Search for B_s^0 oscillations at DØ

Tulika Bose

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY
2006

ABSTRACT

Search for B_s^0 oscillations at DØ

Tulika Bose

Measurement of the B_s^0 oscillation frequency via B_s^0 mixing analyses provides a powerful constraint on the CKM matrix elements. A search for B_s^0 oscillations was performed using data collected by the DØ detector during the period 2002-2005 at the Fermilab Tevatron. Approximately 610 pb^{-1} of data was analyzed to reconstruct a large set of B_s^0 mesons in different semileptonic decay modes. Opposite-side flavor tagging algorithms that were tested on semileptonic B_d^0 decays with the measurement of the B_d^0 mixing frequency were used to determine the initial state flavor of the reconstructed B_s^0 meson.

No significant signal for any particular value of the oscillation frequency was found. A 95% confidence level limit on the B_s^0 oscillation frequency $\Delta m_s > 7.3$ ps^{-1} and a sensitivity of 9.5 ps^{-1} were obtained.

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisor, Hal Evans, for his guidance, wisdom and encouragement these past five years. Without him this thesis would not be what it is today. I find it hard to imagine a better advisor.

Special thanks go to all the past and present members of the Columbia-DØ group for being such a fantastic group of people; Mike Tuts for taking care of all the students (and the innumerable free dinners) and for his meticulous reading of this thesis; John and Gustaaf for their advice on different issues; Georg for helping me get started with DØ software, for his dedication to the STT, and for giving me the opportunity to spend my first American Thanksgiving with his family; Christos, the mixing guru, for getting me started with B Physics and for being such an inspiration; Jovan for his friendship over the years, for planning a most memorable cross-country road-trip, and for visiting India with me; Yannis and Sabine for making the office an enjoyable place and for many wonderful lunch time discussions; David for taking over my STT responsibilities so that I could actually get time to write this thesis.

I'd also like to thank: the members of my dissertation committee for their input which helped improve this thesis; Vivek Jain, Rick Van Kooten, Brad Abbott, Brendan Casey, Sergey Burdin, Andrei Nomerotski, Guennadi Borrisov, and others in the B Physics group for many valuable suggestions; Nayeem and Cano for taking on the B_s^0 mixing challenge with me; the members of my Editorial Board for their comments and for approving this analysis for public dissemination; John, Meenakshi, Ulrich and the entire STT group for making it finally work; Tom Fitzpatrick for being the engineering genius that he is and for making my life so much easier; everybody in the DØ Control Room who literally worked day and night to ensure that we took data; the Columbia Applied Physics professors, and Marlene and Lydia for their help

and guidance; Jerry and Bill for many engaging lunch breaks at Nevis; An Qi for introducing me to AHDL and George Brandenburg for making me feel welcome at Harvard.

Thanks also to all the friends that I have made over the years; Sinjini, Cristina, Alex, Lorenzo, Michel, Eva and Paul for all the fun evenings at Fermilab and for making me less "home-sick" for New York; Raphe, Dominik, and Jovan for all the crazy week-night partying in New York, debates in the Nevis Van, and for exploring the city with me; Suparna and Ben for their friendship during my Plasma Physics years and Johannes for encouraging me to do what I wanted; Irene for being such a good listener and for introducing me to many new things ranging from simple home-cooked pasta meals to Skype; Natasha and Zeeya for being my friends since our Cambridge days, for our endless phone conversations, and for being there for me through good times and bad.

And finally, I would like to thank my family. I would not be writing this dissertation if it weren't for my parents who let go of me and let me travel far across the world so that I could pursue my dreams. For always supporting me and loving me and for making me what I am. My sister, for being the most loving and kind person on earth, and for being a dear friend as well. My brother-in-law, Aakash, for promising to take care of her for the rest of their lives. And Kevin - for inspiring me with his passion for physics, for being incredibly patient and supportive these last six months, for having a loving family, and for being both my fiance and my best friend.

Chapter 1

Introduction

The Standard Model (SM) of particle physics is a cohesive amalgamation of our knowledge of elementary particles and their interactions. The theoretical framework provided by it has passed many rigorous tests till now and the SM predictions have been matched by experimental data with amazing precision [1]. However, the SM does not incorporate gravity and this indicates that although it is an excellent approximation, it cannot be the ultimate theory of fundamental particles and their interactions.

The SM also does not explain or predict the masses of the different particles. Furthermore, there is no real motivation for the method whereby the original electroweak symmetry of the theory appears to spontaneously break into the distinct electromagnetic and weak forces that we observe. At least one scalar field is required to accommodate the observed non-zero masses of the quarks, charged leptons, and weak gauge bosons. The couplings to this field generate flavor violation enabling interactions between different quark families. The complex phase in these couplings leads to violation of the $\mathcal{CP}$ symmetry (or $\mathcal{CP}$ violation), where $\mathcal{C}$ and $\mathcal{P}$ are the charge-

conjugation and parity-transformation operators, respectively. The above forms an important connection between electroweak symmetry breaking (and the fundamental question of how particles acquire mass) and $\mathcal{CP}$ violation. The SM can therefore be vigorously tested by adopting a two-prong approach:

- Understanding the exact mechanism of electroweak symmetry breaking and mass generation;
- Understanding and measuring $\mathcal{CP}$ violation in the SM.

The latter can be done by performing stringent tests of the CKM matrix, a unitary matrix that is constructed within the SM to describe quark mixing or the coupling between an *up*-type quark and a *down*-type quark. The elements of the CKM matrix are typically described in terms of four parameters and determining these is an important goal of particle physics today.

This dissertation focuses on a measurement that aims to put tighter constraints on the two most poorly measured parameters that determine the CKM matrix. The analysis was carried out using proton-antiproton collisions at the world's highest energy accelerator, the Tevatron, located at the Fermi National Accelerator Laboratory (Fermilab). The DØ detector sits at one of the interaction regions and stores information from these collisions. This thesis presents an analysis of the data collected by the DØ detector during the period 2002-2005.

The following chapter provides a brief description of the SM and reviews the theoretical motivation behind the analysis. A summary of how non-SM physics could affect the measurement is also given. Chapter 3 provides an overview of the experimental strategy and introduces the main complications affecting the measurement.

The relevant parts of the DØ detector are reviewed in Chapter 4. A detailed description of the experimental techniques used and the results obtained are covered in Chapter 5. Finally, prospects for future measurements are summarized in Chapter 6.

Chapter 2

Theoretical Overview

2.1 The Standard Model

The Standard Model (SM) of particle physics is a mathematical theory which describes the most basic constituents of matter and their interactions. The particles in the SM are divided into two classes: fermions and bosons. Fermions, having odd half-integer spin compose matter and are cataloged by the fundamental forces which affect them. Bosons, having integer spin, mediate the interaction between fermions, and are responsible for binding fermions together to form composite particles. Table 2.1 gives an overview of the fundamental forces and their corresponding mediators or gauge bosons while Table 2.2 lists the fundamental fermions and their important properties [2].

The SM, at present, includes a description of the electromagnetic, weak, and strong forces only. The photon is the carrier of the electromagnetic force, i.e. it mediates the interactions between electrically charged particles. It is massless and has an interaction that is long range and falls off as $1/r^2$. The weak interactions are

Force	Gauge Boson	Symbol	Charge	Spin	Mass (GeV/c^2)
Electromagnetic	Photon	γ	0	1	0
Weak	Z	Z	0	1	91.2
Weak	$W^{\pm}$	$W^{\pm}$	± 1	1	80.4
Strong	Gluon	g	0	1	0
Gravity	Graviton	G	0	2	0

Table 2.1: *The fundamental forces, the Standard Model gauge bosons, and their properties.*

Name	Symbol	Charge	Mass (MeV/c^2)	Interactions
electron	e	-1	0.511	All but Strong
electron neutrino	ν_e	0	< 0.000003	Weak, Gravity
up quark	u	$\frac{2}{3}$	≈ 3	All
down quark	d	$-\frac{1}{3}$	≈ 5	All
muon	μ	-1	105.6	All but Strong
muon neutrino	μ_ν	0	< 0.19	Weak, Gravity
charm quark	c	$\frac{2}{3}$	≈ 1200	All
strange quark	s	$-\frac{1}{3}$	≈ 100	All
tau	τ	-1	1777	All but Strong
tau neutrino	ν_τ	0	< 18.2	Weak, Gravity
bottom quark	b	$-\frac{1}{3}$	≈ 4500	All
top quark	t	$\frac{2}{3}$	$\approx 178,000$	All

Table 2.2: *The fundamental fermions and their properties.*

mediated by the heavy $W^{\pm}$ and Z bosons with an effective interaction that operates at relatively short distances. The electromagnetic and weak interactions are unified in the SM into a single "electroweak" force with the electroweak gauge group being given by $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$. The strong interaction, mediated by gluons, is described by the SU(3) group symmetry of Quantum Chromodynamics or 'QCD'. Analogous to the photon which couples to particles with electric charge, a gluon couples to particles with a "color" charge. There are three color charges referred to as 'red',

'green', and 'blue' ('anti-red', 'anti-green', and 'anti-blue' for anti-particles). The gluons themselves are colored particles and hence interact with each other. The fourth known force is gravity, however, it has not been successfully incorporated in the Standard Model. The gravitational force is much weaker than the strong and electroweak forces at energy scales relevant to particle physics, and hence does not play an important role in the dynamics of the particles studied here. If gravity were to be described by a theory similar to the other forces, a gauge boson called the Graviton is predicted.

The SM groups the fundamental fermions into three families or generations of leptons and quarks. The weak-isospin symmetry[1] of the electroweak theory dictates that there is a weak-isospin doublet of left-handed leptons, a weak-isospin doublet of left-handed quarks, and corresponding doublets of right-handed antiparticles. In addition to this there is a right-handed singlet for each massive (or charged) fermion but these right-handed singlets do not participate in charged weak interactions. Each generation or lepton doublet has a charged and a neutral lepton. The neutral lepton, called neutrino, was until recently believed to be massless. New experimental evidence, however, indicates that neutrinos have non-zero mass [3]. Nevertheless, the two interactions in which neutrinos participate are very weak and hence neutrinos cannot be detected in typical collider experiments. However, since neutrinos do carry momentum and energy, their presence can be inferred by looking for significant 'missing' energy or momentum. An additional point to note regarding leptons in general is that the electroweak force conserves lepton number[2] — it cannot transform leptons from one doublet to those of another. Therefore, while neutral current interactions

[1]Under weak-isopin symmetry, a left-handed charged lepton — electron, muon or tau — and its associated neutrino are viewed as two possible quantum states of the same entity.

[2]There are actually three kinds of lepton number, one for each family, and each one is independently conserved.

via the Z do not change the type of lepton, charged W exchange can only transform leptons within doublets. For instance, the decay $\mu^- \rightarrow e^- \bar{\nu_e} \nu_\mu$ is allowed but the decay $\pi^0 \rightarrow \mu^\pm e^\mp$ is not.

As mentioned above, the second category of fermions, quarks, can also be arranged in three generations of doublets like the leptons. Each quark doublet is composed of one quark with $+2/3$ electrical charge (*up*-type) and another with $-1/3$ electrical charge (*down*-type). Like the leptons, members of each doublet can be transformed into one another by charged W exchange. However, the weak force in this case does not conserve a "quark family number", allowing cross-generational coupling to occur. This is discussed in more detail in Section 2.2. Also, unlike leptons, quarks carry the color charge and hence interact via the strong force. Typically, as the energy of the interaction between quarks increases, the strength of the interaction decreases. For high energy interactions (E $\approx$ 10 GeV) quarks behave almost like free particles (asymptotic freedom). At lower energies the interaction quickly grows in strength and individual quarks are bound into "colorless" composite particles called *hadrons*. Therefore, unlike leptons, except for brief moments in high energy collisions quarks are not seen as free particles. Even in high energy collisions, quarks do not remain free for very long. Within a time scale typical of strong interactions ($\approx 10^{-24}$ s), quark-antiquark pairs are pulled out of the vacuum and they bind with the quarks from the hard scattering. This process, referred to as fragmentation or hadronization, quickly results in a high energy quark producing a large number of quark-antiquark pairs to form composite particles. Hence, in high energy collisions of hadronic particles, although quarks and gluons are the fundamental participants in the interaction, only the composite hadrons are available to the experimenter. Additionally, owing to conservation of momentum and energy, the hadrons which are produced from the quark which initiated the process form a collimated *jet* of hadronic particles. These

hadrons are then classified into baryons and mesons. Baryons are made of three quarks, qqq, as for example the proton, $p \sim uud$, while mesons are made of one quark and one antiquark as for instance the B mesons, $B_s^0 \sim \bar{b}s$ and $B_d^0 \sim \bar{b}d$. Table 2.3 lists some of the baryons and mesons mentioned in this dissertation, their constituent quarks, and some of their properties [2].

Hadron	Constituents	Charge	Mass (MeV)
Proton	uud	1	938
B^+	$\bar{b}u$	1	5279
B_d^0	$\bar{b}d$	0	5279
B_s^0	$\bar{b}s$	0	5370
D^+	$c\bar{d}$	1	1869
D_s^+	$c\bar{s}$	1	1968
K^+	$u\bar{s}$	1	494
π^+	$u\bar{d}$	1	140

Table 2.3: *Some baryons and mesons, and their properties.*

As described above, the SM describes the strong, weak and electromagnetic interactions, and is modeled using a gauge field theory based on the symmetry groups $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$. The $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ group cannot, however, be an exact symmetry since gauge invariance requires the gauge bosons to be massless, and while the photon is indeed massless, the $W^\pm$ and Z bosons are massive. To resolve this apparent contradiction the Higgs Mechanism was introduced [4]. This mechanism spontaneously breaks the $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ symmetry into the observed $\mathrm{SU(3)_C} \otimes \mathrm{U(1)_{EM}}$. A new scalar field, the Higgs field, is introduced. One can then construct interactions with the $W^\pm$ and the Z such that they acquire mass. Fermions acquire mass in a similar way by coupling to this field. Since this process explicitly breaks the symmetry of the interactions, it is known as electroweak symmetry breaking. The Higgs field is a scalar complex weak doublet, which yields one new

physical scalar particle after electroweak symmetry breaking — the Higgs boson. The Higgs boson has not been experimentally observed yet and is much sought after. Present experimental limits indicate that the mass of the Higgs boson is greater than 114.4 GeV/c^2 at the 95% confidence level [2]. An in depth description of the Higgs mechanism and a review of the experimental searches can be found in Refs. [5] and [6], respectively.

2.2 Quark Mixing

Cross-generational coupling (in the quark sector) was first introduced in 1963 by Cabibbo [7]. He suggested that the $d \rightarrow u + W^-$ vertex carries a multiplicative factor of $\cos\theta_c$, whereas the $s \rightarrow u + W^-$ vertex carries a factor of $\sin\theta_c$. The second one is weaker and hence θ_c is small ($\theta_c = 12.7°$ experimentally). This was a fairly successful model except for the fact that it allowed the $K^0 \rightarrow \mu^+\mu^-$ decay via the Feynman diagram in Fig. 2.1(a). According to Cabibbo's model, the width should be $\Gamma(K^0 \rightarrow \mu^+\mu^-) \approx \sin\theta_c \cos\theta_c$. However, this was considerably larger than the experimentally set limit. Glashow, Iliopolos and Maiani came to the rescue of the Cabibbo model in 1970 by postulating the GIM mechanism [8]. This was an extension of the Cabibbo model and included a fourth quark called *charm* (or c-quark) that formed a doublet with the strange quark. In this model the $d \rightarrow c + W^-$ and $s \rightarrow c + W^-$ vertices were associated with factors of $-\sin\theta_c$ and $\cos\theta_c$, respectively, such that the superposition of the Feynman diagrams with the virtual u and c quarks cancel (Fig. 2.1(a),(b)), and the width $\Gamma(K^0 \rightarrow \mu^+\mu^-) \approx 0$.[3]

[3] If the u and c quarks had the exact same mass, the two diagrams would cancel perfectly. Since their masses are not exactly identical, the new diagram suppressed the decay so that the predicted rate was consistent with the experiments.

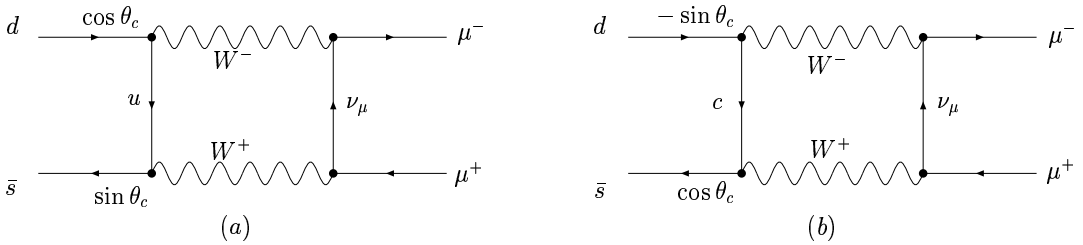


Figure 2.1: *(a) The only first order Feynman diagram for the $K^0 \to \mu^+\mu^-$ decay prior to the formulation of the GIM mechanism. The decay rate was dramatically above the experimentally set limit. (b) The second Feynman diagram cancels most of the contribution of the first; if the up and charm quarks had the same mass, the two diagrams would cancel perfectly. This cancellation is an illustration of the Glashow-Iliopolos-Maiani (GIM) mechanism.*

In general, the GIM mechanism suggested that instead of the physical quarks d and s, the states to use for weak interactions are d' and s', given by

$$d' = (\cos\theta_c)d + (\sin\theta_c)s, \tag{2.1}$$

$$s' = (-\sin\theta_c)d + (\cos\theta_c)s. \tag{2.2}$$

The phenomenon is called quark mixing and Eqs. 2.1 and 2.2 can then be rewritten using the so-called "mixing" matrix which is simply a rotation of the quark basis by the Cabibbo angle θ_c:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}. \tag{2.3}$$

The W's then couple to the "Cabibbo rotated" states

$$\begin{pmatrix} u \\ d' \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} c \\ s' \end{pmatrix}, \tag{2.4}$$

and decays that involve a factor of $\sin\theta_c$ are known as 'Cabibbo suppressed' decays.

Validation of the GIM mechanism came with the discovery of the J/ψ (a bound state of $c\bar{c}$) in 1974 [9, 10]. But even before that, in 1973, Kobayashi and Maskawa [11] added a third generation (the top and bottom quarks) to the model and generalized the GIM mixing matrix to be a general unitary transformation from the flavor states of the *down*-type quarks (d, s, b) to their weak interaction states (d', s', b'). The motivation was to explain $\mathcal{CP}$ violation that had been observed in 1964, by Cronin and Fitch, in the decays of the K^0 meson [12]. With three generations, Kobayashi and Maskawa could incorporate $\mathcal{CP}$ violation through the mixing matrix while with only two generations they could not. A brief discussion is presented in the following section. Experimental proof of the three generation model came with the discovery of the bottom quark in 1977 [13] and that of the top quark in 1995 [14, 15].

2.3 $\mathcal{CP}$ Violation and the CKM Matrix

As mentioned in Section 2.1, the major shortcoming of the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ unified electroweak theory is that the gauge fields (along with the fermions) are left massless. In the SM, this problem is resolved by the Higgs Mechanism. A single SU(2) doublet of complex scalar fields, Φ, is introduced:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{2.5}$$

The most general renormalizable form of the scalar potential for Φ is given by:

$$V(|\Phi|^2) = -\mu^2|\Phi|^2 + \lambda|\Phi|^4 = -\mathcal{L}_{\mathrm{Higgs}}, \tag{2.6}$$

where $|\Phi|^2 = \Phi^\dagger\Phi$. Minimizing V with respect to $|\Phi|^2$ we obtain a ground state that corresponds to $|\Phi|^2 = v^2/2 = \mu^2/2\lambda$. Therefore, the vacuum expectation value of the Φ fields is non-zero, and moreover, since $|\Phi|^2 = |\phi^+|^2 + |\phi^0|^2$, there is a continuous set of allowed ground state values for $|\phi^+|^2$ and $|\phi^0|^2$. By selecting any particular choice of these allowed values of the ground state, the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ symmetry is 'spontaneously broken'. The Higgs field can then be conveniently parametrized by

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \longrightarrow \begin{pmatrix} G^+ \\ \frac{1}{\sqrt{2}}(v + H^0 + iG^0) \end{pmatrix}, \tag{2.7}$$

where H^0 is the Higgs particle and G^+ and G^0 are the Goldstone bosons that become the longitudinal components of the $W^\pm$ and Z bosons, respectively. The charged gauge bosons acquire a mass $M_W = gv/2$ while the neutral gauge boson, Z, acquires mass $M_Z = \sqrt{g^2 + g'^2}v/2$, where g and g' are coupling constants. The mass of the Z boson is related to M_W using the relation $M_Z = M_W/\cos\theta_W$, and this forms the definition of the weak mixing angle $\tan\theta_W = g'/g$.

The SM Lagrangian, $\mathcal{L}_{\mathrm{SM}}$, can then be written as

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{kinetic}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}}, \tag{2.8}$$

where the first term on the right contains the kinetic terms of the fields involved, i.e. left- and right-handed fermions, gauge bosons and scalar Higgs fields ϕ, as well as their gauge interactions, and is always $\mathcal{CP}$ conserving. The second term on the right is the same as in Eq. 2.6 and is $\mathcal{CP}$ -invariant[4] in the SM. This leaves $\mathcal{L}_{\mathrm{Yukawa}}$ which describes the 'Yukawa' couplings of the Higgs field to the fermion fields as the

[4]The reason being that there is only one Higgs doublet in the SM. For an extended scalar sector, such as that of a two Higgs doublet model, $\mathcal{L}_{\mathrm{Higgs}}$ can be $\mathcal{CP}$ violating.

only possible source of $\mathcal{CP}$ violation in the SM. The same Yukawa interactions induce fermion mass terms after spontaneous symmetry breaking.

For the most general case of quarks:

$$\mathcal{L}_{\text{Yukawa}} = -\sum_{i,j=1}^{3} [G_{ij}\overline{Q}_{Li}\Phi d_{Rj} + F_{ij}\overline{Q}_{Li}\tilde{\Phi}u_{Rj}] + h.c. \tag{2.9}$$

where Q_{Li} is a SU(2)_L doublet of weak interaction eigenstates with generation index i, such that

$$Q_{Li} \equiv \begin{pmatrix} u_i \\ d_i \end{pmatrix}, \tag{2.10}$$

and u_{Rj} and d_{Rj} are the right handed SU(2) singlets for the *up*-type and *down*-type quarks, respectively. G_{ij} and F_{ij} are 3×3 Yukawa coupling matrices containing the coupling constants which are parameters of the theory, and $\tilde{\Phi} \equiv i\sigma_2\Phi^*$, where σ_2 is the second Pauli spin matrix. A necessary (but not sufficient) condition for $\mathcal{CP}$ violation is that the Yukawa couplings G_{ij} and F_{ij} are complex[5].

Substituting the vacuum expectation value for ϕ^0 into $\mathcal{L}_{\text{Yukawa}}$ yields *up* and *down* quark mass matrices, $\mathcal{M}^u_{ij}$ and $\mathcal{M}^d_{ij}$:

$$-\mathcal{M}_d\overline{d}_{Li}d_{Rj} - \mathcal{M}_u\overline{u}_{Li}u_{Rj} + h.c. \tag{2.11}$$

[5]The hermiticity of the Lagrangian implies that $\mathcal{L}_{\text{Yukawa}}$ has its terms in the form

$$\mathcal{L}_{\text{Yukawa}} = -\sum_{i,j=1}^{3} [G_{ij}\overline{Q}_{Li}\Phi d_{Rj} + (G_{ij})^*\overline{d}_{Rj}\Phi^\dagger Q_{Li} + \textit{up}\text{-type quark terms...}].$$

The $\mathcal{P}$-transformation exchanges L (left) and R (right) indices, the $\mathcal{C}$-transformation exchanges particles and anti-particles, so that

$$\mathcal{CP}\text{:} \qquad \overline{Q}_{Li}\Phi d_{Rj} \rightarrow \overline{d}_{Rj}\Phi^\dagger Q_{Li}.$$

This implies that $\mathcal{CP}$ is a symmetry of $\mathcal{L}_{\text{Yukawa}}$ if $G_{ij} = G^*_{ij}$. Or in other words, G_{ij} and F_{ij} have to be complex for accommodating $\mathcal{CP}$ violation in the SM.

where $\mathcal{M}_d = \frac{v}{\sqrt{2}} G_{ij}$ and $\mathcal{M}_u = \frac{v}{\sqrt{2}} F_{ij}$. Each mass matrix is composed of nine complex parameters and the off-diagonal elements couple weak interaction eigenstates of different generations. These mass matrices can be transformed into real, diagonal matrices describing the physical mass eigenstates via bi-unitary transformations[6]. One defines four unitary matrices such that:

$$U_{dL}\mathcal{M}_d U_{dR}^{\dagger} = \mathrm{Diag}(\mathrm{m_d}, \mathrm{m_s}, \mathrm{m_b}) \; ; \qquad U_{uL}\mathcal{M}_u U_{uR}^{\dagger} = \mathrm{Diag}(\mathrm{m_u}, \mathrm{m_c}, \mathrm{m_t}) \; . \tag{2.12}$$

The charged current interactions are then given in the mass eigenbasis by

$$\mathcal{L}_W = -\sqrt{\frac{1}{2}} g \overline{u}_{Li} \gamma^u (U_{uL} U_{dL}^{\dagger}) d_{Lj} W_{\mu}^{+} + h.c. \; . \tag{2.13}$$

where

$$V \equiv U_{uL} U_{dL}^{\dagger} \tag{2.14}$$

is the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix for three quark generations. Moreover, since U_{uL} and U_{dL} are unitary by construction, so is V, such that

$$VV^{\dagger} = 1 = V^{\dagger}V. \tag{2.15}$$

The elements of the CKM matrix are given by

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{2.16}$$

and describe the mixing strengths of the quarks, i.e. the strengths of the charged

[6] A complex matrix can be diagonalized by multiplying it on the left and right by the appropriate unitary matrices.

current transitions between *up*-type and *down*-type quarks. These elements are not predicted by the SM, but can, in principle, be determined by measuring the strengths of transitions between quarks. Table 2.4 summarizes the observed strengths of these quark transitions and lists one method of measuring them [2]. A hierarchy can be observed in the sense that transitions between the first and second families are about an order of magnitude stronger than transitions between the second and third. Transitions between the first and third families are even weaker.

Quark transition	Strength	Method of measurement
u, d	0.9739—0.9751	Nuclear β decay
u, s	0.221—0.227	$\bar{K^0} \to \pi^+ e^- \bar{\nu}_e$
u, b	0.0029—0.0045	$B \to \pi l \bar{\nu}_l$
c, d	0.221—0.227	$D^0 \to \pi^- e^+ \nu_e$
c, s	0.9730—0.9744	$D^0 \to K^- e^+ \nu_e$
c, b	0.039—0.044	$B \to X_c l^- \bar{\nu}_l$
t, d	0.0048—0.014	$B^0 - \bar{B}^0$ mixing
t, s	0.037—0.043	$b \to s\gamma$
t, b	0.9990—0.9992	$t \to bW$

Table 2.4: *Quark transitions, their strengths and (some) methods of measuring them.*

The CKM matrix is typically parametrized in some specific way incorporating the unitarity constraints. In general an $n \times n$ complex matrix has $2n^2$ parameters. However, unitarity requires $V^\dagger V = 1$ which halves the number of independent parameters. Therefore, we are left with n^2 free parameters. As the phases are arbitrary, $2n - 1$ of them can be absorbed by phase rotations. We are then left with $(n-1)^2$ physically independent parameters. Furthermore, a unitary matrix is a complex extension of an orthogonal matrix, therefore $n(n-1)/2$ parameters are identified with rotation angles, leaving $(n-2)(n-1)/2$ complex phases. Hence, for three generations ($n = 3$),

the CKM matrix has four independent parameters. Three of them are identified with the real Euler angles, leaving a single complex phase. This complex phase allows for the accommodation of $\mathcal{CP}$ violation. Note that if $n < 3$, as in the original GIM model, there is no phase left in the matrix and consequently no $\mathcal{CP}$ violation. This was the original motivation behind Kobayashi and Maskawa's proposal for a third generation of quarks.

It should also be noted that $\mathcal{CP}$ is not necessarily violated in the three generation SM. If two quarks of the same charge had equal masses, one mixing angle and phase could be removed from the CKM matrix. This leads to a condition on quark mass differences being imposed for $\mathcal{CP}$ violation:

$$F_u \neq 0 \; ; \quad \text{and} \quad F_d \neq 0, \tag{2.17}$$

where

$$\begin{aligned} F_u &= (m_u^2 - m_c^2)(m_c^2 - m_t^2)(m_t^2 - m_u^2), \\ F_d &= (m_d^2 - m_s^2)(m_s^2 - m_b^2)(m_b^2 - m_d^2). \end{aligned} \tag{2.18}$$

Another useful way [16] of representing the above is by re-writing the commutator of the mass matrices, $C = [\mathcal{M}_u \mathcal{M}_u^\dagger, \mathcal{M}_d \mathcal{M}_d^\dagger]$, as

$$C = U_{uL}^\dagger \left[(m_u)^2, V (m_d)^2 V^\dagger \right] U_{uL} \tag{2.19}$$

which shows that det C depends on the physical masses and V. After some algebra[7]

[7]The unitarity of the CKM matrix, $VV^\dagger = 1$, requires that for any choice of $i, j, k, l = 1, 2, 3$

$$J = \mathrm{Im}[V_{ij} V_{kj}^* V_{kl} V_{il}^*] \sum_m \epsilon_{ikm} \sum_n \epsilon_{jln}.$$

one obtains

$$\det C = 2iF_u F_d J \neq 0 \implies \mathcal{CP} \text{ violation}, \tag{2.20}$$

where J is the *Jarlskog parameter* and is given by

$$J = \mathrm{Im}[V_{11}V_{21}^* V_{22}V_{12}^*]. \tag{2.21}$$

The determinant $\det C$ illustrates several essential features of $\mathcal{CP}$ violation in the SM:

- $\det C$ is imaginary, implying that $\mathcal{CP}$ violation originates from a complex coupling.
- There is no $\mathcal{CP}$ violation unless F_u, F_d, and J are non-zero.
- Non-zero F_u and F_d impose conditions on the quark masses (Eq. 2.18).
- The value taken by J is independent of the parametrization (by construction of $\det C$), and gives the strength of $\mathcal{CP}$ violation in the SM.

2.4 Parameterizations of the CKM matrix

The CKM matrix has four quantities with physical significance: three mixing angles and one $\mathcal{CP}$ violating phase. These can be parametrized in many different ways. The Particle Data Group favors the Chau-Keung parametrization [17]:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}c_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \tag{2.22}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ control the mixing between the families and δ_{13} is the $\mathcal{CP}$ violating phase also called the KM phase. The Jarlskog parameter becomes

$$J = c_{12}c_{23}c_{13}^2 s_{12}s_{23}s_{13}\sin\delta_{13}, \tag{2.23}$$

and the parameters must satisfy

$$\delta_{13} \neq 0, \pi \; ; \qquad \theta_{ij} \neq 0, \pi/2; \tag{2.24}$$

otherwise J vanishes. Or in other words, since $\mathcal{CP}$ violation is proportional to J, the CKM matrix must not only have complex entries, but also non-trivial mixing, otherwise the KM phase δ_{13} can be removed.

A convenient parameterization of the CKM matrix was developed by Wolfenstein [18]. He exploited the hierarchy observed in the measured values of the matrix, with diagonal elements close to one, and progressively smaller elements away from the diagonal (see Table 2.4). This hierarchy was formalized by defining λ, A, ρ and η such that

$$\lambda \equiv s_{12}, \quad A \equiv s_{23}/\lambda^2, \quad \rho - i\eta \equiv s_{13}e^{-i\delta_{13}}/A\lambda^3. \tag{2.25}$$

From experiment $\lambda \approx 0.22$, $A \approx 0.8$, and $\sqrt{\rho^2+\eta^2} \approx 0.4$, so every element of the CKM matrix, V, was expanded as a power series in the small parameter $\lambda = |V_{us}|$. Neglecting terms of $\mathcal{O}(\lambda^4)$ resulted in the famous "Wolfenstein parameterization":

$$V = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.26}$$

λ, A, and $\sqrt{\rho^2+\eta^2}$ are real while the phase in question is given by $\arg(\rho,\eta)$. This

parametrization allows for $\mathcal{CP}$ violation if $\eta \neq 0$. The Jarlskog parameter can now be expressed as $J = A^2\lambda^6\eta \sim 10^{-5}$. This indicates that the CKM $\mathcal{CP}$ violation is rather small owing to the smallness of the mixing angles (suppressed by order λ).

This brings forward a very interesting connection to cosmology and new physics. One of the characteristic features of our Universe is the matter-antimatter asymmetry of $\mathcal{O}(10^{-10})$ [19]. In 1967 Sakharov outlined three conditions to explain the dominance of matter over anti-matter in the Universe [20]:

(1) Fundamental interactions violate baryon number conservation (or the proton must decay);

(2) The Universe must have passed through a period of thermal non-equilibrium;

(3) There is $\mathcal{CP}$ violation in the Universe.

Since current experimental results from the CKM sector imply the strength of $\mathcal{CP}$ violation to be at the 10^{-5} level, the $\mathcal{CP}$ violation present in the SM cannot account for the matter-antimatter asymmetry in the Universe. It is, therefore, quite possible that new physics related to new sources of $\mathcal{CP}$ violation is responsible for this large matter-antimatter asymmetry. This stresses the importance of testing the SM via precision measurements of the CKM matrix elements and being on the lookout for new physics that might lead to results that are inconsistent with SM predictions.

2.5 The Unitarity Triangle

Using the unitarity property of the CKM matrix one obtains the following six relations:

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0, \quad (2.27)$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0, \quad (2.28)$$

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \quad (2.29)$$

$$V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0, \quad (2.30)$$

$$V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0, \quad (2.31)$$

$$V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0, \quad (2.32)$$

where the first three relations express the orthogonality of two different columns, and the last three express the orthogonality of two different rows. These relations can be geometrically represented in the complex plane as "unitarity" triangles with rather different shapes. There are only two triangles (Eqs. 2.29 and 2.32) that have their sides approximately equal ($\sim A\lambda^3$) and hence are considered special. The other triangles are long and thin — the triangles in Eqs. 2.28 and 2.31 have sides $(\lambda^2, \lambda^2, A\lambda^4)$ while the triangles in Eqs. 2.27 and 2.30 have sides of the order $(\lambda, \lambda, A\lambda^5)$. Remarkably, all the triangles are equal in area and the area of each triangle equals $|J|/2$.[8]

Taking a closer look at the two special ("non-squashed") triangles, and using the

[8] A rotation of the CKM matrix rotates the unitarity triangle accordingly while leaving its area the same and hence J is also known as the *Jarlskog invariant*.

Wolfenstein parameterization (Eq. 2.26) and neglecting $\mathcal{O}(\lambda^4)$ terms we note that [21]

$$\underbrace{V_{ud}V_{ub}^*}_{(\rho+i\eta)A\lambda^3} + \underbrace{V_{cd}V_{cb}^*}_{-A\lambda^3} + \underbrace{V_{td}V_{tb}^*}_{(1-\rho-i\eta)A\lambda^3} = 0, \tag{2.33}$$

$$\underbrace{V_{ud}V_{td}^*}_{(1-\rho-i\eta)A\lambda^3} + \underbrace{V_{us}V_{ts}^*}_{-A\lambda^3} + \underbrace{V_{ub}V_{tb}^*}_{(\rho+i\eta)A\lambda^3} = 0. \tag{2.34}$$

Or in other words, the above two triangles agree with each other at the λ^3 level, and yield a single relation

$$[(\rho + i\eta) + (-1) + (1 - \rho - i\eta)]A\lambda^3 = 0. \tag{2.35}$$

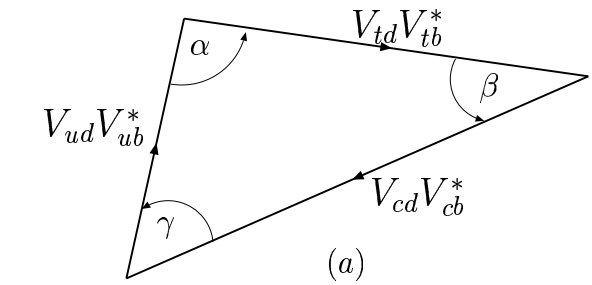
This relation (usually represented using Eq. 2.33) forms "*the* Unitarity Triangle" (Fig. 2.2(a)). Aligning $V_{cd}V_{cb}^*$ with the real axis and dividing all sides by its magnitude $|V_{cd}V_{cb}^*|$ (or $A\lambda^3$), one obtains a rescaled Unitarity Triangle (Fig. 2.2(b)). Two vertices of the rescaled Unitarity Triangle are thus fixed at (0,0) and (1,0) while the coordinates of the third vertex are denoted by the Wolfenstein parameters (ρ, η). The three angles of the triangle are given by:

$$\alpha = \arg[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}], \quad \beta = \arg[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}], \quad \gamma = \arg[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}], \tag{2.36}$$

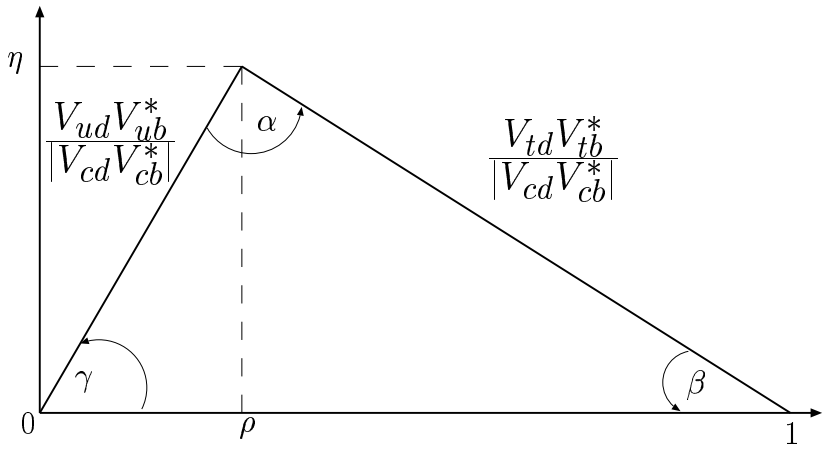
where by construction $\alpha + \beta + \gamma = \pi$.

For higher order corrections ($\mathcal{O}(\lambda^5)$) to the Wolfenstein parameterization, it is common practice to introduce the following definition:

$$\bar{\rho} + i\bar{\eta} \equiv -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \tag{2.37}$$



(a)



(b)

Figure 2.2: *(a) The Unitarity Triangle. (b) The rescaled Unitarity Triangle, all sides divided by* $|V_{cd}V_{cb}^*|$.

where

$$\bar{\rho} = \rho(1 - \lambda^2/2), \qquad \text{and} \qquad \bar{\eta} = \eta(1 - \lambda^2/2). \tag{2.38}$$

Then the rescaled triangle has its apex in the complex plane at $(\bar{\rho}, \bar{\eta})$ and its sides are given by

$$R_b \equiv \sqrt{\bar{\rho}^2 + \bar{\eta}^2} = \frac{1 - \lambda^2/2}{\lambda} \left| \frac{V_{ub}}{V_{cb}} \right|, \qquad R_t \equiv \sqrt{(1 - \bar{\rho})^2 + \bar{\eta}^2} = \frac{1}{\lambda} \left| \frac{V_{td}}{V_{cb}} \right|. \tag{2.39}$$

The angles of the triangle are then easily expressed as:

$$\alpha = \tan^{-1}\left(\frac{\bar{\eta}}{\bar{\eta}^2 + \bar{\rho}(\bar{\rho} - 1)}\right), \qquad \beta = \tan^{-1}\left(\frac{\bar{\eta}}{1 - \bar{\rho}}\right), \qquad \gamma = \tan^{-1}\left(\frac{\bar{\eta}}{\bar{\rho}}\right), \tag{2.40}$$

and we arrive at the unitarity triangle illustrated in Fig. 2.3(a). This is a straightforward generalization of the leading-order case described by Eq. 2.29 where the apex is now simply given by $(\bar{\rho}, \bar{\eta})$ instead of (ρ, η). Since $\bar{\eta}$, $\bar{\rho}$, and $1 - \bar{\rho}$ could potentially be of comparable size, the angles and, hence, the corresponding $\mathcal{CP}$ asymmetries could be large.

Applying the same higher order correction to Eq. 2.32 and again dividing by $|V_{cd}V_{cb}^*|$, we obtain the unitarity triangle drawn in Fig. 2.3(b) [21]. The apex of this triangle is given by (ρ, η) and not $(\bar{\rho}, \bar{\eta})$. In addition, there is a tiny angle ($\delta\gamma \equiv \lambda^2\eta = \mathcal{O}(1^o)$) between the real axis and the base axis which satisfies the relation

$$\gamma = \gamma' + \delta\gamma, \tag{2.41}$$

where γ is the corresponding angle in Fig. 2.3(a).

At present, the term "Unitarity Triangle" is reserved for the leading order case

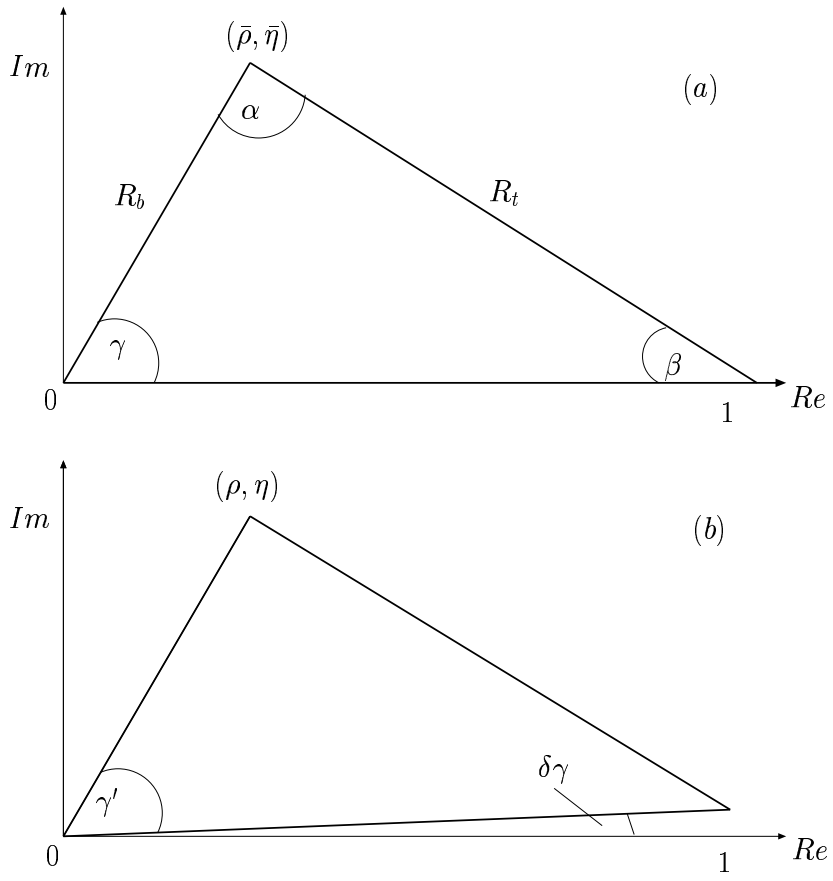


Figure 2.3: *The two "non-squashed" unitarity triangles corresponding to orthogonality relations (a)* $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$ *and (b)* $V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0$.

illustrated in Fig. 2.3(a). In the future (with the start of the next generation machine, the Large Hadron Collider at CERN) when the accuracy of the experimental results improve considerably, the unitarity triangle illustrated in Fig. 2.3(b) will start playing a role and the tiny angle $\delta\gamma$ will be probed directly.

In any case, the Unitarity Triangle described above is extremely useful since it provides a simple and vivid summary of the CKM mechanism. Since separate measurements of the sides and angles of the triangle should fit together, any inconsistency, if found, that is if the triangle is not 'closed', would be a sure sign of physics beyond the three generation Standard Model. Conversely, precise measurements of the CKM matrix elements may put severe bounds on different models of new physics.

It is useful, in this context, to look for inconsistencies by plotting the different experimental constraints in the $\bar{\rho}-\bar{\eta}$ plane. There are four classic types of experiments which constrain the parameters $\bar{\rho}$ and $\bar{\eta}$:

- The determination of $|V_{ub}/(\lambda V_{cb})|$ from $b \to u$ and $b \to c$ decays constrains R_b (see Eq. 2.39). These bounds correspond to circles centered at $(\bar{\rho}, \bar{\eta}) = (0, 0)$.

- The parameter ϵ_K measuring $\mathcal{CP}$ violation in the kaon sector leads to constraints of the type $\bar{\eta}(a - \bar{\rho}) = b$ with suitable constants a and b. These bounds correspond to an hyperbola in the $\bar{\rho} - \bar{\eta}$ plane.

- The measurement of $\sin(2\beta)$ constrains the angle β of the triangle which in turn is related to $\bar{\rho}$ and $\bar{\eta}$ through Eq. 2.40. These bounds correspond to the area between the two lines passing through $(\bar{\rho}, \bar{\eta}) = (1, 0)$.

- Mixing in the $B^0 - \bar{B}^0$ system leads to a constraint on R_t (see Eq. 2.39). These bounds correspond to circles centered at $(\bar{\rho}, \bar{\eta}) = (1, 0)$. This is discussed in more detail in Sec. 2.6.

The combination of the above experimental constraints is shown in Fig. 2.4 [22].

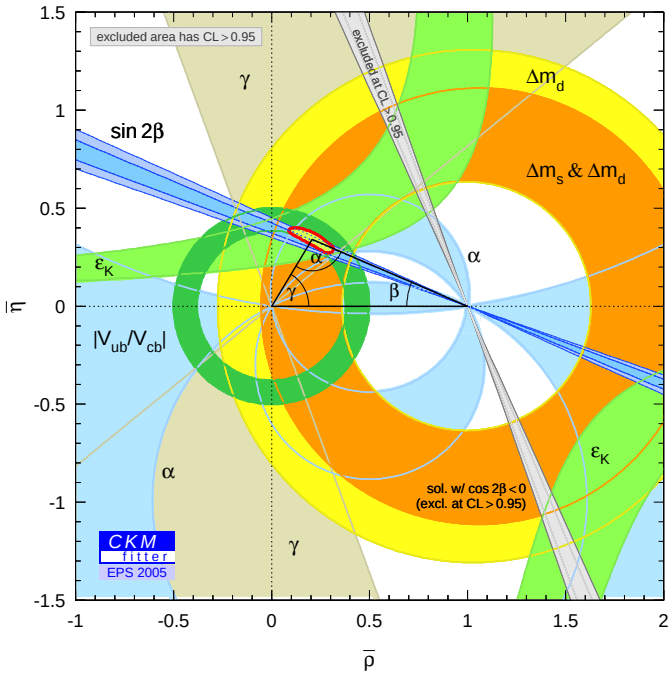


Figure 2.4: *Experimental constraints on the $\bar{\rho}-\bar{\eta}$ plane. The circles centered at (0,0) come from $b \to u$ and $b \to c$ decays while the hyperbolic curves in green come from kaon sector results. The area between the two lines passing through (1,0) are bounds from measurements of* $\sin(2\beta)$*. The circles centered at (1,0) come from $B^0 - \bar{B}^0$ mixing results and the intersection of all constraints (red ellipse) determines the region in the $\bar{\rho}-\bar{\eta}$ plane consistent with these experiments [22].*

From Fig. 2.4 and Table 2.4 one can note that the largest uncertainty in the triangle's shape comes from the side R_t (or the CKM element V_{td}). Therefore, a precise measurement of V_{td} is essential for properly constraining the CKM matrix. Additionally, V_{td} yields information on the $\mathcal{CP}$ violating phase (Eq. 2.26). A useful measurement of V_{td}, namely a SM loop level process, is mixing in the $B^0 - \bar{B}^0$ system — the chief focus of this dissertation.

2.6 $B^0 - \bar{B}^0$ Mixing

Mixing is the process whereby some neutral mesons change from their particle to their anti-particle state, and vice versa. This kind of oscillation of flavor eigenstates into one another was first predicted for the $K^0 - \bar{K}^0$ system in 1955 by Gell-Mann and Pais [23]. They argued for the existence of a long lived neutral strange particle (as a companion to the short lived neutral kaon) which was subsequently observed by Lande in 1956 [24] and is now known as the K^0_L. Mixing has since then been seen for B mesons, first in a mixture of B^0_d and B^0_s by UA1 [25] and then in B^0_d mesons by ARGUS [26][9].

Analogous to the K^0 and $\bar{K}^0$, the B^0 and $\bar{B}^0$ mesons are created by the strong interaction as eigenstates of flavor ($B^0 = \bar{b}q, \bar{B}^0 = b\bar{q}$ where q can be either d or s). These flavor eigenstates are the same weak eigenstates that have been discussed previously. Particle-antiparticle oscillations are made possible by the charged currents of the weak interaction which result in flavor changing transitions as described by the CKM matrix (see Eqs. 2.13 and 2.14). Second order weak interactions couple the flavor eigenstates as shown by the Feynman box diagrams in Fig. 2.5. The physical particles, those with a definite mass and lifetime, become linear combinations of B^0 and $\bar{B}^0$. Consequently, the mass eigenstates are not the same as the flavor eigenstates and this results in $B^0 - \bar{B}^0$ oscillations.

[9] The only hadrons that undergo oscillations are K^0, D^0, B^0, and B^0_s mesons. The π^0 is its own anti-particle and the top quark is so heavy that it decays before forming stable hadrons. Mixing in the D^0 system is expected to be very small and has not been observed so far.

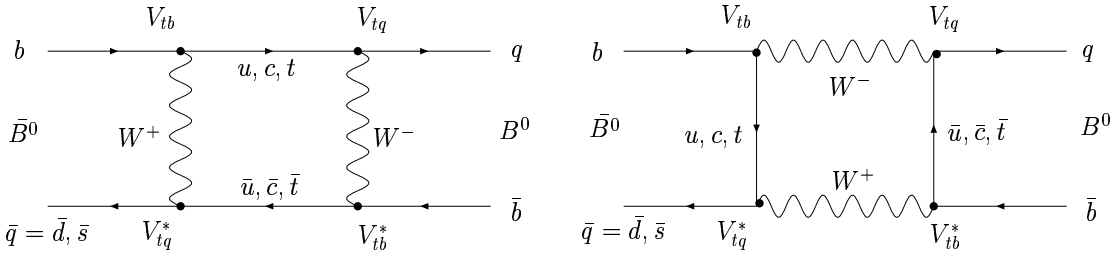


Figure 2.5: *Box diagrams giving rise to neutral B meson oscillations.*

2.7 Mixing Formalism

Consider an initially produced B^0 or $\bar{B}^0$ that evolves in time into a superposition of B^0 and $\bar{B}^0$. Let $|B^0\rangle$ denote the state vector of a B meson which is tagged as a B^0 at time $t = 0$, i.e., $|B^0(t = 0)\rangle = |B^0\rangle$. Similarly, $|\bar{B}^0\rangle$ represents a B meson initially tagged as a $\bar{B}^0$. The time evolution of these states is given by the following Schrodinger equation:

$$i\frac{d}{dt}\begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}\begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix} \tag{2.42}$$

where the mass and decay matrices ($\mathbf{M}$ and $\mathbf{\Gamma}$) are 2×2 t-independent Hermitian matrices. $\mathcal{CPT}$ invariance requires that $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$ so that the particle and anti-particle have the same mass and lifetime. $B^0 - \bar{B}^0$ transitions are induced by non-zero off-diagonal elements where M_{12} represents the virtual transitions (the dominant contribution in Fig. 2.5 is due to the t-quark owing to its larger mass) and Γ_{12} represents the real transitions through common decay modes (i.e. modes that both B^0 and $\bar{B}^0$ mesons can decay to). These common modes are Cabibbo suppressed so that the $B^0 - \bar{B}^0$ mixing amplitude is dominated by virtual transitions [27]. Diagonalization of the Hamiltonian matrix yields the mass eigenstates that can be expressed

in terms of the flavor eigenstates as

$$\begin{aligned} |B_L\rangle &= p|B^0\rangle + q|\bar{B}^0\rangle, \\ |B_H\rangle &= p|B^0\rangle - q|\bar{B}^0\rangle, \end{aligned} \tag{2.43}$$

where B_L and B_H are the light and heavy mass eigenstates, respectively, and the complex coefficients p and q obey the normalization condition $|p|^2 + |q|^2 = 1$. The two eigenvalues are

$$\begin{aligned} \omega_H &= M_H - i\Gamma_H/2, \\ \omega_L &= M_L - i\Gamma_L/2, \end{aligned} \tag{2.44}$$

where $M_{H,L}$ and $\Gamma_{H,L}$ are the masses and widths of the physical states B_H and B_L. The mass difference, Δm, and the width difference, $\Delta\Gamma$, between the neutral B mesons are defined using the convention:

$$\Delta\omega \equiv \omega_H - \omega_L = \Delta m - \frac{i}{2}\Delta\Gamma, \tag{2.45}$$

$$\Delta m \equiv M_H - M_L = \mathcal{R}e(\Delta\omega), \tag{2.46}$$

$$\Delta\Gamma \equiv \Gamma_H - \Gamma_L = -2\mathcal{I}m(\Delta\omega). \tag{2.47}$$

The eigenvalue problem

$$\det|\mathbf{M} - \frac{\mathrm{i}}{2}\mathbf{\Gamma} - \omega| = 0 \tag{2.48}$$

results in the condition

$$\Delta\omega = 2\sqrt{(M_{12}^* - i\Gamma_{12}^*/2)(M_{12} - i\Gamma_{12}/2)}. \tag{2.49}$$

The real and imaginary parts of this equation give

$$\begin{aligned}(\Delta m)^2 - \frac{1}{4}(\Delta\Gamma)^2 &= 4|M_{12}|^2 - |\Gamma_{12}|^2 \qquad \text{and} \\ \Delta m \Delta\Gamma &= 4\mathcal{R}e(M_{12}\Gamma_{12}^*). \end{aligned} \tag{2.50}$$

Using the above two equations, Eqs. 2.46 and 2.47 can be re-written in terms of the matrix elements M_{12} and Γ_{12} as:

$$\Delta m = \sqrt{2}\left(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2 + \sqrt{(|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}^2)^2 + [\mathcal{R}e(M_{12}\Gamma_{12}^*)]^2}\right)^{\frac{1}{2}}, \tag{2.51}$$

$$\Delta\Gamma = 2\sqrt{2}\left(\sqrt{|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2 + [\mathcal{R}e(M_{12}\Gamma_{12}^*)]^2} - (|M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2)\right)^{\frac{1}{2}}. \tag{2.52}$$

Solving for the eigenvalues gives

$$\frac{q}{p} = \frac{-\Delta\omega}{2(M_{12} - \frac{i}{2}\Gamma_{12})} = -\frac{2(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}{\Delta m - \frac{i}{2}\Delta\Gamma}. \tag{2.53}$$

The time evolution of the mass eigenstates is then governed by the two eigenvalues $M_H - i\Gamma_H/2$ and $M_L - i\Gamma_L/2$ such that

$$|B_{H,L}(t)\rangle = e^{-(iM_{H,L} + \Gamma_{H,L}/2)t}|B_{H,L}\rangle, \tag{2.54}$$

where $|B_{H,L}\rangle$ denotes the mass eigenstates at time $t = 0$ (i.e. $|B_{H,L}\rangle = |B_{H,L}(t = 0)\rangle$). Now, inverting Eq. 2.43 to express $|B^0\rangle$ and $|\bar{B^0}\rangle$ in terms of the mass eigenstates

and using their time evolution in Eq. 2.54, we get:

$$\begin{aligned} |B^0(t)\rangle &= \frac{1}{2p}\left[e^{-iM_Lt-\Gamma_Lt/2}|B_L\rangle + e^{-iM_Ht-\Gamma_Ht/2}|B_H\rangle\right], \\ |\bar{B}^0(t)\rangle &= \frac{1}{2q}\left[e^{-iM_Lt-\Gamma_Lt/2}|B_L\rangle - e^{-iM_Ht-\Gamma_Ht/2}|B_H\rangle\right]. \end{aligned} \tag{2.55}$$

Eliminating the mass eigenstates in Eq. 2.55 in favor of the flavor eigenstates we get:

$$\begin{aligned} |B^0(t)\rangle &= g_+(t)|B^0\rangle + g_-(t)\frac{q}{p}|\bar{B}^0\rangle, \\ |\bar{B}^0(t)\rangle &= g_-(t)\frac{p}{q}|B^0\rangle + g_+(t)|\bar{B}^0\rangle, \end{aligned} \tag{2.56}$$

where

$$g_\pm(t) \equiv \frac{1}{2}e^{-iMt}e^{-\frac{\Gamma}{2}t}\left(e^{\frac{\Delta\Gamma}{4}t}e^{i\frac{\Delta m}{2}t} \pm e^{-\frac{\Delta\Gamma}{4}t}e^{-i\frac{\Delta m}{2}t}\right), \tag{2.57}$$

and $M \equiv \frac{1}{2}(M_H + M_L)$ while $\Gamma \equiv \frac{1}{2}(\Gamma_H + \Gamma_L)$.

The above equations indicate that for $t > 0$ there is a finite probability that a $|B^0\rangle$ can be observed as a $|\bar{B}^0\rangle$ and vice versa. This is known as "B^0 mixing".

Let $\mathcal{P}_m^{B^0}(t)$ denote the probability that a particle produced as a B^0 meson oscillated (mixed) and decayed as a $\bar{B}^0$ while $\mathcal{P}_u^{B^0}(t)$ denotes the conjugate probability that this particle did not oscillate, that is, it remained unmixed (similar definitions exist for the initial $\bar{B}$ states). Using Eqs. 2.56 and 2.57 we get the following [28]:

$$\begin{aligned} \mathcal{P}_u^{B^0}(t) &= \frac{e^{-\Gamma t}}{\Gamma\left(\frac{1+|q/p|^2}{\Gamma^2-\Delta\Gamma^2/4} + \frac{1-|q/p|^2}{\Gamma^2+\Delta m^2}\right)}(\cosh\frac{\Delta\Gamma}{2}t + \cos\Delta mt), \\ \mathcal{P}_m^{B^0}(t) &= \frac{|q/p|^2e^{-\Gamma t}}{\Gamma\left(\frac{1+|q/p|^2}{\Gamma^2-\Delta\Gamma^2/4} + \frac{1-|q/p|^2}{\Gamma^2+\Delta m^2}\right)}(\cosh\frac{\Delta\Gamma}{2}t - \cos\Delta mt), \end{aligned} \tag{2.58}$$

$$\begin{aligned}
\mathcal{P}_u^{\bar{B^0}}(t) &= \frac{|q/p|^2 e^{-\Gamma t}}{\Gamma\left(\frac{1+|q/p|^2}{\Gamma^2-\Delta\Gamma^2/4} - \frac{1-|q/p|^2}{\Gamma^2+\Delta m^2}\right)}(\cosh\frac{\Delta\Gamma}{2}t + \cos\Delta m t), \\
\mathcal{P}_u^{\bar{B^0}}(t) &= \frac{e^{-\Gamma t}}{\Gamma\left(\frac{1+|q/p|^2}{\Gamma^2-\Delta\Gamma^2/4} - \frac{1-|q/p|^2}{\Gamma^2+\Delta m^2}\right)}(\cosh\frac{\Delta\Gamma}{2}t - \cos\Delta m t). \qquad (2.59)
\end{aligned}$$

Note that the above formulae are not symmetric for B^0 and $\bar{B}^0$ states (i.e. $\mathcal{P}_m^{B^0}(t) \neq \mathcal{P}_m^{\bar{B^0}}(t)$ and $\mathcal{P}_u^{B^0}(t) \neq \mathcal{P}_u^{\bar{B^0}}(t)$). There are two limiting cases though (given below) for which the B^0 and $\bar{B}^0$ symmetry is regained.

2.7.1 No $\mathcal{CP}$ violation in mixing:

Eq. 2.43 can also be written as

$$\begin{aligned}
|B_L\rangle &= \frac{p+q}{2}\left[(|B^0\rangle + |\bar{B}^0\rangle) + \frac{1-q/p}{1+q/p}(|B^0\rangle - |\bar{B}^0\rangle)\right], \\
|B_H\rangle &= \frac{p+q}{2}\left[(|B^0\rangle - |\bar{B}^0\rangle) + \frac{1-q/p}{1+q/p}(|B^0\rangle + |\bar{B}^0\rangle)\right]. \qquad (2.60)
\end{aligned}$$

The quantity $(1-q/p)/(1+q/p) \equiv \epsilon_B$ is a measure of the amount by which B_H and B_L differ from $\mathcal{CP}$ eigenstates, and according to the SM, is expected to be very small, $\mathcal{O}(10^{-3})$. The current world average for the B_d^0 system is $\mathcal{R}e(\epsilon_B) = 0.002 \pm 0.007$ [2] while there is no measurement yet for the B_s^0 system. $q/p = 1$ thus gives the limit of no $\mathcal{CP}$ violation in mixing and using it in Eqs. 2.58 and 2.59, we obtain $\mathcal{P}_u^{B^0}(t) = \mathcal{P}_u^{\bar{B^0}}(t) = \mathcal{P}_u(t)$ and $\mathcal{P}_m^{B^0}(t) = \mathcal{P}_m^{\bar{B^0}}(t) = \mathcal{P}_m(t)$ where

$$\begin{aligned}
\mathcal{P}_u(t) &= \frac{1}{2}\Gamma e^{-\Gamma t}\left(1 - \frac{\Delta\Gamma^2}{4\Gamma^2}\right)(\cosh\frac{\Delta\Gamma}{2}t + \cos\Delta m t), \\
\mathcal{P}_m(t) &= \frac{1}{2}\Gamma e^{-\Gamma t}\left(1 - \frac{\Delta\Gamma^2}{4\Gamma^2}\right)(\cosh\frac{\Delta\Gamma}{2}t - \cos\Delta m t). \qquad (2.61)
\end{aligned}$$

The above expressions are appropriate for B_s^0 mesons which are not expected to be subject to large $\mathcal{CP}$ violating effects. The ratio $\frac{q}{p}$, as given by Eq. 2.53, is unity to a good approximation, because mixing induced $\mathcal{CP}$ asymmetries vanish due to the fact that $\arg(M_{12}) = \arg(V_{tb}V_{ts}^*) \approx 0$.

2.7.2 Zero lifetime difference:

A simple, approximate solution can be derived when

$$|\Gamma_{12}| \ll |M_{12}|, \qquad \text{and} \qquad \Delta\Gamma \ll \Delta m. \tag{2.62}$$

The above inequalities hold for both B^0 systems [27], and expanding in terms of Γ_{12}/M_{12} and $\Delta\Gamma/\Delta m$ we get

$$\Delta m = 2|M_{12}| \left[1 + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^2\right)\right], \tag{2.63}$$

$$\Delta\Gamma = 2|\Gamma_{12}| \cos\phi \left[1 + \mathcal{O}\left(\left|\frac{\Gamma_{12}}{M_{12}}\right|^2\right)\right], \tag{2.64}$$

where

$$\phi = \arg\left(-\frac{M_{12}}{\Gamma_{12}}\right). \tag{2.65}$$

The lifetime difference, $\Delta\Gamma$, is zero only if $\Gamma_{12} = 0$. In that case Eq. 2.53 gets reduced to

$$\frac{q}{p} = -\frac{M_{12}^*}{|M_{12}|} = -e^{-i\phi_M} \qquad \text{where} \qquad M_{12} = |M_{12}|e^{i\phi_M} \tag{2.66}$$

and ϕ_M is the phase of M_{12}, and hence $|q/p| = 1$.[10] Eq. 2.57 gets modified to

$$\begin{aligned} g_+(t) &= e^{-iMt}e^{-\Gamma t/2}\cos(\Delta m t/2), \\ g_-(t) &= e^{-iMt}e^{-\Gamma t/2} i \sin(\Delta m t/2). \end{aligned} \tag{2.67}$$

The symmetry between the decay probabilities for the B^0 and $\bar{B}^0$ mesons is once again regained and

$$\mathcal{P}_{u,m}(t) = \frac{\Gamma e^{-\Gamma t}}{2}(1 \pm \cos\Delta m t) \tag{2.68}$$

where the + sign corresponds to $\mathcal{P}_u(t)$ while the − sign corresponds to $\mathcal{P}_m(t)$. We also see that the oscillation frequency is equal to Δm, the mass difference between the two mass eigenstates.

Now, while $\Delta\Gamma = 0$ is known to be a good approximation for $B^0_d - \bar{B}^0_d$ mixing, theoretical predictions are more uncertain for $B^0_s - \bar{B}^0_s$ mixing [29]. The latest SM expectation is $\Delta\Gamma/\Gamma = 0.12 \pm 0.05$ for the B^0_s system. Experimentally, a non-negligible $\Delta\Gamma$ would lead to two distinct lifetime distributions for the light and heavy B^0_s mass eigenstates. For $\Delta\Gamma/\Gamma \sim 12\%$ the effect on the fraction of mixed decays versus time is negligible, especially at small times. Therefore, we will asumme $\Delta\Gamma = 0$ for the $B^0_{d,s}$ mixing analyses and use a non-zero $\Delta\Gamma$ as a systematic uncertainty in the B^0_s mixing analysis.

[10] Sometimes it is useful to go beyond the leading approximation for $\frac{q}{p}$ — the relevant expression is:

$$\frac{q}{p} = -\frac{M^*_{12}}{|M_{12}|}\left[1 - \frac{1}{2}\mathcal{I}m\left(\frac{\Gamma_{12}}{M_{12}}\right)\right].$$

2.8 Mass Difference

In principle, the mass difference, Δm, can be calculated from the box diagrams in Fig. 2.5. The dominant contribution is due to the exchange of the virtual top quark and hence in this approximation, the result can be written as [27]:

$$\Delta m_q = 2|M_{12}^q| = \frac{G_F^2}{6\pi^2}\eta_B m_{B_q^0}\mathcal{B}_{B_q^0} f_{B_q^0}^2 M_W^2 S\left(\frac{m_t^2}{M_W^2}\right)|V_{tb}V_{tq}^*|^2. \tag{2.69}$$

where $q = d$ for the B_d^0 system and $q = s$ for the B_s^0 system. G_F is the Fermi constant and (at lowest order) is given by

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2}. \tag{2.70}$$

η_B is a pertubative QCD correction factor, $m_{B_q^0}$ is the B meson mass, and m_t is the mass of the top quark. The parameters $f_{B_q^0}$ and $\mathcal{B}_{B_q^0}$ are the B_q^0 decay constant and the "bag parameter", respectively. $S\left(\frac{m_t^2}{M_W^2}\right)$ is the Inami-Lim function, given by

$$S(x_q) = x_q\left(\frac{1}{4} + \frac{9}{4(1-x_q)} - \frac{3}{2(1-x_q)^2}\right) - \frac{3}{2}\frac{x_q^3 \log x_q}{(1-x_q)^3} \tag{2.71}$$

with $x_q \equiv m_t^2/M_W^2$.

Eq. 2.69 suggests that a measurement of Δm_d should allow the extraction of the CKM element V_{td}. Moreover, Δm_d has been precisely measured and the world average is [2]:

$$\Delta m_d = 0.502 \pm 0.007 \ \text{ps}^{-1}. \tag{2.72}$$

Unfortunately, large theoretical uncertainties in the non-pertubative QCD factors, $f_{B_q^0}$ and $\mathcal{B}_{B_q^0}$ dominate the extraction of V_{td} from Δm_d. Currently, Lattice QCD

calculations give a $15-20\%$ uncertainty [2]. This difficulty, however, is overcome if the B_s^0 mass difference, Δm_s, is also measured. The CKM matrix element, $|V_{td}|$, can then be extracted from the ratio of the oscillation frequencies of the B_s^0 and B_d^0 mesons:

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s^0}}{m_{B_d^0}} \xi^2 | \frac{V_{ts}}{V_{td}} |^2 \tag{2.73}$$

where $m_{B_s^0}$ and $m_{B_d^0}$ are the B_s^0 and B_d^0 masses, respectively, and $\xi^2 \equiv f_{B_s}^2 \mathcal{B}_{B_s} / f_{B_d}^2 \mathcal{B}_{B_d}$. Many of the theoretical uncertainties cancel out in the ratio and ξ has been estimated from Lattice QCD calculations to be $1.21 \pm 0.022^{+0.035}_{-0.014}$ [30]. Therefore, the ratio V_{ts}/V_{td} can be extracted from measurements of Δm_d and Δm_s with a relatively small uncertainty of about 5%.

Furthermore, if we use the approximation $|V_{ts}| \sim |V_{cb}|$ (fairly well known from B decay measurements - see Table 2.4) and Eq. 2.39 we can see that the ratio $\Delta m_d / \Delta m_s$ determines R_t, the right side of the unitarity triangle. The above has motivated many experiments to search for B_s^0 oscillations and though a statistically significant signal has not been observed yet, lower limits on Δm_s have been set. The latest combined world limit is $\Delta m_s > 14.4$ ps^{-1} at 95% C.L. [2]. These lower limits are useful too since Eq. 2.73 indicates that a lower limit on Δm_s corresponds to an upper limit on V_{td}. Rearranging Eq. 2.73 and substituting the Wolfenstein approximations to the CKM matrix elements, we obtain

$$\Delta m_s = \Delta m_d \xi^2 \frac{m_{B_s^0}}{m_{B_d^0}} \frac{1}{\lambda^2} \frac{1}{(1-\bar{\rho}^2) + \bar{\eta}^2}. \tag{2.74}$$

Therefore, the limits on Δm_s can also constrain the allowed region of the $\bar{\rho} - \bar{\eta}$ plane. This can be seen in Fig. 2.6 where the dotted curve shows how the limit on Δm_s can improve on the constraint due to Δm_d [31]. The most recent fits to the Unitarity Triangle result in $\bar{\rho} = 0.214 \pm 0.047$ and $\bar{\eta} = 0.343 \pm 0.028$.

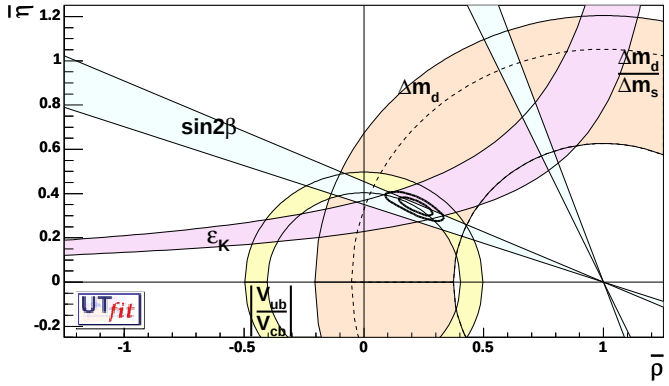


Figure 2.6: *Experimental constraints on the $\bar{\rho} - \bar{\eta}$ plane. The dotted curve corresponds to the 95% upper limit obtained from the experimental study of $B_s^0 - \bar{B}_s^0$ oscillations [31].*

Moreover, by using other measurements and constraints ($|V_{ub}/V_{cb}|$, $|\epsilon_K|$, and Δm_d), it is possible to obtain a SM "prediction" for Δm_s [22]. These results are summarized in Table 2.5. Δm_s values greater than 34.2 ps^{-1} would indicate physics beyond the SM at the 95% confidence level. Some new phyiscs (NP) models with large contributions to $B_s^0 - \bar{B}_s^0$ mixing are discussed in the next section.

Parameter	68%	95%	99%
Δm_s (including Δm_s measurements) ps^{-1}	$17.8^{+6.7}_{-1.6}$	$17.8^{+15.2}_{-2.7}$	$17.8^{+22.1}_{-3.7}$
Δm_s (without Δm_s measurements) ps^{-1}	$16.5^{+10.5}_{-3.4}$	$16.5^{+17.7}_{-5.7}$	$16.5^{+23.9}_{-7.2}$

Table 2.5: *Central values and ranges for Δm_s corresponding to different levels of probability obtained by including (or not) the experimental information from different searches [22].*

2.9 New Physics in the B_s^0 system

There are a large number of models of new physics which can contribute to $B_s^0 - \bar{B}_s^0$ mixing and consequently also affect $\mathcal{CP}$ asymmetries in B decays. The $B_s^0 - \bar{B}_s^0$ mixing amplitude is a CKM-suppressed fourth order weak interaction process and hence very sensitive to the intervention of non-SM physics. Moreover, in the SM, mixing induced $\mathcal{CP}$ asymmetries in B_s^0 decays almost vanish because of the small mixing phase ($\phi^{SM} = \arg(M_{12}) = \arg(V_{tb}V_{ts}^*) \approx 0$). Therefore, any sizable mixing induced $\mathcal{CP}$ violating effects in B_s^0 decays would almost surely come from non-SM processes.

Four of the many models that give new contributions to $B_s^0 - \bar{B}_s^0$ mixing are discussed below. For the purpose of this discussion, it is useful to define the ratio R between the B_s^0 and B_d^0 mixing frequencies, where for the case of the SM

$$R^{\mathrm{SM}} = \frac{\Delta m_s}{\Delta m_d}, \tag{2.75}$$

and is given by Eq. 2.73. Additionally, the new physics contribution to $B_s^0 - \bar{B}_s^0$ mixing can be parameterized using the parameters h_s and σ_s, which denote the magnitude and phase of the new contributions, normalized by the SM amplitudes. The prediction for Δm_s is then modified using

$$\Delta m_s^{\mathrm{NP}} = |1 + h_s e^{29\sigma_s}| \Delta m_s^{\mathrm{SM}} \tag{2.76}$$

- Supersymmetry (SUSY)

  There are several new contributions to $B_s^0 - \bar{B}_s^0$ mixing in the supersymmetric standard model (SSM) [32]. These include box diagrams with internal (i) charged Higgs bosons and charge 2/3 quarks, (ii) charginos and charge 2/3

squarks, (iii) gluinos and charge -1/3 squarks, and (iv) neutralinos and charge -1/3 squarks. In the minimal extension of the standard model (minimal supersymmetric standard model or MSSM) these new contributions have the same phase as in the SM, to a good approximation. The contributions to $B_q^0 - \bar{B}_q^0$ mixing are proportional to $(V_{tb}^* V_{tq})^2$, as in the SM, and $R^{\rm NP} = R^{\rm SM}$. The strongest constraints on the SUSY parameter space come from direct searches. For example, two of these searches put lower bounds of 250 GeV and 94 GeV on squark and chargino masses [2].

On the other hand, in some non-minimal SUSY models [33, 34], new physics effects are much larger and there can be new phases in their contribution to $B_s^0 - \bar{B}_s^0$ mixing. In a class of models described in Ref. [35] up to $\mathcal{O}(30\%)$ new physics effects (relative to the SM) are allowed inspite of no obvious inconsistency in the fit to the CKM triangle (Fig. 2.4). Measurements of $\mathcal{CP}$ asymmetry parameters in $B \to \phi K_S$, $B \to \eta' K_S$, and $B \to K\pi$ decays can, however, put stringent bounds on h_s and σ_s — the magnitude and phase of the contributions (Eq. 2.76). Figure 2.7 shows the allowed region in the $h_s - \sigma_s$ plane using the above constraints along with the lower limit on Δm_s.

- Multi-Higgs-doublet models [32]

  Models with more than one Higgs doublet are generally classified into two types: (i) models with natural flavor conservation (or NFC) in which there are no flavor-changing neutral currents and (ii) models in which flavor-changing interactions are mediated by neutral scalars.

  In models with natural flavor conservation, the new charged scalars contribute significantly to $B_s^0 - \bar{B}_s^0$ mixing if their masses lie in the range of 50 GeV to 1 TeV. There are two types of box diagrams involving charged Higgs bosons —

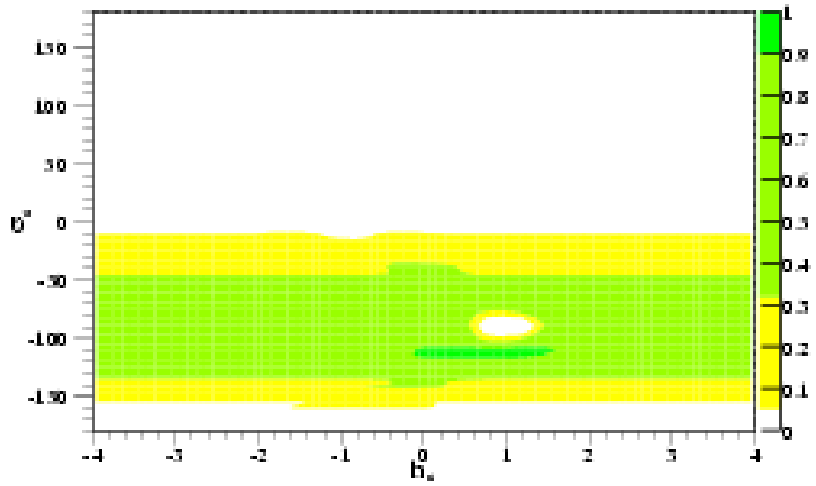


Figure 2.7: *The allowed region for $h_s - \sigma_s$ (magnitude and phase of new physics contributions, respectively) combining the present limits on Δm_s with bounds on σ_s from measurements of $\mathcal{CP}$ asymmetry parameters in $B \to \phi K_S$, $B \to \eta' K_S$, and $B \to K\pi$ decays [35].*

those with one H and one W, and those with two H's. The phase of $B_s^0 - \bar{B}_s^0$ mixing is unaffected by these contributions. Moreover, since the charged scalar couplings involve the same CKM factors as the SM case, $R^{\rm NP} = R^{\rm SM}$. Such models also predict that the branching ratios for $B_s^0 \to l^+ l^-$ and $B \to X_s l^+ l^-$ would be enhanced by up to an order of magnitude or by a factor of two, respectively. Hence, if these branching ratios are found to be consistent with the SM, that would indicate that the effects of such a model to $B_s^0 - \bar{B}_s^0$ mixing would be negligible.

For the second class of models, flavor-changing neutral scalar interactions exist and $B_s^0 - \bar{B}_s^0$ mixing receives large contributions from tree-level neutral Higgs exchange amplitudes which carry new phases. These neutral Higgs contributions can substantially modify the SM predictions for Δm_s. However, in these models

$R^{\rm NP} \approx R^{\rm SM}$ to a good approximation. Constraints on models of this type would come from $b \to \tau^+\tau^-$ and $B_s^0 \to l^+l^-$ decays, for which the branching ratios due to flavor changing neutral scalars are predicted to be a factor of 2-3 times larger than in the SM.

- Four generations [36]

  This is a model with an additional generation of quarks and leptons, including a new charge 2/3 quark, t'. The CKM matrix is a 4×4 matrix, parametrized by 6 angles and 3 phases, and the unitarity triangle gets transformed into a quadrangle. There are new loop-level contributions from t' box diagrams and new phases in the CKM matrix can modify $\mathcal{CP}$ asymmetries. Additionally, the new physics contribution to Δm_s and Δm_d could be different and $R^{\rm NP} \neq R^{\rm SM}$.

  There are severe constraints on this model, however, from measurements at LEP which require the fourth generation neutrino to have a mass $m_\nu > M_Z/2$. Since this is quite unlike the first three generations, this constraint makes the four-generation model much less plausible.

- Z-mediated flavor-changing neutral currents [37]

  These models have an additional charge -1/3 quark which mixes with the ordinary *down*-type quarks. Unitarity is violated and this induces flavor-changing neutral currents (FCNC) which are mediated at tree level by the Z boson. These Z-mediated couplings contribute to mixing and contain new, independent phases. These additional phases provide new sources for $\mathcal{CP}$ violation.

  The contribution of Z-mediated FCNC is different for $B_d^0 - \bar{B}_d^0$ and $B_s^0 - \bar{B}_s^0$ mixing leading to $R^{\rm NP} \neq R^{\rm SM}$. While $B_d^0 - \bar{B}_d^0$ mixing can be dominated by Z-mediated FCNC, the new physics contribution to $B_s^0 - \bar{B}_s^0$ mixing, though non-negligible, is not as significant. Experimental constraints on the FCNC

couplings come from the decay $B \to X_s\mu^+\mu^-$, and imply that their contribution to Δm_s is, at most, 25% of the SM contribution.

If there are indeed contributions to $B_s^0 - \bar{B}_s^0$ mixing from Z-mediated FCNC, there will be large effects in $b \to sl^+l^-$ and $B_s^0 \to l^+l^-$ decays as well as in $b \to s$ penguin processes.

It should be noted that a measurement of Δm_s can not only be used to test SM predictions but can also be used to distinguish between various new physics models. As shown above, different new physics models have different contributions to $B_s^0 - \bar{B}_s^0$ (and $B_d^0 - \bar{B}_d^0$) mixing. Ref. [38] illustrates how the B_s^0 mixing frequency can be used to differentiate amongst different SUSY models, as they have different predictions for the allowed range of Δm_s. The effect of a future measurement of Δm_s on constraining new physics parameter space can also be seen in Fig. 2.8. Using a hypothetical value of $\Delta m_s = 18.3 \pm 0.3$ ps^{-1}, the plot shows a dramatic reduction in the allowed region for $h_s - \sigma_s$ when compared to Fig. 2.7 which used only the lower limit on Δm_s [35]. Once again, this emphasizes the importance of a measurement of Δm_s, both in proving the validity of the SM, and in the search for new physics.

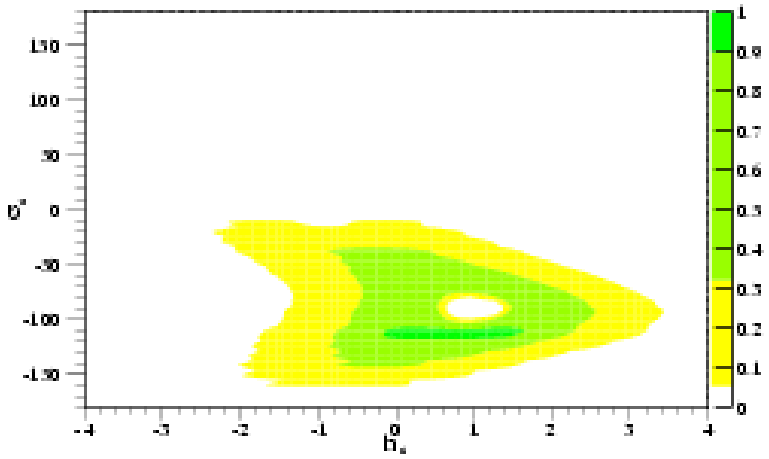


Figure 2.8: *The allowed region for $h_s - \sigma_s$ (magnitude and phase of new physics contributions, respectively) combining a future measured $\Delta m_s = 18.3 \pm 0.3\ ps^{-1}$ with bounds on σ_s from measurements of $\mathcal{CP}$ asymmetry parameters in $B \to \phi K_S$, $B \to \eta' K_S$, and $B \to K\pi$ decays [35].*

Chapter 3

Experimental Overview

Every mixing measurement is, at its heart, an asymmetry measurement. The asymmetry, $\mathcal{A}^{\rm meas}$, as a function of time is written as:

$$\mathcal{A}^{\rm meas}(t_{B_s^0}) = \frac{N^{\rm unm}(t_{B_s^0}) - N^{\rm mix}(t_{B_s^0})}{N^{\rm unm}(t_{B_s^0}) + N^{\rm mix}(t_{B_s^0})} \tag{3.1}$$

where $N^{\rm unm}$ and $N^{\rm mix}$ are the number of unmixed (non-oscillated) and mixed (oscillated) mesons, respectively, and $t^0_{B_s}$ is the time between the production and decay of the B_s^0 meson. A B_s^0 meson is said to have mixed if the sign of the b-quark inferred at the decay point is opposite to that inferred at production while for an unmixed B_s^0 meson the sign of the b-quark at decay is the same as that inferred at its production. Using Eq. 2.68, the above equation can be re-written as:

$$\mathcal{A}^{\rm meas}(t_{B_s^0}) = \frac{N^{\rm unm}(t_{B_s^0}) - N^{\rm mix}(t_{B_s^0})}{N^{\rm unm}(t_{B_s^0}) + N^{\rm mix}(t_{B_s^0})} = \cos(\Delta m_s t_{B_s^0}) \tag{3.2}$$

This analysis, therefore, is based on the simple observation that we can extract the mass difference, Δm_s, from the period of the oscillation.

The different elements essential to a B_s^0 mixing analysis are as follows:

- Production and selection of signal events;
- Proper time determination;
- Flavor tagging or determination of whether a B_s^0 meson has mixed or not and the error rate associated with this estimation;
- Fitting for Δm_s.

The above are discussed in the following sections.

3.1 Production and Selection of Signal Events

3.1.1 Parton model

At the Fermilab Tevatron, protons and anti-protons collide at a center-of-mass energy (usually denoted by $\sqrt{s}$) of 1.96 TeV. In such collisions a proton (or anti-proton) is considered to be composed of free partons — three constituent quarks ('valence quarks'), virtual gluons, and virtual quark-antiquark pairs ('sea quarks'). These partons do not necessarily divide the proton momentum equally amongst themselves. Parton distribution functions, $f_i^a(x)$, are used to denote the probability that a parton i (quark or gluon) carries a fraction x of the momentum of the beam particle a (proton or anti-proton). Moreover, if QCD effects are taken into account, then parton distribution functions also depend on the momentum exchanged in the interaction. Typically, gluons carry about half of the proton's momentum while for quarks $\langle x \rangle \sim$ 10% on average.

3.1.2 $b\bar{b}$ production

In leading order (LO) QCD approximation, heavy quarks are produced by the processes of quark-antiquark annihilation (as shown in Fig. 3.1a) and gluon-gluon fusion (as shown in Figs. 3.1b, 3.1c, and 3.1d). While top quarks are primarily produced via quark-antiquark annihilation at the Tevatron, the gluon-gluon fusion process is the dominant production mechanism for b quarks. In such events, b and $\bar{b}$ quarks are produced back-to-back, and they move away from each other with equal and opposite momenta[1] transverse to the $p\bar{p}$ beam axis.

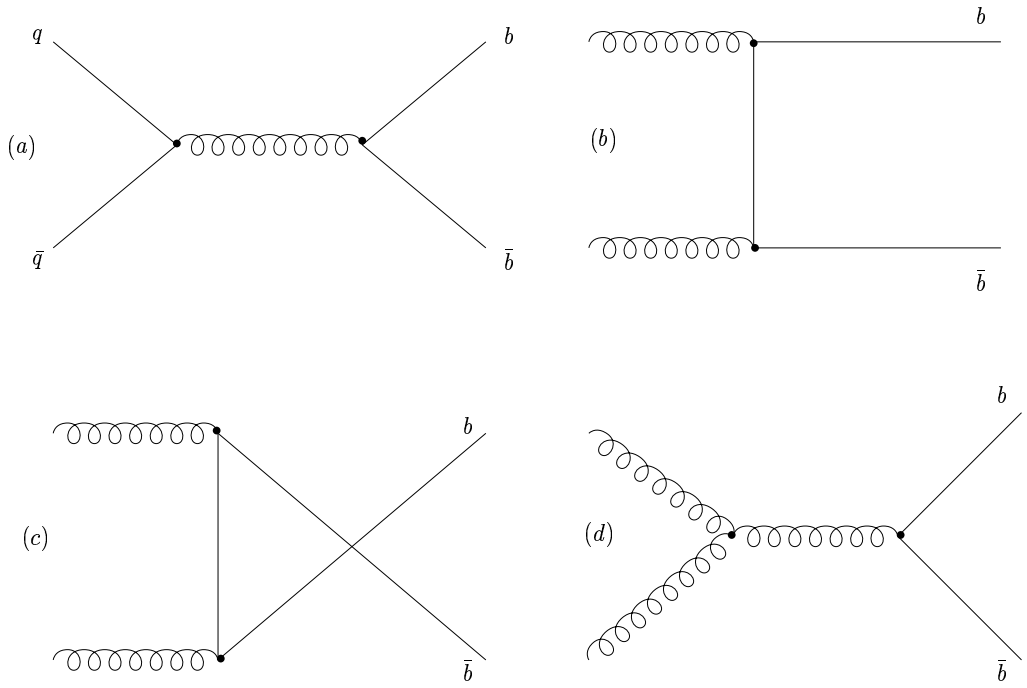


Figure 3.1: *Leading order diagrams for heavy quark production through (a) quark-antiquark annihilation, and (b)-(d) gluon-gluon fusion.*

[1]Actually, due to the motion of partons within the proton and anti-proton $q\bar{q}$ pairs are not necessarily always produced with equal and opposite momenta in the laboratory frame.

There can be significant contributions from higher order diagrams as well. These next-to-leading order (NLO) processes include real emission matrix elements (Fig. 3.2a), interference of virtual matrix elements with the LO diagrams (Fig. 3.2b), gluon splitting (Fig. 3.2c), and flavor excitation processes (Fig. 3.2d) [39]. Understanding of these higher-order production processes is important since the kinematic correlations between b and $\bar{b}$ quarks may be quite different in the two cases. The NLO processes could result in $b\bar{b}$ pairs that are close together in phase space with gluons taking a significant amount of the transverse energy while LO processes usually result in back-to-back pairs and no gluons.

3.1.3 b quark hadronization

After b quarks are produced via the initial hard scattering, the process of forming B hadrons follows and is called *hadronization* or *fragmentation*. The process is described by various theoretical models and a popular approach is the string fragmentation model [40]. Using a simple picture one can imagine a "cloud" of gluons acting as a string between the b and $\bar{b}$ quark pair. As the quark and antiquark move away from each other, the potential energy increases, and the string stretches until it breaks. A new $q'\bar{q}'$ pair is created out of the vacuum to form the new ends of the string and the system splits into two color singlet systems, $b\bar{q}'$ and $q'\bar{b}$. The two new strings also stretch and break until there is no longer sufficient energy for producing new $q'\bar{q}'$ pairs. The particles produced in this hadronization process, along with the B hadron are known as fragmentation particles. They usually show up in the detector as a *jet* of hadrons. An empirical model developed by Peterson *et al.* [41], based on kinematic arguments, parametrizes the fraction of the initial b-quark momentum transferred to

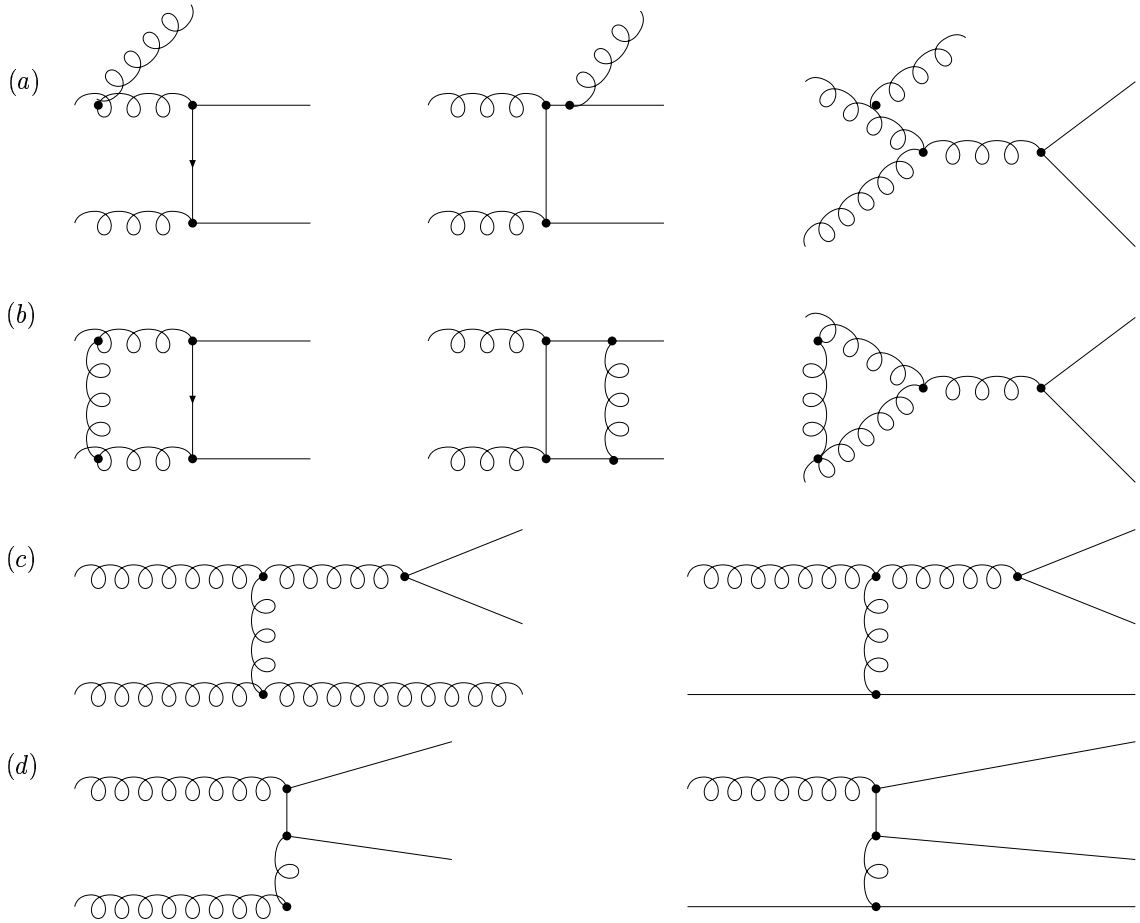


Figure 3.2: *Diagrams for heavy quark production at next-to-leading order: (a) real emission diagrams, (b) virtual emission diagrams, (c) gluon splitting, and (d) flavor excitation.*

the hadrons in the jet as

$$D_q^H(z) \propto \frac{1}{z}\left[1 - \frac{1}{z} - \frac{\epsilon_q}{1-z}\right]^{-2}, \tag{3.3}$$

where D_q^H is the fragmentation function (or the probability for a quark q to produce a hadron H), z is the fraction of the b quark momentum carried by the hadron, and ϵ_q is the so-called Peterson fragmentation parameter. ϵ_q is tuned to experimental

data and has been measured to be ≈ 0.006 for b-quarks and ≈ 0.06 for c-quarks [2]. On average, a B hadron carries approximately 70% of the momentum of its parent b-quark[2]. Charmed hadrons carry a lower fraction of their parent's momentum, $\approx 50\%$, and hence the number of fragmentation particles in a typical jet will be higher than for b jets since more energy is available to create them.

Furthermore, depending on whether the $\bar{b}$ quark initially created a $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$ or a diquark-antidiquark pair, B^+, B^0_d, B^0_s, B^+_c mesons and b-baryons (e.g. Λ_b) are produced, respectively. Since u and d are the lightest quarks, B^+ and B^0_d are produced in the largest numbers. B^0_s production is suppressed due to the larger mass of the s quark. b baryon production is also suppressed since it requires two quarks to be added to the b. B^+_c production is negligible compared to all these. The respective probabilities are given by

$$f_u : f_d : f_s : f_{\text{baryons}} : f_c \tag{3.4}$$

and have been obtained to be $f_u : f_d : f_s : f_{\text{baryons}} : f_c = 39.7 \pm 1.0\% : 39.7 \pm 1.0\% : 10.7 \pm 1.1\% : 9.9 \pm 1.7\%$: negligible, using the assumption that $f_u = f_d$ [2]. This shows that B^0_s mesons are produced approximately four times less often than B^0_d mesons, making a measurement of Δm_s more difficult.

3.1.4 Signal selection

There is a large $b\bar{b}$ production cross-section at the Tevatron ($\sim 100\ \mu$b) but the total inelastic $p\bar{p}$ cross-section is about three orders of magnitude larger. This makes the

[2]A quark and an anti-quark are most likely to combine into a hadron when they have similar velocity. Therefore, the heavy b-quark needs to lose only a small fraction of its energy in order to create a number of light $q'\bar{q}'$ pairs with similar velocity. Hence, when the b-quark combines with a $\bar{q}'$, the newly formed B hadron carries most of the energy of the b-quark.

identification of $b\bar{b}$ events experimentally very challenging. Since leptons are often produced in the decay of B mesons, either directly through the semileptonic decay chain ($b \to cl\nu$, where l is a lepton) or indirectly through sequential decays ($b \to c \to sl\nu$), the presence of high momentum ($p > 3$ GeV/c) muons or electrons can be used to obtain datasets enriched with events from $b\bar{b}$ production (i.e. "trigger" on interesting events). Additionally, these leptons can also be used for flavor tagging (see Sec. 3.3).

Another characteristic of the B meson that comes in handy when differentiating between signal and background is the relatively long lifetime of B mesons: $\tau(B^+) = 1.671 \pm 0.018$ ps, $\tau(B^0_d) = 1.536 \pm 0.014$ ps, and $\tau(B^0_s) = 1.461 \pm 0.057$ ps [2]. This coupled with the Lorentz boost from the initial momentum of the b quark causes the B hadrons to travel several millimeters before decaying. Reconstructing the B decay point or "secondary vertex" and requiring it to be separated from the $p\bar{p}$ interaction point or "primary vertex" further enriches the data with events from $b\bar{b}$ production.

There are a number of B^0_s decays that can be used for mixing analyses and these are broadly categorized into semileptonic decays ($B^0_s \to D^{(*)}_s l\nu X$) and hadronic decays ($B^0_s \to D^{(*)}_s n\pi$ where n is an integer indicating the number of pions in the final state). Note that a B^0_s meson almost always decays to a $D^{(*)}_s$ meson since the branching ratio $\mathcal{B}(B^0_s \to D^{(*)}_s X) \sim 100\%$. The notation $D^{(*)}_s$ here is used to represent both D_s mesons and their excited states like D^*_s and D^*_{s0}. Semileptonic decays have larger branching ratios (implying larger statistics) in comparison to hadronic decays and hence are used in this analysis. They have the additional advantage that the lepton in the final state can be used to select or trigger on the event. However, these decays do suffer from the fact that since there is a neutrino in the final state which escapes detection, the decays cannot be fully reconstructed. This leads to poorer

proper time resolution as will be seen in Sec. 3.2. Hadronic decays will, therefore, start playing a bigger role if Δm_s turns out to be large (> 22 ps^{-1} or so).

Two semileptonic $B_s^0 \to D_s^- \mu^+ X$ decays[3] are used for this analysis — the $D_s^- \to \phi\pi^-$ mode and the $D_s^- \to K^{*0}K^-$ mode. These decays have similar branching ratios ($D_s^- \to \phi\pi^-$: $\mathcal{B} = 3.6 \pm 0.9\%$ and $D_s^- \to K^{*0}K^-$: $\mathcal{B} = 3.3 \pm 0.9\%$) [2] and hence their reconstruction should yield comparable results. Details of the reconstruction and the selection criteria used are presented in Chapter 5.

3.2 Proper Time Determination

The proper lifetime of the B_s^0 meson, $ct_{B_s^0}$, is obtained from the measurement of the distance, L^B, between its production vertex and its decay vertex such that

$$ct_{B_s^0} = \frac{L^B}{\beta\gamma} = L^B \frac{M_{B_s^0}}{p(B_s^0)}. \tag{3.5}$$

β is the speed of the B_s^0 meson, γ is the Lorentz boost factor, and $M_{B_s^0}$ and $p_{B_s^0}$ are the mass and momentum of the B_s^0 meson, respectively. The above is projected in the plane transverse to the beam line since the transverse distance, L_{xy}^B, and the transverse momentum, $p_T(B_s^0)$, are measured more accurately than L^B and $p(B)$:

$$ct_{B_s^0} = L_{xy}^B \frac{M_{B_s^0}}{p_T(B_s^0)}. \tag{3.6}$$

In the case of semileptonic decays the full momentum of the B_s^0 meson cannot be reconstructed since the neutrino is undetected. Instead, the combined momentum of the $D_s\mu$ pair, $p_T(D_s\mu)$, is used to calculate the "visible proper decay length" or x^M

[3]Conjugate modes are implied throughout this dissertation.

given by

$$x^M \equiv L^B_{xy} \frac{M_{B^0_s}}{p_T(D_s\mu)}, \tag{3.7}$$

where L^B_{xy}, the measured transverse decay length, is defined as the displacement $\vec{X}^B_{xy}$ in the transverse plane between the secondary B^0_s vertex (V_B in Fig. 3.3) and the primary vertex (V_P in Fig. 3.3) projected onto the transverse momentum of the $D_s\mu$ system such that

$$L^B_{xy} \equiv \frac{\vec{X}^B_{xy} \cdot \vec{p}_T(D_s\mu)}{|\vec{p}_T(D_s\mu)|}. \tag{3.8}$$

Using Eqs. 3.7 and 3.8 the visible proper decay length or "VPDL" is expressed as

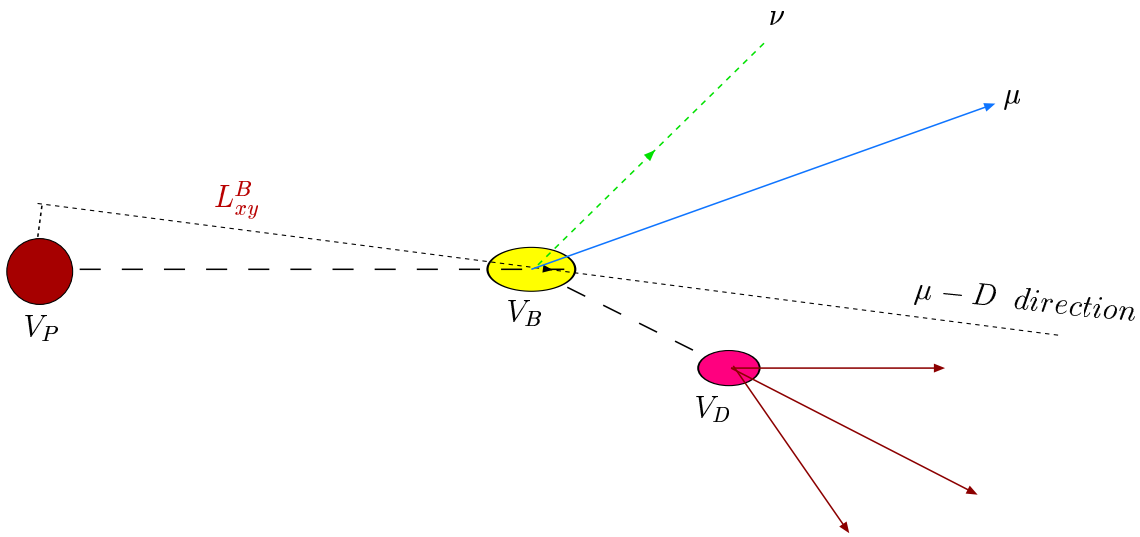


Figure 3.3: *Schematic representation of a* $B \to D\mu\nu$ *decay.*

$$x^M = \frac{\vec{X}^B_{xy} \cdot \vec{p}_T(D_s\mu)}{p_T(D_s\mu)} \cdot \frac{M_{B^0_s}}{p_T(D_s\mu)}. \tag{3.9}$$

The actual proper lifetime is then obtained using Eq. 3.9 and a correction factor such that

$$ct_{B_s^0} = x^M \cdot K. \tag{3.10}$$

$K \equiv p_T(D_s\mu)/p_T(B_s^0)$ and is known as the K-factor. It is essentially a correction factor used in semileptonic decays to account for the missing neutrino (and other neutral or non-reconstructed charged particles). K-factor distributions are obtained from Monte Carlo simulations and are further discussed in Chapter 5.

The uncertainty on the proper decay time, σ_t, can be expressed as:

$$\sigma_t = \sigma(L_{xy}^B) \oplus t \cdot \frac{\sigma(K)}{K}, \tag{3.11}$$

where $\sigma(L_{xy}^B)$ is the uncertainty due to vertexing resolution and $\sigma(K)/K$ is the K-factor resolution. Note that the latter uncertainty scales with the decay time t, while the vertexing resolution is independent of t and only adds a constant uncertainty. In order to resolve the fast B_s^0 oscillations σ_t should be smaller than the oscillation period. Moreover, since the K-factor resolution is significant for semileptonic decays (owing to the undetected neutrino), events with small decay time (or small VPDL) are the most sensitive to oscillations.

3.3 Flavor Tagging

Determination of whether a B_s^0 meson has mixed or not involves the following:

- Tagging the meson flavor (B_s^0 or $\bar{B_s^0}$) at decay time (or final state tagging);
- Tagging the meson flavor (B_s^0 or $\bar{B_s^0}$) at production time (or initial state tagging).

The different techniques used for this purpose are described below.

The flavor of a neutral B meson at decay is given by the charges of its decay products. The most common technique, which is adopted in this thesis, locates a lepton produced in the decay. The charge of the lepton (μ in this case) is used to obtain the flavor of the B_s^0 meson at decay — a negative muon corresponds to a b quark ($b \to \mu^-$), and vice versa. Sequential decays ($b \to c \to \mu^+$) can complicate the issue by making a wrong tag: however, their contribution can be greatly reduced with very simple kinematic cuts. In particular, the momentum spectrum of the sequential leptons is softer than that of direct leptons since they arise further down the decay chain.

The methods for tagging the initial state can be grouped into two categories: the ones that tag the initial charge of the b quark in the B_s^0 candidate itself (same-side tag), and the ones that tag the flavor of the other b quark in the event (opposite side tag).

The same-side tagging techniques use the correlations between the charge of the initial b or $\bar{b}$ quark and the charges of the fragmentation particles around the B meson direction (or the decay products of B^{**} decays). The fragmentation track closest to the B meson has a charge correlated with the charge of the b quark (Fig. 3.4). Positive pions are associated with B^- and B_d^0 mesons while negative pions are associated with B^+ and $\bar{B}_d^0$ mesons (positive kaons are correlated with B_s^0 mesons while negative kaons are associated with $\bar{B}_s^0$ mesons). The correlation is lost if the fragmentation track is neutral.

The opposite side tagger makes use of the fact that the dominant mode of b quark production at the Tevatron is back-to-back $b\bar{b}$ or that the $b/\bar{b}$ quark of interest (reconstructed B meson) is always produced along with another $\bar{b}/b$ quark of opposite

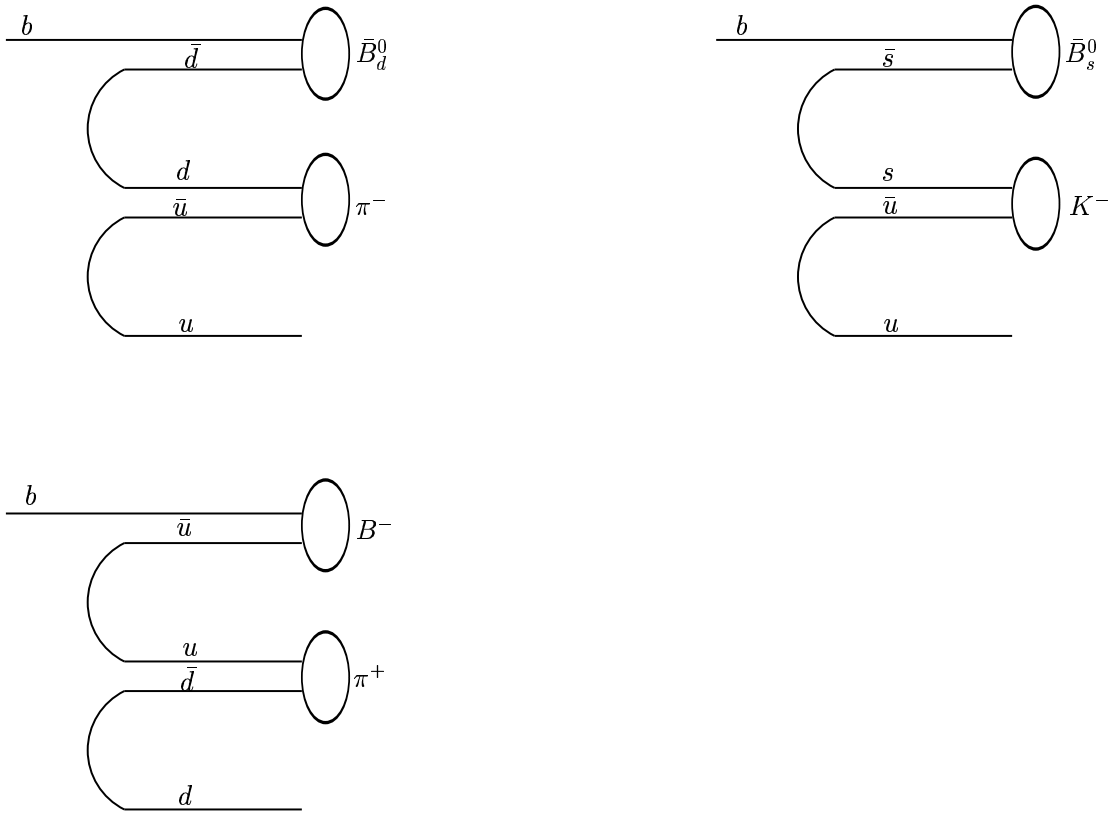


Figure 3.4: *Possible fragmentation scenarios for a b quark.*

charge. The other $\bar{b}/b$ in the event is called the "opposite side b". Determining the flavor of the opposite side B meson allows us to infer the flavor of the reconstructed B meson.

The opposite side lepton tagger, for example, relies on identifying the flavor of the other B in the event using the sign of the lepton it decayed to — a negative lepton corresponds to a b quark, and vice versa. For reconstructed $B^0_s \to D^-_s \mu^+ X$ decays both leptons having the same sign would indicate that one B hadron had oscillated while opposite signs would indicate that neither (or both) had oscillated. Another tagging technique determines the flavor of the opposite-side b hadron by analyzing

the jet associated with it. A momentum-weighted charge distribution (of all tracks in the jet) is used to form a variable to discriminate between b and $\bar{b}$ quarks. Details of the tagging techniques used for this analysis are presented in Chapter 5.

The figure of merit typically used to compare different tagging algorithms is the "tagging power" or $\epsilon\mathcal{D}^2$ where ϵ is the tagging efficiency (or rate), and $\mathcal{D}$ is the "dilution" given by:

$$\epsilon \equiv \frac{N_{\text{correct}} + N_{\text{wrong}}}{N_{\text{correct}} + N_{\text{wrong}} + N_{\text{no tag}}} \quad \text{and} \quad \mathcal{D} \equiv \frac{\mathrm{N}_{\text{correct}} - \mathrm{N}_{\text{wrong}}}{\mathrm{N}_{\text{correct}} + \mathrm{N}_{\text{wrong}}}, \tag{3.12}$$

where N_{correct} (N_{wrong}) is the number of events that have been correctly (incorrectly) tagged and $N_{\text{no tag}}$ is the number of events that do not have a tag.

Eq. 3.12 indicates that a tagging algorithm with a large dilution characterizes a more powerful tagging method than one with a smaller dilution. A large dilution is, therefore, desirable. This makes the term "dilution" counter-intuitive and its use can sometimes be misleading. It is, therefore, better to use the purity, η_s, of the tagging technique instead:

$$\eta_s \equiv N_{\text{correct}}/N_{\text{total tagged events}}, \tag{3.13}$$

where η_s is related to the dilution, $\mathcal{D}$, using the simple formula $\mathcal{D} \equiv 2\eta_s - 1$.

As mentioned earlier, sequential or cascade decays ($b \to c \to l^+$) can result in a wrong tag. Additionally, a misidentified fake lepton can provide a random tag. Therefore, Eq. 3.2 has to be modified to take this into account. Defining η_s as the fraction of correctly tagged events, the number of events (correctly) tagged as unmixed

and mixed are given by:

$$N_{\rm unm}(t) = \eta_s N(t)_{B_s^0 \to B_s^0} + (1-\eta_s) N(t)_{B_s^0 \to \bar{B_s^0}}, \quad (3.14)$$

$$N_{\rm mix}(t) = (1-\eta_s) N(t)_{B_s^0 \to B_s^0} + \eta_s N(t)_{B_s^0 \to \bar{B_s^0}}. \quad (3.15)$$

Using the above equations and Eq. 2.68, the measured asymmetry in Eq. 3.2 gets modified to

$$\mathcal{A}^{\rm meas} = (2\eta_s - 1)\cos\Delta m_s t = \mathcal{D}\cos\Delta m_s t. \quad (3.16)$$

This results in the amplitude of the oscillation being less than one owing to imperfect flavor tagging. In principle, if $\mathcal{A}^{\rm meas}$ is fitted with a cosine, the amplitude gives the dilution while the frequency of the oscillation gives Δm_s. If no obvious oscillations are seen, a different method (described in Sec. 3.4) is used to exclude certain values of the oscillation frequency.

For the analysis presented in this thesis we concentrate on the opposite side tagging algorithms since in that case η_s or $\mathcal{D}$ for the B mesons can be obtained from the (relatively easier) Δm_d measurements and can be used directly for the B_s^0 analysis. This utilizes the fact that the charge of the b quark on the opposite side should not be affected by whether there is a B_d^0 or a B_s^0 meson on the reconstructed side.

3.4 Fitting For Δm_s

First, an expected asymmetry, A^e, is calculated keeping in mind the K factor correction, the VPDL resolution, and the fact that there can be contamination of the B_s^0 sample by mesons that either do not mix ($B^{\pm}$) or mix at a different rate (B_d^0). This is further discussed in Chapter 5.

Then a time dependent asymmetry between unmixed and mixed B mesons is obtained. This is done by producing D_s^- mass distributions for different VPDL bins, for both the unmixed and mixed event samples and then determining the numbers of unmixed and mixed B mesons for each bin by fitting the distributions to functions describing the signal and background contributions.

The experimental observable, asymmetry $A_i^{\rm meas}$ in each VPDL bin, i, is defined as:

$$A_i^{\rm meas} = \frac{N_i^{\rm unm} - N_i^{\rm mix}}{N_i^{\rm unm} + N_i^{\rm mix}}, \tag{3.17}$$

where $N_i^{\rm unm}$ is the number of events tagged as "non-oscillated" in VPDL bin i and $N_i^{\rm mix}$ is the number of events tagged as "oscillated" in the same VPDL bin. The exact procedure for obtaining this asymmetry and the results are discussed in Chapter 5.

For the case no oscillation signal is observed, a technique called the *amplitude fit method* [42] can then be used to set limits on Δm_s. This method is essentially equivalent to a Fourier analysis, in which one searches for peaks in the frequency spectrum of oscillations. This technique also allows results from different experiments to be combined in a straightforward manner. The amplitude fit method works as follows.

The expressions describing the unmixed and mixed probabilities (Eq. 2.68) are modified by introducing an amplitude $\mathcal{A}$:

$$\mathcal{P}_{u,m}(t) = \frac{\Gamma e^{-\Gamma t}}{2}(1 \pm \mathcal{A} \cdot \mathcal{D} \cos \Delta m_s t). \tag{3.18}$$

The frequency of the oscillation is not taken to be a free parameter but is instead fixed to a "test" value ω. The new auxiliary parameter, the amplitude $\mathcal{A}$ of the oscillating term, is left free in the fit. The fitted values of $\mathcal{A}$ (and its error $\sigma_{\mathcal{A}}$) as a function of

ω can then be determined from the minimization of a $\chi^2(\mathcal{A})$ defined as:

$$\chi^2(\mathcal{A}) = \sum_i \frac{(A_i^{\rm meas} - A_i^e(\mathcal{A}))^2}{\sigma^2(A_i^{\rm meas})}. \tag{3.19}$$

When the test frequency is much smaller than the true frequency ($\omega \ll \Delta m_s$), the expected value for the amplitude is $\mathcal{A} = 0$; while at the true frequency ($\omega = \Delta m_s$) the expectation is $\mathcal{A} = 1$ within its total uncertainty $\sigma_{\mathcal{A}}$. All values of the test frequency ω for which $\mathcal{A} + 1.645\sigma_{\mathcal{A}} < 1$ can be excluded at 95% Confidence Level.

It is useful at this point to define the statistical significance (or the sensitivity), $\mathcal{S}$, of a B_s^0 oscillation signal. This can be approximated as [27]:

$$\mathcal{S} = \sqrt{\frac{\epsilon\mathcal{D}^2}{2}} \frac{S}{\sqrt{S+B}} e^{-(\Delta m_s \sigma_t)^2/2}, \tag{3.20}$$

where $\epsilon\mathcal{D}^2$ is the tagging power as described previously, σ_t is the proper time resolution and S and B are the numbers of signal and background events, respectively. The sensitivity $\mathcal{S}$ decreases rapidly as Δm_s increases. This dependence is primarily controlled by σ_t and this reiterates that excellent proper time resolution is needed to resolve high frequency B_s^0 oscillations.

Proper time resolution also results in the error on the amplitude, $\sigma_{\mathcal{A}}$, being an increasing function of the test frequency ω. It is, therefore, expected that individual values of ω can be excluded up to $\Delta m_s^{\rm sens}$, where $\Delta m_s^{\rm sens}$ is called the sensitivity of the analysis defined by $1.645\sigma_{\mathcal{A}}(\Delta m_s^{\rm sens}) = 1$. Since lower limits are susceptible to statistical fluctuations, the sensitivity of an analysis is a better measure of its significance.

An additional nice feature of the amplitude method is that the results from different analyses and experiments can be combined (after accounting for correlations

between systematic errors) by simple averaging of different amplitude spectra.

It should also be noted that deriving Δm_s (or a limit on Δm_s) from a binned asymmetry is not as accurate as using event-by-event information about the decays. However, it is significantly less sensitive to affects from non-oscillating backgrounds and lifetime dependent decay selection criteria and is therefore appropriate for this initial measurement.

Chapter 4

The Experimental Apparatus

4.1 Tevatron

The Fermi National Accelerator Laboratory (Fermilab) houses the most powerful collider in the world. A series of accelerators are used within the laboratory complex to create the world's highest energy particle beams. Figure 4.1 shows an overview of the Fermilab accelerator complex. For a detailed description the reader is referred to Ref. [43], only a brief review is presented here.

The first stage of the acceleration is provided by the Cockcroft-Walton pre-accelerator. Gaseous hydrogen is ionized inside this device to produce H^- ions. These ions are then accelerated by a positive voltage to an energy of 750 keV. The negative hydrogen ions continue on to a linear accelerator, the Linac. This accelerator is 500 feet long and uses oscillating electric fields to accelerate H^- ions to 400 MeV. Before entering the third stage, the Booster, these ions pass through a carbon foil which removes the two electrons, leaving only the positively charged proton.

The Booster, located approximately twenty feet below the ground, is a circular

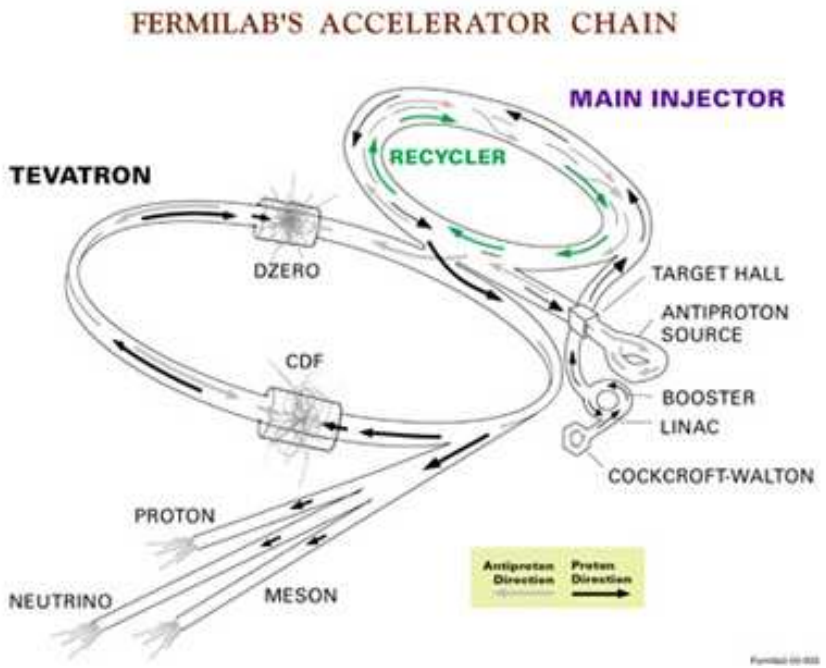


Figure 4.1: *An overview of the Fermilab accelerator complex.*

accelerator that uses magnets to bend the beam of protons in a circular path. The protons travel around the Booster about 20,000 times and repeatedly experience electric fields such that with each revolution they pick up more energy. The protons finally leave the Booster with an energy of 8 GeV and enter the Main Injector.

The Main Injector accelerates protons from 8 GeV to 150 GeV. It also produces 120 GeV protons that it sends to the Antiproton Source for anti-proton production. Inside the Antiproton Source the 120 GeV protons are collided with a nickel target. Anti-protons are among the secondary particles produced in this collision. The anti-protons are transported to the Debuncher where they are focused and reduced to the same low emittance as the protons via a process known as stochastic cooling [44]. They are then stored in the Accumulator ring and when a sufficient number ("stack") has been obtained, they are sent to the Main Injector for further acceleration. The

Main Injector accelerates these anti-protons to 150 GeV and then injects them, in a bunch structure, (along with bunches of 150 GeV protons) into the Tevatron. The Tevatron beam typically consists of 36 proton and anti-proton bunches arranged in three groups of 12, named super-bunches. There is 2 μs between each super-bunch and 396 ns between each bunch inside a super-bunch. The typical number of particles in a proton bunch is $N_p \sim 2.7 \times 10^{11}$, while in an anti-proton bunch it is $N_{\bar{p}} \sim 3 \times 10^{10}$.

The 150 GeV protons and anti-protons enter the Tevatron ring ($\sim$ 4 miles in circumference and buried 20 feet underground) circling in opposing directions and are accelerated to almost 1 TeV (actually 980 GeV) in the 4 Tesla magnetic field of superconducting dipole magnets[1]. The bunches cross each other every 396 ns at six points along the ring and at two of these, B0 and D0, are located two general-purpose detectors — CDF (Collider Detector Facility) and DØ. Quadrapole magnets squeeze the proton and anti-proton beams into a cross-sectional area of $\sigma_a \sim 5 \times 10^{-5}$ cm^2 such that the beams collide in the geometrical center of the two detectors.

The instantaneous luminosity[2] is given by

$$\mathcal{L} = \frac{N_p N_{\bar{p}} n_B f}{4\pi\sigma_a^2}, \tag{4.1}$$

where n_B is the number of bunches and f is the bunch revolution frequency ($\sim$ 50 kHz). The duration during which proton and anti-proton beams circulate in the Tevatron is called a "store" and each store typically lasts for several hours, sometimes

[1]The Tevatron also operates in fixed target mode in which case the proton beam is accelerated to 980 GeV, extracted and sent down the Fixed Target beam line to the experimental areas. In this case, anti-protons are not produced in the Main Injector.

[2]Luminosity is a measure of the "intensity" of the beam, determined by the density of particles in the beam. Instantaneous luminosity gives a measure of the rate of collisions at any given time while integrated luminosity gives a measure of the total number of collisions over a given time period. Integrated luminosity is typically expressed in inverse pico-barns (pb^{-1}) or inverse femto-barns (fb^{-1}), where 1 barn $= 1 \times 10^{-24}$cm^2.

even for more than a day. Figure 4.2 shows the typical length of stores over a period of a few months. The instantaneous luminosity is highest at the beginning of a store and decreases as anti-protons are lost to collisions and beam instabilities. Events collected during each store are grouped into "runs". Most parameters of the detectors' operation (the beam position for example) are stored in databases in run-averaged format. More recently, in order to increase the instantaneous luminosity and the anti-

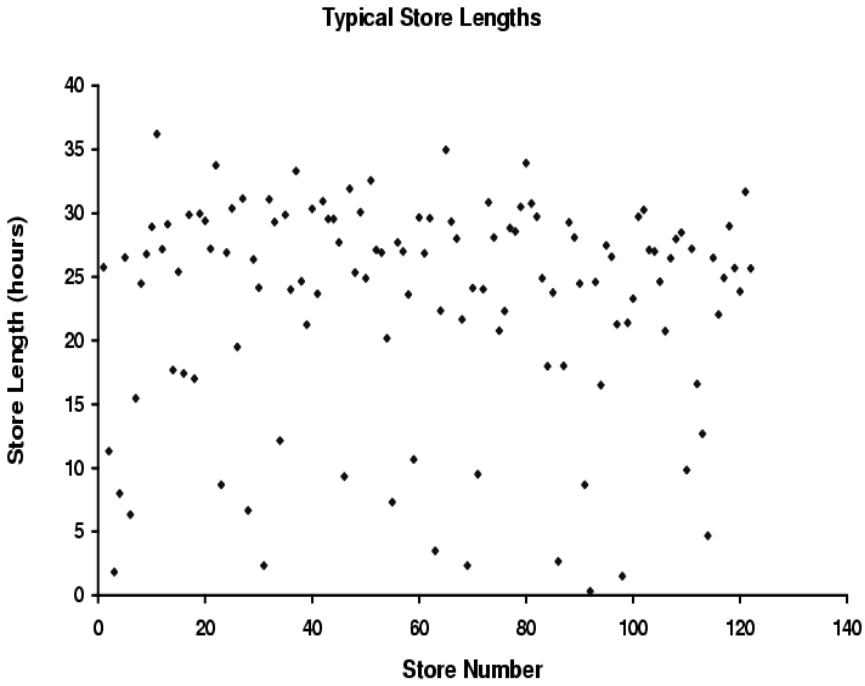


Figure 4.2: *Typical length of stores at the Tevatron.*

proton content, an Antiproton Recycler (inside the Main Injector) is being used to store anti-protons that return from a trip through the Tevatron, and which are then re-injected. The Recycler uses the novel method of electron cooling [45] to cool the anti-protons, i.e. reduce the emittance of the anti-proton beam and hence increase the luminosity of the Tevatron.

4.2 The DØ Detector

This section provides a brief overview of the DØ detector. The contents of this section are primarily derived from Ref. [46].

The DØ detector, like other high energy physics detectors, is composed of different sub-systems that surround each other like the layers of an onion. Figure 4.3 shows a schematic view of the detector. The different sub-systems that are described in the following sections are:

- Central Tracking System (Sec. 4.2.2),
- Calorimeter (Sec. 4.2.3),
- Muon System (Sec. 4.2.4),
- Luminosity Monitor (Sec. 4.2.5).

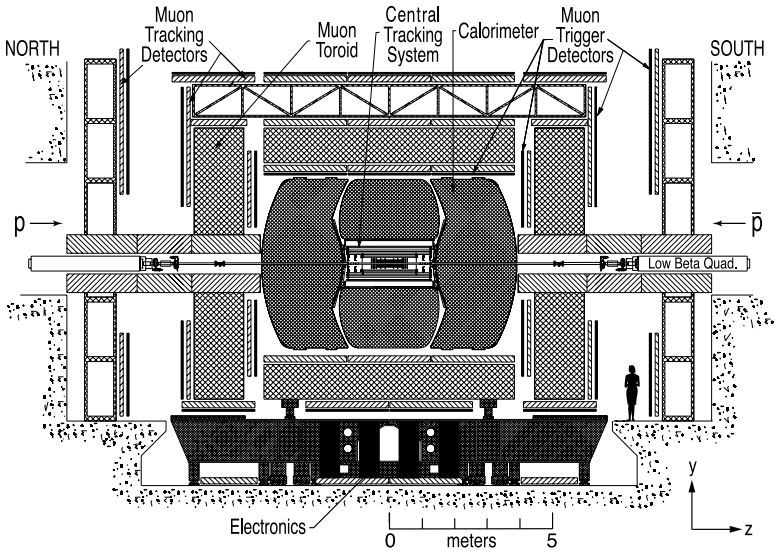


Figure 4.3: *Schematic drawing of the DØ detector [46].*

4.2.1 The coordinate system

DØ uses a right-handed coordinate system. The direction of the proton beam defines the positive z-axis. The y-axis points upwards and the x-axis points towards the center of the Tevatron ring with the DØ detector centered at $(0, 0, 0)$. The azimuthal (ϕ) and polar (θ) angles are also used and are defined as:

$$\begin{aligned} \phi &= \tan^{-1}\left(\frac{y}{x}\right), \\ \theta &= \tan^{-1}\left(\frac{r}{x}\right), \end{aligned} \tag{4.2}$$

where r is the perpendicular distance from the beam axis. Since the partons participating in the collision carry a varying amount of their parent hadron's momentum (determined by the parton distribution function), physics interactions often have large boosts along the beam direction. Moreover, many particles produced in the collision, i.e. the remnants of the proton not participating in the hard-scattering interaction, escape down the uninstrumented beam pipe. Hence the observed momenta projected to the beam axis do not equal zero. However, the total momentum in the plane transverse to the beam direction, (p_x, p_y), does equal zero to a good approximation[3] and its undetectable fraction (owing to limitations in detector coverage) is negligible. In this plane particles are represented using polar coordinates and transverse momentum and energy are defined (in a direction perpendicular to the beam) as:

$$p_T = p \sin\theta \quad \text{and} \quad E_T = E \sin\theta. \tag{4.3}$$

The polar angle θ is usually replaced by the *pseudorapidity*, η, which is defined

[3]This is only approximately true owing to the motion of partons within the proton (and anti-proton) and due to the fact that there can be (very small) beam tilts.

as

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) \tag{4.4}$$

which approximates the true rapidity

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) \tag{4.5}$$

for relativistic particles. Differences in rapidity are invariant under a Lorentz boost in the z direction, hence making it a more convenient coordinate.

4.2.2 Central Tracking System

The central tracking system (Fig. 4.4) consists of an inner high-resolution silicon microstrip tracker (SMT) surrounded by a scintillating central fiber tracker (CFT). Both detectors are immersed in a 2 Tesla solenoidal magnetic field. Surrounding this is a scintillator based preshower detector along with other preshower detectors that are mounted on the inner surfaces of the forward calorimeter cryostats. The tracking system was designed to provide momentum measurement using the magnetic field; good electron identification (and e/π separation); tracking over a large pseudorapidity range ($|\eta| < 3$); secondary vertex detection; triggering; and a fast response for a bunch crossing time of 396 ns.

Silicon Microstrip Tracker

Located closest to the interaction region, the SMT uses the silicon pn junction technology to create a detector that is capable of making position measurements on the

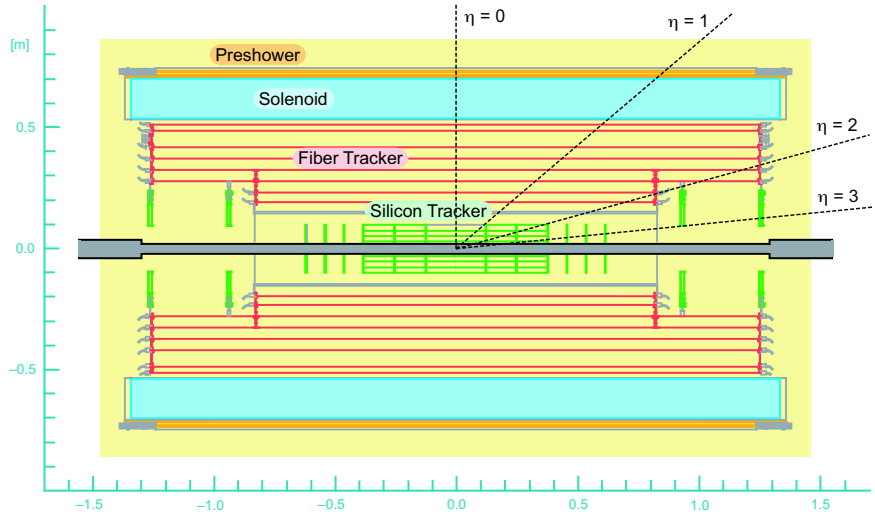


Figure 4.4: *Schematic drawing showing the DØ central tracking system.*

order of 10 μm in resolution. A brief description of the working of such a silicon detector follows (see Fig. 4.5). Finely spaced strips of strongly doped p-type silicon (p^+) are deposited on a lightly doped n-type (n^-) silicon substrate. On the other side, a thin layer of strongly doped n-type (n^+) silicon is deposited. A positive voltage is applied to the n^+ side and this depletes the n^- substrate of free electrons and creates an electric field in the volume of the n^- substrate. A charged particle that passes through the silicon ionizes and leaves a trail of electron and hole pairs. The holes drift to the p^+ strips producing an electric signal and these signals can be read by an integrated circuit at the end of the strip thus enabling the measurement of the position of the particle.

The Silicon Microstrip Tracker at DØ directly surrounds the beam pipe and was designed to maximize tracking for the full η acceptance. The large spread in z for the luminous region of the collisions ($\sigma_z \sim 25$ cm) complicated the task and the resulting design had to be a hybrid with both barrel and disk components (see Fig. 4.6). While the barrel detectors measure the $r-\phi$ coordinates, the disks measure both $r-\phi$ and

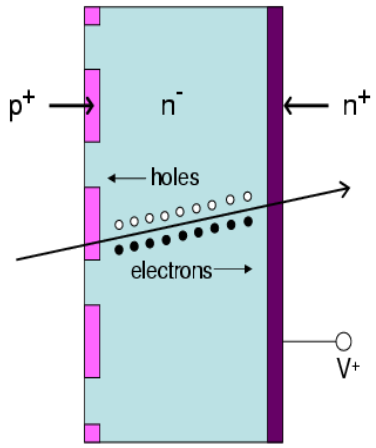


Figure 4.5: *Schematic drawing showing the working a generic silicon detector.*

$r-z$. There are six 12 cm long barrel detectors containing eight layers of rectangular silicon microstrip detectors, referred to as "ladders". Figure 4.7 shows the cross-sectional view of a SMT barrel. In the inner four barrels, layers 1, 2, 5 and 6 are double-sided[4] with axial strips on one side and 90° stereo angle strips on the other side, with pitches of 50 μm and 153.5 μm respectively. The outermost barrels have only single-sided ladders with 50 μm pitch axial strips in these layers, and hence provide no stereo information. Layers 3, 4, 7 and 8 of all the barrels are double-sided with axial strips of 50 μm pitch on one side and 2° stereo angle, 62.5 μm pitch strips on the reverse side. Layers 1-4 consist of 6 ladders each while layers 5-8 have 12 each.

There are 12 so-called "F-disks" made from double sided detectors with 50 μm pitch, $-15°$ stereo angle on one side and 62.5 μm pitch, $+15°$ stereo angle on the other. Four of the disks are sandwiched between the barrels while the remaining

[4]Double-sided detectors have the n^+ and p^+ strips offset at a stereo angle relative to each other allowing the reconstruction of tracks in three dimensions.

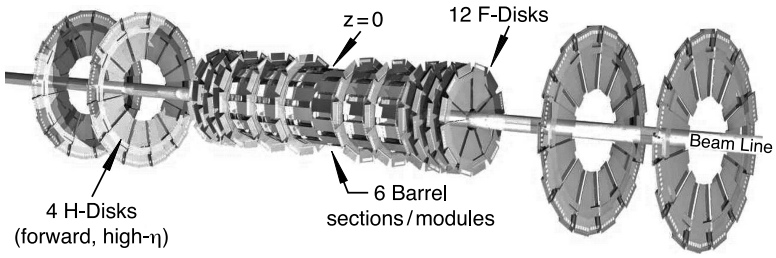


Figure 4.6: *The hybrid disk/barrel design of the Silicon Microstrip Tracker (SMT) [46].*

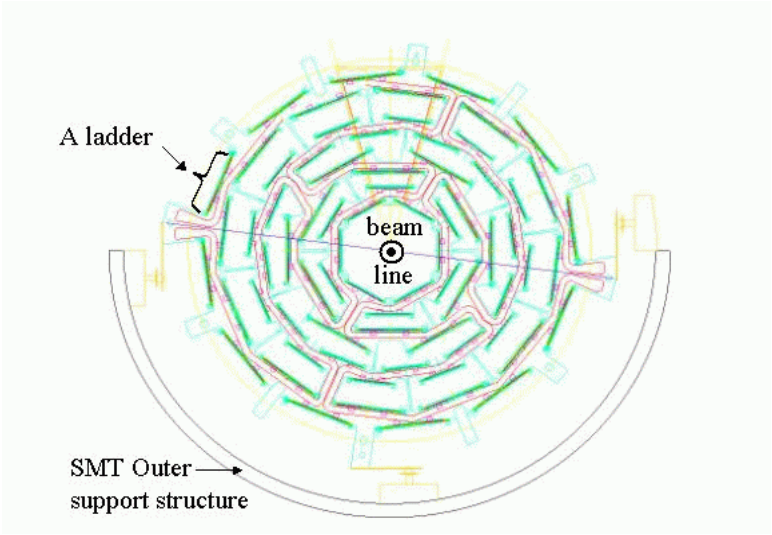


Figure 4.7: *Cross-sectional view of a SMT barrel.*

eight are located at both ends of the barrel section. The disks have an inner radius of 2.6 cm and an outer radius of 10.0 cm.

Towards the end of the interaction region, beyond the F-disks, are two "H-disks" located on each side. These disks are made of single-sided wafers with an inner radius of 9.5 cm and an outer radius of 26.0 cm. These four H-disks help to extend the SMT coverage to $|\eta| \sim 3$. Table 4.1 summarizes some of the salient features of the SMT barrels and disks.

	Barrels	F-Disks	H-Disks
Channels	387072	258048	147456
Modules	432	144	96
Silicon Area	1.3 m^2	0.4 m^2	1.3 m^2
Inner Radius	2.7 cm	2.6 cm	9.5 cm
Outer Radius	10.5 cm	10.0 cm	26 cm

Table 4.1: *An overview of the Silicon Microstrip Detector (SMT).*

The SMT is read out by 128-channel readout chips. These chips called SVXIIe chips are designed to work with double sided detectors and can accept both positive and negative currents as input signals. The chips are mounted on a high density interconnect or HDI. The data passes from the HDI via adaptor cards and interface boards to the sequencer boards. The sequencer boards are connected by means of an optical link to the readout buffer. The whole SMT comprises a total of $792,576$ readout channels.

Central Fiber Tracker

The central fiber tracker (CFT) surrounds the SMT and covers the central pseudo-rapidity region ($|\eta| < 2$). Its chief purpose is to reconstruct tracks and measure the momentum of charged particles within this region. It works on the principle that certain materials produce scintillation light (in or near the visible spectrum) when atomic electrons are excited by knock-on electrons produced by passing charged particles.

The basic element of the CFT is the scintillating fiber, consisting of a polystyrene core clad in an inner thin acrylic layer and an outer layer of fluoroacrylate. The claddings help improve the light trapping and the mechanical strength of the fiber. The nominal diameter of the clad fibers is 835 μm. The polystyrene core is doped with

1% paraterphenyl and 1500 ppm 3-hydroxyflavone. The paraterphenyl helps increase the light yield, while the hydroxyflavone acts as a wavelength shifter to match the transmission properties of polystyrene. The peak emission wavelength of these fibers is around 530 nm (yellow-green part of the visible spectrum). The tracker contains a total of 76,800 such fibers.

The fibers are formed into doublet layers and are mounted on eight concentric cylinders arranged in layers between $r = 20$ cm and $r = 52$ cm. Table 4.2 summarizes some of these parameters concerning the CFT. Each cylinder has an axial fiber doublet layer (providing an $r - \phi$ measurement) and a doublet layer with alternating stereo angles of $\pm 2^\circ$ (allowing three-dimensional reconstruction of tracks). The fibers are grouped into 4.5° sections (called 'sectors') in ϕ. There are a total of 80 such sectors for readout and triggering purposes. A cross-section of the CFT is shown in Figure 4.8.

Layer	Radius (cm)	# fibers	Fiber pitch (μm)
A	20.1	2560	985.606
B	25.0	3200	981.300
C	29.9	3840	978.105
D	34.8	4480	976.101
E	39.7	5120	974.598
F	44.6	5760	973.429
G	49.5	6400	972.297
H	51.5	7040	919.610

Table 4.2: *Summary of important Central Fiber Tracker (CFT) parameters.*

Scintillation light collected in the fibers is transmitted via internal reflection in both directions. At one end an aluminum mirror reflects the light back in the opposite direction. At the other end are optical fiber waveguides that conduct the light

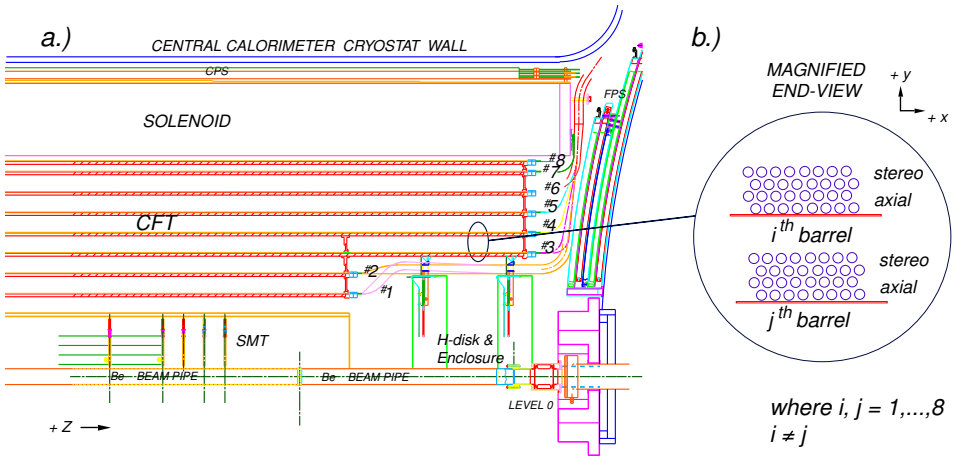


Figure 4.8: *Cross-sectional view of the Central Fiber Tracker (CFT).*

to Visible Light Photon Counters (VLPCs). The VLPCs are arsenic doped silicon diodes operating at temperatures of 8-10 K and they convert the collected photons to an electronic signal via electron-hole pair creation. They have excellent quantum efficiencies (greater than 75%), high gain (between 22,000 to 65,000 electrons per incoming photon), less than 0.1% average noise, and a position resolution of $\sim$ 100 μm.

Solenoid

The solenoid surrounding the tracking region is 2.73 m in length and 1.42 m in diameter. It is superconducting and creates a highly uniform axial magnetic field of 2 Tesla. The trajectory of a charged particle is bent by the magnet thereby allowing its momentum to be measured. The magnet operates at a current of 4749 A and stores 5.3 MJ of energy. The solenoid is constructed of two grades of superconducting multifilamentary Cu:NbTi cables stabilized with pure aluminum and operates at 4.7 K.

Preshower detectors

The preshower detectors help in electron identification and triggering and are also used offline to correct the electromagnetic energy measurement of the calorimeter for losses in the solenoid, lead absorber, and in material such as cables and supports. Figure 4.9 shows a cross-sectional view of the forward tracking region and indicates the location of the forward and central preshower detectors. The central preshower (CPS) is located in the 51 mm gap between the solenoid and the central calorimeter cryostat and covers a region of $|\eta| < 1.3$. There is a flat layer of lead, 5.5 mm thick (corresponding to one radiation length) that lies in front of the CPS and acts as a preradiator.

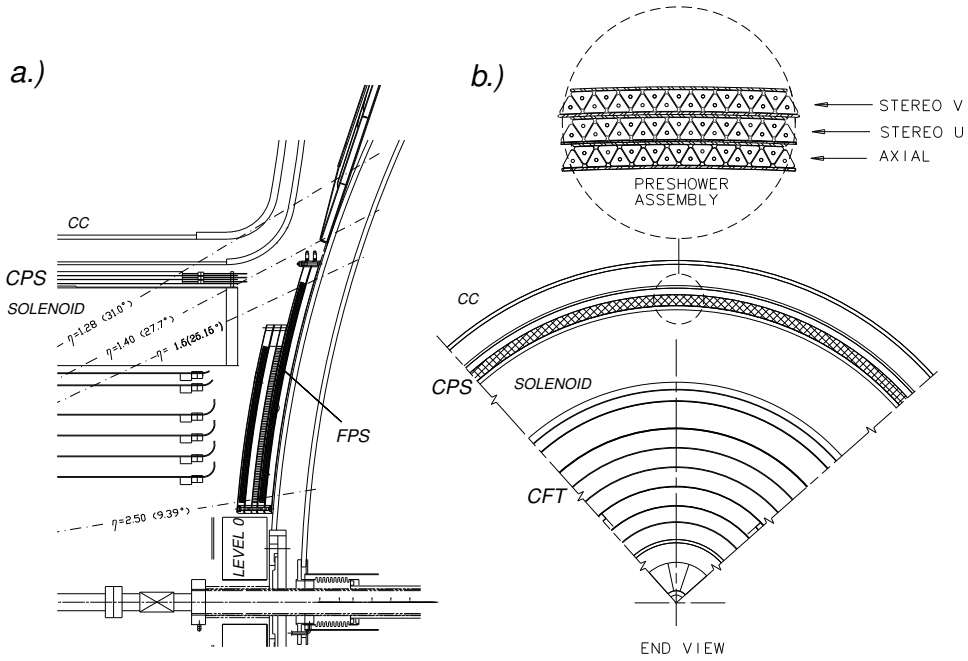


Figure 4.9: *The central and forward preshower detectors.*

The CPS consists of three concentric layers of scintillating strips: one inner axial layer and two outer stereo layers at an angle of $\sim \pm 23°$. The strips have a triangular

cross section with a base of 7 mm and a 1 mm hole in the center. A wavelength shifting fiber passes through this hole and is used for the readout. The light is transmitted by clear waveguides and is read out by VLPCs in a manner similar to the fiber tracker.

The forward preshowers (FPS) are mounted on the faces of the end calorimeters and cover the region $1.5 < |\eta| < 2.5$. The detectors have a design similar to the CPS and use scintillating strips. A layer of lead absorber[5] 11 mm thick (approximately 2 radiation lengths) is sandwiched between two scintillator planes. The inner layer of scintillators helps detect minimum ionizing particles. For the region $1.5 < |\eta| < 1.65$ however, there is only one scintillator plane and no lead layer. Particles in this region pass through a significant thickness of the solenoid and hence the additional layer of lead is not required.

4.2.3 Calorimeter

Calorimeters help measure the energy and position of particles by their total absorption. As particles interact with a medium, they lose energy to it and this energy loss can then be measured. Typically, a high energy ($\gg$ 10 MeV) electron passing through matter will emit photons via *bremsstrahlung*. These photons "pair produce" electron positron pairs, and these in turn emit more photons through bremsstrahlung. This process of forming an electromagnetic "shower" (or cascade) continues until the electron and positron energies are around 10 MeV where they begin to lose energy via ionization rather than bremsstrahlung. These particles at the end of the shower share the energy of the original electromagnetic particle and their presence can be detected via the ionization of atoms. Figure 4.10 depicts the basic concept of such an electromagnetic shower. The quantity X_0 in the figure is the radiation length,

[5]The addition of a layer of lead helps to increase the amount of absorbing material.

typically measured in g cm^{-2}, and is the amount of matter traversed for these interactions. It is defined as the mean distance over which a high energy electron loses all but 1/e of its energy by bremsstrahlung, and also as $\frac{7}{9}$ of the mean free path for pair production by a high energy photon. As an example of a typical value, the radiation length for uranium is about 3.2 mm [2].

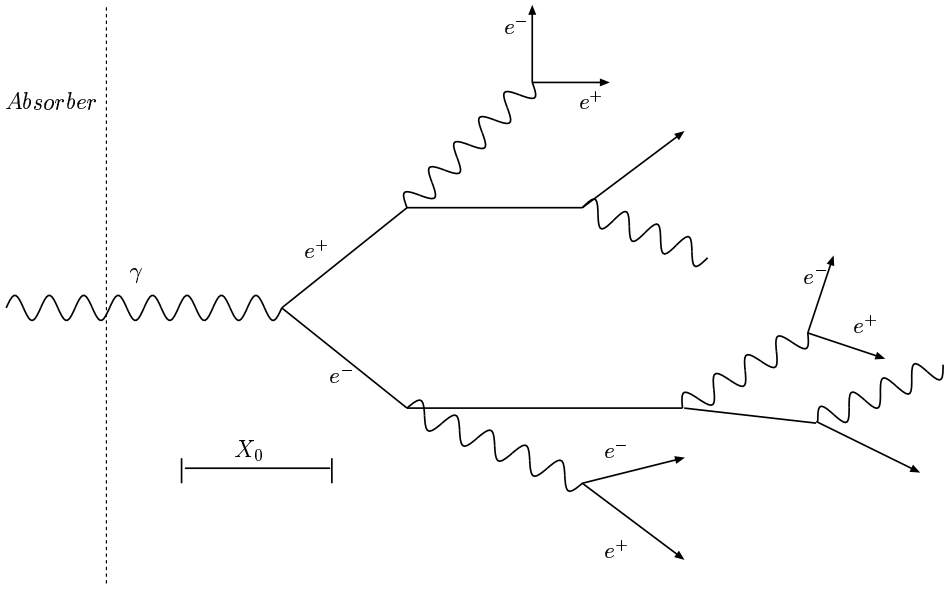


Figure 4.10: *Schematic drawing of the development of an electromagnetic shower.*

Hadronic particles, on the other hand, while passing through a material predominantly interact with the atomic nuclei via the strong force (instead of the electromagnetic force). They also produce secondary particles that go on to produce more particles with decreasing energy thus forming a hadronic shower. The average distance traveled by particles in a hadronic shower before an interaction is typically longer than that in an electromagnetic shower and hence hadronic showers penetrate further. As before, the low energy end particles are detected by their ionization loss. Additionally, neutral hadrons are detected by the production of secondary charged

particles. The hadronic shower shown in Fig. 4.11 shows two distinct components, namely the electromagnetic one (π^0s) and the hadronic one ($\pi^{\pm}$, n, etc.). λ is the nuclear interaction length and is used to describe hadronic shower development. For uranium, $\lambda \approx 10.5$ cm [2]. Therefore, hadronic showers are much more extended in space than electromagnetic showers of similar energy.

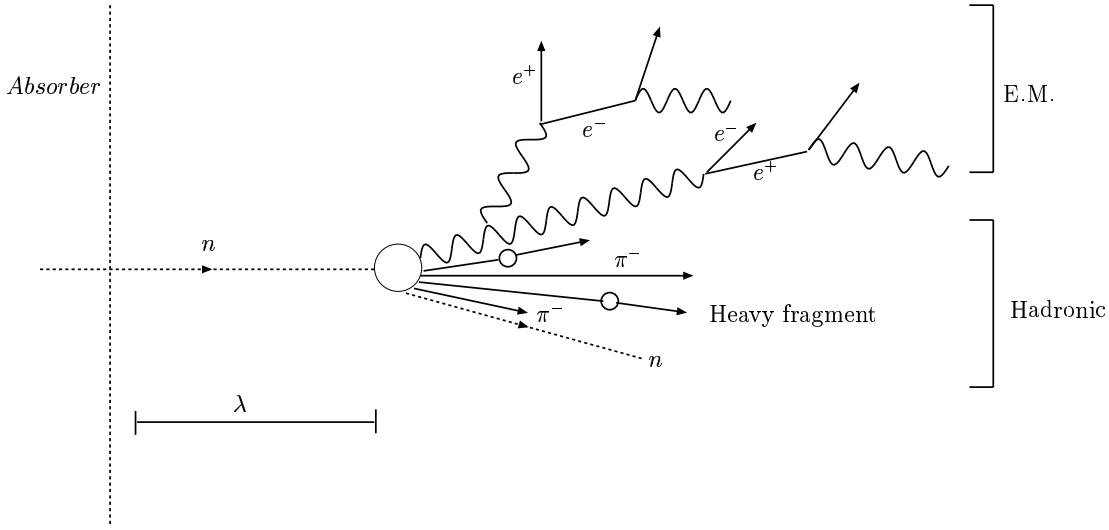


Figure 4.11: *Schematic drawing of the development of a hadronic shower.*

The DØ calorimeter is segmented into cells. Each cell consists of layers of absorbing material to induce shower formation and active layers where atoms are ionized by the passage of charged particles. Such a calorimeter that has alternating absorption and active layers is known as a sampling calorimeter. Depleted uranium, copper and stainless steel are the absorbers used while liquid argon serves as the active layer. The use of liquid argon also makes the design inherently resistant to radiation damage[6]. The ionized charge collects on a copper plate located in each cell and the energy of the

[6]Liquid argon is a simple non-interacting noble element and hence, unlike other ionizing material, it does not degrade over time. Moreover, it does not have a complicated crystalline structure (unlike solid state devices) that it needs to maintain for its functionality. Crystalline structures can degrade over time owing to radiation damage.

traversing particle (or jet) is found by summing the charge collected in all of the cells that it passes through. Calibration of the calorimeter is essential for the observed charge to be accurately converted into an energy measurement.

Now, typically in most calorimeters the fraction of energy deposited by an electron or photon is greater than that deposited by a hadron because much of the low-energy hadronic component is 'hidden' in nuclear binding energy release, etc. and hence is invisible. The use of depleted uranium makes the DØ calorimeter a compensating one where the calorimeter response to electromagnetic and hadronic activity is equal (or in other words $e/h = 1$). Low energy neutrons, from the nuclear breakup in hadronic showers, cause fission in the uranium and energy is converted into charged particles by the β decay of the fission products. This compensation is a useful feature because a hadronic shower contains an electromagnetic component from π^0 decays and hence a non-compensating calorimeter suffers from problems related to an energy dependent response ratio for electrons and hadrons, and a non-linear response to hadrons. Detector performance for a non-compensating calorimeter is compromised owing to fluctuations in the π^0 content of the showers and this results in (a) a skewed signal distribution and (b) an almost constant contribution to detector resolution which is proportional to the degree of non-compensation $(1 - h/e)$. For the DØ calorimeter the ratio of the electromagnetic and hadronic response is close to one and ranges from 1.11 at 10 GeV to 1.04 at 150 GeV.

There are three modules that make up the calorimeter: the Central Calorimeter (CC) and two End Calorimeters (ECs). The CC comprises cylindrical shells around the beam pipe while the ECs are located at either end (Fig. 4.12). The innermost, center, and outermost sections of the CC are, respectively, the electromagnetic (EM), fine hadronic (FH), and coarse hadronic (CH) regions. Each region is subdivided

into a number of layers and segmented in both ϕ and z coordinates. There are four EM layers where electromagnetic showers typically terminate and three FH layers where most of the hadronic showers deposit their energy. Some particles penetrate still further and deposit their energy in the single CH layer. Some important design parameters of the CC are listed in Table 4.3.

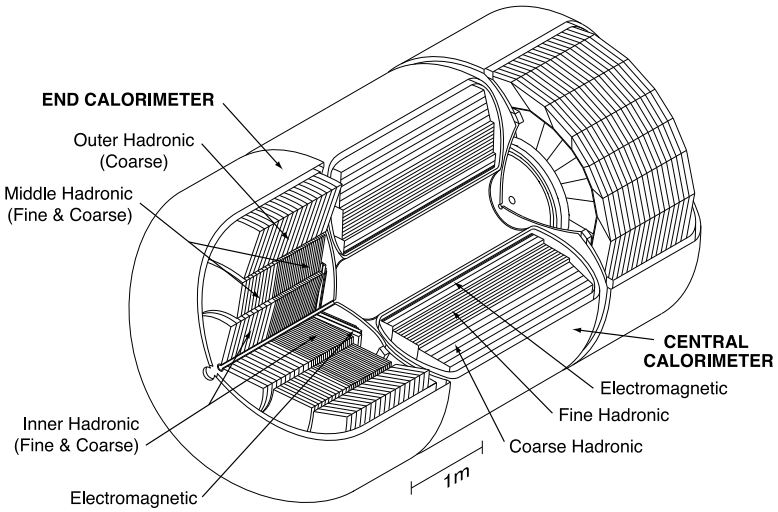


Figure 4.12: *Schematic drawing showing the central and the two end calorimeters [46].*

Module Type	EM	FH	CH
Rapidity Coverage	± 1.2	± 1.0	± 0.6
Number of Modules	32	16	16
Absorber	Ur	U-Nb	Cu
Absorber Thickness (mm)	3	6	46.5
Argon gap (mm)	2.3	2.3	2.3
Total Radiation Lengths	20.5	96.0	32.9
Total Nuclear Absorption Lengths	0.76	3.2	3.2

Table 4.3: *Some important parameters of the Central Calorimeter (CC) [47].*

Like the central calorimeter, the EC is composed of different sections: the forward

electromagnetic (EM), the inner hadronic (IH), the middle hadronic (MH), and the outer hadronic (OH). The IH and MH are additionally split into fine and coarse sections. Each of these is composed of varying numbers of layers and has a segmented structure similar to the CC.

Most of the cells in the CC and EC have a segmentation of 0.1×0.1 in $\eta \times \phi$ space. Since the position resolution of a traversing particle is determined by the size of the cells, the third row in the EM calorimeter (situated at the expected shower maximum[7] for electrons and photons) has a segmentation four times as fine (0.05×0.05), thus improving electron/hadron shower identification. For all layers with $|\eta| > 3.2$ the cell size increases to 0.2×0.2. Figure 4.13 shows a cross-sectional view of one quarter of the calorimeter, where the cell segmentation can be seen.

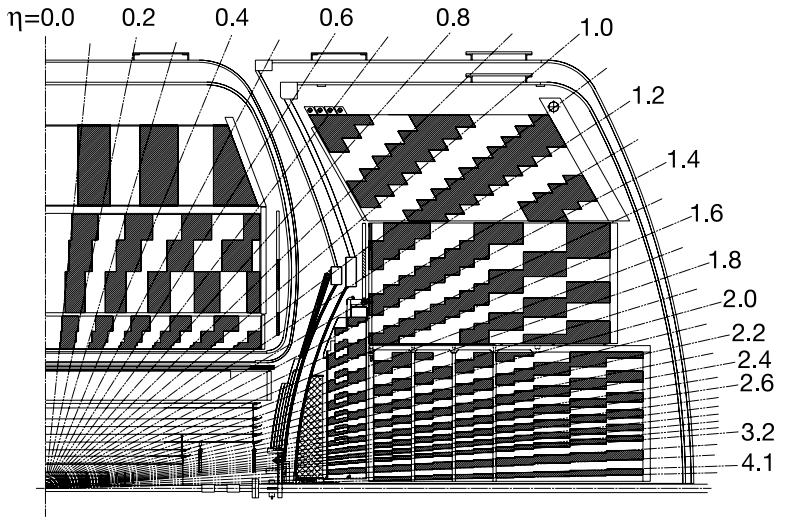


Figure 4.13: *A quarter view of the calorimeter showing the transverse and longitudinal segmentation pattern [46].*

[7]Actually, with the addition of the solenoid and preshower detectors, there is an additional $2X_0$ in front of the calorimeter and this makes the maximum EM shower occur in the front part of the EM3 layer.

In the crossover region from the CC to the EC, there are several areas (gaps) where particles travel mostly through support structures such as cryostat walls. Therefore, there are extra sampling layers on the end faces of the CC and on the inner surfaces of the end cryostats to compensate for the energy losses in these gaps. In addition to this there is also an intercryostat detector (ICD) that consists of a single layer array of 384 scintillating tiles mounted on the surface of both end cryostats. The tile size is chosen to match the calorimeter cell size and the scintillation light is taken by optical fibers to phototubes outside the magnetic field region.

4.2.4 Muon System

The dominant form of energy loss for muons at the Tevatron is through ionization. Owing to the muon's larger mass, bremsstrahlung does not play an important role until the muon has an energy of several hundred GeV. For a 10 GeV muon for example the total energy loss per unit length is more than three orders of magnitude less than that of a 10 GeV electron. As a result, muons with moderate transverse momentum ($p_T > 2.7$ GeV/c) completely penetrate the central calorimeter. Sometimes (high momentum) muons can be detected in the calorimeter by ionization of the liquid argon layer but since they rarely form an electromagnetic shower their total energy cannot be measured. A muon specific detector is therefore used to identify muons and measure their momentum. Such detectors take advantage of the muons' penetrating power: anything that reaches them is most likely a muon.

The DØ muon system surrounds the calorimeter and is arranged around three toroid magnets that bend the muon trajectories in the rz plane. The central toroid is located 318 cm from the beamline, covers the region $|\eta| \leq 1$, and has an internal field of approximately 1.8 T. The two end toroids are located at $454 \leq |z| \leq 610$ cm and

have internal fields of ≈ 1.9 T. The muon detector itself is divided into the central ($|\eta| < 1$) and forward ($1 < |\eta| < 2$) detector regions also referred to as the WAMUS (wide angle muon system) and FAMUS (forward angle muon system), respectively. Both regions are organized in layers of drift tubes and scintillation counters.

Drift tube detectors are gas filled volumes (containers) with an anode wire strung tight through the center of the volume. Cathode pads lie at the top and bottom of the container. Charged particles traverse the volume, ionize the gas, and produce electrons and ions. Electrons move towards the anode wire and as they are accelerated towards the wire, they gain energy and cause further ionization. This produces an 'avalanche' of electrons which amplifies the signal.

The scintillation counters collect scintillation light when a charged particle passes through them. Wavelength shifting fibers are embedded onto the scintillators and are connected to photomultiplier tubes which convert the light to an electronic signal.

The drift tubes and scintillators at DØ are organized into three layers: A, B, and C. The A-layer is located outside the calorimeter and is enclosed by the toroid magnet. The B- and C- layers are situated outside the toroid. The bending of muon tracks due to the magnetic field allows a measurement of the muon momentum. Figures 4.14 and 4.15 show an enlarged view of the drift tubes and scintillators, respectively. Table 4.4 lists some of the important parameters of the drift tubes in the central and forward regions.

The WAMUS region consists of Proportional Drift Tubes (PDTs) and scintillators in each of the three layers. The PDT chambers are formed using extruded aluminum tubes with a rectangular cross-section. There are 94 chambers in total, lying horizontally and vertically around the calorimeter, and they lend the muon system a cuboid geometry. Each drift chamber is typically 2.8×5.6 m^2 and is filled with a gas mixture

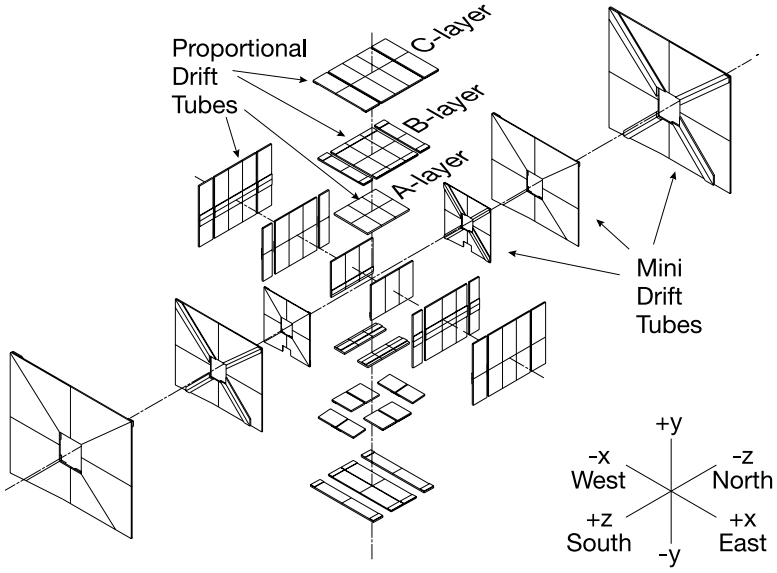


Figure 4.14: *An enlarged view of the muon drift tubes [46].*

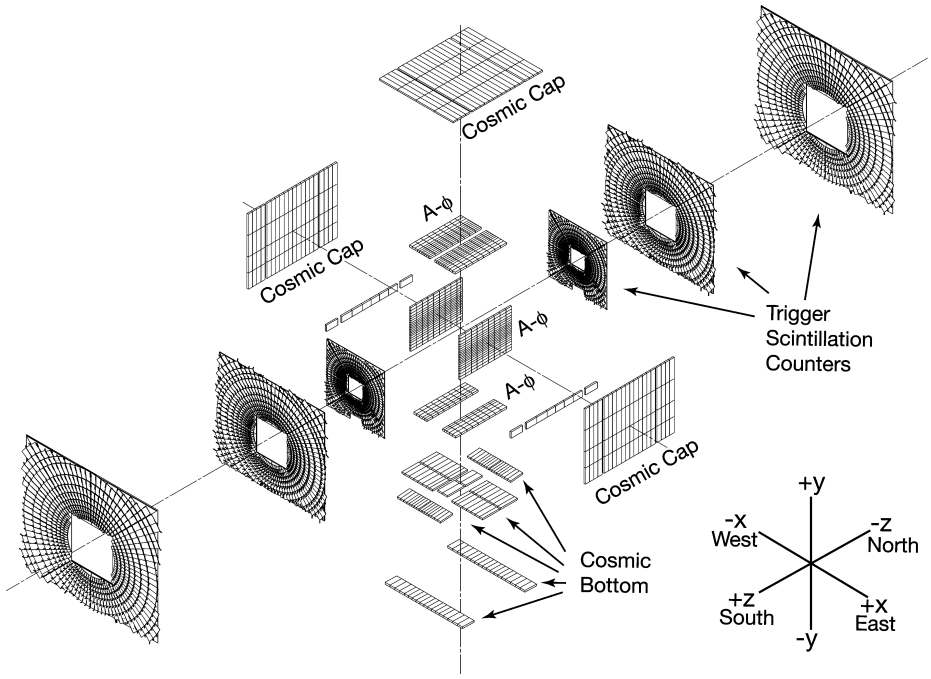


Figure 4.15: *An enlarged view of the muon scintillation counters [46].*

Parameter	Central Drift Tubes	Forward Drift Tubes
Wire Step	130 mm	10 mm
Wire Thickness	0.6 mm	0.6 mm
Cathode Material	Extruded Al	Al, Stainless Steel
Wire Material	W-Au (96% : 4%)	W-Au (96% : 4%)
Wire Diameter	50 μm	50 μm
Gas Material	84% Ar, 8% CH_4, 8% CF_4	90% CF_4, 10% CH_4
Cathode Potential	2300 V	3200 V
Maximum Drift Time	500 ns	60 ns

Table 4.4: *Some important parameters of the Muon Drift Tubes in the central and forward regions.*

of 84% argon, 8% CH_4, and 8% CF_4. The anode wire is gold plated tungsten and lies in the center of each chamber in the direction of the magnetic field lines. The drift time of the ionized gas to the wire gives a position measurement of the distance to the anode in the z direction with a resolution of 1.0 mm. The anodes of neighboring tubes are connected at one end with the readout of both located at the other end. The difference in the time of hits in these tubes gives a position measurement in the x and y direction with a resolution of 10-50 cm, depending on the distance along the wire.

In addition to the drift tubes there are layers of scintillation counters in the WAMUS region. These counters cover the PDT chambers in the A-layer, C-layer and the sides and bottom of the B-layer. There are a total of 1,002 counters having a ϕ segmentation of 4.5°. These scintillation counters provide a fast signal for use in the trigger and are especially useful since the drift time in the tubes (500 ns) is greater than the beam crossing time (396 ns). The scintillators have a time resolution of ~ 2 ns and are used to reject out-of-time backgrounds in the drift tubes.

The FAMUS region consists of mini-drift proportional tubes (MDTs) for track

reconstruction and scintillation counters for triggering. The MDTs are much smaller than the PDTs and have a cross-section of 9.4 mm $\times$ 9.4 mm. There are 6,080 such mini drift tubes which are filled with a gas mixture of 90% CF_4 and 10% CH_4. The MDTs are arranged in six layers of eight octants each. The length of the drift tube, and hence the anode wire, is parallel to the magnetic field (either the x or y direction depending on which octant it is in) and perpendicular to the particle's trajectory. The coordinate resolution for the MDTs is about 0.7 mm per hit. The momentum resolution of the forward muon system is limited by multiple scattering in the toroid and the overall muon momentum resolution is generally defined by the central tracking system for muons with momentum up to $\sim$ 100 GeV/c. The MDT system improves the momentum resolution for higher momentum muons and plays a particularly important role in determining the muon momentum for tracks with $1.6 \leq \eta \leq 2.0$, i.e. those which do not go through all the layers of the CFT.

The drift time in the MDTs is 60 ns, less than the beam crossing time, and so scintillation counters are not needed to match the drift tube hits to events. However, scintillation counters in all three layers in the forward region are still useful for reducing backgrounds coming from sources other than the interaction region (cosmic ray muons for example). These three layers of counters correspond to the three MDT layers and have a ϕ segmentation of 4.5°. There are a total of 4,214 such counters and their timing resolution is $\sim$ 1 ns.

4.2.5 Luminosity Monitor

The chief purpose of the Luminosity Monitor (LM) is to measure the Tevatron luminosity at the DØ interaction region. This requires a measurement of the rate of inelastic $p\bar{p}$ interactions and this is achieved by detecting the charged remnants of the

proton and anti-proton after the collision. The LM detector consists of two arrays of scintillation counters and these are located in front of the end calorimeters and occupy the radial region between the beam pipe and the forward preshower detector at $z \sim \pm 140$ cm (see Fig. 4.16). Each array consists of 24 wedges of scintillating material read out by photomultipliers (PMTs) and covers the pseudorapidity range $2.7 < |\eta| < 4.4$. Figure 4.17 shows the geometry of the LM counters and the location of the PMTs (solid dots).

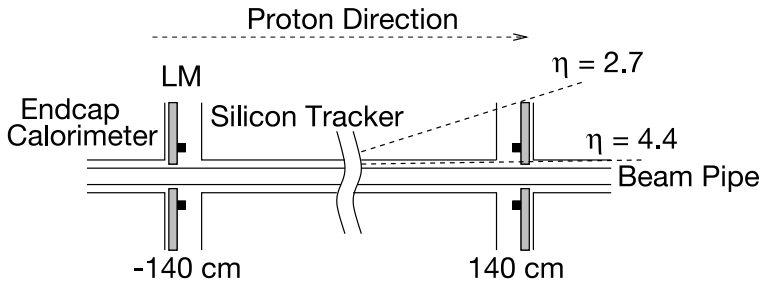


Figure 4.16: *Schematic drawing showing the location of the Luminosity Monitor (LM) detectors [46].*

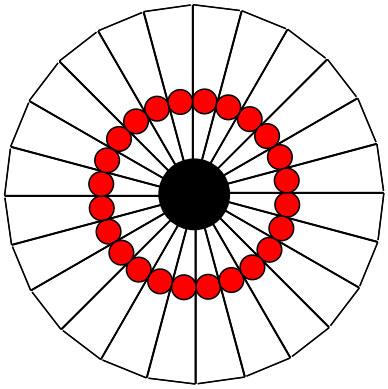


Figure 4.17: *Schematic drawing showing the geometry of the Luminosity Monitor scintillation counters and the location of the photomultiplier tubes (solid dots) [46].*

4.3 The DØ Trigger System

Proton anti-proton collisions take place at the Tevatron every 396 ns (at a rate of $\sim$ 2.5 MHz). This rate is too high for it to be technically feasible to write all events to storage tapes for offline analysis. Moreover, the dominant contribution to the total rate is inelastic $p\bar{p}$ scattering and the processes of interest occur at much smaller rates. Therefore, a technique is developed which enables us to decide in real time which events are interesting enough to store. This is called *triggering* and allows information from the detector sub-systems to be used to decrease the incoming rate to 30-60 Hz.

The DØ trigger is a three tier pipelined system with each tier reducing the rate into the next tier. Each succeeding tier (or level) examines fewer events but in greater detail and increasing complexity. The three tiers are aptly named Level 1, 2 and 3 (L1, L2, L3). An overview of the DØ trigger and data acquisition system is shown in Fig. 4.18. Note that though $p\bar{p}$ collisions take place at the rate of $\sim$ 2.5 MHz, the detector data in the figure is shown to come in at $\sim$ 1.7 MHz. As discussed previously in Sec. 4.1, the Tevatron beam is structured into super-bunches and bunches with 2 μs between each super-bunch and 396 ns between each bunch inside a super-bunch. Therefore, the average data throughput is reduced by the presence of these 2 μs gaps from 2.5 MHz to $\sim$ 1.7 MHz.

The first stage or L1 consists of a collection of hardware trigger elements that provide a trigger accept rate of approximately 2 kHz[8]. Each of the L1 trigger elements report their findings to the Trigger Framework (TFW) upon each beam crossing. The TFW then makes a global decision to either reject the event or accept it for further

[8]At present this rate is typically kept below 1600 Hz in order to minimize deadtime incurred due to the many channels of the SMT and CFT that have to be read out.

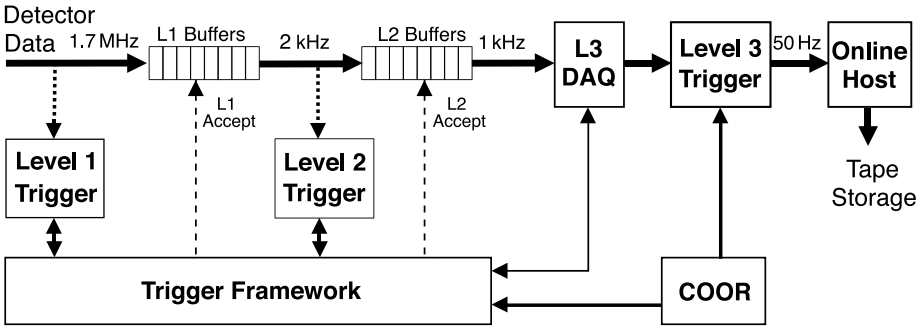


Figure 4.18: *An overview of the DØ trigger and data acquisition systems [46].*

examination. The L1 system can support a total of 128 separate L1 triggers or trigger bits[9]. Each bit is programmed to require a specific combination of trigger terms (or trigger decisions) and the logic is determined by custom hardware (and firmware) built using a series of field programmable gate arrays (FPGAs). The L1 trigger decision has to arrive at the TFW in 3.5 μs or less for it to participate in the trigger decision.

If the TFW issues an accept (called a L1 Accept), the event data is digitized and moved into a series of 16 event buffers for analysis at the next stage (L2). The L2 trigger uses both FPGAs and microprocessor chips; pre-processors (associated with specific sub-detectors) provide information to a global processor (L2Global) which constructs a trigger decision based on individual physics objects as well as object correlations. L2 reduces the trigger rate by a factor of about two and has an accept rate of about 1 kHz[10]. It has 100 μs to either accept or reject the event. On a L2 Accept events are moved into a set of L2 buffers where they await transfer to L3. These L2 (and L1) buffers play an important role in minimizing deadtime by providing FIFO storage for holding event data. A block diagram of the L1 and L2

[9]The trigger bit is said to be set if the event is passed by that particular trigger logic.
[10]At present this rate is typically kept below 900 Hz since there is not enough rejection at L3.

trigger systems is shown in Fig. 4.19. Its components are discussed in more detail in Secs. 4.3.1 and 4.3.2.

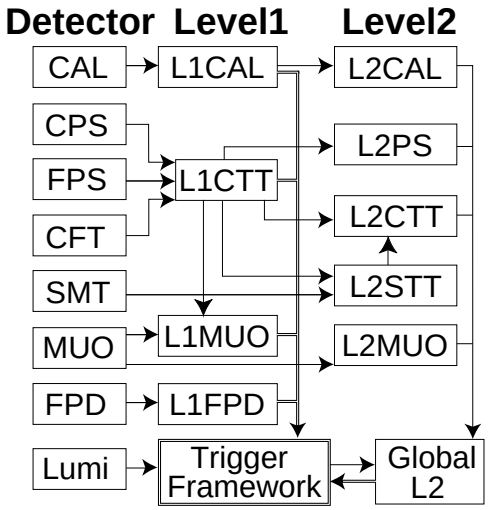


Figure 4.19: *Block diagram of the Level 1 and Level 2 trigger systems. The arrows indicate the flow of data through the system [46].*

Events that pass both L1 and L2 are sent by the data-acquisition system (L3DAQ) to a farm of Level3 (L3) microprocessors. Standard PCs running the Linux operating system refine the physics objects created by the L2 trigger and do a simple reconstruction of the whole event. The trigger has up to 150 ms to make a decision and the rate is reduced to ~ 50 Hz and the accepted events are distributed by the online host for offline reconstruction (written to tape) and monitoring purposes. The overall coordination of triggering and data acquisition is handled by a central coordination program called COOR running on the online host machine.

Table 4.5 summarizes the output rates for the different trigger levels and the time allowed for them to make their decision (latency). Both the design rates and the achieved rates are listed. The limiting factor for the final output rate is the cost of magnetic tape to store the data on, as well as computing resources needed to

reconstruct events in a timely fashion. With additional resources, the triggering and data acquisition systems can be easily upgraded to yield higher rates.

	Output rate		Latency
	Design	Achieved	
Events	1.7 MHz		—
Level 1	2 kHz	1.6 kHz	3.5 μs
Level 2	1 kHz	900 Hz	100 μs
Level 3	30-60 Hz	30-60 Hz	150 ms

Table 4.5: *Output rates and latencies for the different trigger levels.*

Sometimes, owing to the high rate of collisions (high instantaneous luminosity), certain triggers run with a *prescale* on their L1/L2/L3 condition. This implies that the trigger systems only pass the event a certain fraction of the time that the condition is satisfied. For example, if a trigger has a L1 prescale of 25, it will only pass the L1 trigger system in $\frac{1}{25}$th of the events that satisfy the L1 condition. The extent to which triggers are prescaled depends on the instantaneous luminosity.

4.3.1 Level 1 Trigger

The Level 1 trigger uses information from all detector sub-systems barring the SMT. Data from the SMT is sent directly to Level 2. The L1 Muon trigger is special in the sense that it receives input both from the muon detectors as well as the L1 track trigger. These are described in a little more detail in the following sections.

Track Triggers

The Level 1 Central Track Trigger (L1CTT) reconstructs trajectories of charged particles using data provided by three scintillator-based detectors: the central fiber tracker (axial fibers only) and the central and forward preshower detectors. As discussed in Sec. 4.2.2 the ϕ segmentation of the CFT is 4.5° and there are 80 such segments which form its trigger sectors. Hits in each sector are used to search for tracks via pre-programmed look up tables (LUTs). This is done by considering different possible hit patterns and programming those which are consistent with particle tracks into the LUTs. There are approximately 20,000 such pre-defined track equations. If the same hit patterns are seen in data, they generate a track candidate. Each such track candidate is identified by its trigger sector, momentum, and direction of curvature. Additional information is provided by corresponding hits in the preshower detectors. Fig. 4.20 shows a schematic illustration of a single 4.5° sector and a hypothetical track which is overlaid on the eight CFT axial doublet layers and the CPS axial layer. The track equations require a fiber hit on all eight CFT axial layers. The

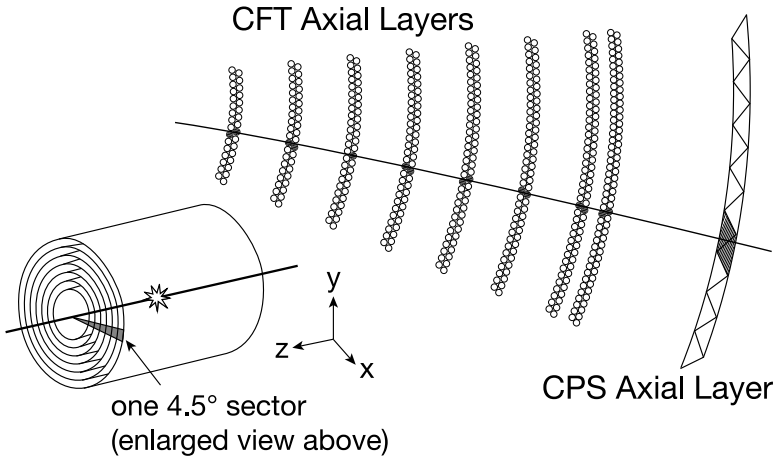


Figure 4.20: *Schematic drawing of a single* 4.5° *sector. A hypothetical track is overlaid on the eight CFT axial layers and the CPS axial layer [46].*

L1 track candidates that are generated are passed on to the L1 Muon trigger. The candidates, remapped onto the geometry of the SMT, are also used as seeds for the L2STT (Silicon Track Trigger).

Muon Triggers

The L1 muon trigger (L1MUO) combines information from muon scintillation counters, wire chamber hits, and the L1CTT to build muon objects. The hit information in the wire chambers is used to form track stubs which are then used to confirm scintillator hits in each layer. Triggers are formed by matching confirmed scintillator hits between layers. In addition, tracks from the L1 track trigger are matched to hits in the muon scintillator system and muon candidates are formed using combinatorial logic performed in FPGAs. Moreover, since cosmic rays originating in the atmosphere produce muons which penetrate the DØ detector, tracks with large transverse momentum are also required to pass cosmic ray veto scintillation counters. Cosmic rays are rejected based on their timing information relative to the beam crossing. Most cosmic rays pass through the detector at oblique angles and do not pass through the center of the interaction region.

Calorimeter Triggers

The L1 calorimeter trigger (L1CAL) adds up the energy in a *tower* of cells and triggers are formed by requiring that the energy deposited in trigger towers be above pre-set levels in one or more towers. Additional trigger terms are constructed from global quantities in the calorimeter, such as total energy, total energy projected in the transverse plane, and missing energy or energy imbalance in the transverse plane.

4.3.2 Level 2 Trigger

The L2 trigger combines and refines the outputs of the Level 1 trigger. The L2 Silicon Track Trigger (STT) is special in the sense that it receives information from two different sources: the detector itself (the SMT in this case) and the Level 1 trigger (L1CTT). Since the STT design is significantly different from the rest of the L2 trigger system, it is discussed separately in a following sub-section.

L2 includes preprocessors for each detector sub-system and a global processor (L2 Global) for combining information from the entire detector. Preprocessor subsystems include tracking, preshower, calorimeter, and muon systems. These systems work in parallel and transform the L1 trigger information into physics objects (tracks, energy clusters, etc.). For example, the calorimeter preprocessor (L2CAL) collects information from all the L1 trigger towers and uses that to build simple jet and electron candidates with the help of clustering algorithms. The L2CTT sorts the list of L1CTT tracks according to transverse momentum. The L2 muon trigger improves muon identification by combining wire and scintillator hits to form muon objects with track quality and transverse momentum information. The preprocessors then pass the above information on to L2Global which in turn correlates the output from these different sub-systems. When L2Global makes a decision, the decision is returned to the TFW which in turn issues a L2 Accept or a L2 Reject. On a L2 Accept, the data is sent to the Level 3 trigger.

The Silicon Track Trigger

The L2 silicon track trigger (STT) performs high precision online reconstruction of tracks found in the CFT by utilizing the much finer spatial resolution of the SMT. Requiring hits in the SMT for example helps reject spurious L1 triggers from acci-

dental track patterns in the CFT. The STT improves the momentum measurement of charged particle tracks at the trigger level and allows a precise measurement of the impact parameter of tracks. This helps tag the decays of long-lived particles, specifically B and D hadrons[11].

Figure 4.21 shows the basic working principle of the STT. For each event, a list of tracks is sent by the L1CTT. A ± 2 mm wide 'road' is defined around each track, and the SMT hits within that road are associated with the track. Only hits in the axial strips of the silicon ladders are used for this purpose. Hits in the innermost and outermost CFT layers are used along with hits in at least three of the four layers of the SMT to fit the track parameters. The results of the track fit are then sent to L2CTT.

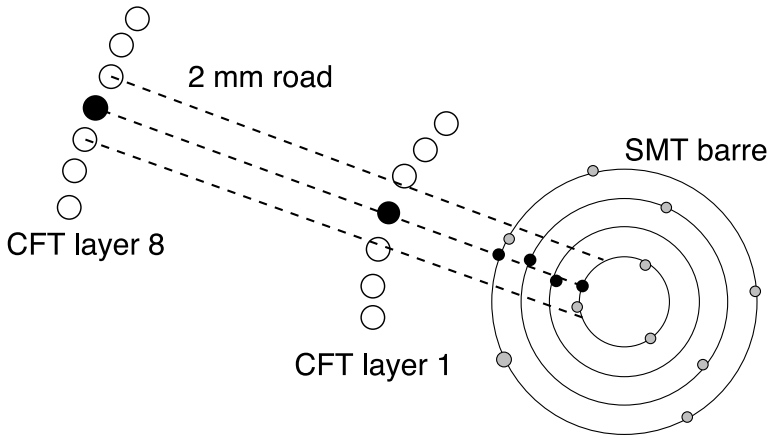


Figure 4.21: *Schematic drawing showing the conceptual design of the Silicon Track Trigger (STT) [46].*

As mentioned earlier, the STT design is significantly different from the rest of

[11] As discussed in Sec. 3.1.4 B (and D) mesons have relatively long lifetimes and typically travel a few millimeters before decaying. The trajectories of particles resulting from these decays therefore do not coincide with the primary vertex. The distance from the primary vertex to the point on the trajectory closest to the primary vertex is called the impact parameter, b. Tracks with large impact parameters thus help detect (or 'tag') B and D meson decays.

the L2 trigger system. The chief difference is dictated by the fact that the large number of readout channels and the speed of readout of the silicon detector prevents it from being used in the L1 trigger. Hence the STT (though a L2 trigger component) must receive and process digitized data coming directly from all the silicon detectors. For this purpose, the SMT barrel ladders are divided into twelve sectors, each one roughly 30° in azimuth. There is slight overlap between sectors but more than 98% of all tracks are contained in a single sector. Each 30° SMT sector is, therefore, treated independently in the STT with negligible loss.

The STT hardware design mainly uses custom-designed digital electronics modules. Logic daughterboards plug into a motherboard, and a common motherboard design is used throughout the system. Data is received from the L1CTT and the SMT via optical fibers which plug into custom receiver cards located in the rear card cage of the VME crate that houses the trigger electronics. The data is then processed in large FPGAs and/or DSPs on the logic daughterboards. There are three different types of daughterboards in the system: the Fiber Road Card (FRC), the Silicon Trigger Card (STC), and the Track Fit Card (TFC). Data communication between these daughterboards is achieved using custom built link transmitter and receiver cards. Additionally, each of the three daughterboards communicates with a common board on its motherboard called the Buffer Controller (BC). The BC is responsible for buffering data for readout through the data acquisition system once an event has been accepted by the trigger system. The daughterboards communicate with the BC, the link transmitter and receiver cards, and the VME backplane via three PCI buses on the motherboard. An additional bus, the VME bus, is used for initialization and monitoring, and to read data out from the Buffer Controllers.

The STT consists of six identical crates, each receiving data from two 30° SMT

sectors. Figure 4.22 shows the flow of data through one such STT crate. Each crate has one FRC, nine STCs, and two TFCs. Their functions are described below.

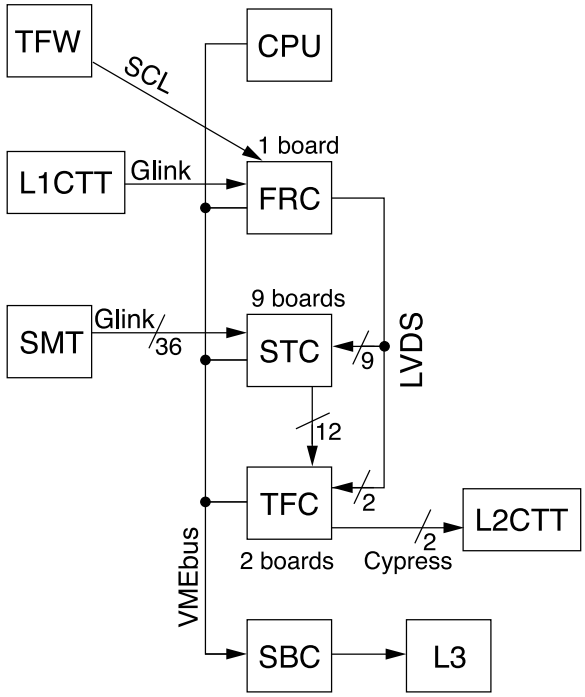


Figure 4.22: *Flow of data through a STT crate [46].*

- Fiber Road Card (FRC):

  The FRC is composed of four main elements implemented in three separate FPGAs. These include the trigger receiver, the road receiver, the trigger/road data formatter, and the buffer manager. The trigger receiver gets DØ-wide synchronization signals from the trigger framework via a special daughter card (the Serial Command Link or SCL card) on the motherboard. The road receiver logic accepts track information from the L1CTT and sends it to the trigger/road data formatter for reformating. The trigger/road data formatter combines the reformated tracks with the relevant SCL data from the trigger framework and

transmits the data to the other logic daughterboards (STCs and TFCs). The buffer manager handles buffering of events and readout to Level 3. Upon every L1 Accept, data is received and processed by all daughterboards. Data for readout is then transferred to the Buffer Controllers for each event the system receives. The buffer manager first assigns specific buffers into which this data is stored. It then decides if an event buffered by the BCs should be sent to Level 3 based on the Level 2 trigger information. If a L2 Accept is issued by the trigger framework, the buffer manager sends control signals to the BCs which prepare the data for readout to L3 via Single Board Computers (SBCs). The FRC thus serves as the STT's main communication link with the rest of DØ.

- Silicon Trigger Card (STC):

  The STC receives data from the axial and stereo strips of the silicon ladders. There are nine STCs per crate, each of which processes the data from eight detectors. The STCs first use downloaded LUTs to mask out noisy and dead silicon strips and then perform a strip by strip gain and offset correction. Next, they execute a fast clustering algorithm on the data. A cluster is defined as a group of contiguous strips with pulse heights above a given threshold (see Sec. 4.4.1 for more details). The cluster centroids are then determined and thereafter they are matched to the tracks broadcasted by the FRC using a second pre-computed LUT. Only information from the axial strips is used for this purpose. If a cluster centroid is within ± 2 mm of a CTT track it is kept, otherwise it is discarded. The list of centroids associated with CTT tracks is then transmitted to the track fitting cards. STC data is also read out to Level 3 via the BC and is used for monitoring the performance of the STC.

- Track Fit Card (TFC):

The TFC receives L1CTT tracks from the FRC and the centroids of the silicon clusters associated with those tracks from the STCs. There are two TFCs in each crate, one for each 30^o sector of the SMT. Processing in the TFC begins with a translation of the silicon hardware coordinates (detector and strip number) in the STC data into $r - \phi$ coordinates which are more suitable for track fitting. This is done by a pre-computed look-up table. For the CTT track, the TFC uses the hits in the inner and outer layers of the CFT (see Fig. 4.21) for the fit. An algorithm then looks at the silicon hits associated with the CTT track and in each layer selects the hit which is closest in ϕ to the center of the road representing the track. A fit is performed only if there are hits in at least three of the four SMT layers. If there are hits in all four layers and the χ^2 of the fit is larger than a pre-determined value, then the hit which contributes the most to the χ^2 is discarded and the track is refit. The track is fit to the linearized function:

$$\phi(r) = \frac{b}{r} + \kappa r + \phi_0, \tag{4.6}$$

where b is the impact parameter with respect to the detector origin, κ is the curvature of the track and ϕ_0 is the direction of the track at the point of closest approach. The TFC also corrects for beam position offsets from the detector origin. Online tracking measures the beam spot and it is downloaded to the TFCs at the beginning of every data-taking run. A correction for the offset is then applied to both the final hit selection in the TFC and the track-fitting. The TFC outputs the track parameters and the fit χ^2 along with additional information to L2CTT, where the tracks are sorted by p_T and impact parameter, and then passed onto L2Global to be used in the global trigger decision. The

data is also sent to the BC for readout to L3.

The STT helps select events with an enhanced heavy-flavor content (i.e. long-lived particles) by measuring the impact parameter b of reconstructed tracks with respect to the beam. Figure 4.23 shows the impact parameter resolution, σ_b, measured in data. The resolution includes the effect of the non-negligible beam spot size ($\sim$35 μm). The total resolution for high p_T tracks is between 45-50 μm with the variation depending primarily on track quality requirements. The value is roughly constant for tracks with $p_T > 3$ GeV/c. The p_T dependence is introduced by multiple scattering and in the trigger, the effect of this dependence is reduced by using the impact parameter significance, $S_b \equiv b/\sigma_b$, instead of the impact parameter b.

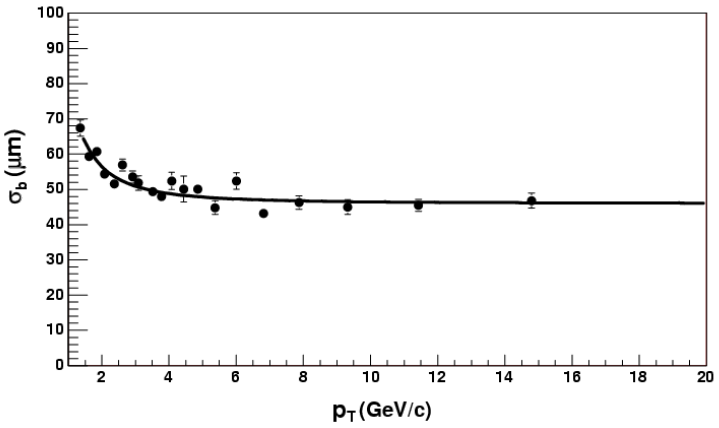


Figure 4.23: *The STT impact parameter resolution versus p_T of the tracks.*

4.3.3 Level 3 DAQ and Trigger

The third and final stage of the trigger is a dedicated computer farm that performs a fast reconstruction using a simpler version of the offline reconstruction code (described

in Sec. 4.4). The final trigger decision is made on high level 'physics' objects (such as electrons, muons, and jets) as well as on the relationships between such objects (such as the azimuthal angle separating the objects or their invariant mass).

Upon a L2 Accept, the data for that event is transferred out from each of the readout crates by a Single Board Computer or SBC (sitting in each crate) via a large ethernet switch. The data is sent to one or more farm nodes specified by routing instructions received from the routing master (RM) process running on an SBC in a special crate containing a hardware interface to the trigger framework (TFW). These farm nodes run two different programs: an event builder (EVB) and an event filter. The EVB process on each node builds a complete event from the event fragments received from the SBCs and makes it available to the event filter processes. The event builder gets an expected crate list from the RM in order to determine when an event is complete and if it does not get a full event, the event is rejected. Complete events are kept in buffers for processing by the filtering processes and the event builder routinely informs the RM of the number of free buffers that it has available.

The second program running on the farm nodes runs the event reconstruction and an event filter. At first software algorithms called *physics tools* are used to generate candidate objects and the relations between them. Individual calls to the tools are made by *filter scripts*. A filter script contains a list of the physics tools that are to be used to process the event, and the parameters (defined by filters) to be passed to the tools. A limited number of parameter sets (called reference sets or *refsets*) are used for each physics tool (e.g. three refsets are used to define an electron, each with different selection criteria). An event is passed by the trigger if all the filters for any of the filter scripts pass. These accepted events are written to tape for offline analysis and status information indicating which scripts passed or failed is passed along as

well.

The supervisor process running on a separate SBC interfaces between the main DØ run control program (COOR) and the L3DAQ system. When a new data-taking run is configured, the supervisor passes on run and trigger information to the routing master, as well as the COOR-provided L3 filter configuration to the farm nodes. Figure 4.24 shows the above flow of information and data through the L3DAQ system. The system's designed bandwidth is 250 MB/s, and this corresponds to an average event size of approximately 200 kB at an L2 Accept rate of 1 kHz.

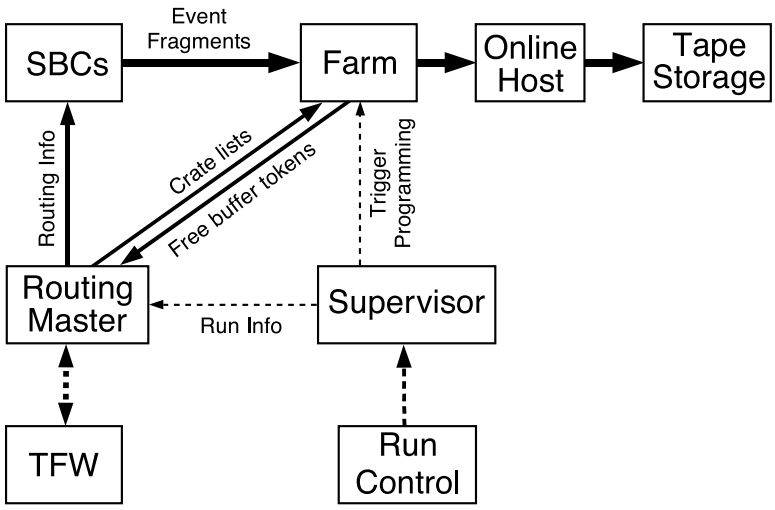


Figure 4.24: *Schematic drawing of the flow of information and data through the Level 3 data acquisition system [46].*

4.4 DØ Offline Event Reconstruction

Events that are written to tape are later processed by a computer farm and the raw data from the different detector systems is converted into a physics oriented format. The offline reconstruction program used for this purpose is called DØRECO and it

organizes information and results for each event using the DØ Event Data Model (EDM), a library of C++ classes and templates. The EDM stores information for each event in blocks called *chunks*. The raw data chunk or RDC is created by a L3 farm node and it provides raw detector signals and the results of trigger processing as input to DØRECO. The reconstruction then proceeds in several different steps:

- Detector-specific processing:

  Detector unpackers process specific detector blocks within the RDC and decode the raw data by associating electronics channels with physical detector elements and applying the necessary calibration constants. This information is then used for different purposes like hit-finding, where the digitized information is converted into hits at definite locations and energies (in the tracking detectors for example) or cluster-finding (in the calorimeters or preshower detectors).

- Track reconstruction:

  This step uses many different tracking algorithms and is the most CPU-intensive part of DØRECO. Hits in the SMT and CFT are used to form global tracks and the results are stored in track chunks that are input to the next step.

- Vertexing:

  Reconstructed tracks in an event are used to find the position of the primary vertex where the $p\bar{p}$ interaction occurred. First, a selection of tracks is performed to find tracks that most likely came from the primary vertex [48]. These tracks are also required to have hits in at least two silicon layers, $p_T > 0.5$ GeV/c, and a transverse impact parameter significance (b/σ_b) less than five[12]. The selected

[12]Tracks with large impact parameter are typically produced by secondary particles which decay after traveling a small distance. Therefore, tracks from the primary vertex are required to have a small transverse impact parameter.

tracks are then fit to a common vertex. If the $\frac{\chi^2}{\text{dof}}$ of the fit is greater than 10, the track which contributes the most is removed. This process is repeated until the $\frac{\chi^2}{\text{dof}}$ is less than 10 or there are less than two tracks left in the fit. The process is then repeated from the beginning using all tracks that have not already been associated with a vertex. Following the identification of primary vertices, displaced secondary vertices (associated with long-lived particles) are obtained. Both the primary and secondary vertex information is stored in vertex chunks and are made available to the next and final step of DØRECO. Additionally, a linear fit to the three-dimensional coordinates of all the primary vertices is performed to obtain the position of the beamspot. The beamspot is then stored in a run-averaged format in a database. This is standard procedure for most DØ analyses. For analyses related to the studies of b-hadrons, however, a slightly different vertexing technique is applied offline post-DØRECO. This technique is especially tuned for B-physics analyses and has been adopted from the standard method of tagging b-hadrons at the DELPHI experiment at CERN [49]. A more detailed description is provided in Sec. 5.1.

- Particle Identification:
  Information from the preceding steps is combined using many different algorithms to create physics object candidates. Tracks and clusters are matched to form electron, muon, neutrino, and jet candidates.

Some details of the above reconstruction steps that are most relevant to this dissertation are discussed in the following subsections.

4.4.1 Central Tracking Reconstruction

When a charged particle passes through the tracking detectors, the SMT and the CFT, several detector elements register the presence of the particle. Hits in the detectors are used to form clusters, and clusters are used as input to the track fitting process.

SMT Clustering

In the silicon detectors charge often accumulates on several contiguous strips when a charged particles passes near one of the strips. After reading out the detector, the first step involves applying gain and offset corrections to each strip in order to correct strip by strip variations in detector performance and electronic readout. The clustering algorithm then proceeds as follows. As each new strip which has an analog-to-digital (ADC) count above a certain threshold (8 ADC counts) is added, the position of the strip is checked to ensure that the strip is located next to the previous one (i.e. they are contiguous strips). If so, the strip is added to the original cluster. If the new strip's ADC count is below the threshold or the strips are not neighbors, a new cluster is started. The position of the cluster is given by the pulse height weighted average ($\bar{n}$) given by [50]:

$$\bar{n} = \frac{\sum n_i w_i}{\sum w_i}, \tag{4.7}$$

where n_i and w_i are the strip numbers and ADC counts for the ith strip in the cluster, respectively. The centroid of the cluster, u, is given by

$$u = u_1 + (\bar{n} - 1)p, \tag{4.8}$$

where u_1 is the position of the first strip in local coordinates, p is the pitch of the strips, and the -1 is needed because the strip numbering starts with 0. The local coordinates (axial and stereo) are combined to produce three-dimensional hits which are then converted to global coordinates.

CFT Clustering

The light yield for each fiber is converted to an ADC count and then calibrated using gain and pedestal information on a fiber by fiber basis. Light emitting diodes with the same light output are used for the calibration, and the light yield is given by [51]:

$$\text{light yield (photoelectrons)} = \frac{\text{ADC} - \text{pedestal}}{\text{gain}}. \tag{4.9}$$

The fiber's light yield obtained in this fashion is compared to a threshold which determines if the fiber has a hit or not. The threshold varies from sector to sector since the wave guides that read out the light vary in length (and longer fibers lose more light) but is typically between 1.4-1.5 photoelectrons. 1 photoelectron corresponds to 15 ADC counts in the axial layers and 7 ADC counts in the stereo layers.

Consecutive fibers with light yields above the threshold are then considered part of the same cluster. A new cluster is started if the fibers are not consecutive. The position of the centroid is taken as the half way point between the two fibers which define the start and end points of the cluster.

Tracking

Cluster information from different layers is used to form track candidates. The chief difficulty arises in identifying which clusters should belong to which tracks. Different

tracking algorithms are used for this purpose [52, 53, 54]. The algorithm most suited to low-p_T B-physics tracks — AA Tracking — is described below.

The first step of the tracking process involves obtaining track candidates. An initial track hypothesis is constructed using three clusters in the SMT barrels or disks. Their selection starts at the innermost layer and proceeds outwards[13]. The first measurement can be a cluster in layers 1-6 of any of the SMT barrels or any F-disk. The second can be selected in any following layer provided the axial angle between the two clusters is less than 0.08°. The third can be a cluster in any following layer as long as the radius of the circle drawn through the three clusters of the track candidate is greater than 30 cm i.e. it corresponds to a track candidate with $p_T > 180$ MeV/c. In addition to this, two other requirements are made: the impact parameter with respect to the beam spot must be less than 2.5 cm and the χ^2 of the track fit must be less than 16.

The track construction process continues into the next layer and an expected crossing region or expectation window is computed. Any cluster within this window is tried as a potential point on the track. If there is more than one cluster in a layer that satisfies all the above requirements, a new track candidate is formed. If no cluster is found in a layer of the expected region, that layer is considered a "missed" layer for that track. Although the trajectory of a charged particle should go through all the layers, because of the finite efficiency of the detectors there can be layers along a track that do not contain hits. A certain number of misses are then allowed in these layers in order to recover efficiency. However, if these selection criteria are very loose, the rate of fake tracks can become very large. Therefore, an optimization of these

[13]Since particles can interact or scatter in the detector material, it is important to know the track parameters at the distance of closest approach. Also, the association of SMT axial and stereo measurements leads to fewer combinatoric possibilities and hence provides motivation for starting the track construction process from the SMT.

criteria is done to ensure the highest possible efficiency for an acceptable fake rate. Three categories of missed layers are defined: inside misses (where the layer with the missing cluster is between the two layers where a cluster was found), forward, and backward misses (where the missed layer is forward or behind the track hypothesis). All track candidates are then required to fulfill the following conditions[14]:

- Clusters on at least four detector layers (SMT or CFT) with both axial and stereo clusters;
- Less than four inside misses;
- Less than seven (forward + backward) misses;
- Less than three inside misses in the SMT;
- Number of clusters should be greater or equal to five times the total number of misses;
- If a track candidate has at least one inside miss, there should be less than five (inside + forward) misses, and less than four (inside + backward) misses.

The second step in the track reconstruction process involves filtering out track candidates which are most likely to be noise or 'ghost' tracks (fake tracks reconstructed from hit patterns of real charged particles due to combinatoric ambiguities).

[14]Different criteria are used for the three kinds of misses since the inside misses are quite different from the forward and backward misses. An inside miss with a 100% efficient detector is a clear indication of a fake track, since a good track should have hits in all detectors. Therefore, the condition on inside misses is relatively strict. Forward misses can, however, occur for real tracks, for e.g. when a track interacts with the material. Also, a backward miss can occur for a track which does not originate from the primary vertex, e.g. a track which comes from the decay of a K_S^0. In both these cases these tracks are good real tracks. Therefore, the conditions on forward and backward misses are softer in comparison to inside misses.

The following algorithm is followed. First, an ordered list of tracks is constructed with tracks having the largest number of clusters placed first. If two tracks have the same number of clusters, the candidate with the fewest total misses is placed first. If the two tracks have the same number of misses as well, the candidate with the lower χ^2 is placed first.

Following the ordering of track candidates, the 'shared hit' criterion is used to select the candidates to be retained. According to the shared hit criterion, two or more tracks are allowed to share the same clusters but the following two criteria have to be satisfied:

- $N_{\text{shared}} \leq \frac{2}{3} N_{\text{total}}$,
- $N_{\text{shared}} \leq \frac{1}{5} N_{\text{total}}$ and $N_{\text{total}} - N_{\text{shared}} > 3$,

where N_{total} is the total number of axial clusters associated with a track candidate and N_{shared} is the number of shared axial clusters.

An additional step helps to further reduce the number of false tracks. The primary vertex is reconstructed using the tracks that survive the above procedure. Each track that has a small impact parameter with respect to any primary vertex has its cluster count artificially increased by two. The tracks are then re-ordered and the track selection procedure repeated. This process helps to ensure that track candidates that are most likely from a primary vertex are kept.

While SMT clusters are needed to start the process of constructing track candidates, in the case of non-functioning and inefficient detectors, the tracking algorithm also considers tracks with CFT only clusters. In this case, the first step involves finding the primary vertices of the event. Then a second round of track finding begins, starting with clusters in the innermost layer of the CFT. This continues on to

the outermost layer with the additional requirement that the impact parameter of the CFT track with respect to the primary vertex be less than 1.5 cm. The CFT track candidate is then extrapolated into the SMT and any SMT clusters that can be associated with this track are kept.

4.4.2 Calorimeter Reconstruction

The calorimeter signal consists of a collection of electrons coming from the ionization of liquid argon. The signal is digitized and then sent through a series of readout electronics. First, a cell by cell correction is performed to take care of intrinsic differences in cell to cell response and electronic readout. The corrected number of ADC counts thus obtained is then converted to an energy deposition in GeV. Calibrations are done using both test beam results and *in-situ*. Test beam calibrations involve particles of known energy which are targeted on portions of the calorimeter [55]. For *in-situ* calibrations the invariant mass of particles whose mass is known to much better precision than the resolution of the calorimeter is reconstructed [56]. Post-calibration, the cell energies are summed in towers of η and ϕ using the deposition in each cell.

Calorimeter objects are reconstructed by forming clusters of neighboring cells containing energy deposits. An electromagnetic cluster is defined as a group of towers in the calorimeter within a cone of radius $\mathcal{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.2$ around the highest energy tower. The total energy within the cone is obtained, and the fraction of cells from EM and hadronic layers is used to identify it as an EM object or a hadronic jet. An EM object is then identified as an electron if it is matched to a central track or as a photon if it is not.

4.4.3 Muon Reconstruction

The process of reconstructing muon objects from hits in the muon detectors involves three different steps [57, 58].

Hit Finding

Hits in the drift tubes and the muon scintillators are used to reconstruct muon trajectories. In the central region the PDTs provide a measurement of the drift time (the time it takes the signal to reach the anode wire) as well as the axial time (the time it takes the signal to move from the wire to be collected). The axial time determines the position of the passing particle along the wire. The drift time, along with the angle of the track, determines the distance of the particle perpendicular to the wire. The scintillator hit position improves the axial resolution of the hits. In the central region, the MDTs provide a single time measurement, which is a sum of the drift and axial times. A matching scintillator hit is then required to determine the axial position of the track. Once that is known, the drift time can be determined and used to calculate the perpendicular distance of the particle trajectory to the wire.

Segment Finding

After the hits have been reconstructed from raw data, straight track segments are reconstructed in each layer of the muon system. This process uses a *linked list* algorithm. Straight lines, called links, are formed between hits that are within 20 cm of each other, are not in the same plane, and are not from the same underlying wire hit. The links are matched recursively to form straight line segments. If two links are found to be compatible with a straight line segment they are merged into a single

new link containing all the hits of the constituent links. The process is repeated until an attempt has been made to match each link with all other links. The resulting segments are then fit as straight lines. Once the fit has converged, the segment is extrapolated to the scintillator position in the drift plane of the wire hits. If a matching hit is found, the segment is re-fitted, this time taking into account the scintillator hit. In addition to this, a new fit is performed to match B- and C-layer hits. Segments in these layers are expected to be part of the same straight line segment since there is no magnetic field between these two layers. The B- and C-layer segments are therefore matched within each octant and region, and a new fit is performed using all hits on both segments. Since this algorithm can find multiple segments from a collection of hits, the best segment is chosen by calculating a χ^2 (assuming a straight line fit) for each segment and selecting the one with the smallest χ^2/n_{dof}. For segments with only two hits, the segment that is best compatible with the primary vertex is chosen.

Track Fitting

A-layer as well as BC-layer segments are fit to obtain a "local" muon track. Starting with the position of the BC-layer segment, the track is propagated step by step through the toroid and on to the A-layer following a helical path. The process takes into account energy loss at each step as well as multiple scattering. The result is a local muon track parameterized by the position and momentum at the A-layer. The local muon track is then matched with a track from the central tracker to obtain a "global" muon track. The procedure takes into account both the magnetic fields (solenoid and toroid) and multiple Coulomb scattering and energy loss in the toroid and the calorimeter. The matching is performed and the distance of closest approach to the beam is calculated.

4.4.4 Particle Identification

Electron Identification

Electrons, being charged particles, produce hits in both the SMT and the CFT. They have a curved trajectory owing to the 2 T magnetic field. They also deposit energy in the calorimeter with bremsstrahlung being the primary cause of energy loss in the inner most part of the calorimeter.

The process of electron identification typically starts in the calorimeter where an electromagnetic cluster is defined as a group of towers in the calorimeter within a cone of radius $\mathcal{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.2$ around the highest energy tower. Additionally, a track reconstructed in the central tracking detectors should point towards the direction of the electromagnetic shower and have a momentum which is consistent with the energy deposition in the calorimeter. Different parameters that rely on the characteristic shape of the shower are also used to differentiate between real electrons and background and are discussed in more detail in Ref. [59]. The chief expected sources of background are (a) π^0 showers (b) photons which convert to e^+e^- pairs, and (c) fluctuations in hadronic shower shapes.

While the above method is efficient for high energy isolated electrons (for example those coming from $W^\pm$ or Z decays), the situation gets more complicated for electrons belonging to hadron jets, such as those arising from semileptonic b-decays that are considered in this analysis. In this case, the jets are more collimated and the hadron and electron showers tend to overlap thus obscuring some of the characteristics of the electron cluster. Fig. 4.25-left illustrates how an electron and a pion shower could overlap in the calorimeter leading to a merged cluster.

This problem is mitigated by adopting an alternate electron identification proce-

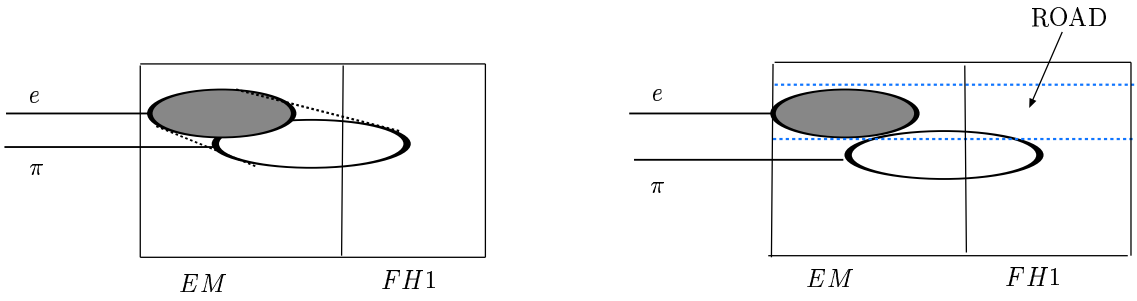


Figure 4.25: *An electron shower and a pion shower overlap in the calorimeter. The two clusters are merged (left) but the pion contribution to the road is reduced (right).*

dure, the *road method*, which takes full advantage of the granularity of the calorimeter and the tracking capabilities of the DØ detector [60]. The method involves extrapolating reconstructed charged particle trajectories into the calorimeter and considering (for a given track) only the energy contained in a narrow *road* (Fig. 4.25-right). This reduces the contribution from neighboring hadronic showers and even from extraneous photons.

The algorithm proceeds in the following manner:

- The trajectory of each selected charged particle track is extrapolated through the calorimeter and a road is constructed. Since typically most of the energy deposited by electrons is contained within the first three layers or floors[15], the construction of the road is optimized using the features of the energy deposits in those floors.

- A list of cells belonging to the road is obtained for every floor of the calorimeter.

[15]The fraction of the calorimeter energy contained in the first three floors is 90% for the central calorimeter and 95% for the end calorimeter, the difference arising mainly due to some cracks in the central calorimeter.

- The transverse energy, E_T, contained in a road is calculated for every floor. E_T is computed as the sum of transverse energies of the cells belonging to the road. Only cells with energy exceeding a given threshold (10 MeV at present) are taken into account in order to avoid including negative energy cells.

Two standard electron identification estimators are then determined for each road using the computed transverse energy, E_T:

$$\begin{aligned} \text{EMF} &= \frac{\sum_{\text{floor number i=1,2,3}} E_T(i)}{\sum_{\text{all floors}} E_T(i)} \\ \text{E/p} &= \frac{\sum_{\text{floor number i=1,2,3}} E_T(i)}{p_T(\text{track})} \end{aligned} \tag{4.10}$$

where $E_T(i)$ is the transverse energy within the road in floor i, and EMF is the electromagnetic fraction. Note that this definition of EMF does not include the energy in the fourth floor.

Selection criteria for the electrons used in this analysis were developed from a study of electrons from conversion decays and pions from K_S decays (fakes). First, the sample was divided into two p_T bins and the cuts were chosen so as to keep the pion rejection at the same level. Then, for electrons with $p_T < 3.5$ GeV/c the following criteria were used: EMF > 0.8 and $0.55 <$ E/p < 1.0. For electrons with $p_T > 3.5$ GeV/c, it was required that EMF > 0.7 and $0.5 <$ E/p < 1.1. Additional details of the selection cuts imposed on electron candidates used in this analysis are provided in Chapter 5.

Muon Identification

As mentioned in Sec. 4.4.3 the muon detector system along with the central tracking system provides an unambiguous way of identifying muons. The muon detector is

especially effective since it covers approximately 90% of the angular acceptance up to $|\eta| < 2$ with the loss in acceptance coming mostly from missing detector coverage at the bottom of the detector. Additionally, the central tracking system is highly efficient in finding tracks from charged particle hits in the muon detector and in fact, has better momentum resolution for the hits. This results from the fact that multiple scattering in the toroids degrades the resolution somewhat for the standalone muon system. Consequently, muon momentum is measured using central tracks matched to the local muon track, i.e. global tracks. Moreover, since muons are minimum ionizing particles in the calorimeter, they also deposit a small amount of energy in the calorimeter. However, the efficiency of identifying muons from calorimeter information alone is $\sim 50\%$ and therefore far less efficient than the other systems.

For the data used in this analysis muons are required to have a BC segment matched to a central track with both hits in the SMT and CFT present. In addition to this, they are required to have transverse momentum $p_T \geq 2.0$ GeV/c, pseudorapidity $|\eta| < 2$, and total momentum $p \geq 3$ GeV/c. More details of the selection criteria used are discussed in Sec. 5.1.

4.5 Event Simulation

Computer simulations of both signal and background events are used to model the response of the DØ detector. These Monte Carlo (MC) simulations proceed through a number of different steps which are described below.

4.5.1 Event Generation

'Event generators' which describe the production mechanism at the $p\bar{p}$ hard scattering level are used to generate simulated events. The typical output is a list of vertices and particles that were produced at those vertices. The PYTHIA event generator [61] is the most widely used at DØ and was used for this analysis. The program generates complete events incorporating our current understanding of the underlying physics. This includes hard and soft sub-processes, parton distribution functions, fragmentation, and decays etc. Monte Carlo techniques are used in addition to properly simulate the quantum mechanical variation between events observed in nature using both average behavior and fluctuations.

While B mesons are created by PYTHIA, a program especially tuned for B physics, EvtGen [62], is used to simulate the decays of B mesons and their daughter particles. Appendix A.1 lists the EvtGen decay files used for generating the different Monte Carlo samples used in this analysis. The generated events are then filtered using the `d0_mess`[16] package [63] and only those events which contain the desired particles satisfying the required kinematic cuts are kept. A few examples are listed in Appendix A.2.

4.5.2 Detector Simulation

The output of the event generation step is passed through a full simulation of the DØ detector. This simulation consists of two programs: DØGSTAR [64] and DØSIM [65]. DØGSTAR[17] is based on the CERN GEANT (v3.21) program [66] which describes the true geometry of a detector by building it up from a library of known shapes.

[16] Abbreviation for DØ **M**onte Carlo **E**vent **S**election **S**ystem.
[17] Abbreviation for DØ **G**EANT **S**imulation of the **T**otal **A**pparatus **R**esponse.

DØGSTAR helps trace particles through the DØ detector, determines where their paths intersect active areas, and simulates their energy deposition and secondary interactions.

The DØSIM program modifies the output of DØGSTAR in order to account for various detector related effects. It simulates the digitization of analog signals from the detector and converts the simulated data to a form that real data takes when processed through the DØ electronics. It also takes into account various detector inefficiencies and noise from both the detector and the electronics. Additionally, it does *pile-up* of any additional interactions that might occur in the same bunch crossing as the signal event[18]. Calorimeter pileup which occurs when significant energy is deposited before the energy from the previous bunch crossing has been read out, is also modeled in the program. The output of DØSIM is in the same format as the raw data and is passed onto the reconstruction program DØRECO (Sec. 4.4). The format of the DØRECO output is identical to that of the data processed offline, but contains additional Monte Carlo information that makes it possible to correlate reconstructed detector data with the original (or "true") generator output.

[18]In $p\bar{p}$ collisions, owing to the composite nature of the incoming particles, several parton pairs can undergo separate hard or semi-hard scattering called "multiple interactions". In addition, other protons and anti-protons can collide in the same bunch crossing, producing an independent $p\bar{p}$ event at another position along the beamline. Such an additional $p\bar{p}$ event is called a pile-up event and there can be more than one in a single beam crossing.

Chapter 5

Analysis

As reviewed in Chapter 3 the different steps essential to a B_s^0 mixing analysis can be summarized as follows:

- Production and selection of signal events;
- Flavor tagging the selected events (i.e. obtaining the number of mixed and unmixed events) and calculating the tagging purity (or the dilution $\mathcal{D}$);
- Calculating an expected asymmetry incorporating the K factor correction, the VPDL resolution and the fact that there can be contamination of the B_s^0 sample by mesons that either do not mix ($B^{\pm}$) or mix at a different rate (B_d^0);
- Obtaining a measured asymmetry between unmixed and mixed B_s^0 mesons as a function of VPDL;
- Fitting for Δm_s using the expected and measured asymmetries.

A general introduction was provided earlier; details specific to this analysis are discussed in the following sections.

A total of 610 pb^{-1} of data collected by the DØ detector during the period April 2002 to May 2005 was used for this analysis.

5.1 Reconstruction and Event Selection

The production of B_s^0 mesons was discussed in detail in Sec. 3.1 and the different steps involved in offline reconstruction of the data were reviewed in Sec. 4.4. A description of the vertexing procedure specific to this analysis and the details of the event selection process are provided in this section.

5.1.1 Vertexing

The primary vertex (PV) for each event was reconstructed using a set of selected tracks and the run-averaged beamspot position from the beamspot database. The beamspot is usually stable within a run, and so can be used as a constraint for the primary vertex fit. The PV position, $\vec{V}$, was therefore determined by minimizing a χ^2 function that depends on all the tracks in the event and on a term that represents the beamspot constraint [49]:

$$\chi^2(\vec{V}) = \sum_a \sum_{\alpha,\beta=1,2} b_\alpha^a (S_\alpha^{-1})_{\alpha\beta} b_\beta^a + \sum_i \frac{(V_i^{sp} - V_i)^2}{(\sigma_i^{sp})^2}. \tag{5.1}$$

$\{b_1^a, b_2^a\} = \{b_T^a, b_L^a\}$ is the 2-dimensional vector of impact parameter (IP) components for each track a entering the fit. b_T^a and b_L^a are the transverse[1] and longitudinal[2] projections of track impact parameter with respect to the primary vertex for each

[1] in the plane perpendicular to the beam direction.
[2] parallel to the beam direction.

track a. S_a is the covariance matrix of the measured quantities $\{\epsilon_T^a, \epsilon_L^a\}$ where ϵ_T^a and ϵ_L^a are equivalent impact parameter components but defined with respect to the origin, rather than with respect to the PV. V_i^{sp} and σ_i^{sp} are the beamspot position and size of the x and y components.

First, the summation in Eq. 5.1 was performed using all the tracks (N_{tr}) and the χ^2 of the fit, $\chi^2(N_{tr})$, was computed. After that each track i was consecutively removed and the corresponding $\chi^2(N_{tr}-1)$ was obtained. The track i that was responsible for the maximal difference $\chi^2(N_{tr})-\chi^2(N_{tr}-1)$ was excluded from the fit if the difference exceeded a threshold value Δ (set to 9). The procedure was repeated until there were no more tracks with a χ^2 difference exceeding Δ.

The error on the transverse impact parameter was obtained to be [49]:

$$\sigma_T^2 = \begin{cases} (\sigma_T^{tr})^2 - (\sigma_T^{pv})^2 & \text{if the track was included in the PV fit} \\ (\sigma_T^{tr})^2 + (\sigma_T^{pv})^2 & \text{otherwise} \end{cases} \tag{5.2}$$

with similar expressions for σ_L^2. In the above equations $\sigma_T^{tr}(\sigma_L^{tr})$ is the error on $\epsilon_T(\epsilon_L)$ coming from the track fit and σ^{pv} is the error from the PV fit obtained using its covariance matrix. Impact parameter significances (S_T and S_L) can then be defined as:

$$S_T = b_T/\sigma_T, \tag{5.3}$$

$$S_L = b_L/\sigma_L. \tag{5.4}$$

These expressions can be used to select events with long-lived particles (b and c-hadrons) which have large impact parameters significantly exceeding the uncertainties σ_T and σ_L.

The next step after reconstructing the primary vertex was to obtain the secondary vertices in the event. This was done by reconstructing two different $B_s^0 \to D_s^- \mu^+ X$ decay modes — the $D_s^- \to \phi\pi^-$ mode and the $D_s^- \to K^{*0}K^-$ mode. The procedure was identical for the two channels with only a few differences in the selection criteria. These differences were motivated by the kinematics and the background levels in the two cases. Sec. 5.1.2 describes the selection process, and the different criteria used are summarized in Table 5.1.

5.1.2 Event Selection

The general strategy for selecting signal $B_s^0 \to D_s^- \mu^+ X$ events involved using criteria that would suppress background events without incurring a significant decrease in the number of signal events. Various kinematic distributions for both signal and background were studied for obtaining the specific criteria described below.

The muon in the $B_s^0 \to D_s^- \mu^+ X$ decay was reconstructed and identified using the standard DØ algorithms as described in Sec. 4.4. It was required to have a BC segment in the muon chambers[3] matched to a central track with at least two SMT hits and two CFT hits. Such requirements on the number of SMT and CFT hits throw away cases where the decay length is poorly reconstructed. Additionally, the muon was required to have transverse momentum $p_T \geq 2.0$ GeV/c, pseudorapidity $|\eta| < 2$, and total momentum $p \geq 3$ GeV/c.

- nSMT ≥ 2; nCFT ≥ 2, $p_T(\mu) \geq 2.0$ GeV/c, $|\eta|(\mu) < 2$, and $p(\mu) \geq 3$ GeV/c.

All the tracks in the event were clustered into jets using the DURHAM clustering algorithm [67] with the cutoff parameter of 15 GeV/c. At least three additional

[3]Or in other words the muon was required to have penetrated the toroid.

charged tracks were required to be from the same jet as the muon, to have at least two hits both in the SMT and the CFT, and to have total charge equal to one in magnitude and opposite to the charge of the muon[4]. These particles were assigned the masses of kaons (K_1 and K_2) and pion (π) requiring the charge combination $\mu^+ K_1^+ K_2^- \pi^-$ (or its charge conjugate). The transverse momenta of these particles were required to be:

- $D_s^- \to \phi\pi^-$:
  $p_T(K_1) > 0.7$ GeV/c, $p_T(K_2) > 0.7$ GeV/c, and $p_T(\pi) > 0.5$ GeV/c,
- $D_s^- \to K^{*0}K^-$:
  $p_T(K_1) > 0.9$ GeV/c, $p_T(K_2) > 1.8$ GeV/c, and $p_T(\pi) > 0.5$ GeV/c,
  assuming that K_1 is the kaon from the $K^{*0} \to K^+\pi^-$ decay.

The values for these cuts were chosen after studying p_T distributions for both signal and background. Imposing these cuts helped reduce the possibility of picking up random tracks (which are typically low in p_T) for reconstructing the D_s^- candidate.

For each particle, the transverse b_T and longitudinal b_L projections of track impact parameter with respect to the primary vertex together with their corresponding errors (σ_T, σ_L) were computed. The combined significance $(b_T/\sigma_T)^2 + (b_L/\sigma_L)^2$ was required to be greater than 4 for K_1 and K_2, while there was no cut on the significance of the pion. These impact parameter significance cuts help reject tracks that come from the primary vertex (instead of the D_s^- vertex).

- $(b_T/\sigma_T)^2 + (b_L/\sigma_L)^2 > 4$ for K_1 and K_2.

[4]These tracks should come from the decay of the D_s^- meson which has a charge opposite to the charge of the muon in the decay $B_s^0 \to D_s^- \mu^+ X$.

Three charged particles were required to come from the same D_s^- vertex with the χ^2 of the vertex fit satisfying $\chi^2 < 16$. The D_s^- candidate produced by their combination was required to have a common B vertex with the muon with the χ^2 of the vertex fit less than nine ($\chi^2 < 9$). These cuts on the quality of the vertex fit reduce the probability of picking up spurious tracks during the vertexing procedure. Spurious tracks typically do not point back to the vertex and hence give large contributions to the χ^2.

- $\chi^2(D_s^-) < 16$,
- $\chi^2(B_s^0) < 9$,

The relatively long lifetime of the D_s^- meson (490 fs) was used to reject random combinatorial background. For this purpose, the decay length significance, $d_T^D/\sigma(d_T^D)$, served as a good discriminating variable. d_T^D is the distance in the axial plane between the D_s^- vertex and the primary interaction point and $\sigma(d_T^D)$ is its associated uncertainty. $d_T^D/\sigma(d_T^D)$ has better discriminating power than the transverse decay length alone since large values of d_T^D can arise from large uncertainties. The D_s^- candidates were therefore required to satisfy $d_T^D/\sigma(d_T^D) > 4$. The angle α_T^D between the momentum direction of the D_s^- candidate and the direction from the primary to the D_s^- vertex in the axial plane was required to fulfill the condition: $\cos(\alpha_T^D) > 0.9$. This requirement ensured consistency between the direction of the decay vertex and the momentum vector of the D_s^- candidate.

- $d_T^D/\sigma(d_T^D) > 4$,
- $\cos(\alpha_T^D) > 0.9$.

For similar reasons, if the distance d_T^B between the primary and B_s^0 vertex in the axial plane exceeded $4 \cdot \sigma(d_T^B)$, the angle α_T^B between the B_s^0 momentum and the direction from the primary to the B_s^0 vertex in the axial plane was required to satisfy the condition: $\cos(\alpha_T^B) > 0.95$. The B_s^0 meson was allowed to travel further than the D_s^- meson or the distance d_T^B was allowed to be greater than d_T^D, provided that the distance between the B_s^0 and D_s^- vertices, d_T^{BD}, was less than $2 \cdot \sigma(d_T^{BD})$, i.e. the D_s^- meson decay length was consistent with zero.

- If $d_T^B > 4 \cdot \sigma(d_T^B)$, then $\cos(\alpha_T^B) > 0.95$,
- If $d_T^B > d_T^D$, then $d_T^{BD} < 2 \cdot \sigma(d_T^{BD})$.

The selection of the B_s^0, ϕ, and K^{*0} candidates was done by requiring their masses to be in a window around the nominal B_s^0 (5.370 GeV/c^2), ϕ (1.020 GeV/c^2), and K^{*0} (0.892 GeV/c^2) masses, respectively.

- $2.6 < M(\mu^+ D_s^-) < 5.4$,
- $D_s^- \to \phi\pi^- : 1.004 < M(K_1K_2) < 1.034$,
- $D_s^- \to K^{*0}K^- : 0.82 < M(K_1\pi) < 0.95$.

Following this initial selection, the final event samples were selected using a likelihood ratio method. The method assumes that a set of discriminating variables $x_1, ...x_n$ can be constructed for a given event. It also assumes that probability density functions $f^s(x_i)$ for the signal and $f^b(x_i)$ for the background can be built for each variable x_i. The combined tagging variable y is then defined as:

$$y = \prod_{i=1}^{n} y_i; \quad y_i = \frac{f_i^b(x_i)}{f_i^s(x_i)}. \tag{5.5}$$

Selection	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$ (differences)
Muons	penetrate toroid	
	matched SMT+CFT track	
	with hits nSMT ≥ 2, nCFT ≥ 2	
	$p_T(\mu) \geq 2.0$ GeV/c	
	$\lvert\eta\rvert(\mu) < 2$	
	$p(\mu) \geq 3$	
D_s^-	3 tracks (K_1, K_2, π) in same jet as μ	
	with hits: nSMT ≥ 2, nCFT ≥ 2	
	charges: $\mu^+ K_1^+ K_2^- \pi^-$ or conjugate	
	$p_T(K_1, K_2) > 0.7$ GeV/c	$p_T(K_1) > 0.9, p_T(K_2) > 1.8$ GeV/c
	$p_T(\pi) > 0.5$ GeV/c	
	K_1,K_2:$(b_T/\sigma_T)^2 + (b_L/\sigma_L)^2 > 4$	
	3-track vertex: $\chi^2 < 16$	
	$d_T^D/\sigma(d_T^D) > 4$	
	$\cos(\alpha_T^D) > 0.9$	
	$1.004 < M(K_1K_2) < 1.034$	$0.82 < M(K_1\pi) < 0.95$
B_s^0	$\mu^+ D_s^-$ vertex: $\chi^2 < 9$	
	$2.6 < M(\mu^+ D_s^-) < 5.4$ GeV/c^2	
	If $d_T^B > 4 \cdot \sigma(d_T^B)$, $\cos(\alpha_T^B) > 0.95$	
	If $d_T^B > d_T^D$, $d_T^{BD} < 2 \cdot \sigma(d_T^{BD})$	
$\log_{10} y$	$\log_{10} y < 0.12$	$\log_{10} y < 0.16$

Table 5.1: A summary of the criteria used for selecting signal $B_s^0 \to D_s^- \mu^+ X$ candidates. All cuts are listed for the $D_s^- \to \phi\pi^-$ sample and only those cuts which differ for the $D_s^- \to K^{*0}K^-$ sample are listed separately.

The variable x_i can be undefined for some events. In this case, the corresponding variable y_i is set to one. The final selection of the signal is done by applying a cut $y < y_0$. For uncorrelated variables $x_1, ...x_n$, signal selection using the combined variable y gives the best possible tagging performance, i.e., maximal signal efficiency for a given background efficiency.

The following discriminating variables were used:

- Helicity angle, defined as the angle between the D_s^- and K_1 momenta in the (K_1, K_2) or $(K_1\pi)$ center of mass system. The (K_1, K_2) and $(K_1\pi)$ center of mass systems are used for the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ modes, respectively. The helicity angle, ψ, for the signal exhibits a distribution $dN/d(\cos\psi) \sim \cos^2\psi$. Combinatorial background, however, has a flat $\cos\psi$ distribution and hence a cut on $|\cos\psi|$ can be used to suppress background.

- Isolation, computed as Iso $= \mathrm{p^{tot}}(\mu^+\mathrm{D_s^-})/(\mathrm{p^{tot}}(\mu^+\mathrm{D_s^-})+\sum \mathrm{p_i^{tot}})$. The sum $\sum p_i^{tot}$ was taken over all charged particles in the cone $\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$, where $\Delta\eta$ and $\Delta\phi$ are the pseudorapidity and the azimuthal angle with respect to the $(\mu^+D_s^-)$ direction. The μ^+, K_1, K_2 and π^- were not included in the sum. Since the fragmentation characteristics of the b quark are such that most of the momentum is carried by the B hadron (Sec. 3.1.3), the number of extra tracks near the B_s^0 candidate tends to be small. Or in other words, signal B_s^0 events should be well isolated, i.e. have larger values of the variable Iso.

- $p_T(K_1K_2)$ for the $D_s^- \to \phi\pi^-$ mode ; $p_T(K_2)$ for the $D_s^- \to K^{*0}K^-$ mode.

- Invariant mass, $M(\mu^+D_s^-)$.

- χ^2 of the D_s^- vertex fit.

- $M(K_1K_2)$ for the $D_s^- \to \phi\pi^-$ mode; $M(K_1\pi)$ for the $D_s^- \to K^{*0}K^-$ mode.

The probability density functions were constructed using real data events. For each channel, three bands B_1, B_2, and S were defined as:

$$\begin{aligned} B_1 &: 1.75 < M(D_s^-) < 1.79 \text{ GeV}/c^2, \\ B_2 &: 2.13 < M(D_s^-) < 2.17 \text{ GeV}/c^2, \\ S &: 1.92 < M(D_s^-) < 2.00 \text{ GeV}/c^2. \end{aligned}$$

The background probability density function for each variable was then constructed using events from the B_1 and B_2 bands. The signal probability density function was constructed by subtracting the background, obtained as a sum of distributions in the B_1 and B_2 bands, from the distribution of events in band S. The distributions for all variables obtained in this manner are shown in Figs. 5.1 and 5.2 for the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples, respectively. The final selection of events for the analysis was done by applying a cut on the combined variable $\log_{10} y$. The actual value of the cut was selected by requiring the maximal value of $S/\sqrt{S+B_1+B_2}$ since this quantity directly feeds into the sensitivity formula (Eq. 3.20).

- $D_s^- \to \phi\pi^-$: $\log_{10} y < 0.12$,
- $D_s^- \to K^{*0}K^-$: $\log_{10} y < 0.16$.

Figures 5.3 and 5.4 show the $\log_{10} y$ distributions for both signal and background for the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples, respectively.

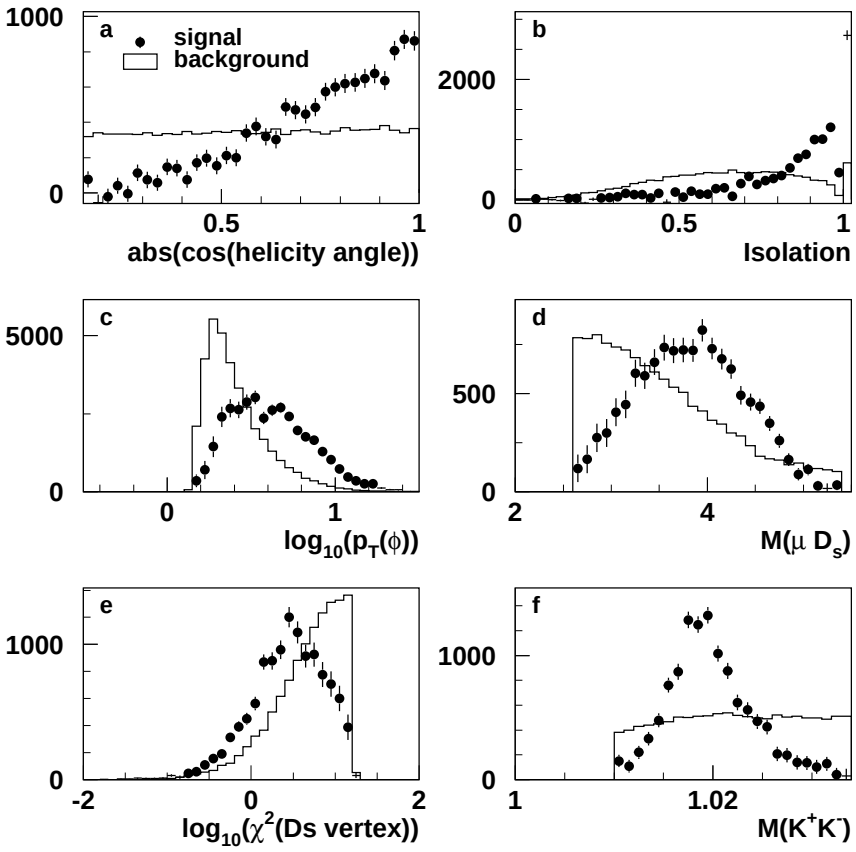


Figure 5.1: $D_s^- \to \phi\pi^-$ *: Probability density functions for both signal (black points) and background (black histogram) for the six discriminating variables:(a) helicity angle, (b) isolation, (c)* $p_T(K_1K_2)$*, (d)* $M(\mu^+D_s^-)$*, (e)* χ^2 *of the* D_s^- *vertex and (f)* $M(K_1K_2)$*.*

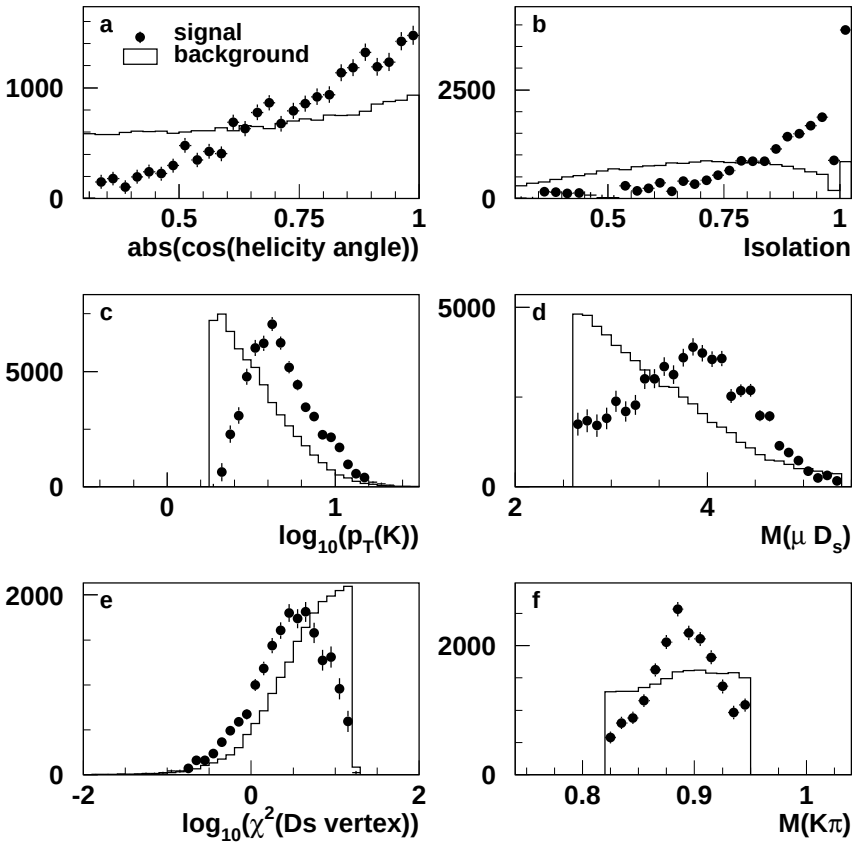


Figure 5.2: $D_s^- \to K^{*0}K^-$: *Probability density functions for both signal (black points) and background (black histogram) for the six discriminating variables:(a) helicity angle, (b) isolation, (c)* $p_T(K_2)$*, (d)* $M(\mu^+ D_s^-)$*, (e)* χ^2 *of the* D_s^- *vertex and (f)* $M(K_1\pi)$*.*

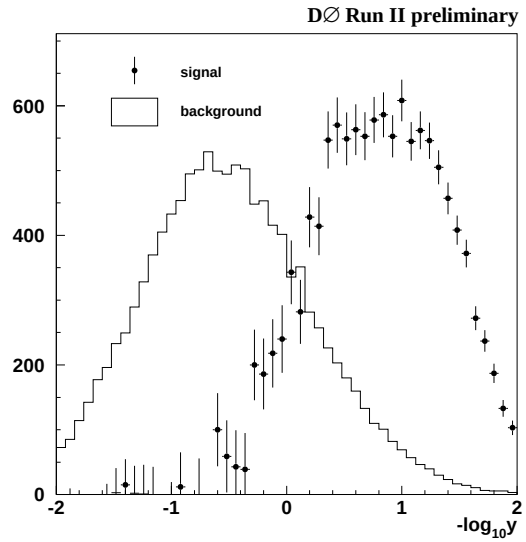


Figure 5.3: $D_s^- \to \phi\pi^-$ *:* $\log_{10} y$ *distributions for both signal (black points) and background (black histogram).*

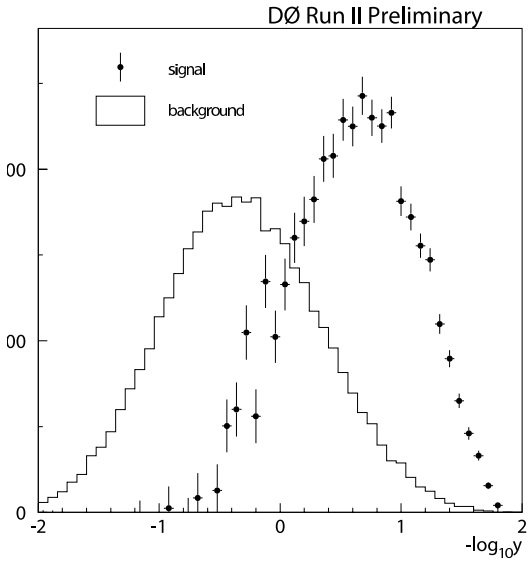


Figure 5.4: $D_s^- \to K^{*0}K^-$ *:* $\log_{10} y$ *distributions for both signal (black points) and background (black histogram).*

5.1.3 Mass Fitting Procedure

$D_s^- \to \phi\pi^-$

A Gaussian was used to describe the $D_s^- \to \phi\pi^-$ signal and an exponential was used to model the background due to "wrong-sign combinations" $(Q_\mu * Q_\pi > 0)$[5]. The $D^- \to \phi\pi^-$ decay has a final state identical to the signal and hence can reflect onto the signal region. It was, therefore, modeled using another single Gaussian. The means and widths (σ) of the two single Gaussians were extracted from the final fit.

Figure 5.5 shows the D_s^- mass distribution for the full untagged B_s^0 sample. The two peaks correspond to the $\mu^+ D^-$ (left) and $\mu^+ D_s^-$ (right) candidates. The total number of signal candidates in the mass peak on the right is 15636 ± 193, while the number of D^- candidates is 4349 ± 152.

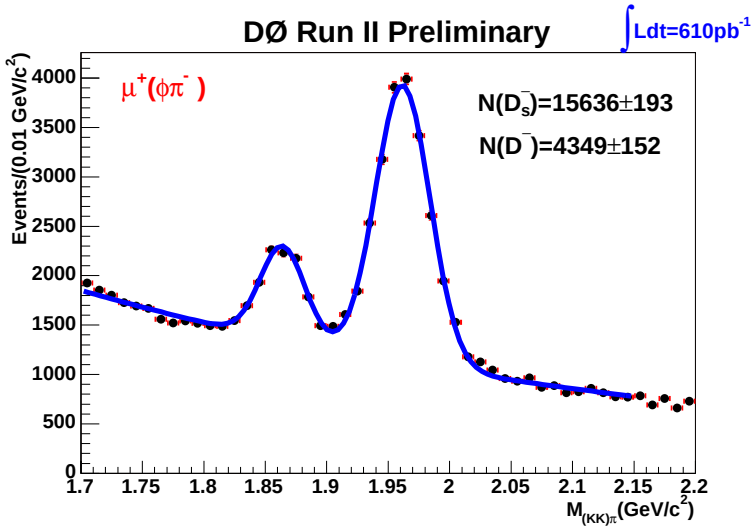


Figure 5.5: *$D_s^- \to \phi\pi^-$: The D_s^- mass distribution for the full B_s^0 sample. The peaks correspond to $\mu^+ D^-$ (left) and $\mu^+ D_s^-$ (right) candidates.*

[5]The 'wrong charge' events are mainly due to combinatorial background, but they can also receive contributions from cases where there is a real D_s^- meson but a fake lepton.

$D_s^- \to K^{*0}K^-$

Figure 5.6 shows the $K^{*0}K^-$ invariant mass distribution after all the selection cuts. Distributions for both the right-sign $D_s^-\mu^+$ combinations ($Q_\mu * Q_\pi < 0$) and the wrong-sign $D_s^-\mu^-$ combinations ($Q_\mu * Q_\pi > 0$) are shown. This plot may look very similar to Fig. 5.5, however, fitting it for the number of signal candidates is complicated by a large reflection under the signal peak. A more complicated fitting procedure is, therefore, used for extracting the number of D_s^- events and is described below.

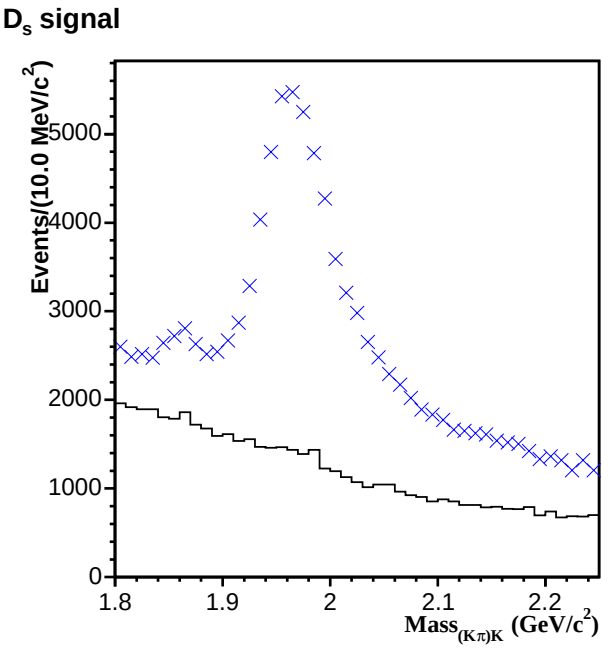


Figure 5.6: *Distribution of the mass of $D_s^- \to K^{*0}K^-$ candidates. Both "right-sign" (crosses) and "wrong-sign" (histogram) combinations are shown.*

Similar to the $D_s^- \to \phi\pi^-$ case a single Gaussian was used to describe the $D_s^- \to K^{*0}K^-$ signal and an exponential was used to model the background due to the wrong-sign combinations. Additionally, to extract an accurate normalization, the following potential reflections were studied and parameterized in the total fit:

- $D^- \to K^+\pi^-\pi^-$ ($\mathcal{B} = 9.2 \pm 0.6\%$):

  A Monte Carlo sample for the above decay which took into account all the underlying amplitudes — $K^{*0}(892)\pi^-$, $K^{*0}(1430)\pi^-$and $K^{*0}(1680)\pi^-$ resonances and $K^+\pi^-\pi^-_{\text{non resonant}}$ — and the interference between them was generated and used for the study. The mode which poses the biggest background is $K^{*0}(892)\pi^-$ ($\mathcal{B} = 1.30 \pm 0.13\%$), where the pion could be mis-identified as a kaon. The other resonances should have smaller impact since we imposed a cut on the mass of the $K^+\pi^-$ combination. Given that we accept all $K^+\pi^-$ combinations in the K^{*0} mass region, the non-resonant contribution could also make an impact since its branching fraction is quite large ($\mathcal{B} = 8.8 \pm 0.9\%$). However, the mass region occupied by the K^{*0} is much smaller than the phase space available to the $K^+\pi^-$ combinations, and thus the impact should be small. Moreover, the helicity cut should further reduce the non-resonant contribution[68].

  Figure 5.7 shows the invariant mass distribution obtained after reconstruction. Two Gaussians were used to model the decay and the means and widths of these Gaussians were extracted from this fit and then used to fit the data. When fitting data the means of these two Gaussians were fixed relative to the D_s^- mean and the two widths were allowed to be scaled by a single factor.

- $D^- \to K^{*0}K^-$ ($\mathcal{B} = 2.9 \pm 0.4 \times 10^{-3}$):

  This decay has a final state identical to the $D_s^- \to K^{*0}K^-$ signal. Its contribution was parameterized by a Gaussian with its mean and sigma fixed relative to the D_s^- Gaussian. The number of $D^- \to K^{*0}K^-$ events was obtained from the final fit. The reconstruction and candidate selection efficiencies from Monte Carlo along with the branching fractions were used to estimate the ratio of $D^- \to K^+\pi^-\pi^-$ and $D^- \to K^{*0}K^-$ events and this ratio was used as a fixed

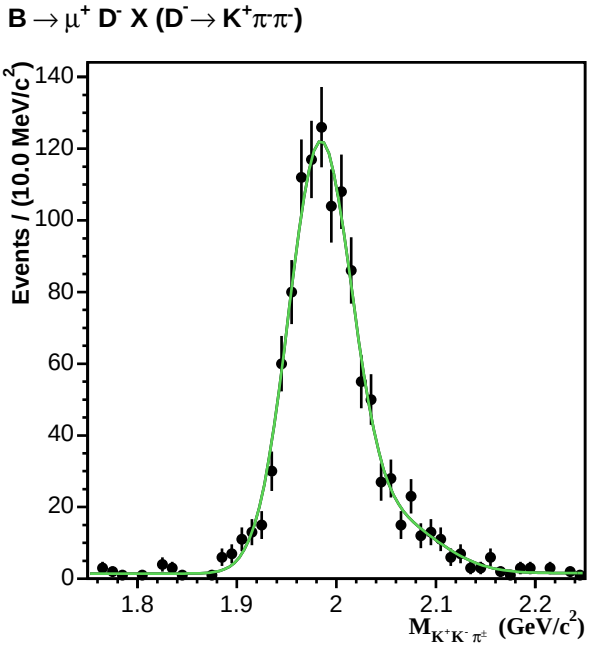


Figure 5.7: *Invariant mass distribution ($D^- \to K^+\pi^-\pi^-$ Monte Carlo).*

parameter when fitting the D_s^- mass distribution in data.

Figure 5.8 shows the same invariant mass distribution as was shown in Figure 5.6. In this plot, the contribution of all the reflections discussed above are shown separately. The total fit gave an estimated 18780 ± 782 $D_s^- \to K^{*0}K^-$ signal events centered at 1.964 GeV/c^2, and a width of 28 MeV. A total of 3233 ± 208 $D^- \to K^{*0}K^-$ events and 14112 ± 910 $D^- \to K^+\pi^-\pi^-$ events were obtained.

5.2 Flavor Tagging

The second B meson (or baryon) in the event was used to tag the initial flavor of the reconstructed B meson. The tagging technique utilized information from identified leptons (muons and electrons) and reconstructed secondary vertices. For re-

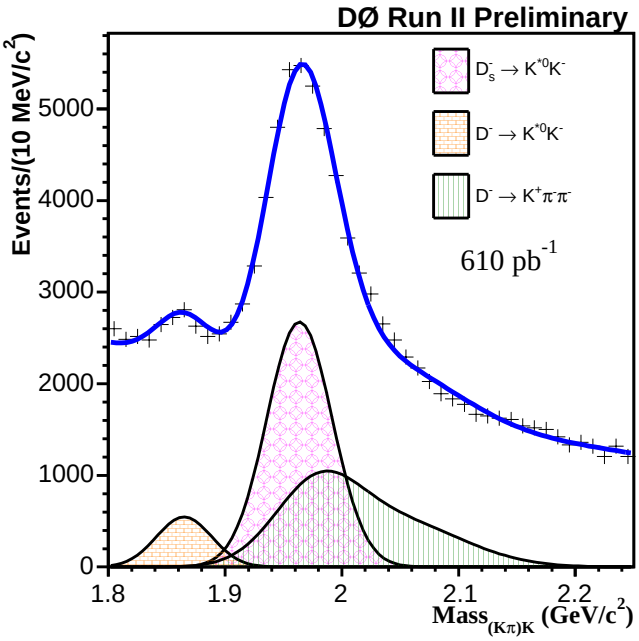


Figure 5.8: *Distribution of the mass of* $D_s^- \to K^{*0}K^-$ *candidates. There is a clear* $D_s^- \to K^{*0}K^-$ *signal peak at* 1.964 GeV/c^2. *The various reflections and their individual contribution to the total fit are given by the legend description.*

constructed $B_s^0 \to D_s^- \mu^+ X$ decays both leptons having the same sign would indicate that one B hadron had oscillated while opposite signs would indicate that neither (or both) had oscillated. The tagging information from the leptons and secondary vertices was combined in an optimal way and the details of the combination can be found in Ref. [69]. The combined tagger was then validated by measuring the B_d^0 mixing frequency, Δm_d. A brief description of the tagging algorithms and the combination procedure (and its validation) is presented here.

5.2.1 Event Selection

The B_d^0 mixing analysis used approximately 460 pb^{-1} of data accumulated by the DØ detector during the period from April 2002 to August 2004. B_d^0 and B^+ mesons were selected using their semileptonic decays $B \to \mu^+ \nu \bar{D}^0 X$ and were classified into two exclusive groups: the "D^*" sample, containing all events with reconstructed $D^{*-} \to \bar{D}^0 \pi^-$ decays, and the "D^0" sample, containing all the remaining events. Both simulation and available experimental results showed the D^* sample to be dominated by $B_d^0 \to \mu^+ \nu D^{*-} X$ decays, while the $\bar{D}^0$ sample was dominated by $B^+ \to \mu^+ \nu \bar{D}^0 X$ decays. Exploiting isotopic invariance and the fact that semileptonic B decays are typically saturated by decays to D, D^* and D^{**} , it was determined that B_d^0 (85%) and B^+ (15%) decays were the main contributors to the "D^*" sample. The D^0 sample correspondingly had a 85% contribution from B^+ and a 15% contribution from B_d^0 decays.

Figure 5.9 shows the $(K\pi)$ invariant mass distribution using the D^0 sample. The narrow peak corresponds to $\bar{D}^0 \to K^+ \pi^-$ decays while the wider peak in the lower mass region originates from the partially reconstructed decay $\bar{D} \to K^+ \pi^- X$. The $(K\pi)$ invariant mass distribution was fit with two Gaussians and a background function, and a total of 81912 ± 511 D^0 events were obtained.

The mass difference $\Delta M = M(\bar{D}^0 \pi) - M(\bar{D}^0)$ for events with $1.75 < M(\bar{D}^0) < 1.95$ GeV/c^2 in the D^* sample is shown in Fig. 5.10. The signal (peak corresponding to the production of $\mu^+ D^{*-}$ events) was modeled by the sum of two Gaussian functions while the background was described by the sum of exponential and first-order polynomial functions. The fit to the mass distribution yielded a total of 39735 ± 341 D^{*-} events.

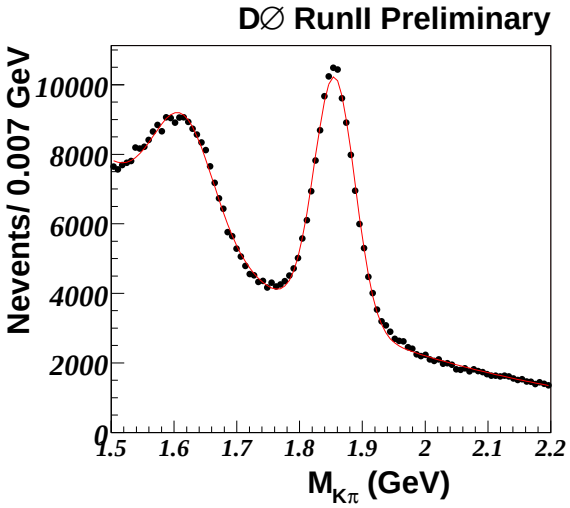


Figure 5.9: *Invariant mass of the $K\pi$ system for selected $\mu^+K^+\pi^-$ candidates. A total of* 81912 ± 511 *$\bar{D}^0$ candidates were obtained in 460 pb^{-1} of data [69].*

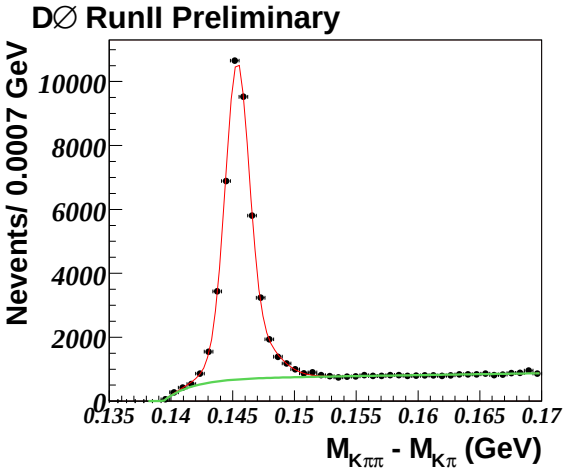


Figure 5.10: *The mass distribution corresponding to D^{*-} events. A total of* 39735 ± 341 *D^{*-} events were obtained [69].*

5.2.2 Flavor Tagging Method

Typically many different techniques are used to identify the original flavor (b or $\bar{b}$) of a heavy quark producing a reconstructed B meson (see Sec. 3.3). While some techniques

perform well by themselves, others provide only a weak separation between flavors. The result can be significantly improved by combining the different tagging techniques in an optimal way. Such a combination is done using a likelihood ratio method which is very similar to the one used for the event selection process in Sec. 5.1.2.

According to this method a set of discriminating variables $x_1, ..., x_n$ can be constructed for a given event. These discriminating variables, by definition, have different distributions for b and $\bar{b}$ flavors. Moreover, they can either be continuous, like the average charge of a jet, or discrete, like the charge of an electron or muon. The combined tagging variable y^t is defined as:

$$y^t = \prod_{i=1}^{n} y_i^t; \quad y_i^t = \frac{f_i^{\bar{b}}(x_i)}{f_i^b(x_i)}, \tag{5.6}$$

where $f_i^b(x_i)$ is the probability density function (PDF) for a given variable x_i for an initial b quark and $f_i^{\bar{b}}(x_i)$ is the corresponding PDF for an initial $\bar{b}$ quark. If a given variable x_i is undefined for some events (for eg. some events do not have an identified muon on the opposite side), then the corresponding variable y_i^t is set to one. For $y^t < 1$ an initial b flavor is more probable while $y^t > 1$ implies that the $\bar{b}$ flavor is more probable. Consequently, an event with $y^t < 1$ is tagged as a b quark while an event with $y^t > 1$ is tagged as a $\bar{b}$ quark. A more convenient tagging variable is defined as:

$$d = (1 - y^t)/(1 + y^t) \tag{5.7}$$

and

$$d > 0 \Longrightarrow b \text{ quark},$$
$$d < 0 \Longrightarrow \bar{b} \text{ quark}.$$

The variable d assumes values between -1 and +1 and larger $|d|$ values correspond to a higher tagging purity. For uncorrelated variables and with perfect modeling in the PDFs, d is the best possible discriminant, and its absolute value gives the dilution for a given event.

The discriminating variables used for this study were constructed using properties of the b quark opposite to the reconstructed B hadron, and are described below. The PDFs for these discriminating variables were obtained using a data sample of $B \to \mu^+\nu D^{*-}$ events. The selected events were required to have visible proper decay length less than 500 μm. Consequently, non-oscillating $B^0_d \to \mu^+\nu D^{*-}$ decays dominate (since the B^0_d oscillation frequency is smaller) and the initial state of the b-quark can be determined using the charge of the signal muon. According to Monte Carlo estimates, the purity of such a sample is 0.956 ± 0.007, where the error stems from the uncertainty in the branching ratios of the decays involved [69].

Opposite Side Muon Tagging

Opposite side muons were used for flavor tagging if $\cos\phi(\mathbf{p}_\mu, \mathbf{p}_B) < 0.8$, where $\mathbf{p}_B$ is the three momentum of the reconstructed B hadron. If more than one muon satisfied this condition, the muon with the highest number of hits in the muon chambers was used. If the selected muons had the same number of hits as well, the muon with the highest p_T was used.

For each muon satisfying the above criteria, a *muon jet charge* Q_J^μ, was constructed:

$$Q_J^\mu = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}, \tag{5.8}$$

where q^i and p_T^i are the charge and transverse momentum of particle i. The sum was taken over all charged particles (including the muon) satisfying the condition $\Delta\mathcal{R} = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$, where $\Delta\phi$ and $\Delta\eta$ were computed with respect to the muon direction. The daughters of the reconstructed B hadron were explicitly removed from the sum. Figure 5.11(a) shows the muon jet charge distribution for the case the muon has hits in three layers of the muon detector while Fig. 5.11(b) shows the same distribution for the case the muon has less than three hits in the muon chambers.

An additional discriminating variable used is the muon's *transverse momentum relative to the nearest jet* or $p_T^{\rm rel}$. This variable is good for discriminating against muons from cascade decays ($b \to c \to \mu$) which have a smaller $p_T^{\rm rel}$ on average. Cascade muons have the same charge as the b quark from the reconstructed side ($b^{\rm rec}$) and hence can wrongly tag the initial flavor, thereby decreasing the tagging purity. Figures 5.11(c) and (d) show the normalized $p_T^{\rm rel}$ distributions for events with $q(\mu) \times q(b^{\rm rec}) > 0$ (i.e. muons from cascade decays) and $q(\mu) \times q(b^{\rm rec}) < 0$. As was expected, the p_T spectrum for the $q(\mu) \times q(b^{\rm rec}) > 0$ decays is softer.

Opposite Side Electron Tagging

Opposite side electrons with $\cos\phi(\mathbf{p}_e, \mathbf{p}_B) < 0.5$ were used for flavor tagging. These electrons were required to have $p_T > 2.0$ GeV/c and $|\eta| < 1.1$. Additionally, the electron track needed to have at least two SMT hits and two CFT hits. The elec-

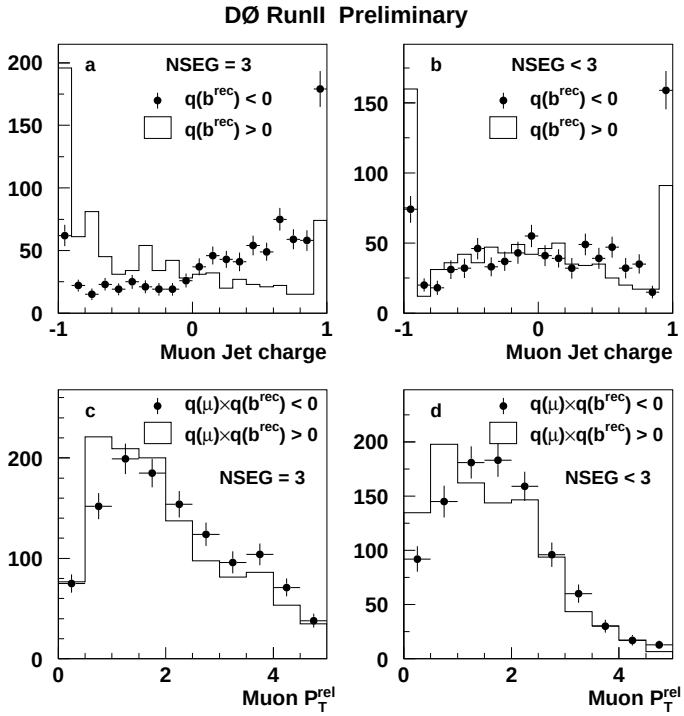


Figure 5.11: *Normalized distributions of muon jet charge and muon* p_T^{rel}*. Figures (a) and (c) correspond to a tagging muon with hits in three layers of the muon detector, and Figs. (b) and (d) correspond to a tagging muon with hits in less than three layers of the muon detector [69].*

tron candidate was also required to satisfy the electron selection criteria described in Sec. 4.4.4. The charge of the electron was used as a discriminating variable.

Opposite Side Secondary Vertex Tagging

Secondary vertices on the opposite side were also used for flavor tagging. The secondary vertex was required to have at least two particles associated with it that have a transverse impact parameter significance greater than three. The distance d_{xy} from the primary to the secondary vertex was required to satisfy the condition:

$d_{xy} > 4\sigma(d_{xy})$. The momentum of the secondary vertex, $\mathbf{p}_{SV}$, was defined as the sum of momenta of all particles associated with the secondary vertex, and was required to fulfill the condition: $\cos\phi(\mathbf{p}_{SV}, \mathbf{p}_B) < 0.8$. Secondary vertices containing any tracks from the reconstructed B hadron were explicitly removed. A *secondary vertex charge* Q_{SV} was defined as:

$$Q_{SV} = \frac{\sum_i (q^i p_L^i)^k}{\sum_i (p_L^i)^k}, \tag{5.9}$$

where the sum is taken over all particles included in the secondary vertex and p_L^i is the longitudinal momentum of a given particle with respect to the direction of the secondary vertex momentum. The value of k was optimized by studies of the Q_{SV} variable and $k = 0.6$ was chosen. Figures 5.12(a) & (b) show the Q_{SV} distributions for events with and without an identified muon, respectively.

The *transverse momentum of the secondary vertex*, p_T^{SV}, was also used as a discriminating variable. This variable can potentially help discriminate against events with fake vertices. Such events are not sensitive to the charge of the reconstructed B meson and their contribution consequently decreases the tagging purity. Usually these events are constructed from low momentum tracks and hence their p_T^{SV} spectrum is softer. Figure 5.12 shows the $\log_{10}(p_T^{SV})$ spectrum and events with $q(Q_{SV}) \times q(b^{\text{rec}}) > 0$ have a somewhat smaller p_T^{SV} distribution owing to the larger fractions of fake vertices. The discriminating power of this variable is somewhat small. It does, however, help by increasing the purity of the combined tagging.

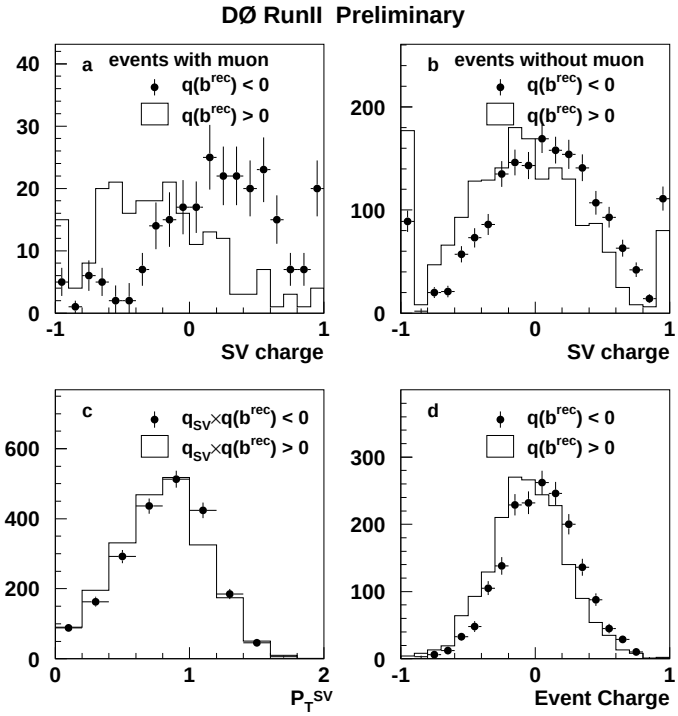


Figure 5.12: *Normalized distributions for (a) the secondary vertex charge for events with a tagging muon, (b) the secondary vertex charge for events without a tagging muon, (c) the secondary vertex p_T, and (d) the event charge [69].*

Event Charge Tagging

The *event charge*, Q_{EV}, defined as:

$$Q_{EV} = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i} \tag{5.10}$$

is another discriminating variable used. The sum is taken over all charged particles with $p_T > 0.5$ GeV/c and with $\Delta\mathcal{R} = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} > 1.5$, where $\Delta\phi$ and $\Delta\eta$ are computed with respect to the reconstructed B hadron direction. This variable is strongly correlated with the muon jet charge and hence is not used for events with

an identified muon. The Q_{EV} distribution is shown in Fig. 5.12(d).

Combined Tagger

The following procedure was followed for combining the different taggers:

A) If a muon was found:

- The muon jet charge, muon $p_T^{\rm rel}$, and the secondary vertex charge were used (*muon tagger*).

B) If a muon was not found:

- If an electron was found, the electron charge was used (*electron tagger*).

Otherwise

- The secondary vertex charge, the p_T^{SV}, and the event charge were used (*SV tagger*).

The resulting distribution of the combined tagging variable d shows good discrimination between b and $\bar{b}$ quarks and is shown in Fig. 5.13.

5.2.3 B_d^0 Mixing and Tagger Calibration

Events tagged following the procedure described above were then divided into seven groups according to their measured visible proper decay length or VPDL. The number of oscillating ($N_i^{\rm mix}$) and non-oscillating ($N_i^{\rm unm}$) $\mu^+ D^{*-}$ events in each VPDL bin i were determined by fitting the D^{*-} signal in the $M(\bar{D^0}\pi) - M(\bar{D^0})$ distributions of the D^* sample. Similarly, the number of oscillating and non-oscillating $\mu^+ \bar{D^0}$ events

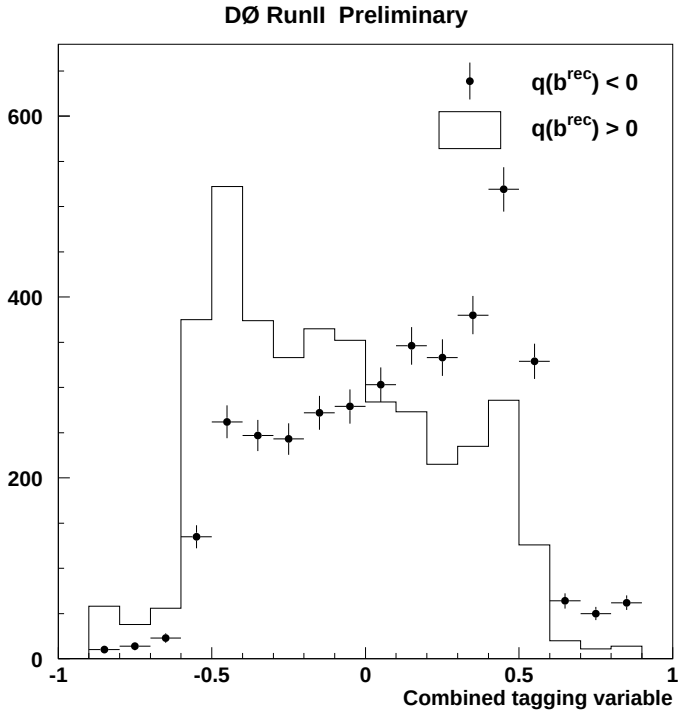


Figure 5.13: *Normalized distributions of the combined tagging variable d for reconstructed b and $\bar{b}$ quarks [69].*

were determined by fitting the $\bar{D^0}$ signal in the $K\pi$ invariant mass distributions of the D^0 sample.

The measured asymmetry, $A_i^{\rm meas}$, in each VPDL bin was then calculated using Eq. 3.17 for both the D^* and D^0 samples. An expected asymmetry, A_i^e, was also computed for each VPDL bin (see Sec. 5.3 for details of the procedure), and the mixing frequency, Δm_d, and the tagging purity, η_s, were determined from a minimization of $\chi^2(\Delta m_d, \eta_s)$ where

$$\chi^2(\Delta m_d, \eta_s) = \sum_i \frac{(A_i^{\rm meas} - A_i^e(\Delta m_d, \eta_s))^2}{\sigma^2(A_i^{\rm meas})}. \tag{5.11}$$

The fit was performed for different intervals of $|d|$: four samples of events having $|d| > 0.3$, $0.22 < |d| < 0.3$, $0.3 < |d| < 0.45$, and $|d| > 0.45$ were used for this study. The sample with $|d| > 0.3$ yielded the highest tagging power (or $\epsilon\mathcal{D}^2$) and hence was used as the default sample. Figure 5.14 shows the measured asymmetry for this sample for both the D^* (top) and D^0 (bottom) events. A nice oscillation pattern can be seen for the D^* sample (since this sample largely consists of oscillating B^0_d mesons with a small contribution from the non-oscillating B^+ mesons) while the oscillation signal is highly suppressed for the D^0 sample (owing to the fact that it largely consists of non-oscillating B^+ mesons). The dilutions for B^0_d and B^+ were found to be consistent with each other within their statistical error and hence were constrained to be equal in the final fit.

The results of the fits are listed below:

$$\mathcal{D}(B^0_d) = 0.414 \pm 0.023(\text{stat.}) \pm 0.017(\text{syst.}) \tag{5.12}$$

$$\mathcal{D}(B^+) = 0.368 \pm 0.016(\text{stat.}) \pm 0.008(\text{syst.}) \tag{5.13}$$

$$\mathcal{D} = 0.384 \pm 0.013(\text{stat.}) \pm 0.008(\text{syst.}) \quad [\mathcal{D}(\mathrm{B^0_d}) = \mathcal{D}(\mathrm{B^+})] \tag{5.14}$$

$$\epsilon\mathcal{D}^2 = 1.94 \pm 0.14(\text{stat.}) \pm 0.09(\text{syst.}) \tag{5.15}$$

$$\Delta m_d = 0.501 \pm 0.030(\text{stat.}) \pm 0.016(\text{syst.})\ \text{ps}^{-1} \tag{5.16}$$

The measured B^0_d mixing frequency, Δm_d, was found to be in good agreement with the world average of $\Delta m_d = 0.502 \pm 0.007$ ps^{-1} [2]. The central value of the dilution $\mathcal{D}$ (or the purity value $\eta_s = 0.692$) thus obtained was used as an input to the B^0_s mixing fit with the variation of dilution between B^0_d and B^+ events included in the total systematic error evaluation.

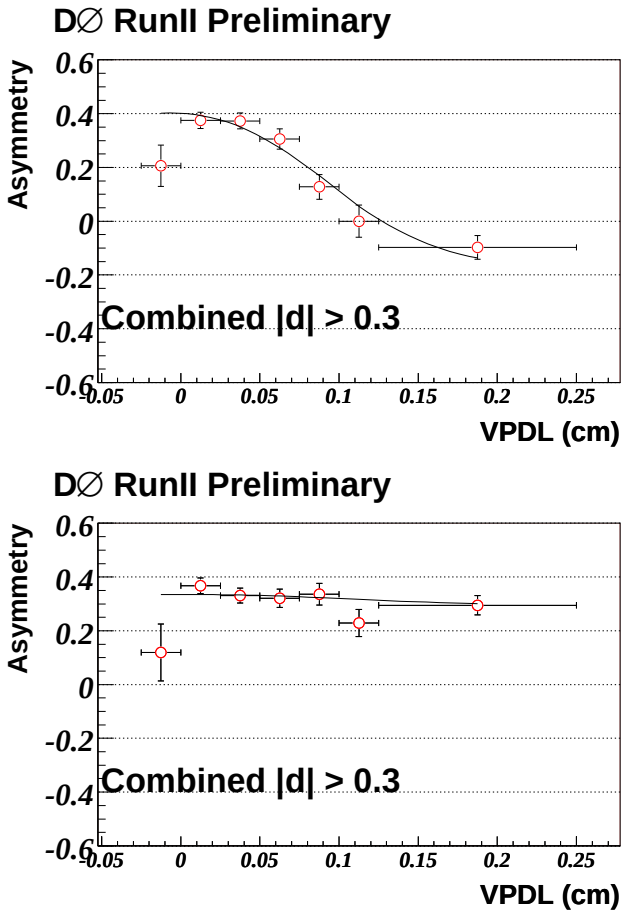


Figure 5.14: *The measured asymmetry between non-oscillated and oscillated events for the D^* sample (top) and the D^0 sample (bottom). The expected oscillation curves are superimposed on the plots [69].*

5.2.4 Tagged B_s^0 Samples

Both the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples were flavor tagged using the exact same combined tagging procedure as described above ($|d| > 0.3$). Figures 5.15 and 5.16 show the invariant mass distributions of tagged $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ candidates, respectively. The fits returned 1917 ± 66 tagged $D_s^- \to \phi\pi^-$ and 2247 ± 316 tagged $D_s^- \to K^{*0}K^-$ candidates indicating a $\sim 12\%$ tagging efficiency (when compared to fit results of Figs. 5.5 and 5.8).

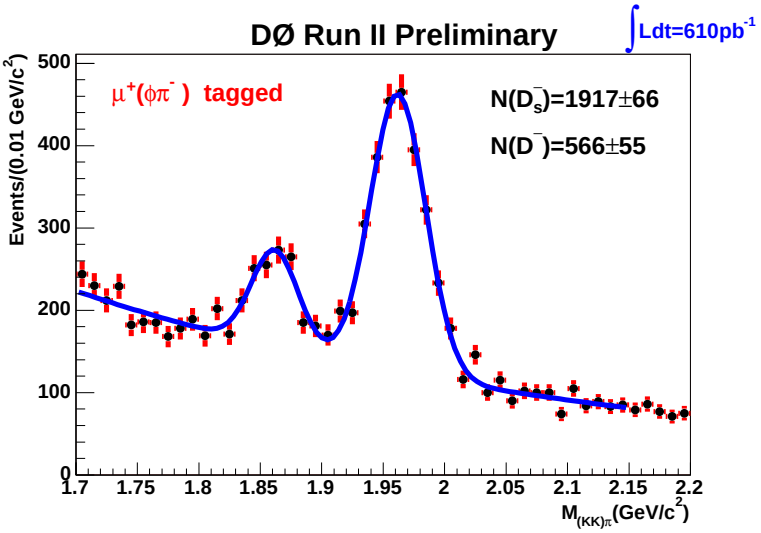


Figure 5.15: *Distribution of the mass of tagged $D_s^- \to \phi\pi^-$ candidates. The peaks correspond to $\mu D^\pm$ (left) and μD_s (right) candidates.*

5.3 Expected Asymmetry

The D_s^- sample is composed mostly of B_s^0 mesons with some contributions from B_u and B_d^0 mesons. Different species of B mesons behave differently with respect to oscillations; neutral B_d^0 and B_s^0 mesons oscillate while charged B_u mesons do not os-

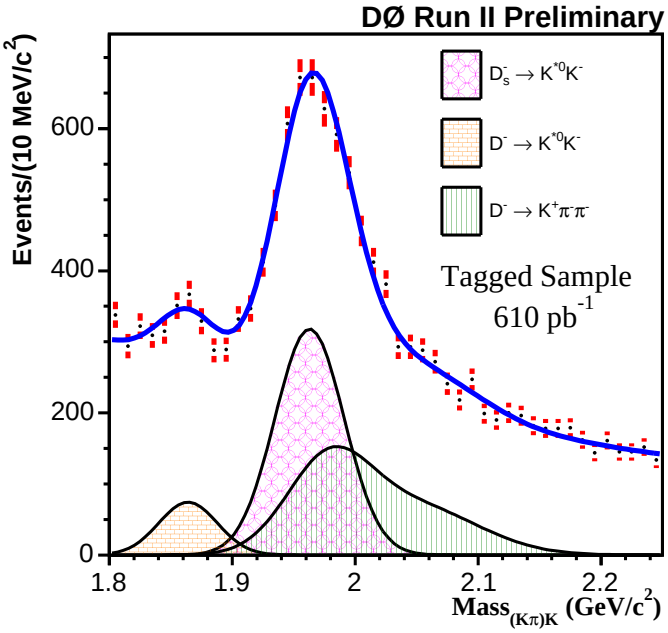


Figure 5.16: *Distribution of the mass of tagged $D_s^- \to K^{*0}K^-$ candidates. The different reflections and their individual contribution to the total fit are given by the legend description.*

cillate. Moreover, contributions of b-baryons to the sample composition are expected to be small and so are neglected. For a given type of B-hadron (i.e. s, d, u), the distribution of the visible proper decay length x for unmixed ("unm") and mixed ("mix") events is given by:

$$n_s^{unm/mix}(x) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5 \cdot [1 \pm (2\eta_s - 1)\cos(\Delta m_s \cdot Kx/c)], \quad (5.17)$$

$$n_d^{unm/mix}(x) = \frac{K}{c\tau_{B_d}} \exp\left(-\frac{Kx}{c\tau_{B_d}}\right) \cdot 0.5 \cdot [1 \mp (2\eta_s - 1)\cos(\Delta m_d \cdot Kx/c)], \quad (5.18)$$

$$
\begin{aligned}
n^{unm}_{DsDs}(x) &= n^{mix}_{DsDs}(x) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5, && (5.19)\\
n^{unm}_{u}(x) &= \frac{K}{c\tau_{B_u}} \exp\left(-\frac{Kx}{c\tau_{B_u}}\right) \cdot (1-\eta_s), && (5.20)\\
n^{mix}_{u}(x) &= \frac{K}{c\tau_{B_u}} \exp\left(-\frac{Kx}{c\tau_{B_u}}\right) \cdot \eta_s.
\end{aligned}
$$

where K is the K-factor as described before, and τ is the lifetime of the B-hadron taken from the PDG [2]. The D_s^- charge has different correlations with the b-quark flavor in the B_u or B_d^0 decays with respect to the B_s^0 semileptonic decays, and Eqs. 5.20 and 5.18 take this into account.

The transition to the measured VPDL, x^M, is achieved by the integration over the K-factors and resolution functions:

$$
.N^{mix,\ unm}_{(d,u,s),\ j}(x^M) = \int dx\ Res_j(x-x^M, x) \cdot \mathit{Eff}_j(x) \cdot \theta(x) \int dK\ D_j(K) \cdot n^{mix,\ unm}_{(d,u,s),\ j}(x, K). \tag{5.21}
$$

$Res_j(x-x^M, x)$ is the detector resolution of the VPDL and $\mathit{Eff}_j(x)$ is the reconstruction efficiency for a given decay channel j of this type of B meson. The step function $\theta(x)$ takes into account that only positive values of x are possible (x^M can have negative values owing to resolution effects) and the function $D_j(K)$ denotes the normalized K-factor distribution for a given channel j. The functions $D(K)$, $\mathit{Eff}(x)$, and $Res(x)$ are determined from Monte Carlo simulations (see Secs. 5.4.2, 5.4.3, and 5.4.4).

The expected number of mixed/unmixed events in the i-th bin of VPDL is equal to

$$
N_i^{e,mix/unm} = \int_i dx^M \left(\sum_{f=u,d,s} \sum_j (Br_j \cdot N^{mix/unm}_{f,\ j}(x^M)) \right). \tag{5.22}
$$

The integration $\int_i dx^M$ is taken over a given interval i, the sum $\sum_j$ is taken over all decay channels $B \to \mu^+ \nu D^{*-} X$ and Br_j is the branching ratio of a given channel j.

Finally, the expected value of the asymmetry, A_i^e, for the i-th VPDL bin is given by:

$$A_i^e(\Delta m, \eta_s) = \frac{N_i^{e,unm} - N_i^{e,mix}}{N_i^{e,unm} + N_i^{e,mix}}. \tag{5.23}$$

For the computation of A_i^e, the B meson lifetimes and the branching rates, Br_j, were taken from Ref. [2]. The functions $D_j(K)$, $Res_j(x)$ and $Eff_j(x)$ were obtained from Monte Carlo simulations and are discussed in more detail in the following sections. Uncertainties in all these inputs contribute to the total systematic error and are discussed in Sec. 5.7.

Zero-lifetime events ($c\bar{c}$) are expected to cancel out in the asymmetry and their treatment is discussed separately in Sec. 5.4.1.

5.4 Inputs For Expected Asymmetry

We used the following measured parameters for B mesons from the PDG [2] as inputs to the fitting procedure:

- $c\tau_{B_u} = c\tau_{B^+} = 501$ μm
- $c\tau_{B_d} = c\tau_{B^0} = 460$ μm
- $c\tau_{B_s} = 438$ μm
- Δm_d= 0.502 ps^{-1}

The tagging purity, η_s, was measured in data using the opposite side combined tagging algorithm (see Sec. 5.2.3): $\eta_s = 0.692$.

5.4.1 Sample composition

To determine the composition of the selected $\mu^+ D_s^-$ sample, the following B meson decay channels were considered:

(1) $B_s^0 \to \mu^+ \nu D_s^-$

(2) $B_s^0 \to \mu^+ \nu D_s^{-*} \to \mu^+ \nu D_s^- X$

(3) $B_s^0 \to \mu^+ \nu D_{s0}^{*-} \to \mu^+ \nu D_s^- X$

(4) $B_s^0 \to \mu^+ \nu D_{s1}^{'-} \to \mu^+ \nu D_s^- X$

(5) $B_s^0 \to \tau^+ \nu D_s^- X \;\; (\tau^+ \to \mu^+ \nu \bar{\nu})$

(6) $B_s^0 \to D_s^- D_s^+ X \;\; (D_s^+ \to \mu^+ \nu X)$

(7) $B_s^0 \to D_s^- D X \;\; (D \to \mu^+ \nu X)$

(8) $B^- \to D_s^- D X \;\; (D \to \mu^+ \nu X)$

(9) $\bar{B_d^0} \to D_s^- D X \;\; (D \to \mu^+ \nu X)$.

The different Monte Carlo samples that were studied are listed in Appendix A.1. The latest PDG values [2] were used for the branching fractions ($\mathcal{B}$) of the above decays, and EvtGen [62] inputs were used for the branching fractions that were not listed in the PDG.

(1-4) $\mathcal{B}(B_s^0 \to \mu^+ \nu D_s^- X) = (7.9 \pm 2.4)\%$

- $\mathcal{B}(B_s^0 \to \mu^+ \nu D_s^-) = 2.0\%$
- $\mathcal{B}(B_s^0 \to \mu^+ \nu D_s^{-*}) = 5.3\%$

- $\mathcal{B}(B_s^0 \to \mu^+ \nu D_{s0}^{*-}) = 0.19\%$
- $\mathcal{B}(B_s^0 \to \mu^+ \nu D_{s1}^{'-}) = 0.35\%$

The total semileptonic $B_s^0 \to \mu^+ \nu D_s^- X$ branching fraction was taken from the PDG while the fractions of exclusive channels were obtained from EvtGen.

(5) $\mathcal{B}(B_s^0 \to \tau^+ \nu D_s^- X) = 2.86\%$ (EvtGen)

(6) $\mathcal{B}(B_s^0 \to D_s^- D_s^+ X) = 23^{+21}_{-13}\%$

(7) $\mathcal{B}(B_s^0 \to D_s^- D X) = 15.4\%$ (EvtGen)

(8) $\mathcal{B}(B^- \to D_s^- D X) = 10.5\%$ (EvtGen)

(9) $\mathcal{B}(B_d^0 \to D_s^- D X) = 10.5\%$ (EvtGen)

Other useful branching fractions are listed below:

- $\mathcal{B}(D_s^{-*} \to D_s^- X) = 100\%$
- $\mathcal{B}(D_{s0}^{*-} \to D_s^- X) = 100\%$
- $\mathcal{B}(D_{s1}^{'-} \to D_s^- X) = 100\%$
- $\mathcal{B}(D_s^+ \to \mu^+ \nu X) = 6.3\%$
- $\mathcal{B}(D^0 \to \mu^+ \nu X) = 6.5\%$
- $\mathcal{B}(D^+ \to \mu^+ \nu X) = 17.2\%$
- $\mathcal{B}(\bar{b} \to B^0) = 39.7 \pm 1.0\%$
- $\mathcal{B}(\bar{b} \to B^+) = 39.7 \pm 1.0\%$

- $\mathcal{B}(\bar{b} \to B_s^0) = 10.7 \pm 1.1\%$

Taking into account the reconstruction efficiencies of the different decay channels and their corresponding branching rates, contributions from the different processes were determined separately for the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples. These are listed in Table 5.2.

Decay	$D_s^- \to \phi\pi^-$ (%)	$D_s^- \to K^{*0}K^-$ (%)
$B_s^0 \to \mu^+\nu D_s^-$	21.7 ± 1.2	22.8 ± 1.7
$B_s^0 \to \mu^+\nu D_s^{-*} \to \mu^+\nu D_s^- X$	60.7 ± 3.2	55.1 ± 4.0
$B_s^0 \to \mu^+\nu D_{s0}^{*-} \to \mu^+\nu D_s^- X$	1.4 ± 0.5	1.2 ± 0.1
$B_s^0 \to \mu^+\nu D_{s1}^{'-} \to \mu^+\nu D_s^- X$	3.1 ± 0.5	3.0 ± 0.2
$B_s^0 \to \tau^+\nu D_s^- X \ \ (\tau^+ \to \mu^+\nu\bar{\nu})$	1.3 ± 0.5	1.6 ± 0.5
$B_s^0 \to D_s^- D_s^+ X \ \ (D_s^+ \to \mu^+\nu X)$	$2.9^{+1.1}_{-2.1}$	$4.2^{+1.4}_{-3.3}$
$B_s^0 \to D_s^- DX \ \ (D \to \mu^+\nu X)$	0.9 ± 0.5	0.9 ± 0.3
$B^- \to D_s^- DX \ \ (D \to \mu^+\nu X)$	4.0 ± 1.3	5.6 ± 1.9
$\bar{B}_d^0 \to D_s^- DX \ \ (D \to \mu^+\nu X)$	4.1 ± 1.4	5.7 ± 1.9
$c\bar{c} \to \mu^+ D_s^- X$	3.5 ± 2.5	3.5 ± 2.5

Table 5.2: *The contributions of different decays to the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples.*

To determine the uncertainties (shown in Table 5.2) we varied the branching fractions used as inputs by their corresponding uncertainties one by one and used the maximum deviation as an estimate of the sample composition uncertainty. The Monte Carlo statistical uncertainty was not taken into account. For the $D_s^- D$ branching fractions taken from EvtGen we used a 25% relative uncertainty [70]. The relative fractions of the signal $B_s^0 \to \mu^+\nu D_s^- X$ were not varied since this contribution to the systematic error is taken into account by the variation of the corresponding K-factors (Sec. 5.7). The $B_s^0 \to D_s^- D_s^+ X$ branching fraction was varied from its PDG value (23%) to its EvtGen value (4.5%). For most of the processes the maximum

deviation occured when the semileptonic signal branching fraction was varied within its uncertainty $(7.9 \pm 2.4)\%$.

The reconstruction efficiencies used in determining the sample composition did not include any lifetime cuts at this point. The effects of the lifetime cuts are discussed separately in Section 5.4.3.

In addition to the above decays, $c\bar{c}$ pairs originating from gluon splitting can potentially contaminate the B_s^0 data sample. In principle tagging on the opposite side should suppress the contamination since a tag implies the presence of a b-quark on that side. The number of $c\bar{c}$ events in each VPDL bin were taken from the study done in Ref. [71] to be $3.5 \pm 2.5\%$ after tagging. A zero asymmetry was assigned to these events. The $c\bar{c}$ contribution was increased to 6.0% for systematic studies.

5.4.2 K-factors

Using the definition from Sec. 3.2, the K-factors for all the above B meson decays were obtained using Monte Carlo simulations. The K-factor was defined as:

$$K \equiv \frac{p_T(D_s\mu)}{p_T(B_s^0)}, \tag{5.24}$$

where p_T denotes the absolute value of transverse momentum, and generator level information was used for its computation. We also verified that using reconstructed values for p_T did not cause any bias. Following the definition in Eq. 5.24 the K-factors for all considered B decays were calculated before applying lifetime biasing cuts. Addition of these cuts did not bias the distributions.

Figures 5.17(a) and (b) show the K-factor distributions for the $B_s^0 \to D_s^- \mu^+ \nu X$ decays for the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples, respectively. Note that since the K-factors in Eq. 5.24 were defined as the ratio of transverse momenta, they can exceed one. Table 5.3 lists the *means* and *root-mean-squared* values (RMS) of the K-factor distributions for the above decays. As expected, the K-factors for decays involving D_s^{-*}, D_{s0}^{*-} and $D_{s1}^{'-}$ have lower mean values because more decay products are missing.

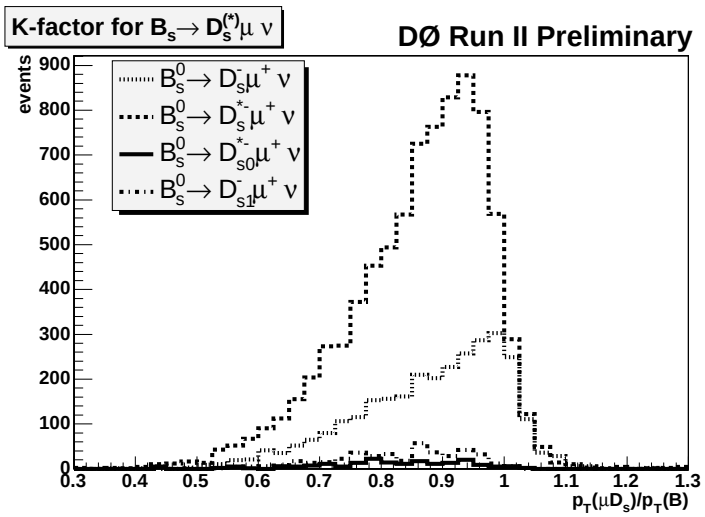
K-factor distributions for some of the background B decays listed in Sec. 5.4.1 are shown in Figs. 5.18(a)-(d). The means and RMS of all these decays for both the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples are given in Table 5.4. These processes have considerably smaller average value of the K-factor than those in Table 5.3.

Decay	$D_s^- \to \phi\pi^-$		$D_s^- \to K^{*0}K^-$	
	Mean	RMS	Mean	RMS
$B_s^0 \to \mu^+\nu D_s^-$	0.881	0.121	0.885	0.115
$B_s^0 \to \mu^+\nu D_s^{-*} \to \mu^+\nu D_s^- X$	0.861	0.113	0.854	0.102
$B_s^0 \to \mu^+\nu D_{s0}^{*-} \to \mu^+\nu D_s^- X$	0.815	0.127	0.823	0.098
$B_s^0 \to \mu^+\nu D_{s1}^{'-} \to \mu^+\nu D_s^- X$	0.830	0.105	0.812	0.101

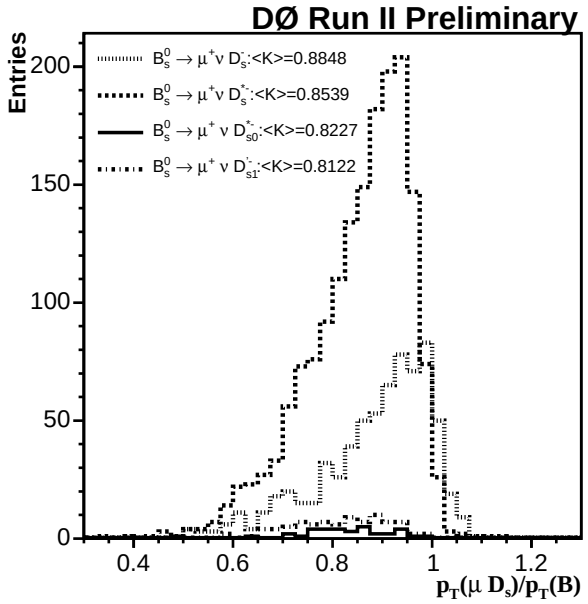
Table 5.3: *The means and root-mean-squared (RMS) values of the K-factor distributions for the semileptonic $B_s^0 \to D_s^-\mu^+\nu X$ decays for both the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples.*

Decay	$D_s^- \to \phi\pi^-$		$D_s^- \to K^{*0}K^-$	
	Mean	RMS	Mean	RMS
$B_s^0 \to \tau^+\nu D_s^- X \ \ (\tau^+ \to \mu^+\nu\bar{\nu})$	0.720	0.128	0.724	0.117
$B_s^0 \to D_s^- D_s^+ X \ \ (D_s^+ \to \mu^+\nu X)$	0.762	0.135	0.782	0.120
$B_s^0 \to D_s^- DX \ \ (D \to \mu^+\nu X)$	0.674	0.105	0.703	0.104
$B^- \to D_s^- DX \ \ (D \to \mu^+\nu X)$	0.710	0.111	0.755	0.096
$\bar{B_d^0} \to D_s^- DX \ \ (D \to \mu^+\nu X)$	0.718	0.108	0.752	0.104

Table 5.4: *The means and root-mean-squared (RMS) values of the K-factor distributions for the background B decays for both the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples.*

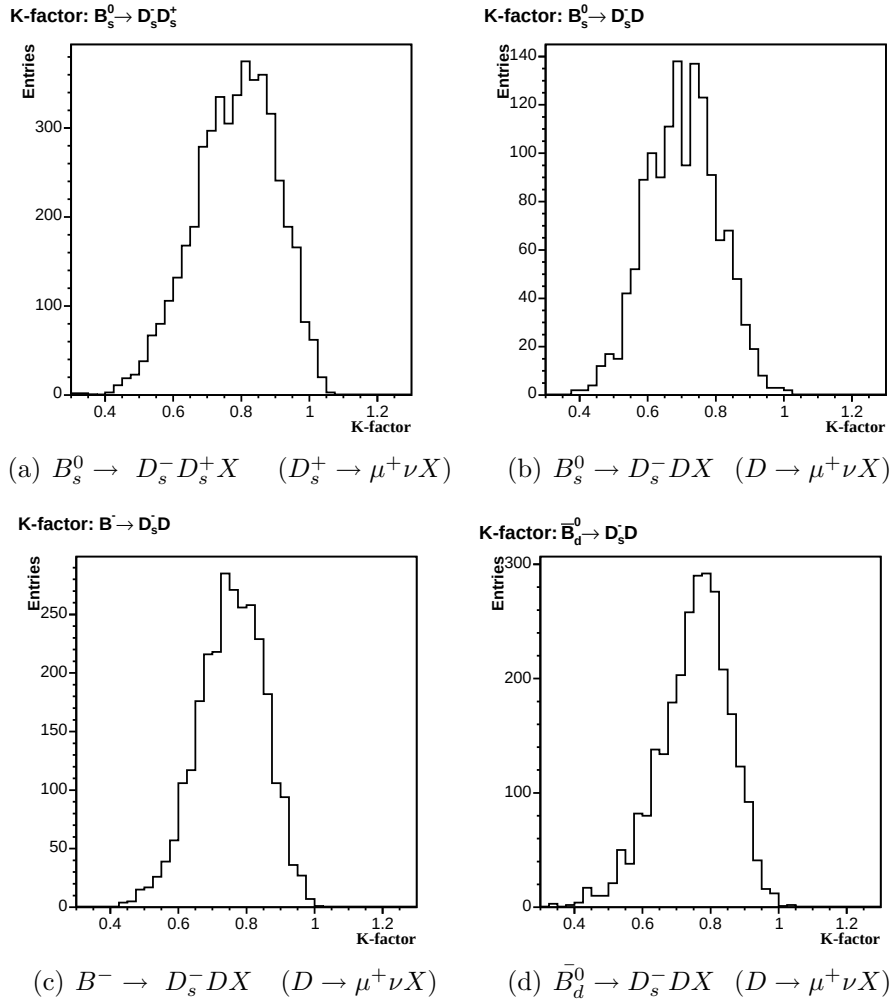


(a) $D_s^- \to \phi\pi^-$



(b) $D_s^- \to K^{*0}K^-$

Figure 5.17: *K-factor distributions for the* $B_s^0 \to \mu^+\nu D_s^-$*;* $B_s^0 \to \mu^+\nu D_s^{-*} \to \mu^+\nu D_s^- X$*;* $B_s^0 \to \mu^+\nu D_{s0}^{*-} \to \mu^+\nu D_s^- X$*;* $B_s^0 \to \mu^+\nu D_{s1}^{'-} \to \mu^+\nu D_s^- X$ *processes.*



(a) $B_s^0 \to D_s^- D_s^+ X \quad (D_s^+ \to \mu^+ \nu X)$ (b) $B_s^0 \to D_s^- DX \quad (D \to \mu^+ \nu X)$

(c) $B^- \to D_s^- DX \quad (D \to \mu^+ \nu X)$ (d) $\bar{B}_d^0 \to D_s^- DX \quad (D \to \mu^+ \nu X)$

Figure 5.18: $D_s^- \to K^{*0} K^-$ *: K-factor distributions for some of the background processes.*

5.4.3 Reconstruction Efficiencies

The reconstruction efficiencies of the different processes contributing to the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples were determined using Monte Carlo simulations. The efficiency strongly depends on the decay length due to the lifetime dependent selection criteria. Therefore, we determined the efficiency as a function of generated VPDL since it enters the computation of number of expected events before it is smeared with the resolution function. The lifetime bias largely cancels in the asymmetry. The effects of the resolution are taken into account separately (Sec. 5.4.4).

Figures 5.19(a) and (b) show the efficiency of the lifetime dependent cuts as a function of VPDL for the signal processes in the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples, respectively. The curves are fit using the functional form $f(x)$ given by:

$$f(x) = \mathrm{p0} \cdot \left[1 - \exp\left(-\frac{(\mathrm{x} - \mathrm{p1})}{\mathrm{p2}}\right)\right], \tag{5.25}$$

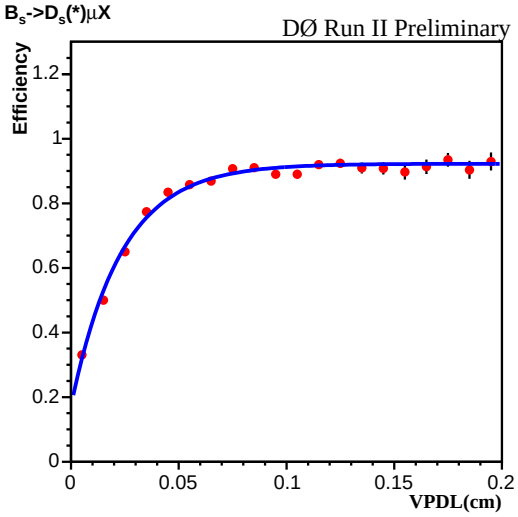
and the parameters of the fit, p1, p2, and p3, are listed in Table 5.5. The efficiency

Decay	$D_s^- \to \phi\pi^-$			$D_s^- \to K^{*0}K^-$		
	p1	p2	p3	p1	p2	p3
$B_s^0 \to D_s^-\mu^+\nu X$	0.9225	-0.0049	0.0233	0.9938	-0.0046	0.0144
$B_s^0 \to \tau^+\nu D_s^- X$ $(\tau^+ \to \mu^+\nu\bar{\nu})$	0.8838	-0.0052	0.0218	0.8666	-0.0070	0.0099
$B_s^0 \to D_s^- D_s^+ X$ $(D_s^+ \to \mu^+\nu X)$	0.9140	-0.0062	0.0258	0.9636	-0.0051	0.0149
$B_s^0 \to D_s^- DX$ $(D \to \mu^+\nu X)$	0.8791	-0.0046	0.0247	0.9089	-0.0044	0.0150
$B^- \to D_s^- DX$ $(D \to \mu^+\nu X)$	0.9181	-0.0045	0.0246	0.9760	-0.0119	0.0184
$\bar{B}_d^0 \to D_s^- DX$ $(D \to \mu^+\nu X)$	0.9255	-0.0101	0.0341	0.9388	-0.0081	0.0175

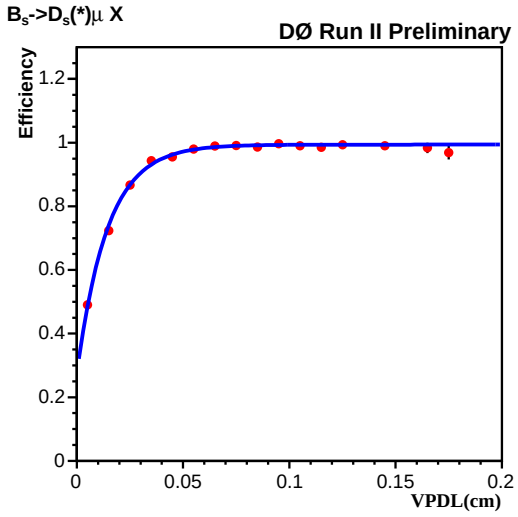
Table 5.5: *The parameters of the fits to the efficiency curves for both the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples.*

curves for some of the other B processes (background contributions) are shown in

Figs. 5.20(a)-(d), and a summary of the fit parameters for all the decays for the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples is provided in Table 5.5.

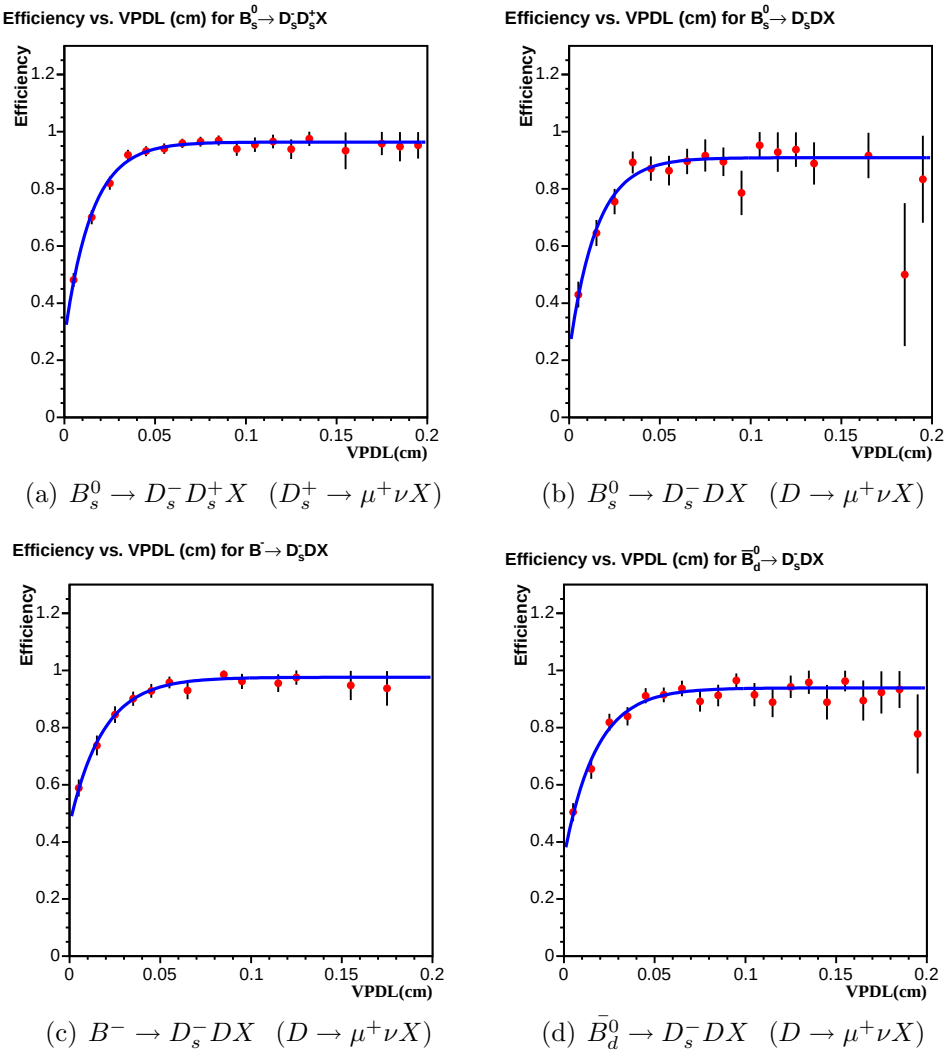


(a) $D_s^- \to \phi\pi^-$



(b) $D_s^- \to K^{*0}K^-$

Figure 5.19: *Efficiency of the lifetime dependent cuts as a function of VPDL for* $B_s^0 \to \mu^+\nu D_s^- X$ *decays.*



(a) $B_s^0 \to D_s^- D_s^+ X \quad (D_s^+ \to \mu^+ \nu X)$

(b) $B_s^0 \to D_s^- DX \quad (D \to \mu^+ \nu X)$

(c) $B^- \to D_s^- DX \quad (D \to \mu^+ \nu X)$

(d) $\bar{B}_d^0 \to D_s^- DX \quad (D \to \mu^+ \nu X)$

Figure 5.20: $D_s^- \to K^{*0} K^-$: *Efficiency of the lifetime dependent cuts as a function of VPDL for some of the background processes.*

5.4.4 VPDL resolution

The visible proper decay length resolution was defined as the difference between the reconstructed (or measured) VPDL and the true (or generated) VPDL. The resolution for the different samples was obtained using Monte Carlo simulation and was parameterized using Gaussians.

5.4.5 Scale factor for VPDL resolution

However, Monte Carlo (MC) simulations do not perfectly model the uncertainties on the track parameters. Therefore, a special procedure was developed to tune the track impact parameter resolution [72]. This procedure changed the track impact parameters and their errors in MC to produce a resolution similar to that in data. Signal MC was used to determine how the above tuning procedure changed the VPDL resolution function.

Figures 5.21(a) and (b) show the VPDL resolution for the $B_s^0 \to \mu^+\nu D_s^- X$ decays in the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples, respectively. Both plots show the VPDL resolution *before* and *after* tuning. The dashed line denotes the VPDL resolution from the untuned Monte Carlo and is modeled using three Gaussians. The solid line denotes the resolution from the tuned Monte Carlo. The fractions and widths of the three Gaussians used to model the untuned MC were used as fixed parameters when fitting the VPDL resolution from tuned MC. Table 5.6 lists these fractions (f_i) and widths (σ_i). The change in VPDL resolution manifested itself in an overall scale factor ($\mathcal{S}.\mathcal{F}.$) such that $\sigma_i(\text{tuned}) = \mathcal{S}.\mathcal{F}. \cdot \sigma_{\text{i}}(\text{untuned})$. The scale factors for the

$D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples were found to be:

$$D_s^- \to \quad \phi\pi^- \quad : \mathcal{S}.\mathcal{F}. = 1.168 \pm 0.020, \tag{5.26}$$

$$D_s^- \to \quad K^{*0}K^- \quad : \mathcal{S}.\mathcal{F}. = 1.142 \pm 0.024. \tag{5.27}$$

The fits in Figs. 5.21(a) and (b) show good agreement with the assumption that the scale factor is the same for all components.

Decay	f_1	σ_1	f_2	σ_2	f_3	σ_3
$D_s^- \to \phi\pi^-$	0.39	0.0028	0.52	0.0065	0.09	0.0187
$D_s^- \to K^{*0}K^-$	0.16	0.0019	0.68	0.0039	0.16	0.0115

Table 5.6: *The fractions (f_i) and widths (σ_i) of the three Gaussians used to model the VPDL resolution (before tuning) for the $B_s^0 \to D_s^- \mu^+ \nu X$ decays in both the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples.*

VPDL resolutions were found to depend on the visible proper decay length itself. This is expected since the VPDL is not fully corrected for the boost of the B-meson. Larger VPDL is correlated with larger boost and therefore more collimated decay products and worse VPDL resolution. This effect was modeled using a variable scale factor (that depends on VPDL resolution). First, all the B candidates were grouped according to their VPDL values. Then in each VPDL group (or bin), the VPDL resolutions were fitted using a single scale factor for all three Gaussians. The widths and the fractions of the Gaussians were fixed to the values in Table 5.6 with the change in VPDL resolution manifested in the scale factor. The resulting plots are shown in Figs. 5.22(a) and (b). This dependence was modeled in the asymmetry fitting procedure.

The resolution for the background samples was also obtained using Monte Carlo

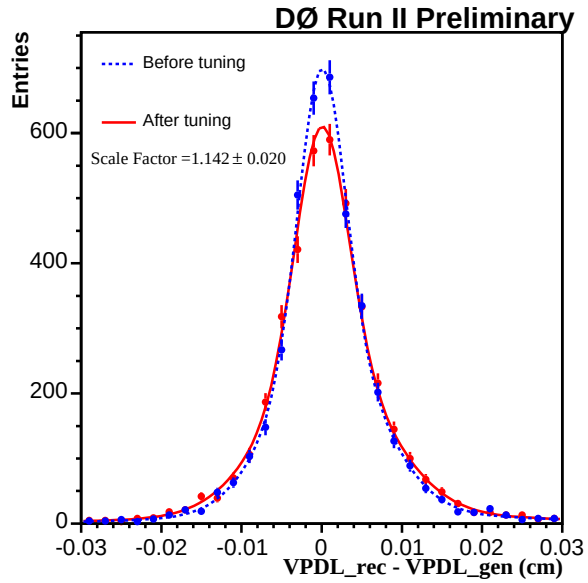
simulation and was parameterized using Gaussians. Since the resolution depends on VPDL and is somewhat better for small values of VPDL (see Fig. 5.22), we imposed a cut of VPDL > 500 μm for generating the resolution plots for the background samples. Some of these plots are shown in Figs. 5.23(a)-(d), while Tables 5.7 and 5.8 list the fractions and widths of the Gaussians used for parameterizing the resolution. These widths were scaled using the constant scale factors listed in Eqs. 5.26 and 5.27 and incorporated into the fitting procedure. Additionally, a larger scale factor (= 2) was used when evaluating systematic uncertainties.

Decay	f_1	σ_1	f_2	σ_2
$B_s^0 \to \tau^+\nu D_s^- X \;\; (\tau^+ \to \mu^+\nu\bar{\nu})$	1.00	0.0099	-	-
$B_s^0 \to D_s^- D_s^+ X \;\; (D_s^+ \to \mu^+\nu X)$	0.34	0.0038	0.66	0.0107
$B_s^0 \to D_s^- DX \;\; (D \to \mu^+\nu X)$	0.30	0.0051	0.70	0.0154
$B^- \to D_s^- DX \;\; (D \to \mu^+\nu X)$	0.35	0.0050	0.65	0.0145
$\bar{B}_d^0 \to D_s^- DX \;\; (D \to \mu^+\nu X)$	0.42	0.0050	0.58	0.0147

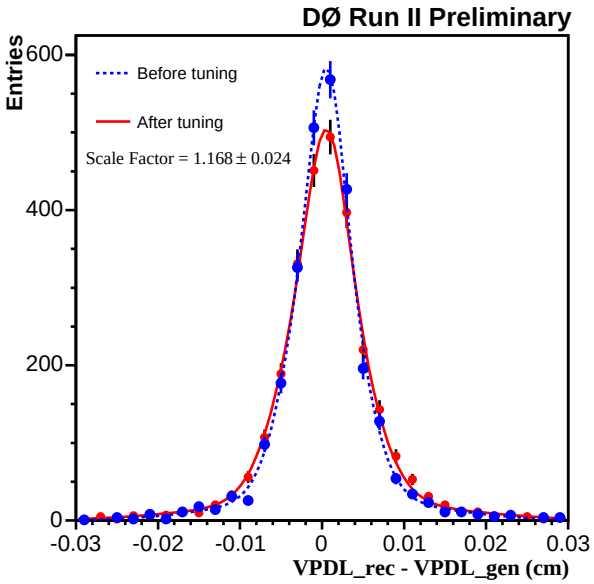
Table 5.7: *The fractions and widths of the Gaussians used to model the VPDL resolution for the background decays in the $D_s^- \to \phi\pi^-$ sample.*

Decay	f_1	σ_1	f_2	σ_2
$B_s^0 \to \tau^+\nu D_s^- X \;\; (\tau^+ \to \mu^+\nu\bar{\nu})$	1.00	0.0096	-	-
$B_s^0 \to D_s^- D_s^+ X \;\; (D_s^+ \to \mu^+\nu X)$	0.64	0.0052	0.36	0.0149
$B_s^0 \to D_s^- DX \;\; (D \to \mu^+\nu X)$	0.22	0.0035	0.78	0.0140
$B^- \to D_s^- DX \;\; (D \to \mu^+\nu X)$	0.51	0.0045	0.49	0.0121
$\bar{B}_d^0 \to D_s^- DX \;\; (D \to \mu^+\nu X)$	0.49	0.0044	0.51	0.0144

Table 5.8: *The fractions and widths of the Gaussians used to model the VPDL resolution for the background decays in the $D_s^- \to K^{*0}K^-$ sample.*

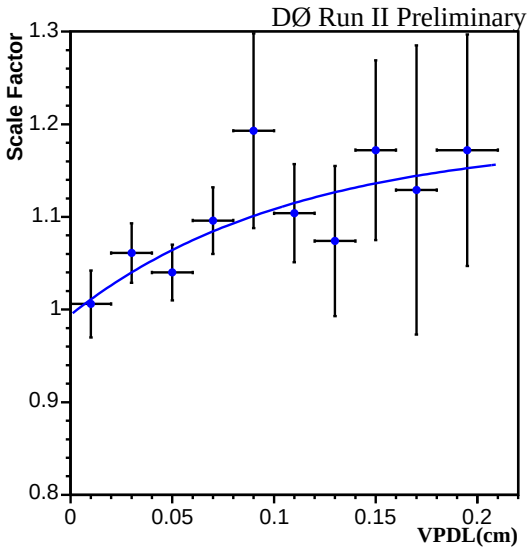


(a) $D_s^- \to \phi\pi^-$

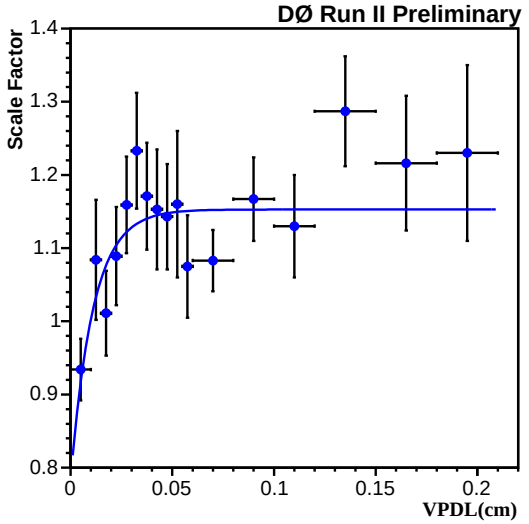


(b) $D_s^- \to K^{*0}K^-$

Figure 5.21: *VPDL resolution before (dashed line) and after (solid line) tuning for* $B_s^0 \to \mu^+\nu D_s^- X$ *decays.*

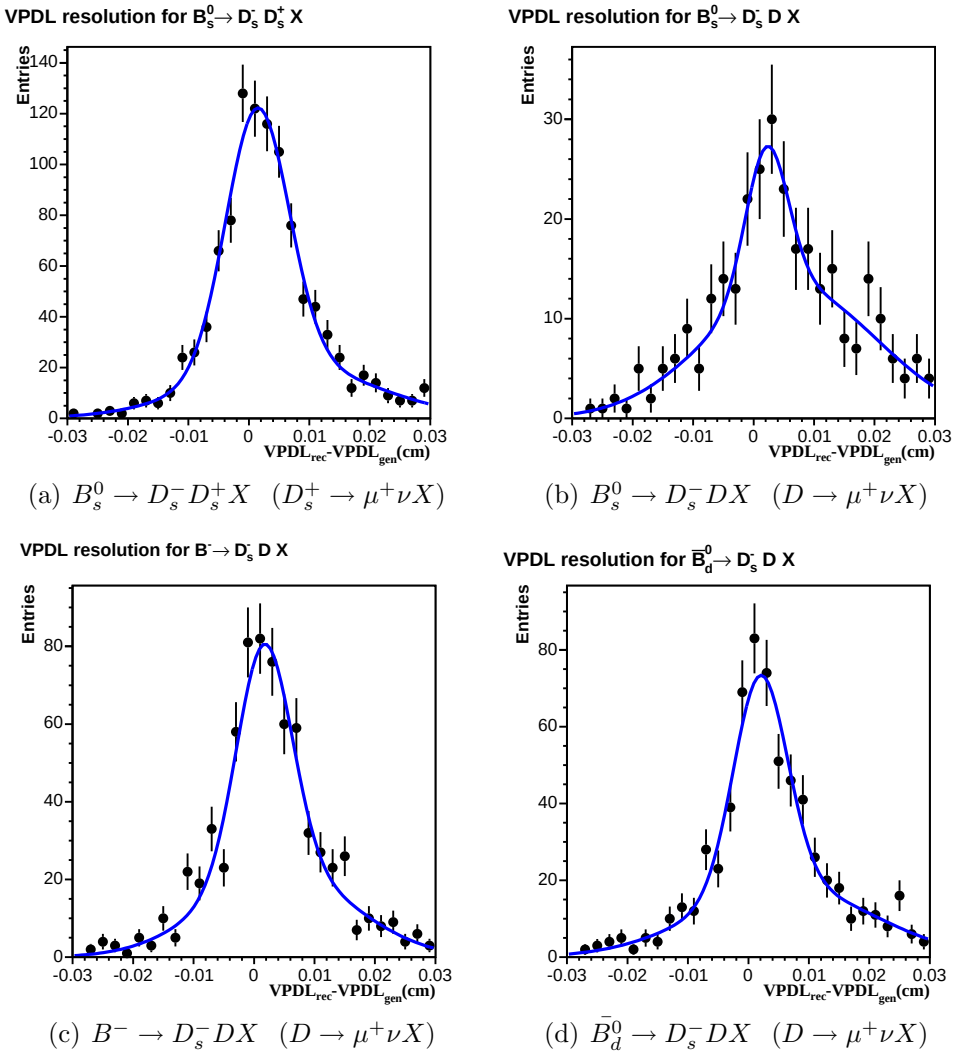


(a) $D_s^- \to \phi\pi^-$



(b) $D_s^- \to K^{*0}K^-$

Figure 5.22: *The VPDL resolution scale factor as a function of VPDL. Bigger VPDL bins are used at larger values of VPDL owing to smaller statistics in those bins.*



(a) $B_s^0 \to D_s^- D_s^+ X \quad (D_s^+ \to \mu^+ \nu X)$

(b) $B_s^0 \to D_s^- D X \quad (D \to \mu^+ \nu X)$

(c) $B^- \to D_s^- D X \quad (D \to \mu^+ \nu X)$

(d) $\bar{B}_d^0 \to D_s^- D X \quad (D \to \mu^+ \nu X)$

Figure 5.23: $D_s^- \to K^{*0} K^-$: *VPDL resolution plots for some of the background processes.*

5.5 Measured Asymmetry

This section describes how a time dependent asymmetry between unmixed and mixed B_s^0 mesons was obtained. First, D_s^- mass distributions for different VPDL bins, for both the unmixed and mixed event samples were produced and then the numbers of unmixed ($N_i^{\rm unm}$) and mixed ($N_i^{\rm mix}$) B_s^0 mesons for each bin were determined. This was done by fitting the D_s^- mass distributions to functions describing both the signal and background contributions. Following this, an asymmetry, $A_i^{\rm meas}$, was calculated for each VPDL bin using the definition given by Eq. 3.17:

$$A_i^{\rm meas} = \frac{N_i^{\rm unm} - N_i^{\rm mix}}{N_i^{\rm unm} + N_i^{\rm mix}}.$$

5.5.1 Mass Fitting Procedure

The number of B_s^0 candidates in the flavor untagged sample is quite large and allows us to fit a large statistics sample. However, once the data is flavor tagged into mixed and unmixed samples, and then separated into bins of VPDL the statistics in each bin are very much reduced. To improve on the fitting, we first fit the full untagged sample and then fixed the mass and width of the D_s^- from that sample when the flavor tagged samples were fit. We verified that fits to the untagged and tagged samples yielded consistent values for the Gaussian parameters. The procedure for fitting the untagged sample as well as the mixed and unmixed sub-samples is described below.

$\mathbf{D_s^- \to \phi\pi^-}$

As was described in Sec. 5.1.3 the untagged sample was fit using single Gaussians to describe the $D_s^- \to \phi\pi^-$ and $D^- \to \phi\pi^-$ decays and the background was modeled by

an exponential (see Fig. 5.5). The tagged sample was fit similarly and is shown in Fig. 5.15.

The D_s^- mass and sigma along with the D^- mass and sigma obtained from fitting the full untagged mass distribution were used as fixed parameters when fitting for the number of tagged (mixed/unmixed) events in the individual VPDL bins. We verified using data that the tagging procedure did not change the D_s^- mean and width. This was also confirmed using simulated events.

For fitting the background in the sub-samples (individual VPDL bins), the slope of the (background) exponential used was fixed to the value obtained from fitting the untagged mass distribution in the corresponding VPDL bin. A potential change in the background shape due to the tagging procedure was taken into account as a correction factor determined from the full untagged, and tagged unmixed and mixed samples. The correction factors for the unmixed and mixed samples were determined to be:

$$\begin{aligned} \frac{\text{bkg slope (unmixed)}}{\text{bkg slope (untagged)}} &= 1.166 \pm 0.079 \qquad \text{(Unmixed)}, \\ \frac{\text{bkg slope (mixed)}}{\text{bkg slope (untagged)}} &= 1.179 \pm 0.086 \qquad \text{(Mixed)}. \end{aligned} \tag{5.28}$$

The numbers of mixed (oscillated) and unmixed (non-oscillated) events obtained by fitting the D_s^- mass distributions in 19 different VPDL bins are listed in Table 5.9. Also listed for each VPDL bin are the uncertainties on the numbers of fitted events ($\sigma(N_i^{\text{unm}})$, $\sigma(N_i^{\text{mix}})$), and the calculated asymmetry, A_i^{meas}, along with its uncertainty, $\sigma(A_i^{\text{meas}})$. The asymmetry is also plotted as a function of VPDL in Fig. 5.24. Superimposed on that plot is the curve for the expected asymmetry corresponding to $\Delta m_s = 8$ ps^{-1}.

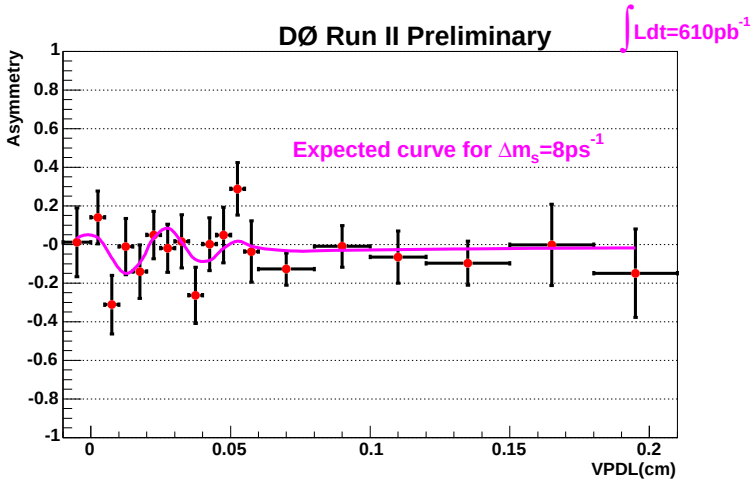


Figure 5.24: *The measured asymmetry in the $D_s^- \to \phi\pi^-$ sample as a function of the visible proper decay length (VPDL). Superimposed on the plot is the curve for the expected asymmetry corresponding to $\Delta m_s = 8\ ps^{-1}$.*

$D_s^- \to K^{*0}K^-$

A procedure identical to that described in Sec. 5.1.3 was followed for fitting the untagged sample. The untagged sample was fit over the range $-0.01 <$ VPDL $<$ 0.2 cm. Single Gaussians were used to describe the $D_s^- \to K^{*0}K^-$ and $D^- \to K^{*0}K^-$ decays and the background was modeled by an exponential. The initial seed for this exponential was provided by the slope of the exponential used to fit the wrong-sign combinations ($Q_\mu = Q_\pi$). The $D^- \to K^+\pi^-\pi^-$ reflection was modeled as described in the same section.

The following additional assumptions were made when fitting the tagged samples. The D_s^- mass and sigma that were obtained from fitting the above untagged mass distribution were used as fixed parameters when fitting for the number of tagged (mixed/unmixed) events in the individual VPDL bins. The D^- sigma was constrained relative to the D_s^- sigma as before. The sigmas of the two Gaussians used to model

Bin	VPDL range (cm)	$N_i^{\rm unm}$	$\sigma(N_i^{\rm unm})$	$N_i^{\rm mix}$	$\sigma(N_i^{\rm mix})$	$A_i^{\rm meas}$	$\sigma(A_i^{\rm meas})$
1	[−0.01, 0.000]	35.50	8.89	34.71	8.81	0.011	0.178
2	[0.000, 0.005]	49.53	8.96	37.35	7.96	0.140	0.137
3	[0.005, 0.010]	27.99	8.16	53.37	9.06	-0.312	0.152
4	[0.010, 0.015]	44.34	9.34	45.32	9.06	-0.011	0.145
5	[0.015, 0.020]	41.85	9.32	55.52	9.66	-0.140	0.139
6	[0.020, 0.025]	57.25	9.97	51.89	9.04	0.049	0.123
7	[0.025, 0.030]	53.83	10.01	55.96	9.34	-0.019	0.125
8	[0.030, 0.035]	45.47	9.08	44.06	8.32	0.016	0.137
9	[0.035, 0.040]	30.91	7.96	53.03	9.33	-0.264	0.145
10	[0.040, 0.045]	44.00	8.65	43.89	8.30	0.001	0.136
11	[0.045, 0.050]	40.77	8.39	36.99	7.54	0.049	0.144
12	[0.050, 0.055]	51.77	9.09	28.64	6.83	0.288	0.136
13	[0.055, 0.060]	32.44	7.71	34.92	7.44	-0.037	0.159
14	[0.060, 0.080]	99.42	13.16	128.51	13.43	-0.128	0.083
15	[0.080, 0.100]	72.64	11.46	74.09	10.89	-0.010	0.108
16	[0.100, 0.120]	46.73	9.53	53.22	9.52	-0.065	0.135
17	[0.120, 0.150]	54.31	9.56	66.04	9.81	-0.097	0.114
18	[0.150, 0.180]	23.13	6.96	23.23	6.81	-0.002	0.210
19	[0.180, 0.210]	15.21	5.56	20.54	6.00	-0.149	0.229

Table 5.9: *$D_s^- \to \phi\pi^-$: For each of the 19 VPDL bins the measured number of unmixed ($N_i^{\rm unm}$) and mixed ($N_i^{\rm mix}$) D_s^- events, their statistical uncertainties, $\sigma(N_i^{\rm unm})$ and $\sigma(N_i^{\rm mix})$, the measured asymmetry, $A_i^{\rm meas}$, and its uncertainty, $\sigma(A_i^{\rm meas})$, all determined from the fits to the corresponding D_s^- mass distributions.*

the $D^- \to K^+\pi^-\pi^-$ reflection were also used as fixed parameters along with the relative fraction of events in those two Gaussians.

For fitting the background in these sub-samples, the slope of the exponential used was fixed to the value obtained from fitting the wrong-sign combinations in the mixed and unmixed samples for each bin. The parameters thus left free in these fits in the individual VPDL bins were: the number of D_s^- events and the number of D^- events and the background normalization. The background model was varied and the

dependence of the fit results on the model was studied for systematic error evaluation.

The numbers of unmixed and mixed events thus obtained, along with the asymmetry for each VPDL bin are listed in Table 5.10. The measured asymmetry is also plotted as a function of VPDL in Fig. 5.25. Furthermore, Fig. 5.26 shows the measured asymmetry overlaid with the curve for the expected asymmetry for two different Δm_s values.

Bin	VPDL range (cm)	$N_i^{\rm unm}$	$\sigma(N_i^{\rm unm})$	$N_i^{\rm mix}$	$\sigma(N_i^{\rm mix})$	$A_i^{\rm meas}$	$\sigma(A_i^{\rm meas})$
1	[−0.01, 0]	26.80	16.28	30.01	15.38	−0.056	0.396
2	[0, 0.005]	65.23	17.62	56.92	14.17	0.068	0.183
3	[0.005, 0.01]	34.75	18.95	40.99	15.08	−0.082	0.327
4	[0.01, 0.015]	83.21	18.67	77.99	13.71	0.032	0.142
5	[0.015, 0.02]	80.50	19.13	92.78	16.68	−0.071	0.148
6	[0.02, 0.025]	57.21	19.10	47.84	14.86	0.0892	0.226
7	[0.025, 0.03]	27.89	17.15	40.99	15.65	−0.190	0.349
8	[0.03, 0.035]	95.93	18.08	48.44	14.55	0.329	0.158
9	[0.035, 0.04]	49.17	16.68	29.69	14.08	0.247	0.274
10	[0.04, 0.045]	75.29	16.42	39.94	13.94	0.307	0.186
11	[0.045, 0.05]	39.81	14.79	33.24	13.19	0.090	0.270
12	[0.05, 0.055]	12.56	13.16	34.82	9.64	−0.470	0.422
13	[0.055, 0.06]	20.20	13.05	30.61	9.45	−0.205	0.343
14	[0.06, 0.08]	178.99	23.74	128.69	21.14	0.163	0.103
15	[0.08, 0.1]	90.90	20.84	57.68	16.46	0.224	0.174
16	[0.1, 0.12]	63.54	16.43	57.91	13.78	0.046	0.175
17	[0.12, 0.15]	12.70	14.80	76.76	16.08	−0.716	0.288
18	[0.15, 0.18]	43.96	12.07	30.24	10.77	0.185	0.217
19	[0.18, 0.21]	36.49	8.94	15.39	9.30	0.407	0.272

Table 5.10: *$D_s^- \to K^{*0}K^-$: For each of the 19 VPDL bins the measured number of unmixed ($N_i^{\rm unm}$) and mixed ($N_i^{\rm mix}$) D_s^- events, their statistical uncertainties, $\sigma(N_i^{\rm unm})$ and $\sigma(N_i^{\rm mix})$, the measured asymmetry, $A_i^{\rm meas}$, and its uncertainty, $\sigma(A_i^{\rm meas})$, all determined from the fits to the corresponding D_s^- mass distributions.*

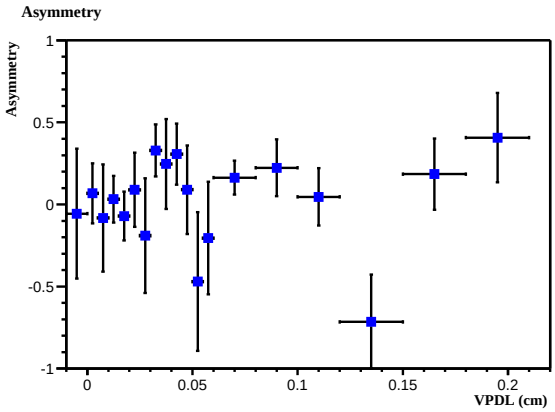


Figure 5.25: *The measured asymmetry in the $D_s^- \to K^{*0}K^-$ sample as a function of the visible proper decay length (VPDL).*

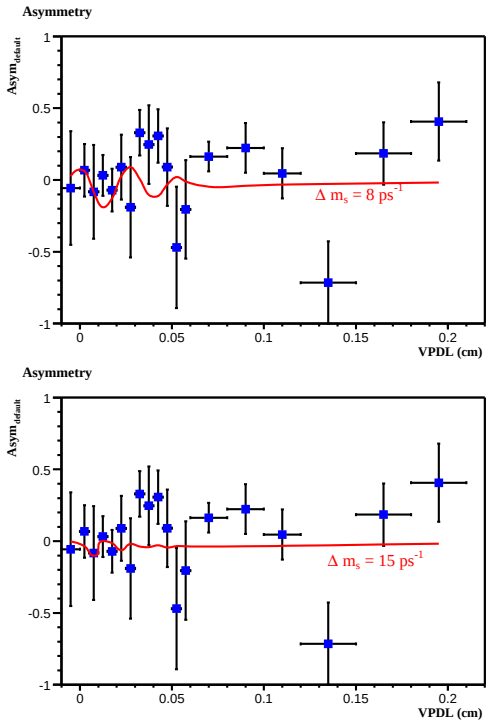


Figure 5.26: *$D_s^- \to K^{*0}K^-$: The measured asymmetry overlaid with the curve for the expected asymmetry for $\Delta m_s = 8\ ps^{-1}$ (top) and $\Delta m_s = 15\ ps^{-1}$ (bottom).*

5.6 Fitting For Δm_s

Figures 5.24 and 5.25 show that no B_s^0 oscillations can be resolved at the moment. Therefore, as described in Sec. 3.4, in order to set a lower limit on Δm_s , Eq. 5.17 was modified to the following form:

$$n_s^{unm/mix}(x) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5 \cdot [1 \pm (2\eta_s - 1)\cos(\Delta m_s \cdot Kx/c) \cdot \mathcal{A}], \quad (5.29)$$

where $\mathcal{A}$ is the only remaining fit parameter. Different choices of Δm_s in the range 1 ps^{-1} to 20 ps^{-1} were input and the fitted values of $\mathcal{A}$ (and its error $\sigma_{\mathcal{A}}$) were determined from the minimization of a $\chi^2(\mathcal{A})$ defined as:

$$\chi^2(\mathcal{A}) = \sum_i \frac{(A_i^{\rm meas} - A_i^e(\mathcal{A}))^2}{\sigma^2(A_i^{\rm meas})}. \quad (5.30)$$

The fitted value of $\mathcal{A}$ was then plotted as a function of the input value of Δm_s , and we searched for a peak of $\mathcal{A}$=1 to obtain a measurement of Δm_s . Since no peak was found, limits were set using this method. The limit was determined by calculating the probability that a fitted value of $\mathcal{A}$ could fluctuate to $\mathcal{A} = 1$. This occurred at $\mathcal{A} + 1.645\sigma_{\mathcal{A}} = 1$ (95% C.L.), where $\sigma_{\mathcal{A}}$ is the uncertainty associated with $\mathcal{A}$. The sensitivity of the measurement was determined by calculating the probability that $\mathcal{A}$=0 could fluctuate to $\mathcal{A}$=1. This occured as $1.645\sigma_{\mathcal{A}} = 1$ (95% C.L.).

Figure 5.27 shows the *amplitude scans* (also called *spectra*) for the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples. The points and error bars in the plots represent the measurements of the amplitude $\mathcal{A}$ and their statistical and systematic uncertainties, $\sigma_{\mathcal{A}}$. The "data $\pm$ 1.645 $\sigma_{\mathcal{A}}$ (stat.)" curve denotes the statistical uncertainties only while the "data $\pm$ 1.645 $\sigma_{\mathcal{A}}$" curve includes both the statistical and systematic uncertainties. Values of Δm_s where the solid curves are below 1 are excluded at 95% C.L.

Additionally, the dotted and dashed curves show 1.645 $\sigma_{\mathcal{A}}$ with statistical only, and with both statistical and systematic uncertainties, respectively, and are a measure of the sensitivity of the analysis. The following results were obtained including statistical uncertainties only and at the 95% confidence level. Systematic uncertainties are discussed in the next section.

$$\begin{aligned} D_s^- \to \phi\pi^- \qquad & \text{Limit} \quad : \Delta m_s > 7.0 \text{ ps}^{-1} \\ & \text{Sensitivity} \quad : 8.1 \text{ ps}^{-1} \end{aligned} \tag{5.31}$$

$$\begin{aligned} D_s^- \to K^{*0}K^- \qquad & \text{Limit} \quad : \Delta m_s > 5.1 \text{ ps}^{-1} \\ & \text{Sensitivity} \quad : 7.8 \text{ ps}^{-1} \end{aligned} \tag{5.32}$$

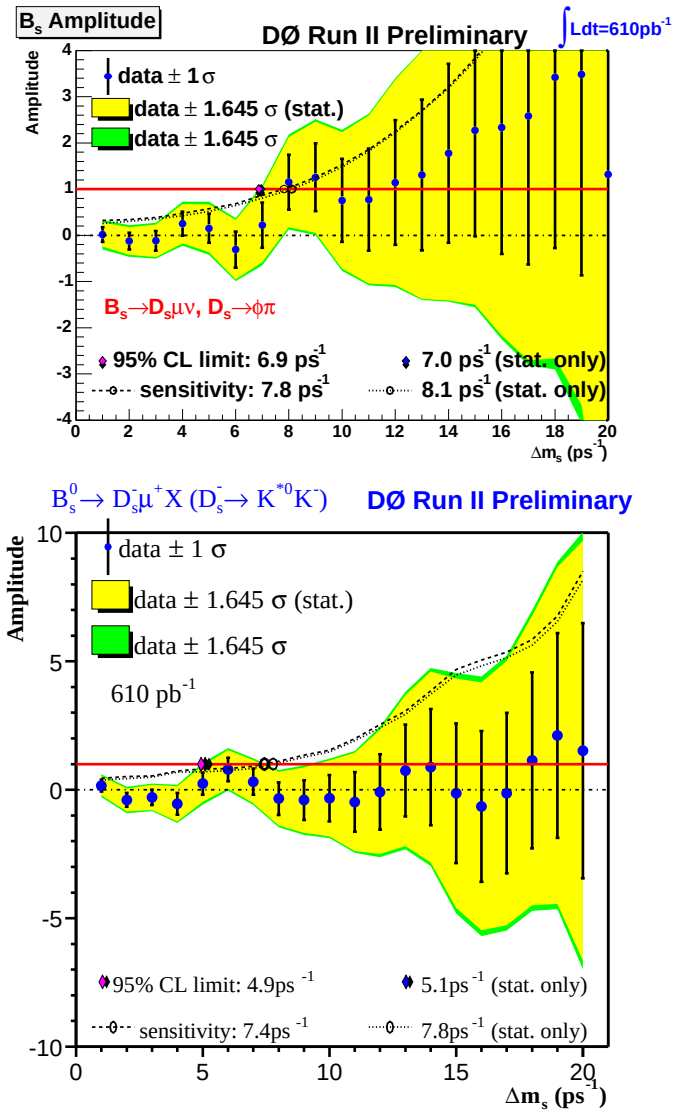


Figure 5.27: *The B_s^0 oscillation amplitude spectra for the $D_s^- \to \phi\pi^-$ (top) and $D_s^- \to K^{*0}K^-$ (bottom) samples. The points and error bars denote the measurements of the amplitude $\mathcal{A}$ and their statistical and systematic uncertainties, $\sigma_{\mathcal{A}}$. The "*data $\pm$ 1.645 $\sigma_{\mathcal{A}}$ (stat.)*" curve represents the statistical uncertainties only while the "data $\pm$ 1.645 $\sigma_{\mathcal{A}}$" curve includes both the statistical and systematic uncertainties. Values of Δm_s where the solid curves are below 1 are excluded at 95% C.L. The dotted and dashed curves show 1.645 $\sigma_{\mathcal{A}}$ with statistical only, and with both statistical and systematic uncertainties, repsectively, and are a measure of the sensitivity of the analysis.*

5.7 Systematic Uncertainties

We considered the following sources of systematic uncertainty on the measurement:

- Tagging Purity
- Sample Composition
- Non-zero $\Delta\Gamma/\Gamma$
- Mass Fitting Procedure
- Efficiency and its VPDL dependence
- Resolution
- K-factors

Their specific contributions are described in more detail in the following subsections. The total systematic uncertainty was estimated using the formula [42]:

$$\sigma_{\mathcal{A}}^{sys} = \Delta\mathcal{A} + (1 - \mathcal{A})\frac{\Delta\sigma_{\mathcal{A}}}{\sigma_{\mathcal{A}}} \tag{5.33}$$

where $\mathcal{A}$ and $\sigma_{\mathcal{A}}$ are the measured amplitude and its uncertainty as defined previously. $\Delta\mathcal{A}$ and $\Delta\sigma_{\mathcal{A}}$ are the changes in the amplitude and its error between a new fit (for a systematic error evaluation), and the fit using the default parameter values. The resulting total systematic uncertainty, $\sigma_{\mathcal{A}}^{sys}$, was then summed in quadrature with other contributions. The individual contributions along with the total systematic uncertainty are listed in Tables 5.11 and 5.12 for the $D_s^- \to \phi\pi^-$ sample, and in Tables 5.13 and 5.14 for the $D_s^- \to K^{*0}K^-$ sample.

5.7.1 Tagging Purity

The tagging purity used as an input to the B_s^0 mixing analysis ($\eta_s = 0.692$) was obtained by constraining the purities (or dilutions) of the B_d^0 and B^+ mesons to be equal i.e. $\mathcal{D}(B_d^0) = \mathcal{D}(B^+)$ (see Eq. 5.14). However, dilutions had been obtained separately for the B_d^0 and B^+ mesons as well (Eq. 5.12 and 5.13). The smallest value, $\mathcal{D}(B^+) = 0.368$, corresponding to a tagging purity of $\eta_s = 0.684$ was used for the systematic error evaluation.

5.7.2 Sample Composition

- The $c\bar{c}$ contribution was changed from 3.5% to 6.0% (one sigma from its central value) (see Sec. 5.4.1).
- The $B_s^0 \to D_s^- D_s^+ X$ branching ratio was changed from its PDG value, 23%, to its EvtGen value, 4.7%.
- The semileptonic signal $B_s^0 \to \mu^+ \nu D_s^- X$ branching ratio was varied by one sigma to 5.5%.
- The B_s^0 lifetime was changed by one sigma to 455 μm.
- Variation of other input parameters like $c\tau(B^+)$, $c\tau(B_d^0)$, and Δm_d produced negligible differences.

Additionally, for the $D_s^- \to \phi\pi^-$ sample, a contribution from $D^- \to K^+\pi^-\pi^-$ decays to the D_s^- signal was estimated at the level of 1.6% and the corresponding systematic error was determined. The sample composition was also determined with a tighter muon p_T cut of 4 GeV/c, and the corresponding contribution to the systematic uncertainty was obtained.

5.7.3 Non-zero $\Delta\Gamma/\Gamma$

The $B^0_{d,s}$ mixing analyses were carried out assuming $\Delta\Gamma = 0$. While this holds to a very good approximation for the B^0_d system, the latest SM expectation is $\Delta\Gamma/\Gamma = 0.12 \pm 0.05$ for the B^0_s system (see Sec. 2.7.2). Also, recent DØ results indicate that $\Delta\Gamma/\Gamma = 0.25^{+0.14}_{-0.15}$ [73]. The effect of such a non-zero value of $\Delta\Gamma/\Gamma$ was studied by using a value of $\Delta\Gamma/\Gamma = 0.2$ for estimating the systematic uncertainty due to it on the B^0_s amplitude.

5.7.4 Mass Fitting Procedure

Variations in the assumptions on signal and background shapes employed in the mass fits were used for determining the systematic uncertainties. Since the mass fitting procedure differs between the $D^-_s \to \phi\pi^-$ and $D^-_s \to K^{*0}K^-$ samples, these are listed separately:

$D^-_s \to \phi\pi^-$ **:**

- The D^-_s mass and width were varied by one sigma.
- The D^- mass was varied by one sigma.
- The slope of the exponential background by changed by one sigma.
- The background was parameterized by a straight line instead of an exponential.
- The bin width of the histograms used for fitting for the number of events was reduced by 50%.

- Assuming the presence of a reflection due to $D^- \to K^+\pi^-\pi^-$ decays in the D_s^- signal region, a fit which included an additional wide Gaussian was performed and the contribution to the systematic error thus estimated.

$D_s^- \to K^{*0}K^-$ **:**

- The D_s^- mass and width were changed by one sigma.
- The widths of the $D^- \to K^+\pi^-\pi^-$ Gaussians and their relative fraction were varied by one sigma.
- The ratio of $D^- \to K^+\pi^-\pi^-$ and $D^- \to K^{*0}K^-$ events used as a fixed parameter in the fits was increased by 25%.
- The bin width of the histograms used for fitting for the number of mixed and unmixed events was reduced by 25%.
- The background was parameterized by a straight line instead of an exponential.
- The background in the different VPDL bins was modeled similar to the $D_s^- \to \phi\pi^-$ case (see Eq. 5.28) by taking the slope from the untagged sample in each bin and applying a correction factor determined from the full untagged, unmixed and mixed samples.

5.7.5 Efficiency and its VPDL dependence

- The signal efficiency versus VPDL parametrization was used for the background samples.

- The efficiency parametrization for the signal was varied by changing the slope by one sigma.

5.7.6 Resolution

- A constant scale factor was used for both the signal and background. The new values were obtained by changing the central value of the scale factor ($\mathcal{S}.\mathcal{F}.$) in Eqs. 5.26 and 5.27 by one sigma.
- A larger scale factor ($= 2$) was used for the background samples.

5.7.7 K-factors

- The K-factors were varied by 2%. This was motivated by a study done in Ref. [74] which investigated the model dependence of simulated K-factor distributions using an alternative model of semileptonic B decays and variation of p_T cuts. In all cases that were studied, the variation of the average value of K-factors did not exceed 2%.
- The K-factor distributions were obtained using reconstructed information instead of generator level information.
- The K-factor histograms were smoothed before using them in the final fit.

Osc. frequency		1 ps^{-1}	2 ps^{-1}	3 ps^{-1}	4 ps^{-1}	5 ps^{-1}	6 ps^{-1}	7 ps^{-1}	8 ps^{-1}	9 ps^{-1}	10 ps^{-1}
$\mathcal{A}$		0.014	−0.125	−0.120	0.251	0.152	−0.310	0.222	1.150	1.257	0.755
Stat. uncertainty		0.158	0.182	0.214	0.258	0.313	0.391	0.489	0.593	0.731	0.898
$\eta_s = 0.684$	$\Delta\mathcal{A}$	+0.000	−0.004	−0.004	+0.012	+0.007	−0.013	+0.010	+0.050	+0.054	+0.033
	$\Delta\sigma$	+0.007	+0.008	+0.009	+0.011	+0.014	+0.017	+0.021	+0.026	+0.032	+0.039
$c\bar{c}:$ 6%	$\Delta\mathcal{A}$	+0.003	−0.009	−0.010	+0.021	+0.004	−0.076	−0.032	+0.093	+0.098	−0.021
	$\Delta\sigma$	+0.009	+0.007	+0.008	+0.013	+0.018	+0.026	+0.037	+0.049	+0.067	+0.091
$\mathcal{B}(B_s^0 \to D_s^- D_s^+) =$ 4.7%	$\Delta\mathcal{A}$	+0.000	+0.003	+0.002	−0.005	−0.003	+0.006	−0.005	−0.024	−0.027	−0.018
	$\Delta\sigma$	−0.004	−0.004	−0.005	−0.006	−0.007	−0.008	−0.011	−0.013	−0.016	−0.019
$\mathcal{B}(B_s^0 \to \mu^+ \nu D_s^- X) =$ 5.5%	$\Delta\mathcal{A}$	+0.003	−0.020	−0.015	+0.001	−0.001	−0.016	+0.007	+0.049	+0.056	+0.033
	$\Delta\sigma$	+0.007	+0.008	+0.010	+0.012	+0.014	+0.018	+0.022	+0.027	+0.033	+0.040
$c\tau_{B_s^0} = 455$ μm	$\Delta\mathcal{A}$	+0.002	−0.001	−0.001	+0.003	−0.002	−0.007	+0.004	+0.010	+0.002	−0.009
	$\Delta\sigma$	+0.000	+0.000	+0.000	+0.001	+0.001	+0.002	+0.002	+0.003	+0.004	+0.006
$D^- \to K^+\pi^-\pi^-$ contribution	$\Delta\mathcal{A}$	−0.012	+0.005	+0.006	+0.011	+0.010	+0.004	+0.011	+0.019	+0.021	+0.020
1.6%	$\Delta\sigma$	+0.002	+0.003	+0.003	+0.004	+0.005	+0.006	+0.007	+0.009	+0.010	+0.013
$p_{T_\mu} > 4$ GeV/c	$\Delta\mathcal{A}$	+0.003	+0.015	+0.009	−0.000	+0.001	+0.006	−0.008	−0.030	−0.033	−0.026
	$\Delta\sigma$	−0.004	−0.005	−0.006	−0.007	−0.008	−0.011	−0.013	−0.016	−0.019	−0.024
$\Delta\Gamma/\Gamma = 0.2$	$\Delta\mathcal{A}$	−0.000	−0.001	+0.000	−0.000	−0.000	+0.001	+0.001	+0.001	+0.001	+0.002
	$\Delta\sigma$	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001
D_s^- mass changed to	$\Delta\mathcal{A}$	+0.002	−0.001	+0.000	−0.000	−0.002	−0.005	−0.003	+0.002	+0.001	+0.001
$M_s + 1\sigma$	$\Delta\sigma$	−0.000	−0.000	+0.000	−0.000	−0.000	+0.000	+0.000	+0.000	+0.000	−0.000
D_s^- width changed to	$\Delta\mathcal{A}$	+0.001	+0.002	−0.001	−0.001	+0.002	−0.000	−0.006	−0.010	−0.005	+0.008
$\sigma_s - 1\sigma$	$\Delta\sigma$	+0.000	+0.000	+0.000	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.002
D^- mass changed to	$\Delta\mathcal{A}$	+0.001	+0.000	−0.000	+0.001	+0.001	−0.000	−0.001	−0.002	−0.003	−0.005
$M_- - 1\sigma$	$\Delta\sigma$	−0.000	−0.000	+0.000	−0.000	−0.000	+0.000	−0.000	−0.000	−0.000	−0.000
Slope for bkg.	$\Delta\mathcal{A}$	−0.000	−0.001	+0.001	−0.002	−0.004	−0.002	−0.003	−0.002	−0.003	−0.006
changed by 1σ	$\Delta\sigma$	−0.001	−0.002	−0.002	−0.002	−0.002	−0.003	−0.003	−0.004	−0.005	−0.006
Bkg. parametrized	$\Delta\mathcal{A}$	+0.004	+0.009	−0.005	+0.011	+0.017	−0.001	−0.014	−0.025	−0.007	+0.044
by straight line	$\Delta\sigma$	+0.008	+0.009	+0.010	+0.011	+0.012	+0.015	+0.018	+0.021	+0.026	+0.030
Mass bin width	$\Delta\mathcal{A}$	+0.010	−0.007	−0.000	+0.013	+0.047	+0.053	+0.024	+0.005	+0.011	−0.025
smaller by 50%	$\Delta\sigma$	−0.001	−0.001	−0.000	−0.002	−0.003	−0.002	−0.005	−0.007	−0.003	−0.002
Wide Gaussion	$\Delta\mathcal{A}$	+0.026	+0.009	+0.016	−0.054	−0.058	+0.033	+0.049	+0.002	+0.066	+0.130
added in mass fit	$\Delta\sigma$	−0.001	−0.001	−0.000	−0.001	−0.004	−0.007	−0.010	−0.015	−0.016	−0.016
Same eff. dependence	$\Delta\mathcal{A}$	+0.009	−0.001	−0.004	−0.001	−0.001	−0.008	−0.001	+0.012	+0.012	+0.003
for signal and bkg	$\Delta\sigma$	+0.001	+0.001	+0.002	+0.003	+0.003	+0.004	+0.005	+0.006	+0.008	+0.010
Eff. slope varied	$\Delta\mathcal{A}$	+0.006	+0.001	−0.001	+0.001	−0.000	−0.007	−0.003	+0.006	+0.003	−0.011
by 1σ for signal	$\Delta\sigma$	+0.001	+0.001	+0.001	+0.002	+0.002	+0.003	+0.003	+0.004	+0.005	+0.006
Resolution $\mathcal{S}.\mathcal{F}. = 1.162$	$\Delta\mathcal{A}$	+0.001	−0.001	−0.004	+0.015	+0.024	−0.023	+0.009	+0.140	+0.188	+0.115
	$\Delta\sigma$	+0.001	+0.003	+0.006	+0.012	+0.020	+0.035	+0.055	+0.079	+0.112	+0.152
Resolution $\mathcal{S}.\mathcal{F}. = 2$	$\Delta\mathcal{A}$	−0.006	+0.004	+0.007	+0.006	+0.000	+0.002	+0.002	−0.005	−0.007	−0.006
for background	$\Delta\sigma$	−0.000	+0.000	−0.000	−0.001	−0.001	−0.002	−0.003	−0.004	−0.005	−0.006
K-factor variation	$\Delta\mathcal{A}$	+0.003	+0.004	−0.009	−0.026	+0.043	+0.016	−0.134	−0.110	+0.064	+0.087
2%	$\Delta\sigma$	−0.001	−0.001	−0.001	−0.003	−0.008	−0.008	−0.011	−0.017	−0.022	−0.031
K-factor using	$\Delta\mathcal{A}$	+0.000	−0.000	+0.000	−0.001	+0.001	+0.006	+0.001	−0.001	+0.014	+0.032
reconstructed values	$\Delta\sigma$	+0.000	+0.000	+0.001	+0.001	+0.001	+0.002	+0.003	+0.004	+0.006	+0.007
Using smoothed	$\Delta\mathcal{A}$	+0.006	+0.001	−0.001	−0.001	+0.000	−0.004	−0.006	+0.001	+0.011	+0.011
K-factor histograms	$\Delta\sigma$	+0.001	+0.001	+0.001	+0.001	+0.002	+0.002	+0.003	+0.004	+0.005	+0.006
Total syst.	σ_{tot}^{sys}	0.113	0.102	0.099	0.129	0.153	0.141	0.200	0.197	0.210	0.240
Total	σ_{tot}	0.194	0.208	0.235	0.289	0.348	0.416	0.528	0.625	0.761	0.929

Table 5.11: *$D_s^- \to \phi\pi^-$: Systematic uncertainties on the amplitude for the range $\Delta m_s = 1\ ps^{-1} - 10\ ps^{-1}$. The shifts of both the measured amplitude, $\Delta\mathcal{A}$, and its statistical uncertainty, $\Delta\sigma_{\mathcal{A}}$, are listed.*

Osc. frequency		11 ps^{-1}	12 ps^{-1}	13 ps^{-1}	14 ps^{-1}	15 ps^{-1}	16 ps^{-1}	17 ps^{-1}	18 ps^{-1}	19 ps^{-1}	20 ps^{-1}
$\mathcal{A}$		0.773	1.141	1.304	1.774	2.272	2.334	2.581	3.417	3.482	1.316
Stat. uncertainty		1.106	1.349	1.629	1.936	2.295	2.737	3.214	3.696	4.346	5.418
$\eta_s = 0.684$	$\Delta\mathcal{A}$	+0.035	+0.052	+0.060	+0.082	+0.104	+0.108	+0.121	+0.158	+0.163	+0.075
	$\Delta\sigma$	+0.048	+0.059	+0.071	+0.084	+0.100	+0.119	+0.140	+0.161	+0.189	+0.236
$c\bar{c}:$ 6%	$\Delta\mathcal{A}$	−0.039	+0.028	+0.047	+0.158	+0.299	+0.337	+0.497	+0.952	+1.266	+1.035
	$\Delta\sigma$	+0.123	+0.162	+0.213	+0.270	+0.341	+0.442	+0.556	+0.649	+0.783	+1.155
$\mathcal{B}(B_s^0 \to D_s^- D_s^+) =$ 4.7%	$\Delta\mathcal{A}$	−0.018	−0.025	−0.030	−0.040	−0.051	−0.054	−0.061	−0.078	−0.081	−0.039
	$\Delta\sigma$	−0.024	−0.029	−0.035	−0.043	−0.049	−0.058	−0.068	−0.079	−0.094	−0.117
$\mathcal{B}(B_s^0 \to \mu^+\nu D_s^- X) =$ 5.5%	$\Delta\mathcal{A}$	+0.024	+0.029	+0.028	+0.041	+0.058	+0.052	+0.052	+0.076	+0.060	−0.084
	$\Delta\sigma$	+0.049	+0.060	+0.072	+0.089	+0.102	+0.121	+0.142	+0.165	+0.196	+0.247
$c\tau_{B_s^0} = 455\ \mu$m	$\Delta\mathcal{A}$	−0.007	−0.002	−0.004	+0.003	+0.008	+0.004	+0.010	+0.033	+0.035	−0.025
	$\Delta\sigma$	+0.008	+0.011	+0.014	+0.017	+0.022	+0.028	+0.035	+0.041	+0.053	+0.078
$D^- \to K^+\pi^-\pi^-$ contribution	$\Delta\mathcal{A}$	+0.027	+0.041	+0.053	+0.065	+0.076	+0.089	+0.104	+0.119	+0.136	+0.156
1.6%	$\Delta\sigma$	+0.015	+0.018	+0.022	+0.026	+0.030	+0.035	+0.040	+0.046	+0.053	+0.063
$p_{T_\mu} > 4$ GeV/c	$\Delta\mathcal{A}$	−0.025	−0.025	−0.027	−0.041	−0.049	−0.049	−0.062	−0.076	−0.051	+0.039
	$\Delta\sigma$	−0.029	−0.035	−0.042	−0.050	−0.059	−0.069	−0.080	−0.088	−0.112	−0.138
$\Delta\Gamma/\Gamma = 0.2$	$\Delta\mathcal{A}$	+0.002	+0.001	+0.002	+0.002	+0.001	+0.002	+0.003	+0.000	−0.000	−0.002
	$\Delta\sigma$	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.002	+0.001
D_s^- mass changed to	$\Delta\mathcal{A}$	+0.006	+0.004	−0.010	−0.016	−0.002	+0.013	+0.017	+0.020	+0.026	+0.034
$M_s + 1\sigma$	$\Delta\sigma$	+0.001	−0.000	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.005	−0.005
D_s^- width changed to	$\Delta\mathcal{A}$	+0.016	+0.019	+0.023	+0.027	+0.032	+0.044	+0.048	+0.034	+0.021	+0.027
$\sigma_s - 1\sigma$	$\Delta\sigma$	+0.002	+0.002	+0.003	+0.003	−0.000	+0.004	+0.005	+0.005	+0.003	+0.006
D^- mass changed to	$\Delta\mathcal{A}$	−0.008	−0.009	−0.010	−0.011	−0.013	−0.014	−0.017	−0.022	−0.028	−0.036
$M_- - 1\sigma$	$\Delta\sigma$	−0.000	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001	−0.002	−0.002
Slope for bkg.	$\Delta\mathcal{A}$	−0.011	−0.013	−0.015	−0.020	−0.022	−0.023	−0.024	−0.023	−0.019	−0.008
changed by 1σ	$\Delta\sigma$	−0.008	−0.009	−0.011	−0.015	−0.022	−0.018	−0.021	−0.024	−0.029	−0.036
Bkg. parametrized	$\Delta\mathcal{A}$	+0.106	+0.125	+0.128	+0.161	+0.213	+0.268	+0.291	+0.272	+0.257	+0.269
by straight line	$\Delta\sigma$	+0.038	+0.046	+0.045	+0.065	+0.075	+0.088	+0.101	+0.115	+0.136	+0.199
Mass bin width	$\Delta\mathcal{A}$	−0.036	−0.028	−0.064	−0.069	+0.032	+0.060	−0.144	−0.429	−0.657	−0.840
smaller by 50%	$\Delta\sigma$	−0.004	−0.006	−0.001	+0.004	+0.003	−0.000	−0.009	−0.021	−0.036	−0.057
Wide Gaussion	$\Delta\mathcal{A}$	−0.005	−0.145	−0.078	−0.104	−0.323	−0.383	−0.254	−0.245	−0.304	−0.235
added in mass fit	$\Delta\sigma$	−0.019	−0.020	−0.016	−0.017	−0.023	−0.040	−0.066	−0.082	−0.087	−0.094
Same eff. dependence	$\Delta\mathcal{A}$	−0.001	−0.002	−0.008	−0.008	−0.008	−0.019	−0.029	−0.029	−0.046	−0.111
for signal and bkg	$\Delta\sigma$	+0.012	+0.015	+0.018	+0.020	+0.025	+0.030	+0.035	+0.041	+0.048	+0.060
Eff. slope varied	$\Delta\mathcal{A}$	−0.022	−0.033	−0.051	−0.069	−0.093	−0.130	−0.172	−0.211	−0.272	−0.379
by 1σ for signal	$\Delta\sigma$	+0.007	+0.009	+0.010	+0.010	+0.011	+0.011	+0.011	+0.009	+0.006	+0.073
Resolution $\mathcal{S}.\mathcal{F}. = 1.162$	$\Delta\mathcal{A}$	+0.114	+0.170	+0.156	+0.202	+0.243	+0.126	−0.031	−0.059	−0.232	−0.541
	$\Delta\sigma$	+0.203	+0.212	+0.339	+0.427	+0.531	+0.657	+0.804	+0.965	+1.096	+1.147
Resolution $\mathcal{S}.\mathcal{F}. = 2$	$\Delta\mathcal{A}$	−0.008	−0.008	−0.004	−0.004	−0.003	+0.005	+0.017	+0.029	+0.047	+0.061
for background	$\Delta\sigma$	−0.007	−0.008	−0.009	−0.011	−0.011	−0.010	−0.009	−0.008	−0.006	+0.008
K-factor variation	$\Delta\mathcal{A}$	−0.059	−0.061	−0.054	−0.168	−0.069	+0.039	−0.182	−0.317	+0.220	+1.269
2%	$\Delta\sigma$	−0.043	−0.053	−0.065	−0.077	−0.099	−0.131	−0.139	−0.150	−0.262	−0.421
K-factor using	$\Delta\mathcal{A}$	+0.021	+0.009	+0.033	+0.036	+0.037	+0.056	+0.048	−0.024	−0.050	+0.022
reconstructed values	$\Delta\sigma$	+0.009	+0.012	+0.014	+0.016	+0.018	+0.018	+0.018	+0.023	+0.027	+0.004
Using smoothed	$\Delta\mathcal{A}$	+0.003	−0.012	−0.015	−0.012	−0.003	−0.012	−0.015	−0.019	−0.023	−0.057
K-factor histograms	$\Delta\sigma$	+0.008	+0.135	+0.013	+0.015	+0.018	+0.022	+0.025	+0.029	+0.034	+0.037
Total syst.	σ_{tot}^{sys}	0.225	0.261	0.213	0.263	0.401	0.532	0.650	1.058	1.468	2.002
Total	σ_{tot}	1.129	1.374	1.643	1.953	2.330	2.788	3.279	3.844	4.587	5.776

Table 5.12: *$D_s^- \to \phi\pi^-$: Systematic uncertainties on the amplitude for the range $\Delta m_s = 11\ ps^{-1} - 20\ ps^{-1}$. The shifts of both the measured amplitude, $\Delta\mathcal{A}$, and its statistical uncertainty, $\Delta\sigma_{\mathcal{A}}$, are listed.*

Osc. frequency		1 ps^{-1}	2 ps^{-1}	3 ps^{-1}	4 ps^{-1}	5 ps^{-1}	6 ps^{-1}	7 ps^{-1}	8 ps^{-1}	9 ps^{-1}	10 ps^{-1}
$\mathcal{A}$		0.162	-0.395	-0.292	-0.546	0.239	0.794	0.316	-0.343	-0.401	-0.331
Stat. uncertainty		0.228	0.269	0.299	0.419	0.428	0.455	0.508	0.637	0.773	0.900
$\eta_s = 0.684$	$\Delta\mathcal{A}$	+0.005	-0.016	-0.013	-0.021	+0.013	+0.035	+0.016	-0.010	-0.013	-0.012
	$\Delta\sigma$	+0.010	+0.012	+0.013	+0.019	+0.019	+0.020	+0.022	+0.028	+0.034	+0.039
$c\bar{c}:$ 6%	$\Delta\mathcal{A}$	-0.000	-0.046	-0.025	-0.066	+0.018	+0.094	+0.037	-0.058	-0.059	-0.038
	$\Delta\sigma$	+0.012	+0.011	+0.010	+0.022	+0.024	+0.029	+0.035	+0.052	+0.065	+0.070
$\mathcal{B}(B_s^0 \to D_s^- D_s^+) =$ 4.7%	$\Delta\mathcal{A}$	-0.004	+0.011	+0.010	+0.015	-0.009	-0.023	-0.011	+0.006	+0.008	+0.008
	$\Delta\sigma$	-0.007	-0.009	-0.010	-0.013	-0.013	-0.014	-0.015	-0.019	-0.023	-0.027
$\mathcal{B}(B_s^0 \to \mu^+\nu D_s^- X) =$ 5.5%	$\Delta\mathcal{A}$	+0.028	-0.032	-0.022	-0.057	-0.007	+0.041	-0.004	-0.062	-0.065	-0.047
	$\Delta\sigma$	+0.015	+0.018	+0.021	+0.028	+0.029	+0.030	+0.033	+0.041	+0.050	+0.059
$c\tau_{B_s^0} = 455$ μm	$\Delta\mathcal{A}$	+0.002	+0.001	-0.002	-0.005	+0.005	+0.006	-0.004	-0.011	-0.006	-0.005
	$\Delta\sigma$	+0.000	+0.001	+0.000	+0.002	+0.002	+0.002	+0.003	+0.005	+0.006	+0.007
$\Delta\Gamma/\Gamma = 0.2$	$\Delta\mathcal{A}$	+0.001	-0.001	-0.002	-0.001	+0.001	-0.000	+0.000	+0.002	-0.000	-0.001
	$\Delta\sigma$	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.000	+0.000	+0.000	+0.001
D_s^- mass changed to	$\Delta\mathcal{A}$	+0.002	+0.019	+0.008	+0.017	-0.016	-0.032	-0.026	-0.022	-0.028	-0.022
$M_s + 1\sigma$	$\Delta\sigma$	-0.004	-0.004	-0.005	-0.008	-0.008	-0.008	-0.010	-0.014	-0.017	-0.019
D_s^- width changed to	$\Delta\mathcal{A}$	-0.005	+0.003	+0.000	+0.004	-0.001	+0.007	+0.008	+0.007	+0.003	+0.002
$\sigma_s - 1\sigma$	$\Delta\sigma$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	-0.001
$K^+\pi^-\pi^-$ G1 width	$\Delta\mathcal{A}$	+0.012	-0.004	+0.002	-0.006	-0.018	-0.017	-0.013	-0.005	+0.022	+0.042
decreased 1σ	$\Delta\sigma$	-0.003	-0.004	-0.004	-0.006	-0.005	-0.005	-0.004	-0.005	-0.006	-0.008
$K^+\pi^-\pi^-$ G2/G1 frac	$\Delta\mathcal{A}$	-0.024	-0.061	-0.019	-0.039	+0.045	+0.077	+0.060	+0.052	+0.017	-0.044
decreased 1σ	$\Delta\sigma$	+0.015	+0.018	+0.020	+0.033	+0.029	+0.028	+0.028	+0.038	+0.049	+0.057
D^- ratio	$\Delta\mathcal{A}$	-0.007	-0.039	-0.023	-0.014	+0.069	+0.073	+0.043	+0.036	+0.035	+0.026
increased by 25%	$\Delta\sigma$	+0.011	+0.013	+0.014	+0.024	+0.021	+0.020	+0.020	+0.025	+0.034	+0.039
Mass bin width	$\Delta\mathcal{A}$	+0.008	-0.015	+0.037	+0.044	+0.050	+0.062	+0.029	+0.008	-0.023	-0.085
smaller by 25%	$\Delta\sigma$	-0.001	-0.000	-0.001	-0.000	-0.001	-0.000	-0.001	-0.005	-0.007	-0.011
Bkg. parametrized	$\Delta\mathcal{A}$	-0.022	-0.002	+0.011	+0.035	-0.068	-0.044	-0.037	-0.076	-0.160	-0.171
by straight line	$\Delta\sigma$	-0.006	-0.005	-0.007	-0.010	-0.009	-0.010	-0.008	+0.001	-0.008	-0.011
Using corrected	$\Delta\mathcal{A}$	-0.015	-0.027	+0.040	+0.040	-0.039	+0.020	+0.075	+0.137	+0.154	+0.060
slopes for bkg	$\Delta\sigma$	-0.013	-0.018	-0.018	-0.025	-0.020	-0.017	-0.013	-0.019	-0.032	-0.033
Same eff. dependence	$\Delta\mathcal{A}$	+0.000	-0.005	-0.004	-0.005	-0.003	-0.001	-0.001	-0.001	-0.001	-0.001
for signal and bkg	$\Delta\sigma$	+0.001	+0.001	+0.002	+0.002	+0.002	+0.002	+0.002	+0.001	+0.002	+0.002
Eff. slope varied	$\Delta\mathcal{A}$	+0.000	-0.002	-0.002	-0.004	-0.001	+0.001	-0.001	-0.005	-0.008	-0.009
by 1σ for signal	$\Delta\sigma$	+0.000	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.002	+0.003	+0.003
Resolution $\mathcal{S}.\mathcal{F}. = 1.192$	$\Delta\mathcal{A}$	+0.003	-0.004	-0.007	-0.042	+0.003	+0.071	+0.029	-0.080	-0.121	-0.125
	$\Delta\sigma$	+0.001	+0.003	+0.006	+0.017	+0.022	+0.031	+0.042	+0.069	+0.106	+0.139
Resolution $\mathcal{S}.\mathcal{F}. = 2$	$\Delta\mathcal{A}$	-0.008	-0.002	-0.003	+0.010	+0.008	-0.002	+0.004	+0.017	+0.017	+0.012
for background	$\Delta\sigma$	-0.001	-0.000	-0.000	-0.002	-0.003	-0.005	-0.006	-0.008	-0.010	-0.012
K-factor variation	$\Delta\mathcal{A}$	+0.016	-0.022	+0.021	-0.015	-0.106	+0.003	+0.102	+0.077	-0.028	-0.009
2%	$\Delta\sigma$	+0.000	-0.004	-0.001	-0.007	-0.006	-0.002	-0.010	-0.024	-0.017	-0.031
K-factor using	$\Delta\mathcal{A}$	-0.000	+0.002	+0.000	+0.001	+0.003	+0.001	-0.005	-0.006	+0.008	+0.004
reconstructed values	$\Delta\sigma$	-0.000	-0.000	-0.000	-0.000	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002
Using smoothed	$\Delta\mathcal{A}$	-0.000	-0.000	+0.000	+0.000	-0.000	+0.000	+0.001	+0.001	-0.001	-0.001
K-factor histograms	$\Delta\sigma$	-0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.001
Total syst.	σ_{tot}^{sys}	0.142	0.163	0.133	0.161	0.249	0.218	0.223	0.234	0.285	0.276
Total (stat.+ syst.)	σ_{tot}	0.269	0.314	0.327	0.449	0.495	0.504	0.555	0.678	0.824	0.942

Table 5.13: $D_s^- \to K^{*0}K^-$ *: Systematic uncertainties on the amplitude for the range* $\Delta m_s = 1\ ps^{-1} - 10\ ps^{-1}$. *The shifts of both the measured amplitude,* $\Delta\mathcal{A}$*, and its statistical uncertainty,* $\Delta\sigma_{\mathcal{A}}$*, are listed.*

Osc. frequency		11 ps^{-1}	12 ps^{-1}	13 ps^{-1}	14 ps^{-1}	15 ps^{-1}	16 ps^{-1}	17 ps^{-1}	18 ps^{-1}	19 ps^{-1}	20 ps^{-1}
$\mathcal{A}$		-0.473	-0.083	0.750	0.886	-0.134	-0.645	-0.130	1.141	2.117	1.524
Stat. uncertainty		1.162	1.468	1.785	2.261	2.713	2.934	3.126	3.425	3.982	4.967
$\eta_s = 0.684$	$\Delta\mathcal{A}$	-0.020	-0.008	+0.022	+0.024	-0.016	-0.032	-0.003	+0.049	+0.092	+0.074
	$\Delta\sigma$	+0.051	+0.064	+0.078	+0.098	+0.118	+0.128	+0.116	+0.149	+0.191	+0.240
$c\bar{c}:$ 6%	$\Delta\mathcal{A}$	-0.072	-0.022	+0.100	+0.128	-0.064	-0.181	-0.066	+0.237	+0.534	+0.594
	$\Delta\sigma$	+0.097	+0.124	+0.144	+0.201	+0.277	+0.319	+0.326	+0.389	+0.469	+0.687
$\mathcal{B}(B_s^0 \to D_s^- D_s^+) =$ 4.7%	$\Delta\mathcal{A}$	+0.013	+0.005	-0.013	-0.015	+0.010	+0.019	+0.005	-0.033	-0.060	-0.047
	$\Delta\sigma$	-0.035	-0.044	-0.053	-0.067	-0.080	-0.087	-0.110	-0.101	-0.118	-0.137
$\mathcal{B}(B_s^0 \to \mu^+\nu D_s^- X) =$ 5.5%	$\Delta\mathcal{A}$	-0.039	+0.027	+0.136	+0.174	+0.080	-0.006	+0.006	+0.083	+0.137	+0.038
	$\Delta\sigma$	+0.077	+0.097	+0.119	+0.151	+0.181	+0.194	+0.187	+0.227	+0.265	+0.346
$c\tau_{B_s^0} = 455\ \mu$m	$\Delta\mathcal{A}$	-0.005	+0.005	+0.018	+0.013	-0.008	-0.014	+0.003	+0.023	+0.036	+0.017
	$\Delta\sigma$	+0.010	+0.011	+0.013	+0.018	+0.021	+0.021	+0.005	+0.030	+0.039	+0.070
$\Delta\Gamma/\Gamma = 0.2$	$\Delta\mathcal{A}$	+0.001	+0.000	-0.001	-0.002	+0.001	+0.001	+0.005	+0.000	-0.002	-0.001
	$\Delta\sigma$	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	-0.018	+0.001	+0.001	+0.002
D_s^- mass changed to	$\Delta\mathcal{A}$	-0.012	-0.030	-0.070	-0.074	+0.006	+0.093	+0.126	+0.108	+0.088	+0.104
$M_s + 1\sigma$	$\Delta\sigma$	-0.027	-0.034	-0.041	-0.057	-0.075	-0.079	-0.097	-0.081	-0.088	-0.093
D_s^- width changed to	$\Delta\mathcal{A}$	+0.002	+0.002	+0.003	+0.005	+0.007	+0.010	+0.012	+0.011	+0.013	+0.019
$\sigma_s - 1\sigma$	$\Delta\sigma$	-0.001	-0.000	+0.000	+0.000	-0.001	-0.002	-0.022	-0.005	-0.007	-0.002
$K^+\pi^-\pi^-$ G1 width	$\Delta\mathcal{A}$	+0.044	+0.010	-0.028	-0.063	-0.085	-0.107	-0.130	-0.174	-0.227	-0.294
decreased 1σ	$\Delta\sigma$	-0.013	-0.019	-0.026	-0.037	-0.041	-0.034	-0.048	-0.030	-0.033	-0.033
$K^+\pi^-\pi^-$ G2/G1 frac	$\Delta\mathcal{A}$	-0.100	-0.035	+0.151	+0.290	+0.208	+0.074	+0.057	+0.139	+0.267	+0.378
decreased 1σ	$\Delta\sigma$	+0.082	+0.114	+0.148	+0.212	+0.261	+0.247	+0.208	+0.230	+0.263	+0.272
D^- ratio	$\Delta\mathcal{A}$	-0.013	-0.058	-0.046	-0.000	+0.011	+0.030	+0.061	+0.107	+0.162	+0.182
increased by 25%	$\Delta\sigma$	+0.058	-0.054	+0.106	+0.151	+0.176	+0.159	+0.126	+0.150	+0.172	+0.224
Mass bin width	$\Delta\mathcal{A}$	-0.201	-0.335	-0.401	-0.402	-0.217	-0.081	-0.127	-0.250	-0.330	-0.353
smaller by 25%	$\Delta\sigma$	-0.018	-0.031	-0.044	-0.073	-0.101	-0.094	-0.095	-0.065	-0.061	-0.055
Bkg. parametrized	$\Delta\mathcal{A}$	-0.143	-0.113	-0.059	+0.107	+0.357	+0.465	+0.489	+0.538	+0.601	+0.685
by straight line	$\Delta\sigma$	-0.019	-0.031	-0.041	-0.048	-0.037	-0.024	-0.040	-0.026	-0.029	-0.015
Using corrected	$\Delta\mathcal{A}$	-0.030	-0.054	-0.008	-0.182	-0.438	-0.479	-0.409	-0.382	-0.432	-0.571
slopes for bkg	$\Delta\sigma$	-0.038	-0.034	-0.044	-0.080	-0.107	-0.101	-0.106	-0.083	-0.096	-0.203
Same eff. dependence	$\Delta\mathcal{A}$	+0.001	+0.003	+0.005	+0.005	+0.008	+0.010	+0.010	+0.001	-0.005	-0.001
for signal and bkg	$\Delta\sigma$	+0.002	+0.003	+0.004	+0.003	+0.001	-0.001	-0.020	-0.002	-0.004	+0.002
Eff. slope varied	$\Delta\mathcal{A}$	-0.014	-0.016	-0.018	-0.032	-0.038	-0.033	-0.042	-0.081	-0.128	-0.158
by 1σ for signal	$\Delta\sigma$	+0.005	+0.005	+0.003	-0.003	-0.016	-0.015	-0.051	-0.040	-0.056	-0.084
Resolution $\mathcal{S}.\mathcal{F}. = 1.192$	$\Delta\mathcal{A}$	-0.236	-0.210	-0.004	-0.070	-0.607	-0.919	-0.786	-0.273	+0.063	-0.492
	$\Delta\sigma$	+0.215	+0.310	+0.391	+0.529	+0.687	+0.807	+0.907	+1.075	+1.246	+1.474
Resolution $\mathcal{S}.\mathcal{F}. = 2$	$\Delta\mathcal{A}$	+0.013	-0.002	-0.030	-0.040	-0.015	+0.008	+0.013	-0.005	-0.009	+0.029
for background	$\Delta\sigma$	-0.016	-0.021	-0.026	-0.033	-0.039	-0.043	-0.063	-0.048	-0.049	-0.038
K-factor variation	$\Delta\mathcal{A}$	+0.026	-0.159	-0.169	+0.082	+0.302	+0.037	-0.265	-0.430	-0.220	+0.424
2%	$\Delta\sigma$	-0.064	-0.062	-0.089	-0.135	-0.105	-0.041	-0.080	-0.123	-0.239	-0.423
K-factor using	$\Delta\mathcal{A}$	-0.009	+0.004	+0.016	-0.004	-0.016	-0.035	+0.015	+0.065	+0.048	-0.008
reconstructed values	$\Delta\sigma$	-0.002	-0.004	-0.005	-0.005	-0.007	-0.010	-0.028	-0.008	-0.006	+0.025
Using smoothed	$\Delta\mathcal{A}$	+0.003	-0.006	-0.003	-0.012	+0.027	-0.003	+0.019	-0.009	-0.037	+0.015
K-factor histograms	$\Delta\sigma$	+0.001	+0.001	+0.002	+0.002	+0.002	+0.002	-0.016	+0.002	+0.003	+0.012
Total syst.	σ_{tot}^{sys}	0.323	0.477	0.533	0.612	0.859	0.906	0.906	0.956	1.017	1.441
Total (stat.+ syst.)	σ_{tot}	1.206	1.543	1.863	2.342	2.845	3.070	3.254	3.556	4.110	5.172

Table 5.14: $D_s^- \to K^{*0}K^-$ *: Systematic uncertainties on the amplitude for the range* $\Delta m_s = 11\ ps^{-1} - 20ps^{-1}$. *The shifts of both the measured amplitude,* $\Delta\mathcal{A}$*, and its statistical uncertainty,* $\Delta\sigma_{\mathcal{A}}$*, are listed.*

5.8 Results

The amplitude scans for the $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ samples incorporating the systematic uncertainties are shown in Fig. 5.27. The limits and sensitivities obtained at the 95% confidence level are listed below. The combined result is discussed in the next Chapter.

$$\begin{aligned} D_s^- \to \phi\pi^- \qquad & \text{Limit} \quad : \Delta m_s > 6.9 \text{ ps}^{-1} \\ & \text{Sensitivity} \quad : 7.8 \text{ ps}^{-1} \end{aligned} \tag{5.34}$$

$$\begin{aligned} D_s^- \to K^{*0}K^- \qquad & \text{Limit} \quad : \Delta m_s > 4.9 \text{ ps}^{-1} \\ & \text{Sensitivity} \quad : 7.4 \text{ ps}^{-1} \end{aligned} \tag{5.35}$$

Chapter 6

Conclusions and Outlook

The $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$ results obtained in Sec. 5.8 were combined using a procedure described in Ref. [75]. The measured amplitudes $\mathcal{A}$ at each test value of Δm_s were averaged and the following contributions to the systematic uncertainty were considered as correlated for the two analyses:

- Uncertainty in purity.
- Uncertainty in $c\bar{c}$ contamination.
- Uncertainty in the $B_s^0 \to D_s^- D_s^+ X$ branching ratio.
- Uncertainty in the $B_s^0 \to \mu^+ \nu D_s^- X$ branching ratio.
- Uncertainty in the B_s^0 lifetime.
- Non-zero $\Delta\Gamma_s/\Gamma_s$.

Figure 6.1 shows the combined DØ B_s^0 amplitude spectrum. The following result was obtained at the 95% confidence level:

$$\begin{aligned} \textbf{Combined} \quad \textbf{DØ} \quad & \textbf{Result} \\ \textbf{Limit} \quad : \quad & \boldsymbol{\Delta m_s > 7.3\ ps^{-1}} \\ \textbf{Sensitivity} \quad : \quad & \textbf{9.5 ps}^{-1}. \end{aligned} \tag{6.1}$$

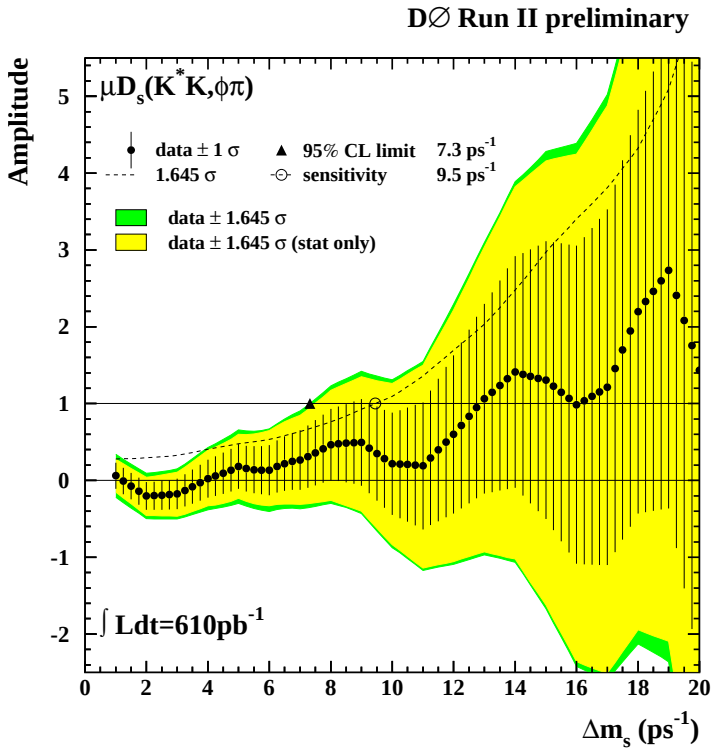


Figure 6.1: *The combined DØ B_s^0 oscillation amplitude spectrum. The points and error bars denote the measurements of the amplitude $\mathcal{A}$ and their statistical and systematic uncertainties, $\sigma_{\mathcal{A}}$. Values of Δm_s where the solid curve ("*data $\pm$ 1.645 $\sigma_{\mathcal{A}}$*") is below 1 are excluded at 95% C.L. The dashed curve shows* 1.645 $\sigma_{\mathcal{A}}$ *and is a measure of the sensitivity of the combined analysis.*

6.1 Combined World Average

The DØ result presented in the previous section was combined with results from other experiments and the new world average at 95% C.L. was obtained [76]:

$$\begin{aligned} \textbf{Combined} \quad & \textbf{World} \quad \textbf{Average} \\ \textbf{Limit} \quad & : \quad \mathbf{\Delta m_s > 14.4\ ps^{-1}} \\ \textbf{Sensitivity} \quad & : \quad \mathbf{18.9\ ps^{-1}}. \end{aligned}$$

Though the DØ limit on Δm_s (Eq. 6.1) has yet to push the world limit, the world average sensitivity increased from 18.5 ps^{-1} (before adding the new DØ result) to 18.9 ps^{-1} (with the inclusion of the new DØ result). Figure 6.2 shows the B_s^0 oscillation amplitude spectrum including results from other experiments and the preliminary DØ result presented in this dissertation[1]. All values of Δm_s below 14.4 ps^{-1} have been excluded at 95% confidence level. The values between 14.4 ps^{-1} and 21.8 ps^{-1} cannot be excluded since the data is compatible with a signal in this region. However, there is no deviation from $\mathcal{A} = 0$, and hence an observation of a signal cannot be claimed.

6.2 Outlook

Both the DØ and the combined world average amplitude spectra show that the B_s^0 mixing analyses are totally dominated by statistical uncertainties at the moment. Adding more data will definitely help push both the limit and the sensitivity up. It is expected that over the course of the next few years, the Tevatron will deliver

[1] At the time of writing this thesis, the CDF experiment presented a new preliminary result that is yet to be included in this world average.

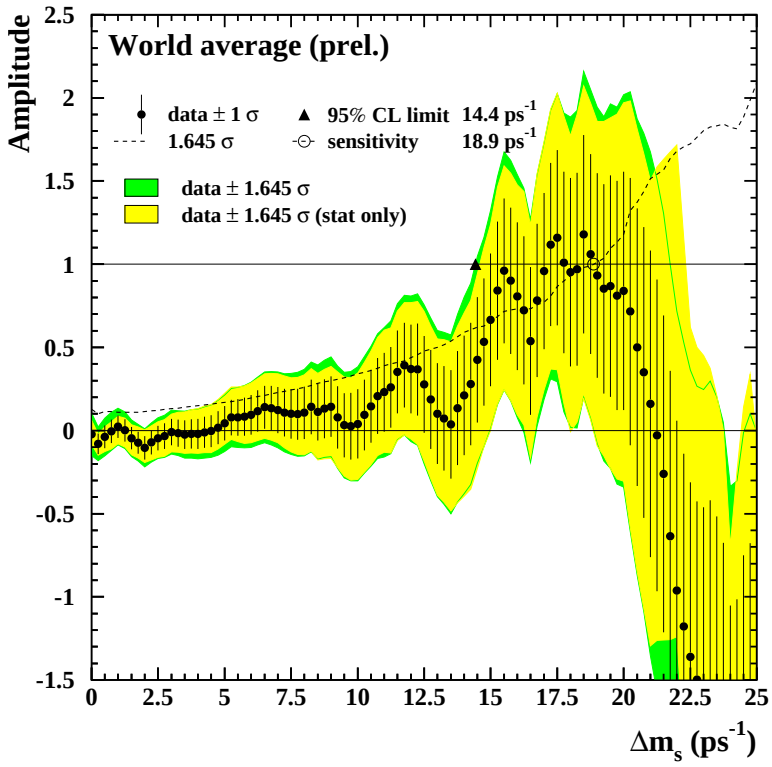


Figure 6.2: *Combined measurements of the B_s^0 oscillation amplitude as a function of Δm_s, including results from other experiments and the new preliminary DØ result.*

$\sim 8-9$ fb^{-1} of data to the experiments. This will certainly help reduce the statistical uncertainties. In addition to this, a number of other improvements have been planned:

- Analysis Improvements: The formula for the statistical significance (or the sensitivity) of a B_s^0 oscillation signal was defined in Sec. 3.4 to be

$$\mathcal{S} = \sqrt{\frac{\epsilon \mathcal{D}^2}{2}} \frac{S}{\sqrt{S+B}} e^{-(\Delta m_s \sigma_t)^2/2}. \tag{6.2}$$

Keeping the above in mind, the following improvements are planned in order to increase the sensitivity:

(i) Additional semileptonic decay channels:
The addition of the following $B_s^0 \to D_s^- \mu^+ X$ channels where $D_s^- \to K_S^0 K^-$, $D_s^- \to K^* K^{*-}$, and $D_s^- \to \pi^+\pi^-\pi^-$ will help increase the signal significantly.

(ii) Better tagging power (or $\epsilon\mathcal{D}^2$):
Improved opposite side flavor tagging with better modeling of probability density functions, and the addition of same side tagging will help increase $\epsilon\mathcal{D}^2$, thereby leading to larger, purer tagged samples.

(iii) Better proper time resolution (or σ_t)
As Eq. 3.11 indicated the K-factor resolution is significant for semileptonic decays (owing to the undetected neutrino) and may very well be the limiting factor in our ability to measure Δm_s using semileptonic decays alone. Studies are ongoing to reduce this resolution by determining the transverse momentum of the B_s^0 meson, $p_T(B_s^0)$, more accurately. One such study [77] estimates the B_s^0 momentum, p_B, by assuming that all of the missing mass can be attributed to the neutrino, i.e., it takes $E_X = E_B - E_{D\mu}$, and $\vec{p}_X = \vec{p}_B - \vec{p}_{D\mu}$, and estimates that $M_X = \sqrt{E_X^2 - |\vec{p}_X|^2}$. Setting $M_X = 0$ (for a neutrino), one ends up with a quadratic equation, and two solutions for p_B. For one particular case, where the two solutions were within 20% of each other for $\sim$ 15% of the data, the $\sigma(K)/K$ component of σ_t (in Eq. 3.11) was reduced by a factor of ~ 0.65. More work is needed in this particular area but the results are promising.

(iv) Improved analysis techniques:

The use of an unbinned likelihood fit and event-by-event resolution and tagging purity will help increase the overall sensitivity.

- Hadronic Modes:

  Hadronic modes like $B_s^0 \to D_s^- \pi^+$ have the distinct advantage that the events can be fully reconstructed leading to better proper time resolution. For selecting these events at DØ one has to rely on triggering on the opposite side lepton. This, coupled with the fact that these decays have smaller branching ratios compared to semileptonic decays leads to a small event sample. Work is ongoing to optimize the selection criteria for such decays for optimal yield. In the future the yield of such a B_s^0 sample can be increased with the help of trigger terms involving the L2STT trigger (Sec. 4.3.2). This trigger can potentially help select B events with large decay length and also reduce the momentum thresholds for the lepton on the opposite side. This will, in turn, help reduce the prescale factors on the single muon triggers used for selecting (semileptonic as well as hadronic) B_s^0 events.

- Layer 0:

  An additional layer of silicon will be placed at a radius of 1.7 cm inside the current detector during the shutdown period starting March 2006. Studies show that this layer will improve decay length resolution for hadronic decays by 30% even if we lose Layer 1 owing to age-related degradation and radiation damage [78]. The impact on semileptonic decays is not as significant owing to the K-factor smearing but will result in some improvement.

- Bandwidth Upgrade:

  A proposal has been submitted to increase the rate to tape from $\sim$ 50 Hz to 100 Hz [77]. The rate to tape is the chief limiting factor affecting triggers used

for B analyses, many of which are heavily prescaled at high luminosities. The rate to tape itself is limited primarily by our ability to reconstruct data online in a timely fashion and the cost of tapes. With larger computing resources enabling an increase in bandwidth, an additional 50 Hz of B physics data would be sent to computing farms at remote institutions for parallel processing. This would allow the prescales on the B-triggers to be lowered thus letting us write additional B events to tape.

Figure 6.3 has been adapted from Ref. [77] and shows the integrated luminosity needed for a 3σ measurement of Δm_s.The three curves (going from left to right) compare the average expected significance (or Δm_s reach) with (a) semileptonic modes and a bandwidth upgrade, (b) hadronic modes and the addition of Layer 0, and (c) hadronic modes, Layer 0, and the bandwidth upgrade. Figure 6.3 also shows that while less luminosity is need to make a measurement with semileptonic modes, a "resolution wall" is reached and hadronic modes are necessary to extend the range. With the $\sim 8-9$ fb^{-1} of integrated luminosity that is expected to be delivered by the Tevatron, DØ's reach covers the entire expected range for Δm_s predicted by the Standard Model (Table 2.5).

Needless to say, the next few years will be an exciting time at the Tevatron with huge amounts of data expected. The foundation for a measurement of Δm_s has been laid by the analysis described in this dissertation, and the stage is set for a measurement. Moreover, if with the expected luminosity, no signal is observed within the range expected by the Standard Model, it could potentially herald the discovery of new physics.

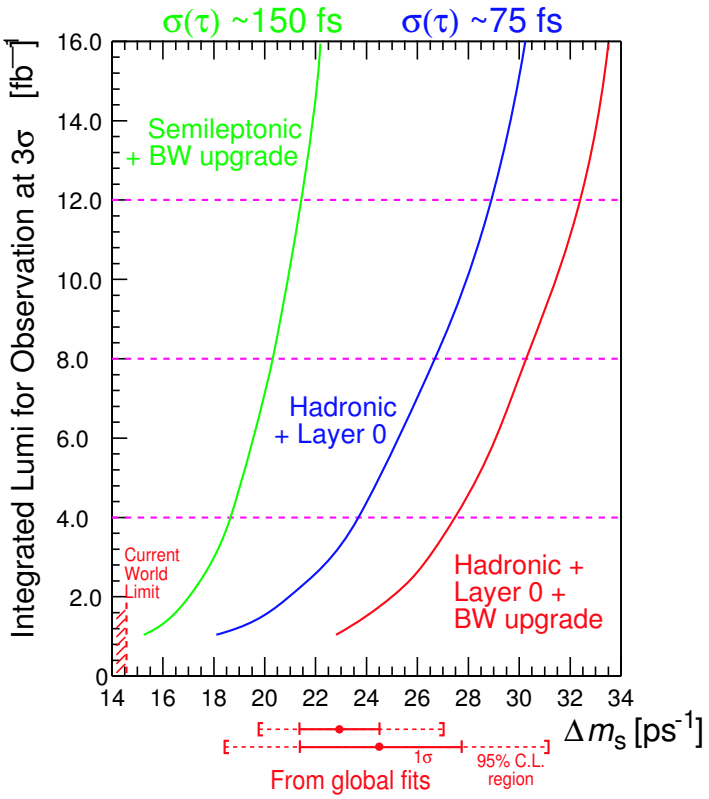


Figure 6.3: *Comparison of the amount of luminosity needed to achieve a measurement of Δm_s at the 3σ level as a function of Δm_s. The solid curves (left to right) are with (a) semileptonic modes and a bandwidth upgrade, (b) hadronic modes and the addition of Layer 0, and (c) hadronic modes, Layer 0, and the bandwidth upgrade. The hatched region shows the current experimental world limit at 95% C.L., and the points with error bars indicate the predicted range, assuming the Standard Model, from global fits to the unitarity triangle, either including or excluding current experimental limits on Δm_s in the fits [31].*

Bibliography

[1] The LEP Electroweak Working Group, arXiv:hep-ex/0511027.

[2] S. Eidelman *et al.*, Phys. Lett. **B592**, 1 (2004).

[3] Y. Fukuda *et al.* [Super-Kamiokande Collaboration], Phys. Rev. Lett. **81**, 1562 (1998) [arXiv:hep-ex/9807003].

[4] P. W. Higgs, Phys. Lett. **12**, 132 (1964).

[5] C. Quigg, "Gauge Theories of the Strong, Weak and Electromagnetic Interactions", Addison-Wesley, Reading, Mass., 1983.

[6] R. Barate *et al.* [LEP Working Group for Higgs boson searches], Phys. Lett. B **565**, 61 (2003) [arXiv:hep-ex/0306033].

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

[8] S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. **D2**, 1585 (1970).

[9] J. J. Aubert *et al.* [E598 Collaboration], Phys. Rev. Lett. **33**, 1404 (1974).

[10] J. E. Augustin *et al.* [SLAC-SP-017 Collaboration], Phys. Rev. Lett. **33**, 1406 (1974).

[11] M. Kobayashi and T. Maskawa, Progr. Theor. Phys. **49**, 652 (1973).

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

[13] S. W. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977).

[14] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **74**, 2626 (1995) [arXiv:hep-ex/9503002].

[15] S. Abachi *et al.* [D0 Collaboration], Phys. Rev. Lett. **74**, 2632 (1995) [arXiv:hep-ex/9503003].

[16] C. Jarlskog, Phys. Rev. Lett. **55**, 1039 (1985).

[17] L. L. Chau and W. Y. Keung, Phys. Rev. Lett. **53**, 1802 (1984).

[18] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[19] S. Burles, K. M. Nollett and M. S. Turner, Astrophys. J. **552**, L1 (2001) [arXiv:astro-ph/0010171].

[20] A. D. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967) [JETP Lett. **5**, 24 (1967 SOPUA,34,392-393.1991 UFNAA,161,61-64.1991)].

[21] R. Fleischer, arXiv:hep-ph/0405091.

[22] J. Charles *et al.* [CKMfitter Group], Eur. Phys. J. C **41**, 1 (2005) [arXiv:hep-ph/0406184].

[23] M. Gell-Mann and A. Pais, Phys. Rev. **97**, 1387 (1955).

[24] K. Lande, E. T. Booth, J. Impeduglia, L. M. Lederman and W. Chinowsky, Phys. Rev. **103**, 1901 (1956).

[25] C. Albajar *et al.* [UA1 Collaboration], Phys. Lett. B **186**, 247 (1987).

[26] H. Albrecht *et al.*, (ARGUS Collaboration), Phys. Lett. **B192**, 245 (1987).

[27] K. Anikeev *et al.*, arXiv:hep-ph/0201071.

[28] C. Gay, Ann. Rev. Nucl. Part. Sci. **50**, 577 (2000) [arXiv:hep-ex/0103016].

[29] A. Lenz, arXiv:hep-ph/0412007.

[30] I. Stewart, "QCD Effects in Weak Decays", Plenary talk at the 22nd International Symposium on Lepton-Photon Interactions at High Energy (LP 2005), Uppsala, Sweden (2005).

[31] M. Bona *et al.* [UTfit Collaboration], `http://www.utfit.org/`.

[32] M. Gronau and D. London, Phys. Rev. D **55**, 2845 (1997) [arXiv:hep-ph/9608430].

[33] A. G. Cohen, D. B. Kaplan and A. E. Nelson, Phys. Lett. B **388**, 588 (1996) [arXiv:hep-ph/9607394].

[34] I. I. Y. Bigi and F. Gabbiani, Nucl. Phys. B **352**, 309 (1991).

[35] K. Agashe, M. Papucci, G. Perez and D. Pirjol, arXiv:hep-ph/0509117.

[36] Y. Nir and D. J. Silverman, Nucl. Phys. B **345**, 301 (1990).

[37] Y. Nir and D. J. Silverman, Phys. Rev. D **42**, 1477 (1990).

[38] A. Ali and D. London, arXiv:hep-ph/0012155.

[39] M. Paulini, Int. J. Mod. Phys. A **14**, 2791 (1999) [arXiv:hep-ex/9903002].

[40] T. Sjostrand,Int. J. Mod. Phys. A **3**, 751 (1988).

[41] C. Peterson, D. Schlatter, I. Schmitt and P. M. Zerwas, Phys. Rev. D **27**, 105 (1983).

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

[43] T. Cole *et al.*, Technical report, Fermilab TM-1909 (1994).

[44] S. van der Meer, CERN/ISR-PO/72-31 (1994).

[45] G. L. Budker, Atomnaya Energiya **22**, 346-8 (1966).

[46] V. M. Abazov *et al.* [DØ Collaboration], submitted to Nucl. Instr. and Methods, hep-physics/0507191, Fermilab-Pub-05/341-E.

[47] S. Snyder, "Measurement of the Top Quark Mass at DØ", Ph.D. Thesis, S.U.N.Y., Stony Brook (1995).

[48] A. Garcia-Bellido *et al.*, "Primary Vertex certification in p14", DØ Note 4320 (2004).

[49] J. Abdallah *et al.* [DELPHI Collaboration], Eur. Phys. J. C **32**, 185 (2004).

[50] R. Illingsworth, "Development of Trigger Software for the Silicon and Fiber Trackers and a study of B meson lifetimes for the DØ experiment", Ph.D. Thesis, Imperial College, London, U.K. (2002).

[51] O. Boeriu, M. Hildreth, S. Jabeen, "CFT Light Yield Studies", DØ Note 4602 (2004).

[52] A. Khanov, "HTF: histogramming method for finding tracks. The algorithm description", DØ Note 3778 (2000).

[53] G. Borrisov, "Technical Details of AA Tracking", to be made a technical DØ Note.

[54] K. M. Black, "A Precision Measurement of the Top Quark Mass", Ph.D. Thesis, Boston University (2005).

[55] S. Abachi *et al.* [D0 Collaboration], Nucl. Instrum. Meth. A **324**, 53 (1993).

[56] E. Varnes, "Measurement of the Top Quark Mass at DØ Experiment", Ph.D. Thesis, U.C. Berkeley (1997).

[57] C. Luo, "Muon Reconstruction and B Physics Studies at the Tevatron Collider Experiment DØ", Ph.D. Thesis, Indiana University (2003).

[58] O. Peters, "Muon Segment Reconstruction: A Linked List Algorithm", DØ Note 3901 (2001).

[59] J. Kozminski *et al.*, "Electron Likelihood in P14", DØ Note 4449 (2004).

[60] F. Beaudette and J. F. Grivaz, "The Road Method", DØ Note 3976 (2002).

[61] T. Sjostrand, L. Lonnblad and S. Mrenna, arXiv:hep-ph/0108264.

[62] D. J. Lange, NIM A 462 (2001) 152; `http://www.slac.stanford.edu/~lange/EvtGen`.

[63] DØ Montecarlo Event Selection System, `http://www-clued0.fnal.gov/d0_mess/d0_mess_home.html`.

[64] Y. Fisyak, J. Womersley, "DØGSTAR: DØGEANT Simulation of the Total Apparatus Response", DØ Note 3191 (1997).

[65] S. Youssef *et al.*, "DØSIM User Manual", DØ Note 407 (1986).

[66] Application Software Group, "Detector Description and Simulation Tool", CERN Program Library Long Writeup W5013, CERN (1995).

[67] S. Catani, Yu. L. Dokshitzer, M. Olsson, G. Turnock, B. R. Webber, Phys. Lett. **B269** (1991) 432.

[68] T. Bose *et al.*, "Reconstruction of the $D_s^- \to K^{*0}K^-$ decay at DØ", DØ Note 4581 (2004).

[69] G. Borissov *et al.*, "B_d mixing measurement using semileptonic sample and combined opposite side tagging technique", DØ Note 4828 (2005).

[70] U. Nierste, *private communication.*

[71] B. Abbott *et al.* "Limits on B_s^0 mixing at DØ using $B_s^0 \to D_s^- \mu^+ X$ with opposite side tagging", DØ Note 4701 (2005).

[72] G. Borisov and C. Mariotti, NIM A 372 (1996) 181-187.

[73] V. M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **95**, 171801 (2005) [arXiv:hep-ex/0507084].

[74] V. M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **94**, 182001 (2005) [arXiv:hep-ex/0410052].

[75] H. F. A. Group (HFAG), "Averages of b-hadron properties as of winter 2005", arXiv:hep-ex/0505100.

[76] H. F. A. Group (HFAG), "Results for the summer 2005 conferences", `http://www.slac.stanford.edu/xorg/hfag/osc/summer_2005/index.html`.

[77] "B Bandwidth Upgrade", `http://www-d0.fnal.gov/Run2Physics/ckm/d0_private/bphys_future/index.html`.

[78] A. Kreymadhi and R. Van Kooten, "Estimated Improvement in Proper Time Resolution with the Addition of Layer 0 Silicon", DØ Note 4418 (2004).

Appendix A

Monte Carlo Samples

A.1 EvtGen User Decay Files

A.1.1 $D_s^- \to \phi\pi^-$:

User Decay File for $B_s^0 \to D_s^-\mu\nu X$ signal sample

```
noMixing
Alias myphi phi
Decay B_s0
0.0210   D_s-     mu+    nu_mu        PHOTOS  ISGW2;
0.0560   D_s*-    mu+    nu_mu        PHOTOS  HQET 0.92 1.18 0.72;
0.0020   D_s0*-   mu+    nu_mu        PHOTOS  ISGW2;
0.0037   D'_s1-   mu+    nu_mu        PHOTOS  ISGW2;
Enddecay
Decay D_s-
1.0000   myphi    pi-                  SVS;
```

```
Enddecay
Decay myphi
1.0000   K+       K-                  VSS;
Enddecay
End
```

User Decay File for $\bar{B}^0 \to D_s^{(*)-} D^{(*)+}$ background sample

```
noMixing
Alias myD_s*- D_s*-
Alias myD_s-  D_s-
Alias myD+    D+
Alias myD*+   D*+
Alias myD_1+  D_1+
Alias myD0    D0
Alias myD*0   D*0
Alias myphi   phi
Decay anti-B0
0.0096   myD+    myD_s-                PHSP;
0.0102   myD*+   myD_s-                SVS;
0.0120   myD_1+  myD_s-                SVS;
0.0094   myD_s*- myD+                  SVS;
0.0200   myD_s*- myD*+                 SVV_HELAMP 0.48 0.0 0.734 0.0 0.48 0.0;
0.0030   myD_s-  myD+  pi0             PHSP;
0.0060   myD_s-  myD0  pi+             PHSP;
0.0030   myD_s*- myD+  pi0             PHSP;
0.0060   myD_s*- myD0  pi+             PHSP;
0.0050   myD_s-  myD+  pi+  pi-        PHSP;
```

```
0.0037   myD_s-  myD+  pi0  pi0          PHSP;
0.0037   myD_s-  myD0  pi+  pi0          PHSP;
0.0050   myD_s*- myD+  pi+  pi-          PHSP;
0.0037   myD_s*- myD+  pi0  pi0          PHSP;
0.0037   myD_s*- myD0  pi+  pi0          PHSP;
Enddecay
Decay myD_1+
0.3333   myD*+   pi0                     VVS_PWAVE  0.0 0.0 0.0 0.0 1.0 0.0;
0.6667   myD*0   pi+                     VVS_PWAVE  0.0 0.0 0.0 0.0 1.0 0.0;
Enddecay
Decay myD*+
0.6830   myD0    pi+                     VSS;
0.3060   myD+    pi0                     VSS;
0.0110   myD+    gamma                   VSP_PWAVE;
Enddecay
Decay myD*0
0.6190   myD0    pi0                     VSS;
0.3810   myD0    gamma                   VSP_PWAVE;
Enddecay
Decay myD_s*-
0.942    myD_s-  gamma                   VSP_PWAVE;
0.058    myD_s-  pi0                     VSS;
Enddecay
Decay myD+
0.0430   anti-K*0   mu+   nu_mu          PHOTOS  ISGW2;
0.0700   anti-K0    mu+   nu_mu          PHOTOS  ISGW2;
0.0036   anti-K_10  mu+   nu_mu          PHOTOS  ISGW2;
0.0038   anti-K_2*0 mu+   nu_mu          PHOTOS  ISGW2;
```

```
0.0064   pi0        mu+   nu_mu         PHOTOS  ISGW2;
0.0028   eta        mu+   nu_mu         PHOTOS  ISGW2;
0.0011   eta'       mu+   nu_mu         PHOTOS  ISGW2;
0.0027   rho0       mu+   nu_mu         PHOTOS  ISGW2;
0.0029   omega      mu+   nu_mu         PHOTOS  ISGW2;
0.0027   K-  pi+    mu+   nu_mu         PHOTOS  PHSP;
0.0008   mu+        nu_mu               PHOTOS  SLN;
Enddecay
Decay myD0
0.0198   K*-   mu+ nu_mu                PHOTOS  ISGW2;
0.0322   K-    mu+ nu_mu                PHOTOS  ISGW2;
0.0014   K_1-  mu+ nu_mu                PHOTOS  ISGW2;
0.0015   K_2*- mu+ nu_mu                PHOTOS  ISGW2;
0.0037   pi-   mu+ nu_mu                PHOTOS  ISGW2;
0.0029   rho-  mu+ nu_mu                PHOTOS  ISGW2;
Enddecay
Decay myD_s-
1.0000   myphi    pi-                   SVS;
Enddecay
Decay myphi
1.0000   K+       K-                    VSS;
Enddecay
End
```

User Decay File for $B^- \to D_s^{(*)-} D^{(*)0}$ background sample

```
noMixing
Alias myD_s*- D_s*-
```

```
Alias myD_s-  D_s-
Alias myD+    D+
Alias myD*+   D*+
Alias myD_10  D_10
Alias myD0    D0
Alias myD*0   D*0
Alias myphi   phi
Decay B-
0.0096   myD0    myD_s-                 PHSP;
0.0102   myD*0   myD_s-                 SVS;
0.0120   myD_10  myD_s-                 SVS;
0.0094   myD_s*- myD0                   SVS;
0.0270   myD_s*- myD*0                  SVV_HELAMP 0.48 0.0 0.734 0.0 0.48 0.0;
0.0060   myD_s-  myD+  pi-              PHSP;
0.0030   myD_s-  myD0  pi0              PHSP;
0.0060   myD_s*- myD+  pi-              PHSP;
0.0030   myD_s*- myD0  pi0              PHSP;
0.0055   myD_s-  myD+  pi-  pi0         PHSP;
0.0055   myD_s-  myD0  pi-  pi+         PHSP;
0.0014   myD_s-  myD0  pi0  pi0         PHSP;
0.0055   myD_s*- myD+  pi-  pi0         PHSP;
0.0055   myD_s*- myD0  pi-  pi+         PHSP;
0.0014   myD_s*- myD0  pi0  pi0         PHSP;
Enddecay
Decay myD_10
0.3333   myD*0 pi0                     VVS_PWAVE  0.0 0.0 0.0 0.0 1.0 0.0;
0.6667   myD*+ pi-                     VVS_PWAVE  0.0 0.0 0.0 0.0 1.0 0.0;
Enddecay
```

```
Decay myD*+
0.6830   myD0    pi+                    VSS;
0.3060   myD+    pi0                    VSS;
0.0110   myD+    gamma                  VSP_PWAVE;
Enddecay
Decay myD*0
0.6190   myD0    pi0                    VSS;
0.3810   myD0    gamma                  VSP_PWAVE;
Enddecay
Decay myD_s*-
0.942    myD_s-  gamma                  VSP_PWAVE;
0.058    myD_s-  pi0                    VSS;
Enddecay
Decay myD+
0.0430   anti-K*0   mu+   nu_mu         PHOTOS  ISGW2;
0.0700   anti-K0    mu+   nu_mu         PHOTOS  ISGW2;
0.0036   anti-K_10  mu+   nu_mu         PHOTOS  ISGW2;
0.0038   anti-K_2*0 mu+   nu_mu         PHOTOS  ISGW2;
0.0064   pi0        mu+   nu_mu         PHOTOS  ISGW2;
0.0028   eta        mu+   nu_mu         PHOTOS  ISGW2;
0.0011   eta'       mu+   nu_mu         PHOTOS  ISGW2;
0.0027   rho0       mu+   nu_mu         PHOTOS  ISGW2;
0.0029   omega      mu+   nu_mu         PHOTOS  ISGW2;
0.0027   K-  pi+    mu+   nu_mu         PHOTOS  PHSP;
0.0008   mu+        nu_mu               PHOTOS  SLN;
Enddecay
Decay myD0
0.0198   K*-   mu+ nu_mu                PHOTOS  ISGW2;
```

```
0.0322   K-    mu+ nu_mu                PHOTOS  ISGW2;
0.0014   K_1-  mu+ nu_mu                PHOTOS  ISGW2;
0.0015   K_2*- mu+ nu_mu                PHOTOS  ISGW2;
0.0037   pi-   mu+ nu_mu                PHOTOS  ISGW2;
0.0029   rho-  mu+ nu_mu                PHOTOS  ISGW2;
Enddecay
Decay myD_s-
1.0000   myphi    pi-                   SVS;
Enddecay
Decay myphi
1.0000   K+       K-                    VSS;
Enddecay
End
```

User Decay File for $B_s{}^0 \to D_s^{(*)-} D_s^{(*)+}$ background sample

```
noMixing
Alias myD_s*- D_s*-
Alias myD_s-  D_s-
Alias myD_s+  D_s+
Alias myD_s*+ D_s*+
Alias myphi   phi
Decay B_s0
0.0086   myD_s-      myD_s+            PHSP;
0.0090   myD_s*+     myD_s-            SVS;
0.0099   myD_s*-     myD_s+            SVS;
0.0197   myD_s*-     myD_s*+           SVV_HELAMP  1.0 0.0 1.0 0.0 1.0 0.0;
Enddecay
```

```
Decay myD_s*+
0.942   myD_s+  gamma                   VSP_PWAVE;
0.058   myD_s+  pi0                     VSS;
Enddecay
Decay myD_s*-
0.942    myD_s-  gamma                  VSP_PWAVE;
0.058    myD_s-  pi0                    VSS;
Enddecay
Decay myD_s+
0.0200    phi       mu+    nu_mu        PHOTOS  ISGW2;
0.0260    eta       mu+    nu_mu        PHOTOS  ISGW2;
0.0089    eta'      mu+    nu_mu        PHOTOS  ISGW2;
0.0027    anti-K0   mu+    nu_mu        PHOTOS  ISGW2;
0.0010    anti-K*0  mu+    nu_mu        PHOTOS  ISGW2;
0.0046    mu+       nu_mu               PHOTOS  SLN;
Enddecay
Decay myD_s-
1.0000   myphi    pi-                   SVS;
Enddecay
Decay myphi
1.0000   K+       K-                    VSS;
Enddecay
End
```

A.1.2 $D_s^- \to K^{*0}K^-$:

User Decay File for $\bar{B}_s^0 \to D_s^+\mu^-\nu X$ signal sample

```
noMixing
Alias mytau- tau-
Alias mykstar anti-K*0
Alias myD_s+ D_s+
Alias myD_s*+ D_s*+
Alias myD_s0*+ D_s0*+
Alias myD'_s1+ D'_s1+
Decay anti-B_s0
0.0210   myD_s+    mu-     anti-nu_mu   PHOTOS  ISGW2;
0.0560   myD_s*+   mu-     anti-nu_mu   PHOTOS  HQET 0.92 1.18 0.72;
0.0020   myD_s0*+  mu-     anti-nu_mu   PHOTOS  ISGW2;
0.0037   myD'_s1+  mu-     anti-nu_mu   PHOTOS  ISGW2;
0.0080   myD_s+    mytau-  anti-nu_tau          ISGW2;
0.0160   myD_s*+   mytau-  anti-nu_tau          ISGW2;
0.0018   myD_s0*+  mytau-  anti-nu_tau          ISGW2;
0.0028   myD'_s1+  mytau-  anti-nu_tau          ISGW2;
Enddecay
Decay myD'_s1+
1.0000  myD_s*+ pi0                    VVS_PWAVE  1.0 0.0 0.0 0.0 0.0 0.0;
Enddecay
Decay myD_s0*+
1.0000  myD_s+  pi0                    PHSP;
Enddecay
Decay myD_s*+
```

```
0.942   myD_s+  gamma                  VSP_PWAVE;
0.058   myD_s+  pi0                    VSS;
Enddecay
Decay myD_s+
1.0000   mykstar    K+                 SVS;
Enddecay
Decay mykstar
1.0000   K-       pi+                  VSS;
Enddecay
End
```

User Decay File for $\bar{B}^0 \to D_s^{(*)-} D$ background sample

```
noMixing
Alias myD_s*- D_s*-
Alias myD_s-  D_s-
Alias myD+    D+
Alias myD*+   D*+
Alias myD_1+  D_1+
Alias myD0    D0
Alias myD*0   D*0
Alias mykstar   K*0
Decay anti-B0
0.0096   myD+    myD_s-               PHSP;
0.0102   myD*+   myD_s-               SVS;
0.0120   myD_1+  myD_s-               SVS;
0.0094   myD_s*- myD+                 SVS;
0.0200   myD_s*- myD*+                SVV_HELAMP 0.48 0.0 0.734 0.0 0.48
```

```
0.0;
0.0030   myD_s-  myD+  pi0              PHSP;
0.0060   myD_s-  myD0  pi+              PHSP;
0.0030   myD_s*- myD+  pi0              PHSP;
0.0060   myD_s*- myD0  pi+              PHSP;
0.0050   myD_s-  myD+  pi+  pi-         PHSP;
0.0037   myD_s-  myD+  pi0  pi0         PHSP;
0.0037   myD_s-  myD0  pi+  pi0         PHSP;
0.0050   myD_s*- myD+  pi+  pi-         PHSP;
0.0037   myD_s*- myD+  pi0  pi0         PHSP;
0.0037   myD_s*- myD0  pi+  pi0         PHSP;
Enddecay
Decay myD_1+
0.3333   myD*+   pi0                    VVS_PWAVE  0.0 0.0 0.0 0.0 1.0 0.0;
0.6667   myD*0   pi+                    VVS_PWAVE  0.0 0.0 0.0 0.0 1.0 0.0;
Enddecay
Decay myD*+
0.6830   myD0    pi+                    VSS;
0.3060   myD+    pi0                    VSS;
0.0110   myD+    gamma                  VSP_PWAVE;
Enddecay
Decay myD*0
0.6190   myD0    pi0                    VSS;
0.3810   myD0    gamma                  VSP_PWAVE;
Enddecay
Decay myD_s*-
0.942    myD_s-  gamma                  VSP_PWAVE;
0.058    myD_s-  pi0                    VSS;
```

```
Enddecay
Decay myD+
0.0430   anti-K*0   mu+   nu_mu          PHOTOS  ISGW2;
0.0700   anti-K0    mu+   nu_mu          PHOTOS  ISGW2;
0.0036   anti-K_10  mu+   nu_mu          PHOTOS  ISGW2;
0.0038   anti-K_2*0 mu+   nu_mu          PHOTOS  ISGW2;
0.0064   pi0        mu+   nu_mu          PHOTOS  ISGW2;
0.0028   eta        mu+   nu_mu          PHOTOS  ISGW2;
0.0011   eta'       mu+   nu_mu          PHOTOS  ISGW2;
0.0027   rho0       mu+   nu_mu          PHOTOS  ISGW2;
0.0029   omega      mu+   nu_mu          PHOTOS  ISGW2;
0.0027   K-  pi+    mu+   nu_mu          PHOTOS  PHSP;
0.0008   mu+        nu_mu               PHOTOS  SLN;
Enddecay
Decay myD0
0.0198   K*-   mu+ nu_mu                PHOTOS  ISGW2;
0.0322   K-    mu+ nu_mu                PHOTOS  ISGW2;
0.0014   K_1-  mu+ nu_mu                PHOTOS  ISGW2;
0.0015   K_2*- mu+ nu_mu                PHOTOS  ISGW2;
0.0037   pi-   mu+ nu_mu                PHOTOS  ISGW2;
0.0029   rho-  mu+ nu_mu                PHOTOS  ISGW2;
Enddecay
Decay myD_s-
1.0000   mykstar   K-                   SVS;
Enddecay
Decay mykstar
1.0000   K+  pi-                        VSS;
Enddecay
```

```
End
```

User Decay File for $B^- \to D_s^{(*)-} D$ background sample

```
noMixing
Alias myD_s*- D_s*-
Alias myD_s-  D_s-
Alias myD+    D+
Alias myD*+   D*+
Alias myD_10  D_10
Alias myD0    D0
Alias myD*0   D*0
Alias mykstar   K*0
Decay B-
0.0096   myD0    myD_s-                PHSP;
0.0102   myD*0   myD_s-                SVS;
0.0120   myD_10  myD_s-                SVS;
0.0094   myD_s*- myD0                  SVS;
0.0270   myD_s*- myD*0                 SVV_HELAMP 0.48 0.0 0.734 0.0 0.48 0.0;
0.0060   myD_s-  myD+  pi-             PHSP;
0.0030   myD_s-  myD0  pi0             PHSP;
0.0060   myD_s*- myD+  pi-             PHSP;
0.0030   myD_s*- myD0  pi0             PHSP;
0.0055   myD_s-  myD+  pi-  pi0        PHSP;
0.0055   myD_s-  myD0  pi-  pi+        PHSP;
0.0014   myD_s-  myD0  pi0  pi0        PHSP;
0.0055   myD_s*- myD+  pi-  pi0        PHSP;
0.0055   myD_s*- myD0  pi-  pi+        PHSP;
```

```
0.0014   myD_s*- myD0  pi0  pi0        PHSP;
Enddecay
Decay myD_10
0.3333   myD*0 pi0                     VVS_PWAVE  0.0 0.0 0.0 0.0 1.0 0.0;
0.6667   myD*+ pi-                     VVS_PWAVE  0.0 0.0 0.0 0.0 1.0 0.0;
Enddecay
Decay myD*+
0.6830   myD0    pi+                   VSS;
0.3060   myD+    pi0                   VSS;
0.0110   myD+    gamma                 VSP_PWAVE;
Enddecay
Decay myD*0
0.6190   myD0    pi0                   VSS;
0.3810   myD0    gamma                 VSP_PWAVE;
Enddecay
Decay myD_s*-
0.942    myD_s-  gamma                 VSP_PWAVE;
0.058    myD_s-  pi0                   VSS;
Enddecay
Decay myD+
0.0430   anti-K*0   mu+   nu_mu        PHOTOS  ISGW2;
0.0700   anti-K0    mu+   nu_mu        PHOTOS  ISGW2;
0.0036   anti-K_10  mu+   nu_mu        PHOTOS  ISGW2;
0.0038   anti-K_2*0 mu+   nu_mu        PHOTOS  ISGW2;
0.0064   pi0        mu+   nu_mu        PHOTOS  ISGW2;
0.0028   eta        mu+   nu_mu        PHOTOS  ISGW2;
0.0011   eta'       mu+   nu_mu        PHOTOS  ISGW2;
0.0027   rho0       mu+   nu_mu        PHOTOS  ISGW2;
```

```
0.0029   omega      mu+   nu_mu        PHOTOS  ISGW2;
0.0027   K-  pi+    mu+   nu_mu        PHOTOS  PHSP;
0.0008   mu+        nu_mu              PHOTOS  SLN;
Enddecay
Decay myD0
0.0198   K*-   mu+ nu_mu               PHOTOS  ISGW2;
0.0322   K-    mu+ nu_mu               PHOTOS  ISGW2;
0.0014   K_1-  mu+ nu_mu               PHOTOS  ISGW2;
0.0015   K_2*- mu+ nu_mu               PHOTOS  ISGW2;
0.0037   pi-   mu+ nu_mu               PHOTOS  ISGW2;
0.0029   rho-  mu+ nu_mu               PHOTOS  ISGW2;
Enddecay
Decay myD_s-
1.0000   mykstar    K-                 SVS;
Enddecay
Decay mykstar
1.0000   K+       pi-                  VSS;
Enddecay
End
```

User Decay File for ${B_s}^0 \to D_s^{(*)-} D_s^{(*)+}$ background sample

```
noMixing
Alias myD_s*- D_s*-
Alias myD_s-  D_s-
Alias myD_s+  D_s+
Alias myD_s*+ D_s*+
Alias mykstar   K*0
```

```
Decay anti-B_s0
0.0086   myD_s-      myD_s+          PHSP;
0.0090   myD_s*+     myD_s-          SVS;
0.0099   myD_s*-     myD_s+          SVS;
0.0197   myD_s*-     myD_s*+         SVV_HELAMP  1.0 0.0 1.0 0.0 1.0 0.0;
Enddecay
Decay myD_s*+
0.942   myD_s+  gamma                SVP_PWAVE;
0.058   myD_s+  pi0                  VSS;
Enddecay
Decay myD_s*-
0.942    myD_s-  gamma               VSP_PWAVE;
0.058    myD_s-  pi0
Enddecay
Decay myD_s+
0.0200    phi       mu+    nu_mu     PHOTOS  ISGW2;
0.0260    eta       mu+    nu_mu     PHOTOS  ISGW2;
0.0089    eta'      mu+    nu_mu     PHOTOS  ISGW2;
0.0027    anti-K0   mu+    nu_mu     PHOTOS  ISGW2;
0.0010    anti-K*0  mu+    nu_mu     PHOTOS  ISGW2;
0.0046    mu+       nu_mu            PHOTOS  SLN;
Enddecay
Decay myD_s-
1.0000   mykstar    K-              SVS;
Enddecay
Decay mykstar
1.0000   K+       pi-               VSS;
Enddecay
```

```
End
```

User Decay File for $B_s^0 \to D_s^{(*)-} D^+$ background sample

```
noMixing
Alias myD_s*- D_s*-
Alias myD_s-  D_s-
Alias myD+    D+
Alias myD*+   D*+
Alias myD0    D0
Alias mykstar K*0
Decay anti-B_s0
0.0096   myD_s-   myD+     K0        PHSP;
0.0096   myD_s*-  myD+     K0        PHSP;
0.0096   myD_s*-  myD0     K+        PHSP;
0.0024   myD_s-   myD+     pi0  K0   PHSP;
0.0048   myD_s-   myD0     pi+  K0   PHSP;
0.0048   myD_s-   myD+     pi-  K+   PHSP;
0.0024   myD_s-   myD0     pi0  K+   PHSP;
0.0024   myD_s*-  myD+     pi0  K0   PHSP;
0.0048   myD_s*-  myD0     pi+  K0   PHSP;
0.0048   myD_s*-  myD+     pi-  K+   PHSP;
0.0024   myD_s*-  myD0     pi0  K+   PHSP;
0.0017   myD_s-   myD+               PHSP;
0.0017   myD*+    myD_s-             SVS;
0.0017   myD_s*-  myD+               SVS;
0.0017   myD_s*-  myD*+              SVV_HELAMP  1.0 0.0 1.0 0.0 1.0 0.0;
Enddecay
```

```
Decay myD*+
0.6830   myD0    pi+                VSS;
0.3060   myD+    pi0                VSS;
0.0110   myD+    gamma              VSP_PWAVE;
Enddecay
Decay myD_s*-
0.942    myD_s-  gamma              VSP_PWAVE;
0.058    myD_s-  pi0                VSS;
Enddecay
Decay myD+
0.0430   anti-K*0   mu+   nu_mu      PHOTOS  ISGW2;
0.0700   anti-K0    mu+   nu_mu      PHOTOS  ISGW2;
0.0036   anti-K_10  mu+   nu_mu      PHOTOS  ISGW2;
0.0038   anti-K_2*0 mu+   nu_mu      PHOTOS  ISGW2;
0.0064   pi0        mu+   nu_mu      PHOTOS  ISGW2;
0.0028   eta        mu+   nu_mu      PHOTOS  ISGW2;
0.0011   eta'       mu+   nu_mu      PHOTOS  ISGW2;
0.0027   rho0       mu+   nu_mu      PHOTOS  ISGW2;
0.0029   omega      mu+   nu_mu      PHOTOS  ISGW2;
0.0027   K-  pi+    mu+   nu_mu      PHOTOS  PHSP;
0.0008   mu+        nu_mu            PHOTOS  SLN;
Enddecay
Decay myD0
0.0198   K*-   mu+ nu_mu            PHOTOS  ISGW2;
0.0322   K-    mu+ nu_mu            PHOTOS  ISGW2;
0.0014   K_1-  mu+ nu_mu            PHOTOS  ISGW2;
0.0015   K_2*- mu+ nu_mu            PHOTOS  ISGW2;
0.0037   pi-   mu+ nu_mu            PHOTOS  ISGW2;
```

```
0.0029   rho-  mu+ nu_mu            PHOTOS  ISGW2;
Enddecay
Decay myD_s-
1.0000   mykstar    K-              SVS;
Enddecay
Decay mykstar
1.0000   K+       pi-               VSS;
Enddecay
End
```

User Decay File for $\bar{B_s}^0 \to D_s^{(*)+} D^-$ background sample

```
noMixing
Alias myD_s*+    D_s*+
Alias myD_s+     D_s+
Alias myD-       D-
Alias myD*-      D*-
Alias myanti-D0  anti-D0
Alias myanti-D*0 anti-D*0
Alias mykstar  anti-K*0
Decay anti-B_s0
0.0150    myD_s*+   myanti-D*0 K-        PHSP;
0.0150    myD_s*+   myD*-      anti-K0   PHSP;
0.0050    myD_s*+   myanti-D0  K-        PHSP;
0.0050    myD_s*+   myD-       anti-K0   PHSP;
0.0050    myD_s+    myanti-D*0 K-        PHSP;
0.0050    myD_s+    myD*-      anti-K0   PHSP;
0.0020    myD_s+    myanti-D0  K-        PHSP;
```

```
0.0020    myD_s+    myD-       anti-K0   PHSP;
0.0030    myD_s*+   myanti-D*0 K*-       PHSP;
0.0030    myD_s*+   myD*-      anti-K*0  PHSP;
0.0050    myD_s*+   myanti-D0  K*-       PHSP;
0.0050    myD_s*+   myD-       anti-K*0  PHSP;
0.0025    myD_s+    myanti-D*0 K*-       PHSP;
0.0025    myD_s+    myD*-      anti-K*0  PHSP;
0.0025    myD_s+    myanti-D0  K*-       PHSP;
0.0025    myD_s+    myD-       anti-K*0  PHSP;
Enddecay
Decay myD*-
0.6830   myanti-D0  pi-                  VSS;
0.3060   myD-       pi0                  VSS;
0.0110   myD-       gamma                VSP_PWAVE;
Enddecay
Decay myD_s*+
0.942    myD_s+     gamma                VSP_PWAVE;
0.058    myD_s+     pi0                  VSS;
Enddecay
Decay myanti-D*0
0.6190   myanti-D0  pi0                  VSS;
0.3810   myanti-D0  gamma                VSP_PWAVE;
Enddecay
Decay myD-
0.0430   K*0   mu-  anti-nu_mu           PHOTOS  ISGW2;
0.0700   K0    mu-  anti-nu_mu           PHOTOS  ISGW2;
0.0036   K_10  mu-  anti-nu_mu           PHOTOS  ISGW2;
0.0038   K_2*0 mu-  anti-nu_mu           PHOTOS  ISGW2;
```

```
0.0064   pi0   mu-  anti-nu_mu          PHOTOS  ISGW2;
0.0028   eta   mu-  anti-nu_mu          PHOTOS  ISGW2;
0.0011   eta'  mu-  anti-nu_mu          PHOTOS  ISGW2;
0.0027   rho0  mu-  anti-nu_mu          PHOTOS  ISGW2;
0.0029   omega mu-  anti-nu_mu          PHOTOS  ISGW2;
0.0027   K+    pi-  mu- anti-nu_mu      PHOTOS  PHSP;
0.0008   mu-   anti-nu_mu               PHOTOS  SLN;
Enddecay
Decay myanti-D0
0.0198   K*+   mu- anti-nu_mu          PHOTOS  ISGW2;
0.0322   K+    mu- anti-nu_mu          PHOTOS  ISGW2;
0.0014   K_1+  mu- anti-nu_mu          PHOTOS  ISGW2;
0.0015   K_2*+ mu- anti-nu_mu          PHOTOS  ISGW2;
0.0037   pi+   mu- anti-nu_mu          PHOTOS  ISGW2;
0.0029   rho+  mu- anti-nu_mu          PHOTOS  ISGW2;
Enddecay
Decay myD_s+
1.0000   mykstar    K+                 SVS;
Enddecay
Decay mykstar
1.0000   K-       pi+                 VSS;
Enddecay
End
```

A.2 d0_mess Files

A.2.1 $B_s^0 \to D_s^- \mu^+ \nu X$ $(D_s^- \to \phi\pi^-)$:

All B_s^0 decays to $D_s^- \mu^+ \nu X$ were used and the D_s^- was forced to decay to $\phi\pi^-$. The ϕ was forced to decay to K^+K^- (see Appendix A.1.1). The `d0_mess` file filtered events using the following criteria:

- Atleast one B_s^0 meson.
- A μ^+ with $p_T > 1.9$ GeV/c and $|\eta| < 2.1$ with a B_s^0 as a parent.
- A D_s^- with $p_T > 1.0$ GeV/c and the following daughters:
  - A π^- with $p_T > 0.6$ GeV/c and $|\eta| < 2.1$.
  - A ϕ with K^+, K^- daughters, each with $p_T > 0.6$ GeV/c and $|\eta| < 2.1$.

```
string PackageName = "d0_mess"
bool d0_mess_on = true
int NumberOfCuts = 7
string Cut1 = "PdgId == 531"
string Cut2 = "PdgId ==  -13 && Pt > 1.9 && AbsEta < 2.1 && ParentId == 531"
string Cut3 = "PdgId == -431 && Pt > 1.0"
string Cut4 = "PdgId == -211 && Pt > 0.6 && AbsEta < 2.1 && ParentId == -431"
string Cut5 = "PdgId ==  333 && ParentId == -431"
string Cut6 = "PdgId ==  321 && Pt > 0.6 && AbsEta < 2.1 && ParentId ==  333"
string Cut7 = "PdgId == -321 && Pt > 0.6 && AbsEta < 2.1 && ParentId ==  333"
```

A.2.2 $\bar{B}_s^0 \to D_s^+ \mu^- \nu X$ $(D_s^+ \to \bar{K}^{*0} K^+)$:

All $\bar{B}_s^0$ decays to $D_s^+ \mu^- \nu X$ were used and the D_s^+ was forced to decay to $\bar{K}^{*0} K^+$. The $\bar{K}^{*0}$ was forced to decay to $K^- \pi^+$ (see Appendix A.1.2). The `d0_mess` file filtered events using the following criteria:

- Atleast one $\bar{B}_s^0$ meson.
- A μ^- with $p_T > 1.9$ GeV/c and $|\eta| < 2.1$ with a $\bar{B}_s^0$ as a parent.
- A D_s^+ with $p_T > 1.0$ GeV/c and the following daughters:
  - A K^+ with $p_T > 1.0$ GeV/c and $|\eta| < 2.1$.
  - A $\bar{K}^{*0}$ with K^-, π^+ daughters, each with $p_T > 1.0$ GeV/c and $|\eta| < 2.1$.

```
string PackageName = "d0_mess"
bool d0_mess_on = true
int NumberOfCuts = 7
string Cut1 = "PdgId == -531"
string Cut2 = "PdgId ==   13 && Pt > 1.9 && AbsEta < 2.1 && ParentId == -531"
string Cut3 = "PdgId ==  431 && Pt > 1.0"
string Cut4 = "PdgId ==  321 && Pt > 1.0 && AbsEta < 2.1 && ParentId ==  431"
string Cut5 = "PdgId == -313 && ParentId ==  431"
string Cut6 = "PdgId == -321 && Pt > 1.0 && AbsEta < 2.1 && ParentId ==  -313"
string Cut7 = "PdgId ==  211 && Pt > 1.0 && AbsEta < 2.1 && ParentId ==  -313"
```