillation counters and absorbers. Single wire, rectangular drift chambers are arranged in arrays with different segmentation in ϕ. Stacks of up to eight layers are overlaid in ϕ to allow coincident particle hits. The chambers are staggered in ϕ to remove ambiguity of the ϕ position. The difference in drift arrival times between neighbouring cells gives a 250 μm position resolution in r-ϕ. Division of charge measured at either end of the sense wires allow measurement of the z position with 1.2 mm resolution. Time measurements from the scintillators are used to associate particle hits in the drift tubes to the correct bunch crossing (the maximum drift time exceeds the inter-bunch crossing time). They are also used to remove background due to secondary interactions in the beam pipe and cosmic rays. A track segment from 3

	CMU	CMP	CMX	IMU
Polar coverage	$\|\eta\| < 0.6$	$\|\eta\| < 0.6$	$0.6 < \|\eta\| < 1.0$	$1 < \|\eta\| < 1.5$
Azimuthal coverage [degrees]	302°	360°	360°	270°
Minimum $p_{\mathrm{T}}(\mu)$ [GeV/c]	1.4	2.2	1.4	1.4 - 2.2

Table 2.1: Design parameters of the muon detectors.

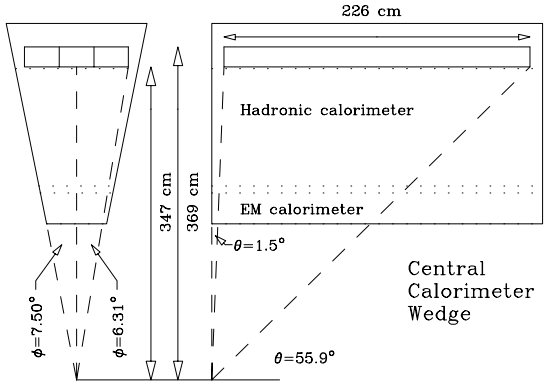


Figure 2.10: Three sections of the CMU embedded at the outer radius of a calorimeter tower [26].

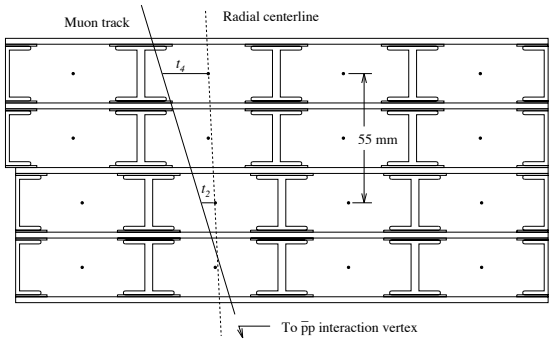


Figure 2.11: One of the three CMU sections in a wedge [26].

matching radial layers in the muon detectors is matched to a track extrapolated out from the COT to give a muon candidate. The four muon systems are listed below. Together they cover a polar range of $|\eta| < 1.5$. The properties of each of the systems are summarised in table 2.1. Further information can be found in [35].

The central muon detector (CMU)

The Central MUon detector (CMU) is outside of the central hadronic calorimeter at a radius of 347 cm. It covers a polar range of $0.03 < |\eta| < 0.63$. The position of the CMU in relation to the calorimeters is shown in figure 2.10. Figure 2.11 shows a cross-section of one section of the CMU.

The central muon upgrade (CMP)

The Central Muon uPgrade (CMP) is located behind an additional 60 cm of steel enclosing the central detector. The CMP covers the gaps in ϕ of the CMU and provides

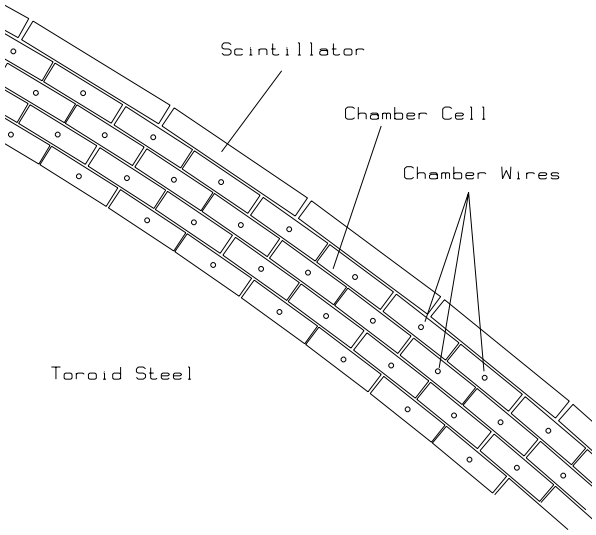


Figure 2.12: Part of the IMU in the r-ϕ plane [26].

enhanced rejection of penetrating high energy hadrons to 1% of total pions and 2-4% of total kaons. Since it covers the same polar range as the CMU (see table 2.1) the CMP does not contain any scintillators. Instead z position information for both the CMP and CMU is obtained from the scintillators in the CMU.

The central muon extension (CMX)

The Central Muon eXtension (CMX) is an arch shaped arrangement of drift tubes and scintillation chambers at a radius of 400 to 600 cm. It extends the polar muon coverage to $0.6 < |\eta| < 1$.

The intermediate muon system (IMU)

The Intermediate MUon system (IMU) extends the polar coverage of the muon systems to $1 < |\eta| < 1.5$ however the ϕ coverage of the IMU is limited by the presence of support structures. Figure 2.12 shows the arrangement of the drift cells and scintillators in the r-ϕ plane for a section of the IMU.

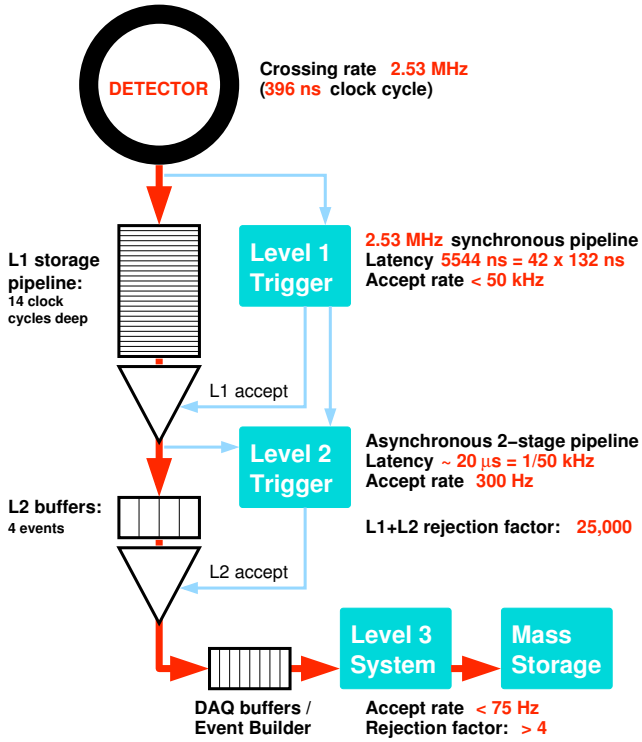


Figure 2.13: Flow diagram of the CDF trigger [26].

2.8 The trigger systems

The Tevatron delivers events to CDF at a crossing rate of 2.53 MHz with bunch crossings separated by 396 ns. This rate is 25300 times greater than the 100 Hz which can be written to tape. Most of the collisions do not contain interesting events. It is necessary to reduce the event rate to the amount which can be written to tape while preserving as many of the interesting events as possible. CDF achieves this using a trigger to select events to be read out from the detector and stored.

The high event rate requires a fast trigger decision which means only simple objects from part of the detector read-out may be used. At a lower input rate, time is available to reconstruct more of the event to make a more sophisticated decision. For this reason the trigger is implemented at 3 levels. A flow diagram of the trigger system is shown in figure 2.13.

Level 1 is a synchronous hardware trigger. It bases a decision to keep an event on

simple objects from the detector. The processing time for level 1 is 5.5 μs (compare this to the 390 ns between bunch crossings). During level 1 processing, event information from each readout channel is stored in a pipeline for 5.5 μs. If level 1 decides to keep an event then at the end of the pipeline it is sent to the level 2 trigger, otherwise the event is lost. Level 1 reduces the event rate from the input 2.53 MHz to an accept rate of less than 50 kHz.

Level 2 is an asynchronous trigger which is partly hardware and mostly software based. The lowered input rate allows time to use greater information to make a more realistic reconstruction of the event. When level 1 accepts an event each readout channel passes the event information from the level 1 pipeline to one of four level 2 buffers where it is stored until level 2 has made a decision about the event. If all four buffers are full when level one accepts an event then the event is lost. This is known as dead-time and the achieved goal of level 2 is for less than 5% dead-time. Level 2 uses 20 μs processing time and has an accept rate of 300 Hz. A decision to accept an event at level 2 triggers read out of the entire detector to be sent to the level 3 trigger.

Level 3 is an entirely software trigger. Full information from the detector is sent to a processor farm of commercial PCs. At this stage nearly full event reconstruction can be performed using a version of the offline reconstruction code. Since the reconstruction is the same as offline it will be discussed further in the next chapter. The output rate of level 3 is $\sim$ 100 Hz which is limited by how fast the events can be written to tape.

The level 1 and 2 triggers are made up of a number of subsystems each of which generate "primitives", simple objects, to base the trigger decision on. At level 1 the objects include tracks from the eXtremely Fast Tracker (XFT, described below), information from single calorimeter towers, track segments from the muon chambers and the sum of the energy in the transverse plane. Level 2, in addition to level 1 information, uses the Silicon Vertex Tracker (SVT) to look for a displaced vertex, has information from clusters of trigger towers and has higher precision tracking.

This analysis uses data selected by the XFT and SVT. The other trigger elements select events containing leptons or jets (clusters of tracks within a cone in solid angle). The ability of CDF to perform fast tracking as part of the trigger decision in particular its ability to trigger on a displaced vertex gives it the unique ability (before the Large Hadron Collider starts) to select hadronic decays at a hadron collider.

2.8.1 The Extremely Fast Tracker

At trigger level 1 the eXtremely Fast Tracker (XFT) identifies tracks in the central outer tracker (COT) [36]. The XFT consists of an axial system and from 2006 a stereo system. The axial system identifies 2-dimensional tracks and is made up of the Finder and the Linker. The Finder looks for track segments in each of the four, axial, super-layers by matching the pattern of signals in the COT to a predefined set of patterns. These track segments are sent to the Linker which also uses pattern matching to find valid tracks which correspond to track segments in three or four axial super-layers. The output from the linker is divided into 288 slices in ϕ for each of which $\Delta\phi = 1.25°$. For each slice the linker returns 1 track corresponding to segments in all four axial superlayers where possible or in the three innermost axial layers if not. In each slice the Linker outputs p_{T} and ϕ for the best track. The XFT identifies tracks with a $p_{\mathrm{T}} > 1.5$ GeV/c with an efficiency greater than 96%. It has a resolution $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 1.7\%$ $(\mathrm{GeV/c})^{-1}$ and $\sigma_{\phi} = 5$ mrad. The XFT allows selection of events containing tracks with a p_{T} above some threshold at level 1.

During 2006 the XFT was upgraded with a stereo system to allow 3 dimensional tracking at level 1 [37], however the data used for this analysis were collected before this upgrade was available.

2.8.2 The Silicon Vertex Trigger

At level 2 the Silicon Vertex Trigger (SVT) provides measurements of impact parameter, azimuth and transverse momentum of track candidates in the transverse plane [38, 39]. It takes as input the ϕ and p_{T} of candidate tracks from the XFT and the digitised hits from the SVX detector. The SVT reconstructs tracks in 4 stages. First, clusters of hits in each silicon layer are identified and combined into "super-hits". Next pattern matching is used to associate these super-hits into low-resolution candidate tracks, or "roads". XFT tracks can be extrapolated to the outer edge of the silicon and binned with a width of 3 mm to be treated as an additional hit for pattern matching. The third step takes the roads and associated full resolution hits and fits using a linear regression for the track parameters and fit chi-square using the coarse roads as a starting point. The final stage corrects the track parameters for the beam position.

The beam position is monitored by the SVT. If the beam moves from its nominal position then nearly all tracks will have an impact parameter with respect to the nominal beam position. The impact parameter distribution as a function of ϕ is sinusoidal and the parameters of the sinusoid relate directly to the beam position. Tracks are sampled from the SVT at a rate of 10^7 tracks per hour and an updated beam position is calculated every 30 s. This beam position is used to correct the SVT measured track parameters in real time. The beam position typically moves by tens of microns in the transverse plane during a Tevatron store.

The SVT provides a resolution on the transverse impact parameter of 35 μm increasing to 50 μm when convoluted with the beam spot resolution. The SVT has an efficiency of 85% for tracks with $p_{\mathrm{T}} > 1.5$ GeV/c that travel the complete depth of the SVX. This efficiency rises to 90% for tracks with $p_{\mathrm{T}} > 2.0$ GeV/c. The SVT allows the level 2 trigger to select events containing tracks with some minimum impact parameter or a vertex corresponding to a long decay length.

2.8.3 The B_PIPI trigger path

A particular combination of trigger requirements at level 1, level 2 and level 3 is known as a trigger path. Many trigger paths are active while CDF is taking data, each tuned to select specific physics signatures such as events containing muons or jets. This analysis uses data collected using the displaced-tracks trigger, in particular the B_PIPI and B_PIPI_HIGHPT trigger paths. Since $B^0_{(s)} \to h^+h'^-$ decays contain no leptons in the final state it is necessary to use tracking to make the trigger decision. B mesons are long lived and can travel ∼1 cm in the detector before decaying. The tracks of the B decay products will, therefore, typically have a significant impact parameter with respect to the interaction point. A displaced decay vertex and two tracks with large impact parameters is the signature used by the B_PIPI trigger paths. The requirements at each trigger level are described below.

Level 1

At level 1 the B_PIPI trigger relies on the XFT so has information available on the $p_{\rm T}$, ϕ and charge of tracks. The requirement is for two tracks each with $p_{\rm T} > 2.04$ GeV/c and with summed transverse momentum, $\sum p_{\rm T} > 5.5$ GeV/c. The difference in azimuthal angle of the two tracks is required to be $0^\circ < \Delta\phi < 135^\circ$. This reduces track pairs from light-quark fragmentation in back-to-back jets. It is also required that the two tracks have opposite charges. At high instantaneous luminosities ($\mathcal{L} > 50 \times 10^{30}$ cm^{-2}s^{-1}) the B_PIPI trigger requirements accept events at a rate which produces dead-time greater that the design limit of 5%. At these high luminosities an alternative trigger path, B_PIPI_HIGHPT, is employed with stricter requirements to reduce the accept rate. While the tighter requirements will reject some interesting events the purity of the sample will be higher than that collected with looser requirements. The tighter requirements are two tracks with $p_{\rm T} > 2.46$ GeV/c and $\sum p_{\rm T} > 6.5$ GeV/c. At the highest Tevatron luminosities even the B_PIPI_HIGHPT trigger rate produces greater

than 5% dead time and the trigger accept rate is reduced further by dynamic prescaling, which is to say the accept rate is reduced by randomly rejecting some fraction of the accepted events. It is dynamic because the fraction to be rejected changes with the instantaneous luminosity.

Level 2

At level 2, SVT information becomes available and so track impact parameters and two track vertex positions become available. The level 2 B_PIPI trigger requires two tracks each with $p_{\mathrm{T}} > 2.0$ GeV/c and $\sum p_{\mathrm{T}} > 5.5$ GeV/c and each with impact parameters in the range 100 μm $< d_0 <$ 1000 μm. The minimum impact parameter cut rejects tracks from the interaction point, vastly reducing the event rate. The maximum impact parameter cut is due to hardware constraints as the SVT pattern matching efficiency drops rapidly beyond this point (although recent upgrades to the SVT have extended its efficiency closer to 1200 μm). The azimuthal opening angle requirement between the two tracks is reduced from the level 1 requirement to be $20^\circ < \Delta\phi < 135^\circ$. The two track vertex is required to have a minimum transverse distance from the interaction point of $L_{\mathrm{T}} > 200$ μm which preferentially selects the decay products of particles that have travelled some distance in the detector. There is also a requirement on the goodness of the track fit with the minimal linearised-fit quality $\chi^2_{SVT} < 15$ [40]. The two trigger tracks can be combined to form a B meson candidate assuming it decayed to two bodies. The impact parameter of the candidate B is required to be $|d_0(\mathrm{B})| < 140$ μm to impose that the candidate came from the primary interaction point.

The B_PIPI_HIGHPT trigger at level 2 has the same requirements as the B_PIPI trigger path except with the tightened requirements: $p_{\mathrm{T}} > 2.5$ GeV/c and $\sum p_{\mathrm{T}} > 6.5$ GeV/c.

Quantity	Units	level 1	level 2	level 3
track p_{T}	GeV/c	>2.04 (2.46)	>2.0 (2.5)	>2.0 (2.5)
$\sum p_{\mathrm{T}}$	GeV/c	>5.5 (6.5)	>5.5 (6.5)	>5.5 (6.5)
$\Delta\phi$	Degrees	[0, 135]	[20, 135]	[20, 135]
track $\lvert d_0\rvert$	μm	—	[100, 1000]	[100, 1000]
$\lvert L_{\mathrm{T}}(\mathrm{B})\rvert$	μm	—	>200	>200
$\lvert d_0(\mathrm{B})\rvert$	μm	—	—	<140
track $\lvert\eta\rvert$	—	—	—	<1.2
$\lvert\Delta z\rvert$	μm	—	—	<5.0
$M_{\pi\pi}$	GeV/c^2	—	—	[4.0, 7.0]

Table 2.2: Requirements of the B_PIPI (B_PIPI_HIGHPT) trigger.

Level 3

Level 3 has available the full event reconstruction. The level 2 requirements are reconfirmed at level 3 using the full event information. A requirement on the two track invariant mass assuming both particles are pions, $4.0 < M_{\pi\pi} < 7.0$ GeV/c^2, is imposed to reduce the level 3 accept-rate while retaining events for background studies. The longitudinal separation of the two tracks at the point of their minimum distance from the beam is required to be $|\Delta z| < 5.0$ cm, rejecting track pairs originating from different interaction points. A requirement of $|\eta| < 1.2$ rejects tracks which are not within the XFT fiducial region. The transverse decay length requirement is now on the absolute value, $|L_{\mathrm{T}}| > 200$ μm. The trigger requirements at all three levels are summarised in table 2.2.

2.9 Summary

$\mathrm{B_s^0}$ mesons are produced in collisions of protons and antiprotons at the Tevatron particle accelerator. CDF detects the products of these collisions with a combination of tracking detectors, calorimeters and muon detectors. This analysis uses the tracking systems which comprises a drift chamber (the COT) and silicon detectors. These tracking systems provide particle trajectory, momentum and decay vertex measure-

ments, necessary for B meson lifetime measurements. Only a small fraction of the events delivered to CDF can be stored and a trigger is employed to retain potentially interesting events. The trigger is implemented in three levels, each level significantly reducing the event rate to allow time for more sophisticated event reconstruction. differing reconstruction algorithms mean that the level 2 and level 3 trigger levels make different measurements of some of the same quantities, for example the impact parameter measured by the SVT at level 2 is different from the impact parameter measured by the nearly full event reconstruction at level 3. The B_PIPI and B_PIPI_HIGHPT trigger paths are used to select $B^0_{(s)} \to h^+h'^-$ candidate events. These triggers look for the signatures of a decay of a long lived particle: events containing two particle tracks with large impact parameters and with a two track vertex which is displaced from the interaction point. Selecting events based (indirectly) on their lifetime will affect the distribution of lifetimes in the data sample. This effect and how to deal with it when measuring a lifetime is the subject of chapter 5. The data selected by the trigger requires further processing and this is the subject of the next chapter.

Chapter 3

Data reconstruction and event simulation

3.1 Introduction

Proton-antiproton collision events at CDF that potentially contain $B^0_{(s)} \to h^+h'^-$ candidate decays are selected by the B_PIPI and B_PIPI_HIGHPT triggers and stored for offline analysis. Section 3.3 describes how particle tracks are reconstructed and candidate $B^0_{(s)} \to h^+h'^-$ events are identified in the data sample. The sample of candidate $B^0_{(s)} \to h^+h'^-$ decays is dominated by background whose contribution can be significantly reduced using the selection requirements described in section 3.4 to produce the signal rich sample used for this analysis.

In order to test the analysis methods in an unbiassed way (see chapter 7), simulated data samples are required. The two types of simulation used for this analysis are described in this chapter. The first (section 3.5) includes a realistic event and detector simulation, the second (section 3.6) is a fast, custom simulation.

3.2 Data sample

This analysis uses data collected by the B_PIPI and B_PIPI_HIGHPT triggers at CDF between February 2002 and March 2006. The data correspond to an integrated luminosity of approximately 1 fb^{-1}. The triggers impose selection requirements on the data which are summarised in table 2.2 on page 41.

3.3 Event reconstruction

In a uniform magnetic field charged particles follow a helical path about the direction of the magnetic field. At CDF the magnetic field is parallel to the z-direction. The path of a charged particle can be described by five parameters. The parameterisation used here uses three parameters describing a circle in the transverse plane and two describing a straight line. The parameters used are:

curvature C: defined as $C = q/(2r)$ where q is the charge of the particle and r is the radius of the circle formed by projecting the helix onto the transverse plane;

impact parameter $\mathbf{d_0}$: the distance of closest approach to the interaction point. It is defined as $d_0 = (\hat{\mathbf{p}} \wedge \mathbf{d}) \cdot \hat{\mathbf{z}}$ where $\mathbf{d}$ is a vector pointing from the interaction point to the point of closest approach and $\hat{\mathbf{p}}$ is a unit vector in the direction of the particles momentum. d_0 is a signed quantity whose sign is determined by the scalar product with a unit vector along the z direction, $\hat{\mathbf{z}}$;

azimuthal angle $\boldsymbol{\phi_0}$: the azimuthal angle of the particle at the distance of closest approach to the z-axis;

the helix pitch $\boldsymbol{\cot(\theta)}$: the cotangent of the angle θ between the z-axis and the momentum of the particle;

$\mathbf{z_0}$: the z-position of the point of closest approach.

Track reconstruction begins in the COT by searching for three aligned axial hits in consecutive layers to form a track segment. These track segments are fitted in two-dimensions with a straight line which is extrapolated to the rest of the super-layer. Hits which are close to the track are added and all the points are then refitted with a straight line. Starting with the outermost super-layer, track segments are extrapolated towards the beam pipe assuming they have zero impact parameter. This defines a region to look for segments in the inner super-layers. Segments within this region that form a tangent to a common circle are linked to form a track candidate and the hits within these segments are fitted to a circle to form a 2-dimensional track. Stereo segments, or if there are none present, stereo hits, are linked to the 2-dimensional track and a helical fit is performed.

The next step is to combine the COT track with silicon hits. A COT track candidate is extrapolated to the outside edge of the silicon. A 4σ wide window around the track is calculated based on the errors of the COT track parameters. If a hit is found in the outermost layer of silicon within this window it is added to the track and the track parameters are refitted giving new parameter errors and a new search window in the next layer. This continues through the silicon layers. If no hit is found in a given layer then the search continues into the next layer. This proceeds via three tracking algorithms, OI, OIS and OIZ. The OI (Outside In) algorithm uses only r-ϕ silicon hits as it steps in from the extrapolated COT track candidate. If at least 3 r-ϕ hits are used then the OIS (Outside In Stereo) and OIZ (Outside In Z) algorithms are used. OIS adds small angle stereo, r-z hits to the r-ϕ hits added by OI. The OIZ tracks contain small angle stereo r-z hits and at least 2 90° r-z hits.

Corrections are applied to the track variables to take in to account energy loss within the volume of the tracking detectors and non-uniformities in the magnetic field. These corrections are dependent on the particle hypothesis. As the final states of the $\mathrm{B}^0_{(s)} \to h^+h'^-$ sample contain a mixture of pions, kaons, protons and electrons and the particle type is not known on an event by event basis all particle tracks are arbitrarily

reconstructed using the pion hypothesis.

Misalignments and noise in the SVX and COT lead to fake and misreconstructed tracks. In order to reduce background caused by these kind of tracks standard requirements are placed on the tracks that they should: come from a converged helix fit with a positive error matrix; have at least 5 hits in each of 2 stereo and 2 axial COT super-layers; have hits in at least 3 r-ϕ SVX layers. This analysis also requires that the tracks came from the OIZ tracking algorithm since these have a higher resolution in the r-z plane than tracks coming from the other two algorithms. This significantly reduces contamination from events where two tracks have come from the decays of two heavy flavours which may give them a large impact parameter though they are separated along the z direction.

The invariant 2-particle mass is calculated for all possible combinations of tracks in an event that have passed the trigger requirements and the selection requirements above, and that have opposite curvature (and therefore opposite charge). The invariant mass is calculated assuming both of the particles are pions. The two tracks are constrained by the vertex fit algorithm to originate from a common 3-dimensional vertex.

3.4 Event selection

Work for this thesis is partly based on the work described in [41] measuring the branching ratios and direct CP asymmetries in $\mathrm{B}^0_{(s)} \to h^+h'^-$ decays using the same data sample as used here. This analysis uses the same selection requirements as [41]. This means models of distributions within the data determined for that analysis may be used here unchanged. It also allows for comparison between components common to both analyses.

This analysis also relies on good knowledge of the trigger requirements placed on

tracks and B candidates. The offline reconstruction is different from the event reconstruction used by the SVT and level-3 triggers. This means that a candidate $B^0_{(s)} \to h^+h'^-$ event which didn't meet the level-2 or level-3 requirements might meet the same requirements applied to the offline quantities. In order to remove these candidates the offline reconstructed tracks are matched to both SVT and level 3 tracks and the trigger requirements are re-applied to those quantities. The trigger decision is also reconfirmed on offline tracks.

An offline track pair is matched to an SVT track pair by requiring proximity of the curvature and azimuthal opening angle of the tracks. offline tracks are matched to the level-3 tracks with the closest $p_{\rm T}$ and ϕ by minimising the measure

$$\chi^2 = \frac{(p_{\rm T}^{\rm off} - p_{\rm T}^{\rm L3})^2}{\sigma^2_{p_{\rm T}^{\rm L3}}} + \frac{(\phi^{\rm off} - \phi^{\rm L3})^2}{\sigma^2_{\phi^{\rm L3}}} \tag{3.1}$$

where "L3" and "off" indicate the level-3 and offline quantities and $\sigma^2_{p_{\rm T}^{\rm L3}}$ and $\sigma^2_{\phi^{\rm L3}}$ are the uncertainties on the level-3 $p_{\rm T}$ and ϕ respectively.

A set of selection requirements were optimised to minimise the statistical uncertainty on the measurement of the direct CP asymmetry in $B^0 \to K^+\pi^-$ decays [41]. These requirements are also suitable for measuring observables in the decays of the other high-yield $B^0_{(s)} \to h^+h'^-$ modes: $B^0_s \to K^+K^-$ and $B^0 \to \pi^+\pi^-$. Requirements on the following variables were optimised:

L_T: the transverse distance to the B decay vertex. Requiring a large $L_{\rm T}$ preferentially selects particles which have travelled in the detector and are therefore long-lived;

d_0(B): the impact parameter of the B. Requiring that the candidate B has a small impact parameter rejects candidates where two tracks not coming from the same B decay but with sizeable impact parameters (such as tracks from 2 different long-lived decays, particles that have been scattered by the beam pipe, or misreconstructed tracks) have been combined that do not point back to the interaction

point;

d_0(tracks): the impact parameters of the 2 candidate tracks. Particles originating from the decay of a long lived particle that has travelled some distance in the detector will typically have a significant impact parameter;

I(B): the "isolation" of the B meson. b-hadrons tend to carry a large fraction of the transverse momentum of particles produced in the fragmentation compared to lighter hadrons [42]. The isolation $I(\mathrm{B})$ is used as an estimator of the fraction of the momentum carried by the b-hadron. It is defined as

$$I(\mathrm{B}) = \frac{p_{\mathrm{T}}(\mathrm{B})}{p_{\mathrm{T}}(\mathrm{B}) + \sum\limits_{i \neq j:\mathrm{B}\to j}^{R} p_{\mathrm{T}}(i)} \tag{3.2}$$

where the sum is over all tracks not coming from the decay of the B, and passing the minimum track quality requirements, that lie within a cone in η-ϕ space whose apex is the primary vertex and whose axis is collinear with the B momentum. The cone is chosen to have a unitary radius $R = \sqrt{\phi^2 + \eta^2} = 1$. Where both of the B decay products travel within this cone the isolation is the fraction of transverse momentum within the cone carried by the B meson. Candidates with large isolation are more likely to be B mesons than candidates with low isolation;

χ^2: the minimum χ^2 of the vertex fit minimisation. Requiring a small vertex χ^2 rejects a large amount of background with a high efficiency for signal.

A complete list of selection requirements are given in table 3.1. Figure 3.1 shows the invariant $\pi\pi$-mass distribution for events meeting these selection requirements. 12969 events are retained by these selection requirements. The signal yield and sample composition are determined from a multivariate likelihood fit which will be described in chapter 4. The composition of this sample has been previously determined [41] and is known to include:

property of track	selection requirement
Axial silicon hits	≥ 3
90° r-z silicon hits	≥ 2
Axial COT superlayers (hits per superlayer)	≥ 2 (≥ 5)
Stereo COT superlayers (hits per superlayer)	≥ 2 (≥ 5)
COT hits	≥ 42
Reconstruction algorithm	OIZ
p_{T}	> 2.0 GeV/c
$\lvert\eta\rvert$	< 1.0
$\lvert d_0\rvert$	$[100, 1000]$ μm
property of B candidate	
$q(1) \times q(2)$	$-1e^2$
$d_0(1) \times d_0(2)$	< 0 $(\mu\mathrm{m})^2$
L_{T}	> 300 μm
$\sum p_{\mathrm{T}}$	> 5.5 GeV
$\lvert d_0\rvert$	< 80 μm
I	> 0.5
Decay vertex fit χ^2	< 15
$\lvert\eta\rvert$	< 1.0
$\Delta\phi$	$[20^\circ, 135^\circ]$
$M_{\pi\pi}$	$[5.0, 5.8]$ GeV/c^2

Table 3.1: $\mathrm{B}^0_{(s)} \to h^+h'^-$ candidate selection requirements.

- combinatorial background: this forms a smooth, slowly decreasing mass spectrum across the whole mass range and is composed of random pairs of oppositely charged particles which are displaced from the beam-line and that happen to meet the selection requirements. Its dominant sources include generic QCD background of light quark decays, lepton pairs from Drell-Yan processes, pairs of mismeasured tracks, combinations of a mismeasured track with a track from a heavy flavour decay ($\mathrm{b\bar{b}}$ or $\mathrm{c\bar{c}}$ production) or combinations of tracks from two heavy flavour decays in the event;

- partially reconstructed heavy flavour decays (referred to as physics background): this is the cause of the enhanced mass distribution below ~ 5.15 GeV/c^2. This consists of misreconstructed multi-body b-hadron decays (such as $\mathrm{B}^0 \to \rho^\pm\pi^\mp$, $\mathrm{B}^0 \to \rho^-\mathrm{K}^+$, $\mathrm{B}^0_\mathrm{s} \to \rho^\pm\pi^\mp$, $\mathrm{B}^0_\mathrm{s} \to \rho^+\mathrm{K}^-$ and many others) where only two of the final state particles were reconstructed. This type of background is suppressed

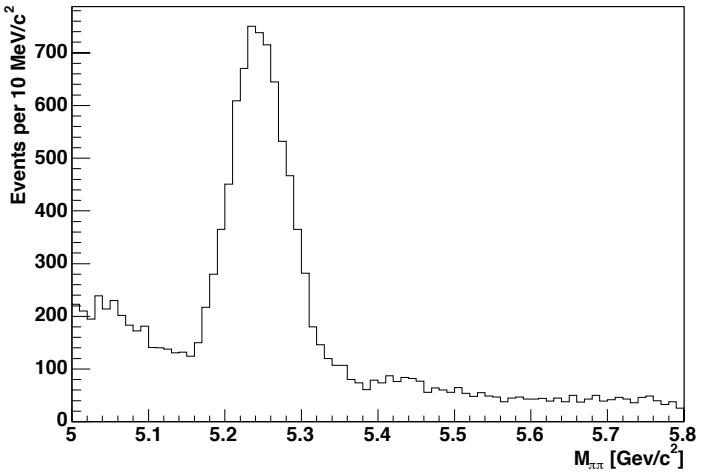


Figure 3.1: Invariant $\pi\pi$-mass distribution of events meeting the selection requirements.

above 5.15 GeV/c^2 because its contribution is kinematically limited to the region $M_{\pi\pi} < M_{\mathrm{B}^0_{(s)}}$;

- $\mathrm{B}^0_{(s)} \to h^+h'^-$decays: The main peak at around 5.2 GeV/c^2 comprises of contributions from the decays $\mathrm{B}^0 \to \mathrm{K}^\pm\pi^\mp$, $\mathrm{B}^0 \to \pi^+\pi^-$, $\mathrm{B}^0_\mathrm{s} \to \mathrm{K}^+\mathrm{K}^-$ and $\mathrm{B}^0_\mathrm{s} \to \mathrm{K}^\pm\pi^\mp$. The mass resolution is such that all these contributions form a single peak in the mass spectrum;

- $\Lambda^0_\mathrm{b} \to \mathrm{p}h^-$decays: these make a small contribution in the mass region just above the main peak and consist of the decays $\Lambda^0_\mathrm{b} \to \mathrm{pK}^-$, $\Lambda^0_\mathrm{b} \to \mathrm{K}^+\overline{\mathrm{p}}$, $\Lambda^0_\mathrm{b} \to \mathrm{p}\pi^-$ and $\overline{\Lambda^0_\mathrm{b}} \to \pi^+\overline{\mathrm{p}}$.

3.5 Realistic simulation

This analysis has been performed without making a measurement of any lifetimes in data until the analysis methods and systematic errors were finalised. Where it has been necessary to test the methods on data the lifetimes have been hidden by adding an unknown random number to the measured lifetime returned by the fits. For this reason it is useful to have a simulated data sample, with known input lifetimes, in order to test the methods. This analysis uses two different data simulations. The first, described

in this section, simulates the detector and trigger response as well as signal event production and shall be referred to as the realistic simulation. The second simulation, described in the next section, is intended to generate large samples of simulated data quickly and does not include a detailed detector simulation. This will be referred to as the fast simulation.

In this thesis the realistic simulation is used to test the methods used in this analysis. It is not used as an input to the final fits in any way. For this reason it is not necessary for the simulation to be tuned exactly to the data. There are three stages to the simulation: event generation, detector simulation and trigger simulation.

b-hadrons are generated using the `Bgenerator` package [43]. This package generates b-hadrons only, without fragmentation products or proton remnants. The decays of the b-hadrons are simulated by the `EvtGen` package [44] which was developed and extensively tested at the BARBAR and CLEO experiments.

The generated events are then passed to a detector and trigger simulation, a detailed description of which can be found in [45]. The material and geometry of the detector are modelled by the `GEANT` package [46]. `GEANT` models the passage of particles through the active and passive detector components accounting for interactions such as multiple scattering, bremsstrahlung, nuclear interactions and photon conversions. For some detector components the `GEANT` simulation is replaced by other packages. Drift time in the COT is simulated using `GARFIELD`, which is tuned on data, [47] and charge deposition in the silicon is simulated using a parametric model, tuned on data [48]. The realistic simulation also includes time dependent inefficiencies and changes in detector configuration. The trigger logic is simulated and the simulation output mimics the data structure such that the same reconstruction software may be used on simulation and data.

Samples of 10^7 events each of $B_s^0 \to K^+K^-$, $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$ and $\overline{B^0} \to \pi^+K^-$ decays were generated. After the trigger simulation and selection requirements

were applied there were 74802 $B_s^0 \to K^+K^-$ events, 84101 $B^0 \to \pi^+\pi^-$ events and 80211 $B^0 \to K^+\pi^-$ and $\overline{B^0} \to \pi^+K^-$ events. The input lifetimes were $c\tau = 462$ μm for B^0 and $c\tau = 438$ μm for $B_s^0 \to K^+K^-$, taken from [49].

3.6 Fast simulation

The realistic simulation described in the previous section includes a detailed model of event decays and detector simulation; however, generating events in this way is slow. It is useful to be able generate large samples quickly, for example to perform the many, repeat, pseudo-experiments described in chapter 7. For this a faster, simpler simulation is required. The fast simulation is used to test the performance of the methods used in this analysis but it is not used directly in the measurement; therefore, while the simulation should replicate the distributions observed in data where possible, it does not need to be an exact match.

The fast simulation generates the kinematic quantities and topology of each event and the lifetime of the $B^0_{(s)}$ or Λ_b^0. The detector resolution effects are simulated by smearing some of these quantities and trigger requirements are applied.

The first step is to consider a B^0, B_s^0 or Λ_b^0 meson decaying to two hadrons (pion, kaon or proton) in the rest frame of the $B^0_{(s)}$ or Λ_b^0. The momenta of the two hadrons in this frame are uniquely determined by the mass of the initial $B^0_{(s)}$ (Λ_b^0). The direction of the two decay products with respect to the z-axis is randomly chosen such that the decays occur uniformly in solid angle. The final rest frame quantity to be generated is the proper decay time, t, which is randomly chosen from an exponential distribution,

$$f(t) = \frac{1}{\tau} e^{\frac{-t}{\tau}}, \tag{3.3}$$

where τ is the mean lifetime of the B (Λ_b^0). The mean B^0 and $B_s^0 \to K^+K^-$ lifetimes are chosen to be the same as those used in the realistic simulation, that is $c\tau(B^0) = 462$

μmand $c\tau(\mathrm{B}_\mathrm{s}^0) = 438$ μm in the decay $\mathrm{B}_\mathrm{s}^0 \to \mathrm{K}^+\mathrm{K}^-$. For the Λ_b^0 the mean lifetime, $c\tau(\Lambda_\mathrm{b}^0) = 369$ μm, taken from [50], is used. This analysis assumes that the average B_s^0 lifetime is the same as the B^0 lifetime so $\mathrm{B}_\mathrm{s}^0 \to \mathrm{K}^+\pi^-$ decays are generated with $c\tau(\mathrm{B}_\mathrm{s}^0) = c\tau(\mathrm{B}^0) = 462$ μm.

In order to simulate the effect of detector resolution the lifetime is smeared according to a Gaussian of width 32.2 μm (determined from a detailed simulation of $\mathrm{B}_\mathrm{s}^0 \to \mathrm{K}^+\mathrm{K}^-$ decays and the CDF detector [51]).

Once rest frame quantities have been generated, laboratory (lab) frame quantities can be determined. The direction of the $\mathrm{B}_{(\mathrm{s})}^0$ (Λ_b^0) in the lab frame is randomly chosen to be uniform in ϕ and η in the range $-1 < \eta < 1$. The absolute momentum of the $\mathrm{B}_{(\mathrm{s})}^0$ (Λ_b^0) in the lab frame, $|\mathbf{p}_\mathrm{B}|$, is randomly selected from a histogram of momenta. How this initial histogram is obtained is discussed below. The momenta of the two decay products in the laboratory frame can be calculated by performing a Lorentz boost on the rest frame quantities. The position of the B decay vertex, $\mathbf{V}_\mathrm{dcy}$, is calculated as

$$\mathbf{V}_\mathrm{dcy} = \mathbf{V}_\mathrm{prod} + \frac{ct}{m_\mathrm{B}}\mathbf{p}_\mathrm{B} \tag{3.4}$$

where m_B is the mass of the parent $\mathrm{B}_{(\mathrm{s})}^0$or Λ_b^0 meson and the production vertex ($\mathbf{V}_\mathrm{prod}$) is chosen to be at the origin of the detector. The impact parameters ($\mathbf{d}$) of the daughter hadrons are calculated as

$$\mathbf{d} = \left((\mathbf{V}_\mathrm{prod} - \mathbf{V}_\mathrm{dcy}) \wedge \frac{\mathbf{p}_h}{|\mathbf{p}_h|} \right) \tag{3.5}$$

where $\mathbf{p}_h$ is the momentum of the daughter hadron.

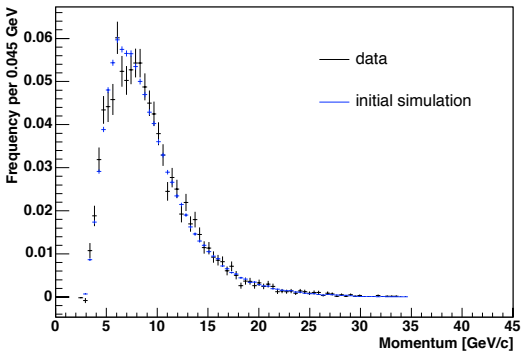
To simulate the behaviour of the SVT the SVT-measured transverse impact parameter is determined to be the transverse component of the offline impact parameter, calculated above, randomly smeared by a Gaussian of width 33 μm. This smeared impact parameter is then shifted to the nearest 10 μm such that it is $d_0^\mathrm{SVT} = 0, \pm 10, \pm 20, \ldots$

μm. This simulates the behaviour of the SVT which is discussed further in section 5.5.

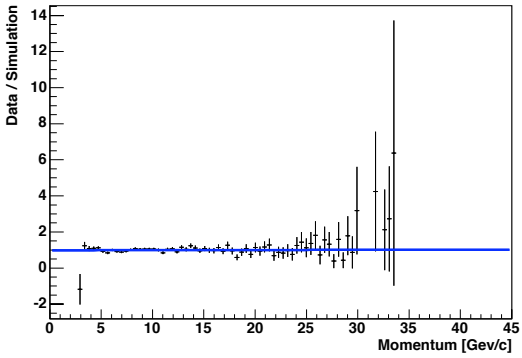
The final stage is to apply the selection and trigger requirements, described earlier, to the SVT and offline, simulated quantities.

The B momentum histogram used to generate the B momentum in the lab frame has been chosen such that the generated distribution of B momenta after the trigger requirements matches that observed in data. The distribution of B momenta in signal is determined from the data sample using a background subtraction technique. Histograms are made of the distribution of momenta in a signal region, defined as events with a reconstructed B mass in the range 5.18 to 5.3 GeV/c^2, and a background region in the mass range 5.6 to 5.8 GeV/c^2. The number of signal and background events in each region is determined by the maximum likelihood fit described in chapter 4. A histogram representative of the momentum distribution in signal is obtained by subtracting the histogram of momentum in the background region, weighted by the ratio of the number of background events in the signal region to the number of events in the background region, from the histogram of momenta in the signal region. The distribution of B momenta from data is shown in figure 3.2a.

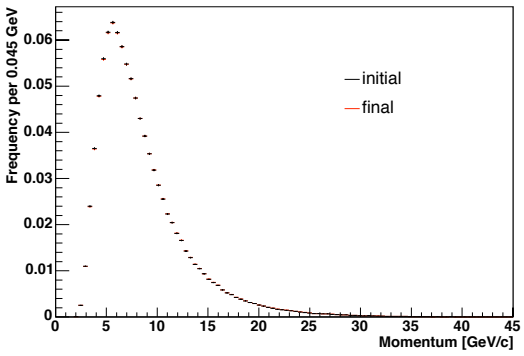
To create the momentum histogram used to generate the B momentum in the lab frame a simulated sample of events is generated using a histogram of the momentum distribution in realistic simulation. The final distribution of momentum in this initial fast simulation is compared to the distribution in data, shown in figure 3.2a. The ratio of the two histograms is fitted with an exponential function, figure 3.2b. This function is used to weight the histogram taken from realistic simulation. This reweighted histogram is the version used to generate the fast simulation. Figure 3.2c shows the initial and the final histograms used to generate the B momentum while figure 3.2d compares the distribution of B meson momenta in signal, in data, to the distribution obtained from the final fast simulation. For the signal simulation the distribution taken from the realistic simulation already generates a distribution matching data.



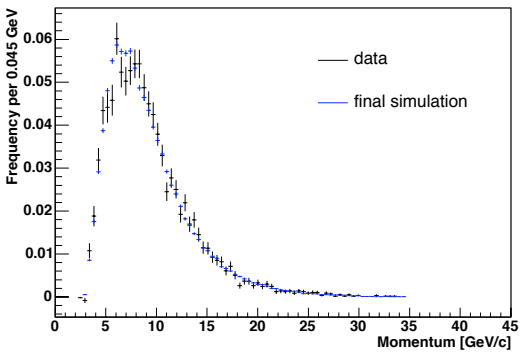
(a) Signal B momentum distribution. Black points are the distribution in data. The blue points are the distribution in the initial fast simulation described in the text.



(b) The ratio of the distribution of B momentum in data to the distribution in the initial fast simulation (black crosses). The line is an exponential function fit to the data points.



(c) The initial (black) and final (red) histograms used to generate the distribution of B momenta for fast simulation.



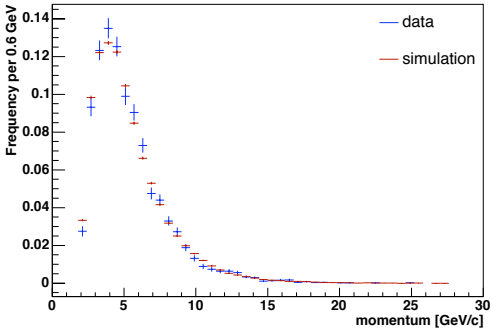
(d) Signal B momentum distribution. Black points are the distribution in data. The blue points are the distribution in the final fast simulation.

Figure 3.2: Obtaining the histogram used to generate the momentum of the B meson in the fast simulation.

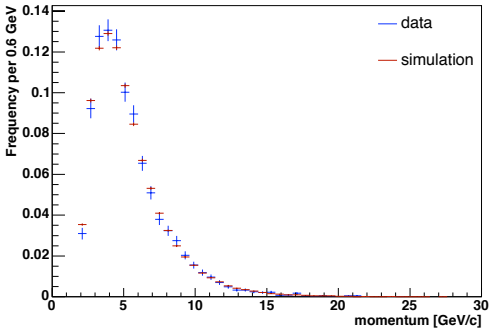
The reconstructed mass of the B (Λ_b^0) may be generated independently of the rest of the kinematics according to the distribution required for any particular study. Figure 3.3 shows a comparison of the distributions of some kinematic quantities of the B meson and the final state hadrons for fast simulation signal decays and signal data. The distributions from the fast simulation match the data very well.

A fast background simulation is required also. To allow fast generation, background events are generated as if they came from the real decay of a B meson with a randomly chosen mass in the range [5.0, 5.8] GeV/c^2. The decay of a B meson to two hadrons, each of which may be pion, kaon, proton or electron, is treated in the same way as for the generation of signal events. The sample is made to look more like background than signal by: using a different histogram to generate momentum; generating lifetimes from a distribution measured in data; generating a different mass distribution. The fractions of final state pions, kaons, protons and electrons used are the same as those observed in data using the fit described in chapter 4.

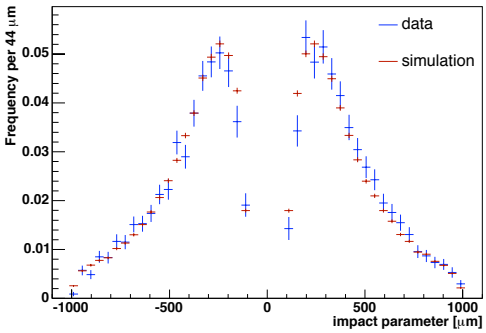
The histogram used to generate the B momentum for background is obtained in a similar way as for signal. An initial simulation is generated using the momentum distribution taken from the realistic simulation. The B momentum distribution resulting from this simulation is compared, this time, to a background sample taken from data in the mass region in the range 5.6 to 5.8 GeV/c^2 (Figure 3.4a). The ratio of the two histograms is fitted with an exponential function (figure 3.4b) and this function is used to reweight the histogram taken from the realistic simulation (figure 3.4c). This reweighted histogram is used to generate the final fast background simulation. Figure 3.4d compares the distribution of B momenta in background from data to that obtained in the fast simulation. The fast simulation reproduces well the momentum distributions of the final state hadrons and the distributions of impact parameter for each to the hadrons. This analysis is particularly concerned with the effects of the trigger which selects events based on impact parameter (amongst other observable quantities). The distribution of impact parameters in signal is observed to be broader



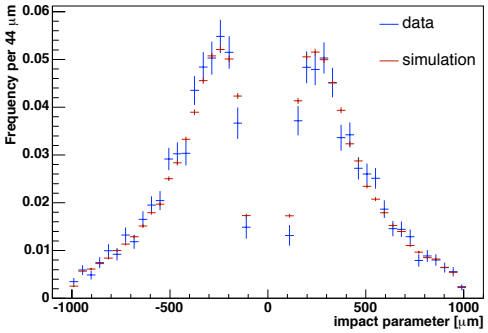
(a) Momentum of positive pion.



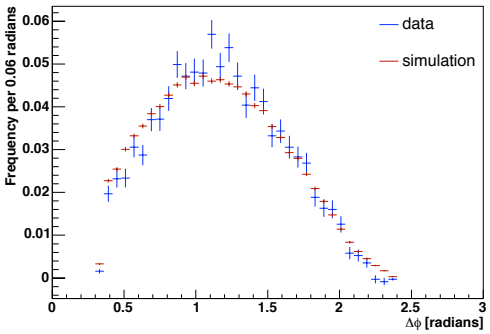
(b) Momentum of negative pion.



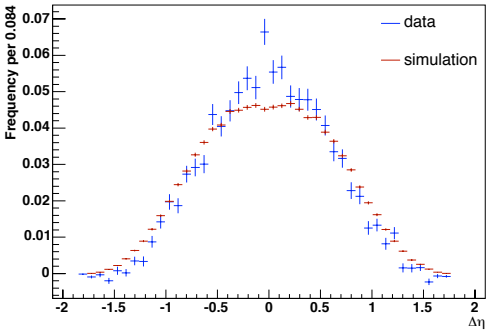
(c) Impact parameter of positive pion.



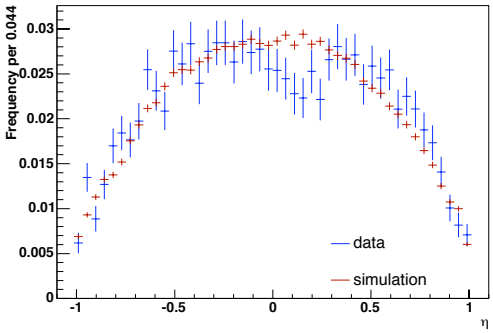
(d) Impact parameter of negative pion.



(e) $\Delta\phi$ of the two pion tracks.

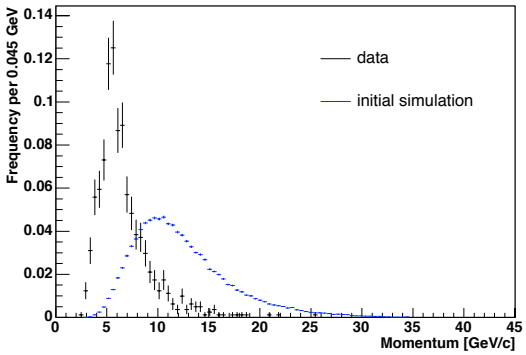


(f) $\Delta\eta$ of the two pion tracks.

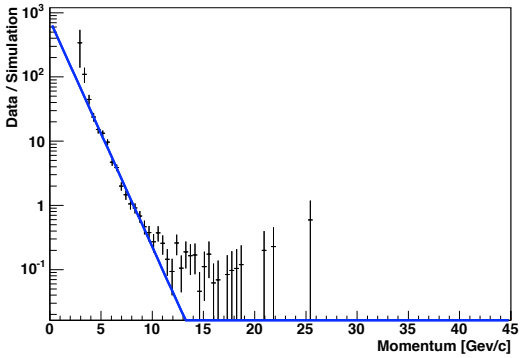


(g) η of the B^0.

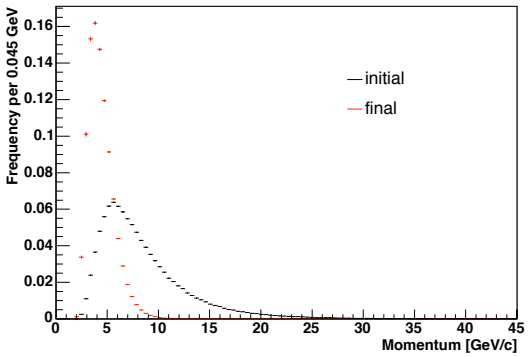
Figure 3.3: Comparing signal from data to the fast simulation of $B^0 \to \pi^+\pi^-$ events. The comparison between the momentum and impact parameter distributions of the two pions is good as is the comparison of the η distribution for B^0. There is only a small discrepancy in the distributions of $\Delta\phi$ and $\Delta\eta$ between the two pions. It should be noted that the signal distribution from data will contain contributions from $B^0 \to K^+\pi^-$, $\overline{B^0} \to \pi^+K^-$ and $B_s^0 \to K^+K^-$ decays, as well as $B^0 \to \pi^+\pi^-$, which may have slightly different distributions.



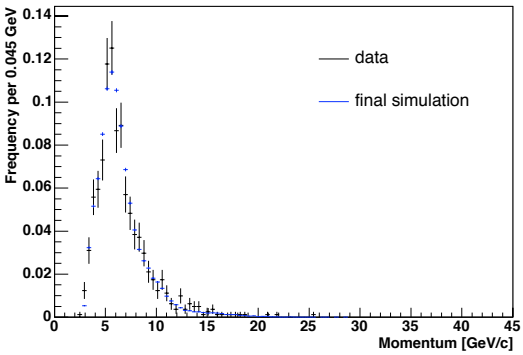
(a) Background B momentum distribution. Black points are the distribution in data. The blue points are the distribution in the initial fast simulation described in the text.



(b) The ratio of the distribution of B momentum in data to the distribution in the initial fast simulation (black crosses) for background. The line is an exponential function fit to the data points.



(c) The initial (black) and final (red) histograms used to generate the distribution of B momenta for fast simulation of background events.



(d) Background B momentum distribution. Black points are the distribution in data. The blue points are the distribution in the final fast simulation.

Figure 3.4: Obtaining the histogram used to generate the momentum of the B meson in the fast simulation of background events.

than for background and this is well reproduced by the fast simulations. The fast simulation of background does not reproduce well the distribution of $\Delta\eta$ seen in data. This is a consequence of assuming the background comes from physical decays of B mesons in the fast simulation while combinatorial background typically does not. The $B^0_{(s)}$ lifetimes are not dependent on $\Delta\eta$ so this difference is not important.

The proper decay time, t, of the background event is not generated according to a simple exponential function but to a general function, $y(t)$, obtained from a fit to data in the upper mass sideband ($5.6 < m_B < 5.8$ GeV/c^2). The function and the fit used to obtain it are described fully in section 5.9. For now it is sufficient to note that it

produces lifetimes typically much shorter than for signal and that it is the differences in the lifetime and momentum distributions which cause the differences in the fast simulations of signal and background.

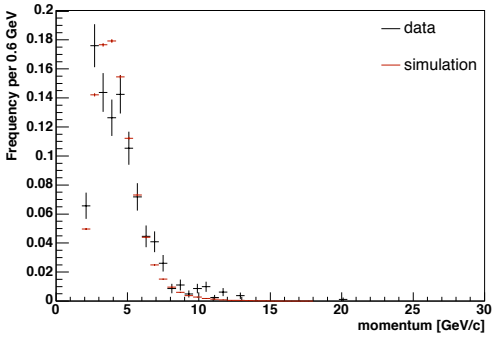
As with signal, The reconstructed mass may be generated independently of the rest of the kinematics of the event to any desired shape. Figure 3.5 shows a comparison of the distributions of some kinematic quantities of background events generated with the fast simulation to those observed in background in data.

3.7 Summary

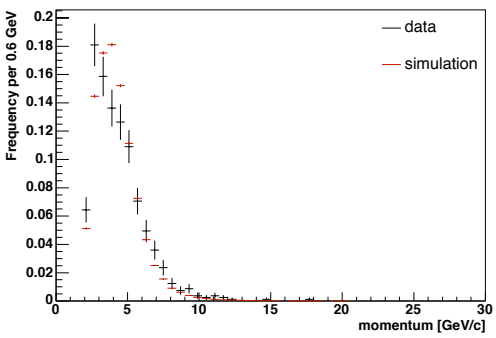
This chapter began by describing the reconstruction and selection of a data sample enriched in $\mathrm{B}^0_{(s)} \to h^+h'^-$ decays but also containing some $\Lambda^0_\mathrm{b} \to \mathrm{p}h^-$ decays and background events. The selection (both by the trigger and offline requirements) was based on the physical properties of the decays, including the impact parameters of the particle tracks and the displacement of the decay vertex from the interaction point.

Also described in this chapter were two types of data simulation. The realistic simulation includes a detailed event and detector simulation. The fast simulation can quickly produce many large datasets. Both types of simulation are used (chapter 7) to test the methods used in this analysis.

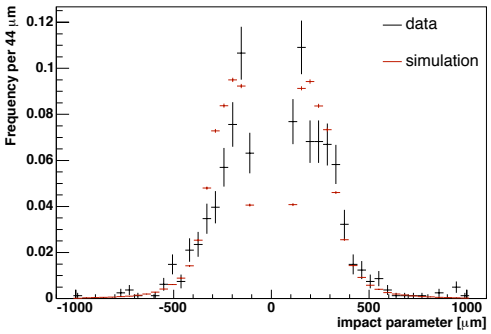
The following chapters describe methods to statistically separate the different contributions to the data sample (chapter 4) and to account for the effect of the selection requirements on the distribution of lifetimes in the data sample (chapter 5).



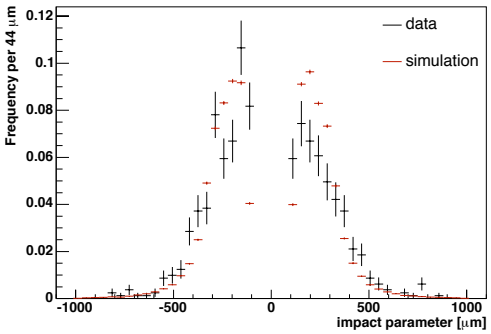
(a) Momentum of positive hadron.



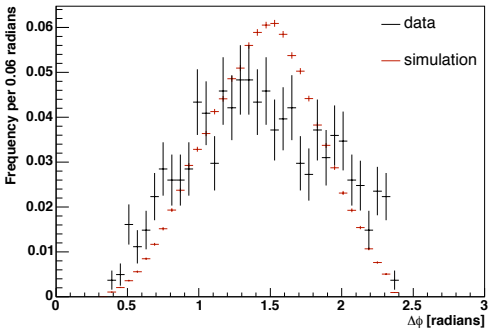
(b) Momentum of negative hadron.



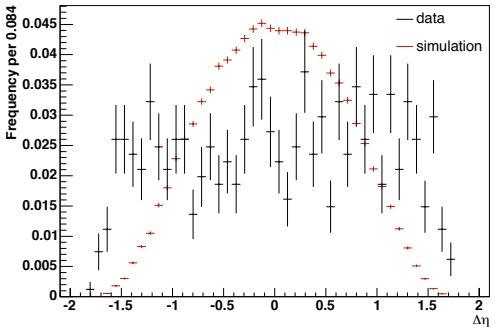
(c) Impact parameter of positive hadron.



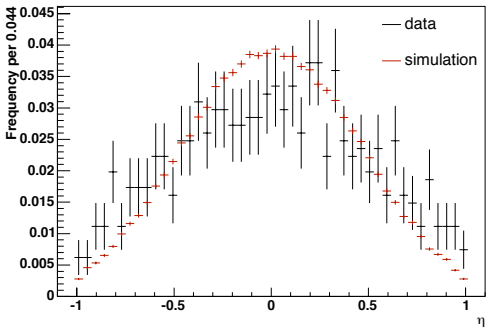
(d) Impact parameter of negative hadron.



(e) $\Delta\phi$ of the two hadron tracks.



(f) $\Delta\eta$ of the two hadron tracks.



(g) η of the B^0 candidate.

Figure 3.5: Comparing background from data to the fast simulation of background events. The comparison between the momentum and impact parameter distributions of the two hadrons is good. There is a small discrepency in the distributions of $\Delta\phi$ and the η distribution of the B candidate and a bigger discrepency in the $\Delta\eta$ distribution. This arises because the simulation simulates the decay of a physical B meson unlike combinatorial background in data.

Chapter 4

Disentangling the contributions to the $\mathrm{B}^0_{(s)} \to h^+h'^-$ signal

4.1 Introduction

Chapter 3 describes the reconstruction and selection of a data sample enriched with $\mathrm{B}^0_\mathrm{s} \to \mathrm{K}^+\mathrm{K}^-$ decays. This sample also contains a number of other two-body decays: $\mathrm{B}^0 \to \mathrm{K}^+\pi^-$, $\mathrm{B}^0 \to \pi^+\pi^-$, $\mathrm{B}^0_\mathrm{s} \to \mathrm{K}^+\pi^-$, $\Lambda^0_\mathrm{b} \to \mathrm{p}\pi^-$, $\Lambda^0_\mathrm{b} \to \mathrm{pK}^-$ and their charge conjugate decays. All of these decays have been observed at CDF in the same 1 fb^{-1} data set used for this analysis [52, 41]. The decays $\mathrm{B}^0_\mathrm{s} \to \pi^+\pi^-$ and $\mathrm{B}^0 \to \mathrm{K}^+\mathrm{K}^-$ could also lie in this region however their contribution has been observed to be consistent with zero for this integrated luminosity [41] and so they are neglected in this analysis. The sample also contains background events consisting of combinatorial background (where two unrelated tracks have been combined to form a decay during the reconstruction) and partially reconstructed B meson decays: $\mathrm{B}^+ \to \rho^0\pi^+(\mathrm{K}^+)$, $\mathrm{B}^0 \to \rho^\pm\pi^\pm(\mathrm{K}^\pm)$ and $\mathrm{B}^0_\mathrm{s} \to \rho^\pm\mathrm{K}^\mp$ [53].

Each of the decay modes in this sample lie so close together in mass that the mass resolution at CDF is not sufficient to separate them. In a histogram of invariant mass

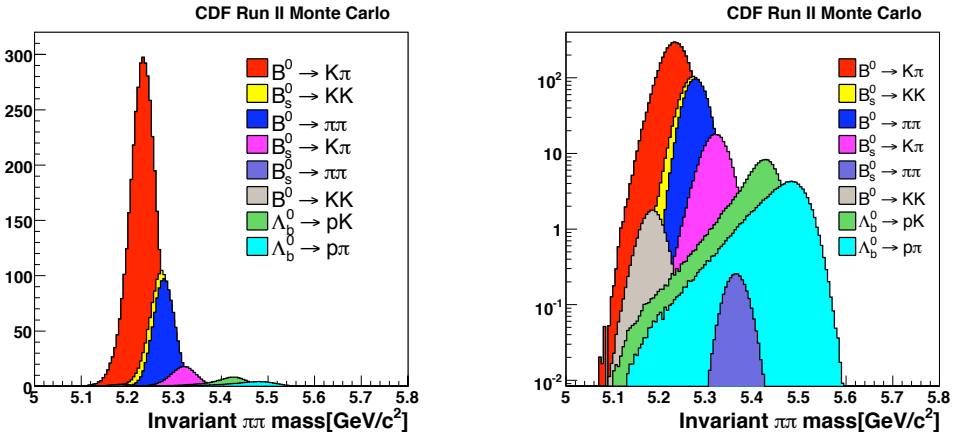


Figure 4.1: Simulation demonstrating the contribution to the invariant $\pi\pi$-mass distribution of the $B^0_{(s)} \to h^+h'^-$ and $\Lambda^0_b \to ph^-$ sample (taken from [41]). Both figures show the same simulated data but using a linear (left) and logarithmic (right) scale to highlight the dominant and rare decay modes respectively.

calculated assuming both decay products are pions the B decays to two hadrons appear as a single broad peak with a wider than expected resolution for a single decay. Figure 4.1 shows a simulation of the distribution of the different decay modes in invariant $\pi\pi$-mass. The distributions are normalised using branching fractions derived from current experimental knowledge and theoretical predictions. Particle identification at CDF is also unable to separate the decay modes on an event-by-event basis.

It is necessary to disentangle the various decay modes in order to measure the $B^0_s \to K^+K^-$ lifetime. This is not possible on an event-by-event basis so statistical methods need to be employed. Particle identification and kinematic information are combined to statistically separate the $B^0_{(s)} \to h^+h'^-$ contributions. This method has been developed at CDF and used for the first observation of $B^0_s \to K^+K^-$ [52], $B^0_s \to K^+\pi^-$, $\Lambda^0_b \to p\pi^-$ and $\Lambda^0_b \to pK^-$ decays; measurements of the various branching ratios, and of the CP asymmetries in these decays [41] as well as a preliminary measurement of the lifetime in $B^0_s \to K^+K^-$ decays in 360 pb^{-1} luminosity [53]. This analysis employs the same statistical separation of sample composition as these earlier analyses. The description of this statistical separation will form the remainder of this chapter.

4.2 Separating modes using kinematic information

It is not possible to separate the $B^0_{(s)} \to h^+h'^-$ decay modes by mass on an event by event basis however the mass distributions for each mode are different so can be used to help statistically separate the contributions. The $B^0_{(s)}$ or Λ^0_b mass is reconstructed as the invariant mass of the two decay products which will be a combination of protons, kaons and pions. Calculating the invariant mass of every combination of decay products would give eight different observables (for $K^\pm\pi^\mp$, K^+K^-, $\pi^+\pi^-$, pK^-, $\bar{p}K^+$, $p\pi^-$ and $\bar{p}\pi^+$) which would be highly correlated with each other and each of these correlations would have to be modelled in the probability density function. However, choosing only one, preferred, mass reconstruction would miss some of the kinematic separation available. Instead, a combination of variables are used to maximise the separation available from kinematic information: the invariant mass calculated assuming both decay products are pions ($M_{\pi\pi}$) and the momenta of the decay products.

Reconstructing all events assuming both decay products are pions broadens the mass distributions of each of the decay modes (except $B^0 \to \pi^+\pi^-$ which has the correct reconstruction). The width of the $M_{\pi\pi}$ distribution for the $B^0 \to \pi^+\pi^-$ mode is approximately 25 MeV/c^2 while for the other modes it is approximately 30 MeV/c^2. The position of each mass peak is also shifted. The $B^0 \to \pi^+\pi^-$ peak is centred on the B^0 mass since this mode is reconstructed correctly. The $B^0 \to K^+\pi^-$ peak is shifted down by 45 MeV/c^2 with respect to the B^0 mass and the $B^0_s \to K^+\pi^-$ peak is shifted up by 45 MeV/c^2 with respect to the B^0_s mass. The $B^0_s \to K^+K^-$ peak is shifted down to lie on top of the $B^0 \to \pi^+\pi^-$peak. This is because the difference in invariant mass due to misreconstructing the kaons as pions is approximately the same as the difference between the B^0_s and B^0 masses. The Λ^0_b decay modes have very wide distributions of $M_{\pi\pi}$ because the proton and pion masses are so different. Figure 4.1 shows the distribution of $M_{\pi\pi}$ for each decay.

All events are reconstructed assuming both decay products are pions, rather than

assuming both tracks are kaons say, to be consistent with the analysis in [41]. In that analysis, which was predominately concerned with the $B^0 \to K^{\pm}\pi^{\mp}$ decays, the π-π reconstruction was chosen over the K-π reconstruction to avoid the ambiguity over which particle should be reconstructed as the pion and which the kaon.

It is possible to calculate the expected central value of each $M_{\pi\pi}$ distribution. When a heavy particle decays to two particles with masses m_1 and m_2 and 3-dimensional momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ the invariant mass of the system is

$$M_{12}^2 = \left(\sqrt{m_1^2 + p_1^2} + \sqrt{m_2^2 + p_2^2}\right)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 . \tag{4.1}$$

If the pion mass were arbitrarily assigned to both decay particles the invariant mass would be calculated as

$$M_{\pi\pi}^2 = \left(\sqrt{m_\pi^2 + p_1^2} + \sqrt{m_\pi^2 + p_2^2}\right)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 . \tag{4.2}$$

Combining equation 4.1 and equation 4.2 gives

$$\begin{aligned} M_{\pi\pi}^2 &= \mathcal{M}^2(p_1, p_2) \\ &= M_{12}^2 + (m_\pi^2 + m_\pi^2) - (m_1^2 + m_2^2) + 2\left(\sqrt{p_1^2 + m_\pi^2}\sqrt{p_2^2 + m_\pi^2} - \sqrt{p_1^2 + m_1^2}\sqrt{p_2^2 + m_2^2}\right). \end{aligned} \tag{4.3}$$

M_{12} is the true mass of the meson in question: B^0, B_s^0 or Λ_b^0. m_1 and m_2 are the pion, kaon or proton masses as appropriate. $\mathcal{M}^2(p_1, p_2)$ gives the expected value of $M_{\pi\pi}$ for each decay hypothesis given the momentum of the two tracks.

There is more separation power available in the kinematic information than given by mass alone. In two body decays the heavier of the two decay products will typically carry the greater momentum. This is useful in the decays with Kπ, pK or pπ in the final state. For example in a decay of a B^0 meson to a kaon plus a pion if the positive particle is carrying the greater momentum then the decay is more likely to be $B^0 \to K^+\pi^-$;

if the negative particle is carrying the greater momentum then it is more likely the decay is $\overline{\mathrm{B}^0} \to \pi^+\mathrm{K}^-$. This information can be included in the PDF by considering the charge weighted momentum imbalance:

$$\alpha = q_1 \left(1 - \frac{p_1}{p_2}\right), \tag{4.4}$$

where p_1 and p_2 are the magnitude of the momenta of the decay products with the smaller and larger momentum respectively and q_1 is the charge of the particle with the smaller momentum. This variable can be used in conjunction with a second momentum dependant variable, the scalar sum of the momenta, $p_{\mathrm{tot}} = p_1 + p_2$. In terms of α and p_{tot}, p_1 and p_2 can be expressed as

$$p_1 = \frac{1-|\alpha|}{2-|\alpha|} p_{\mathrm{tot}} \quad \text{and} \quad p_2 = \frac{1}{2-|\alpha|} p_{\mathrm{tot}}. \tag{4.5}$$

α and p_{tot} now contain all of the momentum information and can replace p_1 and p_2 in equation 4.3 to give the expected $M_{\pi\pi}$ value given α and p_{tot} for each mode hypothesis, that is,

$$\begin{aligned}\mathcal{M}^2(\alpha, p_{\mathrm{tot}}) = M_{12}^2 + (m_\pi^2 + m_\pi^2) - (m_1^2 + m_2^2) \\ + 2\sqrt{\left(\frac{1-|\alpha|}{2-|\alpha|} p_{\mathrm{tot}}\right)^2 + m_\pi^2} \cdot \sqrt{\left(\frac{1}{2-|\alpha|} p_{\mathrm{tot}}\right)^2 + m_\pi^2} \\ - 2\sqrt{\left(\frac{1-|\alpha|}{2-|\alpha|} p_{\mathrm{tot}}\right)^2 + m_1^2} \cdot \sqrt{\left(\frac{1}{2-|\alpha|} p_{\mathrm{tot}}\right)^2 + m_2^2}.\end{aligned} \tag{4.6}$$

Figure 4.2 demonstrates the differences in the average $M_{\pi\pi}$ mass as a function of α for each decay mode. The Average $M_{\pi\pi}$ mass for $\mathrm{B}^0 \to \pi^+\pi^-$and $\mathrm{B}_\mathrm{s}^0 \to \pi^+\pi^-$decays is flat, at the B^0 and B_s^0 masses respectively, since these modes are correctly reconstructed. The average $M_{\pi\pi}$ mass for the other modes is shifted from the nominal $\mathrm{B}_{(\mathrm{s})}^0$ or Λ_b^0 masses. The separation between modes is greatest for large momentum imbalance. Figures 4.3 and 4.4 show the distribution of $M_{\pi\pi}$ and α for each of the decays. It can

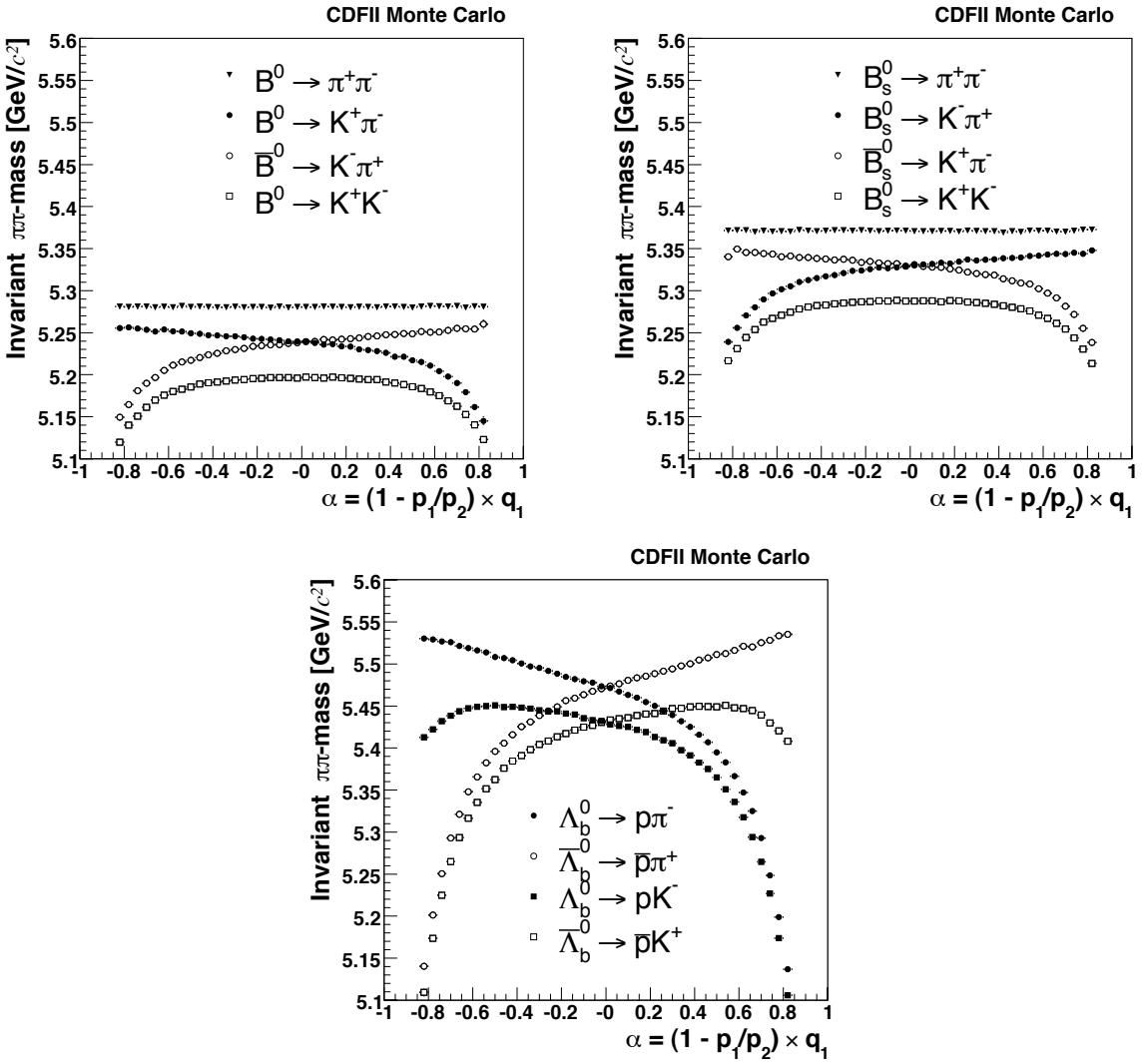


Figure 4.2: Average $M_{\pi\pi}$ as a function of α for simulated $\mathrm{B}^0_{(s)} \to h^+h'^-$ and $\Lambda^0_\mathrm{b} \to \mathrm{p}h^-$ decays. (Taken from [41].)

be seen from these figures that only a statistical separation of modes is possible using the combination of $M_{\pi\pi}$ and α variables because of the mass resolution available.

Kinematic information provides little separation between $B_s^0 \to K^+K^-$ and $B^0 \to \pi^+\pi^-$ decays. Fortunately these decays enjoy a better separation using particle identification than the other modes since both final state particles are different.

4.3 Separating modes using particle identification

There are three methods of particle identification for hadrons available at CDF each of which is useful at a different energy range. The time of flight detector is the only part of the CDF detector dedicated to particle identification. This provides 2σ separation of pions and kaons up to a momentum of 1.6 GeV/c; however, the trigger used to obtain the data sample for this analysis requires tracks to have a transverse momentum of at least 2 GeV/c so time of flight is not useful for identifying pions and kaons in this case. The time of flight detector does provide 2σ separation of kaons and protons up to 2.7 GeV/c and of pions and protons up to 3.2 GeV/c; however, the decay modes with protons in the final state are already well separated by mass and so time of flight is not necessary. Particle identification is also provided by measuring the specific ionisation of particles travelling through the tracking systems. Specific ionisation in the silicon tracker only provides separation up to a momentum of 800 MeV/c so is not useful here. Energy loss through ionisation ($\mathrm{d}E/\mathrm{d}x$) in the drift chamber provides separation of particles with a transverse momentum greater than 2 GeV/c and is used. From now on, $\mathrm{d}E/\mathrm{d}x$ refers to the specific ionisation measured in the drift chamber. Figure 4.5 shows the distribution of $\mathrm{d}E/\mathrm{d}x$ for kaons and pions. The drift chamber provides a 1.4σ separation between kaons and pions for particles with transverse momentum in the range $2 < p_{\mathrm{T}} < 10$ GeV/c.

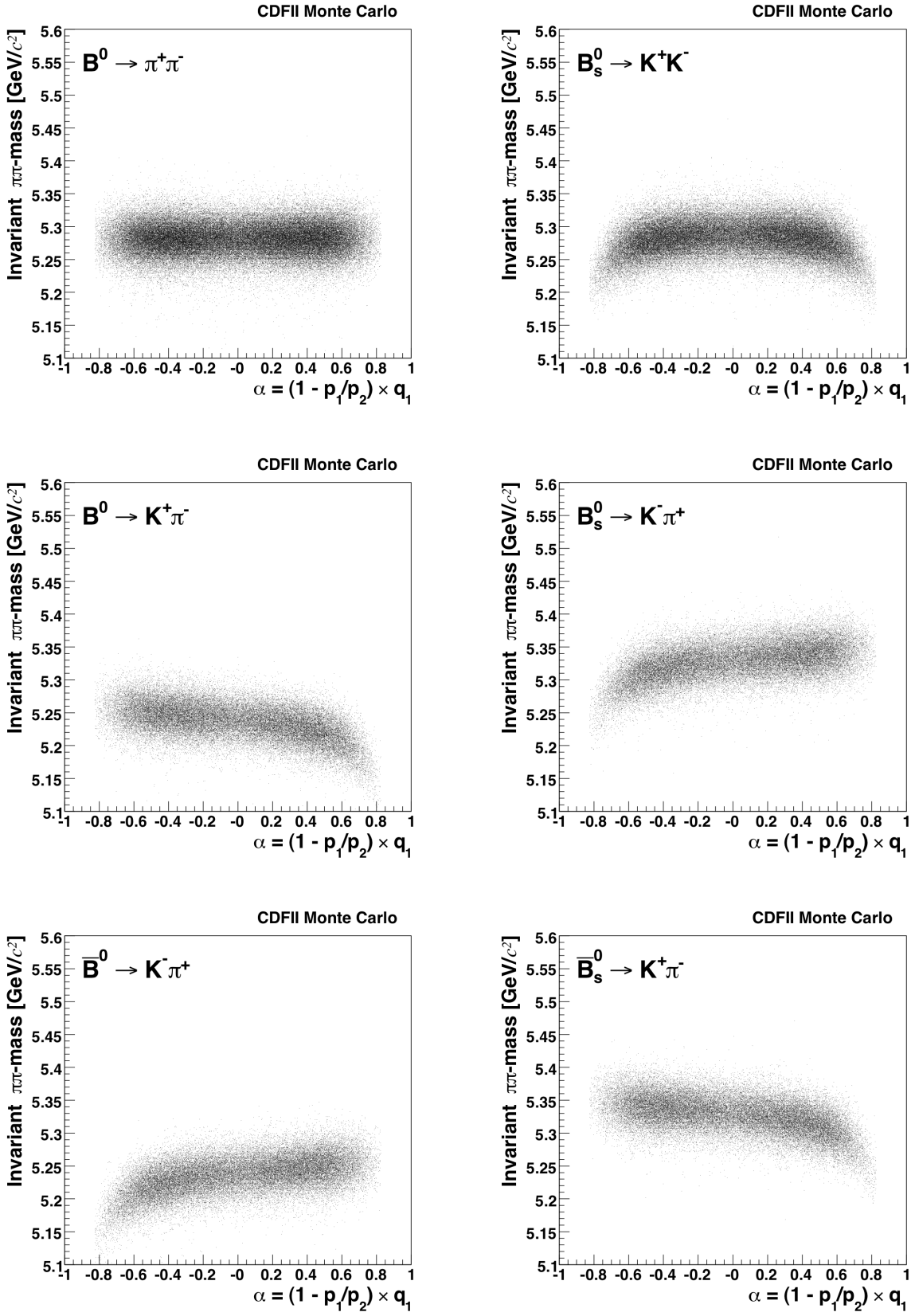


Figure 4.3: Simulated $M_{\pi\pi}$ distribution for the decay modes $B^0 \to \pi^+\pi^-$, $B^0 \to K^+\pi^-$, $\overline{B^0} \to \pi^+K^-$, $B_s^0 \to K^+K^-$, $B_s^0 \to K^+\pi^-$ and $B_s^0 \to \pi^+K^-$ as a function of α. (Taken from [41].)

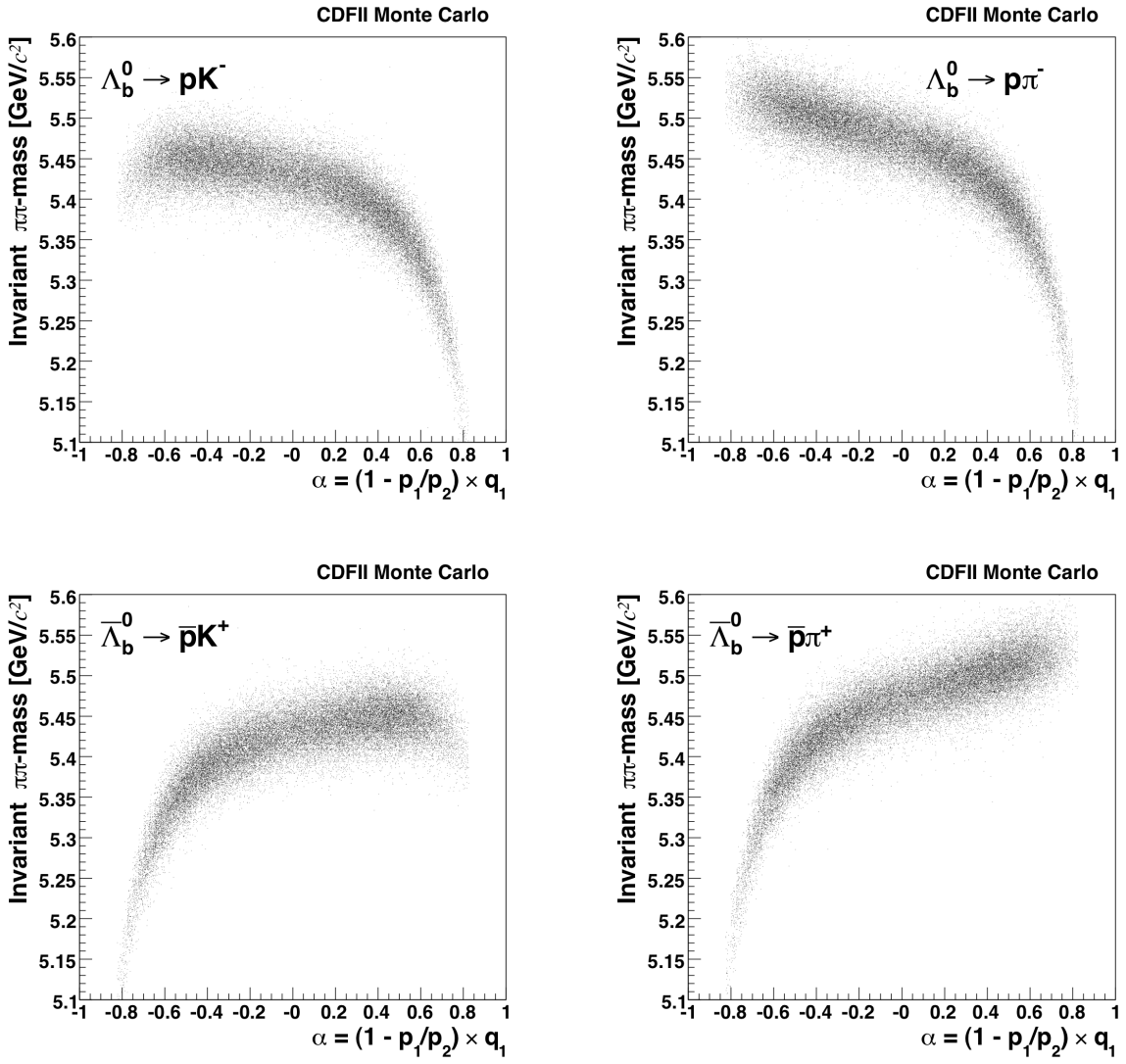


Figure 4.4: Simulated $M_{\pi\pi}$ distribution for the decay modes $\Lambda_b^0 \to pK^-$, $\Lambda_b^0 \to K^+\overline{p}$, $\Lambda_b^0 \to p\pi^-$and $\overline{\Lambda_b^0} \to \pi^+\overline{p}$ as a function of α. (Taken from [41].)

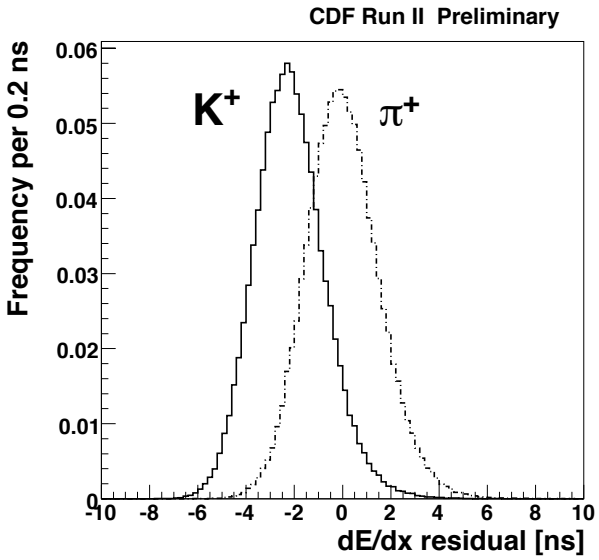


Figure 4.5: The distribution of $\mathrm{d}E/\mathrm{d}x$ around the mean pion response from the calibration samples of pions and kaons from $\mathrm{D}^{*+} \to \mathrm{D}^0\pi_\mathrm{s}^+ \to [\mathrm{K}^-\pi^+]\pi_\mathrm{s}^+$ decays. A 1.4σ separation between kaons and pions is available over the range $2 < p_\mathrm{T} < 10$ GeV/c. (Taken from [41].)

4.4 The separation likelihood

The kinematic and particle identification information is included in the likelihood using five observables. These are:

- $M_{\pi\pi}$, the invariant mass of the two final state particles assuming they are pions;
- $p_\mathrm{tot} = p_1 + p_2$, the scalar sum of the momenta (p_i) of each particle;
- $\alpha = q_1\left(1 - \frac{p_1}{p_2}\right)$, the imbalance between the moduli of the momenta (p_i) of the two particles weighted by the charge (q_1) of the particle with the smaller momentum (p_1);
- ID_1, function of $\mathrm{d}E/\mathrm{d}x$ for the particle with the smaller momentum;
- ID_2, function of $\mathrm{d}E/\mathrm{d}x$ for the particle with the larger momentum.

In section 4.6 the *ID* variable will be defined as

$$ID \equiv \frac{\frac{\mathrm{d}E}{\mathrm{d}x}\big|_{\mathrm{meas}} - \frac{\mathrm{d}E}{\mathrm{d}x}\big|_{\mathrm{exp}-\pi}}{\frac{\mathrm{d}E}{\mathrm{d}x}\big|_{\mathrm{exp}-K} - \frac{\mathrm{d}E}{\mathrm{d}x}\big|_{\mathrm{exp}-\pi}}. \tag{4.7}$$

where $\frac{\mathrm{d}E}{\mathrm{d}x}\big|_{\mathrm{meas}}$ denotes the measured $\mathrm{d}E/\mathrm{d}x$ of the particle and $\frac{\mathrm{d}E}{\mathrm{d}x}\big|_{\mathrm{exp}-\pi(K)}$ the expected $\mathrm{d}E/\mathrm{d}x$ if the particle were a pion (or kaon).

The combined likelihood for separation of the $\mathrm{B}^0_{(s)} \to h^+h'^-$ and $\Lambda^0_\mathrm{b} \to \mathrm{p}h^-$ modes can be written as the product of the likelihoods, $\mathcal{L}_i$, for each event,

$$\mathcal{L} = \prod_{i=1}^{N} \mathcal{L}_i, \tag{4.8}$$

where there are N events in the sample. The likelihood of the i^{th} event can be expressed as

$$\mathcal{L}_i = f_b \cdot \mathcal{L}^{bg} + (1 - f_b) \cdot \mathcal{L}^{sg}, \tag{4.9}$$

where f_b is the background fraction and $\mathcal{L}^{\mathrm{bg(sg)}}$ denotes the likelihood for background (signal).

In terms of probability density functions the signal likelihood is

$$\mathcal{L}^{sg} = \sum_{j=1}^{10} f_j \cdot P(M_{\pi\pi} \mid \alpha, p_{\mathrm{tot}}, s_j) \cdot P(ID_1, ID_2 \mid \alpha, p_{\mathrm{tot}}, s_j) \cdot P(\alpha, p_{\mathrm{tot}} \mid s_j), \tag{4.10}$$

where the sum is over the ten different signal contributions: $\mathrm{B}^0_\mathrm{s} \to \mathrm{K}^+\mathrm{K}^-$ and $\mathrm{B}^0 \to \pi^+\pi^-$; $\mathrm{B}^0 \to \mathrm{K}^+\pi^-$, $\mathrm{B}^0_\mathrm{s} \to \mathrm{K}^+\pi^-$, $\Lambda^0_\mathrm{b} \to \mathrm{pK}^-$, $\Lambda^0_\mathrm{b} \to \mathrm{p}\pi^-$ and the charge conjugate decays. The f_j are the fraction of signal which are of each decay mode. The fractions are all left free in the fit except for the final fraction (the fraction of signal which is $\Lambda^0_\mathrm{b} \to \mathrm{p}\pi^-$) which is completely determined by the other fractions as

$$f_{10} = 1 - \sum_{j=1}^{9} f_j. \tag{4.11}$$

The remaining terms in equation 4.10 are:

- $P(M_{\pi\pi} \mid \alpha, p_{\rm tot}, s_j)$, the mass term gives the probability of measuring $M_{\pi\pi}$ given the signal mode s_j and given the momentum variables α and $p_{\rm tot}$;
- $P(ID_1, ID_2 \mid \alpha, p_{\rm tot}, s_j)$, the particle identification term gives the probability of measuring ID_1 and ID_2 given the signal mode, s_j, and the momentum variables α and $p_{\rm tot}$;
- $P(\alpha, p_{\rm tot} \mid s_j)$, the momentum term gives the probability of measuring α and $p_{\rm tot}$ given the signal mode, s_j. As the mass and particle identification terms depend on α and $p_{\rm tot}$ it is necessary to model the distribution of the momentum variables also. This is discussed further in section 4.7.

These terms will be explained in greater detail below.

The likelihood for background can be similarly expressed:

$$\mathcal{L}^{bg} = \sum_{l=A,{\rm comb}} f_l \cdot P(M_{\pi\pi} \mid \alpha, p_{\rm tot}, b_l) \cdot P(ID_1, ID_2 \mid \alpha, p_{\rm tot}, b_l) \cdot P(\alpha, p_{\rm tot} \mid b_l). \quad (4.12)$$

In this case the sum is over the two contributions to background. The fraction of background which is combinatorial is $f_{\rm comb} = 1 - f_A$ where f_A is the fraction of background which comes from partially reconstructed B meson decays. The remaining terms are:

- $P(M_{\pi\pi} \mid \alpha, p_{\rm tot}, b_l)$, the mass term gives the probability of measuring $M_{\pi\pi}$ given the background type b_l and given the momentum variables α and $p_{\rm tot}$;
- $P(ID_1, ID_2 \mid \alpha, p_{\rm tot}, b_l)$, the particle identification term gives the probability of measuring ID_1 and ID_2 given the background type, b_l, and the momentum variables α and $p_{\rm tot}$;
- $P(\alpha, p_{\rm tot} \mid b_l)$, the momentum term gives the probability of measuring α and $p_{\rm tot}$ given the background type, b_l.

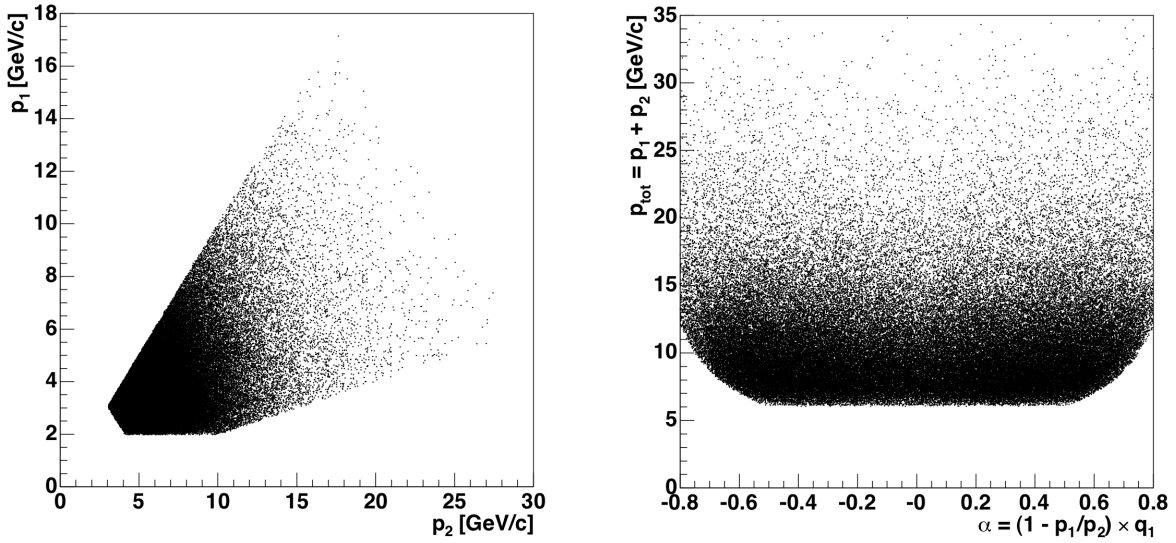


Figure 4.6: The momentum variables α and $p_{\rm tot}$ (right) show less correlation than the momenta, p_1 and p_2, of the two particles (left).

These terms will also be expanded below.

4.5 Kinematic probability density function

Section 4.2 demonstrated that the correlation between $M_{\pi\pi}$ and α can be exploited to provide statistical separation of some of the $\mathrm{B}^0_{(s)} \to h^+h'^-$ and $\Lambda^0_{\rm b} \to {\rm p}h^-$ modes. The shapes of the mass distributions are determined by: the shift in mass from the nominal $\mathrm{B}^0_{(s)}(\Lambda^0_{\rm b})$ mass due to the reconstruction of the decays assuming pions form the final state; the momentum resolution of the detector. To account for both of these effects the PDFs describing the kinematic distributions are written as the probability of measuring $M_{\pi\pi}$ given the momentum of the final state particles. The momenta of the two decay products are encompassed in the two variables α and $p_{\rm tot}$. An advantage of using α and $p_{\rm tot}$ is that they are less correlated than the momenta of the two particles as can be seen in figure 4.6.

4.5.1 Mass distribution for signal

The $M_{\pi\pi}$ distribution is modelled by the sum of two Gaussians with widths σ_{g1} and σ_{g2}. For each mode the mean of each Gaussian is the expected value of $M_{\pi\pi}$ for the measured momenta of the decay products. The expected $M_{\pi\pi}$ is $\mathcal{M}_j(\alpha, p_{\rm tot})$, given by equation 4.6, where the subscript j has been added to indicate that this expression is different for each $\mathrm{B}^0_{(s)} \to h^+h'^-$ decay mode hypothesis. Since the mean of the Gaussian depends on α and $p_{\rm tot}$ it varies event by event. The $M_{\pi\pi}$ distribution is

$$P(M_{\pi\pi} \mid \alpha, p_{\rm tot}, s_j) = f_{g1}\frac{1}{\sigma_{g1}\sqrt{2\pi}}e^{\frac{1}{2}\left(\frac{M_{\pi\pi}-\mathcal{M}_j(\alpha,p_{\rm tot})}{\sigma_{g1}}\right)^2} + (1-f_{g1})\frac{1}{\sigma_{g2}\sqrt{2\pi}}e^{\frac{1}{2}\left(\frac{M_{\pi\pi}-\mathcal{M}_j(\alpha,p_{\rm tot})}{\sigma_{g2}}\right)^2}, \tag{4.13}$$

where f_{g1} is the is the fraction of events with resolution σ_{g1}. In practice it is difficult to fit with σ_{g1}, σ_{g2} and f_{g1} all free in the fit since f_{g1} is highly correlated to the resolutions, σ_{gi}. To avoid this, the fraction is fixed to the value measured in the realistic simulation, $f_{g1} = 0.95$ and the effect of fixing this fraction is considered as a systematic.

$\mathcal{M}_j(\alpha, p_{\rm tot})$ requires the pion and kaon masses and the world average values of these are used [50]. For the B^0, B^0_s and Λ^0_b masses the values measured by CDF are used [54]. Those are:

$$M^{\rm CDF}_{\mathrm{B}^0} = 5279.63 \pm 0.53(\text{stat.}) \pm 0.33(\text{sys.})\ \text{MeV/c}^2;$$
$$M^{\rm CDF}_{\mathrm{B}^0_\mathrm{s}} = 5366.01 \pm 0.73(\text{stat.}) \pm 0.33(\text{sys.})\ \text{MeV/c}^2;$$
$$M^{\rm CDF}_{\Lambda^0_\mathrm{b}} = 5619.7 \pm 1.2(\text{stat.}) \pm 1.2(\text{sys.})\ \text{MeV/c}^2.$$

Using masses measured by the same detector removes common systematic uncertainties. The expected invariant mass for each mode will be further shifted due to misreconstructing kaons and protons as pions. This is different from the shift due to calculating the invariant mass assuming the pion mass for the decay products but rather is due to making this assumption during event reconstruction. As particles travel through the material of the detector they lose energy. The event reconstruction automatically cor-

mode	mass correction (MeV/c^2)
$B^0 \to \pi^+\pi^-$	—
$B^0 \to K^+\pi^-$	1.1
$B_s^0 \to K^+\pi^-$	1.1
$B_s^0 \to K^+K^-$	2.1
$\Lambda_b^0 \to p\pi^-$	1.7
$\Lambda_b^0 \to pK^-$	2.6

Table 4.1: The correction to the invariant $\pi\pi$-mass required to correct for the bias due to reconstructing the event assuming both the decay products are pions.

rects for this energy loss given the mass of the particle involved. This analysis assumes all particles are pions as the particle type is not known at the time of reconstruction. Where one or both particles are not pions an incorrect energy correction is applied, biassing the calculation of invariant mass. This shift has been measured for each of the $B^0_{(s)} \to h^+h'^-$ and $\Lambda_b^0 \to ph^-$ decay modes and can be applied as a correction to the bias [41]. The corrections to be applied for each mode are listed in table 4.1.

To allow for a shift in the global mass scale a parameter Δ_m is included such that the masses used in the fit are:

$$M_{B^0} = M_{B^0}^{CDF} + \Delta_m; \tag{4.14}$$

$$M_{B_s^0} = M_{B^0}^{CDF} + \Delta_m; \tag{4.15}$$

$$M_{\Lambda_b^0} = M_{B^0}^{CDF} + \Delta_m. \tag{4.16}$$

In this way the effect of the uncertainty on the global mass scale is included in the statistical rather than systematic error. This leaves only the uncertainty due to the relative B and Λ_b^0 masses to contribute to the systematic uncertainty.

4.5.2 Mass distribution for background

The background consists of two types of events, combinatorial and physics background, each with a different mass distribution. The $M_{\pi\pi}$ distribution for all background events

is given by

$$P(M_{\pi\pi} \mid b) = (1 - f_A)P_{\rm comb}(M_{\pi\pi}) + f_A P_A(M_{\pi\pi}), \tag{4.17}$$

where f_A is the fraction of background from partially reconstructed decays rather than combinatorial background. The $M_{\pi\pi}$ distributions for combinatorial ($P_{\rm comb}(M_{\pi\pi})$) and physics ($P_A(M_{\pi\pi})$) background are discussed below.

Combinatorial background

The combinatorial background consists of pairs of unrelated particles which pass the selection requirements and are combined to make a decay. In the transverse plane the kinematics of these events look like a $\mathrm{B}^0_{(s)} \to h^+h'^-$ decay however the particle tracks are separated along the z direction. Combinatorial background is distributed across the full mass range.

The $M_{\pi\pi}$ distribution for combinatorial background is described by an exponential function,

$$P_{\rm comb}(M_{\pi\pi}) = \frac{c_1 \cdot e^{(c_1 \cdot M_{\pi\pi})}}{e^{(c_1 \cdot m_{\rm min})} - e^{(c_1 \cdot m_{\rm max})}}, \tag{4.18}$$

which is normalised over the allowed mass range ($m_{\rm min} < M_{\pi\pi} < m_{\rm max}$). The functional form of the PDF was chosen by studying a sample obtained using the same selection cuts as for the analysis sample but with the requirement on χ^2 of the vertex reversed [55]. This produces a sample of combinatorial background with little contamination from $\mathrm{B}^0_{(s)} \to h^+h'^-$ or $\Lambda^0_{\rm b} \to {\rm p}h^-$ decays. The shape of the $M_{\pi\pi}$ distribution is somewhat dependent on χ^2 so this sample can only be used to provide a qualitative model of the mass shape. The slope of the exponential, c_1, is left free to be determined by the fit to data and is dominated by the background events at high mass, above the signal peaks.

Physics background

There is a significant change in the background slope just below the $B^0_{(s)} \to h^+h'^-$ mass peak indicating an additional source of background. This background comes from misreconstructed decays of B-hadrons. These are multi-body decays where only two of the final-state particles have been reconstructed. The $M_{\pi\pi}$ distribution for these decays is suppressed at $M_{\pi\pi} \approx 5.15$ GeV/c^2 for kinematic reasons since $M_{\pi\pi} < M_{B^0_{(s)}} - m_\pi$.

The main contribution to this background comes from the decays $B^+ \to h_1^+ h_2^- h_3^+$, $B^0 \to h_1^+ h_2^- h_3^0$ and $B_s^0 \to h_1^+ h_2^- h_3^0$ where $h = \pi$ or K. These include decays which occur via resonance states of ρ or K^* mesons plus either a kaon or a pion: $B^+ \to \rho^0\pi^+$, $B^+ \to \rho^0 K^+$, $B^0 \to \rho^\pm\pi^\mp$, $B^0 \to \rho^- K^+$, $B_s^0 \to \rho^\pm\pi^\mp$, $B_s^0 \to \rho^+ K^-$ (where $\rho \to \pi\pi$) and $B^+ \to K^{*0}\pi^+$, $B^+ \to K^{*+}\pi^0$, $B^0 \to K^{*+}\pi^-$, $B^0 \to K^{*0}\pi^0$ (where $K^* \to K\pi$) and many others. The branching fractions of many of the B^0 decay modes have been measured at the B-factories BarBar and Belle [56] but the B_s^0 branching fractions are unknown and have large theoretical uncertainties. The shape of the distribution of physics background is fit directly with the data. The functional form of the distribution was determined at CDF using a simulation of some of the decays mentioned above, in particular the decays occurring via the ρ resonance [55]. Figure 4.7 shows the simulated $B \to \rho h$ decays which were generated using the branching fractions in [56] and theoretical predictions for B_s^0 [57]. Similar distributions are expected for all the background decay modes in this region.

The $M_{\pi\pi}$ distribution for the physics background is a convolution of a Gaussian curve and an Argus function [58]:

$$P_A(M_{\pi\pi}) = g_A(M_{\pi\pi}) \otimes h_A(M_{\pi\pi}). \tag{4.19}$$

The Gaussian function, $g_A(M_{\pi\pi})$, accounts for the mass resolution and is centred at zero with a width of 22 MeV/c^2 which is the width of a single mode. The Argus function

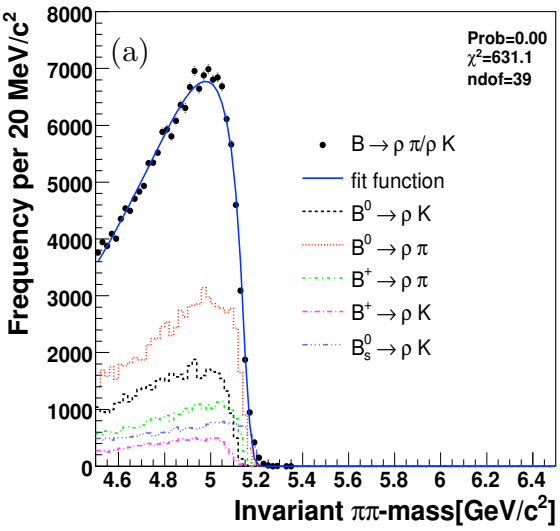


Figure 4.7: The invariant $\pi\pi$-mass distribution of simulated B $\rightarrow$ ρK and B $\rightarrow$ $\rho\pi$ decays. A fit to the combined mass shape is overlaid. (taken from [55].)

is

$$h_{\mathrm{A}} = \begin{cases} \frac{1}{K_A} M_{\pi\pi} \sqrt{1 - \left(\frac{M_{\pi\pi}}{m_A}\right)^2} e^{-c_A \left(\frac{M_{\pi\pi}}{m_A}\right)^2} & \text{if } M_{\pi\pi} \leq m_A \\ 0 & \text{if } M_{\pi\pi} > m_A \end{cases}, \tag{4.20}$$

where K_A is the normalisation constant. The parameters m_A and c_A are free to vary in the fit to data.

4.6 Particle identification probability density function

To construct the probability density function (PDF) for particle identification it is necessary to consider the intrinsic $\mathrm{d}E/\mathrm{d}x$ residuals for each particle and the correlation between the observed $\mathrm{d}E/\mathrm{d}x$ of each particle. The intrinsic residual $\mathrm{d}E/\mathrm{d}x$ of each particle is defined as

$$\delta_i = \left.\frac{\mathrm{d}E}{\mathrm{d}x}\right|_{\mathrm{meas}} - \left.\frac{\mathrm{d}E}{\mathrm{d}x}\right|_{\mathrm{exp}-i}, \tag{4.21}$$

where $\left.\frac{\mathrm{d}E}{\mathrm{d}x}\right|_{\mathrm{meas}}$ denotes the measured $\mathrm{d}E/\mathrm{d}x$ of the particle and $\left.\frac{\mathrm{d}E}{\mathrm{d}x}\right|_{\mathrm{exp}-i}$ the expected $\mathrm{d}E/\mathrm{d}x$ of the particle for a given particle mass hypothesis, i. The particle mass hy-

pothesis i may be kaon, pion, proton or electron. The observed residual is given by the sum of the intrinsic, un-correlated, residual and the common mode shift, c. So for two particles in the same event with mass hypotheses l and m the observed residuals are

$$\delta_l^{obs} = \delta_l + c \quad \text{and} \quad \delta_m^{obs} = \delta_m + c. \tag{4.22}$$

As the observed residuals of the two particles are correlated it is not possible to separate the PDFs for each particle and a combined PDF must be used. The joint PDF for the observed residuals is given by the convolution of the PDFs for the intrinsic residuals and the PDF for the correlation, that is:

$$\begin{aligned} P(\delta_1^{obs}, \delta_2^{obs}) &= (P(\delta_1) \times P(\delta_2)\ ) \otimes P(c) \\ &= \int_{-\infty}^{+\infty} P(\delta_1 - c) P(\delta_2 - c) P(c) \mathrm{d}c. \end{aligned} \tag{4.23}$$

The shapes of the distributions of the intrinsic residuals depend on the mass hypothesis of the particle. It is preferable to use the same observable in all parts of the likelihood. A variable that is independent of the mode hypothesis can be defined as

$$ID = \frac{\frac{\mathrm{d}E}{\mathrm{d}x}\Big|_{\mathrm{meas}} - \frac{\mathrm{d}E}{\mathrm{d}x}\Big|_{\mathrm{exp}-\pi}}{\frac{\mathrm{d}E}{\mathrm{d}x}\Big|_{\mathrm{exp}-K} - \frac{\mathrm{d}E}{\mathrm{d}x}\Big|_{\mathrm{exp}-\pi}}. \tag{4.24}$$

As a function of ID the residual is

$$\delta(ID) = (ID - \langle ID \rangle) \left(\frac{\mathrm{d}E}{\mathrm{d}x}\bigg|_{exp-K} - \frac{\mathrm{d}E}{\mathrm{d}x}\bigg|_{exp-\pi} \right) = (ID - \langle ID \rangle)\, \Delta, \tag{4.25}$$

where $\langle ID \rangle$ is the expected value of ID for a mass hypothesis. Δ is defined as the difference between the expected $\mathrm{d}E/\mathrm{d}x$ for the kaon and pion mass hypotheses and as such is a function of momentum.

The combined, two particle PDF for particle identification in terms of the ID for

each particle (ID_1 and ID_2) is

$$P(ID_1, ID_2) = \int_{-\infty}^{+\infty} P(\delta_1(ID_1) - c)P(\delta_2(ID_2) - c)\frac{\partial(\delta_1, \delta_2)}{\partial(ID_1, ID_2)}P(c)\mathrm{d}c \tag{4.26}$$

$$= \int_{-\infty}^{+\infty} P(\delta_1(ID_1) - c)P(\delta_2(ID_2) - c)\Delta_1\Delta_2 P(c)\mathrm{d}c. \tag{4.27}$$

The difference between the expected dE/dx for kaons and pions is a function of momentum, $\Delta_i(\alpha, p_{\mathrm{tot}})$, so changing variables from residuals to ID has introduced a momentum dependence into the PDF: it is now conditional on α and p_{tot}.

The PID distribution for each signal mode, j, is

$$P(ID_1, ID_2 \mid \alpha, p_{\mathrm{tot}}, s_j) = \frac{1}{\mathrm{K}_j^{\mathrm{PID}}}\int_{-\infty}^{+\infty} P_h(\delta_{1,h} - c)P_{h'}(\delta_{2,h'} - c)\Delta_1\Delta_2 P(c)\mathrm{d}c. \tag{4.28}$$

where $\mathrm{K}_j^{\mathrm{PID}}$ is the normalisation constant and the $\delta_{i,h'}$ are functions of ID and dependent on the particle hypothesis as are the $P_{h'}(\delta_{1,h'} - c)$.

The particle identification PDF for background uses the same joint two-particle PDF as signal. The total PDF is the sum over all possible particle pair combinations of the two-particle PDFs. The PDF for each pair hypothesis, (l, m), is weighted by a factor $w_l w_m$ where each weight, w_l, is proportional to the fraction of that type of particle, l, in the background sample. Contributions from pions, kaons, protons and electrons are considered. The contribution from muons is combined with the pion contribution since they are indistinguishable with the dE/dx resolution of the COT. The PID distribution for background is

$$P(ID_1, ID_2 \mid \alpha, p_{\mathrm{tot}}, b) = \sum_{l,m=\pi,K,p,e} w_l w_m \big(P(ID_1, ID_2 \mid \alpha, p_{\mathrm{tot}}, h, h') + P(ID_1, ID_2 \mid \alpha, p_{\mathrm{tot}}, h', h)\big) \tag{4.29}$$

and the weights w_l are free parameters in the fit.

The intrinsic residuals are modelled by three Gaussians while the correlation is modelled by the sum of two Gaussian distributions. The intrinsic residuals were pa-

rameterised using kaons and pions from $D^{*+} \to D^0\pi^+$ and the conjugate decays where $D^0 \to K^-\pi^+$ [59, 60]. It has been observed that the shape of the correlation is highly sensitive to the sub-sample of runs used and so the parameters describing the correlation were fitted using the $B^0_{(s)} \to h^+h'^-$ sample [41].

4.7 Momentum probability density function

The kinematic and particle identification PDFs are both dependent on α and $p_{\rm tot}$. The dependence is explicit for kinematic information and implicit for particle identification where it comes in because the expected $\mathrm{d}E/\mathrm{d}x$ of a particle is momentum dependent. It is therefore necessary to model the distribution of α and $p_{\rm tot}$ in the likelihood. This is the origin of the term $P(\alpha, p_{\rm tot})$. If the distributions are the same for each signal mode and for background then the term $P(\alpha, p_{\rm tot})$ factorises in the likelihood and so does not contribute when the log of the likelihood is maximised. If the distributions are different for each mode or between signal and background then the term no longer factorises. In this case neglecting this term will bias the fit. Neglecting to model distributions of any variables that vary on an event by event basis may lead to a bias, as will be discussed further in section 5.8

α and $p_{\rm tot}$ are not independent and so a joint PDF, $P(\alpha, p_{\rm tot})$, is used to parameterise their distribution. The selection requirements used to obtain this sample restrict the domain of α as a function of $p_{\rm tot}$. The trigger requirements on the transverse momentum of the final state particles mean that the particle momenta, p_1 and p_2, must be greater than 2 GeV/c and the sum of the momenta, $p_1 + p_2$, must be greater than 5.5 GeV/c. This means that α and $p_{\rm tot}$ are restricted to lie within the domain defined by

$$p_1 = p_{\rm tot}\left(\frac{1-|\alpha|}{2-|\alpha|}\right) > 2 \text{ GeV/c} \tag{4.30}$$

and

$$p_{\rm tot} > 5.5 \text{ GeV/c}. \tag{4.31}$$

With these two constraints the remaining condition that $p_2 = p_{\rm tot}/(2-|\alpha|) > 2$ GeV/c is automatically satisfied. The PDFs for signal were empirically determined from fits to realistic simulations of the signal decays. The PDFs for background were empirically determined from a fit to data in the mass sidebands: $5.000 < M_{\pi\pi} < 5.125$ GeV/c^2 and $5.600 < M_{\pi\pi} < 6.200$ GeV/c^2 [55].

The joint PDF for signal can be written as the product of the probability of measuring $p_{\rm tot}$ given the decay mode hypothesis s_j, $P(p_{\rm tot} \mid s_j)$, and the probability of measuring α given $p_{\rm tot}$ and the mode hypothesis, $P(\alpha \mid p_{\rm tot}, s_j)$. The joint probability density for signal is

$$\begin{aligned} P(\alpha, p_{\rm tot} \mid s_j) &= P(p_{\rm tot} \mid s_j) P(\alpha \mid p_{\rm tot}, s_j) \\ &= \frac{1}{\mathrm{K}_j^{\rm mom}} e^{a_5 p_{\rm tot}} \sum_{i=0}^{4} \left(a_i p_{\rm tot}^i\right) \sum_{k=0}^{6} b_k \alpha^k \left(\frac{p_{\rm tot}-2}{p_{\rm tot}-4}\right)^k . \end{aligned} \tag{4.32}$$

The PDF for $p_{\rm tot}$ is an exponential function multiplied by a 4th-degree polynomial with coefficients a_i. $P(\alpha \mid p_{\rm tot}, s_j)$ is a 6th-degree polynomial, with coefficients b_k, scaled by a factor $(p_{\rm tot}-2)/(p_{\rm tot}-4)$ which comes from the constraint on the domain of α and $p_{\rm tot}$ in equations 4.30 and 4.31. The constants, a_i and b_k, are different for each decay mode. The decays $\mathrm{B}^0 \to \pi^+\pi^-$ and $\mathrm{B}_\mathrm{s}^0 \to \mathrm{K}^+\mathrm{K}^-$ are kinematically symmetric so the odd b_k terms are fixed to zero. The constants were determined using a 2-dimensional, binned, maximum likelihood fit to a detailed simulation for the analysis described in [41]. The normalisation constant, $\mathrm{K}_j^{\rm mom}$, was calculated with a numerical 2-dimensional integration of the PDF in the allowed domain of α and $p_{\rm tot}$. This analysis uses the same data sample and selection requirements as [41] so the same constants are used.

The joint PDF for background is written as the product of the probability of measuring $p_{\rm tot}$ given the event is background b, $P(p_{\rm tot} \mid b)$, and the probability of

measuring α given $p_{\rm tot}$ and given the event is background, $P(\alpha \mid p_{\rm tot}, b)$. The joint probability density for background is

$$\begin{aligned} P(\alpha, p_{\rm tot} \mid b) &= P(p_{\rm tot} \mid b) P(\alpha \mid p_{\rm tot}, b) \\ &= \frac{1}{\mathrm{K}_b^{\rm mom}} \left(1 + \left(\frac{p_{\rm tot} - \lambda}{a}\right)^2\right)^{-m} e^{\left(-\nu \tan^{-1}\left(\frac{p_{\rm tot}-\lambda}{a}\right)\right)} \sum_{k=0}^{6} b_k \alpha^k \left(\frac{p_{\rm tot} - 2}{p_{\rm tot} - 4}\right)^k . \end{aligned} \tag{4.33}$$

The $p_{\rm tot}$ distribution is a Pearson type-IV distribution [61] with free parameters λ, a, m and ν. $P(\alpha \mid p_{\rm tot}, b)$ is a 6th-degree polynomial scaled by the factor $(p_{\rm tot} - 2)/(p_{\rm tot} - 4)$ which comes from the limits on the domain of α and $p_{\rm tot}$. No difference in the momentum distributions of positive and negative particles is expected so the distribution is forced to be symmetric by fixing the odd b_k terms to zero. The constants and the normalisation, $\mathrm{K}_b^{\rm mom}$, were determined from a 2-dimensional binned maximum likelihood fit to a sample of background events from data [41]. The background sample used was taken from the mass range $5.6 < M_{\pi\pi} < 6.2$ GeV/c which lies above the signal mass peak and is combinatorial background and the mass range $5.000 < M_{\pi\pi} < 5.125$ GeV/c which lies below the signal mass peak and contains a mixture of combinatorial background and partially reconstructed B meson decays. Figure 4.8 shows the mass regions used to determine the constants in equation 4.33. The same α, $p_{\rm tot}$ distribution is used for both types of background since little contamination from partially reconstructed decays is expected underneath the signal peak. This analysis uses the same constants as were determined in [41].

4.8 Separation only fit

The composition of the data sample described in sections 3.2 to 3.4 is determined by maximising the likelihood described in section 4.4 for the data sample. Figures 4.9, 4.10 and 4.11 show the projections of the likelihood fit onto the variables used in

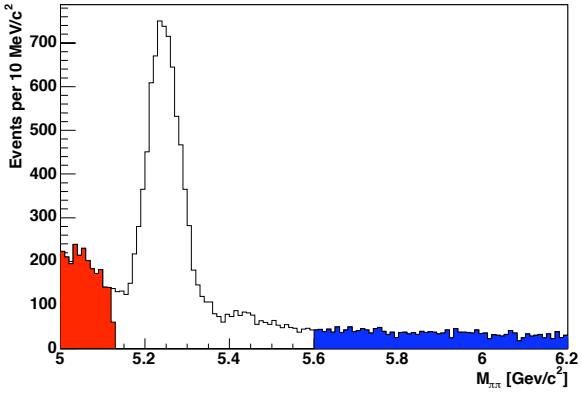


Figure 4.8: The mass regions used to parameterise the distribution $P(\alpha, p_{\text{tot}} \mid b)$ are indicated in red (low mass sideband) and blue (high mass sideband).

the likelihood. Also shown are the residual distributions for each variable, where the residual is the difference between the binned projected fit distribution and the binned distribution in data divided by the statistical uncertainty on the distribution in data for each bin. In addition to the projections for the variables in the likelihood, the projection of the fit onto the sum and difference of the *ID* variables is shown. The quantity $(ID_1 - ID_2)$ is independent of the correlation between ID_i for each particle while $(ID_1 + ID_2)$ has the effect of the correlation doubled so these distributions can be used to check that the both the particle identification resolution and correlation are modelled correctly. The separation fit models the distribution of each of the variables used in the likelihood well.

The fractions of each decay mode measured by the separation likelihood fit are summarised in table 4.2. Also summarised in the same table are the fractions as measured by an earlier analysis that used the same data sample [41]. Both analyses are based on the work in [52] and many improvements were implemented in [41], most of which were included in this analysis. The main difference between this likelihood fit and that in the earlier analysis [41] is in the parameterisation of the signal mass model. In the earlier analysis an invariant mass template was determined using a custom simulation which included the effects of final state radiation as well as decay kinematics and detector resolution effects. The parameters of the template were then determined using $\mathrm{D}^0 \to \mathrm{K}\pi$ decays. Final state radiation leads to a high tail in the mass distributions

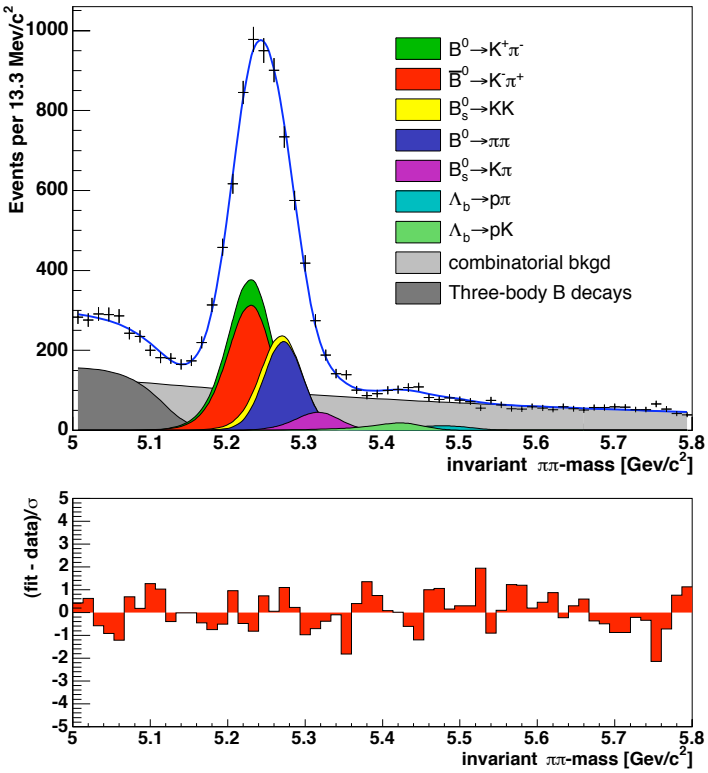
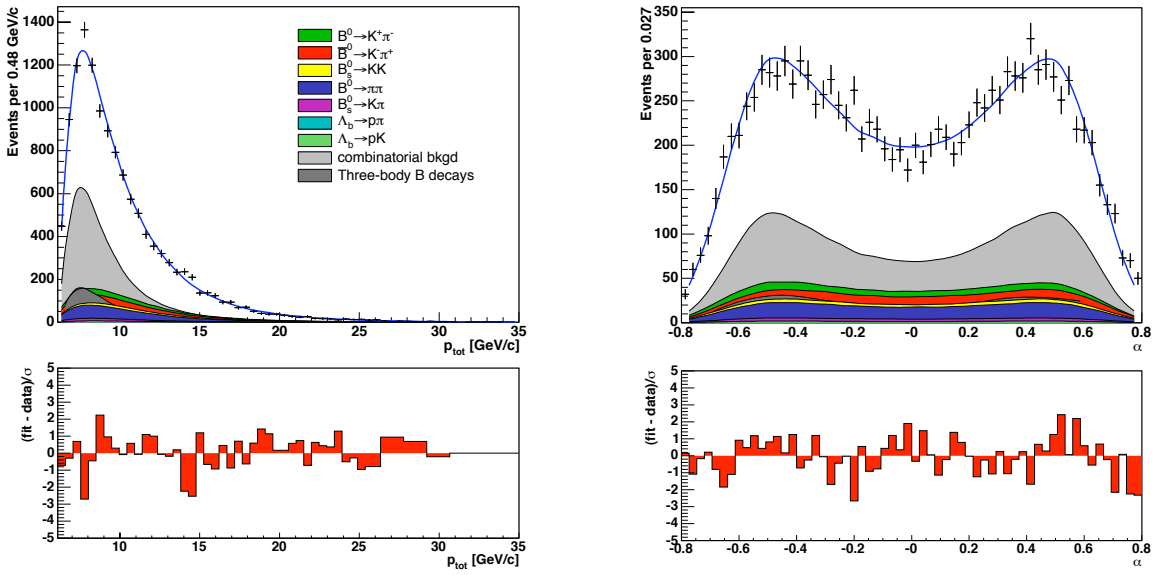
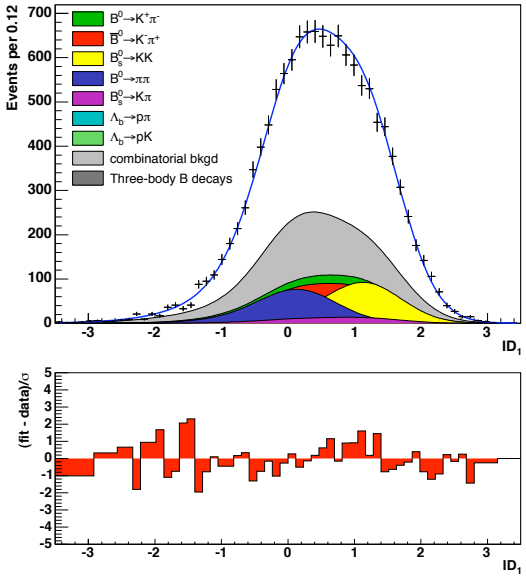


Figure 4.9: $\pi\pi$-invariant mass projection of the separation fit.

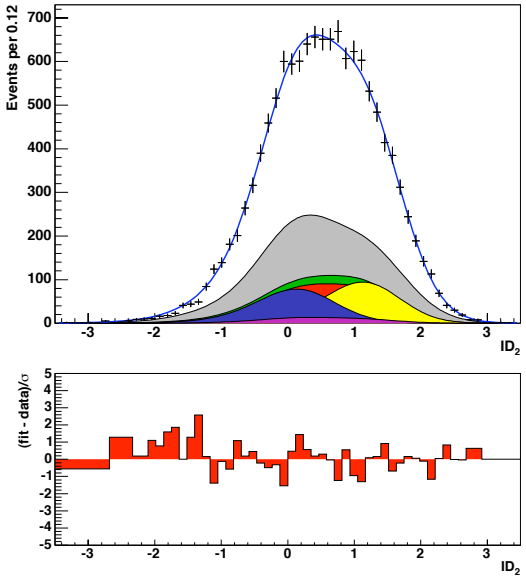


(a) p_{tot} (b) α

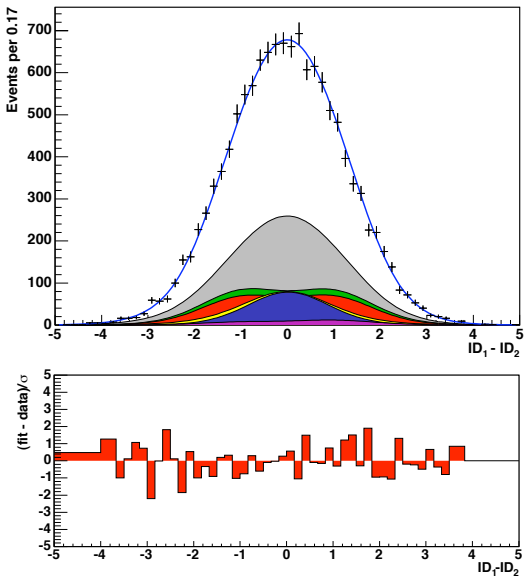
Figure 4.10: Projection of the separation fit onto the α and p_{tot} variables.



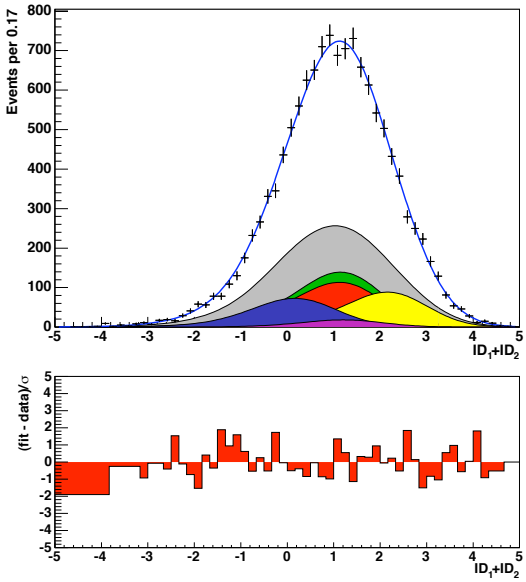
(a) ID_1



(b) ID_2



(c) $ID_1 - ID_2$



(d) $ID_1 + ID_2$

Figure 4.11: Projection of the separation fit onto the ID_i variables.

Parameter	This analysis		Previous analysis [41]	
	fit	error	fit	error
$B^0 \to K^+\pi^-$ fraction	0.580	0.012	0.577	0.010
$B_s^0 \to K^+K^-$ fraction	0.1863	0.0090	0.186	0.009
$B^0 \to \pi^+\pi^-$ fraction	0.1581	0.0093	0.160	0.009
$B_s^0 \to K^+\pi^-$ fraction	0.0378	0.0054	0.035	0.006
$\Lambda_b^0 \to pK^-$ fraction	0.0212	0.0040	0.022	0.003
$\Lambda_b^0 \to p\pi^-$ fraction	0.0169	0.0036	0.015	0.003
$\frac{f_{\overline{B^0}\to\pi^+K^-} - f_{B^0\to K^+\pi^-}}{f_{\overline{B^0}\to\pi^+K^-} + f_{B^0\to K^+\pi^-}}$	-0.095	0.023	-0.092	0.023
$\frac{f_{B_s^0\to K^+\pi^-} - f_{B_s^0\to\pi^+K^-}}{f_{B_s^0\to K^+\pi^-} + f_{B_s^0\to\pi^+K^-}}$	0.42	0.18	0.48	0.19
$\frac{f_{\Lambda_b^0\to K^+\overline{p}} - f_{\Lambda_b^0\to pK^-}}{f_{\Lambda_b^0\to K^+\overline{p}} + f_{\Lambda_b^0\to pK^-}}$	-0.27	0.22	-0.29	0.19
$\frac{f_{\overline{\Lambda_b^0}\to\pi^+\overline{p}} - f_{\Lambda_b^0\to p\pi^-}}{f_{\overline{\Lambda_b^0}\to\pi^+\overline{p}} + f_{\Lambda_b^0\to p\pi^-}}$	0.06	0.21	0.06	0.21
background fraction	0.4839	0.0075	0.481	0.008

Table 4.2: Summary of the fraction of each $B_{(s)}^0 \to h^+h'^-$ and $\Lambda_b^0 \to ph^-$ mode determined by the separation fit in this analysis and also in the previous analysis on the same data sample [41]. "$B^0 \to K^+\pi^-$ fraction" implies the fraction of signal that is either $B^0 \to K^+\pi^-$ or $\overline{B^0} \to \pi^+K^-$ and similarly for the other $B_{(s)}^0 \to h^+h'^-$ and $\Lambda_b^0 \to ph^-$ modes. f_j is the fraction of signal which is is of mode j (not including the conjugate decay).

and neglecting this affect primarily affects the rare decay modes. The second difference between this fit and the earlier analysis is in the parameterisation of the joint distribution of α and $p_{\rm tot}$ in background. The earlier analysis uses only the high mass sideband to parameterise the distribution, neglecting the contribution due to physics-type background. In section 7.10 alternative ways of determining the background momentum distribution will be considered and the final measurement actually uses a parameterisation which is closer to that used in [41]. Both this fit and the earlier analysis use very similar likelihoods for the statistical separation of the $B_{(s)}^0 \to h^+h'^-$ and $\Lambda_b^0 \to ph^-$ decay modes so a close agreement is expected and obtained.

4.9 Summary

The data sample used in this analysis includes contributions from several $\mathrm{B}^0_{(s)} \to h^+h'^-$ and $\Lambda^0_\mathrm{b} \to \mathrm{p}h^-$ decay modes as well as combinatorial background and background from partially reconstructed three-body B meson decays. It is not possible to separate these contributions on an event by event basis; however, statistical separation of the contributions is possible using kinematic information (the invariant $\pi\pi$-mass, scalar sum of the hh' momenta, p_tot, and the charge weighted momentum imbalance, α) and particle identification using $\mathrm{d}E/\mathrm{d}x$ in the COT. Section 4.4 described the likelihood function used to statistically determine the contributions to the sample. This includes terms describing the distributions of $M_{\pi\pi}$ and of $\mathrm{d}E/\mathrm{d}x$ dependent variables which are both, in turn dependent on α and p_tot. This dependence requires that a term describing the joint distribution of α and p_tot is included to complete the likelihood.

This likelihood function is used to determine the composition of the data sample used for this analysis (and described in chapter 3). The fractions of each contribution to the sample obtained here agree well with those obtained by the previous analysis which used the same data sample and a very similar method.

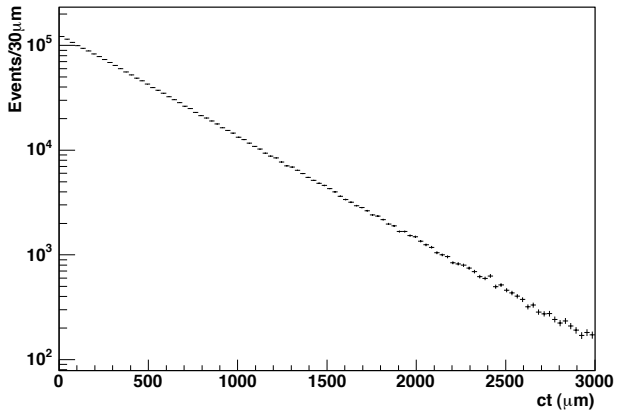
Chapter 5

Measuring an unbiased lifetime

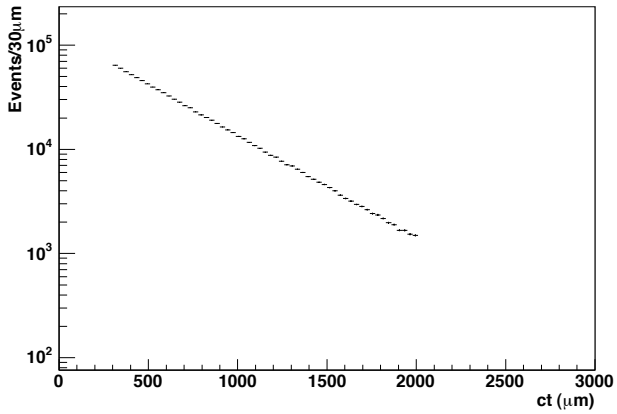
5.1 The effect of a hadronic trigger on a lifetime distribution

An unbiased sample of long-lived particles has a distribution of lifetimes which is exponential (as shown in figure 5.1a). Fitting this distribution with an exponential function to find the mean lifetime is a simple matter. If the range of lifetimes measured is limited by a cut applied directly to the lifetime (figure 5.1b) the mean lifetime can still be found by fitting an exponential function to the distribution within the allowed range of lifetimes. If cuts are applied indirectly to the lifetime they may shape the lifetime distribution as shown in figure 5.1c. If these cuts are not corrected for then the measurement of the mean lifetime in this sample will be biased.

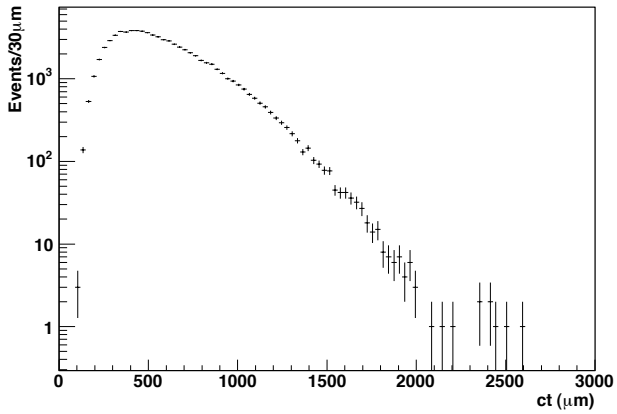
Section 2.8 describes the hadronic trigger used to reject uninteresting events and select events consistent with B decays to two hadrons. This trigger places cuts on the geometry and kinematics of the decays. While the measured lifetime is independent of many of these cuts some of them, such as the cuts on track impact parameter (IP) and on the displacement of the decay vertex from the interaction point in the transverse plane (L_{T}) do affect the measured lifetime distribution. Figure 5.2 demonstrates how



(a) Unbiased lifetime distribution.



(b) Lifetime distribution biased by direct cuts on lifetime.



(c) Lifetime distribution biased by indirect cuts on lifetime.

Figure 5.1: The distribution of measured lifetime, ct, in three samples of $B_s^0 \to K^+K^-$ events from a Monte Carlo simulation. Figure 5.1a shows the lifetime distribution before any event selection is made. Figure 5.1b shows the lifetime distribution after events have been selected by placing a cut directly on the lifetime. Figure 5.1c shows the lifetime distribution after the hadronic trigger has been used to select events, placing cuts which indirectly cut on the lifetime.

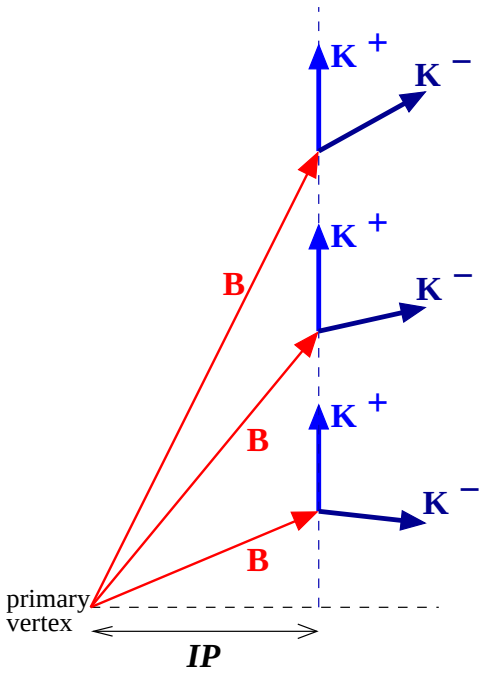


Figure 5.2: The topologies for three decays of a B meson to two kaons are shown. In each case the magnitude of the momentum of the B meson is the same so the length of the red arrow indicates the lifetime of the B. It can be seen that a single value of impact parameter for the K^+ can correspond to a range of lifetimes for the B meson.

a single impact parameter measurement can translate to a range of measured lifetimes. A single measurement of L_T will also correspond to a range of lifetime measurements depending on the momentum of the B meson. Figure 5.1c shows the effect of the hadronic trigger on the distribution of measured lifetimes in a simulated sample of $B_s^0 \to K^+K^-$ events. Averaged over all events the translation between a cut on IP or on L_T to a cut on lifetime is non linear and it is necessary to account for the bias this imposes on a measurement of lifetime.

One method commonly used to correct for the trigger bias requires a detailed simulation of the detector and trigger. By simulating a sample of events and passing them through the detailed detector and trigger simulations a graph, similar to figure 5.1c, of the lifetime distribution for the now biased sample can be made. By comparing the biased distribution to the original, unbiased, lifetime distribution the efficiency of the trigger and detector as a function of lifetime can be obtained. This lifetime efficiency function can be used to correct for the bias imposed by the trigger. This method has

been used successfully in analyses at CDF for example [62, 53]. The disadvantage of this method is that the event, detector and trigger simulation has to very accurately model what happens in data. This not only adds a systematic uncertainty to the measurement but as the size of the data samples increases a much larger and more accurate simulated sample is required to keep the systematic errors smaller than the statistical uncertainty on any measurement. In addition, any change in the selection requirements of the sample or the addition of data produced under different running conditions requires the production of a new, large, simulated sample. This thesis utilises a different method, that does not require a detailed simulation of events, detector and trigger, for correcting for the bias caused to a lifetime measurement by the use of a hadronic trigger. A full description of this method follows.

5.2 A data only approach to account for a trigger bias

It has already been stated that, averaged over all events, there is no linear relationship between a cut on impact parameter or L_{T} and the effective cut on lifetime this creates. If the geometry and kinematics of an event are fixed, allowing only the lifetime of the B meson to vary, then for this event each track impact parameter corresponds to a single lifetime. This is shown in figure 5.3. The figure shows a long-lived B meson decaying to two hadrons. The momenta of all the particles are fixed but the lifetime of the B meson is allowed to vary. Each lifetime corresponds to a single impact parameter for each track (for simplicity this is shown for one track, the K^+, only). Given the kinematics and geometry of the decay the effective cut on lifetime can be calculated from the direct cut on IP for a single event. This is similarly true for the cut on L_{T}.

For each event, the range of lifetimes which could have been accepted by the trigger is calculated by sliding the position of the B along its trajectory, thus varying its

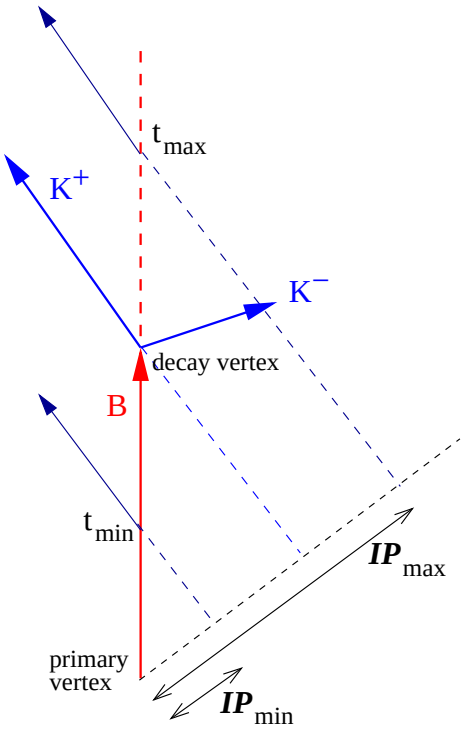


Figure 5.3: The relationship between impact parameter and lifetime. If the geometry and kinematics of the decay are fixed and only the lifetime of the B is allowed to vary then each lifetime corresponds to a single impact parameter (IP) for each decay product track.

lifetime, while keeping the kinematic quantities of the decay fixed. If the decay would have satisfied the trigger requirements at a given point then the lifetime at that point falls within the acceptance. This is demonstrated in figure 5.4. The resulting lifetime acceptance function is a single top hat function with a minimum and a maximum accepted lifetime, t_{min} and t_{max}.

For signal only, neglecting measurement errors for the moment, the probability that an event has a lifetime t can be written as the product of the probability that the event has a lifetime t given that t must lie between t_{min} and t_{max} and the probability that

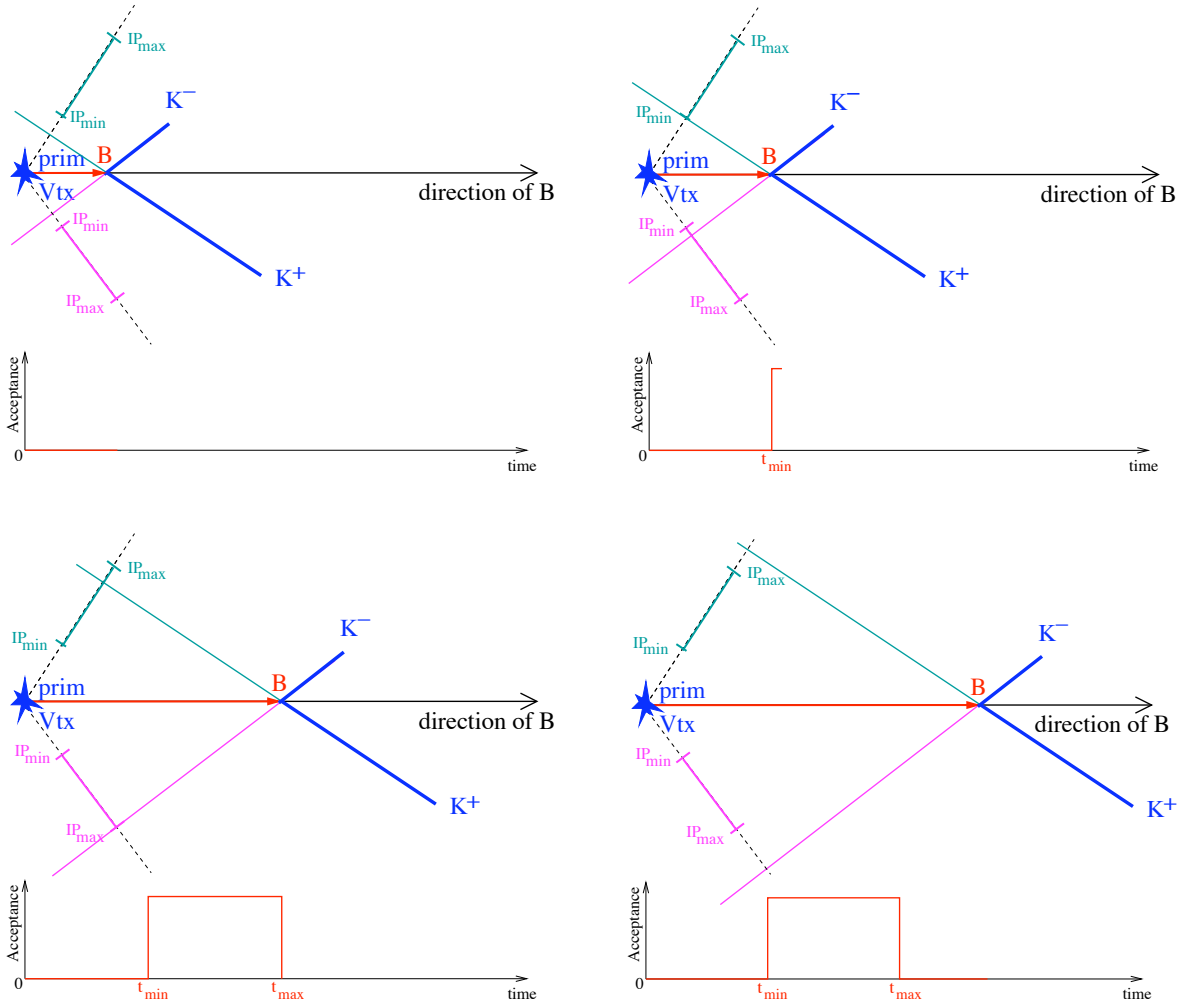


Figure 5.4: A B_s^0 meson decays to two kaons. If the kinematics of the decay are kept fixed and the lifetime of the B_s^0 is increased from zero then at each time t the parameters of the event can be checked to confirm if the trigger conditions are fulfilled for that t. In this way the acceptance function for this event is found. It is a single top hat function with minimum and maximum times $t_{\min}$ and $t_{\max}$. The diagram illustrates this process for impact parameter cuts. L_{T} cuts are included in the same way.

the limits the lifetime is constrained by are $t_{\min}$ and $t_{\max}$, that is,

$$P(t) = P(t \mid t_{\min}, t_{\max}) P(t_{\min}, t_{\max}) \tag{5.1}$$

$$= \frac{\frac{1}{\tau} e^{\frac{-t}{\tau}}}{\int_{t_{\min}}^{t_{\max}} \frac{1}{\tau} e^{\frac{-t'}{\tau}} \mathrm{d}t'} P(t_{\max}, t_{\min}) \tag{5.2}$$

$$= \frac{\frac{1}{\tau} e^{\frac{-t}{\tau}}}{e^{\frac{-t_{\min}}{\tau}} - e^{\frac{-t_{\max}}{\tau}}} P(t_{\max}, t_{\min}) \tag{5.3}$$

for $t_{\min} \leq t \leq t_{\max}$ and $P(t) = 0$ for all other t. The probability, $P(t \mid t_{\min}, t_{\max})$ is an exponential lifetime distribution normalised to lie within $t_{\min}$ and $t_{\max}$. The second term, $P(t_{\max}, t_{\min})$, is the probability an event has minimum and maximum accepted lifetimes $t_{\min}$ and $t_{\max}$ and it depends on impact parameter cuts and on the kinematics of the decay but is independent of measured event lifetime.

For a series of N measurements of lifetime the mean lifetime, τ, of the decay can be estimated by maximising the log of the likelihood for τ. The likelihood for τ can be written as the product of the probabilities for each measured time t_i. For a set of ideal decays (no measurement errors, no background) the likelihood is

$$\mathcal{L} = \prod_{i=1}^{N} P_i(t_i) \tag{5.4}$$

$$= \prod_{i=1}^{N} \big(P(t_i \mid t_{\max\, i}, t_{\min\, i}) \cdot P(t_{\max\, i}, t_{\min\, i}) \big), \tag{5.5}$$

where $t_{\min\, i}$ and $t_{\max\, i}$ are the minimum and maximum lifetimes which satisfy the trigger requirements for the i^{th} event. The log of the likelihood is

$$\log \mathcal{L} = \sum_{i=1}^{N} \log P(t_i \mid t_{\max\, i}, t_{\min\, i}) + \sum_{i=1}^{N} \log P(t_{\max\, i}, t_{\min\, i}). \tag{5.6}$$

$P(t_{\max\, i}, t_{\min\, i})$ is independent of the measured lifetime for the event so the second summation in equation 5.6 is a constant which disappears when the log-likelihood is maximised. Note that this is only true in this simple case of signal only. When

background events are also included this term no longer vanishes. This is discussed further in section 5.8.

Using equation 5.3, the log-likelihood becomes

$$\log \mathcal{L} = -N \log \tau - \sum_{i=1}^{N} \left(\frac{t_i}{\tau} + \log \left(e^{\frac{-t_{\min i}}{\tau}} - e^{\frac{-t_{\max i}}{\tau}} \right) \right). \tag{5.7}$$

Note that this differs from the likelihood for an unbiased sample (without any effective lifetime cuts) only by the addition of the term

$$\log \mathcal{L}_{\text{cuts}} = -\sum_{i=1}^{N} \log \left(e^{\frac{-t_{\min_i}}{\tau}} - e^{\frac{-t_{\max i}}{\tau}} \right). \tag{5.8}$$

Equation 5.3 was derived assuming the lifetime was measured exactly and that the cuts on impact parameter and L_{T} were made precisely. In reality there is an uncertainty on all of these measurements and this needs to be included in the likelihood.

5.3 Incorporating measurement errors

Equation 5.3 gives the probability an event has a true lifetime t given a minimum and maximum accepted lifetime. In fact a lifetime t_0 is measured with a measurement uncertainty σ_t. In this analysis it is assumed that the resolution of the measured lifetime is Gaussian distributed, this is discussed further in section 7.16.2. Assume for the moment that the trigger is applied to the offline reconstructed quantities. The lifetime acceptance function is calculated in exactly the same way as before, as shown in figure 5.4, only now the trigger is applied to offline reconstructed quantities and the acceptance is a function of measured rather than true lifetime. The probability of measuring a lifetime t_0 can be written as an integral over all true decay times t in terms of the following functions:

- The probability that a particle decays with a true lifetime t, given a mean lifetime τ,

$$P(t \mid \tau) = \frac{1}{\tau} e^{\frac{-t}{\tau}}; \tag{5.9}$$

- The probability that the measured lifetime is t_0 given the true decay time t and a measurement uncertainty σ_t,

$$P(t_0 \mid t, \sigma_t) = \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t_0)^2}{2\sigma_t^2}}; \tag{5.10}$$

- The lifetime acceptance as a function of the measured lifetime, t_0,

$$A(t_0).$$

For a single event, the probability of measuring a time t_0 given a mean lifetime τ, an acceptance A and an uncertainty on the measured time σ_t is

$$P(t_0 \mid A, \tau, \sigma_t) = \frac{\int_0^\infty \frac{1}{\tau} e^{-\frac{t}{\tau}} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t_o)^2}{2\sigma_t{}^2}} A(t_0) \mathrm{d}t}{\int_{-\infty}^\infty \int_0^\infty \frac{1}{\tau} e^{-\frac{t}{\tau}} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t_o)^2}{2\sigma_t{}^2}} A(t_0) \mathrm{d}t \, \mathrm{d}t_0}. \tag{5.11}$$

Assuming the trigger cuts are applied to offline quantities, the acceptance is given by

$$A(t_0) = \theta(t_0 - t_{\min}) - \theta(t_0 - t_{\max}) \tag{5.12}$$

where θ is the Heaviside step function. Equation 5.12 is a simple top hat function. It has this simple shape because a linear relationship between impact parameter and lifetime and between L_{T} and lifetime has been established on an event by event basis.

5.4 The signal lifetime probability with different online and offline quantities

In the above it was assumed that the trigger was applied to offline quantities. In fact, the level 2 trigger uses silicon vertex tracker (SVT) quantities that differ from the final, offline quantities used to reconstruct the lifetime. It is now necessary to re-establish a direct link between the online, SVT measured quantities and the offline measured lifetime. For every event the online impact parameters used by the trigger are known so it is possible to calculate the difference between online and offline impact parameter, Δd_0, for each track, in each event. For a single event the acceptance function is calculated by sliding the event along its trajectory, varying its lifetime. The offline impact parameter, $d_0^{\rm off}$, for each track is calculated at each lifetime. The expected SVT measured impact parameter ($d_0^{\rm SVT}$) for each track can be calculated from $d_0^{\rm off}$, as

$$d_0^{\rm SVT} = d_0^{\rm off} + \Delta d_0. \tag{5.13}$$

where it is assumed that Δd_0 is independent of the measured impact parameter. As the B lifetime is varied the event is checked to confirm whether the calculated SVT track impact parameters pass the trigger. In this way cuts on SVT measured quantities are translated into cuts on measured lifetime. This is illustrated in figure 5.5. The distribution of Δd_0 observed in data is shown in figure 5.6. Figure 5.7 demonstrates that Δd_0 is independent of the measured lifetime.

The level 3 trigger uses a version of the offline reconstruction code different from the final offline reconstruction. In the same way that SVT measured quantities must be used to confirm the level 2 trigger it is important to use the track parameters level 3 had to confirm its trigger decision. The level 3 impact parameters and $L_{\rm T}$ are translated into offline impact parameters in the same way as for SVT quantities. For each track in each event the difference between L3 and offline impact parameter, $\Delta d_0^{\rm L3}$, is calculated.

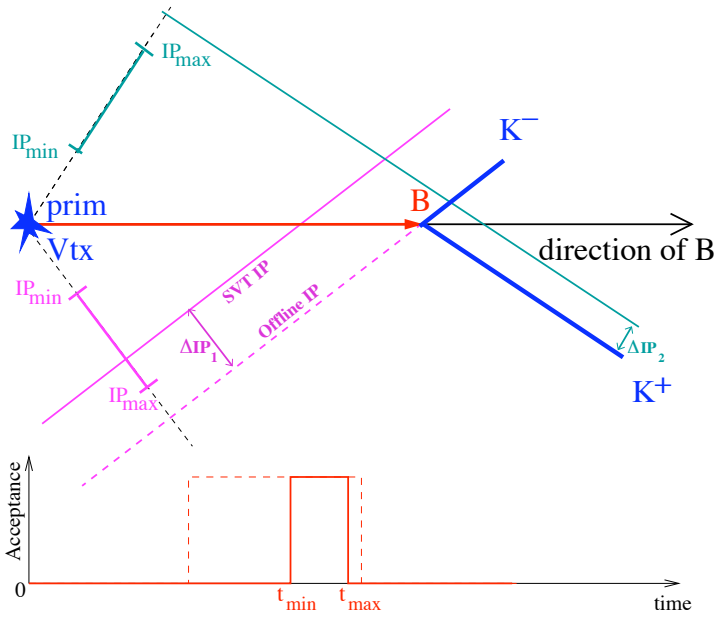


Figure 5.5: Keeping the kinematics of the decay fixed, as the lifetime of the B is varied along its trajectory the impact parameter that would have been measured by the SVT is calculated using the measured difference between offline and SVT measured impact parameter, ΔIP, for that track. In this way, cuts on the SVT measured impact parameter can be translated into cuts on lifetime. The dashed red line indicates the lifetime acceptance function calculated using offline measured impact parameters. The solid red line is the acceptance function calculated using online quantities. ΔIP may be different for each track and each event

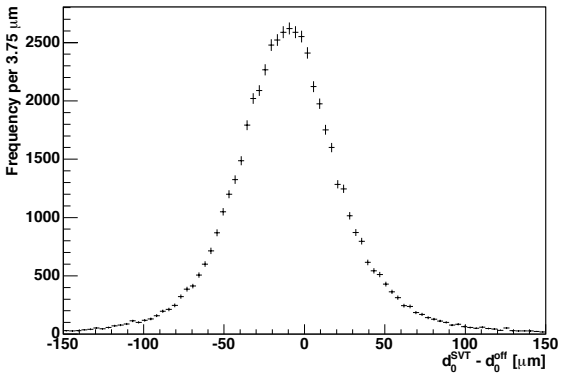


Figure 5.6: The distribution of $\Delta d_0 = d_0^{\rm SVT} - d_0^{\rm off}$ observed in data.

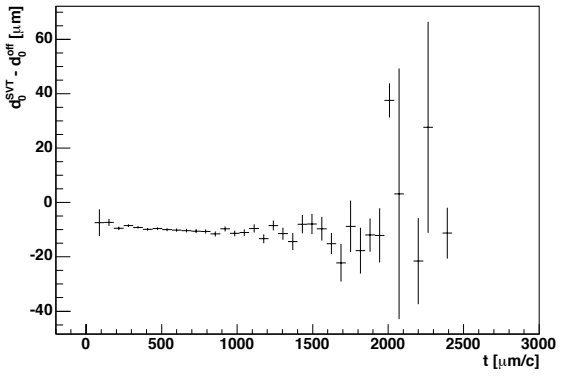


Figure 5.7: The mean value of $\Delta d_0 = d_0^{\rm SVT} - d_0^{\rm off}$ as a function of the measured B lifetime t.

For a single event the acceptance function is calculated by sliding the event along its trajectory, varying its lifetime. As the lifetime is varied so is the offline impact parameter, $d_0^{\rm off}$ of each track. The expected level 3 measured impact parameter for each track, $d_0^{\rm L3}$, is calculated from the offline impact parameter as $d_0^{\rm L3} = d_0^{\rm off} + \Delta d_0^{\rm L3}$. The level 3 trigger is applied to the calculated level 3 impact parameter and $L_{\rm T}$ to obtain the acceptance function for the offline measured lifetime.

5.5 The discrete SVT impact parameter

The SVT uses fast, integer arithmetic to calculate track parameters and returns impact parameters that are multiples of 10 μm ($d_0 = 0$ μm, ± 10 μm, ± 20 μm, ± 30 μm...). Figure 5.8 shows the distribution of SVT measured impact parameter in data collected by the B_PIPI trigger. It has been demonstrated that this discretisation of the SVT impact parameter biases the lifetime fit [63]. A fit to 15k simulated $\mathrm{B}^+ \to \mathrm{D}^0\pi^+$ events passed through a detailed detector and trigger simulation, neglecting the effect of the discretised SVT, yielded a result of $c\tau = 448 \pm 6$ μm compared to an input value of 496 μm. To include this effect in the fit the acceptance function is calculated by sliding the event along its trajectory but instead of calculating $d_0^{\rm SVT}$ simply as $d_0^{\rm SVT} = d_0^{\rm off} + \Delta d_0$ the result is rounded to the nearest 10 μm. Including this correction, the fit to the simulated $\mathrm{B}^+ \to \mathrm{D}^0\pi^+$ events gave a result of 494 ± 7 μm that is in good agreement with the input value of 496 μm.

5.6 The full, signal only, lifetime probability for one decay mode

Section 5.2 established a linear relationship between offline measured impact parameter and lifetime and between offline measured $L_{\rm T}$ and lifetime on an event by event

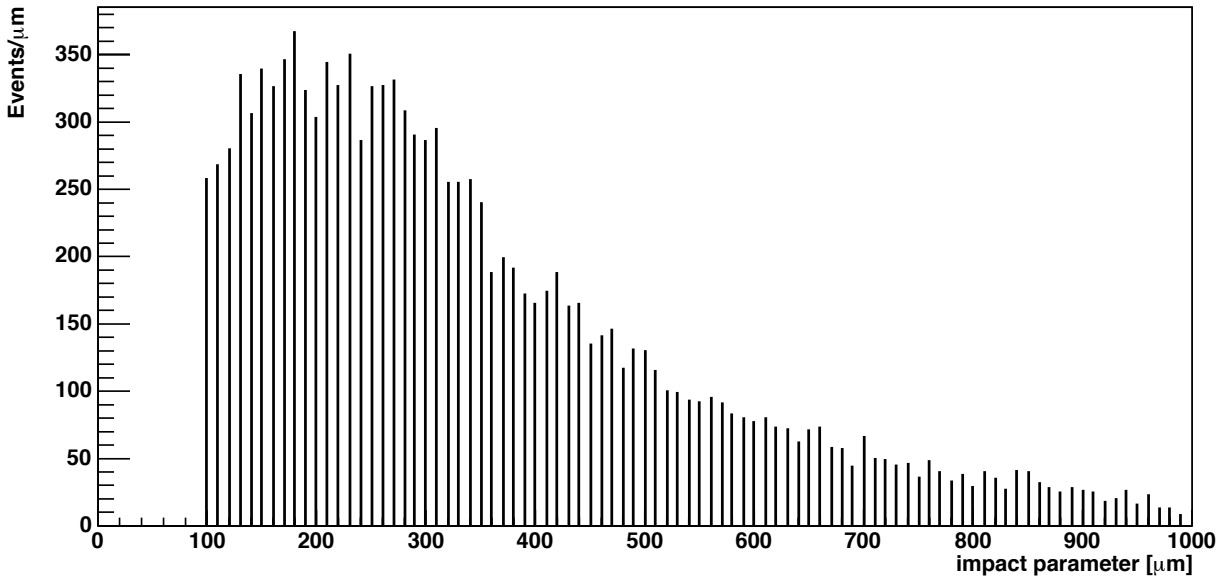


Figure 5.8: The distribution of SVT-measured impact parameter in data. The SVT returns impact parameters that are multiples of 10 μm. The data has been selected by the BPIPI trigger that requires a minimum impact parameter of 100 μm.

basis. Section 5.4 established that there is also a linear relationship between the online measured impact parameters and L_{T} and the offline measured lifetime. This linear relationship means the acceptance function is still a top hat function that can be described by equation 5.12,

$$A(t_0) = \theta(t_0 - t_{\mathrm{min}}) - \theta(t_0 - t_{\mathrm{max}}). \tag{5.14}$$

Using this, the probability of measuring a lifetime t_0 for signal only, given in equation 5.11, becomes

$$P(t_0 \mid \tau, t_{\mathrm{min}}, t_{\mathrm{max}}, \sigma_t) = \frac{\int_0^\infty \frac{1}{\tau} e^{-\frac{t}{\tau}} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t_o)^2}{2\sigma_t{}^2}} \mathrm{d}t}{\int_{t_{\mathrm{min}}}^{t_{\mathrm{max}}} \int_0^\infty \frac{1}{\tau} e^{-\frac{t}{\tau}} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t_o)^2}{2\sigma_t{}^2}} \mathrm{d}t\, \mathrm{d}t_0} \tag{5.15}$$

for $t_{\mathrm{min}} \leq t_0 \leq t_{\mathrm{max}}$ and $P(t_0 \mid \tau, t_{\mathrm{min}}, t_{\mathrm{max}}, \sigma_t) = 0$ for all other t_0.

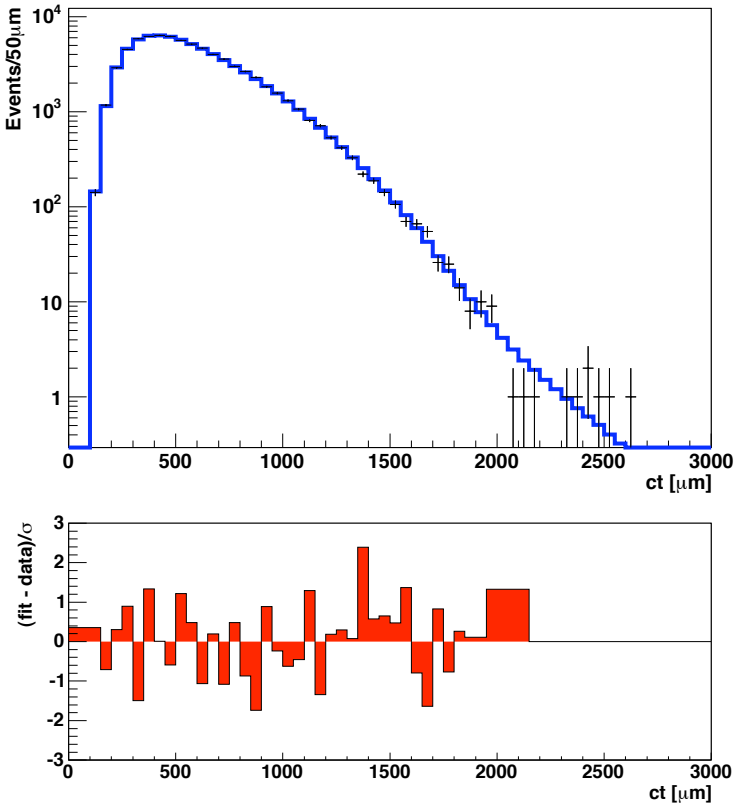


Figure 5.9: *ct* projection for a lifetime fit to a simulated sample of $B_s^0 \to K^+K^-$ events (upper plot). The lower plot shows the residual distribution from this fit.

5.7 Detailed Monte Carlo simulation

To test the basic method the fit was performed on a sample of $B_s^0 \to K^+K^-$ events from the realistic simulation described in section 3.5. The input lifetime was 438 μm/c and the fit returned a lifetime of 441.5 ± 2.9 μm/c which is 1.2σ from the truth value. Figure 5.9 shows the projection of the lifetime fit.

5.8 Including background

The method described above can be used to measure, without using a Monte Carlo simulation, a lifetime using a sample of signal events that have been biased by the hadronic trigger. The PDF so far is the probability of measuring a lifetime, t_0, given the minimum and maximum accepted lifetimes, $t_{\rm min}$ and $t_{\rm max}$, for an event. $t_{\rm min}$ and

$t_{\max}$ are calculated from the kinematics of the decay which translate trigger requirements into limits on measured lifetime and are independent of the lifetime itself. The probability of observing $t_{\min}$ and $t_{\max}$, and hence the decay kinematics, was neglected because the kinematics are independent of lifetime and so the probability factored out of the likelihood. Mathematically, in the expression for the probability of observing a particular combination of lifetime, $t_{\min}$ and $t_{\max}$,

$$P(t_0, t_{\min}, t_{\max}) = P(t_0 \mid t_{\min}, t_{\max}) P(t_{\min}, t_{\max}), \tag{5.16}$$

the term $P(t_{\min}, t_{\max})$ can be neglected because it is a simple factor which is independent of lifetime. In a sample which includes background and signal events, however, the full expression becomes

$$\begin{aligned} P(t_0, t_{\min}, t_{\max}) = {} & P(s) P(t_0 \mid t_{\min}, t_{\max}, s) P(t_{\min}, t_{\max} \mid s) \\ & + P(b) P(t_0 \mid t_{\min}, t_{\max}, b) P(t_{\min}, t_{\max} \mid b), \end{aligned} \tag{5.17}$$

where $P(s)$ is the probability an event is signal and $P(b) = 1.0 - P(s)$ is the probability it is background. Now the terms $P(t_{\min}, t_{\max} \mid s)$ and $P(t_{\min}, t_{\max} \mid b)$ only factor out if they are the same. Neglecting these terms when they are different is equivalent to getting the signal and background fractions wrong. This is more apparent if equation 5.17 is rewritten as

$$\begin{aligned} P(t_0, t_{\min}, t_{\max}) = {} & \big(P(s \mid t_{\min}, t_{\max}) P(t_0 \mid t_{\min}, t_{\max}, s) \\ & + P(b \mid t_{\min}, t_{\max}) P(t_0 \mid t_{\min}, t_{\max}, b) \big) \cdot P(t_{\min}, t_{\max}). \end{aligned} \tag{5.18}$$

In equation 5.18 the last term, $P(t_{\min}, t_{\max})$, is the total probability of observing $t_{\min}$ and $t_{\max}$ irrespective of whether the event was signal or background and this term does factor out of the likelihood and can be neglected. The terms $P(s \mid t_{\min}, t_{\max})$ and $P(b \mid t_{\min}, t_{\max})$ are the probabilities that an event is signal or background respectively given $t_{\min}$ and $t_{\max}$; in effect they give the signal fraction for each combination of $t_{\min}$

and $t_{\rm max}$. There is a choice: calculate the probability of getting $t_{\rm min}$ and $t_{\rm max}$ given that an event is signal or background as in equation 5.17; calculate the probability that an event is signal or background given $t_{\rm min}$ and $t_{\rm max}$ as in equation 5.18. The distribution of $t_{\rm min}$ and $t_{\rm max}$ for signal and background cannot be neglected altogether without, in effect, getting the signal fraction wrong.

This problem of miscalculating the signal fraction arises whenever a quantity that changes event by event, for example event by event acceptance functions or event by event lifetime errors, is used without modelling its distribution correctly. This effect is discussed in [64] for the case of event by event lifetime errors.

5.9 The lifetime distribution for background

It is not necessary to know the physical meaning of the background lifetime distribution for the measurement of the $\rm B_s^0 \to K^+K^-$ lifetime and so it can be described by a general PDF $y(t_0)$. $y(t_0)$ is the distribution of lifetimes in the background before the trigger. It includes all detector resolution effects so does not depend on the measured uncertainty on lifetime, σ_t. The probability of measuring a lifetime, t_0, given that an event is background and given the minimum and maximum accepted lifetimes, $t_{\rm min}$ and $t_{\rm max}$, is

$$P\left(t_0 \mid b, t_{\rm min}, t_{\rm max}\right) = \begin{cases} \dfrac{y(t_0)}{\int_{t_{\rm min}}^{t_{\rm max}} y(t_0){\rm d}t_0} & \text{for } t_{\rm min} \leq t_0 \leq t_{\rm max} \\ 0 & \text{for all other } t_0. \end{cases} \tag{5.19}$$

$y(t_0)$ has been normalised such that t_0 lies within the range of accepted lifetimes.

The background lifetime distribution is parameterised by fitting the height, $y(t_0)$, at different lifetimes, t_j, and interpolating between these points using exponential functions. So

$$y(t_0) = e^{a_j + \left(\frac{a_{j+1}-a_j}{t_{j+1}-t_j}\right)(t_0 - t_j)} \qquad \text{for } t_j \leq t_0 \leq t_{j+1} \tag{5.20}$$

where the a_j are constants to be found. The number, n, and spacing of the points t_j are chosen to use as few parameters as are needed to describe the distribution. The points are spaced most tightly at low lifetimes where the distribution of lifetime varies most rapidly and there are fewer points at large lifetimes where the distribution varies less. The t_j are distributed logarithmically according to

$$t_j = t_{\min} + \left(-c + e^{\left(ln(c) + \frac{ln(c+b)-ln(c)}{n} \cdot j\right)}\right) \cdot \frac{(t_{\max} - t_{\min})}{b}, \tag{5.21}$$

where c and b are constants which affect the scale of the logarithmic spacing. For a given c ($c = 1$ was chosen arbitrarily) decreasing b spaces t_j more equally. Values of $c = 1$ and $b = 4$ were chosen as they give a good fit to the background in the upper sideband.

The B_PIPI trigger ensures that few events pass with very low lifetimes. There is also very little background with very long lifetimes ($t_0 > 1500$ μm). To accommodate the low density of events at the extremes of the lifetime range the first point is set at $t_0 = 0$ and the second point at $t_1 = 130$ μm where the density of events begins to increase. The last point is set to $t_9 = 3000$ μm which is the maximum allowed lifetime in the sample. The penultimate point is placed at $t_8 = 1500$ μm. Between t_1 and t_8 the points are spaced logarithmically as described above and shown in figure 5.10. There are ten parameters in the background lifetime PDF which are free in the final fit. Ten parameters were found to be sufficient for the statistics available. Increasing the number of parameters has little effect on the goodness of the fit but does affect its stability. The systematic effect of the choice of background lifetime distribution is discussed in section 7.12.

To test the background lifetime parameterisation a fit was performed on data from the upper side band region, $5.551 < M_{\pi\pi} < 7.0$ GeV/c^2, above the Λ_b^0 region. The projection of this fit is shown in figure 5.11. This parameterisation gives a good fit to the data. The distribution, $y(t_0)$, returned by the fit is shown in figure 5.10.

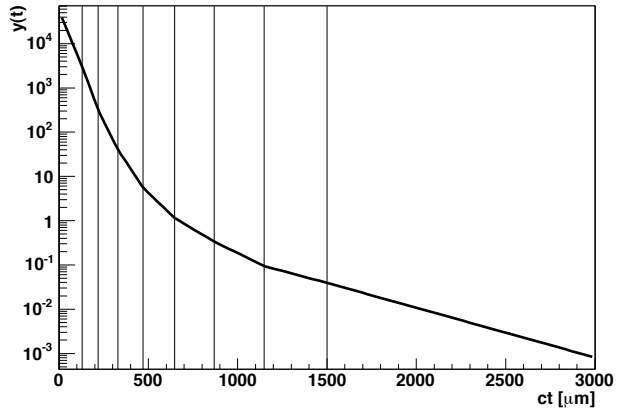


Figure 5.10: The background lifetime distribution, $y(t_0)$, obtained from a fit to a background only sample from the upper sideband in data. This is the distribution before any trigger or selection requirements have been made. The distribution is parameterised by fitting the height of the distribution at times t_j, indicated by the vertical lines, and extrapolating between the points using exponential functions, which appear as straight lines due to the logarithmic scale.

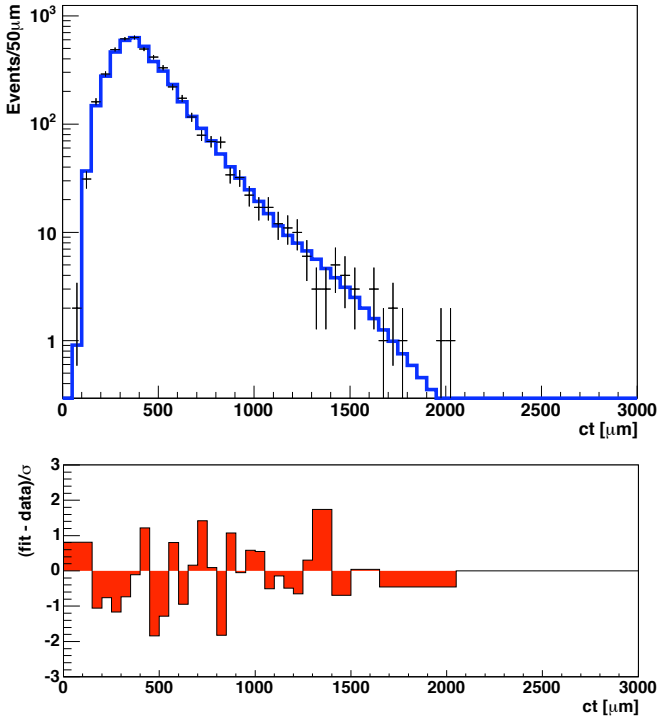


Figure 5.11: Projection of ct for a lifetime fit to a background only sample from the upper sideband in data (upper plot). The lower plot shows the residual distribution from this fit. This distribution has been shaped by the trigger and selection requirements

5.10 The distribution of accepted lifetimes

Including background led to the addition of the terms $P(s \mid t_{\min}, t_{\max})$ and $P(b \mid t_{\min}, t_{\max})$ in the lifetime PDF (in equation 5.18). These terms describe the probability that an event is signal (or background) given $t_{\min}$ and $t_{\max}$. Rather than try to model them directly, these terms are simplified by making the approximation:

$$P(s \mid t_{\min}, t_{\max}) \approx P(s \mid \kappa) \tag{5.22}$$

where κ is a single number. $P(s \mid \kappa)$ is the probability that an event is signal given κ. The two variables, $t_{\min}$ and $t_{\max}$, are reduced to the single variable, κ, using a linear transformation; each co-ordinate $(t_{\min}, t_{\max})$ is projected onto a line in the $t_{\min}$, $t_{\max}$ plane. In order to maximise the information retained the line is chosen which provides the best discrimination between signal and background events. This direction can be found using a Fisher linear discriminant analysis, which is described in the next section. It is important to note that while the term $P(s \mid \kappa)$ may provide some discrimination between signal and background events its purpose is to approximate the term $P(s \mid t_{\min}, t_{\max})$.

The technique of Fisher linear discriminant analysis is described in section 5.11 for the case where there are known samples of signal and background data. Section 5.12 explains how this technique may be used with a signal sample which is contaminated with background. The Fisher analysis is used to convert $t_{\min}$ and $t_{\max}$ to a single variable κ in section 5.13 and this section also describes how the term $P(s \mid \kappa)$ is modelled.

The discussion that follows is for the case where there is one type of signal and one type of background in a sample. In fact the data sample used here contains many decay modes and the ideas here will be revisited in section 6.2 when these are included in the likelihood. Finally this approach of approximating $P(s \mid t_{\min}, t_{\max})$ to $P(s \mid \kappa)$

will be tested in a more complicated scenario in section 7.4.

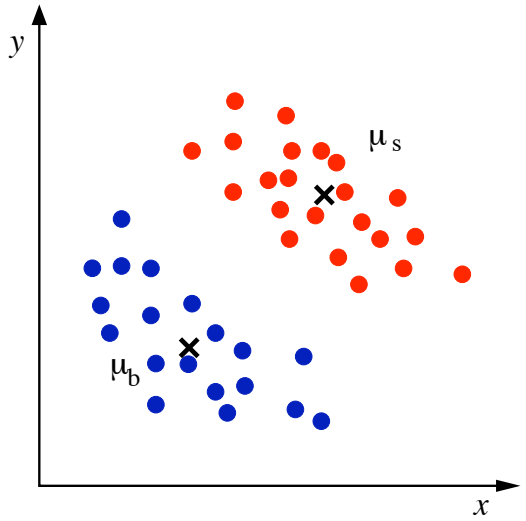
5.11 Fisher linear discriminant analysis

Consider data containing two classes of events, signal and background, where each event can be described by two variables x and y. Figure 5.12a shows an example distribution of such data. The distribution can be reduced from 2 dimensions to 1 dimension by projecting all of the points onto a line. The best line to choose is the one which provides the best separation of the two classes of events once they have been projected onto it. In the example in figure 5.12a the line providing the best separation is approximately the one which joins the means of the two samples, $\boldsymbol{\mu}_s = \left(\begin{smallmatrix}\overline{x}_s \\ \overline{y}_s\end{smallmatrix}\right)$ and $\boldsymbol{\mu}_b = \left(\begin{smallmatrix}\overline{x}_b \\ \overline{y}_b\end{smallmatrix}\right)$. Figure 5.12b shows another example of possible distributions of two classes of events in x and y. In this case projecting the data onto a line along the means of the two distributions would provide poor separation of the two samples. A better separation of the two classes of events would be provided by projecting the events onto the y axis where the variance of each sample is small. In general, the projection which gives the best separation between two sub-samples of events is the one which maximises the separation of the projected means while minimising the spread of each of the projected distributions.

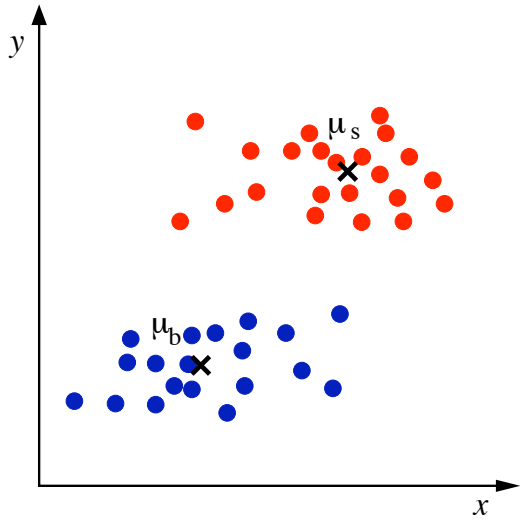
Fisher's linear discriminant analysis [65] finds the direction, $\mathbf{w}$, which provides the best separation by maximising the quantity

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}, \tag{5.23}$$

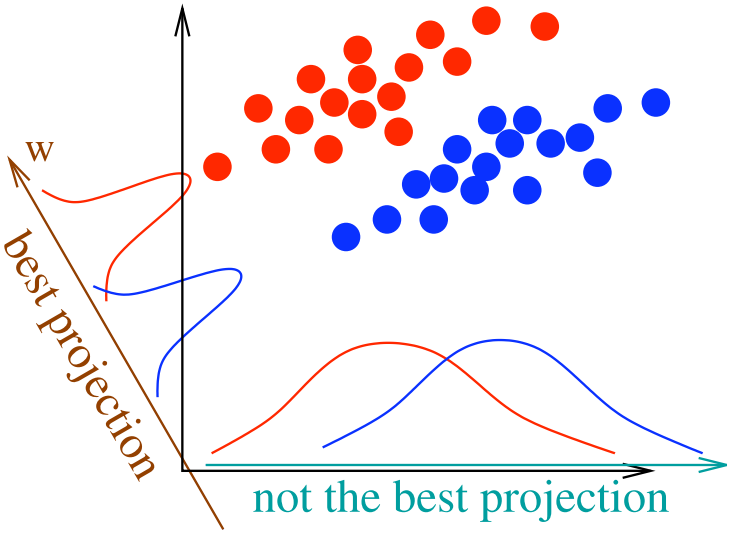
where $\mathbf{S}_B$ can be thought of as the between class scatter matrix and $\mathbf{S}_W$ the within class scatter matrix. $\mathbf{S}_B$ is defined by considering the square of the means of the two



(a) The direction with the best separation between the two samples is the one which joins the means of the two samples.



(b) In this example the separation in the direction of the y axis is better than along the means of the two samples.



(c) The projection along $\mathbf{w}$ is the projection with the best separation of the two samples.

Figure 5.12: Two different samples of events (red and blue) have different distributions of events in the variables x and y. If the distributions are to be projected onto a line the projection direction which will in general have the best separation between the two samples is the one which maximises the distance between the means of the two distributions along the projections while also minimising the spread of each sample about the projected mean.

samples projected onto $\mathbf{w}$:

$$\begin{aligned}\left(\mathbf{w}^T\boldsymbol{\mu}_s - \mathbf{w}^T\boldsymbol{\mu}_b\right)^2 &= \mathbf{w}^T\left(\boldsymbol{\mu}_s - \boldsymbol{\mu}_b\right)\left(\boldsymbol{\mu}_s - \boldsymbol{\mu}_b\right)^T\mathbf{w} \\ &= \mathbf{w}^T\mathbf{S}_B\mathbf{w}.\end{aligned} \tag{5.24}$$

The total within class spread ($\mathbf{S}_W$) is given by the sum of the spread of the signal and background samples around their means projected onto $\mathbf{w}$, $\mathbf{S}_W = \mathbf{S}_{sg} + \mathbf{S}_{bg}$. These are proportional to the covariance matrices. For signal,

$$\begin{aligned}\sum_{i=sg\,\text{events}}\left(\mathbf{w}^T\mathbf{x}_i - \mathbf{w}^T\boldsymbol{\mu}_s\right)^2 &= \sum_{i=sg\,\text{events}}\left(\mathbf{w}^T\left(\mathbf{x}_i - \boldsymbol{\mu}_s\right)\left(\mathbf{x}_i - \boldsymbol{\mu}_s\right)^T\mathbf{w}\right) \\ &= \mathbf{w}^T\mathbf{S}_{sg}\mathbf{w}\end{aligned} \tag{5.25}$$

where $\mathbf{x_i} = \binom{x_i}{y_i}$ and the sum is over all signal events. A similar expression holds for $\mathbf{S}_{bg}$. $\mathbf{w}$ is obtained by maximising $J(\mathbf{w})$ in equation 5.23. That gives

$$\begin{aligned}\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} &= \frac{1}{\left(\mathbf{w}^T\mathbf{S}_W\mathbf{w}\right)^2}\Big(2\left(\mathbf{w}^T\mathbf{S}_W\mathbf{w}\right)\mathbf{S}_B\mathbf{w} - 2\left(\mathbf{w}^T\mathbf{S}_B\mathbf{w}\right)\mathbf{S}_W\mathbf{w}\Big) \\ &= 0\end{aligned} \tag{5.26}$$

and so

$$\mathbf{S}_B\mathbf{w} - \frac{\mathbf{w}^T\mathbf{S}_B\mathbf{w}}{\mathbf{w}^T\mathbf{S}_W\mathbf{w}}\mathbf{S}_W\mathbf{w} = 0 \tag{5.27}$$

or

$$\left(\mathbf{S}_B - \lambda\mathbf{S}_W\right)\mathbf{w} = 0. \tag{5.28}$$

This is an eigenvalue equation. It can be written

$$\mathbf{S}_W^{-1}\mathbf{S}_B\mathbf{w} = \lambda\mathbf{w}. \tag{5.29}$$

From equation 5.24

$$\begin{aligned} \mathbf{S}_B\mathbf{w} &= (\boldsymbol{\mu}_s - \boldsymbol{\mu}_b)(\boldsymbol{\mu}_s - \boldsymbol{\mu}_b)^T \mathbf{w} \\ &\propto (\boldsymbol{\mu}_s - \boldsymbol{\mu}_b). \end{aligned} \tag{5.30}$$

The scale of $\mathbf{w}$ is not important, it is the direction of $\mathbf{w}$ which needs to be found, so it is not necessary to calculate λ. $\mathbf{w}$ can be obtained from $\mathbf{S}_W^{-1}$ and $(\boldsymbol{\mu}_s - \boldsymbol{\mu}_b)$. Combining equations 5.29 and 5.30 gives

$$\mathbf{S}_W^{-1}(\boldsymbol{\mu}_s - \boldsymbol{\mu}_b) \propto \mathbf{w}. \tag{5.31}$$

The Fisher scalar value, κ, for each event is given by the dot product of the event vector, $\mathbf{x}_i$, and $\mathbf{w}$.

5.12 Extracting the Fisher direction from data

A Fisher linear discriminant analysis is used, here, to reduce the information contained in the two variables $t_{\min}$ and $t_{\max}$ into one variable κ. The joint distribution of $t_{\min}$ and $t_{\max}$ is different for signal and background and the Fisher linear discriminant analysis finds the direction in $t_{\min}$-$t_{\max}$ space which provides the best separation of signal and background once all the data have been projected onto that line. Figure 5.13 shows the distributions of events in the signal and background regions (defined as $5.18 \leq M_{\pi\pi} \leq 5.3$ GeV/c^2 and $5.6 \leq M_{\pi\pi} \leq 5.8$ GeV/c^2 respectively). The differences between the distributions in signal and background can be more clearly seen in the projections to $t_{\min}$, $t_{\max}$, $t_{\max} + t_{\min}$ and $t_{\max} - t_{\min}$ in figure 5.14.

For each event the vector $\mathbf{x}_i = \begin{pmatrix} t_{\min\, i} \\ t_{\max\, i} \end{pmatrix}$ can be calculated. The mean values for signal and background are $\boldsymbol{\mu}_s = \begin{pmatrix} \overline{t_{\min}}_{\,\mathrm{s}} \\ \overline{t_{\max}}_{\,\mathrm{s}} \end{pmatrix}$ and $\boldsymbol{\mu}_b = \begin{pmatrix} \overline{t_{\min}}_{\,\mathrm{b}} \\ \overline{t_{\max}}_{\,\mathrm{b}} \end{pmatrix}$ respectively. $\mathbf{w}$ is a two component vector indicating the Fisher direction in $t_{\min}$-$t_{\max}$ space. All of the matrices,

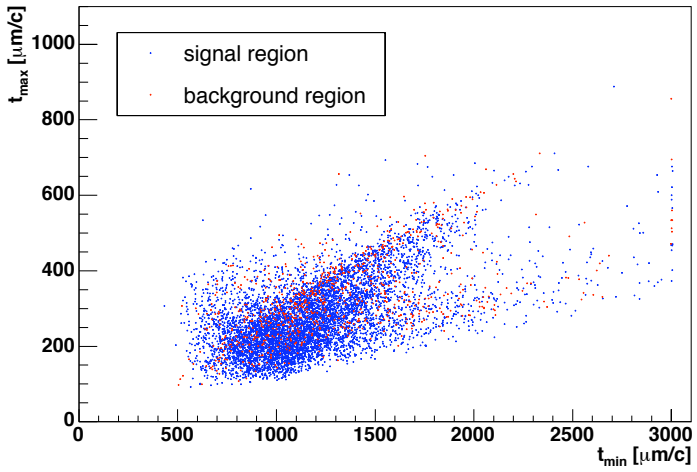
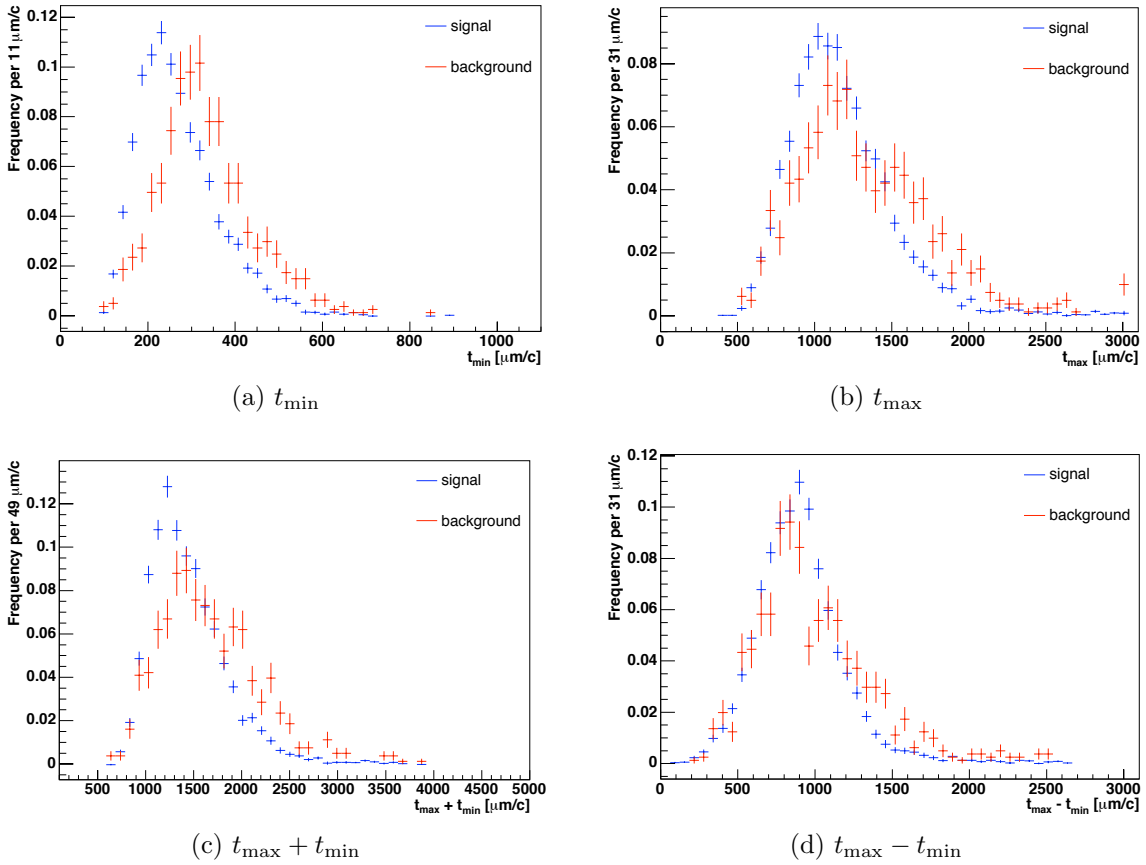


Figure 5.13: The distribution of $t_{\rm min}$ and $t_{\rm max}$ in data for events in the signal region ($5.18 \leq M_{\pi\pi} \leq 5.3$ GeV/c^2, blue points) and background region ($5.6 \leq M_{\pi\pi} \leq 5.8$ GeV/c^2, red points). The differences between the two samples can be seen more clearly in the projections in figure 5.14.

$\mathbf{S}_W$, $\mathbf{S}_B$, $\mathbf{S}_{\rm sg}$ and $\mathbf{S}_{\rm bg}$, are two by two matrices.

Traditionally, Fisher linear discriminant analysis uses known samples of signal and background events, for example from simulated data, to find the Fisher direction. Where there is a sample of known signal events and a sample of known background events it is a simple matter to calculate $(\boldsymbol{\mu}_s - \boldsymbol{\mu}_b)$ and $\mathbf{S}_W$ can be calculated using equation 5.25. The intention of this analysis is to remove the lifetime bias without using simulated data; instead, the data itself is used to find the Fisher direction. How this is done is explained in the remainder of this section.

The mass projection of the separation fit is used to define a signal region, A, and a background region, B, as shown in figure 5.15. Of course the data in this sample contain a number of different decay modes and two types of background events so some consideration is required as to what constitutes a signal or background event. This is discussed further in chapter 6. For the moment though, consider all $\mathrm{B}^0_{(s)} \to h^+h'^-$ type modes to be signal and consider only combinatorial background. How the physics type background and $\Lambda^0_{\rm b} \to {\rm p}h^-$ decays may be dealt with is considered in section 6.4. The background region (B) contains N_B events which can be assumed to be entirely background. The signal region contains N_A events, N_{As} of which are signal and N_{Ab} are



(a) $t_{\min}$ (b) $t_{\max}$

(c) $t_{\max} + t_{\min}$ (d) $t_{\max} - t_{\min}$

Figure 5.14: Comparison of the joint distribution of $t_{\min}$ and $t_{\max}$ in background and background subtracted signal from data projected onto $t_{\min}$, $t_{\max}$, $t_{\max}+t_{\min}$ and $t_{\max}-t_{\min}$. The difference between signal and background is most pronounced when projected onto $t_{\min}$ but barely apparent when projected onto $t_{\max}-t_{\min}$. The signal distribution is made by subtracting from the distribution in the signal region a histogram of the distribution in the background region normalised to the number of background events expected in the signal region as determined by the separation fit described in chapter 4.

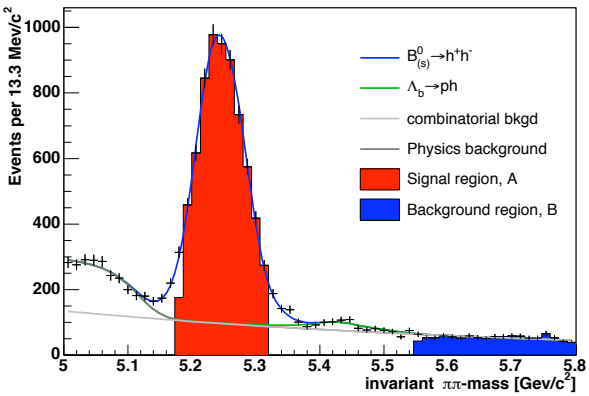


Figure 5.15: The distribution of invariant $\pi\pi$-mass ($M_{\pi\pi}$) with the projection of the separation only fit overlaid. A signal region, A, is defined within the mass region $5.18 < M_{\pi\pi} < 5.32$ GeV/c^2 and a background region, B, in the mass range $5.551 < M_{\pi\pi} < 5.8$ GeV/c^2. These regions are used to determine the Fisher direction as described in the text.

background. The combined sample from regions A and B contains N events, $N_s = N_{As}$ of which are signal and $N_b = N_{Ab} + N_B$ of which are background.

Obtaining $(\boldsymbol{\mu}_s - \boldsymbol{\mu}_b)$

The high mass sideband can be assumed to be entirely background. It is assumed that the events in this region are representative of all the background events. $\boldsymbol{\mu}_b$ can be calculated directly from this sample. The signal region contains a mixture of signal and background events. The fraction of signal events in this region, f_s, can be calculated from the separation fit. The mean of all the events in this region is

$$\boldsymbol{\mu}_A = f_s \boldsymbol{\mu}_s + (1 - f_s)\, \boldsymbol{\mu}_b \tag{5.32}$$

$$\implies \quad \boldsymbol{\mu}_A - \boldsymbol{\mu}_b = f_s \left(\boldsymbol{\mu}_s - \boldsymbol{\mu}_b\right). \tag{5.33}$$

The scale of $\mathbf{w}$ is not important it is sufficient to note that $\boldsymbol{\mu}_s - \boldsymbol{\mu}_b \propto \boldsymbol{\mu}_A - \boldsymbol{\mu}_b$.

Obtaining $\mathbf{S}_W^{-1}$

From equation 5.25,

$$\mathbf{S}_W = \mathbf{S}_{sg} + \mathbf{S}_{bg} \tag{5.34}$$

$$= \sum_{i=0}^{N_s} \left(\mathbf{x}_i - \boldsymbol{\mu}_s\right)\left(\mathbf{x}_i - \boldsymbol{\mu}_s\right)^T + \sum_{j=0}^{N_b} \left(\mathbf{x}_j - \boldsymbol{\mu}_b\right)\left(\mathbf{x}_j - \boldsymbol{\mu}_b\right)^T, \tag{5.35}$$

where the first sum is over all signal events and the second over all background. A matrix $\mathbf{S}_B$ can be calculated directly from events in the background region B as

$$\mathbf{S}_B = \sum_{k=0}^{N_B} \left(\mathbf{x}_k - \boldsymbol{\mu}_b\right)\left(\mathbf{x}_k - \boldsymbol{\mu}_b\right)^T, \tag{5.36}$$

where the sum is over all events in the background region. Another matrix, $\mathbf{S}_{Bs}$, can be defined as

$$\mathbf{S}_{Bs} = \sum_{k=0}^{N_B} (\mathbf{x}_k - \boldsymbol{\mu}_s)(\mathbf{x}_k - \boldsymbol{\mu}_s)^T, \tag{5.37}$$

where once again the sum is over all the events in the background region and $\boldsymbol{\mu}_s$ can be calculated using equation 5.32. For the signal region:

$$\mathbf{S}_A = \sum_{l=0}^{N_A} (\mathbf{x}_l - \boldsymbol{\mu}_s)(\mathbf{x}_l - \boldsymbol{\mu}_s)^T \tag{5.38}$$

$$= \sum_{m=0}^{N_{As}} (\mathbf{x}_m - \boldsymbol{\mu}_s)(\mathbf{x}_m - \boldsymbol{\mu}_s)^T + \sum_{n=0}^{N_{Ab}} (\mathbf{x}_n - \boldsymbol{\mu}_s)(\mathbf{x}_n - \boldsymbol{\mu}_s)^T. \tag{5.39}$$

The sum in equation 5.38 is over all events in region A. The first sum in equation 5.39 is over the signal events in region A and the second sum over the background events in region A. The matrices $\mathbf{S}_B$, $\mathbf{S}_{Bs}$ and $\mathbf{S}_A$ can be calculated directly from the data. By assuming that events in the background region are representative of all background events $\mathbf{S}_W$ can be calculated as

$$\mathbf{S}_W = \mathbf{S}_A - \frac{N_B}{N_{Ab}}\mathbf{S}_{Bs} + \frac{N_b}{N_B}\mathbf{S}_B. \tag{5.40}$$

In practice this procedure is improved by using the event by event signal probability returned by the separation fit to weight each event in the signal region.

As a demonstration, the Fisher direction was found for a sample of simulated $\mathrm{B_s^0} \to \mathrm{K^+K^-}$ and background events using a simple mass fit rather than the full separation fit (a single Gaussian shape for signal plus exponential background distribution was used). The invariant $\pi\pi$-mass distribution and projection of the mass fit are shown in figure 5.16. Figure 5.17 shows the joint distribution of $t_{\min}, t_{\max}$ for this sample and the Fisher direction found.

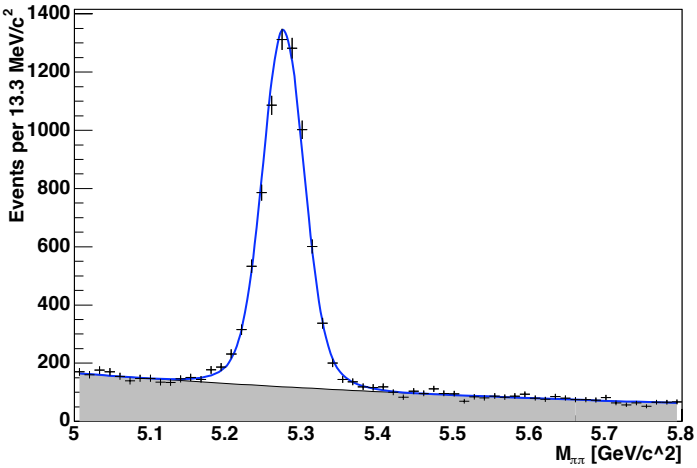


Figure 5.16: The invariant $\pi\pi$-mass distribution for a sample of simulated $B_s^0 \to K^+K^-$ and background events. The blue curve is the projection of the mass fit to this sample.

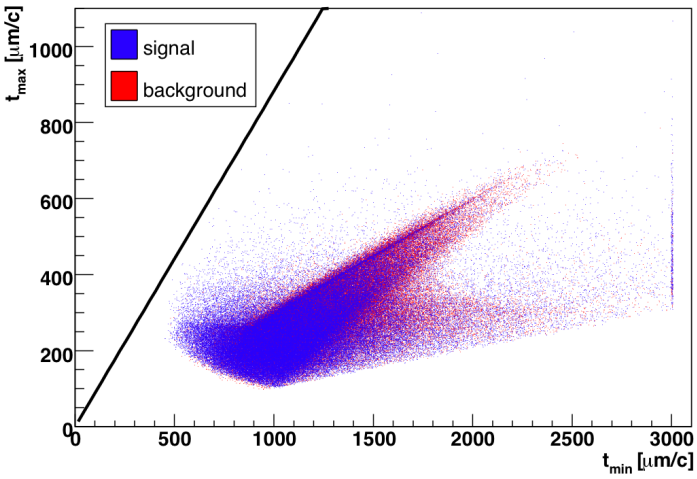


Figure 5.17: The distribution of $t_{\min}$ and $t_{\max}$ for simulated $B_s^0 \to K^+K^-$ (blue points) and background (red points) events. The black line shows the Fisher direction $\mathbf{w}$ found for this sample.

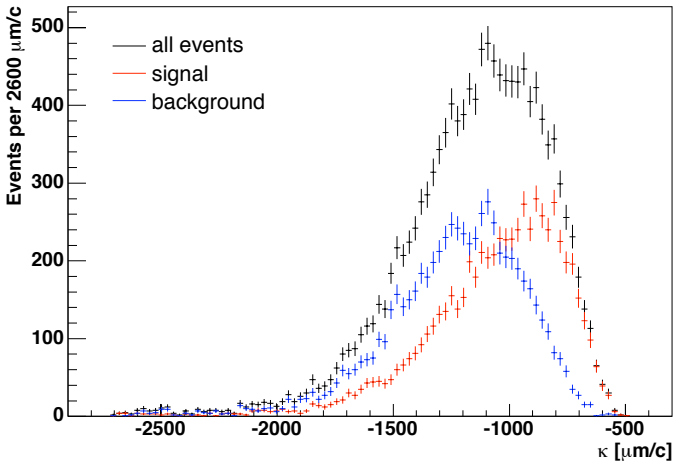


Figure 5.18: The distribution of Fisher scalar κ for a simulated sample of $B_s^0 \to K^+K^-$ (red) and background (blue) events found by projecting the distribution of $(t_{\min}, t_{\max})$ shown in figure 5.18 onto the Fisher direction $\mathbf{w}$ in the same diagram.

5.13 The Fisher scalar distribution

Once the Fisher direction, $\mathbf{w}$ has been found a Fisher scalar variable, κ can be calculated for each event as

$$\kappa_i = \mathbf{w} \cdot \begin{pmatrix} t_{\min\, i} \\ t_{\max\, i} \end{pmatrix}. \tag{5.41}$$

Once κ has been calculated for each event the Fisher direction is no longer needed or used. The distribution of κ for the sample of simulated $B_s^0 \to K^+K^-$ and background events in figure 5.17 is shown in figure 5.18.

The probability $P(s \mid \kappa)$ is modelled using the Lagrange interpolating polynomial [66]. The Lagrange interpolating polynomial is a polynomial of order $\leq n-1$ which passes through n points $x_i, y_i = f(x_i)$ where $i = 1, 2, \ldots, n$. It is defined as

$$P(x) = \sum_{j=1}^{n} P_j(x), \tag{5.42}$$

where

$$P_j(x) = y_j \prod_{\substack{k=1 \\ k \neq j}}^{n} \frac{x - x_k}{x_j - x_k}. \tag{5.43}$$

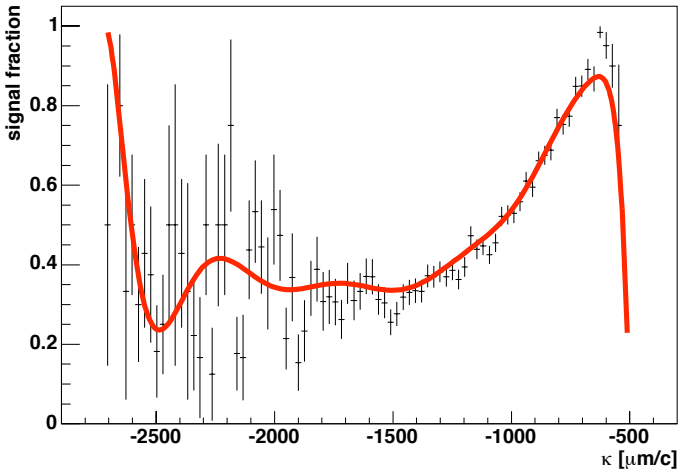


Figure 5.19: The true signal fraction as a function of Fisher scalar, κ, for simulated $B_s^0 \to K^+K^-$ and background events (black points with error bars). The red curve shows the Lagrange interpolating polynomial returned by the lifetime fit; it is not a fit on the data points shown. The fit does not have information about the truth signal fractions yet the returned polynomial models the truth distribution well.

As all events must be either signal or background $P(s \mid \kappa)$ is the signal fraction as a function of κ. Figure 5.19 shows this distribution for a mixture of simulated $B_s^0 \to K^+K^-$ and background events. The distribution is divided into equally sized bins. These bins define the x_i points which the polynomial will pass through. The height of the polynomial (the y_i) are left free in the fit. The distribution in figure 5.19 used fifteen points for the Lagrange interpolating polynomial. The fit does not have the truth information about the signal fractions yet it is able to accurately fit the data in a fit combined with separation and lifetime information.

5.14 The lifetime probability for one decay mode plus background

It is useful to review the lifetime part of the likelihood so far, deriving an exact expression for the PDF. This derivation will use the following notation:

***P*(*A*)** The probability of A;

***P*(*A*, *B*)** The probability of A and B;

***P*(*A* or *B*)** The probability of A or B;

***P*(*A* | *B*)** The probability of A given B.

The following will use these basic rules of manipulating probabilities. The probability of A and B is

$$P(A, B) = P(A)P(B \mid A), \tag{5.44}$$

which leads to Bayes's theorem:

$$P(A \mid B) = \frac{P(A, B)}{P(B)} = \frac{P(A)P(B \mid A)}{P(B)}. \tag{5.45}$$

Also the probability of A or B

$$P(A \text{ or } B) = P(A) + P(B) - P(A, B). \tag{5.46}$$

The likelihood contains three measured quantities: the measured lifetime t_0 and the minimum and maximum allowed lifetimes, $t_{\min}$ and $t_{\max}$. The data sample only contains events which have passed the trigger so the likelihood will be the product of the probabilities of observing t_0, $t_{\min}$ and $t_{\max}$ for each event given that the event passed the trigger:

$$P(t_0, t_{\min}, t_{\max} \mid \text{trigger}). \tag{5.47}$$

This can be separated into a signal and background part using equation 5.46. Using s to denote signal and b to denote background and noting that $P(s, b) = 0$ and $P(s \text{ or } b) = 1$:

$$P(t_0, t_{\min}, t_{\max} \mid \text{trigger}) = P(s, t_0, t_{\min}, t_{\max} \mid \text{trigger}) + P(b, t_0, t_{\min}, t_{\max} \mid \text{trigger}). \tag{5.48}$$

Consider the signal part of the probability, the procedure for background is analogous. Using Bayes's theorem the signal part can be written

$$P(s, t_0, t_{\min}, t_{\max} \mid \text{trigger}) = \frac{P(s, t_0, t_{\min}, t_{\max}) P(\text{trigger} \mid s, t_0, t_{\min}, t_{\max})}{P(\text{trigger})}. \quad (5.49)$$

The second term in the numerator, $P(\text{trigger} \mid s, t_0, t_{\min}, t_{\max})$, is either 0 or a constant ϵ^2 where ϵ is the track finding efficiency. This is because the trigger decision is entirely determined by the lifetime acceptance, the decay time and the track finding efficiency. The denominator, $P(\text{trigger})$, is the probability that the trigger fires. It is preferable to write this in terms of the event by event probability that the trigger fires given $t_{\min}$ and $t_{\max}$, that is $P(\text{trigger} \mid t_{\min}, t_{\max})$. Also it is simpler to calculate this probability separately for signal and background so, using Bayes's theorem,

$$P(\text{trigger}) = \frac{P(\text{trigger} \mid s, t_{\min}, t_{\max}) P(s, t_{\min}, t_{\max})}{P(s, t_{\min}, t_{\max} \mid \text{trigger})}. \quad (5.50)$$

In addition, the first term in the numerator of equation 5.49 can be written

$$P(s, t_0, t_{\min}, t_{\max}) = P(s, t_{\min}, t_{\max}) P(t_0 \mid s, t_{\min}, t_{\max}). \quad (5.51)$$

Substituting equations 5.50 and 5.51 into equation 5.49 gives,

$$P(s, t_0, t_{\min}, t_{\max} \mid \text{trigger}) = \frac{P(s, t_{\min}, t_{\max} \mid \text{trigger}) P(t_0 \mid s, t_{\min}, t_{\max}) P(\text{trigger} \mid s, t_0, t_{\min}, t_{\max})}{P(\text{trigger} \mid s, t_{\min}, t_{\max})}. \quad (5.52)$$

The first term in the numerator may be expanded as

$$P(s, t_{\min}, t_{\max} \mid \text{trigger}) = P(t_{\min}, t_{\max} \mid \text{trigger}) P(s, \mid t_{\min}, t_{\max} \, \text{trigger}). \quad (5.53)$$

The first term, $P(t_{\min}, t_{\max} \mid \text{trigger})$, is independent of lifetime and is identical for signal and background so will be a constant term in the log-likelihood which can be neglected.

It is now necessary to make the assumption that t_0, the measured lifetime, is independent of $t_{\min}$ and $t_{\max}$. This is a reasonable assumption since $t_{\min}$ and $t_{\max}$ depend only on the kinematics of the decay. $P(t_0 \mid t_{\min}, t_{\max})$ is in effect the probability of measuring a lifetime t_0 given the kinematics of the decay which translate the trigger requirements into limits on the measured lifetime before the trigger is applied. This gives:

$$P(t_0 \mid s, t_{\min}, t_{\max}) = P(t_0 \mid s). \tag{5.54}$$

The signal part of the PDF is now

$$P(s, t_0, t_{\min}, t_{\max} \mid \text{trigger}) = \frac{P(s \mid t_{\min}, t_{\max}, \text{trigger}) P(t_0 \mid s) P(\text{trigger} \mid s, t_0, t_{\min}, t_{\max})}{P(\text{trigger} \mid s, t_{\min}, t_{\max})}. \tag{5.55}$$

The terms in this PDF are as follows:

- $P(s \mid t_{\min}, t_{\max}, \text{trigger})$. This is the signal fraction as a function of $t_{\min}$ and $t_{\max}$. This term is approximated to $P(s \mid \kappa)$ using a Fisher linear discriminant analysis and then modelled by the Lagrange interpolating polynomial as described in section 5.13;

- $P(t_0 \mid s)$. This is the probability of measuring a lifetime t_0 given that an event is signal. This is the numerator in equation 5.15:

$$\int_0^\infty \frac{1}{\tau} e^{-\frac{t}{\tau}} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t_o)^2}{2\sigma_t{}^2}} \mathrm{d}t; \tag{5.56}$$

- $P(\text{trigger} \mid s, t_0, t_{\min}, t_{\max})$. The probability that the trigger fires given t_0, $t_{\min}$ and $t_{\max}$. If the trigger were 100% efficient this term would be either 1 or 0 if the event did or didn't pass the trigger cuts. In fact the trigger only finds tracks with some efficiency ϵ. If it is assumed that ϵ is constant for all particles within the trigger acceptance then, because two tracks are required to pass the trigger, this term is either ϵ^2 or 0 for two-body decays. The data sample can only

contain events which have passed the trigger so this term is a constant which can be neglected if ϵ is constant. The affect of a varying track-finding efficiency is considered as a systematic uncertainty in sections 7.6, 7.7 and 7.16.3;

- $P(\text{trigger} \mid s, t_{\min}, t_{\max})$. The probability that the trigger fires given that an event is signal and given $t_{\min}$ and $t_{\max}$. This is the normalisation of the signal lifetime probability and the denominator in equation 5.15:

$$\int_{t_{\min}}^{t_{\max}} \int_{0}^{\infty} \frac{1}{\tau} e^{-\frac{t}{\tau}} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t_o)^2}{2{\sigma_t}^2}} \mathrm{d}t \, \mathrm{d}t_0. \tag{5.57}$$

As was done for the signal part, the background part of equation 5.48 can be expressed as

$$P(b, t_0, t_{\min}, t_{\max} \mid \text{trigger}) = \frac{P(b \mid t_{\min}, t_{\max}, \text{trigger}) P(t_0 \mid b) P(\text{trigger} \mid b, t_0, t_{\min}, t_{\max})}{P(\text{trigger} \mid b, t_{\min}, t_{\max})}, \tag{5.58}$$

where the terms are:

- $P(b \mid t_{\min}, t_{\max}, \text{trigger})$. This is the background fraction as a function of $t_{\min}$ and $t_{\max}$ and $P(b \mid t_{\min}, t_{\max}, \text{trigger}) = 1 - P(s \mid t_{\min}, t_{\max}, \text{trigger})$;

- $P(t_0 \mid b)$. This is the probability of measuring a lifetime t_0 given that an event is background. This is given by the numerator in equation 5.19, $y(t_0)$;

- $P(\text{trigger} \mid b, t_0, t_{\min}, t_{\max})$. The probability that the trigger fires given t_0, $t_{\min}$ and $t_{\max}$. As with the equivalent signal term this term is either ϵ^2 for events that pass the trigger or 0 for all other events if it is assumed that the track finding efficiency, ϵ, is constant. As the data sample can only contain events which have passed the trigger this term is a constant which can be neglected with a systematic uncertainty due to the assumption that ϵ is constant (sections 7.6, 7.7 and 7.16.3);

- $P(\text{trigger} \mid b, t_{\min}, t_{\max})$ The probability that the trigger fires given that an event is

background and given $t_{\rm min}$ and $t_{\rm max}$. This is the normalisation of the background lifetime probability and the denominator in equation 5.19:

$$\int_{t_{\rm min}}^{t_{\rm max}} y(t_0)\mathrm{d}t_0. \tag{5.59}$$

5.15 More than one decay mode

The PDF so far concerns the lifetime probability for a single decay mode with and without background. This needs to be expanded to include several decay modes with different mean lifetimes. Equation 5.15 gives the probability, for a single event, of measuring a lifetime t_0 given a mean lifetime τ, an uncertainty on the measurement σ_t and the minimum and maximum accepted lifetimes $t_{\rm min}$ and $t_{\rm max}$; that is $P(t_0 \mid \tau, \sigma_t, t_{\rm min}, t_{\rm max})$. For a decay to two hadrons, h^+ and h'^-, the lifetime, $t_{\mathrm{hh}'}$, is

$$t_{\mathrm{hh}'} = \frac{L_{\rm T} M_{\mathrm{hh}'}}{p_{\rm T}}, \tag{5.60}$$

where $L_{\rm T}$ is the decay length of the B meson in the transverse plane, $M_{\mathrm{hh}'}$ the invariant mass of the two decay products and $P_{\rm T}$ the transverse momentum of the B.

For a single event, because the lifetime, $t_{\mathrm{hh}'}$, is proportional to the invariant mass of the two tracks, $t_{\mathrm{hh}'}$ is different for each mode hypothesis. So, for example, the probability that an event is $\mathrm{B}^0 \to \pi^+\pi^-$ is dependent on $t_{\pi\pi}$ while the probability that the *same event* is $\mathrm{B}^0_{\rm s} \to \mathrm{K}^+\mathrm{K}^-$ is dependent on a different measured lifetime, t_{KK}. Similarly, for a single event, $t_{\rm min}$ and $t_{\rm max}$ are different for each mode hypothesis. It is preferable to rewrite the PDF so that the same variables are used for each decay mode hypothesis while it is the PDF which varies. To do this, let us define a new time

$$t' = \frac{L_{\rm T} M_{\mathrm{B}^0}}{p_{\rm T}}, \tag{5.61}$$

where M_{B^0} is the B^0 mass from [50]. The $L_{\rm T}$ and $p_{\rm T}$ of the B meson are the same for

all decay mode hypotheses so t' is unchanged however h and h$'$ are assigned. A mass term is included in equation 5.61 to keep the new variable lifetime-like; it could have been left out. The choice of the B^0 mass is arbitrary but means that t' is close to the original value t. In terms of this new t', $t_{hh'}$ is

$$t_{hh'} = t' \frac{M_{hh'}}{M_{B^0}}. \tag{5.62}$$

Substituting t' for t in equation 5.9 gives the probability that a particle decays with a true, modified lifetime t' given a mean lifetime τ,

$$P(t \mid \tau) = \frac{1}{\tau} e^{-\frac{t}{\tau}} \quad \longrightarrow \quad P(t' \mid \tau, M_{hh'}) = \frac{M_{hh'}}{M_{B^0}} \frac{1}{\tau} e^{-\frac{t' M_{hh'}}{\tau M_{B^0}}}, \tag{5.63}$$

where it is assumed there are no cuts on lifetime. This expression is now explicitly dependant on the reconstructed invariant mass, $M_{hh'}$, which is dependant on the mode hypothesis. The invariant mass for each hypothesis needs to be expressed as a function of variables which are independent of the mode hypothesis. This is done using an approach similar to that used for the mass distribution of signal in the separation only fit, section 4.2.

5.16 Expanding the invariant mass

When a heavy particle decays to two particles h_1 and h_2 with masses m_1 and m_2 and 3-dimensional momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ the invariant mass of the system is

$$M_{12}^2 = \left(\sqrt{m_1^2 + p_1^2} + \sqrt{m_2^2 + p_2^2} \right)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 . \tag{5.64}$$

If the pion mass is arbitrarily assigned to both decay particles the obtained invariant mass is

$$M_{\pi\pi}^2 = \left(\sqrt{m_\pi^2 + p_1^2} + \sqrt{m_\pi^2 + p_2^2} \right)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2 . \tag{5.65}$$

Subtracting equation 5.65 from equation 5.64 gives an expression for the mode dependent invariant mass:

$$\begin{aligned} M^2_{\mathrm{hh'}} &= M^2_{12} \\ &= M^2_{\pi\pi} + (m_1^2 + m_2^2) - (m_\pi^2 + m_\pi^2) + 2\left(\sqrt{p_1^2 + m_1^2}\sqrt{p_2^2 + m_2^2} - \sqrt{p_1^2 + m_\pi^2}\sqrt{p_2^2 + m_\pi^2}\right). \end{aligned} \tag{5.66}$$

$M_{\pi\pi}$, p_1 and p_2 are observables of each event which are independent of the mode hypothesis. The masses m_1 and m_2 are dependent on the mode hypothesis (they are either the kaon or pion masses) but they do not change event by event. Equation 5.66 is a mode dependent function of mode invariant variables, $M_{\mathrm{hh'}}(M_{\pi\pi}, p_1, p_2)$, which is what was required in equation 5.63.

In the mode separation PDFs described in chapter 4, rather than using the momenta p_1 and p_2 as variables the sum of the momenta, $p_{\mathrm{tot}} = p_1 + p_2$, and the charge weighted momentum imbalance, $\alpha = q_1\,(1 - p1/p2)$, are used where q_1 and p_1 are the charge and momentum respectively of the track with the smaller momentum. The same variables need to be used throughout the likelihood so p_1 and p_2 are written as functions of α and p_{tot}:

$$p_1 = \frac{1 - |\alpha|}{2 - |\alpha|} p_{\mathrm{tot}}; \tag{5.67}$$

$$p_2 = \frac{1}{2 - |\alpha|} p_{\mathrm{tot}}. \tag{5.68}$$

The invariant masses $M_{\mathrm{hh'}}\big(M_{\pi\pi}, p_1[\alpha, p_{\mathrm{tot}}], p_2[\alpha, p_{\mathrm{tot}}]\big)$, for each mode hypothesis are given in table 5.1.

mode	$M_{\mathrm{hh}'}\big(M_{\pi\pi}, p_1[\alpha, p_{\mathrm{tot}}], p_2[\alpha, p_{\mathrm{tot}}]\big)$ for $\alpha \leq 0$
$\mathrm{B}^0 \to \pi^-\mathrm{K}^+$	$M_{\pi\pi}^2 + m_{\mathrm{K}}^2 - m_\pi^2 + 2\sqrt{p_1^2+m_\pi^2}\left(\sqrt{p_2^2+m_{\mathrm{K}}^2} - \sqrt{p_2^2+m_\pi^2}\right)$
$\overline{\mathrm{B}^0} \to \mathrm{K}^-\pi^+$	$M_{\pi\pi}^2 + m_{\mathrm{K}}^2 - m_\pi^2 + 2\sqrt{p_2^2+m_\pi^2}\left(\sqrt{p_1^2+m_{\mathrm{K}}^2} - \sqrt{p_1^2+m_\pi^2}\right)$
$\mathrm{B}^0 \to \pi^+\pi^-$	$M_{\pi\pi}^2$
$\overline{\mathrm{B}_{\mathrm{s}}^0} \to \pi^-\mathrm{K}^+$	$M_{\pi\pi}^2 + m_{\mathrm{K}}^2 - m_\pi^2 + 2\sqrt{p_1^2+m_\pi^2}\left(\sqrt{p_2^2+m_{\mathrm{K}}^2} - \sqrt{p_2^2+m_\pi^2}\right)$
$\mathrm{B}_{\mathrm{s}}^0 \to \mathrm{K}^-\pi^+$	$M_{\pi\pi}^2 + m_{\mathrm{K}}^2 - m_\pi^2 + 2\sqrt{p_2^2+m_\pi^2}\left(\sqrt{p_1^2+m_{\mathrm{K}}^2} - \sqrt{p_1^2+m_\pi^2}\right)$
$\mathrm{B}_{\mathrm{s}}^0 \to \mathrm{K}^+\mathrm{K}^-$	$M_{\pi\pi}^2 + 2(m_{\mathrm{K}}^2 - m_\pi^2) + 2\left(\sqrt{p_1^2+m_{\mathrm{K}}^2}\sqrt{p_2^2+m_{\mathrm{K}}^2} - \sqrt{p_1^2+m_\pi^2}\sqrt{p_2^2+m_\pi^2}\right)$

mode	$M_{\mathrm{hh}'}\big(M_{\pi\pi}, p_1[\alpha, p_{\mathrm{tot}}], p_2[\alpha, p_{\mathrm{tot}}]\big)$ for $\alpha > 0$
$\overline{\mathrm{B}^0} \to \pi^+\mathrm{K}^-$	$M_{\pi\pi}^2 + m_{\mathrm{K}}^2 - m_\pi^2 + 2\sqrt{p_1^2+m_\pi^2}\left(\sqrt{p_2^2+m_{\mathrm{K}}^2} - \sqrt{p_2^2+m_\pi^2}\right)$
$\mathrm{B}^0 \to \mathrm{K}^+\pi^-$	$M_{\pi\pi}^2 + m_{\mathrm{K}}^2 - m_\pi^2 + 2\sqrt{p_2^2+m_\pi^2}\left(\sqrt{p_1^2+m_{\mathrm{K}}^2} - \sqrt{p_1^2+m_\pi^2}\right)$
$\mathrm{B}^0 \to \pi^+\pi^-$	$M_{\pi\pi}^2$
$\mathrm{B}_{\mathrm{s}}^0 \to \mathrm{K}^+\pi^-$	$M_{\pi\pi}^2 + m_{\mathrm{K}}^2 - m_\pi^2 + 2\sqrt{p_2^2+m_\pi^2}\left(\sqrt{p_1^2+m_{\mathrm{K}}^2} - \sqrt{p_1^2+m_\pi^2}\right)$
$\mathrm{B}_{\mathrm{s}}^0 \to \pi^+\mathrm{K}^-$	$M_{\pi\pi}^2 + m_{\mathrm{K}}^2 - m_\pi^2 + 2\sqrt{p_1^2+m_\pi^2}\left(\sqrt{p_2^2+m_{\mathrm{K}}^2} - \sqrt{p_2^2+m_\pi^2}\right)$
$\mathrm{B}_{\mathrm{s}}^0 \to \mathrm{K}^+\mathrm{K}^-$	$M_{\pi\pi}^2 + 2(m_{\mathrm{K}}^2 - m_\pi^2) + 2\left(\sqrt{p_1^2+m_{\mathrm{K}}^2}\sqrt{p_2^2+m_{\mathrm{K}}^2} - \sqrt{p_1^2+m_\pi^2}\sqrt{p_2^2+m_\pi^2}\right)$

Table 5.1: $M_{\mathrm{hh}'}$ for each mode hypothesis. Upper table: for $\alpha \leq 0$ (the negative particle carries the smaller momentum). Lower table: $\alpha > 0$ (the positive particle carries the smaller momentum). For simplicity the functions are written in terms of p_1 and p_2 where $p_1 < p_2$.

5.17 The full lifetime probability with more than one decay mode

Combining equation 5.63 with the expressions for $M_{\mathrm{hh}'}$ given in table 5.1 gives a mode hypothesis dependent probability for measuring a modified time t' given the mode independent variables, $M_{\pi\pi}$, α and p_{tot}:

$$P(t' \mid \tau, M_{\pi\pi}, \alpha, p_{\mathrm{tot}}) = \frac{M_{\mathrm{hh}'}(\alpha, p_{\mathrm{tot}})}{M_{\mathrm{B}^0}} \frac{1}{\tau} e^{-\frac{t' M_{\mathrm{hh}'}(\alpha, p_{\mathrm{tot}})}{\tau M_{\mathrm{B}^0}}}. \tag{5.69}$$

Lifetime cuts are taken in to account in the same way as in equation 5.3, by normalising over the range of allowed t'. The probability that an event has a modified lifetime t' given that t' is constrained to lie between t'_{min} and t'_{max} is

$$P(t' \mid \tau, t'_{\mathrm{min}}, t'_{\mathrm{max}}, M_{\pi\pi}, \alpha, p_{\mathrm{tot}}) = \frac{\frac{M_{\mathrm{hh}'}(M_{\pi\pi}, \alpha, p_{\mathrm{tot}})}{M_{\mathrm{B}^0}} \frac{1}{\tau} e^{\frac{-t}{\tau} \frac{M_{\mathrm{hh}'}(M_{\pi\pi}, \alpha, p_{\mathrm{tot}})}{M_{\mathrm{B}^0}}}}{\int_{t'_{\mathrm{min}}}^{t'_{\mathrm{max}}} \frac{M_{\mathrm{hh}'}(M_{\pi\pi}, \alpha, p_{\mathrm{tot}})}{M_{\mathrm{B}^0}} \frac{1}{\tau} e^{\frac{-t'}{\tau} \frac{M_{\mathrm{hh}'}(M_{\pi\pi}, \alpha, p_{\mathrm{tot}})}{M_{\mathrm{B}^0}}} \mathrm{d}t'}. \tag{5.70}$$

Finally, the lifetime is still measured with some uncertainty. The probability that, given a true, modified decay time t' and a measurement uncertainty $\sigma_{t'}$, we measure a time t'_0 is

$$\frac{1}{\sqrt{2\pi}\sigma_{t'}} e^{-\frac{(t'-t'_0)^2}{2\sigma_{t'}^2}}. \tag{5.71}$$

Note that $\sigma_{t'}$ is the width of a Gaussian curve which describes the distribution of modified measured lifetime, t'_0, around the modified true lifetime, t', not around the original true lifetime t. This is why the factor $M_{\mathrm{hh}'}/M_{\mathrm{B}^0}$ doesn't appear in equation 5.71.

The total probability of measuring a decay time t'_0 is thus

$$P(t'_0 \mid M_{\pi\pi}, \alpha, p_{\rm tot}, t'_{\rm min}, t'_{\rm max}, s_j) = \frac{\int_0^\infty \frac{M_{\rm hh'}(M_{\pi\pi},\alpha,p_{\rm tot})}{M_{\rm B^0}} \frac{1}{\tau} e^{-\frac{t'}{\tau}\frac{M_{\rm hh'}(M_{\pi\pi},\alpha,p_{\rm tot})}{M_{\rm B^0}}} \frac{1}{\sqrt{2\pi}\sigma_{t'}} e^{-\frac{(t'-t'_o)^2}{2{\sigma_{t'}}^2}} \mathrm{d}t'}{\int_{t'_{\rm min}}^{t'_{\rm max}} \int_0^\infty \frac{M_{\rm hh'}(M_{\pi\pi},\alpha,p_{\rm tot})}{M_{\rm B^0}} \frac{1}{\tau} e^{-\frac{t'}{\tau}\frac{M_{\rm hh'}(M_{\pi\pi},\alpha,p_{\rm tot})}{M_{\rm B^0}}} \frac{1}{\sqrt{2\pi}\sigma_{t'}} e^{-\frac{(t'-t'_o)^2}{2{\sigma_{t'}}^2}} \mathrm{d}t' \, \mathrm{d}t'_0}. \tag{5.72}$$

Comparing this equation to equation 5.15 it can be seen that the old, mode dependant, variables, t_0, σ_t, $t_{\rm min}$ and $t_{\rm max}$ can be replaced with the new, mode invariant, primed, variables by reweighting the mean lifetime τ by $M_{\rm B^0}/M_{\rm hh'}$.

For a sample containing only one signal decay using either equation 5.72 or the original equation 5.15 for the lifetime PDF should return the same fit result. To test this a fit using this modified PDF was performed on the same sample that was used to test the original PDF in section 5.7. The fit returned a lifetime of $441.5 \pm 2.9\,\mu$m$/c$, unchanged from the previous test.

The background lifetime distribution also needs to be written in terms of the new lifetime variable t'_0. The PDF describing the background lifetime distribution in section 5.9 used a general parameterisation with no physical meaning and so the same function can be used with the new variables t'_0, $t'_{\rm min}$ and $t'_{\rm max}$. The probability of measuring a lifetime t'_0 given $t'_{\rm min}$ and $t'_{\rm max}$ and given that an event is background is

$$P\left(t'_0 \mid b, t'_{\rm min}, t'_{\rm max}\right) = \begin{cases} \dfrac{y(t'_0)}{\int_{t_{\rm min}}^{t'_{\rm max}} y(t'_0)\mathrm{d}t'_0} & \text{for } t'_{\rm min} \leq t'_0 \leq t'_{\rm max} \\ 0 & \text{for all other } t'_0 \end{cases}, \tag{5.73}$$

where $y(t'_0)$ is given by equation 5.20.

It now remains to add terms to the likelihood for the background probabilities. Care must be taken however as it will be necessary to consider the distributions of

α, $p_{\rm tot}$ and $M_{\pi\pi}$ and to consider how they are correlated with the other variables in the likelihood just as was done for the distribution of $t_{\rm min}$ and $t_{\rm max}$. In addition some discriminating power between the different decay modes and between signal and background needs to be added; lifetime alone will not be sufficient to disentangle the contributions. The separation likelihood described in chapter 4 can be combined with the lifetime likelihood to simultaneously distinguish between decay modes and fit the lifetimes. Once again correlation between variables needs to be considered. The full likelihood function for this analysis is derived and discussed in the next chapter.

5.18 Summary

A data sample that is selected by making requirements on properties of a B meson decay that are indirectly dependent on the lifetime of the B, will have a distribution of lifetimes that is shaped by those requirements. If a lifetime is measured in that sample, neglecting the effect of those requirements, it will be biassed. The data collected by the B_PIPI trigger at CDF is an example of such a sample. This chapter described a method for correcting for this bias, without the use of simulated data, by calculating the lifetime acceptance due to the selection requirements on an event by event basis.

The method is complicated by the addition of background events and different decay modes. It becomes necessary to properly account for the distributions of all variables which appear in the likelihood when these distributions are different for sections of the data. This chapter has described techniques for dealing with some of these considerations but has not fully addressed how the different contributions to the $\mathrm{B}^0_{(s)} \to h^+h'^-$ signal may be included in the likelihood. In particular the distributions of α, $p_{\rm tot}$ and $M_{\pi\pi}$, and their relation to the other variables, have not been included and additional information is required to disentangle the contributions to the data sample. In the next chapter the method to measure an unbiassed lifetime (described in this chapter) will be combined with the method to statistically separate the contributions to the

$B^0_{(s)} \to h^+h'^-$ sample (described in chapter 4) to form a complete likelihood.

Chapter 6

The combined separation and lifetime likelihood function

6.1 Introduction

Chapter 4 described the likelihood function that uses kinematic and particle identification information to disentangle the various contributions to the $\mathrm{B}^0_{(s)} \to h^+h'^-$ and $\Lambda^0_\mathrm{b} \to \mathrm{p}h^-$ signal. Chapter 5 detailed a likelihood function for determining the lifetime of a signal sample when combined with background. Also discussed was how the lifetime probability density function (PDF) for signal should be modified when there is more than one signal decay mode. All of these parts need to be combined in order to measure the $\mathrm{B}^0_\mathrm{s} \to \mathrm{K}^+\mathrm{K}^-$ lifetime in the combined $\mathrm{B}^0_{(s)} \to h^+h'^-$, $\Lambda^0_\mathrm{b} \to \mathrm{p}h^-$ sample. What follows is a derivation of the complete likelihood combining the separation and lifetime PDFs for all signal modes plus background.

The likelihood will contain the following variables:

- t'_0, the measured lifetime calculated using the world average B^0 mass;
- t'_{min}, the minimum t'_0 that could have been accepted by the trigger for that event;

- $t'_{\rm max}$, the maximum t'_0 that could have been accepted by the trigger for that event;
- $M_{\pi\pi}$, the invariant mass of the two final state particles assuming they are pions;
- $p_{\rm tot} = p_1 + p_2$, the scalar sum of the momenta (p_i) of each particle;
- $\alpha = q_1\left(1 - \frac{p_1}{p_2}\right)$, the imbalance between the moduli of the momenta (p_i) of the two particles weighted by the charge (q_1) of the particle with the smaller momentum (p_1);
- ID_1, function of $\mathrm{d}E/\mathrm{d}x$ for the particle with the smaller momentum;
- ID_2, function of $\mathrm{d}E/\mathrm{d}x$ for the particle with the larger momentum.

It is useful to review the parts of the likelihood discussed so far. Equation 5.55 expresses the lifetime probability for one signal mode 's' in two-body decays as

$$P(s, t_0, t_{\rm min}, t_{\rm max} \mid \text{trigger}) = \frac{P(s \mid t_{\rm min}, t_{\rm max}, \text{trigger}) P(t_0 \mid s) P(\text{trigger} \mid s, t_0, t_{\rm min}, t_{\rm max})}{P(\text{trigger} \mid s, t_{\rm min}, t_{\rm max})}. \tag{6.1}$$

Here t_0 is the lifetime measured using the reconstructed B meson mass and $t_{\rm max}$ and $t_{\rm min}$ are the minimum and maximum lifetimes that could have been accepted by the trigger for that event. Equation 5.58 gives an analogous expression for background 'b'.

In section 5.15 a new lifetime variable, t'_0, was defined. This modified lifetime is invariant as the mode hypothesis is changed. Equation 5.72 gives the probability of measuring a modified lifetime t'_0. This probability is equivalent to the terms

$$\frac{P(t_0 \mid s)}{P(\text{trigger} \mid s, t_{\rm min}, t_{\rm max})} \tag{6.2}$$

in equation 6.1 only now the probabilities explicitly depend on $M_{\pi\pi}$, α, $p_{\rm tot}$ and on the mode hypothesis. Also t_0, $t_{\rm min}$ and $t_{\rm max}$ are replaced with t'_0, $t'_{\rm min}$ and $t'_{\rm max}$ to give

$$\frac{P(t'_0 \mid s_j, \alpha, p_{\rm tot}, M_{\pi\pi})}{P(\text{trigger} \mid s_j, t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot}, M_{\pi\pi})}, \tag{6.3}$$

where s_j indicates signal of decay mode type j.

Chapter 4 discusses the likelihood function for statistically separating the contributions to the $B^0_{(s)} \to h^+h'^-$ signal. The trigger was not explicitly considered during this discussion; however, all the distributions of the separation variables included in the likelihood are the distributions after trigger requirements have been applied to the data. For this reason the expression for the signal likelihood given in equation 4.10 can be written as

$$\begin{aligned}\mathcal{L}_j^{sg} &= P(M_{\pi\pi}, \alpha, p_{\rm tot}, ID_1, ID_2, s_j \mid \text{trigger}) \\ &= P(M_{\pi\pi} \mid \alpha, p_{\rm tot}, s_j, \text{trigger}) \cdot P(ID_1, ID_2 \mid \alpha, p_{\rm tot}, s_j, \text{trigger}) \cdot P(\alpha, p_{\rm tot} \mid s_j, \text{trigger})\end{aligned} \tag{6.4}$$

and for background,

$$\begin{aligned}\mathcal{L}_l^{bg} &= P(M_{\pi\pi}, \alpha, p_{\rm tot}, ID_1, ID_2, b_l \mid \text{trigger}) \\ &= P(M_{\pi\pi} \mid \alpha, p_{\rm tot}, b_l, \text{trigger}) \cdot P(ID_1, ID_2 \mid \alpha, p_{\rm tot}, b_l, \text{trigger}) \cdot P(\alpha, p_{\rm tot} \mid b_l, \text{trigger}),\end{aligned} \tag{6.5}$$

where b_l indicates background of type l which is either combinatorial or physics type background.

A combined likelihood for each decay mode might naively be obtained by taking the product of equations 6.4 and 6.1, replacing expression 6.2 with expression 6.3 and substituting t'_0, $t'_{\rm min}$ and $t'_{\rm max}$ for t_0, $t_{\rm min}$ and $t_{\rm max}$. Care must be taken when expanding a likelihood expression in this way as it is easy to miss a term or neglect a correlation between variables. As discussed in sections 4.7 and 5.8, neglecting to correctly describe a distribution that is different for sub-samples of events and that describes variables that vary event by event may lead to a bias.

6.2 Determining the combined likelihood function

In order to determine whether any terms, or correlation between parameters, are missing in the combined likelihood expression the likelihood function should be derived from first principles in the same way as was done in section 5.14 for the lifetime only probability. The following will use the same notation of probability as section 5.14 and the rules of manipulating probabilities expressed in equations 5.44, 5.45 and 5.46.

The data sample only contains events that have met the trigger requirements. The likelihood will be the product of probabilities of measuring the observables t'_0, $t'_{\rm min}$, $t'_{\rm max}$, $p_{\rm tot}$, α, $M_{\pi\pi}$, ID_1 and ID_2 for each event, given that the event passed the trigger:

$$P(t'_0, t'_{\rm min}, t'_{\rm max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}). \tag{6.6}$$

This can be separated into a signal and background part. Using s_j to denote signal of decay mode type j and b_l to denote background of type l the probability can be written as

$$\begin{aligned} &P(t'_0, t'_{\rm min}, t'_{\rm max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) \\ &\qquad = \sum_j P(s_j, t'_0, t'_{\rm min}, t'_{\rm max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) \\ &\qquad\quad + \sum_l P(b_l, t'_0, t'_{\rm min}, t'_{\rm max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}). \end{aligned} \tag{6.7}$$

Consider the probability for one signal mode,

$$P(s_j, t'_0, t'_{\rm min}, t'_{\rm max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}), \tag{6.8}$$

the procedure for background will be similar.

Using Bayes's theorem, the probability for one signal mode can be written as

$$\begin{aligned} &P(s_j, t'_0, t'_{\min}, t'_{\max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) \\ &\qquad = P(s_j, t'_0, t'_{\min}, t'_{\max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2) \\ &\qquad\quad \times \frac{P(\text{trigger} \mid s_j, t'_0, t'_{\min}, t'_{\max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2)}{P(\text{trigger})}. \end{aligned} \tag{6.9}$$

The second term in the numerator, $P(\text{trigger} \mid s_j, t'_0, t'_{\min}, t'_{\max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2)$, is either 0 or a constant ϵ^2 where ϵ is the track finding efficiency of the trigger. This is because the trigger decision is entirely determined by the lifetime acceptance, the decay time and the track finding efficiency. The denominator, $P(\text{trigger})$, is the probability that the trigger fires. It is preferable to write this in terms of all the variables except lifetime, so an expression in terms of $P(\text{trigger} \mid t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2)$ is required. Also it is simpler to calculate this probability separately for signal and background. Using Bayes's theorem,

$$\begin{aligned} P(\text{trigger}) = &P(\text{trigger} \mid s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2) \\ &\times \frac{P(s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2)}{P(s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger})}. \end{aligned} \tag{6.10}$$

In addition, the left hand term in the numerator in equation 6.9 can be written as

$$\begin{aligned} P(s_j, t'_0, t'_{\min}, t'_{\max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2) = &P(s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2) \\ &\times P(t'_0 \mid s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2). \end{aligned} \tag{6.11}$$

Substituting equations 6.10 and 6.11 into equation 6.9 and abbreviating

$$P(\text{trigger} \mid s_j, t'_{\min}, t'_{\max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2) \quad \text{to} \quad P(\text{trigger} \mid \text{all}),$$

equation 6.9 becomes

$$\begin{aligned} &P(s_j, t'_0, t'_{\min}, t'_{\max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) \\ &\quad = P(t'_0 \mid s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2) \\ &\qquad \times \frac{P(s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) P(\text{trigger} \mid \text{all})}{P(\text{trigger} \mid s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2)} \end{aligned} \tag{6.12}$$

Equation 6.12 was obtained by manipulating probabilities and nothing else. To proceed it is necessary to note the following:

1. t'_0, the measured lifetime, is independent of $t'_{\min}$ and $t'_{\max}$ before the trigger;
2. whether an event passes the trigger is independent of particle identification, ID_1 and ID_2.

Equation 6.12 then becomes

$$\begin{aligned} &P(s_j, t'_0, t'_{\min}, t'_{\max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) \\ &\quad = P(t'_0 \mid s_j, \alpha, p_{\rm tot}, M_{\pi\pi}) \\ &\qquad \times \frac{P(s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) P(\text{trigger} \mid \text{all})}{P(\text{trigger} \mid s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi})}. \end{aligned} \tag{6.13}$$

Compare this expression to the lifetime probabilities in equations 6.1, 6.4 and 6.3. The first term, $P(t'_0 \mid s_j, \alpha, p_{\rm tot}, M_{\pi\pi})$, matches the the lifetime probability in the numerator of expression 6.3. The denominator of equation 6.13 matches the one in expression 6.1. The last term in the numerator, $P(\text{trigger} \mid \text{all})$, is equivalent to the last term in the numerator of 6.1. That leaves the penultimate term in the numerator,

$$P(s_j, t'_{\min}, t'_{\max}, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}). \tag{6.14}$$

This can be broken down in a number of ways. The separation of the $\mathrm{B}^0_{(s)} \to h^+h'^-$ modes requires a term that gives the probability of measuring $M_{\pi\pi}$ given the momentum

variables and given the mode. The separation also requires the probability of measuring ID_1 and ID_2 given the momentum variables and decay mode. Even with both of these terms included there is still some flexibility in how 6.14 is parameterised. Three possibilities are considered below.

Parameterisation 1

$$\begin{aligned} P(t'_{\min}, t'_{\max}, s_j, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2 \,|\, \text{trigger}) = {} & P(s_j \mid \text{trigger}) \\ & \times P(\alpha, p_{\rm tot} \mid s_j, \text{trigger}) \\ & \times P(M_{\pi\pi} \mid s_j, \alpha, p_{\rm tot}, \text{trigger}) \\ & \times P(ID_1, ID_2 \mid s_j, M_{\pi\pi}, \alpha, p_{\rm tot}, \text{trigger}) \\ & \times P(t'_{\min}, t'_{\max} \mid s_j, M_{\pi\pi}, \alpha, p_{\rm tot}, ID_1, ID_2, \text{trigger}). \end{aligned} \tag{6.15}$$

The first term, $P(s_j \mid \text{trigger})$, is the fraction of the data sample that is decay mode s_j. The second and third terms can be recognised as the momentum and mass terms in the separation PDF (equation 6.4). The fourth term, $P(ID_1, ID_2 \mid s_j, M_{\pi\pi}, \alpha, p_{\rm tot}, \text{trigger})$, is the same as the particle identification term in the separation PDF as long as it is noted that

1. ID_1 and ID_2 are independent of $M_{\pi\pi}$.

This was used already to form the separation only PDF. The last term, $P(t'_{\min}, t'_{\max} \mid s_j, M_{\pi\pi}, \alpha, p_{\rm tot}, ID_1, ID_2, \text{trigger})$, is the probability of observing a particular combination of $t'_{\min}$ and $t'_{\max}$ given the decay mode, α, $p_{\rm tot}$, $M_{\pi\pi}$, ID_1 and ID_2. If the distribution of $t'_{\min}$ and $t'_{\max}$ given the separation variables is the same for all signal modes (and background as this term will also occur in the background part of the likelihood) then it is safe to neglect it as it would be a constant term that vanishes when the logarithm of the likelihood is maximised. If this distribution differs between modes, or between signal and background, then neglecting it will introduce a bias. This analysis requires

the assumption that

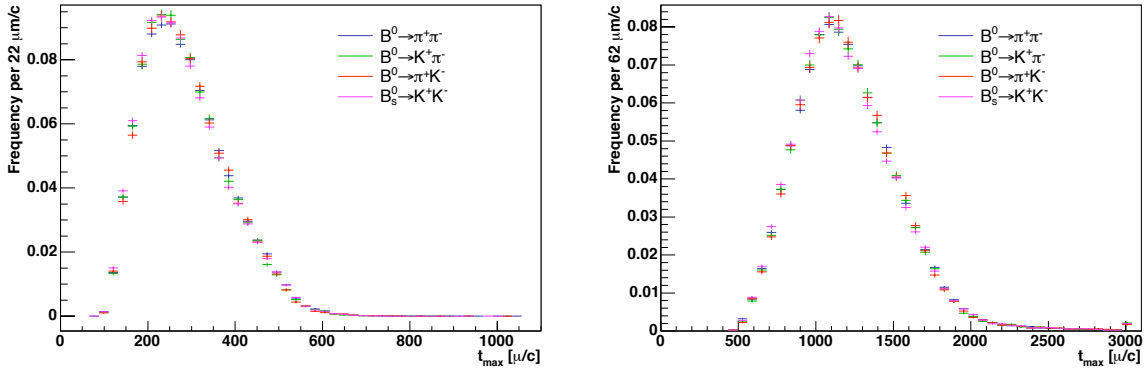
2. the distribution of $t'_{\rm min}$ and $t'_{\rm max}$, given the separation variables, is the same for all signal modes and background,

The last term is neglected and a systematic uncertainty is assigned to account for the bias caused, see section 7.3 for details.

Parameterisation 2

Parameterisation 1 above produced all of the terms necessary for the separation of the different signal contributions, but it also a produced a term that could not easily be modelled and had to be neglected: $P(t'_{\rm min}, t'_{\rm max} \mid s_j, M_{\pi\pi}, \alpha, p_{\rm tot}, ID_1, ID_2, \text{trigger})$. Consider the lifetime only PDF discussed in chapter 5. In section 5.8 the simple lifetime probability in equation 5.17 included a term $P(t_{\rm min}, t_{\rm max} \mid s)$. Rather than try to parameterise this term the PDF was rewritten to include the terms $P(s \mid t_{\rm min}, t_{\rm max})P(t_{\rm min}, t_{\rm max})$. $P(t_{\rm min}, t_{\rm max})$ could be neglected because it is independent of the type of the event and of lifetime and so factors out of the likelihood. The term $P(s \mid t_{\rm min}, t_{\rm max})$ could be modelled as described in sections 5.10 to 5.13. Perhaps the term $P(t'_{\rm min}, t'_{\rm max} \mid s_j, M_{\pi\pi}, \alpha, p_{\rm tot}, ID_1, ID_2, \text{trigger})$ in parameterisation 1 above could be dealt with in a similar way. To attempt this, expression 6.14 is parameterised as:

$$\begin{aligned} P(t'_{\rm min}, t'_{\rm max}, s_j, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) = {} & P(t'_{\rm min}, t'_{\rm max} \mid \text{trigger}) \\ & \times P(s_j \mid t'_{\rm min}, t'_{\rm max}, \text{trigger}) \\ & \times P(\alpha, p_{\rm tot} \mid s_j, t'_{\rm min}, t'_{\rm max}, \text{trigger}) \\ & \times P(M_{\pi\pi} \mid s_j, \alpha, p_{\rm tot}, t'_{\rm min}, t'_{\rm max}, \text{trigger}) \\ & \times P(ID_1, ID_2 \mid s_j, M_{\pi\pi}, \alpha, p_{\rm tot}, t'_{\rm min}, t'_{\rm max}, \text{trigger}). \end{aligned} \tag{6.16}$$



(a) $t_{\min}$ (b) $t_{\max}$

Figure 6.1: Comparison of the distribution of $t'_{\min}$ and $t'_{\max}$ for each $\mathrm{B}^0_{(s)} \to h^+h'^-$ mode in realistic simulation.

The first term, $P(t'_{\min}, t'_{\max} \mid \text{trigger})$, does not depend on lifetime and is identical for all signal and background so can be ignored. The second term, $P(s_j \mid t'_{\min}, t'_{\max}, \text{trigger})$, is similar to the signal probability in equation 6.1. Noting that $P(s \mid s_j) = 1$, this term can be written

$$P(s_j \mid t'_{\min}, t'_{\max}, \text{trigger}) = P(s \mid t'_{\min}, t'_{\max}, \text{trigger}) P(s_j \mid s, t'_{\min}, t'_{\max}, \text{trigger}). \quad (6.17)$$

The first term can be dealt with using the methods described in sections 5.10 to 5.13. The second term can be simplified with the assumption:

1. the probability that a signal event came from a particular decay mode is independent of the observed $t'_{\min}$ and $t'_{\max}$.

Figure 6.1 shows the distributions of $t_{\min}$ and $t_{\max}$ for $\mathrm{B}^0 \to \pi^+\pi^-$, $\mathrm{B}^0 \to \mathrm{K}^+\pi^-$, $\overline{\mathrm{B}^0} \to \pi^+\mathrm{K}^-$ and $\mathrm{B}^0_\mathrm{s} \to \mathrm{K}^+\mathrm{K}^-$ decays from the realistic simulation. There is no observable difference between the distributions so the assumption above is reasonable. With this assumption $P(s_j \mid s, t'_{\min}, t'_{\max}, \text{trigger})$ reduces to $P(s_j \mid s, \text{trigger})$ which is simply f_j, the fraction of signal that is of decay mode type j.

Leaving the third term in equation 6.16 for a moment, the fourth term,

$$P(M_{\pi\pi} \mid s_j, \alpha, p_{\rm tot}, t'_{\rm min}, t'_{\rm max}, \text{trigger}), \tag{6.18}$$

requires another assumption:

2. $M_{\pi\pi}$ is independent of $t'_{\rm min}$ and $t'_{\rm max}$.

Using this assumption this term reduces to $P(M_{\pi\pi} \mid s_j, \alpha, p_{\rm tot}, \text{trigger})$ which is identical to the mass term in equation 6.4. The last term can be reduced by making the assumption:

3. ID_1 and ID_2 are independent of $t'_{\rm min}$, $t'_{\rm max}$ and $M_{\pi\pi}$.

This assumption reduces this term to the particle identification probability in equation 6.4, $P(ID_1, ID_2 \mid s_j, \alpha, p_{\rm tot}, \text{trigger})$. Equation 6.16 is now:

$$\begin{aligned} P(t'_{\rm min}, t'_{\rm max}, s_j, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) = {} & P(s_j \mid s, \text{trigger}) \\ & \times P(s \mid t'_{\rm min}, t'_{\rm max}, \text{trigger}) \\ & \times P(\alpha, p_{\rm tot} \mid t'_{\rm min}, t'_{\rm max}, s_j, \text{trigger}) \\ & \times P(M_{\pi\pi} \mid \alpha, p_{\rm tot}, s_j, \text{trigger}) \\ & \times P(ID_1, ID_2 \mid \alpha, p_{\rm tot}, s_j, \text{trigger}). \end{aligned} \tag{6.19}$$

The term $P(\alpha, p_{\rm tot} \mid t'_{\rm min}, t'_{\rm max}, \text{trigger}, s_j)$ is the probability of measuring this α and $p_{\rm tot}$ given $t'_{\rm min}$ and $t'_{\rm max}$, given that the event passed the trigger and given the signal mode. Note that the joint distribution of α and $p_{\rm tot}$ is assumed to be different for each $\mathrm{B}^0_{(s)} \to h^+h'^-$ decay mode. In separation-only fits, which use equation 6.4, there is a term for the probability of measuring α and $p_{\rm tot}$ given the decay mode and given that the event passed the trigger, that is $P(\alpha, p_{\rm tot} \mid s_j, \text{trigger})$. To use this instead of the third term above it is necessary to assume that:

4. α and $p_{\rm tot}$ are independent of $t'_{\rm min}$ and t'_{max}.

By rewriting the PDF to use the term $P(s \mid t'_{\rm min}, t'_{\rm max}, \text{trigger})$ rather than $P(t'_{\rm min}, t'_{\rm max} \mid s_j, M_{\pi\pi}, \alpha, p_{\rm tot}, ID_1, ID_2, \text{trigger})$ the momentum, mass and particle identification terms have become dependent on $t'_{\rm min}$ and $t'_{\rm max}$. To use this PDF the uncertainty due to neglecting the term $P(t'_{\rm min}, t'_{\rm max} \mid s_j, M_{\pi\pi}, \alpha, p_{\rm tot}, ID_1, ID_2, \text{trigger})$ is replaced by the uncertainty due to neglecting the dependence on $t'_{\rm min}$ and $t'_{\rm max}$ of the momentum, mass and particle identification variables. This is discussed further in section 7.4.

Parameterisation 3

Parameterisation 2 above required assumptions about the momentum dependence of the joint distribution of $t'_{\rm min}$ and $t'_{\rm max}$. An alternative parameterisation of equation 6.14 is

$$\begin{aligned} P(t'_{\rm min}, t'_{\rm max}, s_j, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) = {} & P(t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot} \mid \text{trigger}) \\ & \times P(s_j \mid t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot}, \text{trigger}) \\ & \times P(M_{\pi\pi} \mid \alpha, p_{\rm tot}, t'_{\rm min}, t'_{\rm max}, \text{trigger}, s_j) \\ & \times P(ID_1, ID_2 \mid s_j, M_{\pi\pi}, \alpha, p_{\rm tot}, t'_{\rm min}, t'_{\rm max}, \text{trigger}). \end{aligned} \tag{6.20}$$

The first term, similarly to the first term in parameterisation 2, does not depend on lifetime and is identical for signal and background so can be ignored. The last two terms are identical to the last two terms in parameterisation 2 and can be reduced, using assumptions 2 and 3 in the last section, to $P(M_{\pi\pi} \mid s_j, \alpha, p_{\rm tot}, \text{trigger})$ and $P(ID_1, ID_2 \mid$

$s_j, \alpha, p_{\rm tot}, \text{trigger})$. Equation 6.20 thus reduces to

$$\begin{aligned} P(t'_{\rm min}, t'_{\rm max}, s_j, \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}) = {} & P(s_j \mid t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot}, \text{trigger}) \\ & \times P(M_{\pi\pi} \mid \alpha, p_{\rm tot}, \text{trigger}, s_j) \\ & \times P(ID_1, ID_2 \mid s_j, \alpha, p_{\rm tot}, \text{trigger}). \end{aligned} \tag{6.21}$$

The term, $P(s_j \mid t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot}, \text{trigger})$, looks similar to the signal probability in equation 6.1 and in equation 6.19. In the same way as was done in parameterisation 2, this term can be written:

$$\begin{aligned} & P(s_j \mid t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot}, \text{trigger}) \\ & \qquad = P(s \mid t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot}, \text{trigger}) P(s_j \mid s, t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot}, \text{trigger}). \end{aligned} \tag{6.22}$$

The first term above is similar to the first term in 6.17. In that case the term was dealt with using a linear Fisher discriminant analysis. It would be simple to add α and $p_{\rm tot}$ as discriminating variables and then approximate $P(s \mid t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot}, \text{trigger})$ with $P(s \mid \kappa')$. In this case, however, reducing four variables to one may not retain sufficient information to make this a good approximation, particularly if any correlation between $t'_{\rm min}$, $t'_{\rm max}$ and α, $p_{\rm tot}$ is non-linear.

The second term in 6.22 is similar to the second term in 6.17 but the probability that an event is of decay mode j is now dependant on α and $p_{\rm tot}$. In order to reduce this term to $P(s_j \mid s, \text{trigger})$ it is now necessary to assume not only that the probability that a signal event is from a particular decay mode is independent of $t'_{\rm min}$ and $t'_{\rm max}$ (assumption 1 in the last section) but also that

4. the probability that a signal event is from a particular decay mode is independent of the observed α and $p_{\rm tot}$.

However, the distributions of α and $p_{\rm tot}$ are expected to be different between decay

modes; this is what the term $P(\alpha, p_{\rm tot} \mid s_j, \text{trigger})$ in the separation only probability (equation 6.4) describes.

Parameterisation 3 offers no advantage over parameterisation 2 and is not considered any further. Parameterisation 1 is the version of the PDF used in this analysis. Parameterisation 2 was also tested, but was found to have larger systematic uncertainties than parameterisation 1 (as discussed in section 7.4).

6.3 The full likelihood

The likelihood function for the ith event, combining the determination of signal lifetime and the separation of the different $\mathrm{B}^0_{(s)} \to h^+h'^-$ modes, is given by the sum of the signal and background likelihoods:

$$\mathcal{L}_i = \mathcal{L}^{sg} + \mathcal{L}^{bg}. \tag{6.23}$$

The signal part of the likelihood is given by

$$\mathcal{L}^{sg} = \sum_j P(s_j, t'_0, t'_{\rm min}, t'_{\rm max}, p_{\rm tot}, \alpha, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}). \tag{6.24}$$

Expanded using parameterisation 1 in section 6.2 this gives

$$\mathcal{L}^{sg} = P(\text{trigger} \mid \text{all}) \sum_j P(s_j \mid \text{trigger}) \cdot \mathcal{L}_j^{sep\ sg} \cdot \mathcal{L}_j^{life\ sg}, \tag{6.25}$$

where the term $P(\text{trigger} \mid \text{all})$ has been brought in front of the summation because it is the same for all signal modes. The separation part of the signal likelihood is

$$\mathcal{L}_j^{sep\ sg} = P(\alpha, p_{\rm tot} \mid s_j, \text{trigger}) P(M_{\pi\pi} \mid s_j, \alpha, p_{\rm tot}, \text{trigger}) P(ID_1, ID_2 \mid s_j, \alpha, p_{\rm tot}, \text{trigger}) \tag{6.26}$$

and the lifetime part is

$$\mathcal{L}_j^{life\ sg} = \frac{P(t'_0 \mid s_j, \alpha, p_{\rm tot}, M_{\pi\pi})}{P(\text{trigger} \mid s_j, t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot}, M_{\pi\pi}).} \tag{6.27}$$

The terms in the signal likelihood are:

- $P(\text{trigger} \mid \text{all})$. The probability that the trigger fires given all of the fit variables. As discussed in section 5.14, if it is assumed that the trigger track-finding efficiency, ϵ, is constant for all particles within the trigger acceptance then his term is either ϵ^2 or 0 for two-body decays and can be neglected. The affect of a varying track-finding efficiency is considered as a systematic uncertainty in sections 7.6 7.7 and 7.16.3;

- $P(s_j \mid \text{trigger})$. This is the fraction of data that is of decay mode type j, $P(s_j \mid \text{trigger}) = f_j$;

- $P(M_{\pi\pi} \mid s_j, \alpha, p_{\rm tot}, \text{trigger})$. The mass term gives probability of measuring $M_{\pi\pi}$ given the signal mode s_j and given the momentum variables α and $p_{\rm tot}$. It is given by equation 4.13;

- $P(ID_1, ID_2 \mid s_j, p_{\rm tot}, \alpha, \text{trigger})$. The particle identification term gives the probability of measuring ID_1 and ID_2 given the signal mode s_j and the momentum variables α and $p_{\rm tot}$. It is given by equation 4.28;

- $P(\alpha, p_{\rm tot} \mid s_j, \text{trigger})$. The momentum term gives the probability of measuring α and $p_{\rm tot}$ given the signal mode, s_j. It is given by equation 4.32;

- $P(t'_0 \mid s_j, \alpha, p_{\rm tot}, M_{\pi\pi})$. This is the probability of measuring a lifetime t'_0 given the signal mode, $M_{\pi\pi}$, α and $p_{\rm tot}$. This is the numerator in equation 5.72:

$$P(t'_0 \mid s_j, \alpha, p_{\rm tot}, M_{\pi\pi}) = \int_0^\infty \frac{M_{\rm hh'}(M_{\pi\pi}, \alpha, p_{\rm tot})}{M_{\rm B^0}} \frac{1}{\tau} e^{-\frac{t'}{\tau}\frac{M_{\rm hh'}(M_{\pi\pi},\alpha,p_{\rm tot})}{M_{\rm B^0}}} \frac{1}{\sqrt{2\pi}\sigma_{t'}} e^{-\frac{(t'-t'_o)^2}{2\sigma_{t'}{}^2}} \,{\rm d}t'; \tag{6.28}$$

- $P(\text{trigger} \mid s_j, t'_{\text{min}}, t'_{\text{max}}, \alpha, p_{\text{tot}}, M_{\pi\pi})$. The probability that the trigger fires given an event is signal and given t'_{min}, t'_{max}, $M_{\pi\pi}$, α and p_{tot}. This is the normalisation of the signal lifetime probability and the denominator in equation 5.72:

$$P(\text{trigger} \mid s_j, t'_{\text{min}}, t'_{\text{max}}, \alpha, p_{\text{tot}}, M_{\pi\pi}) = \int_{t'_{\text{min}}}^{t'_{\text{max}}} \int_0^{\infty} \frac{M_{\text{hh}'}(M_{\pi\pi}, \alpha, p_{\text{tot}})}{M_{\text{B}^0}} \frac{1}{\tau} e^{-\frac{t'}{\tau}\frac{M_{\text{hh}'}(M_{\pi\pi},\alpha,p_{\text{tot}})}{M_{\text{B}^0}}} \frac{1}{\sqrt{2\pi}\sigma_{t'}} e^{-\frac{(t'-t'_0)^2}{2\sigma_{t'}{}^2}} \,\mathrm{d}t' \,\mathrm{d}t'_0. \tag{6.29}$$

The background part of the likelihood is given by

$$\mathcal{L}^{bg} = \sum_l P(b_l, t'_0, t'_{\text{min}}, t'_{\text{max}}, p_{\text{tot}}, \alpha, M_{\pi\pi}, ID_1, ID_2 \mid \text{trigger}). \tag{6.30}$$

Expanded this gives

$$\mathcal{L}^{bg} = P(\text{trigger} \mid \text{all}) \sum_l P(b_l \mid \text{trigger}) \cdot \mathcal{L}_l^{sep\,bg} \cdot \mathcal{L}_l^{life\,bg}, \tag{6.31}$$

where

$$\mathcal{L}_l^{sep\,bg} = P(\alpha, p_{\text{tot}} \mid b_l, \text{trigger}) P(M_{\pi\pi} \mid b_l, \alpha, p_{\text{tot}}, \text{trigger}) P(ID_1, ID_2 \mid b_l, \alpha, p_{\text{tot}}, \text{trigger}) \tag{6.32}$$

and

$$\mathcal{L}_l^{life\,bg} = \frac{P(t'_0 \mid b_l, \alpha, p_{\text{tot}}, M_{\pi\pi})}{P(\text{trigger} \mid b_l, t'_{\text{min}}, t'_{\text{max}}, \alpha, p_{\text{tot}}, M_{\pi\pi})}. \tag{6.33}$$

The terms in the background likelihood are:

- $P(\text{trigger} \mid \text{all})$. The probability that the trigger fires given all of the fit variables. As for the corresponding signal term, this term can be neglected if it is assumed that the trigger track-finding efficiency, ϵ, is constant;

- $P(b_l \mid \text{trigger})$. This is the fraction of data that is of background type l, $P(b_l \mid \text{trigger}) = f_l$;

- $P(M_{\pi\pi} \mid b_l, \alpha, p_{\rm tot}, \text{trigger})$. The mass term gives the probability of measuring $M_{\pi\pi}$ given the background type b_l and given the momentum variables α and $p_{\rm tot}$. In fact the background $M_{\pi\pi}$ distribution is considered independent of α and $p_{\rm tot}$. For combinatorial background it is given by equation 4.18 and for physics-type background by equation 4.20;

- $P(ID_1, ID_2 \mid b_l, p_{\rm tot}, \alpha)$. The particle identification term gives the probability of measuring ID_1 and ID_2 given the background type, b_l, and the momentum variables α and $p_{\rm tot}$. It is given by equation 4.29;

- $P(\alpha, p_{\rm tot} \mid b_l, \text{trigger})$. The momentum term gives the probability of measuring α and $p_{\rm tot}$ given that the event is background. The parameterisation used is the same for both combinatorial and physics-type background. It is given by equation 4.33;

- $P(t'_0 \mid b_l, \alpha, p_{\rm tot}, M_{\pi\pi})$. It is assumed that the background lifetime is independent of α, $p_{\rm tot}$ and $M_{\pi\pi}$. This probability then reduces to $P(t'_0 \mid b)$, the probability of measuring a lifetime t'_0 given that an event is background. This is given by the numerator in equation 5.73, $y(t'_0)$. The systematic uncertainty due assuming the background lifetime is independent of α, $p_{\rm tot}$ and $M_{\pi\pi}$ is considered in sections 7.8 and 7.9;

- $P(\text{trigger} \mid b_l, t'_{\rm min}, t'_{\rm max}, \alpha, p_{\rm tot}, M_{\pi\pi})$. The probability that the trigger fires given an event is background and given $t'_{\rm min}$, $t'_{\rm max}$, $M_{\pi\pi}$, α and $p_{\rm tot}$. This is the denominator in equation 5.73:
$$\int_{t'_{\rm min}}^{t'_{\rm max}} y(t'_0)\mathrm{d}t'_0. \tag{6.34}$$

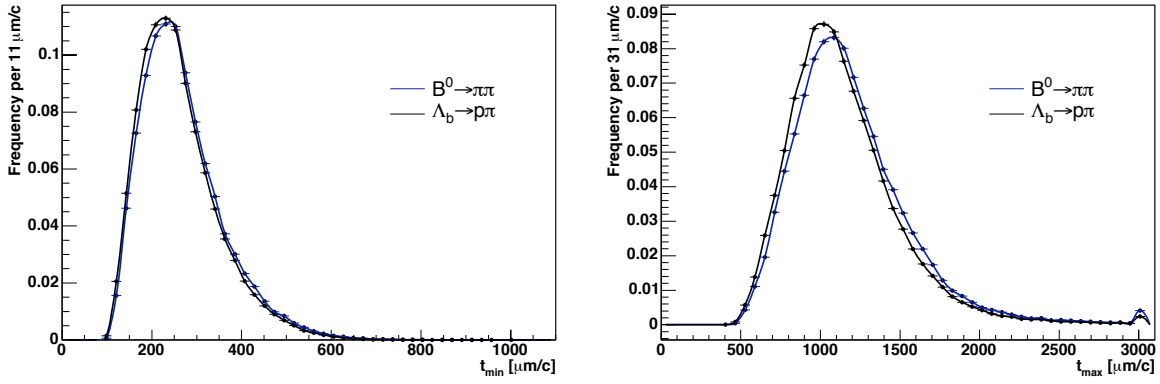
6.4 The fitting procedure

Some consideration is required as to what constitutes signal and background as far as the likelihood is concerned. The background is of two types, combinatorial background

and background from misreconstructed three-body B decays (physics background). As the physics background originates from real long-lived particles the distribution of measured lifetimes in this background is likely to be different to that in the combinatorial background and would need an independent model. The physics-type background is kinematically suppressed above 5.15 GeV/c^2, so by excluding events with $M_{\pi\pi} < 5.16$ GeV/c^2 this type of background can be excluded from the data sample. This mass requirement is below the $\mathrm{B_s^0} \to \mathrm{K^+K^-}$ invariant $\pi\pi$-mass peak and so the physics-type background can be removed without reducing the number of $\mathrm{B_s^0} \to \mathrm{K^+K^-}$ events.

The signal contains a number of $\mathrm{B^0_{(}}_{s)} \to h^+h'^-$ and $\Lambda_\mathrm{b}^0 \to \mathrm{p}h^-$ decay modes. The kaon and pion masses are very close so the kinematic distributions for each of the $\mathrm{B^0_{(}}_{s)} \to h^+h'^-$ modes are very similar. This is why the distributions of $t_{\min}$ and $t_{\max}$ for the $\mathrm{B^0_{(}}_{s)} \to h^+h'^-$ modes look the same (figure 6.1). The proton mass is much larger than either the pion or kaon masses so the kinematic distributions, and hence the distributions of $t_{\min}$ and $t_{\max}$, in the $\Lambda_\mathrm{b}^0 \to \mathrm{p}h^-$ decays will be different from $\mathrm{B^0_{(}}_{s)} \to h^+h'^-$. The likelihood described in section 6.3 required the assumption that the distribution of $t_{\min}$ and $t_{\max}$ was the same for all types of data. Including $\Lambda_\mathrm{b}^0 \to \mathrm{p}h^-$ decays will increase the systematic uncertainty due to this assumption. The $\Lambda_\mathrm{b}^0 \to \mathrm{p}h^-$ decays lie at a higher mass than the $\mathrm{B^0_{(}}_{s)} \to h^+h'^-$ modes and can be removed from the data sample by excluding the region $5.35 < M_{\pi\pi} < 5.55$ GeV/c^2 with little contamination from $\Lambda_\mathrm{b}^0 \to \mathrm{p}h^-$ decays underneath the $\mathrm{B^0_{(}}_{s)} \to h^+h'^-$ mass peak. Figure 6.2 shows the distributions of $t_{\min}$ and $t_{\max}$ for $\Lambda_\mathrm{b}^0 \to \mathrm{p}\pi^-$ and $\mathrm{B^0} \to \pi^+\pi^-$ events from the fast simulation. The mean $t_{\max}$ is lower for $\Lambda_\mathrm{b}^0 \to \mathrm{p}\pi^-$ decays than $\mathrm{B^0} \to \pi^+\pi^-$ but the difference between the two types of decay is much smaller than the difference between signal and background (figure 5.14). Since this difference is small it may be possible to include the $\Lambda_\mathrm{b}^0 \to \mathrm{p}h^-$ mass region in a future analysis with only a small increase in the systematic uncertainty. An additional lifetime parameter would need to be included for the Λ_b^0 lifetime. In this analysis, however, this contribution is excluded.

The full separation and lifetime log-likelihood maximisation is performed on data in



(a) $t_{\min}$ (b) $t_{\max}$

Figure 6.2: The distribution of $t_{\min}$ and $t_{\max}$ for fast simulation $\Lambda_b^0 \to p\pi^-$ and $B^0 \to \pi^+\pi^-$ events.

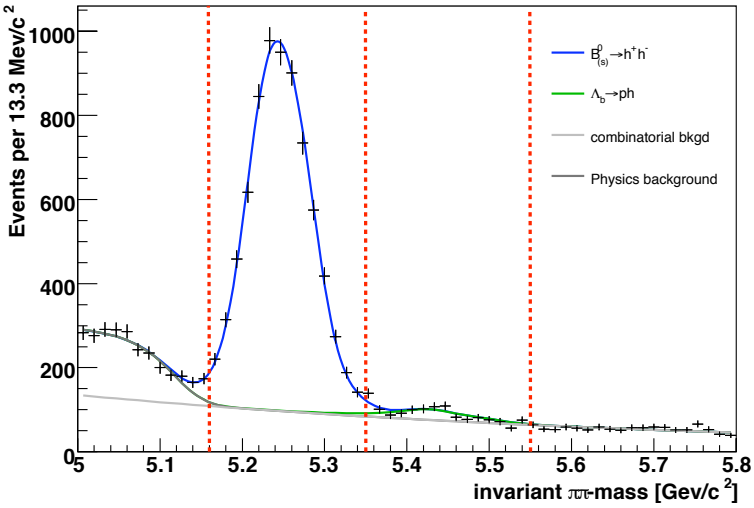


Figure 6.3: The combined separation and lifetime fit is performed for data in the mass regions $5.16 < M_{\pi\pi} < 5.35$ GeV/c^2 and $5.55 < M_{\pi\pi} < 5.8$GeV/c^2 indicated by the dashed lines.

the mass regions $5.16 < M_{\pi\pi} < 5.35$ GeV/c^2 and $5.55 < M_{\pi\pi} < 5.8$ GeV/c^2, indicated in figure 6.3. The performance of this likelihood will be discussed in the next chapter.

The likelihood described in parameterisation 2, section 6.2 is tested in section 7.4. The lifetime measurement using this likelihood is performed in stages. The likelihood requires that, for each event, $t_{\min}$, $t_{\max}$ are translated to a single variable κ. This is done by projecting ($t_{\min}$, $t_{\max}$) onto the Fisher direction chosen using the Fisher linear discriminant analysis explained in section 5.11. The Fisher direction is determined

from the data as described in section 5.12 using the separation only fit from chapter 4 to define the signal and background region. The first step is, therefore, to perform a separation only, maximum log-likelihood fit to the data sample using the likelihood described in chapter 4. This can then be used to find the Fisher direction for the data as described in section 5.12. The Fisher direction is then used to calculate κ for each event as described in section 5.13. After these preliminary steps have been taken, all of the variables used in the alternative combined separation and lifetime likelihood (from parameterisation 2, section 6.2) are known for each event.

6.5 Summary

In this chapter, a complete likelihood has been derived that combines statistical separation of $\mathrm{B}^0_{(s)} \to h^+h'^-$ and background contributions to a data sample, with determination of the mean lifetimes of the B^0 and B^0_s mesons. This can be done in a number of ways and two potentially useful parameterisations of the likelihood have been identified. Both of these likelihood functions require different assumptions to be made in order to be useful in practice. The version of the likelihood described in section 6.3 is the one used for the final measurement and the performance of this likelihood, as well as the systematic uncertainties due to the assumptions which have been made, are assessed in the next chapter. The alternative likelihood is also tested, in section 7.4, and compared to the version of the likelihood which is finally used.

Chapter 7

Performance of the combined separation and lifetime likelihood maximisation

7.1 Introduction

Chapter 6 describes the likelihood used to simultaneously disentangle the various contributions to the $B^0_{(s)} \to h^+h'^-$ signal and measure the mean lifetimes of the long lived components of that signal. This chapter is concerned with testing the performance of that likelihood. This is done by performing ensembles of pseudo-experiments on simulated data samples. For each data sample the lifetime returned by the fit ($\tau_{\rm fit}$) can be compared to the known input lifetime ($\tau_{\rm truth}$) and a pull can be estimated for that event as

$$\text{pull} = \frac{\tau_{\rm fit} - \tau_{\rm truth}}{\sigma(\tau_{\rm fit})}, \tag{7.1}$$

where $\sigma(\tau_{\rm fit})$ is the statistical error on the lifetime returned by the fit. If the fit is unbiased and the statistical errors are correctly estimated then the distribution of sample pulls will be Gaussian with a mean of 0 and width of 1. The mean of the distribution

of residuals ($\tau_{\rm fit} - \tau_{\rm truth}$) can be used to estimate the bias in the measurement.

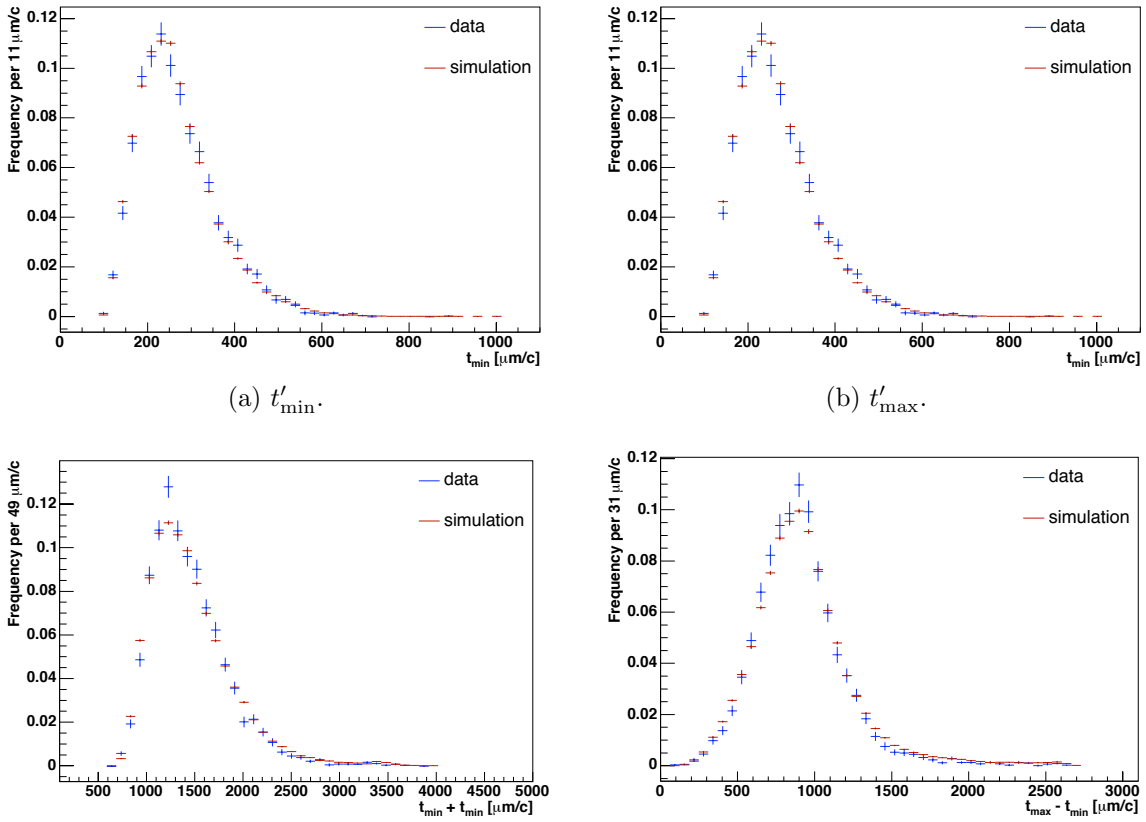
Simulated data samples for pseudo-experiments are generated using the fast simulation described in section 3.6. In that section the distributions of several properties of the events generated were compared to the distributions in data. It is now clear from chapters 4, 5 and 6 that additional variables (such as the minimum and maximum accepted lifetimes, particle momentum imbalance α and particle identification variables ID_1 and ID_2) are needed. How these are generated in the fast simulation and how their distributions compare to data are discussed in section 7.2 below.

The remainder of this chapter then discusses causes of systematic uncertainty on the lifetime measurement and an alternative way to parameterise the likelihood.

7.2 Fast simulation

Section 3.6 described a fast simulation for generating large simulated samples of events quickly. The event by event minimum and maximum accepted lifetimes are calculated for the simulated sample in exactly the same way as for real data. Figure 7.1 compares the distribution of $t'_{\rm min}$ and $t'_{\rm max}$ for the fast simulation of $\mathrm{B}^0 \to \pi^+\pi^-$ events to background subtracted signal events in data; the simulation models the data well. Figure 7.2 compares the distribution of $t'_{\rm min}$ and $t'_{\rm max}$ for the fast simulation of background events to background events taken from the mass region $5.6 < M_{\pi\pi} < 5.8$ GeV/c^2 in data. The simulation does not replicate the data exactly but does approximates the shapes of the distributions observed in data. The simulations of signal and background do introduce a difference between the joint distribution of $t'_{\rm min}$ and $t'_{\rm max}$ in signal and background (see figure 7.3) allowing the effect of this difference to be studied using the fast simulation.

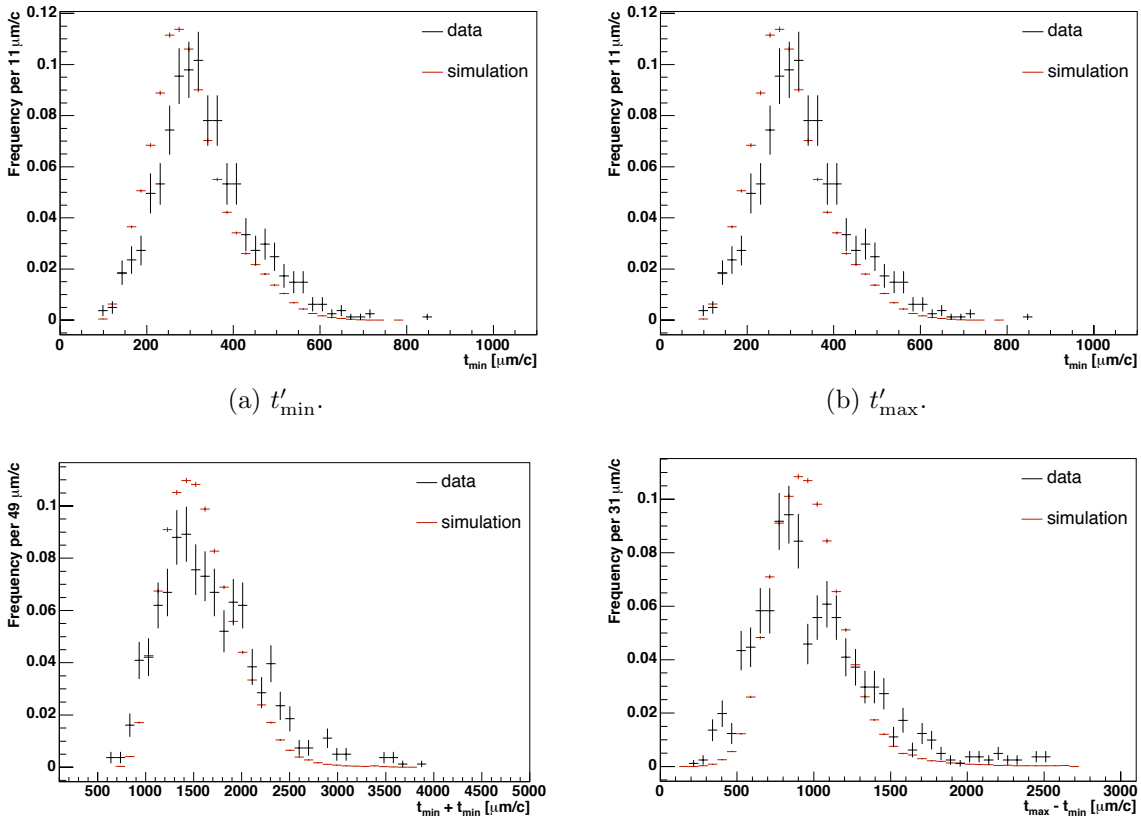
The simulation also needs to generate the variables used by the separation fit described in chapter 4. Those variables are:



(a) $t'_{\min}$. (b) $t'_{\max}$.

(c) $t'_{\max} + t'_{\min}$. (d) $t'_{\max} - t'_{\min}$.

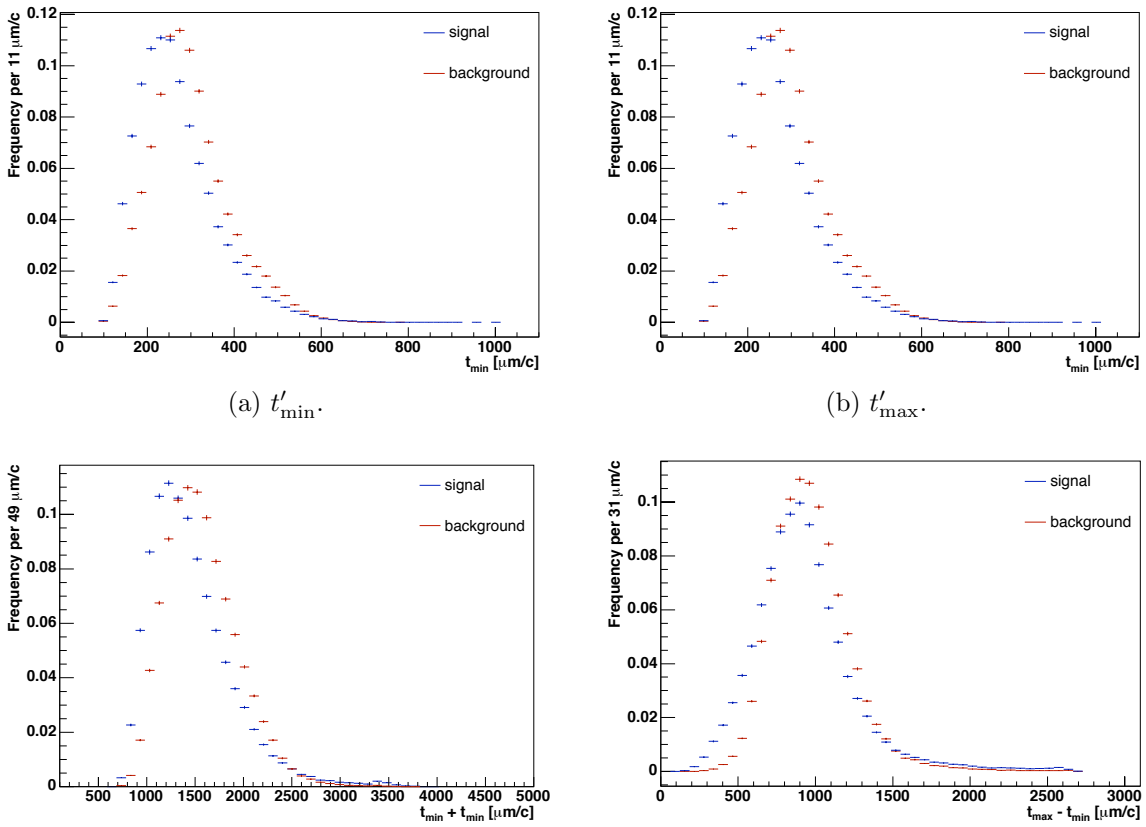
Figure 7.1: Comparison of the joint distribution of $t'_{\min}$ and $t'_{\max}$ in fast simulation $B^0 \to \pi^+\pi^-$ events and background subtracted signal events in data. The joint distribution of $t'_{\min}$ and $t'_{\max}$ has been projected onto 4 directions: $t'_{\min}$; $t'_{\max}$; $(t'_{\max}+t'_{\min})$; $(t'_{\max}-t'_{\min})$. The fast simulation models the data well.



(a) $t'_{\rm min}$. (b) $t'_{\rm max}$.

(c) $t'_{\rm max} + t'_{\rm min}$. (d) $t'_{\rm max} - t'_{\rm min}$.

Figure 7.2: Comparison of the joint distribution of $t'_{\rm min}$ and $t'_{\rm max}$ in fast simulation background events and background events taken from the mass region $5.6 < M_{\pi\pi} < 5.8$ GeV/c^2 in data. The joint distribution of $t'_{\rm min}$ and $t'_{\rm max}$ has been projected onto 4 directions: $t'_{\rm min}$; $t'_{\rm max}$; $(t'_{\rm max} + t'_{\rm min})$; $(t'_{\rm max} - t'_{\rm min})$. The fast simulation does not replicate the data exactly but does approximate the shapes of the distributions in data.



(a) t'_{min}.

(b) t'_{max}.

(c) $t'_{\mathrm{max}} + t'_{\mathrm{min}}$.

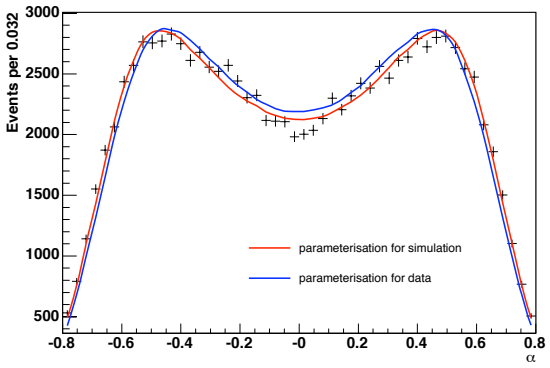
(d) $t'_{\mathrm{max}} - t'_{\mathrm{min}}$.

Figure 7.3: Comparison of the joint distribution of t'_{min} and t'_{max} in fast simulation $\mathrm{B}^0 \to \pi^+\pi^-$ (signal) and background events. The joint distribution of t'_{min} and t'_{max} has been projected onto 4 directions: t'_{min}; t'_{max}; $(t'_{\mathrm{max}} + t'_{\mathrm{min}})$; $(t'_{\mathrm{max}} - t'_{\mathrm{min}})$. The fast simulation produces different distributions for signal and background. The size of the difference compares well to that observed in data (see figure 5.19).

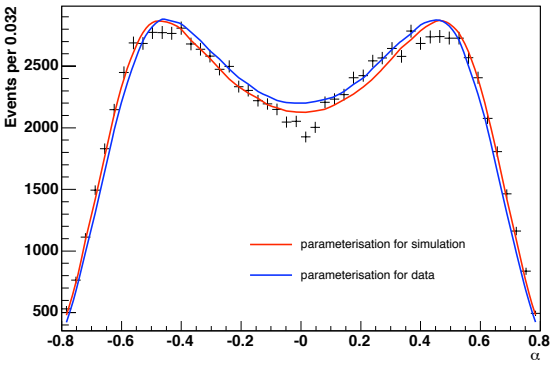
- $M_{\pi\pi}$, the invariant mass of the two final state particles assuming they are pions;
- $p_{\mathrm{tot}} = p_1 + p_2$, the scalar sum of the momenta (p_i) of each particle;
- $\alpha = q_1 \left(1 - \frac{p_1}{p_2}\right)$, the imbalance between the moduli of the momenta (p_i) of the two particles weighted by the charge (q_1) of the particle with the smaller momentum (p_1);
- ID_1, function of $\mathrm{d}E/\mathrm{d}x$ for the particle with the smaller momentum;
- ID_2, function of $\mathrm{d}E/\mathrm{d}x$ for the particle with the larger momentum.

The fast simulation already generates the momenta of the two particles and so α and p_{tot} can be calculated from them. Figures 7.4 and 7.5 show the distributions of α obtained for each of the simulated decay modes. Figures 7.6 and 7.7 show the distributions of p_{tot} for each decay mode. Included in those figures are the projections of the parameterised joint distribution $P(\alpha, p_{\mathrm{tot}} \mid s_j)$, from equation 4.32, used in the final likelihood fit. The fast simulation models this distribution well. Also shown are the projections of the same probability density function but this time with parameters determined from a fit to the fast simulated data. The parameterisation of $P(\alpha, p_{\mathrm{tot}} \mid s_j)$ determined from the fast simulation is used in fits performed on the simulated data in order to remove any effects due to differences between simulation and data from studies of the performance of the method.

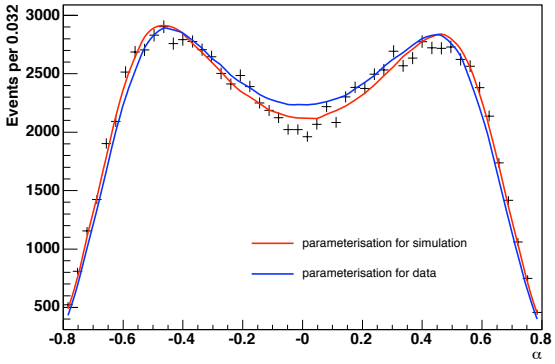
The distributions of α and p_{tot} produced by the fast simulation of background are shown in figure 7.8. Also shown are the projections onto α and p_{tot} of the probability density function $P(\alpha, p_{\mathrm{tot}} \mid b)$, from equation 4.33, used in the final likelihood fit. The fast simulation of background does not approximate this distribution well. It is not desirable to redistribute α and p_{tot} according to $P(\alpha, p_{\mathrm{tot}} \mid b)$ as this would destroy the relationship between these and the lifetime related variables; instead, a different parameterisation of these variables, based on the joint distribution used for signal from equation 4.32, is determined from the simulated data to be used in the studies of



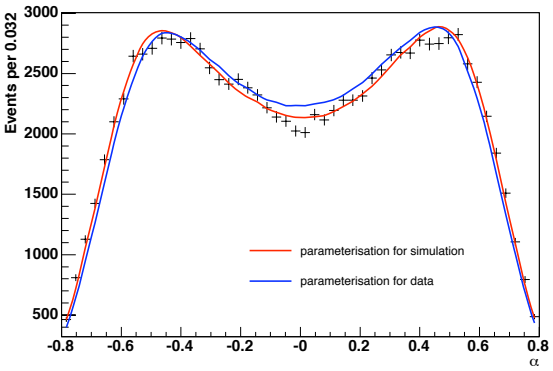
(a) $B_s^0 \to K^+K^-$.



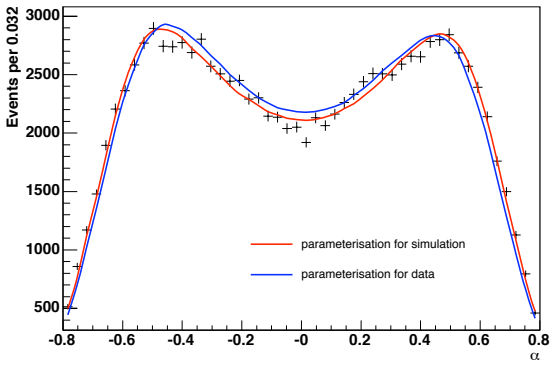
(b) $B^0 \to \pi^+\pi^-$.



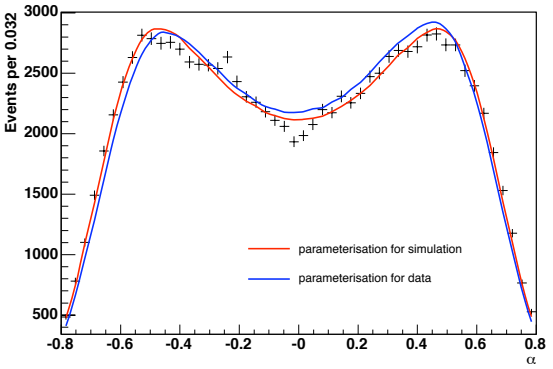
(c) $B^0 \to K^+\pi^-$.



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

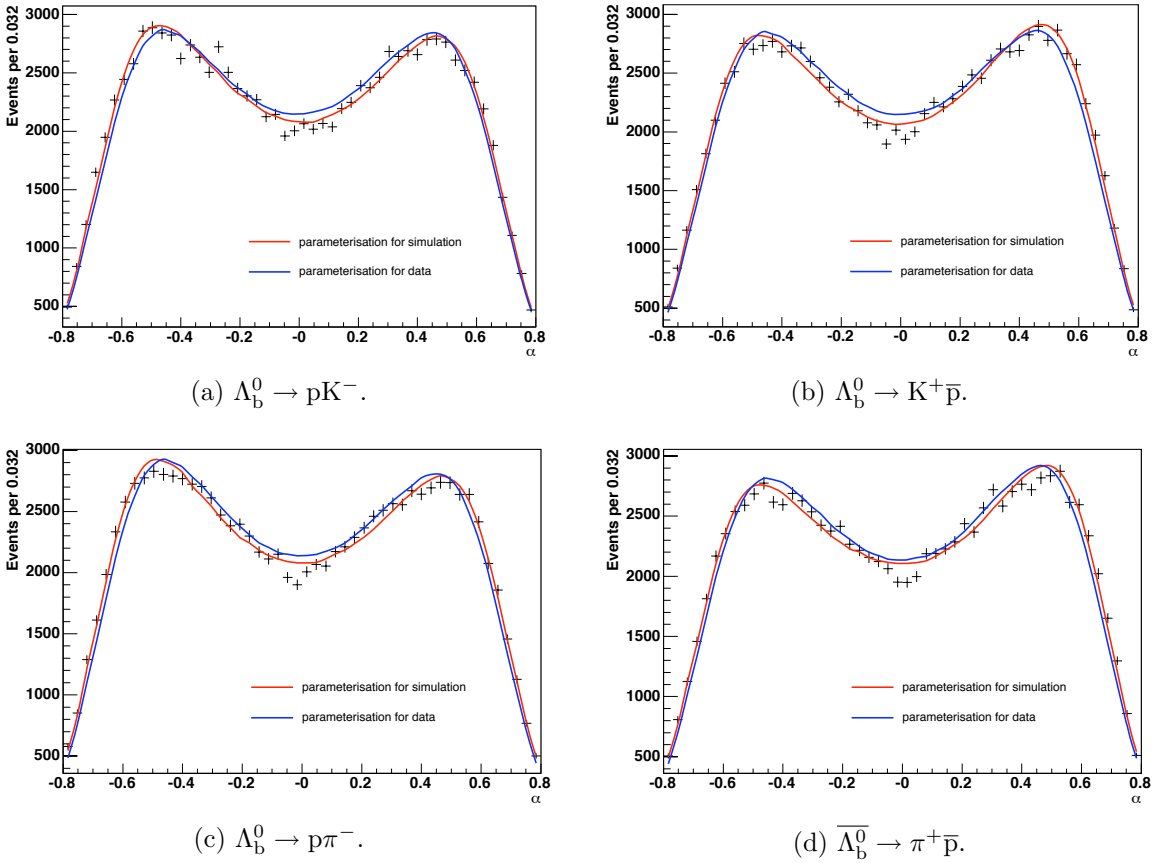


(e) $B_s^0 \to K^+\pi^-$.



(f) $B_s^0 \to \pi^+K^-$.

Figure 7.4: The distributions of α produced by the fast simulation of each of the $B^0_{(s)} \to h^+h'^-$ modes (black points). The blue curves show the projections onto α of the joint distribution of $P(\alpha, p_{\rm tot} \mid s)$ used in the likelihood function describing data. The fast simulation reproduces the distribution expected in data well. The red curves show the projection onto α of the same function reparameterised using the fast simulation data.



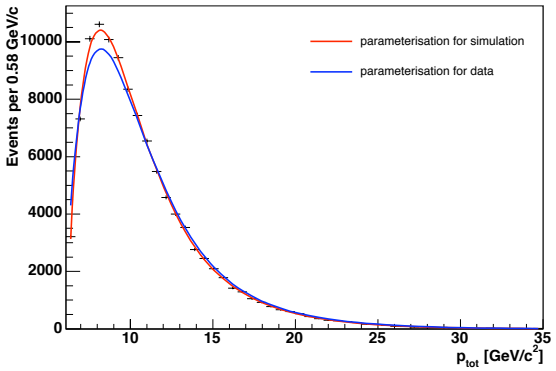
(a) $\Lambda_b^0 \to pK^-$.

(b) $\Lambda_b^0 \to K^+\overline{p}$.

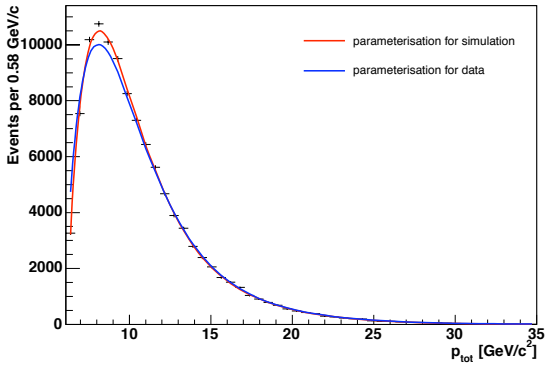
(c) $\Lambda_b^0 \to p\pi^-$.

(d) $\overline{\Lambda_b^0} \to \pi^+\overline{p}$.

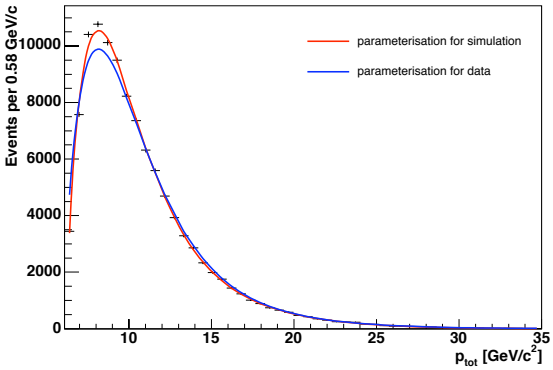
Figure 7.5: The distributions of α produced by the fast simulation of each of the $\Lambda_b^0 \to ph^-$ modes (black points). The blue curves show the projections onto α of the joint distribution of $P(\alpha, p_{\rm tot} \mid s)$ used in the likelihood function describing data. The fast simulation reproduces the distribution expected in data well. The red curves show the projection onto α of the same function reparameterised using the fast simulation data.



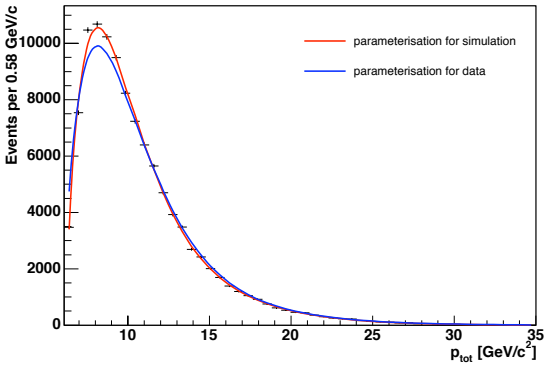
(a) $B_s^0 \to K^+K^-$.



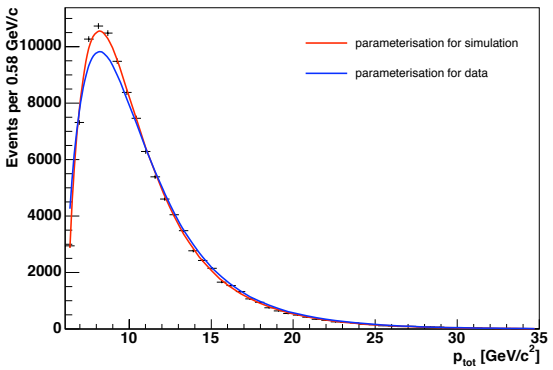
(b) $B^0 \to \pi^+\pi^-$.



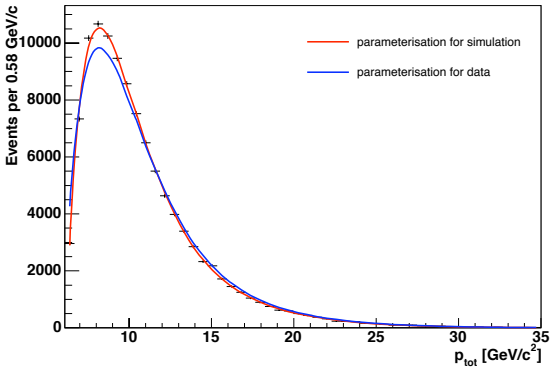
(c) $B^0 \to K^+\pi^-$.



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

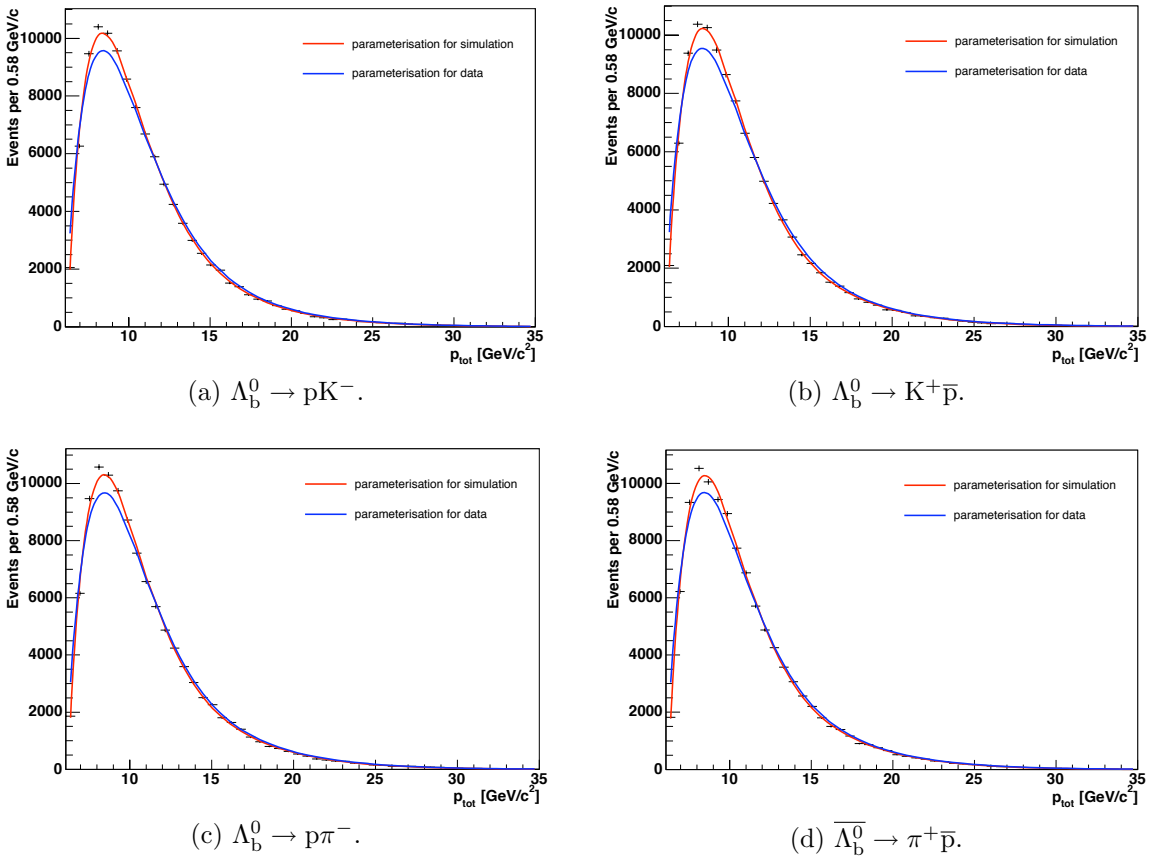


(e) $B_s^0 \to K^+\pi^-$.



(f) $B_s^0 \to \pi^+K^-$.

Figure 7.6: The distributions of $p_{\rm tot}$ produced by the fast simulation of each of the $B^0_{(s)} \to h^+h'^-$ modes (black points). The blue curves show the projections onto $p_{\rm tot}$ of the joint distribution of $P(\alpha, p_{\rm tot} \mid s)$ used in the likelihood function describing data. The fast simulation reproduces the distribution expected in data well. The red curves show the projection onto $p_{\rm tot}$ of the same function reparameterised using the fast simulation data.



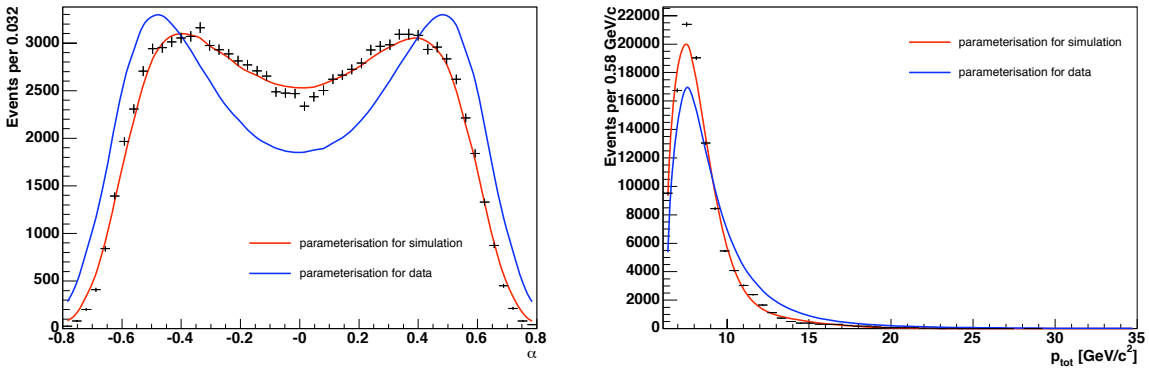
(a) $\Lambda_b^0 \to pK^-$.

(b) $\Lambda_b^0 \to K^+\overline{p}$.

(c) $\Lambda_b^0 \to p\pi^-$.

(d) $\overline{\Lambda_b^0} \to \pi^+\overline{p}$.

Figure 7.7: The distributions of $p_{\rm tot}$ produced by the fast simulation of each of the $\Lambda_b^0 \to ph^-$ modes (black points). The blue curves show the projections onto $p_{\rm tot}$ of the joint distribution of $P(\alpha, p_{\rm tot} \mid s)$ used in the likelihood function describing data. The fast simulation reproduces the distribution expected in data well. The red curves show the projection onto $p_{\rm tot}$ of the same function reparameterised using the fast simulation data.



(a) α. (b) p_{tot}.

Figure 7.8: The distributions of α (left) and p_{tot} (right) produced by the fast simulation of background (black points). The blue curves show the projections onto α and p_{tot} of the joint distribution of $P(\alpha, p_{\mathrm{tot}} \mid b)$ used in the likelihood function describing data. The red curves show the projection onto α and p_{tot} of a parameterisation of the joint distribution of α and p_{tot} fit to the fast simulation as described in the text.

the performance of the method. The projections of the background joint probability density function for α and p_{tot} used for pseudo-experiments are also shown in figure 7.8.

It now remains to simulate the variables $M_{\pi\pi}$, ID_1 and ID_2 used by the separation fit. The probability of observing a particular set of $M_{\pi\pi}$, ID_1 and ID_2 is dependent on α and p_{tot} and the probability density function (PDF) is described by equation 6.4 for signal events and equation 6.5 for background events. These are the PDFs used for the separation fit. Values of $M_{\pi\pi}$, ID_1 and ID_2 are randomly generated according to the PDF for the separation fit using an accept or reject method in the following way. The values of α and p_{tot} produced for an event by the fast simulation are combined with values of $M_{\pi\pi}$, ID_1 and ID_2 randomly chosen from a uniform distribution in the ranges $5.0 < M_{\pi\pi} < 5.8$ GeV/c^2 for mass and $-4.0 < ID_i < 4.0$ for ID_1 and ID_2. The probability of observing this combination of variables is calculated from the separation PDF (equation 6.4 for signal or equation 6.5 for background). A test probability is generated uniformly between zero and a maximum which is slightly larger than the maximum separation probability obtainable. If the test probability is smaller than the probability of observing the particular combination of α, p_{tot}, $M_{\pi\pi}$, ID_1 and ID_2 then

the $M_{\pi\pi}$, ID_1 and ID_2 generated are kept for that event, otherwise they are rejected and new values of $M_{\pi\pi}$, ID_1 and ID_2 are randomly chosen to be combined with the same values of α and $p_{\rm tot}$. The process continues until values for $M_{\pi\pi}$, ID_1 and ID_2 have been accepted for each simulated α and $p_{\rm tot}$.

In this way values of $M_{\pi\pi}$, ID_1 and ID_2 are simulated for all events, consistent with the PDFs describing their distributions and with α and $p_{\rm tot}$. This method of generating $M_{\pi\pi}$, ID_1 and ID_2 leaves them independent of the variables t'_0, $t'_{\rm min}$ and $t'_{\rm max}$.

7.3 Performance of the full fit

The performance of the complete likelihood from section 6.3 is tested by performing $\sim$ 1000 pseudo-experiments on fast-simulated samples. For each pseudo-experiment 13479 events are generated, which is the the number of signal and combinatorial background events observed in the data sample in [41]. The number of each type of $\mathrm{B}^0_{(s)} \to h^+h'^-$, $\Lambda^0_{\rm b} \to {\rm p}h^-$ and combinatorial background events is allowed to vary multi-nomially for each pseudo-experiment. The mean fractions used for each event type are taken from [41]. Figure 7.9 shows the distribution of pulls for the B^0 lifetime returned by 869 pseudo-experiments and a Gaussian curve that has been fitted to the pull distribution. The distribution is Gaussian and has a width of 1.015 ± 0.029, consistent with 1.0, suggesting that the error on the B^0 lifetime is correctly estimated by the fit. The fitted Gaussian has a mean of 0.076 ± 0.036, inconsistent with zero. This indicates an intrinsic bias to the measured B^0 lifetime due to the method used to measure it.

Figure 7.10 shows the distribution of pulls for the $\mathrm{B}^0_{\rm s}$ lifetime returned by the same 869 pseudo-experiments as above and a Gaussian curve that has been fitted to the pull distribution. The distribution is Gaussian and has a width of 0.984 ± 0.028, consistent with 1.0, suggesting that the error on the $\mathrm{B}^0_{\rm s}$ lifetime is correctly estimated by the

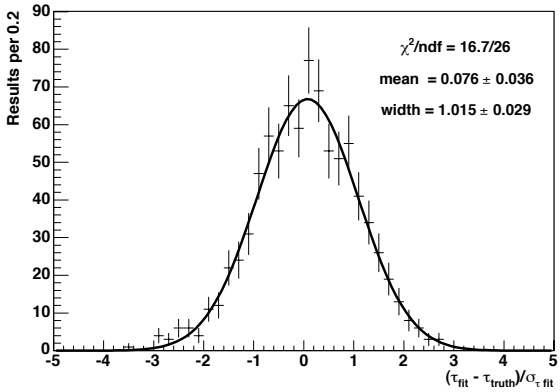


Figure 7.9: The distribution of pulls for fitted B^0 lifetime, $\tau_{\rm fit}$, from 869 pseudo-experiments. The curve is a Gaussian function fit to the points. The mean and width of the Gaussian function and the χ^2 per degree of freedom for the fit are shown

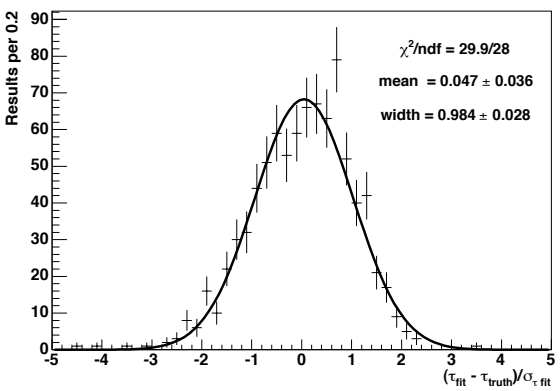


Figure 7.10: The distribution of pulls for fitted B_s^0 lifetime, $\tau_{\rm fit}$, from 869 pseudo-experiments. The curve is a Gaussian function fit to the points. The mean and width of the Gaussian function and the χ^2 per degree of freedom for the fit are shown

fit. The fitted Gaussian has a mean of 0.047 ± 0.036, slightly inconsistent with zero, potentially indicating that the B_s^0 lifetime returned by the fit suffers a bias also.

The biases observed in the measurement of the B^0 and B_s^0 lifetime are intrinsic to the method used to measure them and include the effect of neglecting the term

$$P(t'_{\rm min}, t'_{\rm max} \mid \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2, s_j, \text{trigger}) \tag{7.2}$$

in the likelihood function. Figure 7.11 shows the distribution of the difference between the fitted and input B^0 lifetime for the pseudo-experiments. A Gaussian fit to the distribution gives a width of 14.94 ± 0.40 μm/c and a mean of 1.13 ± 0.52 μm/c. Figure 7.12 shows the distribution of the difference between the fitted and input B_s^0 lifetime for the pseudo-experiments. A Gaussian fit to the distribution gives a width of 32.79 ± 0.85 μm/c and a mean of 4.6 ± 1.2 μm/c. The means of these two fits are taken as the systematic uncertainties on the measured B^0 and B_s^0 lifetimes due to the bias intrinsic to the method. This includes the effect of neglecting the probability of observing $t'_{\rm min}$ and $t'_{\rm max}$ in the likelihood, function 7.2. Table 7.1 summarises the results of the Gaussian fits to the pull and residual distributions.

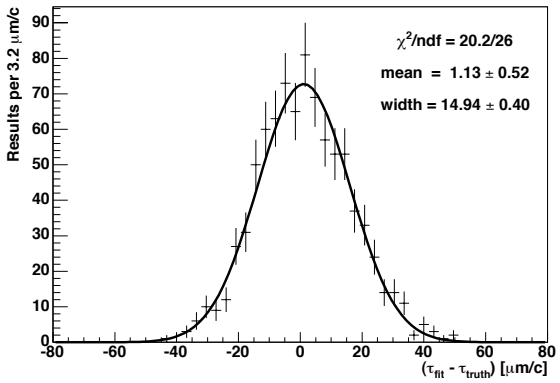


Figure 7.11: Distribution of the difference between the fitted and input B^0 lifetimes ($\tau_{\rm fit} - \tau_{\rm truth}$) for 869 pseudo-experiments. The curve is a Gaussian function fit to the points. The mean and width of the Gaussian function and the χ^2 per degree of freedom for the fit are shown

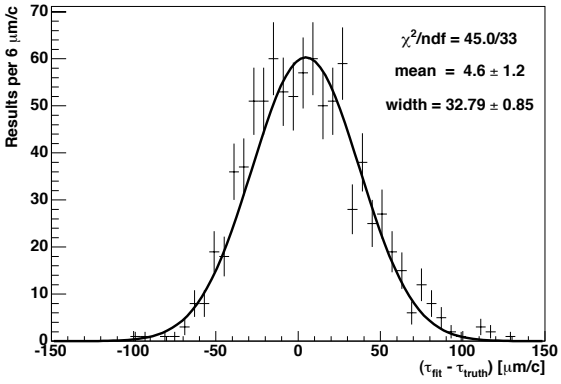


Figure 7.12: Distribution of the difference between the fitted and input B_s^0 lifetimes ($\tau_{\rm fit} - \tau_{\rm truth}$) for 869 pseudo-experiments. The curve is a Gaussian function fit to the points. The mean and width of the Gaussian function and the χ^2 per degree of freedom for the fit are shown

Distribution	Mean	Width
B^0 lifetime pull	0.067 ± 0.036	1.015 ± 0.029
B_s^0 lifetime pull	0.047 ± 0.036	0.984 ± 0.028
B^0 lifetime residual	1.13 ± 0.52 μm/c	14.94 ± 0.40 μm/c
B_s^0 lifetime residual	4.6 ± 1.2 μm/c	32.79 ± 0.85 μm/c

Table 7.1: Parameters of Gaussian fits to the pull and residual distributions shown in figures 7.9 to 7.12 from the study of the performance of the combined likelihood.

7.4 An alternative approach

The likelihood described in section 6.3 included a term that described the distribution of $t'_{\rm min}$ and $t'_{\rm max}$ given the event type and the observables used in the separation:

$$P(t'_{\rm min}, t'_{\rm max} \mid \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2, s_j, \text{trigger}).$$

In the standard fit tested in section 7.3 it has been assumed that this term is the same for all signal modes and background so that the term factors out of the likelihood and can be neglected. This has been shown to be a reasonable assumption between decay modes (see figure 6.1, page 139) but is not true between signal and background (figure 5.14, page 113).

In section 6.2 an alternative parameterisation of the likelihood (parameterisation 2) was determined to try to account for this neglected term. Using this parameterisation the signal part of the likelihood becomes

$$\mathcal{L}^{sg} = P(\text{trigger} \mid \text{all}) P(s \mid t'_{\rm min}, t'_{\rm max}, \text{trigger}) \sum_j P(s_j \mid s, \text{trigger}) \cdot \mathcal{L}_j^{sep\,sg} \cdot \mathcal{L}_j^{life\,sg}. \quad (7.3)$$

The term $P(s_j \mid s, \text{trigger})$ is the fraction of signal that is of decay type j. The term $P(s \mid t'_{\rm min}, t'_{\rm max}, \text{trigger})$ can be approximated to $P(s \mid \kappa, \text{trigger})$ using a Fisher discriminant analysis as described in section 5.10 and then parameterised using the Lagrange interpolating polynomial as described in section 5.13.

The lifetime part of the signal likelihood ($\mathcal{L}_j^{life\,sg}$) is unchanged from the standard likelihood described in section 6.3. The separation part of the signal likelihood using this parameterisation is

$$\begin{aligned} \mathcal{L}_j^{sep\,sg} = & P(\alpha, p_{\rm tot} \mid s_j, t'_{\rm min}, t'_{\rm max}, \text{trigger}) \\ & \times P(M_{\pi\pi} \mid s_j, \alpha, p_{\rm tot}, t'_{\rm min}, t'_{\rm max}, \text{trigger}) \\ & \times P(ID_1, ID_2 \mid s_j, \alpha, p_{\rm tot}, t'_{\rm min}, t'_{\rm max}, \text{trigger}). \end{aligned} \quad (7.4)$$

With this parameterisation all of the separation probability density functions depend on $t'_{\rm min}$ and $t'_{\rm max}$. In order to return to the separation signal likelihood of equation 6.26, section 6.3, it is necessary to assume that α, $p_{\rm tot}$, $M_{\pi\pi}$, ID_1 and ID_2 are independent of $t'_{\rm min}$ and $t'_{\rm max}$. Figure 7.13 shows the relationship between $t'_{\rm min}$ and $t'_{\rm max}$ and the separation variables, $M_{\pi\pi}$ and ID, determined from the fast simulation of $\mathrm{B}^0 \to \pi^+\pi^-$ events. There are no observable correlations in these distributions so $M_{\pi\pi}$, ID_1 and ID_2 are independent of $t'_{\rm min}$ and $t'_{\rm max}$. Figure 7.14 shows the relationship between $t'_{\rm min}$ and $t'_{\rm max}$ and the momentum separation variables, α and $p_{\rm tot}$, determined from the fast simulation of $\mathrm{B}^0 \to \pi^+\pi^-$ events. There is a strong dependence of $t'_{\rm min}$ and $t'_{\rm max}$ on $p_{\rm tot}$ and the minimum values of $t'_{\rm min}$ and $t'_{\rm max}$ are dependent on both α and $p_{\rm tot}$. This dependence arises because the distribution of momentum between the two decay products is dependent on the boost imparted to them by the B meson which is also linked to the impact parameters of the two tracks which are used to determine the lifetime acceptance. The assumption that α and $p_{\rm tot}$ are independent of $t'_{\rm min}$ and $t'_{\rm max}$ is not a good one.

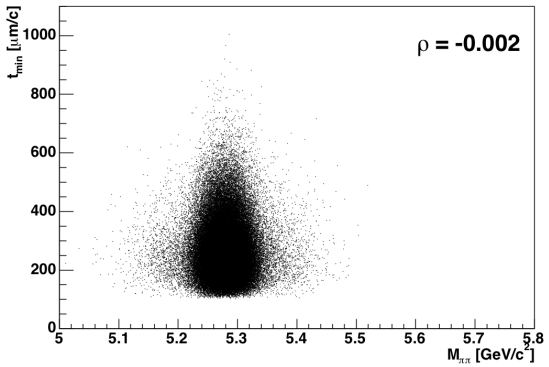
The background part of the likelihood using this parameterisation is

$$\mathcal{L}^{bg} = P(\text{trigger} \mid \text{all}) P(b \mid t'_{\rm min}, t'_{\rm max}, \text{trigger}) \sum_l P(b_l \mid b, \text{trigger}) \cdot \mathcal{L}_l^{sep\,bg} \cdot \mathcal{L}_l^{life\,bg}. \quad (7.5)$$

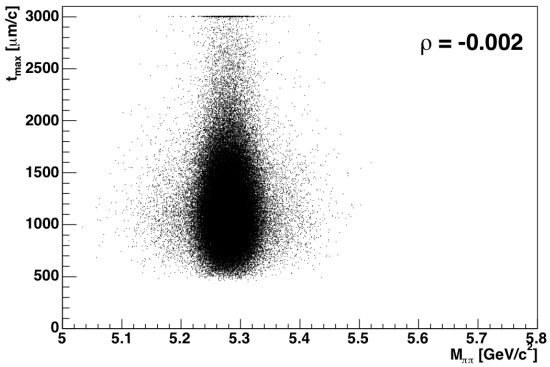
The term $P(b_l \mid b, \text{trigger})$ is the fraction of background that is of type l and the term $P(b \mid t'_{\rm min}, t'_{\rm max}, \text{trigger}) = 1 - P(s \mid t'_{\rm min}, t'_{\rm max}, \text{trigger})$. As for the signal part of the likelihood, the lifetime part of the background likelihood ($\mathcal{L}_l^{life\,bg}$) is unchanged from that in section 6.3 and the separation part of the background likelihood is

$$\begin{aligned} \mathcal{L}_l^{sep\,bg} = {} & P(\alpha, p_{\rm tot} \mid b_l, t'_{\rm min}, t'_{\rm max}, \text{trigger}) \\ & \times P(M_{\pi\pi} \mid b_l, \alpha, p_{\rm tot}, t'_{\rm min}, t'_{\rm max}, \text{trigger}) \\ & \times P(ID_1, ID_2 \mid b_l, \alpha, p_{\rm tot}, t'_{\rm min}, t'_{\rm max}, \text{trigger}). \end{aligned} \quad (7.6)$$

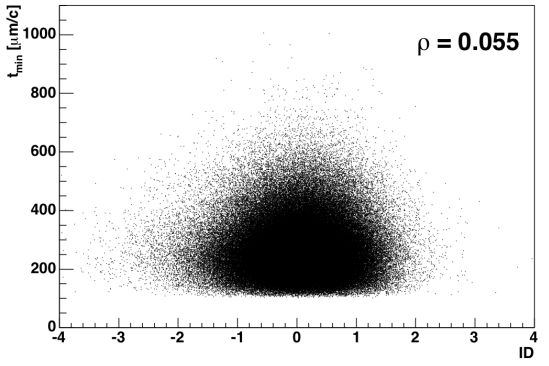
Once again, each of the separation probability density functions are now dependent on



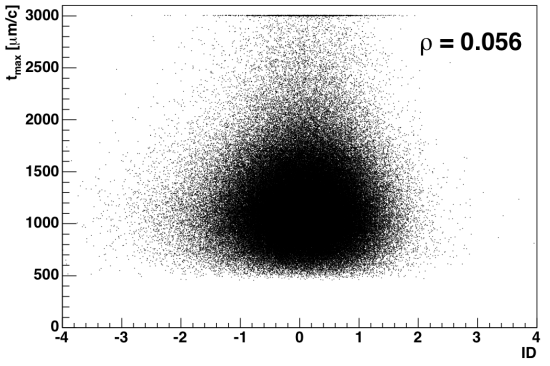
(a) Distribution of $t'_{\rm min}$ and $M_{\pi\pi}$.



(b) Distribution of $t'_{\rm max}$ and $M_{\pi\pi}$.

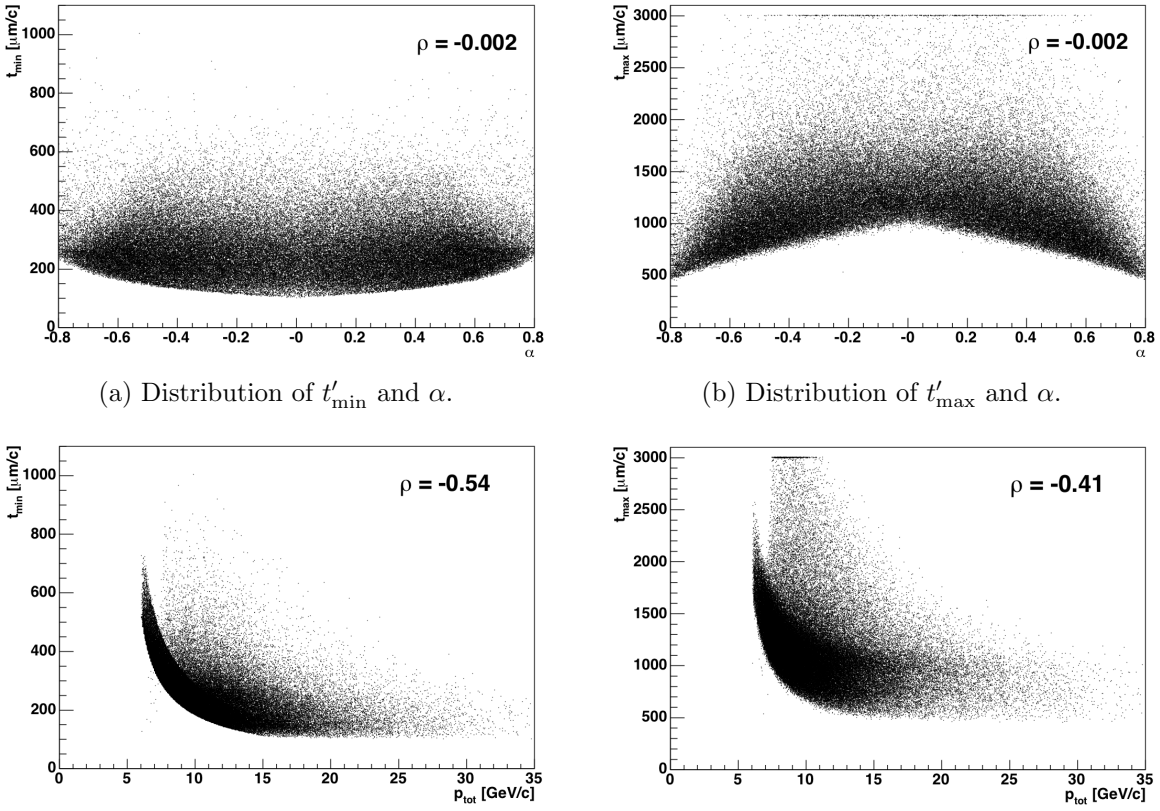


(c) Distribution of $t'_{\rm min}$ and ID.



(d) Distribution of $t'_{\rm max}$ and ID.

Figure 7.13: The relationship between the lifetime limits, $t'_{\rm min}$ and $t'_{\rm max}$, and the separation variables, $M_{\pi\pi}$ and ID, determined from the fast simulation of $\mathrm{B}^0 \to \pi^+\pi^-$. There is no observable correlation. The correlation coefficient, ρ, is given.



(a) Distribution of $t'_{\rm min}$ and α.

(b) Distribution of $t'_{\rm max}$ and α.

(c) Distribution of $t'_{\rm min}$ and $p_{\rm tot}$.

(d) Distribution of $t'_{\rm max}$ and $p_{\rm tot}$.

Figure 7.14: The relationship between the lifetime limits, $t'_{\rm min}$ and $t'_{\rm max}$, and the momentum separation variables, α and $p_{\rm tot}$, determined from the fast simulation of $B^0 \to \pi^+\pi^-$. The correlation coefficient, ρ, is given. There is a strong dependence of $t'_{\rm min}$ and $t'_{\rm max}$ on $p_{\rm tot}$ and the minimum values of $t'_{\rm min}$ and $t'_{\rm max}$ are dependent on both α and $p_{\rm tot}$.

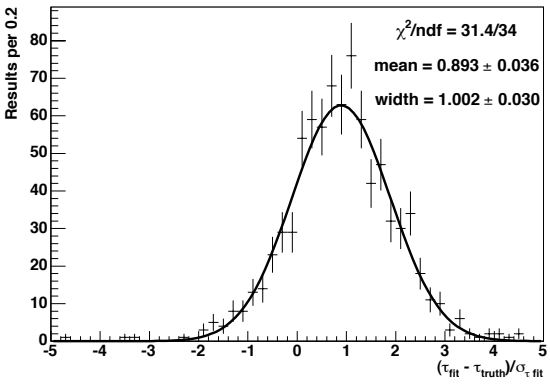


Figure 7.15: The distribution of pulls for fitted B^0 lifetime, $\tau_{\rm fit}$, from 839 pseudo-experiments using the alternative likelihood. The curve is a Gaussian function fit to the points. The mean and width of the Gaussian function and the χ^2 per degree of freedom for the fit are shown

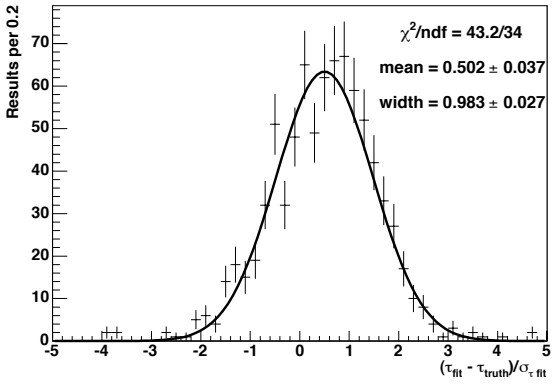


Figure 7.16: The distribution of pulls for fitted B_s^0 lifetime, $\tau_{\rm fit}$, from 839 pseudo-experiments using the alternative likelihood. The curve is a Gaussian function fit to the points. The mean and width of the Gaussian function and the χ^2 per degree of freedom for the fit are shown

$t'_{\rm min}$ and $t'_{\rm max}$ and the same assumptions are needed here as for signal.

This parameterisation attempts to partly deal with the term

$$P(t'_{\rm min}, t'_{\rm max} \mid \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2, s_j, \text{trigger})$$

in the standard likelihood but in so doing adds a dependence on $t'_{\rm min}$ and $t'_{\rm max}$ to the probability of observing α and $p_{\rm tot}$ which is neglected.

To determine which assumptions, associated with each likelihood, produce the least biased result a series of ~ 1000 pseudo-experiments were performed on fast simulated data using this alternative likelihood. Figure 7.15 shows the distribution of pulls for the B^0 lifetime returned by the pseudo-experiments and a Gaussian fit to the distribution. Figure 7.15 shows the distribution of pulls for the B_s^0 lifetime. Figures 7.17 and 7.18 show the distributions of the residual of the lifetime fit $(\tau_{\rm fit} - \tau_{\rm truth})$ for the B^0 and B_s^0 lifetimes respectively. Table 7.2 shows the parameters obtained from the Gaussian fits to each of the pull and residual distributions.

The biases on the measurements on the B^0 and B_s^0 lifetime using this likelihood are

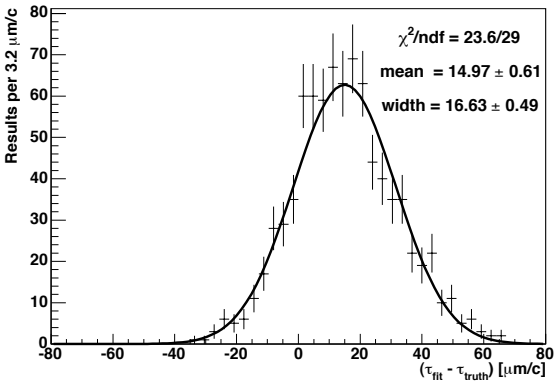


Figure 7.17: Distribution of the difference between the fitted and input B^0 lifetimes ($\tau_{\rm fit} - \tau_{\rm truth}$) for 839 pseudo-experiments using the alternative likelihood. The curve is a Gaussian function fit to the points. The mean and width of the Gaussian function and the χ^2 per degree of freedom for the fit are shown

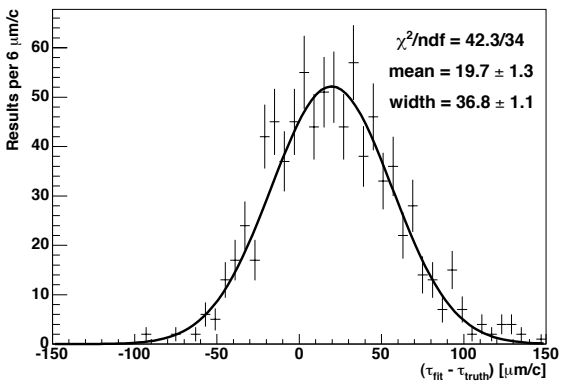


Figure 7.18: Distribution of the difference between the fitted and input B_s^0 lifetimes ($\tau_{\rm fit} - \tau_{\rm truth}$) for 839 pseudo-experiments using the alternative likelihood. The curve is a Gaussian function fit to the points. The mean and width of the Gaussian function and the χ^2 per degree of freedom for the fit are shown

Distribution	Mean	Width
B^0 lifetime pull	0.893 ± 0.036	1.002 ± 0.030
B_s^0 lifetime pull	0.502 ± 0.037	0.983 ± 0.027
B^0 lifetime residual	14.97 ± 0.61 μm/c	16.63 ± 0.49 μm/c
B_s^0 lifetime residual	19.7 ± 1.3 μm/c	36.8 ± 1.1 μm/c

Table 7.2: Parameters of Gaussian fits to the pull and residual distributions shown in figures 7.15 to 7.18 from the study of the performance of the alternative likelihood.

14.97 μm/c and 19.7 μm/c respectively. This compares to biases of 1.13 μm/c on the B^0 lifetime and 4.6 μm/c on the B_s^0 lifetime measured using the standard lifetime in section 7.3.

To determine whether it is the assumption that α, $p_{\rm tot}$, $M_{\pi\pi}$, ID_1 and ID_2 are independent of $t'_{\rm min}$ and $t'_{\rm max}$ that has caused the large biases or if it is the process of approximating $P(s \mid t'_{\rm min}, t'_{\rm max}, {\rm trigger})$ to $P(s \mid \kappa, {\rm trigger})$ using the Fisher linear discriminant analysis, it is necessary to remove the separation part of the likelihood. This can be done for a sample containing only background and $B^0 \to \pi^+\pi^-$ events. In this sample the signal and background can be separated using a simple mass fit (exponential distribution for background and a sum of two Gaussians for signal). $B^0 \to \pi^+\pi^-$ is chosen because in this mode the lifetime probabilities do not depend on α and $p_{\rm tot}$ (as this mode is correctly reconstructed) and so α, $p_{\rm tot}$, ID_1 and ID_2 can be completely removed from the likelihood. The mass fit can be used to define the signal and background regions used to determine the Fisher direction in the same way as the separation fit was used.

Four studies, each of 1000 pseudo-experiments, are performed using $B^0 \to \pi^+\pi^-$ plus background samples. For each pseudo-experiment an average of 6494 background and 6000 $B^0 \to \pi^+\pi^-$ events are generated using the fast simulation with the fractions varied binomially for each sample. The four studies each use a different likelihood:

1. the standard likelihood, including the separation fit;
2. the alternative likelihood using the term $P(s \mid \kappa, {\rm trigger})$ and the separation fit;
3. the standard fit, but with the separation fit replaced by the simple mass fit;
4. the alternative fit using the term $P(s \mid \kappa, {\rm trigger})$ and the mass fit.

For each set of pseudo-experiments the distribution of residuals $(\tau_{\rm fit} - \tau_{\rm truth})$ is fitted with a Gaussian distribution. The mean and width of each of these fits is summarised in table 7.3. The standard likelihood with the separation fit has a bias of 1.39 μm/c.

Likelihood	Mean (μm/c)	Width (μm/c)
1. Standard with separation	1.39 ± 0.40	11.97 ± 0.29
2. Alternative with separation	8.34 ± 0.43	12.72 ± 0.33
3. Standard with mass	−2.75 ± 0.39	11.89 ± 0.30
4. Alternative with mass	−2.40 ± 0.41	12.07 ± 0.33

Table 7.3: Parameters of Gaussian fits to the residual distributions from fits using four likelihood configurations for fast simulated $B^0 \to \pi^+\pi^-$ plus background data.

Using the alternative likelihood with separation fit increases the bias to 8.34 μm/c. This is consistent with the results for all signal modes plus background where an increase in the bias was seen when changing from the standard to the alternative likelihood. The standard likelihood with the separation fit replaced by a mass fit shows a bias of -2.75 μm/c. Using the alternative likelihood with a mass fit slightly reduces this bias to -2.40 μm/c.

The two studies using the mass fit in place of the separation fit demonstrate that, on its own, it is not the process of approximating $P(s \mid t'_{\rm min}, t'_{\rm max}, \text{trigger})$ to $P(s \mid \kappa, \text{trigger})$ using the Fisher linear discriminant analysis that is significantly increasing the bias when using the alternative likelihood. It is only when this likelihood uses the separation fit that a large bias is observed. This suggests that it is the assumption that α, $p_{\rm tot}$, $M_{\pi\pi}$, ID_1 and ID_2 are independent of $t'_{\rm min}$ and $t'_{\rm max}$ that has caused the large bias. It can also be seen from the studies using the mass fit that the alternative likelihood is only making a slight improvement to a bias that is, in any case, small.

The alternative likelihood described above was tested to see if it reduced the bias due to assuming that the term $P(t'_{\rm min}, t'_{\rm max} \mid \alpha, p_{\rm tot}, M_{\pi\pi}, ID_1, ID_2, s_j, \text{trigger})$ was the same for signal and background by partially accounting for the distribution of $t'_{\rm min}$ and $t'_{\rm max}$. In fact the assumptions that had to be made to partly model this term led to a far larger bias so this likelihood is not used for the final measurement.

Distribution	Mean	Width
B^0 lifetime pull	-0.122 ± 0.033	0.998 ± 0.025
B_s^0 lifetime pull	-0.005 ± 0.035	1.032 ± 0.026
B^0 lifetime residual	-0.84 ± 0.44 μm/c	13.28 ± 0.32 μm/c
B_s^0 lifetime residual	2.5 ± 1.1 μm/c	31.29 ± 0.77 μm/c

Table 7.4: Parameters of Gaussian fits to the pull and residual distributions from the study of the performance of the standard likelihood on signal only.

7.5 Performance of the full fit on signal only

Some of the tests of the systematic biases present, discussed below, are performed on signal only. It is useful to know how the standard method performs on signal only to use as a base line. 1000 pseudo-experiments were performed using the signal part of the likelihood on fast simulation signal only samples. Table 7.4 summarises the widths and means of Gaussian fits to the B^0 and B_s^0 lifetime pull and residual distributions. Biases of -0.844 μm/c on the B^0 lifetime and 2.5 μm/c on the B_s^0 lifetime are observed. These biases are intrinsic to the method used and include the effect of neglecting the distribution of $t'_{\rm min}$ and $t'_{\rm max}$. The bias in the signal only fit is already included in the bias observed in the full fit with signal and background in section 7.3. That the bias is smaller in signal only than in the full fit is expected because there was no observed difference in the distributions of $t'_{\rm min}$ and $t'_{\rm max}$ between different signal decay modes so the term describing the distribution of $t'_{\rm min}$ and $t'_{\rm max}$ in the likelihood is a constant factor that can safely be neglected. A difference has been observed between signal and background so neglecting the distribution of $t'_{\rm min}$ and $t'_{\rm max}$ leads to a larger bias when background is included.

7.6 SVT efficiency as a function of impact parameter

This analysis assumes that the particle track finding efficiency of the silicon vertex trigger (SVT) is uniform for all events in the trigger and event selection acceptance. The method will be insensitive to any SVT efficiency dependence on observables that are independent of lifetime; however, it may be very sensitive to an SVT efficiency dependent on variables that are directly or indirectly dependent on lifetime. It is not possible to observe the SVT track finding efficiency using the $B^0_{(s)} \to h^+h'^-$ sample collected by the B_PIPI trigger as both tracks in the event must have been found by the SVT to appear in the sample.

Figure 7.19 shows the SVT efficiency observed in realistic simulation. The selection requirements used by this analysis ensure the data sample only contains events with track impact parameters less than 1000 μm. The simulated data show a drop in efficiency before 1000μm. This drop in efficiency for large impact parameters is expected in data but the exact shape and height of the efficiency distribution is very sensitive to the period of data taking and the realistic simulation does not reproduce the distribution exactly. The efficiency has improved over time and become flatter out to larger impact parameters. The realistic simulated data in figure 7.19 represents early period data and the expected drop in efficiency in real data is not as large.

The SVT efficiency has been measured using the three body decay $B^+ \to D^0\pi^+$ [63] in a sample that was also collected by a two track hadronic trigger using the SVT. If two particular particles, for example the decay products of the D^0, are required to meet the trigger requirements then the third particle will be independent of the trigger and can be used to determine the track finding efficiency of the SVT. The track finding efficiency is determined as the fraction of particle tracks found by the offline reconstruction that can be matched to an SVT track. The SVT efficiency as a

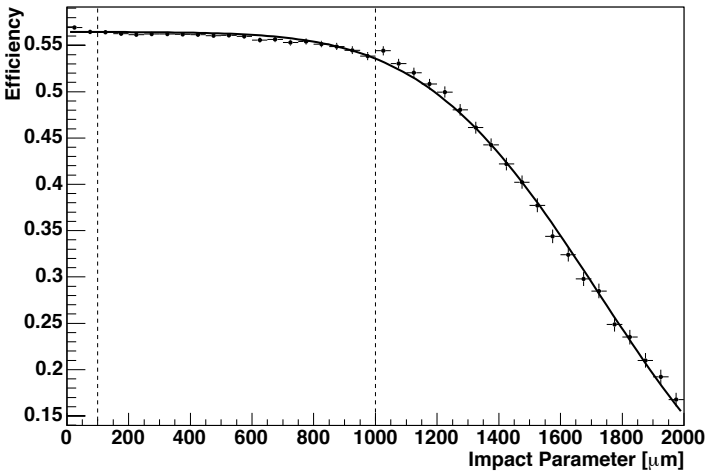


Figure 7.19: The track finding efficiency of the SVT as a function of track impact parameter in realistic simulation. The dashed lines show the minimum and maximum impact parameters accepted by the trigger requirements.

Distribution	Mean	Width
B^0 lifetime pull	-0.331 ± 0.034	1.006 ± 0.026
B_s^0 lifetime pull	-0.227 ± 0.035	0.966 ± 0.025
B^0 lifetime residual	-4.09 ± 0.43 μm/c	13.33 ± 0.31 μm/c
B_s^0 lifetime residual	-5.00 ± 0.96 μm/c	29.42 ± 0.69 μm/c

Table 7.5: Parameters of Gaussian fits to the pull and residual distributions from the study of the effect of non-flat SVT efficiency with respect to impact parameter.

function of impact parameter can be described by the function

$$\epsilon_{\mathrm{SVT}} = c_o \,\mathrm{erfc}\left(\frac{d_0 - c_1}{c_2}\right), \tag{7.7}$$

where erfc is a complimentary error function (defined as 1 minus the error function), d_0 is the track impact parameter and c_0, c_1 and c_2 are parameters determined by a fit to the data.

To test the effect of a varying SVT track finding efficiency with respect to impact parameter 1000 pseudo-experiments were performed on signal only fast simulation. The effect of a non-flat SVT efficiency was modelled in the fast simulation by rejecting particles according to the measured SVT efficiency while the fit was performed assuming the SVT efficiency was still flat. Table 7.5 summarises the results of Gaussian fits to the B^0 and B_s^0 lifetime pull and residual distributions

The pseudo-experiments show biases on the measured B^0 and B_s^0 lifetimes of -4.09 μm/c and 5.00 μm/c respectively. To avoid double counting effects already considered in the systematic uncertainty intrinsic to the method, discussed in section 7.3, the bias observed in these pseudo-experiments is compared to the base-line study performed on signal only using the standard fit (section 7.5) and the difference is taken as the systematic uncertainty. The systematic uncertainty due to assuming the SVT efficiency as a function of impact parameter is flat for $0 < d_0 < 1000$ μm is 3.25 μm/c for the B^0 lifetime and 7.5 μm/c for the B_s^0 lifetime. The difference between the two systematic uncertainties arises because both lifetimes are measured together in the same sample, dominated by B^0 decays. As a check, the study described above was repeated on samples containing only $B_s^0 \to K^+K^-$ or only $B^0 \to \pi^+\pi^-$ decays. The induced biases were -4.38 ± 0.33 μm/c and -4.34 ± 0.36 μm/c respectively, consistent with each other.

7.7 SVT efficiency as a function of transverse momentum

The SVT efficiency as a function of transverse momentum (p_T) in realistic simulation is shown in figure 7.20. The efficiency is constant for most p_T but drops at low p_T. The SVT efficiency as a function of p_T for data has been determined using the three body decay $B^+ \to D^0\pi^+$ in the same way as for the SVT efficiency as a function of impact parameter [63]. The distribution is modelled using a third order polynomial for $0 < p_T \leq 14$ GeV/c and is constant for $p_T > 14$ GeV/c. The effect of a non-flat SVT track finding efficiency as a function of p_T is determined from 1000 pseudo-experiments performed on fast simulated data modified to replicate the shape of the efficiency observed in data. Table 7.6 summarises the means and widths of Gaussian fits to the distributions of pulls and residuals of the B^0 and B_s^0 lifetimes. The residuals indicate biases of 0.05 μm/c for the B^0 and -1.5 μm/c for the B_s^0 lifetime. Comparing

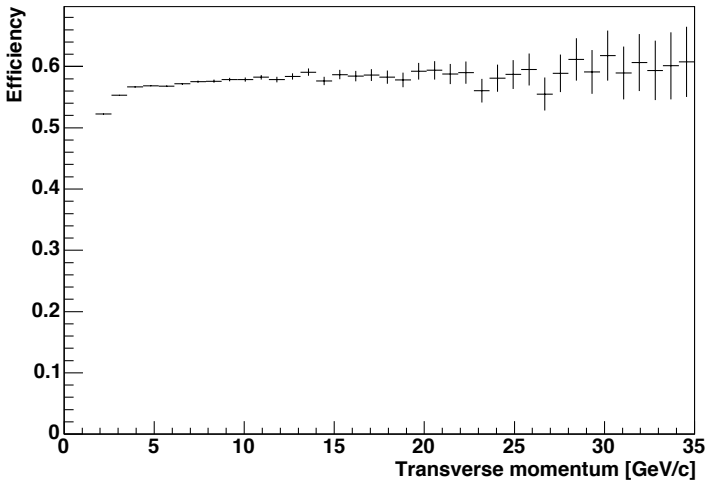


Figure 7.20: The track finding efficiency of the SVT as a function of transverse momentum in realistic simulation.

Distribution	Mean	Width
B^0 lifetime pull	-0.033 ± 0.032	0.977 ± 0.024
B_s^0 lifetime pull	-0.086 ± 0.032	0.969 ± 0.026
B^0 lifetime residual	0.05 ± 0.47 μm/c	14.24 ± 0.37 μm/c
B_s^0 lifetime residual	-1.5 ± 1.1 μm/c	31.42 ± 0.85 μm/c

Table 7.6: Parameters of Gaussian fits to the pull and residual distributions from the study of the effect of non-flat SVT efficiency with respect to transverse momentum.

this to the standard fit on signal only gives a systematic uncertainty due to variations in the SVT track finding efficiency with respect to p_T of 0.89 μm/c for the B^0 lifetime and 4.0 μm/c for B_s^0 lifetime.

7.8 Background mass-lifetime correlation

It has been assumed (see section 6.3) that the measured lifetime for background events is independent of the reconstructed mass. Figure 7.21 shows the mean measured lifetime as a function of reconstructed mass for data events in the mass region $5.6 < M_{\pi\pi} < 6.2$ GeV/c^2. This mass region was chosen to be as broad as possible without containing any $B^0_{(s)} \to h^+h'^-$ or $\Lambda_b^0 \to ph^-$ signal events and so extends beyond the mass range used in the final lifetime fit. Also shown in the figure is a straight line fit to the distribution. The fitted line has a slope of -19 ± 25 μmc/GeV consistent

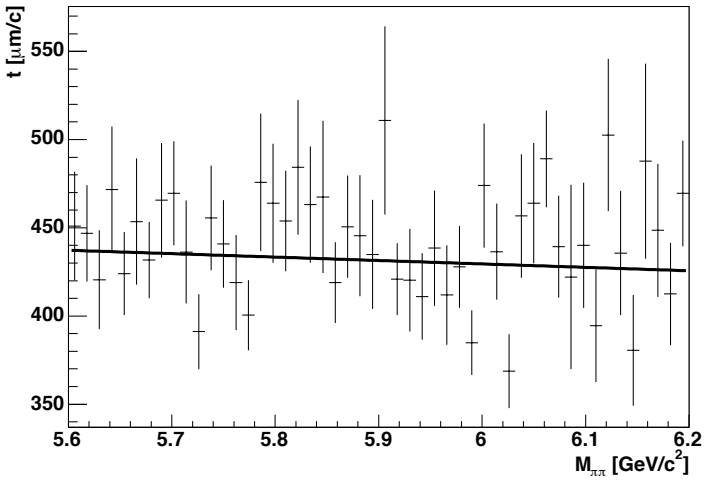


Figure 7.21: Mean measured lifetime t as a function of invariant $\pi\pi$-mass $M_{\pi\pi}$. This mass range contains only combinatorial background events. The line is a straight line fit to the distribution and has a slope -19 ± 25 μmc/GeV consistent with zero

with a gradient of zero and there appears to be no change in the gradient over this mass range. Assuming that the background in this mass region is representative of the background in the mass region $5.16 < M_{\pi\pi} < 5.8$ GeV/c^2 spanned by the lifetime fit, then there is no observable correlation between reconstructed mass and measured lifetime in background events.

7.9 Background momentum-lifetime correlation

In section 6.3 it was assumed that the measured lifetime for background events is independent of the momentum variables α and $p_{\rm tot}$. Figure 7.22 shows the mean measured lifetime as a function of α for background taken from the upper sideband in data ($5.6 < M_{\pi\pi} < 5.8$ GeV/c^2). Also shown is the same distribution in fast simulation background. In data, the distribution is flat for $|\alpha| < 0.5$ with the suggestion that the mean lifetime drops for $|\alpha| > 0.5$. The shape of the distribution is the same in the fast simulation though the average lifetime is lower overall. Figure 7.23 shows the mean measured lifetime as a function of $p_{\rm tot}$ for background taken from the upper sideband in data and the same distribution in fast simulation background. In this case the mean lifetime is strongly dependent on $p_{\rm tot}$, the same dependence is also seen in the fast

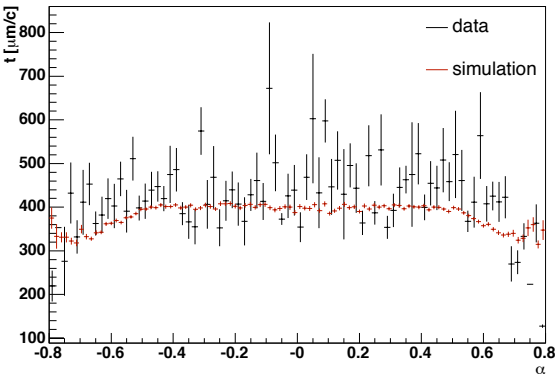


Figure 7.22: Mean measured lifetime t as a function of charge weighted momentum imbalance, α. The black points are for data taken from the upper mass sideband ($5.6 < M_{\pi\pi} < 5.8$ GeV/c^2). The distribution is flat for $|\alpha| < 0.6$. The red points are fast simulated combinatorial background. The distributions in data and fast simulation have the same trend.

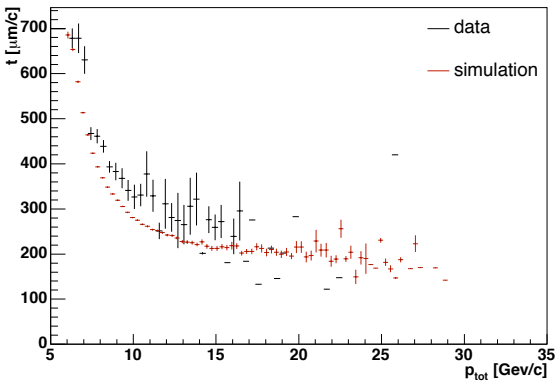


Figure 7.23: Mean measured lifetime t as a function of $p_{\rm tot}$. The black points are for data taken from the upper mass sideband ($5.6 < M_{\pi\pi} < 5.8$ GeV/c^2). The red points are fast simulated combinatorial background. There is a strong dependence on $p_{\rm tot}$ for lifetime. The fast simulation replicates the trend in data.

simulation though, again, the mean lifetime is lower overall.

The assumption that the lifetime of background events is independent of α is reasonable for $|\alpha| < 0.5$, which includes most of the data, though is not good for large $|\alpha|$. The measured lifetime for background events is not independent of $p_{\rm tot}$, so the assumption that it is independent leads to a systematic uncertainty. The fast simulation has the same dependence of lifetime on both α and $p_{\rm tot}$, so the effect of this assumption is already included in the bias observed in the full fit in section 7.3.

7.10 Background momentum model

Section 4.7 describes the probability density function (PDF) for the joint distribution of α and $p_{\rm tot}$ for signal (equation 4.32) and background (equation 4.33) events. Following the approach in [41] it was assumed that the joint distribution of α and $p_{\rm tot}$ is the same for combinatorial and physics type background and an average distribution was used for both background types. Figure 7.24 compares the distributions of α and $p_{\rm tot}$

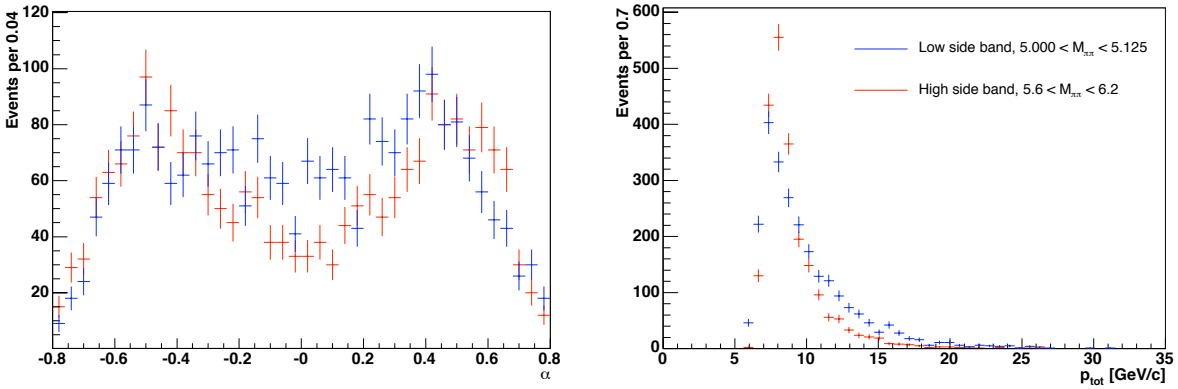


Figure 7.24: Comparison of the distributions of α (left) and p_{tot} (right) in the high and low mass sidebands.

for the low mass sideband ($5 < M_{\pi\pi} < 5.125$ GeV/c^2) and the high mass sideband ($5.6 < M_{\pi\pi} < 6.2$ GeV/c^2). The mass regions used are shown in figure 4.8, page 84. The high sideband contains only combinatorial background events. The low sideband contains an unknown mixture of combinatorial and physics-type background. The two mass sidebands have different distributions of both α and p_{tot}.

To determine the sensitivity of the fit to the parameterisation of $P(\alpha, p_{\mathrm{tot}} \mid b)$ for background alternative parameterisations were determined from fits to data in three alternative mass regions:

- events in the high mass sideband only ($5.6 < M_{\pi\pi} < 6.2$ GeV/c^2), only combinatorial events;
- events from both the high ($5.0 < M_{\pi\pi} < 5.125$ GeV/c^2) and low ($5.6 < M_{\pi\pi} < 6.2$ GeV/c^2) mass sidebands, a mixture of combinatorial and physics type background in approximately the same proportions as in the mass range ($5 < M_{\pi\pi} < 5.8$ GeV/c^2) used in the standard fit;
- events in the low mass sideband only ($5 < M_{\pi\pi} < 5.125$ GeV/c^2), a mixture of both types of background with a higher proportion of physics-type background than in the mass range used in the standard fit.

Figure 7.25 shows the projection onto α and p_{tot} of the joint distribution of $P(\alpha, p_{\mathrm{tot}} \mid b)$

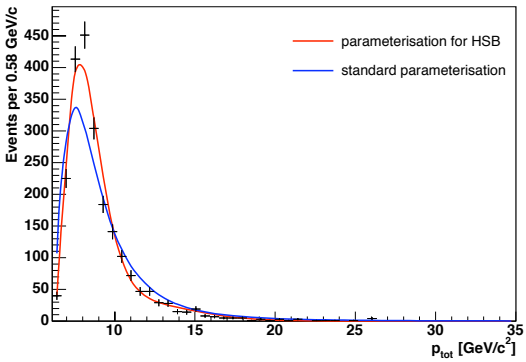
Parameterisation	Shift from standard fit	
	τ_{B^0} (μm/c)	$\tau_{B_s^0}$ (μm/c)
Both	-16	-10
HSB	-39	-24
LSB	$+\ 5$	$+\ 2$

Table 7.7: For fits using $P(\alpha, p_{\rm tot} \mid b)$ determined using the high mass sideband (HSB), low mass sideband (LSB) or both sidebands (both) the shift from the standard fit of the fitted lifetimes is summarised.

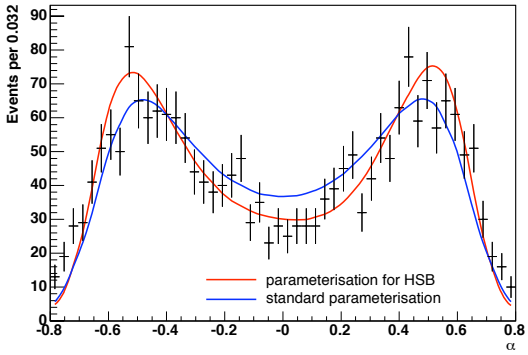
for each of the three samples. The blue curve is the projection of the joint PDF used in the standard fit (equation 4.33, page 83). The red curves are the projections of fits of the joint PDF on each of the three samples. In these fits the PDF used is the same as that for signal (equation 4.32, page 82) and the α distribution was allowed to be asymmetric. Figures 7.25c and 7.25d show the projections of the fitted distribution for both sidebands. The standard and new parameterisations of $P(\alpha, p_{\rm tot} \mid b)$ are very close so the signal PDF provides a satisfactory model for background also.

A full lifetime plus separation fit was performed on the data using each of the three new parameterisations of $P(\alpha, p_{\rm tot} \mid b)$ and the shift in the fitted lifetimes from the standard fit determined. The shift in the fitted lifetimes for each version is summarised in table 7.7. It can be seen that the lifetime is very sensitive to the choice of background parameterisation.

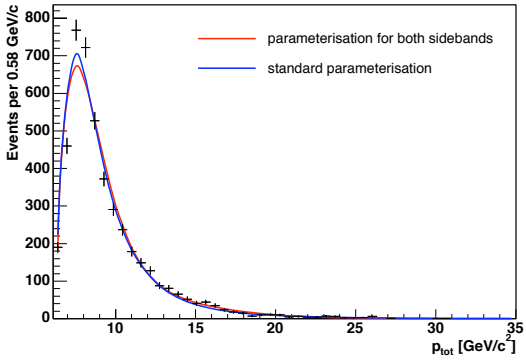
The background underneath the signal peak is predominantly combinatorial background and so is best described using $P(\alpha, p_{\rm tot} \mid b)$ determined from the high mass sideband. This PDF does not describe the physics-type background well, however, and so the background fraction in the low mass sideband will be pulled affecting the total background fraction in the fit. The PDF $P(\alpha, p_{\rm tot} \mid b)$ determined using either the low mass sideband or both sidebands will provide a better (though not complete) description of the physics-type background but will not describe the combinatorial background under the peak well and so, again, the background fraction in this region will be pulled by the incorrect model of $P(\alpha, p_{\rm tot} \mid b)$. The distributions of lifetimes in signal and background are very different from each other and so the lifetime fit is



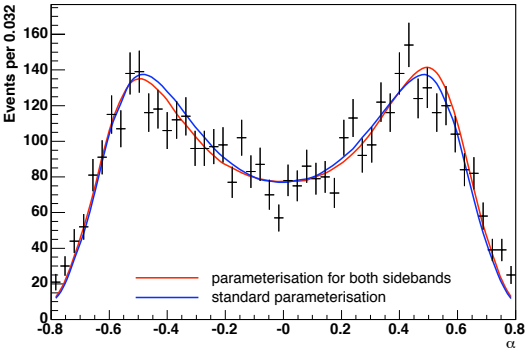
(a) p_{tot} in the high mass sideband (HSB).



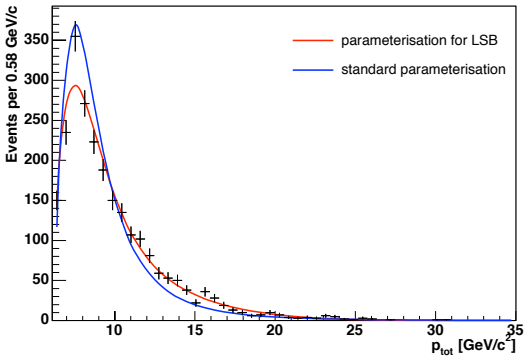
(b) α in the high mass sideband (HSB).



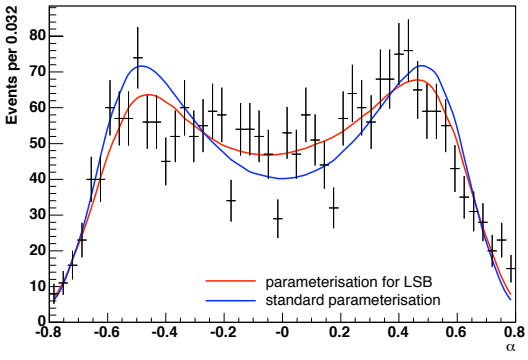
(c) p_{tot} in both mass sidebands.



(d) α in both mass sidebands.



(e) p_{tot} in the low mass sideband (LSB).



(f) α in the low mass sideband (LSB).

Figure 7.25: Projection onto α and p_{tot} of the joint distribution $P(\alpha, p_{\text{tot}} \mid b)$ for low mass sideband (LSB), high mass sideband (HSB) and both sidebands. The blue curves are projections of the standard PDF. The red curves are projections of the parameterised PDFs determined from a fit to the data shown.

very sensitive to background misidentified as signal (and vice-versa) and hence the background fraction.

In order to retain the low mass sideband containing physics type background and reduce the systematic uncertainty due to the model of the joint distribution $P(\alpha, p_{\rm tot} \mid b)$ it is necessary to model the distribution separately for each type of background. One approach would be to try to determine the distribution for physics background using the data itself from the low mass sideband with the distribution due to combinatorial background removed from the total; however the fraction of physics background in that mass region is unknown which could lead to a significant systematic remaining. A second approach could use simulated data to determine $P(\alpha, p_{\rm tot} \mid b)$ for physics-type background in the same way as is done for the signal modes (see section 4.7). As discussed in section 4.5.2, page 77, there are many contributions to this background and the branching ratios are not well known for all of them, also the detector and trigger efficiencies and the momentum distributions would have to be very well modelled in the simulation to in order to reproduce the correct joint distribution of α and $p_{\rm tot}$.

A much simpler approach, adopted here, is to exclude the low mass sideband from the fit entirely; it is already excluded from the lifetime fit. Without a better understanding of the physics-type background this region does not provide additional useful information to the separation fit. With a suitable minimum mass requirement the background is exclusively combinatorial and it is appropriate to use the distribution, $P(\alpha, p_{\rm tot} \mid b)$, determined from the high mass sideband to model all of the background. The physics type background is suppressed by kinematics above $M_{\pi\pi} \approx 5.15$ GeV/c^2 and so restricting the mass region used in the final fit to $M_{\pi\pi} > 5.16$ GeV/c^2 leaves minimal contamination from physics-type background in the sample. This requirement does remove some $\mathrm{B}^0 \to \mathrm{K}^+\pi^-$ and $\overline{\mathrm{B}^0} \to \pi^+\mathrm{K}^-$ events but is below the $\mathrm{B}^0_{\rm s} \to \mathrm{K}^+\mathrm{K}^-$ contribution.

Figure 7.26 shows the mass projection of a separation only fit to the data with the

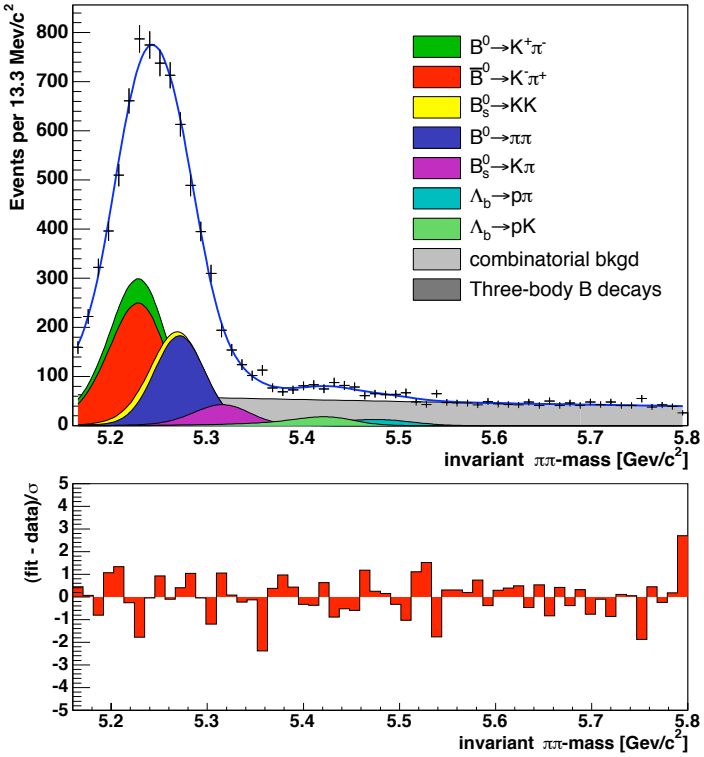


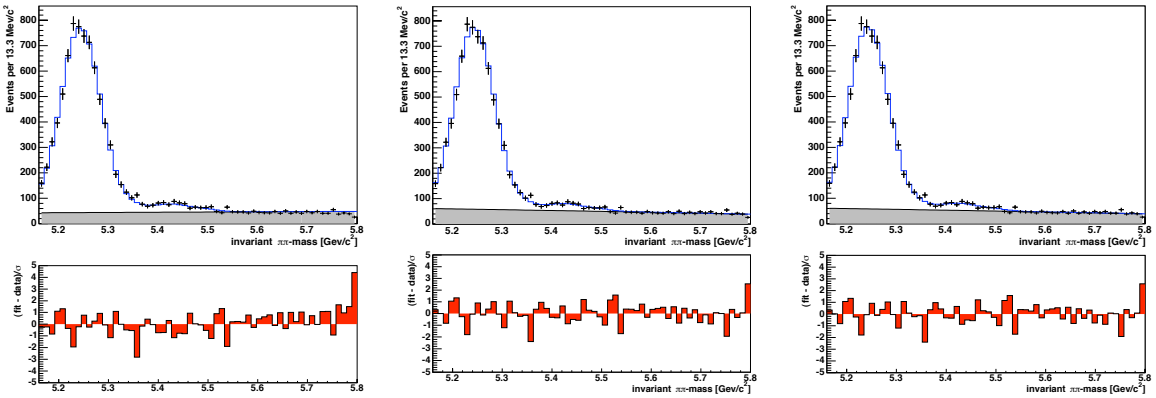
Figure 7.26: $\pi\pi$-invariant mass projection of the separation only fit applied to data with the mass region $M_{\pi\pi} < 5.16$ GeV/c^2 excluded.

low mass sideband excluded from the sample.

7.11 Background mass model

An exponential function was chosen empirically to model the $M_{\pi\pi}$ distribution for combinatorial background. To test the sensitivity of the fit to the choice of background mass shape two alternative parameterisations were tried in the full fit. The alternative distributions were chosen because they produced a reasonable residual in the mass projection of a separation only fit (figures 7.27b and 7.27c). A first order polynomial was also tried but, as can be seen from figure 7.27a, the residual distribution shows that the fit consistently overestimates the population in the high mass region ($M_{\pi\pi} > 5.6$ GeV/c^2) and so a first order polynomial is discounted.

Table 7.8 summarises the shift from the standard fit on data of the B^0 and B_s^0 fitted



(a) First order polynomial. (b) Second order polynomial. (c) Third order polynomial.

Figure 7.27: $\pi\pi$-invariant mass projection of separation only fits using alternative shapes for the combinatorial background.

Parameterisation	Shift from standard fit	
	τ_{B^0} (μm/c)	$\tau_{B_s^0}$ (μm/c)
Polynomial, order 2	-2	0
Polynomial, order 3	-3	-1

Table 7.8: The shift from the standard fit of the fitted lifetimes for fits using a second or third order polynomial to parameterise the background mass shape.

lifetimes when using a second or third order polynomial to describe the combinatorial background mass shape. The B^0 lifetime is shifted more than the B_s^0 lifetime because the $B_s^0 \to K^+K^-$ decay mode mass peak lies at higher mass where there is a tighter constraint on the background fraction from the high mass region. The largest observed shifts in the B^0 and B_s^0 lifetimes are taken as the systematic uncertainties due to the choice of background mass shape.

7.12 Background lifetime model

Section 5.9 describes the probability density function (PDF) for the lifetime distribution in background. This used a general parameterisation of the lifetime as it is not necessary to know the physical meaning of the lifetime distribution. An alternative PDF for the background lifetime distribution is the sum of n exponential distributions

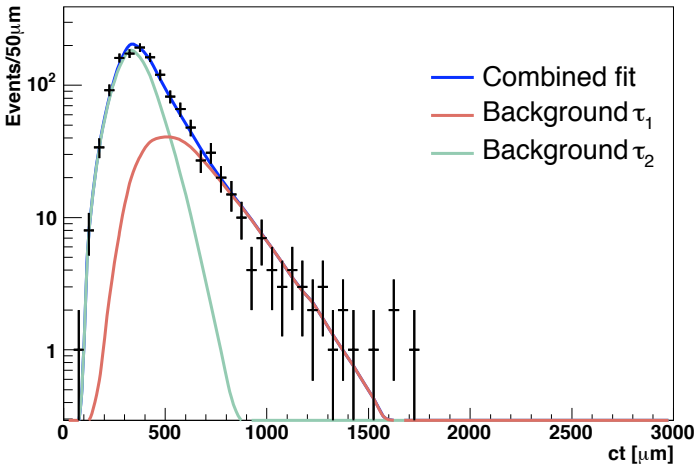


Figure 7.28: The sum of two exponential distributions fit to the lifetime distribution for background events from the high mass sideband in data.

normalised to the region $t_{\min} < t < t_{\max}$:

$$P(t \mid b, t_{\min}, t_{\max}) \begin{cases} = \sum_{i=0}^{n} f_i \frac{\frac{1}{\tau_i} e^{\frac{-t}{\tau_i}}}{e^{\frac{-t_{\min}}{\tau_i}} - e^{\frac{-t_{\max}}{\tau_i}}} & \text{for } t_{\min} < t < t_{\max} \\ = 0 & \text{for all other } t \end{cases}, \tag{7.8}$$

where f_i is the fraction of background with a mean lifetime τ_i. This parameterisation was found to give an unsatisfactory fit to background events from a high statistics sample of $B^+ \to D^0\pi^+$ candidates [63] and so this analysis uses the PDF described in section 5.9 in preference to equation 7.8.

The lifetime of background events in the fast simulation is generated from the same general function $y(t)$ as is used in the background lifetime PDF. To test how robust the PDF is to the true lifetime distribution the method is tested on fast simulated data with background lifetimes generated according to equation 7.8. Two exponential functions are used and the relative fraction and mean lifetimes are determined by fitting the distribution to data taken from the high mass sideband ($5.55 < M_{\pi\pi} < 5.8$ GeV/c^2) which contains only combinatorial background. Figure 7.28 shows the lifetime projection of the fit obtained.

1000 pseudo-experiments were performed on fast-simulated samples generated with

Distribution	Mean	Width
B^0 lifetime pull	-0.412 ± 0.037	1.040 ± 0.030
B_s^0 lifetime pull	-0.090 ± 0.038	1.019 ± 0.032
B^0 lifetime residual	-5.14 ± 0.52 μm/c	14.61 ± 0.39 μm/c
B_s^0 lifetime residual	-1.2 ± 1.2 μm/c	32.63 ± 0.93 μm/c

Table 7.9: Parameters of Gaussian fits to the pull and residual distributions from the study of the sensitivity of the method to the background lifetime distribution.

the alternative background lifetime distribution. Table 7.9 summarises the results of Gaussian fits to the B^0 and B_s^0 lifetime pull and residual distributions.

The residuals indicate biases on the measured B^0 and B_s^0 lifetimes of -5.14 μm/c and 1.20 μm/c respectively. Comparing this to the standard fit gives a systematic uncertainty due to the choice of background model of 6.27 μm/c for the B^0 lifetime and 5.8 μm/c for B_s^0 lifetime.

7.13 Particle identification model

The probability density function for particle identification is described in section 4.6. It includes 48 parameters measured with some statistical uncertainty and fixed in the likelihood. The systematic uncertainty on the measured lifetime due to the statistical uncertainty on the parameters in the particle identification PDF was determined for the previous measurement of the B_s^0 lifetime in $B_s^0 \to K^+K^-$ decays [53]. This was assessed by re-performing the lifetime fit with the parameters randomly varied within a 1σ sphere. This was repeated 100 times with different random shifts in the parameters and the root mean square of the distribution of fitted lifetimes taken as the systematic uncertainty resulting from the statistical uncertainty on the particle identification PDF. This approach neglects the correlation between the parameters of the PDF and so is conservative. The previous analysis assessed a systematic uncertainty due to the statistical uncertainty on the particle identification PDF of 2.0 μm. This is much smaller than the expected statistical uncertainty on the measured $B_s^0 \to K^+K^-$

lifetime and the same uncertainty is quoted for this analysis.

7.14 Input masses

The kinematic part of the likelihood distribution takes as input the masses of the B^0, B_s^0 and Λ_b^0 mesons as measured by CDF (section 4.5). The global mass scale is free to vary in the fit (each of the three input masses is allowed to shift by the same mass, Δ_m) however the difference between the masses is fixed. To test the sensitivity of the fit to the uncertainty on the $B_{(s)}^0$ and Λ_b^0 masses, the likelihood maximisation on data is repeated six times with the input masses individually varied by plus or minus the statistical uncertainty on the CDF measurements. The largest observed shift in the measured lifetime was 0.1 μm; more than an order of magnitude smaller than the largest systematic uncertainties so can be neglected.

7.15 B_s^0 lifetime in the $B_s^0 \to K^+\pi^-$ decay

The $B_{(s)}^0 \to h^+h'^-$ events include contributions from the decay $B_s^0 \to K^+\pi^-$ and its conjugate. This analysis assumes that the B_s^0 lifetime in $B_s^0 \to K^+\pi^-$ is the same as the B^0 lifetime and this mode contributes to the measurement of the B^0 lifetime in the fit. As a check the combined separation and lifetime fit was performed with the lifetime in $B_s^0 \to K^+\pi^-$ included as an additional parameter, free to vary in the fit. The B_s^0 lifetime in $B_s^0 \to K^+K^-$ remained unchanged. The B^0 lifetime shifted by -2 μm/c, much less than than the statistical uncertainty on the B^0 lifetime. The lifetime in $B_s^0 \to K^+\pi^-$ was consistent with the B^0 lifetime however the statistical uncertainty was 79 μm/c. An additional 2 μm/c systematic uncertainty is added to the B^0 lifetime due to assuming the B_s^0 lifetime in $B_s^0 \to K^+\pi^-$ decays equals the B^0 lifetime.

7.16 Other sources of uncertainty

Other potential sources of systematic uncertainty in lifetime measurements have been considered by the first analysis to use the simulation free method to remove the lifetime bias caused by the hadronic trigger [63], as used here. The largest of these uncertainties, such as the effect of the track finding efficiency of the SVT, have already been discussed in this chapter. The remaining ones were found to be less than 0.5 μm/c, an order of magnitude smaller than the other systematic uncertainties and far smaller than the expected statistical uncertainty, so negligible for this analysis. A brief discussion of these sources of uncertainty is included here.

7.16.1 Silicon alignment

Misalignment of the silicon wafers will affect the determination of the p-$\bar{\text{p}}$ beam position and measurements of particle impact parameters and B meson decay vertex positions as well as the selection of events by the hadronic B_PIPI trigger. The silicon wafers are conservatively estimated to be aligned to within 50 μm [67]; this is much smaller than the typical distance travelled by B mesons in the detector (1 cm) and smaller than the minimum transverse distance between the primary interaction and B decay vertices required for the data sample used in this analysis (300 μm). The systematic uncertainty on the measured lifetime due to misalignment of the silicon wafers is estimated to be 0.4 μm/c for data collected by the hadronic triggers [63].

7.16.2 Lifetime resolution

It has been assumed that the resolution of the measured lifetime is Gaussian distributed (section 5.3). Other CDF lifetime analyses have found that the resolution distribution actually has three Gaussian components [68]. In [63] the effect of assuming the lifetime resolution is Gaussian was estimated by performing 1000 pseudo-experiments on fast

B^0 lifetime systematic uncertainty	
Systematic	Uncertainty (μm/c)
Method bias	1.13
SVT-efficiency with respect to d_0	3.25
SVT-efficiency with respect to p_T	0.89
Background mass model	3
Background lifetime model	6.27
Parameterisation of dE/dx	2.0
$B_s^0 \to K^+\pi^-$ lifetime	2
Total	8.30

Table 7.10: Summary of the systematic uncertainties on the measurement of the B^0 lifetime. The total is the sum in quadrature of the individual uncertainties.

simulated samples with measured lifetimes smeared by a three Gaussian resolution function. The experiments then measured the lifetime of the sample assuming a single Gaussian resolution function. A systematic uncertainty of 0.27 μm/c was estimated.

7.16.3 SVT efficiency as a function of η

Just as the SVT track finding efficiency is not uniform as a function of impact parameter (section 7.6) and particle transverse momentum (section 7.7) it is also non-uniform as a function of the η of the particle. The η of the decay products is independent of the B meson lifetime, however, so the method should be insensitive to changes in the SVT efficiency with η. The systematic uncertainty due to the non-uniform SVT efficiency as a function of η was estimated to be 0.3 μm/c [63].

7.17 Summary of systematic uncertainties

Tables 7.10 and 7.11 summarise the systematic uncertainties on the measurements of the B^0 and B_s^0 lifetimes respectively. The totals are the sum in quadrature of all the contributions to the systematic uncertainty.

The systematic uncertainty on the B^0 lifetime is dominated by the uncertainty on

B_s^0 lifetime systematic uncertainty	
Systematic	Uncertainty (μm/c)
Method bias	4.6
SVT-efficiency with respect to d_0	7.5
SVT-efficiency with respect to p_T	4.0
Background mass model	1
Background lifetime model	5.8
Parameterisation of dE/dx	2.0
$B_s^0 \to K^+\pi^-$ lifetime	0
Total	11.5

Table 7.11: Summary of the systematic uncertainties on the measurement of the B_s^0 lifetime in its decay to $B_s^0 \to K^+K^-$. The total is the sum in quadrature of the individual uncertainties.

the parameterisation of the background lifetime distribution and the non-uniformity of the SVT track finding efficiency with respect to particle impact parameter. The largest systematic uncertainty on the B_s^0 lifetime is the non-uniform SVT efficiency with respect to impact parameter. The efficiency with respect to transverse momentum, the background lifetime model and the general bias in the method are also significant for this lifetime. The total systematic uncertainty on the B_s^0 lifetime is 12 μm/c compared to the expected statistical uncertainty of about 33 μm/c. At the close of CDF it is anticipated that 8 fb^{-1} integrated luminosity will have been collected. With this data sample size the statistical and systematic uncertainties will be of comparable size.

7.18 Projections of the final fit

The lifetime of the B^0 meson and the lifetime of the B_s^0 meson in its decay $B_s^0 \to K^+K^-$ are determined by maximising the likelihood described in section 6.3 for the data sample described in sections 3.2 to 3.4. The projection of the fitted PDF onto the lifetime variable is shown in figure 7.29 along with the distribution of the difference between the binned projection and the binned lifetime distribution in data divided by the statistical uncertainty on the true distribution (the residual). The combined separation and lifetime PDF models the lifetime distribution in data well. Figures 7.30,

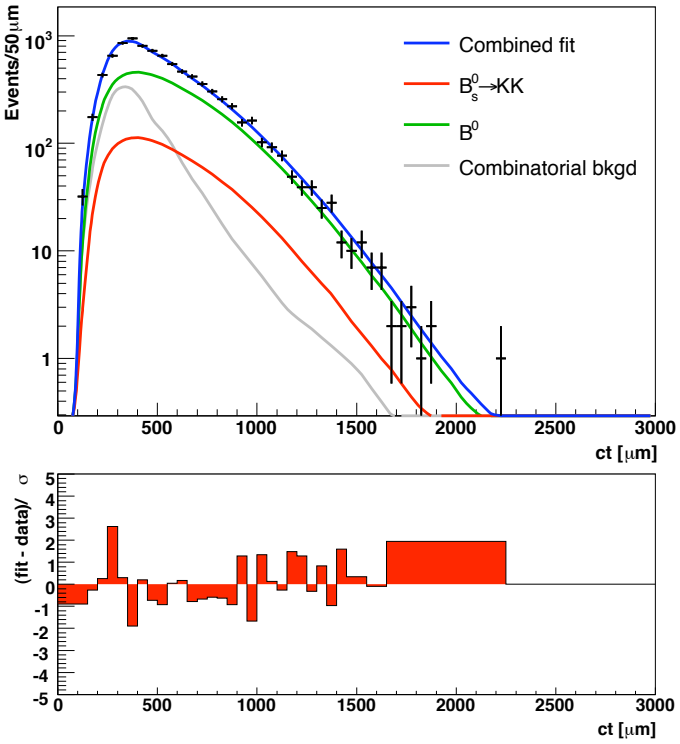


Figure 7.29: Lifetime projection of the final fit.

7.31 and 7.32 show the projection of the fitted PDF onto the other variables in the likelihood and the associated residual distributions. The fitted PDF models all of the distributions well. The lifetimes measured by this analysis will be presented in the next chapter.

7.19 Summary

This chapter assessed the performance of the methods used by this analysis to simultaneously separate the contributions to the $B^0_{(s)} \to h^+h'^-$ signal and measure the mean lifetimes of the B^0 meson and the B^0_s meson in the decay $B^0_s \to K^+K^-$. Two different likelihood functions were tested, each requiring different assumptions. The likelihood which provided the least biassed measurements was chosen. The systematic uncertainties due to the assumptions made in this analysis have been estimated and and the total systematic uncertainty for the $B^0_s \to K^+K^-$ lifetime is 13 μm/c, much smaller

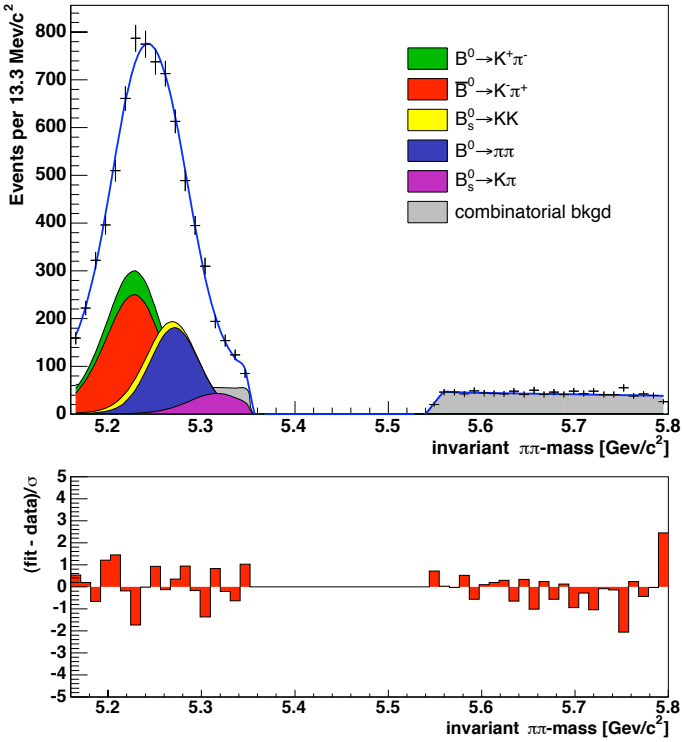
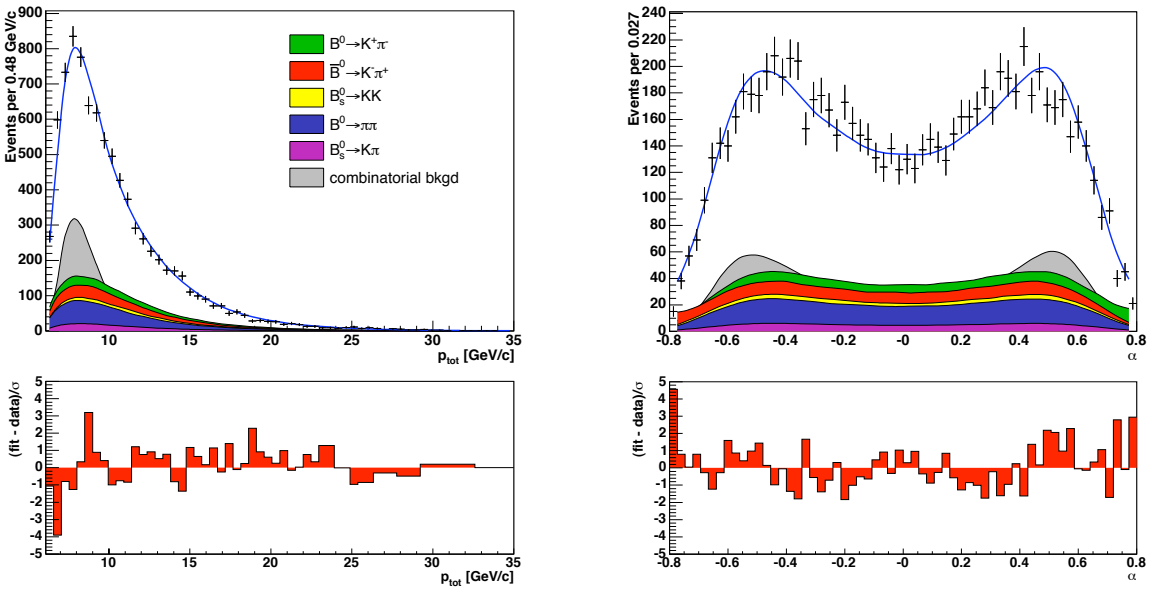
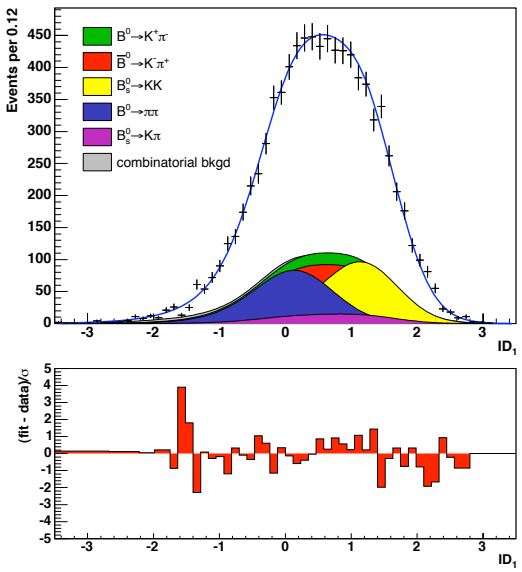


Figure 7.30: $\pi\pi$-invariant mass projection of the final fit.

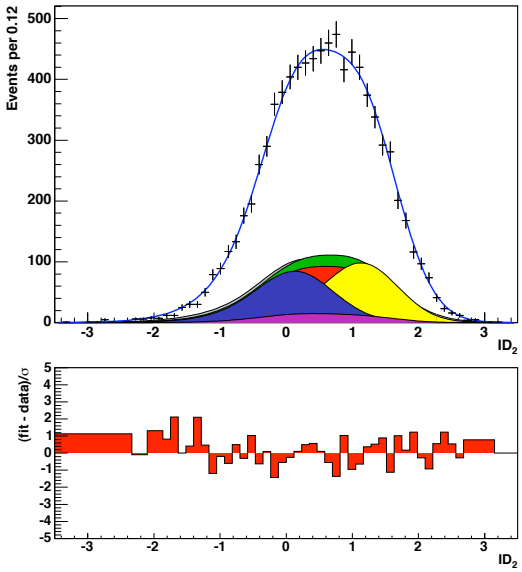


(a) p_{tot} (b) α

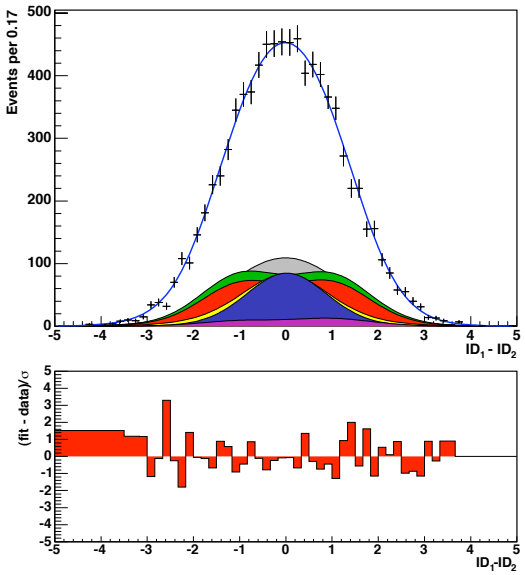
Figure 7.31: Projection of the final fit onto the α and p_{tot} variables.



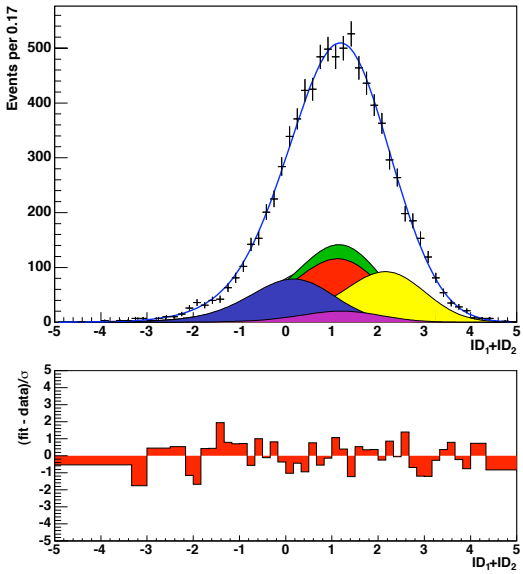
(a) ID_1



(b) ID_2



(c) $ID_1 - ID_2$



(d) $ID_1 + ID_2$

Figure 7.32: Projection of the final fit onto the ID_i variables.

than the expected statistical uncertainty.

Finally, the projections onto each of the separation variables of the final likelihood fit to the data agree well with the distributions in data. The final chapter presents the results of this fit.

Chapter 8

Results and discussion

8.1 Introduction

This thesis describes a method to measure the B_s^0 lifetime in the decay $B_s^0 \to K^+K^-$ in a data sample collected by a hadronic trigger and containing several $B_{(s)}^0 \to h^+h'^-$ decay modes. Section 7.18 introduced the likelihood fit used to measure the B^0 and $B_s^0 \to K^+K^-$ lifetime and the results of that fit are presented in this chapter. These measurements are compared to previous measurements and theoretical predictions.

Section 1.6 discussed how the lifetime measured in $B_s^0 \to K^+K^-$ can be used to search for the effects of new physics, in particular by comparison to the lifetime measured in $B_s^0 \to J/\psi\phi$ and that comparison is made here. Finally, section 8.3 considers the prospects for the methods used in this analysis and for repeating this measurement.

8.2 Results

The analysis described and tested in the previous chapters provides a measurement of the B^0 lifetime of

$$\tau_{B^0} = 1.558^{+0.050}_{-0.047} \pm 0.028 \text{ ps}, \tag{8.1}$$

where the convention is adopted that the first uncertainty is statistical and the second systematic. As the B^0 lifetime is already well measured, comparison of this measurement with the world average measurement provides a good test of the analysis methods. The world average measurement is $\tau_{B^0} = 1.530 \pm 0.009$ ps [2], in good agreement with this measurement.

The B_s^0 lifetime in the decay $B_s^0 \to K^+K^-$ is measured by this analysis to be

$$\tau_{B_s^0 \to K^+K^-} = 1.51^{+0.13}_{-0.11} \pm 0.04 \text{ ps}. \tag{8.2}$$

This is compatible with the previous CDF measurement on the first 360 pb^{-1} of this data sample: $\tau_{B_s^0 \to K^+K^-} = 1.53 \pm 0.18 \pm 0.02$ ps [53] with an improved statistical uncertainty. The quoted systematic uncertainty is larger for this analysis than the earlier analysis. This is primarily due to the drop in SVT track finding efficiency for high impact parameters (see section 7.6) which leads to a 0.025 ps systematic uncertainty and does not impact on the simulation based method to account for the trigger bias used in [53]. It should also be noted, however, that recent studies of the ability of CDF realistic simulation to replicate the SVT track finding efficiency in data have lead to additional systematic uncertainties of 0.014 ps [69] and 0.021 ps [70] being included in the latest lifetime measurements using the simulation based method from CDF.

The current world average of the flavour specific B_s^0 meson lifetime is $\tau_{B_s^0} = 1.417 \pm 0.042$ ps [2]. A recent measurement at CDF using $B_s^0 \to D_s\pi$ decays provides the worlds best measurement of the B_s^0 lifetime: $\tau_{B_s^0} = 1.518 \pm 0.041 \pm 0.27$ ps [69] (which is also closer to the B^0 lifetime than previous measurements). No difference is observed between the flavour specific B_s^0 lifetime and the B_s^0 lifetime measured in $B_s^0 \to K^+K^-$ decays here.

Assuming that $B_s^0 \to K^+K^-$ is a 100% CP-even eigenstate the lifetime measured in the decay $B_s^0 \to K^+K^-$ is the lifetime of the light B_s^0 mass eigenstate: $\tau_{B_s^0} = \tau_{B_{sL}^0}$. The

current world average B^0_{sL} lifetime is $\tau_{B^0_{sL}} = 1.45^{+0.14}_{-0.12}$ ps [2], in agreement with this measurement. No difference between the lifetime of the light B^0_s mass eigenstate and the B^0_s lifetime measured in $B^0_s \to K^+K^-$ decays has been observed. The uncertainty on $\tau_{B^0_s \to K^+K^-}$ from this analysis is competitive with the uncertainty on the world average $\tau_{B^0_{sL}}$ lifetime; however, theory predicts that $B^0_s \to K^+K^-$ is only 95% CP-even [19] so a careful consideration of the CP composition of this mode is needed before a statement on $\tau_{B^0_{sL}}$ can be made, which is beyond the scope of this thesis.

Comparison of the B^0_s lifetime measured in $B^0_s \to K^+K^-$ with the lifetime measured in the CP-even component of $B^0_s \to J/\psi\phi$ can be used to determine the CP phase in the $b \to s\bar{u}u$ decay amplitude. The most recent CDF lifetime measurements in $B^0_s \to J/\psi\phi$ find the lifetime of the light mass eigenstate is $\tau_{B^0_{sL}} = 1.44^{+0.076}_{-0.072} \pm 0.022$ ps [16]. This is consistent with the measurement in $B^0_s \to K^+K^-$ so this analysis is not sensitive to the CP-phase in the $b \to s\bar{u}u$ decay amplitude.

Assuming that $B^0_s \to K^+K^-$ is 100% CP-even and that $\tau_{B^0_s} = \tau_{B^0}$, the width difference in the B^0_s-system for CP-eigenstates can be determined as

$$\frac{\Delta\Gamma^{CP}}{\Gamma} = 0.03^{+0.17}_{-0.15} \pm 0.05 \tag{8.3}$$

using the world average B^0 lifetime. The previous CDF measurement on the first 360 pb^{-1} of this data sample obtained $\frac{\Delta\Gamma^{CP}}{\Gamma} = -0.09 \pm 0.23 \pm 0.03$ using the 2005 world average flavour specific B^0_s lifetime $\tau_{B^0_s} = 1.454 \pm 0.040$ ps [71]. The current world average is $\frac{\Delta\Gamma^{CP}}{\Gamma} = 0.10 \pm 0.05$. The result of this analysis is consistent with the world average.

A recent theoretical calculation predicts $\frac{\Delta\Gamma}{\Gamma} = 0.127 \pm 0.24$ [12]. The measurement of this analysis is within 1σ of this prediction and is also consistent with zero.

8.3 Future prospects

The uncertainties on the measurements in this analysis are dominated by the statistical uncertainty. The current intention is that CDF will continue taking data into 2010 and it is anticipated that 8 fb^{-1} of data will be collected. Assuming that the number of $\text{B}_\text{s}^0 \to \text{K}^+\text{K}^-$ events will scale with integrated luminosity the expected statistical error on the B_s^0 lifetime in $\text{B}_\text{s}^0 \to \text{K}^+\text{K}^-$ decays with 8 fb^{-1} is 0.042 ps. In fact it is likely that the hadronic triggers will be more heavily prescaled at high luminosities so this is estimate is optimistic. Most of the systematic uncertainties will remain unchanged with increased statistics except for the uncertainty due to the drop in the SVT track finding efficiency at large impact parameters (discussed in section 7.6). Improvements in the SVT have reduced the drop in track finding efficiency for impact parameters approaching 1000 μm so this systematic uncertainty, currently the largest, will be smaller for 8 fb^{-1}. The uncertainty on this measurement will continue to be statistics dominated with 8 fb^{-1} of integrated luminosity collected.

Beyond the Tevatron experiments the next place to study B mesons will be LHCb [72], due to begin collecting data in 2009. The LHCb detector will have much better particle identification capabilities than CDF with its RICH detector providing 88% efficient kaon identification for particles with a momentum between 2 and 100 GeV/c with an average pion misidentification rate of only 3% in the same momentum range [1] [73]. Figure 8.1 shows the invariant mass distribution (KK-invariant mass) of simulated $\text{B}_{(s)}^0 \to h^+h'^-$ and $\Lambda_\text{b}^0 \to \text{p}h^-$ events meeting trigger and offline selection requirements (described in [73]). There is very little background from $\text{B}_{(s)}^0 \to h^+h'^-$ or $\Lambda_\text{b}^0 \to \text{p}h^-$ events making this a much easier place to study $\text{B}_\text{s}^0 \to \text{K}^+\text{K}^-$ decays. Statistical separation of the contributions will not be necessary removing the associated systematic uncertainties. The expected annual yield ($\mathcal{L} = 2\text{fb}^{-1}$) of $\text{B}_\text{s}^0 \to \text{K}^+\text{K}^-$ events at LHCb is 35900 with an expected combinatorial background to $\text{B}_\text{s}^0 \to \text{K}^+\text{K}^-$ signal

[1] The pion misidentification rate can be further reduced with a corresponding reduction in the kaon identification efficiency.

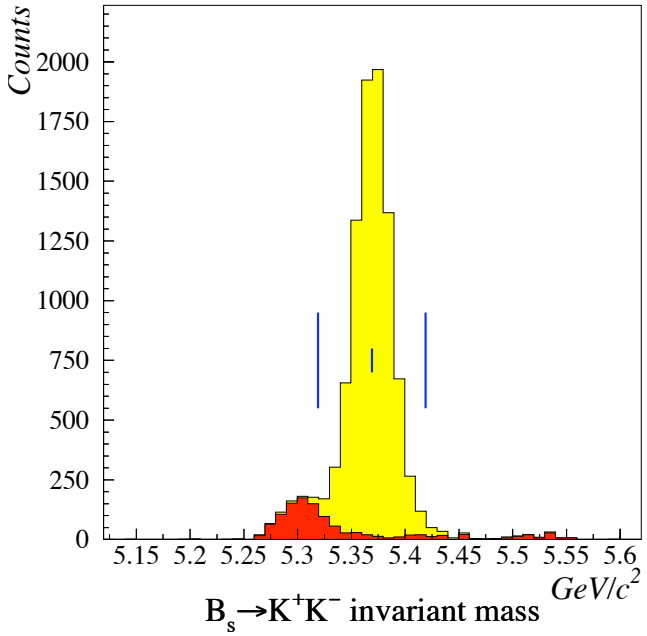


Figure 8.1: KK-Invariant mass distribution of $B_s^0 \to K^+K^-$ candidates selected by trigger and offline requirements from simulation (taken from [73] refer there for details of selection requirements). The yellow histogram contains simulated $B_s^0 \to K^+K^-$ events and the red histogram contains other simulated $B_{(s)}^0 \to h^+h'^-$ and $\Lambda_b^0 \to ph^-$ events.

ratio of < 0.08 and an expected ratio of other $B_{(s)}^0 \to h^+h'^-$ and $\Lambda_b^0 \to ph^-$ events to $B_s^0 \to K^+K^-$ events of 0.02 [74] which is much cleaner than the sample used in this analysis with ~ 27 times the number of $B_s^0 \to K^+K^-$ events. Assuming that the statistical uncertainty assessed here will scale to LHCb the expected statistical uncertainty on the B_s^0 lifetime measured in $B_s^0 \to K^+K^-$ decays would be 0.02 ps.

The lifetime distribution of decays collected using an impact parameter based trigger will be biassed. This thesis describes a method for removing this bias without using detailed simulation to model the effects of the trigger and this analysis is the second measurement to use this method and the first on a two-body decay topology. It is intended that the method will be used at LHCb for data collected by its hadronic triggers [75] and this analysis is an important test of that method at a running experiment.

8.4 Conclusions

This thesis described the statistical separation of a sample containing a number of $B^0_{(s)} \to h^+h'^-$ and $\Lambda^0_b \to ph^-$ decays and the measurement of the lifetimes of the B^0 meson and the B^0_s meson in $B^0_s \to K^+K^-$ decays. The data used in this analysis were collected by an impact parameter based trigger which biasses the lifetime distributions in the data. A method to remove this bias without the need for detailed data and detector simulation was used and this is the second time this method has been tested on data. This is an important test as the LHCb experiment intends to use this method to measure lifetimes in fully hadronic decay modes.

The B^0 lifetime is measured to be

$$\tau_{B^0} = 1.558^{+0.050}_{-0.047} \pm 0.028 \text{ ps}, \tag{8.4}$$

in agreement with the current world average measurement. The B^0_s lifetime in the decay $B^0_s \to K^+K^-$ is measured to be

$$\tau_{B^0_s \to K^+K^-} = 1.51^{+0.13}_{-0.11} \pm 0.04 \text{ ps}. \tag{8.5}$$

which is dominated by statistical the uncertainty. This measurement is not sensitive to any difference between the B^0_s lifetime in $B^0_s \to K^+K^-$ decays and the flavour specific B^0_s lifetime or to the lifetime of the light B^0_s mass eigenstate.

At the completion of running in 2010 it is anticipated that CDF will have a sensitivity of 0.042 ps to the B^0_s lifetime in $B^0_s \to K^+K^-$ decays.

Bibliography

[1] S. L. Glashow, *Partial Symmetries Of Weak Interactions*, Nucl. Phys. **22**, 579 (1961);
A. Salam and J. C. Ward, *Electromagnetic and Weak Interactions*, Phys. Lett. **13**, 168 (1964);
S. Weinberg, *A Model Of Leptons*, Phys. Rev. Lett. **19**, 1264 (1967).

[2] C. Amsler *et al.* [Particle Data Group], *Review of Particle Physics 2008*, Phys. Lett. **B667**, 1 (2008).

[3] Y. Fukuda *et al.* [Super-Kamiokande Collaboration], *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81**, 1562 (1998) [arXiv:hep-ex/9807003].

[4] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12**, 132 (1964).

[5] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13**, 321 (1964).

[6] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10**, 531 (1963).

[7] M. Kobayashi and T. Maskawa, *CP Violation In The Renormalizable Theory Of Weak Interaction*, Prog. Theor. Phys. **49**, 652 (1973).

[8] L. Wolfenstein, *Parametrization Of The Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51**, 1945 (1983).

[9] I. I. Y. Bigi, *The QCD Perspective On Lifetimes Of Heavy Flavor Hadrons*, arXiv:hep-ph/9508408.

[10] E. Franco, V. Lubicz, F. Mescia and C. Tarantino, *Lifetime ratios of beauty hadrons at the next-to-leading order in QCD*, Nucl. Phys. B **633**, 212 (2002) [arXiv:hep-ph/0203089].

[11] I. I. Y. Bigi, *The lifetimes of heavy flavour hadrons: A case study in quark hadron duality*, arXiv:hep-ph/0001003.

[12] A. Lenz and U. Nierste, *Theoretical update of* B_s^0 *-* *anti-*B_s^0 *mixing*, JHEP **0706**, 072 (2007) [arXiv:hep-ph/0612167].

[13] C. Tarantino, *Beauty hadron lifetimes and B-meson CP-violation parameters from lattice QCD*, Eur. Phys. J. C **33**, S895 (2004) [arXiv:hep-ph/0310241].

[14] A. S. Dighe, I. Dunietz and R. Fleischer, *Extracting CKM phases and* $\mathrm{B_s^0}$ *anti-*$\mathrm{B_s^0}$ *mixing parameters from angular distributions of non-leptonic B decays*, Eur. Phys. J. C **6**, 647 (1999) [arXiv:hep-ph/9804253].

[15] A. S. Dighe, I. Dunietz, H. J. Lipkin and J. L. Rosner, *Angular distributions and lifetime differences in* $B_s \to J/\psi\phi$ *decays*, Phys. Lett. B **369**, 144 (1996) [arXiv:hep-ph/9511363].

[16] T. Aaltonen *et al.* [CDF Collaboration], *Measurement of Lifetime and Decay-Width Difference in* $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi$ *Decays*, Phys. Rev. Lett. **100**, 121803 (2008) [arXiv:0712.2348 [hep-ex]].

[17] I. Dunietz, R. Fleischer and U. Nierste, *In pursuit of new physics with* B_s *decays*, Phys. Rev. D **63**, 114015 (2001) [arXiv:hep-ph/0012219].

[18] The CDF Collaboration, *An updated measurement of the CP violating phase* $\beta_s^{\mathrm{J}/\psi\phi}$, CDF public note 9458 (2008), http://www-cdf.fnal.gov/physics/new/bottom/080724.blessed-tagged_BsJPsiPhi_update_prelim/.

[19] R. Fleischer and J. Matias, *Searching for new physics in non-leptonic B decays*, Phys. Rev. D **61**, 074004 (2000) [arXiv:hep-ph/9906274];
R. Fleischer and J. Matias, *Exploring CP violation through correlations in* $B \to \pi K$, $B_d \to \pi^+\pi^-$, $B_s \to K^+K^-$ *observable space*, Phys. Rev. D **66**, 054009 (2002) [arXiv:hep-ph/0204101].

[20] *Accelerator Concepts Rookie Book*, v3.1 (2006), http://www-bdnew.fnal.gov/operations/rookie_books/Concepts_v3.1.pdf.

[21] D. Mohl, G. Petrucci, L. Thorndahl and S. Van Der Meer, *Physics And Technique Of Stochastic Cooling*, Phys. Rept. **58**, 73 (1980).

[22] S. Nagaitsev *et al.*, *Experimental demonstration of relativistic electron cooling*, Phys. Rev. Lett. **96**, 044801 (2006).

[23] Regularly updated plots of CDF integrated and peak luminosity can be obtained at http://www.fnal.gov/pub/now/tevlum.html.

[24] D. Allspach *et al.* [the CDF Collaboration], *Aging in the large CDF axial drift chamber*, IEEE Trans. Nucl. Sci. **52** (2005) 2956.

[25] R. Blair *et al.* [CDF II Collaboration], *The CDF II Detector, Technical Design Report*, FERMILAB-Pub-96/390-E CDF (1996).

[26] Images courtesy of the CDF Collaboration, http://www-cdf.fnal.gov/.

[27] T. Affolder *et al.*, *CDF Central Outer Tracker*, Nucl. Instrum. Methods **A526**, 249 (2004).

[28] A. Sill, *CDF Run II silicon tracking projects*, Nucl. Instrum. Methods A **447** 1 (2000).

[29] C. S. Hill, *Initial experience with the CDF layer 00 silicon detector*, Nucl. Instrum. Methods **A511**, 118 (2003).

[30] A. A. Affolder *et al.*, *Status report of the intermediate silicon layers detector at CDFII*, Nucl. Instrum. Methods **A485**, 6 (2002).

[31] D. Acosta *et al.*, *A Time-of-flight detector in CDF-II*, Nucl. Instrum. Methods **A518**, 605 (2004).

[32] L. Balka *et al.*, *The CDF central electromagnetic calorimeter*, Nucl. Instrum. Methods **A267**, 272 (1988).

[33] S. Bertolucci *et al.*, *The CDF central and endwall hadron calorimeter*, Nucl. Instrum. Methods **A267**, 301 (1988).

[34] M.Albrow *et al.*, *The CDF plug upgrade electromagnetic calorimeter: test beam results*, Nucl. Instrum. Methods **A480**, 524 (2002).

[35] G. Ascoli *et al.*, *CDF central muon detector*, Nucl. Instrum. Methods **A268**, 33 (1988);
C. M. Ginsburg, *CDF run 2 muon system*, Eur. Phys. J. **33**, S1002 (2004).

[36] E. J. Thomson *et al.*, *Online track processor for the CDF upgrade*, IEEE Trans. Nucl. Sci. **49**, 1063 (2002).

[37] A. Abulencia *et al.* [CDF Collaboration], *The CDF II extremely fast tracker upgrade*, Nucl. Instrum. Methods **A572**, 358 (2007).

[38] B. Ashmanskas *et al.* [CDF-II Collaboration], *The CDF silicon vertex trigger*, Nucl. Instrum. Methods **A518**, 532 (2004) [arXiv:physics/0306169].

[39] M. Dell'Orso [CDF Collaboration], *The CDF Silicon Vertex Trigger*, Nucl. Phys. Proc. Suppl. **156**, 139 (2006).

[40] For an exact definition of the χ^2_{SVT} quantity see
S. Belaforte *et al.*, *Silicon Vertex Tracker Technical Design Report*, CDF internal note 3108 (1995), unpublished.

[41] The CDF Collaboration, *Measurement of branching fractions and direct CP asymmetries of* $\mathrm{B}^0_{(s)} \to h^+h'^-$ *decays in 1*fb^{-1}, CDF public note 8579 (2006), http://www-cdf.fnal.gov/physics/new/bottom/060921.blessed-bhh_1fb/.

[42] J. D. Bjorken, *Properties Of Hadron Distributions In Reactions Containing Very Heavy Quarks*, Phys. Rev. D **17**, 171 (1978);
M. Suzuki, *Fifth Quark And Dimuon Production By Neutrinos*, Phys. Lett. B **68**, 164 (1977).

[43] P. Sphincas, *A $b\bar{b}$ Monte Carlo Generator*, CDF internal note 2655 (1994), unpublished;
K. Anikeev, P. Murat and Ch. Paus, *Description of Bgenerator II*, CDF internal note 5092 (1999), unpublished.

[44] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Methods A **462**, 152 (2001);
W. Bell *et al.*, *User Guide for EvtGen @ CDF*, CDF internal note 5618 (2003), unpublished.

[45] E. Gerchtein and M. Paulini, *CDF detector simulation framework and performance*, eConf. **C0303241** TUMT005 (2003) [physics/0306031] (2003).

[46] R. Brun and F. Carminati, *GEANT detector description and simulation tool*, CERN Programming Library Long Writeup **W5013** (1993).

[47] R. Veenhof, *GARFIELD, recent developments*, Nucl. Instrum. Methods A **419**, 726 (1998).

[48] M. Gold, *Description of the Parameterised Charge Deposition Model*, CDF internal note 5871 (2002), unpublished.

[49] K. Hagiwara *et al.* [Particle Data Group], *Review of particle physics 2002*, Phys. Rev. D 66 (2002).

[50] W.-M.Yao *et al.* [Particle Data Group], *Review of particle physics 2006*, J. Phys. G 33, 1 (2006).

[51] M. Donega *et al.*, $B_s^0 \to K^+K^-$ *lifetime measurement and extraction of* $\Delta\Gamma_{CP}/\Gamma_{CP}$, CDF internal note 7978 (2006), unpublished.

[52] A. Abulencia *et al.* [The CDF Collaboration], *Obsevation of* $B_s^0 \to K^+K^-$ *and measurements of branching fractions of two-body decays of* B^0 *and* B_s^0 *mesons in p anti-p collisions at* $\sqrt{s} = 1.96$ TeV, Phys. Rev. Lett. **97**, 211802 (2006).

[53] The CDF Collaboration, *Measurement of the* $B_s^0 \to K^+K^-$ *lifetime and extraction of* $\Delta\Gamma_{CP}/\Gamma_{CP}$, CDF public note (2006),
http://www-cdf.fnal.gov/physics/new/bottom/060126.blessed-BsKK_lifetime/.

[54] D. Acosta *et al.* [The CDF Collaboration], *Measurement of bottom-quark hadron masses in exclusive J/Ψ decays with the* CDF *detector*, Phys. Rev. Lett. 96, 202001 (2006) [hep-ex/0504006].

[55] M. J. Morello, *Measurements of CP asymmetries and branching fractions of two-body charmless decays of B^0 and B_s^0 mesons* (PhD thesis), University of Pisa (2007).

[56] E. Barberio *et al.* [Heavy Flavor Averaging Group], *Averages of b-hadron Properties at the End of 2006* , [hep-ex/0704.3575v1].

[57] M. Beneke and M. Neubert, *QCD factorization for $B \to P\ P$ and $B \to P\ V$ decays*, Nucl. Phys. **B 675**, 333 (2003) [hep-ph/0308039].

[58] H.Albrecht *et al.* [Argus Collaboration], Phys. Lett. B **241**, 2788 (1990); **254**, 288(1991).

[59] R. Carosi *et al.*, *Update of Combined PID in 5.3.4*, CDF internal note 7866 (2005).

[60] M. A. Ciocci *et al.*, *Update of Combined PID in 6.4.1*, CDF internal note 8478 (2006).

[61] K. Pearson, *Contributions to the mathematical theory of evolution – II. Skew variation in homogeneous material*, Philos. Trans. Roy. Soc. London A **186**, 343 (1895);
K. Pearson, *Contributions to the mathematical theory of evolution – X. Supplement to a memoir on skew variation*, Philos. Trans. Roy. Soc. London A 197, 443 (1901);
K. Pearson, *Contributions to the mathematical theory of evolution – XIX. Second supplement to a memoir on skew variation*, Philos. Trans. Roy. Soc. London A **216**, 429 (1916).

[62] The CDF Collaboration, *B+, B0 and Bs meson lifetimes (exclusive reconstruction in hadronic decays)*, CDF public note 7386 (2005),
http://www-cdf.fnal.gov/physics/new/bottom/050303.blessed-bhadlife/blessed-bhadlife.ps.

[63] S. Malde *et al.*, *Measurement of the $B^{\pm}$ lifetime using a new simulation free method to correct for trigger induced bias*, CDF public note 9370 (2008);
S. Malde *et al.*, *A measurement of the $B^{\pm}$ lifetime using the* SVT *based trigger and accounting for the induced bias using a monte-carlo free method*, CDF internal note 9334 (2008).

[64] Giovanni Punzi, *Comments on likelihood fits with variable resolution*, physics/0401045, published in conference proceedings Phystat 2003, SLAC.

[65] R. A. Fisher, *The use of multiple measurements in taxonomic problems*, Annals of Eugenics, 7: 179–188 (1936);
Glen Cowan, *Statistical data analysis*, Oxford Science publications, pg. 51–54 (1998);
W. J. Krazanowski, *The algebraic basis of classical multivariate methods*, The statistician, vol. 20 No. 4 pg. 51–61 (1971).

[66] E. Waring, *Problems concerning interpolations*, Philos. Trans. **69**, 59–67, (1779);
B. Archer, and E. W. Weisstein, *Lagrange Interpolating Polynomial.* From MathWorld — A Wolfram Web Resource, http://mathworld.wolfram.com/LagrangeInterpolatingPolynomial.html.

[67] K. Anikeev *et al.*, *New measurement of lifetimes of B mesons in the exclusively reconstructed decays* $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^{*0}$, $B_s^0 \to J/\psi\phi$, CDF internal note 6551, (2004).

[68] J. Boudreau *et al.*, *Measurement of Exclusive B Lifetimes in the Modes* $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^{*0}$, $B^0 \to J/\psi K_s^0$, $B_s^0 \to J/\psi\phi$ *and* $\Lambda_b^0 \to J/\psi\Lambda$, *and a*

measurement of the B^+/B^0*,* B_s^0/B^0 *and* Λ_b^0/B^0 *lifetime ratios* CDF internal note 8524, (2006).

[69] The CDF Collaboration, *Measurement of the* B_s^0 *Lifetime in Hadronic Decays Using Partially and Fully Reconstructed Modes*, CDF public note 9230, (2008), http://www-cdf.fnal.gov/physics/new/bottom/080207.blessed-bs-lifetime/pubnote_v2.ps.

[70] The CDF Collaboration, Λ_b *Lifetime using 1* fb^{-1} *Data Taken With Two Displaced Track Triggers*, CDF public note 9408 (2008), http://www-cdf.fnal.gov/physics/new/bottom/080703.blessed-lblcpi-ct/lambdab_lifetime_public.pdf.

[71] E. Barberio *et al.* [Heavy Flavor Averaging Group (HFAG)], *Averages of b-hadron properties at the end of 2005*, arXiv:hep-ex/0603003.

[72] A. Augusto Alves Jr *et al.* [The LHCb Collaboration], *The LHCb Detector at the LHC*, JINST **3** S08005 (2008).

[73] The LHCb Collaboration, *LHCb technical design report: Reoptimized Detector Design and Performance Technical Design Report*, CERN/LHCC 2003-030 (2003).

[74] A. Carbone *et al.*, *Charmless Charged Two-body B decays at LHCb*, CERN-LHCb-2007-059 (2007).

[75] V. V. Gligorov and J. Rademacker, *Monte Carlo Independent Lifetime Fitting at LHCb in Lifetime Biassed Channels*, CERN-LHCb-2007-053 (2007).