Flavour Oscillations and CP Asymmetry in Semileptonic B_s^0 Decays

Steven Beale

A Dissertation submitted to the faculty of graduate studies in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Graduate Program in Physics

York University,

Toronto, Ontario

January 2010

Abstract

The B_s^0 meson spontaneously transforms into its antiparticle ($\bar{B}_s^0$). These 'flavour oscillations' occur periodically with a frequency that may be measured. The oscillation frequency is related to the fundamental parameters of the electroweak interaction. Measuring the frequency provides a constraint on the electroweak quark coupling parameter V_{ts} and improves the constraint on V_{td}. Furthermore, the amplitude of the oscillation process may be slightly different in B_s^0 and $\bar{B}_s^0$ mesons due to $\mathcal{CP}$ violating nature of the weak interaction. This 'asymmetry' is expected to be small ($a_{fs}^{SM,s} = (2.06 \pm 0.57) \cdot 10^{-5}$), but may be enhanced ($a_{fs}^s \simeq \mathcal{O}(1\%)$) by new sources of $\mathcal{CP}$ violation.

This thesis describes a search for B_s^0 flavour oscillations and charge asymmetry in the $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ ($D_s^- \to K^{*0} K^-$) decay mode using 5.0 fb^{-1} of D0 data. A lower limit is placed on the oscillation frequency, $\Delta m_s > 9.9$ ps^{-1} with an expected sensitivity to oscillations below 14.8 ps^{-1}. The charge asymmetry is measured to be $a_{fs}^s = 0.018 \pm 0.025(\text{stat}) \pm 0.002(\text{syst})$. A combination of these measurements with other decay modes is also presented.

Acknowledgements

In the past five years, I have learned a great deal about physics and life in general. For this, I am indebted to the many people I have met over the course of this degree. More than anyone else, I would like to thank my supervisor, Wendy Taylor for guiding me in my studies and helping me to overcome obstacles along the way. She has been very generous with her time and patience; I could not have asked for a better supervisor.

I would like to thank members of the D0 B physics group, past and present, with whom I have worked for the past several years. In particular, those that I worked with on the analyses presented in this Dissertation. I inherited this decay channel from Md Naimuddin, who I would like to thank for guidance in understanding the analysis framework that he helped to establish. I would like to thank Hal Evans and Sergey Burdin who have been a great asset in helping me to understand some of the more subtle aspects of these analyses. I would also like to thank Dmitri Tsybychev for his support and Guennadi Borissov for making AATrack so easy to work with.

I would like to thank the members of my examining committee for reviewing, editting and examining this dissertation. In particular, I would like to thank Sampa Bhadra and Scott Menary, who have been very helpful with their yearly questions and criticism during my annual research evaluations.

Finally, I would like to thank all my friends and colleagues both at York and FNAL who have made these years entertaining, interesting, and enlightening.

For everyone who has ever taught me anything.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Background

While theoretical particle physics, it has been argued, has its origins in ancient Greek philosophy, *experimental* particle physics is a far younger field. Greek philosophers, as early as the fifth century B.C, argued that the world must be composed of a combination of indivisible elements (atoms) and empty space. These atomist speculations may bear some resemblance to modern atomic theory, but they were motivated primarily by ancient ideas regarding mathematics and the nature of motion and have little to do with testable properties of matter[1] [2]. These atomist theories were, at best, a lucky guess.

The modern iterative process of postulation and experimentation has more in common with alchemy and its attempts to turn lead into gold than it does with atomist speculation. Though based on the principle of the four Aristotelian elements[2], medieval alchemy set the stage for later molecular theories. 13^{th} century experiments conducted by the physician and alchemist Geber (Jabir) suggested

[1] In fact the atoms were considered to be unobservable.
[2] Fire, air, earth and water.

that metals such as gold, copper and iron were composed of small particles of mercury and sulfur, which in turn were composed of the four Aristotelian elements [3]. While such models may seem laughable to the modern reader, the important point is that these theories were based on experimental tests involving these metals and not on conjecture. With further experimentation, these ideas were refined over the next several centuries into what we now recognize as atomic theory.

As atomic physicists such as Rutherford and Bohr were investigating the atom, early particle physicists such as Victor Hess and Carl Anderson were studying the nature and origin of cosmic rays. It is at this point that particle physics became a branch of physics distinct from atomic physics. Cosmic ray investigations led to the discovery of the positron and the muon [4,5]. The discovery of the charged pion in 1937 was predicted by Hideki Yukawa's model of a nuclear force mediated by mesons, and was followed by the discovery of 'strange' particles (the kaons and Λ^0 baryon). These discoveries were made entirely through the observation of natural cosmic ray showers.

Cyclotrons have been in use since the early 1930s by atomic physicists, but these early machines were not sufficiently energetic to produce the new particles of interest to particle physicists. The first 'high energy' (GeV scale) proton synchrotrons constructed in the early 1950s allowed particle physicists to begin conducting experiments in the laboratory under controlled conditions. The result was an explosion of new particle discoveries (Σ, Ξ, ρ, η and others) [6, 7]. By the 1980s a 'Standard Model' (SM) of particle physics had been developed that described both the electroweak and strong interactions for the quarks and leptons. Now, in the TeV era the focus has switched to testing the SM and probing for new

physics beyond the SM. This chapter will give an overview of the SM and highlight elements of that model that are most relevant to this work.

1.1 The Standard Model

Fermions, particles with half-integer spin, are the basic building blocks of matter. The most fundamental of these are the quarks and leptons. They are arranged into three 'generations,' with each generation having a doublet of quarks and a doublet of leptons. In each quark doublet there is an 'up-type' quark with electric charge $+2/3$ and a 'down-type' quark with electric charge $-1/3$. These fractional charges are with respect to the electron charge $Q_e = 1.6 \times 10^{-19}$ C; this convention will be used throughout this work.

The names 'up' and 'down' are labels for the weak isospin, which is analogous to the intrinsic spin-up ($J_z = +1/2$) and down ($J_z = -1/2$) states. Each lepton doublet includes an electron-like particle with electric charge -1 and neutral massless neutrino[3]. In addition to the electric charge, each quark has "charges" associated with the weak and strong interactions. Similar to electric charge, these are quantum numbers that determine how a particle couples under the associated interaction. The weak "charge" is called flavour; each of the six quarks has its own. These are: up (u), down (d), charm (c), strange (s), top (t) and bottom (b). We will use the notation S, B, etc., to indicate the flavour charge and ΔS, ΔB to indicate the change in flavour charge in a process[4].

[3]The SM neutrino is massless. Current evidence of neutrino oscillations [8] contradicts this assumption.

[4]It should be noted that unlike the electric charge, flavour is not conserved. There is, however, baryon number conservation, which is a weaker form of flavour conservation that requires the

Each generation of leptons has its own associated flavour "charge", for example, the first generation has electron number +1. Similarly there is a muon number and a tau number for the second and third generations, respectively. The strong "charge" is called colour, and is carried only by the quarks. Each quark carries one of three colour "charges" (red, green or blue). Each colour has both positive and negative values (colour and anticolour). The leptons are colourless and do not interact under the strong force. The quarks and leptons are summarized in Table 1.1. In addition to these twelve fermions, each particle has an anti-particle partner with the same mass, but opposite quantum numbers.

	Electric Charge (Q_e)	Generation		
		1^{st}	2^{nd}	3^{rd}
Quarks	+2/3	u	c	t
	−1/3	d	s	b
Leptons	−1	e	μ	τ
	0	ν_e	ν_μ	ν_τ

Table 1.1: List of fermions.

Interactions between the fermions are mediated by integer-spin particles called gauge bosons. Each gauge boson is associated with a particular force. The electromagnetic interaction is mediated by the massless photon (γ) and couples to all of the charged fermions. These processes are described by quantum electrodynamics (QED). The weak interaction is mediated by the $W^{\pm}$ and Z^0 particles. While the Z^0 couples to all fermions, the $W^{\pm}$ couples only to left-handed fermions (or right-handed anti-fermions), that is, to particles with helicity -1 (see Sec. 1.1.1 for a discussion of helicity). The strong interaction is mediated by eight gluons

number of quarks balance the number of antiquarks.

that couple only to the quarks. Each gluon carries a combination of colour and anticolour (i.e., red and antiblue). Strong interactions are described by quantum chromodynamics (QCD). The electromagnetic and weak interactions are closely related and can be unified into a single electroweak theory with the electric and flavour charges combined into a single weak 'hypercharge'[5]. By contrast, the strong interaction appears to be completely independent of the electroweak interaction. The strong interaction is completely 'flavour-blind'; all quarks behave the same under the strong interaction regardless of flavour or electric charge. The reverse is also true; the electroweak interaction is colour-blind, ignoring the colour charge of the quarks.

Concluding the theory at this point would account for all observed SM particles; however, none would have mass. To give mass to the $W^{\pm}$, Z^0 and the charged fermions, we must add a scalar field to the theory which gives us a scalar Higgs boson (see Appendix A for further discussion). Table 1.2 summarizes the (gauge) bosons, with their associated interactions and mass. This work is concerned primarily with the electroweak interaction, which will be described in more detail in Sec. 1.1.2.

1.1.1 Charge-Parity Symmetry

The Charge ($\mathcal{C}$) and Parity ($\mathcal{P}$) operators transform a physical system (or state) under their respective symmetries. $\mathcal{C}$ converts a particle into its antiparticle and $\mathcal{P}$ performs a spatial coordinate inversion ($x, y, z \rightarrow -x, -y, -z$). To understand

[5]The hypercharge is defined in terms of the quark flavours, $Y_W = B + s + c + b + t$, where B is the baryon number ($+1/3$ for quarks, $-1/3$ for antiquarks). It is related to the electric charge and the weak isospin (I_{3W}), $Q = I_{3W} + 1/2Y_W$.

Boson	Interaction	Mass
γ	Electromagnetic	0
$W^{\pm}$	Weak (Charged Current)	80.4 GeV/c^2
Z^0	Weak (Neutral Current)	91.1 GeV/c^2
gluon	Strong	0
Higgs (H^0)	-	> 114 GeV/c^2

Table 1.2: List of (gauge) bosons with their associated forces and masses.

how parity acts on a particle, it is useful to first consider vectors. A polar vector such as position or momentum will be negated under a parity operation ($\mathcal{P}\vec{r} = -\vec{r}, \mathcal{P}\vec{p} = -\vec{p}$). However, a vector product of two polar vectors will not change sign under $\mathcal{P}$. Angular momentum is an example of such a vector ($\mathcal{P}\vec{l} = \mathcal{P}(\vec{r} \times \vec{p}) = (-\vec{r}) \times (-\vec{p}) = \vec{l}$). These are called axial or pseudovectors. A scalar may also transform under parity. A scalar triple product of polar vectors ($\vec{a} \cdot \vec{b} \times \vec{c}$) or an inner product of a polar and axial vector will both change sign under a parity transformation. Helicity, for example, is the sign of the projection of a particle's spin onto its direction of motion ($h = \text{signum}(\vec{v} \cdot \vec{J})$). $\mathcal{P}$ will change the sign of velocity ($\vec{v}$), but not the spin ($\vec{J}$). Consequently, $\mathcal{P}$ will change a right-handed particle ($h = +1$) into a left-handed particle ($h = -1$). In the case of a massless particle (such as the SM neutrino) the velocity cannot be Lorentz transformed into the opposite direction, which makes the helicity a fundamental quantum number. For a massive particle, the helicity is not fundamental but, is a good approximation of the chirality, which is an intrinsic handedness[6].

As discussed above, the $W^{\pm}$ boson couples only to left-handed fermions. Furthermore, there is no right-handed neutrino in the SM (and no left-handed antineu-

[6]For a massless particle the helicity and chirality are the same.

trino). The complete absence of a right-handed neutrino is particularly striking. It indicates that the weak interaction is *maximally* parity violating. If we consider a weak process such as (for example) pion decay, we see that application of $\mathcal{C}$ or $\mathcal{P}$ separately gives a disallowed process due to the handedness of the neutrino (see Fig. 1.1). But the combination of $\mathcal{C}$ and $\mathcal{P}$ gives the proper charge-conjugated process with a right-handed antineutrino, $\bar{\nu}_R$, (Fig. 1.1, lower right). For the moment it seems that $\mathcal{CP}$ is the conserved symmetry in the weak interaction. We will see in Sec. 1.2.2 how this symmetry will be broken.

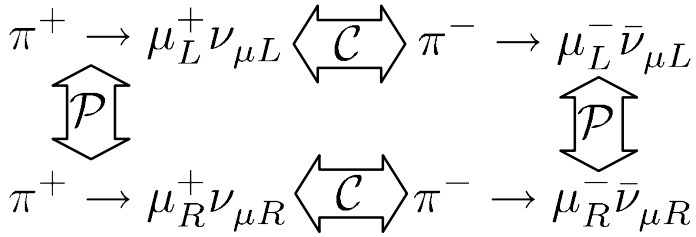

Figure 1.1: $\mathcal{C}$- and $\mathcal{P}$-conjugate processes for pion decay. The top right and bottom left processes are forbidden.

1.1.2 Electroweak Interaction

It is instructive to look at the historical development of weak interaction theory in order to understand the present electroweak formalism. Before the development of the parton model, the proton and neutron were considered to be a doublet of a fundamental nucleon field. They were assigned a spin-like quantum number called isospin with the proton in a $|\frac{1}{2},\ +\frac{1}{2}\rangle$ state[7] and the neutron in a $|\frac{1}{2},\ -\frac{1}{2}\rangle$ state.

[7]We use the Dirac 'bra-ket' notation where $|j,\ m\rangle$ indicates a state with total angular momentum j and projected angular momentum (for example, in the z direction) m.

Isospin is considered the charge associated with the weak interaction. The discovery of strange particles required the introduction of a new charge: strangeness[8]. It was observed that decays involving a change in strangeness ($\Delta S \neq 0$) were less frequent than decays with conserved (or zero) strangeness. For example, compare $n \to p^+ e^- \bar{\nu}_e$ to the equivalent strange-suppressed $\Lambda \to p^+ e^- \bar{\nu}_e$ decay. The neutron decay process has a branching fraction $\mathrm{Br}(n \to p^+ e^- \bar{\nu}_e) = 100\%$[9], whereas the equivalent Λ process has a fraction $\mathrm{Br}(\Lambda \to p^+ e^- \bar{\nu}_e) = (8.32 \pm 0.14) \times 10^{-4}$ [9]. In 1963 Cabbibo proposed an empirically determined scaling factor to suppress these $|\Delta S| > 0$ processes [10]. The amplitude for processes with $\Delta S = 0$ include a factor $\cos\theta_C$ and those with $\Delta S = \pm 1$ include a factor $\sin\theta_C$, where $\theta_C \approx 13.1^\circ$ is the Cabbibo angle [11, 12].

This 'Cabbibo' theory was developed before the more modern quark models of the late 1960s. In a modern theory with quark fields, this can be understood as a splitting of down-type quarks into composite fields (or a rotation in flavour space). In this scheme, the d and s quarks are no longer the 'fundamental' objects. Instead the combination $d_C = \cos\theta_C d + \sin\theta_C s$ becomes the isospin partner to the up quark. The u, d doublet becomes a u, d_C doublet.

The problem with this theory is that it predicts decays such as $K^0 \to \mu^+\mu^-$ (see Fig. 1.2) that are not observed or are heavily suppressed ($\mathrm{Br}(K_S^0 \to \mu^+\mu^-) < 3.2 \times 10^{-7}$, $\mathrm{Br}(K_L^0 \to \mu^+\mu^-) = (6.84 \pm 0.11) \times 10^{-9}$ [9]). A solution to this problem is to postulate a new quark doublet that contains the 'extra' $s_C = -\sin\theta_C d + \cos\theta_C s$[10]

[8] At the time, strangeness was called 'hypercharge.' In the SM, hypercharge is the combination of isospin and the flavour charges of the second and third generation quarks. Isospin plays the same role as flavour in the first generation.

[9] This is the fractional probability that the decay will proceed by this process.

[10] d_C and s_C must be orthogonal, hence the sign of the sine term.

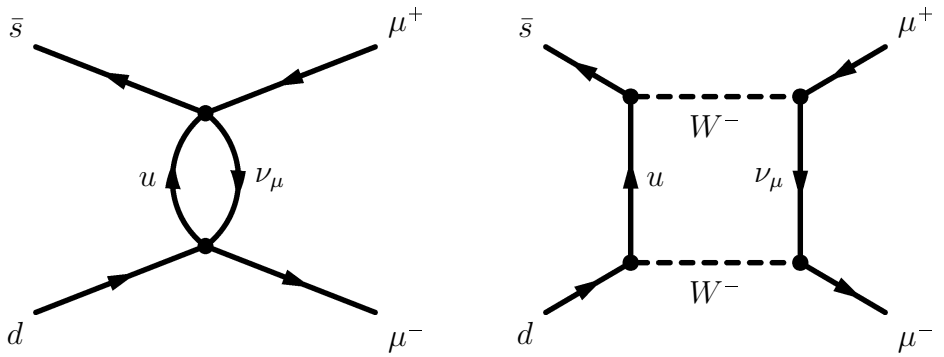

Figure 1.2: Diagram depicting $K^0 \to \mu^+\mu^-$ in Cabbibo theory (left) and the equivalent SM process (right).

quark, as well as a new up-type quark: the charm (c) quark. This refinement of the Cabbibo theory was developed in 1970 by Glashow, Iliopoulos and Maiani and is now referred to as the GIM mechanism [13]. The process in Fig. 1.2 will have an amplitude proportional to $\sin\theta_C \cos\theta_C$, but a similar process with a c quark in place of the u quark will have an amplitude $-\sin\theta_C \cos\theta_C$. The amplitudes of these two processes will cancel giving a suppression of the $K^0 \to \mu^+\mu^-$ branching fraction. Shortly after being proposed, the c quark was confirmed with the discovery of the J/ψ meson ($c\bar{c}$) [14, 15]. Now we can think of the Cabbibo angle θ_C as a rotation in flavour space:

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

The choice of having a composite down-type quarks is completely arbitrary. We could do the same for the u and c quarks instead, if the c quark had been discovered before the s quark.

The extension to three generations of quarks was motivated by the observation of $\mathcal{CP}$ violation in the neutral kaon system [16]. This discovery won J. Cronin and

V. Fitch the Nobel Prize in 1980. In 1973, M. Kobayashi and T. Maskawa argued that to accommodate $\mathcal{CP}$ violation, an additional generation is needed [17]. From our earlier example, the cancellation between the u and c quark terms was perfect. This is true in any two generation model; however, the introduction of a third generation allows interference from another quark term. With this interference, the cancellation is not perfect and can also be $\mathcal{CP}$ asymmetric (that is there can be different rates for K_S^0 and K_L^0[11]). This prediction won Kobayashi and Maskawa the Nobel Prize in 2008.

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

The 3×3 unitary matrix in Eqn. 1.2 is the Cabbibo-Kobayashi-Maskawa (CKM) matrix and it describes the weak couplings between the various quarks. For example, the matrix element V_{us} is the coupling between the s and u quarks and is equivalent to $\sin\theta_C$ from the Cabbibo theory. The CKM matrix will be discussed in more detail in Sec. 1.2.2.

While this model is very successful, it is somewhat artificial. The Cabbibo angle was put in 'by hand' to suppress the $s \leftrightarrow u$ processes, and then generalized to include c, b and t quarks. A more modern approach has the matrix arising naturally from quark couplings to the Higgs boson, this procedure is described in Appendix A. However, this procedure requires that we assume a flavour changing

[11] The experiment conducted by Cronin and Fitch searched for the $\mathcal{CP}$ violating decay $K_L^0 \to \pi^+\pi^-$.

current (i.e., flavour-changing processes for the quarks, the so-called 'quark mixing') , which is basically equivalent to postulating a Cabbibo angle; however, the SM cannot explain the origin of the quark mixing.

1.2 CKM Quark Mixing Matrix

Any complex 3×3 unitary matrix may be parametrized by three real parameters and six complex phases. This parametrization can be applied to the CKM matrix and further reduced by exploiting properties of the quark fields. The fields are only defined up to an arbitrary phase, so the quark terms in the Lagrangian are dependent only on the phase differences. That is, if we make a phase transformation to the quark fields $q_j \rightarrow q_j e^{i\theta_j}$ and $q_k \rightarrow q_k e^{i\theta_k}$, the quark terms in the Lagrangian will transform as

$$\bar{q}_j V_{jk} q_k \rightarrow \bar{q}_j V_{jk} e^{-i(\theta_k - \theta_j)} q_k. \tag{1.3}$$

By suitably choosing the phase difference between the quarks, the phase of V_{jk} can be absorbed. This is equivalent to making a rotation in flavour space. With six quark fields, there are five such independent and arbitrary phase differences. Removing these phases leaves only a single complex phase and three real parameters in the CKM matrix. We will see in Sec. 1.2.2 how this phase is related to the breaking of the Charge-Parity symmetry in the SM. The choice of parametrization is arbitrary, but the Wolfenstein parametrization [9] has become the standard

choice:

$$V = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{1.4}$$

where the real parameters are λ, A and $|\rho + i\eta|$ and the phase is $\arg(\rho + i\eta)$.

The unitarity conditions of the CKM matrix give six complex relations constraining the various matrix elements. Eqns. 1.5 and 1.6 are two such relations related to the B^0 and B_s^0 mesons.

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.5}$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \tag{1.6}$$

If we consider the terms $V_{ij}V_{ik}^*$ as vectors in the complex plane, then these relations may be represented graphically by triangles (see Fig. 1.3). We can represent these using the Wolfenstein parametrization; however, it is convenient to renormalize this parametrization so that the important features are contained within two parameters. We define parameters $\bar{\rho}$ and $\bar{\eta}$ as rescalings of ρ and η by an expansion in powers of λ (for example, $\bar{\rho} = \rho(1 - \lambda^2/2 + \ldots)$, see Ref. [18, 19] for details). In this scheme we have

$$\bar{\rho} + i\bar{\eta} \equiv \frac{-V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \tag{1.7}$$

and the triangle associated with Eqn. 1.5 can be represented with $\bar{\rho}$ and $\bar{\eta}$ determining the apex of the triangle (see Fig. 1.3).

The length of the side $\left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right|$ is constrained by the measurement of the B_s^0 oscillation frequency described herein. The oscillation frequency in the B^0 system

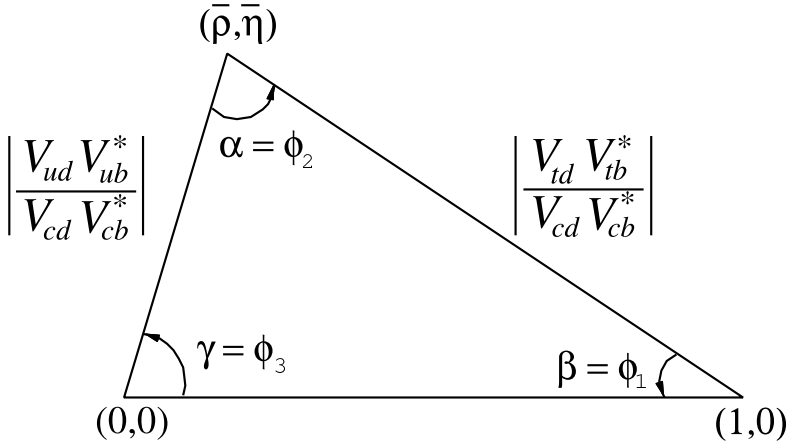

Figure 1.3: The most common representation of the CKM unitarity triangle.

(Δm_d) provides a direct constraint on this side (see Eqn. 1.8). The uncertainty on the CKM parameters constrained by Δm_d is dominated by the theoretical uncertainty on the parameters f_{B_d} and B_d determined from Lattice QCD calculations[12]. The ratio of Δm_d and Δm_s is dependent on the parameter $\xi = \frac{f_{B_s}\sqrt{B_s}}{f_{B_d}\sqrt{B_d}}$, which can be determined with a smaller theoretical uncertainty (see Eqn. 1.9), thus improving the impact of the combined CKM constraint.

$$\Delta m_d \propto f_{B_d}^2 B_d |V_{td}V_{tb}^*|^2 = (f_{B_d}^2 B_d) A^2\lambda^6 \left((1-\bar{\rho})^2 + \bar{\eta}^2\right) + \mathcal{O}(\lambda^{10}) \tag{1.8}$$

$$\frac{\Delta m_d}{\Delta m_s} \propto \xi^2 \frac{|V_{td}V_{tb}^*|^2}{|V_{ts}V_{tb}^*|^2} = \xi^2\lambda^2 \left((1-\bar{\rho})^2 + \bar{\eta}^2\right) + \mathcal{O}(\lambda^6) \tag{1.9}$$

We are also interested in the triangle that results from Eqn. 1.6. This measurement of the B_s^0 semileptonic $\mathcal{CP}$ asymmetry constrains the angle β_s in this triangle. This constraint is explained in Sec. 1.2.2.

[12]These parameters are the decay constant (f_{B_d}) and bag parameter (B_d), which describe the strong contributions to the decay of the B^0 meson (see Ref. [20]).

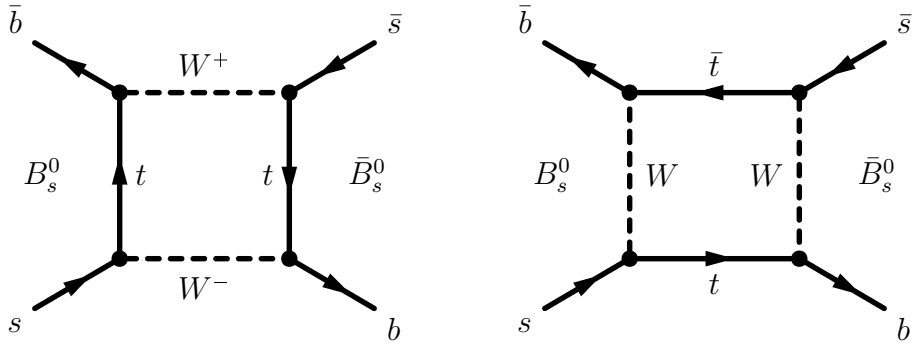

Figure 1.4: $|\Delta B| = 2$ $B_s^0 - \bar{B}_s^0$ mixing diagrams.

1.2.1 $B_s^0 - \bar{B}_s^0$ Mixing

If we allow quark mixing in our model, we must also consider processes such as Fig. 1.4 with $|\Delta B = 2|$. In this case, the Schrödinger equation describing the time evolution of a B_s^0 system is given by Eqn. 1.10.

$$i\frac{d}{dt}\begin{pmatrix} |B_s^0(t)\rangle \\ |\bar{B}_s^0(t)\rangle \end{pmatrix} = \left(\mathbf{M} - i\mathbf{\Gamma}/2\right)\begin{pmatrix} |B_s^0(t)\rangle \\ |\bar{B}_s^0(t)\rangle \end{pmatrix}, \tag{1.10}$$

where $\mathbf{H} = \mathbf{M} - i\mathbf{\Gamma}/2$ is an effective Hamiltonian with hermitian components $\mathbf{M}$ and $\mathbf{\Gamma}$, which are the mass and decay matrices, respectively. The Hamiltonian itself is not, in general, hermitian[13]. The eigenstates of this Hamiltonian are not the B_s^0 and $\bar{B}_s^0$ mesons, but rather a linear combination of these states, which we

[13]A hermitian matrix is equal to its transpose conjugate. A hermitian Hamiltonian would not violate $\mathcal{CP}$ symmetry.

designate by their relative mass (light and heavy)[14].

$$\begin{aligned} \text{Light:} \quad |B_L\rangle &= p|B_s^0\rangle + q|\bar{B}_s^0\rangle, \\ \text{Heavy:} \quad |B_H\rangle &= p|B_s^0\rangle - q|\bar{B}_s^0\rangle. \end{aligned} \tag{1.11}$$

The light and heavy states have masses and decay widths M_L, Γ_L and M_H, Γ_H, respectively. The parameters p and q must satisfy the condition $p^2 + q^2 = 1$. If $|p| = |q|$, then mass eigenstates $|B_L\rangle$ and $|B_H\rangle$ are also $\mathcal{CP}$ eigenstates. The system can be solved to determine the time evolution of an initial B_s^0 or $\bar{B}_s^0$ meson in terms of the parameters p, q, $\Delta m_s = M_H - M_L$ and $\Delta\Gamma_s = \Gamma_L - \Gamma_H$ (see Refs. [21, 22]). From the time dependence we can determine the probability density functions (*pdf*s) for the lifetime of an initial $B_s^0/\bar{B}_s^0$:

$$\mathcal{P}(B_s^0 \to B_s^0; t) = \frac{e^{-\Gamma_s t/2}}{2}\left(\cosh\frac{\Delta\Gamma_s t}{2} + \cos\Delta m_s t\right), \tag{1.12}$$

$$\mathcal{P}(B_s^0 \to \bar{B}_s^0; t) = \left|\frac{q}{p}\right|^2 \frac{e^{-\Gamma_s t/2}}{2}\left(\cosh\frac{\Delta\Gamma_s t}{2} - \cos\Delta m_s t\right), \tag{1.13}$$

$$\mathcal{P}(\bar{B}_s^0 \to \bar{B}_s^0; t) = \frac{e^{-\Gamma_s t/2}}{2}\left(\cosh\frac{\Delta\Gamma_s t}{2} + \cos\Delta m_s t\right), \tag{1.14}$$

$$\mathcal{P}(\bar{B}_s^0 \to B_s^0; t) = \left|\frac{p}{q}\right|^2 \frac{e^{-\Gamma_s t/2}}{2}\left(\cosh\frac{\Delta\Gamma_s t}{2} - \cos\Delta m_s t\right). \tag{1.15}$$

In the relations above, we use 'natural units' where $\hbar \equiv 1$, $c \equiv 1$ and time has units $1/eV$. Or equivalently, the mass difference (Δm_s) has units $1/s$. We interpret this as the oscillation frequency of the cosine term that we are to measure. By substituting the B_d^0 mass and lifetime parameters, we can determine the equivalent

[14] We could just as easily denote the eigenstates by their different lifetimes (as in the K_S^0-K_L^0 system), but the mass difference in this case is more significant.

*pdf*s for the $B_d^0 - \bar{B}_d^0$ system. The parameters p and q will be central to the $\mathcal{CP}$ asymmetry discussed in Sec. 1.2.2, but for the mixing measurement we make the assumption that any $\mathcal{CP}$ violating effects are negligible (i.e., $|p/q| = 1$).

Eqns. 1.12-1.15 describe the probability of decaying at a time t after being produced for each possible initial- and final-state (i.e., $B_s^0/\bar{B}_s^0$). The average decay time, as measured in the B_s^0 rest frame, is called the lifetime (τ). The decay time is determined from the measured displacement of the B_s^0 decay vertex from its point of origin. This displacement is called the decay length. The distribution of the decay length has the same form as Eqns. 1.12-1.15. Provided that we can identify the oscillated decays from the unoscillated decays, Δm_s can be extracted from the distribution of the B_s^0 decay length. The problem is complicated by the high frequency of the oscillation. Δm_s had previously been predicted to be large ($\Delta m_s > 15.0$ ps^{-1} [23]). It has since been limited by the D0 Collaboration ($17 < \Delta m_s < 21$ ps^{-1} [24]) and later measured by the CDF Collaboration ($\Delta m_s = 17.77 \pm 0.10(\text{stat}) \pm 0.07(\text{syst})$ ps^{-1} [25]). The large value of the oscillation frequency requires an excellent decay time resolution in order to resolve the oscillations. With $1/\Gamma_s = 1.425$ ps [9], $\Delta m_s/\Gamma_s \sim 20$, so we expect $\sim 20/2\pi \sim 4$ $B_s^0 - \bar{B}_s^0$ oscillations before decaying. Hence, to resolve the oscillations we need a decay length resolution better than $\sigma \sim c\tau_{B_s}/4 \sim 100$ μm.

1.2.2 $\mathcal{CP}$ Violation in Mixing

We saw in Sec. 1.2 that the CKM matrix can be parametrized by three real values and a complex phase. This ensures that the CKM matrix is irreducibly complex, and hence $\mathcal{CP}$-violating. A quark mixing process, for example, $s \to uW^-$, is

described by a coupling of the form $V_{us}\bar{u}\gamma^{\mu}W_{\mu}^{+}s$, where $\bar{u}$, W_{μ}^{+} and s represent the quark and gauge boson fields and γ^{μ} are the four gamma matrices[15]. The $\mathcal{CP}$-conjugate process ($\bar{s} \rightarrow \bar{u}W^{-}$) is similar, but with a conjugate coupling constant V_{us}^{*}. However, if we perform a $\mathcal{CP}$ transformation on this process, we find that $\mathcal{CP}$ will be violated if V_{us} is complex:

$$\mathcal{CP}(V_{us}\bar{u}\gamma^{\mu}W_{\mu}^{+}s) = V_{us}\bar{s}\gamma^{\mu}W_{\mu}^{-}u \neq V_{us}^{*}\bar{s}\gamma^{\mu}W_{\mu}^{-}u. \tag{1.16}$$

In most cases, the phase of V_{us} (or any other CKM element) will not be relevant, because the observables are dependent on the square of the amplitude (for example, the decay rate $\Gamma \propto |V_{us}|^2$). This conceals the $\mathcal{CP}$-violating phases in the lowest order tree-level processes. Higher order processes that have contributions from several diagrams can have interference effects that reveal the $\mathcal{CP}$-violating behaviour.

One of the mixing processes of Fig. 1.4 is reproduced in Fig. 1.5 with the CKM couplings included for each vertex. The process is shown for both $B_s^0 \rightarrow \bar{B}_s^0$ and $\bar{B}_s^0 \rightarrow B_s^0$ transitions. Similar to the case above, the amplitude of these two processes will be complex conjugates of each other. However, in this case there are contributions where one or both top quarks are replaced by charm quarks[16]. These contributions will add new terms to the total amplitude ($\mathcal{A}$) for the oscillation process.

$$\mathcal{A}(B_s^0 \rightarrow \bar{B}_s^0) \propto S(z_c)\lambda_c^2 + 2S(z_c, z_t)\lambda_c\lambda_t + S(z_t)\lambda_t^2, \tag{1.17}$$

[15]Gamma matrices, also called Dirac matrices, can be thought of as the relativistic versions of the Pauli matrices.

[16]In principle there is a contribution from the up quark as well. However, this is suppressed by the smallness of V_{ub}.

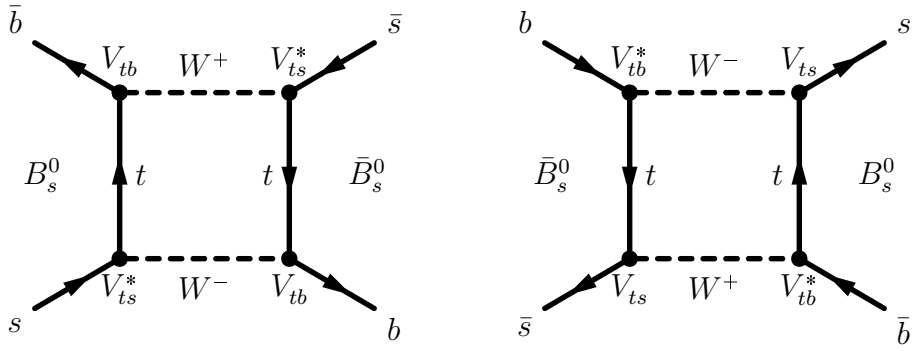

Figure 1.5: $B_s^0 - \bar{B}_s^0$ mixing with quark couplings labeled, $B_s^0 \to \bar{B}_s^0$ on the left and $\bar{B}_s^0 \to B_s^0$ on the right.

where $\lambda_c = V_{cb}V_{cs}^*$, $\lambda_t = V_{tb}V_{ts}^*$ and the functions $S(z)$ and $S(z_c, z_t)$ are QCD form factors [21]. The parameters z_c and z_t are the ratios of the masses of the c and t quarks with respect to the W boson ($z_c = m_c^2/M_W^2$, $z_t = m_t^2/M_W^2$). The phase of the cross term $\lambda_c\lambda_t$ will not vanish when the amplitude is squared. The decay rate will be dependent on the phase difference between λ_c and λ_t.

We can quantify $\mathcal{CP}$-violating effects using the angles of the unitarity triangles described in Sec. 1.2, which are dependent on the phase differences between the various CKM elements. In particular, we will be interested in the angle β_s, which originates from the triangle given by Eqn. 1.6. This angle is related to the phase difference (ϕ_s) between off-diagonal elements of the mass and decay matrices of Eqn. 1.10:

$$\phi_s \equiv \arg\left(-\frac{M_{12}}{\Gamma_{12}}\right) \simeq -2\beta_s \equiv (-2)\arg\left(-\frac{V_{cs}V_{cb}^*}{V_{ts}V_{tb}^*}\right). \tag{1.18}$$

$\mathcal{CP}$-violating effects measured in the B_s^0 system can be described in terms of the phase ϕ_s and subsequently used to extract or limit β_s.

From Eqns. 1.13 and 1.15, we can see immediately that $\mathcal{CP}$ will be violated if $|q/p| \neq 1$. In this case, the lifetime *pdf* for the mixed decays is different for

$B_s^0 \to \bar{B}_s^0$ and $\bar{B}_s^0 \to B_s^0$. We can define a flavour-specific[17], time-dependent decay rate asymmetry (a_{fs}^s) to quantify this difference:

$$a_{fs}^s(t) \equiv \frac{\mathcal{P}(\bar{B}_s^0 \to B_s^0; t) - \mathcal{P}(B_s^0 \to \bar{B}_s^0; t)}{\mathcal{P}(\bar{B}_s^0 \to B_s^0; t) + \mathcal{P}(B_s^0 \to \bar{B}_s^0; t)} = \frac{1 - |q/p|^4}{1 + |q/p|^4}. \tag{1.19}$$

Note that the time-dependent asymmetry is, in fact, independent of time. This allows a time-integrated measurement of the asymmetry (that is, a counting experiment) provided we have good knowledge of the initial-state flavour of the B_s^0 meson (i.e., B_s^0 or $\bar{B}_s^0$, see Sec. 4.2 for a discussion of flavour tagging).

A measurement of a_{fs}^s can be made by counting the number of $B_s^0 \to \bar{B}_s^0$ and $\bar{B}_s^0 \to B_s^0$ decays. The challenge is to identify the B_s^0 and $\bar{B}_s^0$ mesons and determine which oscillation process occurred. In semileptonic decays the process is determined by the charge of the muon from the decays. By looking for pairs of B hadrons (produced from an initial $b\bar{b}$ pair) both decaying to muons, a relatively simple measurement can be made. In this dimuon channel, the charge of the muons is easily determined providing the identification of the decay process. A pair of like-charge muons indicate an oscillated decay; oppositely charged muons indicate a non-oscillated decay (see Sec. 4.2 for a detailed discussion). This analysis has been performed by both the CDF [26] and D0 [27] collaborations. The dimuon selection cannot distinguish between B_s^0 and B^0 mesons, which makes the measured asymmetry a combination of a_{fs}^s and a_{fs}^d. The value of a_{fs}^s can be extracted from these measurements by inputting an a_{fs}^d measurement from B factories. The

[17]Flavour-specific refers to processes whereby only the B_s^0 meson and not the $\bar{B}_s^0$ meson can decay to the final-state. That is, $\bar{B}_s^0 \to f$ is forbidden, as is $B_s^0 \to \bar{f}$. Examples of non-flavour-specific decays are $B_s^0 \to J/\psi\phi$ and $B_s^0 \to D_s^+ D_s^-$.

most recent (May 2009) evaluation from the CDF and D0 measurements yields $a^s_{fs} = 0.025 \pm 0.021 \pm 0.017$ (CDF) and $a^s_{fs} = -0.0061 \pm 0.0058 \pm 0.0077$ (D0), assuming $a^d_{fs} = -0.0047 \pm 0.0046$ [28, 29].

The dimuon requirement in the above analysis sacrifices the large number of events that lack an opposite side muon ($\mathrm{Br}(B \to \mu X) \sim 10\%$). If we drop this requirement, the statistics increase significantly, but we lose knowledge of the initial-state B^0_s flavour. In this case we must generalize Eqn. 1.19 to include possible unmixed events:

$$a'^s_{fs}(t) \equiv \frac{\mathcal{P}(B^0_s/\bar{B}^0_s \to B^0_s; t) - \mathcal{P}(B^0_s/\bar{B}^0_s \to \bar{B}^0_s; t)}{\mathcal{P}(B^0_s/\bar{B}^0_s \to B^0_s; t) + \mathcal{P}(B^0_s/\bar{B}^0_s \to \bar{B}^0_s; t)} = \frac{a^s_{fs}}{2}\left(1 - \frac{\cos\Delta m_s t}{\cosh\Delta\Gamma_s t/2}\right). \tag{1.20}$$

For this 'untagged' asymmetry the time dependence does not vanish. If we again make a time-integrated measurement, the time-dependent term will be vanishingly small and can be neglected. However, due to the factor of 1/2 in Eqn. 1.20 we are only half as sensitive to the more fundamental parameter a^s_{fs}. This analysis has been performed by the D0 collaboration ($a^s_{fs} = 0.0123 \pm 0.0097 \pm 0.0017$ [30]) and is independent of a^d_{fs}[18].

This work represents a combination of these two approaches and uses all available information from the B^0_s decay. In addition, a time-dependent approach is used that does not neglect the time-dependent term (in Eqn. 1.20) and allows better separation of the background asymmetries that are constant in time. The analyses are conducted using a 5.0 fb^{-1} data sample collected by the D0 experiment at the Fermi National Accelerator Laboratory (see Chap. 3). The B^0_s mesons

[18]Reconstructing the B^0_s decay allows separate selection of B^0 and B^0_s.

are produced in $p\bar{p}$ collisions at a centre of mass energy of $\sqrt{s} = 2$ TeV.

1.3 B meson production

At the Tevatron, the $b\bar{b}$ production cross section is approximately 10 μb [19] [31]. With an instantaneous luminosity of order 100 $\mu\text{b}^{-1}\text{s}^{-1}$, the Tevatron produces ~ 1000 $b\bar{b}$ pairs every second.

$b\bar{b}$ quark pairs are produced when partons (i.e., quarks or gluons) within the proton and antiproton collide at high energy. These collisions may include any of the valence quarks[20], the gluons or the 'sea quarks[21].' This allows for a number of QCD processes that give the final-state $b\bar{b}$ pair. Flavour creation is the simplest tree-level process. This includes annihilation of light quarks (see Fig. 1.6(a)) and gluon fusion (see Fig. 1.6(b)). Next-to-leading order (NLO) flavour creation processes may include a gluon radiated from one of the $b\bar{b}$ quarks. These processes make up roughly 35% of the total $b\bar{b}$ cross section [32].

Flavour excitation and parton showering are NLO processes with cross sections comparable to or exceeding the flavour creation cross section. Flavour excitation is the process by which one quark from a sea $b\bar{b}$ pair is scattered by a gluon or other quark into the final-state (see Fig. 1.6(c)). Parton showering is the process by which a $b\bar{b}$ pair is produced from a scattered gluon (see Fig. 1.6(d)). These processes are characterized by lower final-state momentum and small scattering

[19] A barn is a unit of cross section related to area, 1 barn = 10^{-24} cm^2.

[20] Valence quarks are the principle quarks in the hadron that determine the flavour. For example, the proton has valence quarks uud.

[21] In any hadron there are copious amounts of gluons and quark-antiquark pairs. These 'extra' quarks are called sea quarks.

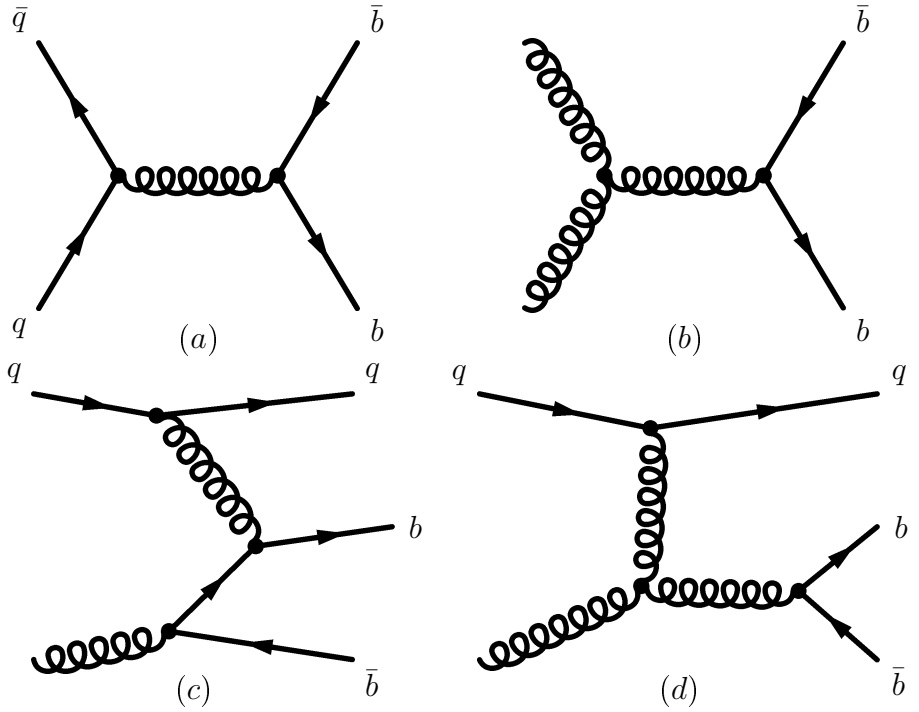

Figure 1.6: Examples of $b\bar{b}$ production diagrams. $q\bar{q}$ annihilation (a) and gluon fusion (b) are flavour creation processes. Flavour excitation (c) and parton showering (d) are NLO processes.

angles for at least one of the $b/\bar{b}$ quarks [32].

1.3.1 Hadronization

Colour confinement in the strong interaction does not permit free quarks. The newly produced $b\bar{b}$ pair must 'hadronize' into semi-stable colourless states (i.e., produce hadrons). From the initial energy of the $b\bar{b}$ pair, additional $q\bar{q}$ pairs are created until all quarks are confined within hadrons. The final result is a pair of B hadrons and a host of other lighter particles. This process cannot be described by perturbative QCD; we instead use a phenomenological model tuned to data.

For this work we use the Lund string model [33]. In this model, the initial $b\bar{b}$

quarks are connected by a string field with energy proportional to the separation of the quarks. As the b and $\bar{b}$ quarks move apart due to their kinetic energy, the string is stretched. Energy is transferred from the quarks to the stretching string, slowing the quarks. When the energy stored in the string field is sufficiently high a new $q\bar{q}$ pair is created, splitting the string. The new separated strings will continue to stretch and create new pairs until the energy drops below some threshold.

The type of B hadron that remains at the end of the hadronization process is determined by the flavour of the last $q\bar{q}$ pair to break a string field connected to the $b/\bar{b}$ quark. The probabilities for producing a B^+, B^0, B_s^0, B_c^+ or a B baryon (Λ_b^0, Σ_b, Ξ_b, Ω_b) are given by the fractions f_u, f_d, f_s, f_c and f_{baryon}. These fractions are measured in $p\bar{p}$ collisions at the Tevatron as well as in e^+e^- collisions at LEP (see Table 1.3). These measurements typically neglect the contributions from heavier B hadrons (B_c^+ and charm B baryons), which have very small contributions ($f_c \approx 0.2\%$) [9].

B fraction	$Z \to b\bar{b}(\%)$	Combined with $p\bar{p}(\%)$
f_u, f_d	40.2 ± 0.9	39.9 ± 1.1
f_s	10.5 ± 0.9	11.1 ± 1.2
f_{baryon}	9.1 ± 1.5	9.2 ± 1.9

Table 1.3: Fractions for producing B hadrons in $e^+e^- \to Z^0 \to b\bar{b}$, and in $p\bar{p}$ collisions at 1.8 TeV [9].

1.4 Motivation

The B sector is a good place to search for new physics. B hadrons are characterized by a long lifetime (~ 1.5 ps^{-1} yielding a decay length of ~ 500 μm) making

them relatively easy to identify. Furthermore, they must decay by flavour-changing processes that are sensitive to new physics. The B_s^0 sector in particular is interesting because of current discrepancies between SM predictions and experimental measurements [34] (see, for example, Fig. 1.8).

Of the matrix elements in Eqn. 1.10, M_{12} is the most sensitive to new physics. The diagonal element of the decay matrix, Γ_{12}, is dominated by tree level processes describing the B_s^0 decay [34]. We can describe the effects of any new physics by how it relates to M_{12}:

$$M_{12}^s \equiv M_{12}^{SM,s} \cdot \Delta_s,\ \Delta_s \equiv |\Delta_s| e^{i\phi_s^\Delta}, \tag{1.21}$$

where $M_{12}^{SM,s}$ is the SM prediction for matrix element. The complex parameter Δ_s describes the contribution from any source of new physics. With this formalism, we have the following relations for Δm_s, $\Delta\Gamma_s$ and a_{fs}^s:

$$\begin{aligned}
\Delta m_s =& 2|M_{12}| = \Delta m_s^{SM} |\Delta_s|, \\
\Delta\Gamma_s =& 2|\Gamma_{12}^{SM,s}| \cos(\phi_s^{SM} + \phi_s^\Delta), \\
a_{fs}^s =& \left| \frac{\Gamma_{12}^{SM,s}}{M_{12}^{SM,s}} \right| \frac{\sin(\phi_s^{SM} + \phi_s^\Delta)}{|\Delta_s|} = \frac{\Delta\Gamma_s^{SM}}{\Delta m_s^{SM}} \frac{\tan(\phi_s^{SM} + \phi_s^\Delta)}{|\Delta_s|}.
\end{aligned} \tag{1.22}$$

In these equations ϕ_s^{SM} is the $\mathcal{CP}$ phase and the source of $\mathcal{CP}$ violation within the SM. ϕ_s^Δ is the phase associated with any new source of $\mathcal{CP}$ violation.

We can now see that both Δm_s and a_{fs}^s are sensitive to the new physics parameters $|\Delta_s|$ and ϕ_s^Δ. The current CDF measurement of the oscillation frequency ($\Delta m_s = 17.77 \pm 0.10 \pm 0.07\ \text{ps}^{-1}$) has a statistical uncertainty significantly smaller

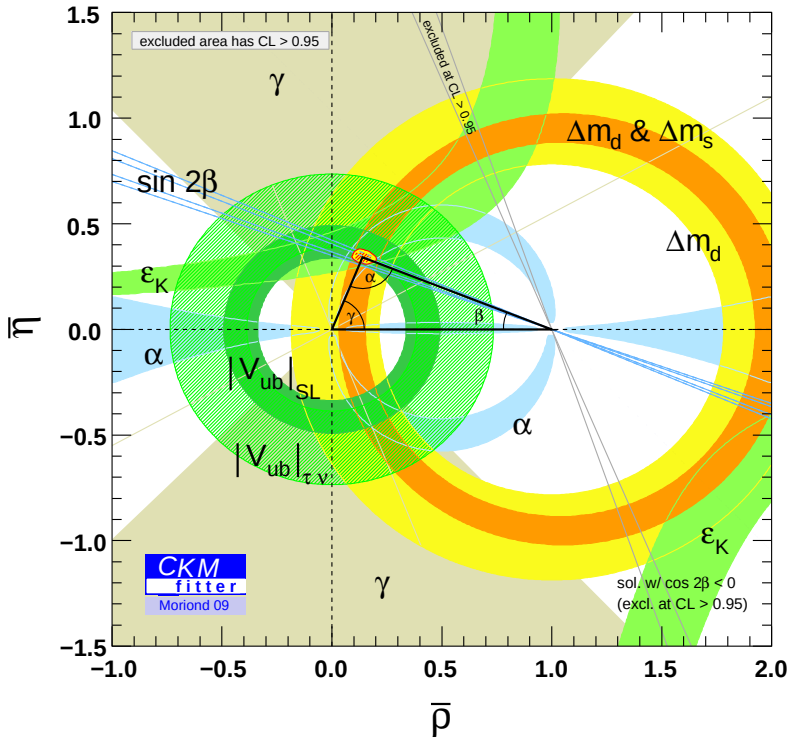

Figure 1.7: The current global fit to the CKM triangle resulting from Eqn. 1.5 [19].

than the theoretical uncertainty on this parameter ($\sim \pm 6$ ps^{-1} [34]). This limits the usefulness of Δm_s as a probe of new physics. The ratio $\Delta m_d/\Delta m_s$ has a smaller theoretical uncertainty and can be used as a measure of the CKM ratio $|V_{td}/V_{ts}|$ (recall Eqn. 1.9). This measurement constrains the global fit to the CKM triangle performed by groups such as CKMfitter [19] and UTfit [35]. The current status of this fit, including the constraint provided by $\Delta m_d/\Delta m_s$ is shown in Fig. 1.7.

The flavour-specific asymmetry is a particularly good measurement for searching for new sources of $\mathcal{CP}$ violation. The SM prediction for this parameter is very

nearly zero and has a small theoretical uncertainty ($a_{fs}^{SM,s} = (2.06 \pm 0.57) \cdot 10^{-5}$). A small increase in ϕ_s could boost the asymmetry parameter by a significant amount.

It is interesting to examine other measurements of the $\mathcal{CP}$ phase. The $\mathcal{CP}$ phase can be measured directly from B_s^0 decays to the $\mathcal{CP}$ eigenstate $J/\psi\phi$. This measurement has been made by the CDF [36, 37] and D0 [38] collaborations. Other constraints on ϕ_s are provided by the B_s^0 flavour-specific lifetime and the branching fraction for the process $B_s^0 \to D_s^- D_s^+ X$ [28, 39, 40]. The B_s^0 asymmetry can also be used to constrain ϕ_s using the relation in Eqn. 1.22. By combining measurements and constraints, a limit can be made on the allowed region for ϕ_s. Fig. 1.8 shows the current limit on ϕ_s as of April 2009. These limits do not include a new 2.8 fb^{-1} result from CDF [36] or the results of this asymmetry measurement.

Fig. 1.8 is striking as both plots show a $\sim 2\sigma$ deviation from the SM point. By improving the constraints, it is hoped that we might gain statistically significant evidence of new physics. The a_{fs}^s measurement in this work, which builds upon this Δm_s measurement, will improve the constraints on ϕ_s and will help to shed new light on possible extensions to the SM.

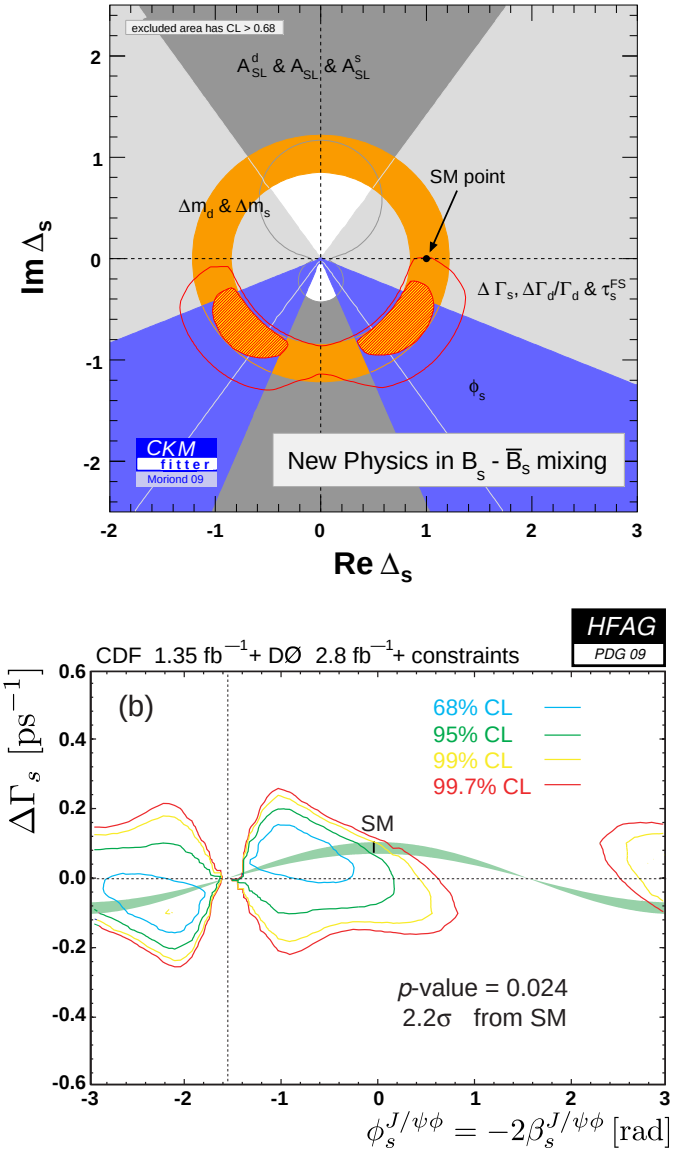

Figure 1.8: Two different representations of limits on new $\mathcal{CP}$-violating physics. The CKMfitter group plots the real and imaginary components of the parameter Δ_s (top) [19]. The HFAG group plots the lifetime difference $\Delta\Gamma_s$ against the $\mathcal{CP}$ phase $\phi_s^{J/\psi\phi} = \phi_s^{SM} + \phi_s^{\Delta}$ (bottom) [39]. The SM prediction is shown in both plots. The band in the bottom plot is the function $\Delta\Gamma_s(\phi_s^{\Delta})$, where $\Delta\Gamma_s(0)$ is the SM prediction. This band shows the allowed region for new physics.

Chapter 2

Analysis Overview

This work includes two independent analyses: a B_s^0 oscillation search and a B_s^0 semileptonic $\mathcal{CP}$ asymmetry search. Both searches are performed by analyzing a common 5 fb^{-1} sample of $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ decays, with $D_s^- \to K^{*0} K^-$ and $K^{*0} \to K^+ \pi^-$(charge conjugate processes are implied throughout this work). The previous chapter explained the physics behind these phenomena; this chapter will explain some of the more practical aspects of making a measurement.

2.1 Analysis Strategy

As explained in the previous chapter, a flavour oscillation in B_s^0 mesons is an oscillation between the B_s^0 and $\bar{B}_s^0$ states. The oscillation occurs in time with some associated frequency, Δm_s. If we could monitor the flavour of a B_s^0 meson during its short life we could make a simple measurement. Similarly, we could measure the $\mathcal{CP}$ asymmetry by measuring the amplitude difference in the oscillation for

B_s^0 and $\bar{B}_s^0$ mesons. However, we can only make (at most) two measurements of the flavour; at the instant the B_s^0 meson is created, and at the instant in which it decays. These two measurements will not tell us at what frequency the oscillation occurs.

In order to determine the oscillation frequency and $\mathcal{CP}$ asymmetry we must observe a large sample of B_s^0 decays. The lifetime of the B_s^0 mesons follows a simple exponential decay law regardless of the flavour. When an initial B_s^0 meson oscillates into a $\bar{B}_s^0$ meson, we are likely to observe a $\bar{B}_s^0$ meson decay and unlikely to observe a B_s^0 decay. As the oscillation proceeds in time, the probability for an initial B_s^0 meson to decay as a $\bar{B}_s^0$ meson will oscillate (see Fig. 2.1). The $\mathcal{CP}$ conjugate process will behave in the same manner but potentially with a different relative probability, compared to the flavour conserving process ($\bar{B}_s^0 \to \bar{B}_s^0$). This difference will result in an asymmetry in the amplitude of the oscillation probability (see Fig, 2.1).

The approach described above does not consider detector effects or background contamination. In reality these are serious problems that cannot be ignored. For example, the uncertainty on our measurement of the lifetime of the B_s^0 mesons is comparable to the period of the oscillation ($\sigma \sim 0.20$ ps and $T \sim 0.35$ ps). This leads to a 'smearing' of the oscillation that makes the measurement much more challenging (see Fig. 2.2). Difficulty in identifying the initial flavour of the B_s^0 meson will lead to a contamination by the flavour-conserving decays that oscillate out of phase with the flavour-changing decays. This will reduce the amplitude of the oscillation, making if harder to observe. Undetected decay products will lead to further uncertainty in the lifetime, and will likewise smear the oscillations. Finally,

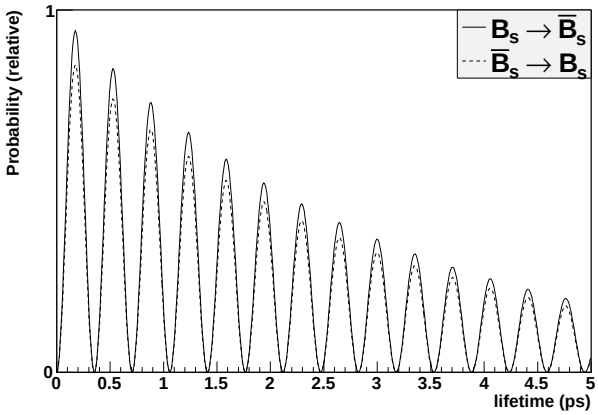

Figure 2.1: An idealization of B_s^0 oscillations. The $\mathcal{CP}$ asymmetry is evident from the difference in amplitude.

background processes that resemble the B_s^0 decay will reduce the significance of the oscillation, making it difficult to detect. Nevertheless, if we account for these factors, the oscillation frequency and the $\mathcal{CP}$ asymmetry can be extracted from the distribution of the B_s^0 decay.

To account for the effects described above, it is necessary to consider more than just the basic lifetime *pdf*. Parameters, such as mass and lifetime resolution, are used to discriminate against the background sources and enhance the significance of the signal events, as well as model the detector effects described above. *Pdf*s for all parameters are combined into a joint conditional *pdf* for the lifetime. This function is then used to define a likelihood function of the oscillation frequency, which can be fitted to determine Δm_s. A power spectrum approach is used to analyze the lifetime function and extract the oscillation frequency. The same approach was used in a previous D0 limit of the B_s^0 oscillation frequency [24], as well as in the CDF measurement [25]. This work is the first to be submitted for publication that

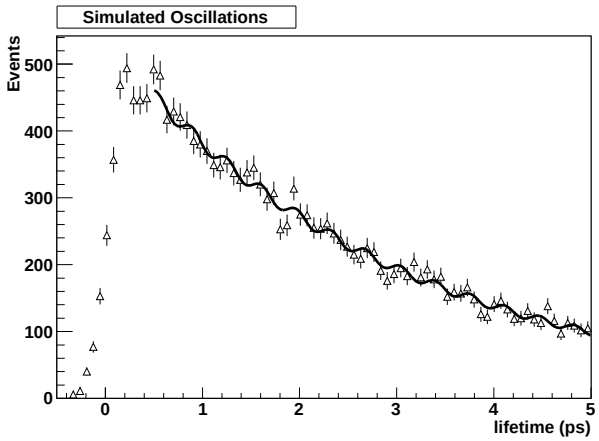

Figure 2.2: A somewhat more realistic example of oscillations. This simulated data includes effects due to the detector resolution.

includes the $D_s^- \to K^{*0}K^-$ decay mode[1].

The B_s^0 $\mathcal{CP}$ asymmetry measurement follows a similar strategy, making use of the B_s^0 lifetime distribution. As above, other parameters are used to discriminate against the background, producing a joint conditional *pdf* for the B_s^0 lifetime distribution as a function of the asymmetry. This is used to create a likelihood function, which is then fitted for the asymmetry. The accuracy of this fit is improved if we can distinguish between different sources of asymmetry and between oscillated and non-oscillated decays[2]. As discussed in Sec. 1.2.2, the inclusive dimuon asymmetry measurement suffers from a dependence on an external measurement of the B^0 charge asymmetry [26, 27]. The untagged measurement uses an exclusive decay, making it independent, but does not include tagging information to separate B_s^0 from $\bar{B}_s^0$ decays [30]. This work is the first to make use of an exclusive decay

[1]Earlier work on this decay mode has been previously presented [41].

[2]The identification of the oscillated decays is referred to as flavour tagging. This is discussed in Sec. 4.2.

with tagging and is the first to make a time-dependent measurement.

The B_s^0 oscillation analysis is completely dependent on our ability to distinguish an oscillated decay from a non-oscillated decay. For the asymmetry analysis, the flavour tag is helpful but not critical. As such, we divide the data into two samples for these two analyses. The 'untagged' data sample includes all events, regardless of the flavour tag. This sample is used to study the charge asymmetry. The 'tagged' data sample is a subset of the untagged sample that excludes all events that lack a flavour tag. This sample is used to study the oscillation frequency. The two samples are the same in all other respects (event selection, reconstruction, etc.).

2.2 Likelihood Fitting Procedure

In order to determine the B_s^0 oscillation frequency and $\mathcal{CP}$ asymmetry, we use a likelihood function. This function gives the probability of obtaining the data sample as a function of either the oscillation frequency or the charge asymmetry. By scanning through a range of Δm_s, we can find the optimal value or place a lower and possibly an upper bound on the frequency (see Sec. 7.1). The charge asymmetry can be fitted directly by maximizing the likelihood (see Sec. 5.6).

The likelihood for each B_s^0 candidate is determined on an event-by-event basis. The total likelihood ($\mathcal{L}$) is the product of likelihoods for each event,

$$\mathcal{L}(\vec{\theta}) = \prod_i^N f_i(\vec{x}_i; \vec{\theta}), \tag{2.1}$$

where the index i runs over all N events. The vectors $\vec{x}_i$ and $\vec{\theta}$ denote the set of measured values for the i^{th} event and the set of fit parameters, respectively. The event-by-event likelihood $f_i(\vec{x}_i;\vec{\theta})$ is the conditional probability of obtaining measured variables $\vec{x}_i$ given the model parameters $\vec{\theta}$. The procedure is simplified by working with the natural logarithm of the likelihood. This converts the likelihood product into a sum and controls the scale of the total. In this scheme, we minimize the total log likelihood L,

$$L(\vec{\theta}) = -2\sum_{i}^{N} \ln f_i(\vec{x}_i;\vec{\theta}). \tag{2.2}$$

The factor of -2 is inserted to simplify the interpretation of the fit parameter uncertainty. With this formalism, the width of the minimum with respect to the model parameter θ_j is related to the uncertainty on the parameter θ_j. The depth of the minimum is related to the statistical significance of the measurement of the parameters $\vec{\theta}$. If we assume a Gaussian-distributed error on the parameter θ_j, $L(\theta_j)$ will have a quadratic form about the minimum: $L(\theta_j) = a\theta_j^2 + b\theta_j + c$. The 1σ uncertainty is then given by the scale (a) of the quadratic term, $\theta^{measure} = \theta_j^{min} \pm \sqrt{1/a}$ (see, for example, Ref. [9] pp. 320-329).

The likelihood function to be used for these analyses is the joint probability density function, f_j^i, for the measured parameters in each event. These parameters will be discussed in detail in the following chapters. They are, in brief:

- The B_s^0 decay length, x^M. See Sec. 6.2.
- The combined selection variable, y. This parameter is used to identify B mesons. See Sec. 4.1.2.

- The uncertainty on the B_s^0 decay length, σ_{x^M}.

- The predicted dilution, d_{pr}. This parameter quantifies our confidence in the identification of the B_s^0 flavour (i.e., B_s^0 or $\bar{B}_s^0$). See Sec. 4.2.

- The mass of the $KK\pi$ system, m. These are the decay product of the D_s^- candidate. See Sec. 6.1.

The *pdf*s for σ_{x^M}, m and d_{pr} are uncorrelated and the *pdf* for x^M is modelled separately for the various $KK\pi$ sources (hence it is independent of m). Thus, we obtain

$$f_j^i = P_j^{x^M}(x_i^M, y_i; \sigma_{x^M}^i, d_{pr}^i) P_j^{\sigma_{x^M}} P_j^m P_j^{d_{pr}}, \tag{2.3}$$

where the index i runs over all events and the index j indicates the $KK\pi$ source (see Sec. 6.1). The total log likelihood is then given by the sum over sources and events:

$$L = -2 \sum_i^N \ln \left(\sum_j Fr_j f_j^i \right), \tag{2.4}$$

where Fr_j is the fractional contribution from the source j. The functions in Eqn. 2.3 must be determined independently before the fit to Eqn. 2.4 can determine Δm_s and a_{fs}^s. The functions P_j^m and $P_j^{x^M}$ are fitted to data and are discussed in Sec. 6.1 and Sec. 6.2, respectively. The functions $P_j^{\sigma_{x^M}}$ and $P_j^{d_{pr}}$ are determined directly from data and are given in Sec. 5.1. The function $P_j^{x^M}$ is dependent on several other input parameters and functions. These dependencies will be described in Chap. 5.

2.3 Structure of this thesis

All of the physics involved in the analyses presented in this work must, in some way, be incorporated into the likelihood function (Eqn. 2.4). Describing this function is no small task. As explained above, there are many ingredients that must be combined before the final fit can be done. Chap. 3 describes the accelerator and the detector used to make these measurements. This includes a brief discussion of the trigger system for data collection and the software tools used to generate simulated data.

Chap. 4 discusses the selection criteria used to identify B_s^0 meson candidates and the techniques used to discriminate against background. In addition, this chapter describes the flavour-tagging technique used in differentiating B_s^0 from $\bar{B}_s^0$ mesons. Flavour tagging is particularly important for the measurement of the B_s^0 oscillation frequency.

A number of parameters are included in the likelihood function that have no direct correlation to the oscillation frequency or $\mathcal{CP}$ asymmetry. These parameters are included to either account for some detector effects, or to give some improvement to the separation of the signal and background sources. These parameters are discussed in Chap. 5.

The two main functions of this analysis are the distributions for the mass of the D_s^- candidate and the lifetime of the B_s^0 candidate. The functions describing these distributions and their fits to the data are described in Chap. 6. This chapter will also explain how the parameters discussed in Chap. 5 are to be applied in the final fit.

The results of the final fits for the B_s^0 oscillation frequency and the $\mathcal{CP}$ asymmetry are shown in Chap. 7. A number of cross checks were performed to verify the fits are not biased in some way. These tests are also described in this chapter. In addition, studies of the systematic uncertainty of the measurements are presented.

In the final chapter, the measurements described in this work are combined with those using other decay modes. These decay modes are studied independently by other D0 collaborators and together with those described in this thesis yield a measurement with higher sensitivity.

Chapter 3

Accelerator and Detector

The Fermi National Accelerator Laboratory (FNAL, or simply 'Fermilab') has been at the frontier of high energy particle physics for the past two decades. The primary accelerator at Fermilab is the Tevatron. As its name suggests, the Tevatron is a 1 TeV proton/antiproton accelerator[1] and storage ring. The Tevatron hosts a pair of collider detectors, CDF and D0 ('D-zero'), which have been in operation since 1987 and 1992, respectively. The extensive physics programs at these two experiments have yielded hundreds of publications and Ph.D. theses.

This chapter will briefly describe the Tevatron and the related accelerators, as well as give a detailed description of the D0 detector and all of its detector subsystems. The Tevatron and the D0 detector were upgraded during the period between 1996 and 2001. The period prior to this upgrade is referred to as Run I. The period following the upgrade is referred to as Run II. Furthermore, Run II is divided by a (smaller) upgrade that took place in 2006, where the periods before

[1]The operational beam energy is 0.98 TeV, but beam tests have achieved energies as high as 1.012 TeV [42].

and after this upgrade are referred to as Run IIa and Run IIb, respectively. Only the Run II Tevatron and D0 detector will be discussed here.

3.1 The Tevatron

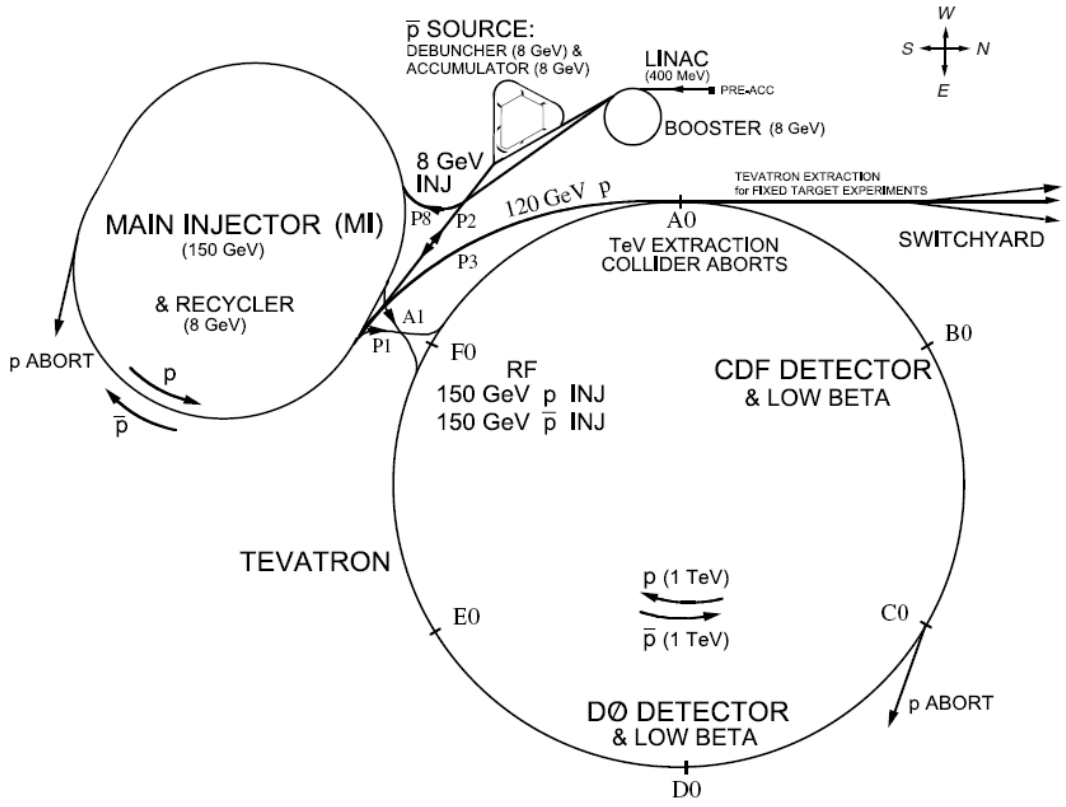

Figure 3.1: The Fermilab accelerator complex. The diagram shows the Tevatron main ring as well as the acceleration stages leading up to the 1 TeV beam energy.

The production of 1 TeV protons and antiprotons starts from a simple tank of hydrogen gas. The acceleration up to 1 TeV as well as the production of antiprotons requires an assortment of additional accelerators apart from the Tevatron itself

(see Fig. 3.1). The various stages of acceleration and the machines involved are described below.

The hydrogen gas is ionized (H^-) inside of the dome of a Cockcroft-Walton voltage multiplier charged to -750 kV. The ionized gas is accelerated up to 750 keV through a column connecting the charged dome to a grounded wall housing the preaccelerator construction. The H^- ions then enter a linear accelerator, where they are accelerated to 400 MeV. The next stage of acceleration is the 8 GeV Booster, a 75 m radius synchrotron. Upon injection into the booster, the ions hit a stripping foil, which removes most of the electrons, leaving a beam of protons. Any remaining H^- ions are deflected out of the accelerator upon reaching the first bending dipole magnet. The protons are then accelerated to 8 GeV and sent via a transfer line to the Main Injector (MI).

The MI is elliptical in shape with a circumference of approximately 3 km. The MI has two operating modes, one for the production of antiprotons and another for injecting protons and antiprotons into the Tevatron. The primary difference between these two modes is the beam energy: 120 GeV for antiproton production and 150 GeV for Tevatron injection. For antiproton production, the beam is directed down a transfer line to a fixed target station housing a nickel-iron target. The resulting collisions produce antiprotons of approximately 8 GeV.

A triangular shaped synchrotron is used to capture and control the antiprotons, which are then transferred to a second similar synchrotron for temporary storage. These machines are referred to as the Debuncher and Accumulator, respectively. Once a sufficient number of antiprotons is collected in the Accumulator, they are returned to the MI by the same transfer line through the target station. Dur-

ing Run IIa, the antiprotons were then accelerated to 150 GeV and injected into the Tevatron. The process was repeated until there was a sufficient number of antiprotons in the ring.

The Run IIb upgrade included the addition of an antiproton storage ring called the Recycler. The Recycler recovers and stores antiprotons remaining in the Tevatron at the end of a period of collisions (called a 'store'). The antiprotons from the Accumulator are added to the Recycler, where they are stored until a sufficiently large 'stash' of antiprotons ($\mathcal{O}(10^{12})$) has been accumulated to be injected into the Tevatron. The Recycler shares the same tunnel as the MI and is solely a storage ring; antiprotons must be transferred to the MI to be accelerated for injection to the Tevatron. This recycling of antiprotons dramatically reduces the downtime in between stores [42, 43].

The final and largest accelerator in the sequence is the Tevatron. 1 km in radius, the Tevatron accelerates protons and antiprotons to their final 0.98 TeV beam energy. The ring is divided into six equal sectors separated by straight sections labeled A0-F0. These coincide with positions around the ring where the proton and antiproton beams can be brought together for collisions. B0 and D0 are collision halls housing the CDF and D0 detectors, respectively. C0 houses the proton abort line used when no antiprotons are in the ring. For example, when terminating a store after the antiprotons have been extracted to the Recycler, the remaining protons can safely be dumped via the proton abort line. A0 contains the proton and antiproton collision aborts (in the event of an unforeseen problem) as well as the connection to the switchyard for the fixed target experiments. E0 houses the apparatus for recovering antiprotons for storage in the Recycler at the end

of a store. F0 houses all RF accelerating cavities for the Tevatron as well as the injection lines from the MI.

Each of the six curved sectors contain 17 FODO cells, each composed of a focusing quadrupole magnet (F), four dipole bending magnets (O), a defocusing quadrupole (D) and four additional dipoles (O). The Tevatron is the only accelerator at Fermilab to use superconducting magnets, which must be housed in a liquid helium cryostat at $4-5$ K, depending on the beam energy. The proton and antiproton beams circulate around the machine in a common beam pipe clockwise and counterclockwise, respectively. The use of a single beam pipe requires that the beams follow separate helical orbits about the beam axis to avoid collisions around the ring. The orbit is made straight upon entering the CDF and D0 detectors to allow the beams to meet for collisions [42].

3.2 The D0 Detector

The D0 detector is a general purpose collider detector. It includes a tracking detector in the centre surrounded by a calorimeter, which is in turn surrounded by a muon spectrometer (see Fig. 3.2). The objective of the detector is to record 'physics objects' resulting from a $p\bar{p}$ collision. Physics objects, such as charged tracks or clusters of deposited energy, are detector signatures that are later interpreted to determine the physics process that occurred in the collision. This chapter will discuss the detector hardware (Secs. 3.2.2-3.2.4) and how physics objects are identified (Sec. 3.2.5). The subsequent interpretation will be discussed in Chaps. 4-5.

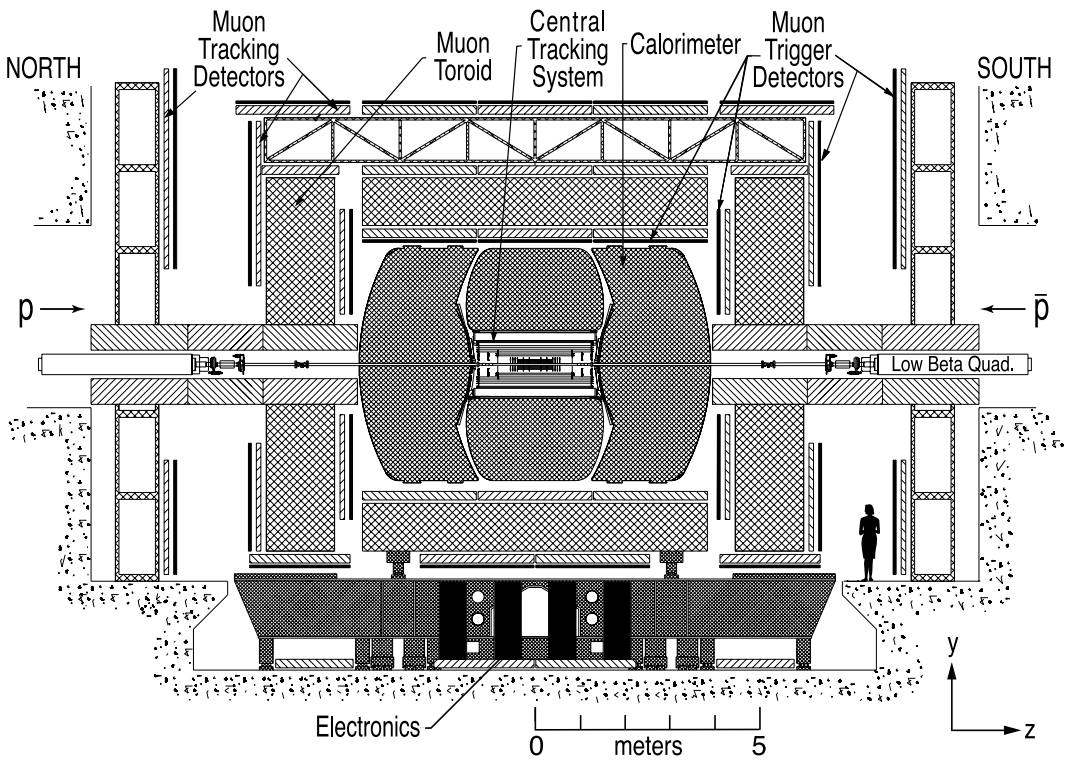

Figure 3.2: The D0 detector viewed from the west side (from inside the Tevatron ring) [44]. The various subdetector systems are labeled and are described separately in the text.

3.2.1 Coordinates

The detector is (mostly) cylindrically symmetric around the beam axis and so a cylindrical coordinate system will be used throughout this work. On occasion, a Cartesian coordinate system will prove to be more useful and will be used instead. Both coordinate systems take the geometric centre of the detector as the origin. The z axis runs down the centre of the beam pipe (north-south), with the proton direction taken to be $+z$ (south). The y axis is upwards and the x axis is horizontal towards the centre of the accelerator ring (west). The r coordinate denotes the

distance perpendicular from the z axis ($r = \sqrt{x^2 + y^2}$). The angles θ and ϕ are the polar and azimuthal angles, respectively.

It should be noted that the $p\bar{p}$ collision point is not, in general, at the origin of the coordinate system. The 'beam spot' describes the point in $r - \phi - z$ where the two beams are focused for collisions and is determined separately for each data collection run (at most four hours long or $1\ \mathrm{pb}^{-1}$ at typical luminosities).

Vector quantities are often broken into components parallel or perpendicular to the z axis. This is particularly true for momentum which is described by a transverse component in the $r - \phi$ plane (p_T), an axial component (p_z) and a scalar total momentum (p). Throughout this work we will make use of decay parameters projected into the transverse plane. The detector resolution is the best in the transverse plane, and so whenever possible we will restrict measured values to this plane. The concept of transverse momentum is extended to energy with the definition of 'transverse energy,' $E_T = E \sin\theta$. This definition is useful in calorimetry where the total transverse energy must be zero. Any imbalance of the transverse energy indicates the presence of a neutrino (or other unreconstructed particle) with a transverse energy equal to the missing energy ($\not\!\!E_T$).

It is useful to use rapidity, y, with respect to the beam axis in place of the polar angle:

$$y = \frac{1}{2} \ln \left[\frac{E + p_z c}{E - p_z c} \right] = \frac{1}{2} \ln \left[\frac{1 + \beta \cos\theta}{1 - \beta \cos\theta} \right]. \tag{3.1}$$

The rapidity difference between two particles is invariant under boosts along the beam axis, but is dependent on the particle energy making it inconvenient for describing detector geometry. The pseudorapidity, $\eta = -\ln\left[\tan\left(\theta/2\right)\right]$, approximates

the true rapidity in the limit $\beta \to 1$ [2]. Unlike true rapidity, the pseudorapidity is purely a geometric parameter. Furthermore, to a good approximation, particle production is uniform in (pseudo)rapidity. These properties make pseudorapidity a useful substitute for the polar angle.

3.2.2 Tracking Detector

The tracking system is designed to provide spatial, momentum and charge determination for charged particles. A charged particle traversing the detector will interact with the detector material, leaving a 'hit' signature. A series of hits in different detector layers forms a 'track' that follows the trajectory of the particle. If a magnetic field is applied parallel to the beam axis, the trajectory will curve in the azimuthal direction. The direction and extent of the curvature provides a measurement of the charge and momentum. Multiple tracks that point back to a common point can be used to determine the position of the decay vertex of their parent particle.

The D0 tracking system is composed of two primary subdetectors: a Silicon Microstrip Tracker (SMT) and a Central Fiber Tracker (CFT). The tracking system is contained within a 2 T solenoid to provide momentum determination from track curvature. The configuration of the detector subsystems and the magnet are shown in Fig. 3.3.

[2] Or equivalently $mc^2/E \to 0$.

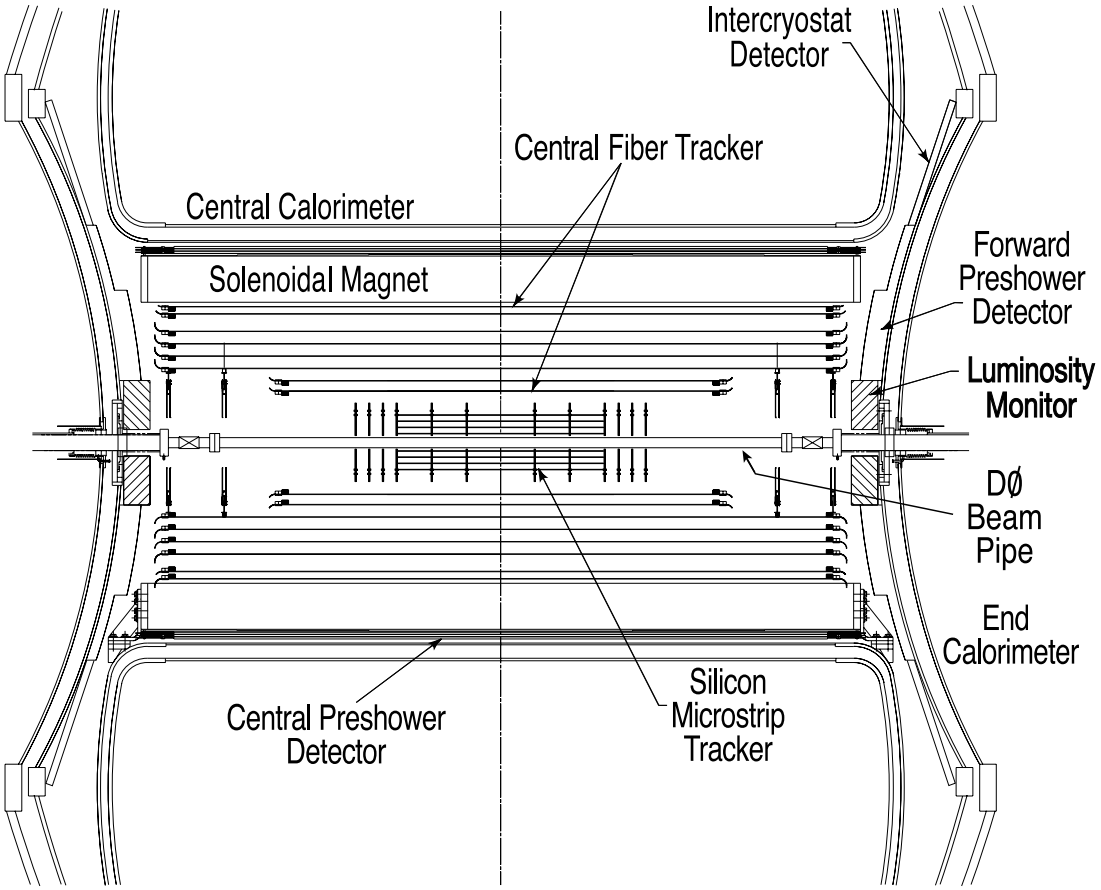

Figure 3.3: Cross-sectional view of the D0 Tracking detector [44].

Silicon Microstrip Tracker

The SMT is designed to provide tracking very close to the primary interaction point. In order to minimize the amount of detector material traversed by a charged particle, the detector surfaces need to be kept nearly perpendicular to the tracks for all η. So far as is possible, this should be true for particles originating anywhere within the interaction region. To achieve this, the detector is composed of a series of sensors arranged cylindrically at small η with $r - \phi$ disks at large η. The intermediate region is composed of a combination of both sensor configurations

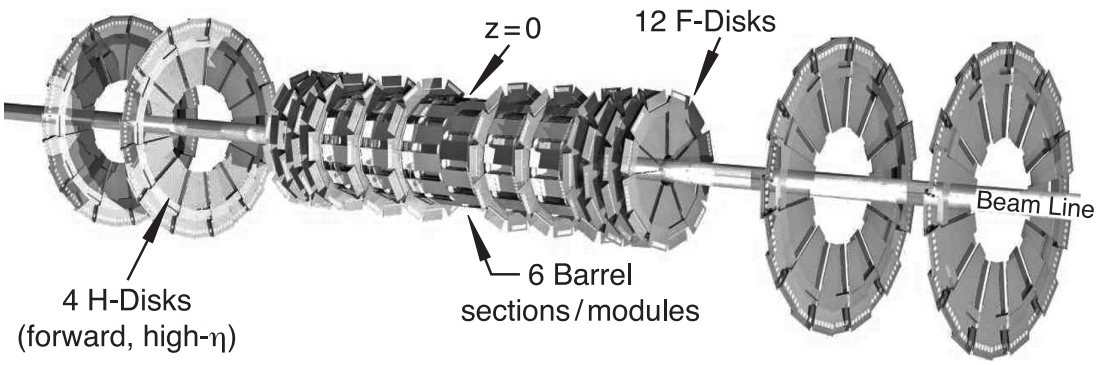

Figure 3.4: 3D view of the Silicon Microstrip Tracker [44].

(see Fig. 3.4).

There are six cylindrical 'barrel' sections in the detector, each 12 cm long (axially) and composed of four layers of sensors with average radii 2.72, 4.55, 7.58 and 10.51 cm. The inner two layers include 12 silicon sensors and the outer layers include 24 sensors. Each barrel is topped on the high $|z|$ side by an 'F-disk' composed of wedge-shaped sensors. The F-disk has an inner (outer) radius of 2.57 (9.96) cm. Three additional F-disks providing high η coverage are located after the barrels at $|z| = 43.1$, 48.1 and 53.1 cm. Two larger disks ('H-disks') with inner (outer) radii of 9.5 (26) cm are located at $|z| = 100.4$ and 121.0 cm. The H-disks provide coverage up to $\eta \sim 3$ in the very forward direction, where there is no CFT coverage.

The various barrels and disks use a variety of different sensor designs, but they operate on the same basic principles. A wafer of n-type silicon (excess e^- donors) is etched with narrow strips of p-type silicon (excess e^- acceptors). The spatial hit resolution of the sensors is determined by the spacing of the strips on the silicon surface (the 'pitch'). Aluminum strips are overlaid on the p-type strips, and the

n-type side of the silicon is mounted on an aluminum plate[3]. The interface between the n-type and p-type silicon forms a p-n junction with conduction-band electrons drifting to the p-side and valence band holes drifting to the n-side. A net negative charge develops on the p-side and a positive charge on the n-side of the junction. The region in between is left in a 'depleted' state, free of electrons and holes. A high voltage is applied to the aluminum strips and plate causing the electrons and holes to move further away from the junction and widening the depletion region. An ionizing particle passing through the depletion region will create electron/hole pairs, which immediately move apart and drift due to the applied electric field in the silicon. The electrons drift to the p-side and holes to the n-side of the silicon, creating a current in the nearest strip. The orientation of the strips determines the plane in which a measurement can be made. The barrel detectors with axial strips provide measurement of hits primarily in the $r-\phi$ plane. Strips can also be added to the n-side of the wafer at a stereo angle to provide some measurement in a second dimension. The F- and H-disks use stereo angles to make measurements in both $r-z$ and $r-\phi$.

The sensors used are as follows:

- Single-Sided Ladder Sensors (SS). These are the most basic of the silicon sensors with a 50 μm pitch. 6 cm long individual sensors were bonded together to form a 12 cm ladder. These sensors are used only in the two outer barrels in layers 1 and 3.

- Double-Sided Ladder Sensors (DS). Double sided sensors have p-type strips

[3]In the case of a double-sided sensor, there are aluminum strips on both sides.

on one side with a 50 μm pitch, and n-type strips on the other with a 62.5 μm pitch. The strips have a 2° stereo angle, which allows some limited measurement in the $r - z$ plane. As with the SS sensors, these are composed of two 6 cm sensors bonded together. All barrels use DS sensors for layers 2 and 4.

- Double-Sided Double Metal Ladder Sensors (DSDM). Similar to the DS sensors, but the second layer has a 90° stereo angle, which requires an additional layer of metal strips ('Double Metal') on the n-type side arranged orthogonally to allow readout at the end of the ladder (rather than at the side). The p-type strips have a pitch of 50 μm, and the double metal n-type strips have a pitch of 153.5 μm. The DSDM ladders are the only ones that are constructed from a single 12 cm sensor. These sensors are used in the four inner barrels in layers 1 and 3.

- Single-Sided Wedge Sensors. The H-disks are composed of single-sided half-wedge sensors, which are attached back-to-back to provide a makeshift double-sided detector. The two sides have an effective stereo angle of 15°. Similar to the SS and DS ladders, the H-disk wedges are composed of a bonded pair of smaller sensors. The inner sensor is 7.63 cm long and the outer sensor is 6.33 cm long. Both sensors have a pitch of 40 μm with every second strip being read out.

- Double-Sided Wedge Sensors. The F-disks use double-sided sensors similar to the DS ladder sensors. The two sides have a 30° stereo angle and are composed of a single sensor 7.93 cm in length. The p-type side has a 50μm

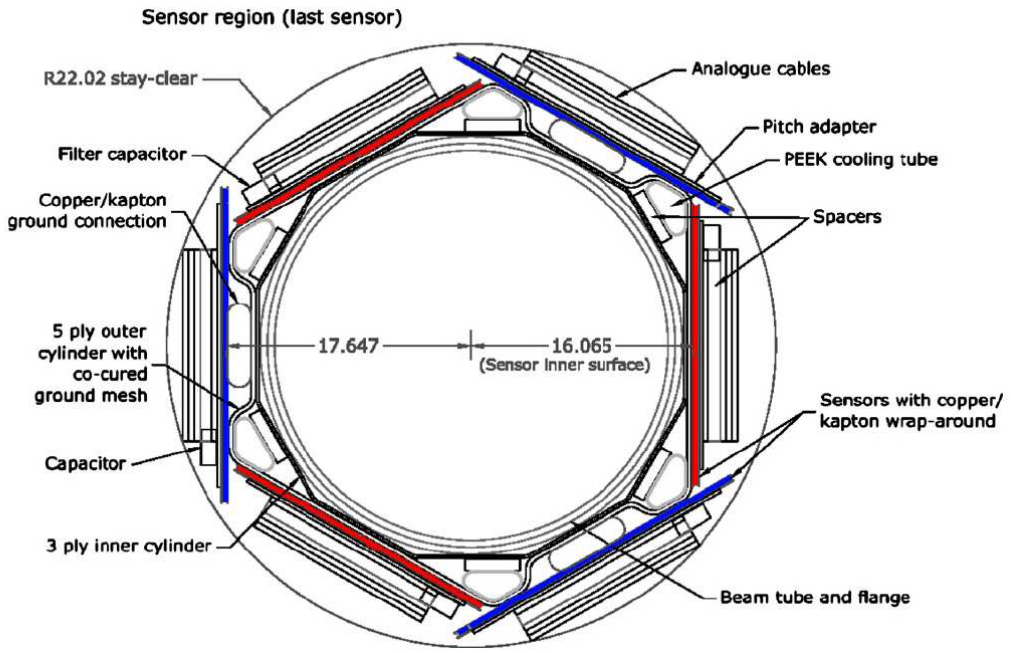

Figure 3.5: Cross section of the Layer 0 detector [45]. The A-layer sensors are at the right and upper/lower left, the B-layer sensors are at the left and upper/lower right.

pitch, and the n-type side has a 62.5 μm pitch.

The silicon detector was upgraded for Run IIb with an additional layer of sensors close to the beam pipe. The beam pipe itself was replaced with a new smaller diameter Beryllium pipe (29.5 mm compared to 38.1 mm), providing additional space that has been exploited by the new silicon layer ('Layer 0' or L0). The detector occupies the space between $r = 15.34$ mm and $r = 22.02$ mm (see Fig. 3.5). It is 77 cm long and arranged with a pair of overlapping sub-layers, each with sensors at three equally spaced azimuthal angles. The inner sub-layer (A-layer) uses single-sided sensors with a 71 μm pitch, and the outer sub-layer (B-layer) uses single-sided sensors with an 81 μm pitch. Each of the six azimuthal detector panels

is composed of eight sensors, the inner four being 7 cm long, with the outer four being 12 cm long. Cosmic ray muons were used to estimate the improvement to the track resolution. Tracks with a L0 hit were studied, and the impact parameter resolution[4] was determined to be 26.5 μm. Using the same tracks, but excluding the L0 hit from the reconstruction, the resolution was found to be 41.2 μm [44–46].

Central Fiber Tracker

The CFT is composed of eight cylinders of scintillating fibers. With the exception of the inner two, the cylinders are 2.52 m in length. The inner two cylinders are shorter (1.66 m) to accommodate the SMT H-disks. The cylinders are typically separated by $\sim$ 4.9 cm and occupy the volume from 20 to 52 cm radially. Each cylinder holds two doublet layers of fibers, one mounted axially and the other with a $\pm 3^\circ$ stereo angle. Odd numbered cylinders use a $+3^\circ$ stereo angle, while even cylinders use a -3° stereo angle. The doublet layers are separated by $\sim$ 2 mm. The first layer of fibers in a doublet is set in V-shaped grooves in the support structure with a separation determined by the number of fibers and the radius of the cylinder ($928 - 993$ μm). The second layer is set over top and offset by half of the fiber spacing such that the fibers are centred on the gap of the preceding layer. In order to achieve the desired tracking resolution (100 μm [47]), the position of the fibers must be known with an accuracy of 50 μm.

The scintillating fibers are composed, primarily, of polystyrene (PS). An ionizing particle passing through the fiber deposits energy, which causes excitations

[4]Determined from the width of the muon impact parameter distribution for tracks with $p_T < 5$ GeV/c.

in the PS (4.8 eV per excitation). The decay of the PS excited state produces a photon that can be detected; however, the decay time of PS is far too long to be useful for a high collision rate. By adding a second organic material, paraterphenyl (PTP), the PS excitations are transferred to the PTP by a non-radiative process. The PTP has a decay time of only a few nanoseconds and emits UV light (340 nm) in the process. To improve transmission through the PS fiber, a wavelength shifter (3-hydroxyflavone) is added, which absorbs the UV and re-emits with a 530 nm wavelength. The fibers are coupled at one end to a clear PS fiber waveguide, and the opposite end is mirrored with an aluminum coating. The fibers are 835 μm in diameter, including the 50 μm cladding. The waveguide fibers are coupled to solid-state photosensors ('visible light photon counters,' or VLPCs), which have a high quantum efficiency and are capable of detecting single photons. The VLPCs are housed in a liquid helium cryostat at 9 K. Taking into account the initial energy deposition in the fiber and all sources of losses en route to the VLPCs, typically $4 - 10$ photoelectrons are detected in a VLPC per fiber hit [44, 48, 49].

Preshower Detectors

The preshower detectors are intermediate detectors between the tracking system and the calorimeter. The addition of a solenoid magnet in the tracking volume for Run II increased the amount of material in front of the calorimeter. Energy loss in this material degrades the performance of the calorimeter system and must be corrected by including a sampling layer within the tracking volume (see Sec. 3.2.3 for a discussion of calorimetry). The preshower detectors serve as a final layer for tracking, but also as the early sampling layer for the calorimeter. It is composed of

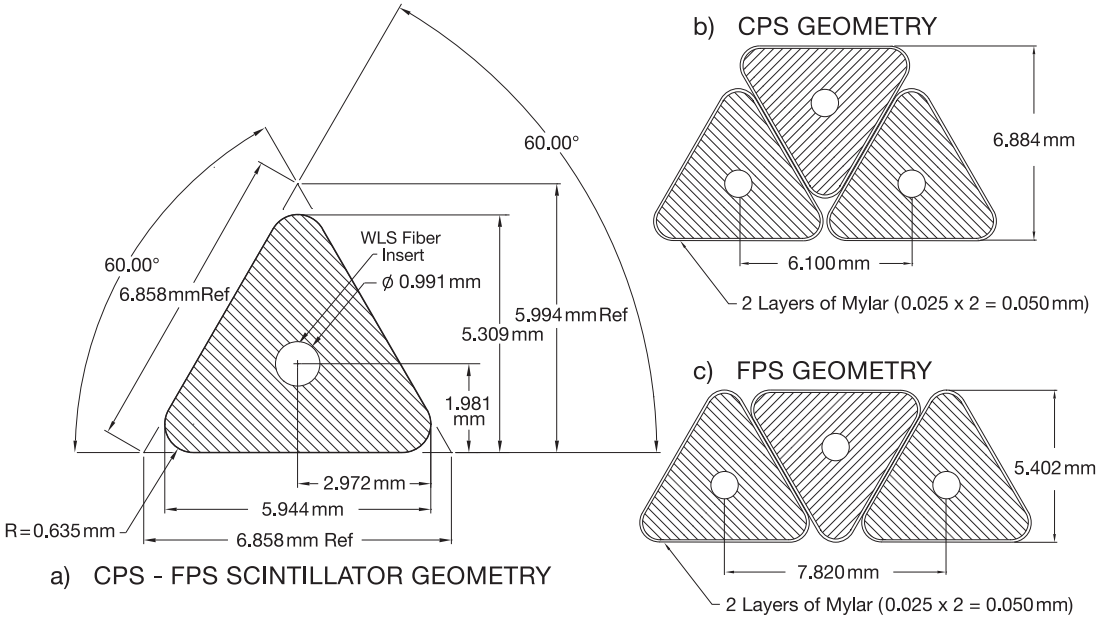

Figure 3.6: *a*) Cross sectional view of the triangular scintillator strips used in the preshower detector. The strip configurations for the central and forward detectors are shown in *b*) and *c*), respectively [44].

triangular strips of scintillator wrapped in aluminized mylar, as shown in Fig. 3.6. The triangular geometry of the scintillator strips is chosen to allow good position resolution without a high channel count [**?**, **?**]. The thickness of the scintillator allows sufficient energy to be deposited for the preshower to act as a calorimeter sampling layer. However, this shape is not well suited to act as a light guide. A wavelength shifting fiber (WLS fiber) is embedded in the centre of the strips in order to collect and carry light to the end, where it is collected and transmitted to VLPCs by waveguide fibers [44, 50, 51].

The central preshower (CPS) is arranged in three cylindrical layers. Each layer is a doublet of triangular strips as shown in Fig. 3.6*b*). The strips of the first CPS layer strips are aligned with the beam axis. The second and third layers are set

at a $+24^\circ$ and -24° stereo angle, respectively. A lead absorber ~ 5.6 mm thick (~ 1 radiation length (X_0)) is placed between the solenoid and the first CPS layer, providing a total $\sim 2X_0$ of material in front of the CPS. The forward preshower (FPS) detectors are composed of four layers with a $2X_0$ lead absorber between the second and third layers to match the radiation thickness of the solenoid and absorber used in the CPS. Each layer, including the absorber layer, consists of eight wedge-shaped modules. The first and second layers, as well as the absorber layer, have a smaller outer radius than the third and fourth layers (note the change in thickness in Fig. 3.3). The inner portion covers the region $1.65 < |\eta| < 2.50$ and the outer portion covers the region $1.50 < |\eta| < 1.65$. A particle hitting the outer portion of the FPS will have traversed the solenoid at a large incident angle, providing up to $3X_0$ of material and so no additional absorber layer is required. The first and last pairs of scintillator layers are arranged with a 22.5° stereo angle [50–53].

3.2.3 Calorimeter

While a calorimeter and a tracking detector function on similar principles (ionizing radiation), the measurement philosophies are very different. A tracking detector attempts to passively observe a passing particle without interrupting its path. A calorimeter not only seeks to stop the particle, but to actively cause it to decay into a cascade of daughter particles. By absorbing the energy of the entire cascade, the energy of the initial particle may be determined.

The D0 calorimeter is shown in Fig. 3.7. It includes a central calorimeter (CC) covering the region $|\eta| \lesssim 1$ and a pair of end calorimeters (ECN -north, ECS

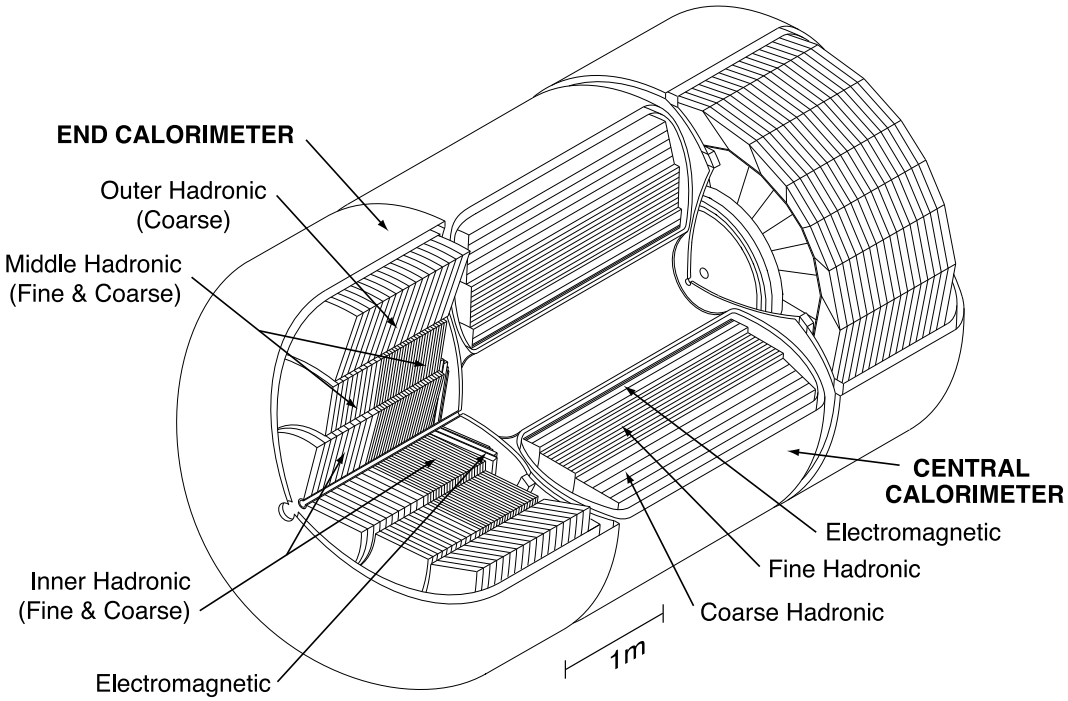

Figure 3.7: The D0 Calorimeter, with labeled central and end sections [44].

-south) extending the coverage to $|\eta| \approx 4$, as shown in Fig. 3.8.

Calorimetry

The cascade process (or shower) is slightly different for electrons and photons compared to heavier hadrons such as pions and kaons. The primary mechanism for energy loss in matter for electrons at the GeV scale is bremsstrahlung radiation of photons. For photons, the primary process is e^+e^- pair production. Hadrons, on the other hand, lose energy primarily by inelastic scattering, producing other secondary hadrons. In each case, the primary incident particle creates secondary particles with some fraction of the initial energy. The secondary particles behave in

a similar fashion, creating another generation of secondary particles propagating the shower deeper into the detector. Each new generation has a lower average energy than the one before. The shower continues until the energy drops below a threshold for the process to propagate further ($\sim$ 10 MeV for electromagnetic (EM) showers and $\sim$ 2 GeV for hadronic showers [54]). The depth of the shower is different for EM and hadronic showers. It is determined by the mean free path for the shower processes (or the cross section) and varies depending on the detector material. Propagation of electrons and photons is determined by the radiation length of the material (for example, $X_0 \sim 0.32$ cm in U^{238}). An EM shower will typically penetrate $14 - 16$ X_0 into the detector. The nuclear absorption cross-section determines the mean free path relevant for hadronic showers characterized by the nuclear interaction length λ ($\sim$ 10 cm in U^{238}). The depth of a hadronic shower is somewhat more complicated, but can be approximated by the relation:

$$L(95\%) = 0.56\lambda(\ln(E) + 4.1), \tag{3.2}$$

where E is the energy of the incident particle in GeV, and $L(95\%)$ is the depth for containment of 95% of the shower energy [54].

In addition to energy losses by the shower processes, small losses due to ionization allow the measurement of the propagation of the shower. A sampling calorimeter is constructed from alternating layers of absorber and active detector material in which ionization may be measured. The absorber layers promote the propagation of the shower while simultaneously limiting its depth. The detector layers sample the energy deposition at regular intervals allowing for determination

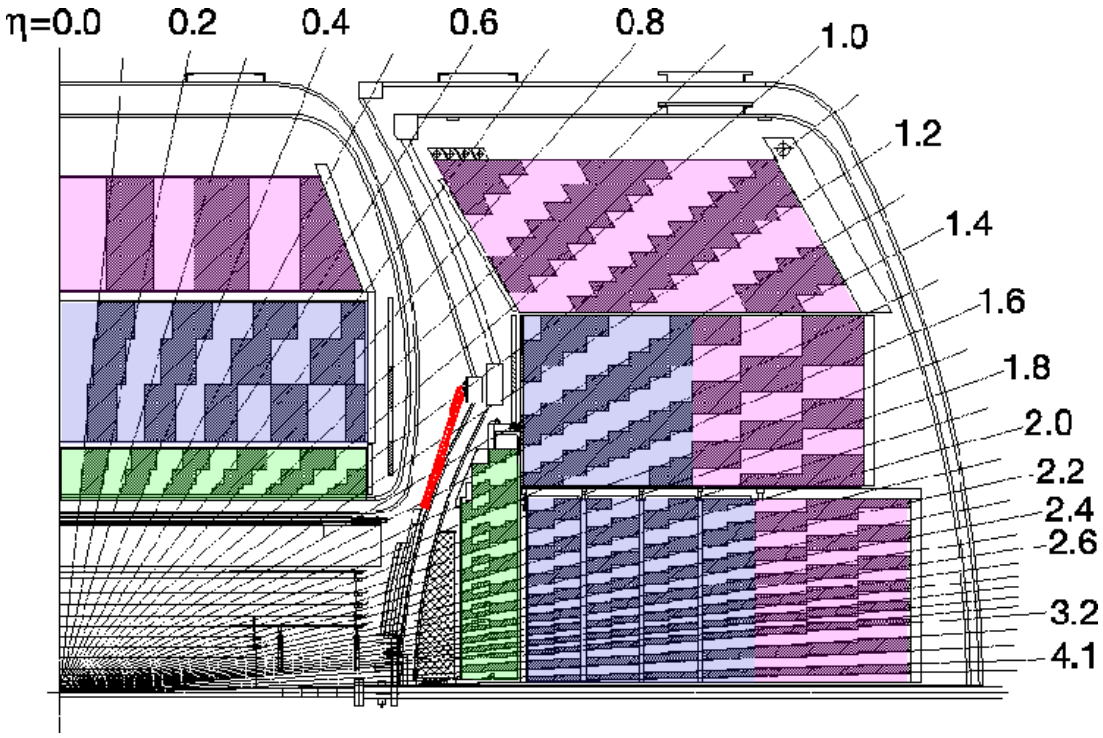

Figure 3.8: A quadrant view of the calorimeter system showing the pseudorapidity coverage [44]. The alternating dark and light regions indicate calorimeter cells which are grouped into readout 'towers.'

of the total energy. The D0 calorimeters primarily use uranium as an absorber material and liquid argon as the active medium. A copper pad is positioned between absorber layers in the liquid argon gap. The pads are grounded and the absorber plates maintained at ~ 2 kV. Ionization electrons will drift towards the pads producing a signal [55].

The shorter shower depth for EM showers requires a finer granularity in the initial part of the calorimeter. Each of the three D0 calorimeters (CC, ECN and ECS) is separated into an EM section with the finest granularity, an intermediate fine hadronic section (FH), and a final coarse hadronic section (CH). The pri-

mary difference in these modules is the thickness of the absorber material. The construction of each module is described below.

Central Calorimeter

The EM, FH and CH calorimeter sections in the CC are arranged as concentric cylindrical rings. The inner ring is the EM section and the outer ring is the CH section. The rings are divided into 32 (EM)/16 (FH,CH) long axial modules housing stacks of alternating absorber and signal boards. The module boundaries in each layer are rotated azimuthally so that no particle should pass through two module boundaries. The EM module uses 3 mm depleted uranium absorber plates while the FH uses 6 mm uranium-niobium alloy plates and the CH uses 46.5 mm copper plates. Between the absorber plates is a signal board segmented into tiles $\eta \approx 0.1$ and $\phi \approx 0.1$ in size. A 2.3 mm liquid argon gap separates the signal board from the absorber on either side. Tiles at similar η and ϕ are grouped together to form a 'tower' of cells with a single readout channel (See Fig. 3.8).

The EM modules have four layers of readout towers with depths $1.4X_0$, $2.0X_0$, $6.8X_0$ and $9.8X_0$ for a total thickness of $\sim 20X_0$ or 0.76λ. The FH modules have three layers of towers with depths 1.3λ, 1.0λ and 0.9λ. The CH module includes a single layer of towers 3.2λ in depth [44, 55].

End Calorimeter

Each of the end calorimeters includes an EM calorimeter disk in front of a hadronic system composed of a small angle cylinder with a pair of concentric cylindrical rings around it (see Fig. 3.8). The EM disk and the inner hadronic section are each a

single module composed of alternating absorber plates and signal boards in the $\phi - z$ plane. The middle hadronic ring is composed of 16 wedge shaped modules arranged azimuthally around z with detector surfaces in the $\phi - z$ plane. Both the inner and middle hadronic sections include both FH and CH segmentations. The outer cylinder is composed of 16 CH wedge modules inclined with a 60° angle towards the interaction region. The EM module uses 4 mm depleted uranium absorber plates. FH sections use 6 mm uranium-niobium plates and the CH sections use 46.5 mm stainless steel plates. Cell construction and segmentation is the same as in the CC.

As in the CC, there are four layers of EM readout towers with depths $1.6X_0$[5], $2.6X_0$, $7.9X_0$ and $9.3X_0$. The FH sections of the inner and middle cylinders include four equal layers of readout towers 1.1λ and 0.9λ deep, respectively. The CH sections of the inner, middle and outer cylinders all contain a single layer of towers 4.1λ, 4.4λ and 6.0λ (maximum) deep, respectively [44, 55].

Intercryostat Detector

The region between the modules of the CC and EC cryostats ($0.8 < |\eta| < 1.4$) is not well instrumented, leading to a loss in energy resolution. In an effort to improve the sampling in this region, sampling layers are added in the air gap between the cryomodules as well as on the wall inside the cryostat. The intercryostat detector (ICD) is a scintillator detector attached outside of the EC. The ICD is a ring segmented into wedge tiles $\eta \times \phi \approx 0.1 \times 0.1$ and coupled to waveguide fibers leading to PMTs for readout. Inside the CC and EC cryostat, a single additional

[5]Primarily due to the cryostat wall, the absorber plates make up only $0.3X_0$ of this.

layer of calorimeter readout cells is positioned next to the cryostat wall. These cells lack absorber material, and are referred to as the 'massless gaps' [44, 55].

3.2.4 Muon System

Determination of muon 4-vectors is done primarily by the central tracking system. However, an outer muon system provides muon identification, as well as a fast muon trigger. The D0 muon system is composed of an array of drift tubes and scintillator counters surrounding a 1.8 T toroidal magnet. The toroid system allows a rough determination of the momentum of the muon candidate in order better match it to a central track. Drift tubes before and after the toroid provide tracking information for muons curving in the magnetic field, but the relatively slow drift time ($\sim$ 500 ns) excludes their use in triggering. The scintillator counters are much faster by comparison ($\lesssim$ 100 ns) with good timing resolution ($\sim$ 2 ns), allowing accurate drift-time determination in the drift tubes.

Central Muon System

The central muon system ($|\eta| \lesssim 1.0$) includes three layers of proportional drift tube detectors (PDTs), two layers of scintillator counters and a 1 m thick iron toroid magnet. Of the three PDT layers, one is positioned before the toroid (A layer) and two are after the toroid (B and C layers). Successive layers are separated by $\sim$ 1 m, allowing measurement of the curvature due to the toroid field. Each layer is composed of three or four planes of drift cells, allowing the determination of the muon direction in each layer in addition to the position. Each cell is $191 - 579$ cm long with a 5.5×10.1 cm^2 rectangular cross section. An anode wire is held

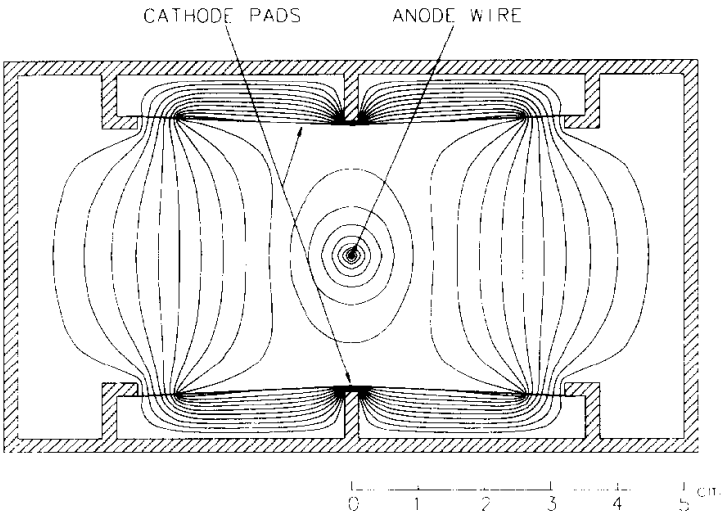

Figure 3.9: Cross sectional view of a drift tube cell for the central muon system. Equipotential lines are shown surrounding the anode and cathode [55].

in the cell centre with cathode pads at the top and bottom (See Fig. 3.9). The anode wire is maintained at +4.7 kV, the cathode at +2.3 kV and the cell wall is grounded.

The gas mixture in the cell ($Ar(84\%)/CF_4(8\%)/CO_2(8\%)$) is ionized by a passing muon, with anions drifting to the anode wire and cations drifting to the cathode. The drift time of anions to the anode wire is used to measure the transverse coordinate of the hit. The further the muon from the anode wire, the longer the drift time. The anode wires of each cell are read out from one end, with the other end connected to the neighbouring cell (azimuthally). A hit in one cell will be registered in both cells with some time difference determined by the longitudinal distance from the connection between the cells. Each cathode pad is composed of two separate electrodes with the boundary forming a repeating diamond pattern

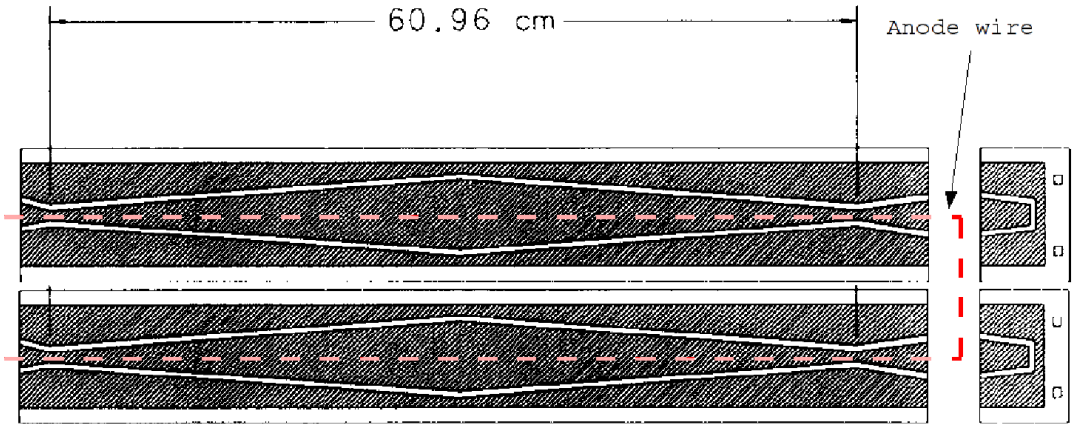

Figure 3.10: $\phi - z$ cross section of a drift tube cell showing the shape of the cathode electrodes and the connection of the anode wires in neighbouring cells [55].

$\sim$ 61 cm long (see Fig. 3.10). The ratio of the charge collected on the inner and outer electrodes determines the longitudinal hit coordinate within the diamond pattern. When used together, the anode time difference and the cathode charge ratio work similar to a vernier scale to measure the longitudinal hit coordinate in the tube. The time difference determines the position to within $10 - 20$ cm, and the charge ratio refines this to 5 mm. The transverse resolution due to the drift time is $\sim$ 1 mm [55–57].

The scintillation counters form the first and last detector layers for the muon system. They are used for triggering muons and providing a time reference for determining drift time in the drift tubes. Both layers are composed of scintillator tiles with photomultipliers mounted directly on the scintillator modules. This construction allows for very short waveguide fibers, reducing the light attenuation. The design of each layer is slightly different, but both use Bicron 404A commercial scintillator. The outer layer is composed of tiles of scintillator 63.5 cm wide and

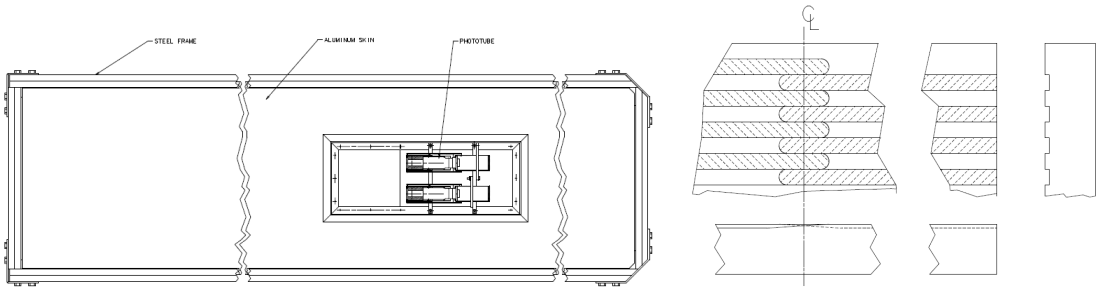

Figure 3.11: Top view of a muon scintillation counter (left) and a top, side and end schematic of the tile scintillator (right) [58]. The cross-hatched regions show grooves cut into the surface for waveguides.

$207 - 287$ cm long. Each tile has shallow grooves cut on the surface, 4 mm wide and 8 mm apart. The grooves run parallel to the long edge from either end to the tile centre (see Fig. 3.11). Each groove if fitted with four wavelength shifting fibers meeting at the tile centre, where they are coupled to a photomultiplier tube. Grooves on either half of the tile are offset so that fibers do not overlap at the centre. The inner scintillator layer is constructed in the same fashion using tiles $23.1 - 36.7$ cm wide and 84.5 cm long. Fiber grooves in this case are cut 6 mm deep, 1 mm wide, and 45 mm apart with six wavelength shifting fibers stacked in the groove [44, 56, 58].

Forward Muon System

The two forward muon systems ($1 \lesssim |\eta| \lesssim 2$) are similar to the central muon, but use mini drift tube detectors (MDTs) in place of PDTs and include three layers of scintillator counters. The MDTs are essentially just miniaturized versions of the PDTs. The smaller size means shorter drift time for ions in the cell (maximum 60 ns). As with the PDTs, each MDT layer is composed of three or four planes of drift

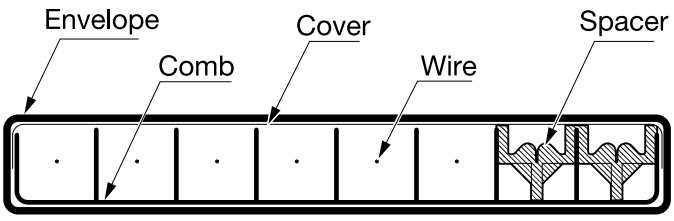

Figure 3.12: Cross sectional view of an eight cell mini drift tube [44].

tubes. Each layer forms a square planar ring divided into octants with drift tubes oriented parallel to the toroidal field. Each MDT layer is separated by ~ 1 m and MDT planes within a layer are ~ 3 cm apart. Tubes are constructed in modules of eight drift cells separated by a comb-shaped aluminum spacer (see Fig. 3.12). Each cell has a 1 cm^2 cross section with a gold-plated tungsten anode wire running the length of the cell. The cell uses a gas mixture composed of $CF_4(90\%)/CH_4(10\%)$ with the cathode wall maintained at -3.2 kV and the anode wire grounded. The transverse resolution due to the drift time is ~ 1 mm; however, in contrast to the PDTs, the MDTs have no means of determining the hit coordinate along the wire (i.e., ϕ coordinate) [44, 56, 59].

The scintillation counters in the forward region serve the same basic function as the central counters. The system is composed of three layers of trapezoidal panels ('pixels') arranged in concentric rings in the $r - \phi$ plane. Each layer is mated to a MDT layer and provides the ϕ hit coordinate required for the track reconstruction. The pixel segmentation varies from $\Delta\eta \times \Delta\phi \approx 0.12 \times 4.5^\circ$ for small r, to $\Delta\eta \times \Delta\phi \approx 0.07 \times 4.5^\circ$ for large r [6]. Each pixel is composed of a Bicron 404A scintillator plate with a pair of wavelength shifter bars on two sides. The

[6]The smaller segmentation at large r is due to the nonlinearity of η, the pixels are actually larger in this region.

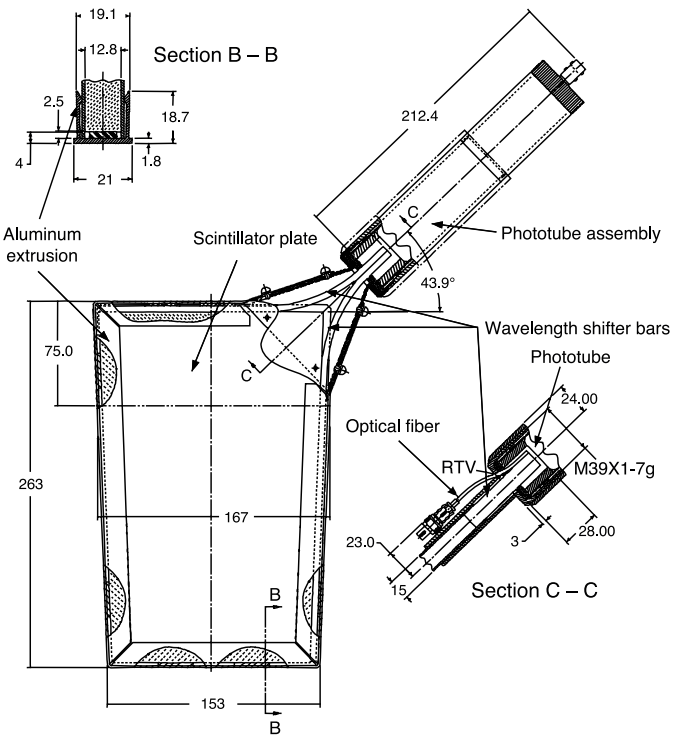

Figure 3.13: An example of a scintillator 'pixel' counter. Dimensions are in mm [44].

bars meet at the corner of the plate and are coupled to a photomultiplier tube (see Fig. 3.13) [44, 56].

3.2.5 Trigger and Data Acquisition

The average collision rate at the Tevatron is ~ 1.7 MHz[7]. The bandwidth, storage and processor time required to manage such a data rate is far greater than we could realistically achieve. Furthermore, the majority of these collisions are uninteresting

[7]The minimum crossing time is 396 ns. However the 36 proton/antiproton bunches are arranged into three super bunches (of 12 bunches each) separated by 2.6 μs, giving an average collision rate of 1.7 MHz.

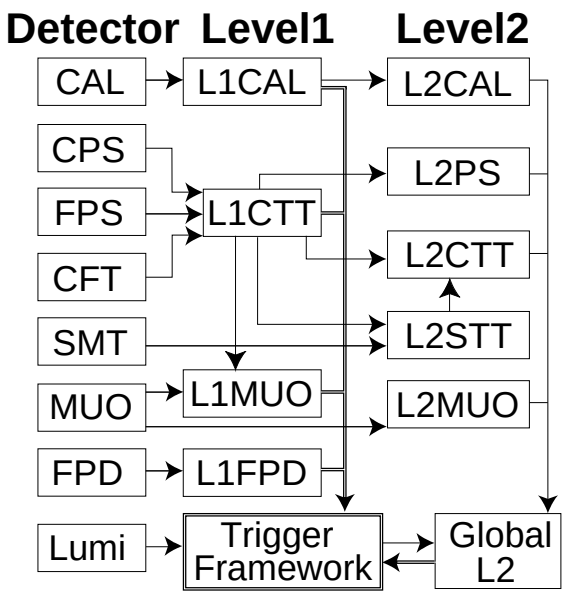

Figure 3.14: Diagram of the D0 L1 and L2 trigger system showing data flow between the various components [44].

processes such as soft scattering or light flavour creation. In order to pick out the interesting physics events, we 'trigger' the data acquisition system on specific decay signatures. These events will be recorded and considered further. The analyses documented herein do not require that any specific trigger be fired; however, most events are collected by muon triggers.

At D0, we use a three level triggering system (L1-L3). Events passed by the L1 trigger are passed to the L2 trigger for further consideration. Events accepted at L2 are then passed to the L3 trigger and, if accepted at L3, are then recorded. At each successive level, the data rate is further reduced. This allows time for increasingly complex algorithms in the event analysis at each trigger level. The most stringent constraints are on the L1 trigger, which must make a decision within

3.5 μs of the beam crossing and reduce the rate to 2 kHz. The L2 system further refines the criteria, reducing the rate to 1 kHz. The L3 system is similar to the offline reconstruction and has an output rate of 50 Hz.

There are three main L1 trigger systems: L1CAL (Calorimeter), L1CTT (Central Track Trigger) and L1MUO (Muon) (see Fig. 3.14)[8]. L1CAL groups hardware towers into 2×2 'trigger towers' with dimensions $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$. The trigger looks for high energy deposits in these towers as well as considering the total energy deposition in all towers. The Run IIb trigger includes new tower clustering algorithms for finding jets that deposit energy in several trigger towers [60].

The L1CTT system combines data from the CFT, the CPS and the FPS. Cluster hits from these systems are used to construct tracks. Isolated and high-p_T tracks are triggered and the track list is provided to the L1MUO system for track matching. Run IIb includes a new L1CalTrk system that matches L1CTT tracks to towers in L1CAL for triggering [60].

Most B_s^0 candidates are triggered by the muon system. The configuration of the muon drift tubes allows the determination of directional vectors (called 'stubs') in each layer. L1MUO matches the stubs to hits in the tile scintillators. These candidates are then matched to tracks from L1CTT to form a L1 muon candidate. In addition, the scintillator hit timing is required to coincide with the beam collision to reject cosmic rays.

The conditions in the L1 systems can be combined in a variety of ways to optimize the trigger performance. For example, we can trigger on a single high-p_T muon without a matched track, or we can trigger on a low-p_T muon that has

[8]There is also an L1 trigger for the forward proton detector (L1FPD).

been matched. These conditions are called 'trigger terms,' and are programmed in custom hardware for each system. There are 256 trigger terms that are combined in AND-OR combinations to create triggers for specific physics signatures. A Trigger Framework (TFW) receives the trigger term signals of each system and performs the logical AND-OR to determine if an event has passed some trigger condition [44].

The L2 trigger system is similar to the L1 trigger, but implements more sophisticated algorithms for reconstructing tracks, jets and leptons. The L2 trigger includes the L2CAL, L2CTT and L2MUO systems, which function in a similar fashion as their L1 counterparts. In addition, there are L2PS (PreShower), L2STT (Silicon Track Trigger) and L2Global systems. The L2Global system receives the fired L1 triggers and matches these to L2 triggers that must be passed for an accept to be issued[9].

The L3 trigger system is implemented entirely in software and is based on reconstructed physics objects, rather than signals in the detector subsystems. For example, there are algorithms for L3 jets, missing energy and electrons, but no general L3CAL system. Physics objects are analyzed using almost the entirety of the detector data. More precise tracking algorithms make better determinations of the momentum and positions of tracks. More precise timing information is used by the muon system to reject cosmic ray events. Events passing the L3 trigger are written to tape for storage and later offline processing.

[9]Note that this means that an L2 trigger cannot be passed if a related L1 trigger did not fire. For example, if there is no L1MUO trigger, L2MUO will not even be considered.

3.3 Monte Carlo Tools

The simulation of physics processes is an important part of this analysis. Simulations of the production, decay and detection of $b\bar{b}$ events is the only way to determine many of the input parameters for this work. A variety of software tools are employed to acquire the simulated data required for these studies.

The $b\bar{b}$ production process is simulated by Pythia 6.4 [61]. Pythia is an 'event generator' that simulates the strong interactions between the colliding partons in order to create a $b\bar{b}$ pair with momentum and rapidity consistent with the process in question. In addition, it simulates the hadronization process by which the $b\bar{b}$ pair produce a B_s^0 and/or other B hadrons. The simulation is potentially very complicated and involves a large number of hadronization particles (pions and kaons, mostly) that are produced alongside the B hadrons.

The primary particles (i.e., particles produced directly in the hard scatter of the $p\bar{p}$ collision) generated by Pythia are taken as input by the EvtGen event generator [62]. This is a specialized tool for the simulation of decays of B hadrons. EvtGen simulates the decay sequences of all particles in the event, correctly modeling their lifetimes and branching fractions to various requested final states. EvtGen uses measured branching fractions where possible and uses theoretically predicted fractions when measurements are not available.

In many cases this is all the simulation that is required. However, it is sometimes necessary to simulate how the detector will respond to the decay process. Geant is a simulation of how particles interact with matter [63]. A simulation of

the D0 detector properties (material, geometry etc.) is implemented in Geant[10] so that it models how the decay products of the $b\bar{b}$ event will interact with the detector material. The response of the detector to the interactions is modelled (for example, charge deposition in the silicon, showering in the calorimeter, ionization in the drift tubes, etc.) to generate a simulated data event, in the same format as the data read out from the D0 detector.

In a real $p\bar{p}$ collision it is common to have multiple interactions occur simultaneously[11]. These additional events are typically light flavour production or elastic scattering and are not of interest. They do, however, add background that makes reconstruction of the interesting events more challenging. To model these underlying events in the detector simulation, we overlay 'minimum bias' data over the simulated $b\bar{b}$ process. The minimum bias events are real data events that are collected with a non-biasing trigger (i.e., all events are triggered).

The simulated detector response is then processed by the standard D0 event reconstruction suite (D0Reco) in the same manner as real data. The reconstructed Monte Carlo (MC) still carries along with it additional 'truth' information (i.e., the particle type and four momenta) that can be analyzed to compare the detector response to the real event properties.

[10]D0 has its own version of Geant with the simulation of the D0 detector called d0gstar (**D0** **G**eant **S**imulation of the **T**otal **A**pparatus **R**esponse). This is based on Geant 3.

[11]The total $p\bar{p}$ cross section at $\sqrt{s} = 2$ TeV is ~ 80 mb [9] and the instantaneous luminosity is $\sim 100\ \mu\text{b}^{-1}\text{s}^{-1}$ giving an event rate of $\sim 80\ 000\ \text{s}^{-1}$. A crossing time of 396 ns means ~ 3 interactions per crossing.

Chapter 4

Event Selection

4.1 Overview

Once physics objects have been reconstructed from the raw detector data, candidate B_s^0 events must be selected and isolated from background processes. The selection criteria must not only be efficient at accepting signal events, but must also reject or discriminate against background events. There are two categories of background that we must consider: physics background consisting of specific decay processes that mimic the B_s^0 signal and combinatoric background consisting of misreconstructed decays.

This chapter will describe the signal selection techniques used in these analyses and also outline the method of identifying mixed decays. This analysis uses a combination of linear cuts and non-linear selection criteria for signal selection and background rejection. The linear cuts described in Sec. 4.1.1 serve as a rough filter to reduce the contribution from decay processes with final-states similar to the

desired B_s^0 decay. The non-linear likelihood ratio method described in Sec. 4.1.2 discriminates between B hadrons and the combinatoric background.

4.1.1 Selection Criteria

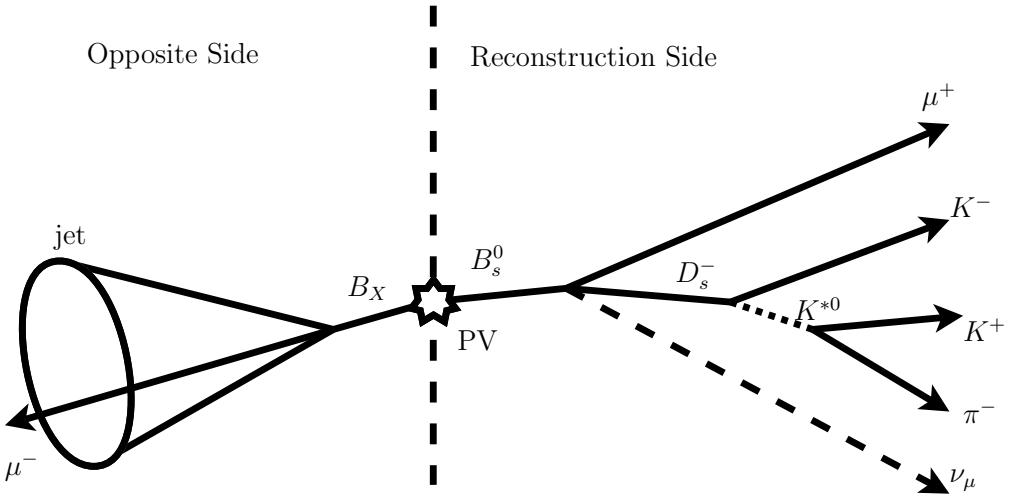

Figure 4.1: Diagram depicting the decay of the B_s^0 meson on the 'reconstruction side,' and the semi-muonic (in this example) decay of the other B hadron on the 'opposite side' of the primary vertex (PV). The opposite side decay process is used to determine the initial flavour of the B_s^0 meson (see Sec. 4.2).

These analyses search for B_s^0 mesons in the $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ decay process with $D_s^- \to K^{*0} K^-$ and $K^{*0} \to K^+ \pi^-$ (see Fig. 4.1). The final-state particles of this process ($K^+ K^- \pi^-$) are used to reconstruct the momentum and trajectories of the intermediate D_s^- and K^{*0} mesons. These, in combination with the final-state muon, are used to reconstruct the momentum and decay length of the B_s^0 meson. The procedure for selecting these events is outlined below.

The clearest signature of a semi-muonic B_s^0 decay is the muon. As such, we begin our data sample selection by including all recorded D0 events with at least

one muon. We require no specific trigger be fired; however, most of the data are collected by single muon triggers. These triggers typically require a muon have $p_T > 3$ GeV/c; however, the final reconstructed momentum may be lower.

We require that the reconstructed muon have a hit in the first muon drift tube layer and a hit in either the second or third layers (or both). These hits are used to construct a muon 'stub' that must be matched to a track in the central tracking system. To guarantee good momentum resolution, the central track is required to have at least two hits in the axial SMT strips and the axial CFT layers[1]. It is required to have $p_T > 2$ GeV/c and $p > 3$ GeV/c.

We next look for the three final-state particles from the $D_s^- \rightarrow K^{*0}K^- \rightarrow K^+\pi^-K^-$ decay. The candidates are assigned kaon and pion masses based on their charges and momenta, but these not be identified with certainty[2] Misidentified pions (mistaken for kaons) are one of the largest contributions to the background and will be discussed in detail in Sec. 6.1.1. The candidate (K_1^+) with the same charge as the muon is assigned a kaon mass and required to have $p_T > 0.9$ GeV/c. The remaining kaon is required to have $p_T > 1.8$ GeV/c and the pion $p_T > 0.5$ GeV/c. In the event that both tracks have $p_T > 1.8$ GeV/c, the particle ID is ambiguous. In this case both mass assignments are considered with one likely to be ruled out by later cuts. Very occasionally ($\sim 0.2\%$ of events) this ambiguity remains after all cuts are applied and is resolved by choosing one candidate at random. If for any other reason an event has two B_s^0 candidates, one is chosen at random.

[1]Only axial strips/layers are used to determine momentum.
[2]The D0 detector does not have the capacity to distinguish kaons from pions.

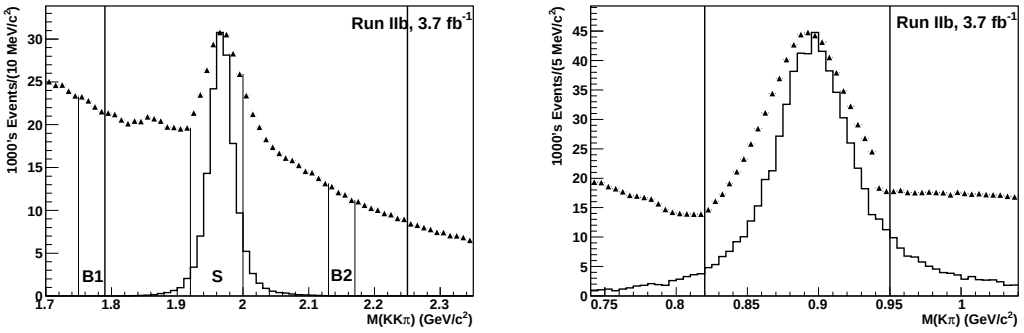

Figure 4.2: Distribution of the $KK\pi$ mass (left) and the $K\pi$ mass (right). The points represent Run IIb data and the histogram is a normalized MC simulation of the signal events. The bands labeled S, $B1$ and $B2$ are regions used to model the likelihood variable (see Sec. 4.1.2). The mass cuts are indicated by the vertical lines. The background from misidentified pions causes the broadening of the D_s^- mass peak.

The remaining criteria are used to control event quality and reject physics and combinatoric backgrounds. Hadronization pions and kaons from the primary vertex may give a pseudo-vertex that can be mistaken for a D_s^- decay vertex. This contribution can be reduced by imposing an impact parameter[3] significance cut on the final-state particles. We require that the kaons satisfy the condition $(\epsilon_T/\sigma_T)^2 + (\epsilon_L/\sigma_L)^2 > 4$, where ϵ_T/ϵ_L and σ_T/σ_L are the transverse/longitudinal projections of the impact parameter and its error, respectively. In addition, the three charged particles are required to form a D_s^- vertex with a vertex fit $\chi^2 < 16$. The subsequent B_s^0 vertex with the D_s^- candidate and the muon is required to have a $\chi^2 < 9$.

Mass cuts are used to eliminate combinatoric background and non-resonant physics background processes. The $KK\pi$ system is required to have a mass in

[3]Unless stated otherwise, the impact parameter is with respect to the primary vertex.

the interval $1.79 < M(K^+K^-\pi^-) < 2.25$ GeV/c^2, consistent with a D_s^- meson. The mass window is considerably wider than the width of the D_s^- mass peak so that the wider physics backgrounds will be properly described in the fit to the D_s^- mass (see Fig. 4.2). The invariant mass of the $K^+\pi^-$ system provides a strong discriminator against the non-resonant $D^- \to K^+\pi^-\pi^-$ physics background. The requirement $0.82 < M(K^+\pi^-) < 0.95$ GeV/c^2, consistent with the K^{*0} mass, reduces this background by $\sim 80\%$. The $\mu^+ D_s^-$ system is required to have a mass in the interval $2.6 < M(\mu^+ D_s^-) < 5.4$ GeV/c^2, consistent with a B_s^0 meson. Missing energy from the decays of the B_s^0 and D_s^- candidates lead to a broadening of the B_s^0 mass distribution (see Sec. 5.4). The cut on $M(\mu^+ D_s^-)$ is chosen to accept as much of the B_s^0 signal as possible (98.9% efficient).

Decay length significance cuts are used to reduce the contribution from 'prompt' background, where the muon originates from the decay of a c or other light quark. These muons originate close to the primary vertex and when matched with the D_s^- candidate yield a B_s^0 candidate with a very short decay length (hence the name 'prompt'). The transverse separation of the D_s^- vertex from the primary vertex (d_T^D, see Fig. 4.3) is required to satisfy the condition $d_T^D/\sigma_T^D > 4$, where σ_T^D is the uncertainty on the separation (event-by-event). In addition, we require that the D_s^- and B_s^0 trajectories be in the same general direction. Specifically, we require that the angle α_T^D between the D_s^- momentum and the displacement d_T^D satisfy the condition $\cos\alpha_T^D > 0.9$. A similar requirement is applied to the B_s^0 decay vertex. We define an angle α_T^B between the B_s^0 momentum and the vector from the primary vertex to the B_s^0 decay vertex. We require this angle satisfy the condition $\cos\alpha_T^B > 0.95$, but only in events with a highly significant B_s^0 decay length (i.e.,

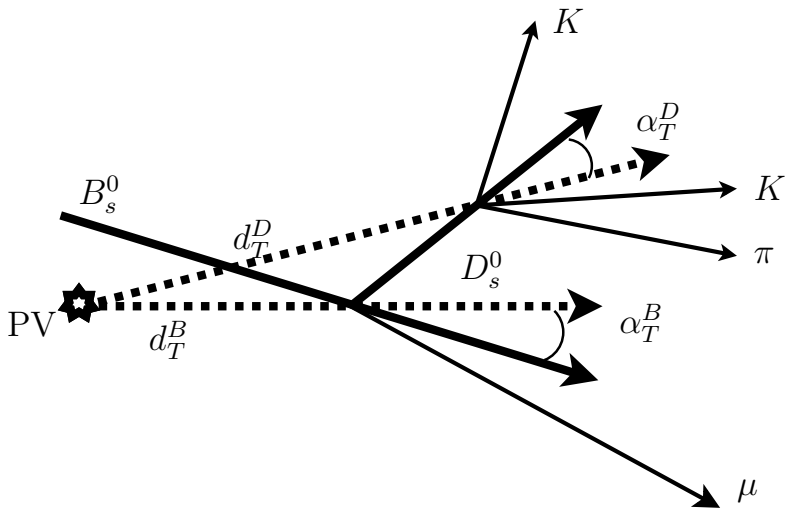

Figure 4.3: Diagram showing the transverse projection of the B_s^0 decay. The vectors for the B_s^0 and D_s^- mesons have been extended beyond their decay vertices and the primary vertex has been connected with the decay vertices (dashed lines). The angles subtended between these lines are the angles α_T^D and α_T^B. The impact parameter of the B_s^0 meson has been exaggerated to make the angle clear.

$d_T^B/\sigma_T^B > 4$). Finally, we require that $d_T^D > d_T^B$, unless the separation between the B_s^0 and D_s^- vertices has a low significance (i.e., $d_T^{BD}/\sigma_T^{BD} < 2$).

4.1.2 Likelihood Ratio

The likelihood ratio selection is intended to identify B hadrons and reject combinatoric background. However, its ability to differentiate B_s^0 mesons from other B hadrons is very limited. This technique is based upon properties of B hadrons that are shared between B_s^0 and B^0 mesons making the distinction very difficult (see Fig. 4.5). The likelihood ratio is constructed from a series of discriminating variables $(x_1, \ldots, x_n)$ that have different *pdf*s for B hadrons ($f_i^s(x_i)$) and background

($f_i^b(x_i)$). These *pdf*s are used to define a combined b-tagging variable

$$y = \prod_{i=1}^{n} \frac{f_i^b(x_i)}{f_i^s(x_i)}. \tag{4.1}$$

The best signal selection can be made by cutting on the tagging variable $y < y_0$, for some value of y_0. Provided the discriminating variables x_i are independent, this cut will provide the best possible selection of the signal [64]. The discriminating variables are:

- The helicity angle between the D_s^- and K^+ meson from the K^{*0} decay, in the K^{*0} rest frame. Because the K^{*0} is produced in a polarized state, the K^+ and π^- are produced preferentially along the D_s^- axis.

- Isolation of the $\mu^+ D_s^-$ system, defined as $Iso = p(\mu D_s)/(p(\mu D_s) + \sum p_i)$, where the sum is over all tracks in the cone $\sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.5$ around the $\mu^+ D_s^-$ direction. The $\mu^+ K^+ K^- \pi^-$ tracks are excluded from the sum.

- Transverse momentum of the K^- candidate from the D_s^- decay.

- Invariant mass of the $\mu^+ D_s^-$ system.

- χ^2 of the D_s^- vertex fit.

- Invariant mass of the $K^+\pi^-$system.

The functions $f_i^s(x_i)$ and $f_i^b(x_i)$ are constructed from a subset of the Run IIa data (the first 610 pb^{-1}, see Fig. 4.4). Events passing the selection criteria described in Sec. 4.1.1 are divided into three bands according to the reconstructed

$KK\pi$ mass (see Fig. 4.2):

$$\begin{aligned}
&\text{Signal Band} \quad S \quad : 1.92 < M(KK\pi) < 2.00 \text{ GeV}/c^2, \\
&\text{Side-Band 1} \quad B1 \quad : 1.75 < M(KK\pi) < 1.79 \text{ GeV}/c^2, \\
&\text{Side-Band 2} \quad B2 \quad : 2.13 < M(KK\pi) < 2.17 \text{ GeV}/c^2.
\end{aligned}$$

Candidates in the two side bands are use to create the f_i^b *pdf*s. The signal f_i^s *pdf*s are constructed from events in the signal band after subtraction of the normalized f_i^b distributions. The cut value, y_0, is chosen to maximize the signal to background ratio $N_S/\sqrt{N_S + N_{B1} + N_{B2}}$, where N_S, N_{B1} and N_{B2} are the number of candidates in the S, $B1$ and $B2$ bands, respectively (background subtracted in the case of N_S). By convention, we define the combined tagging variable as $ctag = -\log_{10} y$ and make the cut $ctag > 0.16$. The distribution of this variable in signal and background is shown in Fig. 4.5. In order to simplify the notation, we will refer to the *ctag* parameter as y, with the logarithm implied.

4.2 Flavour Tagging

Of central importance to the B_s^0 mixing frequency measurement is the ability to distinguish an oscillated from a non-oscillated B_s^0 decay. This is important (though not critical) for the asymmetry measurement as well. A B_s^0 candidate is considered to have oscillated if its production flavour (b or $\bar{b}$) does not match its decay flavour. The decay flavour can be easily obtained from the charge of the muon from the B_s^0 decay. The production flavour cannot be determined directly and must be

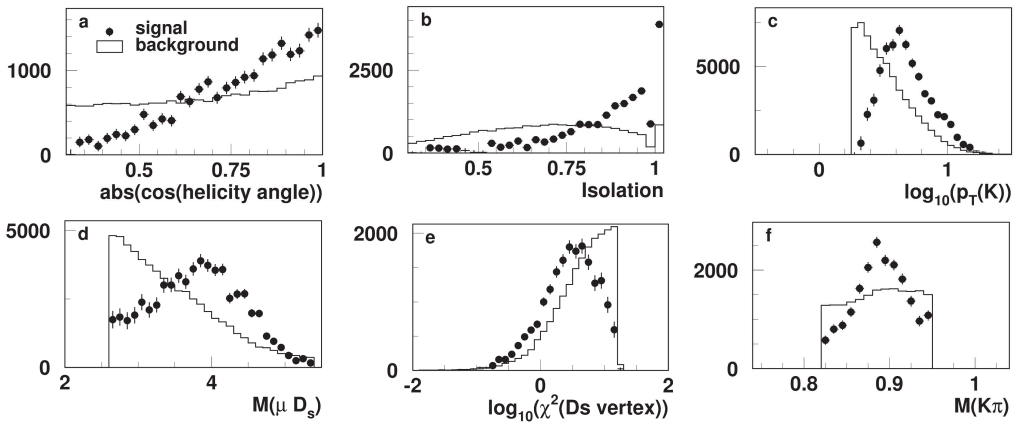

Figure 4.4: Distributions for the discriminating variables for signal (points) and background (histogram) used in the likelihood ratio: *a*) helicity angle, *b*) isolation, *c*) $p_T(K^-)$, *d*) the invariant mass of the μD_s system, *e*) the D_s^- vertex χ^2 and *f*), the invariant mass of the $K^+\pi^-$ system [1].

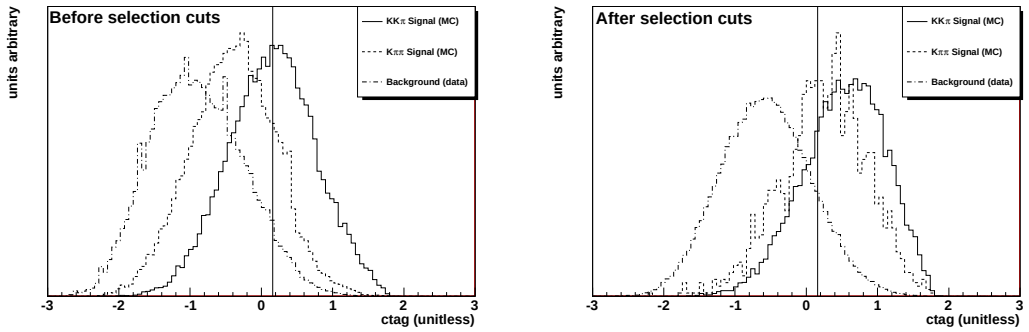

Figure 4.5: The distribution of the *ctag* variable before any selection cuts (left) and after the kinematic cuts (right). The vertical line represents the *ctag* cut.

inferred. Algorithms for determining the production flavour are referred to as 'flavour taggers'.

There are two classes of flavour-tagging algorithms. Opposite-Side Tagging (OST) makes use of information from the decay of the B hadron produced from the other b quark from the primary vertex (recall $b/\bar{b}$ production, Sec. 1.3). An

initial B_s^0 meson ($\bar{b}s$) must be accompanied on the opposite-side of the detector by some b hadron, while a $\bar{B}_s^0$ meson ($b\bar{s}$) must be accompanied by a $\bar{b}$ hadron. This method has the advantage of being completely independent of the B meson we are attempting to tag, since it is determined entirely from the opposite-side decay. Another method, called Same-Side Tagging (SST), makes use of the other hadronization particles from the primary vertex that accompany the B meson. A $\bar{B}_s^0$ meson ($b\bar{s}$), for example, will be produced with a 'spare' s quark, which is likely to produce a K^- or $\bar{K}^0$ meson. Similarly, a B_s^0 meson ($\bar{b}s$) will be produced with a leftover $\bar{s}$, producing a K^+ or K^0 meson. In the case of the charged kaon, the charge tells us the initial-state flavour of the B_s^0 meson.

The performance of a tagger is measured by its efficiency to make a flavour tag and the efficiency to make a correct flavour tag from the tagged events (called tagging purity). The tagging efficiency is $\epsilon = N_{tagged}/N_{total}$, and the tagging purity is defined as $\eta = N_{correct}/N_{tagged}$, where $N_{correct}$, N_{tagged} and N_{total} are the number of correctly tagged events, the number of tagged events and the total number of reconstructed B candidates, respectively. The dilution ($\mathcal{D}$) is then defined as the asymmetry between the number of correctly and incorrectly tagged events:

$$\mathcal{D} = \frac{N_{correct} - N_{incorrect}}{N_{correct} + N_{incorrect}} = 2\eta - 1. \tag{4.2}$$

The dilution quantifies the relative performance of the tagger, but does not account for the tagging efficiency. The 'tagging power' ($\epsilon\mathcal{D}^2$) measures the overall performance of the tagger. The dilution (and power) can be measured in some calibration data where we know the correct tag and used to predict the accuracy

when the tagger is applied to some other data. We also define a parameter that quantifies the confidence in the tag on an event-by-event basis. We call this parameter the *predicted dilution* and denote it d_{pr}. The precise definition of this parameter depends on the tagging algorithm. The calibration data are used to determine $\mathcal{D}$ in bins of the predicted dilution. In this context, d_{pr} is a variable returned by the tagging algorithm that predicts the tagging accuracy for a given event and $\mathcal{D}$ is the calibrated accuracy of the tagging algorithm. In general, the dependence of $\mathcal{D}$ on the parameter d_{pr} is not linear.

These tagging algorithms have been developed elsewhere [65] and are described below. In this work we will only make use of the OST method. The SST technique is problematic in this analysis due to the large background from charged B^+ meson decays where the charge-flavour correlation is opposite. The 'spare' $\bar{u}$ quark from the B^+ meson ($\bar{b}u$) decay produces a π^- or π^0. The SST algorithm would incorrectly tag this as a b hadron. This background cannot be completely separated from the B_s^0 signal and will give an false B_s^0 oscillation signal.

4.2.1 Opposite-Side Tagging

The OST algorithms look at the charge of the objects on the opposite-side of the detector from the B hadron we are reconstructing. The muon tagger looks for a muon where the angle between the directions of the muon and the B candidate satisfy the condition $\cos\phi_\mu^B < 0.8$. The muon jet charge (Q_J^μ) is then defined using all tracks within a cone around the muon ($\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$). We define an electron tagger in a similar manner using an electron jet charge (Q_J^e). In both

cases, the jet charge is defined as

$$Q_J^{\mu/e} = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}, \tag{4.3}$$

where the sum is over all tracks within a cone around (and including) the muon/electron with $\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$. Tracks from the reconstructed B candidate are excluded. The muon and electron taggers are the highest performing OST taggers[4], but are restricted to semileptonic decays and hence suffer from a low tagging efficiency [65].

Opposite-side hadronic decays lack a muon or electron and therefore require some limited reconstruction in order to make a tag. A secondary vertex associated with the decay of the opposite-side B hadron is sought from tracks with a transverse impact parameter significance $\epsilon_T/\sigma_{\epsilon_T} > 3$. The transverse distance (d_T) of this vertex from the primary vertex is required to have a significance $d_T/\sigma_{d_T} > 4$. The vertex momentum $\vec{p}_{SV}$ is defined as the vector sum of the momenta of all tracks included in the vertex. If the angle between this momentum and the momentum of the B candidate satisfies $\cos\phi_{SV}^B < 0.8$, then the vertex will be considered in the tagger. A secondary vertex charge (Q_{SV}) is then defined from all daughters of the vertex:

$$Q_{SV} = \frac{\sum_i (q^i p_L^{i\prime})^k}{\sum_i p_L^{i\prime}}, \tag{4.4}$$

where $p_L^{i\prime}$ is the longitudinal momentum of the i^{th} particle with respect to $\vec{p}_{SV}$ [65]. The value $k = 0.6$ is taken from previous studies at LEP [66].

The final tagger is the event-charge tagger, which makes use of all tracks with

[4]That is, they have the highest dilution $\mathcal{D}$.

momentum $0.5 < p_T < 50$ GeV/c and satisfying $\cos\phi_i^B < 0.8$. The event charge is defined as

$$Q_{EV} = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}. \tag{4.5}$$

The tagging variables Q_J^μ, Q_J^e, Q_{SV}, and Q_{EV} are used as discriminating variables in a likelihood ratio to construct a combined tagger. The combined tagging variable r is defined as

$$r = \prod_{i=1}^{4} r_i,\ r_i = \frac{f_i^{\bar{b}}(Q_i)}{f_i^b(Q_i)}, \tag{4.6}$$

where $f_i^{b/\bar{b}}(Q_i)$ is the *pdf* for the i^{th} charge variable given a reconstruction side $b/\bar{b}$ quark. These *pdf*s are measured in a sample of $B^+ \to \mu^+\nu_\mu\bar{D}^0X$ decays. The B^+ meson does not mix and the OST is not correlated to the reconstruction-side B hadron, so this sample may be used to determine the OST *pdf*s without the flavour ambiguity associated with mixed decays. Distributions for each tagging variable (jet charge, SV charge and event charge) are shown in Fig. 4.6 and the distribution for the combined tagging variable is shown in Fig. 4.7.

Only the most relevant taggers are used for a given event to construct r. An event with an opposite-side muon will use the muon and secondary-vertex taggers, but not the electron or event-charge taggers (r_i for these taggers is set to 1). Similarly, an event with an electron, but no muon, will use the electron and secondary-vertex taggers. An event with neither a muon nor an electron will use the secondary-vertex tagger and the event-charge tagger, provided a secondary vertex exists. An event with no identified secondary vertex will not be tagged.

An event with $r > 1$ is likely to have an initial $\bar{b}$ quark on the reconstruction side and an event with $r < 1$ is likely to have an initial b quark on the reconstruction

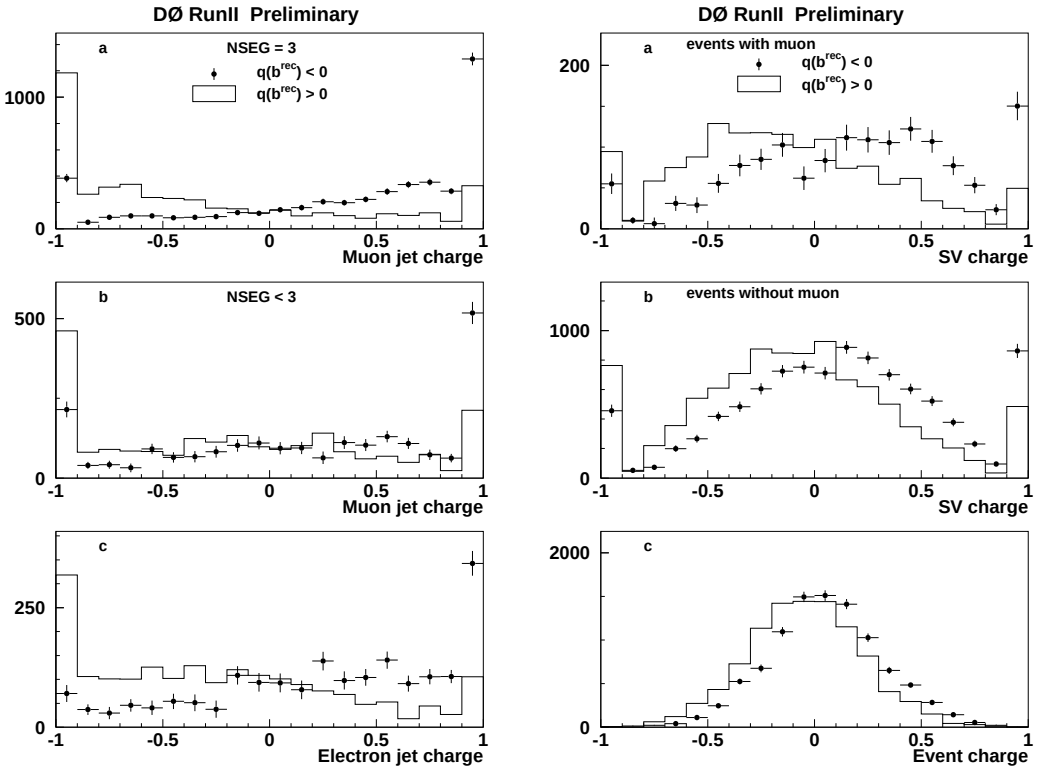

Figure 4.6: Distributions for the tagging variables Q_J^μ (plots a) and b) on the left), Q_J^e (plot c) on the left), Q_{SV} (plots a) and b) on the right) and Q_{EV} (plot c) on the right). In the Q_J^μ plots, NSEG= 3 designates a muon with a hit in the A and either the B or C layers of the muon system. NSEG$<$ 3 indicates a lower quality muon [67]. In all cases, a charge greater than zero indicates a b quark on the reconstruction side and a charge less than zero indicates a $\bar{b}$ quark on the reconstruction side.

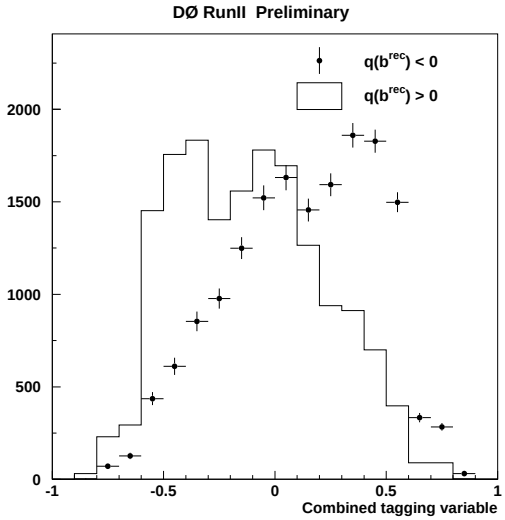

Figure 4.7: Distribution of the combined tagging variable d_{pr}.

side. It is convenient to define a linear combined tagging variable, which we identify as the predicted dilution d_{pr}:

$$d_{pr} = \frac{1-r}{1+r}. \tag{4.7}$$

By definition, the variable d_{pr} is defined in the interval $[-1, 1]$. An event with $d_{pr} > 0$ is tagged as an initial b quark and an event with $d_{pr} < 0$ is tagged as an initial $\bar{b}$ quark. A larger magnitude $|d_{pr}|$ corresponds to a higher tagging purity (or confidence).

The tagging algorithm returns a value for the predicted dilution (d_{pr}). This must be converted into a calibrated dilution value that represents our actual confidence in the flavour tagging. For example, the calibrated dilution accounts for the possibility that there was a flavour oscillation in both B hadrons ($\mathcal{D}(d_{pr})$ does not reach 1). To determine the calibration of the measured dilution ($\mathcal{D}$) as a function

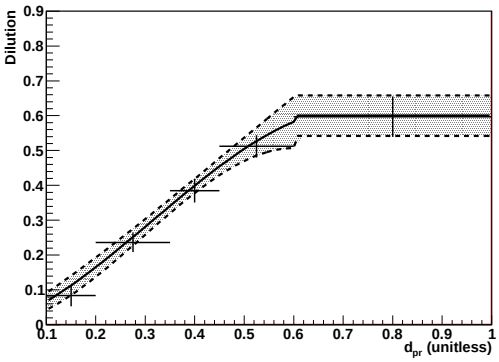

Figure 4.8: Dilution calibration function from $B_d^0 \to \mu^+ \nu_\mu D^{*-} X$ data. The highlighted area represents the $\pm 1\sigma$ error bound.

of the predicted dilution (d_{pr}), we use $B_d^0 \to \mu^+ \nu_\mu D^{*-} X$ decays. Events are divided into bins in d_{pr} and the measured dilution and mixing frequency (Δm_d) are fitted in a manner similar to the amplitude fitting method described in Sec. 7.1. The binned $\mathcal{D}$ is then fitted to a cubic function (see Fig. 4.8):

$$\mathcal{D}(d_{pr}) = \begin{cases} 0.457|d_{pr}| + 2.349|d_{pr}|^2 - 2.498|d_{pr}|^3 & d_{pr} < 0.6 \\ 0.6 & d_{pr} \geqslant 0.6 \end{cases} . \qquad (4.8)$$

Chapter 5

Input Parameters and Functions

A variety of parameters and their associated *pdf*s are required in order to properly account for detector effects. In addition, these parameters provide some discriminating power against the combinatoric background. Many of these parameters have been introduced in Chap. 2, and are explained in the following sections. Their application to the analysis are detailed in Chap. 6.

5.1 Measured *pdf*s

The behaviour of the 'nuisance' parameters, d_{pr}, σ_{x^M} and y, in the signal (or physics background) differs from that in the combinatoric background. By including separate signal and background *pdf*s for these variables we can improve the separation of the signal and background. Distributions for each of the parameters are measured in the $M(KK\pi)$ signal band ($1.91 < M(KK\pi) < 2.01$ GeV/c^2) and separately in the side bands ($1.79 < M(KK\pi) < 1.84$ GeV/c^2 and

$2.08 < M(KK\pi) < 2.13$ GeV/c^2). The distributions from the two side bands are summed to give the *pdf*s for the combinatoric background. The summed side band distributions are subtracted from the signal band distributions for each parameter. These background-subtracted distributions are the signal *pdf*s for each parameter. To ensure good statistics, the distributions from the untagged sample are used for the tagged data[1]. The signal *pdf*s ($P^{d_{pr}}_{Signal}$ and $P^{\sigma_{x^M}}_{Signal}$) are used to model the physics background as well. The background *pdf*s ($P^{d_{pr}}_{Bkgnd}$ and $P^{\sigma_{x^M}}_{Bkgnd}$) are used to model the combinatoric background.

Because the selection variable is dependent on the kinematic variables and the vertex quality, its distribution will be different in long-lived and short-lived (i.e., prompt) decays. The distribution for the prompt decays is generated in the manner described above, but using only events with negative decay length ($x^M < 0.0$) cm [2]. Similarly, the distribution for the long-lived decays is generated using only events with $x^M > 0.03$ cm. The intermediate lifetime region is composed of a mixture of long- and short-lived particles and is not used for these distributions. Four distributions are generated, each modelling the selection variable (y) in some part of the data: the long-lived contributions to the signal ($P^{y,long}_{Signal}$), the prompt contribution to the signal ($P^{y,prompt}_{Signal}$), the long-lived contribution to the background ($P^{y,long}_{Bkgnd}$) and the prompt contribution to the background ($P^{y,prompt}_{Bkgnd}$). The distributions for the nuisance parameters for Run IIa and Run IIb are shown in Fig. 5.1 and 5.2, respectively.

[1]The opposite-side tag is independent of the reconstruction-side parameters, so there is no expected difference in the tagged/untagged distributions.

[2]Due to the detector resolution, it is possible to have a negative lifetime, where the B^0_s decay vertex is 'behind' the primary vertex. In this case the direction from the PV to the B^0_s decay vertex is opposite the direction of the B^0_s momentum.

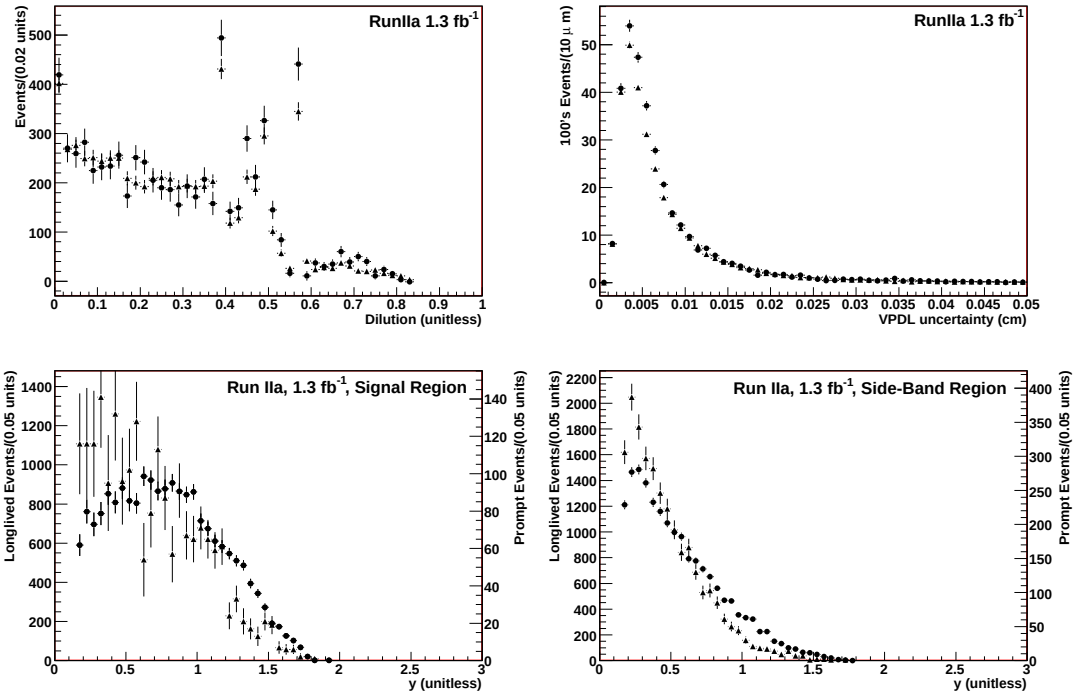

Figure 5.1: Distributions for the nuisance parameters, d_{pr} (top left), σ_{x^M} (top right) and y (bottom) in Run IIa data. The d_{pr} and σ_{x^M} plots show the distributions for the signal (circles) and background (triangles). The parameter y is determined separately in long-lived events (circles) and prompt events (triangles); the scale for the prompt events is on the right vertical axis.

5.2 Resolution Scale Factor

The uncertainty on the decay length is estimated by the vertex fitting procedure. A resolution scale factor (SF) is introduced to take into account a possible bias. The scale factor multiplies the uncertainty returned by the vertex fit to better represent the detector lifetime resolution. The SF is determined on an event-by-event basis, based on the distribution of hits in the silicon vertex detectors. The details of this procedure are given in Ref. [68]. A global scale factor to be applied

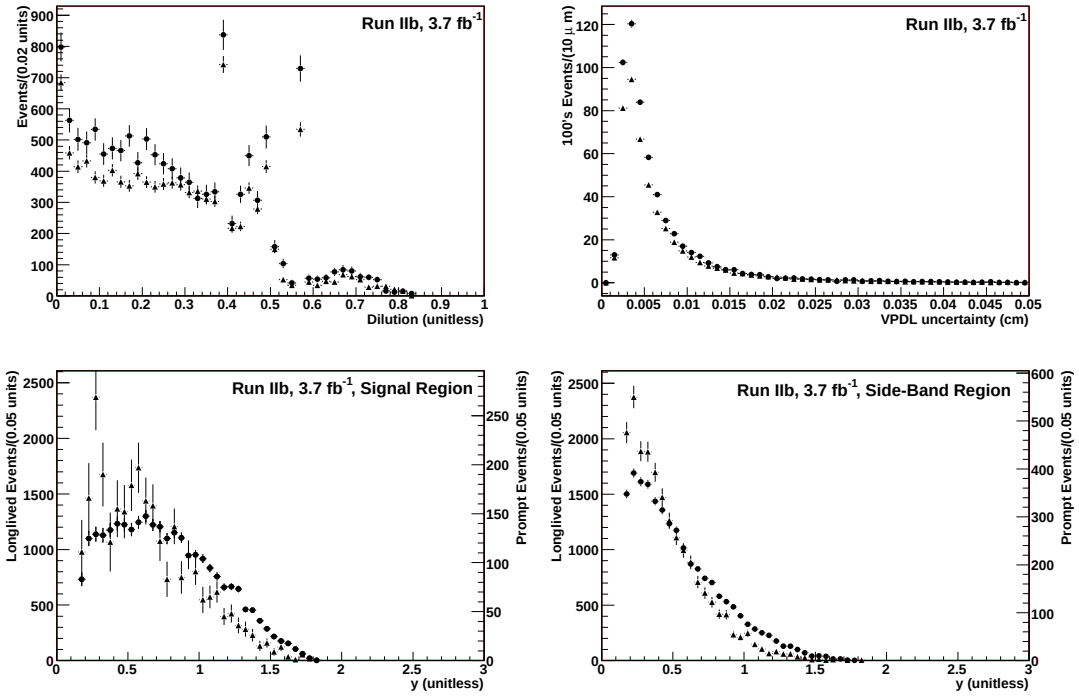

Figure 5.2: Distributions for the nuisance parameters, d_{pr} (top left), σ_{x^M} (top right) and y (bottom) in Run IIb data. The d_{pr} and σ_{x^M} plots show the distributions for the signal (circles) and background (triangles). The parameter y is determined separately in long-lived events (circles) and prompt events (triangles); the scale for the prompt events is on the right vertical axis.

equally to all events is also considered.

The global scale factor tunes the event-by-event scale factor to the data. The event by event scale factor can be verified with $J/\psi \to \mu^+\mu^-$ decays. The J/ψ meson decays almost instantly to a $\mu^+\mu^-$ pair, which can be accurately reconstructed. The vertex of the J/ψ will coincide with the primary vertex and any deviation from this will provide a measure of the vertex resolution. We plot the pull distribution of the J/ψ vertex with respect to the primary vertex[3] and fit the negative side to

[3]The pull is the ratio of the decay length its uncertainty σ_{x^M}.

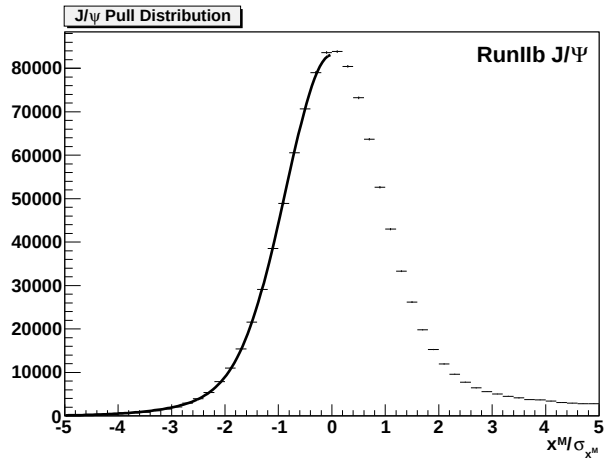

Figure 5.3: J/ψ Pull distribution for the Run IIb data with the fit to the negative side shown. A similar distribution is used for the Run IIa data.

a double Gaussian function. The widths of the Gaussians give the required scaling of the decay length uncertainty to correctly describe the detector resolution (see Fig. 5.3). The positive side of the pull distribution is ignored as it is biased towards larger values due to J/ψ mesons from real B meson decays. The Run IIa/IIb scale factors are $s_1 = 1.79 \pm 0.11/1.657 \pm 0.023$ with fraction $0.772 \pm 0.043/0.804 \pm 0.008$ and $s_2 = 0.957 \pm 0.022/0.845 \pm 0.004$ for the remainder.

5.3 Sample Composition

There are a number of D_s^- decays that resemble the $D_s^- \to K^{*0}K^-$ decay that we are interested in. Furthermore, there are decays of B^0 and B^- mesons that include a D_s^- meson; these events will end up in the D_s^- signal as B_s^0 candidates. These sources have different mixing and asymmetry behaviour and must be modelled

separately. For example, some of these sources have the muon originating from an intermediate D meson, or the D_s meson decaying from an excited D_s^{*-}. These variations of the direct decay lead to different tagging behaviour and/or higher missing energy. These effects must be modelled. For the sample composition of the μD_s signal, the following sources are considered:

- $B_s^0 \to D_s^- \mu^+ \nu_\mu X$
- $B_s^0 \to D_s^{*-} \mu^+ \nu_\mu X;\ D_s^{*-} \to D_s^- X$
- $B_s^0 \to D_{s0}^{*-} \mu^+ \nu_\mu X;\ D_{s0}^{*-} \to D_s^- X$
- $B_s^0 \to D_{s1}^{*-} \mu^+ \nu_\mu X;\ D_{s1}^{*-} \to D_s^- X$
- $B_s^0 \to D_s^- \tau^+ \nu_\tau X;\ \tau^+ \to \mu^+ \bar{\nu}_\tau \nu_\mu$
- $B_s^0 \to D_s^- D_s^+ X;\ D_s^\pm \to \mu^\pm X$
- $B_s^0 \to D_s^- D X;\ D \to \mu^+ X$
- $B_s^0 \to D_s^+ D X;\ D \to \mu^- X$
- $B^+ \to D_s^- D X;\ D \to \mu^+ X$
- $B^0 \to D_s^- D X;\ D \to \mu^+ X$

The latest PDG values [9] are used to determine the branching fractions of the decays listed above. When PDG values are not available, HFAG [39] or EvtGen [62] values are used[4]. The exclusive branching fractions for semileptonic B_s^0 decays to

[4]HFAG values are used when new results are not yet published in the PDG (for example, $\mathrm{Br}(B_s^0 \to D_s^- D_s^+ X)$). EvtGen values are used when no measurement exists, these are theoretically expected values.

D_s^{*-}, D_{s0}^{*-} and D_{s1}^{*-} are not measured; instead they are calculated from those for similar decays of B^0 mesons assuming the spectator model. Branching fractions that contribute to the processes listed above are given in Appendix B. These fractions are multiplied to determine the relative fractions for each source.

μD_s Source	Run IIa (%)	Run IIb (%)
$B_s^0 \to D_s^- \mu^+ \nu_\mu X$	23.03 ± 7.07	23.96 ± 7.27
$B_s^0 \to D_s^{*-} \mu^+ \nu_\mu X$	60.35 ± 18.33	59.32 ± 18.02
$B_s^0 \to D_{s0}^{*-} \mu^+ \nu_\mu X$	1.79 ± 0.54	1.47 ± 0.44
$B_s^0 \to D_{s1}^{*-} \mu^+ \nu_\mu X$	3.01 ± 0.91	3.26 ± 0.99
$B_s^0 \to D_s^- \tau^+ \nu_\tau X$	0.26 ± 0.08	0.23 ± 0.07
$B_s^0 \to D_s^- D_s^+ X$	2.53 ± 0.91	2.44 ± 0.88
$B_s^0 \to D_s^- D X$	0.27 ± 0.13	0.31 ± 0.15
$B_s^0 \to D_s^+ D X$	0.28 ± 0.14	0.28 ± 0.14
$B^+ \to D_s^- D X$	2.84 ± 0.52	3.18 ± 0.59
$B^0 \to D_s^- D X$	5.32 ± 1.09	5.59 ± 1.13

Table 5.1: Average sample composition of the μD_s signal in the entire mass interval for Run IIa and Run IIb.

The relative contribution of each source varies with the reconstructed mass of the B_s^0 candidate (i.e., the μD_s mass). For example, an event with a high μD_s mass is more likely to originate from a direct $B_s^0 \to \mu^+ \nu_\mu D_s^-$ decay, rather than involving an intermediate D_s^{*-} state that will include additional sources of missing energy. A more accurate model can be obtained by binning the sample composition by μD_s mass. MC for each source is used to predict the relative event yield in bins of μD_s mass. This is used to scale the total fraction for each source, requiring the total fraction of all sources be 100% in each mass bin. The fraction for source i in mass bin j ($\mathcal{F}_j^i$) is given by

$$\mathcal{F}_j^i = \frac{f_j^i \cdot F^i}{\sum_i (f_j^i \cdot F^i)}, \tag{5.1}$$

where f_j^i is the fraction of MC events from source i that fall into mass bin j, and F^i is the overall (unbinned) contribution of source i (from Table. 5.1). f_j^i is determined with the selection cuts simulated in MC. In addition, f_j^i includes a model of the muon trigger efficiency dependence on muon p_T (see Sec. 5.4). The binned sample composition is given in Table 5.2 for Run IIa and in Table 5.3 for Run IIb.

To determine the uncertainties, we add in quadrature the uncertainties on branching fractions that go into each mode. The MC statistical uncertainty was not taken into account. The systematic uncertainty associated with the sample composition is obtained by changing individual branching fractions one at a time and recomputing the entire sample composition table.

5.4 K Factor

The B_s^0 decay includes an undetected neutrino and possibly other undetected or unreconstructed particles. To account for this missing energy, the momentum of the B_s^0 meson (and hence its lifetime) is corrected by a scaling factor,

$$K \equiv \frac{p_T^{\mu D_s}}{p_T^{B_s}}. \tag{5.2}$$

This definition requires knowledge of the true B_s^0 momentum and so we use simulated data to acquire a *pdf* of the K factor. The K factor *pdf* is determined independently for each source listed in Sec. 5.3 as well as for the D^- and Λ_c^+ reflections. Insufficient MC was available to model the K factor in the $D^- \to K^- K^{*0}$

Sample Composition for the 10 mass($D_s^- \mu^+$) bins for Run IIa										
	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6	bin 7	bin 8	bin 9	bin 10
$B_s^0 \to D_s^- \mu X$	12.1%	16.4%	19.0%	21.0%	22.9%	24.9%	28.3%	32.2%	39.3%	58.0%
$B_s^0 \to D_s^{*-} \mu X$	46.1%	57.6%	61.1%	62.7%	63.8%	64.6%	64.3%	63.5%	58.8%	41.8%
$B_s^0 \to D_{s0}^{*-} \mu X$	1.4%	1.9%	2.1%	2.1%	2.1%	1.9%	1.6%	1.1%	0.6%	0.1%
$B_s^0 \to D_{s1}^- \mu X$	2.3%	2.5%	3.0%	3.5%	3.9%	3.9%	3.1%	1.8%	0.6%	0.04%
$B_s^0 \to D_s^- \tau X$	1.0%	0.6%	0.4%	0.2%	0.1%	0.07%	0.03%	0.0%	0.0%	0.0%
$B_s^0 \to D_s^- D_s X$	3.3%	3.0%	2.6%	2.1%	1.4%	1.0%	0.7%	0.6%	0.4%	0.0%
$B_s^0 \to D_s^- D X$	1.9%	0.6%	0.2%	0.1%	0.06%	0.03%	0.01%	0.0%	0.0%	0.0%
$B_s^0 \to D D_s^+ X$	2.1%	0.6%	0.1%	0.04%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
$B^+ \to D_s^- D X$	10.3%	6.1%	4.3%	2.9%	1.8%	0.9%	0.3%	0.0%	0.0%	0.0%
$B^0 \to D_s^- D X$	19.2%	10.4%	6.9%	5.1%	3.7%	2.5%	1.5%	0.5%	0.1%	0.0%

Table 5.2: Contributions of each decay chain to the total Run IIa B_s^0 signal, for each mass($D_s^- \mu^+$) bin.

Sample Composition for the 10 mass($D_s^- \mu^+$) bins for Run IIb										
	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6	bin 7	bin 8	bin 9	bin 10
$B_s^0 \to D_s^- \mu X$	12.2%	16.8%	19.5%	21.7%	23.6%	25.7%	29.1%	33.2%	40.4%	59.1%
$B_s^0 \to D_s^{*-} \mu X$	44.6%	56.4%	60.0%	61.7%	62.9%	63.6%	63.3%	62.6%	57.8%	40.8%
$B_s^0 \to D_{s0}^{*-} \mu X$	1.2%	1.5%	1.7%	1.8%	1.7%	1.6%	1.3%	0.9%	0.4%	0.1%
$B_s^0 \to D_{s1}^- \mu X$	2.5%	2.7%	3.2%	3.8%	4.2%	4.2%	3.5%	2.0%	0.7%	0.04%
$B_s^0 \to D_s^- \tau X$	0.8%	0.5%	0.3%	0.2%	0.1%	0.06%	0.02%	0.0%	0.0%	0.0%
$B_s^0 \to D_s^- D_s X$	3.1%	2.9%	2.5%	2.0%	1.4%	0.9%	0.7%	0.6%	0.4%	0.0%
$B_s^0 \to D_s^- D X$	1.8%	0.6%	0.2%	0.1%	0.06%	0.03%	0.01%	0.0%	0.0%	0.0%
$B_s^0 \to D D_s^+ X$	2.4%	0.7%	0.2%	0.05%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
$B^+ \to D_s^- D X$	11.4%	6.8%	4.8%	3.2%	2.0%	1.0%	0.3%	0.06%	0.0%	0.0%
$B^0 \to D_s^- D X$	19.6%	10.7%	7.1%	5.3%	3.8%	2.6%	1.5%	0.5%	0.2%	0.0%

Table 5.3: Contributions of each decay chain to the total Run IIb B_s^0 signal, for each mass($D_s^- \mu^+$) bin.

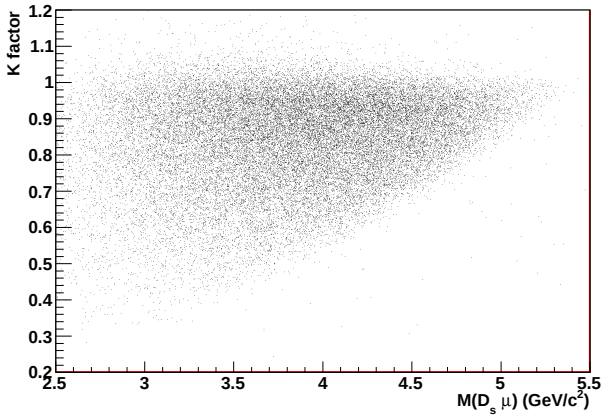

Figure 5.4: Scatter plot of the K factor plotted against the μD_s invariant mass. Plotted from $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ MC.

channel and the combinatoric background; these sources instead use a fixed value. The $D^- \to K^- K^{*0}$ source uses $K = 0.81$ determined from the average of the small MC sample available and the combinatoric background uses $K = 1.0$. The average lifetime of the combinatoric background is fitted to account for any missing momentum inconsistent with the fixed K factor assumption.

Events with an invariant μD_s mass close to the actual B_s^0 mass (5.37 GeV/c^2) are expected to have less missing energy than events with a much lower mass. As such, we expect that the K factor distribution will depend on the reconstructed μD_s mass. Events with a smaller mass are more likely to have a larger amount of missing energy and hence a wider K factor *pdf*. Conversely, events with a high mass are more likely to have a small amount of missing energy and will have a narrow K factor distribution. The K factor distribution is integrated when included in the lifetime *pdf*. The width of the distribution gives an effective uncertainty in the VPDL. A smaller width in the distribution will result in a smaller effective

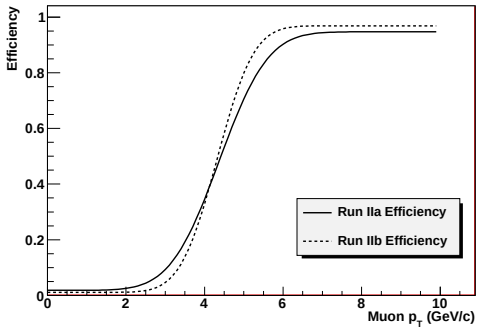

Figure 5.5: The solid and dashed lines show the efficiency of the muon triggers in Run IIa and Run IIb, respectively.

uncertainty, and hence a better measurement. The dependence of the K factor distribution on the μD_s mass is demonstrated in Fig. 5.4. To account for the mass dependence, the MC samples are divided into ten mass bins in the same manner as described in Sec. 5.3. The K factor distribution is determined in each bin for each source.

We model the muon trigger efficiency in the K factor as well. The trigger efficiency has a dependence on the muon p_T. The trigger efficiency is determined by comparing the momentum spectrum of muons in the single muon sample with the spectrum from a sample of events unbiased with respect to the trigger (that is, every event is triggered). The single muon sample includes triggered muons, while the minimum bias sample also includes muons that did not fire a trigger. The normalized ratio of their muon p_T spectra will give the trigger efficiency as a function of p_T. The normalization assumes that the trigger is 100% efficient for muons with $p_T > 6$ GeV/c. This procedure is described in detail in Ref. [69, 70]. The resultant efficiency curves are shown in Fig. 5.5. The trigger efficiency

difference between Run IIa and Run IIb is the only difference between the K factor *pdf*s for these two data sets. Examples of K factor distributions are shown in Figs. 5.6, 5.7 and 5.8. These distributions are convoluted with the lifetime *pdf* (see Eqn. 6.22) and represent an asymmetric uncertainty on the lifetime of the B_s^0 candidate (x^M). As expected, the events with high $D_s\mu$ mass have a narrower K factor distribution and hence a better determination of the B_s^0 lifetime.

5.5 Selection Efficiency

The selection criteria listed in Sec. 4.1 are dependent on the lifetime of the B_s^0 meson. The impact parameter significance criteria, for example, are less likely to be satisfied for short-lived B_s^0 candidates. This results in a lower efficiency for reconstructing short-lived candidates. This inefficiency must be accounted for in the modeling of the B_s^0 lifetime distribution.

The selection cuts are simulated in MC samples for each of the signal sources and in the physics background sources. The selection efficiency is determined in bins of the B_s^0 candidate lifetime, where the efficiency is defined as the ratio of the number of MC events passing the selection cuts to the number of MC events in that lifetime bin. In order to determine analytically the convolution integral (Eqn. 6.22), we require an analytic function for the efficiency. The function chosen for this has the following form:

$$Eff(x^M) = p_0 \cdot \left(1 - \left(p_2 + p_3 \cdot x^M + p_4(x^M)^2 + p_5(x^M)^3\right) \cdot \exp\left(-(x^M)^2/p_1\right)\right). \tag{5.3}$$

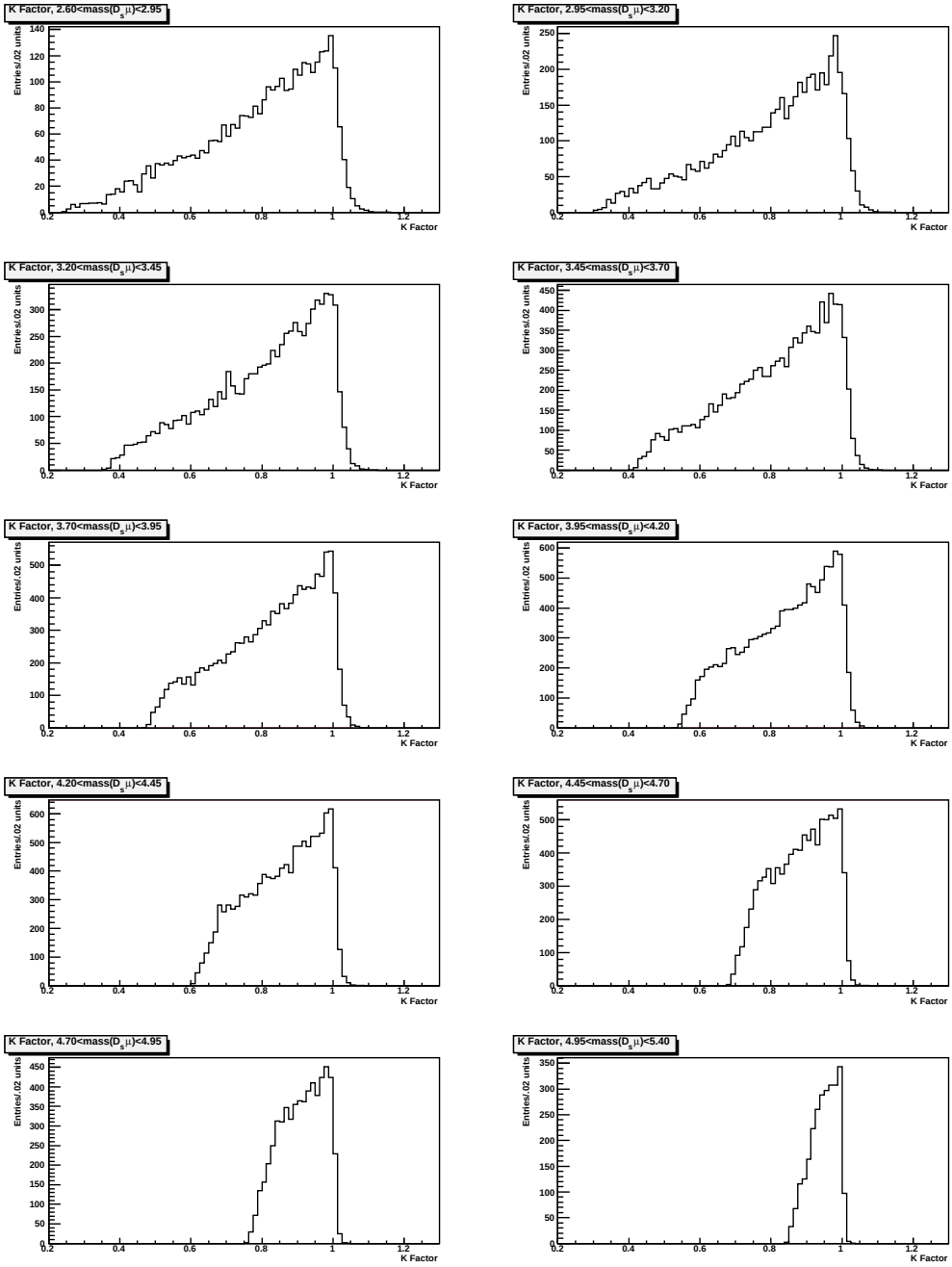

Figure 5.6: K factor distributions in ten mass bins for the $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ source.

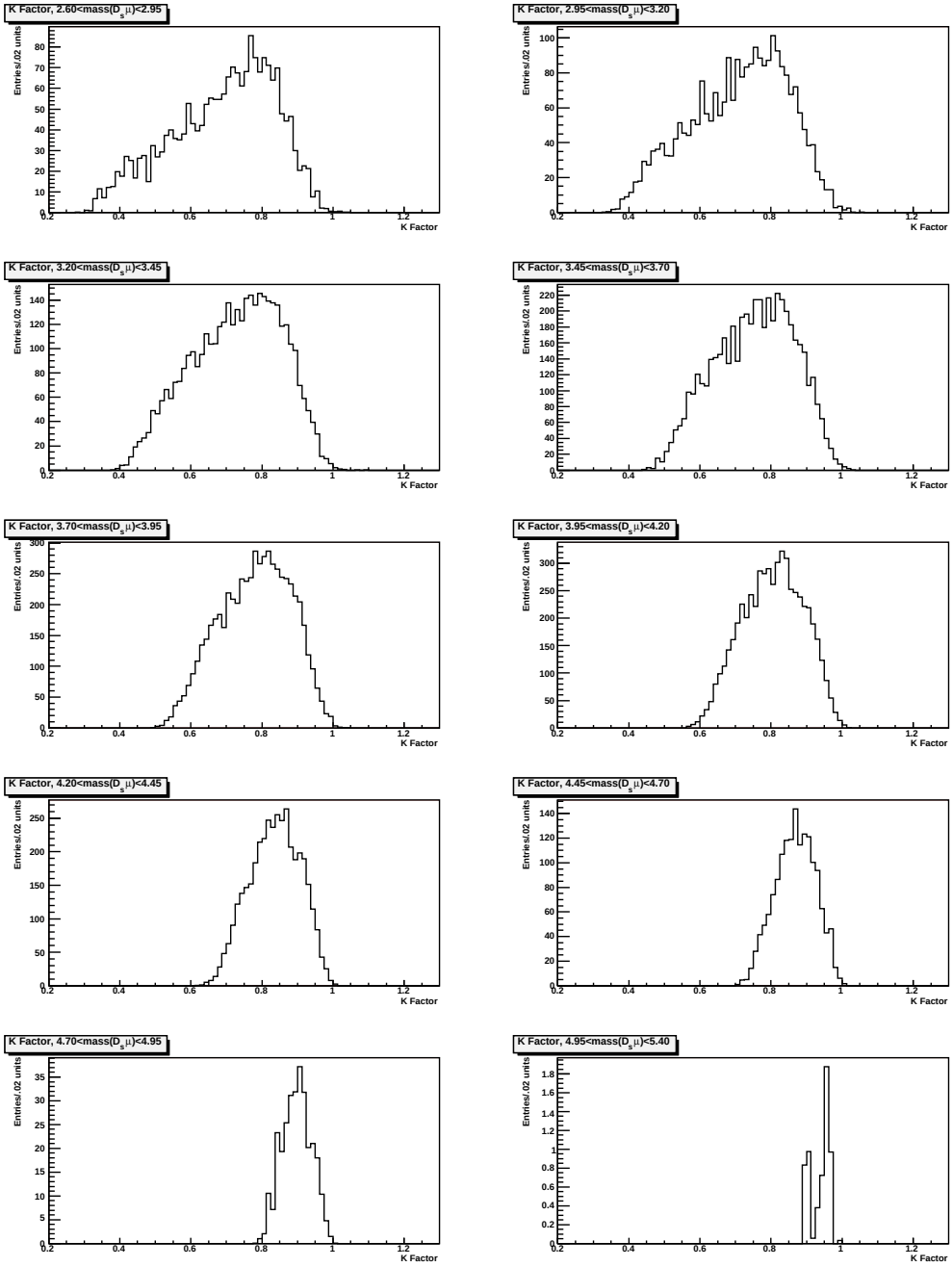

Figure 5.7: K factor distributions in ten mass bins for the $B_s^0 \to D_{s1}^{*-}\mu^+\nu_\mu X$ source.

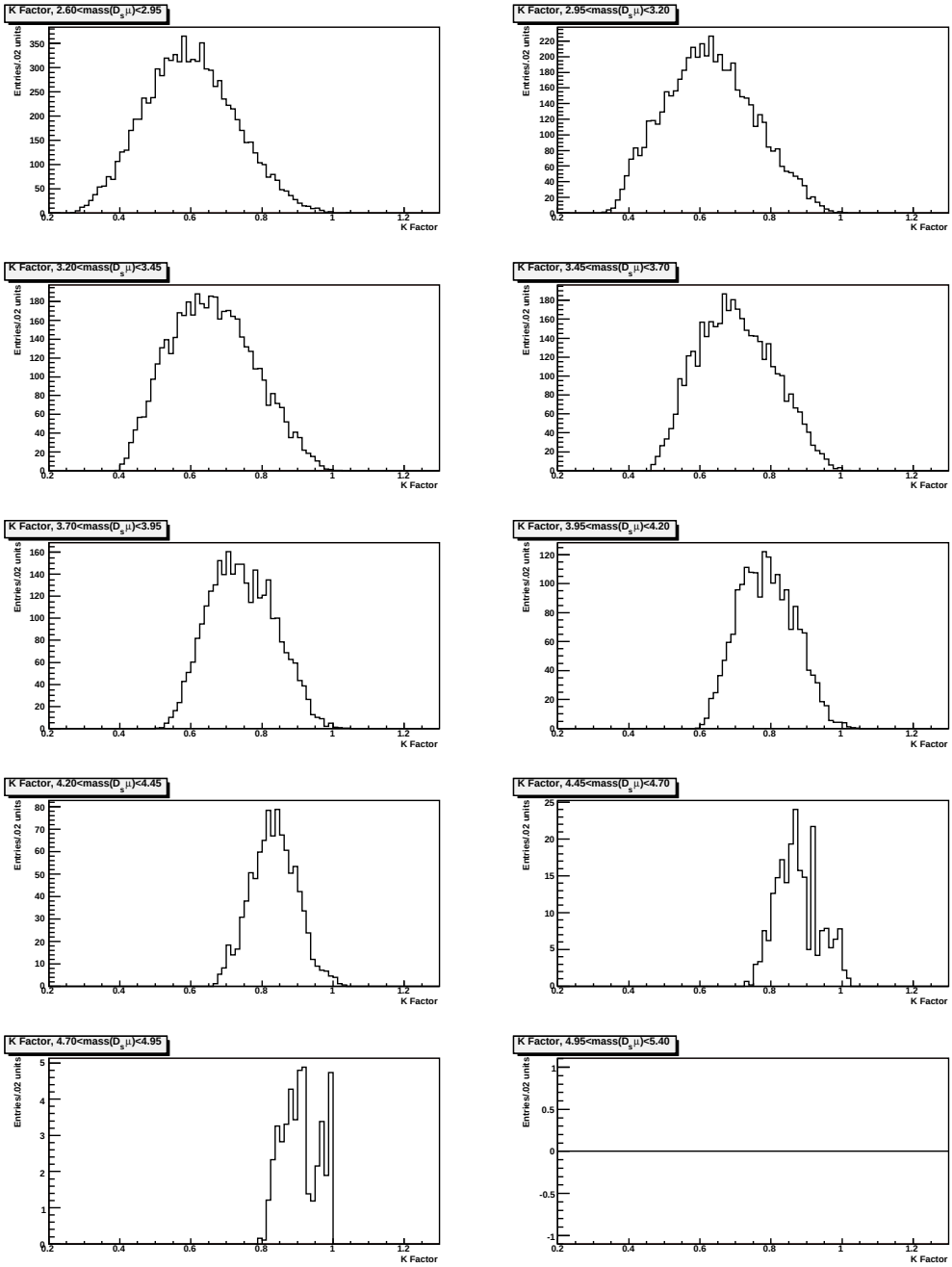

Figure 5.8: K factor distributions in ten mass bins for the $B^0 \to D_s^- DX$ source.

The binned efficiency is determined from MC simulations of each source and fitted to Eqn. 5.3. The results are shown in Figs. 5.9, 5.10 and 5.11. Parameter p_5 is characterized by a very large uncertainty because it is constrained primarily by the negative VPDL region with poor statistics. The same is true (to a lesser extent) for the parameter p_4. A lack of simulated inclusive $b\bar{b}$ MC prevents the efficiency from being determined in the combinatoric background. Instead, the efficiency is fitted directly to the data (see Sec. 6.3).

5.6 Background Asymmetries

In addition to the asymmetries associated with the signal (a^s_{fs}) and backgrounds ($a^d_{fs}$and a^{bkg}), there are asymmetries associated with the detector that must be described. The detector itself is not symmetric, due primarily to the toroid field in the muon spectrometer. These effects are constant in time and uniform over all sources, making them relatively easy to isolate in a time-dependent analysis.

Being made of matter, the detector response to particles and antiparticles is also slightly asymmetric. K^- mesons interact with nucleons to produce hyperons[5], however, these reactions do not occur for the K^+ meson. As a result, the K^- has a shorter average track length and hence a lower detection efficiency than the K^+ meson. These sources of asymmetry are described in detail below.

[5]A hyperon is a baryon with one or more strange quarks but no charm or bottom, that is, the Λ, Σ, Ξ and the Ω^- baryons.

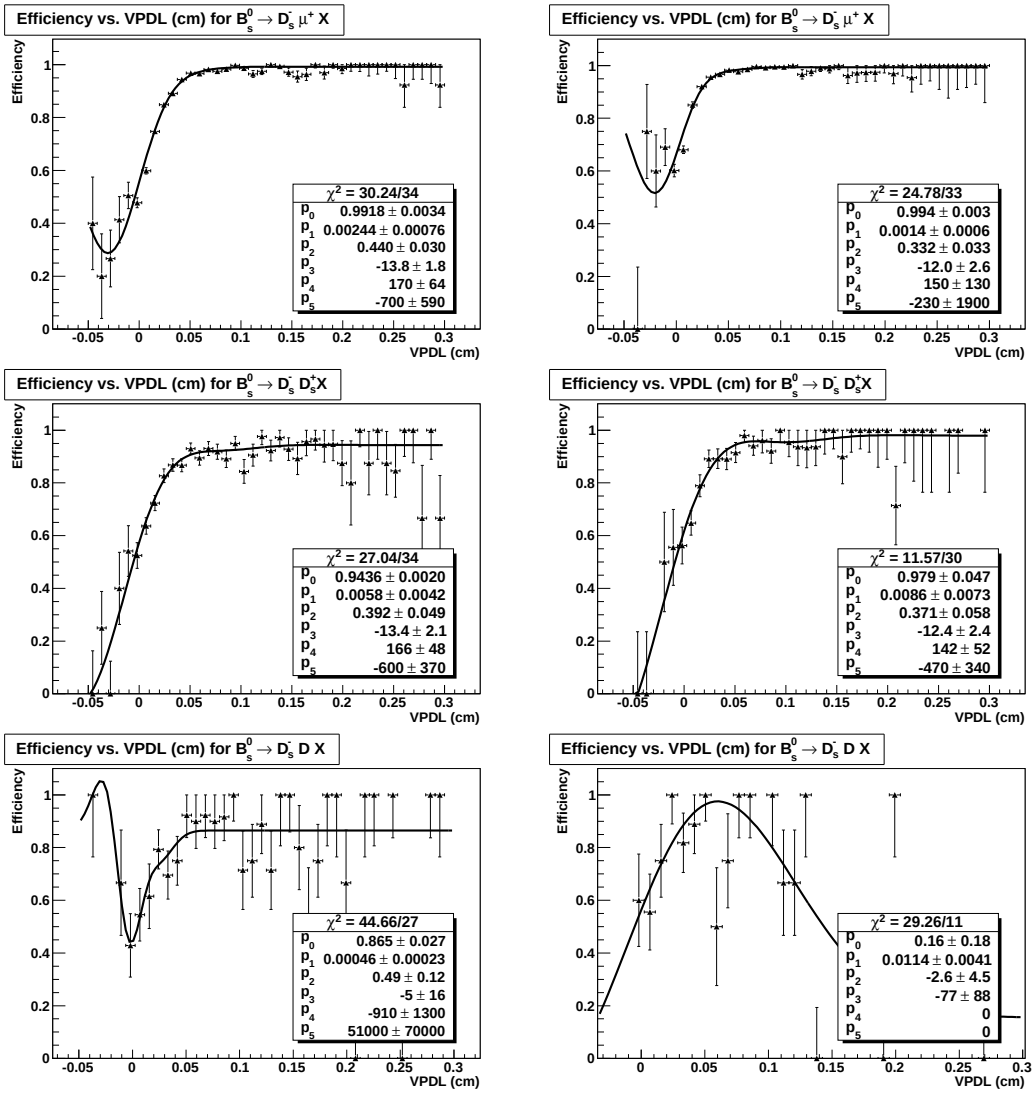

Figure 5.9: Selection efficiency in $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ decays, including D_s^{*-} modes (top), $B_s^0 \to D_s^- D_s^+ X$ (middle) and $B_s^0 \to D_s^- D X$ (bottom). Plots for Run IIa are on the left and Run IIb plots are on the right.

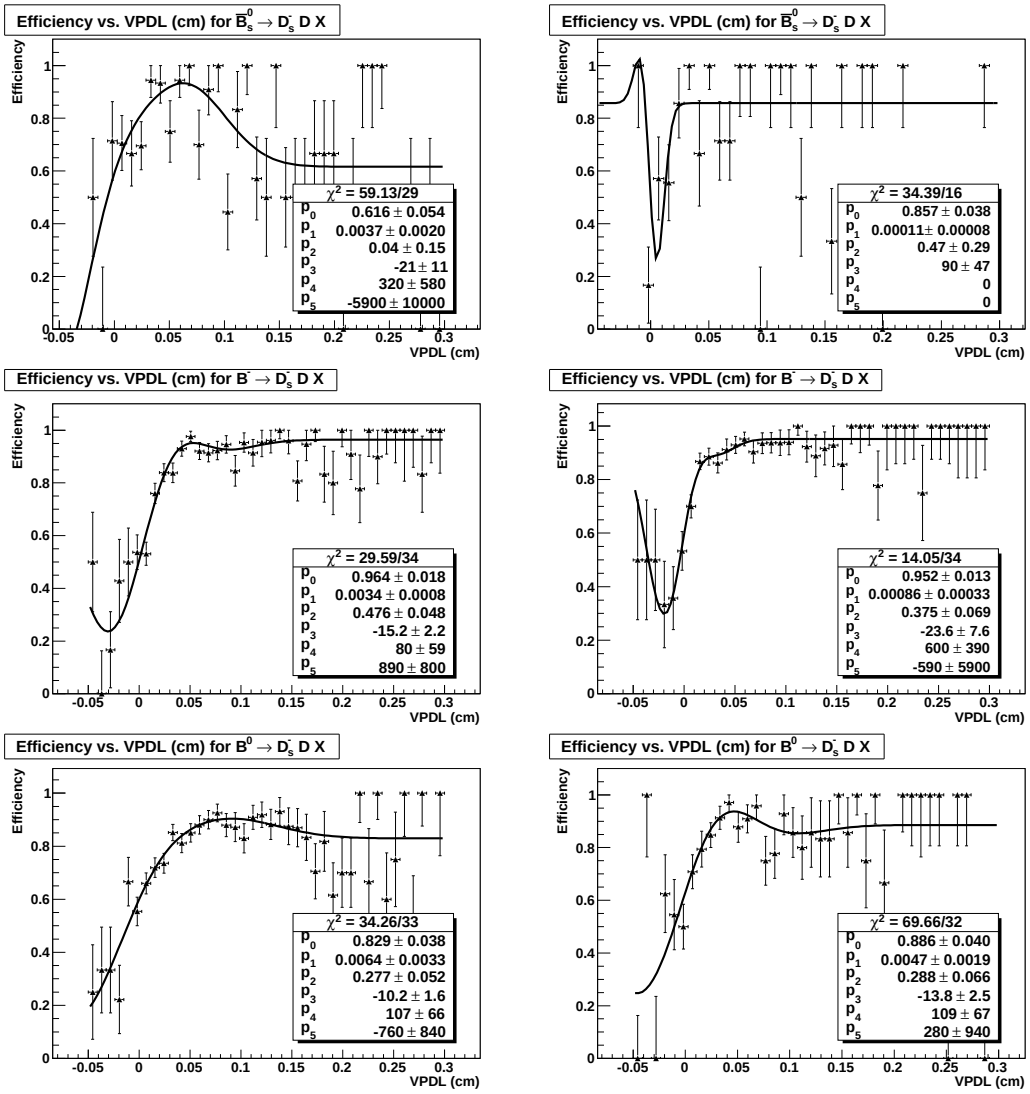

Figure 5.10: Selection efficiency in $B^0_s \to D^+_s DX$ (top), $B^+ \to D^-_s DX$ (middle) and $B^0 \to D^-_s DX$ (bottom). Plots for Run IIa are on the left and Run IIb plots are on the right.

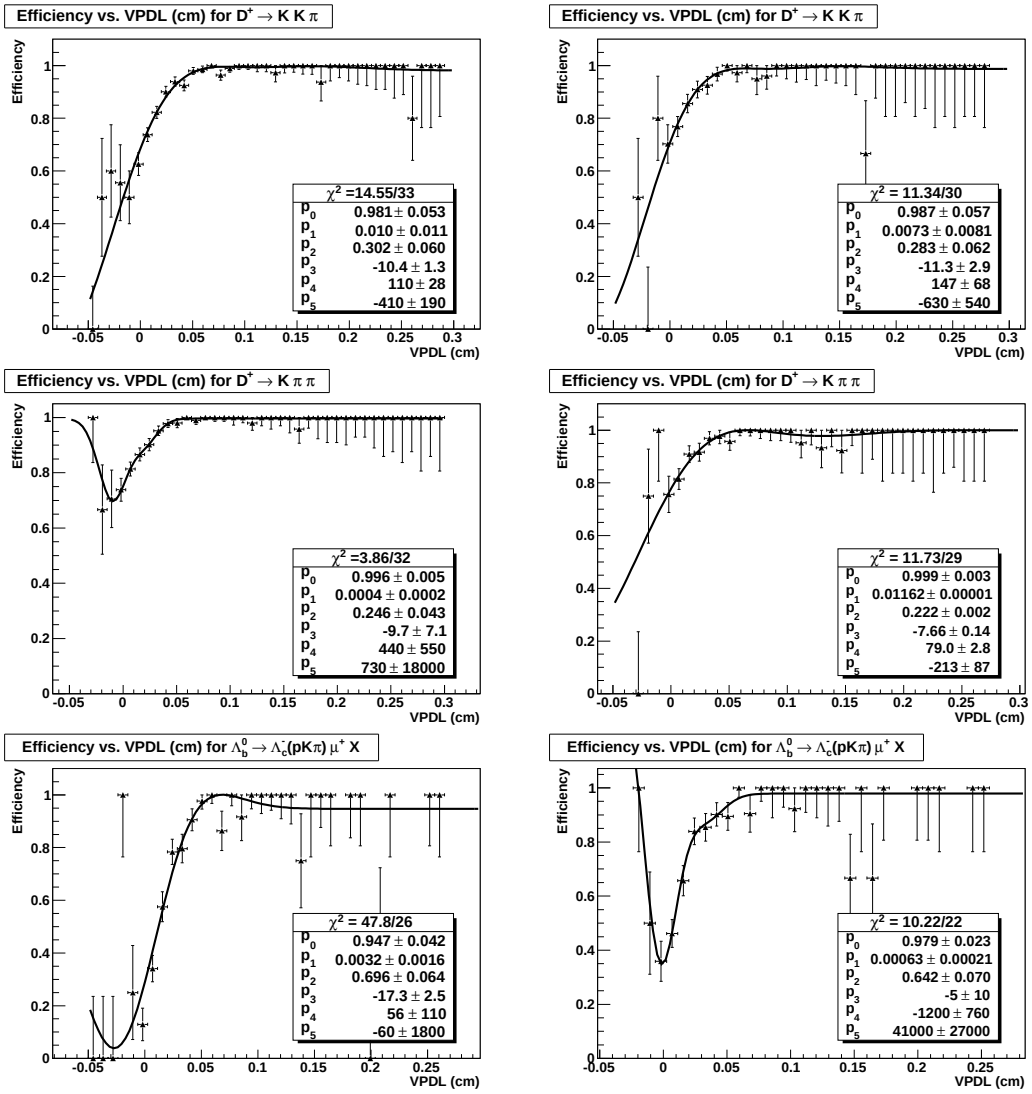

Figure 5.11: Selection efficiency in $D^- \to K^+\pi^-K^-$ (top), $D^- \to K^+\pi^-\pi^-$ (middle) and $\Lambda_c^- \to \bar{p}K^+\pi^-$ (bottom). Plots for Run IIa are on the left and Run IIb plots are on the right.

5.6.1 Detector Asymmetries

To determine the detector asymmetries we consider all independent variables that will not be conserved by a $\mathcal{CP}$ transformation. We model all possible asymmetries that are dependent on these variables. The procedure outlined here has been developed in a previous D0 analysis [27]. There are three relevant variables: the toroid polarity (β), the sign of the muon pseudorapidity (γ_μ) and the final-state flavour tag (q_μ). These variables are combined to give three single variable asymmetries, three double-variable asymmetries and a single triple-variable asymmetry. The asymmetries and their origins are listed below:

- A_{q_μ} : The charge asymmetry, represented by a^s_{fs}, a^d_{fs} and a^{bkg}. This is the measured asymmetry and is not considered as a detector asymmetry.

- A_β : The toroid asymmetry, measured directly as the fraction of data collected with $+/-$ toroid polarity.

- A_{γ_μ} : The geometric detector asymmetry due to potential differences in the north/south sides of the detector. This asymmetry is denoted A_{det}.

- $A_{q_\mu\gamma_\mu}$: The Forward-Backward asymmetry related to the proton-antiproton beam directions. Positive/Negative charge particles tend to go in the proton/antiproton direction. This asymmetry is denoted A_{fb}.

- $A_{q_\mu\beta}$: Accounts for possible changes to efficiency when the toroid polarity is reversed. This asymmetry is expected to be zero.

- $A_{\beta\gamma_\mu}$: Accounts for a possible forward-backward asymmetry related to the toroid polarity. This asymmetry is expected to be zero.

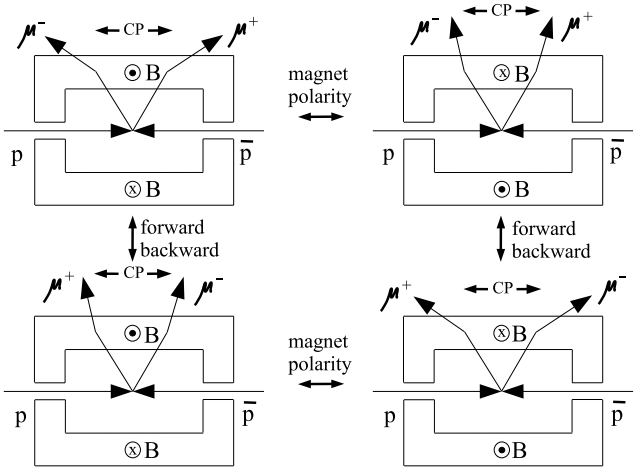

Figure 5.12: Schematic representation of how muons bend in the toroid field.

- $A_{q_\mu \beta \gamma_\mu}$: The Range-Out asymmetry due to different detection efficiency for muons in the toroid bending toward/away from the beam axis. Denoted A_{ro}, this asymmetry is discussed in greater detail below.

All of these detector asymmetries are expected to be small, or in some cases, zero. The one exception is the range-out asymmetry, which is expected to be large. This asymmetry is caused by muons passing through the corners of the muon system where the forward toroid meets the central toroid. The change in the configuration of the PDTs and scintillator counters in this region (see Fig. 3.2) gives a reduction in the detection efficiency. A muon that bends in the toroid towards this region will have a lower probability of being detected compared to a muon that bends away. On one side of the detector a muon will bend towards the beam axis, while on the other side a muon of the same charge will bend away. This is shown schematically in Fig. 5.12. The result is a bias for detecting positive or negative charge muons

on one side of the detector or the other. This causes an asymmetry in the muon acceptance that depends on the muon charge and toroid polarity. Positive muons will be favoured in (for example) the forward direction, while negative muon are favoured in the backward direction. This behaviour will be reversed when the toroid polarity changes.

To model the detector asymmetries, we determine the net detector asymmetry for each event, given by Eqn. 5.4.

$$A_{net} = (1 + \beta A_{\beta})(1 + \gamma_{\mu} A_{det})(1 + q_{\mu}\gamma_{\mu} A_{fb})\cdot$$
$$(1 + q_{\mu}\beta A_{q_{\mu}\beta})(1 + \beta\gamma_{\mu} A_{\beta\gamma_{\mu}})(1 + q_{\mu}\beta\gamma_{\mu} A_{ro}). \quad (5.4)$$

The net asymmetry scales the total likelihood from Eqn. 2.4. The values of each detector asymmetry are fitted to the data simultaneously with the signal asymmetries.

5.6.2 Kaon Asymmetry

The kaon asymmetry is dependent on the kaon momentum. Slower kaons are more likely to interact with a nucleon and so have a higher asymmetry. This dependence has been measured elsewhere [71, 72] and is shown in Fig. 5.13.

We must account for both kaons in the $(K^+\pi^-)K^-$ final-state, and account for their momentum. The *pdf* for the μD_s sources is scaled by the asymmetry for these two kaons. For the μD^- and $\mu\Lambda_c^-$ reflections there is only one kaon giving

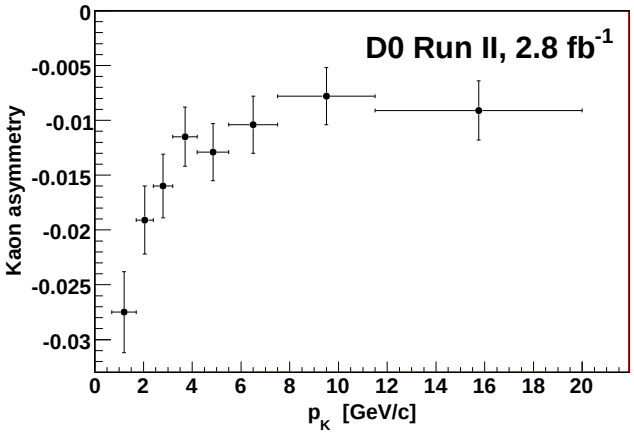

Figure 5.13: The kaon asymmetry as a function of kaon momentum.

an asymmetry. The relations for these corrections are given in Eqn. 5.5:

$$\begin{aligned} P_{\mu D_s^-} \rightarrow & \quad P_{\mu D_s^-}(1 - A^K(p_{K^-}))(1 + A^K(p_{K^+})), \\ P_{D^-(KK\pi)} \rightarrow & \quad P_{D^-(KK\pi)}(1 - A^K(p_{K^-}))(1 + A^K(p_{K^+})), \\ P_{D^-(K\pi\pi)} \rightarrow & \quad P_{D^-(K\pi\pi)}(1 + A^K(p_{K^+})), \\ P_{\Lambda_c^-(Kp\pi)} \rightarrow & \quad P_{\Lambda_c^-(Kp\pi)}(1 + A^K(p_{K^+})). \end{aligned} \tag{5.5}$$

Chapter 6

Mass and Lifetime Functions

Of the various input parameters and functions to the total likelihood function, the most important are the *pdf*s for the mass of the D_s^- candidate and the lifetime of the B_s^0 candidate. The D_s^- mass is the primary means of discriminating B_s^0 mesons from B^0 and other hadrons. The B_s^0 lifetime distribution is where we will observe the flavour oscillations and the time dependence of the charge asymmetry. The lifetime behaviour of background sources must be understood before we can hope to observe the B_s^0 signal oscillations (or the asymmetry in the oscillation). This chapter will describe these functions and the fitting procedure.

6.1 Fit to the D_s^- Mass

The primary means of distinguishing the various physics backgrounds from the B_s^0 signal is the invariant mass of the $KK\pi$ system. By fitting the mass distribution we obtain mass *pdf*s, which are used to weight events as either signal or

background. Additionally, the fit provides the number of signal, physics background and combinatoric background candidates, which are used to determine the fractional contribution from each source. These fractions are important for properly normalizing the *pdf* for each source. The physics backgrounds include 'reflection' sources where a particle is misidentified as a kaon. The difference in mass for the misidentified particle introduces a shift in the reconstructed D_s^- mass. For example, misidentifying a pion ($M_\pi = 139.6$ MeV/c^2) as a kaon ($M_K = 493.7$ MeV/c^2) shifts the reconstructed D_s^- mass to a higher value. Five $KK\pi$ sources are considered:

- $D_s^- \to K^{*0}K^-$ signal with fraction Fr_{sig}.
- $D^- \to K^+\pi^-\pi^-$ reflection, where a π^- is mistaken for a K^-. This includes primarily the resonant decay $D^- \to K^{*0}(K^+\pi^-)\pi^-$, but also includes some contribution from non-resonant $D^- \to K^+\pi^-\pi^-$. This source has a fraction Fr_{Dr}.
- $\Lambda_c^- \to \bar{p}K^+\pi^-$ reflection, where the $\bar{p}$ is mistaken for a K^-. This includes both the resonant $\Lambda_c^- \to \bar{p}K^{*0}(K^+\pi^-)$ decay and non-resonant contributions. This source has a fraction Fr_{Lr}.
- $D^- \to K^{*0}K^-$ with fraction Fr_{Dp}.
- Combinatoric background with fraction $(1 - Fr_{sig} - Fr_{Dr} - Fr_{Lr} - Fr_{Dp})$.

The mass *pdf*s for the signal and physics backgrounds are each described by a Gaussian function whose mean and width will be determined by the fit. The

reflection sources (D^- and Λ_c^+) must be treated differently to account for the shift in mass. This procedure is described below.

6.1.1 Mass Reflection

The energy of a particle is determined by its momentum and the mass it is assigned ($E = \sqrt{P^2 + M_{ID}^2}$). We have no means of differentiating kaons and pions and so we must simply assign the track a kaon mass and take into account the contribution from pions (reflections). The reflection sources are modelled by Gaussian functions, whose means are a function of the momenta of the three final-state particles. The mean is shifted from the true mass due to a higher or lower mass assignment to the misidentified particle. We can predict the extent of the $KK\pi$ mass shift from the final-state momenta and the difference between the true particle mass and the kaon mass. The D^- reflection has a pion assigned the kaon mass, giving a positive mass shift. The Λ_c^- reflection has a proton assigned the kaon mass, giving a negative mass shift. With these corrections, we can then fit to the true D^- and Λ_c^- mass.

The invariant mass of the $K\pi\pi$ or $Kp\pi$ system, assuming the $KK\pi$ hypothesis, is given by the following expression:

$$\begin{aligned} M_{KK\pi}^2 &= (E_K + E_\pi + E_{\pi/\bar{p}}^K)^2 - (P_K + P_\pi + P_{\pi/\bar{p}})^2, \\ M_{KK\pi}^2 &= M_{D^-/\Lambda_c^-}^2 + \left(1 + 2\frac{(E_K + E_\pi)(E_{\pi/\bar{p}}^K - E_{\pi/\bar{p}})}{M_K^2 - M_{\pi/\bar{p}}^2}\right)\left(M_K^2 - M_{\pi/\bar{p}}^2\right), \end{aligned} \quad (6.1)$$

where $E_{\pi/\bar{p}}^K$ is the energy of the misidentified π or $\bar{p}$ when assigned the kaon mass and $E_{\pi/\bar{p}}$ is the energy with the correct mass assignment. M_{D^-/Λ_c^-} is the reconstructed mass with the correct particle identification in the $K\pi\pi$ and $Kp\pi$ case.

This is the mass distribution that will be fitted as a Gaussian function. We then define a pair of reflection variables (R_{D^-} and $R_{\Lambda_c^-}$) to simplify the transformation between the two mass hypotheses. The reflection variables are defined in terms of the momentum of the $KK\pi$ candidates (via the energy):

$$R_{D^-/\Lambda_c^-} = \frac{(E_K + E_\pi)(E^K_{\pi/\bar{p}} - E_{\pi/\bar{p}})}{M_K^2 - M^2_{\pi/\bar{p}}}. \tag{6.2}$$

The corrected mass is then defined as

$$\begin{aligned} M^2_{D^-} &= M^2_{KK\pi} - (1 + 2R_{D^-})\,(M_K^2 - M_\pi^2)\,, \\ M^2_{\Lambda_c^-} &= M^2_{KK\pi} - \left(1 + 2R_{\Lambda_c^-}\right)\left(M_K^2 - M_{\bar{p}}^2\right). \end{aligned} \tag{6.3}$$

The corrected mass (M_{D^-/Λ_c^-}) has a Gaussian distribution, which is fitted for the central value and width. The mass range $1.79 < M(KK\pi) < 2.25$ GeV/c^2 is used for the fit, which is wide enough to include the $KK\pi$ Gaussian and most of the distorted *pdf* for the reflections.

6.1.2 Signal Fraction

The fraction of signal (or physics background) and combinatoric background is related to the energy scale of the decay (i.e., the total momentum). It is convenient to parametrize the fraction of each source using the D^- reflection variable (which we will call R), which is a function of the momentum of all three final-state particles. The fraction of the signal and physics backgrounds will then be determined

event-by-event. The fraction of source j for event i is given by

$$f_i^j = C(R) f_0^j N_i^j,\ C(R) = 1 + p_0 R + p_1 R^2. \tag{6.4}$$

The function $C(R)$ parametrizes the base fraction f_0^j. N_i^j adjusts the normalization to account for any portion of the mass *pdf* that falls outside of the mass interval:

$$N_i^j = \int_{1.79}^{2.25} dM P_{i,j}^m(m), \tag{6.5}$$

where $P_{i,j}^m(m)$ is the Gaussian for the source j for the event i (only the reflection variables depend on i). The parameters f_0^j, p_0 and p_1 are determined from the fit.

6.1.3 Combinatoric Background

The mass *pdf* for the combinatoric background is modelled by an exponential function with a decay constant that is parametrized by the D^- reflection variable R:

$$P_i^{bkg}(M) = \frac{1}{N_{bkg}} \exp\left(-\frac{M}{M_0(R)}\right), \tag{6.6}$$

where N_{bkg} normalizes the function in the mass interval. The parametrization is given by the function

$$M_0(R) = p_2(1.0 + p_3 R + p_4 R^2 + p_5 R^3 + p_6 R^4), \tag{6.7}$$

where the parameters $p_2 - p_6$ are determined from the fit. The background shape is distorted for small R due to the threshold of the kinematic cuts (in Sec. 4.1.1)

and cannot be modelled by an exponential. We remove this distortion by cutting on the reflection variable, $R > 0.22$. This is demonstrated in Fig. 6.1.

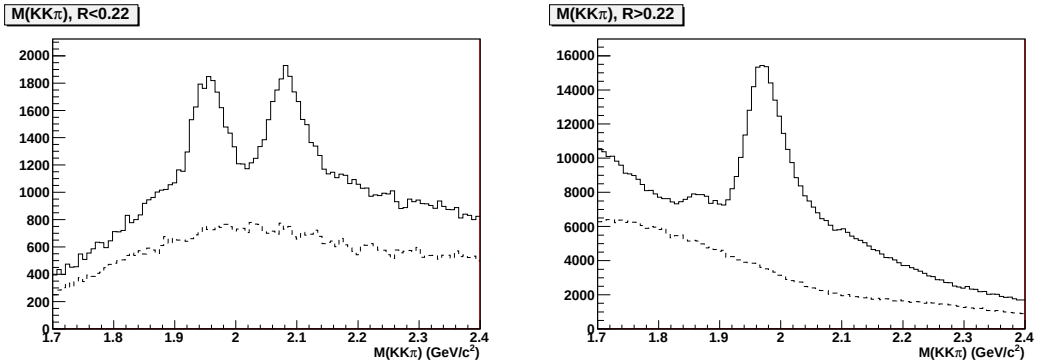

Figure 6.1: An example (Run IIb) $M(KK\pi)$ distribution before kinematic cuts for $R < 0.22$ (left) and $R > 0.22$ (right). The dashed line shows events with track charges inconsistent with a B_s^0 meson (i.e., $\mu^- K^+ K^- \pi^-$). The solid line shows events with the correct charge combination.

6.1.4 Fit and Results

The total log likelihood function to be minimized is given by

$$L = -2 \sum_i \ln \left(\sum_j f_i^j P_{i,j}^m(m_i) + \left(1 - \sum_j f_i^j \right) P_i^{bkg}(M_i) \right). \tag{6.8}$$

Fits are performed separately for the tagged and untagged data samples. The results are tabulated in Tables 6.1 - 6.4. The number of candidates in each channel is determined by averaging the fractions f_i^j over all events and multiplying by the sample size. These values are not used any further in the analysis, but are illustrative of the statistics involved (see Table 6.4). Plots of the $KK\pi$ mass distribution are shown in Figs. 6.2 and 6.3.

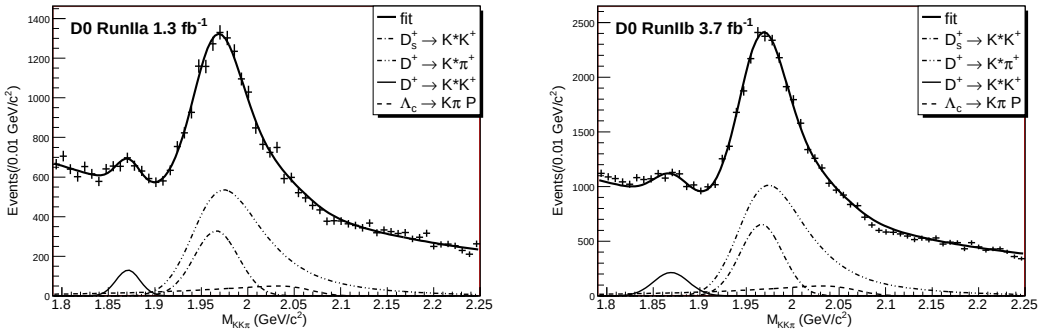

Figure 6.2: $M(KK\pi)$ distribution in the tagged data sample in Run IIa (left) and Run IIb (right).

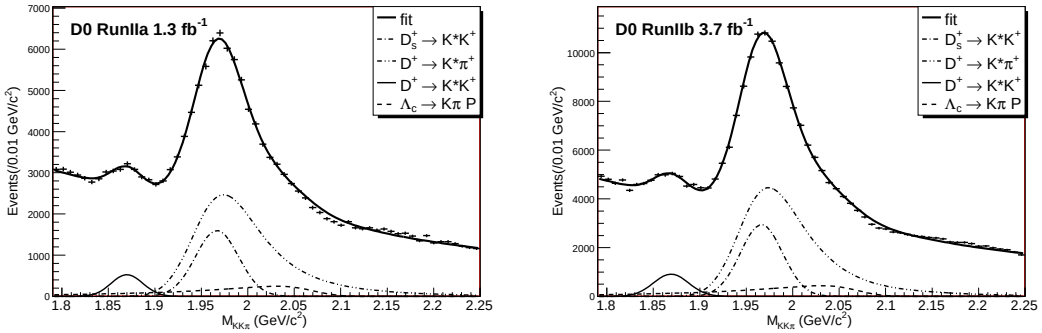

Figure 6.3: $M(KK\pi)$ distribution in the untagged data sample in Run IIa (left) and Run IIb (right).

6.2 B_s^0 Lifetime *pdf*

The B_s^0 lifetime (or decay length) *pdf* is the cornerstone of this work. All of the physics we are seeking to measure is described by this function and its fit. The joint probability density function $P^{x^M}(x^M, y; \sigma_{x^M}, d_{pr})$ for the signal and physics backgrounds is separated into two components to better model the prompt background that contaminates these sources. These are the same prompt decays described in Sec. 4.1.1. As previously discussed, these events have the D_s^- candidate matched

Channel	Mass (GeV/c^2)	Width (MeV/c^2)	f_0
$D_s^- \to K^{*0}K^-$	1.9661 ± 0.0014	23.7 ± 1.4	0.0394 ± 0.0046
$D^- \to K^+\pi^-\pi^-$	1.8625 ± 0.0005	25.2 ± 0.6	0.1170 ± 0.0124
$D^- \to K^{*0}K^-$	1.8711 ± 0.0021	12.9 ± 1.9	0.0085 ± 0.0016
$\Lambda_c^- \to \bar{p}K^+\pi^-$	2.2781 ± 0.0018	20.6 ± 1.9	0.0197 ± 0.0286
$D_s^- \to K^{*0}K^-$	1.9647 ± 0.0009	23.4 ± 0.9	0.0440 ± 0.0018
$D^- \to K^+\pi^-\pi^-$	1.8639 ± 0.0004	24.2 ± 0.4	0.1245 ± 0.0022
$D^- \to K^{*0}K^-$	1.8681 ± 0.0026	20.8 ± 3.1	0.0119 ± 0.0019
$\Lambda_c^- \to \bar{p}K^+\pi^-$	2.2818 ± 0.0016	18.8 ± 2.4	0.0205 ± 0.0015

Table 6.1: Mass parameters from the fit to the tagged data. The top four entries are for the Run IIa data and the bottom four are for the Run IIb data.

Channel	Mass (GeV/c^2)	Width (MeV/c^2)	f_0
$D_s^- \to K^{*0}K^-$	1.9669 ± 0.0006	23.2 ± 0.6	0.0379 ± 0.0019
$D^- \to K^+\pi^-\pi^-$	1.8628 ± 0.0003	25.3 ± 0.3	0.1069 ± 0.0051
$D^- \to K^{*0}K^-$	1.8701 ± 0.0013	16.9 ± 1.3	0.0091 ± 0.0008
$\Lambda_c^- \to \bar{p}K^+\pi^-$	2.2796 ± 0.0009	22.0 ± 1.0	0.0192 ± 0.0014
$D_s^- \to K^{*0}K^-$	1.9664 ± 0.0005	22.8 ± 0.5	0.0372 ± 0.0018
$D^- \to K^+\pi^-\pi^-$	1.8631 ± 0.0002	24.5 ± 0.2	0.1008 ± 0.0045
$D^- \to K^{*0}K^-$	1.8691 ± 0.0010	18.8 ± 1.1	0.0096 ± 0.0008
$\Lambda_c^- \to \bar{p}K^+\pi^-$	2.2799 ± 0.0007	21.8 ± 0.8	0.0181 ± 0.0011

Table 6.2: Mass parameters from the fit to the untagged data. The top four entries are for the Run IIa data and the bottom four are for the Run IIb data.

to a muon from some other light quark that decays nearby the primary vertex. The *pdf* is split into two components to separately model the prompt and long-lived decays:

$$P_j^{x^M}(x^M, y; \sigma_{x^M}, d_{pr}) = Fr_{Signal}^{prompt} P^{x^M,prompt}(x^M) P_{Signal}^{y,prompt} + (1 - Fr_{Signal}^{prompt}) P_j^{x^M,long}(x^M; \sigma_{x^M}, d_{pr}) P_{Signal}^{y,long}, \quad (6.9)$$

	Tagged		Untagged	
	Run IIa	Run IIb	Run IIa	Run IIb
p_0	2.098 ± 0.534	2.195 ± 0.244	2.261 ± 0.248	3.051 ± 0.272
p_1	-0.981 ± 0.296	-1.046 ± 0.147	-0.980 ± 0.136	-1.406 ± 0.148
p_2	0.641 ± 0.022	0.639 ± 0.105	0.784 ± 0.020	0.779 ± 0.018
p_3	-0.744 ± 0.019	-0.912 ± 0.154	-1.043 ± 0.028	-1.261 ± 0.028
p_4	0.391 ± 0.005	0.765 ± 0.026	0.702 ± 0.009	1.042 ± 0.012
p_5	-0.117 ± 0.004	-0.310 ± 0.010	-0.231 ± 0.007	-0.394 ± 0.059
p_6	0.016 ± 0.001	0.049 ± 0.004	0.030 ± 0.002	0.055 ± 0.002

Table 6.3: Parameters for the signal fraction and background shape.

Channel	Tagged		Untagged	
	Run IIa	Run IIb	Run IIa	Run IIb
$D_s^- \to K^{*0}K^-$	2537 ± 295	4867 ± 317	$12\,096 \pm 595$	$21\,461 \pm 1042$
$D^- \to K^+\pi^-\pi^-$	7456 ± 788	$13\,643 \pm 820$	$33\,835 \pm 1625$	$57\,722 \pm 2590$
$D^- \to K^{*0}K^-$	548 ± 105	1316 ± 236	2921 ± 260	5519 ± 435
$\Lambda_c^- \to \bar{p}K^+\pi^-$	1058 ± 153	1914 ± 181	5162 ± 367	8970 ± 529
Total Events	35 855	60 817	169 500	268 872

Table 6.4: The number of candidate events for each source in each sample. The size of the combinatoric background is not fitted, but rather is the balance of the above sources and the total sample size.

where Fr_{Signal}^{prompt} is the prompt fraction, and $P_{Signal}^{y,prompt}$ and $P_{Signal}^{y,long}$ are the *pdf*s for the selection variable for prompt and long-lived events, respectively. The long-lived events are those originating from a properly reconstructed B candidate. The lifetime distribution for the prompt background $P_{Signal}^{x^M,prompt}(x^M)$ is modelled by a pair of Gaussian functions and is assumed to be independent of the resolution (σ_{x^M}) and dilution (d_{pr}). The widths of the two Gaussians and their relative fractions are determined from Monte Carlo simulations of $c\bar{c}$ events, which represent the

majority of this background. The prompt *pdf* is determined to be

$$P^{x^M}_{Signal}(x^M) = 0.486 \cdot G(x^M, 0.003) + 0.514 \cdot G(x^M, 0.013), \tag{6.10}$$

where $G(x^M, \sigma)$ is a Gaussian function[1] in x^M centred at zero with width σ (in cm). The fraction Fr^{prompt}_{Signal} is fitted to the data. The four *pdf*s $P^{y,prompt}_{Signal}$, $P^{y,long}_{Signal}$, $P^{y,prompt}_{Bkgnd}$ and $P^{y,long}_{Bkgnd}$ are determined directly from the data (see Sec. 5.1).

The form of the lifetime *pdf* $P^{x^M,long}_j(x^M; \sigma_{x^M}, d_{pr})$[2] varies depending on the source j. The general *pdf* for a B (or $\bar{B}$) hadron decaying to final-state f (or $\bar{f}$) is given by one of the equations

$$\Gamma_j(B \to f) = N_j \frac{e^{-\Gamma_j t}}{2} \left\{ \cosh\left(\frac{\Delta\Gamma_j t}{2}\right) + \cos(\Delta m_j t) \right\}, \tag{6.11}$$

$$\Gamma_j(\bar{B} \to f) = N_j (1 + a_j) \frac{e^{-\Gamma_j t}}{2} \left\{ \cosh\left(\frac{\Delta\Gamma_j t}{2}\right) - \cos(\Delta m_j t) \right\}, \tag{6.12}$$

$$\Gamma_j(B \to \bar{f}) = N_j (1 - a_j) \frac{e^{-\Gamma_j t}}{2} \left\{ \cosh\left(\frac{\Delta\Gamma_j t}{2}\right) - \cos(\Delta m_j t) \right\}, \tag{6.13}$$

$$\Gamma_j(\bar{B} \to \bar{f}) = N_j \frac{e^{-\Gamma_j t}}{2} \left\{ \cosh\left(\frac{\Delta\Gamma_j t}{2}\right) + \cos(\Delta m_j t) \right\}, \tag{6.14}$$

where each source j has an average decay rate Γ_j and a normalization factor N_j. Mixing sources (B^0 and B^0_s) will have an oscillation frequency Δm_j and decay width difference $\Delta\Gamma_j$. The parameter a_j is the flavour-specific asymmetry in source j. Three asymmetry parameters are considered: the asymmetry in the

[1] We will use the notation $G(x, \sigma)$ to indicate a Gaussian function in x with zero mean and width σ. We can also use this notation to indicate a non-zero mean, for example, $G(x - x^M, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-x^M)^2}{2\sigma^2}\right)$.

[2] We will henceforth drop the designation 'long.'

B_s^0 sources (a_{fs}^s), the asymmetry in the B^0 sources (a_{fs}^d) and the asymmetry in the combinatoric background (a^{bkg}). We assume no direct $\mathcal{CP}$ violation in the semileptonic decay process.

The time, t, is the proper decay time of the B_s^0 candidate[3]. A Lorentz transformation must be made to determine the proper decay time from the measured decay length in the lab frame:

$$t = \frac{t^{lab}}{\gamma} = \frac{L^{lab}}{\gamma u} = \frac{L^{lab} M^{B_s}}{c p^{B_s}} = \frac{L_T^{lab} M^{B_s}}{c p_T^{B_s}}, \tag{6.15}$$

where γ is the Lorentz factor and t^{lab} is the decay time in the lab frame determined from the decay length L^{lab} and the speed u of the B_s^0 candidate. γ and u are, in turn, determined from the momentum of the B_s^0 candidate, p^{B_s}, which has been assigned the B_s^0 mass ($M^{B_s} = 5.366$ GeV/c^2). Measurements can be made more accurately in the transverse plane, hence we used the relationship $\frac{L^{lab}}{L_T^{lab}} = \frac{p^{B_s}}{p_T^{B_s}}$.

This definition of the proper time requires precise knowledge of the momentum of the B_s^0 candidate. However, the semileptonic decay of the B_s^0 meson includes a neutrino and possibly other particles that will not be detected. The momentum carried by these particles will not be included in the final reconstruction of the B_s^0 candidate, hence $p_T^{B_s}$ cannot be determined. Instead, we make a Lorentz transformation into the frame of the μD_s system. Furthermore, because the direction of the boost does not necessarily coincide with the true direction of the B_s^0 candidate,

[3]The proper decay time corresponds to the decay time in the B_s^0 frame of reference.

we project the decay length onto the transverse momentum of the μD_s system:

$$ct' = \frac{\vec{\mathbf{L}}_T^{lab} \cdot \vec{\mathbf{p}}_T^{\mu D_s} M^{B_s}}{(p_T^{\mu D_s})^2} = x^M. \tag{6.16}$$

The quantity x^M is referred to as the visible proper decay length (VPDL).

In order to determine the proper time in the B_s^0 frame, we must account for the missing energy. We use a correction factor, K, determined from MC simulations to scale the VPDL, $t = Kx^M/c$. The modelling of the K factor is discussed in Sec. 5.4. Substituting the VPDL expression into Eqns. 6.11-6.14 gives us the *pdf* in terms of measured parameters. This is demonstrated for Eqn. 6.11:

$$\Gamma_j(B \to f) = \frac{K}{c\tau_j} \exp\left(\frac{-Kx^M}{c\tau_j}\right) \frac{1}{2} \left\{ \cosh\left(\frac{\Delta\Gamma_j K x^M}{2c}\right) + \cos\left(\frac{\Delta m_j K x^M}{c}\right) \right\}, \tag{6.17}$$

where we have replaced the average decay rate Γ_j with the average lifetime τ_j ($\Gamma_j = 1/\tau_j$). The function has been normalized so that the integral is unity ($N_j = K/c\tau_j$). Similar substitutions can be made for Eqns. 6.12-6.14.

Because we cannot be certain of the initial-state flavour tag, we must make a weighted sum of the mixed and unmixed *pdf*s. The relative weight of the two is determined by the flavour tag and dilution. The average *pdf*s are analogous to Eqn. 6.11 - 6.14 and are denoted Γ_j^{+-}, Γ_j^{++}, Γ_j^{-+} and Γ_j^{--}, respectively. These functions are the *hypothesized pdf*s, and should not be confused with the 'truth' *pdf*s, $\Gamma_j(\bar{B} \to \bar{f})$. In the four $\Gamma_j^{\pm\pm}$ functions, the first and second superscripts indicate the sign of the final-state and initial-state flavour tags, respectively. Like-sign *pdf*s ($++$ and $--$) describe oscillated processes ($B \to \bar{f}$ and $\bar{B} \to f$) and opposite-

sign *pdf*s ($+-$ and $-+$) describe non-oscillated processes ($B \to f$ and $\bar{B} \to \bar{f}$). For example, Γ_j^{+-} indicates a final-state μ^+ and an initial-state μ^- (assuming a muon OST). Given a perfect flavour tagger, we would have $\Gamma_j^{+-} = \Gamma_j(B \to f)$. However, because there is some uncertainty in the initial-state tag, there is also a contribution from $\Gamma_j(\bar{B} \to f)$. The contribution from each tagging hypothesis (i.e., correct and incorrect) is scaled by the dilution ($\mathcal{D}$). These functions are given by the expressions

$$\Gamma_j^{+-}(x^M; K, d_{pr}) = \Gamma_j(B \to f)\frac{1+\mathcal{D}(d_{pr})}{2} + \Gamma_j(\bar{B} \to f)\frac{1-\mathcal{D}(d_{pr})}{2}, \quad (6.18)$$

$$\Gamma_j^{++}(x^M; K, d_{pr}) = \Gamma_j(\bar{B} \to f)\frac{1+\mathcal{D}(d_{pr})}{2} + \Gamma_j(B \to f)\frac{1-\mathcal{D}(d_{pr})}{2}, \quad (6.19)$$

$$\Gamma_j^{-+}(x^M; K, d_{pr}) = \Gamma_j(B \to \bar{f})\frac{1+\mathcal{D}(d_{pr})}{2} + \Gamma_j(\bar{B} \to \bar{f})\frac{1-\mathcal{D}(d_{pr})}{2}, \quad (6.20)$$

$$\Gamma_j^{--}(x^M; K, d_{pr}) = \Gamma_j(\bar{B} \to \bar{f})\frac{1+\mathcal{D}(d_{pr})}{2} + \Gamma_j(B \to \bar{f})\frac{1-\mathcal{D}(d_{pr})}{2}. \quad (6.21)$$

To correctly model the measured VPDL in the equations above, we convolute the lifetime *pdf* with a Gaussian function describing the VPDL uncertainty and integrate over the *pdf* for the K factor momentum correction. The VPDL uncertainty is tuned to the data to ensure it is properly estimated. This tuning introduces parameters s_1 and s_2 that scale the uncertainty σ_{x^M}. This procedure is described in Sec. 5.2. Therefore,

$$P_j^{q_f q_i}(x^M; \sigma_{x^M}, d_{pr}) = \int_{K_{min}}^{K_{max}} dK\ D_j(K) \cdot Eff_j(x^M) \cdot$$
$$\int_0^\infty dx\ \left(fr_1 \cdot \frac{G(x - x^M, s_1\sigma_{x^M})}{N_j(K, s_1\sigma_{x^M}, d_{pr})} + (1 - fr_1) \cdot \frac{G(x - x^M, s_2\sigma_{x^M})}{N_j(K, s_2\sigma_{x^M}, d_{pr})} \right)$$
$$\cdot \Gamma_j^{q_f q_i}(x; K, d_{pr}), \quad (6.22)$$

where $G(x-x^M, s_{1/2}\sigma_{x^M})$ is a Gaussian function describing the detector resolution for the x^M and fr_1 is the fractional contribution of the s_1 scaling parameter. We will refer to this function as the lifetime resolution function. $Eff_j(x^M)$ is the selection efficiency for channel j (see Sec. 5.5). The function $D_j(K)$ gives the normalized distribution for the K factor in channel j (see Sec.5.4). The charges q_f and q_i are the signs of the final-state and initial-state flavour tags, respectively. Note that we have replaced the notation $P_j^{x^M}$ with $P_j^{q_f q_i}$. This divides $P_j^{x^M}$ into four distinct *pdf*s for each of four tagging outcomes ($++$, $+-$, $-+$ and $--$). However, this still corresponds to the function $P_j^{x^M,long}$ of Eqn. 6.9. The normalization factor N_j is calculated by integrating over the entire VPDL region using the average lifetime function,

$$N_j(K, \sigma_{x^M}, d_{pr}) = \int_{-\infty}^{\infty} dx^M \; Eff_j(x^M) \int_0^{\infty} dx \; G(x - x^M, \sigma_{x^M}) \cdot \frac{K}{c\tau_j} \exp\left(-\frac{Kx}{c\tau_j}\right). \tag{6.23}$$

Each of the five sources listed in Sec. 6.1 has contributions from a number of different channels. The B_s^0 signal, for example, has a number of different D_s^- decays that contribute to the overall signal. The lifetime distributions for these channels have some differences, which will be outlined in the following sections.

6.2.1 Lifetime *pdf* for $D_s^- \to K^{*0}K^-$ Signal

The μD_s source is composed mostly of B_s^0 mesons with some contributions from B^+ and B^0 mesons. The different species of B mesons have different $\mathcal{CP}$ asym-

metries and behave differently with respect to oscillations. B^0 contributions are assigned a separate asymmetry parameter from the B_s^0 signal, while the charged B^+ meson has no mixing asymmetry. Similarly the neutral B^0 meson oscillates with a frequency different from that of the B_s^0 meson, while the B^+ meson does not oscillate.

The complete list of contributions to the μD_s signal is given in Sec. 5.3. The majority of these sources use the standard *pdf* described in Sec. 6.2. However, some sources require that the *pdf* be modified. The $B_s^0 \to D_s^- D_s^+ X$ final-state is indistinguishable from the similar decay for $\bar{B}_s^0$, and so the *pdf*s must be modified to average the final-state tags:

$$\begin{aligned}
\Gamma^{+-}_{D_sD_s}(x;K,d_{pr}) &= \Gamma^{--}_{D_sD_s}(x;K,d_{pr}) = (\Gamma^{+-}_{B_s} + \Gamma^{--}_{B_s})/2, \\
\Gamma^{++}_{D_sD_s}(x;K,d_{pr}) &= \Gamma^{-+}_{D_sD_s}(x;K,d_{pr}) = (\Gamma^{++}_{B_s} + \Gamma^{-+}_{B_s})/2,
\end{aligned} \tag{6.24}$$

where the functions $\Gamma^{\pm\pm}_{B_s}$ correspond to Eqn. 6.18-6.20 with B_s^0 parameters inserted.

There is a contribution from $B_{u/d} \to D_s^- DX$ and $\bar{B}_s^0 \to D_s^- DX$ decays (where $D \to \mu^+$). In this case, the correlation of the charge of the final-state muon to the final-state flavour is opposite ($b \to \mu^+ X$ rather than $b \to \mu^- X$). This is taken into account by flipping the final-state tag for these contributions. The B^0 meson oscillates with frequency Δm_d and is assumed to have a negligible width difference ($\Delta\Gamma = 0$). Furthermore, the B^+ contribution does not oscillate, therefore, the truth same-sign functions (Eqns. 6.12 and 6.13) vanish and the effective same-sign functions (Eqns.6.19 and 6.21) only contribute due to dilution effects.

In total, there are ten contributions to the μD_s signal that are considered. The

total VPDL *pdf* is a sum over all of these sources that give the D_s mass peak:

$$P^{q_f q_i}_{\mu D_s}(x^M; \sigma_{x^M}, d_{pr}) = \sum_k \mathcal{F}_k \left(M\left(\mu D_s\right)\right) \cdot P^{q_f q_i}_k(x^M; \sigma_{x^M}, d_{pr}), \tag{6.25}$$

where $\mathcal{F}_k \left(M\left(\mu D_s\right)\right)$ is the fractional contribution from the k^{th} channel given the invariant mass of the μD_s system.

6.2.2 Lifetime *pdf* for $D^- \to K^{*0} K^-$ Signal

The $\mu D^{\pm}$ signal, which forms a small peak on the left of the signal peak (see Sec. 6.1), was considered in the final fit. This peak is composed mainly of B^0 and B^+ decays and is modelled in the same manner as the signal contributions. These contributions have $\Delta\Gamma = 0$ and only the B^0 contribution includes oscillation effects. A single average K factor value is used for this source, determined from MC simulations ($K = 0.81$).

The B meson lifetime is reconstructed assuming a B^0_s meson mass (see Eqn. 6.16), which must be corrected in this source. The VPDL is scaled by the ratio M^{B^0}/M^{B_s}. Both the B^0 and B^+ contributions share a reconstruction efficiency function determined from a joint sample of $B_{d/u} \to \mu D^{\pm}(KK\pi)X$ MC (see Sec. 5.5). The B^0 decay forms 80% of this peak, with B^+ making up the remainder. This fraction is determined from the number of B^0 and B^+ candidates passing the selection cuts in the MC sample. The total *pdf* is given by

$$P^{q_f q_i}_{D^+(KK\pi)}(x^M; \sigma_{x^M}, d_{pr}) = 0.8\, P^{q_f q_i}_{B_d}(x^M; \sigma_{x^M}, d_{pr}) + 0.2\, P^{q_f q_i}_{B_u}(x^M; \sigma_{x^M}, d_{pr}). \tag{6.26}$$

6.2.3 Lifetime *pdf* for Reflections

The $\mu D^{\pm}$ reflection is the largest physics contribution and overlaps the μD_s signal peak. This contribution is modelled in the same manner as the $\mu D^{\pm}$ signal with the reconstruction efficiency replaced with a curve determined from a $B_d/B_u \to \mu D^{\pm}(K\pi\pi)X$ MC sample. In addition, the VPDL has a shift applied to account for the mass difference between the B_s^0 and B^0 mesons. The total lifetime *pdf* is

$$P^{q_f q_i}_{D^+(K\pi\pi)}(x^M;\sigma_{x^M},d_{pr}) = 0.8\ P^{q_f q_i}_{B_d}(x^M;\sigma_{x^M},d_{pr}) + 0.2\ P^{q_f q_i}_{B_u}(x^M;\sigma_{x^M},d_{pr}). \quad (6.27)$$

The $\mu\Lambda_c^{\pm}$ reflection forms a low but broad peak. The lifetime *pdf* is similar to the B^+ contribution ($\Delta\Gamma = 0$, no mixing), but with a reconstruction efficiency curve determined from a $\Lambda_b^0 \to \Lambda_c^- \mu^+ \nu_\mu X$ MC sample. The VPDL is shifted to account for the mass difference between the B_s^0 and Λ_b^0 hadrons. The total lifetime *pdf* is an exponential function.

6.2.4 Lifetime *pdf* for Combinatoric Background

The following contributions to the combinatoric background were considered:

1. Prompt background with the μD_s vertex close to the PV, described as a Gaussian with a variable width determined by the lifetime resolution on an event-by-event basis (σ_{x^M}, discussed in Sec. 6.2).

2. Background with vertices distributed around the PV. This background is similar to the prompt background. It is described as a Gaussian function, however, it has a constant width $\sigma_{c\bar{c}}$ that is fitted to the data.

3. Long-lived background, described by an exponential function with a constant decay length $c\tau_{long}$ and convoluted with the lifetime resolution function.

4. A positive exponential convoluted with the lifetime resolution function with a constant decay length $c\tau_{bkg}$ to take into account the outliers in the positive tail.

The long-lived background is further divided into three subsamples:

1. Events insensitive to the tagging. These are events that have a single lifetime *pdf* to describe any tagging state, making them completely insensitive to the tag (for example, $B_s^0 \to D_s^- D_s^+ X$).

2. Events sensitive to the tagging and non-oscillating. These are events that require separate lifetime *pdf*s to describe the same-sign ($++$ and $--$) and opposite-sign ($+-$ and $-+$) tagging states, but do not include any oscillation (for example, B^+ processes).

3. Events sensitive to the tagging and oscillating with frequency Δm_d. These are B^0 decays that oscillate and must be modelled in a similar fashion as the signal *pdf*.

The fractions of these contributions and their parameters were determined from fits to the data. Tagging-sensitive background sources were taken to have a dilution with the same tagging-variable (d_{pr}) dependence as exists in signal sources. The

background lifetime *pdf* has the following form:

$$P_{bkg}(x^M, y; \sigma_{x^M}, d_{pr}) = \mathcal{F}_{c\bar{c}} \cdot G(0 - x^M, \sigma_{c\bar{c}}) \cdot P_{Bkgnd}^{y,prompt} + (1 - \mathcal{F}_{c\bar{c}}) \cdot P_{bkg}^{res}(x^M, y; \sigma_{x^M}, d_{pr}), \quad (6.28)$$

$$P_{bkg}^{res}(x^M, y; \sigma_{x^M}, d_{pr}) = \frac{Eff_{bkg}(x^M)}{N} \int_0^\infty dx\ G(x - x^M, s_{bkg}\sigma_{x^M}) \Big(\mathcal{F}_0 P_{Bkgng}^{y,prompt} \delta(x) + (1 - \mathcal{F}_0) P_{Bkgng}^{y,long} \big(\mathcal{F}_{long} \cdot \Gamma_{long}^{q_f q_i} + (1 - \mathcal{F}_{long}) \cdot \Gamma_{bkg}^{q_f q_i} \big) \Big), \quad (6.29)$$

$$\Gamma_{long}^{q_f q_i}(x; d_{pr}) = \Gamma_{notag,nomix}^{q_f q_i} \cdot (1 - \mathcal{F}_{tsens}) + \mathcal{F}_{tsens} \left(\Gamma_{tag,nomix}^{q_f q_i} \cdot (1 - \mathcal{F}_{osc}) + \Gamma_{tag,mix}^{q_f q_i} \cdot \mathcal{F}_{osc} \right), \quad (6.30)$$

where N is a normalization constant. The functions $\Gamma_{bkg}^{q_f q_i}$ and $\Gamma_{notag,nomix}^{q_f q_i}$ are exponential functions with lifetimes τ_{bkg} and τ_{long}, respectively and model the tagging-insensitive backgrounds. The functions $\Gamma_{tag,nomix}^{q_f q_i}$ and $\Gamma_{tag,mix}^{q_f q_i}$ model the tagging-sensitive backgrounds and are defined along the same lines as Eqns. 6.18-6.20. $\Gamma_{tag,nomix}^{q_f q_i}$ describes the unmixed background (as for the B^+ meson discussed above), and $\Gamma_{tag,mix}^{q_f q_i}$ describes the mixing background ($\Delta m_{bkg} = \Delta m_d$, $\Delta\Gamma_{bkg} = 0$). These contributions have lifetime τ_{long}.

The fit parameters are $\mathcal{F}_{c\bar{c}}$, $\sigma_{c\bar{c}}$, $\mathcal{F}_0$, $\mathcal{F}_{tsens}$, $\mathcal{F}_{long}$, $\mathcal{F}_{osc}$, $c\tau_{long}$, $c\tau_{bkg}$ and s_{bkg}. The efficiency function $Eff_{bkg}(x^M)$ is fitted to the data. The form of this function is described in Sec. 5.5.

6.3 Lifetime Fit

We fit the lifetime distribution of the combinatoric background before performing the fits to determine the oscillation frequency or the asymmetry. We simultaneously fit the lifetime of the B_s^0 meson in order to allow for any lifetime bias due to the selection cuts. The effect of any bias will be checked and included in the systematic uncertainty on the final measurement. In addition the selection efficiency function (see Sec. 5.5) in the combinatoric background will be fitted with the other parameters.

To ensure these fits do not bias the later results, we randomize the $\mathcal{CP}$-dependent variables (i.e., muon pseudorapidity and toroid polarity, see Sec. 5.6) and set the dilution to zero in the signal and physics backgrounds. Furthermore, the detector asymmetry parameters (see Sec. 5.6.1) are fixed to zero. This removes any dependence on actual B_s^0 or B^0 oscillations or asymmetry in these sources.

The fit is performed separately in the tagged and untagged data samples for both the Run IIa and Run IIb data sets. Results of the fits are shown in Figs. 6.4 and 6.5. The fitted parameters are given in Table 6.5. The fit in the tagged sample will be used for the fits to determine the oscillation frequency and the fit in the untagged sample will be used for the fit to the $\mathcal{CP}$ asymmetry.

The untagged fits in Fig. 6.5 undershoot the data in the negative VPDL region. This is due to the modelling of the selection efficiency (see Sec. 5.5). The efficiency function suppresses events near zero, but becomes 100% efficient for $VPDL \lesssim -0.1$ cm. This function was chosen because it can be expressed in an analytic manner and is easily integrable, however, it does not correctly model the efficiency in the

negative VPDL region. To verify that this does not impact the result, a cross-check is performed to test a 'worst case' senario. The largest deviation of the efficiency would be a 0% efficiency for the $VPDL \lesssim -0.1$ cm. We test this senario by repeating the analysis[4] with this worst case efficiency applied to all channels. We find no noticeable difference in the measured oscillation frequency in this senario and conclude that our results are not sensitive to this discrepancy.

	Tagged		Untagged	
Parameter	Run IIa	Run IIb	Run IIa	Run IIb
$\mathcal{F}_{c\bar{c}}$	0.035 ± 0.005	0.045 ± 0.005	0.051 ± 0.003	0.054 ± 0.003
$\sigma_{c\bar{c}}$	0.0071 ± 0.0005	0.0047 ± 0.0003	0.0066 ± 0.0002	0.0059 ± 0.0001
$\mathcal{F}_0$	0.059 ± 0.006	0.083 ± 0.004	0.117 ± 0.003	0.139 ± 0.002
$c\tau_{bkg}(\mu m)$	317 ± 19	305 ± 17	321 ± 11	294 ± 7
$\mathcal{F}_{long}$	0.62 ± 0.03	0.65 ± 0.03	0.55 ± 0.03	0.58 ± 0.02
$\mathcal{F}_{tsens}$	1.00 ± 0.06	0.93 ± 0.06	1.00 ± 0.01	1.00 ± 0.02
$\mathcal{F}_{osc}$	0.63 ± 0.04	0.61 ± 0.03	0.60 ± 0.04	0.63 ± 0.03
$c\tau_{long}(\mu m)$	666 ± 12	643 ± 12	659 ± 10	636 ± 7
$c\tau_{B_s}(\mu m)$	411 ± 17	414 ± 13	422 ± 8	411 ± 6
Fr^{prompt}_{Signal}	0.011 ± 0.004	0.006 ± 0.003	0.039 ± 0.002	0.035 ± 0.002
scale factor	1.71 ± 0.03	1.66 ± 0.02	1.74 ± 0.01	1.69 ± 0.01
$Eff_{bkg}\sqrt{p_1}$	0.021 ± 0.002	0.0135 ± 0.0009	0.0184 ± 0.0005	0.0149 ± 0.0003
$Eff_{bkg}p_2$	0.63 ± 0.02	0.77 ± 0.01	0.684 ± 0.007	0.728 ± 0.005
$Eff_{bkg}p_3$	-11.6 ± 1.7	-17.7 ± 1.9	-9.0 ± 0.7	-11.0 ± 0.7
$Eff_{bkg}p_4$	54 ± 290	19 ± 460	43 ± 73	100 ± 110
$Eff_{bkg}p_5$	-1760 ± 6900	-1225 ± 9517	-3472 ± 1721	-7717 ± 2982

Table 6.5: Results from the fits to the B_s^0 candidate lifetime distribution.

[4]We redo the lifetime fit and the scan in amplitude (see Sec. 7.1) and look for the oscillation frequency.

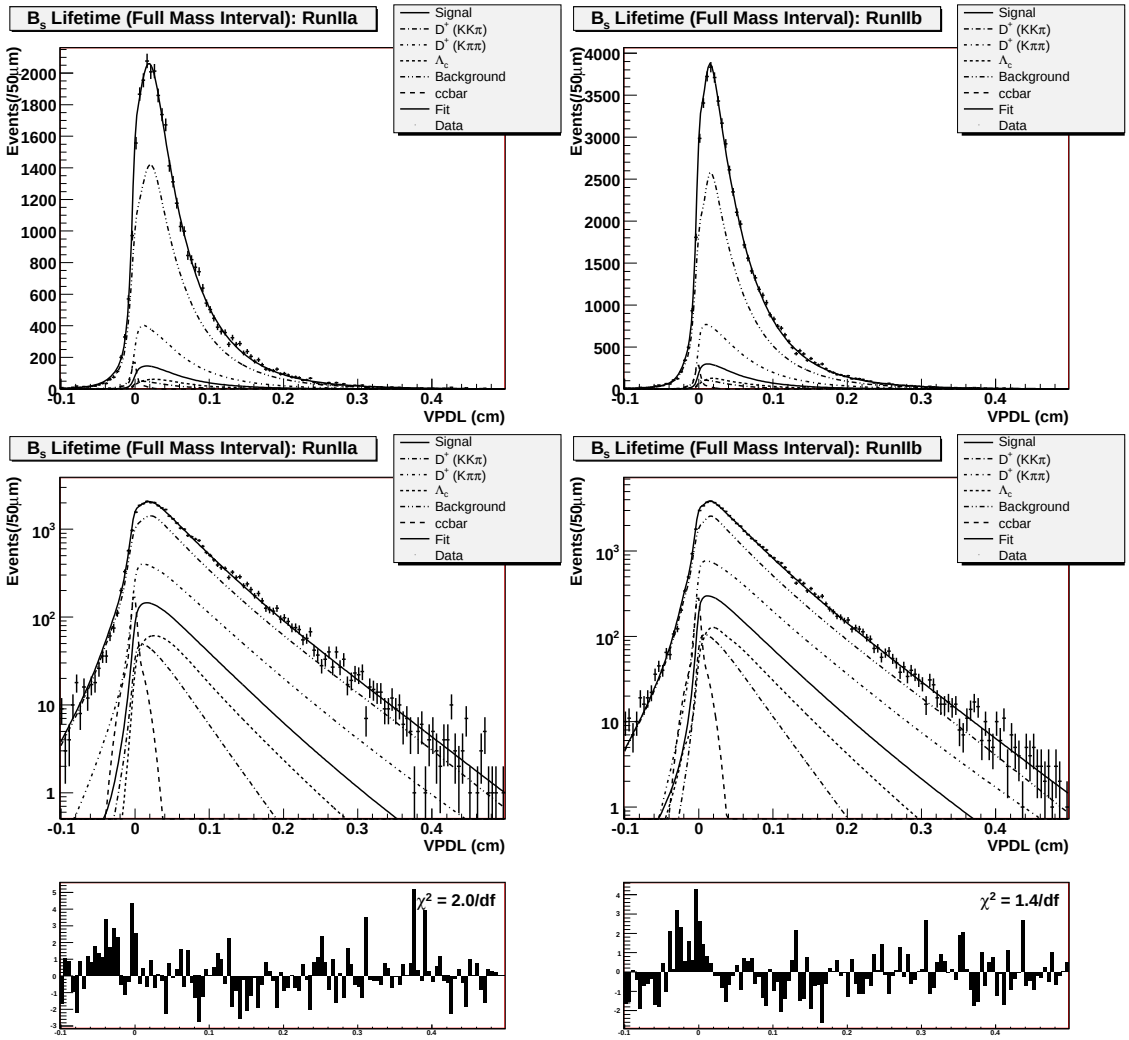

Figure 6.4: Fit to the lifetime distribution in the tagged data sample on a linear scale (top), a log scale (centre). The normalized difference ($\frac{\text{fit}-\text{data}}{\sigma}$) is show at bottom. Run IIa plots are on the left and Run IIb plots are on the right.

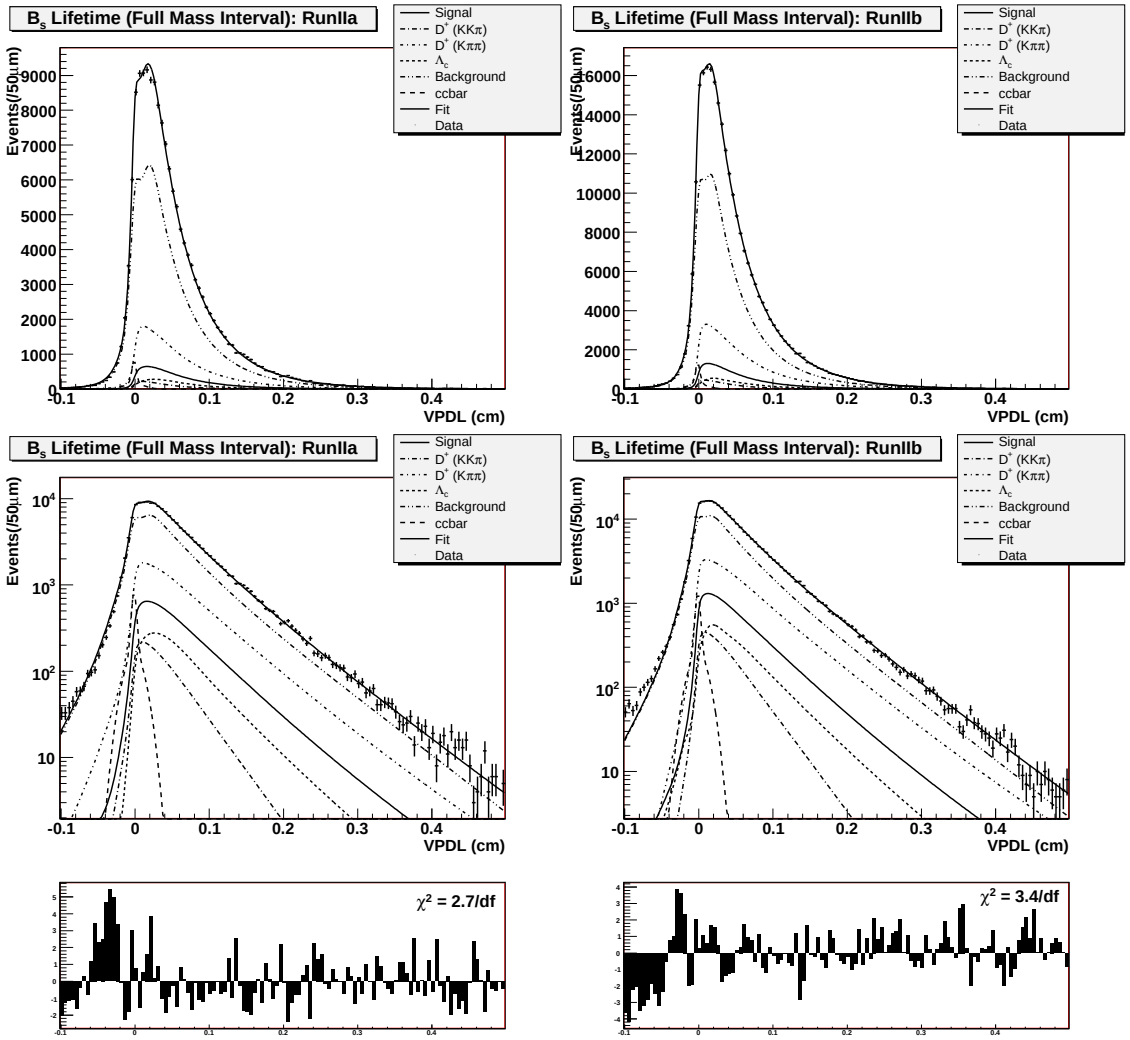

Figure 6.5: Fit to the lifetime distribution in the untagged data sample on a linear scale (top), a log scale (centre). The normalized difference ($\frac{\text{fit}-\text{data}}{\sigma}$) is show at bottom. Run IIa plots are on the left and Run IIb plots are on the right.

Chapter 7

Results

7.1 Oscillation Frequency

We can attempt to make a measurement of the oscillation frequency by scanning through a range of Δm_s values, measuring the log likelihood $(\ln L)$[1] at each Δm_s point. The minimal value of $-2\ln L$ will correspond to the most probable value of Δm_s. The extraction of the Δm_s uncertainty from the likelihood has been discussed in Sec. 2.2. The statistical significance of the measurement is determined by the square root of the depth of the minimum with respect to the likelihood plateau at high Δm_s. The analysis is not sensitive to such high oscillation frequencies, and so this region determines the likelihood of the null hypothesis (i.e., no flavour oscillations). The result of the likelihood scan is shown in Fig. 7.1. We plot the change in the log likelihood with respect to the global minimal value $(\Delta(-2\ln L))$. There is a global minimum in the Run IIb scan from which we ex-

[1]We will use this notation to remind ourselves that this is a *log*likelihood. This should not be confused with the notation used in Eqn. 2.2 where L is the log likelihood.

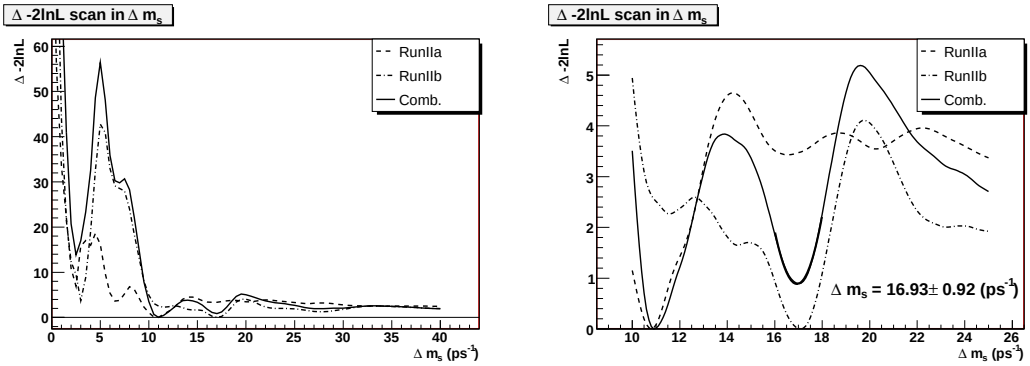

Figure 7.1: Likelihood scans in Δm_s showing the large scale behaviour and the behaviour around the minimum. A fitted value is extracted from the combined log likelihood minimum at ~ 17 ps^{-1}.

tract $\Delta m_s = 16.99 \pm 0.99$ ps^{-1}. This minimum becomes a local minimum in the combined fit, from which we extract $\Delta m_s = 16.93 \pm 0.92$ ps^{-1}. Given that the minimum is local and not global, the $\sim 1.2\sigma$ significance of the extracted Δm_s value is questionable.

We can also use a method similar to a power spectrum analysis to measure the oscillation frequency of the B_s^0 meson. We insert a parameter ($\mathcal{A}$) into the B_s^0 lifetime *pdf* that scales the oscillating term (i.e., the cosine of Eqns. 6.11-6.14):

$$\Gamma_s^{\pm,\pm} \propto \left\{ \cosh\left(\frac{\Delta\Gamma_j t}{2}\right) \pm \mathcal{A}\cos(\Delta m_j t) \right\}. \tag{7.1}$$

We then scan through a range of Δm_s values and fit for $\mathcal{A}$ at each Δm_s value. The amplitude parameter, $\mathcal{A}$, may be interpreted as the power of the oscillation at a particular value of Δm_s. The amplitude should fluctuate around zero until we strike upon the 'correct' value of Δm_s, where $\mathcal{A}$ should peak with $\mathcal{A} = 1$. This formalism is convenient when it is not possible to make a direct measurement as

it allows us to place limits on Δm_s. The sensitivity of the search is determined by finding the Δm_s value at which we can no longer rule out an amplitude value of zero from fluctuating to one (at 95% CL). We will not be sensitive to a true Δm_s value above this point. The sensitivity of the measurement is determined primarily by the statistics. The Δm_s lower limit, in the absence of a Δm_s measurement, is set by the lowest Δm_s value for which the fitted amplitude can fluctuate to one. In the event that there is a higher value of Δm_s for which we can rule out such a fluctuation, we can set an upper limit for Δm_s (provided this value is below the sensitivity bound).

The results of the amplitude scans with statistical uncertainties are shown in Fig. 7.2. The scan is done separately for the Run IIa and Run IIb data samples and also a combined fit to both samples is done. The combined fit sums the total log likelihood of each data sample to obtain a combined log likelihood. The results of these fits yield a Δm_s lower limit, $\Delta m_s > 10.0$ ps^{-1}, with a sensitivity to $\Delta m_s < 15.5$ ps^{-1}.

7.1.1 Cross-Checks

The amplitude fitting code must be cross-checked to ensure it can correctly determine the oscillation frequency. We use the well established oscillations in the B^0 system as a cross-check. To determine Δm_d, we set the dilution, $D(d_{pr})$, to zero for the B_s^0 contributions. This averages the mixed and unmixed *pdf*s, effectively removing any sensitivity to the oscillations in the B_s^0 signal. We then perform a log likelihood scan and extract a measurement of Δm_d. In addition, we perform an amplitude scan with the amplitude parameter associated with the B^0 sources.

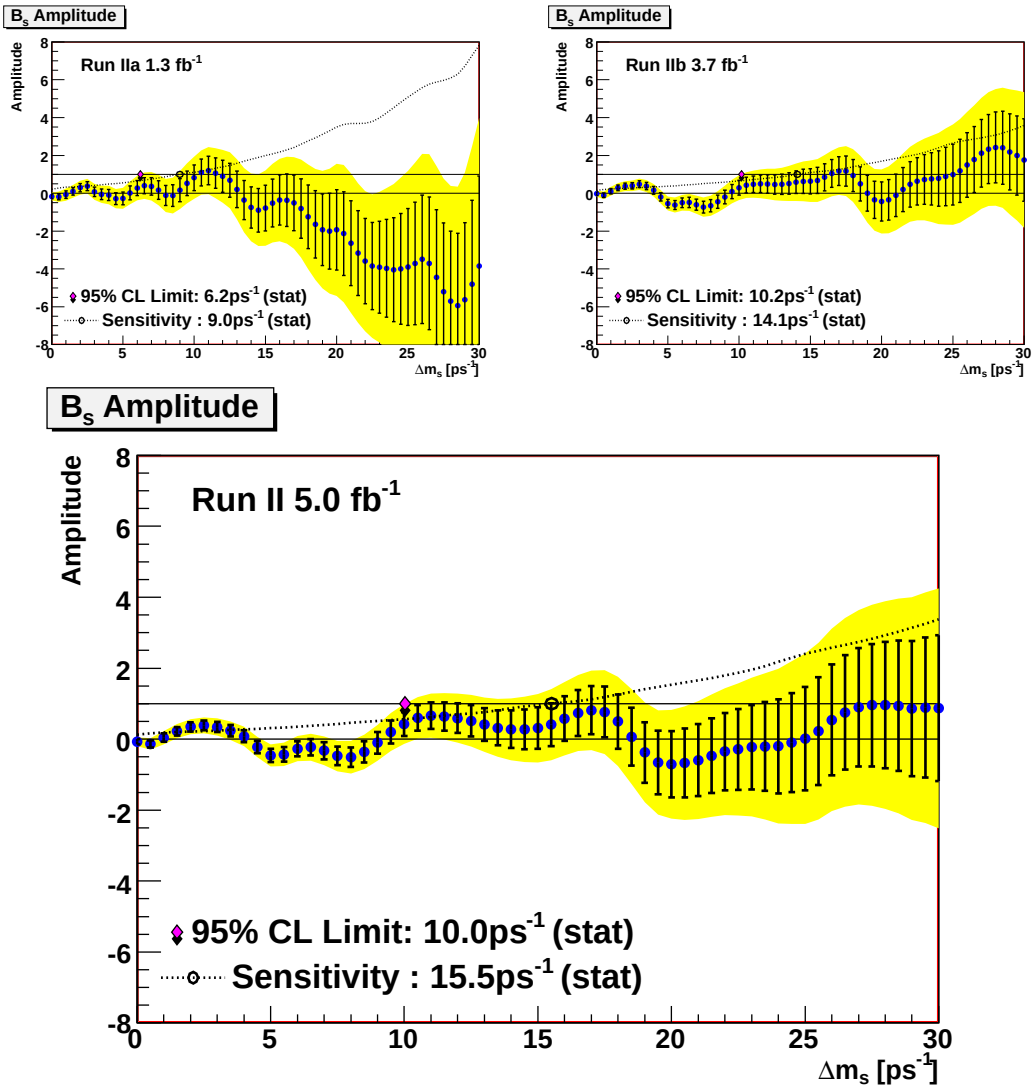

Figure 7.2: B_s^0 amplitude scan results from Run IIa (top left), Run IIb (top right) and the combined scan (bottom). Points are the fitted amplitude value with 1σ error bars. The band represents the 95% confidence interval over the range of Δm_s.

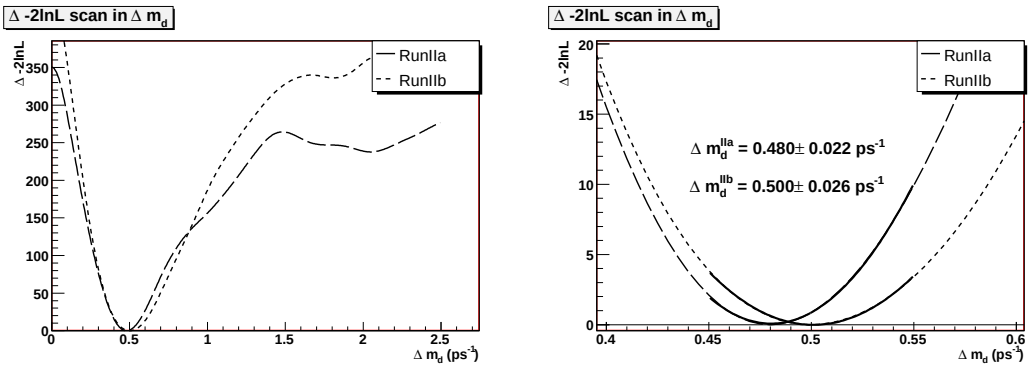

Figure 7.3: Likelihood scans in Δm_d showing the correct oscillation frequency for the B_d^0 meson.

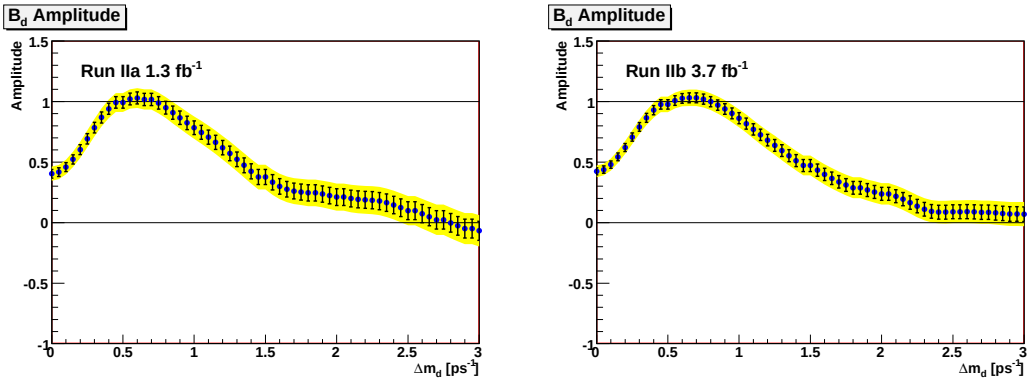

Figure 7.4: Amplitude scans in Δm_d showing the peak at ~ 0.5 ps^{-1}.

Fig. 7.3 shows the log likelihood scan from which we extract $\Delta m_d = 0.480 \pm 0.022$ ps^{-1} from the Run IIa data sample and $\Delta m_d = 0.500 \pm 0.026$ ps^{-1} from the Run IIb data sample. These results are consistent with the world average value $\Delta m_d = 0.507 \pm 0.005$ ps^{-1} [9]. Fig. 7.4 shows the results of the Δm_d amplitude scan. Δm_d is scanned with a resolution of 0.05 ps^{-1} and peaks are found at $\Delta m_d = 0.55$ ps^{-1} (Run IIa) and $\Delta m_d = 0.60$ ps^{-1} (Run IIb) with peak values $\mathcal{A}_d = 1.03 \pm 0.05$ (Run IIa) and $\mathcal{A}_d = 1.03 \pm 0.04$ (Run IIb), consistent with unity.

The Δm_d peak amplitude check is also useful for ensuring that the calibration curve for the predicted dilution is correct. The dilution effectively scales the oscillating term of the lifetime *pdf* in the same manner as the amplitude parameter. If the dilution has been mis-calibrated and is too low/high, the fitted peak amplitude will compensate by shifting up/down from unity. The peak amplitude values listed above are consistent with unity and so we conclude that the dilution is properly calibrated.

We also check that the fitting procedure is not somehow biased to find an oscillation where none exists regardless of the data. To test this, we perform an ensemble test of amplitude fits at some fixed value of Δm_s. For each fit the flavour tag is randomized for all events. The randomized tag obscures any real oscillations in all sources so that the only contributions to an amplitude peak are from statistical fluctuations in the lifetime and flavour tag. The ensemble of fits should verify that the amplitude is distributed around zero with fluctuations consistent with the statistical uncertainty on the fitted amplitude. Approximately 600 fits are conducted for each of the Run IIa and Run IIb data samples. The resultant distribution of the fitted amplitude should be a Gaussian function centred at zero

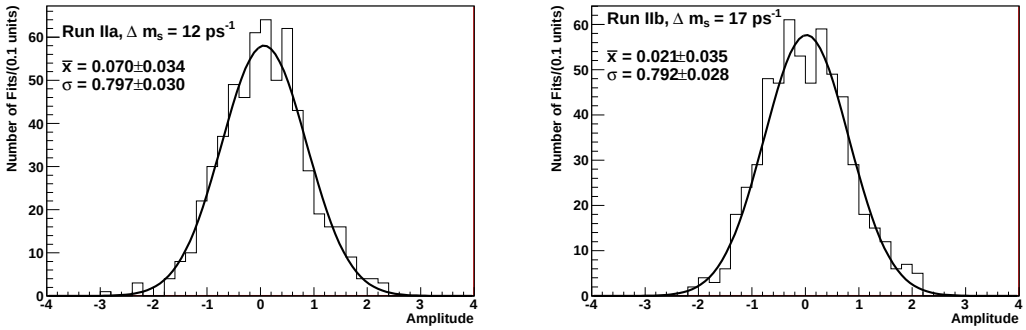

Figure 7.5: Histograms showing the results of the ensemble test of the amplitude fit.

with a width consistent with the statistical uncertainty of the fitted amplitude at the given value of Δm_s. We choose to test Δm_s values $\Delta m_s = 12$ ps^{-1} (Run IIa) and $\Delta m_s = 17$ ps^{-1} (Run IIb) near the principle amplitude peaks of Fig. 7.2. Gaussian functions are fitted to the amplitude distribution. The results of this test are shown in Fig. 7.5. The means of the fitted Gaussians ($\bar{x}$) are consistent with zero and the widths (σ) are consistent with the amplitude uncertainty (± 0.83 and ± 0.77 in Run IIa and Run IIb, respectively). Therefore, we conclude that there is no discernible amplitude bias.

7.2 Asymmetry Fit

The fit to the asymmetry requires that the value for Δm_s be fixed. We take the CDF measured value $\Delta m_s = 17.77$ ps^{-1} [25]. The signal asymmetry (a^s_{fs}) is fitted simultaneously with a^d_{fs}, a^{bkg} and the detector asymmetries described in Sec. 5.6.1. We make separate fits to the Run IIa and Run IIb data samples and

make a combined fit in both samples by summing the log likelihoods. The results are shown in Table 7.1. Note that because all events are equally sensitive to all of the detector asymmetries, they share equal statistical uncertainties. We can verify the fit by examining the log likelihood scan in the asymmetry. This testing is done in a similar manner as for the oscillation frequency; however, we must fix the background and detector asymmetries to their measured values. The result shows the likelihood minimum at the correct value with a width consistent with the statistical uncertainty determined from the direct fit (see Fig. 7.6).

Parameter	Run IIa	Run IIb	Combined
a^s_{fs}	0.089 ± 0.042	-0.020 ± 0.030	0.018 ± 0.025
a^d_{fs}	0.035 ± 0.032	0.056 ± 0.024	0.049 ± 0.019
a^{bkg}	-0.007 ± 0.023	0.007 ± 0.017	0.002 ± 0.014
A_{fb}	-0.0022 ± 0.0024	-0.0021 ± 0.0019	-0.0021 ± 0.0015
A_{det}	0.0050 ± 0.0024	0.0014 ± 0.0019	0.0028 ± 0.0015
$A_{q_\mu \beta}$	-0.0006 ± 0.0024	0.0034 ± 0.0019	0.0019 ± 0.0015
$A_{\beta \gamma_\mu}$	-0.0018 ± 0.0024	0.0015 ± 0.0019	0.0002 ± 0.0015
A_{ro}	-0.0275 ± 0.0024	-0.0317 ± 0.0019	-0.0301 ± 0.0015

Table 7.1: Results of the asymmetry fits in Run IIa, Run IIb and their combination.

7.2.1 Cross-Checks

In a similar manner as above, we cross-check the asymmetry code to verify that it is not biasing the result. We perform two ensemble tests. The first test is performed on data with the muon charge randomized in an ensemble of 500 asymmetry fits. This test randomizes all sources of charge asymmetry in the data. The distribution of the fitted asymmetry should be Gaussian around zero with a width consistent with the statistical uncertainty on the asymmetry. This test is conducted on the

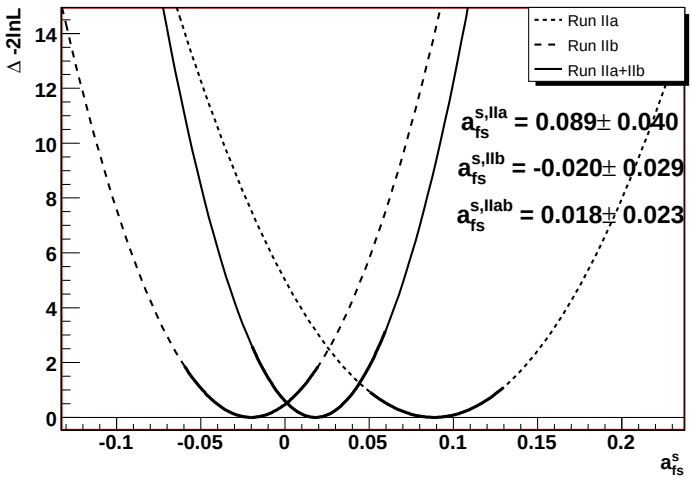

Figure 7.6: Likelihood scans in a_{fs}^{s} showing the fitted values. These are determined by fitting a quadratic function to the minima and do not correspond exactly with the direct fits shown in Table 7.1.

combined Run IIa and Run IIb samples. The results are shown in Fig. 7.7. The fitted Gaussian has a mean $\bar{x} = -0.0008 \pm 0.0009$, consistent with zero and a width $\sigma = 0.026 \pm 0.001$, consistent with the statistical uncertainty on the asymmetry ($\sigma_{measure} = 0.025$). The ensemble distributions for a_{fs}^{d} and a^{bkg} also have means consistent with zero and widths consistent with the measured uncertainty.

This test is useful to verify that the fitting procedure does not introduce any asymmetry bias. However, we cannot use this test to verify that we can recover a correct asymmetry. Unlike the oscillation measurement where we could test Δm_d, there is no easily detected asymmetry to verify. If we wish to perform this sort of test, we must put the asymmetry in 'by hand.' We can randomize the flavour tag in the data in such a way as to introduce an asymmetry; however, we cannot introduce a different B_s^0, B^0 or background asymmetry. To conduct such a test,

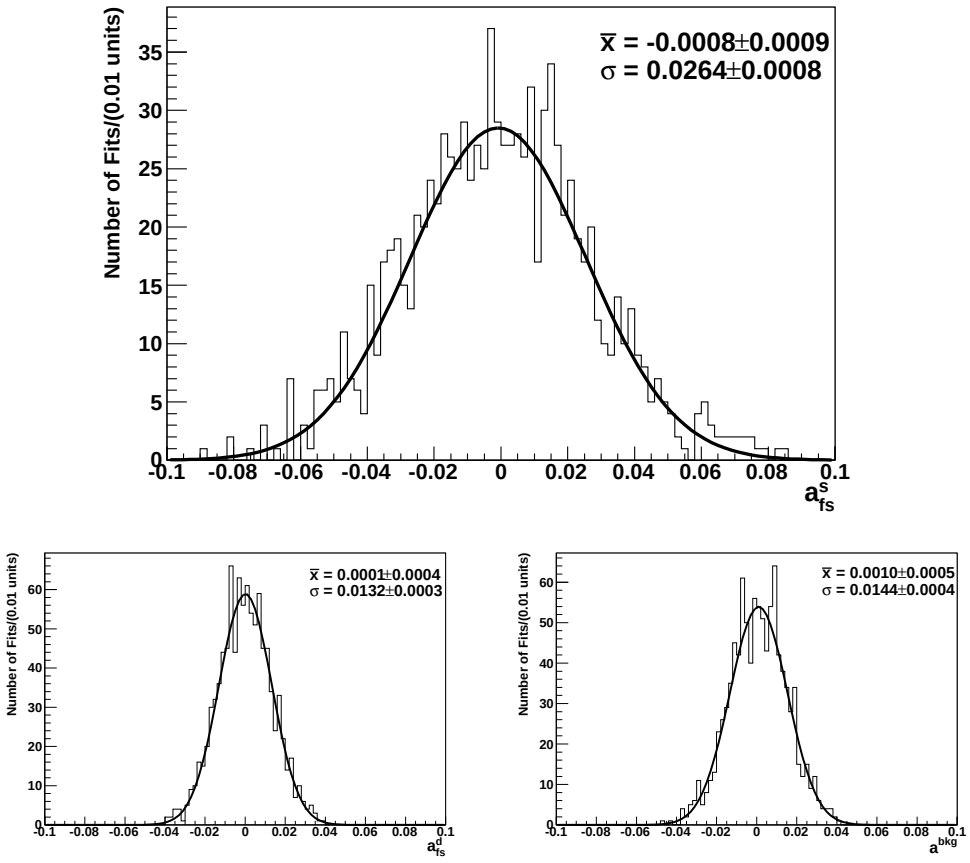

Figure 7.7: An ensemble test of 1000 asymmetry fits showing the distribution of a^s_{fs}, a^d_{fs} and a^{bkg}.

we must turn to MC simulations.

To generate a test sample, we use generator-level MC samples of $B^0_s \to D^-_s \mu^+ \nu_\mu X$ (5500 events) and $B_{d,u,s} \to D^-(K\pi\pi)\mu^+\nu_\mu X$ (1980 events). To make a sample that resembles the sample composition of the data, we use the 1980 D^- events to generate *pdf*s of the lifetime and D^-_s mass[2]. These *pdf*s are used to quickly generate

[2]Recall that the D^- candidates will have a D^-_s mass that is a function of the reflection variable, see Sec. 6.1.1.

a larger sample of 15 500 D^- candidates. To model the combinatoric background, we take 77 458 data events where the charge of the muon is not consistent with a B_s^0 decay (given the $KK\pi$ charges). These three sets of events (B_s^0 MC, D^- MC and wrong sign data) each have an asymmetry inserted that will model one of the three physics asymmetries (a_{fs}^s, a_{fs}^d and a^{bkg}, respectively). The asymmetry is inserted by randomizing the muon charge in each sample in such a way as to introduce the desired asymmetry ($a_{fs}^s = 0.0069$, $a_{fs}^d = 0.0073$ and $a^{bkg} = 0.0088$).

The MC does not include mixing effects, and so we can only insert an integrated asymmetry; the fitting code is modified to accommodate this[3]. We conduct an ensemble of 2000 fits with different random seeds generating an asymmetry. We keep track of the fitted asymmetry and the 'true' asymmetry. The true asymmetry is the literal charge asymmetry that results from randomly choosing the flavour tag. It is a statistical fluctuation of the input asymmetry. Because of the nature of the randomization, there can be a substantial difference between the 'true' asymmetries and the input asymmetry. We also randomize the muon rapidity and toroid polarity to remove the detector asymmetries.

From the results we can verify our ability to correctly determine a small signal asymmetry in a large background environment. Fig. 7.8 shows how the fitted asymmetry value is related to the true charge asymmetry. Because we input an integrated charge asymmetry, the fitted asymmetry must be halved (See Eqn. 1.20). The dependence of the average fitted asymmetry on the true asymmetry is found to be linear ($a_{fs}^s = 0.0016 + 0.765 \cdot a_{fs,true}^s$) with no significant asymmetry bias.

[3]The lifetime-dependent input functions are integrated. This is effectively an integrated asymmetry analysis.

There does, however, appear to be some deviation from a 1 : 1 relationship between the fitted and true asymmetry because the slope is not unity. This could indicate a reduced sensitivity to the asymmetry. However, due to the approximate nature of this test, the deviation should not be taken as problematic; we are interested primarily in the lack of a bias.

This test also shows the relationship between the signal and background asymmetries (a^d_{fs} and a^{bkg}). Fig. 7.9 shows that there is some correlation between a^s_{fs} and a^d_{fs}; however, there is no correlation between a^s_{fs} and a^{bkg}.

7.3 Systematic Uncertainties

The systematic uncertainties on the oscillation amplitude and the signal asymmetry are determined by changing the various input parameters within their respective statistical uncertainties. The variation of the fitted amplitude due to the changes to the inputs is used to determine the total uncertainty on the amplitude at each value of Δm_s. Similarly, the variation in the charge asymmetry due to the changes to the inputs is used to determine the total uncertainty on the signal asymmetry.

7.3.1 Amplitude Scan Systematic Uncertainties

The total uncertainty on the amplitude is determined using the formula [73]

$$\sigma^2_{\mathcal{A},sys} = \sum_i \left(\Delta\mathcal{A}^i + (1-\mathcal{A})\frac{\Delta\sigma^i_{\mathcal{A}}}{\sigma_{\mathcal{A}}} \right)^2, \tag{7.2}$$

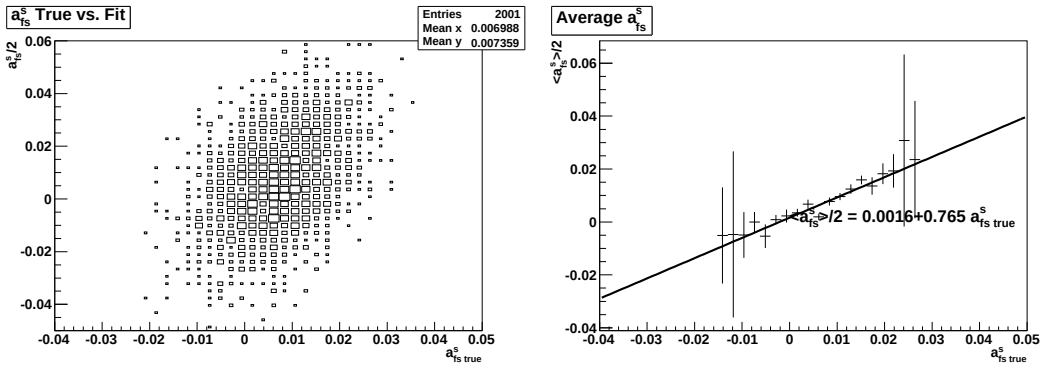

Figure 7.8: The results of an ensemble test with an input asymmetry. The plots show the fitted asymmetry against the true asymmetry (left) and the average fitted asymmetry as a function of the true asymmetry (right).

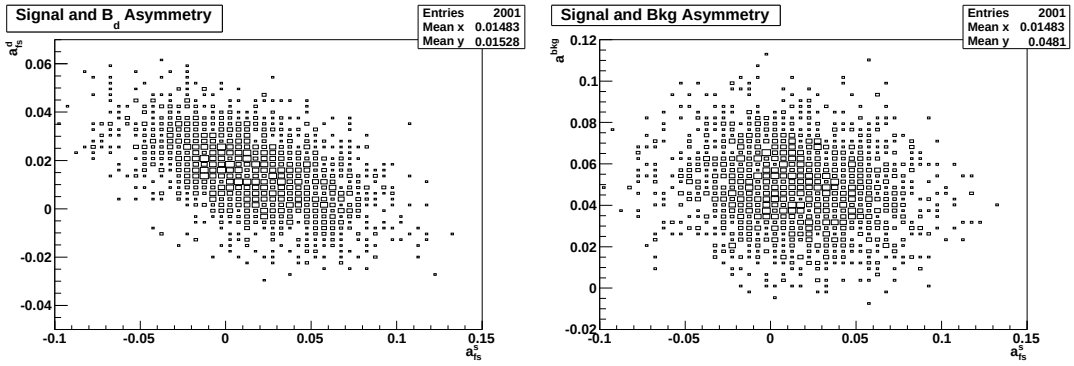

Figure 7.9: The results of an ensemble test with an input asymmetry. The plots show the dependence of the fitted a^s_{fs} on a^d_{fs} (left) and on a^{bkg} (right).

where the index i runs over the input parameters being varied. For each of the considered systematic uncertainties i, we conduct an amplitude scan and determine the shift in both the amplitude ($\Delta\mathcal{A}^i$) and the amplitude uncertainty ($\Delta\sigma^i_{\mathcal{A}}$) as functions of Δm_s. The parameters and their fluctuations are described below.

The primary contribution to the systematic uncertainty is the variation of the dilution calibration curve. The amplitude is refitted using the upper and lower variations of this curve (see Fig. 4.8). Changing the dilution calibration affects the B^0_d oscillations and causes a large shift in the amplitude for small Δm_s (near Δm_d).

The second largest contribution is the uncertainty in the number of D^-_s and D^- candidates determined from the mass fit. We vary $N_{D^-_s}$ down and N_{D^-} up separately in order see the largest variation in the amplitude due to these uncertainties.

The sample composition is recomputed with the most significant branching fractions varied by their statistical uncertainty in the direction to give a smaller signal contribution. We vary the inclusive fraction $Br(B^0_s \to D^-_s \mu^+ \nu_\mu X)$ down, $Br(B^0_s \to D^-_s D^+_s X)$ up, $Br(B^+ \to D^-_s DX)$ up and $Br(B^0 \to D^-_s DX)$ up.

The exclusive branching fractions for the B^0_s decays to excited D^{*-}_s states are not measured (see Sec. 5.3). In place of a measurement, we use the corresponding measured B^0 fractions to infer the fractions in B^0_s mesons assuming the spectator model. The higher excited states include more unreconstructed particles, and hence have wider K factor distributions. Uncertainty in the balance of the exclusive fractions leads to an uncertainty in the fitted amplitude. To account for uncertainty in the exclusive fractions, we propagate the uncertainty on the B^0 fractions into

the computation for the B_s^0 fractions. Using these uncertainties we compute two new sample compositions. The first has a reduced contribution from the dominant exclusive mode ($B_s^0 \to D_s^{*-}\mu^+\nu_\mu X$), and the second has the contribution from this mode increased. In both cases, the total inclusive fraction is kept constant.

The final uncertainty related to the sample composition is due to the uncertainty in the relative fraction of B^0/B^+ in D^- sources. To account for this, the ratio is changed from 80%/20% to 90%/10% and is applied to the lifetime *pdf*s given in Eqn. 6.26 and 6.27.

The K factor distributions are determined from generator-level MC and may be slightly different from those obtained with a full detector simulation. To estimate this contribution to the uncertainty we generate a small MC sample of $B_s^0 \to D_s^-\mu^+\nu_\mu X$ events including the detector simulation. We generate two sets of K factor distributions: one using only the generator information and the other using only the reconstructed information (but for the same MC events). The amplitude difference that results using these two distribution is applied as the systematic uncertainty for the K factors.

The lifetime parameters determined in Sec. 6.3 are each varied by $\pm\sigma$ and included in the total uncertainty. The fitted B_s^0 lifetime values are lower than the world average B_s^0 lifetime ($c\tau_{B_s} = 441$ μm [9], compare to the results in Sec. 6.3). We account for a possible bias by fixing the B_s^0 lifetime to the world average.

The selection efficiency curves are refit with the data points varied up and down by $\pm\sigma$. The new curves are then renormalized to have the same plateau efficiency; an example is shown in Fig. 7.10.

The uncertainty in the muon trigger efficiency is included by fitting the ampli-

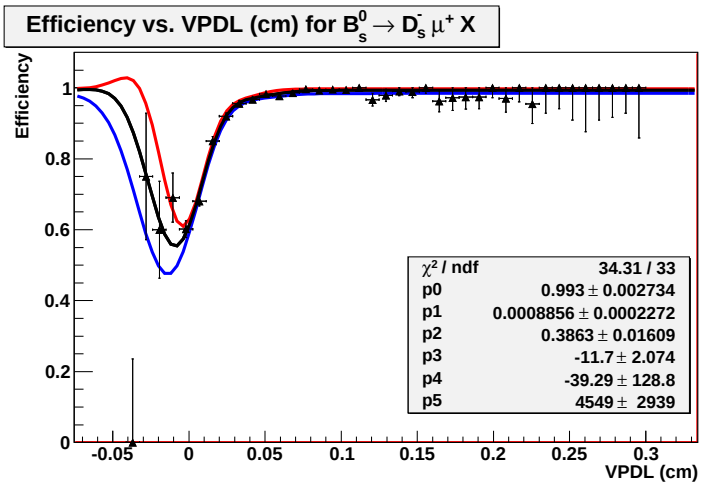

Figure 7.10: An example of selection efficiency curves used to estimate the contribution to the systematic uncertainty. The central curve is the normal efficiency. The upper and lower curves are the $+\sigma$ and $-\sigma$ curves, respectively.

tude with positive and negative variations of the curves shown in Fig. 5.5. This effect is very small. We vary $\Delta\Gamma_s$ within its statistical uncertainty but find the resulting difference in amplitude to be negligible.

The amplitude scans with the total uncertainty are shown in Fig. 7.11. The Δm_s limit, including the systematic uncertainty, is 9.9 ps^{-1} with an expected sensitivity of 14.8 ps^{-1}. Tables showing the magnitude of each contribution to the systematic uncertainty are included in Appendix D.

7.3.2 Asymmetry Systematic Uncertainties

The asymmetry measurement is affected by many of the same systematic uncertainties as the oscillation analysis. These contributions to the uncertainty are estimated in a similar manner as above. Input parameters are varied within their

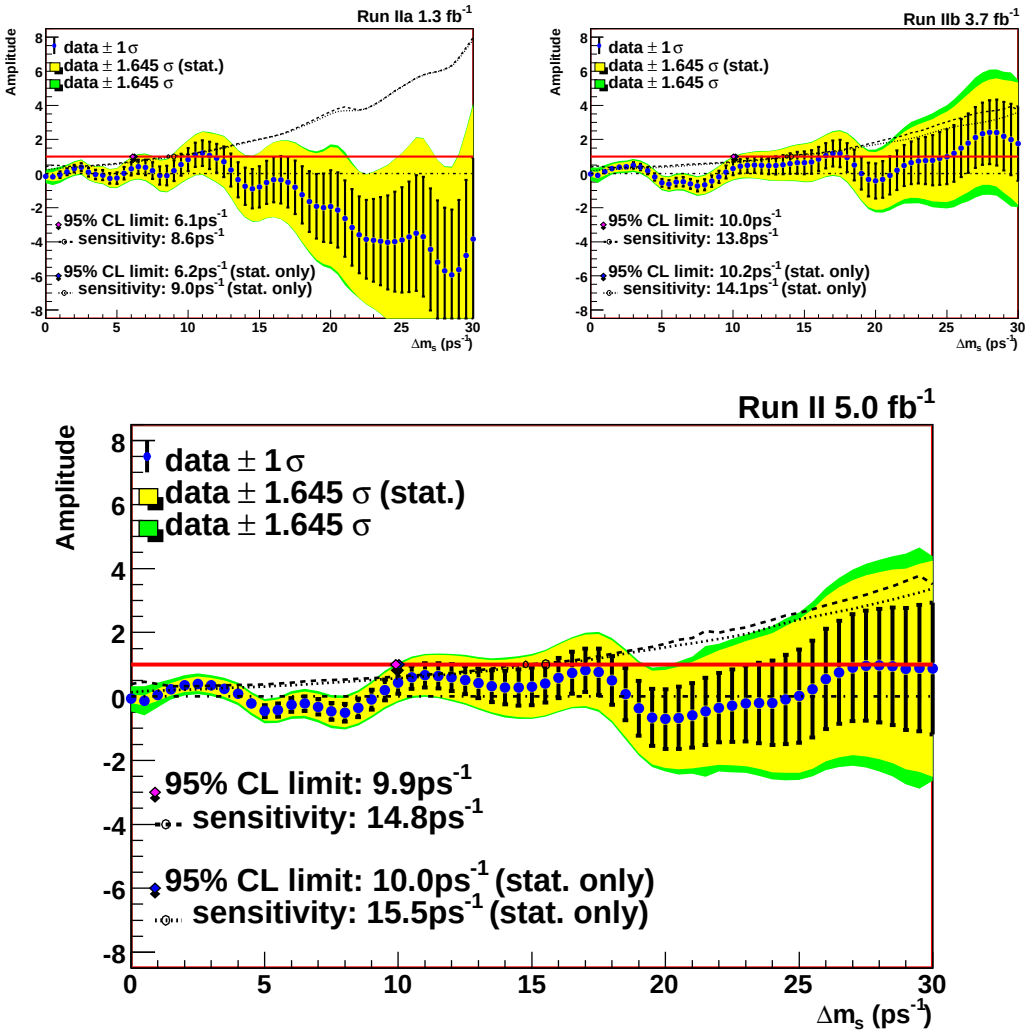

Figure 7.11: B_s^0 amplitude scan results from Run IIa (top left), Run IIb (top right) and the combined scan (bottom). The outer band represents the systematic uncertainty on the fitted amplitude. The limit and sensitivity are recomputed to include the total uncertainty.

statistical uncertainty with the asymmetry being refitted. The positive and negative variations of the asymmetry are added in quadrature to give the final upper and lower systematic uncertainties. The detector asymmetries are the same for all sources and are independent of the lifetime *pdf* and thus are not sensitive to these variations. The correlation coefficients in the physics asymmetries are determined to be $C_{s,d} = -0.266$, $C_{s,bg} = -0.076$ and $C_{d,bg} = -0.280$, where the subscripts s, d and bg denote the asymmetries being correlated. The correlation between a^s_{fs} and a^d_{fs} would be problematic if we were attempting to extract both parameters. We are, however, only interested in a^s_{fs}, hence the correlation is not as important. The correlation is taken into account in determining the statistical uncertainty in the fitted parameters, and is also considered as a systematic uncertainty when combining this measurement with other decay modes (see Chap. 8).

In this analysis, the most significant contribution to the systematic uncertainty is from the number of D_s^- and D^- candidates determined in the mass fit. This is primarily due to the asymmetry related to the B^0_d source of the D^- candidates. Enhancements to the $D^- \to K^+\pi^-\pi^-$ reflection or the combinatoric background will change the fitted signal asymmetry due to the correlations between these parameters. The largest variations that give a reduced sensitivity to the signal asymmetry are tested. The number of signal candidates is reduced by 1σ while simultaneously increasing the number of D^- reflection candidates by 1σ. A second test reduces both the number of signal candidates and the number of D^- reflection candidates. The first test increases the contribution from a^d_{fs} and the second test increases the contribution from a^{bkg}.

The most significant uncertainties related to the sample composition are the

branching fractions $Br(B_s \to \mu^+ D_s^-)$ and $Br(B_s \to D_s^- D_s^+)$. Within their measured uncertainties, $Br(B_s \to \mu^+ D_s^-)$ is varied down and $Br(B_s \to D_s^- D_s^+)$ is varied up. In each case the sample composition is recomputed and used to test for changes to the asymmetry. The relative fractions of $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ and $B_s^0 \to D_s^{*-} \mu^+ \nu_\mu X$ are treated in the same manner as described in Sec. 7.3.1.

Curves for the dilution calibration, the selection efficiency, and the muon trigger turn on were varied up and down in the same manner as is done in the oscillation analysis. Each of the lifetime parameters is varied by its uncertainty as before. The consistency of the K factor distributions in generator and reconstructed MC is treated the same as above. The uncertainty due to the ratio $\Delta\Gamma_s/\Gamma_s = 0.069^{+0.058}_{-0.062}$ is taken from the PDG [9]. Changes to the asymmetry are tested for both upper and lower variations of this ratio.

The contributions to the total systematic uncertainty are listed in Table 7.2. Only the total contribution to the variation of the lifetime parameters is listed. The measured B_s^0 asymmetry with systematic uncertainties is $a_{fs}^s = 0.018 \pm 0.025 \pm 0.002$.

Systematic Uncertainty	$\Delta a^s_{fs} \cdot 10^2$	$\Delta a^d_{fs} \cdot 10^2$	$\Delta a^{bkg} \cdot 10^2$
$\Delta\Gamma_s + \sigma$	0.001	0.004	−0.002
$\Delta\Gamma_s - \sigma$	−0.001	−0.002	0.001
$N_{D_s} - \sigma, N_{D^+} + \sigma$	−0.213	−0.089	0.046
$N_{D_s} - \sigma, N_{D^+} - \sigma$	0.059	0.140	−0.038
$Br(B_s \to \mu D_s) - \sigma$	0.143	0.018	−0.003
$Br(B_s \to D_s D_s) + \sigma$	0.055	−0.008	−0.002
Ratio $D_s/D_s^* + \sigma$	0.024	−0.001	−0.002
Ratio $D_s/D_s^* - \sigma$	−0.006	0.002	0.001
Dilution $+\sigma$	−0.015	−0.034	0.041
Dilution $-\sigma$	−0.006	0.033	−0.024
Selection Eff. $+\sigma$	−0.018	−0.010	0.004
Selection Eff. $-\sigma$	0.018	0.010	−0.004
Trigger Efficiency $+\sigma$	−0.001	−0.001	0.000
Trigger Efficiency $-\sigma$	0.000	0.001	0.000
Lifetime Fit $+\sigma$	0.007	0.098	0.031
Lifetime Fit $-\sigma$	0.057	0.098	0.031
K factor	$\pm$0.037	$\pm$0.012	$\pm$0.001
Total $+\sigma$	0.173	0.201	0.082
Total $-\sigma$	0.227		

Table 7.2: Systematic uncertainties for the asymmetry parameters. Only the combined Run IIa and Run IIb result is computed.

Chapter 8

Conclusions

The weak interaction is a fundamental force of nature and an important component of the SM. The better we understand this interaction, the better equipped we are to search for physics beyond the SM. The work presented in this thesis represents a contribution to this understanding as well as a search for new physics.

A lower limit has been placed on the B_s^0 oscillation frequency,

$$\Delta m_s > 9.9 \text{ ps}^{-1} \text{ at } 95\% \text{ C.L.} \tag{8.1}$$

with sensitivity to oscillations up to 14.8 ps^{-1}. This measurement is consistent with the world average B_s^0 oscillation frequency ($17.77 \pm 0.10 \pm 0.07$ ps^{-1}). This result may be used to constrain the side of the CKM unitarity triangle presented in Sec. 1.2 $\left(\left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right|\right)$. In practice, the CDF measurement of Δm_s [25] dominates the world average and other measurements are not generally included in the unitarity triangle constraint[1]. Nevertheless, an independent confirmation of the oscillation

[1]The CKMFitter group uses the CDF numerical result $\Delta m_s = 17.77 \pm 0.12$ ps^{-1} [19] and the

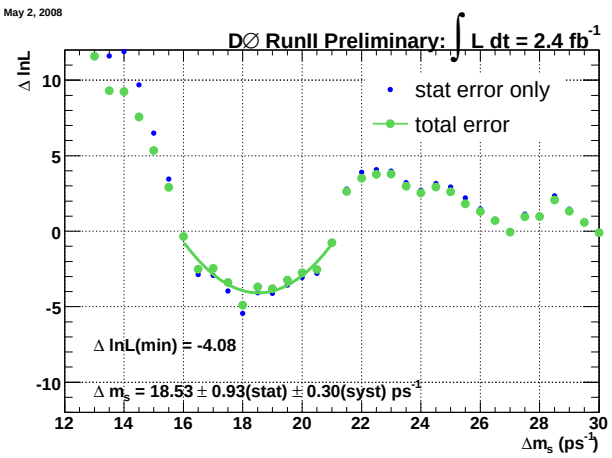

Figure 8.1: Likelihood scan in Δm_s combining several decay modes.

frequency is important.

The measurement presented here can be combined with a variety of other decay modes to make a more significant measurement. These modes include the decays $B_s^0 \to \mu^+ D_s^- (\phi\pi^-)X$, $B_s^0 \to e^+ D_s^- (\phi\pi^-)X$, $B_s^0 \to \mu^+ D_s^- (K_s^0 K^-)$ and $B_s^0 \to \pi^+ D_s^- (\phi\pi^-)$. These measurements are made independently and are documented elsewhere [74–76]. The combination of these decay modes has been completed using 2.4 fb^{-1} of D0 data [77]. The work presented in this thesis is included in the combination using the smaller data set, but lacking some of the refinements made to the analysis in the intervening period. The details of that analysis are documented elsewhere [78]. The likelihood plot of the combined measurement is shown in Fig. 8.1. This combination yields a 2.9σ significance measurement of Δm_s:

$$\Delta m_s = 18.53 \pm 0.93(\text{stat}) \pm 0.30(\text{syst}) \text{ ps}^{-1}. \tag{8.2}$$

UTfit collaboration uses the CDF likelihood [35].

Asymmetries $\times 10^3$	$\mu^+\phi\pi^-$	$\mu^+K^{*0}K^-$	Combined
a^s_{fs}	-7.0 ± 9.9	17.8 ± 24.6	-1.7 ± 9.1
a^d_{fs}	-21.4 ± 36.3	49.9 ± 19.4	40.5 ± 16.5
a^{bkg}	-2.2 ± 10.6	0.2 ± 13.9	-3.1 ± 8.3
A_{fb}	-1.8 ± 1.5	-2.1 ± 1.5	-1.9 ± 1.1
A_{det}	3.2 ± 1.5	2.8 ± 1.5	3.1 ± 1.1
$A_{q_\mu\beta}$	4.3 ± 1.5	1.9 ± 1.5	3.1 ± 1.1
$A_{\beta\gamma_\mu}$	1.1 ± 1.5	2.4 ± 1.5	0.6 ± 1.1
A_{ro}	-36.7 ± 1.5	-30.1 ± 1.5	-33.3 ± 1.1

Table 8.1: Results of the combined fit to the B^0_s asymmetry [79].

A combination measurement using the 5 fb^{-1} data sample is in progress but has not been completed at the time of writing.

A similar combined measurement can be made for the B^0_s asymmetry. The B^0_s asymmetry has been measured in the $B^0_s \to \mu^+ D^-_s(\phi\pi)$ decay mode, documented elsewhere [80]. The combination, documented elsewhere [79, 81], is conducted by joining the likelihood sums of the two analyses (see Eqn. 2.4). The results are presented in Table 8.1. Systematic uncertainties similar to those discussed in Sec. 7.3.2 are studied in the combined analysis. In addition, correlations between a^s_{fs} and a^d_{fs} were taken into account. The combined result gives a measured B^0_s asymmetry:

$$a^s_{fs} = -0.0017 \pm 0.0091(\text{stat})^{+0.0013}_{-0.0023}(\text{syst}). \tag{8.3}$$

This measurement is competitive with the world average value, $a^s_{fs} = -0.0037 \pm 0.0094$ [39]. This average includes measurements of the asymmetry in the dimuon channel (see Sec. 1.2.2). These measurements require a value of a^d_{fs} be input

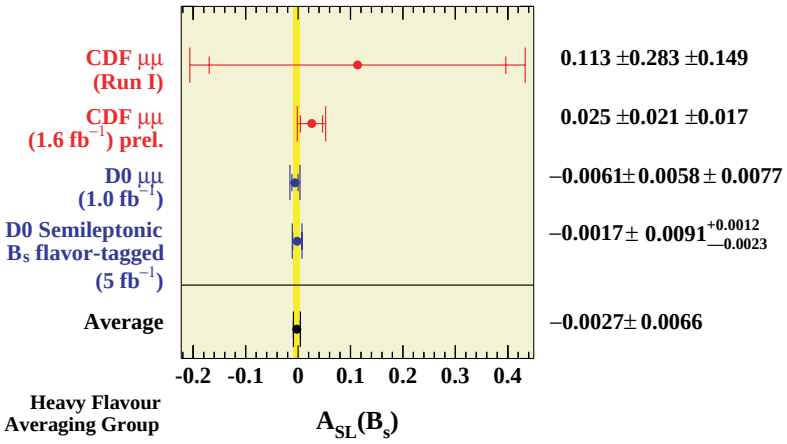

Figure 8.2: The new world average a^s_{fs} value, including the result presented in this work.

to separate the B^0_s asymmetry. HFAG uses the current world average, $a^d_{fs} = -0.0047 \pm 0.0046$ [29]. The new world average B^0_s asymmetry value is $a^s_{fs} = -0.0027 \pm 0.0066$ [29] (see Fig. 8.2) and is consistent with the SM expectation, $a^{SM,s}_{fs} = (2.06 \pm 0.57) \cdot 10^{-5}$.

As discussed in Sec. 1.4, the B^0_s asymmetry is a good probe for new physics. The relationship in Eqn. 1.22 relates the asymmetry to the weak $\mathcal{CP}$ phase ($\phi_s = \phi^{SM}_s + \phi^{\Delta}_s$). The phase can be measured directly in $B^0_s \to J/\psi\phi$ decays (we denote the measured value $\phi^{J/\psi\phi}_s$). These measurements were shown previously in Fig. 1.8. D0 has updated these measurements to include the constraints on ϕ_s from the new world average a^s_{fs} value. The constraint imposed by a^s_{fs} is shown in Fig. 8.3. This combination includes only D0 results.

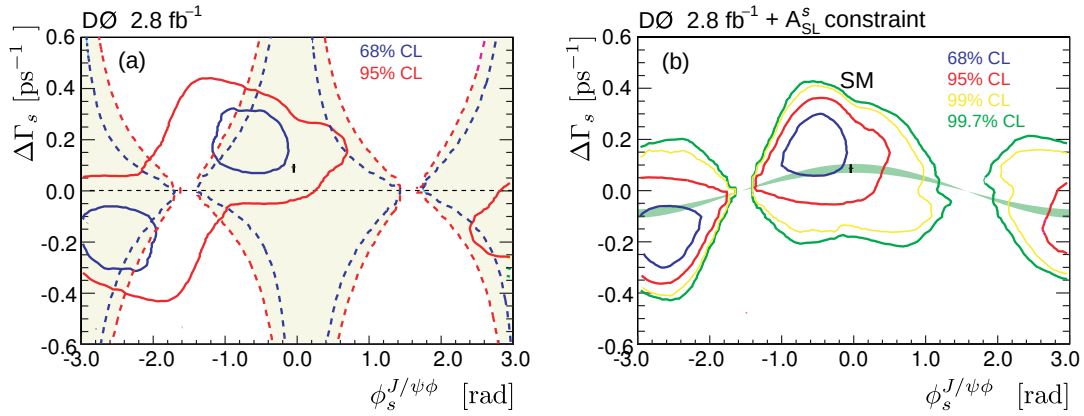

Figure 8.3: The D0 direct measurements of $\phi_s^{J/\psi\phi}$ (left plot, solid contours) and the constraint imposed by the a_{fs}^s world average (left plot, dashed contours). The unshaded region is excluded to 68% CL. The combination is shown at right [28].

8.1 Contributions

High energy particle physics is a very collaborative field of study. Much of the work presented in this thesis is borrowed from other sources, or is developed by collaborators. Furthermore, as in any research, a large portion of the work is 'logistical' in nature and goes undocumented. Similarly, promising leads sometimes turn into dead ends and are likewise undocumented. When I began work on the B_s^0 oscillation frequency measurement, I inherited a mature analysis framework. Selection criteria had already been developed, fitting code was already written, and inputs already studied. There was, however, significant room for improvement. In this section I outline my personal contributions to this analysis, including some aspects not described in the previous chapters or not included in the final analysis.

The $D_s^- \to K^{*0}K^-$ decay considered in this analysis has a large background, making the fit to the B_s^0 lifetime distribution challenging (see Sec. 6.2). In par-

ticular, there is no significant side band to the D_s^- mass distribution that does not contain some contribution from the physics backgrounds (see Figs. 6.2 and 6.3). This mass is our primary means of distinguishing signal from background; overlapping signal and background in the invariant mass distribution leads to a poor separation of the two when modelling the B_s^0 candidate lifetime. This is also true in describing the behaviour of the nuisance parameters (described in Chap. 5) in the combinatoric background. These problems make modelling the shape of the lifetime distribution of the combinatoric background very difficult.

In considering the issues described above, I made a significant effort to improve the quality of the fit to the B_s^0 lifetime distribution by modifying the functional form of the combinatoric background as well as the manner in which the fit is performed. The quality of the final fit has improved significantly due to these efforts on my part. The final function (Eqns. 6.28-6.30) is not so different from when I began these studies, however, the fit is conducted in a more consistent manner. For example, previously, the fit could only be done in stages. First a fit was performed in the signal-free side band to fix the parameters in the combinatoric background. The fit was then repeated in the signal region with these parameters fixed. The changes I made to the implementation of the function now allow the fit to be done in the entire mass region at once. In this manner, the treatment of the background is more consistent.

The improvement to the fit described above is due in part to an improved description of the signal contributions. This, in turn, is due to better descriptions of the various input parameters and distributions. I have improved the treatment of the composition of the D_s^- signal (see Sec. 5.3), the missing energy (see Sec. 5.4)

and the selection efficiency (see Sec. 5.5) to better model the data. The K factor and sample composition are now binned by μD_s mass, and are generated from a significantly increased sample of simulated data. New sources have been added to the sample composition and have associated K factor distributions in order to better model the D_s^- signal. The modelling of the selection efficiency has been improved by using separate MC simulations for each D_s^- and D^- source, rather than using joint MC to model several sources at once. Furthermore, the efficiency in the combinatoric background was improved by being fitted directly to the data. Previously this efficiency was simply assumed to behave in the same manner as the $D^- \to K^+\pi^-\pi^-$ source; direct fits were attempted but were non-convergent. The improvements to the fitting described above have increased the stability of the convergence and have allowed this efficiency to be fitted directly. This improvement provides a much better description of the short-lived backgrounds and improves the fit quality in this high statistics region.

In addition to the improvements to the analysis process described above, I have extended this analysis to include the Run IIb data. As described in Chap. 3, the Run IIb detector includes an additional layer in the silicon detector. This change, though it may seem minor, makes a substantial difference to the tracking performance. This improvement primarily impacts the VPDL uncertainty, but can have an indirect influence on the tagging, selection variable (y), D_s^- sample composition and the selection efficiency. Modelling these changes required the generation of new samples of simulated data for each decay channel (see Sec. 5.3) using the new detector configuration. I generated these samples and repeated the Run IIa studies for use with the Run IIb data sample.

I have also made an effort to study other variables correlated to the K factor in the hope of obtaining a more accurate prediction of the missing energy. I examined the possibility of binning by the transverse momentum component in the direction of the B_s^0 meson. This binning yielded a similar improvement as binning by mass, but was found to be directly correlated to the mass binning and hence gave no overall improvement. Other attempts at improving the analysis include the study of a neural network to replace the selection criteria discussed in Sec. 4.1.1. This study did not yield any improvement over the standard selection criteria. This is due to the difficulty in modelling inclusive $b\bar{b}$ background, which has not been simulated (at D0). The lack of a reasonable simulation of include $b\bar{b}$ background lead to difficulties in training the network. In addition, I have made an attempt at an implementation of the same-side flavour tagging algorithms (see Sec. 4.2). SST is difficult to implement in a B_s^0 analysis (particularily in this decay mode) due to contributions from B^+ mesons; this has been described in Sec. 4.2. By modifying the calibration function applied to the B^+ contributions it was hoped that these problems could be addressed. However, it is not possible to isolate the B^+ contributions, hence this method was not successful.

The work described above was in the process of developing the B_s^0 oscillation analysis. One of the most significant contributions I made was to adapt the B_s^0 oscillation analysis to include a $\mathcal{CP}$ asymmetry measurement as well. This effort required completely rewriting much of the analysis software to function in a new framework that could easily be combined with other decay modes. This new framework required extensive testing to ensure the previous oscillation frequency measurement could be reproduced. Further testing was required to ensure the new

framework was capable of correctly measuring an asymmetry. This is particularly important in the $D_s^- \to K^{*0}K^-$ decay mode, where the large backgrounds make an asymmetry measurement very difficult. I conducted extensive tests using simulated data to ensure that the correlation between the B^0 and B_s^0 asymmetries would not bias the measurement. As a side effect of this new framework, the combined measurement of the B_s^0 oscillation frequency benefits from the new code framework, which allows a joint fit in primary decay modes ($\mu\phi\pi$ and $\mu K^* K$). This joint fit was not possible in the previous framework. The joint fit is also applied in the $\mathcal{CP}$ asymmetry combination. Here the joint fit is particularly important because there are several correlated fit parameters that are only properly considered in a simultaneous fit.

8.2 Final words

The oscillation measurement presented in this thesis represents the first analysis in the $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ ($D_s^- \to K^{*0}K^-$) decay mode where a Δm_s measurement has been possible, even if of low significance. This measurement will be an important contribution to the combined 5.0 fb^{-1} B_s^0 oscillation measurement currently in progress. Furthermore, the development of the oscillation analysis in this decay mode is an important step towards developing the asymmetry measurement.

This work represents the first measurement of the flavour-specific B_s^0 charge asymmetry using the $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ ($D_s^- \to K^{*0}K^-$) decay mode, and the first[2] time-dependent measurement in any decay mode. The measurement in this

[2]This analysis has been performed in parallel with the $B_s^0 \to D_s^- \mu^+ \nu_\mu X$ ($D_s^- \to \phi\pi^-$) decay mode.

mode contributes to the combined measurement by reducing the total statistical uncertainty by $\sim 10\%$.

Appendices

Appendix A

The Higgs Field

In a gauge field theory (such as the SM), the vector bosons arise naturally from gauge symmetries that are related to physically conserved quantities. In electromagnetism, for instance, if we require the theory be invariant under a *local* phase transformation we get a massless photon and charge conservation out of the free Lagrangian for the fermions. However, the fermions, under this scheme, are also massless. A similar procedure can be applied to get the weak gauge bosons, but they too will be massless. The simplest way to make these particles massive is to add a complex scalar field (spin 0) to the theory. This new field will have couplings to the fermions and the weak gauge bosons, but not the neutrinos, the photon or the gluons (we want those massless). We call this new field the Higgs field. The form of the self couplings of the field will determine the potential for the field and hence its ground state. The simplest form that we can imagine for this field is a quartic potential; $V(\phi^\dagger\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$.

The field must have a nonzero ground state in order to give masses to the

free particles, so we require that $\lambda > 0$ and $\mu^2 < 0$. The nonzero ground state corresponds to the vacuum obtaining a nonzero energy. The energetic vacuum can then interact with the fermions and gauge fields. These couplings will give mass terms in the Lagrangian (Eqn. A.1):

$$\mathcal{L} = \ldots - g\bar{\psi}\langle\phi\rangle\psi + g' A^{\dagger}_{\mu}\langle\phi\rangle A^{\mu}, \tag{A.1}$$

where ψ and A_μ are fermion and gauge fields, respectively, g and g' are coupling constants and $\langle\phi\rangle$ is the vacuum expectation of the Higgs field. These terms correspond to masses for the fields: $M_\psi = g\langle\phi\rangle$ and $M_{A_\mu} = g'\langle\phi\rangle$.

In the general example above, we did not identify ψ or A_μ with any particular field. Now, consider the left-handed quark doublet ($q^i_L = (u^i_L, d^i_L)$) and the right-handed quark singlets (d^i_R,u^i_R), where the generation index $i = 1, 2, 3$. In this case we need a doublet of two complex scalar fields and its associated anti-doublet and the vacuum expectation for both:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix},\ \tilde{\phi} = \begin{pmatrix} \bar{\phi}^0 \\ -\phi^- \end{pmatrix},\ \phi_{vac} = \begin{pmatrix} 0 \\ \langle\phi\rangle \end{pmatrix},\ \tilde{\phi}_{vac} = \begin{pmatrix} \langle\phi\rangle \\ 0 \end{pmatrix}, \tag{A.2}$$

and if we make the assumption that the quarks can mix between generations[1],

[1]Certainly this is reasonable, otherwise the Kaons (for example) would be stable. This contrasts with the leptons where no mixing is allowed (in the SM), for example there is no $\mu^- \to \nu_e W^-$ vertex.

then the Lagrangian becomes

$$\begin{aligned}\mathcal{L}^{mass} &= -D_{ij}(\bar{q}_L^i\phi)d_R^j - U_{ij}(\bar{q}_L^i\tilde{\phi})u_R^j, \\ \mathcal{L}_{vac}^{mass} &= -D_{ij}\bar{d}_L^i\langle\phi\rangle d_R^j - U_{ij}\bar{u}_L^i\langle\phi\rangle u_R^j,\end{aligned} \tag{A.3}$$

where D_{ij} and U_{ij} are the Higgs coupling constants and form 3×3 matrices $\mathbf{D}$ and $\mathbf{U}$. In order to give definite masses to the quarks, $\mathbf{D}$ and $\mathbf{U}$ must be made diagonal to remove the cross generation couplings. This is equivalent to making a rotation in flavour space and defining new linear combinations of the three generations which have definite mass and do not mix. So we define four unitary matrices ($\mathbf{\Lambda}_L^U$, $\mathbf{\Lambda}_R^U$, $\mathbf{\Lambda}_L^D$ and $\mathbf{\Lambda}_R^D$) such that $\hat{\mathbf{D}}$ and $\hat{\mathbf{U}}$ are diagonal,

$$\mathbf{\Lambda}_L^D\mathbf{D}\mathbf{\Lambda}_R^{\dagger D} = \hat{\mathbf{D}},\ \mathbf{\Lambda}_L^U\mathbf{U}\mathbf{\Lambda}_R^{\dagger U} = \hat{\mathbf{U}}. \tag{A.4}$$

We insert these matrices into the Lagrangian and regroup terms such that the quark vectors are rotated and the coupling matrix is diagonal. The result for the down type quarks is shown in Eqn. A.5:

$$\begin{aligned}\mathcal{L}_{vac}^d &= -\langle\phi\rangle\bar{\mathbf{d}}_L\mathbf{D}\mathbf{d}_R \\ &= -\langle\phi\rangle\bar{\mathbf{d}}_L(\mathbf{\Lambda}_L^{\dagger D}\mathbf{\Lambda}_L^D)\mathbf{D}(\mathbf{\Lambda}_R^{\dagger D}\mathbf{\Lambda}_R^D)\mathbf{d}_R \\ &= -\langle\phi\rangle(\bar{\mathbf{d}}_L\mathbf{\Lambda}_L^{\dagger D})(\mathbf{\Lambda}_L^D\mathbf{D}\mathbf{\Lambda}_R^{\dagger D})(\mathbf{\Lambda}_R^D\mathbf{d}_R) \\ &= -\langle\phi\rangle\bar{\mathbf{d}}'_L\hat{\mathbf{D}}\mathbf{d}'_R,\end{aligned} \tag{A.5}$$

where $\mathbf{d}'_{L/R}$ are the flavour rotated quarks. Now we use these same rotated states

in the weak coupling between the quarks and the $W^{\pm}$ bosons. The result for the W^{+} boson is shown in Eqn. A.6:

$$\begin{aligned}
\mathcal{L}_{Q-W^{+}} &= \bar{\mathbf{u}}_L \gamma^{\mu} W_{\mu}^{+} \mathbf{d}_L = \bar{\mathbf{u}}_L (\mathbf{\Lambda}_L^{\dagger U} \mathbf{\Lambda}_L^{U}) \gamma^{\mu} W_{\mu}^{+} (\mathbf{\Lambda}_L^{\dagger D} \mathbf{\Lambda}_L^{D}) \mathbf{d}_L \\
&= (\bar{\mathbf{u}}_L \mathbf{\Lambda}_L^{\dagger U}) \mathbf{\Lambda}_L^{U} \gamma^{\mu} W_{\mu}^{+} \mathbf{\Lambda}_L^{\dagger D} (\mathbf{\Lambda}_L^{D} \mathbf{d}_L) \\
&= \bar{\mathbf{u}}'_L \underbrace{\mathbf{\Lambda}_L^{U} \mathbf{\Lambda}_L^{\dagger D}} \gamma^{\mu} W_{\mu}^{+} \mathbf{d}'_L
\end{aligned} \tag{A.6}$$

The combination $\mathbf{\Lambda}_L^{U} \mathbf{\Lambda}_L^{\dagger D}$ describes the coupling strength between the various quarks and the $W^{\pm}$ bosons and we identify this as the CKM matrix [12].

While this procedure may appear somewhat more natural than plugging in a Cabbibo angle, it is not terribly different. We were required to make the assumption that the quarks mix without any real justification. The quark mixings are an empirical fact without any theoretical motivation (in the SM). By contrast, there is no such mixing in the leptons[2], for example there is no $\mu^{-} \rightarrow Z^{0} e^{-}$ process. The fact that we get a mixing matrix if we assume mixing is hardly surprising, but it is not explained by the SM.

[2]Neutrinos do mix, but this is not a SM phenomenon.

Appendix B

Branching Fractions

Branching fractions used in determining the sample composition (see Sec.5.3).

- Br($B_s^0 \to D_s^- \mu^+ \nu_\mu X$)= $(7.9 \pm 2.4)\%$, The total fraction is taken from the PDG, relative fractions of the exclusive decays are taken from EvtGen.
  - Br($B_s^0 \to D_s^- \mu^+ \nu_\mu X$)= 2.0%
  - Br($B_s^0 \to D_s^{*-} \mu^+ \nu_\mu X$)= 5.3%
  - Br($B_s^0 \to D_{s0}^{*-} \mu^+ \nu_\mu X$)= 0.19%
  - Br($B_s^0 \to D_{s1}^{*-} \mu^+ \nu_\mu X$)= 0.35%
- Br($B_s^0 \to D_s^- \tau^+ \nu_\tau X$)= 2.9%, from EvtGen
- Br($\tau \to \mu \nu_\tau \nu_\mu$)= $(17.36 \pm 0.06)\%$, from PDG
- Br($B_s^0 \to D_s^- D_s^+ X$)= $(4.6 \pm 2.2)\%$, from HFAG
- Br($B_s^0 \to D_s^- DX$)= 8.0%, from EvtGen

- $\mathrm{Br}(B_s^0 \to D_s^+ DX) = 6.44\%$, from EvtGen
- $\mathrm{Br}(B^+ \to D_s^- DX) = 7.9^{+1.4}_{-1.3}\%$, from PDG
- $\mathrm{Br}(B^0 \to D_s^- DX) = 10.3^{+2.1}_{-1.8}\%$, from PDG
- $\mathrm{Br}(D_s^{*0} \to D_s^- X) = 100\%$, from PDG
- $\mathrm{Br}(D_{s0}^{*0} \to D_s^- X) = 100\%$, from PDG
- $\mathrm{Br}(D_{s1}^{*0} \to D_s^- X) = 100\%$, from PDG
- $\mathrm{Br}(D_s^- \to \mu^- X) = (6.3 \pm 0.8)\%$, from PDG
- $\mathrm{Br}(D^0 \to \mu^+ X) = (6.6 \pm 0.6)\%$, from PDG
- $\mathrm{Br}(D^+ \to \mu^+ X) = (16.1 \pm 0.4)\%$, from PDG
- $\mathrm{Br}(\bar{b} \to B^0) = (39.7 \pm 1.0)\%$, from PDG
- $\mathrm{Br}(\bar{b} \to B^+) = (39.7 \pm 1.0)\%$, from PDG
- $\mathrm{Br}(\bar{b} \to B_s^0) = (10.7 \pm 1.1)\%$, from PDG

Appendix C

List of Acronyms and Technical Terminology

- Baryon - A particle composed of three quarks.
- CAL - Abreviation of Calorimeter. A detector that measures energy deposition, see Sec. 3.2.3.
- CFT - Central Fiber Tracker. A scintillating fiber detector, see Sec. 3.2.2.
- CKM - Short for 'Cabibbo-Kobayashi-Maskawa.' Refers to the weak quark coupling matrix (CKM matrix). The term CKM is commonly used to refer to weak phenomenology, for example a process may be said to be 'CKM suppressed,' indicating that it is suppressed by the CKM matrix element involved in the interaction.
- $\mathcal{CP}$- Charge-Parity operator/symmetry, see Sec. 1.1.1.

- CPS - Central PreShower Detector. A scintillator based detector in front of the central calorimeter, see Sec. 3.2.2.
- Decay Length - The distance a particle travels before decaying. Determined from the separation of the production and decay vertices (see VPDL). The decay length is related to the lifetime of the particle.
- Flavour Tagging - The process of determining the production flavour of a B hadron ($b/\bar{b}$).
- FPD - Forward Proton Detector.
- FPS - Forward PreShower Detector. A scintillator based detector in front of the end cap calorimeters, see Sec. 3.2.2.
- Hadron - Any particle composed quarks.
- Impact Parameter - The closest approach of a particle (or track) to a vertex. Unless stated otherwise, the primary vertex is used.
- L1CAL/L2CAL - Level 1 and level 2 calorimeter trigger.
- L1CTT/L2CTT - Level 1 and level 2 central track trigger. Based on tracks in the CFT, CPS and FPS.
- L1FPD - Level 1 forward proton detector trigger.
- L1MUO/L2MUO - Level 1 and level 2 muon trigger.
- L2PS - Level 2 preshower trigger.

- L2STT - Level 2 silicon track trigger. Based on tracks in the SMT.
- Lifetime - The time interval in which a particle lives before decaying, determined from the decay length. Typically it is assumed that the particle is travelling at approximately the speed of light, hence the lifetime can be expressed as a distance ct.
- MUO - Short for Muon System, see Sec. 3.2.4.
- Meson - A particle composed of a quark and antiquark.
- NLO - Next-to-leading order.
- OST - Opposite Side Tagging. A flavour tagging technique using the opposite side B hadron decay, see Sec. 4.2.
- *pdf* - Probability density function.
- PV - Primary Vertex. The proton-antiproton collision point.
- QCD - Quantum Chromodynamics. The quantum field theory of the strong interaction.
- QED - Quantum Electrodynamics. The quantum field theory of the electromagnetic interaction.
- Run IIa/IIb - Refers to two separate data taking periods at D0, see Chap. 3.
- SF - Scale Factor. A scaling parameter that tunes the VPDL uncertainty, see Sec. 5.2.

- SM - The Standard Model of particle physics, see Sec. 1.1.
- SST - Same Side Tagging. A flavour tagging technique using light hadrons from the reconstruction side of the B hadron decay, see Sec. 4.2.
- Tagging - See Flavour Tagging.
- TFW - The Trigger Framework, see Sec. 3.2.5.
- VPDL - Visible Proper Decay Length, see Sec. 6.2.

Appendix D

Amplitude Systematic Uncertainty Table

Table D.1: Amplitude scan systematic uncertainties in Run IIa.

Osc. frequency		0 ps^{-1}	3 ps^{-1}	6 ps^{-1}	9 ps^{-1}	12 ps^{-1}	15 ps^{-1}	18 ps^{-1}	21 ps^{-1}
$\mathcal{A}$		-0.167206	+0.077029	+0.276141	+0.166829	+0.899565	-0.780893	-1.246470	-2.638330
σ		+0.126979	+0.254359	+0.392457	+0.608489	+0.845564	+1.212110	+1.679230	+2.242430
$c\tau_{B_s} = 441\mu m$	$\Delta\mathcal{A}$	-0.006466	-0.004052	+0.010835	+0.011648	+0.026543	-0.015426	-0.001223	-0.046390
	$\Delta\sigma$	-0.000102	+0.002127	+0.005190	+0.009723	+0.014671	+0.023020	+0.033630	+0.042490
$c\tau_{B_s} \pm \sigma$	$\Delta\mathcal{A}$	-0.004763	+0.003448	-0.008416	-0.008747	-0.019450	-0.011534	+0.003304	+0.035260
	$\Delta\sigma$	+0.000086	+0.001544	-0.003786	-0.007096	-0.010713	-0.016785	+0.024480	-0.033190
$Fr^{Sig}_{peak} \pm \sigma$	$\Delta\mathcal{A}$	-0.008735	-0.011647	+0.020567	+0.024702	+0.049470	-0.027393	-0.021067	-0.069050
	$\Delta\sigma$	-0.000318	+0.002925	+0.006762	+0.012815	+0.019392	+0.031143	+0.046660	+0.066270
$sf_{bkg} \pm \sigma$	$\Delta\mathcal{A}$	-0.000050	-0.000169	-0.000359	-0.000434	-0.000445	+0.000066	+0.000201	+0.001090
	$\Delta\sigma$	-0.000005	-0.000014	-0.000027	-0.000051	-0.000076	+0.000149	+0.000240	+0.000400
$Fr_{prompt} \pm \sigma$	$\Delta\mathcal{A}$	+0.001198	-0.002281	-0.003005	-0.003505	-0.006933	-0.001424	+0.003139	-0.011560
	$\Delta\sigma$	-0.000076	-0.000004	-0.000120	-0.000552	-0.001119	-0.002648	+0.004660	-0.008920
$c\tau_{long} \pm \sigma$	$\Delta\mathcal{A}$	+0.001616	-0.000248	-0.001145	-0.000247	-0.000877	+0.002034	+0.003950	-0.009040
	$\Delta\sigma$	-0.000008	+0.000216	+0.000457	+0.000798	+0.001106	+0.001748	+0.002660	+0.003000
$Fr_{tsens} \pm \sigma$	$\Delta\mathcal{A}$	+0.015320	+0.000943	+0.000368	+0.000684	-0.000173	-0.000782	-0.002295	-0.003050
	$\Delta\sigma$	-0.000027	-0.000035	-0.000043	-0.000044	-0.000023	-0.000012	+0.000110	+0.000350
$Fr_{osc} - \sigma$	$\Delta\mathcal{A}$	+0.003666	-0.006905	+0.008843	+0.012941	+0.028452	-0.013060	+0.001975	-0.042220
	$\Delta\sigma$	-0.000091	+0.002205	+0.005299	+0.009910	+0.014888	+0.023283	+0.033570	+0.043110
$c\tau_{bkg} \pm \sigma$	$\Delta\mathcal{A}$	-0.001002	+0.001208	+0.000767	-0.000624	+0.005103	-0.001911	-0.003391	-0.014240
	$\Delta\sigma$	-0.000067	-0.000199	-0.000748	+0.001625	+0.002507	+0.004175	+0.006150	+0.008290
$Fr_{Long} \pm \sigma$	$\Delta\mathcal{A}$	+0.014398	-0.003553	-0.002067	+0.001699	+0.005455	+0.002124	+0.001924	+0.003050
	$\Delta\sigma$	-0.000010	-0.000323	+0.000728	+0.001368	+0.001911	+0.002944	+0.003700	+0.005760
$N_{D^+Refl} + \sigma$	$\Delta\mathcal{A}$	-0.066727	-0.002954	+0.012224	+0.026471	+0.063002	-0.032022	+0.077844	+0.139530
	$\Delta\sigma$	+0.003516	+0.006829	+0.009579	+0.015228	+0.025226	+0.036465	+0.065320	+0.159850
$N_{D_s} - \sigma$	$\Delta\mathcal{A}$	-0.098073	+0.007705	+0.039566	+0.057092	+0.135732	-0.033187	+0.043143	+0.020650
	$\Delta\sigma$	+0.010127	+0.019382	+0.028798	+0.045151	+0.065436	+0.096797	+0.149560	+0.268490
Fr $D_s/D_s^* \pm \sigma$	$\Delta\mathcal{A}$	+0.000046	+0.002498	-0.003837	-0.009111	+0.029560	-0.014167	+0.034240	+0.064980
	$\Delta\sigma$	+0.000078	-0.000511	-0.001334	-0.002704	-0.003814	-0.005328	-0.006430	-0.003850
Trigger Turn On	$\Delta\mathcal{A}$	-0.000100	+0.000085	-0.000427	-0.000633	-0.001138	-0.001527	+0.002471	+0.004110
	$\Delta\sigma$	-0.000030	-0.000082	-0.000160	-0.000256	-0.000356	-0.000488	-0.000590	-0.000530
Br($B_s \to D_s\mu X$)=5.5%	$\Delta\mathcal{A}$	-0.030141	-0.006566	+0.016780	+0.018366	+0.057126	-0.008007	+0.053706	+0.065100
	$\Delta\sigma$	+0.003688	+0.006480	+0.008554	+0.012958	+0.020089	+0.029758	+0.056830	+0.140470
Br($B_s \to D_s D_s$)=6.8%	$\Delta\mathcal{A}$	-0.000952	-0.001285	+0.000932	+0.001729	+0.005827	+0.002209	+0.004776	-0.003400
	$\Delta\sigma$	+0.000764	+0.001303	+0.001734	+0.002408	+0.003000	+0.004078	+0.005270	+0.006060
Br($B^+ \to D_s X$)=9.3%	$\Delta\mathcal{A}$	+0.003850	-0.001868	-0.000373	+0.001306	+0.002677	+0.001690	+0.004073	+0.000100
	$\Delta\sigma$	+0.000392	+0.000619	+0.000749	+0.000989	+0.001123	+0.001509	+0.001860	+0.002030
Br($B^0 \to D_s X$)=12.4%	$\Delta\mathcal{A}$	+0.004264	-0.003719	-0.001090	+0.002261	+0.005351	+0.003209	+0.007818	-0.000640
	$\Delta\sigma$	+0.000818	+0.001329	+0.001630	+0.002184	+0.002531	+0.003395	+0.004250	+0.004820
Dilution	$\Delta\mathcal{A}$	-0.146061	-0.027956	+0.012874	-0.017451	+0.090123	+0.098889	+0.239657	+0.423200
	$\Delta\sigma$	+0.009745	+0.020119	+0.030452	+0.046154	+0.068308	+0.099237	+0.156100	+0.282250
Reco Eff vs. VPDL	$\Delta\mathcal{A}$	-0.038819	+0.001745	+0.014408	+0.008543	+0.041272	-0.024251	+0.035882	+0.059450
	$\Delta\sigma$	+0.001776	+0.003506	+0.005277	+0.009725	+0.017546	+0.027326	+0.054560	+0.137610
Ratio $B_d/B_u = 90\%/10\%$	$\Delta\mathcal{A}$	+0.064193	+0.003572	+0.005938	+0.005451	+0.003900	+0.002103	-0.008299	-0.014020
	$\Delta\sigma$	-0.000318	-0.000081	-0.000104	-0.000154	-0.000241	-0.000352	-0.000670	-0.001550
Pythia vs. Reco MC KF	$\Delta\mathcal{A}$	+0.000017	+0.000228	-0.004783	-0.004469	+0.009366	+0.011031	+0.020870	+0.122620
	$\Delta\sigma$	-0.000010	+0.001879	+0.004527	+0.008925	+0.012041	+0.016201	+0.023250	+0.018135
Total Syst.	σ^{sys}_{tot}	+0.206132	+0.115050	+0.146588	+0.173101	+0.222818	+0.182979	+0.380172	+0.791179
Total	σ_{tot}	+0.242103	+0.279169	+0.418940	+0.632632	+0.874429	+1.225843	+1.721727	+2.377910

Table D.2: Amplitude scan systematic uncertainties in Run IIb.

Osc. frequency		0 ps^{-1}	3 ps^{-1}	6 ps^{-1}	9 ps^{-1}	12 ps^{-1}	15 ps^{-1}	18 ps^{-1}	21 ps^{-1}
$\mathcal{A}$		-0.022277	+0.477748	-0.499657	-0.196703	+0.490526	+0.631818	+0.939469	-0.109186
σ		+0.095643	+0.180564	+0.255173	+0.359505	+0.487201	+0.666072	+0.849860	+1.118220
$c\tau_{B_s} = 441\mu m$	$\Delta\mathcal{A}$	-0.002531	+0.015609	+0.005419	+0.005789	+0.025563	-0.006378	-0.005837	-0.070830
	$\Delta\sigma$	-0.000159	+0.003601	+0.007528	+0.013073	+0.019640	+0.033155	+0.044470	+0.063820
$c\tau_{B_s} \pm \sigma$	$\Delta\mathcal{A}$	+0.001297	+0.007070	-0.002267	-0.003137	-0.012148	+0.002583	-0.003332	-0.029680
	$\Delta\sigma$	+0.000098	+0.001630	-0.003440	-0.006005	-0.009016	-0.015192	-0.020330	-0.029340
$Fr^{Sig}_{peak} \pm \sigma$	$\Delta\mathcal{A}$	-0.003413	-0.011261	-0.005594	-0.013180	-0.029404	-0.037671	-0.040729	-0.034000
	$\Delta\sigma$	-0.000261	-0.001105	-0.001681	-0.002939	-0.004404	-0.008596	-0.011590	-0.018630
$sf_{bkg} \pm \sigma$	$\Delta\mathcal{A}$	-0.000048	+0.000128	+0.000187	+0.000340	+0.000465	-0.000446	-0.000294	-0.000630
	$\Delta\sigma$	-0.000003	+0.000018	+0.000022	+0.000038	-0.000054	-0.000116	-0.000150	-0.000260
$Fr_{prompt} \pm \sigma$	$\Delta\mathcal{A}$	-0.001617	-0.003993	-0.002413	-0.004139	-0.008466	-0.011695	-0.015049	-0.012910
	$\Delta\sigma$	-0.000086	-0.000095	-0.000132	-0.000440	-0.000839	-0.002537	-0.003560	-0.006870
$c\tau_{long} \pm \sigma$	$\Delta\mathcal{A}$	+0.001149	+0.000486	+0.000311	+0.001255	-0.000578	-0.000493	-0.002010	-0.004470
	$\Delta\sigma$	-0.000013	+0.000245	+0.000463	+0.000740	+0.001115	+0.001738	+0.002440	+0.003060
$Fr_{tsens} \pm \sigma$	$\Delta\mathcal{A}$	-0.013397	+0.000271	-0.000086	-0.000101	+0.001552	-0.002513	-0.003636	+0.005050
	$\Delta\sigma$	-0.000047	-0.000047	-0.000045	-0.000036	-0.000027	-0.000037	-0.000030	-0.000110
$Fr_{osc} - \sigma$	$\Delta\mathcal{A}$	+0.020069	-0.002432	-0.003766	+0.002612	+0.003482	+0.001869	-0.000185	+0.013120
	$\Delta\sigma$	-0.000005	+0.000198	+0.000331	+0.000262	+0.000390	+0.000460	+0.000550	+0.000770
$c\tau_{bkg} \pm \sigma$	$\Delta\mathcal{A}$	-0.000262	+0.003828	+0.000201	+0.000078	+0.004521	-0.000582	+0.000726	-0.001080
	$\Delta\sigma$	-0.000081	-0.000225	-0.000680	+0.001281	+0.002349	+0.004164	+0.005910	+0.008130
$Fr_{Long} \pm \sigma$	$\Delta\mathcal{A}$	-0.012600	-0.000320	+0.002815	-0.001209	-0.005053	-0.001602	+0.000951	-0.006830
	$\Delta\sigma$	-0.000031	-0.000383	-0.000813	-0.001169	-0.001909	-0.003002	-0.004190	-0.005580
$N_{D^+Refl} + \sigma$	$\Delta\mathcal{A}$	-0.055610	+0.021205	+0.006533	+0.000007	-0.001758	+0.032170	+0.031955	+0.011639
	$\Delta\sigma$	+0.002971	+0.005559	+0.007386	+0.010299	+0.013998	+0.023227	+0.031476	+0.053060
$N_{D_s} - \sigma$	$\Delta\mathcal{A}$	-0.071751	+0.042986	-0.002613	+0.012653	+0.040957	+0.063791	+0.056219	-0.034040
	$\Delta\sigma$	+0.007931	+0.015153	+0.020961	+0.029680	+0.040936	+0.064246	+0.089892	+0.125600
Fr $D_s/D_s^* \pm \sigma$	$\Delta\mathcal{A}$	+0.001163	+0.003360	+0.003162	-0.010863	-0.007344	+0.019595	-0.016861	+0.004550
	$\Delta\sigma$	+0.000234	+0.000449	+0.000584	-0.001505	-0.001821	-0.002946	-0.003960	-0.003860
Trigger Turn On	$\Delta\mathcal{A}$	+0.000505	+0.000270	-0.000499	-0.000170	-0.001579	-0.000452	-0.002355	-0.008420
	$\Delta\sigma$	-0.000019	+0.000015	+0.000066	+0.000145	+0.000183	+0.000289	+0.000570	+0.001560
Br($B_s \to D_s\mu X$)=5.5%	$\Delta\mathcal{A}$	-0.017032	+0.005137	+0.003657	-0.002059	-0.014216	+0.017087	+0.015412	-0.004990
	$\Delta\sigma$	+0.003134	+0.005359	+0.006649	+0.009032	+0.012005	+0.019286	+0.028261	+0.047520
Br($B_s \to D_s D_s$)=6.8%	$\Delta\mathcal{A}$	-0.000181	+0.001553	-0.001175	-0.001187	+0.004038	+0.001822	+0.001196	-0.003170
	$\Delta\sigma$	+0.000746	+0.001262	+0.001535	+0.002079	+0.002637	+0.003740	+0.004370	+0.005570
Br($B^+ \to D_s X$)=9.3%	$\Delta\mathcal{A}$	+0.003919	+0.000621	-0.000409	-0.000581	+0.001093	-0.000089	-0.000075	-0.001140
	$\Delta\sigma$	+0.000391	+0.000616	+0.000660	+0.000808	+0.000933	+0.001290	+0.001350	+0.001740
Br($B^0 \to D_s X$)=12.4%	$\Delta\mathcal{A}$	+0.005065	+0.000894	-0.002584	-0.002058	+0.002120	-0.000992	-0.001016	-0.004300
	$\Delta\sigma$	+0.000856	+0.001396	+0.001599	+0.002060	+0.002458	+0.003477	+0.003900	+0.005020
Dilution	$\Delta\mathcal{A}$	-0.134527	+0.012328	+0.066861	-0.042326	-0.071468	+0.025982	+0.046391	+0.061844
	$\Delta\sigma$	+0.007948	+0.016277	+0.022756	+0.032356	+0.045311	+0.069135	+0.097232	+0.135940
Reco Eff vs. VPDL	$\Delta\mathcal{A}$	-0.038216	-0.004358	+0.023460	+0.025635	+0.019777	+0.043528	+0.079261	+0.084051
	$\Delta\sigma$	+0.001227	+0.002581	+0.005076	+0.010012	+0.016629	+0.029214	+0.044592	+0.067100
Ratio $B_d/B_u = 90\%/10\%$	$\Delta\mathcal{A}$	+0.056126	+0.006606	+0.007743	+0.008769	+0.009396	+0.009785	+0.004751	-0.000630
	$\Delta\sigma$	-0.000383	+0.000111	+0.000186	+0.000165	+0.000138	+0.000186	+0.000180	+0.000310
Pythia vs. Reco MC KF	$\Delta\mathcal{A}$	-0.000112	+0.000035	-0.004166	+0.007126	+0.008881	-0.000893	+0.051132	-0.037150
	$\Delta\sigma$	+0.000074	-0.000886	+0.000062	+0.000501	+0.005061	+0.008312	+0.020560	+0.025210
Total Syst.	σ^{sys}_{tot}	+0.212631	+0.144097	+0.161992	+0.146048	+0.136450	+0.182108	+0.251592	+0.545179
Total	σ_{tot}	+0.233151	+0.231014	+0.302249	+0.388039	+0.505948	+0.690518	+0.886319	+1.244040

Table D.3: Amplitude scan systematic uncertainties in the combined data .

Osc. frequency		0 ps^{-1}	3 ps^{-1}	6 ps^{-1}	9 ps^{-1}	12 ps^{-1}	15 ps^{-1}	18 ps^{-1}	21 ps^{-1}
$\mathcal{A}$		-0.074859	+0.342939	-0.268535	-0.102516	+0.592354	+0.308424	+0.498157	-0.590181
σ		+0.076407	+0.147123	+0.214277	+0.309466	+0.421599	+0.586102	+0.765744	+1.011010
$c\tau_{B_s} = 441\mu m$	$\Delta\mathcal{A}$	-0.004607	+0.004339	+0.010639	+0.009943	+0.026167	-0.015779	-0.006523	-0.063320
	$\Delta\sigma$	-0.000090	+0.001982	+0.004507	+0.008142	+0.012321	+0.019899	+0.028068	+0.037610
$c\tau_{B_s} \pm \sigma$	$\Delta\mathcal{A}$	+0.003091	+0.001812	-0.004980	-0.005535	-0.015525	-0.005932	+0.001887	+0.031510
	$\Delta\sigma$	+0.000065	+0.001121	-0.002553	-0.004609	-0.006926	-0.011296	+0.015657	-0.022090
$Fr^{Sig}_{peak} \pm \sigma$	$\Delta\mathcal{A}$	-0.006332	-0.009902	-0.008417	+0.014327	+0.036495	+0.023491	+0.032060	-0.028750
	$\Delta\sigma$	-0.000207	+0.001459	+0.002904	+0.005214	+0.007682	+0.013559	+0.018700	+0.029000
$sf_{bkg} \pm \sigma$	$\Delta\mathcal{A}$	-0.000023	-0.000106	-0.000193	-0.000320	-0.000448	-0.000217	-0.000079	-0.000690
	$\Delta\sigma$	-0.000003	-0.000011	+0.000016	-0.000030	-0.000046	-0.000093	-0.000132	-0.000230
$Fr_{prompt} \pm \sigma$	$\Delta\mathcal{A}$	+0.001378	+0.003133	+0.002912	+0.004093	+0.008180	+0.006863	+0.008691	+0.002000
	$\Delta\sigma$	-0.000058	+0.000043	+0.000105	+0.000379	+0.000728	+0.001924	+0.002951	+0.005800
$c\tau_{long} \pm \sigma$	$\Delta\mathcal{A}$	+0.002794	+0.003329	+0.003128	+0.004488	+0.008451	+0.007814	+0.011351	+0.007790
	$\Delta\sigma$	-0.000063	+0.000194	+0.000415	+0.000891	+0.001471	+0.003065	+0.004619	+0.007720
$Fr_{tsens} \pm \sigma$	$\Delta\mathcal{A}$	+0.015824	+0.003651	+0.002861	+0.004076	+0.008655	+0.008333	+0.011696	+0.005560
	$\Delta\sigma$	-0.000081	+0.000059	+0.000120	+0.000376	+0.000701	+0.001841	+0.002843	+0.005660
$Fr_{osc} - \sigma$	$\Delta\mathcal{A}$	+0.012579	-0.000413	-0.000133	+0.007953	+0.022231	+0.012090	+0.016087	-0.002730
	$\Delta\sigma$	-0.000093	+0.000936	+0.002028	+0.003653	+0.005473	+0.009917	+0.013351	+0.020580
$c\tau_{bkg} \pm \sigma$	$\Delta\mathcal{A}$	+0.001881	+0.005465	+0.003007	+0.004174	+0.012540	+0.007220	+0.010008	+0.008550
	$\Delta\sigma$	-0.000107	+0.000167	+0.000593	+0.001368	+0.002398	+0.004781	+0.007084	+0.011380
$Fr_{Long} \pm \sigma$	$\Delta\mathcal{A}$	+0.014916	+0.004739	+0.005061	+0.005178	+0.012930	+0.008467	+0.008917	+0.003420
	$\Delta\sigma$	-0.000069	+0.000277	+0.000630	+0.001236	+0.002027	+0.003922	+0.005612	+0.009550
$N_{D^+Refl} + \sigma$	$\Delta\mathcal{A}$	-0.060177	+0.011425	+0.008790	+0.008913	+0.019880	+0.010431	+0.050082	+0.054630
	$\Delta\sigma$	+0.002269	+0.004259	+0.005757	+0.008246	+0.011445	+0.017302	+0.023430	+0.039520
$N_{D_s} - \sigma$	$\Delta\mathcal{A}$	-0.082031	+0.029213	+0.011620	+0.026961	+0.072905	+0.033058	+0.055732	-0.010392
	$\Delta\sigma$	+0.006251	+0.011871	+0.016912	+0.024576	+0.033810	+0.051458	+0.071806	+0.101330
Fr $D_s/D_s^* \pm \sigma$	$\Delta\mathcal{A}$	+0.001765	+0.005771	+0.005922	-0.006042	+0.024880	+0.009715	+0.029272	+0.031750
	$\Delta\sigma$	+0.000024	+0.000245	-0.000549	+0.001270	+0.001711	+0.003209	+0.003758	+0.006060
Trigger Turn On	$\Delta\mathcal{A}$	+0.001544	+0.003153	+0.002660	+0.003747	+0.007382	+0.005962	+0.009264	-0.002270
	$\Delta\sigma$	-0.000069	+0.000007	+0.000057	+0.000319	+0.000642	+0.001772	+0.002895	+0.005950
Br($B_s \to D_s\mu X$)=5.5%	$\Delta\mathcal{A}$	-0.022460	+0.000266	+0.008685	+0.004918	+0.009752	+0.008496	+0.029707	+0.015789
	$\Delta\sigma$	+0.002387	+0.004084	+0.005155	+0.007110	+0.009342	+0.013779	+0.019947	+0.033090
Br($B_s \to D_s D_s$)=6.8%	$\Delta\mathcal{A}$	+0.000852	+0.003159	+0.002395	+0.003929	+0.012614	+0.008813	+0.011631	-0.001750
	$\Delta\sigma$	+0.000480	+0.000941	+0.001244	+0.001928	+0.002667	+0.004594	+0.006200	+0.009690
Br($B^+ \to D_s X$)=9.3%	$\Delta\mathcal{A}$	+0.005296	+0.002362	+0.002243	+0.004075	+0.009575	+0.007548	+0.010592	+0.000970
	$\Delta\sigma$	+0.000222	+0.000469	+0.000587	+0.000980	+0.001405	+0.002820	+0.003938	+0.006910
Br($B^0 \to D_s X$)=12.4%	$\Delta\mathcal{A}$	+0.006044	+0.001496	+0.000862	+0.003796	+0.011373	+0.007684	+0.011742	-0.001100
	$\Delta\sigma$	+0.000536	+0.000994	+0.001231	+0.001847	+0.002448	+0.004263	+0.005694	+0.009060
Dilution	$\Delta\mathcal{A}$	-0.139297	+0.010190	+0.047673	-0.033983	-0.043033	+0.040220	+0.108639	+0.173269
	$\Delta\sigma$	+0.006158	+0.012601	+0.018201	+0.026248	+0.036763	+0.054518	+0.077191	+0.143470
Reco Eff vs. VPDL	$\Delta\mathcal{A}$	-0.038566	-0.003387	+0.021268	+0.021489	+0.025658	+0.012748	+0.054042	+0.065132
	$\Delta\sigma$	+0.000970	+0.001866	+0.003449	+0.006998	+0.011457	+0.018335	+0.028195	+0.043190
Ratio $B_d/B_u = 90\%/10\%$	$\Delta\mathcal{A}$	+0.061876	+0.007911	+0.009881	+0.011106	+0.014596	+0.012396	+0.007210	-0.006250
	$\Delta\sigma$	-0.000301	+0.000037	+0.000120	+0.000360	+0.000661	+0.001777	+0.002681	+0.005150
Pythia vs. Reco MC KF	$\Delta\mathcal{A}$	+0.001390	+0.004960	-0.002172	+0.006831	+0.016702	+0.015263	+0.026785	-0.079860
	$\Delta\sigma$	-0.000036	+0.000455	+0.001667	+0.003432	+0.006465	+0.010428	+0.018235	+0.020930
Total Syst.	σ^{sys}_{tot}	+0.222777	+0.133374	+0.155933	+0.155140	+0.151550	+0.191045	+0.260577	+0.466516
Total	σ_{tot}	+0.235515	+0.198579	+0.265009	+0.346176	+0.448010	+0.616453	+0.808866	+1.113453

Bibliography

[1] (for the D0 Collaboration), 'A Search for B_s^0 Oscillations at D0 using $B_s^0 \rightarrow D_s \mu \nu X (D_s \rightarrow K^* K)$', D0 Internal Note 4863.

[2] C.C.W Taylor (ed.), *From the Beginning to Plato*. Routledge (1997).

[3] William R. Newman, *Atoms and Alchemy*. University of Chicago Press (2006).

[4] Seth H. Neddermeyer, Carl D. Anderson, *Science* **76**, 238 (1932).

[5] Seth H. Neddermeyer, Carl D. Anderson, *Phys. Rev.* **51**, 884 (1937).

[6] Brown, Hoddeson, *The Birth of Particle Physics*. Cambridge University Press (1983).

[7] Andrew Pickering, *Constructing Quarks*. University of Chicago Press (1984).

[8] Q. R. Ahmad *et al.*, *Phys. Rev. Lett.* **87**, 071301 (2001).

[9] C. Amsler *et al.* (Particle Data Group), *Phys. Lett. B* **667** (2008).

[10] Nicola Cabibbo, *Phys. Rev. Lett.* **10**, 531 (1963).

[11] David Griffiths, *Introduction to Elementary Particles*. John Wiley and Sons (1987).

[12] Elliot Leader, Enrico Predazzi, *An introduction to gauge theories and modern particle physics*, vol. 1. Cambridge University Press (1996).

[13] S. L. Glashow, J. Iliopoulos, L. Maiani, *Phys. Rev. D* **2**, 1285 (1970).

[14] J. J. Aubert *et al.*, *Phys. Rev. Lett.* **33**, 1404 (1974).

[15] J. E. Augustin *et al.*, *Phys. Rev. Lett.* **33**, 1406 (1974).

[16] J. H. Christenson *et al.*, *Phys. Rev. Lett.* **13**, 138 (1964).

[17] Makoto Kobayashi, Toshihide Maskawa, *Progress of Theoretical Physics* **49**, 652 (1973).

[18] Andrzej J. Buras, Markus E. Lautenbacher, Gaby Ostermaier, *Phys. Rev. D* **50**, 3433 (1994).

[19] J. Charles *et al.*, *Eur. Phys. J. C* **41** (2005), http://ckmfitter.in2p3.fr.

[20] R.Gupta (1997).

[21] Elliot Leader, Enrico Predazzi, *An introduction to gauge theories and modern particle physics*, vol. 2. Cambridge University Press (1996).

[22] K. Anikeev *et al.*, *B Physics at the Tevatron: Run II and Beyond*. hep-ph/0201071 (2002).

[23] M. Ciuchini *et al.* hep-ph/0012308 (2000).

[24] V. M. Abazov *et al.* (D0 Collaboration), *Phys. Rev. Lett.* **97**, 021802 (2006).

[25] A. Abulencia *et al.* (CDF Collaboration), *Phys. Rev. Lett.* **97**, 242003 (2006).

[26] (CDF Collaboration), *CDF public note* **9015** (2007).

[27] V. M. Abazov *et al.* (D0 Collaboration), *Phys. Rev. D* **74**, 092001 (2006).

[28] (D0 Collaboration), *D0 conference note* **5933**.

[29] E. Barberio *et al.* (2009), http://lepbosc.web.cern.ch/LEPBOSC/combined_results/mid_2009_

[30] V. M. Abazov *et al.*, *Phys. Rev. Lett.* **98**, 151801 (2007).

[31] B. Abbott *et al.* (D0 Collaboration), *Phys. Lett. B* **487**, 264 (2000).

[32] R. D. Field, *Phys. Rev. D* **65**, 094006 (2002).

[33] B. Andersson, G. Gustafson, C. Peterson, *Z. Phys. C* **1**, 105 (1979).

[34] Alexander Lenz, Ulrich Nierste, *JHEP* **6**, 72 (2007).

[35] M. Bona *et al.* (UTfit Collaboration), *JHEP* **07**, 28 (2005), http://www.utfit.org.

[36] (CDF Collaboration), *CDF public note* **9458** (2008).

[37] T. Aaltonen *et al.* (CDF Collaboration), *Phys. Rev. Lett.* **100**, 161802 (2008).

[38] V. M. Abazov *et al.* (D0 Collaboration), *Phys. Rev. Lett.* **101**, 241801 (2008).

[39] E. Barberio *et al.*, 'Averages of b-hadron and c-hadron Properties at the End of 2007' (2008), arXiv:0808.1297.

[40] V. M. Abazov *et al.* (D0 Collaboration), *Phys. Rev. Lett.* **102**, 091801 (2009).

[41] B.Abbott *et al.*, 'B_s Mixing studies with $B_s \to D_s \mu X(D_s \to K^*K)$ Decay using unbinned Fit', D0 Internal Note 5172.

[42] 'Accelerator Division Rookie Books', http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html.

[43] 'Fermilab Proton Source Department', http://www-bd.fnal.gov/proton/proton.html.

[44] V.M. Abazov *et al.* (D0 Collaboration), *Nucl. Instrum. and Meth. A* **565**, 463 (2006).

[45] Dmitri Tsybychev (for the D0 Collaboration), *Nucl. Instrum. and Meth. A* **582**, 701 (2007).

[46] W.E. Cooper (for the D0 Collaboration), *Nucl. Instrum. and Meth. A* **598**, 41 (2009).

[47] B. Baumbaugh *et al.*, *IEEE Trans. Nucl. Sci.* **43**, 1146 (1996).

[48] D. Adams, 'The D0 Upgrade Central Fiber Tracker - Technical Design Report', D0 Internal Note 4164.

[49] Oana Boeriu, Mike Hildreth, Shabnam Jabeen, 'CFT Light Yeild Studies', D0 Internal Note 4602.

[50] P. Baringer *et al.* (for the D0 Collaboration), *Nucl. Instrum. and Meth. A* **469**, 295 (2001).

[51] A. Gordeev *et al.*, 'Technical Design Report of the Forward Preshower Detector for the D0 Upgrade', D0 Internal Note 3445.

[52] M. Adams *et al.* (for the D0 Collaboration), *Nucl. Instrum. and Meth. A* **366**, 263 (1995).

[53] M. Adams *et al.* (for the D0 Collaboration), *Nucl. Instrum. and Meth. A* **378**, 131 (1996).

[54] Claus Grupen, Boris Shwartz, *Particle Detectors*. Cambridge University Press, 2nd edn. (2008).

[55] S. Abachi *et al.* (D0 Collaboration), *Nucl. Instrum. and Meth. A* **338**, 185 (1994).

[56] V.M. Abazov *et al.* (D0 Collaboration), *Nucl. Instrum. and Meth. A* **552**, 372 (2005).

[57] B. Baldin *et al.*, 'Technical Design of the Central Muon System', D0 Internal Note 3365.

[58] B. S. Acharya *et al.* (for the D0 Collaboration), *Nucl. Instrum. and Meth. A* **401**, 45 (1997).

[59] G. Alexeev *et al.*, 'Technical Design for the D0 Forward Muon Tracking Detector Based on Mini-drift Tubes', D0 Internal Note 3366.

[60] M. Abolins *et al.* (for the D0 Collaboration), *Nucl. Instrum. and Meth. A* **584**, 75 (2008).

[61] T. Sjöstrand, S.Mrenna, P.Skands (2006), arXiv:hep-ph/0603175v2].

[62] David J. Lange, *Nucl. Instrum. and Meth. A* **462**, 152 (2001), see http://www.slac.stanford.edu/~lange/EvtGen.

[63] S. Agostinelli *et al.*, *Nucl. Instrum. and Meth. A* **506**, 250 (2003).

[64] G. Borisov, *Nucl. Instrum. and Meth. A* **417**, 384 (1998).

[65] V. M. Abazov *et al.* (D0 Collaboration), *Phys. Rev. D* **74**, 112002 (2006).

[66] J.Abdallah *et al.* (DELPHI Collaboration), *Eru. Phys. J. C* **35**, 35 (2004).

[67] The D0 Collaboration, 'Bd mixing measurement using Opposite-side Flavor Tagging', D0 Internal Note 5029.

[68] M.Anzelc, S.Burdin, 'Decay Length Resolution Studies: Application to B_s^0 Mixing and p17 Calibration', D0 Internal Note 5336.

[69] SungWoo Youn, Brendan Casey, 'Inclusive Single Muon Trigger Efficiency Study for RunIIa', D0 Internal Note 5311.

[70] SungWoo Youn *et al.*, 'Trigger Efficiency Study for RunII at D0 Using Inclusive Single Muon Sample', D0 Internal Note 5522.

[71] K. Holubyev, G. Borissov, 'Measurement of Kaon Asymmetry', D0 Internal Note 5339.

[72] V. M. Abazov *et al.* (D. Collaboration Collaboration), *Phys. Rev. Lett.* **100**, 211802 (2008).

[73] H. G. Moser, A. Roussarie, *Nucl. Instrum. and Meth. A* **384**, 491 (1997).

[74] M.S. Anzelc *et al.*, 'Study of B_s Mixing Using the Semileptonic Decay $B_s \to D_s\mu\nu X$, Where $D_s \to \phi\pi$', D0 Internal Note 5676.

[75] T. Moulik, 'Update on B_s Mixing in $B_s \to D_s e\nu_e X, D_s \to \phi\pi$ decay mode', D0 Internal Note 5682.

[76] G. Borissov *et al.*, 'A Search For Bs Oscillations Using $B_s \to D_s\mu X(D_s \to K_S^0 K)$ Decays', D0 Internal Note 5254.

[77] (for the D0 Collaboration), 'D0 B_s Oscillation Combination for Summer 2007', D0 Internal Note 5618.

[78] S.Beale, M.Naimuddin, W.Taylor, 'B_s Mixing studies with $B_s \to D_s\mu\nu X(D_s \to K^*K)$ Decay using Unbinned Fit and Same Side Tagging', D0 Internal Note 5446.

[79] V. M. Abazov *et al.* (D0 Collaboration) arXiv:0904.3907v1 [hep-ex] (submitted to Phys. Rev. Lett.).

[80] (for the D0 Collaboration), 'Study of the B_s semileptonic decay rate asymmetry using a time-dependent tagged analysis', D0 Internal Note 5693.

[81] (for the D0 Collaboration), 'Combined Search for CP Violation in Semileptonic B_s Decays', D0 Internal Note 5886.