Measurement of the CP violating phase β_s in $B_s^0 \to J/\psi\phi$ decays

Louise Oakes
St Cross College

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

Trinity Term, 2010

Abstract

The CP violating phase $\beta_s^{J/\psi\phi}$ is measured in decays of $B_s^0 \to J/\psi\,\phi$. This measurement uses 5.2 fb^{-1} of data collected in $\sqrt{s} = 1.96$ TeV $p\bar{p}$ collisions at the Fermilab Tevatron with the CDF Run-II detector. CP violation in the B_s^0-$\bar{B}_s^0$ system is predicted to be very small in the Standard Model. However, several theories beyond the Standard Model allow enhancements to this quantity by heavier, New Physics particles entering second order weak mixing box diagrams. Previous measurements have hinted at a deviation from the Standard Model expectation value for $\beta_s^{J/\psi\phi}$with a significance of approximately 2σ. The measurement described in this thesis uses the highest statistics sample available to date in the $B_s^0 \to J/\psi\,\phi$ decay channel, where $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$. Furthermore, it contains several improvements over previous analyses, such as enhanced signal selection, fully calibrated particle ID and flavour tagging, and the inclusion of an additional decay component in the likelihood function. The added decay component considers S-wave states of KK pairs in the $B_s^0 \to J/\psi K^+K^-$ channel. The results are presented as 2-dimensional frequentist confidence regions for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$(the width difference between the B_s^0 mass eigenstates), and as a confidence interval for $\beta_s^{J/\psi\phi}$ of [0.02,0.52] $\cup$ [1.08, 1.55] at the 68 % confidence level. The measurement of the CP violating phase obtained in this thesis is complemented by the world's most precise measurement of the lifetime $\tau_s = 1.53 \pm 0.025$ (stat.) ± 0.012 (syst.) ps and decay width difference $\Delta\Gamma = 0.075 \pm 0.035$ (stat.) $\pm$ 0.01 (syst.) ps^{-1} of the B_s^0 meson, with the assumption of no CP violation.

Contents

Introduction

The reason for the dominance of matter over antimatter is one of the most fundamental and unanswered questions in physics today. It is expected that equal amounts of matter and antimatter were produced during the Big Bang; all experimental evidence however shows that matter dominates in the current universe. This contradiction is known as the *baryon asymmetry* problem. Matter particles, called baryons, and anti-baryons (anti-matter particles) annihilate into photons when they meet. An initially equal number of bayrons and anti-baryons would leave almost uniquely photons in the Universe rather than the matter observed in the Universe now. Currently there is no model which satisfactorily explains this puzzle, but there are certain conditions under which it can be understood. These are called the *Sakharov conditions* [1], named after Andrei Sakharov, who first described the necessary physical requirements for a matter excess to occur regardless of a specific mechanism. One of these conditions is the non-conservation of Charge-Parity, referred to as CP violation, which will be described in Chapter 1, the Theoretical Overview of this thesis. The Standard Model of particle physics predicts CP violation in certain types of particle interactions [2]. The amount of CP violation predicted by the Standard Model is, however, too small to explain the baryon asymmetry of the Universe. This has lead to the development of theories beyond the Standard Model to explain potential enhancements to the magnitude of CP violation. Thus, systems where there is a very small expected value of CP violation provide key areas to test the Standard Model and look for larger than predicted values.

This thesis describes a measurement of the amount of CP violation in $B_s^0 \to J/\psi\,\phi$ decays, in terms of the CP violating phase $\beta_s^{J/\psi\phi}$, using 5.2 fb^{-1} integrated luminosity of data collected with the CDF Run-II detector at the Fermilab Tevatron. According to the Standard Model, $\beta_s^{J/\psi\phi}$ is expected to be close to zero [3], thus a large observed quantity for this parameter would be clear evidence for New Physics. In the past three years, published measurements of this parameter by the CDF and DØ collaborations [4, 5] have generated excitement by showing a hints of a disagreement with the Standard Model expectations. A combined result from the two Tevatron collider experiments [6] found a 2.2σ deviation from the expected value. This thesis presents an updated and improved measurement of $\beta_s^{J/\psi\phi}$, with more than twice the data sample of the previous CDF result [7], enhanced signal selection, particle ID and flavour tagging, and the inclusion of an additional $B_s^0 \to J/\psi K^+K^-$ component in the likelihood fit function. In addition to the main results measuring $\beta_s^{J/\psi\phi}$, this analysis also yields the world's most precise measurement of the B_s^0 meson lifetime, τ_s, and decay width difference $\Delta\Gamma_s$.

This document is structured as follows: The development and fundamental components of the Standard Model are described in Chapter 1, followed by a derivation of the time development of $B_s^0 \to J/\psi\,\phi$ decays and a review of the current experimental status of this measurement. Chapter 2 describes the experimental apparatus used to collect the data analysed for this measurement. The process of selecting a high quality data sample is discussed in Chapter 3. The combined probability density function used to form the likelihood to fit the data is laid out in Chapter 4, and its performance is tested in Chapter 5. Finally, the results are presented in Chapters 6 and 7, describing fitted values for some parameters of physical interest with the assumption of no CP violation ($\beta_s^{J/\psi\phi}$=0.0), and the measurement of $\beta_s^{J/\psi\phi}$respectively.

Chapter 1

Theoretical Overview

This chapter describes the development of the Standard Model, introduces some of the areas which are not fully explained by this model, and in this context gives a theoretical overview of the topic of this dissertation. The parameter of interest for the measurement described in this thesis, $\beta_s^{J/\psi\phi}$, is introduced, and a model for the time development of the B_s^0 meson decay channel in which $\beta_s^{J/\psi\phi}$ is measured is developed. In the final section of this chapter, the current experimental status for this measurement is presented.

1.1 Standard Model - historical overview

At the end of the 19th Century, the atom was considered to be a solid object made up of positively charged matter, with negatively charged "corpuscules" arranged non-randomly throughout to balance the charge [8]. This idea was disproved by one of the first particle physics experiments, directed by Ernest Rutherford, which showed by analysing the scattering angles of α particles fired at a thin gold foil that the atom must have a nucleus several orders of magnitude smaller than its total size [9]. Today the prevailing theory describing the structure of matter and the forces governing its interaction is called the Standard Model. Experimentalists are continuously probing the predictions of this theory, as Rutherford did over a century ago; the apparatus has changed significantly, and with it the depth of knowledge of physical phenomena. To date, no experiment has found evidence incompatible with the Standard Model.

The Standard Model (SM) describes fundamental components of matter and their interactions. The elementary particles making up all matter in this model are fermions, which can be divided into two categories: quarks and leptons. There are three generations of fermions, and those of which have been well measured are observed to increase in mass with each generation. Leptons have integer electric charge and the electric charge of quarks is fractional. The charge and mass properties of the twelve Standard Model fermions are summarised in Table 1.1. Fermion interactions are mediated by bosons, particles which can be thought of as force carriers. The Standard model describes three of the four forces: strong, electromagnetic and weak; the gravitational force is not incorporated into this model. The bosons and their properties are listed in Table 1.2

Particle	Charge	Generation					
		I	$mass/c^2$	II	$mass/c^2$	III	$mass/c^2$
Leptons	-1	e	0.511 MeV	μ	$105.658 MeV$	τ	1776.84 MeV
	0	ν_e	< 225 eV	ν_μ	< 0.19 MeV	ν_τ	< 18.2 MeV
Quarks	+2/3	u	$1.5 - 3.3$ MeV	c	$1.27^{+0.07}_{-11}$ MeV	t	171.3 ± 1.63 GeV
	-1/3	d	$3.5 - 6.0$ MeV	s	105^{+25}_{-35} MeV	b	$4.2^{+0.27}_{-0.07}$ GeV

Table 1.1: Properties of fermions, the fundamental particles of matter [10]

1.1.1 Quarks and leptons

The fundamental fermions are spin-1/2 particles, and for each there is an antiparticle with equal mass and lifetime, but opposite electric charge and magnetic moment. The existence of antiparticles was predicted by Dirac in 1931 to solve the problem of potential negative energy solutions to the relativistic relation between energy, momentum and rest mass of a particle. The 20th Century saw the prediction and observation of almost all of the fundamental particles of the Standard Model, a notable exception being the Higgs boson, which is postulated to interact with all massive particles. The search for the Higgs boson is one of the main goals of the LHC [11] at CERN and will be an important test of the standard model.

Normal matter is made up of the lightest, first generation fermions. The electron (e) is a lepton of the first generation, which was discovered in 1896 by J.J Thompson in cathode ray experiments [12], and was an integral component of his early model of the atom referred to at the beginning of this section.

Leptons exist as free particles, and have been experimentally observed as such, however quarks appear to exist only in bound states. This is due to a property which is not shared by the leptons, namely, in addition to carrying electric charge quarks carry colour charge which means that they are acted on by the strong force. The strong force binds quarks in hadron states, currently known hadrons fall into two categories: mesons and baryons. Mesons contain a quark and an antiquark, baryons contain three quarks, both types combine quarks such that the hadron has integer electric charge and neutral colour charge.

The first generation quarks, up and down, are the constituents of neutrons and protons. Other hadrons can be formed from combinations of the heavier quarks, but these are unstable and decay with order of picosecond lifetimes to lighter stable states. Hadrons containing top quarks have not been observed, the t is so much more massive than other elementary particles that its lifetime is several orders of magnitude shorter than even the next most massive quark, the bottom quark, and it decays too fast to hadronise.

Strange particles (now known to contain 2nd generation, strange quarks) were first detected in cloud chamber tracks of cosmic rays in 1946, with the decay of an neutral Kaon ($\bar{s}d$ or $s\bar{d}$) into two charged pions ($u\bar{d}$, $\bar{u}d$) [13]. The "strangeness" of these particles came from the fact that the forces under which they were produced and decayed differed greatly in time scale. It was found that their behaviour could be explained if a new quantum number (similar to charge) was assigned to them; the strange particles were observed as being produced in pairs, and if the pair has opposite "strangeness number" (S=+1, S=-1) then "strangeness" is conserved.

As more hadrons were discovered throughout the middle of the 20th century, physicists tried to fit them into a pattern, as Mendeleev had done with the periodic table of elements. In 1961, Murray Gell-Mann suggested the Eight Fold Way, which placed the known hadrons into octets according to their charge and strangeness quantum numbers [14]; he later explained their behaviour by suggesting that they were in fact composed of even smaller, elementary particles, which he called quarks [12].

The discovery of charmonium, a $c\bar{c}$ resonance, in e^+e^- collisions in 1974 at both SLAC [15] and BNL [16] brought the quark model into good agreement with the lepton sector. There were at that time four known leptons, and with the addition of charm, four quarks. This symmetry was spoiled with the detection of the τ lepton, which introduced a third generation of leptons, that had no equivalent quarks. However, a third generation of quarks was predicted by Kobayashi and Maskawa [17], as described in Section 1.2.1. In 1977 a resonance similar to charmonium, but heavier, was observed in a proton on fixed target experiment at Fermilab [18]. The new particle was a $b\bar{b}$ bound state, "bottomonium", and the bottom quark fitted into the quark model as a third generation, down-type quark. The *top* quark, the up-type counter part to the b quark was expected, but not discovered until 1995, again at Fermilab [19], because it's large mass required a much higher energy to produce than any other quark.

No evidence has yet been produced for further generations beyond the three observed, and the Standard Model explains the need for three generations (see section 1.2.1). However, some theories predict an even heavier fourth generation including t' an b' quarks, which would not be observed at the energies of the current experimental limits [20].

1.1.2 Four forces

The Standard Model describes the interactions of the fundamental fermions, which are governed by the exchange of mediators called bosons. This section summarises the fundamental forces which cause these interactions, the following section will focus in more detail on the weak interaction which is of most significance for the measurement described in this thesis.

Force	Mediator boson			Coupling strength
	Name	Mass	Spin	
Strong	gluon (x8)	0	1	0.1-1
Weak	$W^{\pm}$	$80GeV/c^2$	1	10^{-5}
	Z^0	$91GeV/c^2$	1	
EM	photon	0	1	1/137

Table 1.2: The fundamental forces [21]

Electromagnetic (EM) interactions bind electrons with nuclei in atoms and molecules, and are responsible for intermolecular forces. The mediator boson of the electromagnetic force is the photon, which is massless and interacts with all electrically charged particles. The theory describing electromagnetic interactions is Quantum Electro Dynamics (QED), which is an abelian, gauge invariant field theory with symmetry group U(1). The coupling strength of the EM force is given by the constant α, in terms of

the electric charge e and Plank's constant h:

$$\alpha = \frac{e^2}{4\pi hc} = \frac{1}{137.0360...}. \tag{1.1}$$

The classical EM potential between elementary charges at distance r (the Coulomb potential) is

$$V_{em} = -\frac{\alpha}{r}. \tag{1.2}$$

This shows that the range of the EM force is infinite, but it decreases rapidly with distance.

The Strong force binds neutrons and protons in atomic nuclei, and quarks within hadrons, it is responsible for the confinement of quarks. Gluons, which are massless bosons, mediate this force. Analogous to photons in EM interactions, gluons act on charge, but in strong interactions it is colour charge rather than electric charge. Quantum Chromodynamics (QCD) is the theory of strong interactions, and it contains six charges, called "colours", where colour is simply a label for an internal degree of freedom. It is a non-abelian gauge theory with symmetry group SU(3). Quarks carry one of red, blue or green charge and antiquarks carry the equivalent anticolours. Because of this, colour neutral objects can be formed from two quarks as $q_c\bar{q}'_{\bar{c}}$ (mesons) or three quarks as $q_r q'_b q''_g$ (baryons). An important difference between QED and QCD is that gluons themselves carry colour charge, whereas photons are not charged. Gluons carry one colour and one anticolour, and as they act on colour charge they can interact between themselves. With three colour charges and three anti colours, it would be expected that there could be 3^2 gluons, but one is a colourless singlet state, so there are eight interacting gluons.

Quark confinement is caused by the gluon self-interactions; if an attempt is made to separate two quarks, there comes a point when it would take less energy to form a new $Q\bar{Q}$ pair (a meson) than to continue pulling apart. This can be better understood by looking at the static QCD potential, a convenient classical approximation to the full field theory treatment:

$$V_s = -\frac{4}{3}\frac{\alpha_s}{r} + kr \tag{1.3}$$

The second term in Eqn. 1.3 grows linearly with distance, where k is a constant, which means that the force between them grows stronger as the quarks are pulled apart. The strong force conserves *flavour*, the quark type such as u or b, leading to production of quark anti-quark pairs under this interaction.

Both quarks and leptons can undergo weak interactions, they carry the weak charge g. The strength of the weak interaction is about 3 orders of magnitude smaller than the EM interaction; weak interactions are heavily disfavoured compared to EM and strong interactions. In spite of this, weak interactions are experimentally accessible. For example, neutrinos can interact due to the weak force, but not the EM or strong forces as they carry no electric or colour charge. Weak interactions can also cause flavour change between quarks, which is forbidden under the strong force. The "weakness" of these interactions comes from the fact that they are governed by massive $W^{\pm}$ and Z^0 bosons. The large mass of these vector bosons gives rise to very short range interactions. Figure 1.1.2 shows examples of interactions governed by the EM, strong and weak forces.

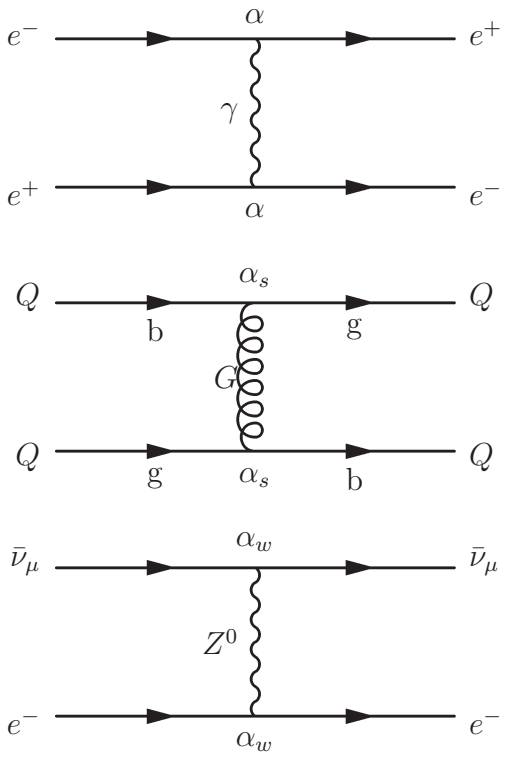

Figure 1.1: Comparison of (top) EM, (centre) strong and (bottom) weak boson exchanges

Electroweak theory unifies the weak and electromagnetic forces. Glashow, Weinberg and Salam suggested in the late 1960s that weak and EM interactions could have the same origins [21]. They postulated a set of four massless bosons as a triplet: $\mathbf{W}_\mu$ belonging to the group SU(2) and a singlet B_μ of the group U(1). In group theory terms, this model is described as SU(2) $\otimes$ U(1). In the Standard Model, the Higgs mechanism [22] gives mass to the weak bosons via spontaneous symmetry breaking. This mechanism introduces an extra field, the *Higgs field*, and due to the shape of the *Higgs potential* interactions with this field result in a mass term in the Lagrangian. This leads to three massive bosons, W^+_μ, W^-_μ and Z^0_μ, with the photon, A_μ remaining massless. The Higgs mechanism also predicts at least one new scalar particle, the *Higgs boson* which has not yet been detected. The weakly interacting massive bosons are linear combinations of the triplet and singlet states in the model. The Standard Model Lagrangian incorporating the electroweak and strong forces, is described as being invariant under the SU(3) $\otimes$ SU(2) $\otimes$ U(1) transformation. Three forces have been discussed in this section. The fourth, gravity, has so far not been integrated into the Standard Model.

1.1.3 Conservation laws

The Standard Model is built on symmetries (invariance properties) and their associated conservation laws. This subsection describes some of the invariance properties of the interactions introduced in Section 1.1.2, which are of importance to the measurement presented in this thesis.

Parity

The parity operator, P, transforms a wave function as

$$P\psi(\mathbf{r}) = \psi(-\mathbf{r}) \tag{1.4}$$

where $\mathbf{r}$ is the spatial position vector. It is clear from Eqn. 1.4 that applying P again brings back the original function, so P is a unitary operator, $P^2 = 1$, and it has eigenvalues $P = \pm 1$.

Until 1956, it was assumed that parity invariance was a fundamental law, but there had been no specific experimental tests of this. Lee and Yang found that there was existing evidence from experiments for parity invariance in strong and EM interactions, but not for weak processes. The status of parity invariance in weak interactions was then tested by C.S. Wu using radioactive Cobalt 60 nuclei [12] in an experiment suggested by Lee and Yang [23]. In these experiments the nuclei had their spins aligned with a magnetic field, and when they underwent beta decay, the direction of the emitted electron was measured. If the electrons were emitted uniformly with respect to the spin direction it would imply that parity was conserved. By applying the parity operator to the system, the spatial coordinates are inverted, but spin is invariant under parity, so a preferential direction for the emitted electrons would be flipped in the parity inverted Universe thus parity would be violated as this would be distinguishable from the original Universe. Wu's experiment showed that the majority of electrons were emitted opposite to the direction of the spin of the nuclei and therefore provided the first evidence of parity violation in the weak interaction.

The violation of parity symmetry in weak interactions is most evident in the neutrino system. As was previously mentioned, neutrinos carry no EM or colour charge, so they only interact via the weak force. Treating neutrinos as massless particles (their mass has been measured to $< 20MeV$), it can be stated that neutrinos with positive helicity (spin aligned with velocity) are right handed, and with negative helicity (spin anti-parallel to velocity) are right handed. The parity operation inverts the handedness of the particle, so if parity was conserved it would be expected that left and right handed neutrinos would behave in the same way. However, when this was tested experimentally, it was discovered that all observed neutrinos are left handed, and all observed antineutrinos are right handed.

Charge conjugation

The action of the charge conjugation operator C is to reverse the sign of the charge and magnetic moment of a particle. Classical electrodynamics is invariant under this operation,for example, Maxwells's equations are equivalent if all charges, potentials and fields reverse their signs. In terms of quantum mechanics, the charge conjugation operator also changes the sign of all internal quantum numbers such as lepton number and strangeness, so it has the effect of converting a particle into its antiparticle.

$$C\,|p\rangle = |\bar{p}\rangle \tag{1.5}$$

The case of neutrinos under the parity operation was discussed previously, but it is also of interest to see what happens under charge conjugation. The C operator trans-

forms a particle into its antiparticle, so a left handed neutrino becomes a left handed antineutrino. It was already explained that only right handed antineutrinos interact, so it is clear that the weak interaction is not invariant under charge conjugation.

Charge-Parity conjugation

The effects on neutrinos under P and C operations discussed in the two previous sections indicate that the weak interaction does not conserve either symmetry. However, when the C and P operators (CP) are applied together a left handed neutrino is transformed into a right handed antineutrino, and this is what is observed experimentally. It was thought that CP symmetry was conserved in weak interactions, until 1964, when neutral Kaons were observed decaying to two different CP eigenstates, therefore violating CP invariance [21]. This has significant implications, which will be discussed in section 1.4.

Charge-Parity-Time

CPT theorem states that under the operation of time reversal, parity, and charge conjugation, there is an exact symmetry for any interaction. This is based on very generalised principles of quantum field theory, it is impossible to construct a quantum field theory where CPT invariance is violated. A consequence of this theorem is that if either time reversal (T) or charge-parity (CP) invariance is violated, the other must also be broken.

Summary

The symmetries and related conservation laws which particularly affect the measurement in this dissertation have been introduced in the preceding sub-sections, and are summarised in Table 1.3.

Conserved quantity	strong	electromagnetic	weak
Parity (P)	yes	yes	no
Charge conjugation (C)	yes	yes	no
CPT	yes	yes	yes
CP (or T)	yes	yes	violation of $O(10^{-5})$

Table 1.3: Summary of conserved quantities under strong, EM and weak interactions in the Standard Model

1.2 Weak Interactions

In order to maintain flavour conservation in the lepton sector, and yet allow cross-generational interactions between quarks, there has to be a significant difference between the actions of the weak force on leptons and quarks. From experimental observations, lepton flavour seems to be largely conserved for charged leptons; in charged current weak interactions ($W^{\pm}$ exchange), transitions such as $s \rightarrow u$ are not uncommon, although they are rarer than transitions within one generation such as $u \rightarrow d$.

In the early 1960s, when this problem was first under consideration, only u, d and s quarks had been discovered. At this time, Cabibbo suggested that the strength of the interaction governing the process $d \rightarrow u + W^-$ carries a factor $\cos(\theta_c)$ and the process $s \rightarrow u + W^-$ a factor $\sin(\theta_c)$ [24]. The relative rates of these two examples implies

that θ_c must be small. This solution was successful in resolving the rates of many interactions, but one significant problem indicated that this was not the full answer: the rate of $K^0 \rightarrow \mu^+\mu^-$ measured experimentally was far below the calculated rate under Cabibbo's theory. In 1970, Glashow, Iliopoulos and Maiani (GIM) proposed that the way to fix this was to introduce a fourth quark [25] (this was four years before the observation of the $J/\psi = c\bar{c}$ resonance). This quark would couple more strongly to s than to d quarks, and importantly, introduced a new diagram in K^0 decays which cancelled the equivalent for the u quark, therefore reducing the expected decay amplitude.

Combining the GIM mechanism with Cabibbo theory leads to the conclusion that the quark states acted on by the charged weak current are not the physical states, but states rotated by the Cabibbo angle θ_c

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{1.6}$$

Eqn. 1.6 shows that the d' and c' states which interact via the charged weak current are linear combinations of the physical d and s states. The discovery of the c quark in 1974 gave strong evidence to this model, and the extension by Kobayashi and Maskawa to three generations allowed for third generation mixing when the b quark was later discovered.

1.2.1 CKM matrix

The proposed extension by Kobayashi and Maskawa [17] is called the CKM matrix, it describes the weak currents between the three generations of quarks. This was suggested before the discovery of the b quark, to incorporate the CP violation observed in the neutral kaon system (introduced in Section 1.1.3). When considering the number of generations in the Standard Model, and in the matrices in Equations 1.6 and 1.7 it should be noted that an $N \times N$ matrix has $N(N-1)/2$ real parameters and $(N-1)(N-2)/2$ non-trivial phases. For the $2D$ case, this means that there is one real parameter, the Cabibbo angle θ_c, for the $3D$ case there are three real angles and one phase. This phase in the the 3×3 case is crucial, because it introduces the possibility of removing invariance under time reversal due to the way it enters the wave function. The principal of CPT-invariance, and CP-violation will be discussed in section 1.4.

The CKM matrix is a 3×3 unitary matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.7}$$

which can be parameterised in terms of three mixing angles and a CP-violating phase. This parameterisation can be represented as

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{232}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{1.8}$$

where $s_{ij} = \sin\theta_{\rm ij}$, $c_{ij} = \cos\theta_{\rm ij}$ and δ is the CP-violating phase. The indices $\{ij\}$ give the generation numbers involved in the interaction. Experimental evidence shows that transitions between generations have very small probability amplitudes, so it is reasonable to use a small angle approximation to simplify the above matrix. The Wolfenstein parameterisation [26] uses this principle to write the matrix as

$$V_{CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4) \tag{1.9}$$

where the following substitutions have been made

$$\begin{aligned} s_{12} &= \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}} \\ s_{23} &= A\lambda^2 = \lambda \left| \frac{|V_{cb}|}{|V_{us}|} \right| \\ s_{13}e^{i\delta} &= V_{ub}^* = A\lambda^3(\rho + i\eta) = \frac{A\lambda^3(\bar{\rho} + i\bar{\eta})\sqrt{1 - A^2\lambda^4}}{\sqrt{1-\lambda^2}[1 - A^2\lambda^4(\bar{\rho} + i\bar{\eta})]} \end{aligned} \tag{1.10}$$

This parameterisation of the CKM matrix illustrates that transitions within one generation (diagonal elements) are favoured over cross-generational transitions which are given by the smaller, off diagonal elements.

Determining the values of the CKM matrix elements gives an important constraint on the Standard Model, and has the potential to reveal evidence of physics beyond the Standard Model, for example in processes which are dominated by loop contributions. The combined experimentally measured values for the CKM matrix elements give the following matrix, [27].

$$V_{CKM} = \begin{pmatrix} 0.97419 \pm 0.00022 & 0.2257 \pm 0.0010 & 0.00359 \pm 0.00016 \\ 0.2256 \pm 0.0010 & 0.97334 \pm 0.00023 & 0.0415^{+0.0010}_{-0.0011} \\ 0.00874^{+0.00026}_{-0.00037} & 0.0407 \pm 0.0010 & 0.9991333^{+0.000044}_{-0.000043} \end{pmatrix} \tag{1.11}$$

1.2.2 Unitarity conditions

The unitarity condition of the CKM matrix can be used to formulate a geometric representation of the CKM parameters. This allows observables to be identified, which can be measured experimentally to test the Standard Model assumptions going into the matrix. This section explains how unitarity triangles can be formed from the relationship between elements of the matrix, and why these are of interest. The unitarity of the CKM matrix gives the relation

$$\sum_{i=1}^{3} |V_{ij}|^2 = 1; \qquad j = 1, 2, 3 \tag{1.12}$$

Information about the CP violating weak phases can be gained from the relations

$$\sum_{i=1}^{3} V_{ji}V_{ki}^{*} = \sum_{i=1}^{3} V_{ij}V_{ik}^{*} = 0; \qquad j,k = 1,2,3 \quad j \neq k \tag{1.13}$$

These form triangles in the complex plane, which are not affected by changes in phase convention (beyond a possible rotation of the triangle). The triangles are of varied shapes but all have the same area, equal to half of the Jarlskog invariant [28]:

$$J = Im[V_{km}^{*}V_{lm}V_{kn}V_{ln}^{*}] = Im[V_{mk}^{*}V_{ml}V_{nk}V_{nl}^{*}] \tag{1.14}$$

This can be derived from the relations in Eqn. 1.13, and is a measure of CP violation in the Standard Model which is independent of phase convention. Experimental measurements combined to estimate J give a value of $O(10^{-5})$ [27].

Focusing on the three triangles which are relevant in neutral meson mixing selects the triangles formed from products of the columns of the CKM matrix

$$V_{ud}^{*}V_{us} + V_{cd}^{*}V_{cs} + V_{td}^{*}V_{ts} = \delta_{ds} = 0 \tag{1.15}$$

which corresponds to CP violation in K meson decay,

$$V_{us}^{*}V_{ub} + V_{cs}^{*}V_{cb} + V_{ts}^{*}V_{tb} = \delta_{sb} = 0 \tag{1.16}$$

corresponding to CP violation in B_s^0 meson decay, and

$$V_{ud}V_{ub}^{*} + V_{ts}V_{cs}^{*} + V_{tb}V_{cb}^{*} = \delta_{db} = 0 \tag{1.17}$$

which describes CP violation in B_d^0 meson decay.

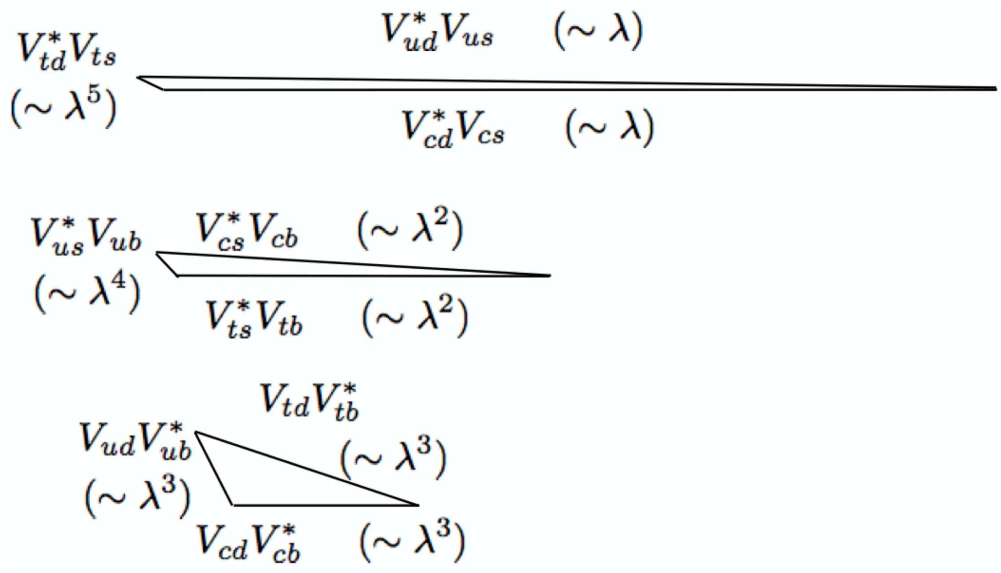

Figure 1.2: Three unitarity triangles of the CKM matrix. Corresponding to Eqn. 1.15 (top). Eqn. 1.16 (middle) and Eqn. 1.17 (bottom).

The triangles parameterised in Eqn.s 1.15-1.17 are shown in figure 1.2. This thesis

describes a measurement of the smallest angle of the B_s^0 mixing triangle, the second in Figure 1.2.

Measurements of all parameters (sides and angles) of these triangles is an over-constraint, and can therefore be used to test the unitarity condition. Experimentally, it is easier to measure the larger interior angles, so the best known triangle is the one corresponding to 1.17, which is generally referred to as "The Unitarity Triangle". This has been well constrained by studies of the $B^0 - \bar{B}^0$ system, and is a good reference point for the measurement in this thesis. This canonical unitarity triangle is shown in its usual representation, normalised form, in Figure 1.3. In this representation, one vertex is set at $(0, 0)$ and the sides are normalised so that a second vertex is at $(1, 0)$ and the third is at $(\bar{\rho}, \bar{\eta})$, which is defined as

$$\begin{aligned} \bar{\rho} &\equiv \left(1 - \frac{\lambda^2}{2}\right)\rho \\ \bar{\eta} &\equiv \left(1 - \frac{\lambda^2}{2}\right)\eta \end{aligned} \tag{1.18}$$

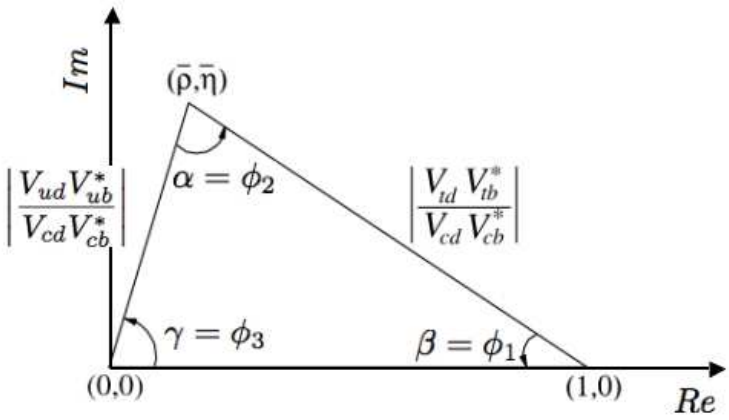

Figure 1.3: Normalised unitarity triangle of the $B^0 - \bar{B}^0$ system

The angles of the canonical triangle shown in Figure 1.3 are defined as

$$\begin{aligned} \beta \equiv \phi_1 &= arg\left(-\frac{V_{cb}^* V_{cd}}{V_{tb}^* V_{td}}\right) \\ \alpha \equiv \phi_2 &= arg\left(-\frac{V_{tb}^* V_{td}}{V_{ub}^* V_{ud}}\right) \\ \gamma \equiv \phi_3 &= arg\left(-\frac{V_{ub}^* V_{ud}}{V_{bb}^* V_{cd}}\right) \end{aligned} \tag{1.19}$$

This thesis focuses on the measurement of the angle β_s, which is the equivalent in the B_s system of the angle β in 1.19. This angle is proportional to the phase of the transition amplitude in $B_s - \bar{B}_s$ mixing in the phase convention chosen in Equations 1.9 and 1.10,

which will be discussed further in Section 1.4.1. It is defined as

$$\beta_s \equiv arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) \tag{1.20}$$

1.3 The production and decay of B hadrons

Before moving to the main topic of this thesis, CP violation in neutral B mesons, it is important to understand the mechanisms through which the studied B hadrons are produced and decay.

1.3.1 B production in $p\bar{p}$ collisions

The data analysed in this thesis are collected in $p\bar{p}$ collisions at the Fermilab Tevatron. The experimental equipment and data acquisition will be described in detail in Chapter 2. In $p\bar{p}$ collisions, b quarks are primarily produced through QCD interactions [29]. To first order, the majority are pair produced in either of

- **Flavour creation**
  This can occur by two gluons from the colliding beam protons undergoing hard scattering to produce a $b\bar{b}$ pair, or the annihilation of a quark anti-quark pair can produce the same outcome. This is also called *direct production* and is shown in the upper two diagrams in Figure 1.4.

- **Gluon splitting**
  After hard scattering, a $b\bar{b}$ pair can be produced from a gluon during fragmentation, shown in the lower left diagram in Figure 1.4.

and a smaller number are produced in

- **Flavour excitation**
  A b quark which is one of the *sea* quarks of the colliding proton or anti proton undergoes hard scattering with another parton from the other colliding beam particle, shown in Figure 1.4.

After production, the b quarks undergo hadronisation. This occurs when the $b\bar{b}$ pair moves apart, feeling the effect of the strong force described in Section 1.1.2; the colour field of a b quark creates a $q\bar{q}$ pair from the vacuum and forms a $b\bar{q}$ meson with the $\bar{q}$. In the case of B_s^0 mesons, the $q\bar{q}$ pair is $s\bar{s}$, a B_s^0 is made up of $\bar{b}s$, and a $\bar{B}_s^0$ is $b\bar{s}$.

1.3.2 B meson lifetimes

Measurement of the B_s lifetime is an important component of this analysis. As part of the simultaneous measurement of the B_s^0 lifetime τ_s, the width difference between the heavy and light B_s^0 eigenstates $\Delta\Gamma_s$, and the CP violating phase $\beta_s^{J/\psi\phi}$, the averaged lifetime of the the heavy and light B_s states is presented. Not only is this an important cross check of the likelihood fitter, being a well constrained value in the Standard Model, it is also a significant measurement in its own right.

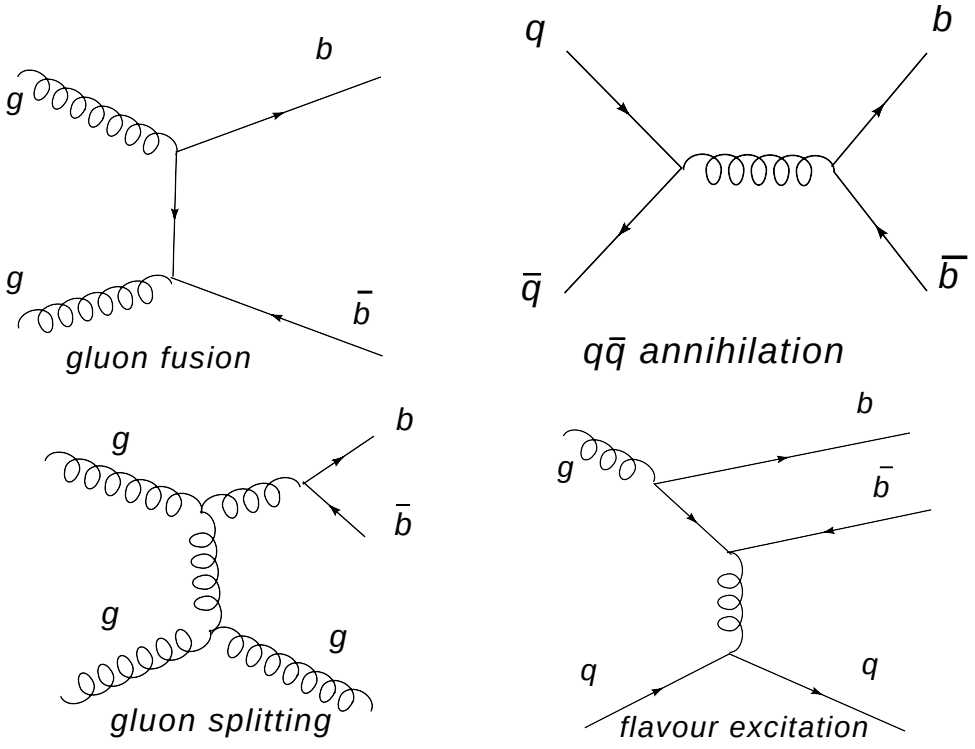

Figure 1.4: Lowest order diagrams for b production processes in $p\bar{p}$ interactions

This section gives a brief introduction to the theoretical concepts of B meson lifetimes, but does not go into further detail as this thesis primarily describes a CP violation measurement.

The spectator quark model is the simplest model of B meson lifetimes. It states that in mesons with unbalanced valance quark masses, the lifetime will be dominated by the decay of the heavier quark. In the case of the B_s meson, the b quark is about $400\times$ heavier than the s quark mass, so the s can be treated as a "spectator"in the decay. Following from the assumption that the lighter quark can be essentially ignored, is the prediction that the lifetimes of all flavours of B mesons, independent of the mass of the lighter quark, should be equal. This is not in agreement with experimental observations, which give a hierarchy of

$$\tau_{B_c} < \tau_{B_s} \approx \tau_{B_d} < \tau_{B_u} \tag{1.21}$$

The favoured decay of the b quark in the Standard Model is to a c quark via a virtual W boson, and it can decay (with far less frequency) to a u quark. The b quark predicted decay width, Γ_b ($\equiv 1/\tau_b$) can be written in terms of the muon decay width [13], which is well known.

$$\Gamma_\mu = \frac{G_F^2}{192\pi^3} m_\mu^5 \Longrightarrow \Gamma_b \sim \frac{G_F^2}{192\pi^3} m_b^5 |V_{cb}|^2 \times (2 \times 3 + 3) \tag{1.22}$$

where G_F is the Fermi coupling constant, $|V_{cb}|$ is the CKM matrix element for the $b \to c$ transition and the final multiplicative factor comes from the extra available decay channels to the b quark: 3 quarks, 3 antiquarks and 3 colours of quarks. From

Equation 1.22 it is possible to construct a relation between the μ and b lifetimes:

$$\tau_b \sim \tau_\mu \left(\frac{m_\mu}{m_b}\right)^5 \frac{1}{9} \frac{1}{|V_{cb}|^2} \tag{1.23}$$

With the central value of the current measurement of $|V_{cb}| = 0.0415$ this predicts a lifetime of around 1.4 ps, which is of the order of the measured B meson lifetimes.

To understand the hierarchy of the B mesons lifetimes it is necessary to look at the decay channels open to the individual mesons according to the light quark content. Three effects which can alter different flavours of B meson lifetimes differently are Pauli Interference (PI), Weak Annihilation (WA) and Weak Exchange (WE). Pauli Interference is specific to B_u mesons, and occurs because both the external W and internal (colour suppressed) W decays have the same final state and can therefore undergo interference. The neutral B mesons have different final states for these two decay types, so there is no interference. Figure 1.5 shows the effect of PI in B_u and B_d (equivalent to B_s) mesons.

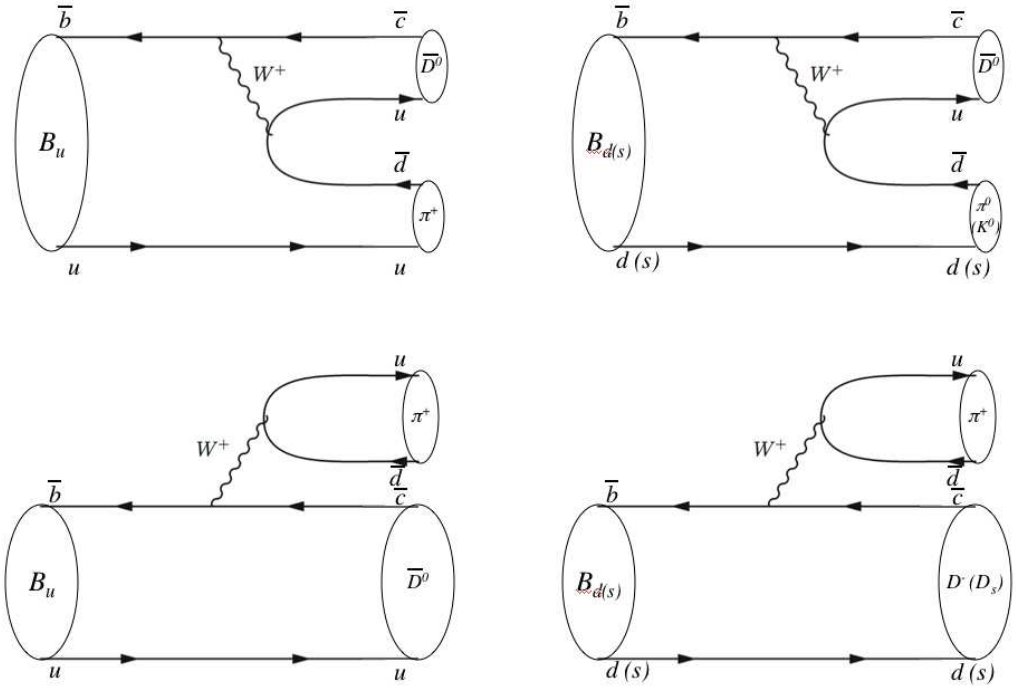

Figure 1.5: Pauli Interference

The effect of Pauli Interference on the charged B mesons is to lengthen the lifetime relative to a decay without interference. Weak Annihilation (charged mesons) and Weak Exchange (neutral mesons), shown in Figure 1.6, have the effect of reducing the lifetime of the meson, by providing an additional decay channel. However, WE is helicity suppressed as the B meson spin of zero means that its q and $\bar{q}$ have opposite helicity. WA is a smaller effect than PI, so the overall contribution to the B^+ lifetime is a net increase in length. Considering these effects brings good agreement with the experimentally observed hierarchy of B meson lifetimes.

To go beyond the prediction of the hierarchy of B meson lifetimes, it is necessary

to look to a QCD based model, Heavy Quark Expansion (HQE) theory. This model expands the decay width, Γ_B in powers of the inverse b quark mass, $1/m_b$ [30]

$$\Gamma_B \sim |V_{ij}|^2 \sum_n c_n(\mu) \left(\frac{1}{m_b}\right)^n \langle H_b|o_b|H_b\rangle \tag{1.24}$$

where c_n are the Wilson coefficients and $|V_{ij}$ is the CKM matrix element for the b decay transition. To first order, this gives the Spectator Quark model described previously. The terms in $1/m_b^2$ separate mesons and baryons, describing the helicity suppression in mesons. The effects of PI and weak interference are brought in with the terms on $1/m_b^3$. This theory predicts lifetime ratios as

$$\frac{\tau(B^+)}{\tau(B^0)} = 1.06 \pm 0.02, \qquad \frac{\tau(B_s^0)}{\tau(B^0)} = 1.00 \pm 0.01, \qquad \frac{\tau(\Lambda_b^0)}{\tau(B^0)} = 0.88 \pm 0.05. \tag{1.25}$$

1.4 CP violation and mixing in neutral mesons

It has been shown in the preceding sections that CP violation arises in weak interactions via the quark mixing matrix. This section will look at the generalised case of CP violation in mixing and decay of neutral mesons (K, D, B_d and B_s), before focusing specifically on the $B_s - \bar{B}_s$ system which is under investigation in this thesis.

One of the most significant unsolved questions in particle physics today, is the discrepancy between the small fraction of CP violation predicted by the Standard Model, and the observed Baryon Asymmetry of the Universe (BAU). As matter and antimatter particles are pair produced in the strong interaction (due to flavour conservation), it would be expected that these would be present in equal quantities, but the present matter-antimatter asymmetry implies that a large source of CP violation should be present to account for the prevalence of matter over antimatter. It is therefore important to constrain the SM predictions with experimental measurements, as a measurement of CP violation significantly over the expected level could be an indicator of New Physics.

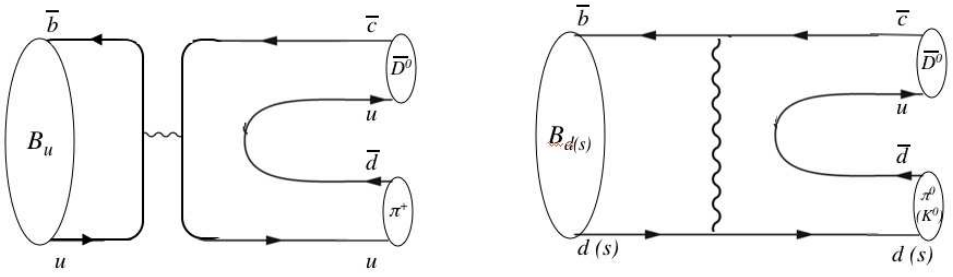

Figure 1.6: Weak annihilation/ exchange

Mixing in neutral mesons

To consider meson mixing, take a neutral meson system with flavour eigenstates X^0 and $\bar{X}^0$. An initial state which is a superposition of these states is

$$|\psi(0)\rangle = a(0)|X^0\rangle + b(0)|\bar{X}^0\rangle \tag{1.26}$$

it evolves in time as

$$|\psi(t)\rangle = a(t)|X^0\rangle + b(t)|\bar{X}^0\rangle \tag{1.27}$$

The time evolution of this system can be calculated using an approximate formalism developed by Weisskopf and Wigner [31]. This uses an effective Hamiltonian $\mathbf{H}$, which can be written in terms of $\mathbf{M}$ and $\mathbf{\Gamma}$

$$\mathbf{H} = \mathbf{M} - \frac{\mathbf{i}}{\mathbf{2}}\mathbf{\Gamma} = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{22} \end{pmatrix} - \frac{\mathbf{i}}{\mathbf{2}} \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{22} \end{pmatrix} \tag{1.28}$$

The diagonal elements of $\mathbf{M}$ and $\mathbf{\Gamma}$ are flavour conserving, the off-diagonal elements change the flavour of quarks being acted on. Eigenvectors of $\mathbf{H}$ are mass eigenstates of the X^0 system, the heavy and light states X_H and X_L

$$\begin{aligned} |X_L\rangle &= p|X^0\rangle + q|\bar{X}^0\rangle \\ |X_H\rangle &= p|X^0\rangle - q|\bar{X}^0\rangle \end{aligned} \tag{1.29}$$

which is normalised with $|q|^2 + |p|^2 = 1$.

The real and imaginary parts of the eigenvalues $e_{L,H}$ of $\mathbf{H}$ which correspond to $|X_{L,H}\rangle$ give the mass and decay width of the mass states. The mass and width differences between the light and heavy states are given by:

$$\begin{aligned} \Delta m &\equiv m_H - m_L = \Re(e_H - e_L) \\ \Delta\Gamma &\equiv \Gamma_H - \Gamma_L = -2\Im(e_H - e_L) \end{aligned} \tag{1.30}$$

Solving to find the eigenvalues gives

$$e_{H,L} = \left(m - \frac{i}{2}\Gamma\right) \pm \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)} \tag{1.31}$$

and

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^* - (i/2)\Gamma_{12}^*}{M_{12} - (i/2)\Gamma_{12}} \tag{1.32}$$

Time development

The amplitudes of decay of the meson X ($\bar{X}$) to the final state f ($\bar{f}$) are

$$A_f = \langle f|\mathcal{H}|X\rangle, \qquad \bar{A}_f = \langle f|\mathcal{H}|\bar{X}\rangle \tag{1.33}$$

$$A_{\bar{f}} = \langle \bar{f}|\mathcal{H}|X\rangle, \qquad \bar{A}_{\bar{f}} = \langle \bar{f}|\mathcal{H}|\bar{X}\rangle \tag{1.34}$$

where $\mathcal{H}$ is the weak interaction Hamiltonian. If CP is conserved, the amplitudes of A_f and $\bar{A}_{\bar{f}}$ should be identical.

With relatively small $\Delta\Gamma/\Gamma$ it is not possible to experimentally separate the mass eigenstates in time evolution. Consider the time evolution of an initially pure state $|X^0\rangle$ or $|\bar{X}^0\rangle$ after time t:

$$|X(t)\rangle = g_+(t)|X^0\rangle - \frac{q}{p}g_-(t)|\bar{X}^0\rangle$$
$$|\bar{X}(t)\rangle = g_+(t)|\bar{X}^0\rangle - \frac{q}{p}g_-(t)|\bar{X}^0\rangle$$

where

$$g_\pm = \frac{1}{2}\left(e^{-im_H t-\frac{1}{2}\Gamma_H} \pm e^{-im_L t-\frac{1}{2}\Gamma_L}\right) \tag{1.35}$$

Combining Equations 1.33- 1.35 leads to the decay rates $|\langle f|X^0(t)\rangle|^2$, $|\langle f|\bar{X}^0(t)\rangle|^2$

$$\begin{aligned}|\langle f|X^0(t)\rangle|^2 & \\ &= (|A_f|^2 + |(q/p)\bar{A}_f|^2)\cosh(\Delta\Gamma t/2) \\ &+ (|A_f|^2 - |(q/p)\bar{A}_f|^2)\cos(\Delta M t) \\ &+ 2\Re((q/p)A_f^*\bar{A}_f)\sinh(\Delta\Gamma t/2) \\ &- 2\Im((q/p)A_f^*\bar{A}_f)\sin(\Delta m t)\,.\,e^{-\Gamma t}\end{aligned} \tag{1.36}$$

$$\begin{aligned}|\langle f|\bar{X}^0(t)\rangle|^2 & \\ &= (|(p/q)A_f|^2 + |\bar{A}_f|^2)\cosh(\Delta\Gamma t/2) \\ &- (|(p/q)A_f|^2 - |\bar{A}_f|^2)\cos(\Delta M t) \\ &+ 2\Re((q/p)A_f\bar{A}_f^*)\sinh(\Delta\Gamma t/2) \\ &- 2\Im((q/p)A_f\bar{A}_f^*)\sin(\Delta m t)\,.\,e^{-\Gamma t}\end{aligned} \tag{1.37}$$

and their complex conjugates which correspond to substituting the amplitudes from Eqn. 1.34 for those of Eqn. 1.33 in the above equations. The specific case of the time development of the B_s^0 meson system will be derived in section 1.5.

Categories of CP violation

CP violation occurs in meson decays through three effects [31], which are explained below focussing on their relevance in the $B_s - \bar{B}_s$ system:

1. **CP violation in decay**
   Also called *direct CP violation*, this occurs when the amplitude of decay to a final state is not the same as the amplitude of the CP conjugate of the initial state decaying to the CP conjugate of the final state:

   $$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| \neq 1 \tag{1.38}$$

   Direct CP violation is the only possible CP violating effect in charged meson decays, which cannot undergo mixing. In the $B_s^0 \to J/\psi\,\phi$ channel, the Standard Model CP-violating weak phase in decay is strongly Cabibbo suppressed by a factor of λ^2 [32] (where λ is the Wolfenstein parameter of the CKM matrix -

Section 1.2.1). This means that to a high order approximation, it is possible to make the assumption that there is no direct CP violation in $B_s^0 \to J/\psi\,\phi$ decays, and treat the amplitudes as $|A_f| = |\bar{A}_{\bar{f}}|$.

2. CP **(and T) violation in mixing**
This type of CP violation is defined as

$$\left|\frac{q}{p}\right| \neq 1 \tag{1.39}$$

In the B_s meson system, the CKM model predicts [33] $|q/p| = 1 + \mathcal{O}(10^{-3})$. In semileptonic B_s decays this leads to a charge asymmetry in the decay products, but in $B_s^0 \to J/\psi\,\phi$, the channel used in this measurement, the factor $|q/p|$ is not isolated, therefore CP violation in mixing is not directly measured in this analysis.

3. CP **violation due to interference between decays with and without mixing**

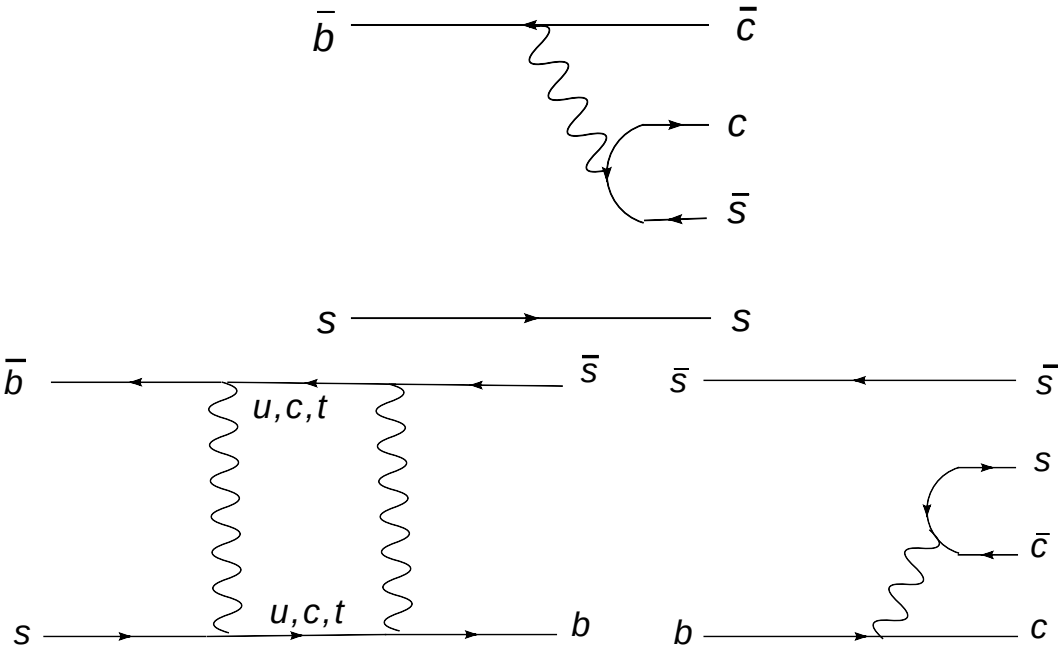

Figure 1.7: [upper] $B_s^0 \to J/\psi\,\phi$ decay without mixing; [lower] $B_s^0 \to J/\psi\,\phi$ decay via mixing

In decays to a common CP eigenstate, CP violation can occur independently of the previous two cases, because of interference between decays without mixing: $X^0 \to f$ and via mixing: $X^0 \to \bar{X}^0 \to f$ to the same final state. This is defined by

$$\Im(\lambda_f) \neq 0 \tag{1.40}$$

where

$$\lambda_f \equiv \frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_f} \tag{1.41}$$

This type of CP violation is possible in the $B_s^0 \to J/\psi\,\phi$ mode because $J/\psi\phi$ is a common final state for B_s^0 and $\bar{B}_s^0$.

1.4.1 CP violation measurements in neutral meson systems

$K^0 - \bar{K}^0$ mixing

The K^0 system was the first neutral meson system to be comprehensively studied. The K^0 contains an anti-strange quark and therefore carries a strangeness of $S = +1$. The antiparticle neutral kaon state, $\bar{K}^0$ differs by strangeness $\Delta S = 2$ from the particle state; applying the CP operator to $|\bar{s}d\rangle$ gives $|s\bar{d}\rangle$. The existence of the two discrete states was shown by Gell-Mann and Pais in 1954 [34], under the assumption that CP invariance was conserved. Mixing between the particle and antiparticle states was observed by selecting a pure K^0 beam, which, after a few metres was found to combine both K^0 and $\bar{K}^0$ states. The weak eigenstates of the neutral Kaon system are

$$\begin{aligned} |K_S >&= \frac{1}{\sqrt{2}}\left(|K^0\rangle + |\bar{K}^0\rangle\right) \\ |K_L >&= \frac{1}{\sqrt{2}}\left(|K^0\rangle - |\bar{K}^0\rangle\right) \end{aligned} \tag{1.42}$$

where K_S is a short lived, state and K_L is long lived. The CP of these states can be determined by applying the CP operator

$$\begin{aligned} \mathbf{CP}|K_S\rangle &= \frac{1}{\sqrt{2}}\left(|\bar{K}^0\rangle + |K^0\rangle\right) = +|K_S\rangle \\ \mathbf{CP}|K_L\rangle &= \frac{1}{\sqrt{2}}\left(|\bar{K}^0\rangle - |K^0\rangle\right) = -|K_L\rangle \end{aligned} \tag{1.43}$$

The neutral Kaon system was also the source of the first observation of CP violation. After the discovery of parity violation, it was thought that the combined effect of CP conjugation would be a symmetry of the weak interaction. However, this was disproved in 1964, when it was found that the Kaon eigenstate labeled as CP odd, K_L, could decay to a CP even final state with a branching ratio of $\mathcal{O}(10^{-3})$. The CP even state, K_S, decays to a two pion final state which is CP even ($\mathbf{CP}|\pi\pi\rangle = +|\pi\pi\rangle$), the favoured decay mode of the K_L is to three pions ($\mathbf{CP}|\pi\pi\pi\rangle = -|\pi\pi\pi\rangle$), as expected for a CP odd state. Cronin and Fitch discovered [35] that the K_L state could also decay to $|\pi\pi\rangle$, with a rate of

$$\frac{\Gamma(K_L \to \pi^+\pi^-)}{\Gamma(K_S \to \pi^+\pi^-)} = (2.0 \pm 0.4) \times 10^{-3} \tag{1.44}$$

This implies that the weak eigenstates are not exact CP eigenstates, and that there is a small fraction of *indirect* CP violation in this system.

$D^0 - \bar{D}^0$ mixing

Evidence for mixing in the $D^0 - \bar{D}^0$ system has only recently been observed by the BaBar [36] and CDF [37] collaborations. This is a very small effect, and its observation relies on the highly accurate flavour tagging of the initial state which is achievable from knowledge of the D^0 meson production process. Current investigations focus on a search for New Physics in this system rather than a constraint of Standard

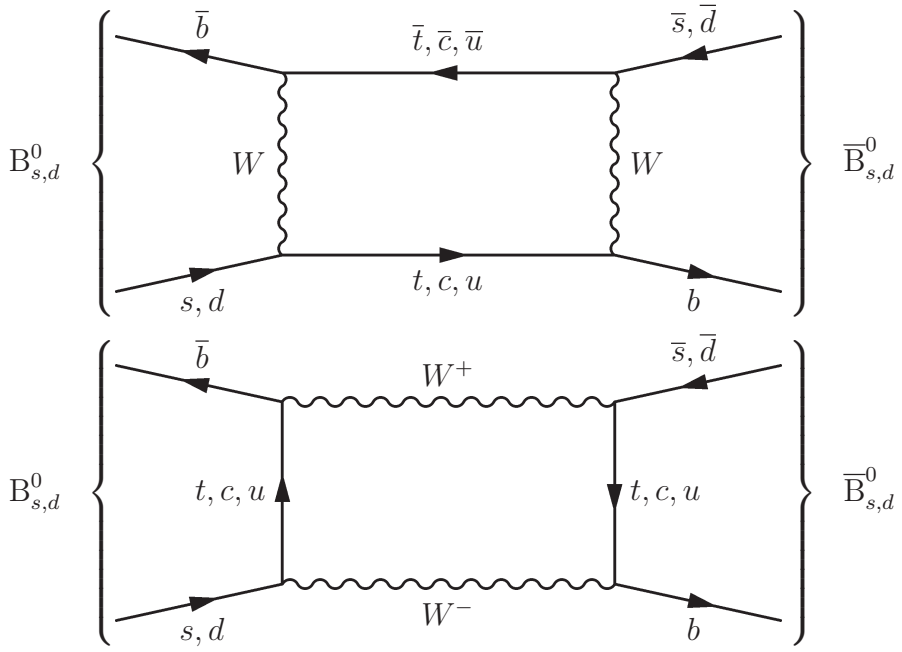

Figure 1.8: Weak mixing box diagrams in the neutral B meson system.

Model CKM parameters. CP violation in the neutral D system is predicted to be negligibly small in the Standard Model [38], so a significant observed rate would be a strong signal of New Physics. The main decay modes used in the present searches are $D \to K^+K^-$ and $D \to K^\pm\pi^\mp$. There is so far no evidence for *indirect* CP violation in this system, with the combined measurement [39] (assuming no direct CP violation) of

$$1 - \left|\frac{q}{p}\right| = +0.06 \pm 0.14 \tag{1.45}$$

Neutral B meson mixing

Observation of CP violation in the neutral B meson system came later than in the Kaon system, because the more massive particles take a higher energy collider to produce. There is also the challenge in that the lifetime difference between the weak eigenstates of the $B^0 - \bar{B}^0$ system is almost negligible. However, this system has already produced some very significant results to constrain the CKM parameters, and shows evidence of much larger CP violation than the neutral kaon system.

Mixing in B^0 and B_s mesons occurs through W boson exchange, shown in the weak mixing box diagrams in Figure 1.8.

For B^0 mesons ($|b\bar{d}\rangle, |\bar{b}d\rangle$), predictions from studies of CP violation in semileptonic B decays can be used to estimate the level of CP violation in $B^0 - \bar{B}^0$ mixing through the CP asymmetry in the semileptonic decay rates, $\mathcal{A}_{SL}$:

$$\mathcal{A}_{SL} = -\left|\frac{\Gamma_{12}}{M_{12}}\right| \sin(\phi_M - \phi_\Gamma). \tag{1.46}$$

From current measurements [31]

$$\mathcal{A}_{SL} = (-0.4 \pm 5.6) \times 10^{-3}. \tag{1.47}$$

The measure of CP violation in meson mixing, $|q/p$, was described in section 1.4, and can be related to $\mathcal{A}_{SL}$ by

$$\left|\frac{q}{p}\right| \approx 1 - \frac{\mathcal{A}_{SL}}{2} = 1.0002 \pm 0.0028. \tag{1.48}$$

This sub-percentage effect is treated as negligible in B^0 mixing measurements.

The first measurements of mixing in the B^0 system were time integrated analyses, first published in 1987 by UA1 [40], and relied on counting same-sign and opposite-sign lepton pairs from semileptonic decays of produced $b\bar{b}$ pairs. Increased sensitivity could be obtained in time dependent measurements, directly observing the oscillation frequencies, and many high energy collider and B factory experiments have produced such results. The combined result of these measurements of the mixing frequency in the B^0 system is [41]

$$\Delta m_d = 0.507 \pm 0.003 \text{ (stat.)} \ \pm 0.003 \text{ (syst.)} ps^{-1} \tag{1.49}$$

The B_s system is more difficult for mixing measurements. The mixing frequency,Δm_s is significantly higher than Δm_d, and the decay width difference, $\Delta\Gamma_s$ is non-negligible, whereas in the B^0 system it can be treated as $\Delta\Gamma_d \approx 0$. The first significant (5σ) observation of mixing in the B_s system was by CDF in 2006 [42], and this remains the world's best measurement to date:

$$\Delta m_s = 17.77 \pm 0.1 \text{ (stat.)} \ \pm 0.07 \text{ (syst.)} ps^{-1} \tag{1.50}$$

Δm_s is an important parameter in the CP violation measurement described in this thesis, and this precision measurement is a direct input to the likelihood fit.

The CP asymmetry measured in semileptonic B_s decays is [43]

$$\mathcal{A}^s_{SL} = -0.003 \pm 0.0101 \Rightarrow \left|\frac{q}{p}\right|_s = 1.0015 \pm 0.0051. \tag{1.51}$$

which gives weight to the assumption from the SM prediction that CP violation in mixing in the B_s system is minimal.

1.5 Phenomenology of $B^0_s \to J/\psi\,\phi$ decays

This section introduces the phenomenology of the $B^0_s \to J/\psi\,\phi$ decay and the experimental method used to measure the CP violating parameter $\beta_s^{J/\psi\phi}$ in this channel. Firstly, the generalised case of time development of a neutral meson system will be adjusted for the specific case of B^0_s decays. Next, the concept of *flavour tagging* is introduced in Section 1.5.1, which is essential for separating the decays of B^0_s and $\bar{B}^0_s$ mesons. The final state in the $B^0_s \to J/\psi\,\phi$ channel is an admixture of CP odd and even states, which must be separated in order to study the effect of CP violation. The method used to do this is described in Section 1.5.2.

The effective Hamiltonian describing neutral meson time development was introduced in Equation 1.28. Here, the elements M_{12} and Γ_{12} are of interest. Experimental evidence summarised in section 1.4.1 showed that $\Delta\Gamma_s \gg \Delta m_s$, which implies that $|\Gamma_{12}| \ll |M_{12}|$. Using this relation, it is possible to write $|q/p|, \Delta m_s$ and $\Delta\Gamma_s$ in terms of Γ_{12} and M_{12}:

$$\Delta m_s = 2|M_{12}|, \qquad \Delta\Gamma_s = 2|\Gamma_{12}|\cos\phi_s \tag{1.52}$$

and from Equation 1.32, with the above approximation

$$\frac{q}{p} = -\frac{M_{12}^*}{|M_{12}|} = -\frac{V_{ts}V_{tb}^*}{V_{ts}^*V_{tb}} \tag{1.53}$$

where

$$\phi_s = \phi_M - \phi_\Gamma = \arg\left(-\frac{M_{12}}{\Gamma_{12}}\right) \tag{1.54}$$

and higher order terms in Γ_{12}/M_{12} are ignored. In the Standard Model, ϕ_s is predicted to be very small [44]:

$$\phi_s^{SM} = 0.0041 \pm 0.0008 \tag{1.55}$$

CP violation in this channel arises due to interference between decays with and without net mixing to the common final state (as defined in Section 1.4). The convention independent measure of this type of CP violation is given in Equation 1.41 as $\lambda_i = \frac{q}{p}\frac{\bar{A}_i}{A_i}$. The λ_i quantities for a set of final states $\{i\}$ in the Standard Model can be written in terms of CKM matrix elements, in some cases with the addition of strong phases. The quantity $|q/p|$ was defined in this way in Equation 1.53, and similarly, from the fact that quark relative weak interaction strengths are defined by the CKM elements, the ratio of decay amplitudes can be written as

$$\frac{\bar{A}_f}{A_f} = -\eta_{CP}\frac{V_{cs}^*V_{cb}}{V_{cs}V_{cb}^*}. \tag{1.56}$$

Combining equations 1.53 and 1.56 gives the SM prediction for λ_i

$$\lambda_i = \eta_i^{CP}\lambda = \eta_i^{CP}\frac{V_{ts}V_{tb}^*}{V_{ts}^*V_{tb}}\frac{V_{cs}^*V_{cb}}{V_{cs}V_{cb}} = \eta_i^{CP}e^{i2\beta_s} \tag{1.57}$$

where $\eta_i^{CP} = \pm 1$, with the positive case for CP even eigenstates and negative for CP odd states. The phase β_s is the small angle of the CKM unitarity triangle for the B_s^0 system described in Section 1.2.2, and is predicted in the Standard Model to be

$$2\beta_s \equiv 2\arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) = 0.037 \pm 0.002 \tag{1.58}$$

These definitions can be related to the description of the parameters of the effective Hamiltonian in terms of the phase ϕ earlier in this section. Both ϕ_s and β_s are predicted to be small in the Standard Model. If there is New Physics (NP) present in mixing, it will contribute to both phases, and they can be written as combinations of the SM and NP contributions according to the chosen phase conventions

$$\phi_s = \phi_s^{SM} + \phi_s^{NP}, \qquad \text{and} \qquad 2\beta_s = 2\beta_s^{\mathrm{SM}} - \phi_s^{\mathrm{NP}}. \tag{1.59}$$

With current experimental resolution and statistics, the SM contributions to both of these phases can be treated as negligibly small, and the NP phase would be expected to dominate. In this case, the measured phases can be related as

$$2\beta_s \approx -\phi_s. \tag{1.60}$$

An enhancement due to NP could come from additional massive particles such as extra generations of quarks [20] entering the B mixing box diagrams (Figure 1.8).

1.5.1 Flavour tagging principles

Knowledge of the B meson flavour at production, is a key component of this analysis or any involving CP-violation in neutral B systems. The process used in distinguishing between initial B_s^0 and $\bar{B}_s^0$ is called *flavour tagging*. Properties of the $p\bar{p} \to b\bar{b}$ production process, and b quark hadronisation and fragmentation are used in two flavour tagging algorithms, Same Side and Opposite Side Taggers (SST, OST). Figure 1.9 illustrates the kinematics exploited by the two tagger types in the production and decay of a $b\bar{b}$ pair.

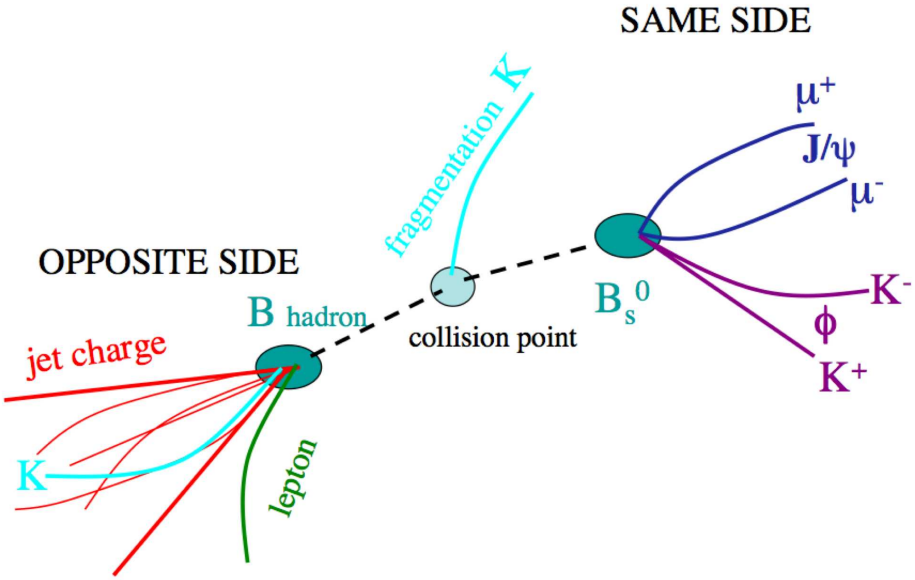

Figure 1.9: Production and decay of a $b\bar{b}$ pair, showing the components of an event used for Same Side and Opposite Side flavour tagging. The *Same Side* part refers to the decay of the candidate B meson, the *Opposite Side* consists of the decay of a B meson containing the pair produced partner of the b ($\bar{b}$) quark in the candidate B meson.

The SST uses the fragmentation tracks of the B meson of interest (the *candidate* B meson) to determine its flavour. In the case of a candidate B_s^0 meson, this means identifying the flavour (q or $\bar{q}$) of the s quark in a kaon produced alongside the B_s^0 ($\bar{B}_s^0$). As this is the pair produced partner of the s ($\bar{s}$) in the candidate B meson, its flavour tags the B meson as containing the opposite $\bar{s}$ (s) quark at production and thus identifies the B meson flavour when it was produced. For charged kaons, the quark content is readily identifiable. The OST takes advantage of the pair production of b quarks and uses information from the hadronisation and decay of the other b quark from the pair (the *opposite side* B). If the opposite side B meson decays to charged leptons, the

charge of its b quark can be identified, which tags the flavour of the candidate B meson at production because it is known to contain the opposite b or $\bar{b}$. For example, an opposite side B decaying to a negative lepton implies that the candidate B contained a $\bar{b}$ at the time of production. Similarly, in hadronic decays of the opposite side B meson, the charge of a b jet identified as coming from this decay can tag the quark content of the decaying meson, from which the flavour of the candidate B can be inferred.

The algorithms used for B flavour tagging in this analysis are described in Chapter 3. The use of flavour tagging to determine the flavour of a B_s^0 or$\bar{B}_s^0$ meson at production means that decays of the particle and antiparticle states can be followed separately. This will be utilised in the development of the equations in the following subsections.

1.5.2 Angular analysis of $B_s^0 \to J/\psi\,\phi$

Eigenvectors of **H** were defined for a generic neutral meson in Equation 1.29, for the B_s^0 system, the mass eigenstates are:

$$|B_s^H\rangle = p|B_s^0\rangle - q|B_s^0\rangle \text{ and } |B_s^L\rangle = p|B_s^0\rangle + q|B_s^0\rangle \tag{1.61}$$

With negligible direct CP violation, such is the case for the $B_s^0 \to J/\psi\,\phi$ channel, these can be treated as CP eiegnstates. The heavy, long-lived mass eigenstate, $|B_s^H\rangle$ is CP odd, following the arguement set out for the Kaon system in Section 1.4.1, and the light, short lived state, $|B_s^L\rangle$ is CP even.

The decay $B_s^0 \to J/\psi\,\phi$ is a pseudo-scalar (B_s) decaying to a two vector final state ($J/\psi, \phi$). The B_s has spin 0, the vector particles each have spin 1 and are both C odd. This means that the CP of the final state can be determined from only the P of this state. Conservation of momentum for this system leads to three possible relative angular momentum values, $L = 0, 1$ or 2, in order for the vector particles to respect conservation of total angular momentum, $J \equiv S + L$.

There is a corresponding decay amplitude for each relative angular momentum value. Mathematically, it is more straight forward to work with polarisation states of the vector particles, so the *transversity basis* is used. This allows separation of CP odd and even components of the decay by using angular distributions of their decay products. There are three potential polarisation states of the final state vector particles,

- $|\mathcal{P}_0\rangle$: spins are polarised longitudinally with respect to their momentum and parallel to each other, CP even
- $|\mathcal{P}_\perp\rangle$: spins are polarised transversely with respect to their momentum and perpendicular to each other, CP odd
- $|\mathcal{P}_\parallel\rangle$: spins are polarised transversely with respect to their momentum and parallel to each other, CP even

The CP odd and even B_s states described earlier in this section decay to the respective CP states listed. The CP odd state corresponds to the $L = 1$ angular momentum state, the $L = 0, 2$ states are linear combinations of $|\mathcal{P}_0\rangle$ and $|\mathcal{P}_\parallel\rangle$.

The transversity basis is defined in terms of angles in the rest frames of the final state particles, as shown in Figure 1.10. The $x-$axis is determined by the direction

of the ϕ meson in the J/ψ rest frame, the K^+K^- plane defines the xy plane, where $p_y(K^+) > 0$. The three angles are defined as:

- θ is the angle between $p(\mu^+)$ and the xy plane, in the J/ψ meson rest frame
- ϕ is the angle between the $x-$ axis and $p_{xy}(\mu^+)$, the projection of the μ^+ momentum in the xy plane, in the J/ψ meson rest frame
- ψ is the angle between $p(K^+)$ and $-p(J/\psi)$ in the ϕ meson rest frame

The predicted angular distributions of the CP odd and even final states are fully developed in [45]. Here it suffices to define the unit vector in the direction of the positive lepton in the J/ψ rest frame

$$\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{1.62}$$

and the amplitudes to decay into each of the described CP states. The decay amplitudes are $\mathcal{A}_i$ and $\bar{\mathcal{A}}_i$, the amplitudes to the final states $i = \{\|, \perp, 0\}$

$$\mathcal{A}_i = \langle B_s|H|\mathcal{P}_i\rangle, \qquad \bar{\mathcal{A}}_i = \langle \bar{B}_s|H|\mathcal{P}_i\rangle \tag{1.63}$$

which are normalised such that $|\mathcal{A}_0|^2 + |\mathcal{A}_\||^2 + |\mathcal{A}_\perp|^2 = 1$. The complex vector $\mathbf{A}$ is defined as

$$\mathbf{A} = \left(\mathcal{A}_0\cos\psi, -\frac{\mathcal{A}_\|\sin\psi}{\sqrt{2}}, i\frac{\mathcal{A}_\perp\sin\psi}{\sqrt{2}}\right). \tag{1.64}$$

The probability density governing the angular distributions is defined in terms of Equations 1.64 and 1.62

$$P(\theta, \phi, \psi) = \frac{9}{16\pi}|\mathbf{A} \times \hat{n}|^2 \tag{1.65}$$

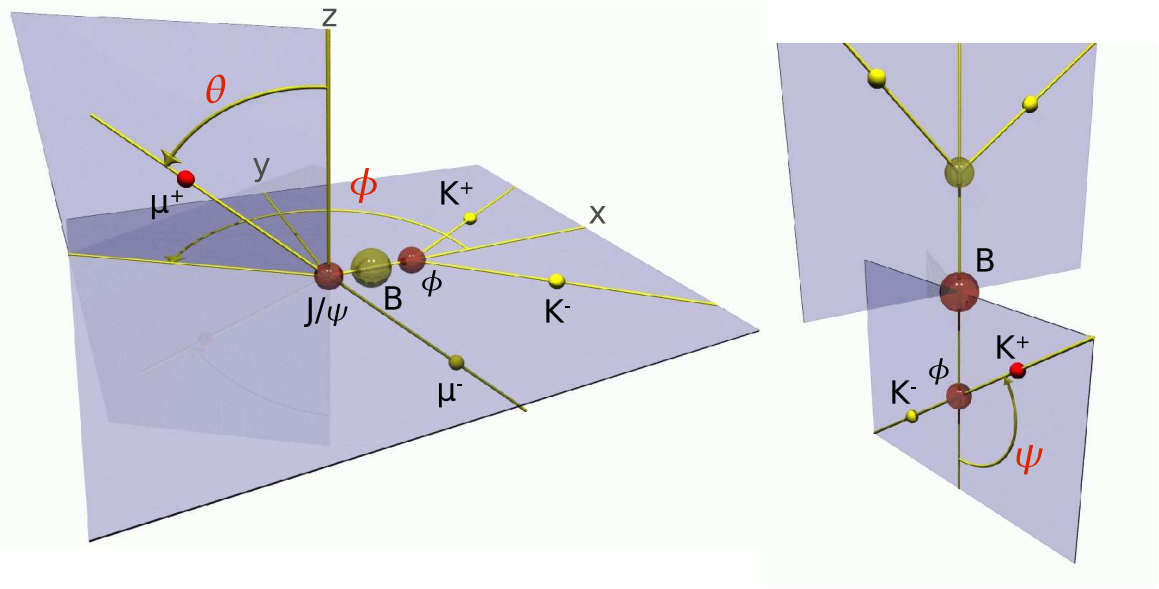

Figure 1.10: The transversity basis: defined in terms of the angles θ and ϕ in the j/ψ rest frame (left) and ψ in the ϕ meson rest frame (right).

which is normalised such that

$$\iiint \frac{9}{16\pi}|\mathbf{A}\times\hat{n}|^2 \sin\theta d\theta \sin\psi d\psi = 1. \tag{1.66}$$

Time development of B_s^0 and $\bar{B}_s^0$ decay rates to a single CP eigenstate

Combining equations 1.35 and 1.35, and re-writing in terms of B_s^0 mesons, gives

$$\begin{aligned} |B_s^0(t)\rangle &= e^{-imt}e^{-\Gamma t/2}\left[E_+(t)|B_s^0\rangle + \frac{q}{p}E_-(t)|\bar{B}_s^0\rangle\right], \\ |\bar{B}_s^0(t)\rangle &= e^{-imt}e^{-\Gamma t/2}\left[\frac{p}{q}E_-(t)|B_s^0\rangle + E_+(t)|\bar{B}_s^0\rangle\right] \end{aligned} \tag{1.67}$$

where $E_\pm$ are defined as

$$E_\pm = \frac{1}{2}\left[e^{+(\frac{\Delta\Gamma}{4}+i\frac{\Delta m}{2})t} \pm e^{-(\frac{\Delta\Gamma}{4}+i\frac{\Delta m}{2})t}\right]. \tag{1.68}$$

To study the time development of the $B_s^0 \to J/\psi\,\phi$ decay, first a simplified case of decay to a single final state, a CP eigenstate, will be considered. Again, following the steps laid out in Section 1.4, the time dependent decay to the common final state $|f\rangle$ can be written as

$$\begin{aligned} \langle f|B_s^0(t)\rangle &= e^{-imt}e^{-\Gamma t/2}\left[E_+(t)|B_s^0\rangle + \frac{q}{p}\frac{\bar{A}}{A}E_-(t)|\bar{B}_s^0\rangle\right]A \\ \langle f|\bar{B}_s^0(t)\rangle &= e^{-imt}e^{-\Gamma t/2}\left[\frac{p}{q}\frac{A}{\bar{A}}E_-(t)|B_s^0\rangle + E_+(t)|\bar{B}_s^0\rangle\right]\bar{A}. \end{aligned} \tag{1.69}$$

Substituting Equation 1.41 into these definitions gives

$$\begin{aligned} \langle f|B_s^0(t)\rangle &= e^{-imt}e^{-\Gamma t/2}\left[E_+(t)|B_s^0\rangle + \lambda E_-(t)|\bar{B}_s^0\rangle\right]A \\ \langle f|\bar{B}_s^0(t)\rangle &= e^{-imt}e^{-\Gamma t/2}\left[\lambda^{-1}E_-(t)|B_s^0\rangle + E_+(t)|\bar{B}_s^0\rangle\right]\bar{A}. \end{aligned} \tag{1.70}$$

These can be converted into time-dependent decay rates for the B_s^0 and $\bar{B}_s^0$ mesons

$$\begin{aligned} R &= \langle f|B_s^0(t)\rangle^2 = e^{-\Gamma t}\left\{\cosh\frac{\Delta\Gamma}{2}t - \Re\left\{\lambda^*\left(\sinh\frac{\Delta\Gamma}{2}t + i\sin\Delta m t\right)\right\}\right\}, \\ \bar{R} &= \langle f|\bar{B}_s^0(t)\rangle^2 = e^{-\Gamma t}\left\{\cosh\frac{\Delta\Gamma}{2}t - \Re\left\{\lambda\left(\sinh\frac{\Delta\Gamma}{2}t + i\sin\Delta m t\right)\right\}\right\}, \end{aligned} \tag{1.71}$$

It is necessary to normalise these rates, when considering a probability distribution as will be done in the maximum likelihood developed in Chapter 4. The normalisation

can be calculated for the sum of the rates in Equation 1.71, and can be written as

$$\begin{aligned}\int (R+\bar{R})dt &= \Gamma_H^{-1}+\Gamma_L^{-1}+\Re(\lambda)\left(\Gamma_L^{-1}-\Gamma_H^{-1}\right)\\ &= \tau_H+\tau_L+\Re(\lambda)(\tau_L-\tau_H).\end{aligned} \tag{1.72}$$

This normalisation factor can be applied to the rates in Equation 1.71 to give

$$\begin{aligned}R &= \frac{e^{-\Gamma t}}{\tau_H+\tau_L+\Re(\lambda)(\tau_L-\tau_H)}\left\{\cosh\frac{\Delta\Gamma}{2}t-\Re\left\{\lambda^*\left(\sinh\frac{\Delta\Gamma}{2}t+i\sin\Delta mt\right)\right\}\right\},\\ \bar{R} &= \frac{e^{-\Gamma t}}{\tau_H+\tau_L+\Re(\lambda)(\tau_L-\tau_H)}\left\{\cosh\frac{\Delta\Gamma}{2}t-\Re\left\{\lambda\left(\sinh\frac{\Delta\Gamma}{2}t+i\sin\Delta mt\right)\right\}\right\},\end{aligned} \tag{1.73}$$

From the expressions for the decay rates, the CP asymmetry can be written

$$A_{CP}(t)=\frac{(\bar{R}-R)}{(\bar{R}+R)}=\frac{\Im(\lambda)\sin\Delta mt}{\cosh\frac{\Delta\Gamma t}{2}-\Re(\lambda)\sinh\frac{\Delta\Gamma t}{2}} \tag{1.74}$$

and the averaged rate (applicable in the case of no flavour tagging, for example) is

$$\bar{R}+R=\frac{e^{-\Gamma t}}{\tau_H+\tau_L+\Re(\lambda)(\tau_L-\tau_H)}\left\{\cosh\frac{\Delta\Gamma}{2}t-\Re(\lambda)\sinh\frac{\Delta\Gamma}{2}t\right\}. \tag{1.75}$$

By adding flavour tagging information (introduced in Section 1.5.1) the decay rates of the B_s^0 and $\bar{B}_s^0$ mesons can be written separately as

$$\begin{aligned}\bar{R} &= \frac{e^{-\Gamma t}}{\tau_H+\tau_L+\Re(\lambda)(\tau_L-\tau_H)}\left\{\cosh\frac{\Delta\Gamma}{2}t-\Re(\lambda)\sinh\frac{\Delta\Gamma}{2}t\right\}\left(1+\frac{\Im(\lambda)\sin\Delta mt}{\cosh\frac{\Delta\Gamma t}{2}-\Re(\lambda)\sinh\frac{\Delta\Gamma t}{2}}\right)\\ R &= \frac{e^{-\Gamma t}}{\tau_H+\tau_L+\Re(\lambda)(\tau_L-\tau_H)}\left\{\cosh\frac{\Delta\Gamma}{2}t-\Re(\lambda)\sinh\frac{\Delta\Gamma}{2}t\right\}\left(1-\frac{\Im(\lambda)\sin\Delta mt}{\cosh\frac{\Delta\Gamma t}{2}-\Re(\lambda)\sinh\frac{\Delta\Gamma t}{2}}\right)\end{aligned} \tag{1.76}$$

As the CP asymmetry in this system is predicted to be very small, an approximation can be made using the fact that λ is almost real. Defining the components $\Im(\lambda)\approx\delta$ and $\Re(\lambda)\approx\pm(1-\delta^2/2)$, where the plus (minus) sign is for CP even (odd) final states, the rates can be approximated to first order in δ as:

$$\begin{aligned}\bar{R} &= \frac{e^{-\Gamma_{L(H)}t}}{2\tau_{L(H)}}+\frac{e^{-\Gamma t}}{2\tau_{L(H)}}\delta\sin\Delta mt,\\ R &= \frac{e^{-\Gamma_{L(H)}t}}{2\tau_{L(H)}}-\frac{e^{-\Gamma t}}{2\tau_{L(H)}}\delta\sin\Delta mt,\end{aligned} \tag{1.77}$$

for CP even (odd) final states.

Separating the definition of λ from Equation 1.57 into its real and imaginary com-

ponents, Equation 1.76 can be written as

$$\bar{R}_{E,O} = \frac{e^{-\Gamma t}}{\tau_H + \tau_L \mp \cos 2\beta_s(\tau_L - \tau_H)} \left\{ \cosh\frac{\Delta\Gamma}{2}t \mp \cos 2\beta_s \sinh\frac{\Delta\Gamma}{2}t \right\} \left(1 \pm \frac{\sin 2\beta_s \sin\Delta m t}{\cosh\frac{\Delta\Gamma t}{2} \mp \cos 2\beta_s \sinh\frac{\Delta\Gamma t}{2}} \right)$$
$$R_{E,O} = \frac{e^{-\Gamma t}}{\tau_H + \tau_L \mp \cos 2\beta_s(\tau_L - \tau_H)} \left\{ \cosh\frac{\Delta\Gamma}{2}t \mp \cos 2\beta_s \sinh\frac{\Delta\Gamma}{2}t \right\} \left(1 \mp \frac{\sin 2\beta_s \sin\Delta m t}{\cosh\frac{\Delta\Gamma t}{2} \mp \cos 2\beta_s \sinh\frac{\Delta\Gamma t}{2}} \right) \tag{1.78}$$

where the upper sign applies to CP even final states (E) and the lower to CP odd (O).

Recalling that δ is the imaginary part of λ above, 1.77 can also be re-written in terms of β_s:

$$\bar{R} = \frac{e^{-\Gamma_{L(H)} t}}{2\tau_{L(H)}} + \frac{e^{-\Gamma t}}{2\tau_{L(H)}} \sin(2\beta_s)\sin(\Delta m t),$$
$$R = \frac{e^{-\Gamma_{L(H)} t}}{2\tau_{L(H)}} - \frac{e^{-\Gamma t}}{2\tau_{L(H)}} \sin(2\beta_s)\sin(\Delta m t), \tag{1.79}$$

Time development of angular amplitudes in the case of multiple CP final states

Analogous to the decay rates which have been developed here, the time dependence of the angular amplitudes for the admixture of CP odd and even final states is of the form of Equation 1.71, with the normalisation from Equation 1.72 applied. The time dependent amplitudes to the $i = \{\|, \perp, 0\}$ states are

$$\begin{aligned} \mathcal{A}_i &= \frac{e^{-imt} e^{-\Gamma t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s(\tau_L - \tau_H)}} \left[E_+(t) \pm e^{2i\beta_s} E_-(t) \right] \mathcal{A}_i(0) \\ \bar{\mathcal{A}}_i &= \frac{e^{-imt} e^{-\Gamma t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s(\tau_L - \tau_H)}} \left[\pm E_+(t) + e^{2i\beta_s} E_-(t) \right] \mathcal{A}_i(0) \end{aligned} \tag{1.80}$$

where the upper sign applies to CP even and the lower to CP odd states, and the λ terms of 1.71 have been written in terms of β_s. The probability in Equation 1.65 can now be written time-dependently as

$$P_B(\theta, \phi, \psi, t) = \frac{9}{16\pi} |\mathbf{A}(t) \times \hat{n}|^2$$
$$P_{\bar{B}}(\theta, \phi, \psi, t) = \frac{9}{16\pi} |\bar{\mathbf{A}}(t) \times \hat{n}|^2 \tag{1.81}$$

where the complex vector $\mathbf{A}(t)$ is the time dependent equivalent of Equation 1.64. The normalisation is analogous to Equation 1.66

$$\iiiint \sum_{j=B,\bar{B}} P_j(\theta, \phi, \psi, t) \sin\theta d\theta \sin\psi d\psi = 1. \tag{1.82}$$

With these equations defined, it is possible to build a function to fit the decays of

$B_s^0 \to J/\psi\,\phi$ with arguments which are the observables of the transversity angles, time, flavour tagging information and the mixing and CP violation parameters β_s, $\Gamma, \Delta\Gamma$ and the angular amplitudes. However, in order to take into account the resolution and smearing effects of the detector, it is mathematically easier to expand the probabilities in terms of separate decay rates so that each can be separately convoluted with analytic resolution functions, which will be discussed in Chapter 4. This means separating the time development of the three $\mathcal{A}_i$s. Conveniently, the time development of the two CP even amplitudes, $\mathcal{A}_0(t)$ and $\mathcal{A}_\parallel(t)$ is identical, so they can be treated simultaneously. The time dependent vector $\mathbf{A}$ can be decomposed as

$$\mathbf{A}(t) = \mathbf{A}_+(t) + \mathbf{A}_-(t), \tag{1.83}$$

where the CP even component is

$$\mathbf{A}_+(t) = \mathbf{A}_+ f_+(t) = \left(\mathcal{A}_0 \cos\psi, -\frac{\mathcal{A}_\parallel \sin\psi}{\sqrt{2}}, 0\right) \cdot f_+(t), \tag{1.84}$$

and the CP odd component is

$$\mathbf{A}_-(t) = \mathbf{A}_- f_-(t) = \left(0, 0, i\frac{\mathcal{A}_\perp}{\sqrt{2}}\right) \cdot f_-(t). \tag{1.85}$$

The probabilities in Equation 1.81 can then be re-written in terms of the separate decay components

$$\begin{aligned}
P_B(\theta,\phi,\psi,t) &= \frac{9}{16\pi}\{|\mathbf{A}_+(t)\times\hat{n}|^2 + |\mathbf{A}_-(t)\times\hat{n}|^2 \\
&\quad +2\Re((\mathbf{A}_+(t)\times\hat{n})\cdot(\mathbf{A}_-^*(t)\times\hat{n}))\} \\
&= \frac{9}{16\pi}\{|\mathbf{A}_+\times\hat{n}|^2|f_+(t)|^2 + |\mathbf{A}_-\times\hat{n}|^2|f_-(t)|^2 \\
&\quad +2\Re((\mathbf{A}_+\times\hat{n})\cdot(\mathbf{A}_-^*\times\hat{n})\cdot f_+(t)\cdot f_-^*(t))\} \\
P_{\bar{B}}(\theta,\phi,\psi,t) &= \frac{9}{16\pi}\{|\bar{\mathbf{A}}_+(t)\times\hat{n}|^2 + |\bar{\mathbf{A}}_-(t)\times\hat{n}|^2 \\
&\quad +2\Re((\bar{\mathbf{A}}_+(t)\times\hat{n})\cdot(\bar{\mathbf{A}}_-^*(t)\times\hat{n}))\} \\
&= \frac{9}{16\pi}\{|\mathbf{A}_+\times\hat{n}|^2|\bar{f}_+(t)|^2 + |\mathbf{A}_-\times\hat{n}|^2|\bar{f}_-(t)|^2 \\
&\quad +2\Re((\mathbf{A}_+\times\hat{n})\cdot(\mathbf{A}_-^*\times\hat{n})\cdot\bar{f}_+(t)\cdot\bar{f}_-^*(t))\}.
\end{aligned} \tag{1.86}$$

The CP even and odd decay rates to a single CP eigenstate were calculated in Equation 1.78, so rearranging these and discarding higher order terms in δ gives $|f_+(t)|^2$ and $|f_-(t)|^2$, which are written here in terms of the observables Γ_L and Γ_H instead of the difference $\Delta\Gamma$

$$|\bar{f}_{\pm}|^2 = \frac{1}{2}\frac{(1 \pm \cos 2\beta_s)e^{-\Gamma_L t} + (1 \mp \cos 2\beta_s)e^{-\Gamma_H t} \pm 2\sin 2\beta_s e^{\Gamma t}\sin \Delta m t}{\tau_L(1 \pm \cos 2\beta_s) + \tau_H(1 \mp \cos 2\beta_s)},$$
$$|f_{\pm}|^2 = \frac{1}{2}\frac{(1 \pm \cos 2\beta_s)e^{-\Gamma_L t} + (1 \mp \cos 2\beta_s)e^{-\Gamma_H t} \mp 2\sin 2\beta_s e^{\Gamma t}\sin \Delta m t}{\tau_L(1 \pm \cos 2\beta_s) + \tau_H(1 \mp \cos 2\beta_s)}. \quad (1.87)$$

The interference terms for B_s^0 and $\bar{B}_s^0$ are

$$\bar{f}_+(t)\bar{f}_-^*(t) = \frac{-e^{-\Gamma t}\cos \Delta m t - i\cos 2\beta_s e^{-\Gamma t}\sin \Delta m t + i\sin 2\beta_s(e^{-\Gamma_L t} - e^{-\Gamma_H t})/2}{\sqrt{[(\tau_L - \tau_H)\sin 2\beta_s]^2 + 4\tau_L\tau_H}},$$
$$f_+(t)f_-^*(t) = \frac{-e^{-\Gamma t}\cos \Delta m t + i\cos 2\beta_s e^{-\Gamma t}\sin \Delta m t + i\sin 2\beta_s(e^{-\Gamma_L t} - e^{-\Gamma_H t})/2}{\sqrt{[(\tau_L - \tau_H)\sin 2\beta_s]^2 + 4\tau_L\tau_H}}. \quad (1.88)$$

Now that the decay terms for the different components are separated, these can be more easily included in a likelihood fitting function with smeared exponentials replacing the exponentials, and smeared sine exponentials replacing the sine exponentials, to incorporate the detector resolution effects. The full development of the likelihood function is discussed in Chapter 4, and includes a re-working of the normalisation to account for the angle dependent efficiency of the detector.

1.5.3 $B_s^0 \to J/\psi K^+K^-$ in $B_s^0 \to J/\psi\,\phi$ signal sample

A significant addition to this analysis over earlier investigations in the $B_s^0 \to J/\psi\,\phi$ channel is the consideration of a potential S-wave contamination of the signal ϕ meson ($P-$wave state) in $B_s^0 \to J/\psi\,\phi$. It has been suggested [46] that a 5-10% contamination from an S-wave KK final state due to $B_s^0 \to J/\psi K^+K^-$ (non-resonant) or $B_s^0 \to J/\psi f_0$ could be present in the in the $B_s^0 \to J/\psi\,\phi$ signal. The effect of this potential level of contamination is predicted in [47], with the conclusion that a 5-15% S-wave KK contamination could bias the measurement of $\beta_s^{J/\psi\phi}$ significantly towards zero, thus favoring the Standard Model value. In order to include the S-wave KK component in the likelihood fit, it is necessary to combine the $B_s^0 \to J/\psi KK(f_0)$ amplitude with the equations developed in the previous section for the ϕ meson P-wave state, and calculate a normalisation factor for the full decay rate.

The decay to the final state including the KK S-wave follows the time dependence of the CP odd component of the $B_s^0 \to J/\psi\,\phi$ decay, $|f_-(t)|$, because the combined $J/\psi K^+K^-$ or $J\psi f_0$ state is a CP odd eigenstate. Contributions from non-resonant KK and f_0 are treated simultaneously, as they would be indistinguishable in the current data sample size and over the mass range considered. The f_0 component which lies under the signal ϕ peak is the upper tail of the broad f_0 resonance, which is treated as a flat distribution within the narrow ϕ mass window defined in the data selection process for this analysis (Chapter 3). From this point onwards, both the $B_s^0 \to J/\psi K^+K^-$ and $B_s^0 \to J/\psi f_0$ will be written as $B_s^0 \to J/\psi KK$.

First, a probability density for the pure $B_s^0 \to J/\psi KK$ process is calculated, anal-

ogous to Equation 1.81 for the $B_s^0 \to J/\psi\,\phi$ decay

$$\begin{aligned} Q_B(\theta,\phi,\psi,t) &= \frac{3}{16\pi}|\mathbf{B}(t)\times\hat{n}|^2 \\ Q_{\bar{B}}(\theta,\phi,\psi,t) &= \frac{3}{16\pi}|\bar{\mathbf{B}}(\mathbf{t})\times\hat{n}|^2, \end{aligned} \tag{1.89}$$

where (c.f. Equation 1.84, 1.85)

$$\begin{aligned} \mathbf{B}(t) &= (\mathcal{B}(t),0,0) \\ \bar{\mathbf{B}}(\mathbf{t})(t) &= (\bar{\mathcal{B}}(t),0,0) \end{aligned} \tag{1.90}$$

and the time dependent amplitudes for B_s^0 and $\bar{B}_s^0$ to the CP odd final state are

$$\begin{aligned} \mathcal{B}(t) &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H+\tau_L+\cos 2\beta_s\,(\tau_L-\tau_H)}}\left[E_+(t)-e^{2i\beta_s}E_-(t)\right], \\ \bar{\mathcal{B}}(t) &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H+\tau_L+\cos 2\beta_s\,(\tau_L-\tau_H)}}\left[-E_+(t)+e^{-2i\beta_s}E_-(t)\right]. \end{aligned} \tag{1.91}$$

To combine the probability amplitudes for the $J/\psi KK$ and $J/\psi\phi$ final states, it is necessary to sum and square equations 1.89 and 1.81. It is also essential to account for the relative phases and mass shapes of the two components. The ϕ meson mass shape is described here by a non-relativistic (mass dependent) Breit Wigner distribution, and the S-wave is treated as a flat (non-resonant) mass distribution, as discussed earlier. Only the relative phase between the two components can be measured, and the S-wave can take any phase with respect to the P-wave part.

Defining the parameters: F_{sw} as the fraction of $J/\psi KK$; μ_ϕ as the ϕ mass (1019 MeV/c^2); Γ_ϕ as the ϕ width (4.26 MeV/c^2) [10], and δ_s as the relative phase of the $J/\psi KK$ component, the normalised probability for the combination of the final states can be written as

$$\begin{aligned} \rho_B(\theta,\phi,\psi,t,\mu) &= \frac{9}{16\pi}\left|\left[\sqrt{1-F_{sw}}h(\mu)\mathbf{A}(t)+e^{i\delta_s}\sqrt{F_{sw}}\frac{g(\mu)}{\sqrt{3}}\mathbf{B}(t)\right]\times\hat{n}\right|^2 \\ \rho_{\bar{B}}(\theta,\phi,\psi,t,\mu) &= \frac{9}{16\pi}\left|\left[\sqrt{1-F_{sw}}h(\mu)\bar{\mathbf{A}}(t)+e^{i\delta_s}\sqrt{F_{sw}}\frac{g(\mu)}{\sqrt{3}}\bar{\mathbf{B}}(\mathbf{t})\right]\times\hat{n}\right|^2, \end{aligned} \tag{1.92}$$

when the reconstructed ϕ mass μ lies within a window $\mu_{lo}<\mu<\mu_{hi}$. In Equation 1.92, the ϕ meson mass propagator, $h(\mu)$, is of the relativistic Breit Wigner form:

$$h(\mu)=\sqrt{\frac{\Gamma_\phi/2}{\Delta\omega}}\cdot\frac{1}{\mu-\mu_\phi+i\Gamma_\phi/2}, \tag{1.93}$$

where

$$\omega_{hi}=\tan^{-1}\frac{2(\mu_{hi}-\mu_\phi)}{\Gamma_\phi} \qquad \omega_{lo}=\tan^{-1}\frac{2(\mu_{lo}-\mu_\phi)}{\Gamma_\phi}. \qquad \Delta\omega=\omega_{hi}-\omega_{lo} \tag{1.94}$$

The S-wave KK state's flat mass distribution is given by $g(\mu)$:

$$g(\mu) = \frac{1}{\sqrt{\Delta\mu}}, \tag{1.95}$$

where the KK invariant mass window is

$$\Delta\mu = \mu_{hi} - \mu_{lo} \tag{1.96}$$

In order to simplify the expansion of Equation 1.92, the mass dependent function $\mathcal{F}(\mu)$ can be defined

$$\mathcal{F}(\mu) \equiv \sqrt{\frac{F_{sw}(1-F_{sw})\Gamma_\phi}{2\Delta\mu\Delta\omega}} \cdot \frac{e^{i\delta_s}}{\mu - \mu_\phi - i\Gamma_\phi/2} \tag{1.97}$$

In this analysis, for the final fit function the KK mass will be integrated over, so the quantity of interest in this case is the integral of $\mathcal{F}(\mu)$

$$\mathcal{I}_\mu \equiv \int \mathcal{F}(\mu)d\mu = \sqrt{\frac{F_{sw}(1-F_{sw})\Gamma_\phi}{2\Delta\mu\Delta\omega}} \cdot e^{i\delta_s} \cdot \log\frac{\mu_{hi} - \mu_\phi - i\Gamma/2}{\mu_{lo} - \mu_\phi - i\Gamma/2}. \tag{1.98}$$

Following the procedure as for the $B_s^0 \to J/\psi\,\phi$ decay, the time dependent equivalent of 1.89, the CP odd final state which incorporates the S-wave KK, is defined as

$$\begin{aligned} Q_B(\theta,\psi,\phi,t) &= \frac{3}{16\pi}|\mathbf{B}(\mathbf{t}) \times \hat{n}|^2 \\ &= \frac{3}{16\pi}|\mathbf{B} \times \hat{n}|^2 |f_-(t)|^2. \end{aligned} \tag{1.99}$$

for the B_s^0 initial state and

$$\begin{aligned} Q_{\bar{B}}(\theta,\psi,\phi,t) &= \frac{3}{16\pi}|\bar{\mathbf{B}}_+(t) \times \hat{n}|^2 \\ &= \frac{3}{16\pi}|\mathbf{B}_+ \times \hat{n}|^2 |\bar{f}_-(t)|^2. \end{aligned} \tag{1.100}$$

for the $\bar{B}_s^0$ initial state.

Now, the combined time dependence can be expanded, integrating over the ϕ meson mass and using the identity defined in Equation 1.98, as

$$\begin{aligned} \rho_B(\theta,\psi,\phi,t,F_{sw},\delta_{sw}) &= (1-F_{sw}) \cdot P_B(\theta,\psi,\phi,t) + F_{sw}Q_B(\theta,\psi,\phi,t) \\ &+ 2\frac{\sqrt{27}}{16\pi}\Re[\mathcal{I}_\mu((\mathbf{B}\times\hat{n})\cdot(\mathbf{A}_-\times\hat{n})\cdot|f_-(t)|^2 \\ &+ (\mathbf{B}\times\hat{n})\cdot(\mathbf{A}_+\times\hat{n})\cdot f_+(t)\cdot f_-^*(t))] \end{aligned} \tag{1.101}$$

$$\begin{aligned}\rho_{\bar{B}}(\theta,\psi,\phi,t,F_{sw},\delta_{sw}) &= (1-F_{sw})\cdot P_{\bar{B}}(\theta,\psi,\phi,t)+F_{sw}Q_{\bar{B}}(\theta,\psi,\phi,t)\\ &+ 2\frac{\sqrt{27}}{16\pi}\Re[\mathcal{I}_{\mu}((\mathbf{B}\times\hat{n})\cdot(\mathbf{A}_{-}\times\hat{n})\cdot|\bar{f}_{-}(t)|^{2}\\ &+ (\mathbf{B}\times\hat{n})\cdot(\mathbf{A}_{+}\times\hat{n})\cdot\bar{f}_{+}(t)\cdot\bar{f}_{-}^{*}(t))]. \qquad (1.102)\end{aligned}$$

The equations necessary to construct a likelihood fit function with the aim of measuring the CP violating parameter β_s in the channel $B_s^0 \to J/\psi\,\phi$ (written from this point as $\beta_s^{J/\psi\phi}$) have been defined in this section. These will be developed in the context of the experimental method, to include detector effects, and combined into a full maximum log likelihood fit in Chapter 4.

There exist several symmetries under transformations of the equations describing the $B_s^0 \to J/\psi\,\phi$ decay, which will also be defined along with a discussion of their implications in Chapter 4. These can be reduced by adding flavour tagging information to follow the time dependence of an initial B_s^0 or $\bar{B}_s^0$ state separately, and by the inclusion of the additional S-wave $B_s \to J/\psi K^+K^-$ or $B_s \to J/\psi$ contamination, but with the currently available statistics it is not possible to determine between two values of $\beta_s^{J/\psi\phi}$ which are equivalent under the symmetry described in Section 4.6, preventing the possibility of measuring an exact point value of $\beta_s^{J/\psi\phi}$ at this stage.

1.6 Review of current experimental status

The analysis documented in this thesis uses the largest sample of $B_s^0 \to J/\psi\,\phi$ decays available to date, as well as improvements to the analysis method over previous studies, which are described in Chapter 4. Both the CDF and DØ collaborations have published measurements of the parameters of interest in $B_s^0 \to J/\psi\,\phi$ decays, using datasets of about 25% of the size of the sample described in this thesis. Additionally, these collaborations have produced more recent preliminary updates which have been combined to make the most of the statistics available from the Tevatron collider experiments. So far these are the only two experiments to analyse this decay channel, although the B factories, BaBar and Belle have published measurements for the B^0 system in the kinematically similar $B^0 \to J/\psi\,K^{*0}$ channel [48], [49].

Using a data sample of around 2500 $B_s^0 \to J/\psi\,\phi$ signal events, CDF found [50]:

$$\begin{aligned} c\tau_s &= 456 \pm 13\ (\text{stat.}) \pm 7\ (\text{syst.})\ \mu\text{m}\\ \Delta\Gamma &= 0.076^{+0.059}_{-0.063}\ (\text{stat.}) \pm 0.006\ (\text{syst.})\ ps^{-1}\\ |A_{\parallel}(0)|^2 &= 0.230 \pm 0.021\ (\text{stat}) \pm 0.007\ (\text{syst.})\\ |A_0(0)|^2 &= 0.530 \pm 0.027\ (\text{stat}) \pm 0.009\ (\text{syst.})\end{aligned}$$

with the assumption of no CP violation ($\beta_s^{J/\psi\phi}$=0.0) and no flavour tagging used in the fit. From this analysis, with $\beta_s^{J/\psi\phi}$ floating, there was not sufficient sensitivity to exclude any range of values for $\beta_s^{J/\psi\phi}$ at the 68% confidence level.

The CDF analysis published in 2008 [4] using around 2000 $B_s^0 \to J/\psi\,\phi$ signal events included flavour tagging of the initial B_s^0 meson state. This found $\beta_s^{J/\psi\phi}$ to lie

within [0.16,1.41] at the 68% confidence level.

For comparison, the DØ collaboration published a measurement in 2007 [5] using a non-flavour tagged analysis with the hypothesis of no CP violation, with 1040 signal $B_s^0 \to J/\psi\,\phi$ events giving

$$\begin{aligned} c\tau_s &= 456 \pm 24 \text{ (stat.)}^{+3}_{-9} \text{ (syst.) } \mu\text{m} \\ \Delta\Gamma &= 0.12^{+0.08}_{-0.10} \text{ (stat.)} \pm 0.02 \text{ (syst.) } ps^{-1} \end{aligned}$$

In a separate paper [51], DØ present the angular amplitudes

$$\begin{aligned} |A_\parallel(0)|^2 &= 0.244 \pm 0.032 \text{ (stat)} \pm 0.014 \text{ (syst.)} \\ |A_0(0)|^2 &= 0.555 \pm 0.027 \text{ (stat)} \pm 0.006 \text{ (syst.)} \end{aligned}$$

from a fit without flavour tagging, assuming $\beta_s^{J/\psi\phi}$=0.0.

The most recent updates prior to the measurement described in this thesis were presented as preliminary results by the CDF and DØ in 2008 and 2009 [7] [52], each with about 50% of the statistics of the current dataset. These analyses were combined [6], finding $\beta_s^{J/\psi\phi}$ in the range $[0.27, 0.59] \cup [0.97, 1.30]$ at the 68% confidence level, which gives a probability of 3.4%, or 2.1σ for the Standard Model expected values of $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$. This hint at a deviation from the Standard Model expectation adds further incentive to produce a more precise measurement of $\beta_s^{J/\psi\phi}$.

Chapter 2

Experimental apparatus

The data analysed and presented in this thesis were collected at the Collider Detector Facility (CDF-II detector) from 2002-2009 of the ongoing Run-II data taking period at Fermilab. This detector is at one of two collision points in the Tevatron accelerator, which collides protons with antiprotons at 1.96 TeV centre of mass energy. This chapter first describes the components of the Fermilab accelerator complex, and the methods used to produce and accelerate protons and antiprotons, then gives details of the CDF-II detector, with particular focus on those detector components used for this and other B-physics analyses.

2.1 The Tevatron accelerator complex

The Tevatron is a super-conducting synchrotron, of 1 km radius, which was the world's highest energy particle accelerator until December 2009 when it was surpassed by the energy of the Large Hadron Collider (LHC) at CERN. The current Tevatron accelerator is an upgrade of the original, and is referred to as the Run-II Tevatron. Its name derives from the term *Tera electron Volt*, as it accelerates protons and antiprotons to 980 GeV (0.98 TeV) in standard running for Run-II.

The Tevatron is the final stage of acceleration for beams of protons and antiprotons produced at the Fermilab accelerator complex. The full chain of accelerators is shown in figure 2.1. In addition to providing colliding beams for the two detectors positioned around the Tevatron, CDF and DØ [53], the complex can send high energy beams to fixed target experiments and test beam facilities at Fermilab.

2.1.1 Proton production and acceleration

There are three components of the accelerator system which are collectively known as the Proton Source: the Cockroft Walton (Pre-accelerator), Linear accelerator (Linac) and Booster.

Cockroft Walton Pre-accelerator

The Pre-accelerator is the first stage of the acceleration chain, there are two electrostatic accelerators of the Cockroft Walton design [54] which provide negatively charged hydrogen ions to the Linac. Only one of the two is used at any time. Each consists of a H^- ion source contained in an electrically charged dome. The H^- ions are linearly accelerated to 750 keV by the potential of the negatively charged dome, and then travel through a transfer line to the Linac.

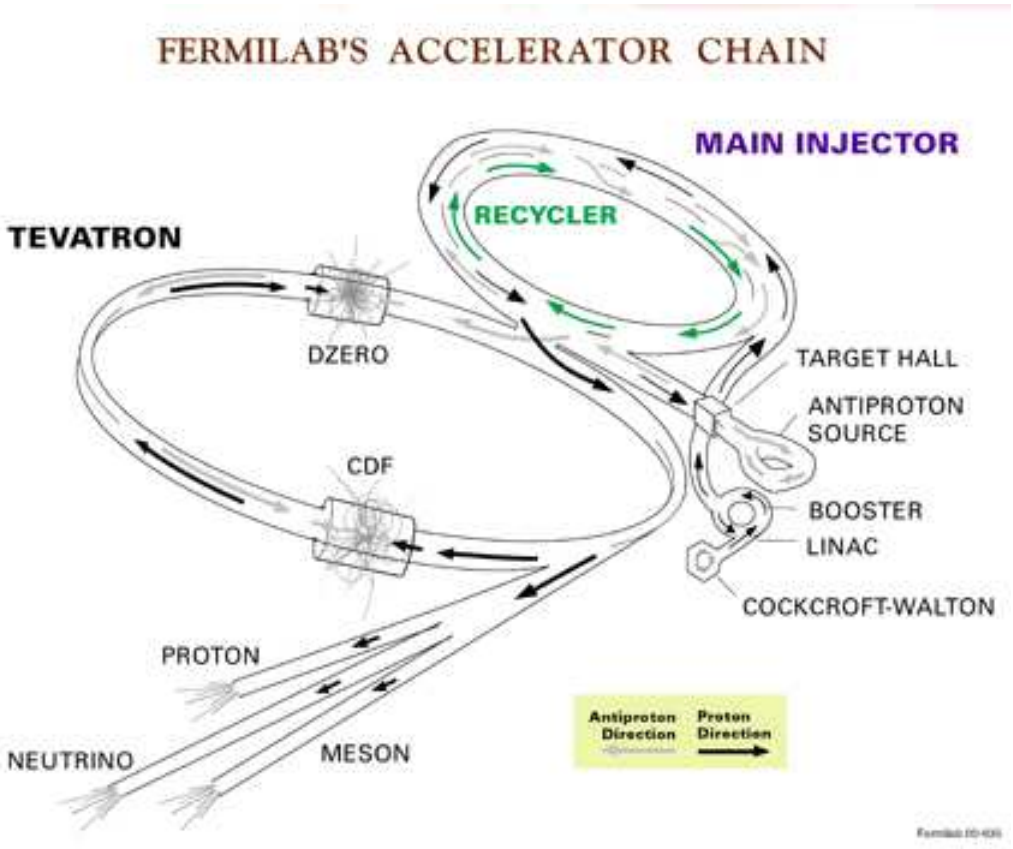

Figure 2.1: The Fermilab accelerator chain

Linac

The Linear Accelerator (Linac) [55] accelerates the 750 keV H^- ions to 400 MeV. It consists of two sections, the first is made up of five Alvarez drift tubes which are named for the Professor of Physics, Luis Alvarez, who designed the Linac at Berkley in 1947 on which the original Fermilab linac design was based. This section of the chain accelerates the ion beam from the input 750 keV to 116 MeV. The second part is a set of seven radio frequency (RF) cavity modules, which operate at 805 MHz and accelerate the beam to 400 MeV.

Booster

At the booster stage, the 400 MeV H^- ions are striped of their electrons, leaving only the protons which are accelerated to 8 GeV. The Booster [56] is a synchrotron of 75 m radius, with 19 RF cavities interspersed with 96 bending and focusing magnets about its ring. Before the Fermilab accelerator complex was built, most proton colliders used protons directly injected from the linac into the main accelerator, however, a Linac capable of producing 8 GeV protons required for the Tevatron would have had to be about 6 km long, which was too expensive and impractical to be implemented. The solution was to add a booster synchrotron, a circular accelerator which accelerates the protons in *bunches*. The term *bunch* refers to the portion of a proton beam in a stable phase space area of the RF accelerator, which can be captured and accelerated by the application of RF. The beam is accelerated in cycles in the Booster, gaining energy

each time it circulates, the full process of acceleration to 8 GeV takes approximately 33 ms, then the beam can be transferred to the Main Injector.

Main Injector

The Main Injector (MI) [57] is a second synchrotron accelerator, it has several modes of operation, of which the ones of relevance to the high energy collisions will be discussed here. The MI has 18 accelerating RF cavities, and increases the energy of beams to 120 GeV or 150 GeV according to the mode of use.

In antiproton (Pbar) production mode, the MI takes protons at 8 GeV energy from the Booster, combines two Booster batches in a process called *slip stacking* and accelerates them to 120 GeV, before sending them to the Antiproton source which will be described in the following section.

For injecting beams directly to the Tevatron, the energy reached is 150 GeV. For this mode, the MI can take antiproton beams from the Antiproton source as well as protons from the Booster. For injecting protons the MI takes seven bunches from the Booster and combines them into a single bunch, by a process called *coalescing*. This is repeated 36 times to load the protons necessary for a standard store in the Tevatron. For the antiprotons, the MI takes four antiproton bunches and accelerates them to 150 GeV, repeating this four times to acquire the 36 bunches equivalent to the proton bunches for collision in the Tevatron.

2.1.2 Antiproton production, storage and acceleration

Antiprotons are not abundant in any terrestrial source, so their production is a challenge for any proton-antiproton collider [58]. At Fermilab they are produced in collisions of protons with a fixed target and selected from the large range of collision products. Storage and control of the antiprotons also requires sophisticated techniques, which will be described in the following sections.

Target

The antiproton source is a nickel alloy fixed target, which is struck by high energy protons from the MI. The product of these collisions is a spray of particles, a small fraction of which are antiprotons. The antiprotons at 8 GeV are selected using a pulsed dipole magnet to control the momentum and charge, and sent to the Debuncher.

Debuncher

The antiprotons directly from the antiproton source have a wide spread of momentum and transverse phase space. For high energy collisions, it is advantageous to have a focused, narrow beam, as will be discussed further in Section 2.1.3. For the proton beams, this can be achieved by discarding the particles which are outside the required phase space, because of the abundance of protons and their relative ease of production. The antiprotons, however, are too valuable to be thrown away like this, so they have to be treated using more efficient methods.

The first stage of controlling the antiproton momentum spread is carried out by RF bunch rotation and adiabatic debunching. The Debuncher RF system takes bunches from the target, which are short in time spread, and have a large momentum spread; it rotates them in phase space resulting in bunches with small momentum spread and

a larger time spread. By lowering the RF voltage, the beam is then adiabatically debunched.

The beam is further reduced in momentum spread and transverse beam size by *stochastic cooling*. In this context, cooling refers to narrowing the spread in phase space of the beam. Stochastic cooling was pioneered by Simon van der Meer [59] in the UA1 experiment at CERN in the early 1980s. This technique was essential in the discovery of the W and Z bosons, for which van der Meer was awarded the Nobel Prize in 1984, along with the physicist Carlo Rubbia [60]. The method relies on feedback from pickup sensors at one point in the ring, to kicker electrodes at a further point. The sensor detects the spread in the transverse or longitudinal motion of the antiprotons, and sends this information to the kicker, a magnet which applies a force to correct it, as demonstrated in Figure 2.2. Stochastic cooling is also used in the Accumulator stage of the antiproton source, and in the Recycler.

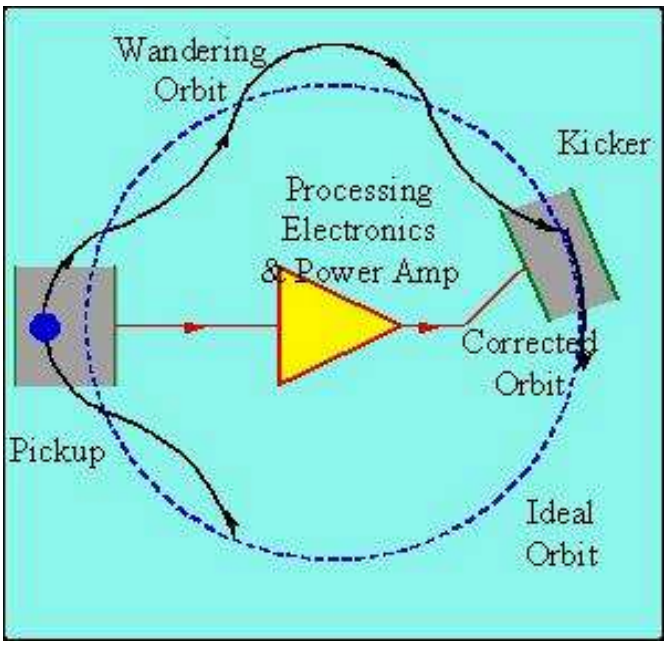

Figure 2.2: The stochastic cooling method [61]

Accumulator

The second synchrotron of the antiproton source is the Accumulator, which is of similar dimensions to the Debuncher. It is a storage ring for 8 GeV antiprotons, which undergo further cooling there before being sent to the MI for acceleration.

Recycler

Another antiproton storage ring is the Recycler. It's original purpose was to recycle remaining antiprotons from a Tevatron store and re-cool them along with the new antiprotons from the antiproton source [54]. However, problems with this use in early in Run-II caused a decision to change the system so that it is now only used to store and further cool antiprotons direct from the antiproton source.

In addition to stochastic cooling, as described for use in the debuncher, the recycler uses Electron cooling. Stochastic cooling loses effectiveness at higher intensities, when

there are more than 200×10^{10} antiprotons in the Recycler electron cooling is more suitable. Electron cooling was developed by G.I. Budker at Novsibirsk [58]. It works on the same principle as a heat exchanger, in that the "warm" (wide momentum spread) beam of antiprotons travels parallel to a "cool" beam of electrons, and as the beams interact, the energy of the warmer antiproton beam is transferred to the cooler electron beam, thereby reducing the spread of the antiproton beam.

2.1.3 Tevatron collisions

The final stage in the Fermilab acceleration process is the Tevatron [62]. This accelerator stands apart from the rest of the chain in several aspects; it is the largest ring, the only cryogenically cooled accelerator and the only device capable of operating in colliding beams mode. The Tevatron takes both protons and antiprotons from the MI at 150 GeV and accelerates them to 980 GeV energy. The beams are collided at two points along the circumference of the accelerator, corresponding to the two colliding beam experiments at Fermilab, CDF and DØ.

The super-conducting magnets used by the Tevatron require cryogenic cooling to maintain their super-conductivity, which allows very high currents to pass through them in order to create a strong magnetic field. Liquid helium is used to keep the magnets at around 4.6 K. There are three main uses of magnets in the Tevatron: bending, focusing and correcting the beam. Dipole magnets are used to bend the beams into a circular trajectory. These produce a radial force on particles which travel through the beam pipe as it passes through the magnet. Focusing magnets are quadropoles, they focus the beam by applying equal forces about the beam pipe, such that a particle in the centre of the beam will feel no net force, but one which is out of alignment will be pushed towards the centre by a force in the opposite direction to the one in which it is off-centre, thus correcting any misalignments. Quadropoles are also used to focus the beam to the collision (interaction) points.

Luminosity and energy

Two features in particular characterise a particle accelerator: its luminosity and energy. The energy reached by the machine determines the type of physics which can be investigated with it, by way of the range of the particle masses which can be produced in the resulting collisions. The luminosity, however, determines the rate of data acquisition. This affects the likelihood of observing rarely occurring events, and the statistical uncertainties on measurements made with the detectors. This can be understood by considering the definition of luminosity. The probability of a beam particle colliding with another from the opposite beam is given by the interaction cross-section, σ_{int}. The rate of these interactions in a detector is [62]

$$R = \sigma_{int} L \tag{2.1}$$

where the luminosity, L is a measure of how likely the particles are to interact between the two beams. It is given in the Tevatron by

$$L = \frac{f n N_p N_{\bar{p}}}{A} = \frac{f n N_p N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} F\left(\frac{\sigma_l}{\beta^*}\right) \tag{2.2}$$

where f is the revolution frequency, n is the number of bunches per beam, $N_{p,\bar{p}}$ is the number of particles per bunch and A is the cross sectional area of the beams. In the right hand side of Equation 2.2, A is written in terms of the standard deviation of the beams at the interaction point ($\sigma_{P,\bar{p}}$) and a form factor, $F\left(\frac{\sigma_l}{\beta^*}\right)$ which is defined by the bunch length, σ_l and the interaction point beta function, β^*. From Equation 2.2 it is clear that focusing the beams, as discussed in previous sections, increases the luminosity by reducing the denominator size. Increasing the luminosity increases the rate of interactions, therefore making greater statistics available for the experiments. The Tevatron Operations group have been consistently increasing the luminosity provided to the experiments throughout the run period of the detector. Figure 2.3 shows the rate of increase in peak luminosity week by week since the start of Run-II, where peak luminosity is the highest luminosity point, occurring at the beginning of a store of protons and antiprotons injected into the Tevatron. Integrated luminosity is the measure of total data delivered to the experiments, which is shown for Run-II in Figure 2.4.

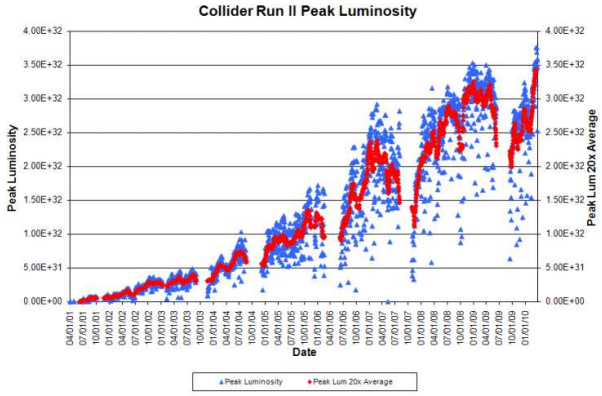

Figure 2.3: Tevatron Run-II peak instantaneous luminosity [63]

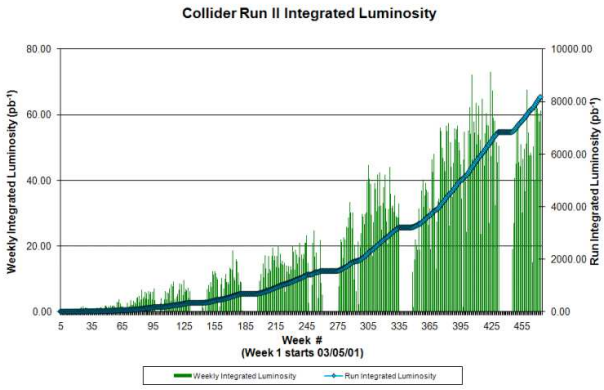

Figure 2.4: Tevatron Run-II integrated luminosity [63]

2.2 The CDF detector

The CDF detector saw its first $p\bar{p}$ collisions in 1985, since then it has been upgraded to meet new physics goals and to keep up with the increases in Tevatron luminosity. The

latest version of the detector, CDF-II, was completed in 2002. It is a general purpose solenoidal detector with strengths in charged particle tracking, calorimetry and muon detection. The design is symmetric both forward-backward, and azimuthally, and as a multi-purpose detector it is optimised to extract the maximum information about the properties of a wide range of particles produced in $p\bar{p}$ collisions. A full description of the various components which make up the CDF Run-II detector is given in [64].

Before introducing the sub-detectors of CDF, some terms used in the detector descriptions will be defined.

Coordinate system definition

The three coordinate systems of the CDF detector [65] each have their origin at the collision point in the centre of the detector. In Cartesian coordinates, the z-axis is along the beam direction, with the positive direction following the motion of the proton beam, the x-axis is in the horizontal plane pointing radially outwards from the centre of the Tevatron ring, and the y-axis points vertically upwards from this. Due to the cylindrical shape of the detector, it is practical to also use both polar (ρ, ϕ, θ) and cylindrical (r, ϕ, z) coordinates. In these cases, the ρ and r coordinates define the radial distance from the origin. The angle ϕ is the azimuthal angle, and the plane $r-\phi$, is like the $x-y$ plane, transverse to the beamline, the polar angle, θ is defined with respect to the x-axis.

In addition to the coordinates described, it is useful to define the angular variable η, the pseudorapidity

$$\eta \equiv -\ln \tan\left(\frac{\theta}{2}\right) \tag{2.3}$$

which is equivalent in the ultrarelativistic, massless limit, to the rapidity of a particle

$$y \equiv \frac{1}{2} \ln \frac{E + p_z}{E - p_z}. \tag{2.4}$$

Overview of the CDF detector

A cut-away view of the CDF detector is shown in Figure 2.5. The beampipe lies at the centre of the detector. The tracking system, consisting of a silicon microstrip detector surrounded by a multi-wire drift chamber, surrounds the beampipe. The trackers have a cylindrical geometry about the beamline. Immediately outside of the tracking is the Time of Flight (TOF) detector which is used to gain particle identification information. A super-conducting solenoid surrounds these inner systems, providing a 1.4 Tesla magnetic field longitudinal to the beamline. The calorimeters lie outside of the solenoid, to absorb the energy of particles which interact with them. The muon detectors form the outer layer of the CDF detector, because muons are minimally ionising particles, which lose little energy in the inner parts of the detector and can therefore be identified by reaching the outlying muon chambers.

The following sections give an overview of the functions of the detector and its components.

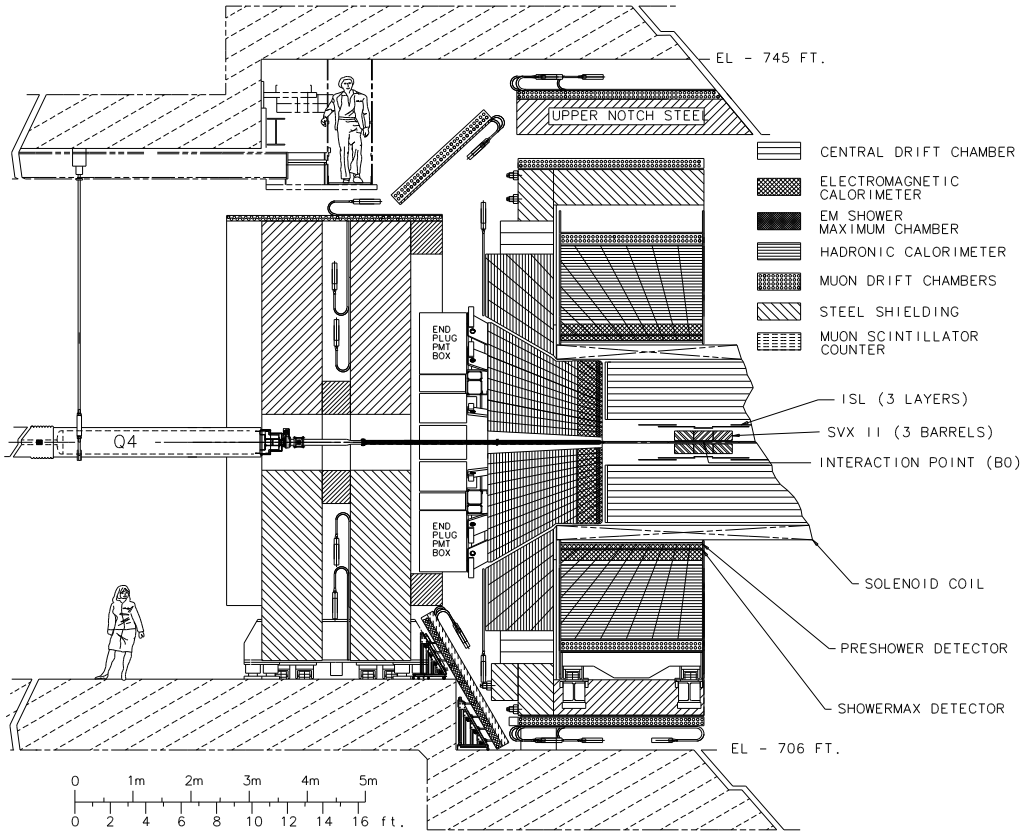

Figure 2.5: Elevation view of half of the CDF Run-II detector

2.2.1 Tracking detectors

Tracking refers to the measurement of the trajectories of charged particles traversing the detector volume. This allows the measurement charge and momenta of particles and the calculation of kinematical quantities. Through the process of ionisation, charged particles leave deposits of charge and energy in the detector, which are called *hits*. The pattern of these hits can be combined to reconstruct *tracks* of particles. In addition, the tracks can then be put together to calculate the points in space (*vertices*) where a particle has decayed to daughter particles.

The CDF Run-II tracking system is based on a similar design to the Run-I detector, which was operational 1987-1996. An upgrade was necessary to cope with the higher luminosities expected in Run-II, and to improve the silicon tracking with newer technology. The Run-I Central Tracking Chamber (CTC) has been replaced with the Central Outer Tracker (COT), a small cell, higher luminosity capable drift chamber, with significantly better stereo tracking than its predecessor. The second generation silicon detector for Run-II has greatly improved coverage and tracking power.

The layout of the CDF Run-II tracking system is shown in Figure 2.6, and the sub-components described in more detail in the following subsections.

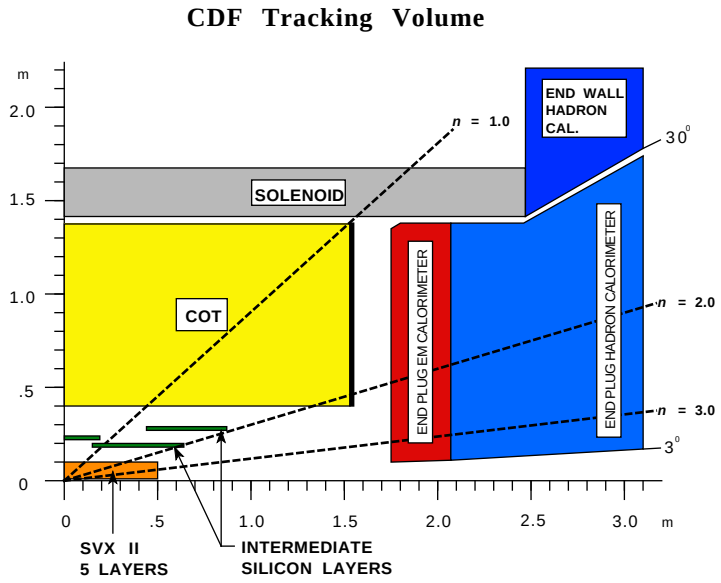

Figure 2.6: Layout of the CDF Run-II tracking system (side view)

Silicon trackers

A silicon strip detector, such as those used in CDF, is a reverse-biased p-n junction, with the p or n side segmented into fine strips. When an ionising particle passes through the detector, it leaves a trail of electron-hole pairs, which are separated by the electric field. The purpose of the reverse biased voltage is to increase the gap between the conduction and valance bands of the p-n junction, thereby reducing thermal currents. By reading out the charge deposited on individual strips, a localised position for the hit can be obtained. CDF uses both single and double sided silicon microstrip detectors. Single sided detectors have strips for only one side of the junction, for double sided detectors both sides are segmented. With double sided strip detectors, one side can be segmented into strips at an off-set angle to the other. In the case of CDF, the p side strips are parallel to the z axis, and the n side is segmented at a stereo angle to the z direction. This means that in addition to getting a precision measurement in the ϕ direction on the p side, z position information can be recorded on the n side.

The silicon tracking system consists of the Silicon Vertex Detector (SVX II), Intermediate Silicon Layers (ISL) and the innermost part, Layer 00 (L00). These are illustrated in Figure 2.7.

The SVX II [66] is the main component of the silicon system. It is made up of 3 barrels, each 29 cm in length, divided into 12 wedges in ϕ. Each wedge is constructed of 5 silicon sensor layers radially located 2.45-10.65 cm from the beam pipe. The sensors are mounted in units called *ladders*, in groups of four, which are fixed to a bulkhead containing cooling channels at the end of each barrel. This sub-detector uses double sided silicon strip sensors, with three layers of 90° stereo sensors and two of small angle stereo sensors, giving 3D vertex reconstruction capability and pattern matching with the outer tracking systems. The resolution in $r-\phi$ for SVX II hits is 9μm.

Between the beam pipe and the SVX II detector lies the L00 silicon detector [67]. This single-sided, radiation hard microstrip detector is designed to withstand the high

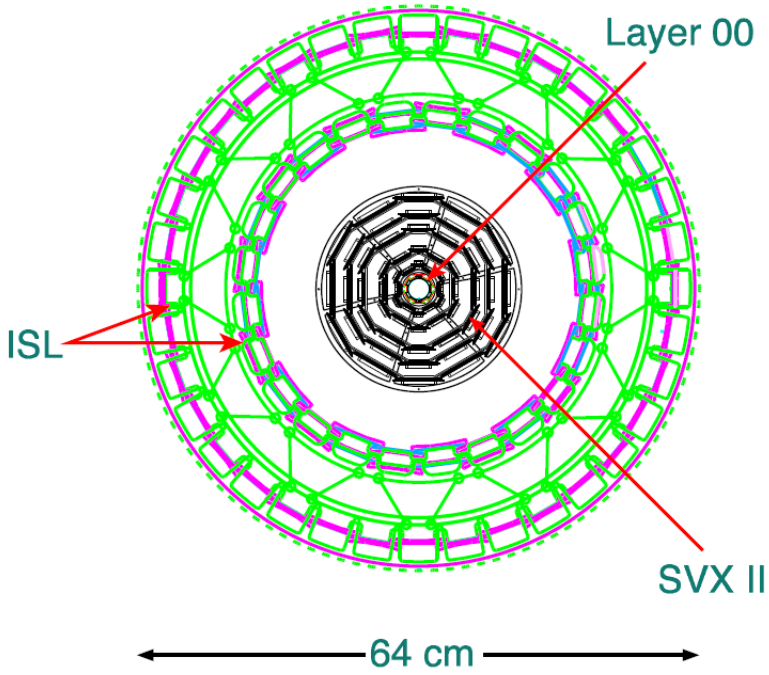

Figure 2.7: Cross section view of the CDF silicon detectors

radiation environmental conditions and extend the useful lifetime of the CDF silicon system. The inner layers of the SVX will cease to be operational as radiation damage occurs and the double sided sensors can no longer be depleted at the maximum bias voltage, but L00 is built to take higher bias voltages, and therefore remain functional as radiation damage occurs. It is positioned at 2 cm radially from the centre of the beampipe, has full azimuthal coverage and extends to $|z| < 47$ cm.

The final silicon sub-detector is the ISL, which provides coverage between the SVX II and the COT. It is divided into three layers, a central one at 23 cm radially, one at 20 cm and one at 28 cm. The purpose of the central layer is to provide tracking in conjunction with the COT, the other layers cover the region of $1.0 < |\eta| < 2.0$ which is not available to the COT. The ISL is a double layer microstrip detector with small angle stereo sensors.

Central Outer Tracker

The sub-detector radially outside the ISL is the COT [68], a cylindrical open-cell drift chamber, which spans the radial space 44-132 cm from the beamline. The COT is mainly used for charged particle tracking in the central pseudo-rapidity region, $|\eta| < 1$, it has full azimuthal coverage, and spans an axial range of 310 cm. The COT has excellent tracking resolution: the hit resolution is 140 μm and the transverse momentum (p_T) resolution is σ_{pT}/p_T $0.15\% \cdot p_T$.

There are 30240 sense wires running end to end of the detector, half of which are axial, and half are stereo with an angle of 2°, giving an accurate $r - \phi$ measurement for calculating transverse momentum, p_T, but less precise information in the $r - z$ plane. The sense wires are arranged in 96 layers radially, which are grouped into 8 *super layers*, as shown in Figure 2.8, numbered 1-8 from the centre. The super layers are divided into

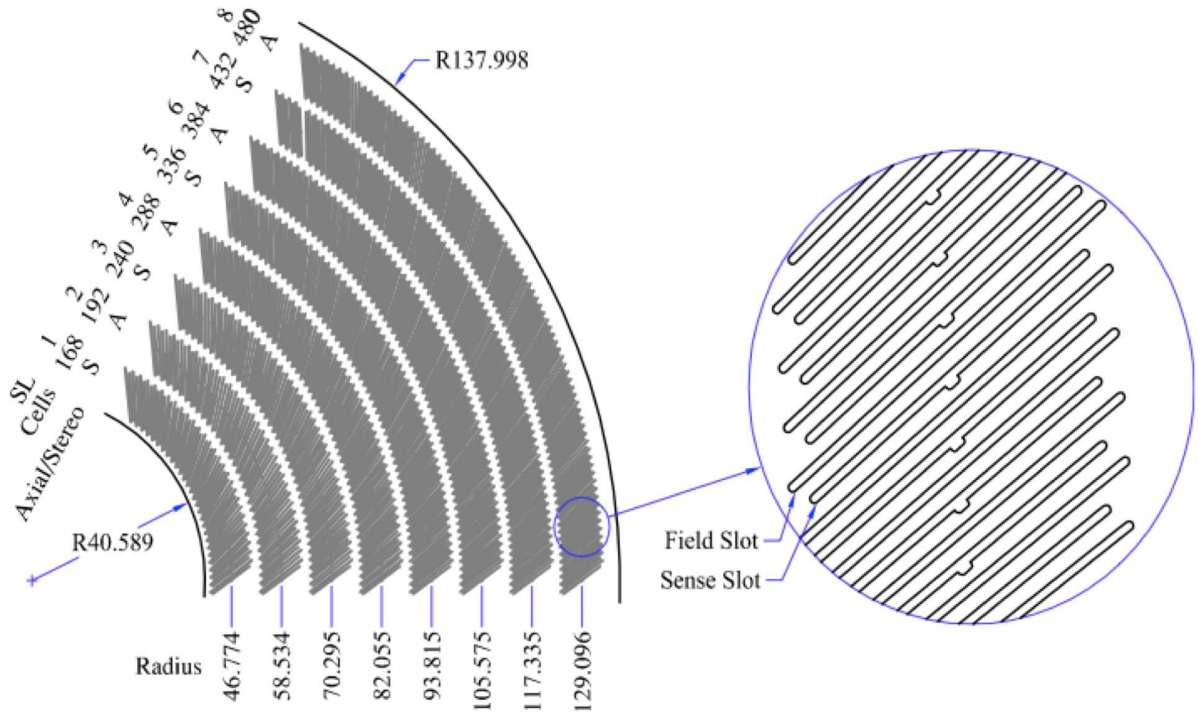

Figure 2.8: Layout of the COT super layers, showing 1/6 of the endplate, dimensions in cm. Sense and field slot format is shown in enlargement. [68]

supercells, consisting of 12 sense wires, the layout of these is shown in Figure 2.9. The number of supercells in a super layer scales with distance from the beampipe, i.e. the cells are of constant size. In each supercell there is a wire plane which contains as well as sense wires, potential wires which are used to shape the field. Both types of wires are made of gold plated tungsten with 40 μm diameter, and each supercell is surrounded by a 6.35 μm Mylar cathode with vapour deposited gold on each side, which is shared between adjacent cells.

The COT is filled with a 50:50 mix of argon and ethane gases plus isopropyl alcohol. Charged particles pass through the chamber ionise the gas leaving a trail of electrons, which are attracted towards the sensor wires by the electric field between the potential wires and the cathodes. In order to compensate for the Lorentz angle of electrons drifting in the magnetic filed of the solenoid, each supercell is angled 35° with respect to the radial direction.

In addition to tracking, the COT is used to measure charged particle ionisation energy loss. As drifting electrons get closer to the surface of the wires, they are accelerated by the local electric field, which causes an avalanche due to secondary ionisation. The effect of this is a signal on the sensor wire which is processed out by the attached readout electronics, an ASDQ (amplifier, shaper, discriminator, charge encoded) chip [69]. The width of the digital pulse encodes the amount of charge collected, which after calibration to remove any kinematic or environmental dependences, gives the ionisation energy loss of a particle, dE/dx. This is an important parameter in particle identification (PID) and will be discussed in more detail in Chapter 3. The leading edge of the pulse gives timing information for a track.

The Tevatron has surpassed the Run-II luminosity expectations, and in 2007 it was

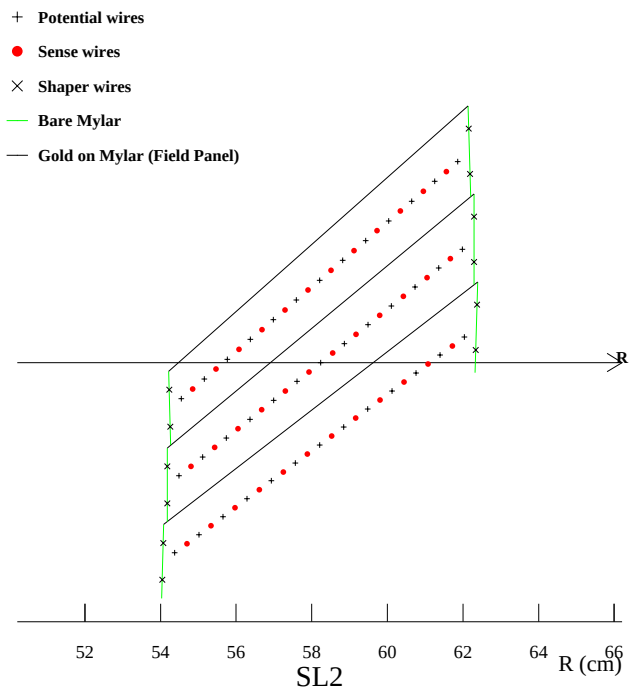

Figure 2.9: Sketch of three supercells in super layer 2, looking along the Z direction

necessary to take the effect of the high luminosity regime on the COT sub-detector. In order to minimise the radiation effects, and avoid overwhelming the readout electronics (event selection and triggers to be discussed later in this section), the dE/dx readout functionality of the inner two super layers of the COT was switched off. The dE/dx measured from the COT before and after this change needed to be carefully studied and recalibrated, the full details of which are detailed in Chapter 3.

2.2.2 Time of Flight

The use of the COT in particle identification has been introduced; the detector component radially outside of the COT is the Time of Flight (TOF) system, which exists to enhance PID of low momentum particles. The primary purpose of the TOF detector is to identify charged kaons in order to improve B meson flavour tagging, it is also used in certain applications in the CDF trigger system, such as cosmic ray, highly ionizing particle and high multiplicity triggers. The triggers used in CDF are the first level of event selection, and will be discussed in detail later in this section. The flight time of a particle combined with its momentum can give an estimate of its mass, which can be used to identify the particle type.

The TOF detector consists of a single barrel of 216 Bircon scintillator bars, 279 cm long with a cross section of 4×4 cm, within the magnetic field of the solenoid. The bars are arranged radially around the COT, and run parallel to the beampipe. The TOF system lies 138 cm from the beampipe, which is approximately 5 ns flight time for the fastest particle types, and covers the pseudorapidity region $|\eta| < 1$. Time of flight is defined as the time of arrival at the TOF scintillator, minus the collision time, t_0. Each end of every bar is read out by a fine mesh photomultiplier tube (PMT), with a total of 438 PMTs. When a particle passes through the scintillators, it leaves a deposit of energy which causes a flash of light which is detected and amplified by the PMTs.

The timing difference read out by PMTs at two ends of a scintillator bar indicates the z position of the particle. As well as a timing measurement, the PMTs readout a measurement of charge, which is used in triggering, as well as in calibrating the TOF. The resolution is of order 100 ps, which gives a 2σ separation between charged, low p_T pions and kaons. The use of time of flight as a particle ID variable will be described in Chapter 3. As a PID tool, TOF is complementary to dE/dx, as its strength lies in separating low momentum particles, in the range $p < 1.6$ GeV/c, which is poorly covered by dE/dx.

2.2.3 Solenoid

The tracking and time of flight detectors are immersed in a 1.4 Tesla magnet which is generated by a superconducting solenoid. Charged particles moving in a magnetic field travel in a helix, the curvature of which can determine the momentum of a particle from its tracks, and the sign of the curvature gives the charge of the particle. The solenoid coil is made from single layer aluminium stabilised monolithic NbTi/Cu, it is supported by an aluminium cylinder which lies outside of the coil [70]. To achieve superconductivity, the solenoid is cooled with liquid helium to around 4.7 K.

2.2.4 Calorimeters

The detector components outside of the solenoid are not used in tracking, therefore do not need the magnetic field to bend charged particle trajectories. The next parts of the detector, are the calorimeters which are used to measure the energy of particles stopped by them, and separate electrons and photons from hadrons. Muons, which are minimally ionising and therefore deposit only a small fraction of energy passing through the calorimeter material, leave little trace in the calorimeters and pass through to the muon detectors which are described in the next section. Calorimeters are particularly important in identifying neutral particles, which do not leave tracks in the inner detectors. They are also essential in deducing the likelihood of the presence of a neutrino in an event, as neutrinos do not interact with the detector but can be observed by the absence of energy that should be present due to energy-momentum conservation, called *missing energy*.

CDF has two types of calorimeter: electromagnetic (EM) which mainly absorb photon and electron energy through EM interactions and hadronic (HA) which absorb energy from hadrons via the strong interaction. The central EM and HA calorimeters, the CEM and CHA, surround the tracking system radially, and cover the pseudo-rapidity range $|\eta| < 1.1$ [71]. Coverage in η is extended by the plug calorimeters, PEM and PHA, to $|\eta| < 3.6$, and the region between the central and plug areas is bridged by the WHA [72].

The calorimeters are made up of alternating layers, of a scintillator which absorbs energy and emits light, and a passive metal which the incident particles interact with and lose energy to. In the central calorimeter these are arranged in wedges consisting of lead-scintillator layers for the EM section and a steel-scintillator hadron section. The wedges are segmented into towers, each of which covers about 15° in ϕ and 0.11 in η. The plug calorimeters are also segmented in towers, these vary in $\eta - \phi$ coverage. The PEM, like the CEM, uses lead interspersed with the plastic scintillator layers, the PHA consists of iron-scintillator layers and the WHA layers of steel and scintillator. Read out is done via PMTs which amplify the signal from the absorbed particles.

The EM calorimeters rely on Bremsstrahlung and photon absorption in the lead absorber to create showers of particles; the energy of these showers is proportional to the energy of the initial particle. Each lead layer absorbs a fraction of the shower, creating a further cascade of EM interacting particles which deposit energy in the next scintillator layer.

Particles are stopped in the hadron calorimeter by interactions with nuclei in the iron layers. As the nuclear interaction cross section is significantly lower than the EM cross section, the hadron calorimeters need much more material to stop particles which travel through it. The main fraction of the calorimeter size consists of the hadron sections, which lie outside the EM calorimeter sections.

The EM calorimeters are enhanced by *showermax* detectors, CES and PES, which are gas filled wire and strip chambers. These give position measurements which can be matched to tracks, and a transverse profile of the shower, to separate photons from neutral pions. Additionally, preshower scintillator tile chambers (CPR and PPR) are positioned on the front of the central calorimeter wedges and the first layer of the PEM. These improve soft (low momentum) photon and electron identification.

2.2.5 Muon detectors

The furthest detector component from the beampipe is the muon system. It was explained in the previous section that as minimally ionising particles, muons pass through the rest of the detector losing little energy, and if they have sufficient momentum they reach the muon drift chambers. A particle entering the muon chambers leaves a track which is registered as a *muon stub*. As they are charged, muons leave tracks in the COT; if a COT track is matched to a muon stub these can be combined to make up a muon candidate. The design of the detector is such that other types of particle than muons should be absorbed by the material between the beam pipe and the first of the muon detectors.

Muon identification is important for the analysis described in this thesis; the dimuon (J/ψ) trigger is used to select $J/\psi \to \mu\mu$ events, which make up half of the final state particles of the $B_s \to J/\psi\phi$ decay. The CDF triggers will be explained in the final part of this chapter.

The CDF muon system [73] consists of several subcomponents, the main part is the Central Muon detector (CMU) which was the initial muon system of the Run I detector. During Run I, the muon system was upgraded by adding the central muon extension (CMX) and central muon upgrade (CMU) components. These components were improved and finalised for Run II. An additional section, the Intermediate Muon detector (IMU), extends the coverage in η to the forward region. The $\eta - \phi$ coverage of the different muon detector components is shown in Figure 2.2.5. Features and properties of the muon detectors are summarised in Table 2.2.5. The CMU, CMP, CMX and IMU are drift chambers, there are also three scintillators, the CSP, CSX and BSU, which are close to each drift chamber. These serve two main purposes, they are used for triggering and deliver timing information to reduce background by identifying which beam crossing produced a specific muon.

The CMU is made up of 144 modules each with 16 cells filled with the same mixture of gases as the COT. In the centre of each cell there is a 50 μm stainless steel sense

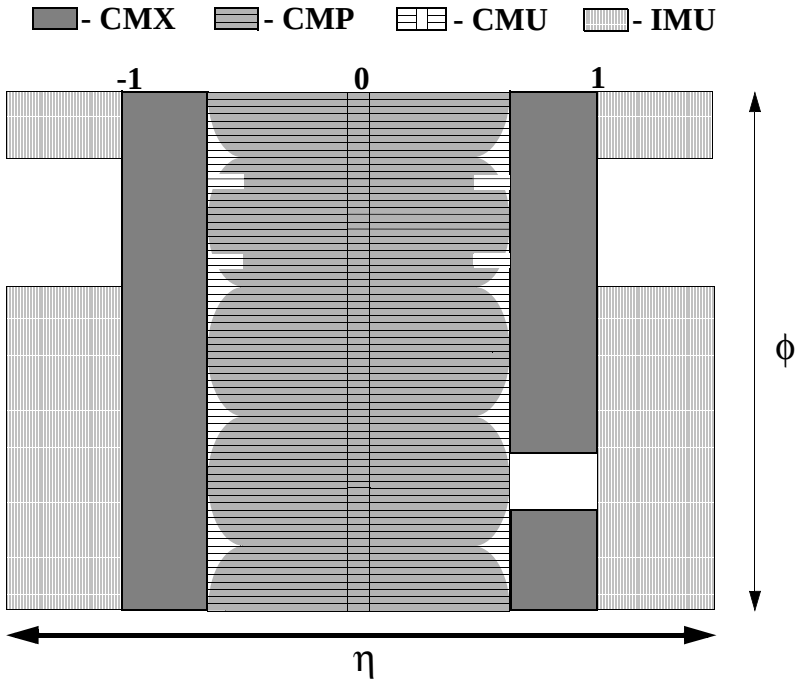

Figure 2.10: Layout of the CDF muon detector components in azimuth and pseudorapidity

	CMU	CMP	CMX	IMU
Pseudo-rapidity range	$\|\eta\| < 0.6$	$\|\eta\| < 0.6$	$0.6 < \|\eta\| < 1.0$	$1.0 < \|\eta\| < 1.5$
Drift tube cross section	2.68 x 6.35 cm	2.5 x 15 cm	2.5 x 15 cm	2.5 x 8.4 cm
Drift tube length	226 cm	640 cm	180 cm	363 cm
Minimum muon p_T	1.4 GeV/c	2.2 GeV/c	1.4 GeV/c	1.4-2.0GeV/c
Pion interaction lengths	5.5	7.8	6.2	6.2-20

Table 2.1: Some important design parameters of the CDF II muon detectors [73]

wire. The signal from these wires is read out in pairs which are slightly offset in ϕ, providing timing and amplitude measurements which give the z and ϕ position of the muon. The CMU is shielded by the CHA, which is 5.5 pion interaction lengths of absorbing material, to minimise the non-muon particles which reach this component. However, it is possible for some non-muon particles, such as high momentum pions, to reach the muon chambers, these are referred to as *punch-throughs* which result in *fake muons*.

The first upgraded muon detector, the CMP, is behind an additional 60 cm of steel, further reducing the chance of fake muon events being recorded. CMP measurements therefore improve the muon identification, and a muon candidate with hits in both the CMU and CMP (called a CMUP muon) has higher precision than one with CMU information only. The CMP is rectangular in design, therefore has variable η coverage with respect to ϕ, as shown in Figure 2.2.5. Sense wires in the CMP are read out individually, and groups of hits in nearby wires can be combined with CMU information for triggering.

The CMX is made up of conical sections with four layers of twelve drift tubes, and

is associated with the CSX set of scintillation counters. These detectors extend the η coverage of the muon system, lying at either end of the central detector components. There is no additional shielding for the CMX, but the large angle from the interaction point means that to leave a stub in the CMX, particles travel through significantly more material than those which reach the CMU.

The furthest reaching muon detector in η is the IMU, which has trigger capability for muons with $|\eta| < 1.5$ and can be used in offline reconstruction of up to $|\eta| < 2.0$ muons. The drift chambers of the IMU of the same type as in the CMP, and are arranged in barrel structures. The readout from the IMU is linked with scintillator timing information for the trigger and muon identification.

2.2.6 Luminosity detectors

CDF has two dedicated luminosity detectors, the Cherenkov Luminosity Counters (CLCs), which are innovative devices for making precision measurements in the high luminosity regime. Prior to CDF Run II, luminosity measurement at hadron machines was usually carried out with scintillating counters which recorded the number of bunch crossings with no interactions. For high luminosities this technique is not practical as the number of bunch crossings with no interactions is minimal, so it is necessary to directly measure the number of interactions. High precision luminosity measurements are essential for analyses, such as cross-section measurements, which require knowledge of the total integrated luminosity in a dataset. From the CLCs, the average number of particles per bunch crossing is measured by the amount of Cherenkov light collected, and this can be used to estimate the number of inelastic $p\bar{p}$ interactions in each bunch crossing.

The CLCs are placed at either end of the CDF detector, in the end plug calorimeters, covering the pseudo-rapidity range $3.7 < |\eta| < 4.7$ [74]. They are each made up of 48 long conical Cherenkov counters, filled with isobutane gas, arranged in 3 concentric layers about the beampipe. Small, fast PMTs are used to collect the Cherenkov light, and backgrounds such as secondary particles are excluded by setting suitable light thresholds.

2.3 Triggers and Data Acquisition

The collision rate in a hadron collider such as the Tevatron is orders of magnitude higher than the rate at which data can be recorded. The CDF Run II data acquisition system (DAQ) can write to tape at around 75-100 Hz, whereas the collision rate is 7.6 MHz, so it is essential to filter out only the interesting physics events at an early stage in the data taking. Removing minimum bias events and selecting only processes of interest reduces the amount of data to be stored to a manageable rate, for example the cross section for B physics events is about 1/1000 the total cross section for hadronic processes. This selection process is implemented via a *trigger* system using hardware and software components, and is done in real-time as data is collected.

The CDF trigger system [75] has three levels, each making a tighter selection, the data flow through each level is shown in Figure 2.11. The first level (L1) is a buffered, hardware trigger. The level 2 (L2) trigger is a combined hardware and software system. Level 3 (L3) is a software trigger which applies a more complex selection and reconstruction to events passing the hardware trigger levels. The system aims to be

deadtimeless because the buffering means that the output rate of each level matches the rate processable by the next level.

An event can be selected and written to tape by passing any of about 100 different sets of requirements in the L1, L2 and L3 triggers, these sets are called *trigger paths*. The trigger path used to select events for the analysis described in this thesis is the dimuon trigger which sets requirements for events with two muon candidates from a J/ψ decay vertex.

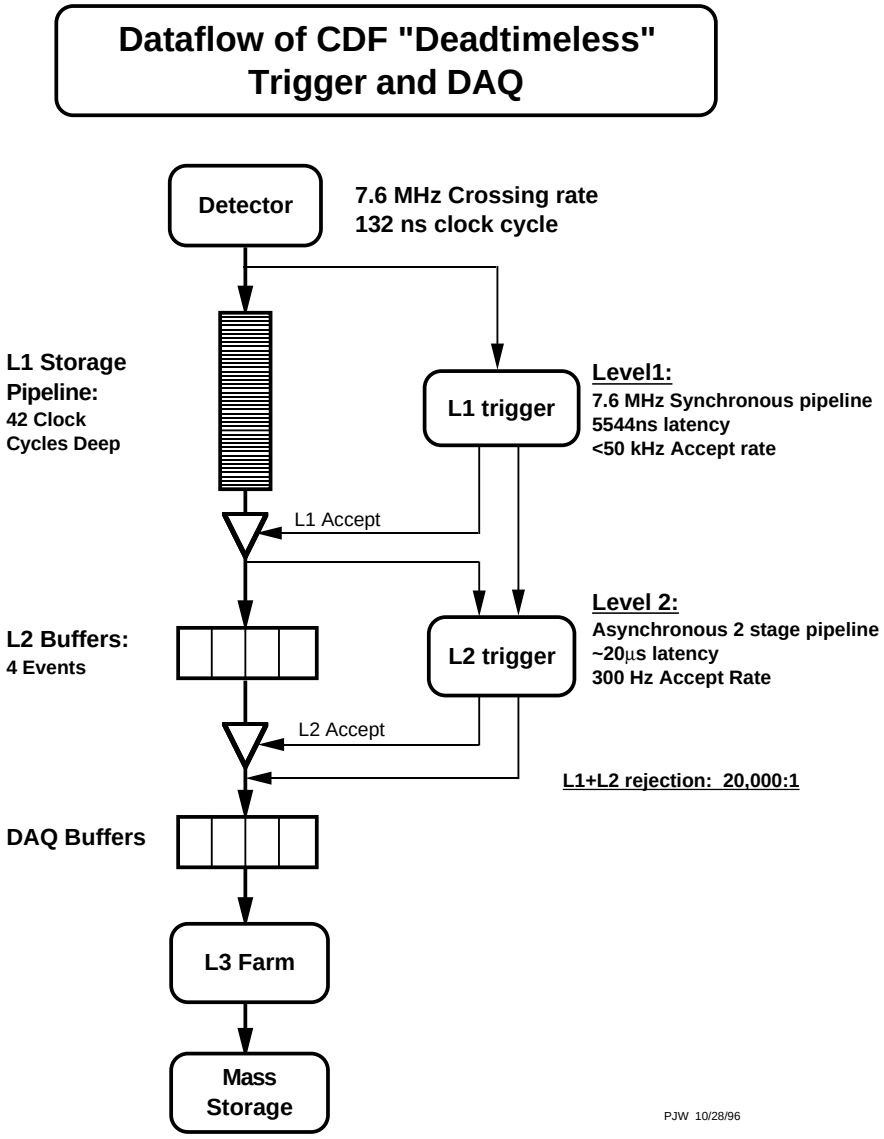

Figure 2.11: Data flow through the CDF Run II trigger system, showing the acceptance rate for each stage.

Level 1 trigger

The Level 1 trigger reads events from every bunch crossing, its decision is based on a subset of detector information which is used to find and count physics events. Figure 2.12 shows the input from each detector component to the two hardware level triggers. The hardware for the Level 1 trigger is synchronous; it takes inputs from three parallel streams, calorimeter based, muons and central tracking. Trigger paths

selecting muon and electron events require track matching between the COT and outer detector components, so information from the tracking is passed to the calorimeter and muon streams in addition to the track only stream. The three Level 1 streams are synchronised to the same clock, and a global Level 1 decision is made every 132 ns.

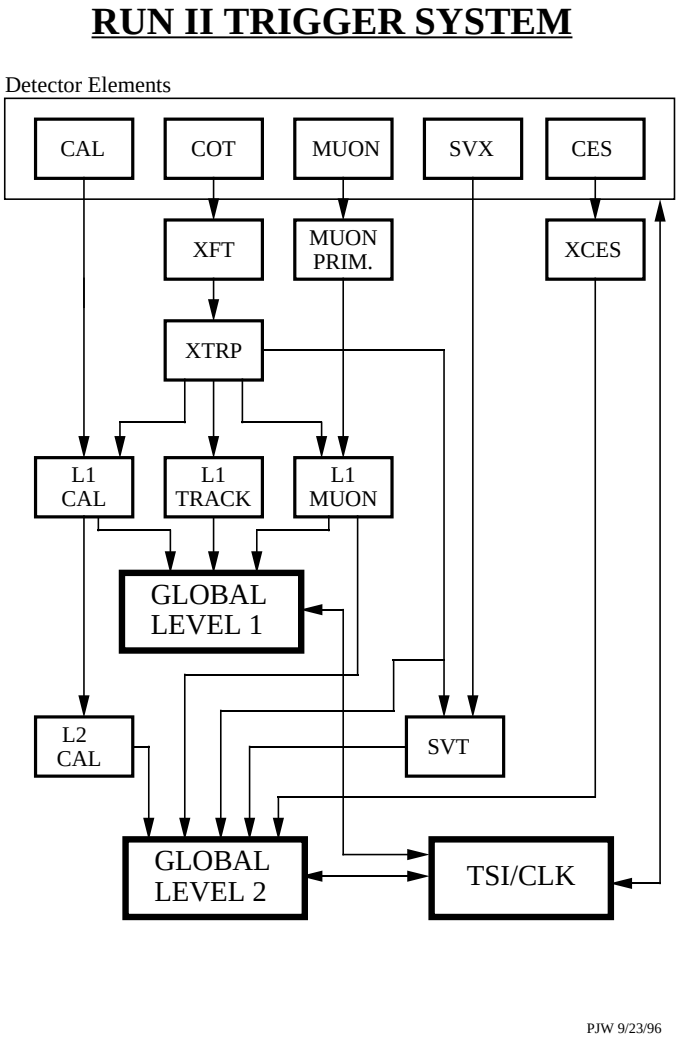

Figure 2.12: Block diagram of the CDF Run II trigger system

The Level 1 calorimetry trigger has two roles, triggering on specific objects (electrons, photons, hadron jets) and on global energy (total or missing energy). The global triggers sum information from all calorimeter towers to calculate the total energy of an event, or the missing energy according to energy/momentum conservation. There are two types of Level 1 calorimeter object triggers: single object, where a single object such as an electron from $W \rightarrow e\nu$ is accepted, and di-object where a tighter selection is required due to a higher rate for events such as $J/\psi \rightarrow ee$. The Level 1 muon trigger selects single and dimuon objects, starting with track *primitives* in each muon detector element which are then developed into a trigger decision from the combined muon system for each beam crossing. The muon (dimuon) primitives are derived from patterns of hits in the muon wire chambers, and single or coincident hits from the muon scintillators.

Triggering on tracks at Level 1 is important in identifying high momentum leptons and low momentum charged tracks. The L1 track based trigger uses the eXtremely Fast Tracker (XFT) [76] to identify in real time tracks in the COT. In order to deliver results in time for each global Level 1 trigger decision, the XFT has < 1.5 μs to find tracks. The XFT processor has three stages, using programmable logic devices (PLDs) to implement pattern recognition algorithms for hit classification, segment finding and segment linking. The Finder algorithm selects track segments with high transverse momentum (p_T) in each of the four super layers of the COT. This information is used by the Linker algorithm to match segments across the four super layers, aiming to find a match in four out of four of the layers, or three out of three of the innermost layers, consistent with a track coming from a high p_T charged particle. The minimum track p_T found by the XFT is 1.5GeV/c, this level is motivated by the fact that muons with $p_T < 1.5$GeV/c are stopped by the calorimeters before reaching the muon chambers. XFT output is used in conjunction with the EM calorimeter to identify electrons and matched to stubs in the muon chambers to select muon candidates.

The global Level 1 trigger decision is issued by the Level 1 Decision card which combines the single-bit trigger signals from the sub-components of the trigger to form the final Level 1 trigger.

Level 2 trigger

A set of asynchronous subsystems make up the Level 2 trigger hardware, giving input to the programmable Level 2 processors which perform some initial, limited, event reconstruction. A Level 1 accept leads an event to be read into one of four Level 2 buffers to be analysed by the more detailed Level 2 trigger. While Level 2 is analysing the event, that buffer cannot take further data from Level 1, so if all 4 buffers are filled it causes *deadtime*. To maintain an acceptable level of deadtime with a Level 1 rate of 50 kHz, the Level 2 processing time would need to be <16 μs, however the actual Level 2 processing time is around 20 μs. To get around this, Level 2 is pipelined in two 10 μs stages. Data is taken from the Level 1 XFT and Level 1 muon trigger systems and the showermax detector of the calorimetry system, and calorimeter data and tracking is processed simultaneously. The final stage is to check whether the processed data passes any of the trigger paths at Level 2.

The Level 2 calorimeter trigger component finds clusters of calorimeter towers with energy above a threshold value to form jets. The EM showermax detectors are used at Level 2 to reduce the background from fake electrons and photons. By matching COT tracks to showermax clusters, combinatorial backgrounds for electron triggers are greatly reduced as the resolution in the showermax is finer than the main calorimeter wedges. The muon trigger at Level 2 uses finer resolution angular bins than at Level 1, and improves on the muon selection precision.

The Silicon Vertex Tracker (SVT) has particular importance for B physics, and for certain electroweak measurements with b-jet signatures. A displaced secondary vertex in the Silicon detectors indicates the decay of a relatively long lived particle such as a meson containing a b quark. Reconstructing the secondary vertex is too time consuming at the trigger level, but this can be approximated by selecting events with a large impact parameter with respect to the $p-\bar{p}$ interaction point. The capacity to trigger on impact-parameter at Level 2 opens up many B decay channels which would be virtually unseen in other triggers. However as discussed previously the decay channel of interest in this

thesis, $B_s^0 \to J/\psi\phi$, is selected via the dimuon trigger; the displaced vertex trigger is not used to select the dataset for this analysis.

The structure of the SVT reflects that of the silicon detectors themselves [77], having a 12-fold azimuthal symmetry, thus each phi segment of the detector is processed separately. For events which pass the Level 1 XFT trigger, the SVT links the COT hits with deposits of charge in the silicon detector by comparison with a large set of predefined *patterns* which indicate the most probable coincidences. The track output is combined to produce an impact parameter with 50 μm resolution when convoluted with the beam spot.

Level 3 trigger

The output rate of Level 2 is about 350 Hz, which is reduced to 100 Hz after Level 3 selection is completed. After being accepted by the Level 2 trigger, event fragments are assembled by the DAQ system and processed by the software trigger Level 3 which is a farm of 256 dual processor Linux PCs [78]. The Level 3 trigger performs event reconstruction using algorithms which are almost the same as those used offline in order to give an accept/reject decision on every possible trigger path for each event.

Chapter 3

Data selection and particle ID

This chapter describes the selection and reconstruction of the data sample used in the measurement of $\beta_s^{J/\psi\phi}$. Also explained is the use of particle ID, and the comprehensive calibration of dE/dx, which is used in the separation of pions and kaons. PID is of particular importance for selecting kaons from $\phi \to K^+K^-$, and in kaon based flavour tagging. The Neural Network used to make the main selection of events, its training, and optimisation are explained. The flavour tagging algorithms used to separate B_s^0 and $\bar{B}_s^0$ events, and their calibration are also described. Finally, the Monte Carlo simulated data sample used in several parts of the analysis is described.

3.1 Data sample

The data used in this measurement was collected using the CDF Run-II detector between February 2002-July 2009, and corresponds to an integrated luminosity of $\mathcal{L}$=5.2 fb^{-1}. The modes reconstructed are:

- $B_s^0 \to J/\psi\,\phi$, $J/\psi \to \mu^+\mu^-$, $\phi \to K^+K^-$,
  for the measurement of $\beta_s^{J/\psi\phi}$
- $B^+ \to J/\psi K^+$, $J/\psi \to \mu^+\mu^-$
  for the validation and calibration of the Opposite Side Tagger (OST) used in flavour tagging of the initial B meson flavour, which will be described in Chapter 4.

The online selection is done using the dimuon triggers, which select any event passing the trigger level muon ID cuts for events coming from $J/\psi \to \mu\mu$. This allows a large proportion of events from non-B decays to enter the sample, such as prompt J/ψ from $p\bar{p}$ directly, which are not of interest for this analysis. These are called *background* events, while the $B_s^0 \to J/\psi\,\phi$ events are referred to as *signal*. It is not possible to remove all background events from the sample, but a much tighter selection is required than the trigger in order to perform the analysis. The selection is applied in two stages; after event reconstruction the events are passed through preselection, a loose set of cuts designed to initially focus on the signal region, then the preselected data is fed into a Neural Network to perform high level background suppression.

The following subsections describe the event reconstruction and selection processes used to obtain the analysis sample, and the optimisation and calibrations performed at each stage.

3.2 Trigger requirements

The dimuon, or J/ψ, trigger sample is selected based on trigger requirements at Level 1 and Level 3. To pass Level 1, the event must have two muon stubs (see Section 2.2.5), either both in the CMU, or one in the CMU and one in the CMX. The muon stubs have to be matched to an XFT track, for the CMU stubs this must have p_T> 1.5 GeV/c, for CMX $P_T > 2.0$GeV/c. The Level 2 and 3 triggers further impose that the pair of muons must have opposite charge and ensure that track matching requirements are met. Additionally, to check that the two muons are from the same decaying particle, there is a condition on the position in z of the two tracks of $|z_{track1} - z_{track2}| < 5$ cm at their closest point to the origin. The invariant mass of the dimuon pair is selected in a window around the J/ψ mass, of $2.7 < m_{\mu\mu} < 4$ GeV/c^2.

3.3 Event reconstuction and variables of interest

Events which pass the trigger requirements described in the previous section are written to tape, and can then be analysed in more detail offline. At this stage it is useful to define certain quantities which are used in the selection process.

Charged particles in the CDF tracking volume move with helical trajectories due to the magnetic field of the solenoid. There are five parameters defined to reconstruct the path, or *track* of a particle at CDF

- **C**
  half-curvature of the helix, $C \equiv \frac{q}{2R}$ where q is the charge of the particle and R is the radius of the helix.

- $\cot\theta$
  helix pitch, θ is the polar angle at the point where the track is closest to the z-axis, which can be related to momentum as $\cot\theta = p_z/p_T$

- $\boldsymbol{\phi_0}$
  direction in ϕ of the track where it is closest to the z-axis

- $\mathbf{d_0}$
  signed impact parameter of the track, this is the point of closest approach to the z axis, given by $q[\hat{p} \times d] \cdot \hat{z}$, where q is the charge, $\hat{p}$ and $\hat{z}$ are the unit vectors in the momentum and z directions, and d is the vector from the interaction point to the closest point in the track.

- $\mathbf{z_0}$
  the position in z at the point where the track is closest to the z-axis.

The transverse momentum, p_T is given by $p_T = cB/(2|C|)$, where c is the speed of light and B is the magnetic field of the solenoid. The parameters above are extracted from a fit to hits in the COT and SVX. The *primary vertex* is the point of the $p\bar{p}$ collisions. The *secondary vertex* is a point away from the $p\bar{p}$ interaction region where tracks intersect.

J/ψ reconstruction

As explained in the previous section, events are selected using the dimuon trigger, which passes only events with opposite sign dimuon pairs. These pairs are analysed and fitted to a common vertex to obtain a χ^2, estimated vertex position and vertex-constrained tracks [79]. Important quantities such as the J/ψ invariant mass and p_T are then estimated from the refitted tracks.

ϕ reconstruction

To find ϕ candidates, oppositely charged pairs of non-muon tracks coming from a displaced vertex fitted by a kinematic fitting algorithm [79] are examined if they are within events containing a J/ψ. The two tracks are initially assumed to be kaons. At a later stage, a probability for them to actually be kaons is assigned based on dE/dx and TOF.

Variables of interest

After the 4-track vertex fit is performed, the best fit values of the B_s^0 momenta and its daughter particles are obtained. These are used to calculate the angular variables described in Section 1.5 which are used in the separation of the CP eigenstates. One of the most important calculated variables in the study of $B_s^0 \to J/\psi\,\phi$ is the B-meson proper decay length, which is the time to decay in the hadron's rest frame $\times$ c (the speed of light). This is constructed from the primary vertex (the production point of the B meson) and the secondary vertex or decay point. The event tracks are used to locate the secondary vertex and reconstruct the B meson and its four-momentum.

The proper decay length is calculated in the transverse plane, using the transverse decay length, L_{xy} of the B meson

$$L_{xy} = \frac{\vec{V} \cdot \vec{p}_T}{|\vec{p}_T|} \tag{3.1}$$

where $\vec{V}$ is the 2D vector from the primary to the secondary vertex, and $\vec{p}_T$ is the transverse momentum vector. From this, the proper decay length is

$$c\tau = \frac{M L_{xy}}{p_T} \tag{3.2}$$

Where M is the world average B meson mass. Associated with the proper decay length is its event by event uncertainty $\sigma_{c\tau}$ which is obtained from the error on L_{xy}, the components from the other parameters being treated as negligible [80].

The transversity angles were defined in Section 1.5.2. Using those definitions they can be calculated by first boosting the four-momenta of the decay particles into the B_s^0 meson rest frame. The world average B_s^0 mass and the reconstructed momentum of the B_s^0 are used to calculate this boost vector. To then boost into the J/ψ frame the world average J/ψ mass is used, together with the reconstructed J/ψ momentum, and to calculate the boost vector into the ϕ meson rest frame the reconstructed K^+K^- mass and momentum are used. The use of the world average or the reconstructed mass for each particle type is motivated by the natural width of the particle. For the case of the ϕ meson, its natural width is close to the resolution of the CDF detector, so

the reconstructed mass is used, but the B_s^0 and J/ψ have widths far smaller than the resolution so it is more accurate to use the world average mass.

The helicity angle ψ, of the K^+ is defined as

$$\cos\psi = -\frac{\vec{p}^{\,\phi}_{K^+} \cdot \vec{p}^{\,\phi}_{J/\psi}}{|\vec{p}^{\,\phi}_{K^+}| \cdot |\vec{p}^{\,\phi}_{J/\psi}|}. \tag{3.3}$$

where $\vec{p}_A^{\,B}$ is the three momentum of particle A in the rest frame of particle B. A coordinate system is defined in order to calculate the other two angles

$$\begin{aligned}
\hat{x} &= \frac{\vec{p}^{\,J/\psi}_{\phi}}{|\vec{p}^{\,J/\psi}_{\phi}|}, \\
\hat{y} &= \frac{\vec{p}^{\,J/\psi}_{K^+} - \left(\vec{p}^{\,J/\psi}_{K^+} \cdot \hat{x}\right)\hat{x}}{|\vec{p}^{\,J/\psi}_{K^+} - \left(\vec{p}^{\,J/\psi}_{K^+} \cdot \hat{x}\right)\hat{x}|}, \\
\hat{z} &= \hat{x} \times \hat{y}.
\end{aligned} \tag{3.4}$$

This is used to calculate ψ and $\cos\theta$

$$\cos\theta = \frac{\vec{p}^{\,J/\psi}_{\mu^+}}{|\vec{p}^{\,J/\psi}_{\mu^+}|} \cdot \hat{z}, \tag{3.5}$$

$$\phi = \tan^{-1}\left(\left(\frac{\vec{p}^{\,J/\psi}_{\mu^+}}{|\vec{p}^{\,J/\psi}_{\mu^+}|} \cdot \hat{y}\right) / \left(\frac{\vec{p}^{\,J/\psi}_{\mu^+}}{|\vec{p}^{\,J/\psi}_{\mu^+}|} \cdot \hat{x}\right)\right), \tag{3.6}$$

The signs of $\vec{p}^{\,J/\psi}_{\mu^+} \cdot \hat{x}$ and $\vec{p}^{\,J/\psi}_{\mu^+} \cdot \hat{y}$ are used to resolve the ambiguity of the angle ϕ.

Both the preselection and final NN selection require the following variables to assess the probability of a specific particle type hypothesis:

- $\chi^2_{r\phi}$ **(p)**
  The χ^2 for the kinematic fit for particle **p** in the transverse plane, including the relevant mass and topology constraints.

- **P(χ^2,p)**
  The χ^2 probability of the kinematic fit, based on the full χ^2 with the z-direction included.

- **LL$_\mu$(p)**
  Likelihood based quantity for muon ID, developed for the soft muon tagger [81], which uses track-stub matching parameters and calorimeter information to discriminate real from fake muons.

- **LL$_K$(p)**
  Likelihood based combined particle ID (PID) discriminant, using TOF and dE/dx

information specifically for kaon selection. Constructed as:

$$LL_K(p) = \frac{P^K_{dE/dx}(p)P^K_{TOF}(p)}{\sum_{j=\pi,K,p} f_j P^j_{dE/dx}(p)P^j_{TOF}(p)} \tag{3.7}$$

where $P^i_{dE/dx,TOF}$ are the probabilities of observing the measured dE/dx or TOF of the particle under the hypothesis that $i = \pi, K, p$. The fractions, f_i are related to the production of pions, kaons and protons in CDF. The particle ID variables will be discussed in more detail in Section 3.4

3.4 Particle ID

Particle identification (PID) plays an important role in two aspects of this analysis, as a component in the discriminating variables of the neural network selection (Section 3.6), and in selecting kaon tracks for flavour tagging (Chapter 4). The two quantities used as PID to distinguish pions from kaons at CDF were introduced in Chapter 2: Time of Flight, using the dedicated TOF detector, and dE/dx using the COT.

3.4.1 Charged particle ionisation energy loss: dE/dx

The ionisation energy loss of a charged particle as it moves through matter is given by the Bethe-Bloch formula [21]

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi N e^4}{m_e c^2 \beta^2} q^2 \left[\ln \frac{2m_e c^2 (\beta\gamma)^2}{I^2} - \beta^2 \right] \tag{3.8}$$

where N is the number density of electrons in the material of interest, e is the electron charge, m_e the electron mass, q the particle's charge, βc the particle's speed,$\gamma = 1/\sqrt{1-\beta^2}$, and I is the mean excitation energy for atoms in the material. For a material with known properties, such as the gas in the COT, this equation can yield a measurement of the particle's mass when combined with a momentum measurement, which can then be compared to known particle masses in order to estimate the particle type.

The Bethe-Bloch formula is empirically modified to better model the CDF detector, as [82]

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{1}{\beta^2} \left[c_1 \ln \left(\frac{\beta\gamma}{b + \beta\gamma} \right) + c_0 \right] + a_1(\beta - 1) + a_2(\beta - 1)^2 + C \tag{3.9}$$

where a_i, b, c_j and C are free parameters which float when fitting the data. From Equation 3.9, a *universal curve* can be plotted against $\beta\gamma$. An example of the CDF universal curve for several particle types is shown in Figure 3.1, next to a plot of the momentum dependence of measured dE/dx for different particles which demonstrates the ability to separate particle types using dE/dx and momentum. This separation when dE/dx is plotted against momentum occurs because the dependence on $\beta\gamma$ is an implicit function of mass and momentum, as $\beta = v/c$ and momentum $= mv$.

The amount of ionisation charge produced by a charged particle near a COT sense

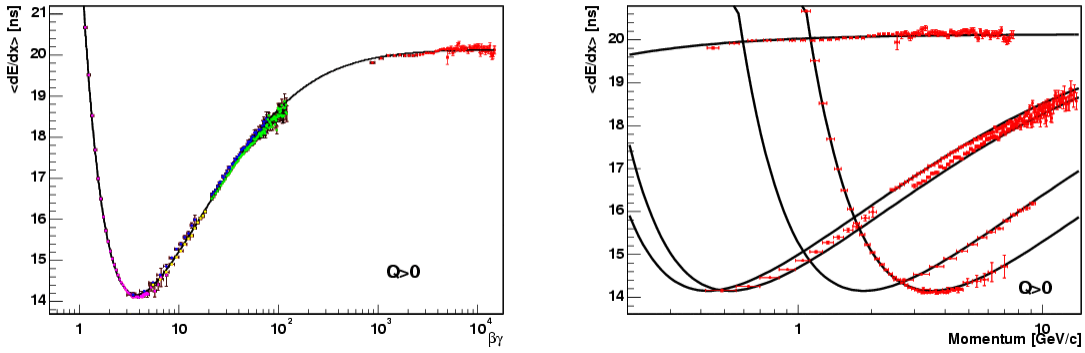

Figure 3.1: [left]Universal Curve showing combination of pions, kaons, protons and muons, [right] momentum dependence for (from left) muons, pions, kaons, protons and (top) electrons at CDF [83]

wires affects the signal strength in the wire. dE/dx is measured as the amount of charge, above a threshold value, which is proportional to the width (Δt) of the pulse from the readout chip. Thus dE/dx values are given in nanoseconds rather than a unit of energy.

dE/dx calibration

While the measured dE/dx should only depend on the boost, $\beta\gamma$, of a particle (as shown in Equation 3.8) the measurement capability of the COT for dE/dx is not perfect and introduces effects due to both environmental and kinematic variables. These dependences reduce the power of the dE/dx variable to separate between particle types, so in order to optimise PID at CDF the measured dE/dx must be calibrated to remove or minimise these dependences. The steps taken to correct the detector and kinematic effects on measured dE/dx in the COT are described in the following subsections.

Calibration data sample

To calibrate the dE/dx, pure samples of pions and of kaons are used [84]. These are obtained from D^0 decays, where the flavour of the D^0 meson is unambiguously identified by the sign of the D^* which produces it. The D^* charge is tagged by the sign of the *soft* pion π_s, a low-momentum pion, in the decay chain $D^{*+} \to D^0\pi_s^+, D^0 \to K^-\pi^+$ and its charge conjugate.

The sample is reconstructed from an impact parameter based trigger, using data taken between February 2002-April 2008. This is a dataset with high statistics, the following cuts are applied to obtain a signal sample

- $sgn(d_0(K) \times d_0(\pi))$ <0 to ensure opposite charge kaon and pion
- $L_{xy}(D^0) > 300\mu$m
- $|d_0(D^0)| < 140\mu$m
- $p_T(\pi_s) > 0.4$ GeV/c
- $d_0(\pi_s) < 500\mu$m

- $m(D^0)$ within $\pm$ 25 MeV/c^2 of world average
- $\Delta m(D^*) \equiv m(D^*) - m(D^0) - m(\pi_s)$ within 1.5 MeV/c^2 of World average

The variable $\Delta m(D^*)$ is very powerful in reducing background coming from real D^0 combined with random (wrong) soft pion tracks to form a D^* candidate. The yield, and distribution of $\Delta m(D^*)$ with all cuts applied are shown in Figure 3.2. The $\Delta m(D^*)$ distribution demonstrates the power of this variable in distinguishing between real and fake D^*, as the accidental combinations of random tracks with a real D^0 which do not come from D^* have a flat, non-peaking distribution in $\Delta m(D^*)$ and can thus be largely eliminated with a tight restriction on this value.

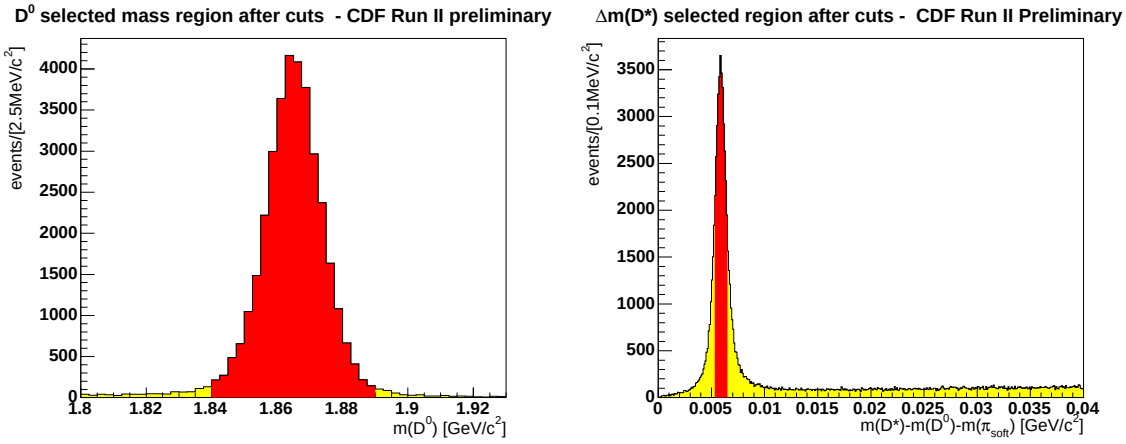

Figure 3.2: [left] D^0 invariant mass distribution after selection cuts applied [right] $m(D^*) - m(D^0) - m(\pi_s)$ distribution with cuts applied. The red part shows the region selected by the cuts.

Calibration procedure

Using the obtained high purity samples of kaons and pions, studies were carried out to assess the variation in measured dE/dx response as a function of kinematic and environmental parameters. After selecting the most significant effects, these are checked for factorisability - whether they could be corrected independently or must be treated simultaneously due to interdependence. The dependences are corrected for by calculating and applying multiplicative correction factors, either independently or simultaneously as determined in the previous step. For a useful PID variable, it is necessary to know the predicted value dependent on momentum for a given particle hypothesis, so after the calibration factors have been calculated the Universal Curve is fitted on the data after calibration, using Equation 3.9. Also used when dE/dx is included in a likelihood parameter for event selection are the resolution functions, which parameterise the residual $dE/dx_{measured} - dE/dx_{predicted}$, so these must be fitted.

Correcting the effects

The measured dE/dx shows significant dependence on six variables, to varying extents of severity. In principle these can be corrected, and so were studied to find which could be treated independently (were factorisable) or had to be simultaneously corrected, and which, if any had some momentum dependence. The size of each effect is estimated from the variations in the uncalibrated dE/dx distributions in Figure 3.3.

The decision to treat an effect independently, or to correct it simultaneously with other correlated parameters was motivated by studying the dE/dx dependence of each parameter, in slices of every other parameter in the study, as shown in Appendix A. If the distribution of dE/dx with respect to a particular parameter varied across the different slices of another parameter (other than a simple up or down shift in the whole distribution), the dE/dx dependence on those two parameters would need to be corrected for simultaneously.

The following effects are corrected for

- **Time**
  A variation in dE/dx of $\sim$ 5 ns as a function of time (indicating a variation with run number).
  The dependence of dE/dx on time is understood to be due to several properties of the drift chamber which have varied since the start of Run-II. One example of this is that in 2006, the two inner layers of the COT had dE/dx read out switched off in order to maintain tracking capabilities at high luminosity. The fact that the delivered luminosity from the Tevatron has increased significantly over the run period also affects measured dE/dx. Additionally, aging of the COT can affect the amount of charge collected for dE/dx measurement. This effect exhibits correlations with luminosity, number of COT hits, and track density (Figures in Appendix A); variations in these parameters are calibrated for in a 4-dimensional simultaneous correction.

- **Luminosity**
  A variation of $\sim$ 4 ns with respect to instantaneous luminosity.
  This interdependence is to be expected given the time dependence of luminosity itself, and the increase in occupancy of the COT in higher luminosity running will affect the number of hits, and the density of tracks. Higher occupancy in the COT means a larger number of tracks which could be in the region of the track of interest, and could contribute to the measured charge deposit for that track. This causes an artificially raised dE/dx measurement for higher luminosity events.

- **COT hits**
  An effect of $\sim$ 4 ns depending on the number of hits in the COT for a given track
  Past dE/dx calibrations [85] found a significant dependence on the number of associated COT hits for a track. This is itself correlated with other occupancy related variables.

- **Secance (track density)**
  A $\sim$ 3 ns variation dependent on the track density in the vicinity of the candidate, measured as the number of $r-\phi$ intersections of the candidate track with other tracks within the COT.
  There is no longitudinal segmentation in the wires of the COT, so the charge deposited by any hit axially in the region of the candidate track can be counted towards the dE/dx of that track. Secance is a variable which has been constructed to count the number of tracks intersecting the candidate track, which was found to be a more effective quantity than a more traditional approach such as track isolation. An assumption has to be made in correcting this parameter, that the

number of reconstructed tracks for an event is proportional to the total number of tracks, as only the reconstructed tracks are accessible in the dataset.

- η
  An effect of $\sim$ 2.5 ns according to the track pseudorapidity
  The pseudorapidity is a function of transverse momentum, and therefore there will be some intrinsic variation in measured dE/dx with η as the momentum dependence is a physical effect, η is the only variable which exhibits a significant momentum dependence. This parameter demonstrated independence from the other parameters in the study shown in Appendix A, so can be corrected for separately.

- ϕ_0
  A variation of $\sim$ 2 ns as a function of the azimuthal angle of the track
  This parameter can be corrected for independently as the effect has been shown to factorise with the other studied variables, modulo a global shift due to the change in mean dE/dx measured, in Appendix A. Variations in the measured dE/dx for particles traversing different sectors of the COT can be explained by the temperature gradient and the flow of gas within the chamber. The variable chosen to measure this effect is the track ϕ_0, which gives the initial azimuthal direction within the COT and therefore is expected to show clearly any variations which occur due to the effects mentioned.

The size of the dE/dx variations with the above parameters is shown in Figure 3.3, before and after the corrections have been applied. The size of these dependences is comparable to the average separation in dE/dx between kaons and pions of just 1.5 ns.

Correction function

In order to find the correction factors necessary to reduce the dE/dx dependence on the parameters introduced in the previous subsection, a correction function of the form:

$$dE/dx = f(\phi_0) \times g(\eta|p) \times h(\mathcal{H}, L, T, S) \tag{3.10}$$

where the segments on the right of the equation are the separate correction functions for ϕ_0 (f), pseudorapidity - conditional on momentum (g) and the 4-dimensional simultaneous correction for number of COT hits, $\mathcal{H}$, luminosity $\mathcal{L}$, time (run number) $\mathcal{T}$, and secance $\mathcal{S}$ (h).

To calculate the correction factors which go into equation 3.10 the data is divided up into bins of the above quantities, and the dE/dx dependence on each filled into histograms. The factorisable corrections use 1D histograms, the four interdependent parameters use a 4-dimensional array which is treated in the same way as a histogram. The binning is arranged such that each bin has comparable statistics. The mean dE/dx value for the combined kaon-pion sample was found to be 17.7 ns, so this value is used in re-weighting the bins. The content of bin i, $\langle dE/dx \rangle_i$ is reweighted by a multiplicative weighting factor of $w_i = 17.7/\langle dE/dx \rangle_i$ resulting in flattened histograms in each variable.

The red points in Figure 3.3 show the flattened variations in measured dE/dx after these corrections have been applied to an independent test sample, which was separated from the calibration sample to avoid effects from statistical correlations. These results

show that the variations of dE/dx with the studied parameters is reduced to < 0.5 ns after calibration. As η is dependent on momentum, there is some residual dE/dx depen-

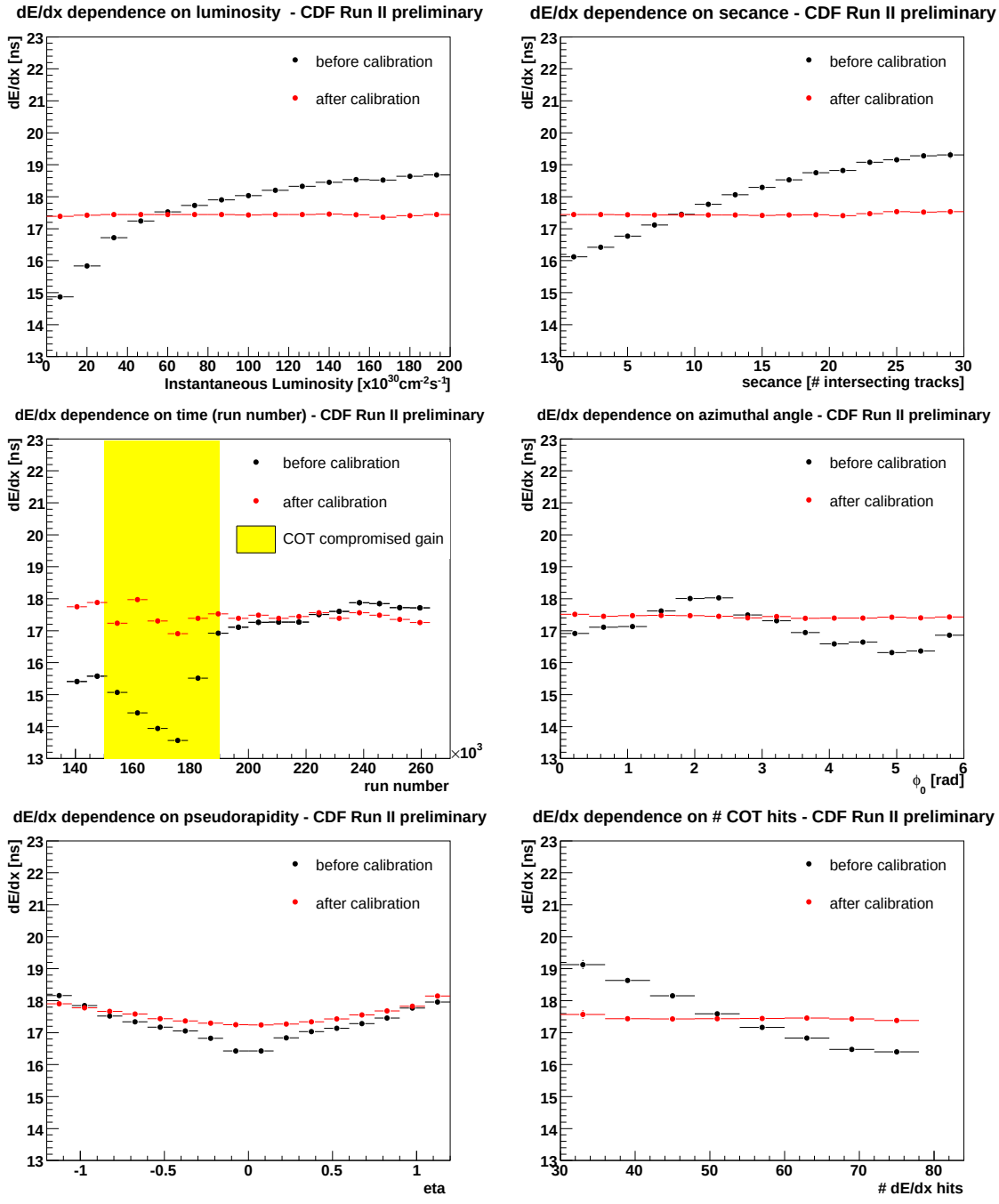

Figure 3.3: variations in measured dE/dx with several environmental and kinematic parameters before and after calibration

dence on η after the calibrations have been applied when looking at the full momentum spectrum as in Figure 3.3. This is expected, and when dE/dx is plotted against η for slices of momentum the distribution is flat; examples are shown in Figure 3.4 for the momentum ranges $2 < p < 3$ GeV/c and $5 < p < 6$ GeV/c. The remaining variations in these plots are due to the reduced statistics from slicing the dataset into momentum segments.

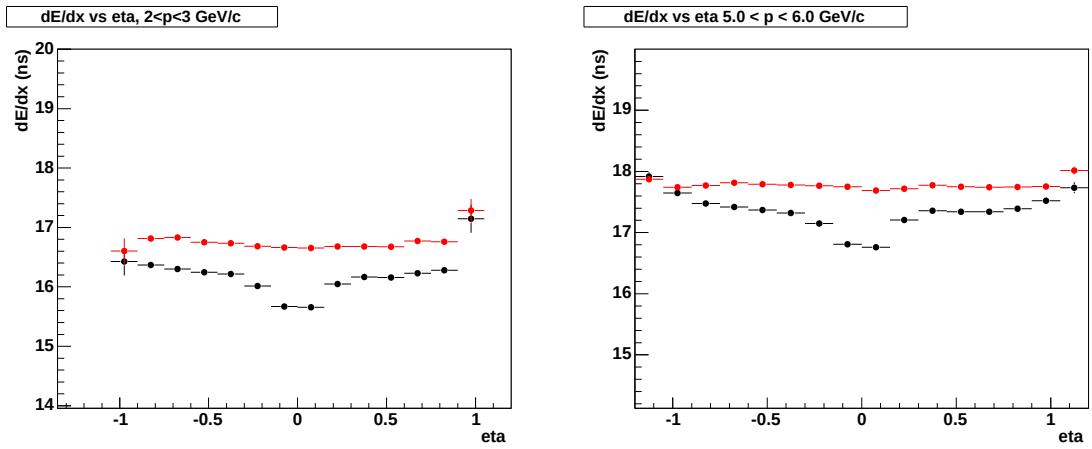

Figure 3.4: variations in measured dE/dx with η in slices of momentum, to show the affect of calibrations with momentum dependence removed. [Left] 2 < p < 3 GeV/c [Right] 5 < p < 6 GeV

Predicted dE/dx curves

To extract the curves of expected dE/dx as a function of momentum, the purity of the sample was further enhanced by applying tighter cuts on the D^0 mass (within 5 MeV of the world average) and $\Delta m(D^*)$ (within 1 MeV of the world average). This reduces the potential for any bias to enter the curves from unknown backgrounds. The dE/dx dependence on momentum curves for π^+, π^-, K^+ and K^- were fitted using the CDF empirical modification of the Bethe-Bloch formula, Equation 3.9. Figure 3.5 shows the projections of these fits onto the data points for the different particle types, and demonstrates good agreement between the fit and data across the whole momentum range.

The curves for the four studied particle types, positive and negative kaons and pions, can be compared in Figure 3.6, showing the separation between the particle species. As is clear from this plot, there is a systematic difference in the average dE/dx response of the COT for positive and negatively charged particles, and this difference is momentum dependent. The change in separation of the curves shown causes a drop-off in performance with momentum for dE/dx as a PID tool, but for the momentum ranges used in the $B_s^0 \to J/\psi\,\phi$ analysis described in this thesis.

Using the predicted dE/dx curves, it is possible to construct dE/dx residuals, $(dE/dx_{meas} - dE/dx_{predict})$ which are of interest in checking the flattening effect of the calibrations on the variations described previously. The comparison of dE/dx residuals before and after the calibrations are applied is shown in Figure 3.7. The slight improvement in the calibrated residual distribution for η compared to the calibrated dE/dx_{meas} is due to the fact that subtracting the predicted dE/dx removes any remaining momentum dependent contributions.

Distribution of dE/dx residuals

These predicted dE/dx values are also used to prepare a set of functions used in PID simulation for likelihood calculations, which are the fitted dE/dx residual distributions. These are parameterised under the assumption that the particle is a pion (the *pion*

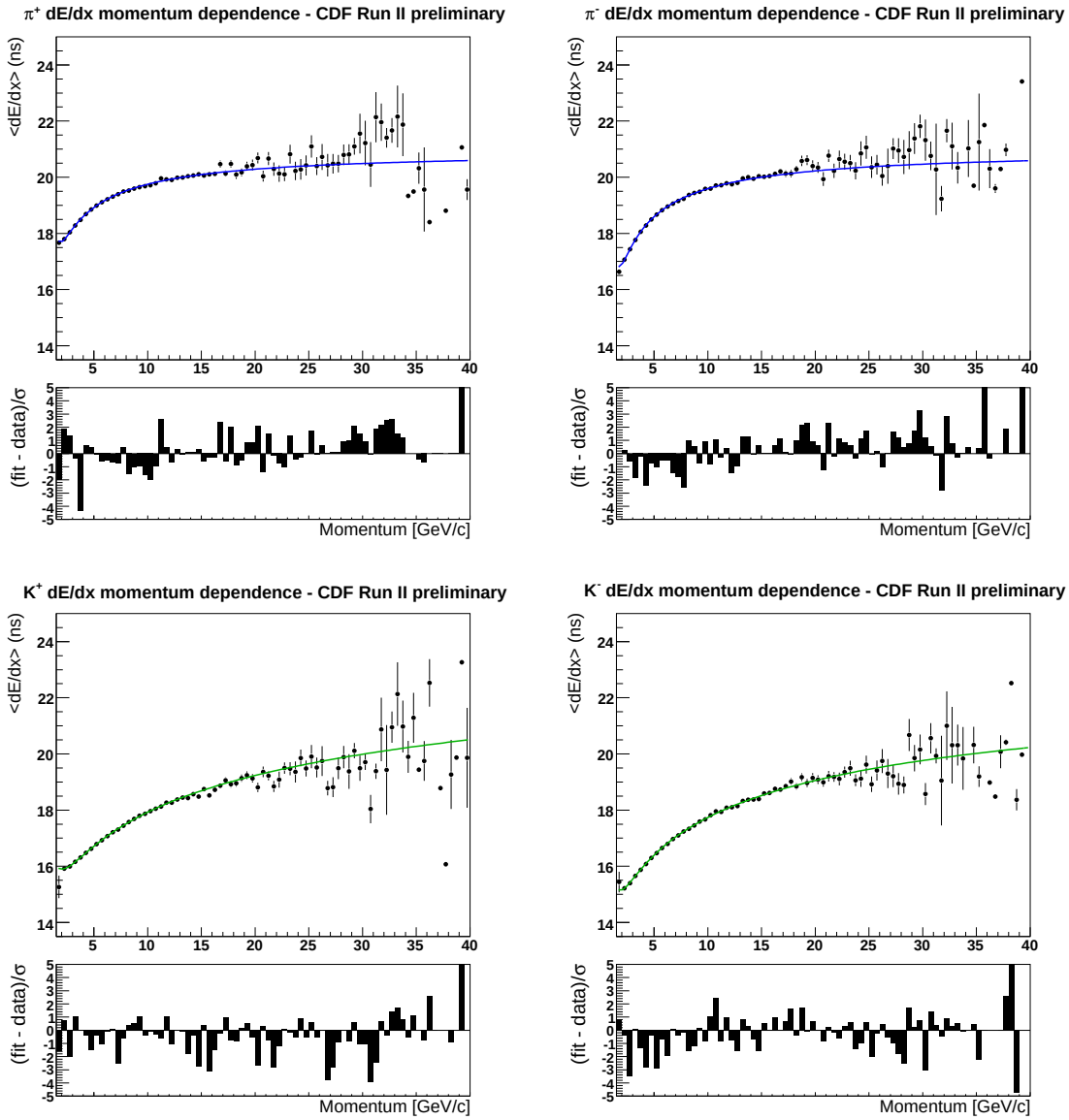

Figure 3.5: dE/dx dependence on momentum for different particle types after calibration.

hypothesis), meaning that the measured dE/dx is shown with respect to the average pion response. The functions are fitted using data for pions and kaons separately, with the predicted dE/dx value for pions subtracted from the measured, calibrated, dE/dx for each particle type. The distributions are dependent on the number of hits in the COT and the particle momentum. To account for this, the data is divided into momentum bins to fit the residual distributions, and an additional function consisting of the fitted dependence on COT hits is calculated as a multiplicative correction. The distributions of $dE/dx_{K,\pi meas} - \langle dE/dx_{\pi} \rangle$ for all momentum bins combined are shown in Figure 3.8.

Performance

To assess the effect of calibrating the dE/dx on the PID performance of this variable, the main figure of merit used is the pion-kaon separation power. This is studied by

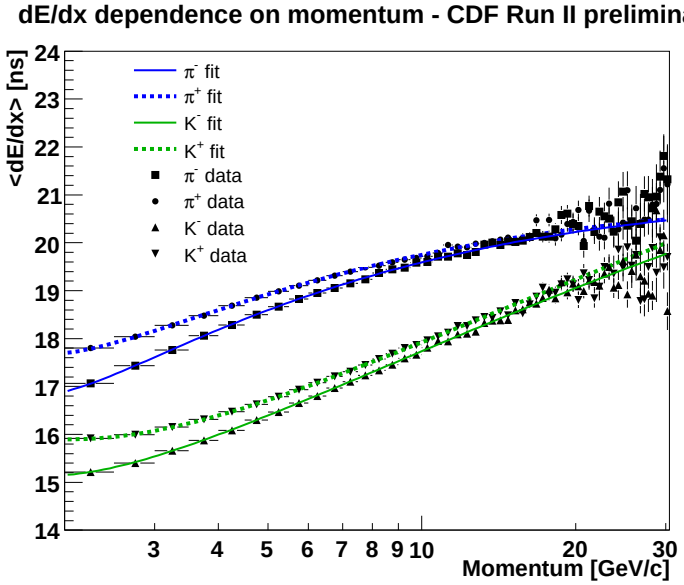

Figure 3.6: dE/dx dependence on momentum for different particle types after calibration.

looking at the separation between residual dE/dx distributions for pions and kaons, where perfect separation would mean no overlap in the distributions. The separation value is traditionally quoted in terms of Gaussian sigma, but for non-perfect Gaussian distributions this method is not appropriate. Instead, for a sample of N events, where f is the fraction of pions and $1-f$ the fraction of kaons, the statistical error on the estimate of the pion fraction, σ_f is used to determine the separation power in terms of the precision, $(1/\sigma_f)$ using [84]:

$$\sigma_f^2 = \frac{1}{N}\left[\int \frac{(p_\pi(x) - p_K(x))^2}{f p_\pi(x) + (1-f)p_K(x)} dx\right]^{-1} \tag{3.11}$$

where x is the dE/dx residual and $p_{\{\pi,K\}}$ are the probability distributions of x for pions and kaons. The upper bound on this quantity is $1/\sigma_f^{best} = \sqrt{N/(f(1-f))}$, corresponding to the ideal case where there is no overlap between the residual distributions for pions and kaons.The separation power is then quoted as a fraction of the ideal case, $s = \sigma_f^{\text{best}}/\sigma_f$:

$$s = \sqrt{f(1-f)\int \frac{(p_\pi(x) - p_K(x))^2}{f p_\pi(x) + (1-f)p_K(x)} dx}. \tag{3.12}$$

The separation, s can take any value in the range 0-1, where 1 corresponds to the perfect separation described previously. The probability density functions which enter Equation 3.12 are constructed from data residual distributions shown in Figure 3.9. The value of separation for uncalibrated dE/dx is $s = 48.8\%$, which increases after calibration to $s = 56.4\%$, a 22% improvement.

3.4.2 Time of Flight

Time of Flight is a complementary quantity to dE/dx, as it is most powerful at separating pions from kaons at low momentum, $p < 1.5$ GeV/c. Particle identification

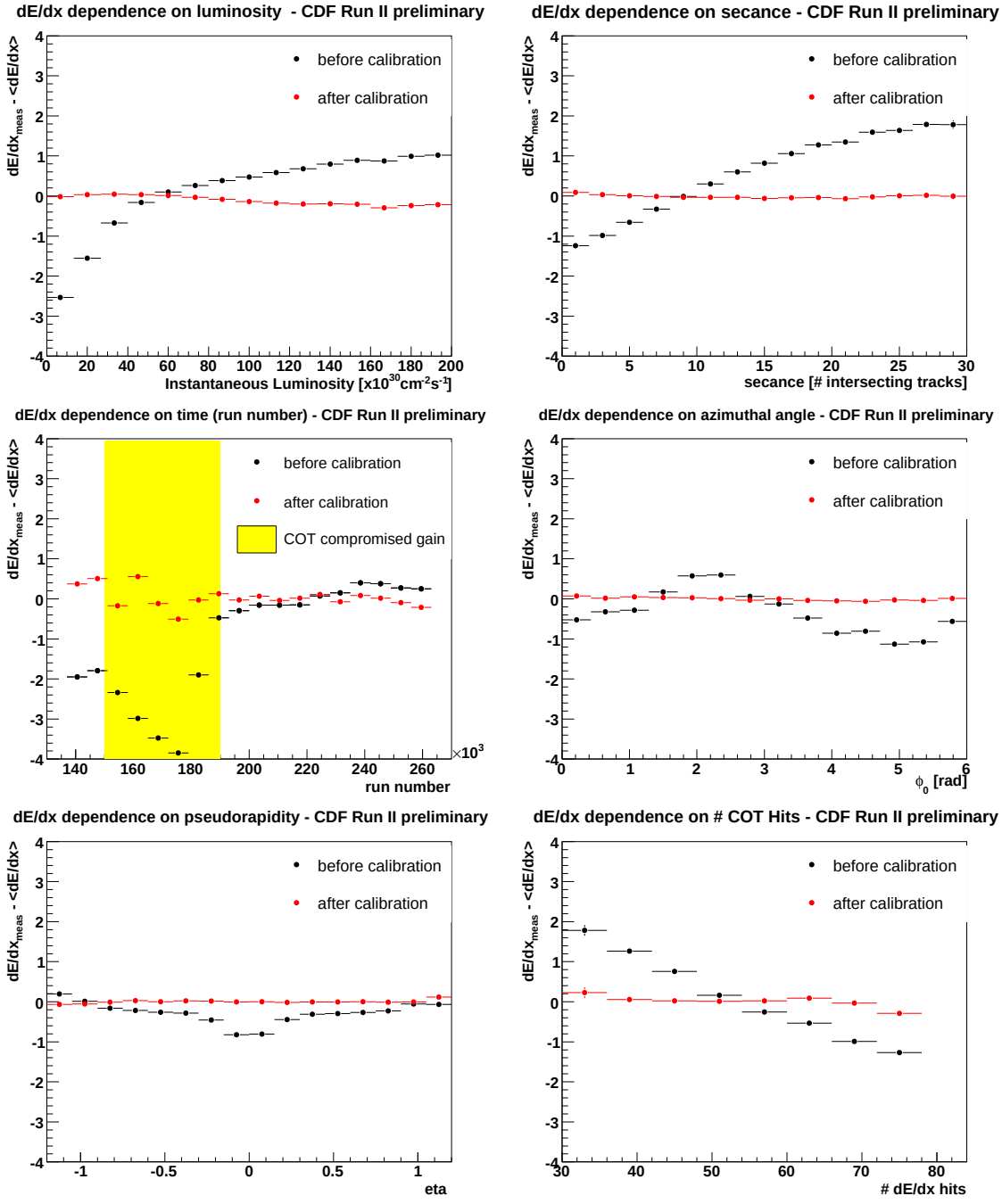

Figure 3.7: dE/dx residuals before and after calibration

using TOF utilises the relation

$$m = \frac{p}{c}\sqrt{\frac{c^2 t_{flight}^2 - 1}{L}} \tag{3.13}$$

where m is the predicted mass (which can be used to identify a particle by comparison with known particle masses), c is the speed of light, L is the track length and p is the particle momentum. The time of flight, t_{flight} is measured as the difference between the arrival time in the TOF detector scintillators, t_{TOF} and the beam crossing time, t_0.

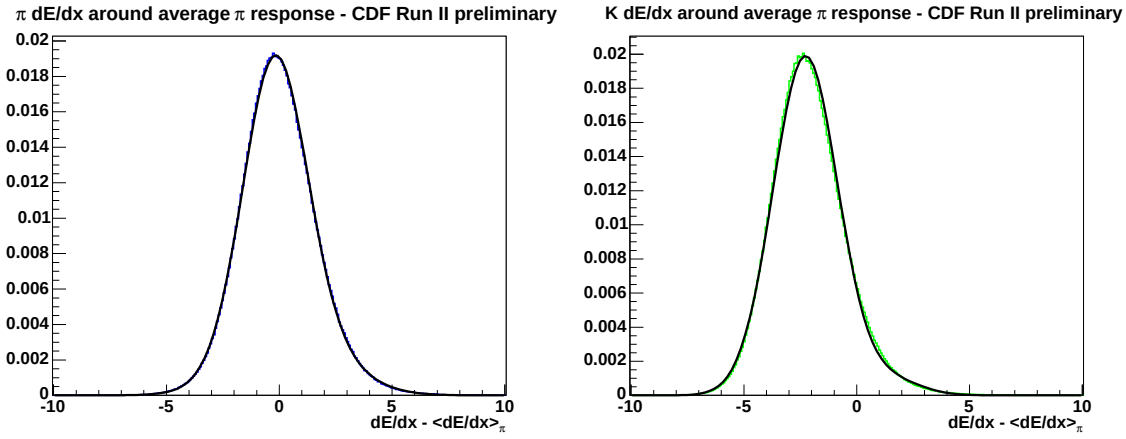

Figure 3.8: dE/dx dependences before and after calibration

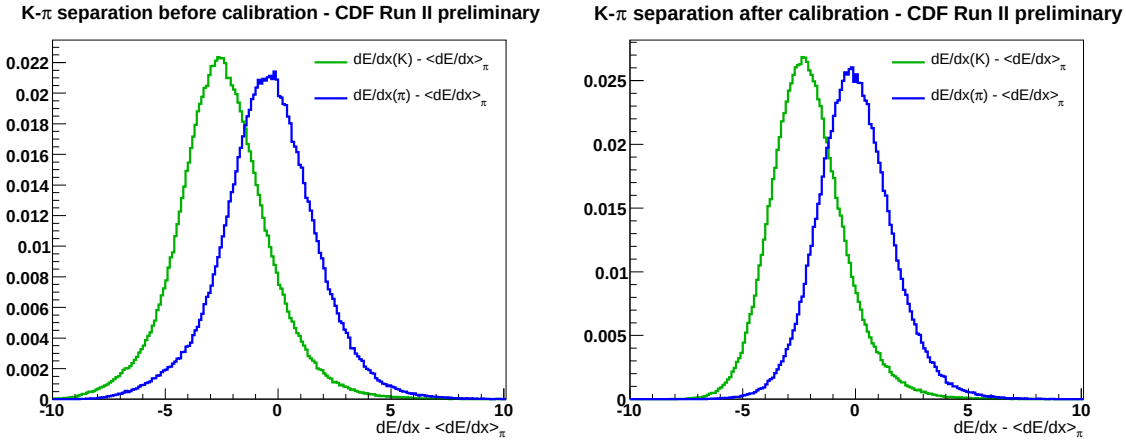

Figure 3.9: Data distributions for dE/dx residuals before and after calibration

Figure 3.10 shows an example of the TOF distribution for different particles, demonstrating the separation according to mass, and the separation power for several particle types compared to the $K - \pi$ separation power of dE/dx.

Like dE/dx, TOF should be dependent only on the mass and momentum of a particle, however the detector and event kinematics can also affect this quantity. The measured TOF therefore must be calibrated in order to obtain optimal separation between particle species. This is done using pure samples of each particle type separately; the full calibration method for this variable is described in [87].

3.5 Preselection

Loose selection requirements are applied to reduce the sample size before running the neural network and ensure reasonable candidates are included. These *preselection cuts* are:

- $5.1 < J/\psi\phi$ invariant mass < 5.6 GeV/c^2, this window is set to be 100 MeV around the world average B_s^0 mass, a range large enough to avoid throwing away

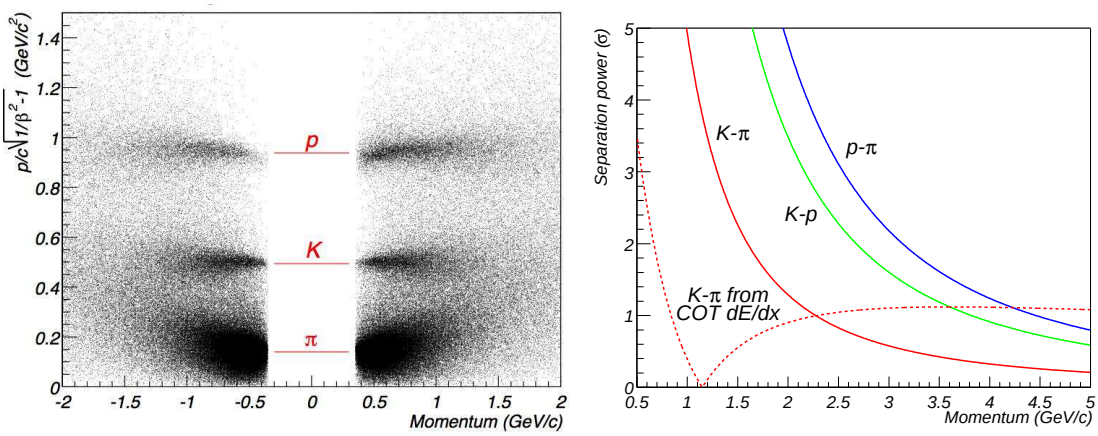

Figure 3.10: [left] TOF distribution of different particles, [right] separation power for TOF compared to that of dE/dx [86]

any $B_s^0 \to J/\psi\,\phi$ signal events

- $p_T(B_s^0) \geq 4.0$ GeV/c
- $\chi^2_{r\phi} < 50$ for kinematic fit of four tracks constrained to come from secondary vertex
- ≥ 10 axial and ≥ 10 stereo hits in the COT per track
- ≥ 3 axial hits per track in the silicon detector
- $p_T(K) \geq 400$ MeV/c for each kaon
- $p_T(\phi) \geq 1$ GeV/c
- $3.04 <$ $\mu\mu$ invariant mass < 3.14 GeV/c^2 to constrain to the J/ψ mass
- $1.009 <$ KK invariant mass < 1.028 GeV/c^2 to constrain to ϕ mass

3.6 Neural Network

Rectangular cuts, such as those used in the preselection, veto any events which do not fall into an accepted range for any given variable. This means that an individual variable can determine whether an event is signal or background. An improved selection method is to use an artificial neural network (NN), which takes information from all input variables and combines it into a single decision variable in which the inputs are weighted proportionally to their power in separating signal from background. The weighting is determined by training the network on pure signal and background samples, to indicate what characteristics are signal-like and background-like, and which variables have the most power to distinguish these.

The artificial neural network used to make the final candidate selection was constructed using the NeuroBayes package [88] and trained using Monte Carlo simulated events corresponding to 3.9 fb^{-1} data of a pure signal sample, and events from the

B_s^0 mass sidebands as pure background. The signal sample consists of 350000 simulated events, the background 300000 events. The mass sideband regions are defined as $5.2 < \mathrm{M}(J/\psi\phi) < 5.3$ GeV/c^2 on the low side of the B_s^0 mass, and $5.45 < \mathrm{M}(J/\psi\phi) < 5.55$ GeV/c^2 on the high side. The simulated signal data is described in detail in Section 3.8

The quantities used as inputs to the NN training were defined in Section 3.3, Table 3.1 shows the ranking and significance of each input value [80]. The rightmost three columns in the table show variables relating to the importance of the input quantity to the overall performance of the NN. The *added* quantity (4th column) shows how much power is added to that already achieved by the inputs above in the table by adding a given input. The *only* (5th column) variable shows the information level if only that input is used. The *loss* (6th column) variable describes how much information is lost if that input is removed. The input variables are ranked by importance to the performance of the NN, showing that, after the preselection cuts, the muon transverse momentum and kaon likelihoods are of most importance to the signal selection.

Variable	rank	index	added	only	loss
$p_T(\phi)$	1	5	438.14	438.14	157.97
$LL_K(K_2)$	2	11	159.28	379.61	121.44
$LL_K(K_1)$	3	10	106.55	376.28	99.49
$LL_\mu(\mu_2)$	4	9	78.00	147.66	44.83
$\chi^2_{r\phi}(B_s)$	5	2	67.52	163.62	27.32
$p_T(B_s)$	6	3	39.68	270.93	33.19
$LL_\mu(\mu_1)$	7	8	22.78	140.72	22.67
$P(\chi^2, B_s)$	8	4	21.85	132.42	18.80
$P(\chi^2, \phi$	9	6	13.18	23.95	12.66
$P(\chi^2, J/\psi)$	10	7	7.36	51.21	7.36

Table 3.1: NN inputs for $B_s^0 \to J/\psi\,\phi$ signal selection, ranked according to their importance to the NN decision.

As can be seen from Table 3.1, not all of reconstructed parameters for the data are used as inputs to the NN. For sets of highly correlated parameters, only one parameter is used because after the first of the set, the subsequent parameters would add little discrimination power to the NN. The correlations between the inputs listed in Table 3.1 are shown in Figure 3.11 where the first column is the truth (signal or background) value, and the other columns have the same ordering as in the table.

The output of the NN for signal and background using the training samples is shown in Figure 3.12. The network returns a value between -1 and +1, where +1 implies a high degree of confidence that the event is signal, and -1 the same for background. In the same figure, the signal purity as a function of NN output is shown, where the ideal case is a gradient of 1 showing good correlation between the network output and the signal purity.

3.6.1 NN optimisation

Having trained the NN and studied its output for pure signal and background, it is necessary to select the best output value to cut on in order to achieve a high degree

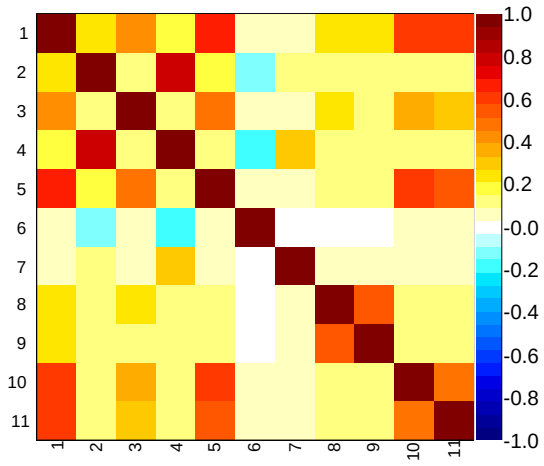

Figure 3.11: Correlation matrix of NN inputs and truth (signal or background) values from pure signal and background NN training samples. First column gives truth value, the others show different inputs where the numbering corresponds to that in Table 3.1.

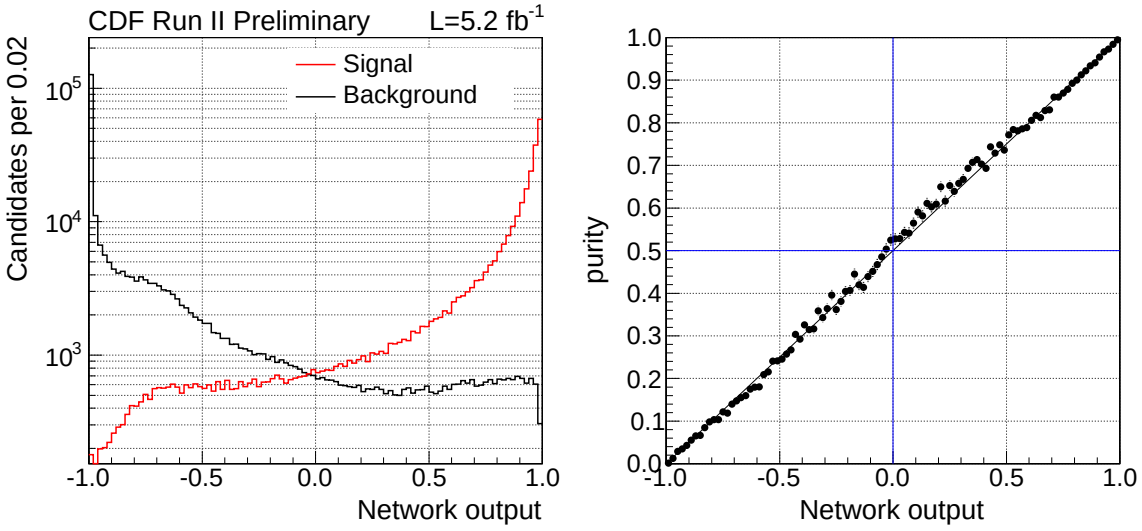

Figure 3.12: [left] NN output distribution for training sample, red line is signal, black is background. [right] signal purity as a function of NN output

of signal purity as well as a good signal yield. Selecting a high NN output value as a cutoff will produce a very pure signal, but will restrict the number of signal events available therefore reducing the statistical power of the sample. A threshold which is too low will increase the statistics, but the separation between signal and background will be inefficient. The NN must be studied on the real data sample size to optimise the cut value.

A common way of selecting the NN cut value is to maximise the quantity $S/\sqrt{(S+B)}$, where S is the number of signal events and B is the number of background. However, in this statistically limited data sample it is important to gain all possible signal yield, so instead the quantity studied to optimise the cut is the sensitivity to $\beta_s^{J/\psi\phi}$ in terms

of statistical errors. This is done by investigating the size of the statistical errors on $\beta_s^{J/\psi\phi}$ in different samples of pseudo experiments relative to the NN cut. The fast Monte Carlo simulation used to generate these pseudo experiments is described in Chapter 5.

Pseudo experiments are generated corresponding to different NN cut values by choosing the associated signal and background numbers for each cut value. Studies are carried out for three potential true values of $\beta_s^{J/\psi\phi}$; 0.02, 0.3 and 0.5 as the NN cut should be optimised for a full range of possible $\beta_s^{J/\psi\phi}$ measurements. For each of these $\beta_s^{J/\psi\phi}$ values, the decay width difference $\Delta\Gamma$ is generated according to the theoretical relationship between these parameters, and all other variables are generated according to their best fit values from the previous $B_s^0 \to J/\psi\,\phi$ analysis [7]. About 700 pseudo experiments are generated and fit for each case, and the $\beta_s^{J/\psi\phi}$ statistical uncertainty is checked at each NN cut value. The error distribution is fitted with a Landau function to determine the most probable value of the statistical error for $\beta_s^{J/\psi\phi}$ at each NN cut value for the three input values of $\beta_s^{J/\psi\phi}$, these fitted values are plotted against the NN cut level and shown in figure 3.13. From this study, it was determined that the optimal NN cut value is 0.2, which minimises $\beta_s^{J/\psi\phi}$ errors by maximising signal yield despite being a looser cut than would be selected by the traditional $S/\sqrt{(S+B)}$ optimisation. The signal yield, number of background events, signal to background ratio and $S/\sqrt{(S+B)}$ are shown as functions of the NN cut value in Figure 3.14.

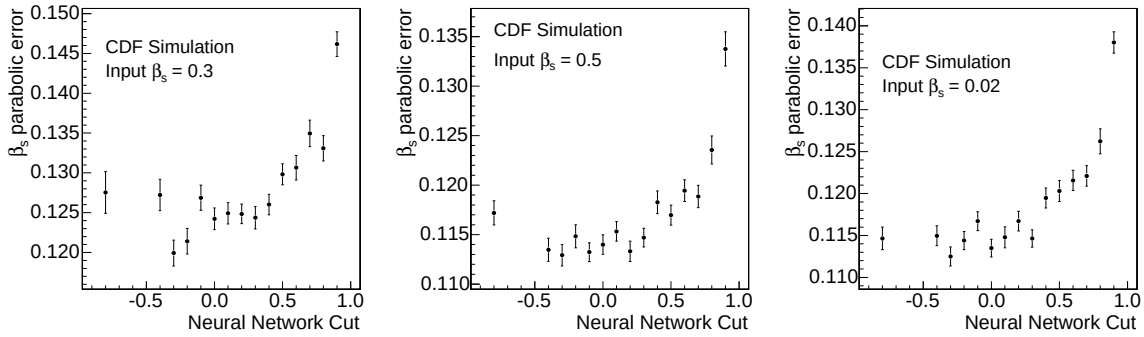

Figure 3.13: Fitted $\beta_s^{J/\psi\phi}$ statistical errors vs. NN cut value for three potential true values of $\beta_s^{J/\psi\phi}$ in pseudo experiments, from which a value of 0.2 is selected as the NN cut threshold.

The $B_s^0 \to J/\psi\,\phi$ yield after NN selection has been applied is shown in terms of the B_s^0 mass distribution in Figure 3.15, fitted as described in Section 4.3.

3.7 Flavour tagging

The flavour taggers used in this analysis were developed for the CDF B_s^0 mixing measurement, and have been re-calibrated and optimised for the $B_s^0 \to J/\psi\,\phi$ analysis. The following subsections describe the algorithms used for these two tagging methods, but first it is useful to describe some common principles of all flavour taggers.

The tag decision, ξ is a discrete variable, which can have the value -1,0 and 1. A value of $\xi = -1$ implies that the initial B meson state was tagged as $\bar{B}_s^0$, for a B_s^0 meson initial state, the tag decision is ξ =1, and when the tagger was unable to determine the initial state the value is ξ=0.

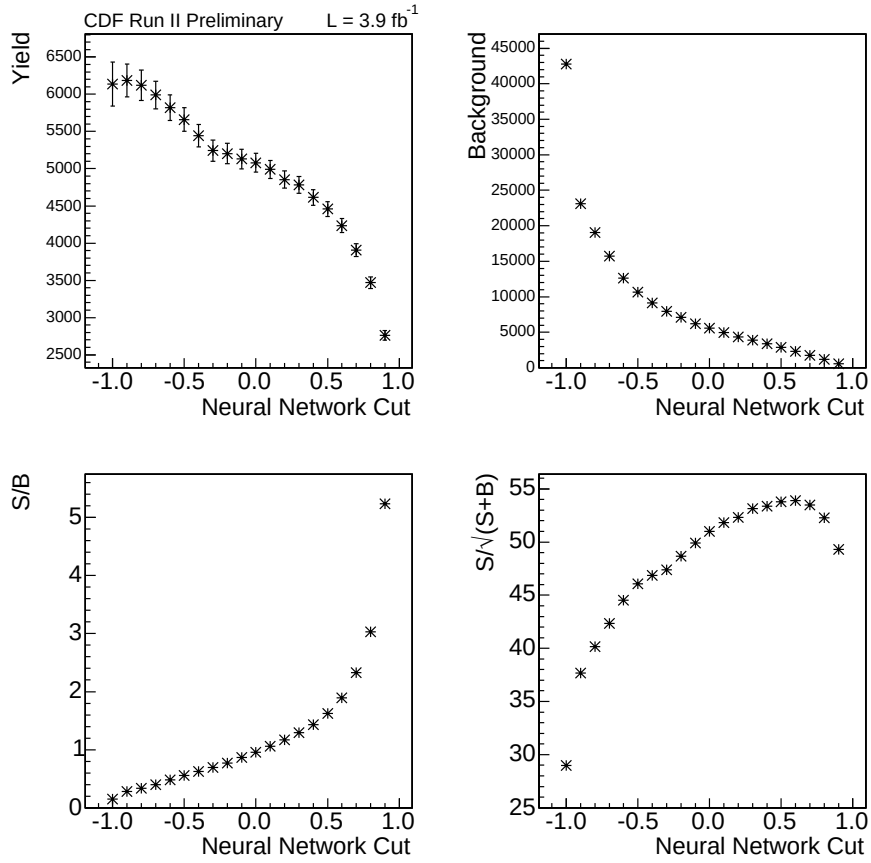

Figure 3.14: Quantities determining signal purity as a function of NN cut value: [top left] signal yield, [top right] background yield, [bottom left] signal to background ratio, [bottom right] $S/\sqrt{S+B}$

For both tagger types, the necessary information may not be available in every event to make a flavour decision; the fraction of events for which a decision can be made is called the *tagging efficiency*. The efficiency is defined as

$$\epsilon = \frac{N_{tagged}}{N_{untagged} + N_{tagged}}. \tag{3.14}$$

The decision of the tagger is not always correct, misidentification of a track type or selecting a track which is not from the correct decay vertex can lead to a mis-tag. The quantity defined to characterise the rate of mis-tagging for a particular algorithm is called *dilution*, D, which is measured as

$$D = 1 - 2P_W \equiv \frac{N_R - N_W}{N_R + N_W} \tag{3.15}$$

where N_R is the number of correctly tagged events in a sample, N_W is the number of wrongly tagged events, and P_W is the probability of an event being wrongly tagged. The dilution variable is constructed such that a tagging algorithm which randomly assigns decisions, therefore having a wrong tag probability of 0.5, will have $D = 0$, and a perfect tagger with P_W=0 will have $D = 1$. For each tagger, the probability of

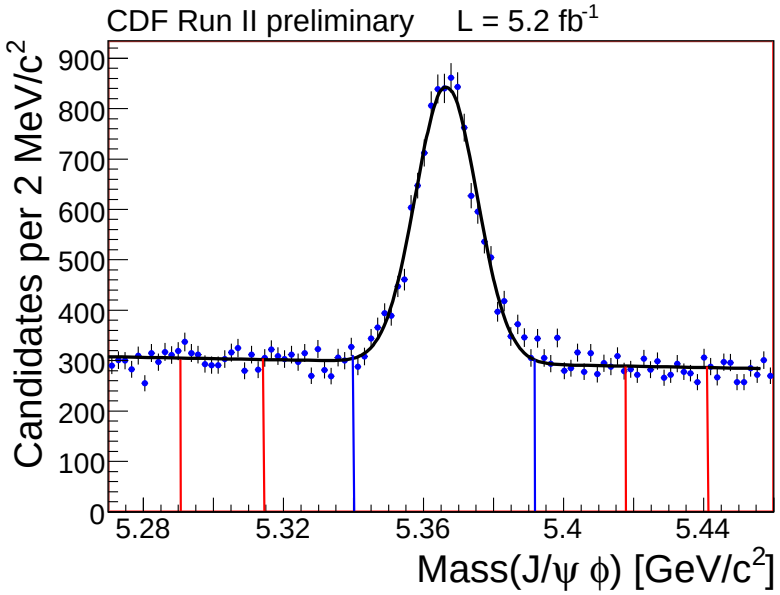

Figure 3.15: B_s^0 mass distribution after preselection and NN cuts applied, yielding ~ 6500 signal $B_s^0 \to J/\psi\,\phi$ events, fitted with a single Gaussian for the signal and a first order polynomial for the background component, as described in Section 4.3.

wrongly tagging the flavour of an event is parameterised, so that the dilution can be predicted on an event by event basis. This prediction can be used to weight events in a likelihood function so that events with higher confidence in the flavour tagging can have a larger contribution. Additionally, the predicted dilution allows porting of the tagging algorithm from the dataset on which it was developed or calibrated, to other samples.

The measure of performance of a flavour tagger is called the *effective tagging efficiency*, and is defined as ϵD^2, where in this case D is the average dilution over the whole sample.

3.7.1 Opposite side tagger

The principle of the OST was introduced in Section 1.5.1. Using the knowledge that a B meson of interest hadronised from one b quark of a $b\bar{b}$ pair, the opposite side tagging algorithms get the initial flavour of that meson from the hadronisation of the other $b(\bar{b})$ quark. This is done in two ways

- by identifying the charge of the lepton from semileptonic B decays (soft electron and soft muon taggers, SET and SMT).
- by identifying the charge of the opposite side b jet (jet charge tagger, JQT).

These processes are completely independent from the candidate side hadronisation and decay, so it is possible to transfer opposite side tagging algorithms between samples of different B types. This means that the algorithms can be developed on light B meson samples, which have significantly higher statistics due to their larger branching ratios, than B_s^0 samples, then applied to tag B_s^0 mesons. Events from the l+SVT trigger sample are used for tagger development and parameterisation of the predicted dilution

in high statistics samples. The taggers are then calibrated on di-muon samples, as described later in this section, finding a global scale factor S_D which can be applied to the event by event dilution to account for kinematic differences in the two samples.

The opposite side lepton taggers exploit the semi-leptonic decay of the opposite side B meson, where the b flavour is tagged as

$$\begin{aligned} b &\to c\, l^- \bar{\nu}_l X \\ \bar{b} &\to \bar{c}\, l^+ \nu_l X \end{aligned} \tag{3.16}$$

so that a positive lepton implies a $\bar{b}$, meaning that the quark in the B_s^0 was a b.

Separate algorithms are used for electrons and muons as opposite side lepton tags. The SET is described in detail in [89] and the SMT in [90]. The efficiency of these taggers is rather low, of order 20%, which is similar to the branching fraction of B to semileptonic decays. The tagging dilution is worsened by mis-identification of leptons, and B^0 mesons oscillating to the opposite of their production flavour and therefore giving an incorrect tag. The predicted tagging dilution for the lepton taggers is a function of p_T and the lepton likelihood (the confidence in the lepton identification derived from calorimeter and dE/dx data), and is given in full in [89]. Well identified leptons with high p_T lead to a good dilution. It is clear that a good lepton probability will improve the tagging quality, the improvement with higher momentum leptons is due to the fact that leptons from b decays are likely to have higher momentum in the transverse plane than those from lighter quarks, due to a larger available phase space.

The jet charge tagger [91] infers the flavour of the candidate B_s^0 from the charge of the opposite side b jet. The jet charge is calculated as the momentum weighted sum of all charges in the jet. This is done by selecting tracks isolated from the candidate meson, as it is important to look at jets only from the opposite side, then using neural networks to find the jet most likely to come from a b quark. One NN is used to assess the probability of an individual track as coming from a B meson, P_{trk}, which is highly dependent on the impact parameter of the tracks, as a displaced vertex is a key feature of b decays. A second NN is used to find P_{jet}, the probability of a jet as containing a b quark. This quantity is based on the number of tracks in a jet with a high P_{jet}, and whether the jet is tagged as being from a displaced vertex with a high decay length significance. The most probable b jet is then selected, and the weighted charge calculated as

$$Q_{jet} = \frac{\sum_i^{Ntracks} Q^i p_T^i (1 + P_{trk}^i)}{\sum_i^{Ntracks} p_T^i (1 + P_{trk}^i)} \tag{3.17}$$

The dilution for this tagger is parameterised as a linear function of the jet charge and jet probability, and is described in more detail in [91]..

The three opposite side taggers described here are not independent, as they can share tracks, which must be accounted for when using them together. They are combined, to give a single tagging decision from the opposite side, using a neural network [92] which exploits correlations between the decisions of these taggers.

As a cross check for the predicted dilution in the $B_s^0 \to J/\psi\,\phi$ sample, an equivalent di-muon sample is selected, $B^+ \to J/\psi\,K^+$, which has comparable kinematics to the sample of interest, but has the advantage that the flavour of the candidate meson is known, as it is tagged by the charge of the daughter kaon. This gives a truth value to compare with the tagging decision, allowing the true dilution to be measured in the sample. From this sample, a scale factor, S_D to apply to the predicted dilution is calculated. If the predicted dilution is suitable for this sample, the scale factor should be equal to 1.0 within errors. It is valid to apply the S_D measured in $B^+ \to J/\psi\,K^+$ to $B_s^0 \to J/\psi\,\phi$ because, as discussed, only the opposite side information is used so the type of candidate meson is not important. Two scale factors are calculated, for B^+ and B^- separately, to account for any charge related asymmetry in the tagging algorithms. The measured dilution is plotted against the predicted dilution for the b and $\bar{b}$ events. To validate the use of the OST developed on l+SVT data for di-muon samples, the dependence should be consistent with a straight line, with gradient of order unity. The fitted gradients of these dependences are taken as the scale factors for the dilution. Figure 3.16 shows the fitted distributions of measured against predicted dilution for the $B^+ \to J/\psi\,K^+$sample.

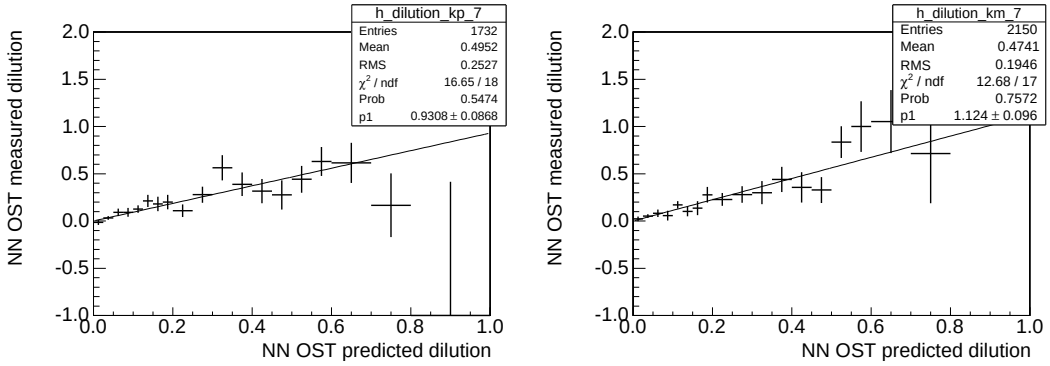

Figure 3.16: Measured vs. predicted dilution for calibration of OST tagging dilution using $B^+ \to J/\psi\,K^+$ data

The measured scale factors are

$$S_D^+ = 0.93 \pm 0.09$$
$$S_D^- = 1.12 \pm 0.10. \tag{3.18}$$

In the same sample, the measured tagging efficiency is 94.3±0.3%, and the mean predicted dilution is found to be $6.9 \pm 0.1\%$.

3.7.2 Same side kaon tagger

Same side tagging algorithms use correlations between the flavour of the b quark making up the candidate meson, and associated tracks. Fragmentation tracks of a particular type are specific to both the b quark flavour in the B meson, and the other quark type in the meson which defines the B meson species. In the hadronisation of a b quark, described in Section 1.3.1, a $q\bar{q}$ pair is produced from the vacuum, and the $\bar{q}$ forms a $b\bar{q}$ meson with the b. The meson formed by the pair-produced q partner of the $\bar{q}$, if identified, can be used to tag the quark content of the candidate B meson. For example, a $\bar{b}$ hadronising with an s quark to form B_s^0 is often produced alongside a

K^-, which contains the $\bar{s}$ partner of the s quark produced as an $s\bar{s}$ pair.

The SSKT used in this analysis is the tagger which was developed for the CDF B_s^0 mixing measurement [42]. It uses a neural network to combine the kaon particle identification likelihood, with kinematic variables of the kaon, to produce a single tagging decision per event. The NN is trained on realistic MC simulations. The dilution of the tagger can be parameterised as a function of the NN output.

For the SST, unlike the OST, the behaviour in different B species is expected to vary, eliminating the possibility of developing and calibrating the SSKT in the same high statistics light B samples as used for the OST. Previously, the SSKT used in CDF B physics analyses has been calibrated on MC simulated data, as the fast oscillations of B_s^0 mesons complicate the estimation of the production flavour. The predicted dilution is still calculated in this way, however, for the first time at CDF this tagger has been calibrated on real B_s^0 data. This data driven calibration uses the measurement of the B_s^0 mixing amplitude to deduce the dilution scale factor. The SSKT calibration is fully documented in [93], and summarised here.

The scale factor enters the probability function for the measurement of B_s^0 mixing as damping on the mixing amplitude, in the term $\xi\mathcal{A}D \cdot \cos(\Delta m_s)$ where ξ is the tag decision, $\mathcal{A}$ is the mixing amplitude and Δm_s is the mixing frequency. The mixing amplitude is expected to be unity at the true mixing frequency, so the measured value at this point gives the required scale factor for D. For calibration purposes, this measurement is made using only the SSKT, the published B_s^0 mixing measurement [42] used a combination of SST and OST tools.

The data sample used in this calibration is made up of four modes collected using the displaced vertex trigger. These decays are

- $B_s^0 \to D_s^- \pi^+, D_s^- \to \phi\pi^-, \phi \to K^+K^-$
- $B_s^0 \to D_s^- \pi^+, D_s^- \to K^*K^-, K^* \to K^+\pi^-$
- $B_s^0 \to D_s^- \pi^+, D_s^- \to \pi^+\pi^+\pi^-,$
- $B_s^0 \to D_s^- \pi^+\pi^+\pi^-, D_s^- \to \phi\pi^-, \phi \to K^+K^-$

where about half of the statistics come from the first channel listed, and the total number of events in the sample is 12877 ± 113.

If the measured amplitude is ≈ 1 that implies that the dilution and time resolution are accurately predicted for the tagger. A measured amplitude of < 1 suggests that the tagger over estimates its accuracy, and > 1 implies an underestimation. Figure 3.17 shows the amplitude scan, which is created by fixing Δm_s to each point along the x-axis, and fitting the amplitude which is plotted as the y variable. It is essential to have a good knowledge of the time resolution in this measurement; a phenomenological function for the resolution is developed in sideband subtracted signal data, and sideband data separately, then applied in the main fit [93, 94].

The measured amplitude is

$$\mathcal{A} = 0.94 \pm 0.15 \text{ (stat.)} \pm 0.13 \text{ (syst.)} \tag{3.19}$$

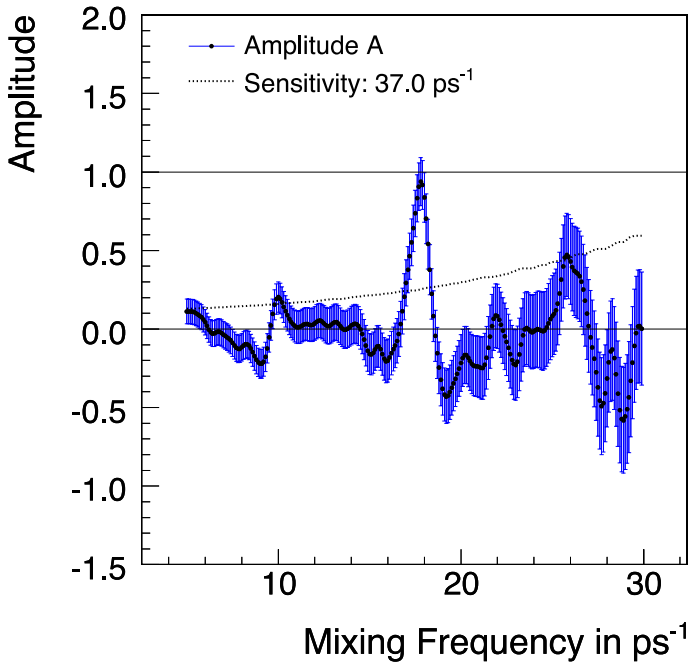

Figure 3.17: Amplitude scan for Δm_s, for calibration of SSKT.

which is consistent with unity, and acts as the scale factor S_D for the tagging dilution. Another interesting result of this calibration measurement is the mixing frequency

$$\Delta m_s = 17.79 \pm 0.07 \text{ (stat. only) ps}^{-1} \tag{3.20}$$

which is in good agreement with the CDF published measurement of Δm_s, of $\Delta m_s = 17.77 \pm 0.10$ (stat.) ± 0.07 (syst.) [42], increasing confidence in this calibration.

3.8 Monte Carlo simulation

Monte Carlo (MC) simulated data is used for two parts of this analysis, the pure signal sample for training of the Neural Network, and to model the detector sculpting on the angular parameters caused by the non-hermiticity of the CDF detector. Realistic $B_s^0 \to J/\psi\,\phi$ MC is generated using the BGenerator MC generator [95], which contains full detector simulation, according to a phase space decay model. The MC sample of 100 million events is modelled using input parameters from the first 1 fb^{-1} of CDF Run-II data, and as would be expected has some inconsistencies with the current dataset which can be corrected for. The disagreements are in the p_T spectra of the candidate particles, which can affect the measurement of the transversity angles, an essential component in this analysis. The original weighting of the MC was trigger related, and as the Tevatron luminosity has increased during the collection of this data there have been changes in the trigger scaling which make it necessary to re-do the re-weighting to match the current data. For training the Neural Network, there is no re-weighting applied to the MC, but for the detector sculpting modelling an event by event re-weighting is applied. The re-weighting factors are calculated according to the trigger path (CMU-CMU or CMU-CMX) of the dimuon pair, and the p_T of the muons. By fitting the distribution of ratios of data to MC for p_T of the B_s^0 meson a p_T dependent re-weighting factor is obtained.

Comparisons of the p_T distributions for sideband subtracted signal events from the full 5.2 fb^{-1} data sample and re-weighted MC are shown in Figure 3.18, which demonstrate good agreement after the corrections have been applied. The MC distributions have been normalised to the number of events in the data sample.

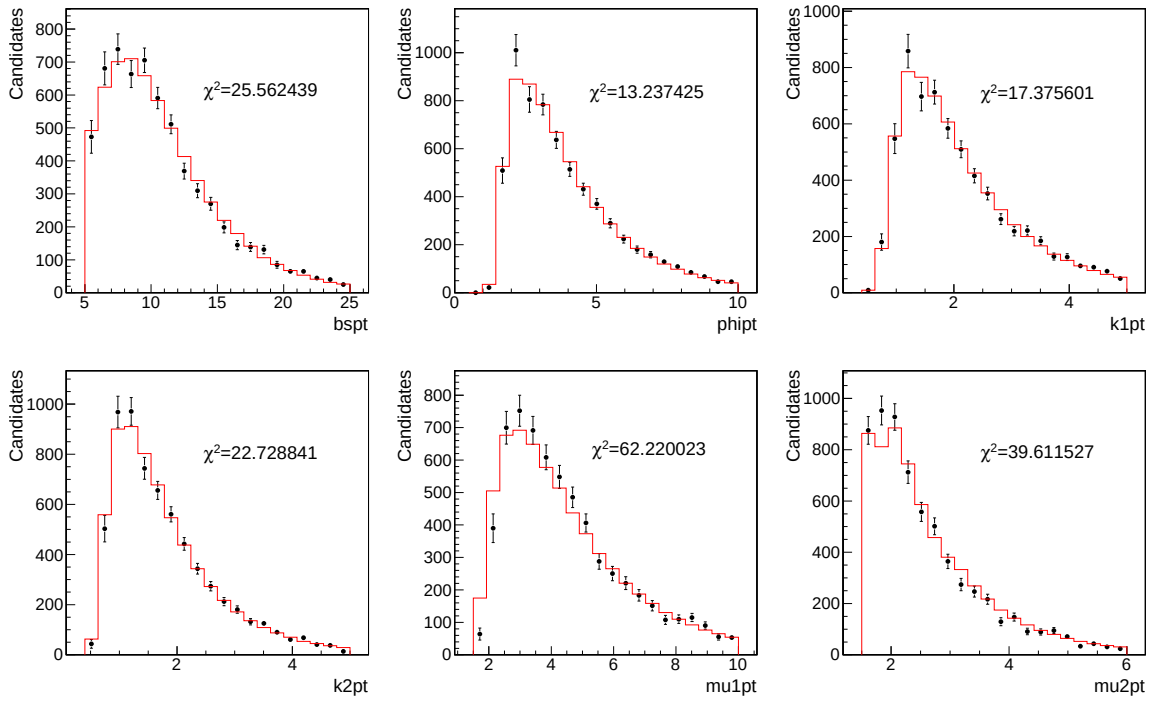

Figure 3.18: p_T distributions for 100 million re-weighted MC simulated events (red lines) and 6500 signal data events (black points)

Chapter 4

Maximum likelihood fit

This chapter describes the full multivariate likelihood function used to extract the physical parameters of interest in particular $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$. Correctly normalised probability density functions for signal and background are constructed in each of the observed variables. These are combined to make one multivariate probability density function from which a likelihood is constructed. The likelihood has degenerate minima corresponding to symmetries in its parameterisation which must be interpreted appropriately; this treatment is also explained.

4.1 Maximum likelihood method

An unbinned maximum likelihood fit is used to extract the values of interesting physics parameters in the decay $B_s^0 \to J/\psi\,\phi$. This technique maximises a likelihood with respect to a set of parameters which are assumed to describe the data. A set of probability density functions (PDFs), $P(\vec{x}|\vec{\mu})$, gives the probability density of observing the measured variables for an event, $\vec{x}_i$ given a set of unknown parameters, $\vec{\mu}$. The likelihood function for the dataset of N events is

$$\mathcal{L}(\mu) = \prod_{i=1}^{N} P(\vec{x}_i|\vec{\mu}). \tag{4.1}$$

The process of maximising the likelihood is carried out by numerically minimising the negative log of the likelihood (NLL) $-\log\mathcal{L}(\vec{\mu})$ as

$$\log\mathcal{L} = \sum_{i=1}^{N} \log P(\vec{x}_i|\vec{\mu}) \tag{4.2}$$

This minimisation is carried out using the MINUIT minimiser package [96] which is implemented in the ROOT programming environment [97].

4.2 Components of the likelihood function

The following sections describe each PDF necessary to construct the likelihood function $\mathcal{L}$. These are:

- The B_s^0 mass PDFs, $P_s(m)$ and $P_b(m)$

- The signal decay time and angular PDF, $\rho'(\theta, \phi, \psi, t, F_s, \delta_s)$
- The background decay time PDF $P_b(t)$
- The background angular model PDFs, $P(\theta)P(\psi)P(\phi)$
- The lifetime error PDFs $P_s(\sigma_{ct})$ and $P_b(\sigma_{ct})$
- PDFs for the flavour tagging variables, $P_s(\xi)$,$P_b(\xi)$, $P_s(D)$ and $P_b(D)$, and the modified signal time and angular dependence including flavour tagging decision $P_s(\theta, \phi, \psi, t, F_s, \delta_s, \xi|\sigma_{ct})$

where the parameters introduced above are defined in the following discussion.

Combining these conditional probabilities, the full likelihood is constructed for all events as in Equation 4.2. The full likelihood function including flavour tagging, is

$$\begin{aligned}\mathcal{L} &= \prod[f_s \cdot P_s(m) \cdot P_s(\xi) \cdot P_s(\theta, \phi, \psi, t, F_{SW}, \delta_{SW}, \xi, \mathcal{D}|\sigma_{ct}) \cdot P_s(\sigma_{ct}) \cdot P_s(\mathcal{D}) \\ &+ (1 - f_s) \cdot P_b(m) \cdot P_b(\xi) \cdot P_b(t|\sigma_{ct}) \cdot P_b(\theta) \cdot P_b(\phi) \cdot P_b(\psi) \cdot P_b(\sigma_{ct}) \cdot P_b(\mathcal{D})]\end{aligned} \tag{4.3}$$

For the case of the fit without flavour tagging, this reduces to

$$\begin{aligned}\mathcal{L} &= \prod[f_s \cdot P_s(m) \cdot P_s(\theta, \phi, \psi, t, F_{SW}, \delta_{SW}|\sigma_{ct}) \cdot P_s(\sigma_{ct}) \\ &+ (1 - f_s) \cdot P_b(m) \cdot P_b(t|\sigma_{ct}) \cdot P_b(\theta) \cdot P_b(\phi) \cdot P_b(\psi) \cdot P_b(\sigma_{ct})]\end{aligned} \tag{4.4}$$

where $P_s(\theta, \phi, \psi, t, F_{SW}, \delta_{SW})$ simply corresponds to the flavour tagged case where the tag decision $\xi = 0$.

4.3 B_s^0 mass PDF

The mass distribution of events is a clear way of separating signal candidates from background as the B_s^0 is a well defined resonance. The following subsections describe the models used to incorporate the signal and background mass distributions into the likelihood fit.

4.3.1 Signal model, $P_s(m)$

The signal mass distribution is modelled using a single Gaussian function. To form the probability density function (PDF), $P_s(m)$, this is smeared with an event-by-event mass resolution which is scaled using a single scale factor to account for a general mis-estimation of the mass errors. The signal region is taken to be $5.342 < m(B_s^0) < 5.392$.

The PDF is, normalised over the range $5.2 < M(B_s^0) < 5.6$:

$$P_s(m|M, \sigma_m) = \frac{\frac{1}{\sqrt{2\pi}s_m\sigma} e^{\left[-\frac{1}{2}\left(\frac{m-M}{s_m\sigma}\right)^2\right]}}{\frac{1}{2}[\mathrm{erf}\left(\frac{\mathrm{M_{max}-M}}{\sqrt{2}\mathrm{s_m}\sigma}\right) - \mathrm{erf}\left(\frac{\mathrm{M_{min}-M}}{\sqrt{2}\mathrm{s_m}\sigma}\right)]} \tag{4.5}$$

where erf is the error function, M_{max} and M_{min} are the upper and lower limits of the mass range, the scale factor on the mass error is s_m.

4.3.2 Background model, $P_b(m)$

The background mass model is a first order polynomial function, which is smeared with the same error distribution as the signal mass, giving the background mass PDF $P_b(m)$. The B_s^0 mass side band regions, used in some of the background studies discussed in the following sections, are $5.290 < m(B_s^0) < 5.315$ (lower) and $5.418 < m(B_s^0) < 5.442$ (upper).

The background mass PDF, normalised in the same mass range as the signal mass PDF, is

$$P_b(m) = p_1 \cdot m + \frac{1}{M_{max} - M_{min}} \left[1 - \frac{p_1}{2}(M_{max}^2 - M_{min}^2)\right] \tag{4.6}$$

where p_1 is the slope of the 1st order polynomial.

4.4 PDF in angular and time variables

The angular and time dependent decay functions derived in Chapter 1 describe the phenomenology of the $B_s^0 \to J/\psi\,\phi$ decay, and the additional contribution from $B_s^0 \to J/\psi KK$, as would be recorded by a perfect detector. In the likelihood fit, the effects of the CDF detector on the distributions of variables must be accounted for. Including the detector effects, the distributions are different than what would be predicted by the PDFs in various respects:

- The distributions of the transversity angles defined in Chapter 1 are expected to be flat in $\cos\theta$, $\cos\psi$ and ϕ, but these distributions undergo some sculpting by the detector which is discussed in the following subsection.

- The time dependence of the various signal and background components of the probability density functions are modified by the detector resolution and hence must be convoluted with an estimated resolution to take this into account.

As a consequence of the detector efficiency effects the corresponding PDFs will have to be modified, and normalised, taking the derived efficiency into account.

4.4.1 Detector sculpting of transversity angles for signal events, $\rho'(\theta, \psi, \phi, t, F_s, \delta_s)$

Initially, the effect of the detector efficiency will be described for the $B_s^0 \to J/\psi\,\phi$ signal component (Equation 1.81) only, then the $B_s^0 \to J/\psi K^+K^-$ component will be added. The equations for the signal $B_s^0 \to J/\psi\,\phi$ probability functions in Chapter 1 were normalised such that

$$\iiiint \sum_{j=B,\bar{B}} P_j(\theta, \phi, \psi, t) \sin\theta d\theta \sin\psi d\psi = 1. \tag{4.7}$$

Including the efficiency function, $\epsilon(\theta, \phi, \psi)$ in the time and angular decay PDF, $P(\theta, \phi, \psi, t)$, gives a new PDF, $P'(\theta, \phi, \psi, t)$

$$P'(\theta, \phi, \psi, t) = \frac{P(\theta, \phi, \psi, t)\epsilon(\theta, \phi, \psi)}{N} \tag{4.8}$$

and the equivalent for $\bar{P}'$, which are both normalised by the factor

$$N = \iiiint \sum_{i=B,\bar{B}} P_i(\theta, \phi, \psi, t)\epsilon(\theta, \phi, \psi)d(\cos\psi)d(\cos\theta)d\phi dt. \tag{4.9}$$

In Equation 4.8, the $\epsilon(\theta, \phi, \psi)$ term in the numerator does not affect the maximum likelihood, as it does not depend on the fit variables. However, the addition of this term does affect the normalisation, so it is this which is focused on in the following discussion.

The detector efficiency function is parameterised using a set of real spherical harmonics and Legendre polynomials as basis functions, in 3-dimensions, with ranges $0 < \psi < \pi$, $0 < \theta < \pi$ and $0 < \phi < 2\pi$.

$$\epsilon(\psi, \theta, \phi) = \sum_{lmk} a^k_{lm} P_k(\cos\psi) Y_{lm}(\theta, \phi), \tag{4.10}$$

The normalisation factor, N, can be derived analytically, as shown in Appendix B, the result being

$$\begin{aligned}
N &= \frac{3}{8\sqrt{\pi}}\left[\frac{4a^0_{00}}{3}(|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2) + \frac{4a^2_{00}}{15}(2|A_0|^2 - |A_\parallel|^2 - |A_\perp|^2)\right] \\
&+ \frac{3}{8\sqrt{5\pi}}\left[\frac{2a^0_{20}}{3}(|A_0|^2 + |A_\parallel|^2 - 2|A_\perp|^2) + \frac{4a^2_{20}}{15}(|A_0|^2 - \frac{1}{2}|A_\parallel|^2 + |A_\perp|^2)\right] \\
&- \frac{9}{16\sqrt{15\pi}}\frac{1}{\sqrt{1 + \frac{4\tau_L\tau_H}{((\tau_L - \tau_H)\sin 2\beta_s)^2}}}\left[(A^*_\parallel A_\perp + A_\parallel A^*_\perp)(\frac{4}{3}a^0_{2-1} - \frac{4}{15}a^2_{2-1})\right] \\
&+ \frac{9}{16}\frac{\sqrt{2}}{\sqrt{15\pi}}\frac{1}{\sqrt{1 + \frac{4\tau_L\tau_H}{((\tau_L - \tau_H)\sin 2\beta_s)^2}}}\left[(A^*_0 A_\perp + A_0 A^*_\perp)(\frac{\pi a^1_{21}}{8} - \frac{\pi a^3_{21}}{32} + ...)\right] \\
&+ \frac{9}{8\sqrt{15\pi}}\left[\frac{2a^0_{22}}{3}(-|A_0|^2 + |A_\parallel|^2) - \frac{4a^2_{22}}{15}(|A_0|^2 + \frac{1}{2}|A_\parallel|^2)\right] \\
&+ \frac{9}{16}\frac{\sqrt{2}}{\sqrt{15\pi}}\left[(A^*_0 A_\parallel + A_0 A^*_\parallel)(\frac{\pi a^1_{2-2}}{8} - \frac{\pi a^3_{2-2}}{32} + ...)\right],
\end{aligned} \tag{4.11}$$

The detector efficiency enters N through the coefficients a^k_{lm}, which are fit using realistic Monte Carlo simulated data. This simulated set of 100 million $B^0_s \to J/\psi\,\phi$ signal events is described in Section 3.8. The variables $\cos\theta$, ϕ and $\cos\psi$ are generated flat then reconstructed using the same procedure as for real data, thus the distributions represent the detector efficiencies in each variable. The Monte Carlo distributions of ϕ, $\cos\theta$ and $\cos\psi$ are filled into a 3-dimensional histogram with 20 bins in each variable,

these distributions of the transversity angles are shown in Figure 4.1.

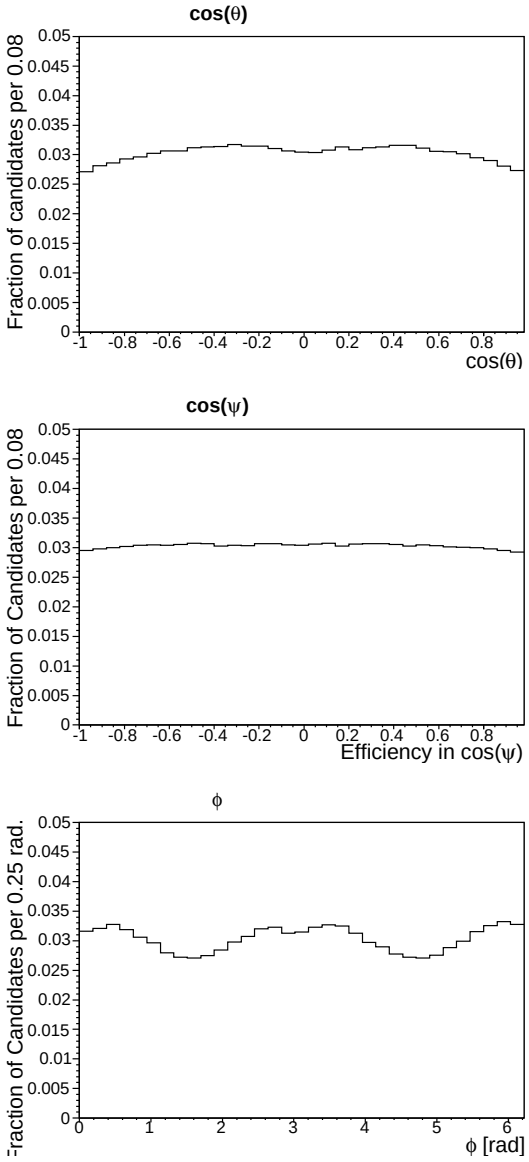

Figure 4.1: Detector sculpting of transversity angles in Monte Carlo simulations. Each is generated flat in the variable shown, and sculpted by the detector using the same reconstruction as for data.

To obtain the coefficients a^k_{lm}, the real spherical harmonics in (θ, ϕ) are expanded according to the Laplace series

$$Y_{lm}(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=0}^{l}[C_{lm}\cos(m\phi) + S_{lm}\sin(m\phi)]P_l^m(\cos\theta) \tag{4.12}$$

where each term is expanded as a function of a Legendre polynomial used to fit ψ

$$S_{lm} = \sum_{k=0}^{\infty} S_{lm}^{k} \sqrt{\frac{(2k+1)}{2}} P_k(\cos\psi)$$
$$C_{lm} = \sum_{k=0}^{\infty} C_{lm}^{k} \sqrt{\frac{(2k+1)}{2}} P_k(\cos\psi) \tag{4.13}$$

The series in Equation 4.12 can be related to the set of orthonormal basis functions used in Appendix B for the calculation of the normalisation factor N, by

$$\begin{aligned} Y_{lm} &= Y_l^m \quad (m=0), \\ Y_{lm} &= \frac{1}{\sqrt{2}}(Y_l^m + (-1)^m Y_l^{-m}) = \sqrt{2} C_{lm} \cos(m\phi) P_l^m(\cos\theta) \quad m>0, \\ Y_{lm} &= \frac{1}{i\sqrt{2}}(Y_l^{|m|} - (-1)^{|m|} Y_l^{-|m|}) = \sqrt{2} S_{l|m|} \sin(|m|\phi) P_l^m(\cos\theta) \quad m<0. \end{aligned} \tag{4.14}$$

from which the final coefficients a_{lm}^k can be taken directly as C_{lm}^k and S_{lm}^k according to Equation 4.13.

The 3-dimensional fit results are given in Appendix C, Figure 4.2 shows the fit and residuals for a 2-dimensional fit to the (θ, ϕ) distribution integrated over $\cos\psi$.

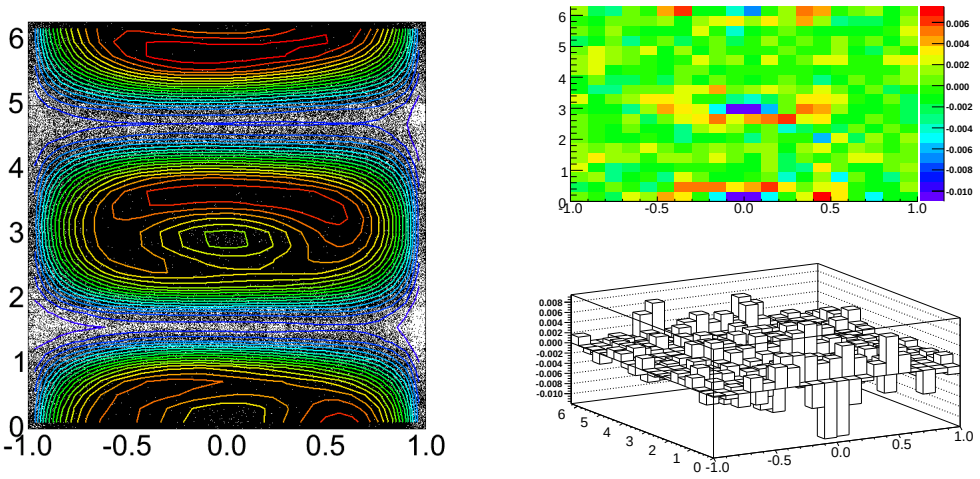

Figure 4.2: Two dimensional fit to $(\cos\theta, \phi)$ transversity angles integrated over $\cos\psi$ to find detector efficiency coefficients for normalisation of $B_s^0 \to J/\psi\,\phi$ signal time and angular decay PDF.

To add in the S-wave KK part, and normalise the angular sculpted modification of Equations 1.101 and 1.102 requires two additional normalisation factors, one for the S wave KK term and one for the S-wave - P-wave interference term

$$\mathcal{N} = (1 - F_s) \cdot N + 2Re\left[\mathcal{I}_\mu \cdot N'\right] + F_s \cdot N'' \tag{4.15}$$

Calculating this analytically [98], using the model for the ϕ meson mass shape described

in Chapter 1 gives

$$\begin{aligned} N' &= \sqrt{3}A_0^*(\frac{1}{6\sqrt{\pi}}a_{00}^1 + \frac{1}{12\sqrt{5\pi}}a_{20}^1 - \frac{1}{4\sqrt{15\pi}}a_{22}^1) \\ &+ \frac{3}{16}\sqrt{\frac{2}{5\pi}}A_\parallel^*(\frac{\pi}{2}a_{2-2}^0 - \frac{\pi}{8}a_{2-2}^2 + ...) \\ &+ \frac{3}{16}\sqrt{\frac{2}{5\pi}}A_\perp^*\frac{\sin 2\beta_s(\tau_L - \tau_H)}{\sqrt{((\tau_L - \tau_H)\sin 2\beta_s)^2 + 4\tau_L\tau_H}}(\frac{\pi}{2}a_{21}^0 - \frac{\pi}{8}a_{21}^2 + ...) \\ N'' &= \frac{1}{2\sqrt{\pi}}a_{00}^0 + \frac{1}{4\sqrt{5\pi}}a_{20}^0 - \frac{3}{4\sqrt{15\pi}}a_{22}^0 \end{aligned} \tag{4.16}$$

where the $a_{2,1}^k$, $a_{2,-2}^k$ were defined in the calculation of N for the $B_s^0 \to J/\psi\,\phi$ only normalisation.

This shows the full analytic normalisation of the probability function including detector sculpting of the angular distributions for the full decay including the S-wave KK component. However, the true mass model used for the ϕ meson is somewhat more complex than the Breit-Wigner function used so far. The model was enhanced to better fit the distribution seen in realistic MC, using an asymmetric relativistic Breit-Wigner with mass dependent width

$$BW = \frac{m}{m_\phi} \cdot \Gamma_1 \cdot \frac{k^*(B_s, m, J/\psi)}{k^*(B_s, m_\phi, J/\psi)} \cdot \frac{1}{(m_\phi^2 - m^2)^2 + m_\phi^2 \cdot \Gamma_{tot}^2}. \tag{4.17}$$

where the particle momentum is given by the k^* terms, and its mass is m in the B_s^0 rest frame. This assumes a two body decay, where the other daughter particle is the J/ψ, and the total decay width Γ_{tot} is

$$\Gamma_{tot} = \Gamma_1 + \Gamma_2 + \Gamma_3, \tag{4.18}$$

where $\Gamma_{1,2,3}$ are the partial decay widths for the decays $\phi \to K^+K^-$ $(48.8 \pm 0.5\%)$, $\phi \to K_L^0 K_S^0$ $(34.2 \pm 0.4\%)$ and $\phi \to \rho\pi + \pi^+\pi^-\pi^0$ $(15.32 \pm 0.32\%)$ [10] respectively. The agreement between this function and the ϕ mass shape in MC is shown in Figure 4.3.

With this improved description of the ϕ mass incorporated into the probability for the $B_s^0 \to J/\psi\,\phi$ plus $B_s^0 \to J/\psi KK$ decay, the normalisation factors N' and N'' are calculated numerically during the fit. The improvement in accuracy of the model is considered worth the slight increase in computing time.

The normalised full PDFs for the time and angular dependence, including angular sculpting of the detector, and decays to both ϕ and $KK(f_0)$ final states, are

$$\rho'(\theta, \phi, \psi, t, F_s, \delta_s) = \frac{1}{N''}\rho(\theta, \phi, \psi, t, F_s, \delta_s)\epsilon(\theta, \phi, \psi) \tag{4.19}$$

and

$$\bar{\rho}'(\theta, \phi, \psi, t, F_s, \delta_s) = \frac{1}{N''}\bar{\rho}(\theta, \phi, \psi, t, F_s, \delta_s)\epsilon(\theta, \phi, \psi) \tag{4.20}$$

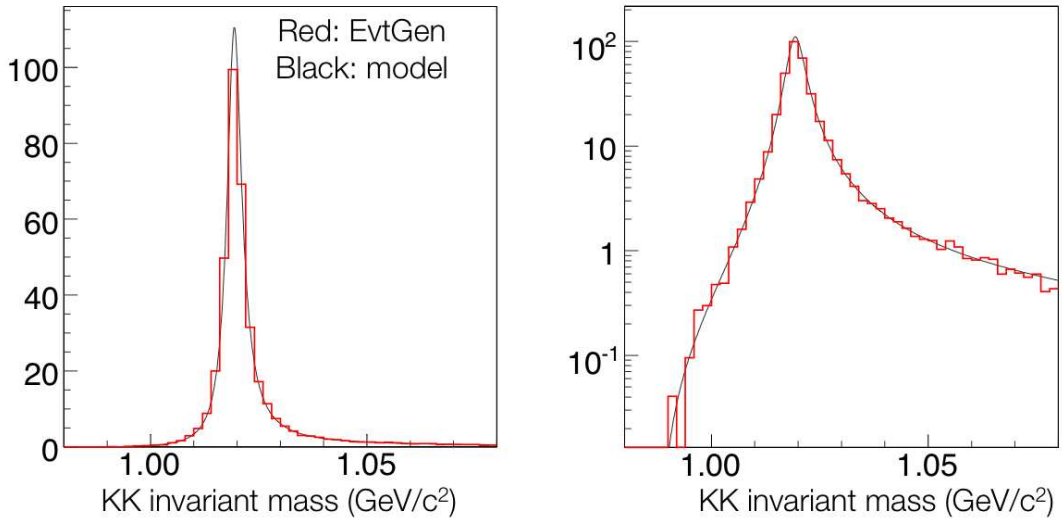

Figure 4.3: Agreement between the asymmetric, mass dependent Breit Wigner parameterisation and realistic Monte Carlo simulation

4.4.2 Detector resolution smearing of signal $P(\sigma_{ct})$

The detector resolution was discussed in Chapter 1, where the time and angular dependent probabilities were developed. These were constructed such that smearing could be simply applied to the separated exponential functions. The exponential and sin(exponential) functions are smeared with two separate Gaussian functions, according to the event-by-event error σ_t, multiplied by scale factors which are floated in the fit. The scale factors for the two Gaussians are different, and they are included to account for an overall mis-estimation of the decay time resolutions.

For each of the signal lifetime components, there is a PDF of the form

$$\begin{aligned} P'_s(ct,\sigma_{ct}|c\tau,s_{ct1,2}) &= P(ct|c\tau)\otimes(f_{s_{ct1}}G_1(ct,\sigma_{ct}|s_{ct1})+(1-f_{s_{ct1}})G_2(ct,\sigma_{ct}|s_{ct2})) \\ &= \frac{1}{c\tau}e^{-\frac{ct}{c\tau}}\otimes\left(f_{s_{ct1}}\frac{1}{\sqrt{2\pi}S_1\sigma_{ct}}e^{-\frac{c^2t^2}{2s^2_{ct1}\sigma^2_{ct}}}+(1-f_{s_{ct1}})\frac{1}{\sqrt{2\pi}s_{ct2}\sigma_{ct}}e^{-\frac{c^2t^2}{2s^2_{ct2}\sigma^2_{ct}}}\right) \end{aligned} \tag{4.21}$$

where s_{ct1} and s_{ct2} are the σ_{ct} scale factors, and f_1 is the fraction of the first smearing Gaussian.

When using event-by-event errors, which are not distributed identically for the signal and background events, it is necessary to include a PDF for the separate error distributions [99]. The error distributions differ for signal and background because combinatorial background events have a $c\tau$ which is fitted from a random combination of four tracks and therefore tend to have a worse resolution σ_t than the better fitted signal events.

The decay time error PDF $P_s(\sigma_{ct})$ is constructed from Gamma functions

$$\Gamma(x) \equiv \frac{x^a e^{-x/b}}{b^{a+1}\Gamma(a+1)} \tag{4.22}$$

where a and b define the mean and width of the distribution. Each function has different values of a and b, and these values are found in a separate lifetime only fit before running the full angular analysis. The projections of this preparatory fit on data are shown in Figure 4.4

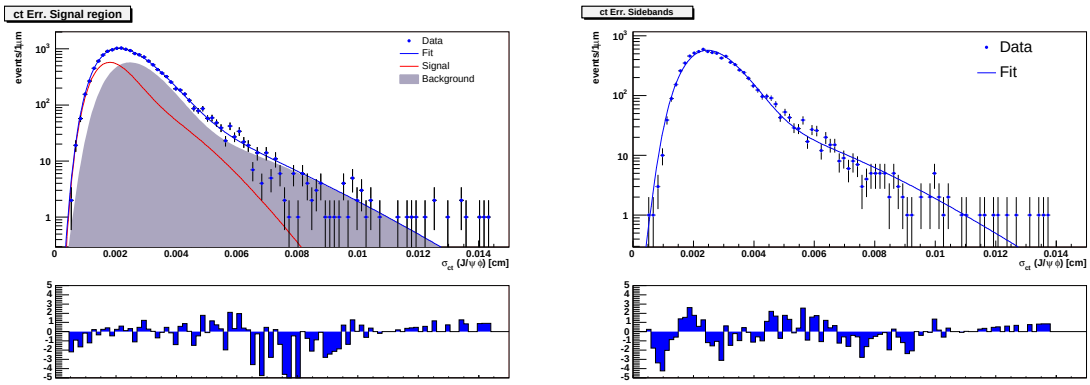

Figure 4.4: Fit projection of the B_s^0 proper decay length error distributions for signal (left) and sideband (right) regions

4.4.3 Background lifetime and angular PDFs, $P_b(t)$ and $P_b(\theta)$, $P_b(\phi)$, $P_b(\psi)$

For the background, the angular distributions are not used in separating lifetime components; they will be presented individually in this subsection.

The background proper decay time function, $P_b(t)$ is parameterised as a prompt peak modelled by a Gaussian distribution, two positive exponentials and a negative exponential. This function is smeared with the same resolution function as the signal decay time dependence. The components of this parameterisation are motivated as follows: the prompt peak models the majority of the combinatorial background events, which are expected to have no significant lifetime, the positive exponentials account for a small fraction of longer lived background such as real kaons, and the negative exponential is to take into account events where the vertex reconstruction of the proper decay length results in a τ below zero. A negative τ measurement can occur because in the vertex constrained fit to calculate $c\tau$, there is no guarantee that L_{xy} and p_T will be collinear, therefore Equation 3.2 may not be positive. Equation 4.23 shows the mathematical form of this distribution.

$$\begin{aligned} P_b(t,\sigma_{ct}) &= \left\{ f_g + (1-f_g)\left(f_{++}\frac{1}{\lambda_{++}}e^{-\frac{t}{\lambda_{++}}} + (1-f_{++})\left(f_-\frac{1}{\lambda_-}e^{\frac{t}{\lambda_-}} + (1-f_-)\frac{1}{\lambda_+}e^{-\frac{t}{\lambda_+}} \right) \right) \right\} \\ &\otimes \{ f_{s_{ct1}} G_1(\sigma_{ct1}) + (1-f_{s_{ct1}}) G_2(\sigma_{ct2}) \} \end{aligned} \tag{4.23}$$

The background lifetime resolution function $P_b(\sigma_{ct})$ is handled in the same way as described for the signal distribution, using another set of Gamma functions. The fitted distribution of background σ_{ct} is shown in Figure 4.4.

There are no predictions for the distribution of the background transversity angles, the following functions are found to describe the data

$$f(\cos\theta) = \frac{a_0 - a_1 cos^2(\theta)}{2a_0 - 2a_1/3}$$
$$f(\phi) = \frac{1 + b_1 cos(2\phi + b_0)}{2\pi}$$
$$f(\cos\psi) = \frac{c_0 + c_1 cos^2(\theta)}{2c_0 + 2c_1/3}. \quad (4.24)$$

They are initially fitted on data from only the B_s^0 mass sidebands, to find sensible starting values of $a_{0,1,2}$, $b_{0,1,2}$ and $c_{0,1,2}$, then allowed to float freely in the full minimisation. The preparatory fits to the sideband data are shown in Figure 4.5 projected on data points.

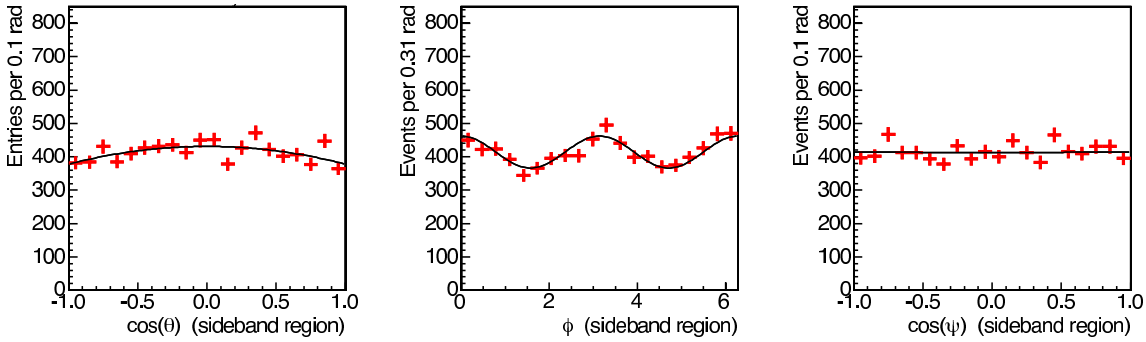

Figure 4.5: Distribution of transversity angles for background events, with fit projections from initial fit to sideband only data

The PDFs for the background angles are treated as factorisable, as they have minimal correlations between them thus they can be modelled separately. The assumption of factorisability is checked in Section 6.2.6, where it is determined that the actual correlations in the angles lead to a small systematic effect. The small correlations can be seen in Figure 6.2. After normalising the functions in Equation 4.24 the PDF for the background angular distributions is $P(\theta, \phi, \psi) = P(\theta) \cdot P(\phi) \cdot P(\psi)$.

4.5 Flavour tagging PDFs

Flavour tagging was introduced in Section 1.5.1; details of the opposite side (OST) and same side (SST) tagging algorithms and their calibrations were given in Section 3.7. Now, flavour tagging can be incorporated into the probability function for the signal and background components.

4.5.1 Combining tagging algorithms in the signal decay PDF, $P_s(\theta, \phi, \psi, t, F_s, \delta_s, \xi, D)$, $P_s(\xi)$, $P_s(D)$

For each tagger, three additional variables are brought into the likelihood: the tag decision, ξ, the tagging dilutions D, and the dilution scale factor, S_D. As discussed previously (Section 3.7), the tag decision ξ can be +1 (B_s^0), -1 ($\bar{B}_s^0$) or 0 (no decision) and the predicted dilution D gives a weighting to each tagged event, signifying the quality of the tagging. The dilution is scaled by S_D, a scale factor determined by calibrating the taggers on suitable data samples. The scale factors are allowed to float

within Gaussian constraints to their errors in the main $B_s^0 \to J/\psi\,\phi$ fit, to account for any variations in performance between the calibration and analysis samples.

Analogous to the inclusion of the proper decay length error, $\sigma_{c\tau}$, it is necessary to include a probability density function, $P(D)$, for the predicted tagging dilution in the full likelihood. This probability density function is taken from a separate histogram for signal and background. The signal histogram is filled with background-subtracted signal region data, the background dilution histograms are taken from sideband region data. These distributions are shown in Figure 4.6.

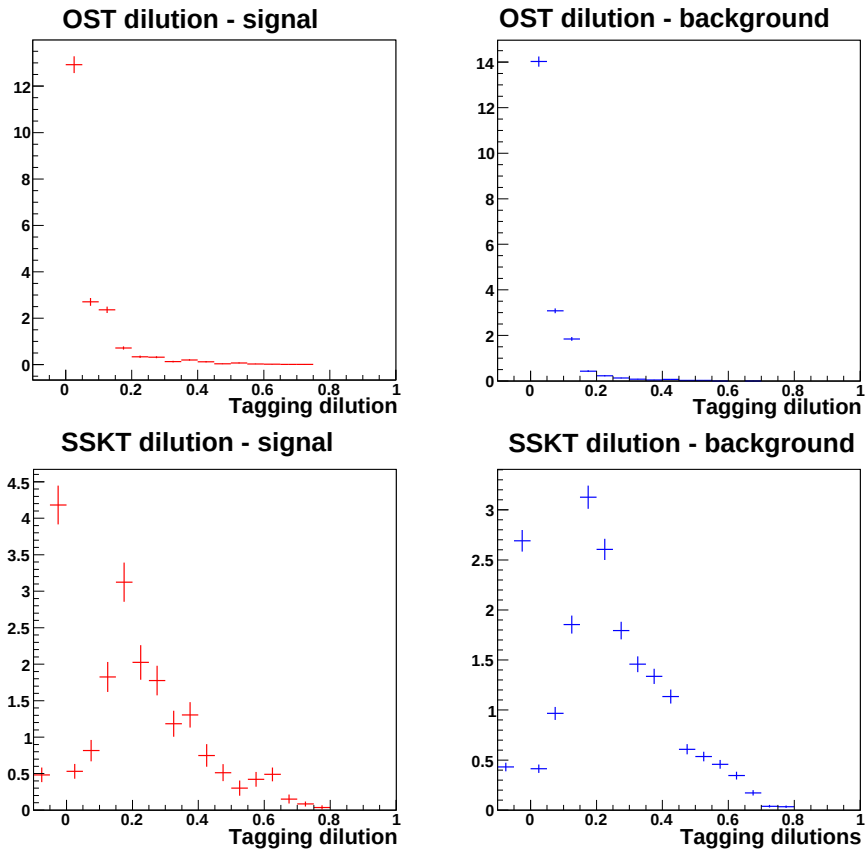

Figure 4.6: Normalised predicted tagging dilutions for the OST (top) and SSKT (bottom)

The probability density function for the tag decision, $P(\xi)$ must take into account the efficiency of the tagging algorithm to produce a tag decision. If the tagging efficiency for a single tagger is ϵ, then $P(\xi)$ will be

$$P(\xi) = (1-\epsilon)\cdot\delta(\xi - 0) + \epsilon\cdot\delta(|\xi| - 1) \tag{4.25}$$

With two independent flavour taggers, with ϵ_1 and ϵ_2 as tagging efficiencies, the probability can be written as

$$P(\xi) \equiv P(\xi_1)P(\xi_2) = \left(1 - \sum_{j=1}^{2}\epsilon_j\right)\cdot\delta\left(\sum_{j=1}^{2}\xi_j - 0\right) + \sum_{j=1}^{2}\delta\left(|\xi_j| - 1\right) \tag{4.26}$$

The modified decay probability for the time and angular dependence can now be

constructed to include the tag decision variables; it is presented here initially for a single tagger for clarity, then extended to the case with two flavour taggers. The lifetime and angular dependent probability function was modified in Section 4.4 to incorporate detector sculpting of the angular distributions, giving $\rho'(\theta, \phi, \psi, t, F_s, \delta_s)$.

Using a single tagger, this becomes

$$\begin{aligned} T(\theta,\phi,\psi,t,F_s,\delta_s,D,\xi) = \frac{1-sD}{2}\rho'(\theta,\phi,\psi,t,F_s,\delta_s) + \frac{1+sD}{2}\bar{\rho}'(\theta,\phi,\psi,t,F_s,\delta_s) \text{ for } \xi &= -1 \\ T(\theta,\phi,\psi,t,F_s,\delta_s,D,\xi) = \rho'(\theta,\phi,\psi,t,F_s,\delta_s) + \bar{\rho}'(\theta,\phi,\psi,t,F_s,\delta_s) \text{ for } \xi &= 0 \\ T(\theta,\phi,\psi,t,F_s,\delta_s,D,\xi) = \frac{1+sD}{2}\rho'(\theta,\phi,\psi,t,F_s,\delta_s) + \frac{1-sD}{2}\bar{\rho}'(\theta,\phi,\psi,t,F_s,\delta_s) \text{ for } \xi &= 1 \end{aligned} \tag{4.27}$$

This notation can be made more compact by the inclusion of the decision variable as a multiplication factor. For the inclusion of tagging decisions from two separate flavour taggers, the probability can be written in this compact form as

$$\begin{aligned} T(\theta,\phi,\psi,t,F_{SW},\delta_{SW},D_1,D_2,\xi_1,\xi_2) &= \frac{1+\xi_1 s_1 D_1}{1+|\xi_1|}\frac{1+\xi_2 s_2 D_2}{1+|\xi_2|}\rho'(\theta,\phi,\psi,t) \\ &+ \frac{1-\xi_1 s_1 D_1}{1+|\xi_1|}\frac{1-\xi_2 s_2 D_2}{1+|\xi_2|}\bar{\rho}'(\theta,\phi,\psi,t) \end{aligned} \tag{4.28}$$

However one of the taggers, the OST, has two separate dilution scale factors, for mesons tagged as containing b and $\bar{b}$, notated as S^- and S^+. This modifies the probability to

$$\begin{aligned} T(\theta,\phi,\psi,t,F_{SW},\delta_{SW},D_1,D_2,\xi_1,\xi_2) &= \frac{1+\xi_1 s_1^+ D_1}{1+|\xi_1|}\frac{1+\xi_2 s_2 D_2}{1+|\xi_2|}\rho'(\theta,\phi,\psi,t) \\ &+ \frac{1-\xi_1 s_1^- D_1}{1+|\xi_1|}\frac{1-\xi_2 s_2 D_2}{1+|\xi_2|}\bar{\rho}'(\theta,\phi,\psi,t) \end{aligned} \tag{4.29}$$

if the OST is taken to be tagger 1.

The final decay PDF including the ct resolution, σ_{ct}, as well as the detector angular efficiencies and combined flavour tagging is

$$P_s(\theta,\phi,\psi,t,F_{SW},\delta_{SW},\xi,\mathcal{D}|\sigma_{ct}) \cdot P_s(\xi) \cdot P_s(\mathcal{D}) \cdot P_s(\sigma_{ct}) \tag{4.30}$$

In this analysis, some results are produced using a likelihood fitter without flavour tagging (the *untagged fit*), as a cross check that the tagging does not introduce a bias in the measured values. In this fit configuration, the likelihood can be seen to reduce to the $\xi = 0$ case.

4.5.2 PDFs for background flavour tagging $P_b(\xi)$ and $P_b(D)$

Tagging information for background events is essentially without physical meaning, as they are mostly combinations of unrelated tracks which pass the signal selection criteria. Despite this, the effect of flavour tagging has to be accounted for in the likelihood construction for background events. The background dilution PDF $P_b(D)$ is,

like the signal equivalent, taken from the normalised histogram of tagging dilution for background events (Figure 4.6).

The background tag decision PDF $P_b(\xi)$ contains a term to correct for any charge asymmetry in the taggers, a disparity in the fraction of background events tagged as B_s^0 over $\bar{B}_s^0$. This term is notated as f_{b+}, for the fraction of background events which are tagged as containing a $\bar{b}$ quark (positively charged) and is floated in the full likelihood fit. This gives a PDF of the form

$$P_b(\xi) = \delta(\xi - 0) \cdot 1 + \delta(\xi - 1) \cdot f_{b+} + \delta(\xi + 1) \cdot (1 - f_{b+}) \tag{4.31}$$

4.6 Symmetries in the likelihood function

Considering first the decay function for $B_s^0 \to J/\psi\,\phi$, ignoring the S-wave $KK[f_0]$ component, the likelihood has symmetry properties under certain transformations. These can be seen clearly from the explicit form of the PDF of the decay. From Equation 1.86, the components for $\mathbf{A}_+$ and $\mathbf{A}_-$ can be written out using Equations 1.84 and 1.85

$$\begin{aligned} |\mathbf{A}_+(0) \times \hat{n}|^2 &= |A_0(0)|^2 \cos^2\psi(\cos^2\theta + \sin^2\theta\sin^2\phi) \\ &+ \frac{1}{2}|A_\parallel(0)|^2 \sin^2\psi(\cos^2\theta + \sin^2\theta\cos^2\phi) \\ &+ \frac{1}{2\sqrt{2}}|A_0(0)||A_\parallel(0)|\cos(\phi_\parallel)\sin 2\psi \sin^2\theta \sin 2\phi) \end{aligned} \tag{4.32}$$

and

$$|\mathbf{A}_-(0) \times \hat{n}|^2 = \frac{1}{2}|A_\perp(0)|^2 \sin^2\psi \sin^2\theta \tag{4.33}$$

where by convention, $\phi_\parallel = \arg(A_\parallel)$, $\phi_\perp = \arg(A_\perp)$ and A_0 is real. The B_s^0-$\bar{B}_s^0$ interference term becomes

$$\begin{aligned} (\mathbf{A}_+ \times \hat{n}) \cdot (\mathbf{A}_-^* \times \hat{n}) &= \frac{i}{4}|A_\parallel(0)||A_\perp(0)|e^{i(\phi_\parallel - \phi_\perp)} \sin^2\psi \sin 2\theta \sin\phi \\ &+ \frac{i}{4\sqrt{2}}|A_0(0)||A_\perp(0)|e^{-i\phi_\perp} \sin 2\psi \sin 2\theta \cos\phi \end{aligned} \tag{4.34}$$

and analogously for the $\bar{B}_s^0$. At this stage it is useful to use some trigonometric relations to simplify the equations. The angular functions in Eqn. 4.32 can be re-written as

$$\begin{aligned} \cos^2\theta + \sin^2\theta\sin^2\phi &= 1 - \sin^2\theta + \sin^2\theta(1 - \cos^2\phi) \\ &= 1 - \sin^2\theta\cos^2\phi, \end{aligned} \tag{4.35}$$

$$\begin{aligned} \cos^2\theta + \sin^2\theta\cos^2\phi &= 1 - \sin^2\theta + \sin^2\theta(1 - \sin^2\phi) \\ &= 1 - \sin^2\theta\sin^2\phi. \end{aligned} \tag{4.36}$$

Defining a set of functions

$$\begin{aligned} g_1(\vec{\omega}) &= \cos^2\psi(1-\sin^2\theta\cos^2\phi) \\ g_2(\vec{\omega}) &= \sin^2\psi(1-\sin^2\theta\sin^2\phi) \\ g_3(\vec{\omega}) &= \sin^2\psi\sin^2\theta \\ g_4(\vec{\omega}) &= \sin 2\psi\sin^2\theta\sin 2\phi \\ g_5(\vec{\omega}) &= \sin^2\psi\sin 2\theta\sin\phi \\ g_6(\vec{\omega}) &= \sin 2\psi\sin 2\theta\cos\phi, \end{aligned}$$

where $\vec{\omega} \equiv (\theta, \phi, \psi)$, enables the re-writing of 1.86

$$\begin{aligned} P(\theta,\phi,\psi,t) &= \frac{9}{16\pi}\{|A_0(0)|^2|f_+(t)|^2 g_1(\vec{\omega}) + \frac{1}{2}|A_\parallel(0)|^2|f_+(t)|^2 g_2(\vec{\omega}) + \frac{1}{2}|A_\perp(0)|^2|f_-(t)|^2 g_3(\vec{\omega}) \\ &- \frac{1}{2\sqrt{2}}|A_0(0)||A_\parallel(0)|\cos(\phi_\parallel)|f_+(t)|^2 g_4(\vec{\omega}) \\ &- \Re(\frac{i}{2}|A_\parallel(0)||A_\perp(0)|e^{i(\phi_\parallel-\phi_\perp)}f_+(t)f_-^*(t))g_5(\vec{\omega}) \\ &+ \Re(\frac{i}{2\sqrt{2}}|A_0(0)||A_\perp(0)|e^{i\phi_\perp}f_+(t)f_-^*(t))g_6(\vec{\omega})\}. \end{aligned} \tag{4.37}$$

To take the real part of the last two terms in Equation 4.37, the explicit time dependence of the interference terms must be written out, using Equation 1.88.

$$\begin{aligned} P(\theta,\phi,\psi,t) &= \frac{9}{32\pi}\{2|A_0(0)|^2|f_+(t)|^2 g_1(\vec{\omega}) + |A_\parallel(0)|^2|f_+(t)|^2 g_2(\vec{\omega}) + |A_\perp(0)|^2|f_-(t)|^2 g_3(\vec{\omega}) \\ &-\frac{1}{\sqrt{2}}|A_0(0)||A_\parallel(0)|\cos(\phi_\parallel)|f_+(t)|^2 g_4(\vec{\omega}) \\ &-\frac{1}{N}|A_\parallel(0)||A_\perp(0)|[\sin(\phi_\parallel-\phi_\perp)e^{-\Gamma t}\cos(\Delta m t) \\ &+\cos(\phi_\parallel-\phi_\perp)(\cos(2\beta_s)e^{-\Gamma t}\sin(\Delta m t)+\sin(2\beta_s)(e^{-\Gamma_L t}-e^{-\Gamma_H t})/2)]g_5(\vec{\omega}) \\ &+\frac{1}{N\sqrt{2}}|A_0(0)||A_\perp(0)|[\sin(\phi_\perp)e^{-\Gamma t}\cos(\Delta m t) \\ &-\cos(\phi_\perp)(\cos(2\beta_s)e^{-\Gamma t}\sin(\Delta m t)+\sin(2\beta_s)(e^{-\Gamma_L t}-e^{-\Gamma_H t})/2)]g_6(\vec{\omega}), \end{aligned} \tag{4.38}$$

where $N = \sqrt{[(\tau_L-\tau_H)\sin 2\beta_s]^2 + 4\tau_L\tau_H}$.

With these explicit definitions in hand, the invariant behaviour of the functions can be observed. Transforming the vector $\mathbf{A}(t)$ (Equation 1.83) into its complex conjugate is an invariant operation on the PDF. This can be shown by looking at the simultaneous transformation of the dependent parameters required to carry out this operation:

- $\beta_s \to \pi/2 - \beta_s$
- $\Delta\Gamma \to -\Delta\Gamma$
- $\phi_\perp \to \pi - \phi_\perp$
- $\phi_\parallel \to 2\pi - \phi_\parallel$.

These simultaneous transformations perform the complex conjugation as follows

- The terms in square brackets in Equation 1.80 are transformed into their complex conjugates, as the terms $E_\pm \to E_\pm^*$ (where $E_\pm$ are defined in Equation 1.68), and $e^{-2i\beta_s} \to -e^{+2i\beta_s}$ and vice versa.

- The denominator in Equation 1.80 is invariant under the simultaneous transformation of the parameters described, as both $\cos 2\beta_s$ and $\tau_L - \tau_H$ change sign, so their product remains the same, therefore the original value is equal to its complex conjugate.

- The amplitudes $A_\parallel \to A_\parallel^*$ and $iA_\perp \to -iA_\perp^*$ by the definition of the phases $\phi_\perp$ and $\phi_\parallel$, and A_0 is equal to its own complex conjugate as it is a real value.

Thus the combined transformation has converted $A(t)$ to $A^*(t)$. The consequence of this invariance, is that there are two minima of the likelihood function which cannot be distinguished between.

In the case of the fit without flavour tagging information, there is an additional irresolvable symmetry under the simultaneous transformation

- $\beta_s^{J/\psi\phi} \to \beta_s^{J/\psi\phi} - \pi/2$

- $\Delta\Gamma_s \to -\Delta\Gamma_s$

The addition of flavour tagging removes this symmetry, as the ability to follow separately the decay of B_s^0 and $\bar{B}_s^0$ introduces sensitivity to the sign of the parameters transformed above. This effect can be seen in the difference between Equations 1.75 and 1.76

Now returning to the probability function which takes into account $B_s^0 \to J/\psi KK$ decays, it can be observed that the addition of this component to the $B_s^0 \to J/\psi\,\phi$ signal could break the symmetry described for the fit including flavour tagging information. First, it is necessary to find the transformation which performs the complex conjugation of the combined S- and P-wave amplitudes which appear in ϕ meson mass dependent Equation 1.92 in the form of

$$\sqrt{1 - F_{SW}}h(\mu)\mathbf{A}(t) + e^{i\delta_s}\sqrt{F_{SW}}\frac{g(\mu)}{\sqrt{3}}\mathbf{B}(t). \tag{4.39}$$

From the form of $h(\mu)$, such a complex conjugation would require the transformation of the ϕ meson decay width, Γ_ϕ, to $-\Gamma_\phi$, which would be unphysical. This situation can be avoided by instead transforming Equation 4.39 to its negative complex conjugate, which is carried out by in addition to performing the previously described transformation of the terms in $\beta_s, \Delta\Gamma, \phi_\parallel$, and $\phi_\perp$, transforming

- $\delta_{SW} \to \pi - \delta_{SW}$

- $(\mu - \mu_\phi) \to -(\mu - \mu_\phi)$.

In this transformation, the last part flips from one side of the ϕ meson mass peak to a point symmetrically on the opposite side. Considering this fact in terms of the ϕ meson mass integrated function, which is the case used in the main likelihood fit in this analysis, the final combined transformation which gives a symmetry in the likelihood is as described above, with the μ terms omitted.

$$\{\beta_s, \Delta\Gamma, \phi_\perp, \phi_\parallel, \delta_{SW}\} \Rightarrow \{\pi/2 - \beta_s, -\Delta\Gamma, \pi - \phi_\perp, 2\pi - \phi_\parallel, \pi - \delta_{SW}\} \tag{4.40}$$

However, this symmetry only holds as long as the ϕ mass distribution is symmetric, and the interval of integration is symmetric about the central value of the ϕ meson mass. As was explained in Section 4.4.1 of this chapter, the ϕ mass distribution is in fact an asymmetric function, so this symmetry is broken. In Chapter 7 it will be shown that this effect is actually rather small, and not significant enough to isolate a single solution to the likelihood function with this level of statistics.

4.7 Conversion from α_{CPOdd} and $\alpha_\parallel$ to $|A_0(0)|^2$ and $|A_\parallel(0)|^2$

In the likelihood fit, only two of the three angular amplitude parameters are determined, the third is calculated as $|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2 = 1$. The parameters used in the minimisation are $\alpha_{\mathrm{CPOdd}} \equiv |A_\perp|^2$ and $\alpha_\parallel \equiv |A_\parallel|^2/(1 - |A_0|^2)$. This change of variables is imposed in order to prevent unphysical values for the angular amplitudes being found by the fitter. By using $\alpha_{parallel}$ and α_{CPOdd}, the fit is prevented from finding $|A_0|^2 + |A_\parallel|^2 > 1$ without having to set any artificial constraints on the amplitudes.

In order to quote values of the angular amplitudes which can be compared with existing measurements, it is necessary to follow the convention of presenting the initial amplitudes $|A_{\{0,\parallel,\perp\}}(0)|^2$. These can be obtained from the fitted parameters, which give the integrated rates to each polarisation state, as follows:

$$\begin{aligned} |A_0(0)|^2 &= \frac{(1-\alpha_\parallel)(1-\alpha_{\mathrm{CPOdd}})}{1+(y-1)\alpha_{\mathrm{CPOdd}}} \\ |A_\perp(0)|^2 &= \frac{\alpha_{\mathrm{CPOdd}}}{1+(y-1)\alpha_{\mathrm{CPOdd}}} \\ |A_\parallel(0)|^2 &= \frac{\alpha_\parallel(1-\alpha_{\mathrm{CPOdd}})}{1+(y-1)\alpha_{\mathrm{CPOdd}}} \end{aligned} \tag{4.41}$$

where $z \equiv \cos(2\beta_s)\Delta\Gamma/(2\Gamma)$ and $y \equiv (1-z)/(1+z)$. As β_s is fixed to 0.0 for the measurement of these values, z reduces to $z = \Delta\Gamma/(2\Gamma)$. It is only necessary to quote final values for two of the three amplitudes because of the $|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2 = 1$ relationship.

To properly treat the errors for the angular amplitudes, it is necessary to propagate the statistical and systematic errors of the measured quantities taking into account

correlations between the parameters [100], as shown in Equation 4.42.

$$\sigma^2_{|A_0(0)|^2} = \sum_{i=1}^{4} \left(\frac{\partial |A_0(0)|^2}{\partial P_i} \right) \sigma^2_{P_i} + \sum_{i=1}^{4} \sum_{j=1}^{4} \frac{\partial |A_0(0)|^2}{\partial P_i} \frac{\partial |A_0(0)|^2}{\partial P_j} \rho(P_i, P_j)$$
$$\sigma^2_{|A_\parallel(0)|^2} = \sum_{i=1}^{4} \left(\frac{\partial |A_\parallel(0)|^2}{\partial P_i} \right) \sigma^2_{P_i} + \sum_{i=1}^{4} \sum_{j=1}^{4} \frac{\partial |A_\parallel(0)|^2}{\partial P_i} \frac{\partial |A_\parallel(0)|^2}{\partial P_j} \rho(P_i, P_j) \qquad (4.42)$$

where P_i are the dependent parameters $\alpha_\parallel, \alpha_{\mathrm{CPOdd}}, c\tau_s$ and $\Delta\Gamma_\mathrm{s}$, ρ is the correlation matrix for the statistical errors and σ_{P_i} are the errors for each parameter in the relation. The full error propagation relations, and error matrix values are given in Appendix D.

Chapter 5

Fitter validation

This chapter presents results from fits using pseudo experiments to check the likelihood and evaluate the behaviour of the fit parameters, especially to assess biases in parameters of interest. A full validation of the likelihood has already been performed in [101] but special attention is paid here to testing the significantly modified components such as the inclusion of the S-wave KK component in the likelihood.

5.1 Validation of the fitter

The aim of this section is to establish whether biases are present in the fitted values across a range of physically possible situations. A technique called *pull studies* can be used to check the behaviour of a fitter by looking at the distribution of normalised deviations (*pulls*) in pseudo experiments from the generated value of each parameter [102]. For a variable which has a Gaussian distribution with mean μ and width σ, the pull

$$g = \frac{x - \mu}{\sigma} \tag{5.1}$$

is expected to be Gaussian by definition. However, the central limit theorem shows that this principle can in fact be extended to non-Gaussian parameters such as a lifetime which would be expected to have a Gaussian distribution of measured values if a suitably large number of pseudo experiments were carried out. In this case, the pull

$$g = \frac{\theta_m - \theta_g}{\sigma_m} \tag{5.2}$$

is expected to be a Gaussian of mean=0 and σ=1, where θ_m and θ_g are the measured and generated (true) values of the parameter being studied, and σ_m is the error on the measurement. A deviation from the Gaussian distribution in this case can indicate either a bias in the fitting technique, or a mis-calculation of the errors which enter in the denominator.

Pseudo experiments are necessary for the analysis of biases in the analysis, because in data the true value is not known. In the study of $B_s^0 \to J/\psi\,\phi$, a fast Monte Carlo simulation has been developed which allows the generation of large numbers of events for different values of the physics parameters and limited detector effects. In this simulation, the following components are generated:

- **Mass and mass error**

Following the mass PDF described in Section 4.3, the signal mass is generated according to a Gaussian smeared by event-by-event mass errors as in Equation 4.5 where the parameters are taken from data. If background is included, the background mass is generated according to a first order polynomial distribution as in Equation 4.6 with parameters and error distribution from the relevant data fit.

- **Lifetime and distribution of decay angles**
  The B_s^0 and $\bar{B}_s^0$ proper decay time and angular amplitudes are generated according to the PDFs developed in Chapter 1. When background is included, the combinatorial background lifetimes are constructed using the three exponential plus one Gaussian model described in Section 4.4.3, given by Equation 4.23. The background distributions for the transversity angles are generated according to Equation 4.24, described in the same section.
- **Detector acceptance**
  The PDFs for the lifetime and angular distribution are multiplied by acceptance (efficiency) functions fitted from realistic Monte Carlo which uses full detector simulation, as explained in Section 4.4.1.
- **Detector resolution**
  The lifetime for each generated event is smeared with event by event errors taken from data distributions to simulate the decay time resolution of the detector, this effect is shown in Equation 4.21 for signal and included in Equation 4.23 for background events.
- **Flavour tagging**
  The flavour tagging efficiencies, decisions and dilution are generated using distributions measured in data for each tagger, according to the PDFs laid out in Section 4.5.1 for signal events and Section 4.5.2 for background.
- $B_s^0 \to J/\psi KK$ **component**
  A fraction of the signal component is generated with $B_s^0 \to J/\psi KK$ (non-resonant) instead of $B_s^0 \to J/\psi\,\phi$. The fraction of this component, and its phase relative to the P-wave ϕ can be varied to study the effects of different values for these parameters.

The fast Monte Carlo simulations described here are used extensively throughout the fitter validation process, and in estimating systematic and statistical effects in Chapters 6 and 7. The following sections describe the use of pull studies to validate the likelihood fitter for various configurations of the fit.

5.2 Checks of the S-wave KK parameters

The improved likelihood used in this analysis is based on a fitter used in previous measurements which was extensively validated in [101] and shown to be unbiased in the limit of very high statistics. After the modifications to improve the fitter, such as the inclusion of $B_s^0 \to J/\psi KK$ component, the focus has been on checking the new functionality rather than a full re-validation with high statistics tests.

The parameters of particular interest in this section, the first part of the validation, are the S-wave fraction of the $B_s^0 \to J/\psi\,\phi$ signal, and relative phase, A_{SW} and δ_{SW}. For the high statistics studies, 16 ensembles of 500 pseudo experiments, each with 100000 signal events (with no background), were generated using the fast Monte Carlo generator described in the previous section. These pseudo experiments were generated

with input values for all parameters taken from the data fit values (Table 7.1) except for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ which are generated at their SM predicted values, and the S-wave KK component fraction and relative phase, which are varied for each ensemble of experiments. Sets of pseudo experiments are generated at with the S-wave fraction, A_{SW}, at 5%, 10%, 25% and 50%, with δ_{SW} generated at 0, $\pi/4$, $\pi/2$ and $3\pi/4$.

The pulls for the S-wave fraction, shown in Tables 5.1 and 5.2 are not badly shifted from zero mean, and of approximately unit width, for all generated values of the S-wave fraction and phase. For the S-wave relative phase, there are some statistically significant biases present in Tables 5.3 and 5.4 of 1-10σ significance. In terms of absolute magnitude these are small shifts, and this parameter is not included in the set of best fit values quoted as final results because there is not sufficient sensitivity to it, as shown in Figure 7.1. However, it is important to investigate these biases in case they indicate a problem with the fitter for other parameters.

Mean	5%	10%	25%	50%
0	0.22 ± 0.06	0.13 ± 0.05	-0.02 ± 0.04	-0.05 ± 0.05
$\pi/4$	-0.07 ± 0.04	-0.09 ± 0.06	-0.02 ± 0.05	-0.06 ± 0.05
$\pi/2$	-0.07 ± 0.04	-0.08 ± 0.05	-0.08 ± 0.04	-0.10 ± 0.05
$3\pi/4$	-0.001 ± 0.05	-0.07 ± 0.04	-0.15 ± 0.05	-0.09 ± 0.05

Table 5.1: Mean of the S-wave fraction pull from high statistics pseudo experiments with A_{SW} and δ_{SW} generated at the values in each row and column.

σ	5%	10%	25%	50%
0	0.83 ± 0.05	0.87 ± 0.04	0.95 ± 0.03	0.99 ± 0.04
$\pi/4$	0.91 ± 0.03	0.93 ± 0.06	0.97 ± 0.04	1.02 ± 0.04
$\pi/2$	0.87 ± 0.03	0.91 ± 0.04	0.95 ± 0.03	0.99 ± 0.04
$3\pi/4$	0.93 ± 0.04	0.89 ± 0.03	0.95 ± 0.04	1.01 ± 0.04

Table 5.2: Width of the S-wave fraction pull from high statistics pseudo experiments with A_{SW} and δ_{SW} generated at the values in each row and column.

Mean	5%	10%	25%	50%
0	0.10 ± 0.06	0.18 ± 0.05	0.42 ± 0.05	-0.47 ± 0.04
$\pi/4$	0.14 ± 0.10	0.24 ± 0.05	0.25 ± 0.04	0.27 ± 0.04
$\pi/2$	-0.06 ± 0.6	0.04 ± 0.04	-0.03 ± 0.04	-0.05 ± 0.05
$3\pi/4$	-0.19 ± 0.11	-0.21 ± 0.05	-0.31 ± 0.04	-0.28 ± 0.04

Table 5.3: Mean of the S-wave relative phase pull from high statistics pseudo experiments with A_{SW} and δ_{SW} generated at the values in each row and column.

The larger biases seen in the S-wave fraction, A_{SW}, for low input values are thought to be due to the truncated nature of its likelihood profile, which is cut off at zero, and can be seen in Figure 7.1. The biased behaviour of the relative phase δ_{SW} can not be understood from the studies shown so far, so further investigation has been carried out. From Table 5.3 the case with the worst pull, $A_{SW} = 0.5$, $\delta_{SW} = 0$, was taken. For these input values, $\sim$500 pseudo experiments with 1000000 signal events

σ	5%	10%	25%	50%
0	0.87 ± 0.06	0.91 ± 0.04	1.01 ± 0.04	0.94 ± 0.03
$\pi/4$	0.96 ± 0.09	0.86 ± 0.04	0.92 ± 0.03	0.93 ± 0.04
$\pi/2$	0.80 ± 0.05	0.87 ± 0.03	0.92 ± 0.04	0.97 ± 0.04
$3\pi/4$	0.98 ± 0.11	0.89 ± 0.04	0.86 ± 0.03	0.82 ± 0.03

Table 5.4: Width of the S-wave relative phase pull from high statistics pseudo experiments with A_{SW} and δ_{SW} generated at the values in each row and column.

were generated. Figures 5.1-5.2 show that the main parameters of interest, $\beta_s^{J/\psi\phi}$, c/τ_s and $\Delta\Gamma_s$ exhibit no biases despite the biased pulls for the phases. The parameters which have a moderate biases of about 0.3 ± 0.05, α_{CPodd} and $\alpha_{\parallel}$, are closely related to the phases, and examining the residual shifts for these parameters shows that the actual size of the shift is very small, of order 0.001. For the phases themselves, $\phi_{\parallel}$, $\phi_{\perp}$ and δ_{SW} all have significantly biased pulls in this test, of around 2.0 ± 0.05. The non-Gaussian behaviour of $\phi_{\parallel}$ is understood as being due to the proximity of a symmetry point to the generated value (details are given in Section 5.3.1) but this does not fully explain the bias exhibited here.

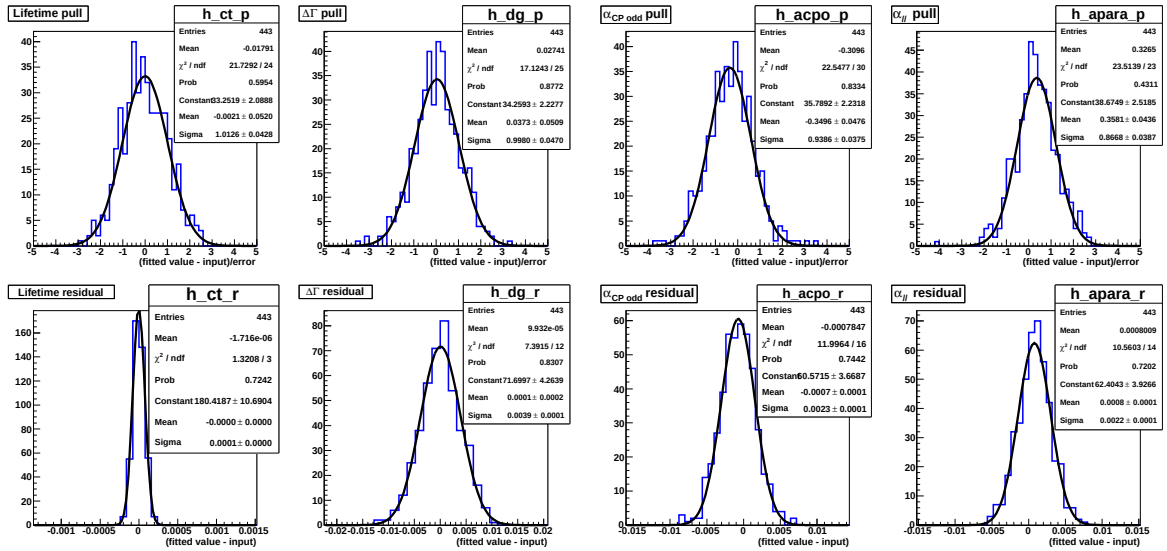

Figure 5.1: Pull distributions and residuals for $\Delta\Gamma$, $c\tau$ and the transversity amplitudes from high statistics tests with A_{SW}=0.5, $\delta_{SW} = 0.0$

To further investigate this problem, two additional sets of pseudo experiments were generated and fitted. As it is understood that $\phi_{\parallel}$ is not fitted well close to its reflection point at π (see section 5.3.1), these sets were generated with values of $\phi_{\parallel} = 1.5$ and $\phi_{\parallel} = 1.0$, rather than the fitted value from data of $\phi_{\parallel} = 3.08$, to check whether the biases in $\phi_{\perp}$ and δ_{SW} could be explained by the behaviour of $\phi_{\parallel}$. The results of these experiments are shown in Figures 5.4-5.8. These show a significant improvement over those generated with $\phi_{\parallel}$ close to π, but the biases are not removed entirely.

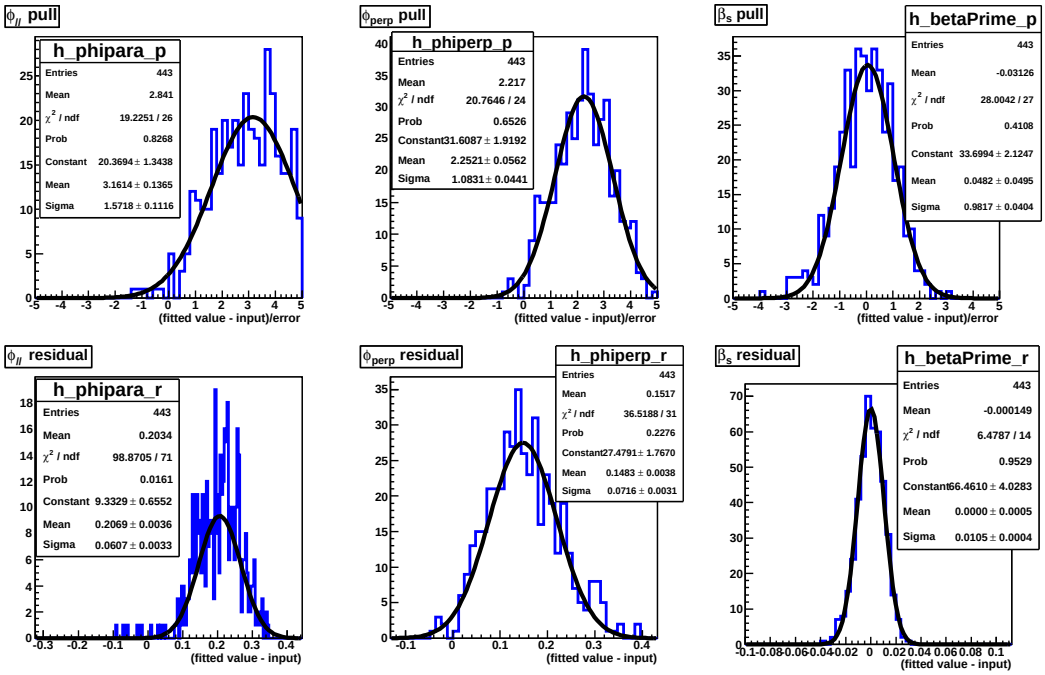

Figure 5.2: Pull distributions and residuals for $\phi_{\parallel}$, $\phi_{\perp}$ and $\beta_s^{J/\psi\phi}$ from high statistics tests with A_{SW}=0.5, $\delta_{SW} = 0.0$

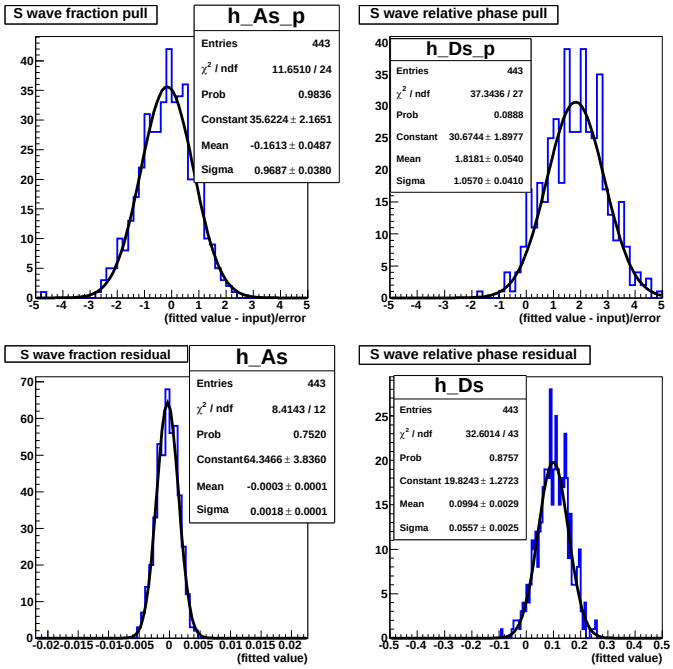

Figure 5.3: Pull distributions and residuals for S-wave fraction and relative phase from high statistics tests with input A_{SW}=0.5, $\delta_{SW} = 0.0$

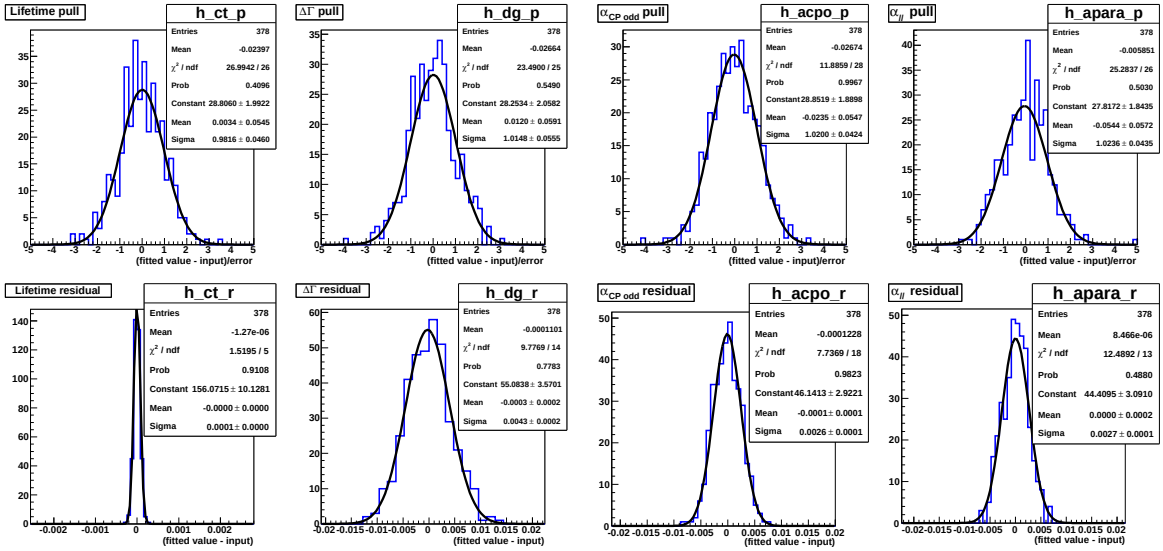

Figure 5.4: Pull distributions and residuals for $\Delta\Gamma$, $c\tau$ and the transversity amplitudes from high statistics tests with A_{SW}=0.5, $\delta_{SW} = 0.0$ and the input value of $\phi_{\parallel}$ changed to 1.5 to avoid its reflection point at π.

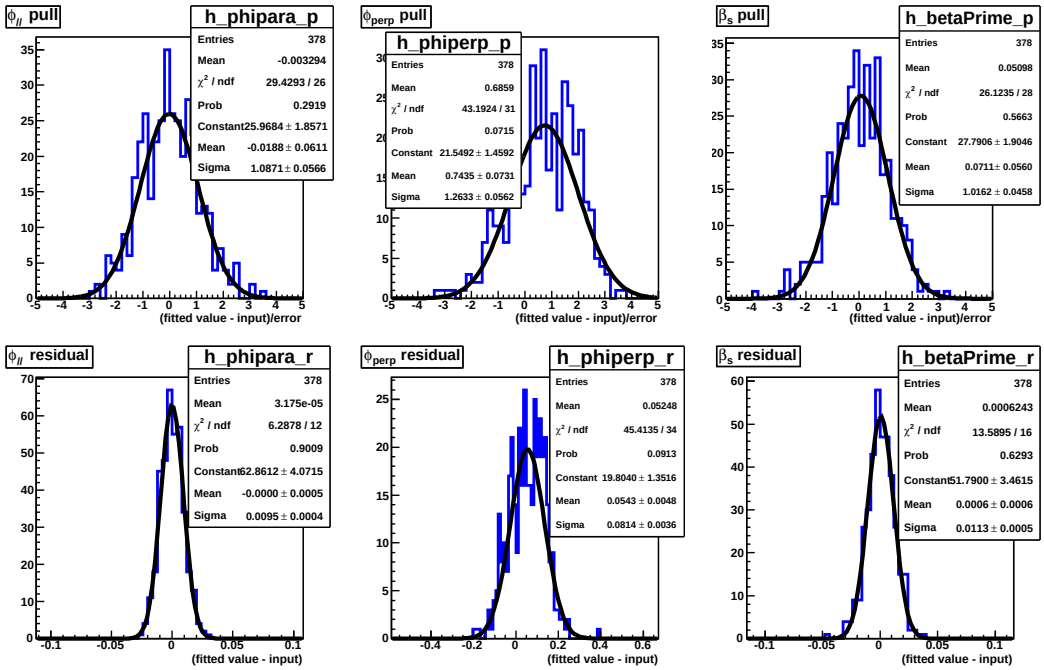

Figure 5.5: Pull distributions and residuals for $\phi_{\parallel}$, $\phi_{\perp}$ and $\beta_s^{J/\psi\phi}$ from high statistics tests with A_{SW}=0.5, $\delta_{SW} = 0.0$ and the input value of $\phi_{\parallel}$ changed to 1.5 to avoid its reflection point at π.

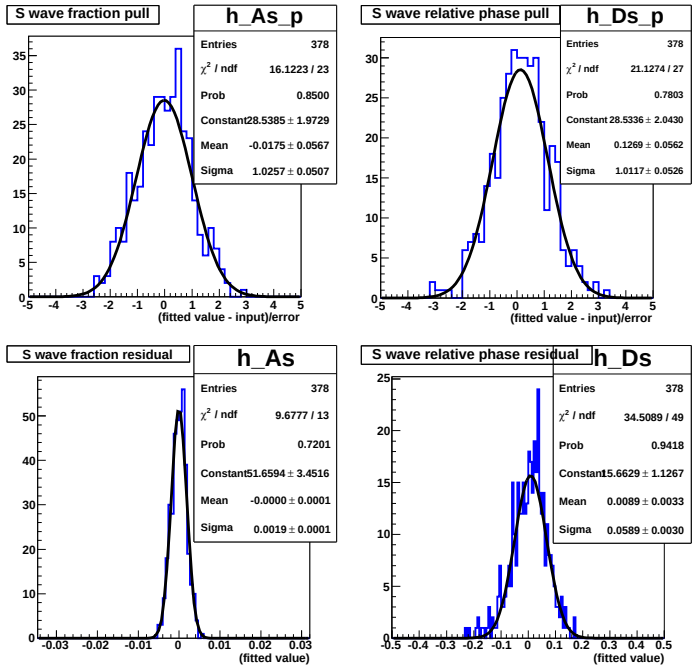

Figure 5.6: Pull distributions and residuals for S-wave fraction and relative phase from high statistics tests with input A_{SW}=0.5, $\delta_{SW} = 0.0$ and the input value of $\phi_{\|}$ changed to 1.5 to avoid its reflection point at π.

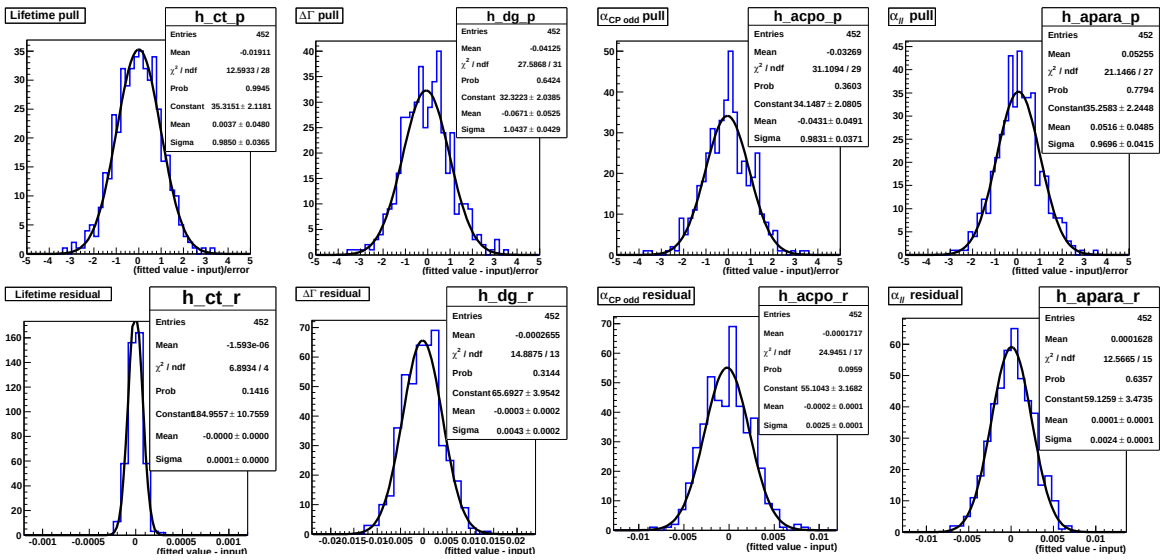

Figure 5.7: Pull distributions and residuals for $\Delta\Gamma$, $c\tau$ and the transversity amplitudes from high statistics tests with A_{SW}=0.5, $\delta_{SW} = 0.0$ and the input value of $\phi_{\|}$ changed to 1.0 to avoid its reflection point at π.

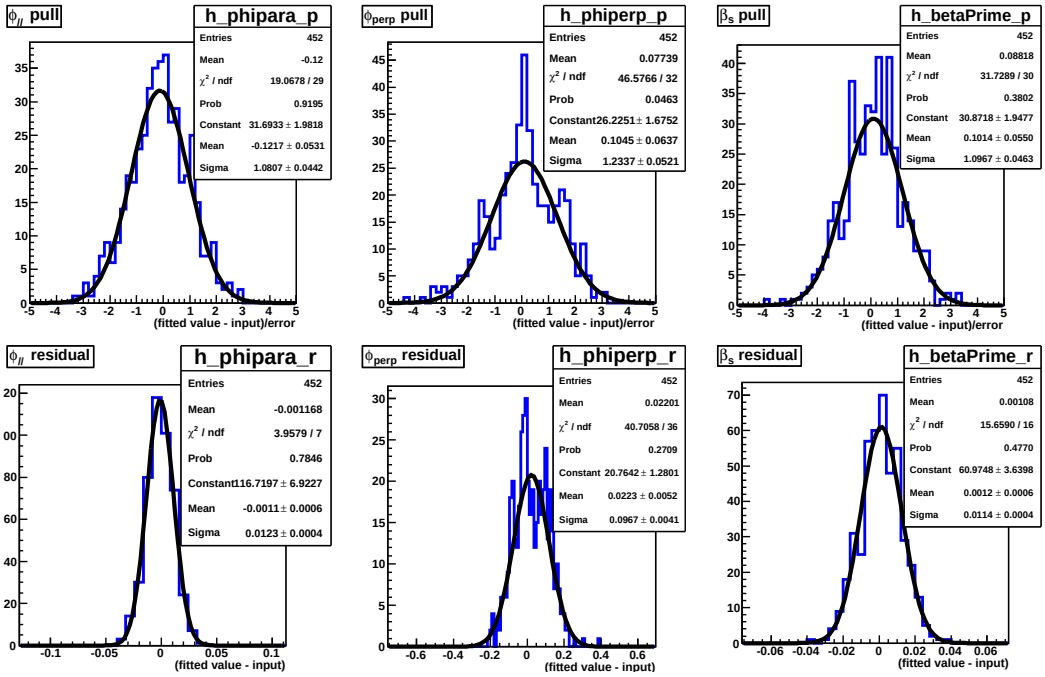

Figure 5.8: Pull distributions and residuals for $\phi_{\|}$, $\phi_{\perp}$ and $\beta_s^{J/\psi\phi}$ from high statistics tests with A_{SW}=0.5, $\delta_{SW} = 0.0$ and the input value of $\phi_{\|}$ changed to 1.0 to avoid its reflection point at π

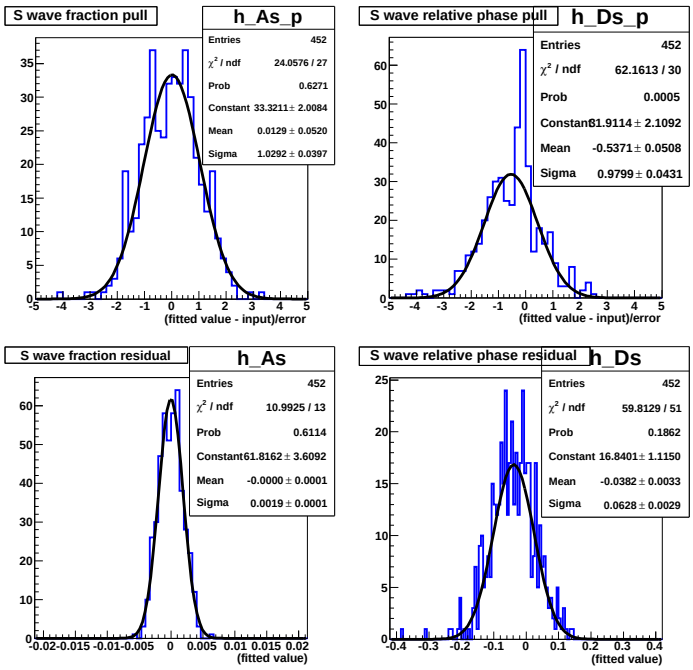

Figure 5.9: Pull distributions and residuals for S-wave fraction and relative phase from high statistics tests with input A_{SW}=0.5, $\delta_{SW} = 0.0$ and the input value of $\phi_{\parallel}$ changed to 1.0 to avoid its reflection point at π

Two cross checks have been carried out to complement the pull studies described so far in this section. To check the correlation observed between $\phi_{\perp}$ and δ_{SW}, the correlation matrix was investigated. Appendix E gives the full correlation matrix for the fit to a signal only pseudo experiment with 1000000 events, and for the full fit to data. These confirm the strong correlation between δ_{SW} and the phases $\phi_{\parallel}$ and $\phi_{\perp}$, which both exhibit bad pulls for some generated values when there is a large S-wave component, but not between δ_{SW} and the main parameters of interest. Finally, likelihood scans are presented for the S-wave parameters and the main parameters of interest for the fit to data (Figure 5.10) and for one example high statistics pseudo experiment with 1000000 signal events (Figure 5.11). These highlight the behaviour of $\phi_{\parallel}$ close to its reflection point, and demonstrate the non-parabolic errors of δ_{SW}. The asymmetry of the likelihood scan for δ_{SW} is particularly evident in Figure 5.11. Perhaps the most revealing of these plots is the scan for $\phi_{\perp}$ in the high statistics study in Figure 5.11, which shows distinctly non-parabolic errors due to a reflection about π similar to that observed for $\phi_{\parallel}$. In the data scans (Figure 5.10), this behaviour for $\phi_{\perp}$ is not evident, but in the large A_{SW}, high statistics sample this explains the biased pull for this parameter seen in Figure 5.2.

The conclusion of this high statistics fitter validation study is that despite persistent biases in the phases $\phi_{\parallel}$, $\phi_{\perp}$ and δ_{SW} for certain values of the S-wave fraction and relative phase, the main parameters of interest are unbiased. The biases in the phases are likely to be due to the non-parabolic nature of the likelihood distributions for these parameters shown in figure 5.11, and would be expected to disappear in the case of infinite statistics. Comprehensive studies have confirmed that these do not affect $\beta_s^{J/\psi\phi}$, $\Delta\Gamma$ or the B_s^0 lifetime, and the measurement of the S-wave fraction, A_{SW}, is largely unbiased.

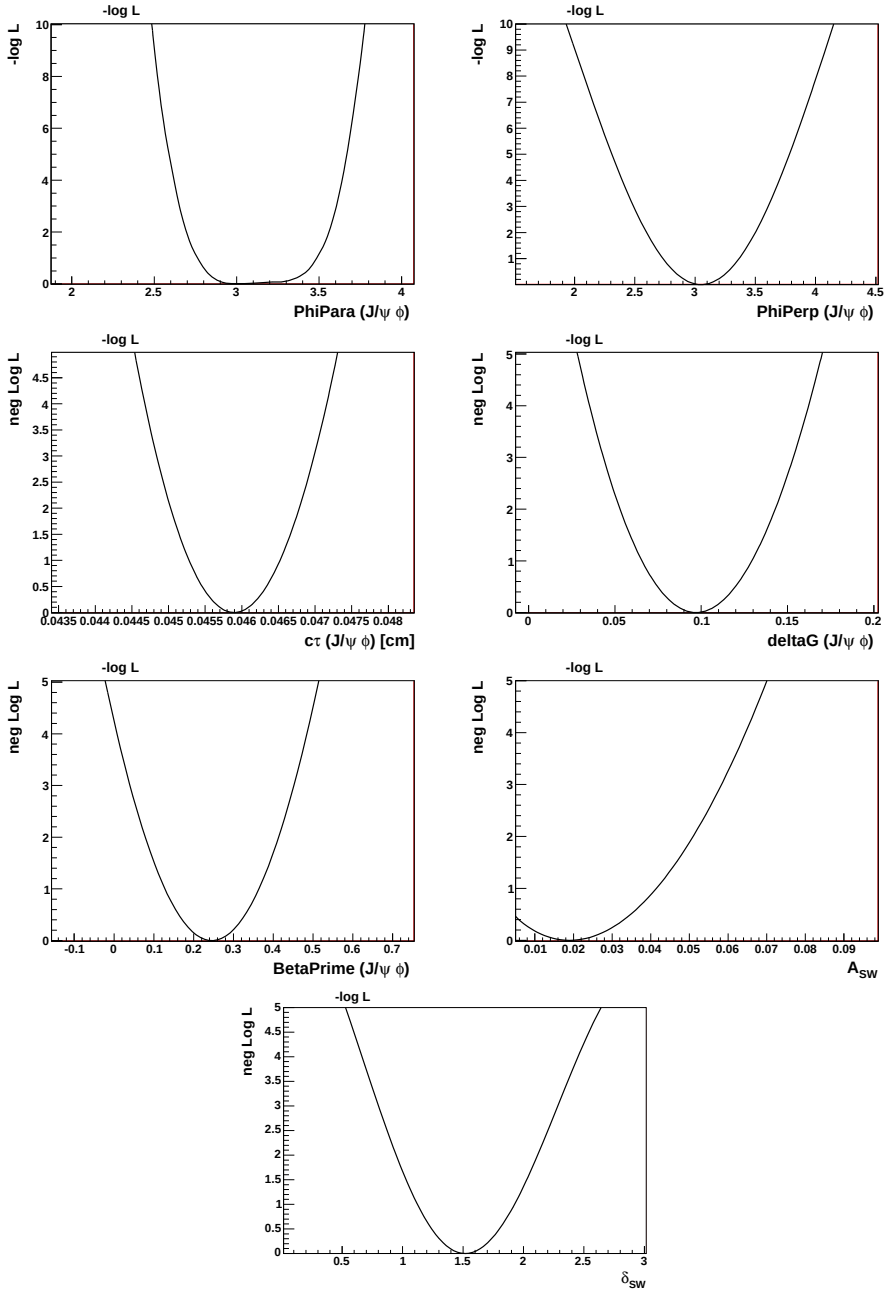

Figure 5.10: Likelihood scans for main parameters of interest from fit to data sample, with full flavour tagging

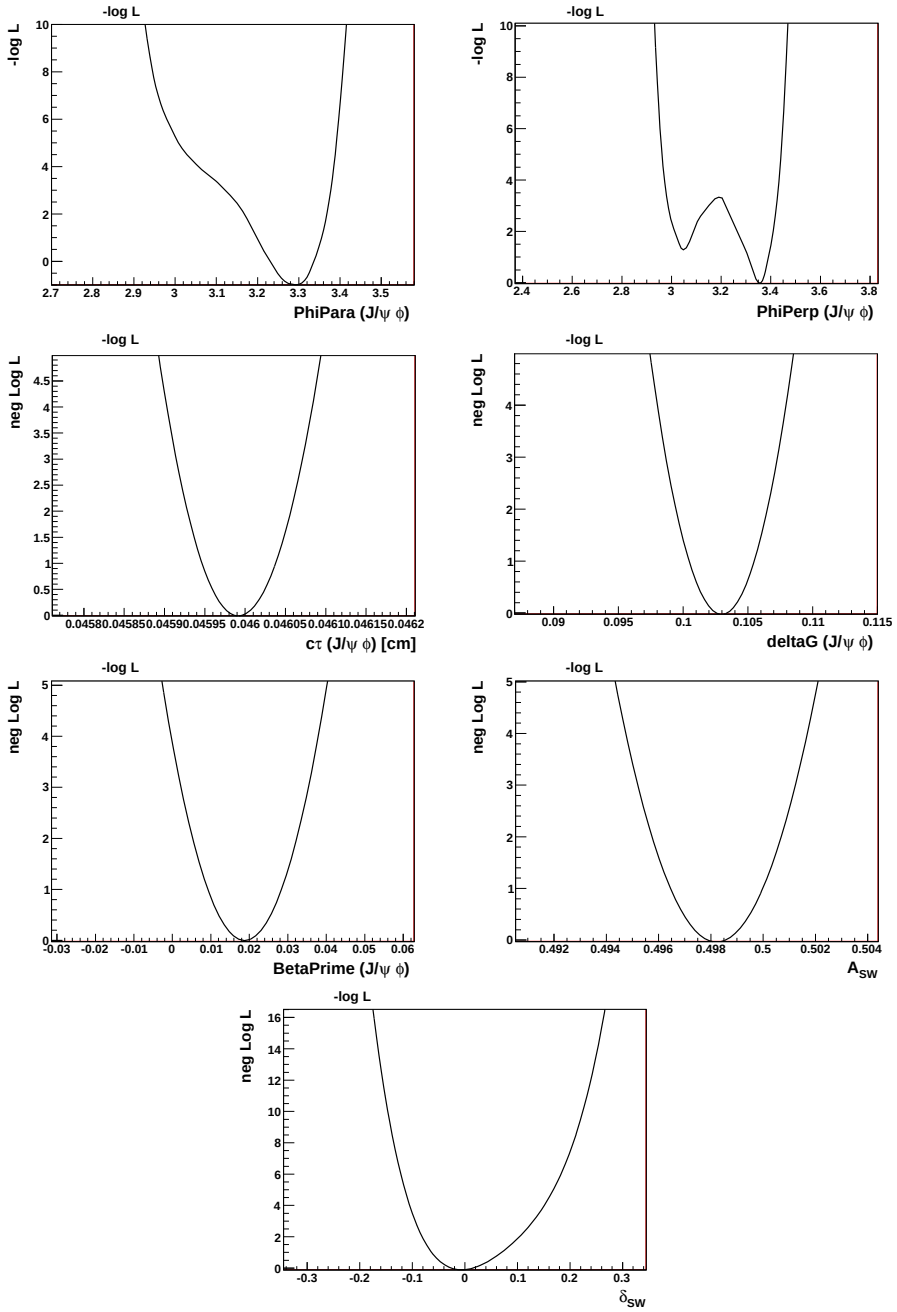

Figure 5.11: Likelihood scans for main parameters of interest from fit to 1000000 signal only simulated events generated with $A_{SW} = 0.5$, $\delta_{SW} = 0.0$, $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ at their SM values and all other parameters from a fit to data.

5.3 Realistic sample tests of the fitter

The pull study technique described at the start of this chapter is used to assess the biases for the parameters of interest in several configurations of the fit with sample sizes generated to match the current data sample available:

- the full fit with flavour tagging and $\beta_s^{J/\psi\phi}$ floating,
- the fit with flavour tagging but with $\beta_s^{J/\psi\phi}$ fixed to 0.0 ($\sim$SM value),
- the fit with $\beta_s^{J/\psi\phi}$=0.0 and no flavour tagging.

Previous $B_s^0 \to J/\psi\,\phi$ analyses at CDF have presented best fit values for physics parameters of interest using a non-flavour tagged fit [7, 4], because of persistent biases in the likelihood fitter at low statistics when flavour tagging was included. With the extended dataset of $L = 5.2$ fb^{-1} the feasibility of publishing values for the flavour tagged likelihood fit has been studied. Ideally, fit values for all physics parameters including $\beta_s^{J/\psi\phi}$ would be presented using the full flavour tagged fit with $\beta_s^{J/\psi\phi}$ floating, but in earlier iterations this configuration of the fitter showed bad biases in realistic data sample sizes [103]. This has been checked for the new fitter with the present, larger data sample, but non-negligible biases in several key variables remain present. These studies are discussed and shown in Section 5.3.2. Using flavour tagged fit with $\beta_s^{J/\psi\phi}$ fixed to 0.0 is a good compromise, as it exhibits unbiased pulls (shown in Section 5.3.1) and still provides a good cross check of the main fitter used in the full $\beta_s^{J/\psi\phi}$ measurement as explained in Chapter 6.

Generally, for the main parameters of interest the input values for generation of pseudo experiments are randomised across their expected ranges, these parameters and their ranges are shown in Table 5.5. The exceptions are $\phi_\parallel$ and $\phi_\perp$; $\phi_\parallel$ exhibits worse behaviour in the fit for values close to π (see Figure 5.14 and discussion in following section), which is a point of reflection for this parameter, and this effect would be washed out if the input values were randomised across its expected range. For this reason, $\phi_\parallel$ and $\phi_\perp$ are generated at the values fitted from data for each configuration of the fit.

Parameter	Minimum	Maximum
ct	$400\mu m$	$500\mu m$
$\Delta\Gamma$	0.0ps^{-1}	0.6ps^{-1}
$\alpha_\perp$	0.0%	90.0%
$\alpha_\parallel$	0.0%	90.0%
$S-$wave fraction	0.0%	10.0%
$S-$wave relative phase	$-\pi$	π

Table 5.5: Ranges of randomized inputs for realistic sample size pull studies.

This simultaneous randomisation technique was shown in [103] to give an equivalent level of accuracy to the more standard approach of varying each parameter individually with the remaining parameters fixed, and has the advantage of covering the full parameter space whilst minimising the necessary CPU time.

For the pull studies, 1000 pseudo experiments are generated with signal and background events corresponding to the full dataset, for each fit configuration being tested. The input values for generation of all parameters other than those discussed so far in this section are taken from the corresponding fit values from data, with $\beta_s^{J/\psi\phi}$ fixed or floating, and with or without flavour tagging. Tables 7.1, 6.1 and 6.2 show the relevant values from data for each of the fit configurations.

5.3.1 Studies at the Standard Model point

When $\beta_s^{J/\psi\phi}$ is fixed to 0.0, which is approximately the Standard Model expected value [3], the likelihood function is simplified, as can be seen from Equation 4.38. Studies with this condition applied are shown in this section.

Flavour tagged fit with $\beta_s^{J/\psi\phi}$ fixed to the SM prediction

For the flavour tagged likelihood fit with $\beta_s^{J/\psi\phi}$ fixed to 0.0, the pull distributions show little significant deviation from unit pulls. The parameters which do show small biases in the pull distributions are dealt with by adding a systematic uncertainty to their errors to account for the effect.

Figures 5.12 to 5.13 show the pull distributions for the flavour tagged fit with $\beta_s^{J/\psi\phi}$=0.0, while Table 5.6 lists the mean and standard deviations for the pulls distributions for the parameters of interest. Full tables of pulls for all parameters in the fit are given in Appendix F. The systematic error for parameters with biased pulls is included by adding the fraction of the statistical error according to the size of the mean shift of the pull. Any parameter with a pull bias of $> 2\sigma$ is treated in this way; for $\alpha_{\parallel}$ and α_{CPOdd} 15% of the statistical error is included as a systematic effect and for $\Delta\Gamma$ 8% of the statistical error is added. These percentages correspond to the shifts in the pulls for the parameters effected, as shown in Table 5.6.

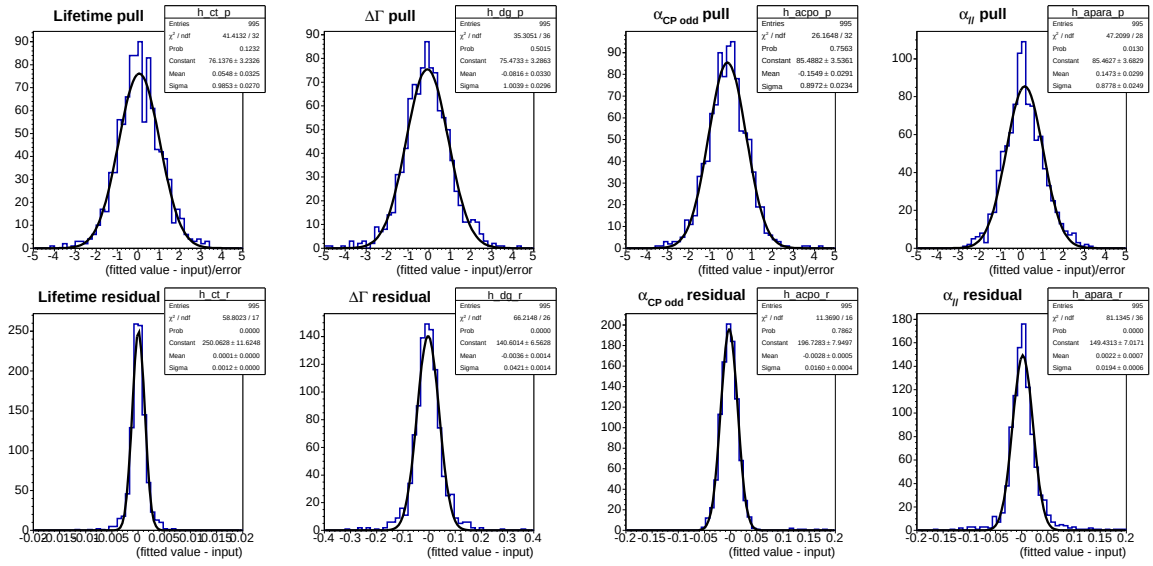

Figure 5.12: Pull distributions and residuals for $\Delta\Gamma$, $c\tau$ and the transversity amplitudes α_{CPOdd} and $\alpha_{\parallel}$ from tagged fit with fixed $\beta_s^{J/\psi\phi}$=0.0 . All fit parameters shown in these plots are summarised in Table 5.6.

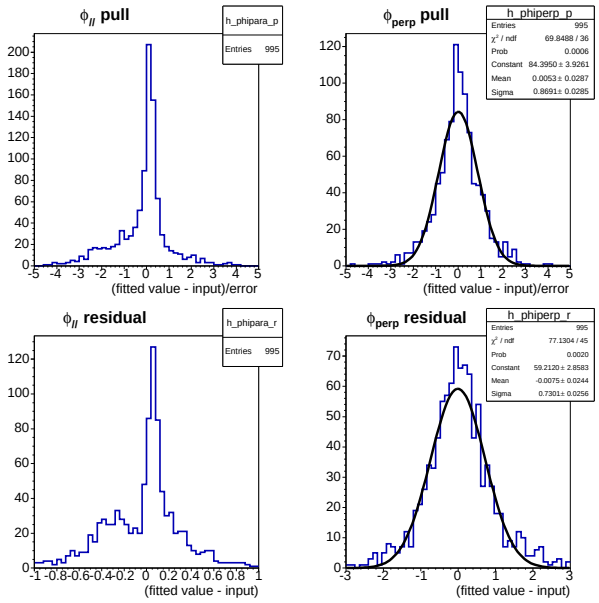

Figure 5.13: Pull distributions and residuals for $\phi_{\|}$ and $\phi_{\perp}$ from tagged fit with fixed $\beta_s^{J/\psi\phi}$=0.0 . All fit parameters shown in these plots are summarised in Table 5.6.

Parameter	Pull mean	Pull σ	Residual mean	Residual σ
$c\tau$	0.05 ± 0.03	0.99 ± 0.03	0.00007 ± 0.00004	0.001 ± 0.00004
$\Delta\Gamma$	-0.08 ± 0.03	1.00 ± 0.03	-0.004 ± 0.001	0.042 ± 0.0014
α_{cpo}	-0.15 ± 0.03	0.90 ± 0.02	-0.004 ± 0.001	0.016 ± 0.0004
$\alpha_{\|}$	0.15 ± 0.03	0.88 ± 0.03	0.002 ± 0.001	0.019 ± 0.00064
$\phi_{\|}$	non-Gaussian		non-Gaussian	
$\phi_{\perp}$	0.01 ± 0.02	0.87 ± 0.03	-0.008 ± 0.024	0.73 ± 0.026

Table 5.6: Pull study fit results for main parameters of interest using tagged fit with fixed $\beta_s^{J/\psi\phi}$=0.0. Details for all fit parameters are given in Appendix F

The pulls for $\phi_{\|}$ show non-Gaussian behaviour, as can be seen in Figure 5.13, so it is not possible to quote a value for this parameter with unbiased errors. There is a reflected symmetry about π for $\phi_{\|}$, and for values close to π the the fit cannot always determine clearly between the two cases. This results in a tendency for the fit to return a value of π, the point exactly between them, as the fitted value. In previous analyses at CDF the fit has been *boxed* for $\phi_{\|}$ values above or below π, but as the minimum is now so close to π the boxing has been removed. The effect is illustrated in Figure 5.14, which shows the distribution of fitted values for an input $\phi_{\|}$=3.08, with a peak in the fitted values at π, and the likelihood profile for this parameter. This effect also appears in the plot of the residual for $\phi_{\|}$ in Figure 5.13 which instead of a symmetric distribution as would be ideal, has a significant peak at $\phi_{\|fit}$-$\phi_{\|input}$=0.08=π-3.08.

As a cross check, this pull study was repeated with a generated value of $\phi_{\|}$ = 1.5 in an ensemble of 500 pseudo experiments. The results are shown in Figure 5.15; this parameter demonstrates unbiased behaviour when the generated value is far from the reflection point at π.

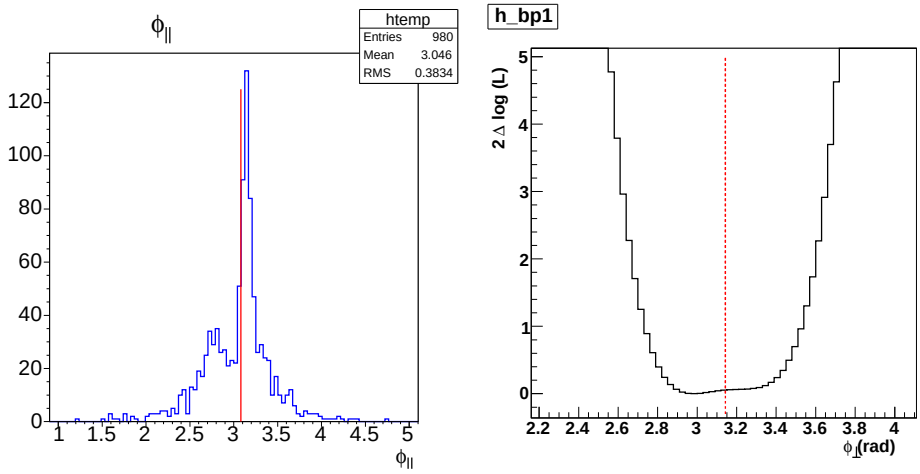

Figure 5.14: [Left]Fitted values for $\phi_{\parallel}$ with input value for pseudo experiments marked in red [Right] Likelihood profile (minimised at each value of $\phi_{\parallel}$) for $\phi_{\parallel}$ with reflection point marked as a dashed red line at π

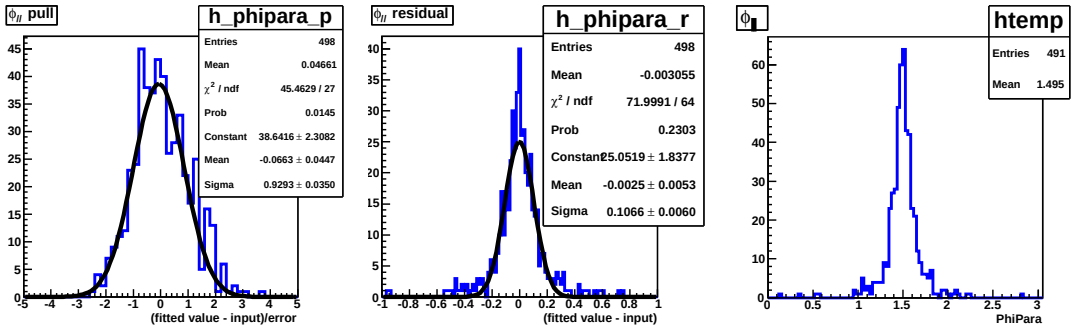

Figure 5.15: [Left] Unbiased pull for $\phi_{\parallel}$ from pseudo experiments with input value 1.5 [centre] $\phi_{\parallel}$ residual [right] fitted distribution of $\phi_{\parallel}$

Fit without flavour tagging, with fixed $\beta_s^{J/\psi\phi}$=0.0

The untagged fit pull distributions here are included as a cross check. Figures 5.16 to 5.17 show the untagged fit pull distributions, Table 5.7 lists the fitted mean and standard deviations for these pulls, the pulls for all fit parameters are given in Appendix F.

Parameter	Pull mean	Pull σ	Residual mean	Residual sigma
$c\tau$	-0.02 ± 0.03	1 ± 0.02	-3e-06 ± 4e-05	0.001 ± 4e-05
$\Delta\Gamma$	0.06 ± 0.03	1 ± 0.02	0.002 ± 0.002	0.05 ± 0.002
α_{cpo}	-0.1 ± 0.03	1 ± 0.03	-0.003 ± 0.0005	0.02 ± 0.0004
$\alpha_{\parallel}$	0.1 ± 0.03	1 ± 0.02	0.001 ± 0.0007	0.02 ± 0.0008
$\phi_{\parallel}$	non-Gaussian		non-Gaussian	

Table 5.7: Pull study fit results for main parameters of interest from untagged fit with fixed $\beta_s^{J/\psi\phi}$=0.0 . Full details of the pulls for all fit parameters are given in Appendix F

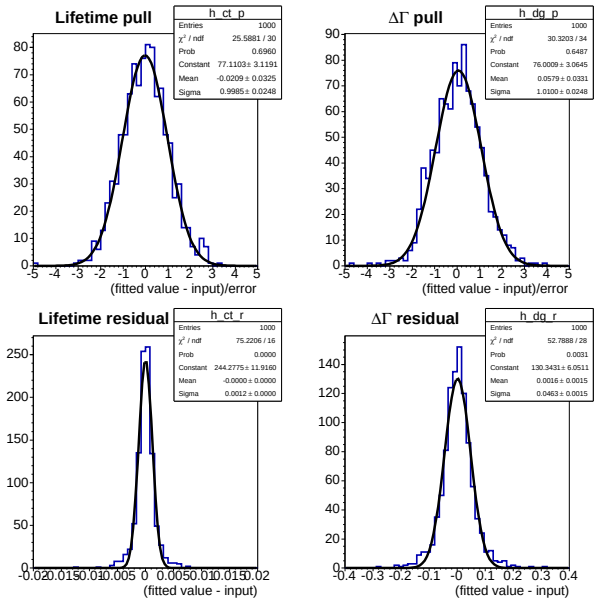

Figure 5.16: Pull distributions and residuals for $\Delta\Gamma$ and $c\tau$ from untagged fit with fixed $\beta_s^{J/\psi\phi}$=0.0 . All fit parameters shown in these plots are summarised in Table 5.7.

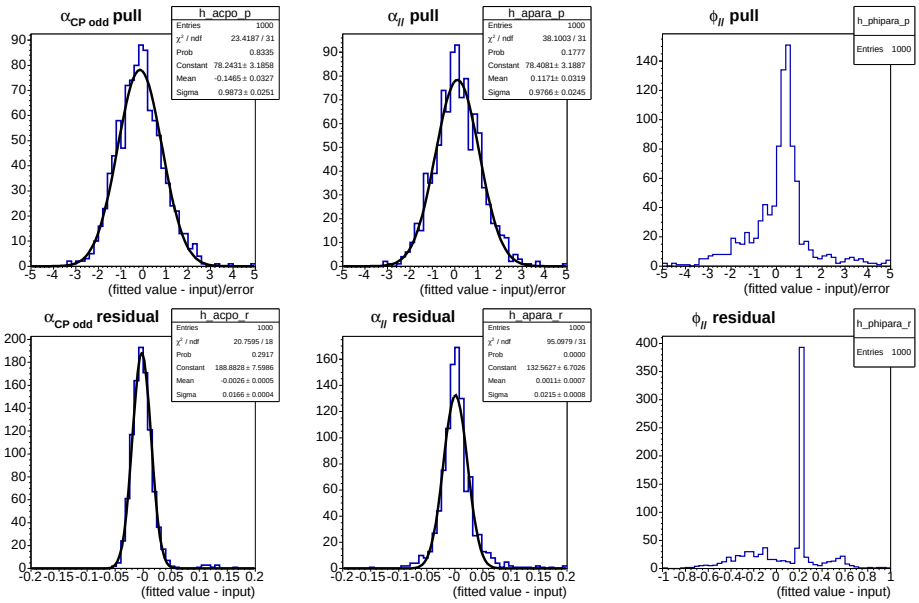

Figure 5.17: Pull distributions and residuals for α_{CPOdd}, $\alpha_{\parallel}$ and $\phi_{\parallel}$, from untagged fit with fixed $\beta_s^{J/\psi\phi}$=0.0 . All fit parameters shown in these plots are summarised in Table 5.7

5.3.2 Study of flavour tagged fit with $\beta_s^{J/\psi\phi}$ floating

In the following discussion, $\beta_s^{J/\psi\phi}$ is allowed to float in the fit. Pseudo experiments for this study are generated with at $\beta_s^{J/\psi\phi}$=0.24 (the central value found in the full data fit) which is relevant for judging the feasibility of presenting best fit values for the parameters of interest from the fit with $\beta_s^{J/\psi\phi}$ floating with the current sample size.

There are some significant biases in these results, especially for $\beta_s^{J/\psi\phi}$ and the angular

amplitudes. Figures 5.18- 5.19 show the pull distributions, the fitted mean and standard deviation for the pulls and residual shifts are listed in Table 5.8 for the parameters of interest, and in Appendix F for all fit parameters. These shifts are less significant than those seen in the previous $B_s^0 \to J/\psi\,\phi$ analyses at CDF with smaller data samples [7, 101], which is encouraging for the prospects of a study of $\beta_s^{J/\psi\phi}$ utilising the full CDF dataset at the end of data taking. It would be necessary to better understand the behaviour of the fit across the full range of $\beta_s^{J/\psi\phi}$ to ensure that any biases have been accounted for before quoting a single value within errors for this parameter. The level of non-Gaussian behaviour exhibited here combined with the unresolved symmetries in the likelihood function prevent the possibility of presenting a point value for $\beta_s^{J/\psi\phi}$ at this stage.

Parameter	Pull mean	Pull σ	Residual mean	Residual sigma
$c\tau$	-0.23 ± 0.03	0.91 ± 0.025	-0.00035 ± 4.7e-05	0.0014 ± 4.4e-05
$\Delta\Gamma$	0.24 ± 0.032	0.96 ± 0.026	0.016 ± 0.0017	0.048 ± 0.0017
α_{cpo}	-0.18 ± 0.03	0.92 ± 0.024	-0.0029 ± 0.0005	0.016 ± 0.00041
$\alpha_{\parallel}$	0.1 ± 0.031	0.96 ± 0.026	0.0017 ± 0.0007	0.021 ± 0.00075
$\phi_{\parallel}$	non-Gaussian		non-Gaussian	
$\phi_{\perp}$	0.056 ± 0.04	1.1 ± 0.049	0.05 ± 0.02	0.61 ± 0.017
$\beta_s^{J/\psi\phi}$	0.34 ± 0.032	0.95 ± 0.037	0.026 ± 0.0028	0.072 ± 0.0032

Table 5.8: Pull study fit results for tagged fit pulls with input $\beta_s^{J/\psi\phi}$=0.24 and $\beta_s^{J/\psi\phi}$ floating in the fit . Full details of the pulls for all fit parameters are given in Appendix F.

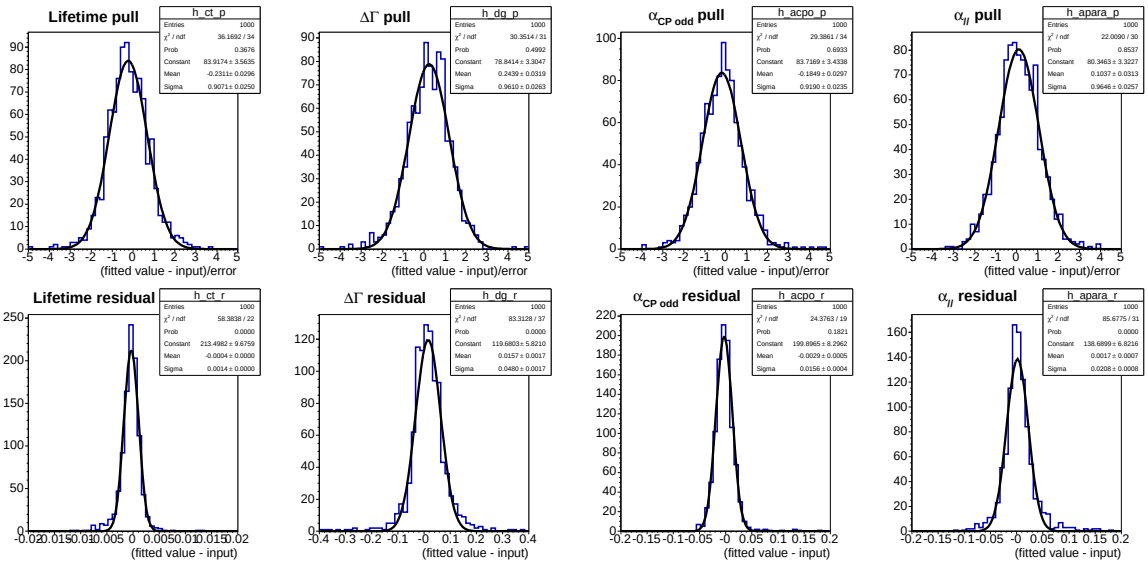

Figure 5.18: Pull distributions and residuals for $\Delta\Gamma$, $c\tau$, α_{CPOdd} and $\alpha_{\parallel}$ with $\beta_s^{J/\psi\phi}$ floating in the flavour tagged fit. All fit parameters shown in these plots are summarised in Table 5.8.

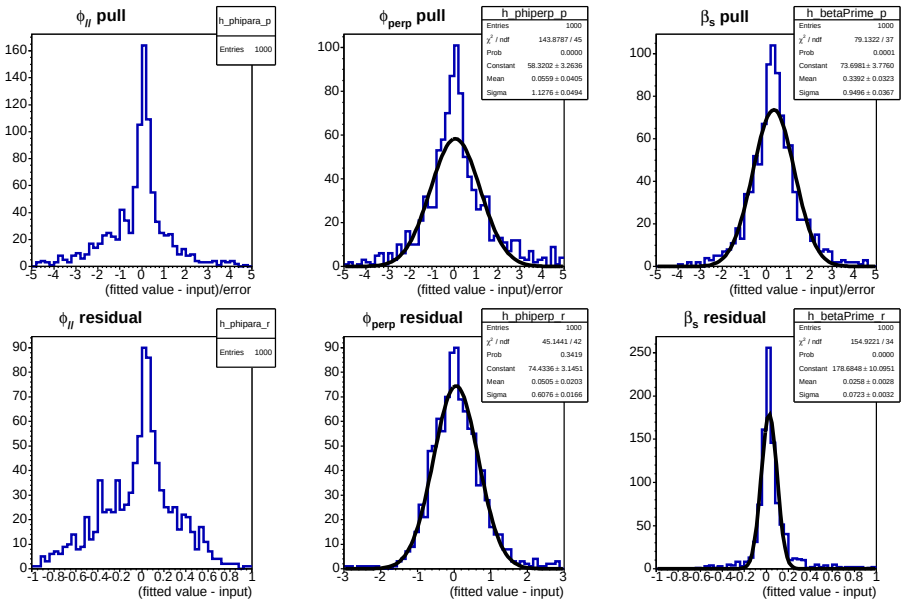

Figure 5.19: Pull distributions and residuals for $\phi_{\parallel}$, $\phi_{\perp}$and $\beta_s^{J/\psi\phi}$, with $\beta_s^{J/\psi\phi}$ floating in the flavour tagged fit. All fit parameters shown in these plots are summarised in Table 5.8.

Chapter 6

Measurement of B_s^0 lifetime and $\Delta\Gamma_s$ assuming no CP violation

In this chapter the results for five parameters of interest are presented assuming $\beta_s^{J/\psi\phi} = 0.0$, i.e. no CP violation. These parameters are the B_s^0 lifetime, τ_s, decay width difference $\Delta\Gamma$, transversity amplitudes, $|A_0(0)|^2$ and $|A_\parallel(0)|^2$, and the strong phase $\phi_\perp$. A full study of the effects of systematic uncertainties on each parameter is presented. Projections of the fit results onto individual parameters in data are shown, to demonstrate the quality of the fit. These measurements are the world's best values for τ_s and $\Delta\Gamma_s$, comparisons with previous measurements and world averages are discussed.

6.1 Fitted values at the Standard Model point

As was demonstrated in Chapter 5, the maximum likelihood fit produces minimally biased measurements of key physics parameters with $\beta_s^{J/\psi\phi}$ fixed to 0.0, approximately the Standard Model expected value [3]. Aside from constraining $\beta_s^{J/\psi\phi}$, the likelihood fitter used to produce these SM point estimates is identical to that used in the final measurement of $\beta_s^{J/\psi\phi}$, including use of flavour tagging variables. Thus, in addition to giving the world's most precise measurements of τ_s, $\Delta\Gamma_s$ and the angular amplitudes, comparing these values to theoretical predictions and other measurements provides a good cross check of the full analysis technique.

The fit parameters for which unbiased SM point values can be quoted are the proper decay length, $c\tau_s$ (from which can be calculated the B_s^0 lifetime, τ_s), the decay width difference, $\Delta\Gamma_s$, the angular amplitudes $|A_0(0)|^2|$ and $|A_\parallel(0)|^2|$, and the strong phase $\phi_\perp$. The other strong phase, $\phi_\parallel$ was shown to demonstrate non-Gaussian behaviour even with $\beta_s^{J/\psi\phi}$ fixed to zero as explained in Section 5.3. For the fit without flavour tagging, $\phi_\perp$ is not in the likelihood function when $\beta_s^{J/\psi\phi}$ is set to 0.0, so in the comparison of results at the end of this chapter, $\phi_\perp$ is only given for the flavour tagged fit.

Table 6.1 shows the best fit values for the 34 floating parameters of the PDFs described in Chapter 4 where $\beta_s^{J/\psi\phi}$ is fixed to zero. As a further check of the stability of the fit, the results at the Standard Model point are assessed without flavour tagging, the 22 floating variables for this case are shown in Table 6.2.

Parameter	Definition	Value
A_{sw}	fraction of S-wave KK component in signal	0.019 ± 0.027
$\alpha_\perp$	CP odd fraction	0.266 ± 0.014
$\alpha_\parallel$	$A_\parallel$ fraction in CP even states	0.306 ± 0.015
p_1	mass background slope	-2.3 ± 0.562
$S_\mathcal{D}(SST)$	SST dilution scale factor	0.924 ± 0.0847
$S_\mathcal{D}(OST+)$	OST dilution scale factor +	1.12 ± 0.092
$S_\mathcal{D}(OST-)$	OST dilution scale factor -	0.901 ± 0.175
δ_{sw}	Relative phase of S-wave KK component	1.37 ± 0.77
$\epsilon_b(OST)$	OST tagging efficiency for background	0.87 ± 0.002
$\epsilon_b(SST)$	SST tagging efficiency for background	0.719 ± 0.00269
$\mathcal{A}^+(OST)$	OST background positive tag asymmetry	0.495 ± 0.00319
$\mathcal{A}^+(SST)$	SST background positive tag asymmetry	0.496 ± 0.00349
$\epsilon_s(OST)$	OST tagging efficiency for signal	0.943 ± 0.00335
$\epsilon_s(SST)$	SST tagging efficiency for signal	0.522 ± 0.00681
FCN	NLL value	-5.29e+05 ± 0
f_p	Prompt fraction of background	0.884 ± 0.00519
f_-	Fraction of bkg which decays w/λ_-	0.173 ± 0.0336
f_{++}	Fraction of bkg which decays w/λ_{++}	0.662 ± 0.0452
f_s	Signal Fraction	0.181 ± 0.00236
λ_-	Effective background lifetime, neg. comp.	0.0382 ± 0.00391
λ_+	Effective background lifetime, pos. comp. 1	0.0407 ± 0.00364
λ_{++}	Effective background lifetime, pos. comp. 2	0.0112 ± 0.00117
$c\tau$	average of $c\tau_H$ and $c\tau_L$	0.0459 ± 0.000754
s_m	Mass error scale factor	1.73 ± 0.0216
m	B hadron mass [MeV/c^2]	5.37 ± 0.000131
N_{events}	Number of signal events	6.5e+03 ± 84.8
ϕ_1	First parameter in bkg fit to ϕ	0.139 ± 0.00837
$\phi_\parallel$	arg($A_\parallel A_0$) asymmetry parameter	3.08 ± 0.632
$\phi_\perp$	arg($A_\perp A_0$) asymmetry parameter	2.95 ± 0.637
$s_{c\tau}1$	Lifetime error scale factor 1	1.27 ± 0.0138
$s_{c\tau}2$	Lifetime error scale factor 2	3.32 ± 0.19
f_{sf1}	fraction of 1st lifetime error scale factor	0.882 ± 0.0119
$\cos(\psi)_1$	First parameter in bkg fit to cos(ψ)	0.0043 ± 0.0201
$\cos(\theta)_1$	First parameter in bkg fit to cos(θ)	0.161 ± 0.0176
$\Delta\Gamma$	CP asymmetry parameter [ps^{-1}]	0.0746 ± 0.035
Δm_s	B_s^0 mixing frequency	17.7 ± 0.109

Table 6.1: Fit results for the fit with flavor tagging with $\beta_s^{J/\psi\phi}$ fixed to 0.0 (the SM point)

6.2 Estimation of systematic uncertainties

Several assumptions or effects that are not fully accounted for in the likelihood function are considered to be sources of systematic uncertainty. Systematic uncertainties can occur due to mis-parameterisation in the fit model, assumptions in the model, or physical effects which are not well known or fully incorporated into the model. Every effort is made to minimise biases present in the fit model, but it is important to quantify any remaining systematic uncertainties in order to produce a full and accurate description of the overall uncertainty for each measured parameter.

Parameter	Description — Value	
A_{sw}	fraction of S-wave KK component in signal	0.00984 ± 0.0263
$\alpha_\perp$	CP odd fraction	0.267 ± 0.0146
$\alpha_\parallel$	$A_\parallel$ fraction in CP even states	0.309 ± 0.0157
p_1	mass background slope	-2.28 ± 0.562
δ_{sw}	Relative phase of S-wave KK component	1.57 ± 0.602
FCN	NLL value	-5.29e+05 ± 0
f_p	Prompt fraction of background	0.885 ± 0.00515
f_-	Fraction of bkg which decays w/λ_-	0.168 ± 0.0328
f_{++}	Fraction of bkg which decays w/λ_{++}	0.654 ± 0.0469
f_s	Signal Fraction	0.181 ± 0.00238
λ_-	Effective background lifetime, neg. comp.	0.0384 ± 0.00394
λ_+	Effective background lifetime, pos. comp. 1	0.0407 ± 0.00365
λ_{++}	Effective background lifetime, pos. comp. 2	0.0114 ± 0.00121
$c\tau$	average of $c\tau_H$ and $c\tau_L$	0.0457 ± 0.000769
s_m	Mass error scale factor	1.72 ± 0.0216
m	B hadron mass [MeV/c^2]	5.37 ± 0.000131
N_{events}	Number of signal events	6.51e+03 ± 85.4
ϕ_1	First parameter in bkg fit to ϕ	0.139 ± 0.00837
$\phi_\parallel$	arg($A_\parallel A_0$) asymmetry parameter	2.93 ± 0.321
$\phi_\perp$	arg($A_\perp A_0$) asymmetry parameter	1.27 ± 0.0132
$s_{c\tau}1$	Lifetime error scale factor 1	3.35 ± 0.181
$s_{c\tau}2$	Lifetime error scale factor 2	0.884 ± 0.0111
f_{sf1}	fraction of 1st lifetime error scale factor	0.00462 ± 0.0201
$\cos(\psi)_1$	First parameter in bkg fit to cos(ψ)	0.159 ± 0.0176
$\cos(\theta)_1$	First parameter in bkg fit to cos(θ)	0.572 ± 0
$\Delta\Gamma$	CP asymmetry parameter [ps^{-1}]	0.0705 ± 0.0357

Table 6.2: Fit results for the fit without flavor tagging, with $\beta_s^{J/\psi\phi}$ fixed to 0.0 (the SM point)

To estimate the size of these uncertainties, an effect is identified which could be a source of systematic uncertainty, then a set of 1000 pseudo-experiments are generated with this effect included in the simulations but not in the likelihood fit. The studied uncertainties are described in more detail in the following subsections. A set of reference pseudo-experiments, with no systematic variations, is also produced and fitted in the same way. The systematic error quoted quoted for each effect is the difference between the mean shift of the fitted value of each parameter from its input value for the pseudo-experiments with the systematic alteration included, and the equivalent shift for the reference set of pseudo-experiments which are generated using the default model. For each systematic effect studied, the same seed for randomisation is used as for the reference set in order to minimize any variations between the samples. Using the difference between the shift for the systematically altered experiments and the reference set is a conservative choice of method, which aims to avoid systematic effects being hidden by statistical fluctuations.

The individual systematic uncertainties are summed in quadrature and presented in Table 6.18 to give the total contribution to the errors for each parameter due to sources

of systematic uncertainty. Table 6.3 shows the deviations from the input values of the fitted parameters for pseudo-experiments generated without systematic variations and fitted with the default fit. The *mean*, *significance*, *sigma* and *prob* columns contain the parameters of a Gaussian fit to the pull distribution: (fitted value - input value)/error, "Shift" is the fitted minus the input value. For the tables in the following subsections, the final column ("syst") contains the systematic error for the systematically altered pseudo experiments estimated from this study, which is calculated as Equation 6.1:

$$\text{syst} = (\text{par}_{\text{fit}} - \text{par}_{\text{input}}) - \text{shift}_{\text{ref.}} \tag{6.1}$$

where par_{fit} and $\text{par}_{\text{input}}$ are values from the systematically altered set of pseudo experiments, and $\text{shift}_{\text{ref}}$ is the difference for the reference set of pseudo experiments generated using the default model.

Parameter	Mean	signi	σ	Prob	Shift
$c\tau$	-0.04 ± 0.033	-1.2	1.0	0.72	-0.22 μm
$\Delta\Gamma$	-0.0062 ± 0.035	-0.18	1.1	0.49	-0.00091 ps^{-1}
α_{CPOdd}	-0.24 ± 0.031	-7.7	0.93	0.25	-0.0036
$\alpha_{\parallel}$	0.19 ± 0.034	5.7	1.0	0.48	0.0026
$\phi_{\perp}$	-0.026 ± 0.029	-0.87	0.87	4.4e-10	-0.036 rad

Table 6.3: Summary of shifts from input value for reference pseudo experiments, these values are subtracted from the shifts found in the alternatively modelled pseudo experiments used to estimate the size of systematic uncertainties.

The following subsections describe the systematic effects which are accounted for, with tables detailing the size of each effect on the parameters of interest.

6.2.1 Signal angular efficiency

One source of systematic uncertainty is the modelling of the angular efficiency of the detector. In Section 4.4.1 it was explained that the efficiency is modelled by parameterising the angular efficiency in terms of an expansion using spherical harmonics and Legendre polynomials. The coefficients for this parameterisation are fitted in realistically reconstructed Monte Carlo simulations which are reweighted to match the p_T distributions in data (Section 3.8). If this model does not accurately reflect the true detector efficiency, it would introduce a systematic effect in the fit. This is tested by generating pseudo-experiments with the angular efficiency functions taken directly from 3-dimensional histograms of the realistic Monte Carlo used to fit the polynomial coefficients, instead of using the fitted function, then fitting them with the default model which uses the angular PDF described in Section 4.4.1.

Additionally, any inaccuracies in the p_T re-weighting of the Monte Carlo used to fit the angular efficiency parameters could introduce a systematic uncertainty. This is tested by generating pseudo-experiments as described in the previous paragraph, but using non-reweighted Monte Carlo histograms for the angular efficiency functions.

Table 6.4 shows the divergences of the fitted parameters from the input values for the study of the angular efficiency parameterisation with reweighted Monte Carlo, Table 6.5 shows the values for the test using non-reweighted Monte Carlo inputs.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	0.098 ± 0.034	2.9	1.0	0.34	0.74 μm	0.96 μm
$\Delta\Gamma$	0.058 ± 0.036	1.6	1.1	0.08	0.0015 ps^{-1}	0.0024 ps^{-1}
α_{CPOdd}	-0.5 ± 0.033	-15	1.0	0.71	-0.0073	-0.0037
$\alpha_{\parallel}$	0.89 ± 0.034	26	1.0	0.21	0.013	0.01
$\phi_{\perp}$	-0.028 ± 0.032	-0.87	0.95	0.0017	-0.02 rad	0.016 rad

Table 6.4: Summary of shifts from input value for pseudo experiments with signal angular efficiency functions taken directly from p_T reweighted Monte Carlo histograms, fit with the default parameterisation.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	0.13 ± 0.039	3.3	1.1	0.0038	0.72 μm	0.94 μm
$\Delta\Gamma$	-0.003 ± 0.038	-0.072	1.1	0.11	-0.0001 ps^{-1}	0.00081 ps^{-1}
α_{CPOdd}	-0.094 ± 0.038	-2.5	1.1	0.12	-0.002	0.002
$\alpha_{\parallel}$	1.4 ± 0.039	36	1.1	0.12	0.02	0.017
$\phi_{\perp}$	0.02 ± 0.036	0.55	1.0	0.017	-0.013 rad	0.022 rad

Table 6.5: Summary of shifts from input value for pseudo experiments with signal angular efficiency functions taken directly from non-reweighted Monte Carlo histograms, fit with the default parameterisation.

Another possible source of systematic uncertainty is mis-modelling of the true angular sculpting of the detector in the realistic MC sample; this is checked in two ways. One test of the model is to inspect the transversity angular distributions in background (sideband) data, in comparison with the angular sculpting of the signal distributions, as shown in Figure 6.1. The good agreement exhibited between these two distributions is expected, but not guaranteed as the differing kinematics of the signal and background events could lead to variations in the detector sculpting effects between them.

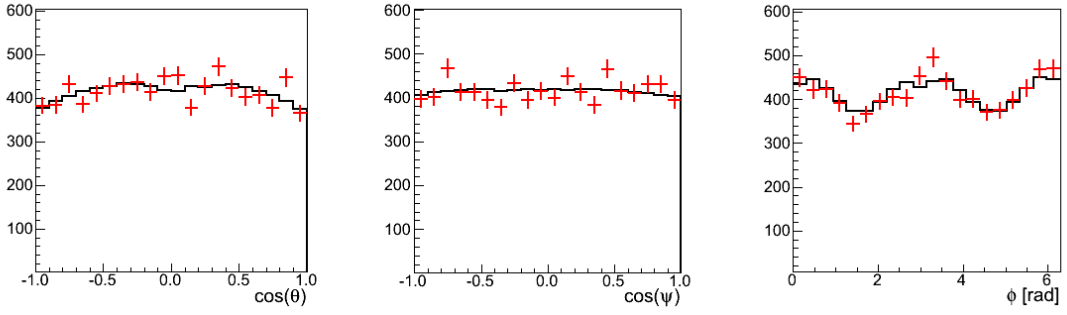

Figure 6.1: Comparison between transversity distributions of B_s^0 mass sideband region data (red points) and phase space generated signal only Monte Carlo (black histograms)

In addition to the positive comparison between the MC transversity angles and those found in sideband data, a further check was carried out to assess the effect of the modelling. This study takes the extreme case of ignoring the angular efficiency in the fit altogether, and fitting pseudo experiments generated with the default angular

model. The results are not intended to be included as a systematic uncertainty in the final results, simply to demonstrate that even in this extreme case of mis-modelling (by treating the distributions as flat), the systematic effects determined are not overwhelmingly large. Table 6.6 shows the results of this study.

Parameter	Mean	signi	*sigma*	Prob	Shift	Syst
$c\tau$	0.21 ± 0.062	3.4	1.2	0.13	1.6 μm	1.9 μm
$\Delta\Gamma$	-0.28 ± 0.057	-4.9	1.1	0.43	-0.016 ps^{-1}	-0.015 ps^{-1}
$\alpha_{\rm CPOdd}$	0.86 ± 0.061	14	1.2	0.2	0.011	0.015
$\alpha_{\parallel}$	1.4 ± 0.057	24	1.1	0.091	0.028	0.025
$\phi_{\perp}$	-0.065 ± 0.046	-1.4	0.87	0.0094	-0.018 rad	0.018 rad

Table 6.6: Shifts measured by fitting pseudo experiments generated with the default angular efficiency model, and fitted with the efficiencies considered flat. This study demonstrates an extreme case of mis-modelling in the realistic MC used to assess the angular efficiencies, which is far from the case shown in Figure 6.1 and is not included in the final sum of systematic errors.

The conclusion of the investigations into potential mis-modelling of the angular efficiencies in the realistic MC sample, is that no systematic uncertainty in the final results is included for this effect. An additional justification for this choice is the inclusion of two conservative estimates of systematic uncertainties due to the angular efficiency parameterisation and the MC re-weighting.

6.2.2 Signal mass model

By default, the signal B_s^0 mass is fitted with a single Gaussian model, as described in Section 4.3.1, using the PDF in Equation 4.5. If this model is not accurate for the true distribution in data, it would introduce a systematic uncertainty in the likelihood fit. To analyze the effect of mismodelling this parameter, a different model is used to generate pseudo-experiments, then they are fitted with the default fit. This alternative model uses a double Gaussian B_s^0 mass parameterisation and two mass error scale factors, shown in Equation 6.2 to simulate the effect on the fitted variables if the data had such a distribution but was fitted with the single Gaussian model.

$$\begin{aligned} P_s(m|M,\sigma_m 1,\sigma_m 2,f_m) &= (f_m)\frac{\frac{1}{\sqrt{2\pi}s_{m1}\sigma_{m1}}e^{\left[-\frac{1}{2}\left(\frac{m-M}{s_{m1}\sigma_{m1}}\right)^2\right]}}{\frac{1}{2}[\mathrm{erf}\left(\frac{\mathrm{M_{max}-M}}{\sqrt{2}\mathrm{s_{m1}}\sigma_{\mathrm{m1}}}\right)+\mathrm{erf}\left(\frac{\mathrm{M_{min}-M}}{\sqrt{2}\mathrm{s_{m1}}\sigma_{\mathrm{m1}}}\right)]} \\ &+ (1-f_m)\frac{\frac{1}{\sqrt{2\pi}s_{m2}\sigma_{m2}}e^{\left[-\frac{1}{2}\left(\frac{m-M}{s_{m2}\sigma_{m2}}\right)^2\right]}}{\frac{1}{2}[\mathrm{erf}\left(\frac{\mathrm{M_{max}-M}}{\sqrt{2}\mathrm{s_{m2}}\sigma_{\mathrm{m2}}}\right)+\mathrm{erf}\left(\frac{\mathrm{M_{min}-M}}{\sqrt{2}\mathrm{s_{m2}}\sigma_{\mathrm{m2}}}\right)]} \end{aligned} \quad (6.2)$$

The parameters for generating the alternative model pseudo-experiments are obtained by fitting the data with two Gaussians and two mass error scale factors. Table 6.7 shows the divergences of the fitted parameters from the input values for this systematic effect.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	-0.014 ± 0.035	-0.38	1.1	0.56	0.045 μm	0.26 μm
$\Delta\Gamma$	-0.053 ± 0.036	-1.5	1.1	0.049	-0.0022 ps^{-1}	-0.0013 ps^{-1}
α_{CPOdd}	-0.27 ± 0.032	-8.4	0.97	0.071	-0.0045	-0.00092
$\alpha_{\parallel}$	0.099 ± 0.034	2.9	1.0	0.19	0.0014	-0.0012
$\phi_{\perp}$	-0.061 ± 0.03	-2	0.86	8.9e-09	-0.044 rad	-0.0086 rad

Table 6.7: Summary of shifts from input value for pseudo experiments with signal mass model systematically altered, fit with the default fitter.

6.2.3 Background mass model

A first order polynomial, normalised over the range $5.2 < m(B_s^0) < 5.6$ GeV is the default model for the background mass distribution, this PDF is given in equation 4.6 and explained in more detail in Section 4.3.2. A systematic uncertainty could be introduced if this model is not sufficient to describe the data. To analyse the effect of a potential mismodelling of the background mass distribution, pseudo experiments are generated with a 2nd order polynomial distribution for the background mass, and fitted with the default 1st order polynomial model. The coefficients for the alternative model experiments are taken directly from fitting the background mass with a 2nd order polynomial in data.

Table 6.8 shows the divergences of the fitted parameters from the input values for the assessment of this systematic uncertainty.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	-0.19 ± 0.035	-5.5	1.1	0.11	-1.6 μm	-1.4 μm
$\Delta\Gamma$	0.045 ± 0.035	1.3	1.1	0.89	0.00002 ps^{-1}	0.00093 ps^{-1}
α_{CPOdd}	-0.23 ± 0.033	-7	1.0	0.23	-0.0033	0.00031
$\alpha_{\parallel}$	0.18 ± 0.034	5.4	1.0	0.34	0.0031	0.00053
$\phi_{\perp}$	-0.057 ± 0.03	-1.9	0.91	0.00054	-0.039 rad	-0.0036 rad

Table 6.8: Summary of shifts from input value for pseudo experiments with background mass model systematically altered, fit with the default fitter.

6.2.4 Lifetime resolution model

A systematic uncertainty could be introduced due to inaccuracies in the parameterisation of the detector resolution for the B_s^0 lifetime. The lifetime resolution for signal events is modelled by convoluting each lifetime exponential term with two Gaussian distributions, each with a different scale factor, as in Equation 4.21, described in Section 4.4.2. To assess the level of systematic uncertainty due to this modelling of the lifetime resolution, pseudo-experiments are generated with an alternative resolution model. The alternative model uses three Gaussians with three separate scale factors, two of which correspond to those fitted in data and the third which has a value between those.

$$\begin{aligned} P'_s(ct, \sigma_{ct}|c\tau, s_{ct1,2,3}) &= P(ct|c\tau) \otimes (f_{s_{ct1}} G_1(ct, \sigma_{ct}|s_{ct1}) \\ &+ (1 - f_{s_{ct1}})\{f_{s_{ct2}} G_2(ct, \sigma_{ct}|s_{ct2}) + (1 - f_{s_{ct2}}) G_3(ct, \sigma_{ct}|s_{ct3})\} \end{aligned} \tag{6.3}$$

These alternative model pseudo experiments are fitted with the default model and the shifts from the input to fitted values are presented in Table 6.9.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	0.018 ± 0.033	0.54	1.0	0.18	0.49 μm	0.69 μm
$\Delta\Gamma$	-0.027 ± 0.034	-0.77	1.0	0.16	-0.0013 ps^{-1}	-0.00041 ps^{-1}
α_{CPOdd}	-0.21 ± 0.032	-6.6	0.98	0.98	-0.0033	0.00033
$\alpha_{\parallel}$	0.15 ± 0.033	4.5	1.0	0.7	0.0024	-0.00022
$\phi_{\perp}$	-0.017 ± 0.029	-0.6	0.87	3.7e-06	-0.015 rad	0.022 rad

Table 6.9: Summary of shifts from input value for pseudo experiments generated with lifetime resolution model systematically altered, fit with the default model.

6.2.5 Background lifetime fit model

The background lifetime model is described in Section 4.4.3, using the PDF in equation 4.23 to fit the background decay distributions with 3 exponentials and one Gaussian function. The systematic uncertainty due to this parameterisation of the background lifetime is assessed by generating pseudo-experiments with background lifetimes taken from a histogram of the lifetime distribution for the sidebands of the B_s mass peak in data and fitting them with the default fit model. This alteration tests the systematic uncertainty due to the model, because the background lifetimes are generated according to their actual distributions in data rather than the parameterisation used in the default fit, so any significant discrepancies between the model and data should show as shifts in the fitted values of the variables.

The shifts of the fitted parameters from the input values are shown in Table 6.10.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	0.23 ± 0.035	6.4	1.1	0.27	1.70 μm	2.00 μm
$\Delta\Gamma$	-0.1 ± 0.038	-2.7	1.1	0.61	-0.0045 ps^{-1}	-0.0036 ps^{-1}
α_{CPOdd}	-0.19 ± 0.034	-5.5	1.0	0.087	-0.0029	0.0007
$\alpha_{\parallel}$	0.26 ± 0.037	7.2	1.1	0.94	0.0033	0.0008
$\phi_{\perp}$	0.036 ± 0.044	0.81	1.2	0.0007	0.022 rad	0.058 rad

Table 6.10: Summary of shifts from input value for pseudo experiments with background lifetime model generated according to B_s^0 mass sideband data histograms, fit with the default fitter.

6.2.6 Angular background model and correlations

The model for the transversity angles in data is given in Equation 4.24 and explained in Section 4.4.3. Analogous to the method used to obtain the systematic uncertainty

deriving from the background lifetime parameterisation (Section 6.2.5), the background angular distribution is taken from the data sidebands to generate pseudo-experiments and then fitted with the default background angular distribution to estimate the systematic uncertainty due to the choice of background angular model. The shifts of the fitted values from the inputs for parameters in this study are shown in Table 6.11.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	-0.048 ± 0.033	-1.5	1.0	0.45	-0.20 μm	0.021 μm
$\Delta\Gamma$	0.012 ± 0.035	0.35	1.1	0.67	-0.00069 ps^{-1}	0.00021 ps^{-1}
α_{CPOdd}	-0.22 ± 0.031	-7.1	0.97	0.93	-0.0036	-0.00003
$\alpha_{\parallel}$	0.16 ± 0.033	4.8	1.0	0.62	0.0024	-0.00013
$\phi_{\perp}$	-0.027 ± 0.029	-0.9	0.86	2e-11	-0.037 rad	-0.00084 rad

Table 6.11: Summary of shifts from input value for pseudo-experiments with background angular model systematically altered, fit with the default fitter.

The model for the background transversity angles assumes the factorisability of the three angular distributions, $P(\theta)$ $P(\psi)$ and $P(\phi)$. However, some small correlations can be observed in these angles, specifically between the angles ϕ and $cos(\theta)$, shown in Figure 6.2. Also, a small effect can be seen between the background transversity angles and σ_{ct}.

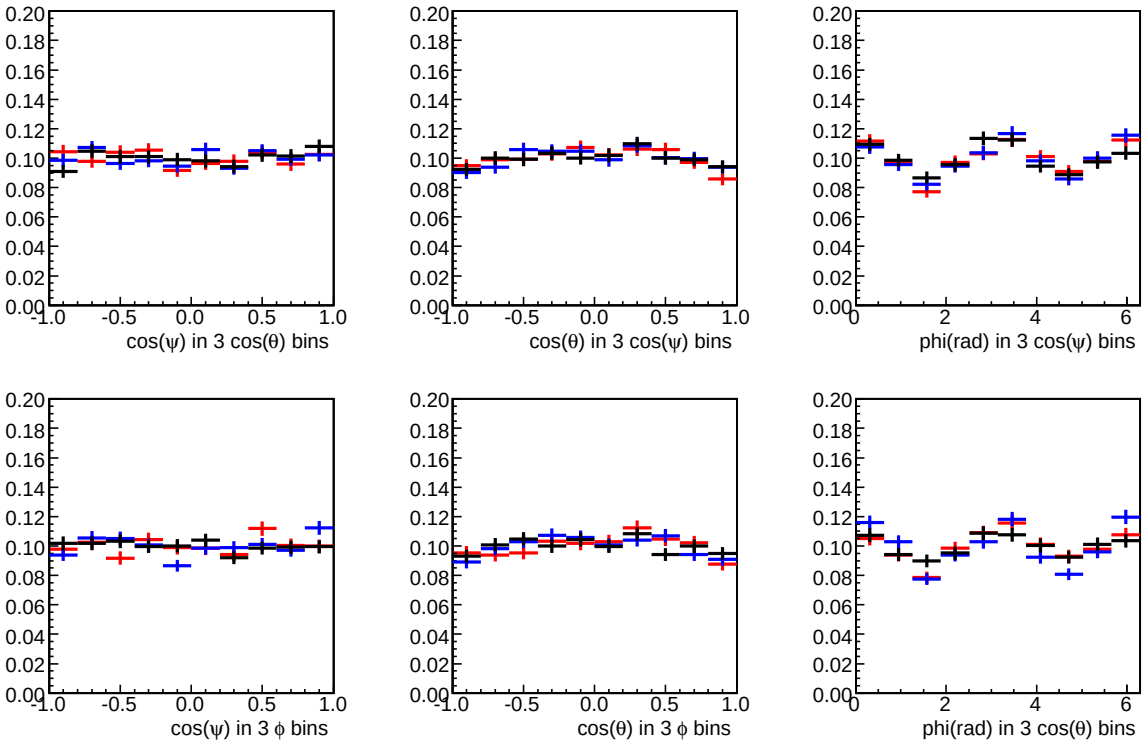

Figure 6.2: Background angular distributions (from B_s^0 sidebands) binned in each angle

The systematic effect of ignoring these small correlations in the default likelihood function is assessed. For the uncertainty due to ignoring angular correlations with $\sigma(c\tau)$,

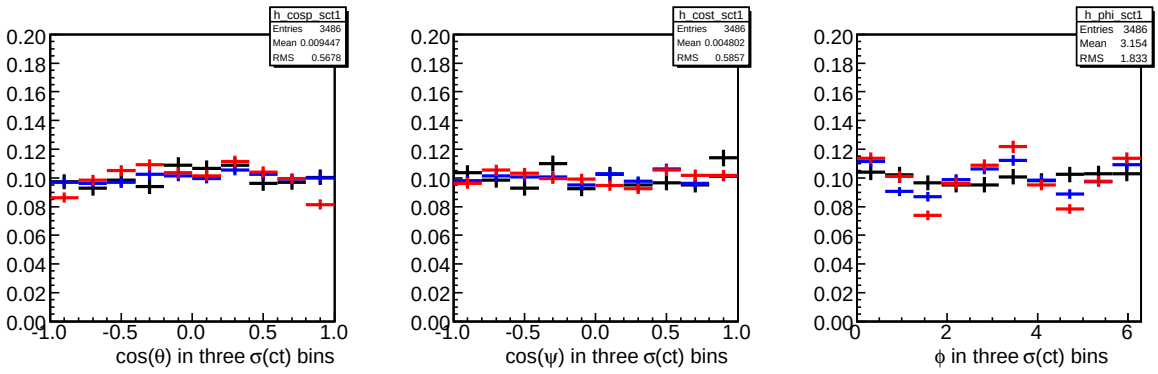

Figure 6.3: Background angular distributions (from B_s^0 sidebands) binned in $\sigma(c\tau)$

pseudo experiments are generated using histograms separated into three bins of $\sigma(c\tau)$ as input for the background angle distributions instead of the model in Equation 4.24. The shifts in the fitted values seen from this test by fitting with the default angular parameterisation can be seen in Table 6.12. To check the systematic effect from assuming the factorisation of ϕ and $cos(\theta)$, pseudo experiments are generated with $cos(\theta)$ and $cos(\psi)$ sampled from the standard background histograms described at the start of this subsection, and ϕ sampled according to the generated $cos(\theta)$ value from one of three histograms of ϕ in different $cos(\theta)$ bins. These are fitted using the default parameterisation in the fit, and the effects are shown in Table 6.13.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	-0.053 ± 0.036	-1.5	1.1	0.59	-0.36 μm	-0.14 μm
$\Delta\Gamma$	-0.012 ± 0.035	-0.34	1.1	0.95	-0.00074 ps^{-1}	0.00017 ps^{-1}
α_{CPOdd}	-0.28 ± 0.033	-8.5	0.94	0.0014	-0.0033	0.00028
$\alpha_{\parallel}$	0.14 ± 0.035	3.9	1.1	0.73	0.0017	-0.00085
$\phi_{\perp}$	-0.05 ± 0.039	-1.3	1.2	0.8	-0.030 rad	0.006 rad

Table 6.12: Summary of shifts from input value for pseudo-experiments with background angular model systematically altered to account for correlations with $\sigma(c\tau)$, fit with the default fitter.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	-0.0074 ± 0.035	-0.21	1.1	0.63	-0.16 μm	0.059 μm
$\Delta\Gamma$	-0.031 ± 0.036	-0.86	1.1	0.22	-0.0009 ps^{-1}	0.00001 ps^{-1}
α_{CPOdd}	-0.27 ± 0.033	-8.2	1.0	0.75	-0.0038	-0.00025
$\alpha_{\parallel}$	0.24 ± 0.036	6.7	1.1	0.82	0.0031	0.00052
$\phi_{\perp}$	-0.064 ± 0.041	-1.6	1.2	1.0	-0.038 rad	-0.0025 rad

Table 6.13: Summary of shifts from input value for pseudo-experiments with background angular model systematically altered to account for non-factorisable contributions to the distributions of angles ϕ and $cos(\theta)$, fit with the default fitter.

6.2.7 B_d crossfeed

It is possible for the signal $B_s^0 \to J/\psi\,\phi$ decays to be contaminated by misreconstructed $B^0 \to J/\psi\,K^{*0}$ events, where K^* decays to $K\pi$, and the π is mis-identified as a K. A small fraction of misidentified $B^0 \to J/\psi\,K^{*0}$ events, referred to as *crossfeed*, lie under the signal B_s^0 mass peak and pass the analysis selection. As this component is not included in the fit model, it could introduce a systematic uncertainty in the fit.

To find the size of this systematic uncertainty for the parameters of interest, the size of the crossfeed must first be estimated. This is done by using measured production fractions of the B_s^0 and B^0 mesons, their relative decay rates to $J\psi\phi$ and $J\psi K^*$ respectively, and the efficiency of each type of event passing the analysis selection cuts when reconstructed in the $B_s^0 \to J/\psi\,\phi$ hypothesis. The production fractions are taken from [29] and the branching fractions from [10], the efficiencies are estimated using realistic Monte Carlo, with both $B_s^0 \to J/\psi\,\phi$ and $B^0 \to J/\psi\,K^{*0}$ reconstructed as $B_s^0 \to J/\psi\,\phi$. The fraction of B^0 crossfeed events in the B_s^0 sample is calculated as:

$$f(B^0 \text{ in } \mathrm{B_s^0} \text{ sample}) = \frac{\mathrm{F}(\bar{\mathrm{b}} \to \mathrm{B}^0)\mathrm{B}(\mathrm{B}^0 \to \mathrm{J}/\psi\,\mathrm{K}^{*0})\epsilon(\mathrm{B}^0)}{\mathrm{F}(\bar{\mathrm{b}} \to \mathrm{B_s^0})\mathrm{B}(\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi)\epsilon(\mathrm{B_s^0})} \tag{6.4}$$

From Equation 6.4 the fraction of B^0 crossfeed in the signal sample used for this analysis is $1.6 \pm 0.6\%$. To make a conservative estimate of the systematic uncertainty this adds to the measurement of the parameters of interest, pseudo experiments are generated with a fraction 2.2% B^0 crossfeed (the estimated value plus 1σ) events, and fitted with the default model which has no knowledge of this component. The crossfeed component is generated by using values of the B^0 lifetime, decay width and transversity amplitudes from the CDF angular analysis of $B^0 \to J/\psi\,K^{*0}$ decays [50]. Shifts from the input values for the fitted value of each parameter for this altered model are shown in Table 6.14.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	-0.017 ± 0.035	-0.48	1.0	0.29	0.025 μm	0.24 μm
$\Delta\Gamma$	-0.034 ± 0.036	-0.94	1.1	0.039	-0.0023 ps^{-1}	-0.0014 ps^{-1}
α_{CPOdd}	-0.29 ± 0.032	-9	0.95	0.75	-0.0045	-0.00091
$\alpha_{\parallel}$	0.14 ± 0.034	4	1.0	0.2	0.0017	-0.00089
$\phi_{\perp}$	-0.04 ± 0.029	-1.4	0.86	0.061	-0.030 rad	0.006 rad

Table 6.14: Summary of shifts from input value for pseudo experiments with B^0 crossfeed included, fit with the default fitter.

6.2.8 SVX alignment

A systematic uncertainty can be introduced by the assumption that the silicon detector (described in Section 2.2.1) is perfectly aligned, when it can actually be misaligned by bowing of the detector layers of up to 50 μm [104]. Precise knowledge of the positions of all elements of the tracking system is of particular importance to the measurement of the lifetime of the B_s^0, as it is measured in terms of the proper decay length, $c\tau_s$, which is calculated from vertex positions measured in the tracking detectors. An in depth study of the effect of the limited knowledge of the CDF silicon detector

alignment was carried out in [104], which concluded that a conservative estimation of the systematic uncertainty on $c\tau$ in CDF lifetime measurements is 2 μm. This was done by fully reconstructing both data and realistic Monte Carlo under different silicon alignment assumptions, with shifts of ± 50 μm in all silicon detector components. The lifetime was fitted in several $B \to J/\psi X$ channels, and the worst shift was taken as the systematic uncertainty on the lifetime due to the assumption of perfect silicon alignment.

The value of 2 μm is taken as the systematic uncertainty on $c\tau_s$ in this analysis, and is used in the assessment of secondary effects on the other parameters of interest. Due to correlations between the lifetime and the other physics parameters, it is expected that an additional uncertainty on the lifetime measurement will also cause uncertainties in the measurement of the other parameters. To quantify the effect on the other parameters, pseudo experiments are generated with the lifetime shifted $\pm$ 2 μm and fitted as usual so that comparisons can be made between the input and fitted values of the parameters of interest. Results of this test are displayed in Table 6.15.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$\Delta\Gamma$	0.0066 $\pm$ 0.038	0.17	1.1	0.95	-0.00036 ps^{-1}	0.00055 ps^{-1}
α_{CPOdd}	-0.21 $\pm$ 0.033	-6.5	0.96	0.36	-0.0035	0.00011
$\alpha_{\parallel}$	0.14 $\pm$ 0.034	4.1	0.98	0.096	0.0026	0.0001
$\phi_{\perp}$	-0.021 $\pm$ 0.03	-0.7	0.88	0.05	-0.033 rad	0.0024 rad

Table 6.15: Summary of shifts from input value for pseudo experiments with B_s^0 lifetime shifted $\pm$ 2 μm according to the expected effect from silicon detector mis-alignment, which are fit with the default fitter assuming perfect alignment.

6.2.9 Mass error distribution

In the construction of the likelihood fit function, the distribution of errors on the mass is assumed to be the same for signal and sideband events. If these distributions are actually not consistent for signal and background, as is the case for the proper decay length errors described in Section 4.4.2, this should be taken into account in the likelihood function.The size of the systematic uncertainty due to this assumption is checked by generating pseudo experiments with mass error distributions modelled by histograms of B_s^0 sideband data for background events and sideband subtracted signal region data for signal events separately, then fitted with the default model. Shifts in the fitted values from the inputs to the pseudo experiments are displayed in Table 6.16.

6.2.10 $c\tau$ error distribution

A source of systematic uncertainty can be introduced through the modelling of $\sigma_{c\tau}$. As shown in Section 4.4.2, the proper decay length error PDFs for signal and background are parameterised separately as sets of Gamma functions given in Equation 4.22, with the coefficients taken from a fit to data. To test the effect of any mismodelling in this parameterisation, pseudo experiments are generated with $\sigma_{c\tau}$ randomly sampled from data histograms of sideband and sideband subtracted data, then fitted with the default model.

It was also observed that there was a small effect arising from $\sigma(c\tau) - mass$ cor-

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	0.0036 ± 0.046	0.078	1.1	0.12	0.26 μm	0.48 μm
$\Delta\Gamma$	0.0065 ± 0.049	0.13	1.2	0.76	-0.001 ps^{-1}	-0.00011 ps^{-1}
α_{CPOdd}	-0.26 ± 0.046	-5.7	1.1	0.89	-0.0035	6.6e-05
$\alpha_{\parallel}$	0.16 ± 0.044	3.7	1.1	0.83	0.002	-0.00053
$\phi_{\perp}$	-0.14 ± 0.049	-2.9	1.2	0.27	-0.038 rad	-0.0023 rad

Table 6.16: Summary of shifts from input value for pseudo experiments with mass error taken from histograms of sideband and sideband subtracted data histograms, fit with the default fitter which assumes a single distribution for signal and background mass errors.

relations in the latest CDF $B \to J\psi X$ lifetimes analysis [105]. Such a potential effect is accounted for in this systematic study by taking the $c\tau$ error distributions from the upper and lower sidebands separately when generating the alternative model pseudo experiments, according to where the generated mass of background events lies. The size of the systematic uncertainties caused by these alterations to the generated experiments are shown in Table 6.17.

Parameter	Mean	signi	σ	Prob	Shift	Syst
$c\tau$	-0.0092 ± 0.046	-0.2	1.1	0.77	-0.045 μm	0.17 μm
$\Delta\Gamma$	-0.075 ± 0.049	-1.5	1.2	0.56	-0.0021 ps^{-1}	-0.0012 ps^{-1}
α_{CPOdd}	-0.22 ± 0.044	-5.1	1.1	0.57	-0.003	0.00058
$\alpha_{\parallel}$	0.16 ± 0.047	3.4	1.1	0.94	0.002	-0.00056
$\phi_{\perp}$	-0.12 ± 0.051	-2.3	1.2	0.87	-0.048 rad	-0.013 rad

Table 6.17: Summary of shifts from input value for pseudo experiments with $c\tau$ error generated from histograms of sideband and sideband subtracted data histograms instead of the usual parameterisation, then fit with the default model.

6.2.11 Combined systematic errors

Table 6.18 shows the systematic errors summed in quadrature for each parameter which will be presented in this measurement. The largest effects on the lifetime measurement come from the parameterisation of the background lifetimes and the silicon detector alignment, however, all of the systematic uncertainties are significantly lower than the statistical error. The modelling of the background lifetimes also has a relatively large effect on the decay width difference, $\Delta\Gamma_s$, which is to be expected as this parameter is dependent on the measurement of the heavy and light B_s^0 lifetimes. The main systematic effect on the angular amplitudes and the strong phase $\phi_{\perp}$ is caused by the knowledge of the signal angular efficiency of the detector. The angular amplitudes shown in Table 6.18 have been converted from those used in the maximum likelihood fit ($\alpha_{\parallel}$ and α_{CPOdd}) to the initial amplitudes $A_0(0)$ and $A_{\parallel}(0)$ using the calculation described in Section 4.7.

The calculation of values for the *Pull bias* entry in Table 6.18 is explained in Section 5.3.1.

Systematic	$\Delta\Gamma(\mathrm{ps}^{-1})$	$c\tau_s$ (μm)	$\lvert A_\parallel(0)\rvert^2$	$\lvert A_0(0)\rvert^2$	$\phi_\perp$(rad)
Signal efficiency:					
Parameterisation	0.0024	0.96	0.0076	0.008	0.016
MC reweighting	0.0008	0.94	0.0129	0.0129	0.022
Signal mass model	0.0013	0.26	0.0009	0.0011	0.009
Background mass model	0.0009	1.4	0.0004	0.0005	0.004
Resolution model	0.0004	0.69	0.0002	0.0003	0.022
Background lifetime model	0.0036	2.0	0.0007	0.0011	0.058
Background angular distribution:					
Parameterisation	0.0002	0.02	0.0001	0.0001	0.001
$\sigma(c\tau)$ correlation	0.0002	0.14	0.0007	0.0007	0.006
Non-factorisation	0.0001	0.06	0.0004	0.0004	0.003
$B^0 \to J\psi K^*$ crossfeed	0.0014	0.24	0.0007	0.0010	0.006
SVX alignment	0.0006	2.0	0.0001	0.0002	0.002
Mass error	0.0001	0.58	0.0004	0.0004	0.002
$c\tau$ error	0.0012	0.17	0.0005	0.0007	0.013
Pull bias	0.0028		0.0013	0.0021	
Totals	0.01	3.6	0.015	0.015	0.07

Table 6.18: Summary of systematic uncertainties assigned to parameters of interest.

6.3 Fit projections on physics parameters

An important check of the fitter performance is carried out by projecting the distributions from the multidimensional fit onto single parameters. Projections of the fit results are shown overlaid with data for the proper decay length and angular amplitudes in Figures 6.4-6.7 for the fit using flavour tagging with $\beta_s^{J/\psi\phi}$ fixed to 0.0, and the untagged case with the same $\beta_s^{J/\psi\phi}$ condition.

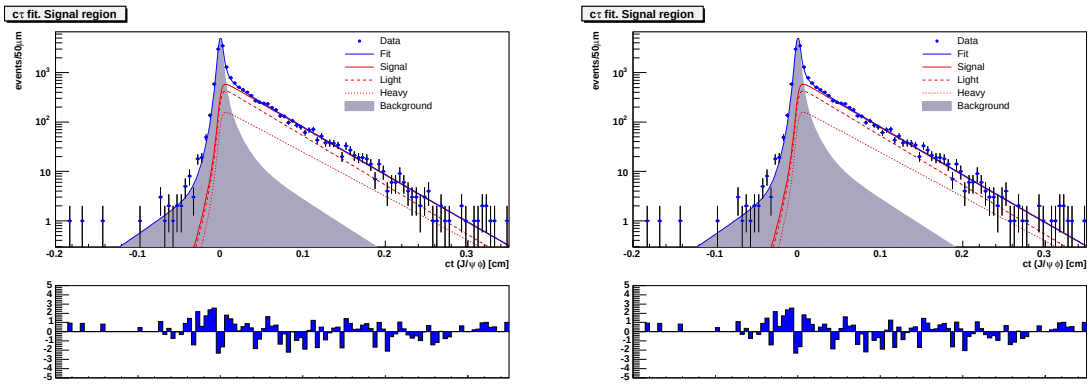

Figure 6.4: Fit projection of the B_s^0 proper decay length distribution in the signal region for the tagged fit (left) and untagged fit (right) with $\beta_s^{J/\psi\phi}$= 0.0

6.4 Results in the hypothesis of no CP violation

Values are presented here for the lifetime (proper decay length), decay width difference, angular amplitudes and strong phase measured in $B_s^0 \to J/\psi\,\phi$ decays at CDF.

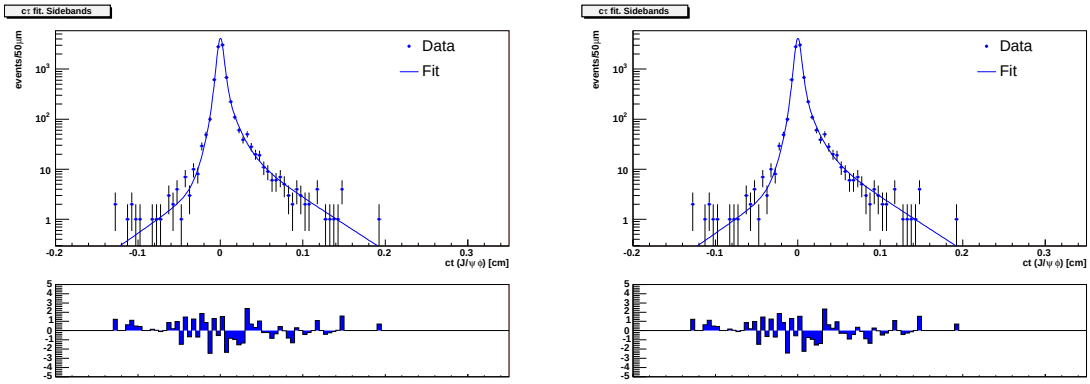

Figure 6.5: Fit projection of the B_s^0 proper decay length distribution in the sideband region for the tagged fit (left) and untagged fit (right) with $\beta_s^{J/\psi\phi}$= 0.0

$$\begin{aligned} c\tau_s &= 458.6 \pm 7.5\ (\text{stat.}) \pm 3.6\ (\text{syst.})\ \mu\text{m} \\ \Delta\Gamma &= 0.075 \pm 0.035\ (\text{stat.}) \pm 0.01\ (\text{syst.})\ \text{ps}^{-1} \\ |A_{\parallel}(0)|^2 &= 0.231 \pm 0.014\ (\text{stat}) \pm 0.015\ (\text{syst.}) \\ |A_0(0)|^2 &= 0.524 \pm 0.013\ (\text{stat}) \pm 0.015\ (\text{syst.}) \\ \phi_{\perp} &= 2.95 \pm 0.64\ (\text{stat}) \pm 0.07\ (\text{syst.})\ \text{rad} \end{aligned} \tag{6.5}$$

which are the world's most precise single determination of these parameters. They are in good agreement with an earlier iteration of this analysis at CDF [7] and for those which have been measured elsewhere, with world averages. The B_s^0 lifetime is calculated from the proper decay length as

$$\tau_s = 1.53 \pm 0.025\ (\text{stat.}) \pm 0.012\ (\text{syst.})\ \text{ps} \tag{6.6}$$

which can be compared to the world average value [10] of $\tau_s = 1.47^{+0.026}_{-0.027}$ps.

For a cross check within this dataset, the results for the fit without flavour tagging information in the Standard Model hypothesis ($\beta_s^{J/\psi\phi}$=0.0), with statistical errors only are

$$\begin{aligned} c\tau_s &= 456.9 \pm 7.7\ (\text{stat.})\ \mu\text{m} \\ \Delta\Gamma &= 0.071 \pm 0.036\ (\text{stat.})\ ps^{-1} \\ |A_{\parallel}(0)|^2 &= 0.233 \pm 0.015\ (\text{stat}) \\ |A_0(0)|^2 &= 0.521 \pm 0.013\ (\text{stat}) \end{aligned} \tag{6.7}$$

Comparing the statistical errors in 6.5 and 6.7 shows that in addition to being able to quote an additional parameter ($\phi_{\perp}$) by using the flavour tagged fit, there is a slight reduction in the statistical errors.

For completeness, and assurance that the inclusion of the S-wave KK component in the fit does not have a large effect on the measurement, SM point values are shown here from the fit without the S-wave included. This provides a clear comparison with

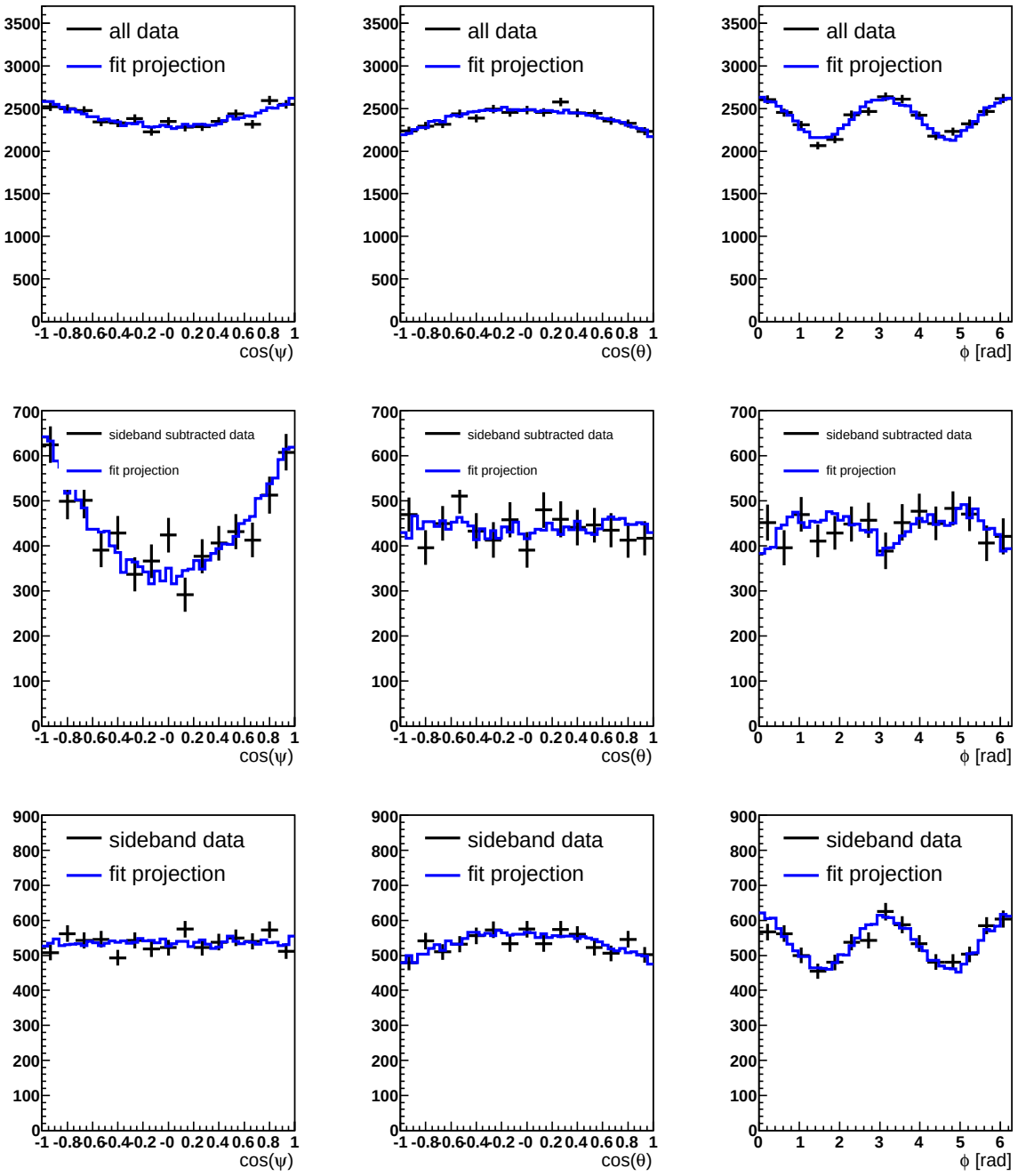

Figure 6.6: Fit projections of the transversity angles for the flavour tagged fit with $\beta_s^{J/\psi\phi}$=0.0 (top) all data (middle) sideband subtracted signal (bottom) background

earlier versions of this analysis as well as checking consistency within the results from this dataset.

For the fit with flavor tagging, the fitted values without the S-wave are in excellent agreement with the results with the S-wave KK component included in the fit:

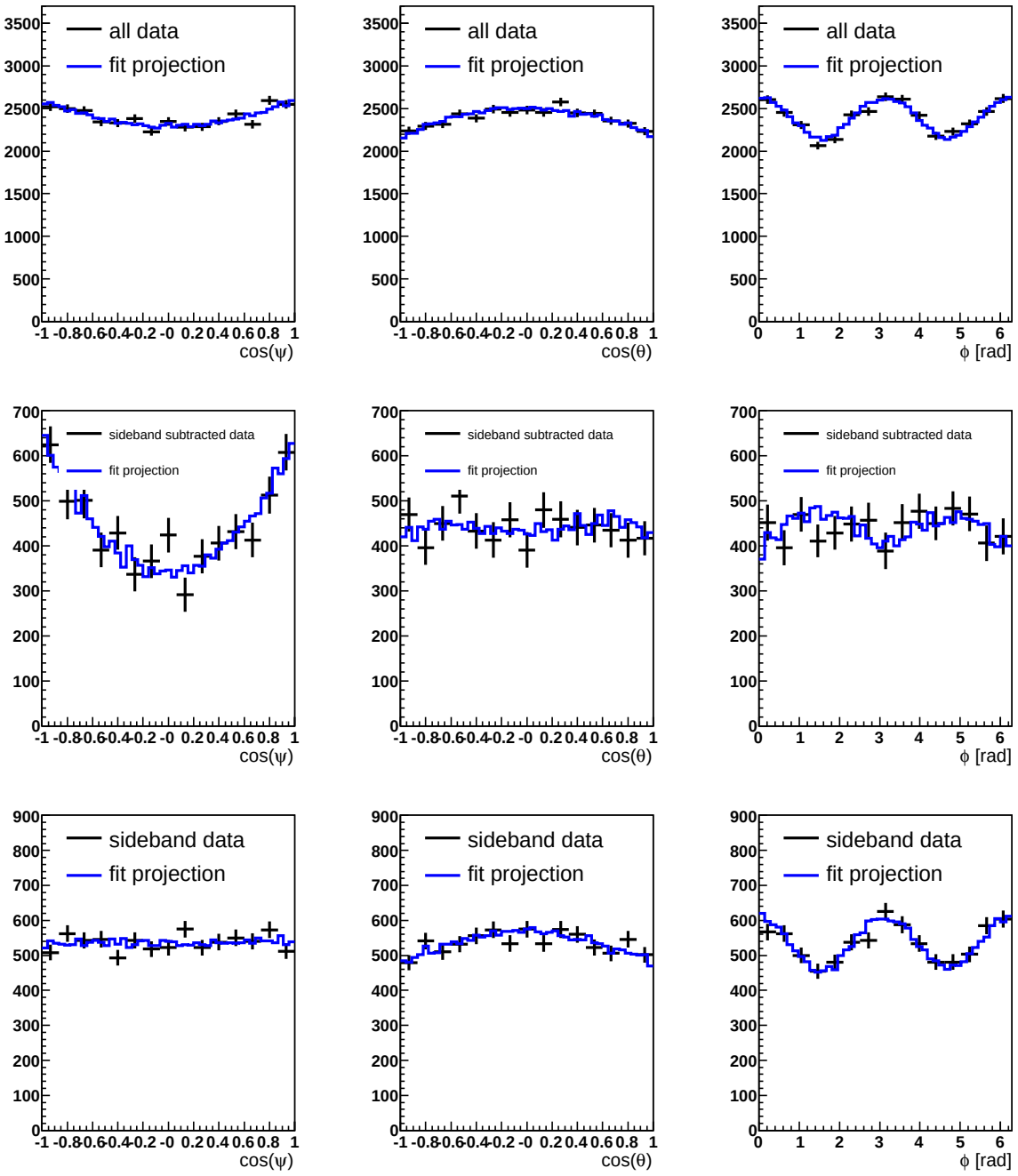

Figure 6.7: Fit projections of the transversity angles for the untagged fit with $\beta_s^{J/\psi\phi}$=0.0 (top) all data (middle) sideband subtracted signal (bottom) background

$$
\begin{aligned}
c\tau_s &= 459.1 \pm 7.7 \text{ (stat.) } \mu\text{m} \\
\Delta\Gamma &= 0.073 \pm 0.03 \text{ (stat.) ps}^{-1} \\
|A_{\parallel}(0)|^2 &= 0.232 \pm 0.014 \text{ (stat)} \\
|A_0(0)|^2 &= 0.523 \pm 0.012 \text{ (stat)} \\
\phi_{\perp} &= 2.80 \pm 0.56 \text{ (stat) rad}
\end{aligned}
\tag{6.8}
$$

The results for direct comparison with the CDF result produced for ICHEP 2008 [7], with 2.8 fb^{-1} data, are from the fit without flavour tagging and the S-wave KK component not included:

$$\begin{aligned} c\tau_s &= 457.2 \pm 7.89 \text{ (stat.) } \mu\text{m} && [459 \pm 12 \pm 3\mu\text{m}] \\ \Delta\Gamma &= 0.070 \pm 0.04 \text{ (stat.) ps}^{-1} && [0.02 \pm 0.05 \pm 0.01 \text{ ps}^{-1}] \\ |A_{\parallel}(0)|^2 &= 0.233 \pm 0.016 \text{ (stat)} && [0.241 \pm 0.019 \pm 0.01] \\ |A_0(0)|^2 &= 0.520 \pm 0.013 \text{ (stat)} && [0.508 \pm 0.024 \pm 0.01] \end{aligned} \tag{6.9}$$

where the values in square brackets are from [7].

The preceding comparisons show that the agreement between the four cases of the fit is good, as is the consistency with external experimental results.

Chapter 7

Results for CP violating parameter β_S

In this chapter the main results of this analysis will be presented. The results are given not as a measurement of a best fit value of $\beta_s^{J/\psi\phi}$ and associated uncertainties but rather as 2-dimensional confidence level contours as a function of $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$. These represent the latest results from CDF in the measurement of $\beta_s^{J/\psi\phi}$, and are the most accurate measurements to date of this parameter. The statistical technique used to create these contours is explained, and the results of several studies to validate the method. Further discussion of the S-wave $KK[f_0]$ component is given; this is the first study of $B_s^0 \to J/\psi\,\phi$ decays to incorporate this additional channel.

7.1 Fitted values with $\beta_s^{J/\psi\phi}$ floating in fit

The maximum likelihood method described in Chapter 4 was used to present fitted point values for several parameters of interest in Chapter 6, with the CP violating phase $\beta_s^{J/\psi\phi}$ fixed to zero which is approximately its standard model value [3]. In Chapter 5 it was shown that non-negligible biases are present in the full fit with $\beta_s^{J/\psi\phi}$ allowed to float, and symmetries exist in the likelihood function which prevent the fitter from selecting a single value of $\beta_s^{J/\psi\phi}$, so a different technique is required to produce a result for $\beta_s^{J/\psi\phi}$ itself. Instead of quoting a single value with errors, results for this parameter are given in the form of confidence intervals, in the 1-dimensional space of $\beta_s^{J/\psi\phi}$ and the 2-dimensional space of $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$. The 2-dimensional contours are of interest as theory predicts a relationship between $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ of [44]

$$\Delta\Gamma = 2|\Gamma_{12}^s|\cos(2\beta_s) \tag{7.1}$$

where Γ_{12}^s is the off-diagonal matrix element of the effective Hamiltonian in Equation 1.28.

The values found by the maximum likelihood fitter in the case with and without the use of flavour tagging are presented in Tables 7.1 and 7.2. Later in this section, the size of the $B_s^0 \to J/\psi KK$ component in the $B_s^0 \to J/\psi\,\phi$ signal is presented and discussed. The variables in these tables are introduced in Chapter 4. The method used for measuring $\beta_s^{J/\psi\phi}$ is described in Section 7.2, and the results in 7.3.1.

Parameter	Description	Value
A_{SW}	fraction of S-wave KK component in signal	0.018 ± 0.0234
$\alpha_\perp$	CP odd fraction	0.264 ± 0.0138
$\alpha_\parallel$	$A_\parallel$ fraction in CP even states	0.307 ± 0.0149
p_1	mass background slope	-2.3 ± 0.562
$\beta_s^{J/\psi\phi}$	CP asymmetry parameter	0.244 ± 0.132
$S_{\mathcal{D}}(SST)$	SST dilution scale factor	0.924 ± 0.0847
$S_{\mathcal{D}}(OST+)$	OST dilution scale factor +	1.12 ± 0.0919
$S_{\mathcal{D}}(OST-)$	OST dilution scale factor -	0.886 ± 0.174
δ_{SW}	Relative phase of S-wave KK component	1.5 ± 0.64
$\epsilon_b(OST)$	OST tagging efficiency for background	0.87 ± 0.002
$\epsilon_b(SST)$	SST tagging efficiency for background	0.719 ± 0.00269
$\mathcal{A}^+(OST)$	OST background positive tag asymmetry	0.495 ± 0.00319
$\mathcal{A}^+(SST)$	SST background positive tag asymmetry	0.496 ± 0.00349
$\epsilon_s(OST)$	OST tagging efficiency for signal	0.943 ± 0.00335
$\epsilon_s(SST)$	SST tagging efficiency for signal	0.522 ± 0.00681
FCN	NLL value	-5.29e+05 ± 0
f_p	Prompt fraction of background	0.884 ± 0.0052
f_-	Fraction of bkg which decays w/λ_-	0.172 ± 0.0334
f_{++}	Fraction of bkg which decays w/λ_{++}	0.661 ± 0.0453
f_s	Signal Fraction	0.181 ± 0.00236
λ_-	Effective background lifetime, neg. comp.	0.0382 ± 0.00392
λ_+	Effective background lifetime, pos. comp. 1	0.0405 ± 0.0036
λ_{++}	Effective background lifetime, pos. comp. 2	0.0112 ± 0.00117
$c\tau$	average of $c\tau_H$ and $c\tau_L$	0.0459 ± 0.000726
s_m	Mass error scale factor	1.73 ± 0.0216
m	B hadron mass [MeV/c^2]	5.37 ± 0.000131
N_{events}	Number of signal events	6.5e+03 ± 84.9
ϕ_1	First parameter in bkg fit to ϕ	0.139 ± 0.00836
$\phi_\parallel$	arg($A_\parallel A_0$) asymmetry parameter	3.02 ± 0.473
$\phi_\perp$	arg($A_\perp A_0$) asymmetry parameter	3.03 ± 0.517
$s_{c\tau}1$	Lifetime error scale factor 1	1.27 ± 0.0138
$s_{c\tau}2$	Lifetime error scale factor 2	3.32 ± 0.19
f_{sf1}	fraction of 1st lifetime error scale factor	0.882 ± 0.0119
$\cos(\psi)_1$	First parameter in bkg fit to cos(ψ)	0.004 ± 0.0201
$\cos(\theta)_1$	First parameter in bkg fit to cos(θ)	0.161 ± 0.0176
$\Delta\Gamma$	CP asymmetry parameter [ps^{-1}]	0.0973 ± 0.0351
Δm_s	B_s^0 mixing frequency	17.7 ± 0.11

Table 7.1: Fit results for the fit with flavor tagging with $\beta_s^{J/\psi\phi}$ floating

7.1.1 Measurement of the fraction of S-wave KK present in $B_s^0 \to J/\psi\,\phi$ sample

Tables 7.1 and 7.2 show that the value of the S-wave KK fraction of the $B_s^0 \to J/\psi\,\phi$ signal is estimated consistently within errors between the fit with and without flavour tagging. However, as discussed in Chapter 5 and above, biases in the maximum likelihood fitter at this level of statistics, and the proximity of the boundary at $A_{SW} = 0$ mean that it is not possible to quote a point value for the S-wave fraction using the

Parameter	Description	Value
A_{SW}	fraction of S-wave KK component in signal	0.0106 ± 0.0182
$\alpha_\perp$	CP odd fraction	0.265 ± 0.0153
$\alpha_\parallel$	$A_\parallel$ fraction in CP even states	0.31 ± 0.0162
p_1	mass background slope	-2.28 ± 0.562
$\beta_s^{J/\psi\phi}$	CP asymmetry parameter	0.575 ± 0.12
δ_{SW}	Relative phase of S-wave KK component	2.74 ± 0.649
FCN	NLL value	-5.29e+05 ± 0
f_p	Prompt fraction of background	0.885 ± 0.00516
f_-	Fraction of bkg which decays w/λ_-	0.166 ± 0.0324
f_{++}	Fraction of bkg which decays w/λ_{++}	0.649 ± 0.0475
f_s	Signal Fraction	0.181 ± 0.00239
λ_-	Effective background lifetime, neg. comp.	0.0384 ± 0.00394
λ_+	Effective background lifetime, pos. comp. 1	0.0405 ± 0.00359
λ_{++}	Effective background lifetime, pos. comp. 2	0.0113 ± 0.00123
$c\tau$	average of $c\tau_H$ and $c\tau_L$	0.0443 ± 0.00123
s_m	Mass error scale factor	1.72 ± 0.0217
m	B hadron mass [MeV/c^2]	5.37 ± 0.000131
N_{events}	Number of signal events	6.52e+03 ± 86.1
ϕ_1	First parameter in bkg fit to ϕ	0.139 ± 0.00837
$\phi_\parallel$	arg($A_\parallel A_0$) asymmetry parameter	3.41 ± 0.236
$\phi_\perp$	arg($A_\perp A_0$) asymmetry parameter	4.36 ± 0.321
$s_{c\tau}1$	Lifetime error scale factor 1	1.27 ± 0.0132
$s_{c\tau}2$	Lifetime error scale factor 2	3.35 ± 0.181
f_{sf1}	fraction of 1st lifetime error scale factor	0.884 ± 0.011
$\cos(\psi)_1$	First parameter in bkg fit to cos(ψ)	0.00452 ± 0.0201
$\cos(\theta)_1$	First parameter in bkg fit to cos(θ)	0.159 ± 0.0176
$\Delta\Gamma$	CP asymmetry parameter [ps^{-1}]	0.174 ± 0.0636

Table 7.2: Fit results for the fit without flavor tagging with $\beta_s^{J/\psi\phi}$ floating

full likelihood fit. The value of this parameter is of great interest, and as introduced in Chapter 1, there has been for some time discussion in the field ([46, 47]) of this contribution to the signal sample. In order to present a limit on the fraction of the $B_s^0 \to J/\psi\,\phi$ signal sample coming from $B_s^0 \to J/\psi KK$ a likelihood scan was carried out, in which the S-wave fraction was fixed at 100 evenly spaced points within its likely range, and all other parameters floated in the fit. The results of this scan are shown in Figure 7.1. Also of interest is the relative phase of this S-wave contribution. The likelihood scan for δ_{SW}, the phase of relative phase of the S-wave KK component, is also shown in Figure 7.1.

From Figure 7.1, it can be seen that there is not enough sensitivity to the S-wave KK relative phase to quote a likelihood interval for this parameter; the minimum value has a significance of $< 1\sigma$. For the S-wave KK fraction, the Gaussian behaviour of the errors means that confidence limits can be presented at the (1-tailed) 68 % and 95 % confidence levels. These are calculated by integrating the probability distribution in Figure 7.2. At the 68% confidence level, the upper limit on the fraction of $B_s^0 \to J/\psi KK$

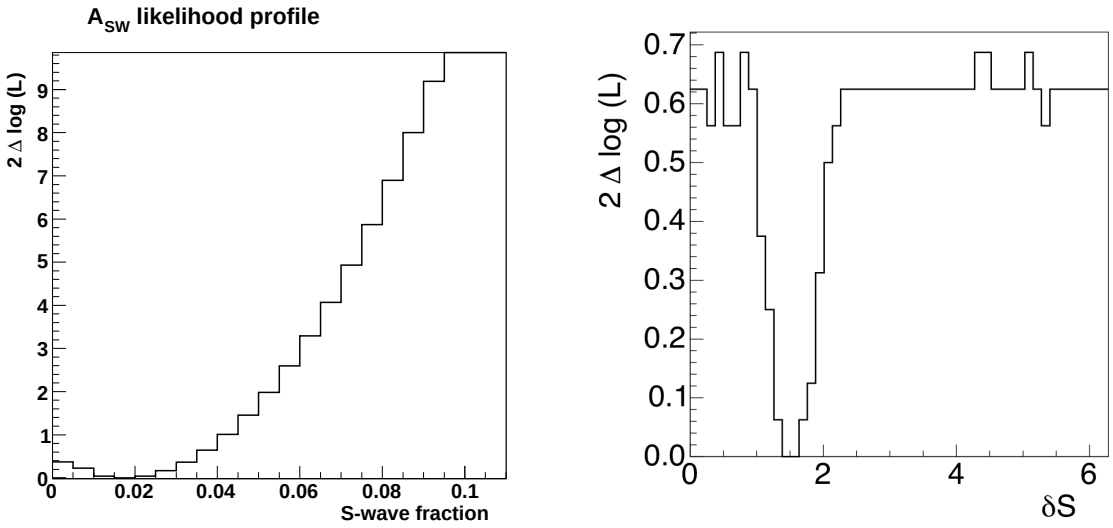

Figure 7.1: [left] Likelihood profile for the fraction of S-wave KK component in the $B_s^0 \to J/\psi\,\phi$ signal, [right] likelihood profile of the relative phase of the S-wave KK component.

in the $B_s^0 \to J/\psi\,\phi$ signal is 3.5%, and at the 95% confidence the upper limit is 6.2%. These estimates are somewhat lower than the predictions in the previously cited papers, which suggest of order 6% as a central value, but are not inconsistent.

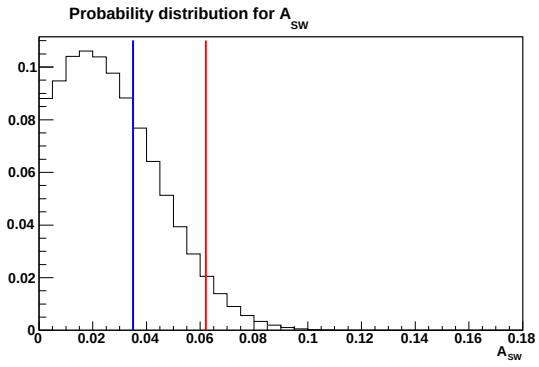

Figure 7.2: Probability distribution for the fraction of the $B_s^0 \to J/\psi\,\phi$ signal made up of $B_s^0 \to J/\psi KK$. The blue line indicates the 68% confidence limit of 3.5%, the red line the 95% limit of $A_{SW} = 6.2\%$

As a cross check of the values measured by the full likelihood fit, an alternative technique was also used to examine the potential non-resonant KK or f_0 contamination of the ϕ meson signal in $B_s^0 \to J/\psi\,\phi$. This was done by looking directly at the invariant KK mass distribution from the $\phi \to KK$ decays, and observing whether a significant additional component could be seen below the ϕ signal. The ϕ meson mass window is expanded from the narrow one used in the main fit, which gives a greater range to investigate this effect, but also allows a significantly larger $B^0 \to J/\psi\,K^{*0}$ crossfeed component than the negligible amount described in Section 6.2.7 of the previous chapter, which must be accounted for. This component is modelled using fully simulated $B^0 \to J/\psi\,K^{*0}$ Monte Carlo which has been reconstructed as $B_s^0 \to J/\psi\,\phi$. The expected non-resonant component is small, it could easily be absorbed if the combinatorial or

$B^0 \to J/\psi\, K^{*0}$ fractions were allowed to vary in a fit to the KK mass, as this flat distribution could simply boost one of those fractions. In order to avoid washing out this small component, instead of fitting the KK invariant mass where the S-wave KK part is expected to be a flat distribution, the invariant $J/\psi K^+K^-$ mass is fitted in a data sample using the enlarged ϕ mass window selection. This fit includes

- **Signal B_s^0 mass**
  Modelled by a single Gaussian distribution as in Section 4.3
- **Combinatorial background**
  Modelled by a first order polynomial distribution as in Section 4.3
- $B^0 \to J/\psi\, K^{*0}$ **crossfeed** Modelled using realistic Monte Carlo simulation of $B^0 \to J/\psi\, K^{*0}$ misreconstructed as $B_s^0 \to J/\psi\, \phi$, this effect is described in Section 6.2.7

There is no non-resonant component included in this fit so far, the idea of the test is to investigate whether there appears to be a significant component missing with only the signal, background and misreconstructed $B^0 \to J/\psi\, K^{*0}$ parts. The fractions of each component found in this fit are then projected onto the KK invariant mass distribution, where the signal ϕ mass shape is modelled using Equation 4.17. The fit to the $J/\psi K^+K^-$ invariant mass, and the projections of the fitted fractions of each component onto the KK invariant mass are shown in Figure 7.3.

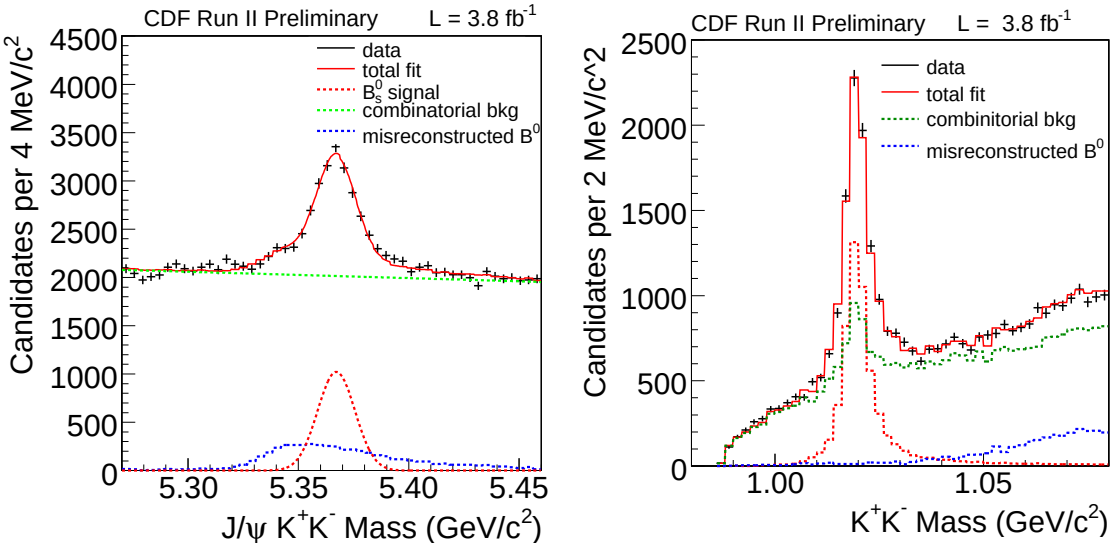

Figure 7.3: [left] Fit to invariant $J/\psi K^+K^-$ mass with enlarged ϕ meson mass window, including signal, combinatorial background and mis-reconstructed $B^0 \to J/\psi\, K^{*0}$components [right] invariant K^+K^- mass distribution, with projected components of signal, combinatorial background, and mis-reconstructed K^*

From the right hand plot in Figure 7.3, there is no clear evidence for a significant missing pedestal component, however, this does not constitute evidence against the existence of the S-wave channel in the sample. While this cross check provides no conclusive evidence for or against the existence of an S-wave KK component under the ϕ meson mass peak, it indicates no disagreement with the small central value measured in the full likelihood fit.

7.1.2 Fit projections

The projections for the fit with and without flavour tagging, with $\beta_s^{J/\psi\phi}$ allowed to vary, are shown in Figures 7.4-7.7. They demonstrate good agreement between the fitted function and the data distributions for these parameters.

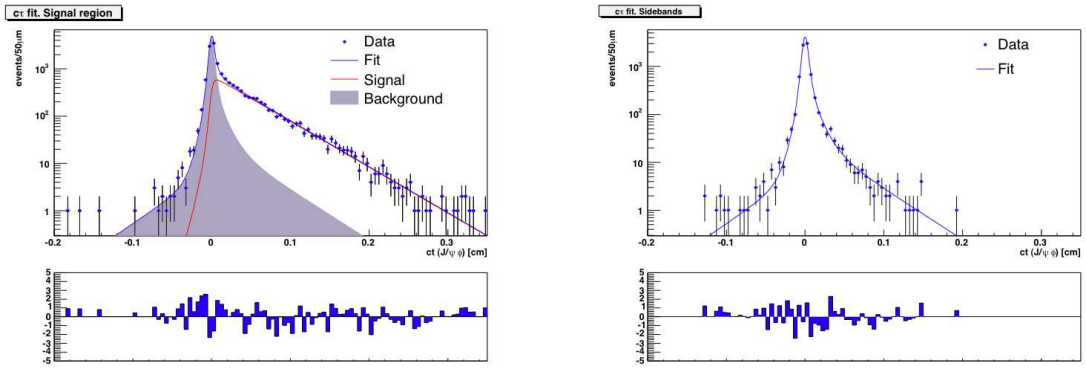

Figure 7.4: Fit projection of the B_s^0 proper decay length distribution in the signal region (left) and sideband region (right) for the flavour tagged fit

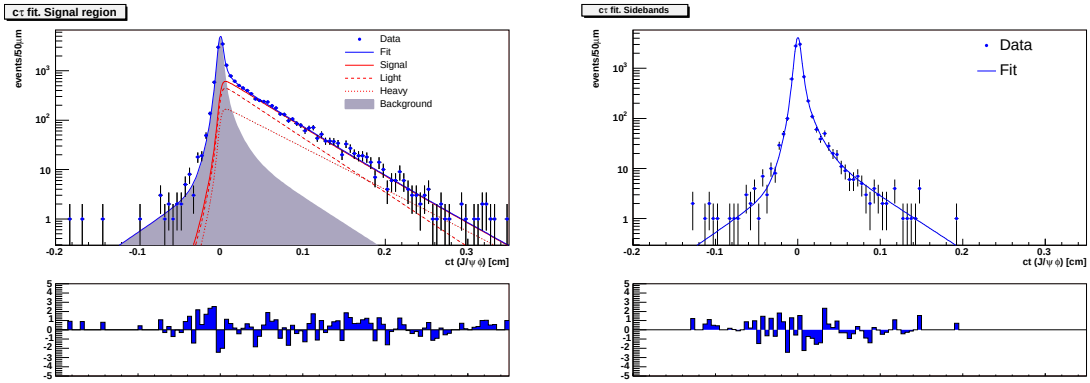

Figure 7.5: Fit projection of the B_s^0 proper decay length distribution in the signal region (left) and sideband region (right) for the untagged fit

7.2 Confidence regions for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$

As, for reasons previously discussed, the maximum likelihood fit is not sufficient to produce an estimate of the value of $\beta_s^{J/\psi\phi}$, instead the results are presented in the form of confidence regions, the construction of which is described in Section 7.2.1.

If the behaviour of a log likelihood ratio is parabolic near its minimum then N-sigma regions can be defined in terms of the area under a Gaussian function within the relevant range of the log likelihood ratio $\Delta(\mathcal{L})$ for a given N, which will be defined in Section 7.2.1. However if the behaviour is not parabolic, these ranges need to be adjusted to ensure that a claimed 1σ interval actually covers the same region as if the errors were Gaussian. This is easier to visualise in 1-dimension, but is also applicable to the 2-dimensional case. In order to ensure that the non-Gaussian behaviour of the statistical errors does not result in under-coverage of the relevant parameter space, and

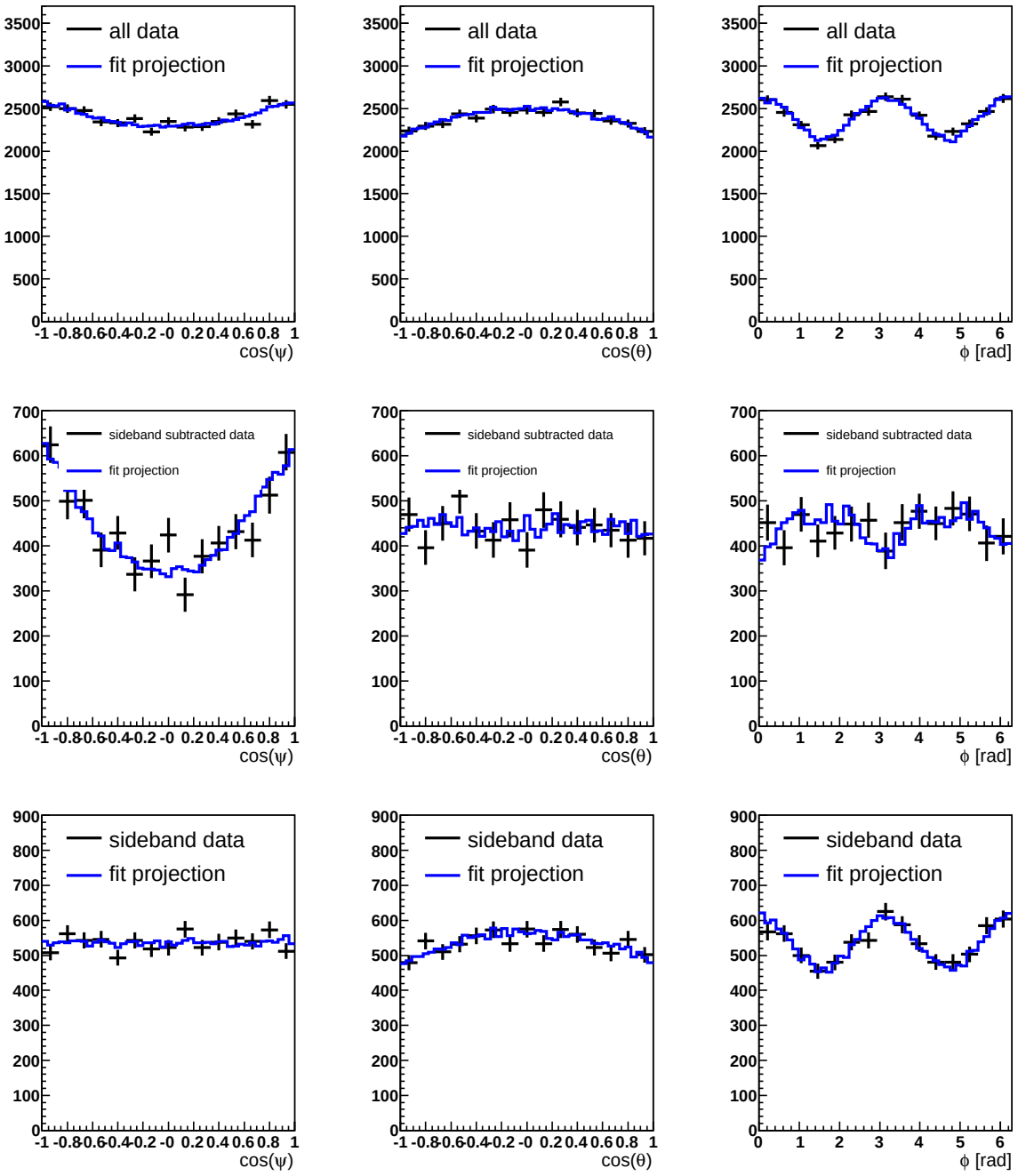

Figure 7.6: Fit projections of the transversity angles for the flavour tagged fit with $\beta_s^{J/\psi\phi}$ floating (top) all data (middle) sideband subtracted signal (bottom) background

to account for systematic effects on the statistical errors, the confidence regions are adjusted using a frequentist method described in Section 7.2.2.

7.2.1 Likelihood profiles

Construction of the likelihood profiles for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ uses a likelihood ratio method. This is carried out both in 1-dimension for $\beta_s^{J/\psi\phi}$ only, and 2-dimensions for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$. Here, the method will be described for the 2-dimensional case,

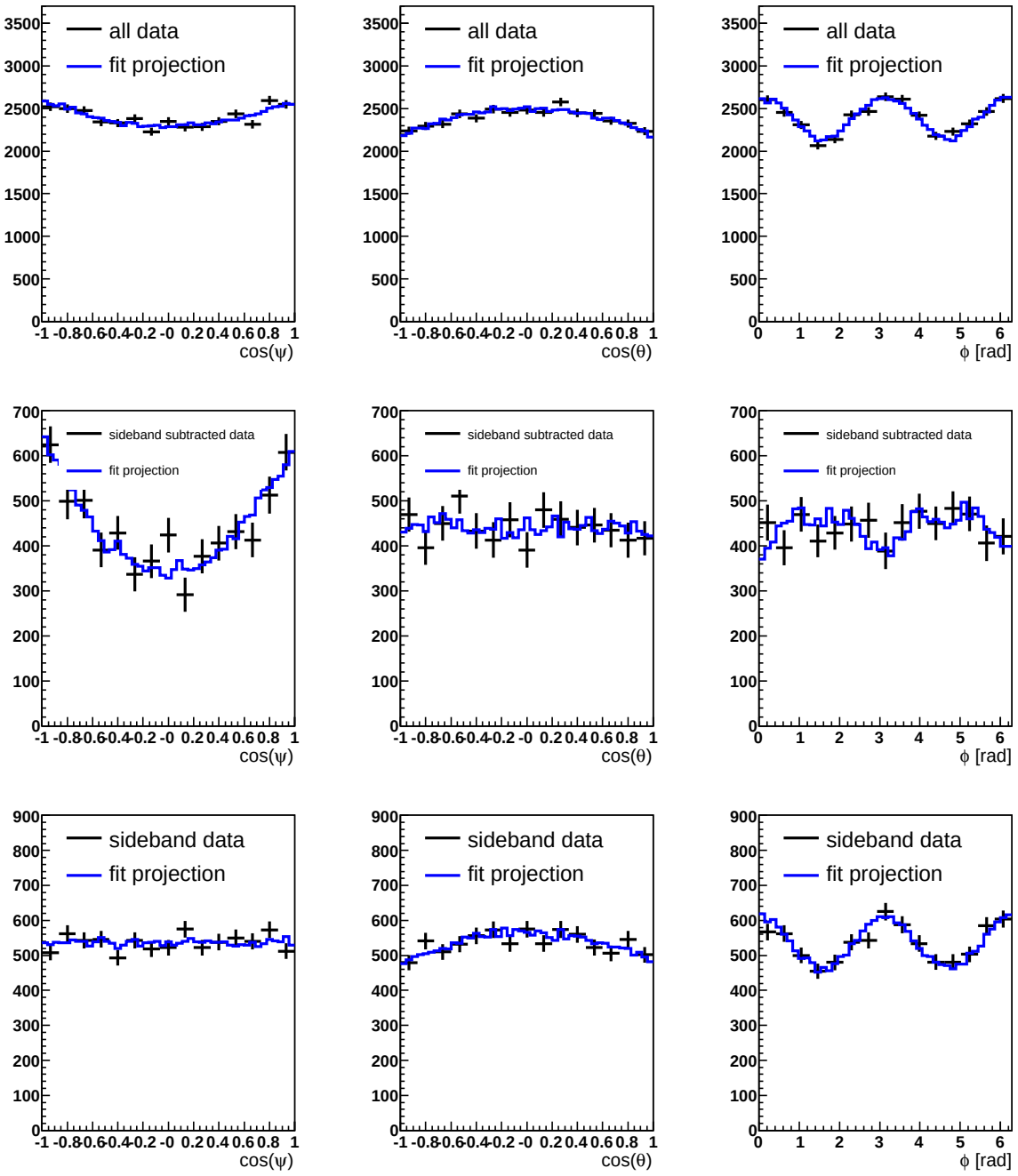

Figure 7.7: Fit projections of the transversity angles for the untagged fit with $\beta_s^{J/\psi\phi}$ floating (top) all data (middle) sideband subtracted signal (bottom) background

and where relevant the modifications to the 1-dimensional case will be clarified. Before describing the technique it is useful to define some quantities:

- μ is the set of all parameters other than those under study ($\beta_s^{J/\psi\phi}$, or $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$), also called *nuisance parameters*.
- $\mathcal{L}(\beta_s^{J/\psi\phi}, \Delta\Gamma, \mu)$ is the likelihood with all parameters floating
- $\mathcal{L}(\beta_{s,i}^{J/\psi\phi}, \Delta\Gamma_j, \mu)$ is the likelihood function with the nuisance parameters floating,

and the parameters of interest fixed to specific points i, j.

The stages to construct the likelihood profile are then as follows:

- **Initial fit with all parameters floating**
  The full maximum likelihood fit is performed (with or without flavour tagging information) to find the values of $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ which maximise the likelihood. This point in $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ space is called the *global minimum*. The likelihood value at this point is recorded and denoted as $\mathcal{L}^G$
- **Fit at each point on a 20×20 grid in $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ space**
  The fit is then performed with $\beta_s^{J/\psi\phi}$ and$\Delta\Gamma$ fixed to each point on an evenly spaced 20×20 grid, with range $-\pi/2 < \beta_s^{J/\psi\phi} < \pi/2$, $-0.7 < \Delta\Gamma < 0.7$. The likelihood at each point is recorded as $\mathcal{L}^{i,j}$
- **Construction of likelihood contours**
  Likelihood contours for different probability regions, or σ levels, can be drawn by connecting points on the grid for which $2\Delta\log\mathcal{L} \equiv 2(\mathcal{L}^G - \mathcal{L}^{i,j})$ corresponds to the relevant value for the desired confidence level.

The modification to this method for the 1-dimensional case is that instead of a grid, the fit uses 10 equally spaced points along the expected range of $\beta_s^{J/\psi\phi}$, between $-\pi/2 - \pi/2$.

The relevant $2\Delta\log\mathcal{L} \equiv 2(\mathcal{L}^G - \mathcal{L}^{i,j})$ values for each confidence level and dimensionality can be found in a look-up table such as is given in [106]. For the 2-dimensional case, the value of $2\Delta\log\mathcal{L}$ corresponding to the 1σ or 68% confidence level is 2.30, and for the 2σ or 95% level is 5.99. The values for $2\Delta\log\mathcal{L}$ corresponding to 1 and 2σ for the 1-dimensional case are 1.0 and 4.0 respectively.

In an attempt to resolve the ambiguity due to the invariance in the likelihood function under certain symmetries described in Section 4.6, the fit can be carried out separately for the strong phase $\phi_\parallel < \pi$ or $\phi_\parallel > \pi$, which removes the exact symmetry in the flavour tagged fit, and one of the symmetries for the fit without flavour tagging. However, with the limited statistics available, certain approximate symmetries remain in the likelihood. These occur because the terms on which the determination of certain parameters depend are very close to zero, limiting the power of the likelihood fitter to find an exact value. With the current dataset, the statistical limitations make it difficult to determine the sign of $\Delta\Gamma$, and the quadrant in which the phases $\beta_s^{J/\psi\phi}$, $\phi_\perp$ and $\phi_\parallel$ lie. This results in a local minimum, in addition to the global minimum found by the NLL minimisation process; the likelihood profiles need to be symmetrised to account for this effect.

Without the S-wave KK component, this symmetrisation can be done by simply reflecting the contours about the symmetry axis, but as was previously described, the inclusion of this component breaks the exact symmetry in $\{\beta_s, \Delta\Gamma, \phi_\perp, \phi_\parallel\} \Rightarrow \{\pi/2 - \beta_s, -\Delta\Gamma, \pi - \phi_\perp, 2\pi - \phi_\parallel\}$. If the addition of the S-wave KK in the likelihood changed the depth of the minima (with $\phi_\parallel < \pi$ or $\phi_\parallel > \pi$), this would mean that the ambiguity in $\beta_s^{J/\psi\phi}$ was resolved, and a single value could be selected, but unfortunately, due to the small size of the measured S-wave KK component this is not the case. The difference in $2\Delta\log\mathcal{L}$ for the two $\phi_\parallel$ options is < 0.05 units, which is not significant.

Thus, the symmetrisation is carried out by constructing the contours as described here, for both $\phi_{\parallel} < \pi$ and $\phi_{\parallel} > \pi$, then producing a single set of likelihood contours by taking the deeper minima of the two at each point on the grid.

This method constitutes a part of the full technique used to produce the final likelihood contours for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$. The final component to complete the statistical integrity of the contours is highly computationally intensive, so it is convenient to carry out some initial studies using only this part of the technique before applying the full coverage adjustment to produce statistically sound final results.

Unadjusted contours for full dataset

First, it is interesting to look at the likelihood profiles for the full dataset before any coverage adjustment. In Figure 7.8 the 2-dimensional profiles are shown from the fit with and without flavour tagging. These demonstrate clearly the symmetries described in section 4.6, and the reduction from the four fold ambiguity seen in the untagged case to the two fold ambiguity by tagging the flavour of the decaying B_s^0 meson is clearly visible. This effect is also seen in the 1-dimensional profile of $\beta_s^{J/\psi\phi}$, in Figure 7.9

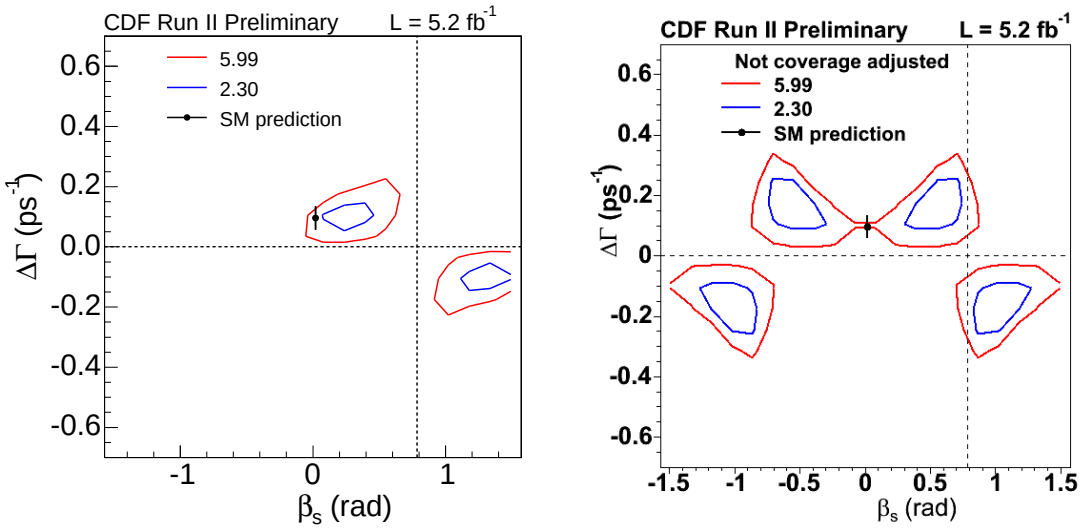

Figure 7.8: Unadjusted 2-dimensional likelihood profiles for $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ in full dataset with (left) flavour tagged fit and (right) fit without tagging information.

Comparison of fit with and without flavour tagging in pseudo experiments

To check the behaviour of the likelihood fitter, it is useful to compare its result in different datasets. In order to do this, pseudo experiments can be generated, with equal statistics to the real dataset, and input values for parameter generation taken from the fit to data, in this case shown in Table 7.1. This gives multiple simulated experiments on which to cross check the fitter behaviour, which have the same underlying *true* values of the fitted parameters. From these experiments it is possible to see whether the fitter performs consistently in the configuration with and without flavour tagging.

In this test, 5 pseudo experiments were generated, then the 2-dimensional likelihood contours in $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ space were generated as described earlier in this section. One simplification was made for the study; the symmetrisation of the contours was carried

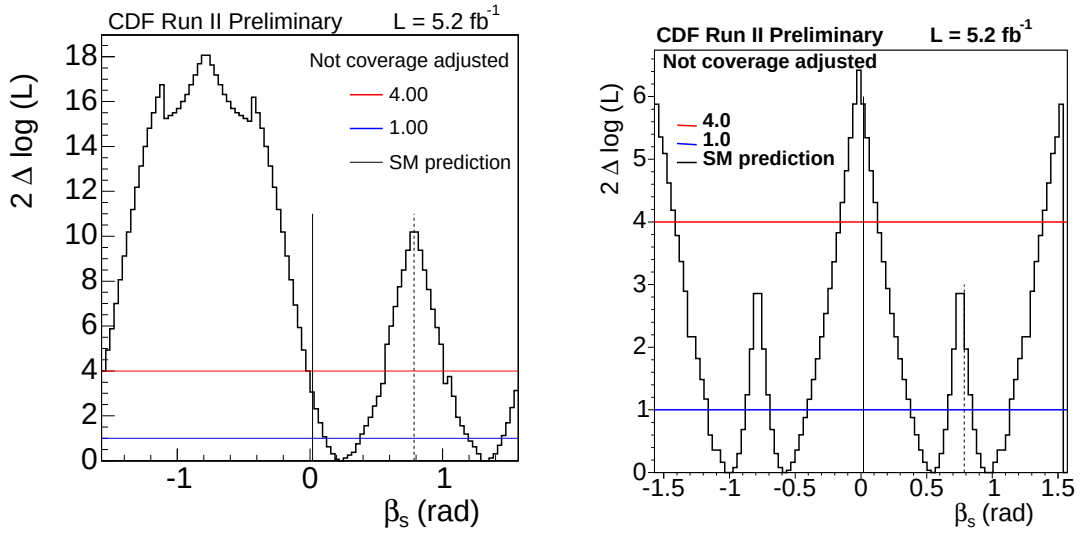

Figure 7.9: Unadjusted likelihood profiles for $\beta_s^{J/\psi\phi}$ in full dataset with (left) flavour tagged fit and (right) fit without tagging information.

out by reflecting about the symmetry axis instead of the full technique discussed above. This approximation is justified for use in this cross check by the observed minimal effect of adding in the S-wave KK component on the breaking of the symmetry in the likelihood function.

These likelihood contours are shown in Figure 7.10. Aside from the additional minima in the fit where tagging of the initial B_s^0 meson flavour is not used, the contours around the minima which coincide with the flavour tagged fit are larger - showing less sensitivity to the value of $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$, as would be expected for a fit with less information.

Comparison of likelihood contours in different parts of the dataset

Perhaps a more significant comparison, is the behaviour of the fit with and without flavour tagging on different samples of real data. This can only be done by dividing up the complete dataset, so the reduction in statistics must be accounted for when looking at these contours next those from the full sample.

The data was divided into three time periods of approximately equal luminosity, which correspond to each of the updates of the $B_s^0 \to J/\psi\,\phi$ analysis. These three subsets were fitted with the flavour tagged and untagged fitters to produce likelihood contours. For ease of comparison with earlier CDF results, these were fitted with the S-wave KK fixed to 0.0 which is consistent with the previous analyses that did not account for the S-wave KK component.

These likelihood contours can be seen in Figure 7.11. At the first instance, the rows in Figure 7.11 appear somewhat varied, for samples which are subsets of the same data. However, they are consistent within statistical errors, and a similar level of variation can be seen in the pseudo experiments in Figure 7.10 so this does not indicate any inconsistency between the data from different periods of time. Another point which can be taken from this study is the agreement between the fit with and without flavour

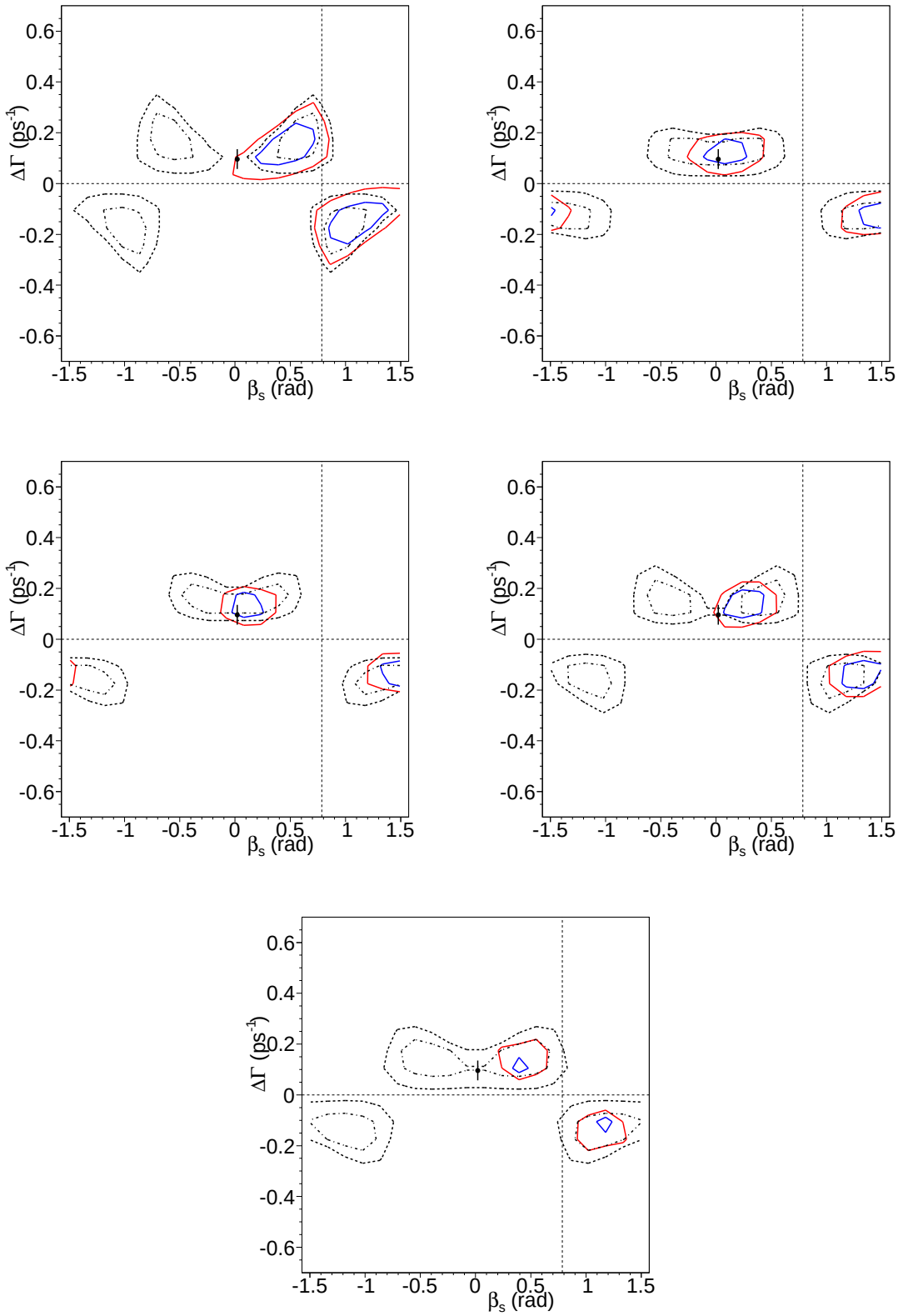

Figure 7.10: Likelihood contours plotted for 5 pseudo experiments generated with input values taken from Table 7.1 and fitted using tagged (solid lines) and untagged (dashed lines) fit configurations.

tagging included, for the different data periods. As in the previous test, the positive $\beta_s^{J/\psi\phi}$ side of the untagged contours can be compared to those with flavour tagging information, and these show similar behaviour for the different parts of the dataset.

7.2.2 Coverage adjustment

The likelihood profile technique described so far does not account for the non-Gaussian distribution of the statistical errors, which can be seen from the shape of the unadjusted contours, or systematic effects on the errors. In order to properly include these effects, which will ensure that the final confidence regions do in fact cover the area in $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ space corresponding to the claimed confidence level, a *Likelihood Ratio Ordering* technique [107] is used.

Firstly considering the coverage adjustment to account for non-Gaussian statistical errors, 1000 pseudo experiments are generated at the Standard Model expected point on the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ grid described in Section 7.2.1 (or at the SM expected value of $\beta_s^{J/\psi\phi}$=0.02 for the 1-dimensional case). These are each fitted twice, once with all parameters floating ($\mathcal{L}(\beta_s^{J/\psi\phi}, \Delta\Gamma, \vec{\mu})$) and once with $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ fixed to their generated values ($\mathcal{L}(\beta_{sSM}^{J/\psi\phi}, \Delta\Gamma_{SM}, \vec{\mu})$). From these, a likelihood ratio can be constructed:

$$\mathcal{R} = 2\log\frac{\mathcal{L}(\beta_s^{J/\psi\phi}, \Delta\Gamma, \vec{\mu})}{\mathcal{L}(\beta_{sSM}^{J/\psi\phi}, \Delta\Gamma_{SM}, \vec{\mu})}. \tag{7.2}$$

Likelihood ratios for all of the pseudo experiments are stored in a histogram, which is normalised and integrated, to give a probability distribution of (1-Confidence Level), where the confidence level (CL) corresponds to the probability that the true value lies within the interval defined for that level. These distributions are shown as the solid black lines in Figure 7.12 for the 1- and 2-dimensional flavour tagged and untagged versions of the fitter.

Studies carried out for the CDF published result for $\beta_s^{J/\psi\phi}$ [4], which are documented in [101] showed that calculating the coverage adjustment at this single point is as conservative as using pseudo experiments generated at every point on the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ grid independently, so the single point approach is chosen for economy of computational time.

The normalised (1-CL) distributions for the pseudo experiments are used to adjust the likelihood profiles described in the previous section, by selecting the appropriate values for $2\Delta\log\mathcal{L}$ to give coverage of 68% and 95% confidence regions. The canonical values, of 2.30 (1σ level) and 5.99 (2σ level) for the 2-dimensional case are only valid for the situation where the errors on the parameters of interest exhibit perfectly Gaussian behaviour. These are the points on the y-axis of the distributions in Figure 7.12 where the Gaussian distribution (shown in green) has a (1-CL) value corresponding to (1-0.95) and (1-0.68). To adjust to the non-Gaussian case, instead of the canonical values the points on the y-axis where the distribution from the pseudo experiments intersects with the relevant (1-CL) line are used. Histogram binning effects are avoided by interpolating from one bin edge to the next at the point where the distribution meets each (1-CL) value.

For the 2-dimensional $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ confidence regions from the fit without flavour tagging information, the adjusted $2\Delta\log\mathcal{L}$ values are

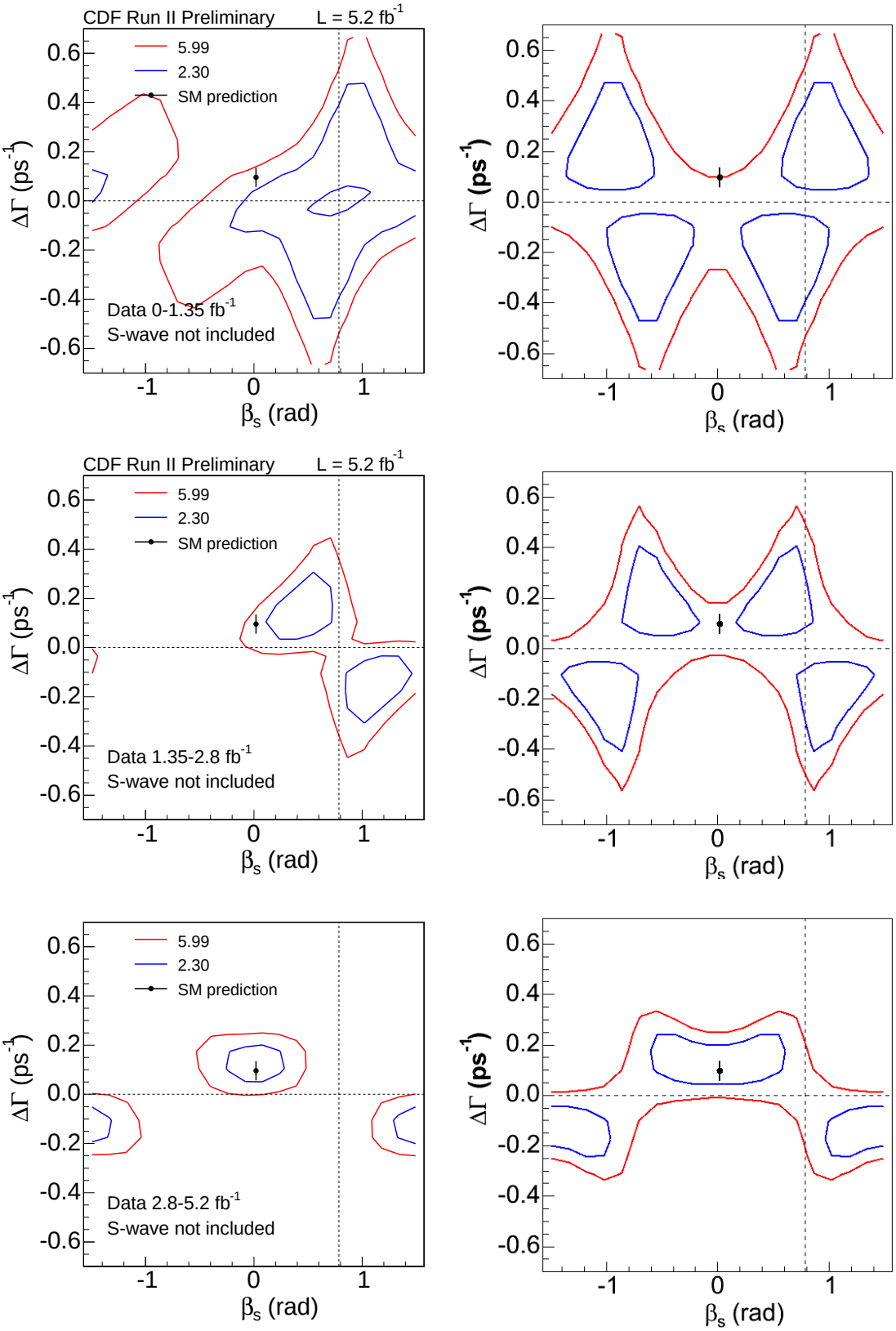

Figure 7.11: Likelihood contours for (left) fit with flavour tagging (right) fit with no flavour tagging information in dataset divided into 3 time periods, integrated luminosity L = 0-1.35 fb^{-1}, L = 1.35-2.8 fb^{-1} and L = 2.8-5.2 fb^{-1}.

- 68% CL (1σ confidence region): $2\Delta \log \mathcal{L} = 3.50$
- 95% CL (2σ confidence region): $2\Delta \log \mathcal{L} = 7.96$

For the 2-dimensional $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ confidence regions where flavour tagging of the initial B_s^0 state is included, the adjusted values are

- 68% CL (1σ confidence region): $2\Delta \log \mathcal{L} = 2.85$
- 95% CL (2σ confidence region): $2\Delta \log \mathcal{L} = 7.34$

For the 1-dimensional $\beta_s^{J/\psi\phi}$ confidence interval where flavour tagging is used, the adjusted values are

- 68% CL (1σ confidence region): $2\Delta \log \mathcal{L} = 1.68$
- 95% CL (2σ confidence region): $2\Delta \log \mathcal{L} = 5.37$

The values of these levels for the Gaussian case were given in Section 7.2.1

Using these adjusted values, the likelihood regions for each case described above can be presented, ignoring systematic effects. Figure 7.13 show the tagged and untagged fit 2-dimensional contours, and 1-dimensional $\beta_s^{J/\psi\phi}$ scan from the tagged fit after this initial stage of adjustment.

Systematic variations

In addition to correcting the likelihood contours due to the non-Gaussian error distributions, effects of systematic variations in the set of nuisance parameters are accounted for using an extension of the technique described so far in this section.

To observe the effects of potential systematic variations in all parameters other than $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ (or $\beta_s^{J/\psi\phi}$ only, in the 1-dimensional case), pseudo experiments are generated in 16 *alternate universes*. The concept of these *universes* is that each represents a physical situation with a different set of measurements for the nuisance parameters; the pseudo experiments used for the non-Gaussian error adjustment can be considered as being in *our universe* as they have values of nuisance parameters taken directly from the measured values in data. For each *alternate universe*, 1000 pseudo experiments are generated, where the input values for the generation of the nuisance parameters have been randomly varied within $\pm 5\sigma$ of their measured value in data. Each experiment within one universe has the same set of input values, however the randomised values are different for every universe. These randomised input values are given for each universe in Appendix G.

The construction of likelihood ratios, and from those, distributions of (1-CL) values for the pseudo experiments is carried out as described earlier in this section for the adjustment to non-Gaussian errors. In Figure 7.12, the dashed lines each represent a different universe. It can be seen from this figure that nearly all of the *alternate universes* lie further from the Gaussian error regime than *our universe*, which for the flavour tagged fit in particular is very close to Gaussian. This is thought to be an effect of the randomisation for inputs with asymmetric limits, namely the S-wave KK fraction,

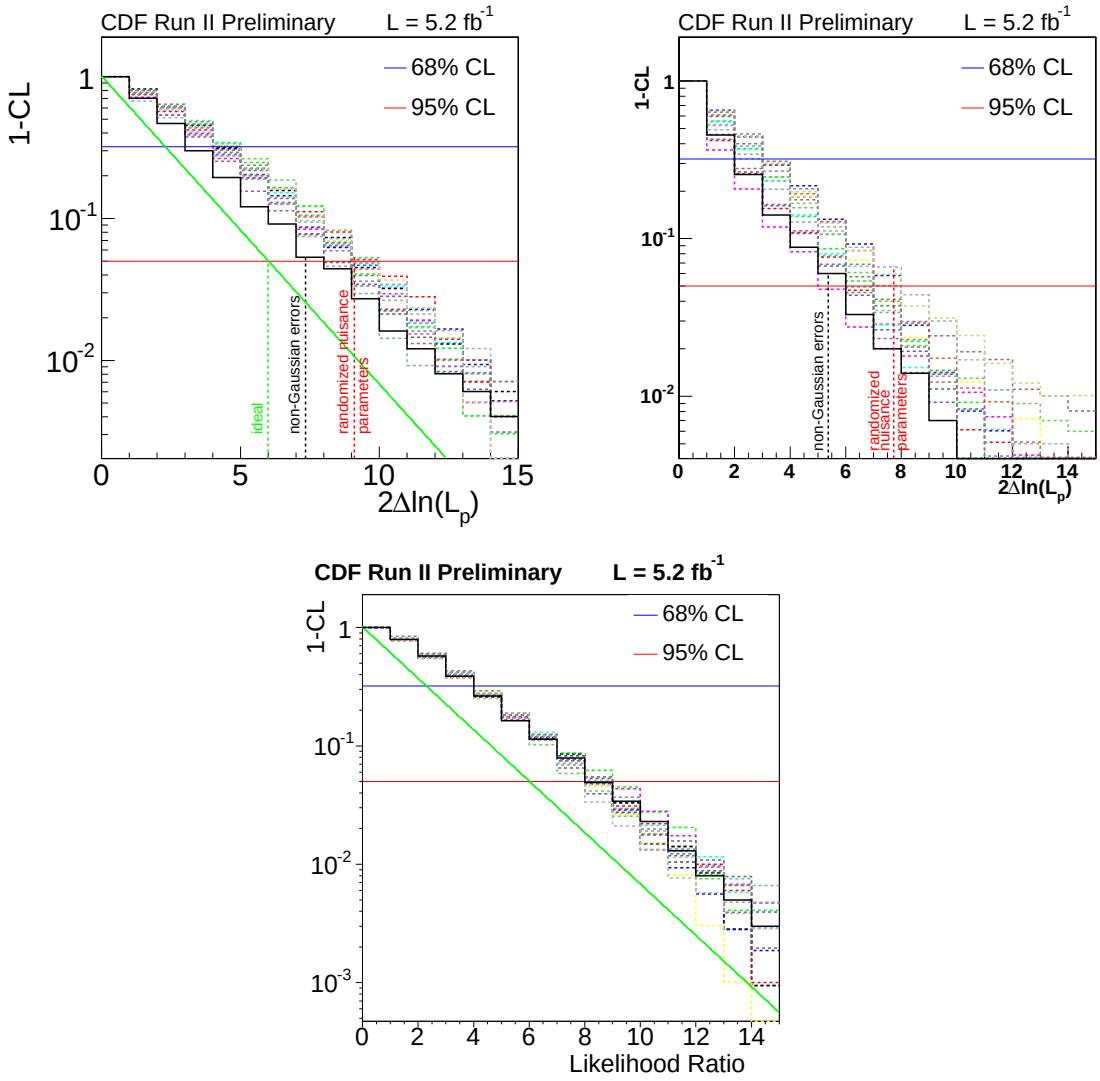

Figure 7.12: Distributions of (1-CL) for pseudo experiments used to check the likelihood coverage of $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ confidence regions. [top left] fit using flavour tagging for 2-dimensional,[top right] fit using flavour tagging for 1-dimensional contours, [bottom right] fit without flavour tagging information for 2-dimensional contours.$\beta_s^{J/\psi\phi}$

A_{SW}. When selecting an input value to generate within $\pm$ 5σ of the central value of A_{SW}, the cut off at 0.0 must be observed as this parameter cannot take a negative value. As 0.0 is less than 1σ below the central value, the majority of parameter space available is above the level found in *our universe*, so the *alternate universes* are likely to have a larger A_{SW} which can push the errors further from the Gaussian regime. The values of A_{SW} for the *alternate universes* can be seen in Appendix G.

The $2\Delta \log \mathcal{L}$ values for this adjustment are taken from the (1-CL) distribution of the universe which lies furthest from the Gaussian regime at each confidence level. These values are, for the 2-dimensional $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ confidence regions fitted without flavour tagging

- 68% CL (1σ confidence region): $2\Delta \log \mathcal{L} = 3.77$

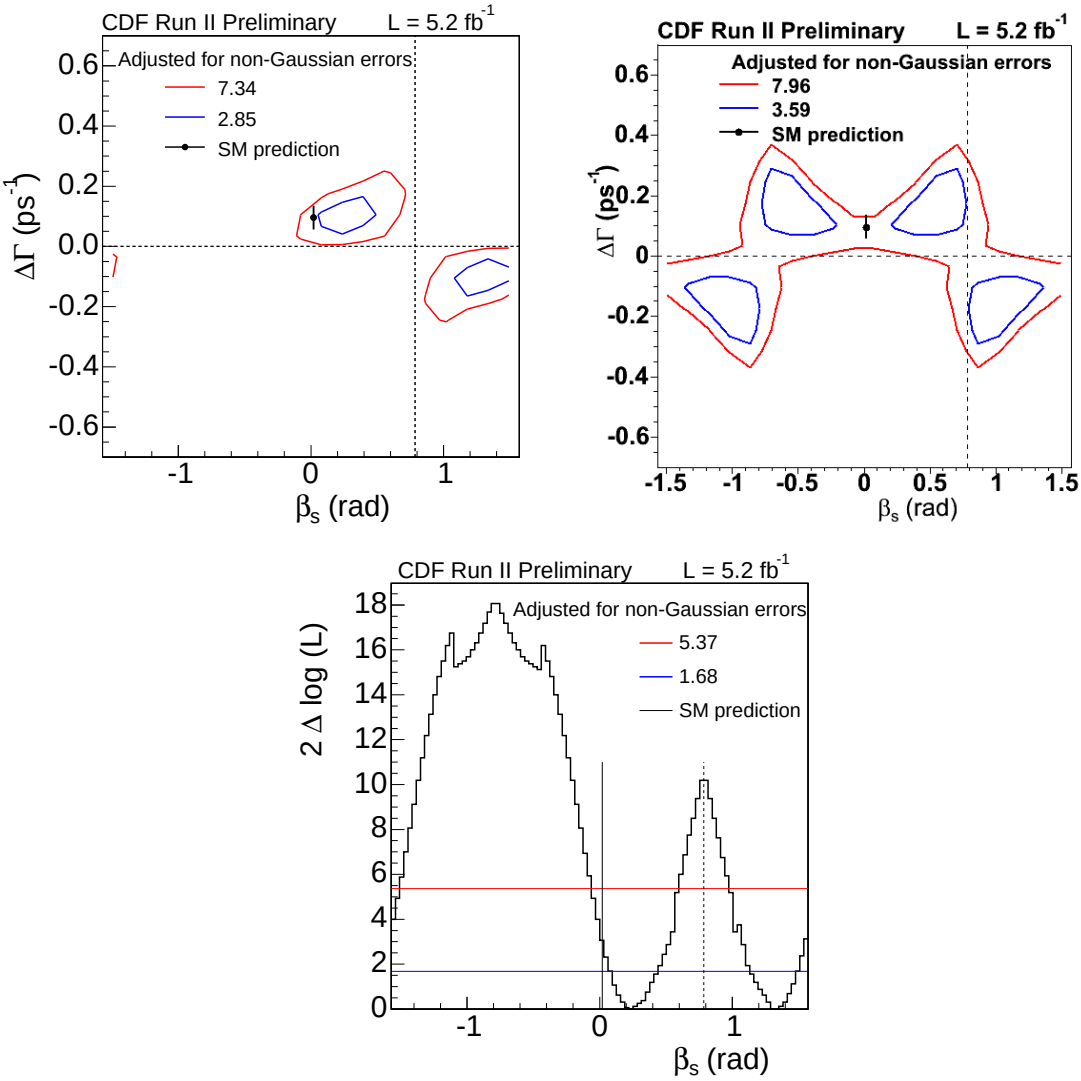

Figure 7.13: Confidence regions for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ after adjustment to account for non-Gaussian statistical errors [top left] fit using flavour tagging, 2-dimensional contours,[top right] fit using flavour tagging, 1-dimensional contours, [bottom right] fit without flavour tagging information, 2-dimensional contours.$\beta_s^{J/\psi\phi}$.

- 95% CL (2σ confidence region): $2\Delta \log \mathcal{L} = 8.67$

For the 2-dimensional $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ confidence regions where flavour tagging of the initial B_s^0 state is included, the adjusted values are

- 68% CL (1σ confidence region): $2\Delta \log \mathcal{L} = 4.27$
- 95% CL (2σ confidence region): $2\Delta \log \mathcal{L} = 9.10$

For the 1-dimensional $\beta_s^{J/\psi\phi}$ confidence interval where flavour tagging is used, the adjusted values are

- 68% CL (1σ confidence region): $2\Delta \log \mathcal{L} = 2.93$

- 95% CL (2σ confidence region): $2\Delta \log \mathcal{L} = 7.73$

These are the numbers used to set the levels of the final confidence regions, which are presented in Section 7.3.1. They can be compared to those calculated to adjust for the non-Gaussian error distributions, and as expected, the adjustment is larger when systematic variations in the nuisance parameters are taken into account. That the shift between these two adjustments is somewhat larger for the fit with flavour tagging included, is expected because there including flavour tagging information adds an extra 12 nuisance parameters into the fit, all of which are varied within $\pm 5\sigma$ for the systematic coverage adjustment.

7.3 Final measurement of $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$

7.3.1 Confidence regions for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$

The final confidence regions with full coverage adjustment are shown from the fit with flavour tagging of the initial B_s^0 state in Figure 7.14, and without flavour tagging in Figure 7.15. These are the first measurements of $\beta_s^{J/\psi\phi}$ to include the $B_s^0 \to J/\psi KK$ component in the fit function, and make use of the largest single data sample for this analysis to date. The Standard Model point on the following plots is marked at $\beta_s^{J/\psi\phi} = 0.02$ [3], with $\Delta\Gamma = 0.096 \pm 0.039$ [44].

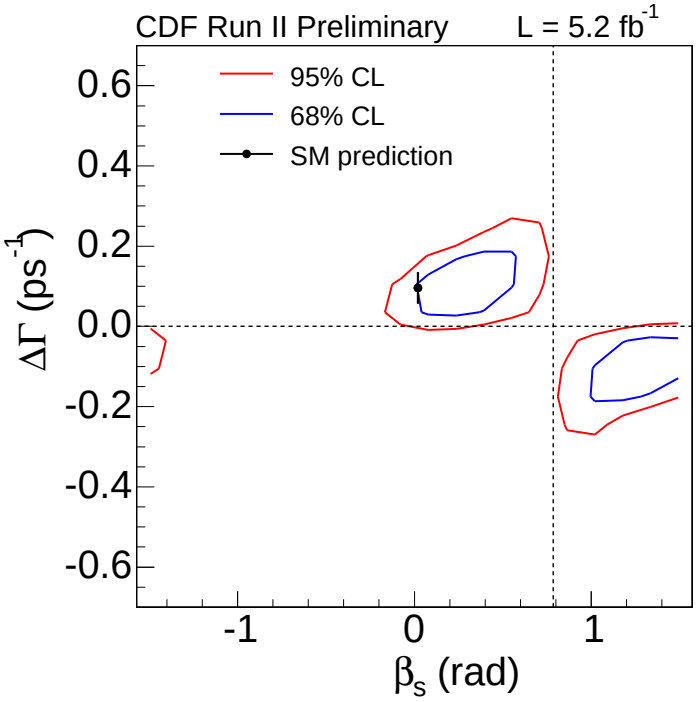

Figure 7.14: 2-dimensional confidence regions in $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ space, with flavour tagging of the initial B_s^0 state. Likelihood regions adjusted to include systematic effects.

The fully adjusted likelihood profile in 1-dimension for $\beta_s^{J/\psi\phi}$ is shown in Figure 7.16, this scan is from the fit including flavour tagging. From this, likelihood intervals for

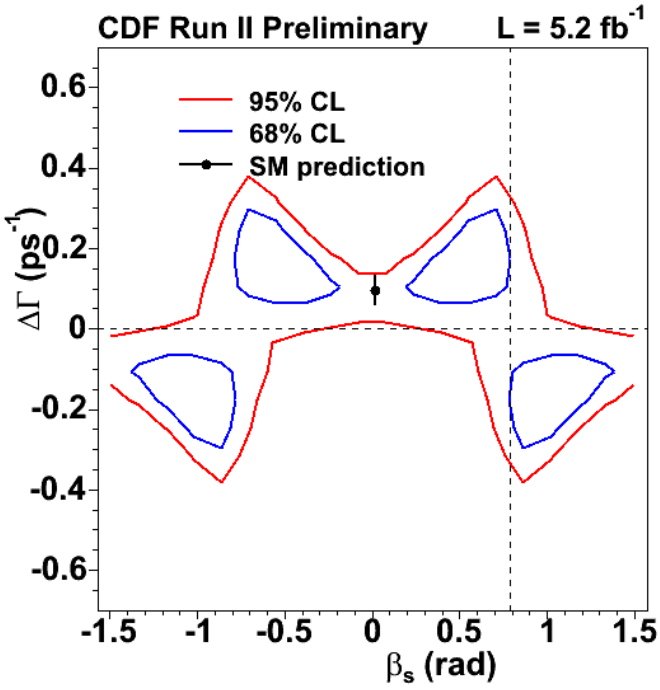

Figure 7.15: 2-dimensional confidence regions in $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ space, without flavour tagging of the initial B_s^0 state, to cross-check the effect of flavour tagging on the measurement. Likelihood regions adjusted to include systematic effects.

$\beta_s^{J/\psi\phi}$ can be presented. At the 68% confidence level, $\beta_s^{J/\psi\phi}$ is within

$$[0.02, 0.52] \cup [1.08, 1.55], \tag{7.3}$$

and at the 95% level

$$[-\pi/2, -1.44] \cup [-0.13, 0.68] \cup [0.89, \pi/2]. \tag{7.4}$$

7.3.2 Standard Model p-values

It is valuable to estimate the probability that the measured values for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ from this analysis are in fact fluctuations from true values which are the Standard Model expectation of $\beta_s^{J/\psi\phi}$=0.02, $\Delta\Gamma$=0.096 [3]. To do this, p-values are calculated by comparing likelihood ratios for the global minimum and Standard Model point in the data sample, and in pseudo experiments generated with $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ at their Standard Model expected values.

The p-value for the Standard Model values of $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ in the dataset is defined as the fraction of pseudo experiments generated in the Standard Model hypothesis which the likelihood fitter measures as having at least as large a fluctuation from the Standard Model as in the data sample. The size of the fluctuation in each experiment is measured the ratio given in Equation 7.2, and the p-value is calculated as

$$\mathcal{P}(\beta_{s_{0.02}}^{J/\psi\phi}, \Delta\Gamma_{s_{0.096}}) = \frac{N_{\mathcal{R}>\mathcal{R}_{\rm data}}}{N_{\rm total}} \tag{7.5}$$

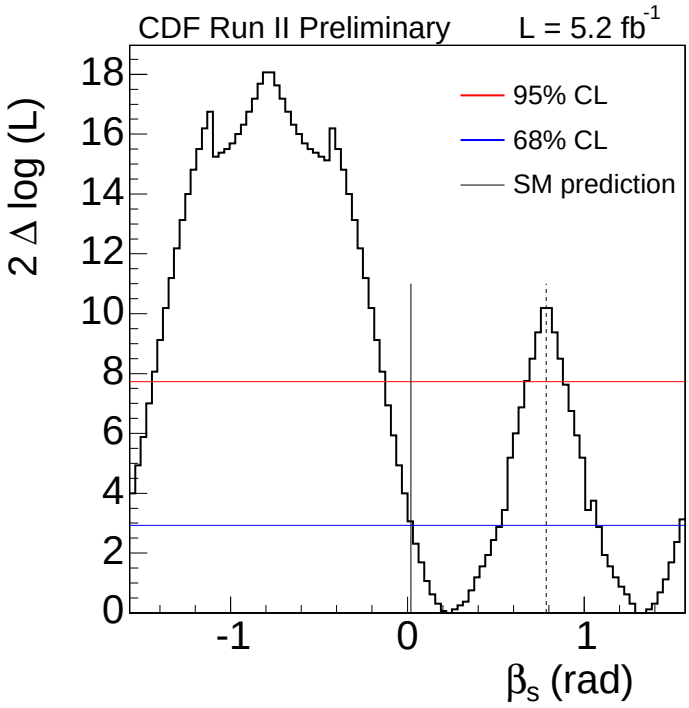

Figure 7.16: 1D confidence intervals for $\beta_s^{J/\psi\phi}$ with full coverage adjustment, using flavour tagged fit.

where N_{total} is the total number of pseudo experiments. The same set of pseudo experiments is used for this estimation as for the coverage adjustment explained in the previous section.

For each coverage adjusted set of confidence regions, a p-value can be calculated for the Standard Model hypothesis. In the 2-dimensional $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ study using flavour tagging information, the p-value is 0.44, equivalent to a 0.8σ fluctuation from the Standard Model expectation. Without flavour tagging, in the 2-dimensional case, the p-value is 0.08, or 1.7σ. The case of the fit without flavour tagging is significantly less sensitive to the true value of $\beta_s^{J/\psi\phi}$, as can be seen from the size of the confidence regions, and this version of the fit has a tendency to find a higher value for $\beta_s^{J/\psi\phi}$ than when the full tagging information is included (as seen in Figure 7.10), so observing a slightly larger deviation from the Standard Model point would be expected in the untagged fit. For the 1-dimensional $\beta_s^{J/\psi\phi}$ confidence intervals using the flavour tagged fit, the p-value is 0.31, which equates to a 1.0σ deviation. In each case, the p-value is well below 3σ, indicating no evidence for a significant deviation from the Standard Model expectation in the data.

Conclusion

In this thesis, the analysis of $B_s^0 \to J/\psi\,\phi$ decays to measure $\beta_s^{J/\psi\phi}$ using the CDF detector at Fermilab has been presented. For the first time, the contribution of $B_s^0 \to J/\psi KK(f^0)$ decays to the $B_s^0 \to J/\psi\,\phi$ signal sample has been assessed and included in the likelihood fit function, where the non-resonant KK, or f^0 is an S-wave state. The resulting measured values for $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ using the full flavour tagged fit are displayed as confidence regions, shown in Figure 7.14, and the 1-dimensional confidence intervals for $\beta_s^{J/\psi\phi}$ in Figure 7.16.

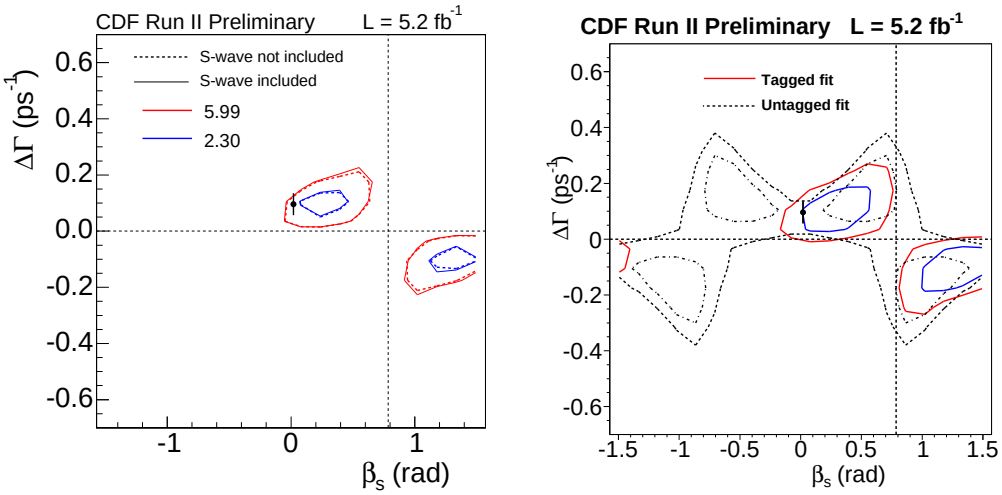

Figure 7.17: [left] Comparison of unadjusted $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ likelihood profile for flavour tagged fit with and without the S-wave KK component included in the likelihood function. [right] Overlay of fully adjusted confidence regions from the fit with and without flavour tagging of the initial B_s^0 meson state.

To cross check the behaviour of the fit, which uses flavour tagging of the initial B_s^0 meson state, the fully coverage adjusted confidence regions were also constructed for the likelihood fit *without* flavour tagging information. This fit has less sensitivity to the value of $\beta_s^{J/\psi\phi}$, but shows no disagreement with the flavour tagged case, as can be seen in the right hand plot of Figure 7.17 where the two versions are overlaid. To check the effect of adding in the $B_s^0 \to J/\psi KK$ component, the unadjusted likelihood profiles from the fit with and without this additional part are shown in the left hand plot of Figure 7.17. The upper limit on the S-wave KK fraction of the signal $B_s^0 \to J/\psi\,\phi$ sample was measured as 6.2% at the 95% confidence level, and the small effect seen in the difference between these two likelihood profiles reflects this small contribution. In addition to this modification of the likelihood function, other improvements to the analysis include re-

calibrated particle ID and flavour taggers, a novel optimisation of the NN cut level for data selection and an enhanced dataset of ~ 6500 signal events.

With a p-value of 0.44, the Standard Model expected values of $\beta_s^{J/\psi\phi}$ and $\Delta\Gamma$ are compatible with the flavour tagged measurements. This p-value indicates a 0.8σ deviation in the measured value from the expectation, which is a smaller shift than observed in previous analyses [4, 7, 6]. The contours can be compared between two earlier CDF measurements, and the Tevatron combined result using data from CDF and DØ in Figure 7.18. The combined result has a similar statistical sample size to this latest measurement, and found a deviation of 2.2σ from the Standard Model expectation. A comparison of the results obtained from the fit with and without taking the S-wave KK component into account is shown in Fig 7.17. This demonstrates that the addition of this component has not caused a shift towards the Standard Model value, as the contours are very similar in both cases of the fit. The studies in pseudo experiments, and in different data periods, presented in Figures 7.10 and 7.11, show that the confidence regions are prone to statistical fluctuations, and demonstrate that the observed differences of the order of 1σ between the latest and previous results are thus to be expected. Within errors, all of the confidence regions in Figure 7.18 are consistent.

On the plots in Figure 7.18, a green band is superimposed, which shows within theoretical errors the relationship $\Delta\Gamma = 2|\Gamma_{12}|\cos(2\beta_s)$. The bottom right plot in this figure shows that the $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ confidence regions from this latest measurement are in very good agreement with this predicted relationship.

The measurement of the CP violating phase $\beta_s^{J/\psi\phi}$ is complimented by further measurements of parameters of interest. These have been assessed with the assumption of no CP violation ($\beta_s^{J/\psi\phi}$=0.0). They include the world's most precise determination of the lifetime and decay width difference of the B_s^0 meson, as well as the transversity amplitudes and the phase $\phi_\perp$:

$$\begin{aligned}
\tau_s &= 1.53 \pm 0.025\ (\text{stat.}) \pm 0.012\ (\text{syst.})\ \text{ps} \\
\Delta\Gamma &= 0.075 \pm 0.035\ (\text{stat.}) \pm 0.01\ (\text{syst.})\ ps^{-1} \\
|A_\parallel(0)|^2 &= 0.231 \pm 0.014\ (\text{stat}) \pm 0.015\ (\text{syst.}) \\
|A_0(0)|^2 &= 0.524 \pm 0.013\ (\text{stat}) \pm 0.015\ (\text{syst.}) \\
\phi_\perp &= 2.95 \pm 0.64\ (\text{stat}) \pm 0.07\ (\text{syst.})
\end{aligned} \tag{7.6}$$

These results use the fully flavour tagged fit with $\beta_s^{J/\psi\phi}$ fixed to 0.0. A detailed study of the influences of systematic effects for these measurements was presented in Chapter 6. The τ_s measurement compares well with the PDG world average value [10] of $\tau_s = 1.47^{+0.026}_{-0.027}$ ps. Theoretical predictions from HQET (also see Chapter 1) for the parameter τ_s suggest that the ratio of the B_s^0 and B^0 lifetimes should be ≈ 1 [30]. The B^0 lifetime was recently measured at CDF in the channel $B^0 \to J/\psi K^{*0}$ [105] to be

$$\tau_{B^0} = 1.50 \pm 0.010\ (\text{stat.}) \pm 0.008\ (\text{syst.})\ \text{ps} \tag{7.7}$$

comparing this with the B_s^0 lifetime measured in $B_s^0 \to J/\psi\,\phi$ shows consistency with HQET predictions.

Looking to the future, it is clear from the progression shown in Figure 7.18 that in-

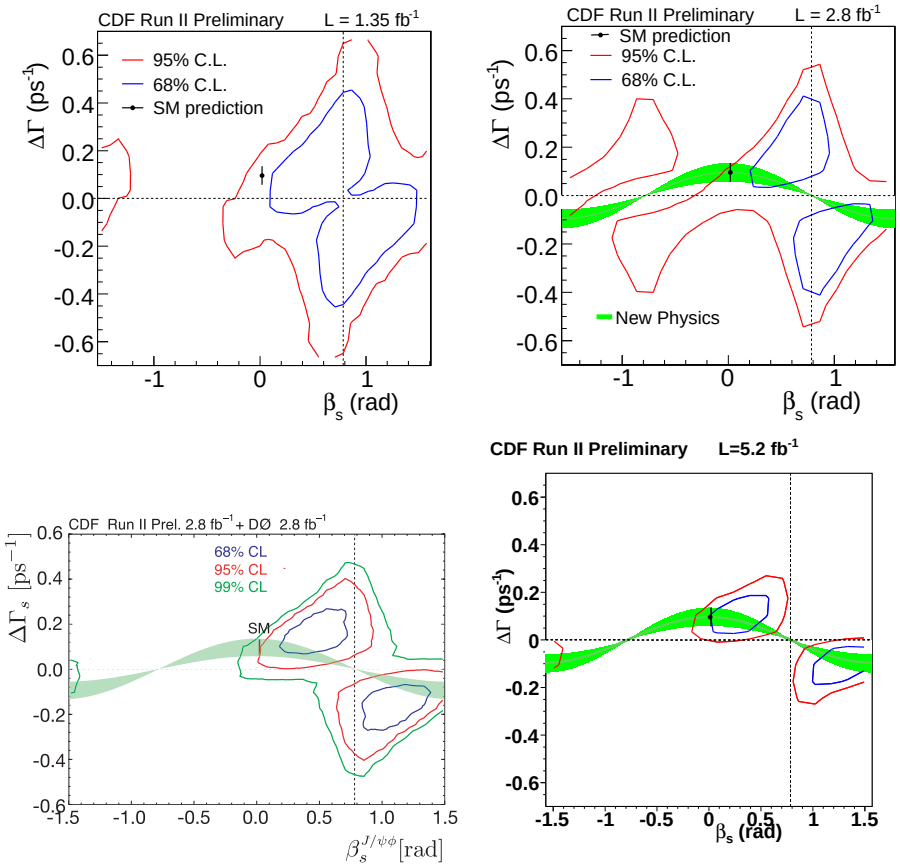

Figure 7.18: Comparison of latest $\beta_s^{J/\psi\phi}$-$\Delta\Gamma$ confidence regions (lower right) with previous results from CDF published analysis [4] (upper left), CDF ICHEP 2008 result [7] (upper right) and Tevatron combined result using 2.8 fb^{-1} datasets from both CDF and D0 experiments [6].

creasing the data sample size will not only shrink the confidence regions, giving a more accurate estimation of the true value of $\beta_s^{J/\psi\phi}$, but also reduce non-Gaussian uncertainties in the measurement, bringing the errors closer to a fully Gaussian regime. Further increases in statistics should make the measurement of a point value for $\beta_s^{J/\psi\phi}$ achievable. The CDF experiment expects a total recorded dataset of > 10 fb^{-1} by the end of 2011 [108]. This would mean approximately doubling again the sample size available for the measurement of $\beta_s^{J/\psi\phi}$. As CDF stops taking data, and the LHC data samples increase, it will be the turn of the LHC experiments to take on this measurement. In particular LHC*b* expects a high precision in a direct measurement of $\beta_s^{J/\psi\phi}$ in $B_s^0 \to J/\psi\,\phi$ [109]. Furthermore the large expected statistics and higher center of mass energy will open up additional channels to access $\beta_s^{J/\psi\phi}$.

Appendix A

Study of dE/dx dependence on correlated parameters

The following figures show the dependence of uncalibrated dE/dx on six variables, in slices of each of the other variables. These were used in conjunction with cross checks at each stage of the calibration process to find which parameters could be corrected independently and which must be treated simultaneously. Those parameters for which the behaviour of dE/dx is comparable in different slices (neglecting a global shift) are considered to be factorisable.

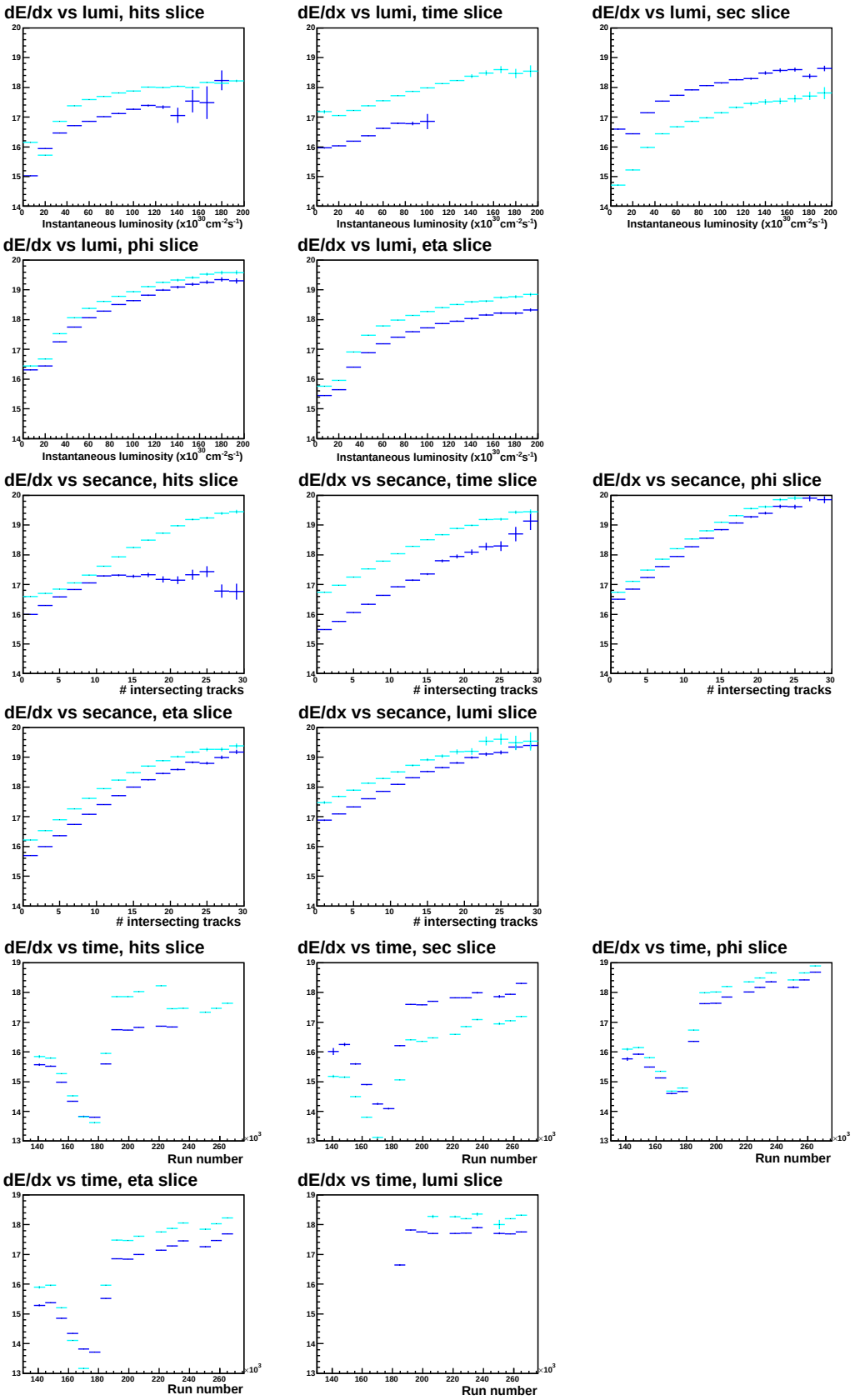

Figure A.1: variations in measured dE/dx with respect to luminosity, track density (secance), time, in slices of each ther parameter

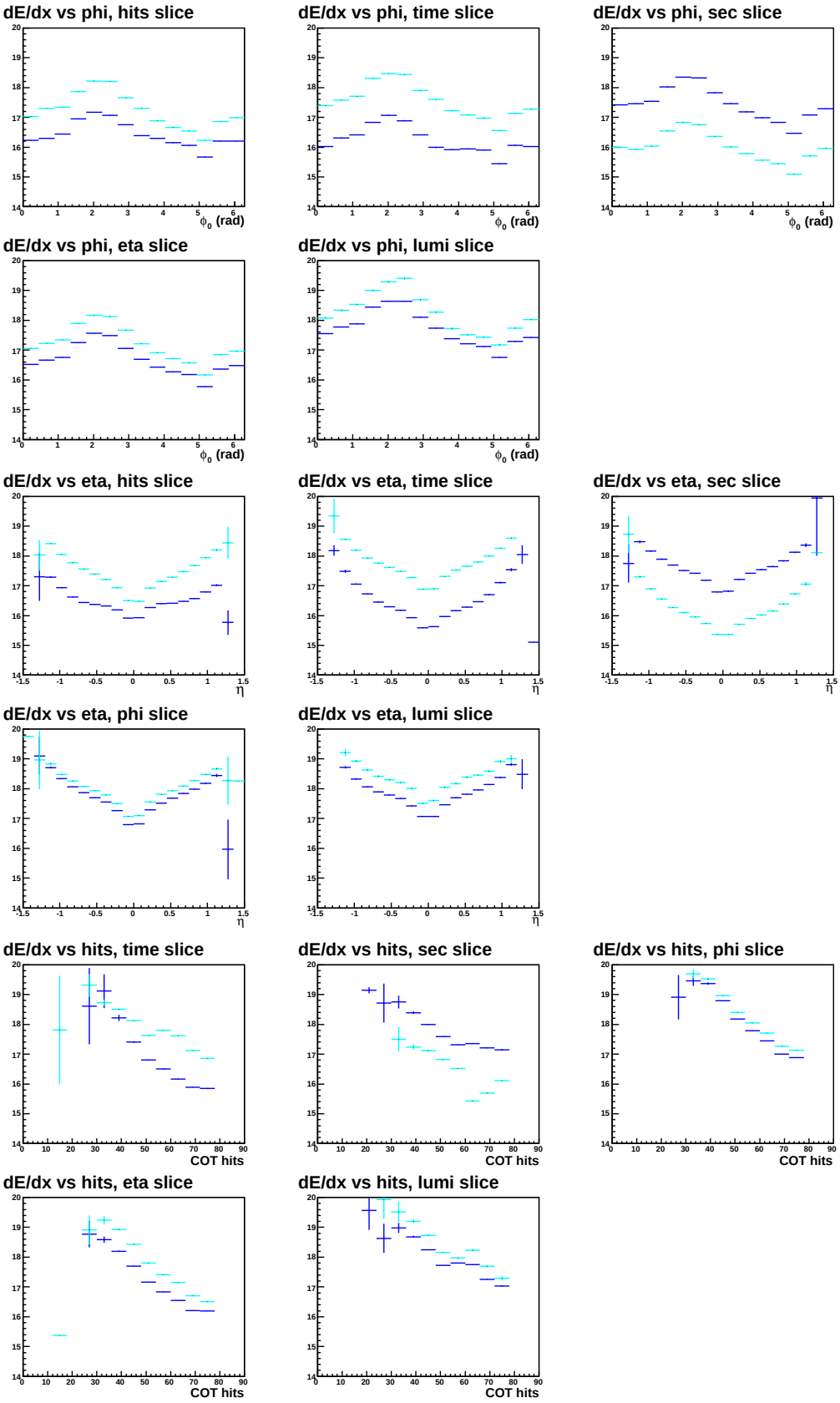

Figure A.2: variations in measured dE/dx with respect to time, eta, phi0 and number of COT hits, in slices of each other parameter

Appendix B

Normalisation of $B_s^0 \to J/\psi\,\phi$ transversity PDF

The PDF for time and transversity angle dependence in $B_s^0 \to J/\psi\,\phi$ decays is developed in Chapter 4, here the analytic normalisation of this component is explained in more detail, following the method described in [101]. Starting from Equation 4.9, the first step is to integrate over time the decay rates

$$\begin{aligned}
\int |f_+(t)|^2 + |\bar{f}_+(t)|^2 dt &= 1, \\
\int |f_-(t)|^2 + |\bar{f}_-(t)|^2 dt &= 1, \\
\int \sum_{i=B,\bar{B}} f_+(t) f_-^*(t) dt &= \frac{i}{\sqrt{1 + \frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta')^2}}},
\end{aligned} \tag{B.1}$$

giving factors which can be substituted in via Equation 1.81 to re-write N as

$$\begin{aligned}
N = \frac{9}{16\pi} \int\int\int d(\cos\psi) d(\cos\theta) d\phi \\
\left[|\mathbf{A}_+ \times \hat{n}|^2 + |\mathbf{A}_- \times \hat{n}|^2 + 2Re((\mathbf{A}_+ \times \hat{n}) \cdot (\mathbf{A}_-^* \times \hat{n}) \frac{i}{\sqrt{1 + \frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta')^2}}}) \right] \\
\epsilon(\psi, \theta, \phi).
\end{aligned} \tag{B.2}$$

For the purposes of simplifying the angular integration, it can be observed that

$$\begin{aligned}
\sin\theta\cos\phi &= \sqrt{\frac{2\pi}{3}}(Y_1^{-1} - Y_1^1), \\
\sin\theta\sin\phi &= \sqrt{\frac{2\pi}{3}} i (Y_1^{-1} + Y_1^1), \\
\cos\theta &= \sqrt{\frac{4\pi}{3}} Y_1^0.
\end{aligned} \tag{B.3}$$

where, recalling Equations 1.62, 1.84 and 1.84, the terms above enter the PDF as products. These can be substituted into N to give

$$\begin{aligned}
N \;=\; & \frac{9}{16\pi}\int\int\int \epsilon(\psi,\theta,\phi)d(\cos\psi)d(\cos\theta)d\phi\frac{|A_\perp|^2}{2}\left[\sqrt{\frac{16\pi}{9}}Y_0^0-\sqrt{\frac{16\pi}{45}}Y_2^0\right]\sin^2\psi \\
& +\frac{|A_\parallel|^2}{2}\left[\sqrt{\frac{4\pi}{9}}Y_0^0+\sqrt{\frac{16\pi}{45}}Y_2^0\right]\sin^2\psi+|A_0|^2\left[\sqrt{\frac{4\pi}{9}}Y_0^0+\sqrt{\frac{16\pi}{45}}Y_2^0\right]\cos^2\psi \\
& +\frac{|A_\parallel|^2}{2}\left[\sqrt{\frac{4\pi}{9}}Y_0^0-\sqrt{\frac{4\pi}{45}}Y_2^0+\sqrt{\frac{2\pi}{15}}(Y_2^{-2}+Y_2^2)\right]\sin^2\psi \\
& +|A_0|^2\left[\sqrt{\frac{4\pi}{9}}Y_0^0-\sqrt{\frac{4\pi}{45}}Y_2^0-\sqrt{\frac{2\pi}{15}}(Y_2^{-2}+Y_2^2)\right]\cos^2\psi \\
& +i\frac{A_\parallel A_0^*+A_\parallel^* A_0}{\sqrt{2}}\left[\sqrt{\frac{2\pi}{15}}(Y_2^{-2}-Y_2^2)\right]\sin\psi\cos\psi+2Re\left[\frac{i}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta')^2}}}\right. \\
& \left.\times\left\{\frac{A_\parallel A_\perp^*}{2}\sqrt{\frac{2\pi}{15}}(Y_2^{-1}+Y_2^1)\sin^2\psi+i\frac{A_0A_\perp^*}{\sqrt{2}}\sqrt{\frac{2\pi}{15}}(Y_2^{-1}-Y_2^1)\sin\psi\cos\psi\right\}\right] \\
& \epsilon(\psi,\theta,\phi)d(\cos\psi). \qquad \text{(B.4)}
\end{aligned}$$

Now, the efficiency function can be dealt with, this was defined in Equation 4.10 in terms of spherical harmonics and Legendre polynomials. Integrating Equation 4.10 gives

$$\begin{aligned}
\int\int Y_l^m(\theta,\phi)\bar{Y}_{l'}^{m'}(\theta,\phi)\sin\theta d\theta d\phi \;&=\; \delta_{mm'}\delta_{ll'}, \\
Y_l^{-m}(\theta,\phi) &= (-1)^m\bar{Y}_l^m(\theta,\phi), \qquad \text{(B.5)}
\end{aligned}$$

Putting this definition into Equation B.4 leads to

$$\begin{aligned}
N &= \frac{9}{16\pi}\int d(\cos\psi)\frac{|A_\perp|^2}{2}\left[\sqrt{\frac{16\pi}{9}}a_0^0(\psi)-\sqrt{\frac{16\pi}{45}}a_2^0(\psi)\right]\sin^2\psi \\
&+\frac{|A_\parallel|^2}{2}\left[\sqrt{\frac{4\pi}{9}}a_0^0(\psi)+\sqrt{\frac{16\pi}{45}}a_2^0(\psi)\right]\sin^2\psi+|A_0|^2\left[\sqrt{\frac{4\pi}{9}}a_0^0(\psi)+\sqrt{\frac{16\pi}{45}}a_2^0(\psi)\right]\cos^2\psi \\
&+\frac{|A_\parallel|^2}{2}\left[\sqrt{\frac{4\pi}{9}}a_0^0(\psi)-\sqrt{\frac{4\pi}{45}}a_2^0(\psi)+\sqrt{\frac{2\pi}{15}}(a_2^{-2}(\psi)+a_2^2(\psi))\right]\sin^2\psi \\
&+|A_0|^2\left[\sqrt{\frac{4\pi}{9}}a_0^0(\psi)-\sqrt{\frac{4\pi}{45}}a_2^0(\psi)-\sqrt{\frac{2\pi}{15}}(a_2^{-2}(\psi)+a_2^2(\psi))\right]\cos^2\psi \\
&-i\frac{A_\parallel A_0^*+A_\parallel^* A_0}{\sqrt{2}}\left[\sqrt{\frac{2\pi}{15}}(a_2^{-2}(\psi)-a_2^2(\psi))\right]\sin\psi\cos\psi+2Re\left[\frac{i}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta')^2}}}\right. \\
&\left.\times\left\{-\frac{A_\parallel A_\perp^*}{2}\sqrt{\frac{2\pi}{15}}(a_2^{-1}(\psi))+a_2^1(\psi)\sin^2\psi+i\frac{A_0A_\perp^*}{\sqrt{2}}\sqrt{\frac{2\pi}{15}}(a_2^{-1}(\psi)-a_2^1(\psi))\sin\psi\cos\psi\right\}\right].
\end{aligned} \tag{B.6}$$

A set of orthonormal basis functions, in terms of spherical harmonics

$$\begin{aligned}
Y_{lm} &= Y_l^m \quad (m=0), \\
Y_{lm} &= \frac{1}{\sqrt{2}}(Y_l^m+(-1)^m Y_l^{-m}) \quad m>0, \\
Y_{lm} &= \frac{1}{i\sqrt{2}}(Y_l^{|m|}-(-1)^{|m|}Y_l^{-|m|}) \quad m<0.
\end{aligned} \tag{B.7}$$

are used to express the coefficients a_{lm} as

$$\begin{aligned}
a_{lm} &= a_l^m \quad (m=0), \\
a_{lm} &= \frac{1}{\sqrt{2}}(a_l^m+(-1)^{|m|}a_l^{-m}) \quad m>0, \\
a_{lm} &= \frac{i}{\sqrt{2}}(a_l^{|m|}-(-1)^{|m|}a_l^{-|m|}) \quad m<0.
\end{aligned} \tag{B.8}$$

Substituting the explicit forms of B.8 into N, and simplifying terms leaves an integral dependent only on ψ:

$$
\begin{aligned}
N &= \frac{3}{8\sqrt{\pi}}\left[(|A_\perp|^2+|A_\parallel|^2)\int a_{00}(\psi)\sin^2\psi d(\cos\psi)+2|A_0|^2\int a_{00}(\psi)\cos^2\psi d(\cos\psi)\right] \\
&+ \frac{3}{8\sqrt{5\pi}}\left[(-|A_\perp|^2+\frac{1}{2}|A_\parallel|^2)\int a_{20}(\psi)\sin^2\psi d(\cos\psi)+|A_0|^2\int a_{20}(\psi)\cos^2\psi d(\cos\psi)\right] \\
&- \frac{9}{16\sqrt{15\pi}}\frac{1}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta')^2}}}\left[(A_\parallel^* A_\perp+A_\parallel A_\perp^*)\int a_{2-1}(\psi)\sin^2\psi d(\cos\psi)\right] \\
&+ \frac{9}{16}\frac{\sqrt{2}}{\sqrt{15\pi}}\frac{1}{\sqrt{1+\frac{4\tau_L\tau_H}{((\tau_L-\tau_H)\sin 2\beta')^2}}}\left[(A_0^* A_\perp+A_0 A_\perp^*)\int a_{21}(\psi)\sin\psi\cos\psi d(\cos\psi)\right] \\
&+ \frac{9}{8\sqrt{15\pi}}\left[\frac{|A_\parallel|^2}{2}\int a_{22}(\psi)\sin^2\psi d(\cos\psi)-|A_0|^2\int a_{22}(\psi)\cos^2\psi d(\cos\psi)\right] \qquad \text{(B.9)} \\
&+ \frac{9}{16}\frac{\sqrt{2}}{\sqrt{15\pi}}\left[(A_0^* A_\parallel+A_0 A_\parallel^*)\int a_{2-2}(\psi)\sin\psi\cos\psi d(\cos\psi)\right]. \qquad \text{(B.10)}
\end{aligned}
$$

Finally, N must be integrated over ψ, by expressing $a_{lm}(\psi)$ as a Fourier-Legendre series:

$$a_{lm}(\psi) = a_{lm}^k P_k(\cos\psi). \qquad \text{(B.11)}$$

The few a_{lm}^kterms remaining after integration are

$$
\begin{aligned}
\int a_{lm}(\cos\psi)\sin^2(\psi)d(\cos\psi) &= \frac{4}{3}a_{lm}^0-\frac{4}{15}a_{lm}^2, \\
\int a_{lm}(\cos\psi)\cos^2(\psi)d(\cos\psi) &= \frac{2}{3}a_{lm}^0+\frac{4}{15}a_{lm}^2, \\
\int a_{lm}(\cos\psi)\cos(\psi)\sin\psi d(\cos\psi) &= \frac{\pi}{8}a_{lm}^1-\frac{\pi}{32}a_{lm}^3+.... \qquad \text{(B.12)}
\end{aligned}
$$

which are substituted into Equation B.10 to give the final normalisation for the $B_s^0 \to J/\psi\,\phi$ PDF which is given in Equation 4.11. The integration of the S-wave KK component PDF and S- and P-wave interference term is carried out numerically in the fit.

Appendix C

Results of 3-dimensional fit to detector angular efficiency

The results of the 3-dimensional fit to the detector angular sculpting of the transversity angles described in Section 4.4.1 is given in Table C.1. These coefficients can be substituted into the normalisation factor in Equation 4.11 to normalise the PDF for B_s^0 decay time and transversity angular distributions.

Coefficient	Fit Result	Coefficient	Fit Result	Coefficient	Fit Result
C_{00}^{0}	843.96 ± 0.778826	C_{22}^{0}	40.2944 ± 6.92456	C_{33}^{0}	-0.0941144 ± 0.45078
C_{00}^{2}	-9.00419 ± 0.78018	C_{22}^{2}	-2.71047 ± 6.93834	C_{33}^{2}	-0.188061 ± 0.451555
C_{11}^{0}	-3.73299 ± 0.797056	S_{22}^{0}	0.468246 ± 7.01755	S_{33}^{0}	2.59885 ± 0.455331
C_{11}^{2}	-0.525782 ± 0.79737	S_{22}^{2}	0.50368 ± 7.02796	S_{33}^{2}	-1.09173 ± 0.456139
S_{11}^{0}	0.630685 ± 0.769779	C_{21}^{0}	0.783937 ± 0.792322	C_{32}^{0}	11.2979 ± 6.11867
S_{11}^{2}	-0.43546 ± 0.772351	C_{21}^{2}	1.26672 ± 0.793161	C_{32}^{3}	-1.26441 ± 6.13092
C_{10}^{0}	0.225966 ± 0.770231	S_{21}^{0}	-0.3033 ± 0.761416	S_{32}^{0}	-0.0258396 ± 6.1991
C_{10}^{2}	1.09715 ± 0.771387	S_{21}^{2}	-1.30095 ± 0.763543	S_{32}^{2}	0.52274 ± 6.20827
		C_{20}^{0}	-32.3269 ± 0.770779	C_{31}^{0}	0.384408 ± 0.775819
		C_{20}^{2}	0.227447 ± 0.771696	C_{31}^{2}	-0.707931 ± 0.776999
				S_{31}^{0}	-0.715725 ± 0.755324
				S_{31}^{2}	-1.18586 ± 0.757484
				C_{30}^{0}	-0.280284 ± 0.785759
				C_{30}^{2}	1.04249 ± 0.786224

Coefficient	Fit Result	Coefficient	Fit Result
C_{44}^{0}	-21.4158 ± 0.781966	C_{55}^{0}	0.213818 ± 0.78413
C_{44}^{2}	-0.690403 ± 0.783678	C_{55}^{2}	-0.674089 ± 0.78570
S_{44}^{0}	-0.109872 ± 0.786434	S_{55}^{0}	6.07759 ± 0.783858
S_{44}^{2}	-0.0355397 ± 0.787626	S_{55}^{2}	-0.945723 ± 0.78524
C_{43}^{0}	-0.738286 ± 0.784317	C_{54}^{0}	2.02852 ± 0.78581
C_{43}^{2}	0.60811 ± 0.78559	C_{54}^{2}	-0.187081 ± 0.78689
S_{43}^{0}	0.697828 ± 0.786366	S_{54}^{0}	0.356674 ± 0.787165
S_{43}^{2}	0.767655 ± 0.787814	S_{54}^{2}	0.253496 ± 0.788684
C_{42}^{0}	3.80235 ± 7.98292	C_{53}^{0}	-0.78978 ± 0.776905
C_{42}^{2}	0.997011 ± 7.9989	C_{53}^{2}	1.1786 ± 0.77847
S_{42}^{0}	-0.579774 ± 8.08883	S_{53}^{0}	-2.59779 ± 0.781566
S_{42}^{2}	-0.54718 ± 8.1008	S_{53}^{2}	0.243988 ± 0.783463
C_{41}^{0}	-0.444079 ± 0.769242	C_{52}^{0}	6.06004 ± 3.66702
C_{41}^{2}	0.760549 ± 0.769936	C_{52}^{2}	0.000679862 ± 3.674
S_{41}^{0}	-0.279113 ± 0.746484	S_{52}^{0}	0.0452646 ± 3.71547
S_{41}^{2}	-0.537755 ± 0.747896	S_{52}^{2}	-0.0236447 ± 3.7209
C_{40}^{0}	-13.4083 ± 0.800189	C_{51}^{0}	-0.913286 ± 0.75785
C_{40}^{2}	-1.16636 ± 0.801091	C_{51}^{2}	-0.0953101 ± 0.7580
		S_{51}^{0}	1.19732 ± 0.738142
		S_{51}^{0}	-1.13507 ± 0.739085
		C_{50}^{0}	-1.89068 ± 0.828197
		C_{50}^{2}	0.545038 ± 0.829863

Table C.1: 3-dimensional fit to angular sculpting of the CDF detector.

Appendix D

Propagation of correlated errors

The partial derivatives for substitution into Equation 4.42, in order to propagate the correlated errors from the $\alpha_{\parallel,\mathrm{CPodd}}$ fitted parameters to the initial transversity amplitudes, $A_{0,\parallel}(0)$ are given here. For $A_0(0)$ these are:

$$\frac{\partial|A_0(0)|^2}{\partial c\tau} = \frac{4\frac{\Delta\Gamma}{c}(1-\alpha_\parallel)(1-\alpha_{\mathrm{CPodd}})\alpha_{\mathrm{CPodd}}}{(1+(y-1)\alpha_{\mathrm{CPodd}})^2(2+\frac{\Delta\Gamma}{c}c\tau)^2} \quad \text{(D.1)}$$

$$\frac{\partial|A_0(0)|^2}{\partial \Delta\Gamma} = \frac{4c\tau(1-\alpha_\parallel)(1-\alpha_{\mathrm{CPodd}})\alpha_{\mathrm{CPodd}}}{(1+(y-1)\alpha_{\mathrm{CPodd}})^2(2+\frac{\Delta\Gamma}{c}c\tau)^2} \quad \text{(D.2)}$$

$$\frac{\partial|A_0(0)|^2}{\partial \alpha_\parallel} = \frac{(1-\alpha_{\mathrm{CPOdd}})}{(1+(y-1)\alpha_{\mathrm{CPodd}})} \quad \text{(D.3)}$$

$$\frac{\partial|A_0(0)|^2}{\partial \alpha_{\mathrm{CP_o dd}}} = \frac{(-\alpha_\parallel(1+(y-1)\alpha_{\mathrm{CPodd}})-(y-1)(1-\alpha_{\mathrm{CPodd}})(1-\alpha_\parallel))}{(1+(y-1)\alpha_{\mathrm{CPodd}})^2}. \quad \text{(D.4)}$$

For $A_\parallel(0)$ the partial derivatives are:

$$\frac{\partial|A_\parallel(0)|^2}{\partial c\tau} = \frac{4\frac{\Delta\Gamma}{c}\alpha_\parallel(1-\alpha_{\mathrm{CPodd}})\alpha_{\mathrm{CPodd}}}{(1+(y-1)\alpha_{\mathrm{CPodd}})^2(2+\frac{\Delta\Gamma}{c}c\tau)^2} \quad \text{(D.5)}$$

$$\frac{\partial|A_\parallel(0)|^2}{\partial \Delta\Gamma} = \frac{4c\tau\alpha_\parallel(1-\alpha_{\mathrm{CPodd}})\alpha_{\mathrm{CPodd}}}{(1+(y-1)\alpha_{\mathrm{CPodd}})^2(2+\frac{\Delta\Gamma}{c}c\tau)^2} \quad \text{(D.6)}$$

$$\frac{\partial|A_\parallel(0)|^2}{\partial \alpha_\parallel} = \frac{-(1-\alpha_{\mathrm{CPOdd}})}{(1+(y-1)\alpha_{\mathrm{CPodd}})} \quad \text{(D.7)}$$

$$\frac{\partial|A_\parallel(0)|^2}{\partial \alpha_{\mathrm{CP_o dd}}} = \frac{-(1-\alpha_\parallel(1+(y-1)\alpha_{\mathrm{CPodd}})-(y-1)(1-\alpha_{\mathrm{CPodd}})(1-\alpha_\parallel))}{(1+(y-1)\alpha_{\mathrm{CPodd}})^2}. \quad \text{(D.8)}$$

	$c\tau$	$\Delta\Gamma$	α_{CPodd}	$\alpha_{\parallel}$
$c\tau$	5.90484e-07	1.46913e-05	-1.89861e-06	1.03847e-06
$\Delta\Gamma$	1.46913e-05	0.00117218	-8.11584e-05	4.39767e-05
α_{CPodd}	-1.89861e-06	-8.11584e-05	0.000189205	-0.000114205
$\alpha_{\parallel}$	1.03847e-06	4.39767e-05	-0.000114205	0.000223853

Table D.1: Correlations in statistical errors for paramters used in transformation from $\alpha_{\{\parallel,\text{CPodd}\}}$ to $A_{\{0,\parallel\}}$ for fit including flavour tagging, with the S-wave KK component in the likelihood fit.

	$c\tau$	$\Delta\Gamma$	α_{CPodd}	$\alpha_{\parallel}$
$c\tau$	6.32898e-07	1.62166e-05	-6.60451e-06	5.53911e-06
$\Delta\Gamma$	1.62166e-05	0.00125199	-0.000151459	0.00010744
α_{CPodd}	-6.60451e-06	-0.000151459	0.000999688	-0.00090279
$\alpha_{\parallel}$	5.53911e-06	0.00010744	-0.00090279	0.00099388

Table D.2: Correlations in statistical errors for paramters used in transformation from $\alpha_{\{\parallel,\text{CPodd}\}}$ to $A_{\{0,\parallel\}}$ for fit without flavour tagging with the S-wave KK component in the likelihood fit.

	$c\tau$	$\Delta\Gamma$	α_{CPodd}	$\alpha_{\parallel}$
$c\tau$	5.90663e-07	1.46933e-05	-1.90319e-06	1.04273e-06
$\Delta\Gamma$	1.46933e-05	0.00117259	-8.13323e-05	4.41253e-05
α_{CPodd}	-1.90319e-06	-8.13323e-05	0.000189682	-0.000114706
$\alpha_{\parallel}$	1.04273e-06	4.41253e-05	-0.000114706	0.000224432

Table D.3: Correlations in statistical errors for paramters used in transformation from $\alpha_{\{\parallel,\text{CPodd}\}}$ to $A_{\{0,\parallel\}}$ for fit including flavour tagging without the S-wave KK component in the likelihood fit.

	$c\tau$	$\Delta\Gamma$	α_{CPodd}	$\alpha_{\parallel}$
$c\tau$	6.22597e-07	1.65845e-05	-2.39201e-06	1.39609e-06
$\Delta\Gamma$	1.65845e-05	0.0012749	-9.84193e-05	5.43697e-05
α_{CPodd}	-2.39201e-06	-9.84193e-05	0.000246075	-0.000166969
$\alpha_{\parallel}$	1.39609e-06	5.43697e-05	-0.000166969	0.000274719

Table D.4: Correlations in statistical errors for paramters used in transformation from $\alpha_{\{\parallel,\text{CPodd}\}}$ to $A_{\{0,\parallel\}}$ for fit without flavour tagging without the S-wave KK component in the likelihood fit.

Appendix E

Correlation matrices

The correlation matrices for the full fit to data, and for an example high statistics pseudo experiment are shown here. These are produced by the fit package Minuit [96]. Table E.1 shows the labeling scheme for the paramters in the fit.

Parameter	label	index	Parameter	label	index
$c\tau_s$	Lifetime	1	$\epsilon_s(OST)$	EffSig	19
$s_{c\tau}1$	ScaleFac	2	$\epsilon_b(OST)$	EffBkg	20
$s_{c\tau}2$	ScaleFac2	3	$\mathcal{A}^+(OST)$	EffBkgp	21
f_{sf1}	ScaleFrac1	4	$\epsilon_s(SST)$	EffSig2	22
m	Mean	5	$\epsilon_b(SST)$	EffBkg2	23
s_m	MassScl	6	$\mathcal{A}^+(SST)$	EffBkgp	24
$\alpha_\perp$	AlphaCPOdd	7	$\cos(\psi)_1$	cosTheta	25
$\alpha_\parallel$	AlphaPara	8	$\cos(\theta)_1$	cosPsi1	26
$\phi_\parallel$	PhiPara	9	ϕ_1	Phi1	27
$\phi_\perp$	PhiPerp	10	f_s	FracSig	28
$\beta_s^{J/\psi\phi}$	BetaPrime	11	p_1	BGmslop	29
A_{SW}	ASWave	12	λ_-	Lambdap	30
δ_{SW}	DeltaS	13	λ_+	Lambdap	31
Δm_s	deltaM	14	λ_{++}	Lambdam	32
$S_\mathcal{D}(OST+)$	DScale	15	f_p	Fgauss	33
$S_\mathcal{D}(OST-)$	DScale2	16	f_-	Fm	34
$S_\mathcal{D}(SST)$	DScale21	17	f_{++}	Fpp	35
$\Delta\Gamma$	deltaG	18			

Table E.1: Labeling of parameters in correlation matrix printout

PARAMETER CORRELATION COEFFICIENTS

	NO.	GLOBAL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Lifetime	1	0.53971	1.000	0.001	-0.004	0.000	-0.003	-0.034	-0.189	0.144	0.157	0.143	-0.313	-0.188	0.148	0.036	-0.001	-0.005
ScaleFac	2	0.81875	0.001	1.000	0.704	-0.798	0.001	0.007	-0.003	0.005	0.009	0.008	0.003	-0.008	0.008	0.002	0.001	-0.002
ScaleFac2	3	0.91144	-0.004	0.704	1.000	-0.848	0.003	0.009	-0.001	0.004	0.005	0.005	0.003	-0.005	0.005	0.001	0.000	-0.001
ScaleFrac	4	0.91988	0.000	-0.798	-0.848	1.000	-0.002	-0.005	0.001	-0.003	-0.005	-0.003	0.000	0.005	-0.004	0.000	0.000	0.001
Mean	5	0.04213	-0.003	0.001	0.003	-0.002	1.000	0.025	0.005	-0.004	-0.003	-0.004	-0.002	0.006	-0.002	-0.002	0.001	-0.001
MassScl	6	0.26237	-0.034	0.007	0.009	-0.005	0.025	1.000	-0.003	0.014	0.013	0.010	0.003	-0.010	0.012	0.002	-0.003	0.001
AlphaCP0	7	0.67924	-0.189	-0.003	-0.001	0.001	0.005	-0.003	1.000	-0.618	-0.430	-0.322	-0.027	0.079	-0.283	-0.025	0.016	0.032
AlphaPar	8	0.66368	0.144	0.005	0.004	-0.003	-0.004	0.014	-0.618	1.000	0.466	0.368	0.041	-0.231	0.349	0.045	-0.017	-0.032
PhiPara	9	0.83610	0.157	0.009	0.005	-0.005	-0.003	0.013	-0.430	0.466	1.000	0.782	0.113	-0.565	0.724	0.086	-0.034	-0.070
PhiPerp	10	0.91799	0.143	0.008	0.005	-0.003	-0.004	0.010	-0.322	0.368	0.782	1.000	0.121	-0.555	0.892	0.250	-0.027	-0.053
BetaPr	11	0.46711	-0.313	0.003	0.003	0.000	-0.002	0.003	-0.027	0.041	0.113	0.121	1.000	-0.049	0.132	-0.027	-0.015	-0.023
ASWave	12	0.69323	-0.188	-0.008	-0.005	0.005	0.006	-0.010	0.079	-0.231	-0.565	-0.555	-0.049	1.000	-0.627	-0.120	0.016	0.028
DeltaS	13	0.90662	0.148	0.008	0.005	-0.004	-0.002	0.012	-0.283	0.349	0.724	0.892	0.132	-0.627	1.000	0.221	-0.023	-0.048
deltaM	14	0.31895	0.036	0.002	0.001	0.000	-0.002	0.002	-0.025	0.045	0.086	0.250	-0.027	-0.120	0.221	1.000	-0.005	-0.009
DScale	15	0.13347	-0.001	0.001	0.000	0.000	0.001	-0.003	0.016	-0.017	-0.034	-0.027	-0.015	0.016	-0.023	-0.005	1.000	0.129
DScale2	16	0.15852	-0.005	-0.002	-0.001	0.001	-0.001	0.001	0.032	-0.032	-0.070	-0.053	-0.023	0.028	-0.048	-0.009	0.129	1.000
DScale21	17	0.23063	0.059	0.002	0.001	-0.001	0.002	-0.002	-0.022	0.007	0.019	-0.043	-0.187	0.029	-0.042	-0.079	-0.023	-0.060
deltaG	18	0.42831	0.280	0.001	0.002	0.000	-0.001	0.000	-0.214	0.136	0.103	0.114	0.148	0.016	0.116	-0.004	-0.007	-0.011
EffSig	19	0.10902	0.015	0.000	0.002	0.001	0.003	-0.004	-0.002	-0.003	-0.002	-0.002	-0.001	0.006	-0.003	-0.001	0.000	0.000
EffBkg	20	0.09566	0.008	0.000	-0.001	0.000	-0.001	-0.017	0.001	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	0.000	0.000	0.000
EffBkgp	21	0.02444	-0.002	-0.001	-0.001	0.001	0.000	0.004	-0.001	0.001	0.003	0.003	0.008	-0.004	0.003	0.000	-0.013	0.016
EffSig2	22	0.10986	-0.025	0.001	0.003	-0.003	-0.004	0.023	0.004	0.005	0.006	0.006	0.004	-0.005	0.006	0.001	-0.001	0.000
EffBkg2	23	0.09507	-0.012	-0.001	-0.001	0.002	0.003	0.023	-0.002	0.003	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
EffBkgp2	24	0.01631	-0.003	0.000	0.001	0.000	0.000	0.001	-0.002	0.002	0.003	0.004	0.014	-0.003	0.004	-0.001	0.000	-0.001
cosTheta	25	0.08023	0.013	-0.001	-0.001	0.001	-0.002	0.004	-0.064	0.033	0.003	-0.002	-0.007	0.036	-0.007	-0.003	0.000	0.000
cosPsi1	26	0.09414	-0.019	0.000	0.001	-0.001	0.000	-0.013	0.031	0.034	-0.019	-0.019	0.009	0.028	-0.021	-0.004	0.001	0.001
Phi1	27	0.08682	-0.020	0.000	0.001	0.000	0.002	0.007	0.027	-0.060	-0.056	-0.050	-0.004	0.073	-0.054	-0.009	0.002	0.003
FracSig	28	0.27424	-0.125	-0.001	0.003	0.000	0.004	0.170	0.002	0.030	0.022	0.018	0.013	-0.013	0.017	0.002	-0.003	0.000
BGmslope	29	0.02624	0.002	0.000	0.000	0.000	-0.025	-0.006	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Lambdapp	30	0.89328	-0.017	-0.023	-0.163	0.138	-0.018	0.000	-0.002	0.005	0.002	0.001	-0.007	-0.001	-0.001	0.001	0.000	0.000
Lambdap	31	0.87332	-0.051	0.004	-0.054	0.034	-0.016	0.003	0.000	0.006	0.000	-0.001	-0.008	0.001	-0.003	0.001	0.000	0.000
Lambdam	32	0.60305	-0.003	0.218	0.370	-0.207	0.002	0.005	0.000	0.002	0.003	0.002	0.002	-0.002	0.002	0.001	0.000	-0.001
Fgauss	33	0.80370	-0.054	0.150	0.147	0.032	-0.003	0.103	0.003	0.019	0.019	0.014	0.007	-0.012	0.012	0.003	-0.002	-0.001
Fm	34	0.87846	-0.011	-0.220	-0.426	0.267	-0.015	0.045	0.001	0.006	0.002	0.001	-0.010	-0.001	-0.001	0.001	-0.001	0.001
Fpp	35	0.91374	0.001	0.024	-0.013	-0.030	-0.018	0.036	0.000	0.004	0.001	0.000	-0.013	-0.001	-0.002	0.002	0.000	0.000

Figure E.1: Correlation matrix for full flavour tagged fit to data, (part 1)

PARAMETER CORRELATION COEFFICIENTS

	NO.	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
DScale21	17	1.000	-0.032	0.000	0.000	-0.001	-0.003	0.001	0.004	0.002	-0.002	0.001	-0.006	0.000	0.001	0.001	0.000	-0.001	0.000	0.002
deltaG	18	-0.032	1.000	-0.002	-0.002	0.001	0.001	0.004	0.003	0.037	-0.036	0.000	0.020	0.000	-0.007	-0.019	0.001	0.000	-0.014	-0.019
EffSig	19	0.000	-0.002	1.000	-0.093	0.000	-0.001	-0.008	0.002	0.001	0.003	0.000	-0.041	0.000	-0.003	-0.001	0.002	-0.025	-0.008	0.000
EffBkg	20	0.000	-0.002	-0.093	1.000	0.000	-0.005	0.000	-0.001	-0.001	0.000	0.000	-0.013	0.000	-0.001	-0.001	-0.001	-0.003	-0.001	-0.002
EffBkgp	21	-0.001	0.001	0.000	0.000	1.000	0.003	-0.001	0.000	0.000	0.000	0.000	0.002	0.000	0.002	0.002	0.000	0.002	0.002	0.002
EffSig2	22	-0.003	0.001	-0.001	-0.005	0.003	1.000	-0.086	0.000	-0.002	-0.001	0.001	0.050	0.000	0.005	0.005	0.002	0.031	0.007	0.001
EffBkg2	23	0.001	0.004	-0.008	0.000	-0.001	-0.086	1.000	0.000	0.001	-0.002	0.000	0.027	-0.001	0.000	0.001	-0.001	0.007	0.004	0.002
EffBkgp2	24	0.004	0.003	0.002	-0.001	0.000	0.000	0.000	1.000	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.000	0.001	0.000	0.001
cosTheta	25	0.002	0.037	0.001	-0.001	0.000	-0.002	0.001	0.000	1.000	-0.004	0.001	0.002	0.000	0.000	0.001	0.000	-0.001	0.001	0.001
cosPsi1	26	-0.002	-0.036	0.003	0.000	0.000	-0.001	-0.002	0.000	-0.004	1.000	-0.005	-0.011	0.000	0.001	0.002	0.000	-0.002	0.000	0.001
Phi1	27	0.001	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.001	-0.005	1.000	0.004	0.000	0.000	0.000	0.000	-0.001	0.000	0.000
FracSig	28	-0.006	0.020	-0.041	-0.013	0.002	0.050	0.027	0.001	0.002	-0.011	0.004	1.000	-0.005	0.015	0.017	0.002	0.115	0.039	0.015
BGmslope	29	0.000	0.000	0.000	0.000	0.000	0.000	-0.001	0.000	0.000	0.000	0.000	-0.005	1.000	0.001	0.001	0.000	-0.002	0.001	0.001
Lambdapp	30	0.001	-0.007	-0.003	-0.001	0.002	0.005	0.000	0.001	0.000	0.001	0.000	0.015	0.001	1.000	0.692	-0.037	0.560	0.670	0.706
Lambdap	31	0.001	-0.019	-0.001	-0.001	0.002	0.005	0.001	0.000	0.001	0.002	0.000	0.017	0.001	0.692	1.000	-0.019	0.264	0.672	0.851
Lambdam	32	0.000	0.001	0.002	-0.001	0.000	0.002	-0.001	0.000	0.000	0.000	0.000	0.002	0.000	-0.037	-0.019	1.000	0.215	-0.394	0.017
Fgauss	33	-0.001	0.000	-0.025	-0.003	0.002	0.031	0.007	0.001	-0.001	-0.002	-0.001	0.115	-0.002	0.560	0.264	0.215	1.000	0.098	0.148
Fm	34	0.000	-0.014	-0.008	-0.001	0.002	0.007	0.004	0.000	0.001	0.000	0.000	0.039	0.001	0.670	0.672	-0.394	0.098	1.000	0.692
Fpp	35	0.002	-0.019	0.000	-0.002	0.002	0.001	0.002	0.001	0.001	0.001	0.000	0.015	0.001	0.706	0.851	0.017	0.148	0.692	1.000

Figure E.2: Correlation matrix for full flavour tagged fit to data, (part 2)

PARAMETER	NO.	GLOBAL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Lifetime	1	0.79187	1.0	-0.004	-0.003	0.002	0.000	0.000	-0.099	0.055	0.069	0.015	0.047	-0.151	-0.048	-0.035	0.000	0.000	-0.005	-0.775
ScaleFac	2	0.78295	-0.004	1.0	0.598	0.767	0.000	0.000	0.010	-0.009	-0.006	0.014	0.003	0.001	0.027	0.020	-0.013	-0.013	0.074	0.001
ScaleFac2	3	0.87934	-0.003	0.598	1.0	0.872	0.000	0.000	0.005	-0.004	-0.004	0.005	0.002	0.001	0.011	0.008	-0.003	-0.003	0.018	0.000
ScaleFrac	4	0.92465	0.002	0.767	0.872	1.0	0.000	0.000	0.007	-0.006	-0.005	0.008	0.002	0.001	0.016	0.012	-0.005	-0.005	0.028	-0.001
Mean	5	0.00136	0.000	0.000	0.000	0.000	1.0	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MassScl	6	0.00136	0.000	0.000	0.000	0.000	0.001	1.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AlphaCPOdd	7	0.78042	-0.099	0.010	0.005	0.007	0.000	0.000	1.0	-0.720	-0.541	-0.133	-0.001	0.069	0.344	0.213	0.000	0.000	0.015	-0.020
AlphaPara	8	0.73566	0.055	-0.009	-0.004	-0.006	0.000	0.000	-0.720	1.0	0.475	0.117	0.000	-0.008	-0.292	-0.181	-0.001	-0.001	-0.010	0.037
PhiPara	9	0.63179	0.069	-0.006	-0.004	-0.005	0.000	0.000	-0.541	0.475	1.0	0.148	0.001	-0.311	-0.183	-0.167	0.009	0.009	-0.036	0.039
PhiPerp	10	0.66918	0.015	0.014	0.005	0.008	0.000	0.000	-0.133	0.117	0.148	1.0	-0.004	-0.042	0.556	0.395	0.011	0.011	-0.007	0.007
BetaPrime	11	0.06569	0.047	0.003	0.002	0.002	0.000	0.000	-0.001	0.000	0.001	-0.004	1.0	-0.005	-0.004	-0.021	0.005	0.011	0.023	-0.054
ASWave	12	0.38903	-0.151	0.001	0.001	0.001	0.000	0.000	0.069	-0.008	-0.311	-0.042	-0.005	1.0	0.040	0.044	-0.005	-0.005	0.017	0.039
DeltaS	13	0.76580	-0.048	0.027	0.011	0.016	0.000	0.000	0.344	-0.292	-0.183	0.556	-0.004	0.040	1.0	0.602	0.017	0.017	-0.008	0.005
deltaM	14	0.61311	-0.035	0.020	0.008	0.012	0.000	0.000	0.213	-0.181	-0.167	0.395	-0.021	0.044	0.602	1.0	-0.012	-0.013	0.022	0.006
DScale	15	0.91141	0.000	-0.013	-0.003	-0.005	0.000	0.000	0.000	-0.001	0.009	0.011	0.005	-0.005	0.017	-0.012	1.0	0.910	-0.502	0.001
DScale2	16	0.91183	0.000	-0.013	-0.003	-0.005	0.000	0.000	0.000	-0.001	0.009	0.011	0.011	-0.005	0.017	-0.013	0.910	1.0	-0.505	0.001
DScale21	17	0.52269	-0.005	0.074	0.018	0.028	0.000	0.000	0.015	-0.010	-0.036	-0.007	-0.023	0.017	-0.008	0.022	-0.502	-0.505	1.0	0.001
deltaG	18	0.78546	-0.775	0.001	0.000	-0.001	0.000	0.000	-0.020	0.037	0.039	0.007	-0.054	0.039	0.005	0.006	0.001	0.001	0.001	1.0

Figure E.3: Correlation matrix for high statistics pseudo experiment

Appendix F

Full tables of pulls for all fit parameters

Parameter	Pull mean	Pull σ
A_{SW}	0.054 ± 0.03	0.88 ± 0.026
$\alpha_{\perp}$	-0.15 ± 0.033	0.98 ± 0.026
$\alpha_{\parallel}$	0.16 ± 0.032	0.94 ± 0.027
p_1	0.024 ± 0.035	1 ± 0.029
δ_{SW}	2.8 ± 0.21	5.5 ± 0.22
f_p	0.66 ± 0.033	1 ± 0.025
f_-	0.21 ± 0.033	0.93 ± 0.027
f_{++}	0.2 ± 0.033	0.96 ± 0.027
f_s	-0.016 ± 0.032	0.96 ± 0.025
Λ_-	0.029 ± 0.035	1 ± 0.029
Λ_+	0.22 ± 0.032	0.9 ± 0.025
Λ_{++}	0.79 ± 0.033	0.96 ± 0.029
$c\tau$	-0.019 ± 0.032	0.97 ± 0.026
s_m	-0.11 ± 0.032	0.96 ± 0.026
m	0.018 ± 0.033	0.98 ± 0.024
ϕ_1	0.023 ± 0.032	0.97 ± 0.025
$\phi_{\parallel}$	0.25 ± 0.048	0.65 ± 0.066
$s_{c\tau}1$	0.08 ± 0.036	1.1 ± 0.028
$s_{c\tau}2$	0.037 ± 0.034	1 ± 0.026
$f_{sf1}1$	-0.42 ± 0.033	0.99 ± 0.026
$cos(\psi)_1$	0.08 ± 0.034	1 ± 0.028
$cos(\theta)_1$	0.04 ± 0.033	1 ± 0.025
$\Delta\Gamma$	0.037 ± 0.034	0.99 ± 0.026

Table F.1: Pull study fit results for untagged fit with fixed $\beta_s^{J/\psi\phi}$= 0.0, realistic sample size

Parameter	Pull mean	Pull σ
A_{SW}	0.063 ± 0.026	0.75 ± 0.024
$\alpha_\perp$	-0.15 ± 0.029	0.89 ± 0.024
$\alpha_\parallel$	0.14 ± 0.029	0.85 ± 0.024
p_1	0.028 ± 0.033	0.97 ± 0.026
$S_D(SST)$	0.01 ± 0.011	0.33 ± 0.008
$S_D(OST+)$	0.012 ± 0.012	0.35 ± 0.0081
$S_D(OST-)$	0.11 ± 0.012	0.38 ± 0.0098
δ_{SW}	0.11 ± 0.012	0.38 ± 0.0098
$\epsilon_b(OST)$	0.0091 ± 0.035	1 ± 0.029
$\epsilon_b(SST)$	0.03 ± 0.035	1 ± 0.03
$\mathcal{A}^+(OST)$	0.024 ± 0.032	0.92 ± 0.025
$\mathcal{A}^+(SST)$	0.047 ± 0.033	1 ± 0.027
$\epsilon_s(OST)$	-0.03 ± 0.036	1.1 ± 0.029
$\epsilon_s(SST)$	0.0073 ± 0.034	1 ± 0.031
f_p	0.67 ± 0.036	1.1 ± 0.03
f_-	0.21 ± 0.031	0.89 ± 0.027
f_{++}	0.14 ± 0.035	1 ± 0.03
f_s	-0.077 ± 0.034	1 ± 0.026
Λ_-	0.024 ± 0.033	0.99 ± 0.029
Λ_+	0.29 ± 0.03	0.89 ± 0.024
Λ_{++}	0.84 ± 0.034	1 ± 0.032
$c\tau$	0.053 ± 0.033	0.97 ± 0.028
s_m	-0.076 ± 0.033	0.99 ± 0.025
m	-0.025 ± 0.033	1 ± 0.027
ϕ_1	0.025 ± 0.032	0.95 ± 0.025
$\phi_\parallel$	-0.29 ± 0.064	1.1 ± 0.075
$\phi_\perp$	0.0072 ± 0.029	0.85 ± 0.029
$s_{c\tau}1$	0.13 ± 0.035	1 ± 0.029
$s_{c\tau}2$	0.053 ± 0.038	1.1 ± 0.031
$f_{sf1}1$	-0.42 ± 0.034	1 ± 0.03
$cos(\psi)_1$	0.038 ± 0.033	1 ± 0.027
$cos(\theta)_1$	0.0039 ± 0.034	1 ± 0.028
$\Delta\Gamma$	-0.087 ± 0.032	0.97 ± 0.029
Δm_s	0.012 ± 0.0067	0.21 ± 0.0061

Table F.2: Pull study fit results for tagged fit with fixed $\beta_s^{J/\psi\phi}$=0.0, realistic sample size

Parameter	Pull mean	Pull σ
A_{SW}	-0.062 ± 0.031	0.88 ± 0.026
$\alpha_\perp$	-0.19 ± 0.03	0.91 ± 0.025
$\alpha_\parallel$	0.11 ± 0.031	0.94 ± 0.025
p_1	0.0089 ± 0.033	0.98 ± 0.024
β_s	0.33 ± 0.031	0.9 ± 0.033
$S_D(SST)$	0.015 ± 0.01	0.31 ± 0.0075
$S_D(OST+)$	0.003 ± 0.011	0.34 ± 0.0077
$S_D(OST-)$	0.052 ± 0.015	0.43 ± 0.011
δ_{SW}	0.052 ± 0.015	0.43 ± 0.011
$\epsilon_b(OST)$	0.017 ± 0.035	1 ± 0.026
$\epsilon_b(SST)$	0.02 ± 0.033	0.99 ± 0.027
$\mathcal{A}^+(OST)$	0.054 ± 0.03	0.92 ± 0.024
$\mathcal{A}^+(SST)$	0.013 ± 0.033	0.99 ± 0.024
$\epsilon_s(OST)$	-0.011 ± 0.033	0.98 ± 0.025
$\epsilon_s(SST)$	0.0051 ± 0.033	1 ± 0.027
f_p	0.63 ± 0.035	1 ± 0.027
f_-	0.2 ± 0.034	0.98 ± 0.032
f_{++}	0.13 ± 0.034	1 ± 0.029
f_s	-0.04 ± 0.036	1 ± 0.028
Λ_-	0.038 ± 0.034	0.97 ± 0.028
Λ_+	0.2 ± 0.031	0.91 ± 0.025
Λ_{++}	0.83 ± 0.033	0.99 ± 0.029
$c\tau$	-0.22 ± 0.03	0.89 ± 0.025
s_m	-0.083 ± 0.033	0.98 ± 0.025
m	0.038 ± 0.033	0.99 ± 0.026
ϕ_1	0.024 ± 0.032	0.96 ± 0.026
$\phi_\parallel$	-0.15 ± 0.05	1.2 ± 0.058
$\phi_\perp$	0.053 ± 0.041	1.1 ± 0.051
$s_{c\tau}1$	0.088 ± 0.036	1.1 ± 0.029
$s_{c\tau}2$	-0.028 ± 0.034	1 ± 0.028
$f_{sf1}1$	-0.45 ± 0.033	0.98 ± 0.026
$cos(\psi)_1$	-0.016 ± 0.034	0.98 ± 0.026
$cos(\theta)_1$	-0.016 ± 0.034	1 ± 0.027
$\Delta\Gamma$	0.24 ± 0.032	0.94 ± 0.026
Δm_s	0.043 ± 0.013	0.4 ± 0.011

Table F.3: Pull study fit results for tagged fit with fixed $\beta_s^{J/\psi\phi}$floating, realistic sample size

Appendix G

Alternative universe inputs

Parameter	Uni 1	Uni 2	Uni3	Uni 4	Uni 5	Uni 6	Uni 7	Uni 8
$c\tau$	0.0485376	0.0431136	0.0467222	0.0479733	0.0435036	0.0470618	0.0481373	0.0432942
$s_{c\tau}1$	1.33138	1.25479	1.2792	1.29586	1.20609	1.26436	1.28245	1.31733
$s_{c\tau}2$	3.35863	2.66445	4.17486	2.7754	4.13995	4.06452	2.61939	2.77312
$f_{sf}1$	0.92797	0.911962	0.888969	0.896533	0.931649	0.848981	0.857911	0.912806
m	5.36673	5.36661	5.36709	5.36632	5.36603	5.36589	5.36625	5.36697
s_m	1.81726	1.72931	1.62249	1.82807	1.80878	1.65638	1.81114	1.8225
$\alpha_\perp$	0.329356	0.246929	0.208379	0.216926	0.203438	0.311751	0.328833	0.21120
$\alpha_\parallel$	0.30832	0.383089	0.34227	0.238227	0.253503	0.316596	0.32298	0.230097
$\phi_\parallel$	3.10287	1.85547	1.54391	2.36212	2.58462	1.40371	1.77574	2.77451
$\phi_\perp$	4.00734	0.780991	2.89364	5.68146	4.07611	3.19112	2.82485	6.02177
$\beta_s^{J/\psi\phi}$	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
A_{SW}	0.0972382	0.0814032	0.11042	0.0592829	0.0505495	0.0652947	0.12677	0.0764773
δ_{SW}	4.86898	4.49777	1.7644	1.4203	5.61291	4.68301	3.98209	0.561049
$\Delta\Gamma$	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096
$cos(\psi)_1$	0.210696	0.13839	0.089126	0.0920404	0.184946	0.211708	0.0774938	0.12088
$cos(\theta)_1$	-0.0737035	-0.0127692	0.0238879	-0.0608936	0.04536	0.018862	0.0687134	0.0629707
ϕ_1	0.17132	0.134837	0.179712	0.151433	0.108297	0.154011	0.154856	0.114698
f_s	0.182083	0.173316	0.179497	0.18552	0.188161	0.190191	0.190268	0.185348
p_1	-3.51058	-4.56139	0.0969551	-0.462068	-1.45836	-0.990054	0.32942	-1.60436
Λ_-	0.0166336	0.00665507	0.0115273	0.016374	0.00736725	0.00726708	0.00869923	0.013702
Λ_+	0.044876	0.0283433	0.0438055	0.030544	0.0488414	0.0435895	0.0585809	0.0539351
Λ_{++}	0.0525794	0.0541871	0.0314057	0.041783	0.0203742	0.0364937	0.0352796	0.0235069
f_p	0.890572	0.888271	0.892361	0.89066	0.897624	0.85977	0.9049	0.908599
f_-	0.156703	0.241715	0.102642	0.130044	0.0854399	0.196921	0.130275	0.076429
f_{++}	0.695682	0.516151	0.425661	0.695347	0.630538	0.559635	0.858216	0.551804
Parameter	Uni 9	Uni 10	Uni 11	Uni 12	Uni 13	Uni 14	Uni 15	Uni 16
$c\tau$	0.0474013	0.0490635	0.0430629	0.0442995	0.0466586	0.04654	0.0467884	0.0476583
$s_{c\tau}1$	1.23399	1.22474	1.26654	1.29705	1.26115	1.29173	1.25797	1.32562
$s_{c\tau}2$	4.03696	3.04103	2.84346	4.23455	3.58774	3.0222	2.60447	4.18949
$f_{sf}1$	0.905735	0.913376	0.857147	0.863108	0.927531	0.9226	0.860075	0.832762
m	5.36586	5.36705	5.36661	5.36651	5.36678	5.36643	5.36702	5.36611
s_m	1.6902	1.81882	1.78622	1.70577	1.75073	1.71912	1.67429	1.76842
$\alpha_\perp$	0.245307	0.211084	0.265799	0.304696	0.324884	0.265	0.283612	0.286257
$\alpha_\parallel$	0.389162	0.286739	0.297919	0.384272	0.328809	0.2719	0.349935	0.304067
$\phi_\parallel$	2.12479	1.74204	2.28866	2.09906	2.12609	2.47416	2.05691	2.14135
$\phi_\perp$	3.0328	5.67102	3.39944	1.30452	2.27201	0.898723	0.454564	5.41598
$\beta_s^{J/\psi\phi}$	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
A_{SW}	0.081127	0.12896	0.138608	0.0482121	0.0925161	0.068460	0.0413977	0.116351
δ_{SW}	3.80558	5.32683	3.49857	2.8309	1.45382	5.61932	0.532978	5.19981
$\Delta\Gamma$	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096
$cos(\psi)_1$	0.134106	0.122312	0.109263	0.133062	0.154713	0.131	0.108625	0.08374
$cos(\theta)_1$	-0.00335732	0.0419133	0.0314992	0.0529829	0.0148029	-0.0011	-0.0472706	0.0471503
ϕ_1	0.135411	0.140905	0.138205	0.101471	0.152161	0.121737	0.136595	0.118178
f_s	0.174087	0.187647	0.176076	0.191371	0.178038	0.17465	0.172549	0.179677
p_1	-3.00459	-3.21415	-2.91335	-0.942078	-0.798538	-0.928	-1.43536	-2.17614
Λ_-	0.00538458	0.014377	0.00531193	0.00753525	0.00973217	0.0118	0.00900288	0.00948485
Λ_+	0.0514337	0.0528967	0.0321764	0.0495291	0.0513255	0.03454	0.0469498	0.0415923
Λ_{++}	0.0553435	0.0232477	0.0245135	0.0549706	0.0481597	0.03701	0.0234051	0.049595
f_p	0.874235	0.881585	0.891028	0.90442	0.873208	0.87696	0.891969	0.865252
f_-	0.0625045	0.0874061	0.156166	0.30064	0.141432	0.273411	0.0811882	0.182302
f_{++}	0.6193	0.791976	0.817742	0.471797	0.745994	0.754898	0.659251	0.603105

Table G.1: Randomised inputs for generating coverage adjustment pseudo experiments in alternative universes for the fit without flavour tagging

Parameter	Uni 1	Uni 2	Uni3	Uni 4	Uni 5	Uni 6	Uni 7	Uni 8
$c\tau$	0.0432173	0.0466444	0.0467537	0.0468813	0.0477369	0.0432672	0.0468162	0.0471502
$s_{c\tau}1$	1.26526	1.28996	1.25998	1.25686	1.32318	1.25374	1.27767	1.26312
$s_{c\tau}2$	2.85197	3.01888	3.547	2.6288	4.10893	2.68481	4.09527	3.99224
$f_{sf}1$	0.855979	0.920072	0.924829	0.858844	0.832126	0.909599	0.887107	0.847992
m	5.36664	5.36647	5.36682	5.36706	5.36615	5.36665	5.36713	5.36593
s_m	1.79724	1.73028	1.76182	1.68554	1.77948	1.74045	1.63384	1.66767
$\alpha_\perp$	0.266234	0.2664	0.318902	0.282112	0.28447	0.249413	0.21505	0.307195
$\alpha_\parallel$	0.296112	0.272845	0.323776	0.342694	0.301618	0.372385	0.33583	0.312838
$\phi_\parallel$	2.58851	2.72587	2.46811	2.41688	2.47941	2.26771	2.03699	1.93317
$\phi_\perp$	3.35407	1.29838	2.42728	0.933268	5.01174	1.2016	2.93828	3.18282
$\beta_s^{J/\psi\phi}$	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
A_{SW}	0.125852	0.0677486	0.0876738	0.0453329	0.107416	0.0784691	0.102503	0.0651266
δ_{SW}	2.37472	2.17957	0.246557	3.7807	1.14988	2.69263	1.51965	0.578822
Δm_s	17.4203	17.558	17.6972	17.4165	17.2649	17.5978	17.2977	18.0444
$S_D(SST)$	1.04103	0.903192	0.970613	0.708831	1.10703	0.854333	1.00893	0.987731
$S_D(OST+)$	1.12093	0.940915	1.27348	1.10333	0.90202	1.08411	1.0665	1.29371
$S_D(OST-)$	0.707744	1.31954	0.685867	1.1632	0.805173	1.02479	0.928644	1.30043
$\Delta\Gamma$	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096
$\epsilon_s(SST)$	0.926215	0.944044	0.951546	0.936306	0.943329	0.927102	0.945549	0.93156
$\epsilon_b(SST)$	0.864963	0.866254	0.866972	0.873009	0.866569	0.874013	0.866101	0.871179
$\mathcal{A}^+(SST)$	0.48363	0.493629	0.50396	0.482743	0.495506	0.485507	0.499488	0.493214
$\epsilon_s(OST)$	0.530538	0.512053	0.53882	0.531774	0.541272	0.536803	0.508727	0.489466
$\epsilon_b(OST)$	0.71848	0.727954	0.71352	0.712421	0.709386	0.716825	0.706522	0.721773
$\mathcal{A}^+(OST)$	0.508251	0.503617	0.493305	0.496565	0.497588	0.488098	0.488457	0.489221
$cos(\theta)_1$	0.17098	0.230247	0.195493	0.088124	0.142495	0.153228	0.21401	0.204084
$cos(\psi)_1$	0.0564802	-0.0371032	-0.0501571	0.0280545	0.0695549	0.0149087	0.0990028	-0.0666806
ϕ_1	0.115723	0.113155	0.118409	0.13401	0.104673	0.134433	0.174716	0.161989
f_s	0.176872	0.171108	0.187184	0.183603	0.176961	0.170759	0.174753	0.171462
p_1	-3.95843	0.272108	-0.646885	-2.29218	-2.585	-3.1459	-4.61713	-4.75646
Λ_-	0.00547831	0.0143018	0.0134633	0.0122385	0.00917823	0.0140637	0.00963167	0.016705
Λ_+	0.0273634	0.041269	0.0344873	0.0312654	0.0312548	0.0479791	0.0321409	0.040338
Λ_{++}	0.047571	0.0203809	0.0454564	0.0422742	0.0352785	0.0334592	0.0363474	0.0234425
f_p	0.901339	0.876286	0.900666	0.892677	0.88786	0.886915	0.876512	0.900318
f_-	0.173189	0.223673	0.214417	0.320843	0.0504921	0.249821	0.130487	0.207871
f_{++}	0.740545	0.539799	0.450041	0.849394	0.637548	0.732377	0.709705	0.84869
Parameter	Uni 9	Uni 10	Uni 11	Uni 12	Uni 13	Uni 14	Uni 15	Uni 16
$c\tau$	0.0482079	0.0474841	0.0481794	0.0473766	0.0457722	0.0480692	0.0470152	0.0438284
$s_{c\tau}1$	1.28086	1.23334	1.20887	1.23675	1.32948	1.28343	1.28782	1.26029
$s_{c\tau}2$	2.64274	3.9665	3.88049	3.7355	3.55683	3.67778	3.22062	3.43151
$f_{sf}1$	0.856727	0.903508	0.89764	0.851654	0.889603	0.871329	0.922651	0.917811
m	5.36629	5.3659	5.36603	5.36711	5.36647	5.36589	5.36621	5.36648
s_m	1.82211	1.70141	1.7232	1.79643	1.83109	1.67919	1.83897	1.76357
$\alpha_\perp$	0.322422	0.247967	0.196461	0.292146	0.275761	0.325867	0.224887	0.28921
$\alpha_\parallel$	0.318562	0.377824	0.285819	0.245541	0.316605	0.374275	0.239617	0.236353
$\phi_\parallel$	2.20867	2.46715	2.05542	1.94557	2.16218	2.77284	2.40864	2.11724
$\phi_\perp$	2.88173	3.05268	1.64	3.97415	0.627106	3.87774	0.458077	4.73695
$\beta_s^{J/\psi\phi}$	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
A_{SW}	0.116046	0.0782403	0.0400511	0.0546117	0.0502018	0.032305	0.096285	0.0499373
δ_{SW}	4.37883	2.45647	2.58395	1.4591	4.85068	2.5789	0.833799	5.89658
Δm_s	17.2268	17.5717	18.069	17.2466	18.2165	17.2011	17.8135	18.0495
$S_D(SST)$	1.19797	0.894026	0.794701	0.92357	0.799718	0.969389	0.790846	0.883213
$S_D(OST+)$	1.30294	1.09039	0.976481	0.89166	1.40406	1.16789	0.796768	1.22927
$S_D(OST-)$	0.767956	0.679629	1.30754	0.848395	0.994162	0.718446	0.823849	1.20071
$\Delta\Gamma$	0.096	0.096	0.096	0.096	0.096	0.096	0.096	0.096
$\epsilon_s(SST)$	0.955862	0.926414	0.946102	0.955104	0.957868	0.944301	0.948173	0.957621
$\epsilon_b(SST)$	0.867326	0.875451	0.870614	0.872044	0.868455	0.865682	0.872922	0.879198
$\mathcal{A}^+(SST)$	0.508537	0.508293	0.507208	0.481608	0.487673	0.499392	0.50279	0.50379
$\epsilon_s(OST)$	0.531218	0.508472	0.527043	0.515008	0.518818	0.524868	0.501134	0.503744
$\epsilon_b(OST)$	0.717395	0.710911	0.731714	0.706675	0.712931	0.71424	0.724085	0.719939
$\mathcal{A}^+(OST)$	0.512965	0.49362	0.508861	0.497643	0.491116	0.483188	0.481367	0.511425
$cos(\theta)_1$	0.176008	0.179567	0.179731	0.104078	0.139358	0.161903	0.165451	0.204246
$cos(\psi)_1$	-0.0013416	0.0747247	0.0280512	0.0680616	-0.017734	0.0499583	-0.0448446	-0.0023898
ϕ_1	0.159026	0.159728	0.154712	0.154244	0.179041	0.153657	0.167918	0.126116
f_s	0.172133	0.170747	0.180496	0.173379	0.174856	0.179267	0.172262	0.18118
p_1	-1.26589	0.240422	-1.94293	-4.70976	-2.44748	-2.24152	-3.99001	-2.99492
Λ_-	0.00639135	0.0163105	0.0077308	0.014159	0.0124066	0.0156009	0.0120454	0.0129415
Λ_+	0.0229175	0.0446909	0.0472462	0.0234318	0.0248084	0.0276943	0.0481009	0.0341138
Λ_{++}	0.0321981	0.0553537	0.020261	0.030657	0.0428638	0.03744	0.0211797	0.050413
f_p	0.863925	0.864661	0.881883	0.86175	0.89634	0.893635	0.858087	0.893034
f_-	0.29982	0.161753	0.164445	0.123794	0.257845	0.322474	0.274869	0.28668
f_{++}	0.4921	0.627799	0.61564	0.569478	0.771953	0.453613	0.825463	0.729675

Table G.2: Randomised inputs for generating coverage adjustment pseudo experiments representing 16 alternative universes for the fit including flavour tagging

Parameter	Uni 1	Uni 2	Uni3	Uni 4	Uni 5	Uni 6	Uni 7	Uni 8
$c\tau$	0.0425647	0.0465308	0.0466573	0.046805	0.0477952	0.0426223	0.0467296	0.0471162
$s_{c\tau}1$	1.26541	1.29109	1.25992	1.25668	1.32562	1.25344	1.27831	1.26319
$s_{c\tau}2$	2.82025	2.99949	3.56662	2.5806	4.17005	2.64074	4.15538	4.04474
$f_{sf1}1$	0.854522	0.92255	0.9276	0.857562	0.829205	0.911434	0.887562	0.846044
m	5.36662	5.36645	5.3668	5.36704	5.36613	5.36663	5.36711	5.36591
s_m	1.79734	1.73037	1.76192	1.68562	1.77958	1.74054	1.63391	1.66774
$\alpha_\perp$	0.26786	0.26803	0.321544	0.284045	0.286448	0.250715	0.215689	0.309612
$\alpha_\parallel$	0.295156	0.271668	0.323083	0.342181	0.300714	0.372154	0.335252	0.312041
$\phi_\parallel$	1.89213	2.24959	1.57884	1.44553	1.60825	1.05735	0.456957	0.186771
$\phi_\perp$	3.40801	0.974674	2.31096	0.542481	5.37022	0.860114	2.91584	3.20531
$\beta_s^{J/\psi\phi}$	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
A_{SW}	0.147781	0.0765247	0.100961	0.0490344	0.125172	0.0896722	0.119147	0.0733091
δ_{SW}	3.93499	5.30209	4.70837	4.73228	6.17646	5.69734	5.52594	5.96337
Δm_s	17.414	17.554	17.6955	17.41	17.2559	17.5944	17.2892	18.0485
$S_D(SST)$	1.04098	0.903156	0.970572	0.708809	1.10698	0.854301	1.00888	0.98769
$S_D(OST+)$	1.12075	0.939993	1.27394	1.10308	0.900937	1.08378	1.0661	1.29424
$S_D(OST-)$	0.705918	1.32384	0.683823	1.16594	0.804323	1.02614	0.92903	1.30454
$\Delta\Gamma$	-0.0942111	0.0912517	0.169297	0.0107584	0.0838168	-0.0849825	0.106911	-0.0386071
$\epsilon_s(SST)$	0.934924	0.937089	0.938293	0.948417	0.937617	0.9501	0.936832	0.945348
$\epsilon_b(SST)$	0.862643	0.868914	0.875393	0.862087	0.870091	0.86382	0.872588	0.868653
$\mathcal{A}^+(SST)$	0.498675	0.490018	0.502553	0.499253	0.503702	0.501609	0.48846	0.479439
$\epsilon_s(OST)$	0.519667	0.543626	0.507125	0.504345	0.496669	0.515483	0.489426	0.527995
$\epsilon_b(OST)$	0.728815	0.725242	0.717289	0.719803	0.720592	0.713274	0.713551	0.714139
$\mathcal{A}^+(OST)$	0.498093	0.509872	0.502964	0.481621	0.492429	0.494562	0.506645	0.504672
$cos(\theta)_1$	0.206751	0.124798	0.113366	0.181858	0.218201	0.170346	0.243989	0.0988962
$cos(\psi)_1$	-0.0515928	-0.0577678	-0.0451335	-0.00761315	-0.0781663	-0.00659732	0.0902824	0.0596752
ϕ_1	0.124393	0.103914	0.161027	0.148306	0.124709	0.102674	0.116865	0.105174
f_s	0.173982	0.191765	0.187902	0.180986	0.179755	0.177398	0.171214	0.170628
p_1	-5.02064	-0.812301	-1.21221	-1.79638	-3.25598	-0.925886	-3.03971	0.333896
Λ_-	0.0068766	0.0113757	0.00918152	0.00813911	0.00813568	0.0135467	0.00842238	0.0110745
Λ_+	0.0494389	0.0238069	0.0474454	0.0444457	0.0378508	0.0361358	0.0388585	0.0266931
Λ_{++}	0.0516094	0.0326157	0.0510995	0.0450426	0.0413907	0.0406741	0.0327876	0.0508354
f_p	0.883785	0.891695	0.890245	0.90692	0.86456	0.895792	0.877094	0.889219
f_-	0.230767	0.0830187	0.0980848	0.310879	0.154962	0.224755	0.208069	0.310361
f_{++}	0.606248	0.592195	0.810423	0.707492	0.518047	0.62914	0.544675	0.476926
Parameter	Uni 9	Uni 10	Uni 11	Uni 12	Uni 13	Uni 14	Uni 15	Uni 16
$c\tau$	0.0483404	0.0475026	0.0483073	0.0473782	0.0455214	0.0481797	0.0469599	0.0432719
$s_{c\tau}1$	1.28163	1.23224	1.2068	1.23578	1.33216	1.2843	1.28885	1.26024
$s_{c\tau}2$	2.59556	4.0171	3.92474	3.76904	3.57717	3.70706	3.21613	3.4426
$f_{sf1}1$	0.855315	0.904969	0.898742	0.849931	0.89021	0.870815	0.925288	0.920151
m	5.36627	5.36588	5.36601	5.36709	5.36645	5.36587	5.36619	5.36646
s_m	1.82222	1.7015	1.72328	1.79654	1.8312	1.67927	1.83908	1.76367
$\alpha_\perp$	0.325132	0.249241	0.196741	0.294272	0.277571	0.328644	0.225716	0.29128
$\alpha_\parallel$	0.31782	0.377644	0.284766	0.244105	0.315844	0.374062	0.238125	0.23483
$\phi_\parallel$	0.903694	1.57634	1.55339	1.40225	0.782724	2.37183	1.42408	0.665793
$\phi_\perp$	2.84891	3.05125	1.11879	0.344115	0.180073	4.02789	-0.0200077	5.04496
$\beta_s^{J/\psi\phi}$	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
A_{SW}	0.135756	0.0893916	0.0425568	0.0604138	0.0550055	0.0330571	0.111521	0.0546812
δ_{SW}	2.74387	2.43678	0.557664	4.26468	4.03498	5.36315	0.833799	1.61601
Δm_s	17.2172	17.5678	18.0735	17.2373	18.2235	17.191	17.8138	18.0538
$S_D(SST)$	1.19791	0.893991	0.794672	0.923533	0.79969	0.969349	0.790818	0.883179
$S_D(OST+)$	1.30352	1.09008	0.975706	0.890534	1.40505	1.16791	0.795249	1.22954
$S_D(OST-)$	0.766733	0.677521	1.31172	0.847977	0.995204	0.716727	0.823186	1.20382
$\Delta\Gamma$	0.214188	-0.0921448	0.11266	0.206307	0.235054	0.0939305	0.134204	0.232485
$\epsilon_s(SST)$	0.938887	0.952512	0.9444	0.946798	0.94078	0.93613	0.948271	0.958795
$\epsilon_b(SST)$	0.878263	0.87811	0.87743	0.861375	0.865179	0.872528	0.874659	0.875286
$\mathcal{A}^+(SST)$	0.498993	0.488341	0.497038	0.491402	0.493186	0.49602	0.484904	0.486126
$\epsilon_s(OST)$	0.516924	0.500527	0.553133	0.489815	0.505636	0.508945	0.53384	0.523356
$\epsilon_b(OST)$	0.732451	0.717532	0.729286	0.720634	0.715601	0.709487	0.708082	0.731263
$\mathcal{A}^+(OST)$	0.499091	0.499798	0.499831	0.484792	0.491805	0.496287	0.496992	0.504704
$cos(\theta)_1$	0.156115	0.222728	0.181855	0.216893	0.14176	0.20104	0.118018	0.155197
$cos(\psi)_1$	0.0525496	0.0542359	0.0421743	0.0410489	0.100682	0.0396354	0.0739332	-0.0265977
ϕ_1	0.107556	0.102631	0.137266	0.111983	0.117232	0.132902	0.108016	0.139697
f_s	0.1853	0.191632	0.182454	0.170824	0.180333	0.181199	0.17385	0.178032
p_1	-4.58517	0.145727	-3.94632	-0.88039	-1.71622	-0.192702	-1.88847	-1.46111
Λ_-	0.00543823	0.0124828	0.0133096	0.00560464	0.00605002	0.00698371	0.0135861	0.0090607
Λ_+	0.034947	0.0567756	0.0236939	0.0334942	0.0450014	0.0398885	0.02456	0.0521181
Λ_{++}	0.0232445	0.0238023	0.0368595	0.021596	0.0478199	0.0457686	0.0188186	0.0453135
f_p	0.903626	0.881993	0.882415	0.876045	0.897049	0.907175	0.857797	0.901567
f_-	0.0479121	0.147787	0.138837	0.104862	0.253883	0.142571	0.275668	0.222767
f_{++}	0.750563	0.612134	0.621314	0.540314	0.784541	0.848978	0.824559	0.859856

Table G.3: Randomised inputs for generating coverage adjustment pseudo experiments representing 16 alternative universes for the fit including flavour tagging for 1D $\beta_s^{J/\psi\phi}$scan

Acknowledgments

I would like to thank all of the CDF collaborators who have helped me so much in my work at Fermilab and in Oxford, in particular my supervisor, Farrukh Azfar, for his guidance, expertise, patience and sense of humour throughout the course of my PhD. I'm grateful to my colleagues in the $\beta_s^{J/\psi\phi}$ analysis group and would like to acknowledge the essential role in the work presented here of Farrukh, Joe Boudreau, Karen Gibson, Gavril Giurgiu, Manfred Paulini, Michal Kreps, Thomas Kuhr, Jan Morlock; and most of all Elisa Pueschel, whose hard work and friendship made it possible to get to this stage.

To name just a few others from the collaboration I'd like to express my gratitude to: Sneha Malde, Nicola Pounder and Todd Huffman for their help when I started out both in Oxford and Chicago; Sam Harper for his detailed introduction to CDF software; Alberto Annovi, Luciano Ristori and Stefano Torre for detailed explanations of the SVT beamline service work; Diego Tonelli for his clear teaching as we worked on the dE/dx calibrations and careful scrutiny as B group convener for the $\beta_s^{J/\psi\phi}$ analysis; and all of the Drug Sniffing Dogs for the great gigs.

Outside of CDF, I am very pleased to have had the opportunity during my time at Fermilab to play with the Chicago Brass Band, and enjoy the friendship and musical talents of Brad, Tab, Colin, and too many others to mention; I could not have asked for a better group of friends and musicians to meet when I moved to the US. Throughout my life, Jane Phipps and Steve Petch have always provided academic inspiration. My parents have given endless encouragement but never pressure, for which I am grateful. My friends in England have put up with my long absences and stayed in contact, I hope to see them all more often in the next few years than in the last. Finally, I'd like to thank Ingo Bloch, who has constantly inspired, supported and encouraged me through the best and worst times.

List of Tables

List of Figures

References

[1] A. Sakharov. "Violation of CP invariance, C asymmetry and baryon asymmetry of the Universe", *JETP Lett.*, **5** (1967) 32–35.

[2] J R Fry. "CP violation and the standard model", *Reports on Progress in Physics*, **63**, 2 (2000) 117.

[3] M. Bona et al. "First Evidence of New Physics in $b \leftrightarrow s$ Transitions", *PMC Phys.*, **A3** (2009) 6.

[4] T. et al Aaltonen. "First Flavor-Tagged Determination of Bounds on Mixing-Induced CP Violation in $B_s \to J/\psi\phi$ Decays", *Phys. Rev. Lett.*, **100** (2008) 161802.

[5] V. M. Abazov et al (DØ collaboration). "Lifetime Difference and CP-Violating Phase in the Bs System", *Phys. Rev. Lett.*, **98**, 12 (2007) 121801.

[6] CDF and DØ$\Delta\Gamma_s, \beta_s$ Combination Working Group, "Combination of DØ and CDF Results on $\Delta\Gamma_s$ and the CP-Violating Phase β_s", CDF public note 9787.

[7] The CDF Collaboration, "An Updated Measurement of the CP Violating Phase β_s using $B_s^0 \to J/\psi\,\phi$ Decays", CDF Public Note 9458.

[8] J.J. Thomson. "On the Structure of the Atom", *Philosophical Magazine*, **6-7** (1904) 237–265.

[9] E. Rutherford. "The scattering of α and β particles and the Structure of the Atom", *Philosophical Magazine*, **6-21** (1911) 669–689.

[10] (Particle Data Group) C. Amsler et al. "The Review of Particle Physics", *Phys. Lett.*, **B667**.

[11] LHC, `http://lhc.web.cern.ch/lhc/`.

[12] D. Griffiths. "Introduction to Elementary Particles, 2nd Edition". Wiley-VCH (2008).

[13] A.I. Sanda I.I. Bigi. "CP Violation, 2nd Edition". Cambridge University Press (2009).

[14] M. Gell-Mann. "Symmetries of Baryons and Mesons", *Phys. Rev.*, **125-3** (1962) 1067.

[15] J.E. Augustin et al. "Discovery of a Narrow Resonance in $e+e-$ Annihilation", *Phys. Rev. Lett.*, **33**, 23 (1974) 1406–1408.

[16] J.J. Aubert et al. "Experimental Observation of a Heavy Particle J", *Phys. Rev. Lett.*, **33**, 23 (1974) 1404–1406.

[17] M. Kobayashi & T. Maskawa. "CP-violation in the renoralizable theory of weak interaction", *Prog. of Th. Phys*, **49**, 2 (1973) 652–657.

[18] S.W. Herb et al. "Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions", *Phys. Rev. Lett.*, **39**, 5 (1977) 252–255.

[19] F. Abe et al. "Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab", *Phys. Rev. Lett.*, **74**, 14 (1995) 2626–2631.

[20] W.S. Hou et al. "Large Time-dependent CP Violation in Bs0 System and Finite D0-D0bar Mass Difference in Four Generation Standard Mode", *Phys. Rev. D*, **76** (2007) 016004.

[21] D.H. Perkins. "Introduction to High Energy Physics, 4th Edition". Cambridge University Press (2000).

[22] F. Halzen & D. Martin. "Quarks and Leptons". Wiley (1984).

[23] T.D. Lee & C. N. Yang. "Question of Parity Conservation in Weak Interactions", *Phys. Rev.*, **104**, 1 (1956) 254–258.

[24] N. Cabibbo. "Unitary symmetry and leptonic decays", *Phys. Rev. Lett*, **10**, 12.

[25] S.L. Glashow et al. "Weak Interactions with Lepton-Hadron Symmetry", *Phys. Rev. D*, **2**, 7 (1970) 1285–1292.

[26] L. Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix", *Phys. Rev. Lett.*, **51**, 21 (1983) 1945–1947.

[27] A. Ceccucci et al, "The CKM quark-mixing matrix", PDG review (2008).

[28] C. Jarlskog et al. "Unitarity polygons and CP violation areas and phases in the standard electroweak model", *Physics Letters B*, **208**, 2 (1988) 268 – 274.

[29] Y. Kwon et al,"Production and decay of B flavored hadrons", PDG review (2009).

[30] M. Paulini, "Properties of heavy B hadrons", *Proceedings of 34th ICHEP* (2008) arXiv:hep-ex/0904.2406.

[31] D. Kirby and Y. Nir,"CP violation in meson decays", PDG review (2009).

[32] P. Ball & R. Fleischer. "An analysis of Bs decays in the left-right-symmetric model with spontaneous CP violation", *Physics Letters B*, **475**, 1-2 (2000) 111.

[33] Isard Dunietz. "$B_s - \bar{B}_s$ mixing, CP violation and extraction of CKM phases from untagged B_s data samples", *Phys. Rev.*, **D52** (1995) 3048–3064.

[34] A. Gell-Mann, M. Pais. "Behavior of Neutral Particles under Charge Conjugation", *Phys. Rev.*, **97**, 5 (1955) 1387–1389.

[35] J. H. et al Christenson. "Evidence for the 2π Decay of the $K2$ Meson", *Phys. Rev. Lett.*, **13**, 4 (1964) 138–140.

[36] B. Aubert et al. "Evidence for $D - \bar{D}$ Mixing", *Phys. Rev. Lett.*, **98**, 21 (2007) 211802.

[37] T. Aaltonen et al. "Evidence for $D - \bar{D}$ Mixing Using the CDF II Detector", *Phys. Rev. Lett.*, **100**, 12 (2008) 121802.

[38] D. Asner, "D0-D0 bar mixing", PDG review (2008).

[39] Heavy Flavour Averaging Group, `http://www.slac.stanford.edu/xorg/hfag/`.

[40] C. Albajar et al. "Search for $B^0 - \bar{B}^0$ oscillations at the CERN $p\bar{p}collider$", *Phys. Lett.*, **B186** (1987) 247.

[41] Barberio, E. et al (HFAG), "Averages of $b-$hadron properties at the end of 2006", FERMILAB-FN-0815-E (2007) arXiv:hep-ex/0704.3575.

[42] A. Abulencia et al. "Observation of $B_s^0 - \bar{B}_s^0$ Oscillations", *Phys. Rev. Lett.*, **97** (2006) 242003.

[43] O. Schneider, "B-B bar mixing", PDG review (2009).

[44] A. Lenz & U. Nierste. "Theoretical update of $B_s^0 - \bar{B}_s^0$ mixing", *J. HEP*, **06** (2006) 072.

[45] A.S. Dighe et al. "Extracting CKM phases and B/s anti-B/s mixing parameters from angular distributions of non-leptonic B decays", *Eur. Phys. J. C*, **6** (1999) 647–662.

[46] S. Stone & L. Zhang. "S-waves and the Measurement of CP Violating Phases in B_s Decays", *Phys. Rev. D*, **79** (2009) 074024.

[47] Y. Xie et al. "Determination of $2\beta_s$ in $B_s \to J/\psi K^+K^-$ Decays in the Presence of a K^+K^- S-Wave Contribution", *JHEP*, **09** (2009) 074.

[48] K. Abe et al (Belle Collaboration). "Measurements of branching fractions and decay amplitudes in $B \to J/\psi K*$ decays", *Physics Letters B*, **538**, 1-2 (2002) 11 – 20.

[49] B. Aubert et al (BaBar collaboration). "Measurement of the $B \to J/\psi K*(892)$ Decay Amplitudes", *Phys. Rev. Lett.*, **87**, 24 (2001) 241801.

[50] G. DiGiovanni et al, "Angular Analysis of $Bs- > J/\psi\phi$ and $B_d- > J/\psi K*$ Decays and Measurement of $\Delta\Gamma_s$ and ϕ_s", CDF public note 7538.

[51] V. M. Abazov et al (DØ collaboration). "Measurement of the Angular and Lifetime Parameters of the Decays $Bd \to J/\psi K*$ and $Bs \to J/\psi\phi$", *Phys. Rev. Lett.*, **102**, 3 (2009) 032001.

[52] The DØ collaboration, "DØ results on $\Delta\Gamma_s$ vs. ϕ_s", DØ conference note 5933.

[53] DØ, `http://www-d0.fnal.gov/`.

[54] C. Gattuso et al,"Concepts Rookie Book", (2010), http://www-bdnew.fnal.gov/operations/rookie_books/Concepts_v3.6.pdf.

[55] B. Worthal et al,"Linac Rookie Book", (2004), http://www-bdnew.fnal.gov/operations/rookie_books/LINAC_v2.pdf.

[56] B. Worthal et al,"Booster Rookie Book", (1998), http://www-bdnew.fnal.gov/operations/rookie_books/Booster_V4.1.pdf.

[57] T. Asher et al,"Main Injector Rookie Book", (2010), http://www-bdnew.fnal.gov/operations/rookie_books/Main_Injector_v1.1.pdf.

[58] J. Morgan et al,"Antiproton Source Rookie Book", (2009), http://www-bdnew.fnal.gov/operations/rookie_books/Pbar_v2.pdf.

[59] S. van der Meer. "Stochastic cooling and the accumulation of antiprotons", *Rev. Mod. Phys*, **57** (1985) 689–697.

[60] The Nobel Prize in Physics 1984, http://nobelprize.org/nobel_prizes/physics/laureates/1984/.

[61] E. Malamud"Accelerator Details: Antiproton Source", (2000), http://www-bd.fnal.gov/public/antiproton.html.

[62] B. Worthal et al,"Tevatron Rookie Book", (2009), http://www-bdnew.fnal.gov/operations/rookie_books/Tevatron_v2.3.pdf.

[63] "Tevatron Luminosity", (2010), http://www.fnal.gov/pub/now/tevlum.html.

[64] The CDF collaboration,"Chapter 1: Overview", CDF Run II TDR (1996), FERMILAB-Pub-96/390-E.

[65] F. Abe et al. "The CDF detector: an overview", *Nucl. Inst. Meth.*, **A271** (1988) 387–403.

[66] S. Worm, "The SXV II Silicon Vertex Detector at CDF", FERMILAB-CONF-96/303-E.

[67] Christopher S. Hill. "Initial experience with the CDF layer 00 silicon detector", *Nucl. Inst. Meth.*, **A511**, 1-2 (2003) 118 – 120. Proceedings of the 11th International Workshop on Vertex Detectors.

[68] T. Affolder et al. "CDF Central Outer Tracker", *Nucl. Inst. Meth.*, **A526**, 3 (2004) 249 – 299.

[69] D. Ambrose. "The Central Outer Tracker for Run II at CDF", *Nucl. Inst. Meth.*, **A518**, 1-2 (2004) 117 – 120.

[70] H. Minemura et al. "Fabrication of a 3 m x 5 m superconducting solenoid for the Fermilab Colider Detector Facility", *Journ. Phys.*, **45** (1984) C1–333.

[71] The CDF collaboration,"Chapter 8: Central Calorimeter", CDF Run II TDR (1996), FERMILAB-Pub-96/390-E.

[72] The CDF collaboration,"Chapter 8: Plug Calorimeter", CDF Run II TDR (1996), FERMILAB-Pub-96/390-E.

[73] The CDF collaboration,"Chapter 10: Muon Detectors", CDF Run II TDR (1996), FERMILAB-Pub-96/390-E.

[74] D. Acosta et al. "The CDF Cherenkov luminosity monitor", *Nucl. Inst. Meth.*, **A461**, 1-3 (2001) 540 – 544.

[75] The CDF collaboration,"Chapter 12: Trigger", CDF Run II TDR (1996), FERMILAB-Pub-96/390-E.

[76] E.J. Thomson et al. "Online Track Processor for the CDF upgrade", *IEEE Trans. Nucl. Sci.*, **49**, 3 (2002) 1063.

[77] B. Ashmanskas et al. "The CDF Silicon Vertex Trigger", *Nucl. Inst. and Meth.*, **A518**, 1-2 (2004) 532 – 536. Frontier Detectors for Frontier Physics: Proceeding.

[78] Y.S. Chung et al. "The Level-3 Trigger at the CDF Experiment at Tevatron Run II", *Nuclear Science, IEEE Transactions on*, **52**, 4 (2005) 1212 – 1216.

[79] K. Burkett et al, "CTVMFT", CDF note 6430.

[80] F. Azfar et al, "An updated measurement of the CP violating phase β_s in $B_s \to J/\psi\phi$ decays using L=5.2fb^{-1}", CDF note 10053.

[81] G. Giurgiu et al, "Muon B flavour tagging, a likelihood approach", CDF note 7043.

[82] S-S Yu et al, "COT dE/dx Measurement and corrections", CDF note 6361.

[83] "COT Track Based Calibrations for Summer Conferences", (2004), `http://www-cdf.fnal.gov/internal/physics/bottom/COTDEDX/COT_dedx_cali.html`.

[84] F.Azfar, L. Oakes & D.Tonelli, "Extended and improved dE/dx calibration for 3/fb analyses", CDF note 9592.

[85] S. D'Auria et al "Track based calibration of the COT specific ionization", CDF Note 6932.

[86] "TOF PR plots page", (2002), `http://www-cdf.fnal.gov/upgrades/TOF/pr_plots/TOF_PR_plots.html`.

[87] M. Feindt et al "Particle Dependent Time-of-Flight Calibration", CDF note 9618.

[88] M. Feindt, "A Neural Bayesian Estimator for Conditional Probability Densities", arXiv:physics/0402093.

[89] V.Tiwari et at, "Likelihood based electron tagger", CDF note 7121.

[90] D. Usynin et al, "B flavor identification using opposite side muons", CDF note 6483.

[91] Bs Mixing Group, "A Neural Network based jet charge tagger", CDF note 7492.

[92] I. Furic et al, "Combined opposite side flavour tagger", CDF note 8460.

[93] M. Feindt, T. Kuhr, M. Kreps, J. Morlock, A. Schmidt, "Calibration of the Same Side Kaon Tagger", CDF public note 10108.

[94] M. Feindt, T. Kuhr, M. Kreps, J. Morlock, A. Schmidt, "Measurement of Mixing Frequency and Amplitude of Neutral B-Mesons", CDF note 10025.

[95] K. Anikeev, P. Murat, Ch. Paus, "Description of Bgenerator II", CDF note 5092.

[96] F. James & M. Roos. "Minuit: A System for Function Minimization and Analysis of Parameter Errors and Correlations", *Comp. Phys. Commun.*, **10** (1975) 343–367.

[97] R. Brun & F.Rademakers. "ROOT: An Object Oreinted Data Analysis Framework", *Nucl. Inst. Meth.*, **A389** (1997) 81 – 86.

[98] J. Boudreau et al, "Formulae for the analysis of flavor tagged $B_s \to J/\psi\phi$", arXiv:hep-ph/1008.4283v1.

[99] G. Punzi, "Comments on Likelihood fits with variable resolution", *Proceedings of PHYSTAT2003* (2004) arXiv:physics/0401045v1.

[100] Frederick James. "Statistical Methods in Experimental Physics, 2nd Edition". World Scientific (2006).

[101] J. Boudreau et al, "Measurement of the CP Violation Paramter beta' in Bs to Jpsi Phi", CDF note 8960.

[102] L. Demortier et al, "Everything you always wanted to know about pulls", CDF Note 5776.

[103] F. Azfar et al, "An Updated Measurement of the CP Violating Phase β_s using $B_s^0 \to J/\psi\,\phi$ Decays", CDF Note 9395.

[104] J. Boudreau et al, "Measurements of the exclusive B lifetimes in the modes $B \to J/\psi X$ and measurement of the B^+/B^0 lifetime ratio ", CDF note 6387.

[105] F. Azfar et al, "Update on the measurement of exclusive B lifetimes and measurement of B lifetime ratios", CDF note 9704.

[106] W. Press et al. "Numerical Recipes, the Art of Scientific Computing, 2nd Edition". Cambridge University Press (1992).

[107] Giovanni Punzi, "Ordering Algorithms and Confidence Intervals in the Presence of Nuisance Parameters", *Proceedings of PHYSTAT05* (2005) arXiv:physics/0511202.

[108] R. Moore, "Tevatron Collider Status and Prospects", FERMILAB-CONF-09-492-AD, (2009) arXiv:hep-ex/0910.3612.

[109] G. Cowan, "Determination of the CP violating phase in $B_s \to J/\psi\phi$ decays at LHCb", *Proceedings of LLWI2010* (2010) arXiv:hep-ex/1006.1041.