

A Search For B_s^0 Oscillations at the Tevatron Collider Experiment DØ

Dan Krop

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Physics
Indiana University
April 2007

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment
of the requirements for the degree of Doctor of Philosophy.

Rick Van Kooten, Ph.D.

Alan Kostelecky, Ph.D.

Jon Urheim, Ph.D.

Andrzej Ziemiński, Ph.D.

April 23, 2007

Acknowledgements

I would like to first thank my adviser, Rick Van Kooten, for his guidance and encouragement over the last two and a half years. His advice has always been insightful and he has helped me through many difficult problems during the formulation of this analysis. He is, in my opinion, an excellent physicist and adviser. I feel very lucky to have had the opportunity to work with him.

I would also like to express my gratitude to Heidi Schellman for giving me numerous opportunities and repeatedly expressing confidence in my abilities. She has been unreasonably kind to me and I am in her debt. There is no way that I would have been able to write this thesis without her seemingly endless support and encouragement.

Brendan Casey deserves a special mention here for giving me the direction to do these analyses, for putting up with me camping out in his office night and day while he taught me how to do physics, and for having a good sense of humor. He has been an extraordinary teacher to me, both for analysis techniques and phenomenology. I can say without hesitation that I could not have done this work without his constant help.

Special thanks go to the members of the B physics group without whom I would not have been able to complete any work: Guennadi for the most robust and easy to use code base I have ever used, Sergey for the likelihood framework and the many analytical insights, Andrei for his approachability, Cano for all of his help generating Monte Carlo samples, and Nayeem for the many discussions and hours of collaboration.

Thanks go to all the faculty and students at Indiana University from whom I have learned in the last six years. I would like to thank John Challi-four for an outstanding E&M course and Alan Kostelecky for the best particle physics course I have ever taken. I am also grateful to Pawel, Ewa, Chris, Nick, and Carrie for helping me have fun when we should have all been doing homework. Thanks also go to Niko for the many excellent operatic performances and to Maciej for the many jokes (as well as occasional physics insight). My DØ office-mates, Jason and Leah, deserve a special mention for putting up with me when I wanted to throw the computer out the window.

I'd also like to thank all the members of the DØ Northwestern group with whom I have spent innumerable hours trying not to get back to work. I couldn't have asked for a better group of people to work with (or near) than Tim, Derek, Michelle, Meghan, Sahal, and Jon.

Finally, I would like to profusely thank my family. My brother has been there for me with encouraging words when I felt like I would never finish, and my parents have supported me throughout the many years and difficult moments. They have never wavered, even when it seemed that things would

not work out. I feel extremely grateful to them. I know I would not be getting this degree without them. Lastly, and most importantly, I would like to thank my beautiful future wife and best friend, Ami. Her patience with me is unbelievable as is her willingness to sit through an hour long practice talk (more than once) that must have bored her to tears. She is a wonderful companion and fiance and I am eternally grateful that I have been blessed with her presence.

Abstract

We present a search for B_s^0 oscillations using semileptonic $B_s \rightarrow D_s \mu X$ ($D_s \rightarrow K_s^0 K$). The data were collected using the DØ detector from events produced in $\sqrt{s} = 1.96$ TeV proton-antiproton collisions at the Fermilab Tevatron. The Tevatron is currently the only place in the world that produces B_s^0 mesons and will be until early 2008 when the Large Hadron Collider begins operating at CERN.

One of the vital ingredients for the search for B_s^0 oscillations is the determination of the flavor of the B_s^0 candidate (B_s^0 or \bar{B}_s^0) at the time of its production, called *initial state flavor tagging*. We develop an likelihood based initial state flavor tagger that uses objects on the side of the event opposite to the reconstructed B meson candidate. To improve the performance of this flavor tagger, we have made it multidimensional so that it takes correlations between discriminants into account. This tagging is then certified by applying it to sample of semimuonic $B^{(0,+)}$ decays and measuring the well-known oscillation frequency Δm_d . We obtain $\Delta m_d = 0.486 \pm 0.021$ ps⁻¹, consistent with the world average. The tagging performance is characterized by the effective efficiency, $\epsilon \mathcal{D}^2 = (1.90 \pm 0.41)\%$. We then turn to the search for B_s^0 oscillations in the above-named channel. A special two-dimensional mass fitting procedure is developed to separate kinematic reflections from signal events. Using this mass fitting procedure in an unbinned likelihood

framework, we obtain a 95% C.L. of $\Delta m_s > 1.10 \text{ ps}^{-1}$ and a sensitivity of 1.92 ps^{-1} . This result is combined with other analyzed B_s^0 decay channels at DØ to obtain a combined 95% C.L. of $\Delta m_s > 14.9 \text{ ps}^{-1}$ and a sensitivity of 16.5 ps^{-1} . The corresponding log likelihood scan has a preferred value of $\Delta m_s = 19 \text{ ps}^{-1}$ with a 90% confidence level interval of $17 < \Delta m_s < 21 \text{ ps}^{-1}$, assuming Gaussian uncertainties. A comparison of the change in the likelihood between $\Delta m_s = 19 \text{ ps}^{-1}$ and $\Delta m_s = \infty$ yields an 8% expectation for a background fluctuation.

Contents

List of Tables	x
List of Figures	xii
1 Introduction	1
1.1 History and Status of Mixing	2
1.2 Outline	3
2 Theoretical Overview	4
2.1 Standard Electroweak Theory	5
2.1.1 The Higgs Mechanism and Mass Generation	8
2.1.2 Lepton Masses	10
2.1.3 Quark Masses	10
2.2 <i>CP</i> Violation and the CKM Matrix	11
2.2.1 Parametrization of the CKM Matrix	13
2.3 Mixing Formalism	16
2.4 Mixing in the Standard Model	20
3 Experimental Overview	24
3.1 Vertices and Impact Parameters	25
3.1.1 Track Impact Parameters	25
3.1.2 Vertexing	26
3.2 Proper Decay Time	28
3.3 Flavor Tagging	29
3.4 Statistical Significance	32
4 The Experimental Apparatus	37
4.1 The Tevatron Accelerator Complex	37

4.2	The DØ Detector	40
4.2.1	Coordinate System	42
4.2.2	Central Tracking System	43
4.2.3	Calorimeter	52
4.2.4	Muon System	57
4.2.5	Luminosity Monitor	64
4.3	The DØ Trigger System	68
4.3.1	Level 1 Trigger	70
4.3.2	Level 2 Trigger	72
4.3.3	Level 3 Trigger and Data Acquisition	76
5	Multidimensional Opposite-Side Flavor Tagging	79
5.1	Likelihood Based Flavor Tagging	80
5.2	Multidimensional Likelihoods	81
5.3	Monte Carlo Samples and Selections	82
5.4	Discriminating Variables	85
5.5	Likelihoods, Binnings, and Tagging Logic	88
5.6	Binned Likelihood Reweighting	93
5.6.1	Adaptive Binning	93
5.6.2	Poisson Reweighting	94
5.7	A Measurement of Δm_d	95
5.7.1	Data Sample and Event Selection for Δm_d Measurement	95
5.7.2	Fitting Procedure for Δm_d Measurement	97
5.7.3	Mass Fit	98
5.7.4	Expected Flavor Asymmetry	102
5.7.5	Sample Composition	105
5.7.6	Results	109
6	B_s^0 Oscillation Analysis	124
6.1	Data Sample and Event Selections	125
6.1.1	Likelihood Ratio Method and Selection	128
6.2	Mass Fitting Procedure	133
6.2.1	Description of Technique	134
6.2.2	Tests of the Technique	136
6.2.3	Application to Data	138
6.3	Initial State Flavor Tagging	139
6.3.1	Tagged Sample Yields	142
6.4	Unbinned Likelihood Fit Method	145

6.4.1	<i>pdf</i> for the μD_s Signal	147
6.4.2	<i>pdf</i> for μD^+ Components	154
6.4.3	<i>pdf</i> for the $\mu \Lambda_C$ Component	154
6.4.4	<i>pdf</i> for the Combinatorial Background	154
6.5	Inputs to the Fit	157
6.5.1	Sample Composition	157
6.5.2	K Factors	159
6.5.3	Reconstruction Efficiencies	163
6.6	Resolution Scale Factor	166
6.7	Results of the Lifetime Fit	170
6.8	Fitting Procedure for the Δm_s Limit	171
6.9	Results	176
6.10	Cross Checks and Systematic Errors	177
6.10.1	Dilution	177
6.10.2	Mass fitting procedure	179
6.10.3	Resolution scale factor	180
6.10.4	Sample Composition	181
6.10.5	K factors	181
6.10.6	Randomized flavor tagging	182
6.10.7	Variations of fit parameters	182
7	Discussion and Conclusions	185
7.1	Combination and Conclusions	185
7.2	Outlook	188
7.2.1	Increasing the Statistics and Signal to Background	189
7.2.2	Improving the Flavor Tagging	189
7.2.3	Improving the Proper Time Resolution	189
A	Luminosity Monitor Calibration	197
A.1	CAFÉ Calibration	198
A.2	Global Time Calibration	199
A.3	Loading Lookup Tables	200
	Bibliography	206

List of Tables

4.1	An overview of the Silicon Microstrip Detector (SMT).	48
4.2	Design parameters of the Central Fiber Tracker (CFT).	50
4.3	Central Calorimeter Parameters.	58
4.4	Muon drift tube parameters.	61
5.1	Correlation matrix for all tagging discriminants	89
5.2	Table showing the χ^2 of different fits to the entire $B^+ \rightarrow \mu^+ \nu \bar{D}^0 X$ sample.	100
5.3	The results of a free fit to $B^+ \rightarrow \mu^+ \nu \bar{D}^0 X$ samples	101
5.4	The results of a free fit to $B_d^0 \rightarrow \mu^+ \nu D^{*-} X$ samples	102
5.5	The measured number of D^* and D^0 events, the measured asymmetry A_i , and the expected asymmetry A_i^e for $0.2 < d < 0.35$	111
5.6	The measured number of D^* and D^0 events, the measured asymmetry A_i , and the expected asymmetry A_i^e for $0.35 < d < 0.45$	112
5.7	The measured number of D^* and D^0 events, the measured asymmetry A_i , and the expected asymmetry A_i^e for $0.45 < d < 0.6$	113
5.8	The measured number of D^* and D^0 events, the measured asymmetry A_i , and the expected asymmetry A_i^e for $ d > 0.6$	114
5.9	Tagging performance for events with a reconstructed B^0 for different subsamples	115
5.10	Tagging performance for events with reconstructed B^+ for different subsamples	115
5.11	Measured value of Δm_d and $f_{c\bar{c}}$ for different subsamples.	115
6.1	Component yields and fractions for the untagged, tagged, and tagged with $ d > 0.3$ samples.	145

6.2	Sample composition calculated before the application of lifetime-biasing cuts.	159
6.3	Parameters determined from a lifetime fit to the full tagged sample.	175
6.4	Systematic uncertainties on the Δm_s oscillation amplitude. . .	184
A.1	CAFÉ calibration constants.	199
A.2	Luminosity monitor global time calibration constants.	201
A.3	Lookup table address bitmap	201
A.4	Arrival time lookup table data bitmap	202
A.5	Charge ranges for storing pulse height information in lookup tables on pulse height CAFÉ cards	203
A.6	Pulse height lookup table bitmap.	203
A.7	Global luminosity monitor calibration constants.	204

List of Figures

2.1	The unitarity triangle.	14
2.2	Probability to observe a mixed or unmixed event assuming perfect flavor tagging.	22
2.3	Lowest-order box diagrams responsible for B - \bar{B} mixing.	23
3.1	Definition of impact parameter components.	33
3.2	Schematic drawing showing the definition of transverse distance X_{xy} and transverse decay length L_{xy}	34
3.3	A diagram of a typical semileptonic B_s^0 event.	35
3.4	Probability to observe a mixed or unmixed event with the effect of flavor tagging taken into account.	36
4.1	The Tevatron accelerator complex.	38
4.2	A simple magnetron source	39
4.3	The Tevatron beam structure.	41
4.4	Schematic of the DØ detector.	42
4.5	Schematic drawing of the DØ tracking system.	44
4.6	A diagram of a generic silicon detector.	45
4.7	The disk/barrel design of the silicon microstrip tracker.	46
4.8	Cross section of an silicon microstrip barrel module.	47
4.9	A view of the top half of the tracking system in the $y - z$ plane.	48
4.10	An end-on view of a central fiber tracker ribbon.	49
4.11	Cross section and layout geometry of the preshowers.	52
4.12	Fractional energy loss per radiation length in lead as a function of electron or positron energy.	53
4.13	Diagram of an electromagnetic shower.	54
4.14	Isometric view of the central and two end calorimeters.	56
4.15	Schematic view of a unit cell for the calorimeter	57

4.16	Schematic view of $\frac{1}{4}$ of the calorimeter showing the segmentation pattern.	59
4.17	Average muon energy loss in various materials.	60
4.18	Exploded view of the muon scintillation detectors.	62
4.19	Exploded view of the muon wire chambers.	63
4.20	Cross-sectional view of proportional drift tube cells.	64
4.21	Cross-sectional of an mini drift tube.	65
4.22	Schematic drawing showing the location of the luminosity monitor.	66
4.23	Schematic drawing showing the geometry of the luminosity monitor counters.	67
4.24	Overview of the $D\bar{O}$ data acquisition system.	68
4.25	Block diagram of the $D\bar{O}$ Level 1 and Level 2 trigger systems.	70
4.26	Transverse schematic view of a single 4.5° Level 1 central track trigger input sector.	72
4.27	Schematic drawing showing the conceptual design of the Level 2 silicon track trigger.	74
4.28	Schematic illustration of the data flow through the Level 3 data acquisition system.	78
5.1	qp_T and impact parameter distributions of opposite-side muons from Monte Carlo simulations.	83
5.2	Dilution of events tagged by Q_{SV} versus the coefficient κ	88
5.3	Normalized distributions of the tagging variable for 3 taggers.	92
5.4	The $M(K\pi\pi) - M(K\pi)$ invariant mass for selected μD^* candidates	97
5.5	The $K\pi$ invariant mass of the μD^0 and μD^* samples.	98
5.6	The tagged $M(K\pi)$ distribution for events in the D^* sample with $ d > 0.37$	103
5.7	The fit to $M(K\pi)$ for tagged events with $ d > 0.37$ in the first 2 VPDL bins.	110
5.8	The fit to $M(K\pi)$ for tagged events with $ d > 0.37$ in the first 3^{rd} and 4^{th} VPDL bins.	118
5.9	The fit to $M(K\pi)$ for tagged events with $ d > 0.37$ in the last 3 VPDL bins.	119
5.10	Asymmetries for the $ d > 0.37$ sample	120
5.11	Asymmetries for the $0.2 < d < 0.35$ and $0.35 < d < 0.45$ samples	121

5.12	Asymmetries for the $0.45 < d < 0.60$ and $ d > 0.60$ samples .	122
5.13	Multidimensional flavor tagger calibration.	123
6.1	Invariant mass of the K_S^0 candidates, $m(\pi^+\pi^-)$	127
6.2	Distributions of $m(K_S^0K)$ for right-sign ($q_\mu \cdot q_K < 0$) and wrong-sign ($q_\mu \cdot q_K > 0$) candidates after the pre-selections. . .	130
6.3	$B_s^0 \rightarrow D_s^- \mu^+ X$, $D_s^- \rightarrow K_S^0 K^-$ signal and background for the discriminants used in the likelihood ratio selections.	131
6.4	Signal significance versus likelihood ratio selection variable, $\log_{10} y$	132
6.5	The distribution of the combined selection variable, $-\log_{10} y$ for signal and background.	133
6.6	Mass spectrum for $m(K_S^0K)$ after all selections, including the likelihood ratio selection.. . . .	134
6.7	Mass spectrum for $m(K_S^0K)$ before and after likelihood ratio selections.	135
6.8	The fit of $D^+ \rightarrow K_S^0 \pi$ Monte Carlo to Eqn. 6.8.	137
6.9	The fit of $\Lambda_c \rightarrow K_S^0 p$ Monte Carlo to Eqn. 6.10.	138
6.10	An example experiment from an ensemble test of the asymmetry mass fitting technique.	139
6.11	The pull of the signal yield, $(N(D_s)^{fitted} - N(D_s)^{true}) / \sigma(D_s)$	140
6.12	A fit to the entire untagged data sample.	141
6.13	Mass fit to the entire tagged sample.	143
6.14	Mass fit to the tagged sample having $ d > 0.3$	144
6.15	Fitted values of the 3 rd order Chebyshev polynomial background in 11 bins of VPDL.	146
6.16	The distribution of VPDL errors for signal and background.	148
6.17	The distribution of predicted dilution for signal and background.	149
6.18	Dilution calibration, taken from DØ Note 4991.	151
6.19	K factors for $B_s^0 \rightarrow D_s^- \mu^+ X$ and $B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu D_s^-$ processes.	160
6.20	K factors for $B_s^0 \rightarrow D_s^- \mu^+ X$ divided into 4 bins of $m(D_s \mu)$	161
6.21	K factors for $B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu D_s^-$ divided into 4 bins of $m(D_s \mu)$	162
6.22	K factor for $B \rightarrow \mu \nu D^+ \rightarrow K_S^0 \pi$	163
6.23	K factor for $B \rightarrow \mu \nu D^+ \rightarrow K_S^0 \pi \pi^0$	164
6.24	K factor for $B^0 \rightarrow D_s D$	164

6.25	K factor for $B^- \rightarrow D_s D$.	164
6.26	K factor for $B^0 \rightarrow D_s D_s$.	164
6.27	K factor for $B_s \rightarrow D_s D$.	164
6.28	K factor for $\Lambda_b \rightarrow \mu\nu\Lambda_c \rightarrow K_S^0 p$.	164
6.29	Efficiency as a function of VPDL for $B_s^0 \rightarrow D_s^- \mu^+ X$.	165
6.30	Efficiency as a function of VPDL for $B \rightarrow \mu\nu D^+ \rightarrow K_S^0 \pi$.	166
6.31	Efficiency as a function of VPDL for $B^0 \rightarrow D_s D$.	166
6.32	Efficiency as a function of VPDL for $B^- \rightarrow D_s D$.	166
6.33	Efficiency as a function of VPDL for $B_s \rightarrow D_s D_s$.	166
6.34	Efficiency as a function of VPDL for $B_s \rightarrow D_s D$.	167
6.35	Efficiency as a function of VPDL for $\Lambda_b \rightarrow \mu\nu\Lambda_c \rightarrow K_S^0 p$.	167
6.36	Distributions of $\Delta M = m(D^*) - m(D^0)$ for $D^{*+} \rightarrow D^0 \pi^+$, $D^0 \rightarrow K_S^0 \pi^- \mu^+ X$ events.	169
6.37	Double Gaussian fits to the negative part of background sub- tracted $PDL(D^*)/\sigma_{PDL}(D^*)$ distributions.	170
6.38	Widths and fractions of double Gaussian fits to pull distribu- tions versus $l_{xy}(K_S^0)$.	171
6.39	Lifetime fit projected on the entire mass range, $1.4 < m(K_S^0 K) <$ $2.4 \text{ GeV}/c^2$.	172
6.40	Lifetime fit projected on the 3σ signal region mass range.	173
6.41	Lifetime fit projected on the right and left sidebands	174
6.42	B_s^0 oscillation amplitude with statistical and systematic errors.	176
6.43	The $B_d^0 - \bar{B}_d^0$ oscillation amplitude scanning the $D^+ \rightarrow K_S^0 \pi$ component.	178
6.44	The $B_d^0 - \bar{B}_d^0$ oscillation amplitude scanning the long-lived os- cillating background component.	179
6.45	The $B_d^0 - \bar{B}_d^0$ oscillation amplitude scanning all three Δm_d com- ponents.	180
6.46	B_s^0 oscillation amplitude with randomized flavor tagging.	182
7.1	Combined $D\bar{O}$ B_s^0 oscillation amplitude.	191
7.2	Log likelihood scan for the combination of all $D\bar{O}$ decay modes.	192
7.3	Experimental constraints on the unitarity triangle before the inclusion of the combined $D\bar{O}$ Δm_s result.	193
7.4	Experimental constraints on the unitarity triangle after the inclusion of the combined $D\bar{O}$ Δm_s result.	194
7.5	Experimental constraints on the unitarity triangle after the inclusion of the CDF observation of B_s^0 mixing.	195

7.6 Model-independent constraints on new phenomena. 196

A.1 Schematic drawing of the operation of lookup tables on TDC
CAFÉ cards. 205

Chapter 1

Introduction

The Standard Model (SM) of particle physics is a mathematical theory that describes the most basic constituents of matter and their interactions. Experimental results have thus far agreed with SM expectations to a high degree of precision across a wide range of different observables [1]. The SM aims to describe the electromagnetic, strong nuclear, and weak nuclear interactions. Part of the description of weak interactions involves an explanation for processes that change the flavor of the quarks involved, called *flavor-changing charged currents*. A crucial element of the formulation of these currents is the Cabbibo-Kobayashi-Maskawa (CKM) matrix, which has four free parameters.

Measuring the frequency of particle-antiparticle oscillations in the $B - \bar{B}$ system, where B can be either B^0 or B_s^0 , allows us to determine some of these free CKM parameters. In the SM, the CKM matrix is supposed to be unitary; we can verify this property by making many independent measurements of its free parameters. If experimental results do not support the CKM unitarity hypothesis, this would indicate new physics beyond the SM. Below we give a history of particle-antiparticle oscillations as well as an outline of the research presented in this dissertation.

1.1 History and Status of Mixing

Neutral particle-antiparticle oscillations were first predicted in 1955 in a paper by Gell-Mann and Pais [2]. A consequence of studies of these oscillations was the discovery of the long-lived neutral kaon, the K_L^0 , by Lederman and his collaborators a year later at Brookhaven [3]. In 1963, when the only known quarks were the *up*, *down*, and *strange* quarks, Cabbibo introduced a mixing angle between *up* and *strange* quarks to account for the measured rates of strange particle decays [4]. This led to the proposal of a fourth quark, the *charm* quark, in 1970 by Glashow, Iliopoulos, and Maini [5] to explain the rate of the weak decay $K_L^0 \rightarrow \mu^+\mu^-$. In 1973 Kobayashi and Maskawa, motivated by the 1964 discovery of CP violation in neutral kaons by Fitch and Cronin [6], proposed a third generation of quarks [7]. The *bottom* and *top* quarks postulated by Kobayashi and Maskawa were discovered in 1977 [8] and 1995 [9], respectively, at Fermilab.

The first evidence for mixing in the B system was from UA1 in 1987 based on a time-independent analysis of the ratio of like-sign to unlike-sign dimuon pairs [10]. Later that year, ARGUS made the first observation of mixing in the B^0 system using an unambiguous identification of $B^0\bar{B}^0$ pairs [11]. This observation was confirmed by CLEO in 1989 [12]. Time-dependent measurements have since then yielded a precise determination of the oscillation frequency of the B^0 system, Δm_d . The current world average, $\Delta m_d = 0.507 \pm 0.005 \text{ ps}^{-1}$ [13], is dominated by measurements from the B -factories [14].

Within a few years of the observation B_d^0 mixing, it was clear that the B_s^0 meson also oscillates. The time-integrated ARGUS and CLEO mixing measurements mentioned above were done at e^+e^- colliders operating at the $\Upsilon(4S)$ where the dominant production is $B^0\bar{B}^0$ pairs. Subsequent measurements of the time-integrated mixing parameter at the LEP experiments, which operated on the Z -pole and therefore were able to access B_s^0 production, led to limits on the oscillation frequency Δm_s [15].

DØ has recently measured the world's first two-sided bound on Δm_s of $17.0 < \Delta m_s < 21.0 \text{ ps}^{-1}$ at the 90% C.L with a most probable value of $\Delta m_s = 19 \text{ ps}^{-1}$ [16]. CDF later measured $\Delta m_s = 17.31_{-0.18}^{+0.33} \text{ (stat.)} \pm 0.07 \text{ (sys) ps}^{-1}$ [17].

1.2 Outline

The search for B_s^0 oscillations in a channel not included in Ref. [16], but to be included in future global oscillation analyses, constitutes the main analysis of this dissertation. A vital component of this search, called *flavor tagging*, is the determination of the b quark flavor of the reconstructed B_s^0 meson at the time of its production i.e., whether it is a B_s^0 or \bar{B}_s^0 . To calibrate and validate the flavor tagger, we first perform a measurement of Δm_d in samples of B_d^0 and B^+ . We then apply this flavor tagging to the B_s^0 analysis and use an unbinned likelihood framework to search for Δm_s .

Chapter 2 provides an overview of the theoretical motivation for the search for Δm_s . Chapter 3 gives an overview of the experimental strategy and issues common to both the Δm_d and Δm_s measurement. A description of the experimental apparatus used for these measurements, the Fermilab Tevatron and DØ detector, is given in Ch. 4. Chapter 5 explains the principles of a flavor tagger and presents the measurement of Δm_d , while the search for Δm_s is described in Ch. 6 and limits on Δm_s are given.

Chapter 2

Theoretical Overview

The fundamental constituents of matter in the Standard Model [18] of particle physics are quarks and leptons. Quarks participate in strong interaction while leptons do not. These spin-1/2 building blocks can be arranged in three generations:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad \begin{pmatrix} t \\ b \end{pmatrix}; \quad (2.1)$$

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}, \quad \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \quad \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}; \quad (2.2)$$

where each successive generation is more massive than the previous. Forces in the SM are mediated by integral-spin gauge bosons – the photon for the electromagnetic interaction, the W^\pm and Z^0 for the weak interaction, and an octet of gluons for the strong interaction. There is one particle remaining in the SM that we have not yet mentioned, a scalar boson called the *Higgs boson* that is responsible for the generation of the masses of the other fundamental particles. We are interested in SM processes that lead to $B - \bar{B}$ mixing. We next turn our attention to the electroweak sector of the Standard Model.

2.1 Standard Electroweak Theory

Muon and charged pion lifetimes are considerably longer than those of particles that decay electromagnetically or through the strong force. For example [13],

$$\begin{aligned}\tau(\pi^+ \rightarrow \mu^+ \nu_\mu) &= (2.6033 \pm 0.0005) \times 10^{-8} \text{ s}, \\ \tau(\mu^+ \rightarrow e^+ \bar{\nu}_e \nu_\mu) &= (2.19703 \pm 0.00004) \times 10^{-6} \text{ s}.\end{aligned}\quad (2.3)$$

The typical lifetime of a particle decaying electromagnetically or by the strong force is much shorter, $\tau(\text{EM}) \approx 10^{-16}$ s and $\tau(\text{strong}) \approx 10^{-23}$ s. Because lifetimes are inversely proportional to the coupling strength of the force responsible for decay, we can deduce that there is an interaction even weaker than electromagnetism that is responsible for the lifetimes in Eq. 2.3. Further evidence for this interaction comes from nuclear β -decays, $n \rightarrow pe^-\bar{\nu}_e$, which are responsible for the instability of the neutron. Again, we deduce that the interaction must be quite weak because the neutron lifetime is so long ($\tau = 886$ s) [13].

We begin the formulation of the electroweak interaction by defining the weak hypercharge in accordance with the Gell-Mann–Nishijima formula,

$$Q = I^3 + \frac{1}{2}Y, \quad (2.4)$$

where Q is the electric charge, I^3 is the third component of isospin, and Y is the weak hypercharge. The gauge group used for the electroweak interaction is $SU(2) \otimes U(1)_Y$, where the subscript Y on $U(1)$ denotes that the group is a symmetry of weak hypercharge. We next construct the fermion fields that participate in the electroweak interactions.

All experimental evidence supports the hypothesis that only right-handed antineutrinos and left-handed neutrinos participate in weak interactions. Consequently, we formulate the electroweak theory so that it treats right-

and left-handed fermions differently. This is accomplished by defining chiral fermion fields through projection operators,

$$\left. \begin{aligned} \psi_L(x) &= P_L \psi(x) \\ \psi_R(x) &= P_R \psi(x) \end{aligned} \right\} \equiv \frac{1}{2} (1 \mp \gamma^5) \psi(x), \quad (2.5)$$

where $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$. The Dirac γ -matrices are defined as $\gamma^\mu \equiv (\beta, \vec{\alpha})$ where the $\vec{\alpha}$ and β matrices are (in the Dirac-Pauli representation):

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad (2.6)$$

with I denoting a unit 2×2 matrix and $\vec{\sigma}$ being the Pauli spin matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.7)$$

We write the right-handed fields as:

$$\begin{aligned} E_R &= (e_R, \mu_R, \tau_R), & Y_e &= -2; \\ U_R &= (u_R, c_R, t_R), & Y_u &= \frac{4}{3}; \\ D_R &= (d_R, s_R, b_R), & Y_d &= -\frac{2}{3}; \end{aligned} \quad (2.8)$$

with the hypercharge Y given. The $SU(2)$ doublets for the left-handed fermions are constructed as follows:

$$\begin{aligned} E_L &= \left(\left(\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \right), & Y_L &= -1; \\ Q_L &= \left(\left(\begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \right), & Y_Q &= \frac{1}{3}. \end{aligned} \quad (2.9)$$

In the above equation, E_L is the upper isospin component ($Y(E_L) = +1/2$) and Q_L is the lower isospin component ($Y(Q_L) = -1/2$). The Lagrangian

density describing the propagations of fermions and the interaction between fermions and gauge bosons is then [19],

$$\begin{aligned} \mathcal{L} = & \bar{E}_L(i\cancel{D})E_L + \bar{E}_R(i\cancel{D})E_R + \bar{Q}_L(i\cancel{D}_{QCD})Q_L + \bar{U}_R(i\cancel{D}_{QCD})U_R \\ & + \bar{D}_R(i\cancel{D}_{QCD})D_R + g(W_\mu^+ J_W^{\mu+} + W_\mu^- J_W^{\mu-} + Z_\mu^0 J_Z^\mu) \\ & + eA_\mu J_{EM}^\mu, \end{aligned} \quad (2.10)$$

where a sum over the 3 generations of fermions is implied, D_{QCD}^μ is the $SU(3)$ covariant derivative, and $\cancel{D} \equiv \gamma^\mu \partial_\mu$. The currents in Eq. 2.10 are:

$$\begin{aligned} J_W^{\mu+} &= \frac{1}{\sqrt{2}} (\bar{\nu}_L \gamma^\mu e_L + \bar{u}_L \gamma^\mu d_L), \\ J_W^{\mu-} &= \frac{1}{\sqrt{2}} (\bar{e}_L \gamma^\mu \nu_L + \bar{d}_L \gamma^\mu u_L), \\ J_Z^\mu &= \left[\bar{\nu}_L \gamma^\mu \left(\frac{1}{2} \right) \nu_L + \bar{e}_L \gamma^\mu \left(-\frac{1}{2} + \sin^2 \theta_W \right) e_L + \bar{e}_R \gamma^\mu (\sin^2 \theta_W) e_R \right. \\ &\quad + \bar{u}_L \gamma^\mu \left(\frac{1}{2} - \frac{2}{3} \sin^2 \theta_W \right) u_L + \bar{u}_R \gamma^\mu \left(-\frac{2}{3} \sin^2 \theta_W \right) u_R \\ &\quad \left. + \bar{d}_L \gamma^\mu \left(-\frac{1}{2} + \frac{1}{3} \sin^2 \theta_W \right) d_L + \bar{d}_R \gamma^\mu \left(\frac{1}{3} \sin^2 \theta_W \right) d_R \right], \\ J_{EM}^\mu &= \bar{e} \gamma^\mu (-1) e + \bar{u} \gamma^\mu \left(+\frac{2}{3} \right) u + \bar{d} \gamma^\mu \left(-\frac{1}{3} \right) d. \end{aligned} \quad (2.11)$$

The parameter θ_W in Eq. 2.11 is the *Weingberg angle* and is defined by $g \sin \theta_W = e$ where g is the weak charged current coupling constant and e is the charge of the electron.

The theory as described by Eqs. 2.10 and 2.11 is incomplete because every particle is massless. If we try, for example, to add electron mass terms like,

$$\mathcal{L}_m = -m (\bar{e}_L e_R + \bar{e}_R e_L), \quad (2.12)$$

we would break the $SU(2) \otimes U(1)_Y$ invariance of the Lagrangian density because the left- and right-handed fields belong to different $SU(2)$ represen-

tations and $U(1)_Y$ hypercharges. We have to introduce another field to this Lagrangian to generate particle masses.

2.1.1 The Higgs Mechanism and Mass Generation

Consider a $SU(2)$ doublet of complex scalar fields,

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \quad (2.13)$$

having the potential,

$$V(|\Phi|^2) = -\mu^2 |\Phi|^2 + \lambda |\Phi|^4, \quad (2.14)$$

with $|\Phi|^2 = \phi^\dagger \phi$ and two new parameters, λ and μ . If $\mu^2 > 0$, then we can choose $V(|\Phi|^2)$ to have a minimum at

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (2.15)$$

with $v = \mu/\sqrt{\lambda}$. Note that this minimum does not respect $SU(2) \otimes U(1)_Y$ gauge symmetry even though the scalar doublet and its interactions do. This phenomenon is called *spontaneous symmetry breaking*.

We use the unitary gauge [20] to write the field $\Phi(x)$ as:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (2.16)$$

The particle associated with the field $h(x)$ is called the *Higgs boson*. It has the Lagrangian density,

$$\mathcal{L}_H = |D_\mu \Phi|^2 + \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2, \quad (2.17)$$

where the covariant derivative is from gauging $SU(2) \otimes U(1)_Y$,

$$D^\mu \Phi = \left(\partial_\mu - \frac{i}{2} g A_\mu^a \sigma^a - \frac{i}{2} g' B_\mu \right) \Phi, \quad (2.18)$$

with σ^a defined in Eq. 2.6. The potential energy terms in Eq. 2.17 are then:

$$\begin{aligned} \mathcal{L}_V &= -\mu^2 h^2 - \lambda v h^3 - \frac{1}{4} \lambda h^4 \\ &= -\frac{1}{2} m_h^2 h^2 - \sqrt{\frac{\lambda}{2}} m_h h^3 - \frac{1}{4} \lambda h^4. \end{aligned} \quad (2.19)$$

We can see that the quantum of the Higgs field $h(x)$ is a scalar particle with mass

$$m_h = \sqrt{2} \mu^2. \quad (2.20)$$

If we explicitly put the covariant derivative of Eq. 2.18 into Eq. 2.17, we can write the kinetic terms as,

$$\mathcal{L}_K = \frac{1}{2} (\partial_\mu h)^2 + \left(1 + \frac{h}{v} \right)^2 \left[m_W^2 W^{\mu+} W_\mu^- + \frac{1}{2} m_Z^2 Z^\mu Z_\mu \right], \quad (2.21)$$

with $m_W = g \frac{v}{2}$ and $m_Z = m_W / \cos \theta_W$. Spontaneous symmetry breaking has therefore made the formerly massless gauge vector bosons massive without violating $SU(2) \otimes U(1)_Y$ invariance.

Fermions couple to the Higgs field through Yukawa interactions¹[21],

$$\mathcal{L}_Y(x) = - \sum_{i,j=1}^G \left[\hat{y}_{ij}^e \bar{E}_L^i \Phi E_R^j + \hat{y}_{ij}^d \bar{Q}_L^i \Phi D_R^j + \hat{y}_{ij}^u \bar{Q}_L^i \tilde{\Phi} U_R^j + \text{h.c.} \right], \quad (2.22)$$

where we have considered G generations of fermions, h.c. stands for hermitian

¹A coupling of the form $\bar{\psi}(x)\phi(x)\psi(x)$ where $\psi(x)$ is a spinor and $\phi(x)$ is a scalar is called a Yukawa interaction.

conjugate, and

$$\tilde{\Phi} \equiv i\sigma_2\Phi^* = \begin{pmatrix} \bar{\phi}^0 \\ -\phi^- \end{pmatrix}. \quad (2.23)$$

We will consider the Yukawa couplings of the lepton and quark fields separately.

2.1.2 Lepton Masses

The non-Yukawa part of the Lagrangian density is invariant under the transformations,

$$\begin{aligned} E_R &\Rightarrow RE_R, & \bar{E}_R &\Rightarrow \bar{E}_R R^\dagger \\ E_L &\Rightarrow SE_L, & \bar{E}_L &\Rightarrow \bar{E}_L S^\dagger, \end{aligned} \quad (2.24)$$

where $R \in U(G)_{E_R}$ and $S \in U(G)_{E_L}$. The Yukawa matrix \hat{y}^e is therefore equivalent to $y^e = S\hat{y}^e R^\dagger$. We can choose to make y^e diagonal, real, and non-negative. This has implications for the lack of CP violation in the lepton sector which we will return to in Sec. 2.2.

2.1.3 Quark Masses

The Higgs coupling to quarks is similar to its interaction with leptons, although slightly more complicated because quarks have three $U(G)$ symmetries instead of the two that the leptons have. The non-Yukawa quark Lagrangian density is invariant under

$$\begin{aligned} D_R &\Rightarrow R_d D_R, & \bar{D}_R &\Rightarrow \bar{D}_R R_d^\dagger, \\ U_R &\Rightarrow R_u U_R, & \bar{U}_R &\Rightarrow \bar{U}_R R_u^\dagger, \end{aligned} \quad (2.25)$$

$$Q_L \Rightarrow S_u Q_L, \quad \bar{Q}_L \Rightarrow \bar{Q}_L S_u^\dagger. \quad (2.26)$$

We transform the Yukawa matrices \hat{y}^d as,

$$S_u \hat{y}^d R_d^\dagger = S_u S_d^\dagger S_d \hat{y}^d R_d^\dagger = V y^d, \quad (2.27)$$

where $y^d = S_d \hat{y}^d R_d^\dagger$ is diagonal, real, and non-negative and

$$V = S_u S_d^\dagger. \quad (2.28)$$

is the Cabibbo-Kobayashi-Maskawa (CKM) matrix.

2.2 CP Violation and the CKM Matrix

Before examining the CKM matrix in detail, we first look at general properties of CP violation. Consider a weak process $ab \rightarrow cd$ with the invariant amplitude²

$$\begin{aligned} \mathcal{M} &\sim J_{ca}^\mu J_{\mu bd}^\dagger \\ &\sim (\bar{\psi}_c \gamma^\mu (1 - \gamma^5) V_{ca} \psi_a) (\bar{\psi}_b \gamma_\mu (1 - \gamma^5) V_{bd} \psi_d)^\dagger \\ &\sim V_{ca} V_{db}^* (\bar{\psi}_c \gamma^\mu (1 - \gamma^5) \psi_a) (\bar{\psi}_d \gamma_\mu (1 - \gamma^5) \psi_b), \end{aligned} \quad (2.29)$$

where V is the CKM matrix. The Hermitian conjugate is,

$$\mathcal{M}' \sim V_{ca}^* V_{db} (\bar{\psi}_a \gamma^\mu (1 - \gamma^5) \psi_c) (\bar{\psi}_b \gamma_\mu (1 - \gamma^5) \psi_d). \quad (2.30)$$

Because the full Hamiltonian is Hermitian and therefore must contain $\mathcal{M} + \mathcal{M}^\dagger$, to verify whether a theory is CP invariant, all we have to do is calculate \mathcal{M}_{CP} from Eq. 2.29 and check

$$\mathcal{M}_{CP} \stackrel{?}{=} \mathcal{M}^\dagger. \quad (2.31)$$

²The invariant amplitude \mathcal{M} is defined so that the differential cross section is $d\sigma = \frac{|\mathcal{M}|^2}{F} dQ$, where the flux is F and dQ is the Lorentz invariant phase space factor.

If they are equal then the theory is *CP* invariant and if they are not then it is *CP* violating.

We apply the *CP* operator to the weak current,

$$(J_{ca}^\mu)_{CP} = -V_{ca}\bar{\psi}_a\gamma^{\mu\dagger}(1 - \gamma^5)\psi_c, \quad (2.32)$$

and obtain

$$\mathcal{M}_{CP} \sim V_{ca}V_{db}^* [\bar{\psi}_a\gamma^\mu(1 - \gamma^5)\psi_c] [\bar{\psi}_b\gamma_\mu(1 - \gamma^5)\psi_d]. \quad (2.33)$$

If the CKM matrix V is real then the theory is *CP* invariant.

The CKM matrix as defined in Eq. 2.28 is a $G \times G$ unitary matrix. Such a matrix has a total of G^2 parameters: $\frac{1}{2}G(G-1)$ real and $\frac{1}{2}G(G+1)$ imaginary. We can rotate the phases of the $2G$ quark states without changing the physics so that V contains

$$G^2 - (2G - 1) - \frac{1}{2}G(G - 1) = \frac{1}{2}(G - 1)(G - 2) \quad (2.34)$$

imaginary parameters, where one phase is omitted as an overall phase change. For $G = 2$, the matrix will contain 1 real and 0 imaginary parameters; if $G = 3$ there will be 3 real and 1 imaginary parameters and the CKM matrix will therefore accommodate *CP* violation. In fact, it was the appearance of a *CP*-violating phase that led Kobayashi and Maskawa to introduce a third generation of quarks to the mixing matrix so their theory would accommodate *CP* violation.

It is important to note that *CP* may still be conserved in a 3×3 CKM matrix if the masses of two of the quarks of equal charges are the same. To see this more clearly, we follow the method of Jarlskog [22] and construct the commutator of the mass matrices,

$$C = S_u^\dagger \left[(m^u)^2, V (m^d)^2 V^\dagger \right] S_u, \quad (2.35)$$

the determinant of which gauges the size of CP violation. This determinant is

$$\det C = -2iF_u F_d J, \quad (2.36)$$

where

$$F_u = (m_u^2 - m_c^2)(m_c^2 - m_t^2)(m_t^2 - m_u^2), \quad (2.37)$$

$$F_d = (m_d^2 - m_s^2)(m_s^2 - m_b^2)(m_b^2 - m_d^2), \quad (2.38)$$

$$J = \text{Im}[V_{11}V_{21}^*V_{22}V_{12}^*]. \quad (2.39)$$

We can see from $\det C$ that F_u , F_d , and J must all be different from zero for there to be CP violation from the CKM mechanism.

2.2.1 Parametrization of the CKM Matrix

The CKM matrix is often written with the indices labelled by the quark flavors as:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \quad (2.40)$$

to emphasize associated flavor transitions, e.g. $b \rightarrow c$ involves the matrix element $|V_{bc}|^2$. If we apply the unitarity constraint on the first and third columns of V , we get

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \quad (2.41)$$

Equation 2.41 defines a triangle in the complex plane which is called the “unitarity triangle”, shown in Fig. 2.1. The sides have length $|V_{ud}V_{ub}^*|$, $|V_{td}V_{tb}^*|$,

and $|V_{cd}V_{cb}^*|$; the angles are,

$$\begin{aligned}\alpha &= \arg \left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right], \\ \beta &= \arg \left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right], \\ \gamma &= \arg \left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right].\end{aligned}\tag{2.42}$$

The triangle's area is a measure of the CP -violation of the theory because it is equal to half the Jarlskog invariant, $A = |J|/2$. On the right of Fig. 2.1, we show a rescaled version of the triangle corresponding to dividing Eq. 2.41 by $|V_{cd}V_{cb}^*|$. Two new parameters, $\bar{\rho}$ and $\bar{\eta}$, are defined in this triangle by,

$$\bar{\rho} + i\bar{\eta} \equiv -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}.\tag{2.43}$$

A global fit using all available data gives $\bar{\rho} = 0.221_{-0.028}^{+0.064}$ and $\bar{\eta} = 0.340_{-0.045}^{+0.017}$ [23].

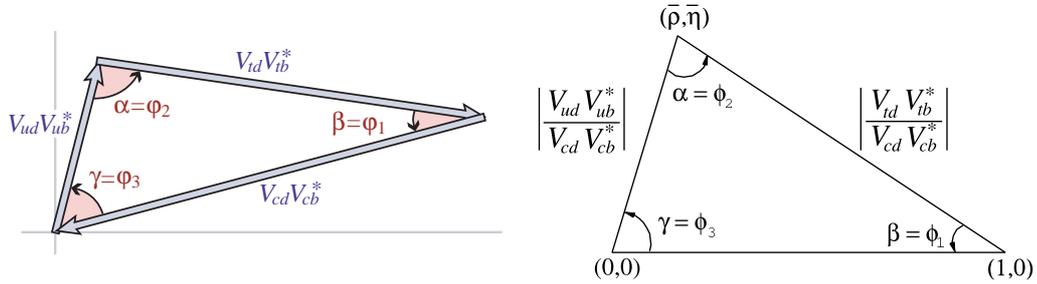


Figure 2.1: The unitarity triangle. The plot on the left corresponds to Eq. 2.41. The plot on the right shows the definition of $(\bar{\rho}, \bar{\eta})$ [13].

The unitarity triangle is a very useful visualization of the CKM mechanism. Separate measurements of the sides and angles of the triangle should be compatible. If we find that the triangle does not “close”, that would indicate that there are new processes not accounted for in the SM or that the CKM matrix has more dimensions than we thought. The unitarity triangle thus provides a summary of our knowledge of the CKM mechanism and CP

violation in the Standard Model.

As we have seen in Sec. 2.2, the matrix has three real and one imaginary parameter. Equation 2.40, however, does not make this clear. We now discuss two parameterizations of the CKM matrix that make the number of free parameters explicit. The first is called the ‘‘Chau-Keung’’ parameterization [24] and is as follows:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (2.44)$$

where $s_{ij} = \sin \theta_{ij}$, $c_{ij} = \cos \theta_{ij}$, and δ is a single complex CP -violating phase. The Jarlskog invariant, Eq. 2.39, becomes

$$J = c_{12}c_{23}c_{13}^2 s_{12}s_{23}s_{13} \sin \delta \quad (2.45)$$

in this parameterization. Since J is proportional to the level of CP -violation, we must have,

$$\delta \neq 0, \pi; \quad \theta_{ij} \neq 0, \pi/2; \quad (2.46)$$

for CP -violating effects to arise through a three-dimensional CKM matrix.

A second parametrization, the ‘‘Wolfenstein’’ parameterization [25], is based on the observation that diagonal elements of the CKM matrix are close to 1 and off-diagonal elements are progressively smaller. We can see this clearly if we look at the results of the global fit to the CKM matrix [13],

$$\begin{pmatrix} 0.97384_{-0.00023}^{+0.00024} & 0.2272 \pm 0.0010 & (3.96 \pm 0.09) \times 10^{-3} \\ 0.2271 \pm 0.0010 & 0.97296 \pm 0.00024 & (42.21_{-0.80}^{+0.10}) \times 10^{-3} \\ (8.14_{-0.64}^{+0.32}) \times 10^{-3} & (41.61_{-0.78}^{+0.12}) \times 10^{-3} & 0.999100_{-0.000004}^{+0.000034} \end{pmatrix}. \quad (2.47)$$

We define the parameters λ , A , ρ , and η as,

$$\lambda \equiv s_{12}, \quad A \equiv s_{23}/\lambda^2, \quad \rho + i\eta \equiv s_{13}e^{i\delta}/A\lambda^3. \quad (2.48)$$

Expanding V in powers of λ we get,

$$V = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (2.49)$$

The Jarlskog invariant can be expressed in terms of the Wolfenstein parameters as $J = A^2\lambda^6\eta \sim (7 \times 10^{-5})\eta$. From this we can see that CP -violation in the CKM matrix is small mainly because the mixing angles are small and not solely because δ is small.

2.3 Mixing Formalism

Consider the states $|B\rangle = |\bar{b}q\rangle$ and $|\bar{B}\rangle = |b\bar{q}\rangle$ with $q = d, s$ that are eigenstates of the strong interaction Hamiltonian \mathcal{H}_0 . If we assume CPT invariance, the masses of the two states must be equal, leading to the Hamiltonian

$$\mathcal{H}_0 = \begin{pmatrix} m_0 & 0 \\ 0 & m_0 \end{pmatrix}. \quad (2.50)$$

When we add the weak interaction Hamiltonian \mathcal{H}_W , the two-state system becomes more complicated because of the new states accessible to $|B\rangle$ and $|\bar{B}\rangle$. We can write the combined Hamiltonian, $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_W$, as

$$\mathcal{H} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}, \quad (2.51)$$

where the Γ 's are decay widths. We assume CPT invariance and hence require $M_{11} = M_{22} \equiv M$ and $\Gamma_{11} = \Gamma_{22} \equiv \Gamma$. The addition of the weak

interaction, \mathcal{H}_W , introduces a slight mass shift in the states so that $M \neq m_0$. It also introduces the off-diagonal real element M_{12} via nonzero mixing amplitudes such $\langle \bar{B} | \mathcal{H}_W | B \rangle$ and off-shell continuum states accessible to $|B\rangle$, and $|\bar{B}\rangle$. The imaginary entries are caused by the accessibility of on-shell continuum states, with the diagonals given by the states common unique to $|B\rangle$ or $|\bar{B}\rangle$ and the off-diagonals those states that are common to both. Since states common to both are Cabbibo-suppressed, B - \bar{B} mixing is dominated by virtual transitions [21].

The eigenstates of \mathcal{H}_W in the terms of the strong eigenstates are,

$$\begin{aligned} |B_L\rangle &= p |B^0\rangle + q |\bar{B}^0\rangle, \\ |B_H\rangle &= p |B^0\rangle - q |\bar{B}^0\rangle, \end{aligned} \quad (2.52)$$

with,

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} = e^{-i\phi}. \quad (2.53)$$

Solving for the eigenvalues of the system, we arrive at,

$$\begin{aligned} M_{H,L} &= M \pm \text{Re} \sqrt{|M_{12}| - \frac{|\Gamma_{12}|^2}{4} - i \text{Re}(M_{12}\Gamma_{12}^*)} \equiv M \pm \Delta m/2, \\ \Gamma_{H,L} &= \Gamma \pm 2\text{Im} \sqrt{|M_{12}| - \frac{|\Gamma_{12}|^2}{4} - i \text{Re}(M_{12}\Gamma_{12}^*)} \equiv \Gamma \pm \Delta\Gamma/2, \end{aligned} \quad (2.54)$$

which satisfy,

$$\begin{aligned} \Delta m - \frac{\Delta\Gamma^2}{4} &= 4 |M_{12}|^2 - |\Gamma_{12}|^2, \\ \Delta m \Delta\Gamma &= 4 \text{Re}(M_{12}\Gamma_{12}^*). \end{aligned} \quad (2.55)$$

From Eq. 2.54 we can see the implied convention for the weak eigenvalues,

$$\begin{aligned} M &= \frac{M_H + M_L}{2}, \quad \Gamma = \frac{\Gamma_H + \Gamma_L}{2}, \\ \Delta m &= M_H - M_L, \quad \Delta\Gamma = \Gamma_H - \Gamma_L. \end{aligned} \quad (2.56)$$

Note that the sign convention for $\Delta\Gamma$ is opposite to the one used in Ref. [21].

We rearrange Eq. 2.52 to express the flavor eigenstate $|B^0\rangle$ in terms of $|B_{H,L}\rangle$ and then apply the time evolution operator in the standard fashion to obtain the oscillation probabilities as a functions of time. The result is,

$$\begin{aligned} |B^0(t)\rangle &= g_+(t) |B^0\rangle + \frac{q}{p} g_-(t) |\bar{B}^0\rangle, \\ |\bar{B}^0(t)\rangle &= \frac{p}{q} g_-(t) |B^0\rangle + g_+(t) |\bar{B}^0\rangle, \end{aligned} \quad (2.57)$$

where

$$\begin{aligned} g_+(t) &= e^{-iMt} e^{-\Gamma t/2} \left[\cosh \frac{\Delta\Gamma t}{4} \cos \frac{\Delta m t}{2} + i \sinh \frac{\Delta\Gamma t}{4} \sin \frac{\Delta m t}{2} \right], \\ g_-(t) &= e^{-iMt} e^{-\Gamma t/2} \left[\sinh \frac{\Delta\Gamma t}{4} \cos \frac{\Delta m t}{2} + i \cosh \frac{\Delta\Gamma t}{4} \sin \frac{\Delta m t}{2} \right]. \end{aligned} \quad (2.58)$$

Let $\mathcal{P}_m^B(t)$ be the probability that a particle produced as a B^0 mixed and decayed as a \bar{B}^0 , $\mathcal{P}_m^B(t) = |\langle \bar{B}^0 | B^0(t) \rangle|^2$. Let $\mathcal{P}_u^B(t)$ be the probability that the particle did not mix, $\mathcal{P}_u^B(t) = |\langle B^0 | B^0(t) \rangle|^2$, with similar definitions for

the $|\bar{B}^0\rangle$ states. We can then evaluate these probabilities as,

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

Note that these general expressions are not symmetric between the B^0 and \bar{B}^0 due to possible CP -violating effects.

If we assume that there is no CP violation in mixing so that the relative phase between M_{12} and Γ_{12} vanishes, this implies that $|q/p| = 1$. In this limit the symmetry between B^0 and \bar{B}^0 is restored so that the decay probabilities for B^0 and \bar{B}^0 are the same,

$$\mathcal{P}_{u,m}(t) = \frac{1}{2} \Gamma e^{-\Gamma t} \left(1 - \frac{\Delta\Gamma^2}{4\Gamma^2} \right) \left(\cosh \frac{\Delta\Gamma}{2} t \pm \cos \Delta m t \right). \tag{2.60}$$

If we further assume that the width difference is zero, $\Delta\Gamma = 0$, we obtain,

$$\mathcal{P}_{u,m}(t) = \frac{1}{2} \Gamma e^{-\Gamma t} (1 \pm \cos \Delta m t). \tag{2.61}$$

Figure 2.2 shows a plot of $\mathcal{P}_{u,m}(t)$ as defined in Eq. 2.61 above with $\Delta m = 17.3 \text{ ps}^{-1}$.

2.4 Mixing in the Standard Model

The lowest order SM diagrams for $B-\bar{B}$ mixing are the Feynman diagrams shown in Fig. 2.3. The external quark \bar{q} can be either \bar{d} or \bar{s} depending on whether we are studying $B^0-\bar{B}^0$ or $B_s^0-\bar{B}_s^0$ oscillations. The internal quark lines Q and F can be u , c , or t . The matrix element corresponding to the diagrams in Fig. 2.3 is $\mathcal{M} = \langle B | \mathcal{H}_W | \bar{B} \rangle$ so that $M_{12} = \mathcal{M}/(2m_B)$ when normalization is taken into account.

Because the internal quarks' coupling is proportional to the square of their masses, the top quark dominates in the loop. Making the assumption that only the top quark contributes and performing a calculation (see Ref. [26] for details), one arrives at,

$$|M_{12}| = \frac{|\mathcal{M}|}{2m_B} = \frac{G_F^2}{12\pi^2} m_{B_q} f_{B_q}^2 B_{B_q} \eta_B m_t^2 f_2(m_t^2/M_W^2) |V_{tq}^* V_{tb}|^2. \quad (2.62)$$

G_F is the Fermi constant which is related to the weak coupling in Eq. 2.10 by $G/\sqrt{2} = g^2/8M_W$. η_B is a QCD correction factor necessary because the quark lines can have an arbitrary number of gluons between them. The calculation of the internal loop of the diagram goes into the function $f_2(m_t^2/M_W^2)$, the Imani-Lim function [27], which is,

$$f_2(x_q) = x_q \left(\frac{1}{4} + \frac{9}{4} \frac{1}{1-x_q} - \frac{3}{2} \frac{1}{(1-x_q)^2} \right) - \frac{3}{2} \frac{x_q^3 \log x_q}{(1-x_q)^3}. \quad (2.63)$$

The remaining terms in Eq. 2.62 are the decay constant, f_B , and the bag parameter, B_B .

To relate $|M_{12}|$ to Δm , we would in general need to evaluate Γ_{12} as Eq. 2.55 indicates. However because Γ_{12} involves on-shell decays only, top quark loops do not contribute. We can then estimate that $\Gamma_{12}/M_{12} \approx m_b^2/m_t^2 \ll 1$ and thereby ignore the contribution of Γ_{12} . Furthermore, be-

cause $\Delta\Gamma = \mathcal{O}(10^{-2})|M_{12}|$ [13], we write,

$$\Delta m \simeq 2|M_{12}|. \quad (2.64)$$

The final formulas are then:

$$\begin{aligned} \Delta m_d &= \frac{G_F^2}{6\pi^2} m_{B_d} f_{B_d}^2 B_{B_d} \eta_B m_t^2 f_2 (m_t^2/M_W^2) |V_{td}^* V_{tb}|^2, \\ \Delta m_s &= \frac{G_F^2}{6\pi^2} m_{B_s} f_{B_s}^2 B_{B_s} \eta_B m_t^2 f_2 (m_t^2/M_W^2) |V_{ts}^* V_{tb}|^2, \end{aligned} \quad (2.65)$$

and

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s} f_{B_s}^2 B_{B_s}}{m_{B_d} f_{B_d}^2 B_{B_d}} \left| \frac{V_{ts}}{V_{td}} \right|^2. \quad (2.66)$$

The CKM matrix element V_{td} can be extracted directly from Δm_d , as Eq. 2.65 shows. However this extraction is dominated by the uncertainty on the hadronic matrix element $f_{B_d} \sqrt{B_{B_d}}$. The result of current lattice QCD calculations is $f_{B_d} \sqrt{B_{B_d}} = 244 \pm 11 \pm 24$ MeV [28]. However, in the ratio of Eq. 2.66 several uncertainties cancel so that current lattice calculations give [28]:

$$\xi = \frac{f_{B_s} \sqrt{B_{B_s}}}{f_{B_d} \sqrt{B_{B_d}}} = 1.210 \pm 0.04_{-0.01}^{+0.04}. \quad (2.67)$$

The primary motivation for measuring Δm_s comes from the above reduction in theoretical error that can be achieved by combining Δm_s with Δm_d . We measure Δm_s so that we can extract $|V_{td}|$ with a higher precision than by measuring Δm_d alone.

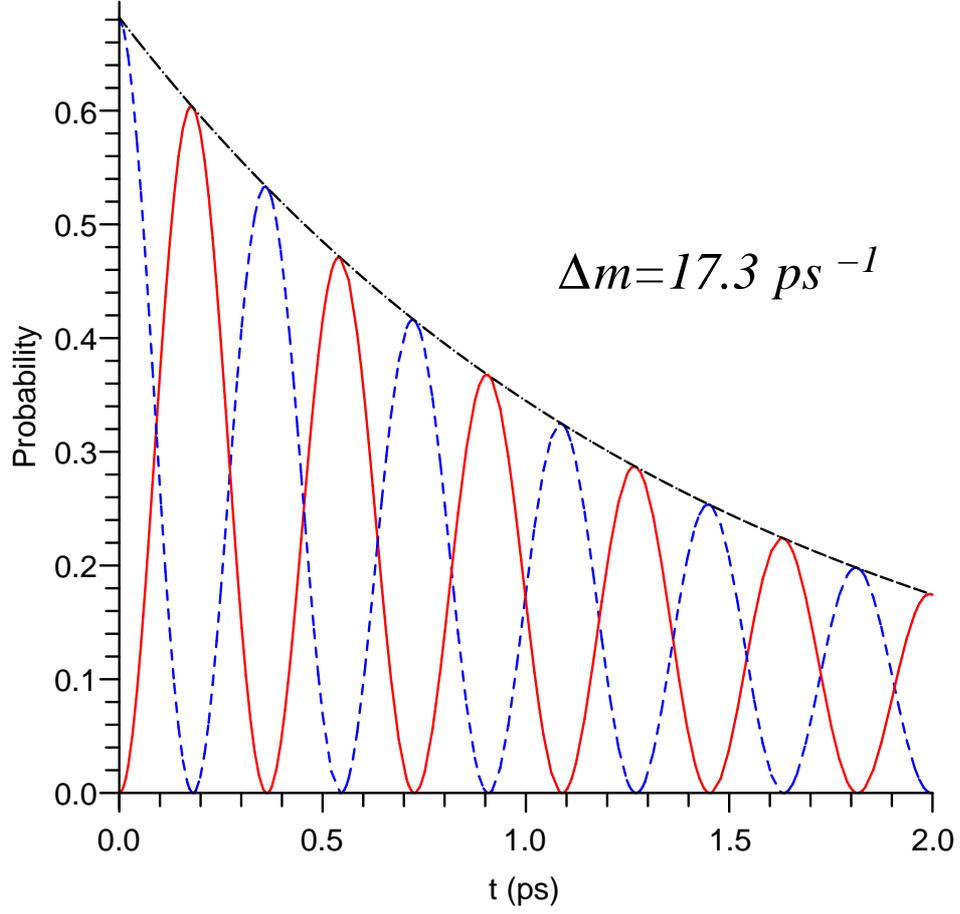


Figure 2.2: Probability to observe a mixed or unmixed event as described by Eq. 2.61. We have used $\Delta m = 17.3 \text{ ps}^{-1}$ for this plot. The blue dashed line is the probability for a non-oscillated event while the red solid line is the probability for an oscillated event. The black dot-dash line is $\Gamma e^{-\Gamma t}$ which is the decay probability function.

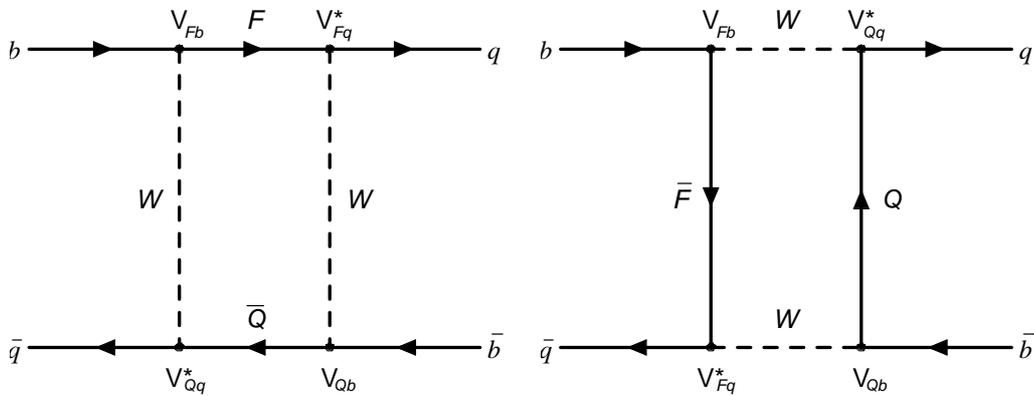


Figure 2.3: Lowest-order box diagrams responsible for B - \bar{B} mixing. The external quark lines, q , are d or s depending on whether the process represents B_d^0 or B_s^0 mixing. The internal quark propagators, Q and F , can be (u , c , or t).

Chapter 3

Experimental Overview

As mentioned in Sec. 1.2, we will perform two separate time-dependent analyses, a calibration of the flavor tagger via a measurement of Δm_d and a search for Δm_s . The data used for these analyses were collected by the DØ detector between April 2002 and February 2006, referred to as *Run IIa*. The crucial elements of these and any other time-dependent oscillation analysis are as follows:

- the proper decay time τ_B of the B meson and,
- the flavor of the B meson (B or \bar{B}) both at the time of its production and decay.

The flavor of the meson at production and decay tells us whether it has oscillated. Combining this information with τ_B enables us to construct a mixing probability in accord with Eq. 2.61. This probability is then cast into a maximum likelihood framework from which the oscillation frequency Δm can be extracted.

For both the Δm_d and Δm_s analysis, we will be studying semileptonic B decays $B \rightarrow D\ell X$ where the D can be D^+ or D_s depending on whether the B is B^0 or B_s^0 . The analyses we present use muons for the lepton in the decay chain. This muon is also the *trigger* object used to collect the majority

of our events, i.e., we use single and dimuon triggers to record semileptonic B events (see Sec. 4.3.1). Note, however, that none of the analyses use an explicit trigger requirement because there are an inconsequential number of lifetime-biasing triggers in Run IIa. We therefore treat trigger effects as systematic errors.

3.1 Vertices and Impact Parameters

We measure the decay length of the B meson, l_B , rather than its decay time because these two quantities are trivially related by the speed of light and can be used interchangeably. The first step to obtain l_B is measuring the distance between the primary vertex (PV), where the B meson is produced, and the secondary vertex (SV), where it decays. We furthermore use secondary vertices as physics object inputs to our flavor tagging algorithm, described in detail in Ch. 5. The impact parameters of tracks are also used as inputs to the flavor tagger, as well as in signal selections. Vertexing and impact parameter determination are closely related; these techniques are briefly described below. For a more detailed exposition, the reader is referred to Ref. [30].

3.1.1 Track Impact Parameters

The 3-dimensional impact parameter (IP) is the minimal distance between the estimated primary interaction point and the track trajectory. We separate the 3-dimensional information into $R\phi$ and Rz components. The $R\phi$ IP component is defined as the minimal distance between the PV and the track trajectory projected onto the plane perpendicular to the beam direction. The point of closest approach (P_C) of the track trajectory to the PV in the $R\phi$ plane is also used to define the Rz component of the IP (see Fig. 3.1).

The parameters of the track trajectory used to determine the IP are the polar and azimuthal angles (θ, ϕ) of the track at the point P_O of the closest approach to the origin O and $(\varepsilon_{R\phi}, \varepsilon_{Rz})$, the equivalent of the IP components but defined with respect to the origin O . The IP components $d_{R\phi}$ and d_{Rz} with respect to the primary vertex position \vec{V} are as follows:

$$d_{R\phi} = \varepsilon_{R\phi} - (\vec{e} \cdot \vec{V}), \quad (3.1)$$

$$\begin{aligned} d_{Rz} &= \varepsilon_{Rz} + \cot \theta (\vec{u} \cdot \vec{V}) - V_z \\ &= \varepsilon_{Rz} - (\vec{l} \cdot \vec{V}), \end{aligned} \quad (3.2)$$

where \vec{u} is the unit vector along the track direction in the $R\phi$ plane, $\vec{u} = \{\cos \phi, \sin \phi, 0\}$, \vec{e} is the unit vector perpendicular to \vec{u} , $\vec{e} = \{\sin \phi, -\cos \phi, 0\}$, and $\vec{l} = \{-\cot \theta \cos \phi, -\cot \theta \sin \phi, 1\}$. These components are shown in Fig. 3.1.

3.1.2 Vertexing

The primary vertex for each event is reconstructed using a set of selected tracks and the beam-spot position. The beam-spot is stable within a run so it can be used as a constraint for the primary vertex fit. The PV position, \vec{V} , is obtained by minimising the following χ^2 function:

$$\chi^2(\vec{V}) = \sum_a \sum_{\alpha, \beta=1,2} d_\alpha^a (S_a^{-1})_{\alpha\beta} d_\beta^a + \sum_i \frac{(V_i^{\text{sp}} - V_i)^2}{(\sigma_i^{\text{sp}})^2}. \quad (3.3)$$

$\{d_1^a, d_2^a\} = \{d_{R\phi}^a, d_{Rz}^a\}$ is the 2-dimensional vector of impact parameter components for each track a entering into the PV fit for each track a and S_a is the covariance matrix of $\{\varepsilon_{R\phi}^a, \varepsilon_{Rz}^a\}$. V_i^{sp} and σ_i^{sp} are the beam-spot position and size for the x and y coordinates.

The summation in Eq. 3.3 is first performed using all the tracks to obtain $\chi^2(N_{tr})$. Each track i is then removed consecutively and $\chi_i^2(N_{tr} - 1)$ is

calculated. The track i giving the maximal difference $\Delta_i = \chi^2(N_{tr}) - \chi_i^2(N_{tr} - 1)$ is excluded from the fit if $\Delta_i > 9$. This procedure continues until only tracks satisfying $\Delta_i < 9$ remain in the PV.

The error on the impact parameter in the $R\phi$ plane is obtained as,

$$\sigma_{R\phi}^2 = \begin{cases} (\sigma_{R\phi}^{\text{tr}})^2 - (\sigma_{R\phi}^{\text{PV}})^2 & \text{if the track is included in the PV} \\ (\sigma_{R\phi}^{\text{tr}})^2 + (\sigma_{R\phi}^{\text{PV}})^2 & \text{otherwise,} \end{cases} \quad (3.4)$$

with similar equations for σ_{Rz}^2 . $\sigma_{R\phi}^{\text{tr}}$ (σ_{Rz}^{tr}) is the error on $\varepsilon_{R\phi}$ (ε_{Rz}) coming from the track fit and σ^{PV} is the error from the PV fit. These primary vertex errors are calculated as follows:

$$(\sigma_{R\phi}^{\text{PV}})^2 = \sum_{ij} e_i S_{ij} e_j, \quad (3.5)$$

$$(\sigma_{Rz}^{\text{PV}})^2 = \sum_{ij} l_i S_{ij} l_j, \quad (3.6)$$

where S_{ij} is the covariance matrix from the primary vertex fit and \vec{e} and \vec{l} are shown in Fig. 3.1.

Secondary Vertices

In every analyzed event, jets are constructed using the DURHAM algorithm [31]. Secondary vertices are made by first removing all tracks with $p_T < 2$ GeV and those identified as products of K_S^0 , Λ^0 decay, or photon conversion. Tracks having transverse impact parameter significance, $d_{R\phi}/\sigma_{R\phi}$, less than 3 are also removed. We then select all possible combinations of pairs of the remaining tracks coming from the same jet having a common vertex with a χ^2 less than 3. After this, all the tracks from the same jet are tested one by one for inclusion into the given SV candidate if $\cos \theta_{\text{SV} \rightarrow \text{trk}} > 0.4$, where $\theta_{\text{SV} \rightarrow \text{trk}}$ is the angle between the secondary vertex direction and the candidate track. Tracks not included in the given jet are added to the list of

SV candidate tracks if $\cos \theta_{\text{SV} \rightarrow \text{trk}} > 0.85$. The track producing the smallest change Δ in of the vertex fit χ^2 is included in the SV candidate if $\Delta < 5$. Finally, SV candidates are discarded if the distance between the PV and the SV, $d_{\text{PV} \rightarrow \text{SV}}$, is greater than 2.5 cm or if $d_{\text{PV} \rightarrow \text{SV}}/\sigma_d < 4$.

3.2 Proper Decay Time

The proper lifetime of a B meson, $c\tau_B$, is obtained from the measurement of the distance X^B between the production and decay vertices as,

$$c\tau_B = \frac{X^B}{\beta\gamma} = X^B \frac{M_B}{p(B)}, \quad (3.7)$$

where β and γ are the usual Lorentz factors, and M_B and $p(B) = |\vec{p}(B)|$ are the mass and momentum of the B meson. This relation is projected onto the plane transverse to the beam line,

$$c\tau_B = X_{xy}^B \frac{M_B}{p_T(B)}, \quad (3.8)$$

because the transverse distance, X_{xy}^B , and the transverse momentum, $p_T(B)$, are measured more precisely than X^B and $p(B)$. The full momentum of a B meson decaying semileptonically cannot be reconstructed because of the presence of an undetected neutrino in the decay chain. Instead, we use the momentum of the $D + \mu$ system to calculate the “visible proper decay length” (VPDL or x^M) as,

$$x^M \equiv L_{xy}^B \frac{M_B}{p_T(D_s\mu)}. \quad (3.9)$$

L_{xy}^B is the transverse decay length, defined as the displacement \vec{X}_{xy}^B between the primary and B vertex projected along the transverse direction of the

$D + \mu$ system (see Fig. 3.2),

$$L_{xy}^B = \frac{\vec{X}_{xy}^B \cdot \vec{p}_T(D\mu)}{|\vec{p}_T(D\mu)|}. \quad (3.10)$$

Using Eqs. 3.9 and 3.10, we can write the visible proper decay length as,

$$x^M = \frac{\vec{X}_{xy}^B \cdot \vec{p}_T(D\mu)}{p_T(D\mu)} \times \frac{M_B}{p_T(D\mu)}. \quad (3.11)$$

The proper lifetime is then obtained by multiplying by a correction factor,

$$c\tau_B = x^M \cdot K, \quad K \equiv \frac{p_T(D\mu)}{p_T(B)}, \quad (3.12)$$

known as the “ K -factor”. The K -factor reflects the difference between the measured $p_T(D\mu)$ and the true $p_T(B)$ momentum of the B meson; it is estimated from Monte Carlo simulations.

The proper decay time uncertainty, σ_t , can be expressed as,

$$\sigma_t = \sigma(L_{xy}^B) \oplus t \cdot \frac{\sigma(K)}{K}, \quad (3.13)$$

where $\sigma(L_{xy}^B)$ is the uncertainty on L_{xy}^B due to vertexing tracks with measurement errors and $\sigma(K)/K$ is the K -factor resolution. It is important to note that the K -factor resolution scales with decay length while the vertexing resolution does not and thus only adds a constant uncertainty. This means that for semileptonic decays, events with short decay length are more sensitive to oscillations.

3.3 Flavor Tagging

To determine whether a particular meson has mixed we must know its flavor at both production and at decay. For semileptonic channels, the flavor

of the meson at decay is trivially given by the charge of the lepton. Determining the flavor at production is much more difficult and is known as *initial state flavor tagging*.

Figure 3.3 shows a drawing of a typical semileptonic B event. Because b quarks are predominantly produced in pairs, i.e. $b\bar{b}$, we can infer that there is another b quark in an event in which we reconstruct a B hadron. Thus, we label the two sides of the event as “Reconstructed” and “Opposite” as shown in Fig. 3.3. The flavor taggers used in the analyses presented here only use properties of the opposite side of the event. These opposite-side taggers (OST’s) have the advantage that there is almost no correlation between the hadronization of a reconstructed-side b quark and an opposite-side b quark. Δm_d is well known, so we can apply the flavor tagging to a measurement of Δm_d to see if we get an answer consistent with the world average and thereby validate the tagger. Furthermore, because the performance of an OST should not depend on the species of reconstructed b hadron (i.e., B^+ , B^0 , or B_s), we can quantify its performance on the Δm_d sample and then use this characterization as an input to the search for Δm_s . Three objects are shown on the opposite side of the event in Fig. 3.3: a secondary vertex, a muon, and an electron. These are the primary objects that are used to discriminate the flavor of the opposite side and thereby infer the flavor of the reconstructed side at production. More details of the construction and implementation of the flavor tagging are given in Ch. 5.

The figure of merit used to compare different taggers is the “effective efficiency”, $\epsilon\mathcal{D}^2$, where ϵ is the tagging efficiency and \mathcal{D} is the “dilution” given by:

$$\epsilon \equiv \frac{N_{\text{tagged}}}{N_{\text{total}}}, \quad \mathcal{D} \equiv \frac{N_{\text{correct}} - N_{\text{wrong}}}{N_{\text{correct}} + N_{\text{wrong}}}, \quad (3.14)$$

where N_{correct} (N_{wrong}) is the number of events that have been correctly (incorrectly) tagged and $N_{\text{tagged}} = N_{\text{correct}} + N_{\text{wrong}}$. We also occasionally refer

to the “tagging purity” η_s , defined as,

$$\eta_s \equiv \frac{N_{\text{correct}}}{N_{\text{tagged}}}. \quad (3.15)$$

Clearly, $\mathcal{D} = 2\eta_s - 1$.

Flavor tagging modifies the oscillation probability given in Eq. 2.61 in an important way. A tagger that has purity η_s will lead to the following probabilities to observe mixed and unmixed events,

$$\mathcal{P}_m^{\text{obs}}(t) = \eta_s \mathcal{P}_m(t) + (1 - \eta_s) \mathcal{P}_u(t), \quad (3.16)$$

$$\mathcal{P}_u^{\text{obs}}(t) = (1 - \eta_s) \mathcal{P}_m(t) + \eta_s \mathcal{P}_u(t). \quad (3.17)$$

Substituting $\mathcal{P}_{u,m}(t)$ from Eq. 2.61 into Eqs. 3.16 and 3.17, we arrive at,

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

Figure 3.4 shows a plot of $\mathcal{P}_{u,m}(t)$ as defined by Eq. 3.18. By comparing Fig. 2.2 to Fig. 3.4 we can see that flavor tagging introduces an offset into the probabilities so that neither $\mathcal{P}_u(t)$ nor $\mathcal{P}_m(t)$ goes to zero.

We can see an intuitive interpretation of the dilution if we look at the mixing asymmetry,

$$A(t) = \frac{\mathcal{P}_u(t) - \mathcal{P}_m(t)}{\mathcal{P}_u(t) + \mathcal{P}_m(t)} = \mathcal{D} \cos \Delta mt; \quad (3.19)$$

the dilution is the maximum value of the asymmetry, i.e. $A(t = 0) = \mathcal{D}$.

3.4 Statistical Significance

The expected signal to noise ratio of an oscillation measurement is [29],

$$\text{Significance}(\Delta m) \approx \frac{S}{\sqrt{S+B}} \sqrt{\frac{\epsilon \mathcal{D}^2}{2}} \exp\left(-\frac{(\Delta m \sigma_t)^2}{2}\right), \quad (3.20)$$

where S is the number of signal events, B is the number of background events, $\epsilon \mathcal{D}^2$ is the effective efficiency as defined in Sec. 3.3, and σ_t is the proper time resolution. We can therefore improve the sensitivity of an oscillation measurement in the following ways:

- increase the statistics of the sample;
- increase the signal to background ratio;
- improve the flavor tagging (increase $\epsilon \mathcal{D}^2$); and/or
- improve the proper time resolution (e.g., use fully reconstructed decays).

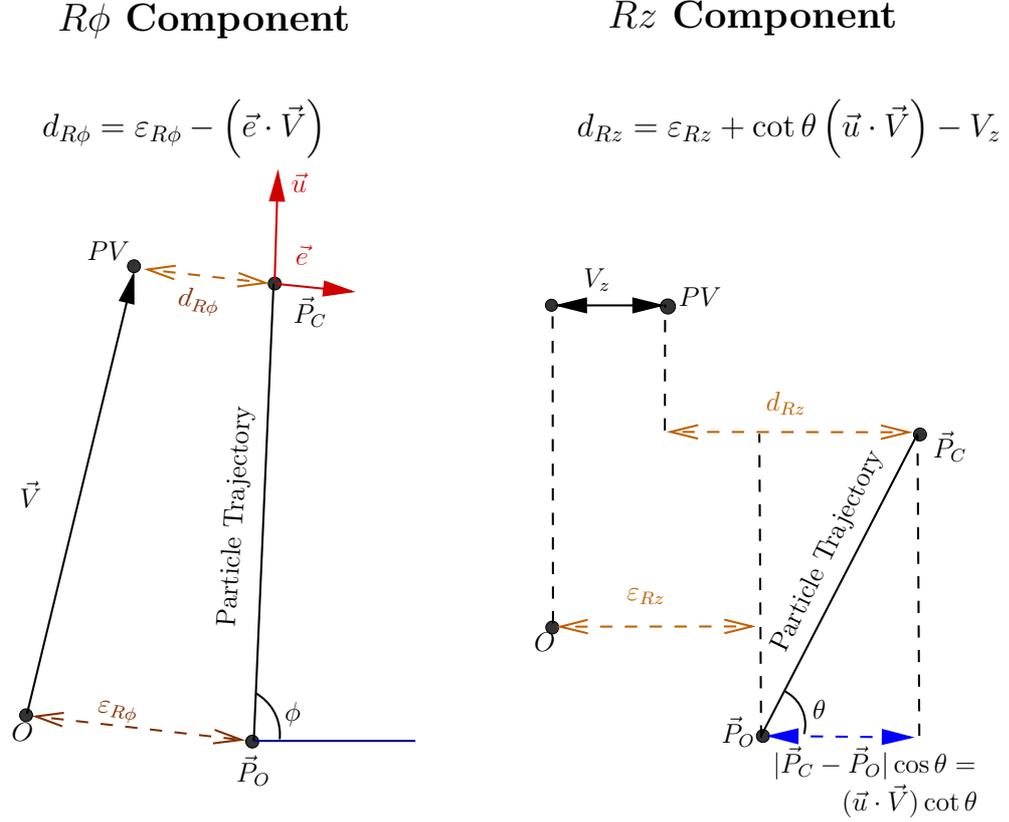


Figure 3.1: Definition of $R\phi$ and Rz impact parameter (IP) components. \vec{u} is a unit vector along the R - ϕ projection of the track direction, and \vec{e} is a unit vector perpendicular to \vec{u} in the $R\phi$ plane. \vec{V} is the vector from the origin O to the primary vertex PV . P_0 and P_c are the points of closest approach in the $R\phi$ plane of track trajectory to O and the PV respectively. The diagrams show the projections onto the $R\phi$ and Rz planes. The IP components are $d_{R\phi}$ and d_{Rz} , while $\varepsilon_{R\phi}$ and ε_{Rz} are the corresponding components from P_0 to the origin. Adapted from Ref. [30].

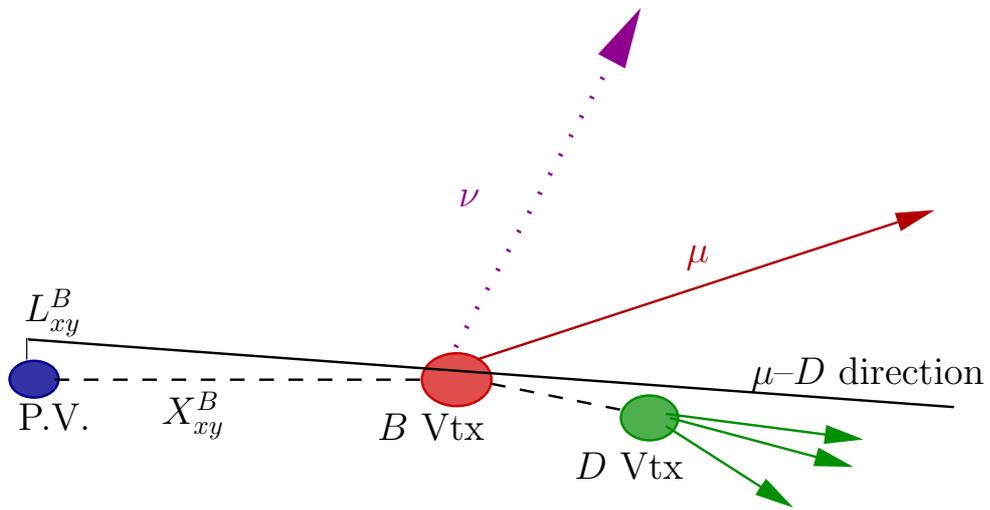


Figure 3.2: Schematic drawing showing the definition of transverse distance X_{xy} and transverse decay length L_{xy} .

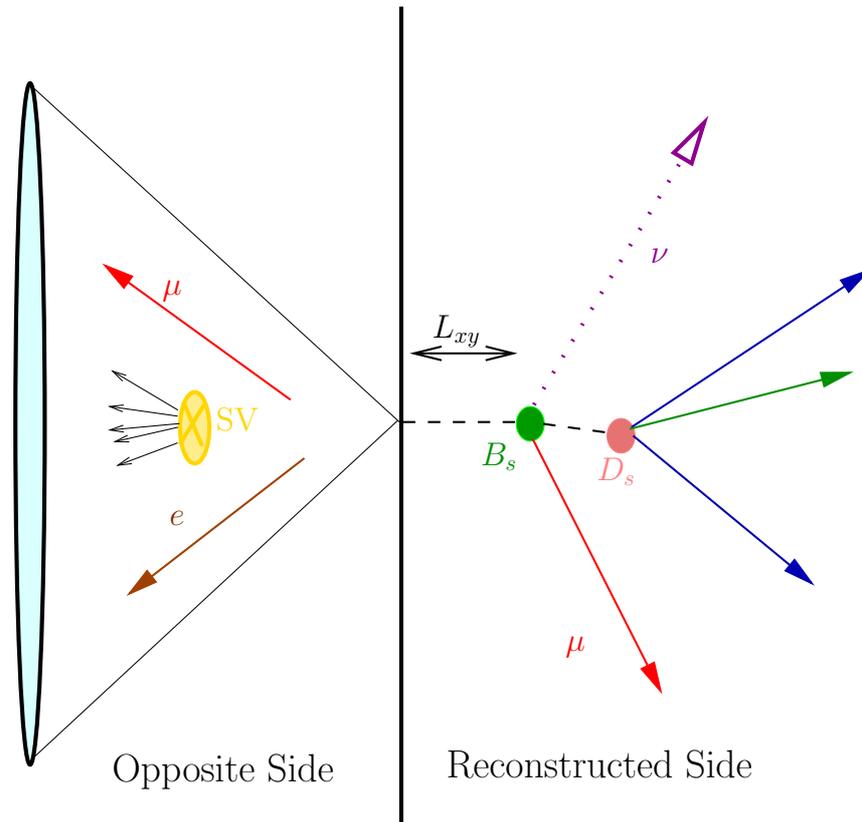


Figure 3.3: A diagram of a typical semileptonic B_s^0 event showing the reconstructed meson as well as objects on the opposite side of the event that are used as discriminants for the flavor tagger.

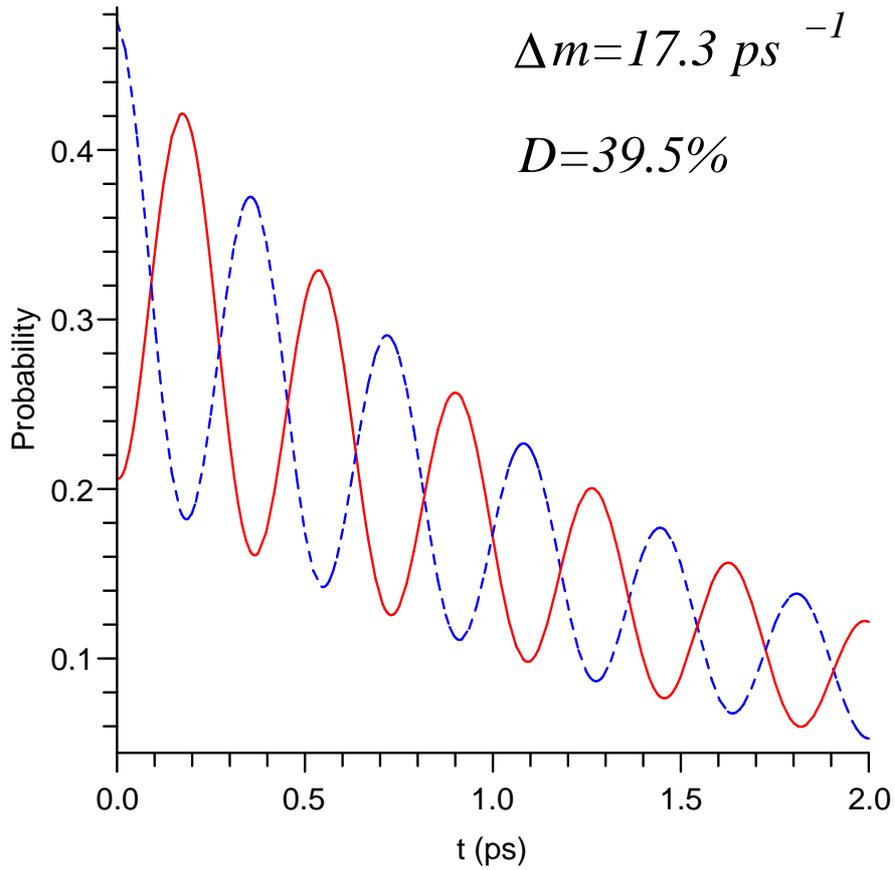


Figure 3.4: Probability to observe a mixed or unmixed event after flavor tagging as described by Eq. 3.18. We have used $\Delta m = 17.3 \text{ ps}^{-1}$ and $\mathcal{D} = 0.395$ for this plot. The blue dashed line is the probability for a non-oscillated event while the red solid line is the probability for an oscillated event.

Chapter 4

The Experimental Apparatus

4.1 The Tevatron Accelerator Complex

The Tevatron, located 40 miles west of Chicago at the Fermi National Accelerator Laboratory (FNAL), is a synchrotron accelerator that collides protons with antiprotons. Collisions occur every 396 nsec at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV, thus making the Tevatron the highest energy collider in the world until the Large Hadron Collider begins operating at CERN. Figure 4.1 shows a schematic view of the Fermilab accelerator complex. A brief review of the complex is presented here; for a more detailed description the reader is referred to Refs. [32] and [33].

The proton beam originates as a pulsed 18 KeV negative hydrogen ion beam from a magnetron surface-plasma source. A schematic of a basic magnetron source is shown in Fig. 4.2. It consists of a cathode surrounded by an anode with a small gap, typically 1 mm, with a magnetic field passing through the apparatus. Hydrogen gas is added to a pressure of a few hundred millitorr and a dense plasma is produced while electrons are confined to spiral in the anode-cathode gap. Energetic particles strike the cathode and sputter off hydrogen atoms which have been absorbed on the surface. A Cesium vapor coating the cathode surface raises the probability that a hy-

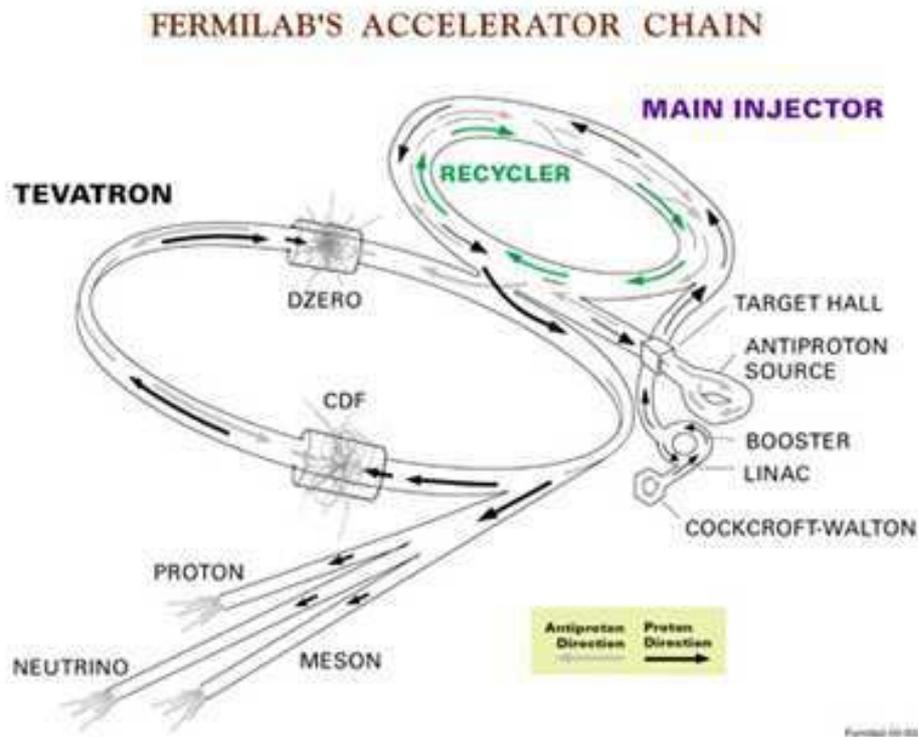


Figure 4.1: The Tevatron accelerator complex.

drogen atom will remove the necessary electrons. After the H^- are formed they are extracted through the anode aperture and accelerated through the extraction plate.

The hydrogen ions are then accelerated through a Cockcroft-Walton generator to an energy of 750 KeV and injected into the Linac, a long line of radio frequency (RF) cavities consisting of drift tubes separated by gaps. A particle travelling down the Linac experiences an accelerating field while in the gap between the drift tubes and is shielded from the decelerating field within the drift tube. The hydrogen ions are thus bunched together in the Linac, accelerated to 400 MeV and injected into the Booster, a synchrotron accelerator that accelerates the protons to 8.9 GeV in 33 msec. The Booster

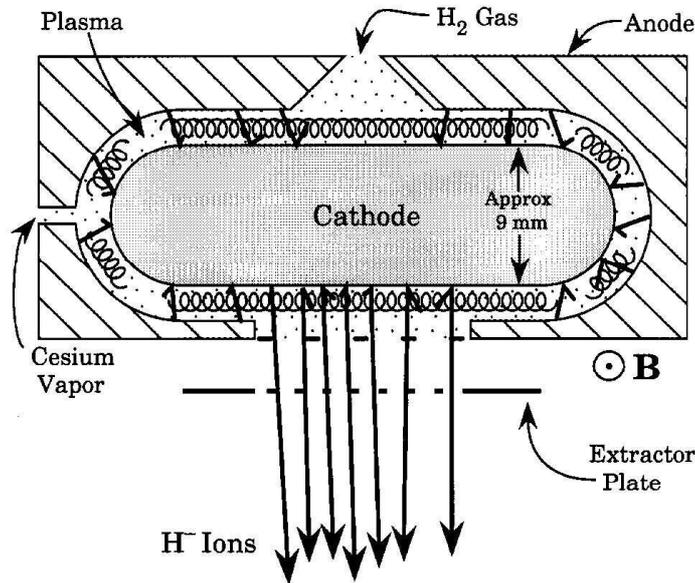


Figure 4.2: A simple magnetron source [32].

uses a charge-exchange injection system in which negative hydrogen ions are brought into a parallel path with protons circulating in a closed orbit in a straight section. The two beams are then merged and passed through a carbon foil which strips the electrons from the H^- 's. The original proton beam is then restored into a closed orbit while unstripped ions are passed to a beam dump.

The proton bunches are then transferred to the Main Injector, a larger synchrotron accelerator that operates in two modes. In the first mode the protons are accelerated to 120 GeV and sent to the \bar{p} production target. In the second mode the protons are accelerated to 150 GeV and are injected into the Tevatron. Antiprotons are created by firing protons onto a nickel target and then focusing the secondaries produced through a lithium lens. These antiprotons are then sent to the Debuncher where their momentum spread is reduced and the transverse profile of the beam is reduced via stochastic

cooling [34]. They are then stored in the Accumulator ring until they are needed for a period of collisions, referred to as a “store”. The collection of antiprotons in the Accumulator is referred to as the “stack”. Because the stacking rate in the Accumulator decreases as the stack size increases, the stack is transferred either to the Main Injector for acceleration or to the Recycler Ring for further storage and cooling. Transfers of stacks to the Recycler enable the Accumulator to stack at a faster rate and thereby increase the number of antiprotons available for collisions. Antiprotons in the Recycler, referred to as the “stash”, are stored until they are needed for a store in which case they are transferred to the Main Injector for further acceleration.

The proton or antiproton beams are then injected into the Tevatron ring where they are accelerated to 980 MeV and steered by superconducting magnets. Each beam has 36 bunches distributed in three groups of 12 called superbunches. The relative position of the bunches is marked by time periods of 132 ns called ticks, of which there are 159 in the ring. Within a superbunch, the bunch spacing is three ticks (396 ns, which corresponds to about 120 m). The tick and turn structure of the Tevatron beams is shown in Fig. 4.3. The p and \bar{p} beams are kept in helical orbits everywhere except the two interaction regions: DØ, where the detector of the same name is located, and BØ where the Collider Detector at Fermilab (CDF) is housed. At these two locations, quadrupole magnets squeeze the beams into a cross-sectional area of $\sigma_a \approx 5 \times 10^{-5} \text{ cm}^2$ such that the beams collide in the geometrical center of each detector.

4.2 The DØ Detector

In this section we present a brief overview of the DØ detector. For a more detailed presentation, the reader is referred to Ref. [35].

The DØ detector is a multi-purpose physics detector comprised of a num-

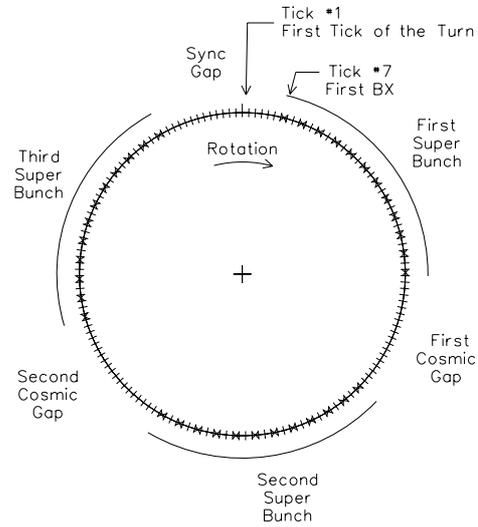


Figure 4.3: The Tevatron beam structure showing 36 bunches distributed in 3 superbunches.

ber of subsystems that enclose each other. Figure 4.4 shows a schematic view of the detector. The following subsystems are presented in the subsections of this section:

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

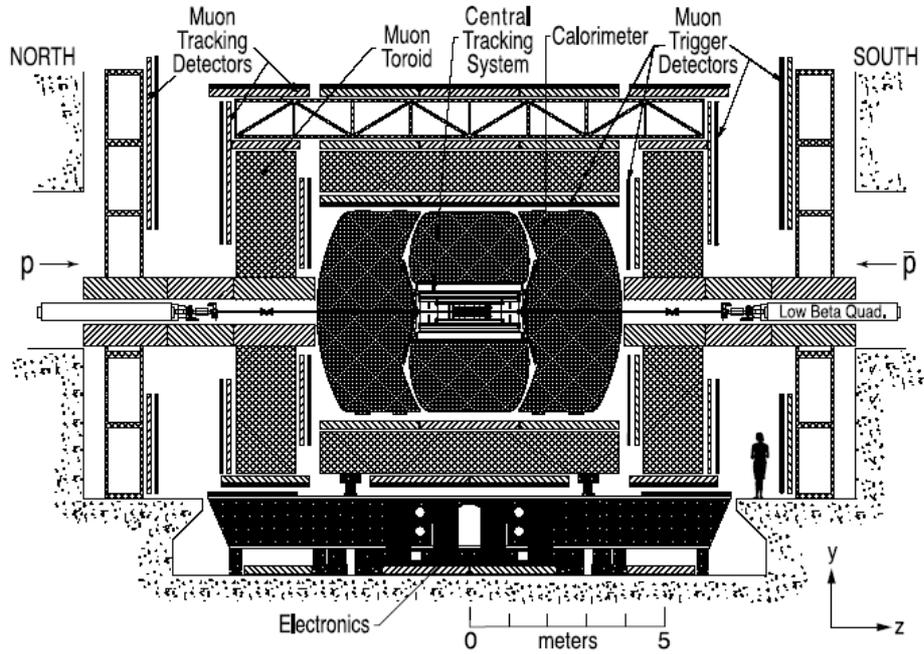


Figure 4.4: Schematic of the DØ detector, as installed in the collision hall and viewed from inside the Tevatron ring [35].

4.2.1 Coordinate System

DØ uses a right-handed coordinate system with the z -axis along the proton direction and the y -axis in the upward direction. The x -axis is therefore in the direction of the center of the Tevatron ring. The angles ϕ and θ are the azimuthal and polar angles, respectively. We use r to denote the perpendicular distance from the z -axis. We define the *pseudorapidity* as,

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]. \quad (4.1)$$

The pseudorapidity is used to approximate the true rapidity,

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (4.2)$$

for finite angles in the limit that $(m/E) \rightarrow 0$, which is almost always valid for the relativistic particles arising from interactions at $\sqrt{s} = 1.96$ TeV. We use the term “forward” to describe regions at large $|\eta|$.

4.2.2 Central Tracking System

The central tracking system consists of the silicon microstrip tracker (SMT) and the central fiber tracker (CFT) surrounded by a 2 T solenoidal magnet. The tracking system encloses the DØ beryllium beam pipe, which has a wall thickness of 0.508 mm, an outer diameter of 38.1 mm, and is 23.7 m long. Outside of the solenoid is a scintillator-based preshower detector along with other preshower detectors that are mounted on the inner surfaces of the forward calorimeter cryostats. A diagram of the tracking system is shown Fig. 4.5.

The tracking detectors can locate the primary interaction vertex with a resolution of approximately $35 \mu\text{m}$ along the beamline. Furthermore, the tracking system can measure the impact parameters of tracks with respect to the primary vertex with a resolution of $\sigma_{IP} = 55 \mu\text{m}$ for tracks having transverse momentum $0.7 < p_T < 2.0$ GeV and a resolution of $\sigma_{IP} = 25 \mu\text{m}$ for tracks with $5.0 < p_T < 10.0$ GeV.

Silicon microstrip tracker

Silicon detectors are p - n junction diodes operated at reverse bias [13]. They are attractive for particle physics because they can be fabricated into small structures leading to very precise position measurements of charged tracks and they have low ionization energy. For example, in silicon one gets

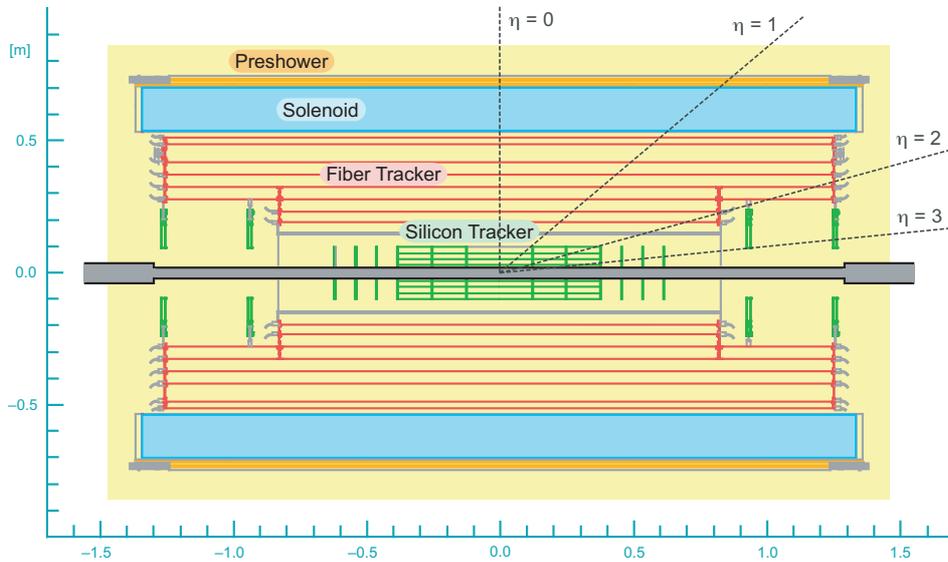


Figure 4.5: Schematic drawing of the DØ tracking system.

an electron-hole pair for every 3.6 eV released by a particle crossing the medium [36]. This is quite low compared to the 30 eV required to ionize a molecule in a gaseous detector or 300 eV to extract an electron from a plastic scintillator coupled to a photocathode. We now briefly explain the basic principles of the operation of a generic silicon detector.

Silicon has four electrons on its valence shell. p - and n materials are obtained by replacing some of the silicon atoms by atoms with five or three valence electrons respectively [36], a process known as “doping”. Figure 4.6 shows a simple diagram of a silicon-based detector. 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, depleting the n^- substrate of free electrons and creating an electric field in the n^- substrate. A charged particle that passes through the silicon ionizes and leaves electron and hole pairs. The holes drift to the p^+ strips producing

an electric signal. These signals are read by an integrated circuit at the end of the strip thereby enabling the measurement of the position of the particle.

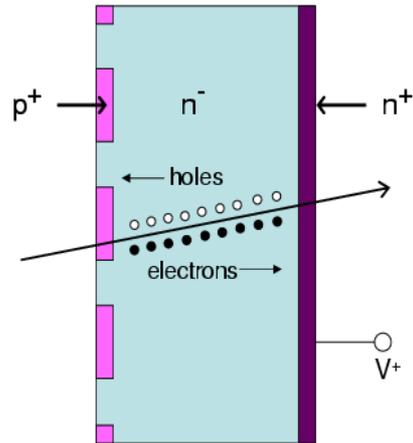


Figure 4.6: A diagram of a generic silicon detector [37].

The design of the DØ SMT is in large part dictated by the long interaction region, $\sigma_z \approx 25$ cm, which makes it difficult to deploy detectors such that tracks are perpendicular to the detector surfaces. The resulting design uses barrel modules interspersed with disks in the central region and assemblies of disks in the forward regions (see Fig. 4.7); the barrel detectors measure the $r - \phi$ coordinate while the disks measure $r - z$ as well as $r - \phi$.

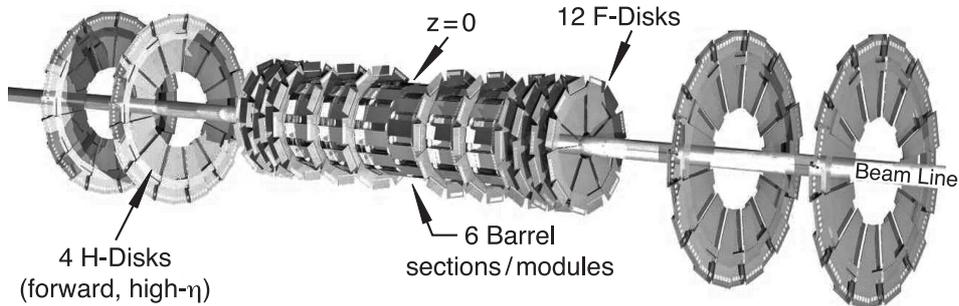


Figure 4.7: The disk/barrel design of the silicon microstrip tracker [35].

The SMT has six 12-cm long barrel detectors. Each barrel has four silicon readout layers composed of an array of silicon modules. These modules are referred to as “ladders”. A cross-sectional view of an SMT barrel is shown in Fig. 4.8. Layers 1 and 2 have twelve ladders each while layers 3 and 4 have twenty-four ladders each, for a total of 432 ladders. In the central four barrels, layers 1 and 3 are composed of double-sided¹ sensors with axial strips on one side and 90° stereo angle strips on the other, with pitches of 50 μm and 135 μm respectively. The two outer barrels have single-sided ladders with 50 μm pitch axial strips for layers 1 and 3. Layers 2 and 4 of all 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 other side.

Each barrel is capped at high $|z|$ with a disk of twelve double-sided wedge detectors, referred to as an “F-disk”. Forward of the three disk/barrel assemblies on each side is a unit consisting of three F-disks. All F-disks have an inner radius of 2.57 cm and an outer radius of 9.96 cm. These disks are composed of double-sided sensors with 50 μm , -15° stereo angle on one side and 62.5 μm pitch, $+15^\circ$ stereo angle on the other.

In the far forward regions, two large-diameter disks, “H-disk”, provide tracking at high $|\eta|$ on both the $+z$ and $-z$ side. These disks are made from

¹Double-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.

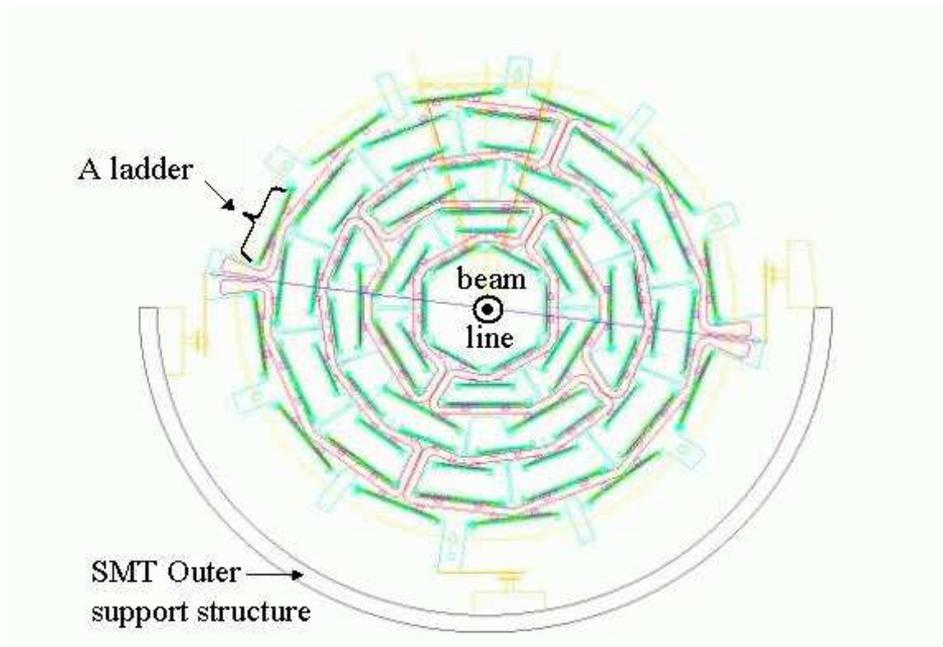


Figure 4.8: Cross section of an SMT barrel module showing the position of “ladders” [37].

single sided sensors with an inner radius of 9.5 cm and an outer radius of 26 cm. These four H-disks help to extend the SMT coverage to $|\eta| \sim 3$.

The centers of the barrels are at $|z| = 6.2, 19.0,$ and 31.8 cm. The F-disks are located at $|z| = 12.5, 25.3, 38.2, 43.1, 48.1,$ and 53.1 cm. Finally, the H-disks are located at $|z| = 100.4,$ and 121.0 cm. Table 4.1 summarizes some of the characteristics of the SMT.

The SMT is read out by 128-channel readout chips called SVXIIe chips [38]. These chips are designed to work with double-sided detectors and are mounted on a high density interconnect (HDI). The data passes from the HDI via adaptor cards and interface boards to sequencer boards. Data is sent to the sequencers via optical link fibers. There are a total of 792,576 readout channels in the SMT.

	Barrels	F-Disks	H-Disks
Channels	387,072	258,048	147,456
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).

Central Fiber Tracker

The Central Fiber Tracker (CFT) surrounds the SMT and is composed of scintillating fibers mounted on eight concentric carbon fiber support cylinders. It occupies the radial space from 20 to 52 cm from the center of the beampipe. Figure 4.9 shows a drawing of the top half of the tracking system in the $y - z$ plane. Note that to accommodate the SMT H-disks, the two innermost cylinders are 1.66 m long while the six outer cylinders are 2.52 m long.

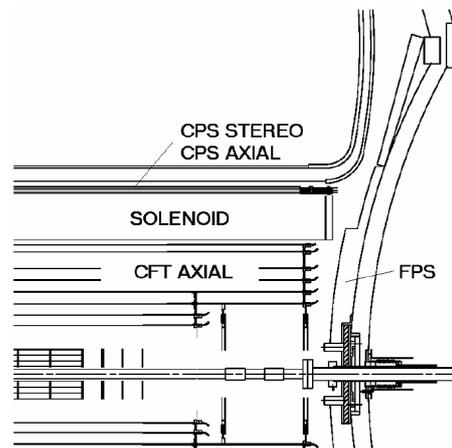


Figure 4.9: A view of the top half of the tracking system in the $y - z$ plane.

The scintillating fibers of which the CFT is composed consist of a polystyrene core clad in an inner acrylic layer and an outer layer of fluoroacrylate. The claddings, both of which are approximately $25\ \mu\text{m}$ thick, improve the attenuation length of light in the fiber to about 5 m. The complete fibers are $835\ \mu\text{m}$ in diameter. The polystyrene core of the fiber is doped with 1% paraterphenyl and 1500 ppm 3-hydroxyflavone. The paterphenyl increases the light yield of the fiber, while the 3-hydroxyflavone shifts the wavelength of light emitted by the paraterphenyl to 530 nm, which matches the peak emission wavelength of the polystyrene. The tracker contains 76,800 such scintillating fibers.

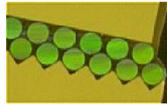


Figure 4.10: An end-on view of a CFT ribbon.

The fibers are formed into doublet layers, as seen in Fig. 4.10, and are mounted on concentric cylinders arranged in layers between $r = 20\ \text{cm}$ and $r = 52\ \text{cm}$. Each cylinder contains one doublet layer of fibers oriented along the beam direction (z) and a second doublet layer at a stereo angle in ϕ of $\pm 3^\circ$. Layers with fibers oriented along the beam axis are referred to as “axial” layers and the layers oriented at $\phi = \pm 3^\circ$ are referred to as “stereo” layers. The axial layers provide an $r - \phi$ measurement while the stereo layers allow three-dimensional reconstruction of tracks. Table 4.2 summarizes some of the important properties of the CFT.

One end of the scintillating fibers is mirrored with a sputtered aluminum coating that provides a reflectivity of about 90%. Clear fiber waveguides attached to the readout end of the scintillating fibers route the light to Visible Light Photon Counters (VLPCs). VLPCs are arsenic doped silicon avalanche photodetectors that operate at temperatures of $8 - 10^\circ\ \text{K}$ and are capable of detecting single photons. They have excellent quantum efficiencies (greater

Layer	Radius (cm)	Fibers/layer	Fiber separation (μm)	Active length (m)
A	20.04	1280×2	982.4	1.66
<i>Au</i>	20.22	1280×2	990.3	1.66
B	24.93	1600×2	978.3	1.66
<i>Bv</i>	25.13	1600×2	985.1	1.66
C	29.87	1920×2	976.1	2.52
<i>Cu</i>	30.05	1920×2	980.9	2.52
D	34.77	2240×2	974.4	2.52
<i>Dv</i>	34.95	2240×2	979.3	2.52
E	39.66	2560×2	971.7	2.52
<i>Eu</i>	39.86	2560×2	976.3	2.52
F	44.56	2880×2	970.0	2.52
<i>Fv</i>	44.74	2880×2	974.3	2.52
G	49.49	3200×2	969.8	2.52
<i>Gu</i>	49.67	3200×2	973.3	2.52
H	51.97	3520×2	926.1	2.52
<i>Hv</i>	52.15	3520×2	927.8	2.52

Table 4.2: Design parameters of the CFT; $u = +3^\circ$, $v = -3^\circ$. A through H correspond to the eight axial layers of the CFT [35].

than 75%), high gain (between 22,000 and 65,000 electrons per incoming photon), and less than 0.1% average noise. Together with the fibers, these provide a position resolution of $\sim 100 \mu\text{m}$.

Solenoidal Magnet

The solenoid that surrounds the tracking system is 2.73 m long and 1.42 m in diameter. It is wound with two layers of 0.848 mm superconducting Cu:NbTi strands stabilized with pure aluminum. The magnet operates at 4.7 K with a current of 4749 A and a corresponding central magnetic field of 2 T.

Preshower Detectors

The preshower detectors aid in electron identification, enhance spatial matching between tracks and calorimeter showers, and correct the electromagnetic energy measurement of the calorimeters for losses in the solenoid and upstream material. Figure 4.9 shows the location of the preshower detectors in the tracking system. The central preshower detector (CPS) covers the region $|\eta| < 1.3$ and is located in the 5 cm gap between the solenoid and the central calorimeter. The forward preshower detectors (FPS) cover $1.5 < |\eta| < 2.5$ and are attached to the faces of the endcap calorimeter.

The CPS consists of three concentric cylindrical layers of triangular scintillator strips with a wavelength-shifting (WLS) fiber embedded in the center of each strip. The WLS fibers transmit the light generated in the scintillator to waveguides and ultimately to VLPC's where readout is performed in a manner similar to the CFT. The three layers of scintillator are arranged in an axial- u - v geometry, where the stereo angles are $\sim \pm 23^\circ$. Between the solenoid and CPS is a lead radiator which is approximately 1 radiation length (X_0) thick. The solenoid is itself $0.9X_0$ thick, providing a total of about two radiation lengths of material at normal incidence, increasing to four radiation lengths at large angles.

The FPS has a design similar to the CPS, using scintillating strips with WLS fibers in the center to transmit the light to VLPC's. Figure 4.11 shows the layout geometry of both the CPS and FPS. The FPS detectors are made from two layers, at different z , of two planes of scintillator strips. A lead-stainless-steel absorber of thickness $2X_0$ separates the two layers. The inner layer helps to detect minimum ionizing particles (MIP). For the region, $1.5 < |\eta| < 1.65$, there is only one scintillator layer and no absorber layer. This region lies in the shadow of the solenoidal magnet coil, which provides up to $3X_0$ of material in front of the FPS, thus rendering MIP and absorber layers unnecessary.

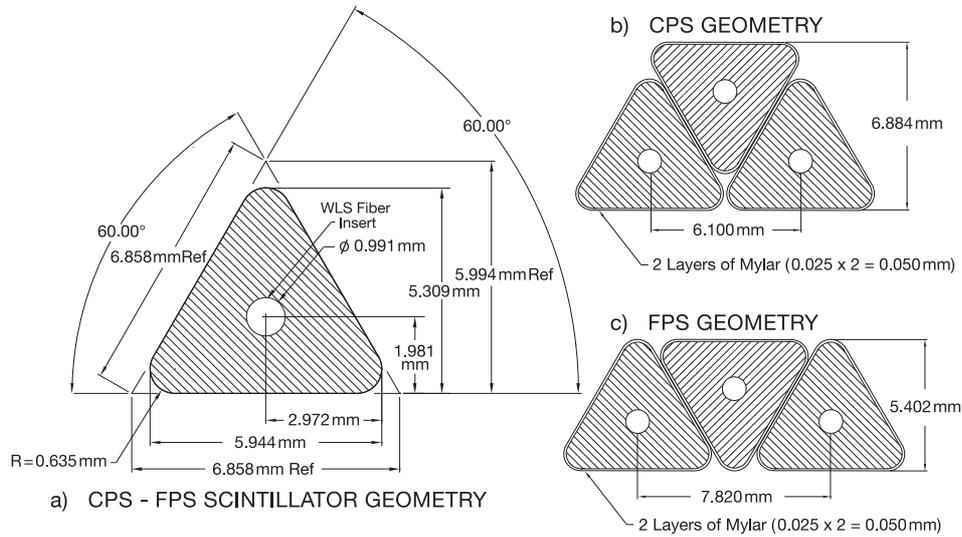


Figure 4.11: Cross section and layout geometry of the CPS and FPS scintillator strips. The circles show the location of the embedded wavelength-shifting fibers [35].

4.2.3 Calorimeter

A detector that measures the amount of energy lost by a particle or group of particles as they traverse through material is called a *calorimeter*. Calorimeters may be either “homogenous” or “sampling”, the difference being that the entire volume of a homogenous calorimeter outputs a signal while only a portion of a sampling calorimeter does. Sampling calorimeters, such as the DØ calorimeter, contain an absorber of high density to cause energy loss and an active medium which generates a signal.

All calorimeters operate by measuring the energy loss of particles through their interaction with a medium, with the interaction being caused by either the electromagnetic or strong nuclear force. We will first discuss electromagnetic energy loss. For electrons this is dominated by one of two processes depending on the energy of the particle: low-energy electrons (< 10 MeV) lose energy primarily via ionization while high-energy electrons (> 10 MeV)

lose energy primarily via brehmsstrahlung, as seen in Fig. 4.12. The radiation length, X_0 , implicit in the y -axis in this figure is the characteristic length scale for electromagnetic energy loss in materials. It represents both the mean distance over which a high-energy electron loses all but $1/e$ of its energy to brehmsstrahlung and $7/9$ of the mean free path for pair production by a high-energy photon [39].

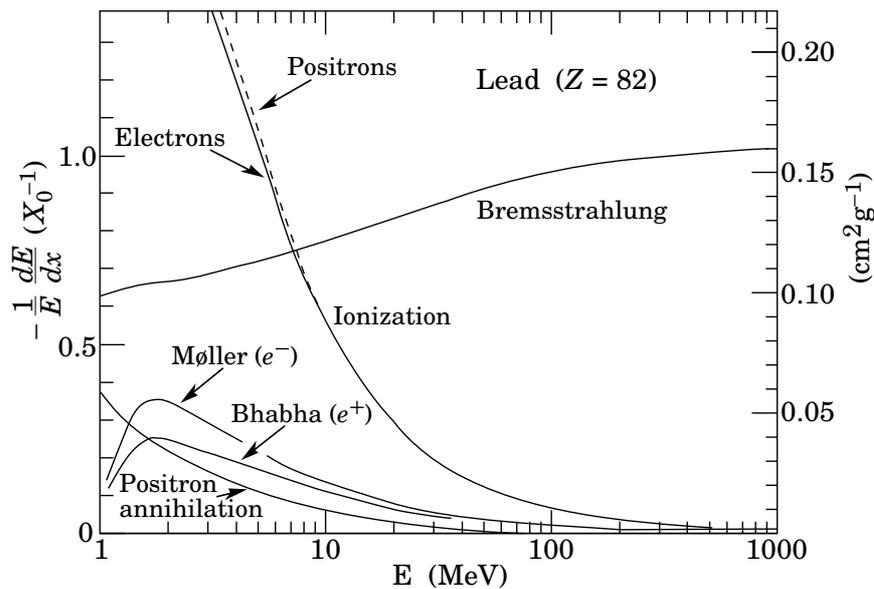


Figure 4.12: Fractional energy loss per radiation length in lead as a function of electron or positron energy [13].

The photons produced by high-energy electron brehmsstrahlung tend to lose energy by pair production. The electrons and positrons that are the products of this pair production go on to radiate more photons which then pair produce more electrons and positrons. Thus, an electromagnetic “shower” is created through the interaction of a single particle in the absorber material. Figure 4.13 depicts the development of such a shower. At some point, the energy of the produced electrons becomes low enough that

they lose energy primarily through ionization rather than through radiation. At that point, the showering stops because photon production ceases.

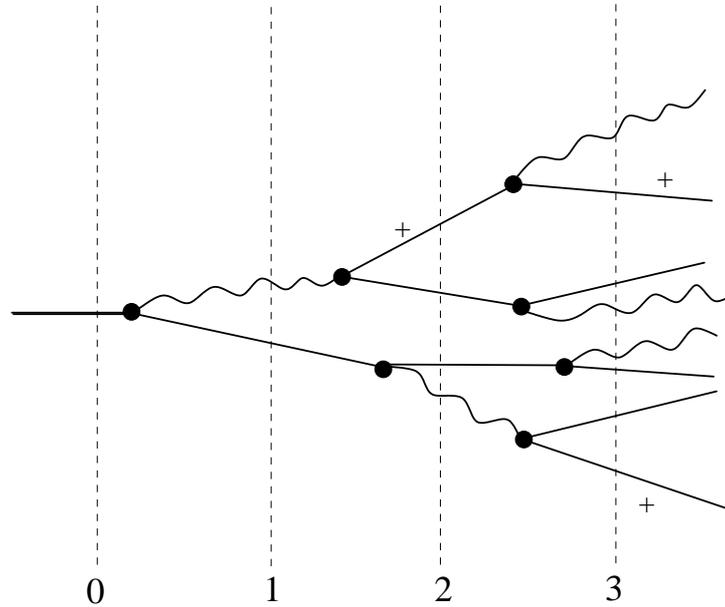


Figure 4.13: Diagram of the development of an electromagnetic shower in a calorimeter. Solid lines (with +) indicate electrons (positrons) and wavy lines indicate photons. The numbers at the bottom show the distance measured in radiation lengths with the absorber beginning at 0. Adapted from Ref. [40].

Hadronic particles interact with material via strong nuclear interactions. A hadron, such as a pion, has an inelastic interaction with the nucleus of the absorber material causing the emission of other hadrons which go on to have other inelastic interactions. Thus, similar to the case of purely electromagnetic energy loss, a cascade of particles is produced from a single incoming hadron. In analogy to the radiation length characterizing electromagnetic energy loss, we define a nuclear interaction length, λ_I . The nuclear interaction length for a given material tends to be larger than the radiation length. For example, uranium has $\lambda_I \approx 10.5$ cm and $X_0 \approx 0.32$ cm. Therefore, hadronic showers tend to be more extended in space than electromagnetic

showers. Furthermore, because some portion of the particles produced in hadronic showers are neutral pions decaying to photons, there is an electromagnetic component to these cascades. We typically see a sharp peak of energy deposited near the first interaction point from these $\pi^0 \rightarrow \gamma\gamma$ decays.

The DØ calorimeter is composed of three modules: a central calorimeter (CC) covering the region $|\eta| \approx 1$ and two end calorimeters (EC) extending the coverage to $|\eta| \approx 4$. Each calorimeter contains an electromagnetic (EM) portion closest to the interaction region followed by fine hadronic (CH) and coarse hadronic (FH) sections. A drawing of the calorimeter system highlighting the relative placement of the various sections is seen in Fig. 4.14. A drawing of a typical calorimeter cell is shown in Fig. 4.15. The active medium for all cells is liquid argon, maintained at approximately 90 K by cryostats. The EM sections use 3(4) mm depleted uranium plates in the CC(EC) as absorbers. The FH absorbers are 6 mm thick uranium-niobium alloy plates and the CH modules use 46.5 mm thick copper plates. A potential difference is maintained by grounding the absorber plates and connecting the resistive surfaces of the signal boards to positive high voltage (typically 2.0 kV).

As mentioned above, hadronic showers contain an electromagnetic component arising from neutral pion decays. The ratio of signal conversion efficiencies for the electromagnetic to the hadronic shower components is called the intrinsic e/h ratio. If $e/h \approx 1.0$ the calorimeter is said to be *compensating* while if e/h differs from unity by more than 5 – 10%, the detector performance is compromised because of fluctuations in the electromagnetic content of the cascades [13]. A non-compensating calorimeter has a constant contribution to the signal resolution proportional to $|1 - h/e|$. The use of depleted uranium plates as absorbers makes the DØ calorimeter compensating. This is because a portion of the incident particle energy goes into the nuclear binding energy and is usually not detected. With uranium, however, some fraction of this energy results in the production of neutrons which induce fission in ^{238}U producing photons, electrons, and fast neutrons which can all

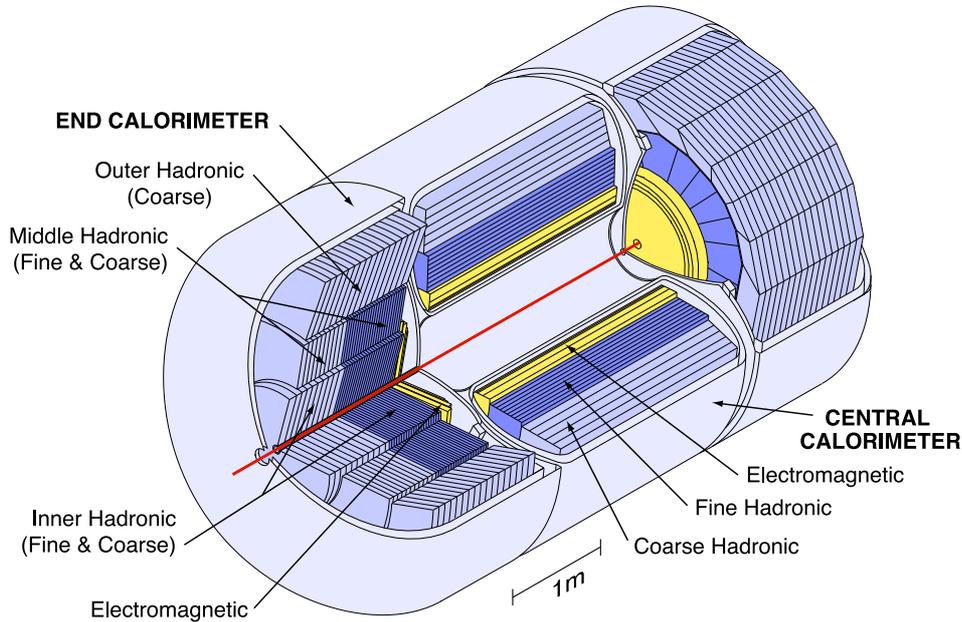


Figure 4.14: Isometric view of the central and two end calorimeters [35].

be detected [40]. The ratio of the electromagnetic to hadronic response is close to one and ranges from 1.11 at 10 GeV to 1.04 at 150 GeV [37].

Each section of the calorimeter is subdivided into a number of layers and segmented in ϕ and z . There are four depth layers for the EM modules, three layers for the FH sections, and one CH layer. As can be seen in Fig. 4.14, the EC hadronic modules are divided into inner, middle, and outer sections. The inner and middle sections are additionally segmented into fine and coarse pieces similar to the segmentation in the CC. Important parameters of the CC are summarized in Table 4.3.

Cells at approximately the same η and ϕ are ganged together to form pseudo-projective towers as shown in Fig. 4.16. Towers in EM and hadronic modules have a transverse size of $\Delta\eta = 0.1$ and $\Delta\phi = 0.1$. The third layer of the EM modules, situated at the shower maximum, is segmented twice as finely in η and ϕ . For layers with $|\eta| > 3.2$ the cell size increases to 0.2×0.2 .

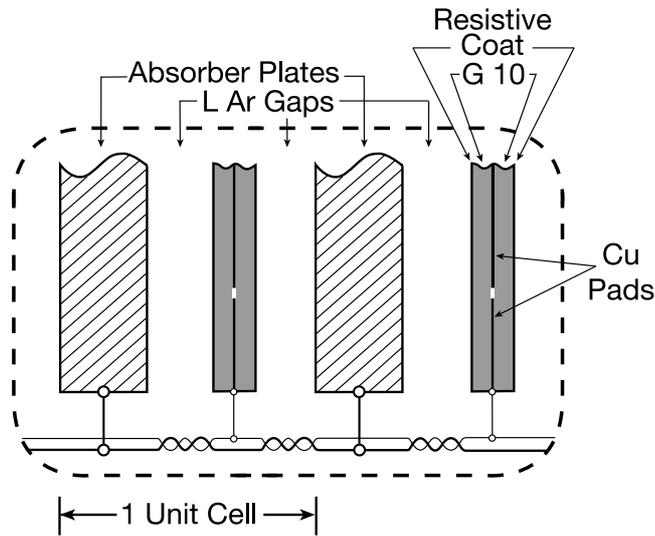


Figure 4.15: Schematic view of a the liquid argon gap and signal board unit cell for the calorimeter [35].

In the region between the CC and EC, there are several gaps where particles travel mostly through support structures such as cryostat walls. Additional layers of sampling have been added to counteract the degradation in energy resolution caused by these insensitive areas. First, there are single-cell calorimeter readout cells in front of the first layer of uranium called *massless 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

Figure 4.17 shows the average energy loss of a muon in hydrogen, iron, and uranium as a function of the muon energy. One can see that ionization processes dominate for muon energies below approximately 200 GeV, which is

Module Type	EM	FH	CH
Rapidity Coverage	± 1.2	± 1.0	± 0.6
Number of Modules	32	16	16
Absorber	DU	DU-Nb	Cu
Absorber Thickness (mm)	3	6	46.5
Argon gap (mm)	2.3	2.3	2.3
Total radiation lengths (X_0)	20.5	96.0	32.9
Total nuclear interaction lengths (λ_I)	0.76	3.2	3.2
Sampling fraction (%)	11.79	6.79	1.45

Table 4.3: Central Calorimeter Parameters [41]. Note that values of X_0 and λ_I are for $\eta = 0$, i.e., normal incidence.

the energy range for the majority of muons produced at TeVatron collisions. Because large energy losses are more probable via radiation than ionization, muons with moderate transverse momentum ($p_T > 2.7$ GeV/c) usually completely traverse the central calorimeter [37]. Therefore, a detector specifically designed to identify muons is situated outside the calorimeter volume.

The DØ Run II muon detector is composed of three main components: scintillators for triggering and cosmic rejection, a toroidal magnet to allow for an independent muon momentum measurement, and drift tubes to measure hit positions. The relative positions of these components is shown in Fig. 4.4. There are actually three toroids used in the muon magnet system—the central iron toroid (CF) and two end iron toroids (EFs). These magnets compose 65% of the 5500 ton weight of the detector [42]. The CF is a 109 cm thick square structure whose inner side is 318 cm from the beamline and is wound with 20 coils of 10 turns each. It covers the region $|\eta| \lesssim 1$ and has an internal field of ≈ 1.8 T. The EF's are 156 cm thick with the surface facing closest to the interaction region at $z = 454$ cm. The EF windings are 8 coils of 8 turns each, producing an internal field of ≈ 1.9 T. There are two regions cut out of the end toroids: a 183 cm square hole in the center to allow for the beam

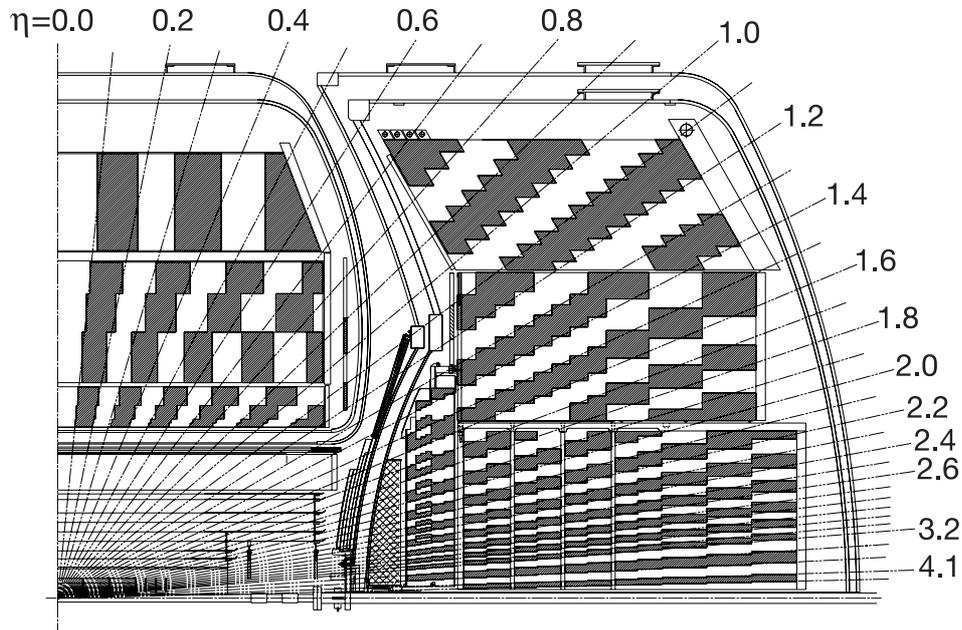


Figure 4.16: A quarter view of the calorimeter showing the segmentation pattern. Groups of cells ganged together for readout are represented by the shading pattern. The rays indicate pseudorapidity intervals from the center of the detector [35].

pipe to pass through and a 30.5 cm circular hole at $x = -33.0$ cm, $y = 206.9$ cm as a remnant from the bypass of the Run I accelerator main ring pipe.

Figures 4.18 and 4.19 show exploded views of the muon system scintillators and drift tubes, respectively. Table 4.4 lists some important parameters for the drift tubes in both the central and forward regions. The drift tubes and counters are organized into three layers: A, B, and C. The A-layer is located in between the calorimeter and the toroid while the B- and C- layers are outside the toroid. We describe the central and forward muon systems separately below.

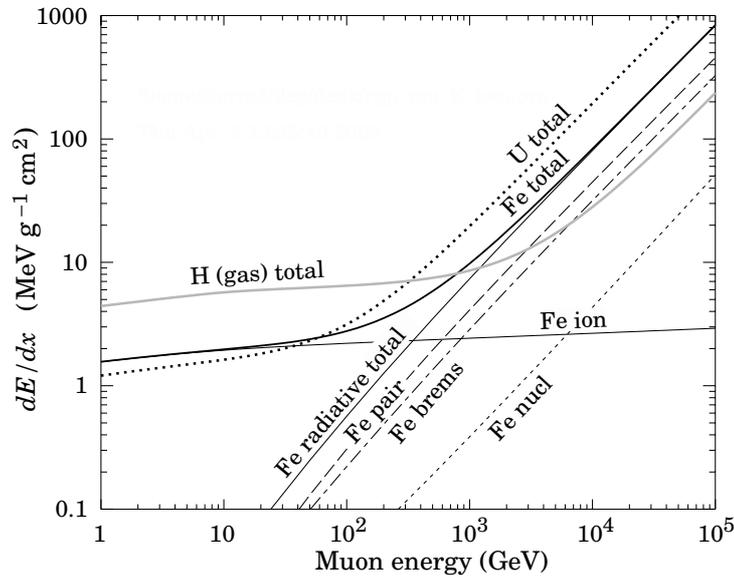


Figure 4.17: The average energy loss of a muon in hydrogen, iron, and uranium as a function of muon energy [13].

Central Muon Detector

The central muon system is composed of $A\phi$ scintillator counters on all sides of the A-layer and the sides and bottom of the B-layer, the CF magnet, proportional drift tube (PDT) chambers in all layers, and the cosmic cap and bottom scintillators in the C-layer. The PDT chambers are formed from three or four decks of aluminum extrusion unit cells as shown in Fig. 4.20. 94 such chambers make up the central muon tracking system, each with a cross section of $2.8 \times 5.6 \text{ m}^2$. The cells which form each chamber consists of a gold plated tungsten anode wire at the center and two cathode pads above and below the anode wire. The wires are ganged together in pairs within a deck and read out by electronics located at one end of each chamber. A gas mixture consisting of 84% argon, 8% CF_4 , and 8% CH_4 is recirculated through the PDT's at a rate of 3 full volume changes per day. A single-wire

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: Muon drift tube parameters [37].

resolution of $\sigma \approx 1$ mm is obtained in the z direction. The difference in arrival time between neighboring cells gives a position measurement in the x and y directions with a resolution of 10 – 50 cm, depending on the location of the hit along the wire.

The cosmic cap scintillators are installed on the outside of the C-layer PDT's while the cosmic bottom scintillators are installed on the outside of the B- and C-layers. There are 240 counters in the cap and 132 in the bottom. They are approximately 4.5° wide in ϕ to match the segmentation of the central track trigger. The $A\phi$ scintillators are located between the calorimeter and the A-layer PDT's. There are 630 $A\phi$ counters, again with a ϕ segmentation of 4.5° .

Forward Muon Detector

The forward muon system covers the region $1.0 \lesssim |\eta| \lesssim 2.0$ and consists of the EF magnets, three layers of mini drift tubes (MDT's), three layers of scintillation counters, and shielding around the beampipe. The MDT system is composed of 6080 mini drift tubes assembled into six layers of eight octants each. Each tube is made from eight drift cells, each having a $9.4 \times$

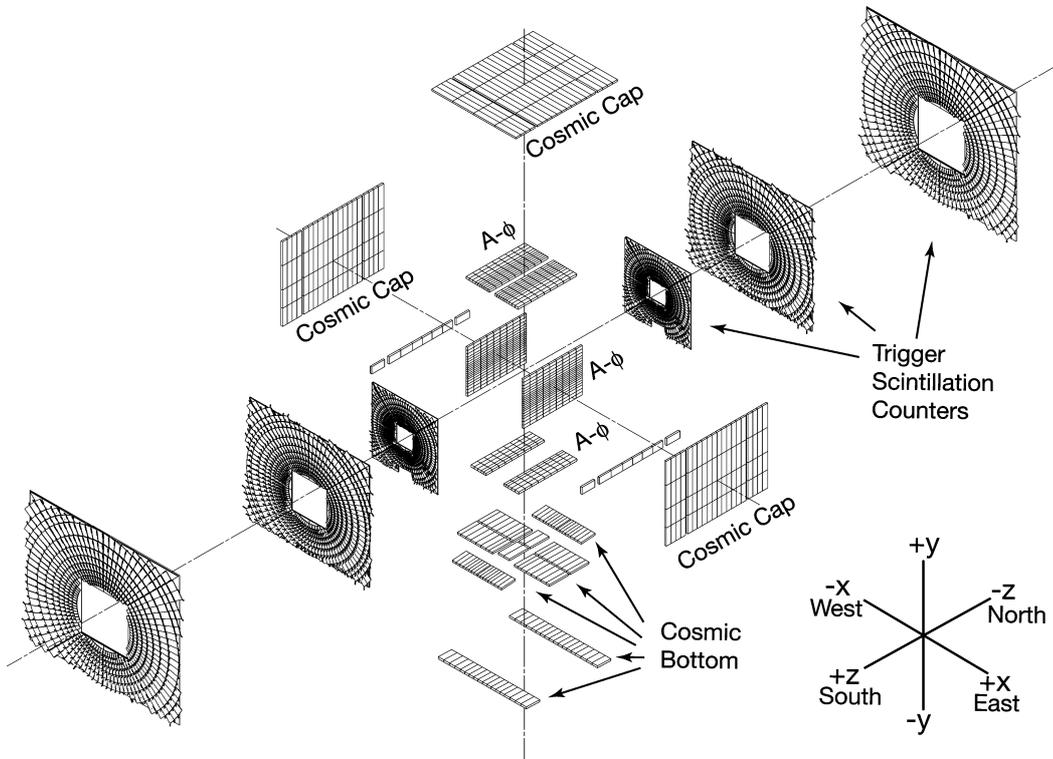


Figure 4.18: Exploded view of the muon scintillation detectors [35].

9.4 mm^2 cross section and a $50 \mu\text{m}$ tungsten-gold wire at the center, as seen in Fig. 4.21. A gas mixture of 90% CF_4 and 10% CH_4 is circulated through the MDT's at a rate of 0.5 volume changes per day. The tubes are oriented with the anode wire parallel to the EF field, and therefore perpendicular to the muon trajectory. The MDT coordinate resolution in the test beam is $\approx 350 \mu\text{m}$, but the large 18.8 ns bin size of the drift time digitization leads to a resolution of $\sigma = (0.8 \pm 0.1) \text{ mm}$, which satisfies the design goal of 1 mm. The forward muon momentum resolution is dominated by the central tracking system for muons with $|\vec{p}| \lesssim 100 \text{ GeV}$, while the forward muon system improves the resolution for higher momentum muons and is very useful for tracks which do not go through all the layers of the CFT, i.e.,

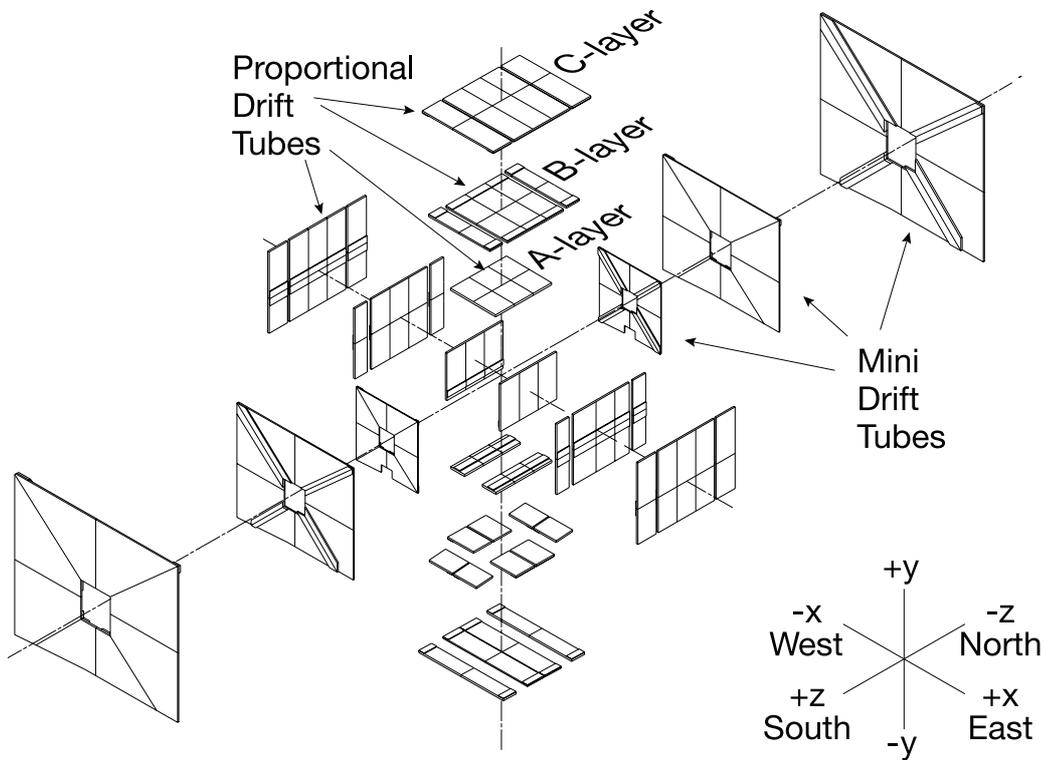


Figure 4.19: Exploded view of the muon wire chambers [35].

those which have $1.6 \lesssim |\eta| \lesssim 2.0$.

A shielding assembly consisting of iron, polyethylene, and lead is located on either side of the beamline. This shielding prevents particles arising from the interaction of beam remnants with the beam pipe, the forward calorimeter, and the accelerator's low-beta quadrupole magnets from traversing the MDT's. The innermost layer of iron absorbs electromagnetic and hadronic showers. Because iron is transparent to slow neutrons, the layer of polyethylene is placed outside the iron to absorb these neutrons. Neutron capture in polyethylene cause emission of gamma rays which are then absorbed by the outside layer of lead.

There are three layers of "pixel" counters placed in front of the corre-

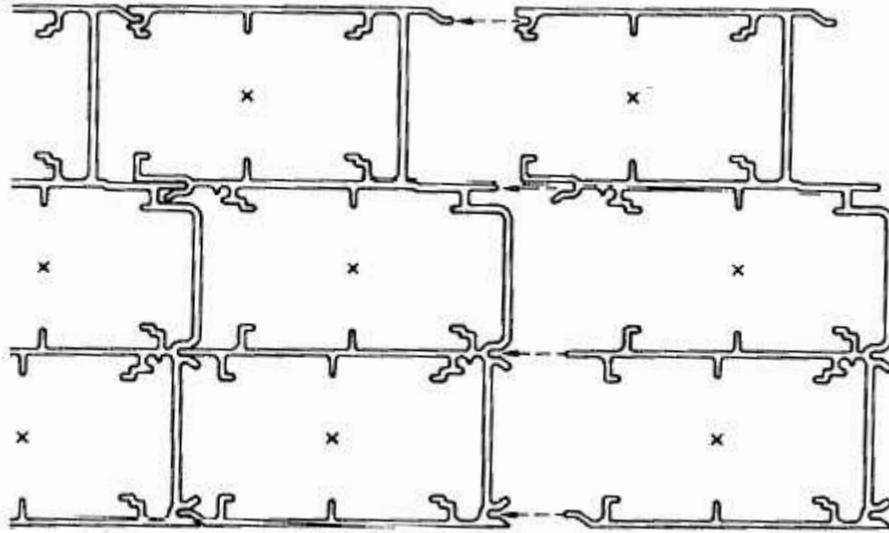


Figure 4.20: Cross-sectional view of three decks of proportional drift tube cells [43]. The “x”’s show the position of anode wires.

sponding layers of MDT’s. In the A-layer, the phototubes attached to the scintillators are placed in mu-metal/iron shields to protect them from the fringe solenoidal and toroidal magnetic fields which can reach up to 300 G. The scintillators have a ϕ segmentation of 4.5° to match the track trigger. There are a total of 4214 counters and their average time resolution is ~ 1 ns. The pixels are useful for reducing backgrounds coming from sources other than the interaction, such as cosmic ray muons.

4.2.5 Luminosity Monitor

The luminosity monitor (LM) determines the luminosity at the DØ interaction region. The LM detector is composed of two arrays of twenty-four plastic scintillator located at $z = \pm 140$ cm as seen in Fig. 4.22. A drawing of a circular array of counters is shown in Fig. 4.23. The counters are

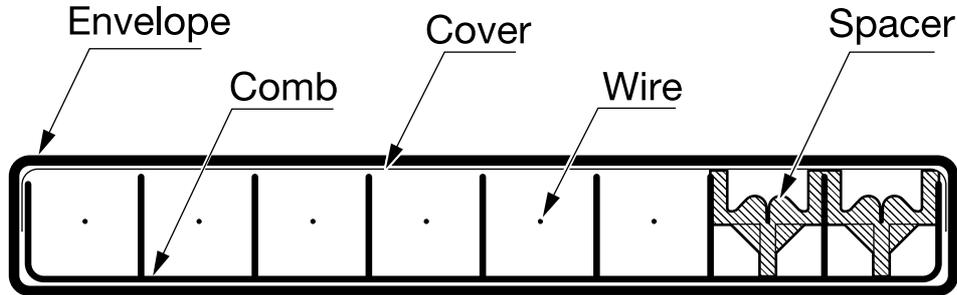


Figure 4.21: Cross-sectional of a mini drift tube [35].

15 cm long and cover the range $2.7 < |\eta| < 4.4$. The arrays are placed in front of the end cap calorimeter and between the beam pipe and the forward preshower. Photomultiplier tubes (PMT's) are mounted on the faces of the scintillator counters with their axes parallel to the z axis. Accurate timing of the PMT signals is achieved by bringing the signals from the detector to the electronics via low-loss cables. Six luminosity monitor time-to-digital converter (LM-TDC) boards digitize the time and charge for each PMT and apply calibration corrections to create time-of-flight measurements. The calibration procedure is further described in Appendix A. A single luminosity monitor vertex (LM-VTX) board takes the measurements coming from the LM-TDC boards and determines the z coordinate of the interaction vertex.

The luminosity \mathcal{L} is determined from the average number of inelastic collisions per beam crossing \bar{N}_{LM} measured by the LM,

$$\mathcal{L} = \frac{f \bar{N}_{LM}}{\sigma_{LM}}, \quad (4.3)$$

where f is the beam crossing frequency and σ_{LM} is the cross section for the LM that takes into account the acceptance and efficiency of the LM detector [44]. The number of inelastic collisions is done by counting the number of crossings with no collisions and using Poisson statistics to determine \bar{N}_{LM} .

The number of collisions is determined by making time-of-flight measure-

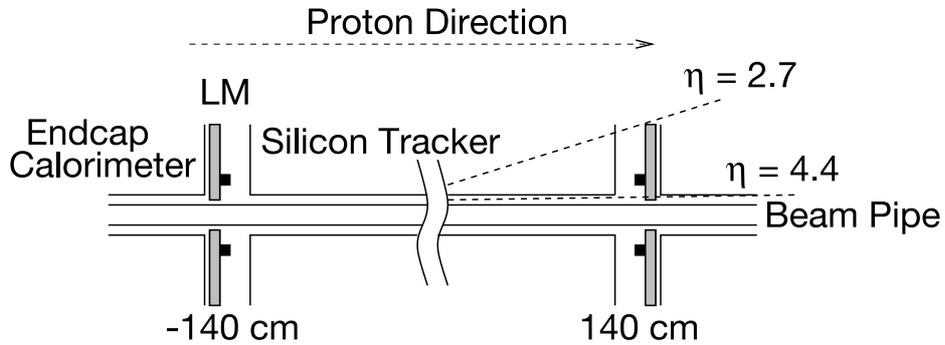


Figure 4.22: Schematic drawing showing the location of the luminosity monitor [35]

ments of particles travelling at small angles with respect to the beam line. The z coordinate of the interaction vertex z_v is estimated from the difference in the hit times of the two ends of the LM detector,

$$z_v = \frac{c}{2}(t_- - t_+). \quad (4.4)$$

The longitudinal width of the interaction region at the DØ is $\sigma_z \approx 30$ cm. Therefore, inelastic collisions are selected by requiring $|z_v| < 100$ cm. Halo particles travelling along with the beam will have $z_v \approx \pm 140$ cm, corresponding to the location of the LM detectors, and are rejected by the $|z_v| < 100$ cm cut.

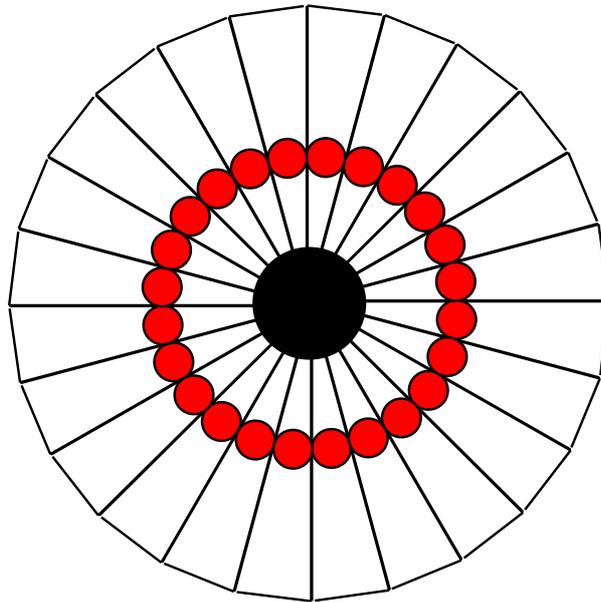


Figure 4.23: Schematic drawing showing the geometry of the LM counters and the location of the PMT's (solid dots) [35].

4.3 The DØ Trigger System

The rate of collisions at the Tevatron, ≈ 2.5 MHz, is too fast to record every event to tape for offline analysis. DØ uses a three stage *triggering* system to store events that have interesting physics signatures. The first stage, Level 1 or L1, is comprised of hardware trigger elements that operate at a rate of 2 kHz. The second stage, Level 2 or L2, utilizes hardware with embedded microprocessors to construct trigger decisions using individual physics objects as well as correlations between objects at a rate of 1 kHz. The final stage, Level 3 or L3, uses a farm of CPU's running more sophisticated algorithms to reduce the rate to 50 Hz which is the rate that events are recorded for offline reconstruction. Note that the Tevatron beam structure is split into super-bunches and bunches within the super-bunches, as shown in Fig. 4.3. There are $2 \mu\text{s}$ between super-bunches and 396 ns between bunches inside the super-bunch. This reduces the average rate of data coming in from the detector from 2.5 MHz to ~ 1.7 MHz.

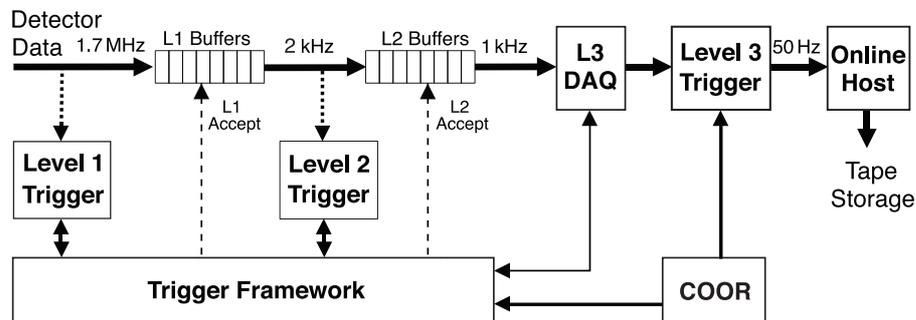


Figure 4.24: Overview of the DØ trigger and data acquisition systems [35].

Figure 4.24 shows an overview of the DØ trigger and data acquisition system. After the beam crossing is read out and L1 processing is complete, the L1 trigger elements report their results to the trigger framework (TFW). The L1 system supports up to 128 specific triggers or trigger bits, the “OR”

of which determines whether a given crossing has a valid trigger. The logic that determines whether a particular bit is set is built into the L1 hardware utilizing a series of field-programmable gate-arrays (FPGA's). L1 trigger decisions must arrive at the TFW within $3.5 \mu\text{s}$ to participate in the global trigger decision for a particular beam crossing.

When the TFW issues an L1 accept, it sends a command to all detector elements to digitize the data and place it into a series of 16 L1 event buffers to prepare for L2 processing. The L2 preprocessor system is composed of FPGA's as well as microprocessors. These preprocessors take their inputs from detector front-ends as well as the L1 system and send their results to a global L2 processor, L2Global, which examines the incoming physics objects as well as correlations between them to make a L2 decision. This system reduces the rate by a factor of 2; it has approximately $100 \mu\text{s}$ to either accept or reject an event. If an L2 accept is issued, event data is transferred to the L2 buffers where they await transfer to the L3 system. The L1 and L2 buffers minimize the experiment's deadtime by providing FIFO storage for event data. A block diagram of the L1 and L2 system data flow is shown in Fig. 4.25.

Events accepted by L2 are sent to the L3 system to be processed by a farm of L3 microprocessors. These microprocessors are standard desktop computers running the Linux operating system; they perform a limited reconstruction of the entire event and reduce the nominal 1 kHz input rate to 50 Hz for transfer to the offline storage system. Overall coordination and run control for the entire trigger system is provided by a program called "COOR" running on an online host machine. The trigger system also supports *prescaling* of trigger conditions at all three levels of processing. Prescaled triggers are accepted by the system only a fraction of the time t that their trigger condition is satisfied. For example, if a certain trigger has an L1 prescale of 2 only $\frac{1}{2}$ of the events that satisfy the trigger condition will be accepted. Triggers are prescaled depending the instantaneous luminosity, allowing data to keep

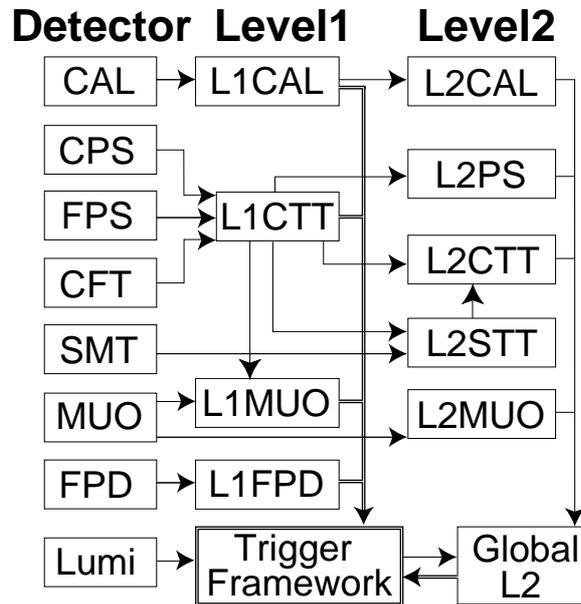


Figure 4.25: Block diagram of the DØ L1 and L2 trigger systems. The arrows show the flow of trigger-related data [35].

flowing even during high luminosity beam conditions.

4.3.1 Level 1 Trigger

The Level 1 trigger uses detector information from all subsystems except for the SMT. Due to bandwidth constraints, SMT information bypasses the L1 system and is sent directly to Level 2. There are three different L1 trigger systems,

- the Level 1 calorimeter trigger,
- the Level 1 central track trigger, and
- the Level 1 muon trigger.

The three triggers are described in greater detail below.

Level 1 Calorimeter Trigger

The Level 1 calorimeter trigger (L1CAL) looks for energy deposition patterns in trigger towers exceeding programmed limits on transverse energy deposits. The standard size of trigger tower is $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$. There are 12 electromagnetic towers and 1280 hadronic towers. Tower energies are converted to missing transverse energy, E_T , on input to L1CAL and are pedestal subtracted and adjusted for energy scaling when necessary. Triggers may be formed from the sum of all transverse energy $\sum E_T$, the missing transverse energy \cancel{E}_T , and the E_T in a localized section of the calorimeter.

Level 1 Central Track Trigger

The Level 1 central track trigger (L1CTT) makes a fast reconstruction of tracks using information from the CFT, the CPS, and the FPS. The L1CTT is optimized to make fast trigger decisions within the $3.5 \mu\text{s}$ L1 decision time, but it also stores more detailed data to be used as an input to L2. The track-finding algorithms of the L1CTT operate on 80 4.5° sectors of the axial CFT and CPS layers. An FPGA compares the pattern of fiber hits to a set of 20,000 predefined track equations and then sends out a list of the six highest- p_T tracks it finds. This list is an input to another FPGA where tracks are matched to axial CPS clusters, the number of tracks is counted, and the total p_T is calculated. Figure 4.26 shows an illustration of a single 4.5° sector with a hypothetical track overlaid across all eight CFT axial doublet layers and the CPS axial layer. The L1CTT equations require hits in all eight CFT axial layers. At this stage, lists of tracks are passed to the Level 1 muon system as well as the next tier of L1CTT processing. This last tier collects and sorts data with an octant (ten sectors) and generates trigger term bits, such as requiring the existence of a track with $p_T > 5 \text{ GeV}/c$, to the TFW. If L1CTT issues an accept, the L1 track candidates are sent as inputs to the Level 2 silicon track trigger.

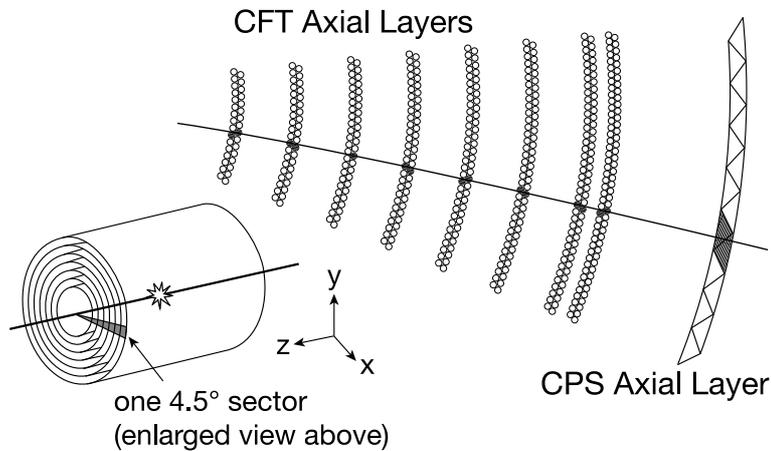


Figure 4.26: Transverse schematic view of a single 4.5° sector. A hypothetical track is overlaid on the eight CFT axial doublet layers and CPS axial layer [35].

Level 1 Muon Trigger

The Level 1 muon trigger (L1MUO) combines information from muon wire chambers, muon scintillation counters, and tracks from L1CTT to form muon objects. Track centroids, also called “stubs”, are formed by matching wire hits to corresponding scintillator hits. In parallel, L1CTT tracks are matched to scintillator hits using FPGA’s to perform combinatorial logic. These two objects, wire/scintillator centroids and L1CTT/scintillator candidates, are then combined before a global muon decision is made. To reduce the contribution from cosmic rays, high- p_T candidates are required to pass cosmic ray scintillator timing vetoes.

4.3.2 Level 2 Trigger

The Level 2 system is composed of an array of detector-specific pre-processing engines as well as a global processor (L2Global) that looks for correlations in physics signatures across subsystems. L2 processors collect

data from detector readout as well as the L1 system to form physics objects. There are six detector-specific processors at L2,

- the Level 2 silicon track trigger (L2STT),
- the Level 2 central track trigger (L2CTT),
- the Level 2 central preshower trigger (L2CPS),
- the Level 2 forward preshower trigger (L2FPS),
- the Level 2 muon trigger (L2Muon), and
- the Level 2 calorimeter trigger (L2CAL).

L2Global is responsible for making trigger decisions based on the objects identified by the processors listed above. Objects can be made directly from the output of those processors or from the combination of the objects of different processors. The individual detector pre-processors are discussed in more detail below.

L2STT

The Level 2 silicon track trigger performs online reconstruction of tracks found in the CFT by utilizing the much finer spatial resolution of the SMT. The L2STT improves the momentum measurement of tracks at Level 2, allows for a precise measurement of the impact parameter of tracks, and helps reject backgrounds due to accidental track patterns in the CFT.

Figure 4.27 shows the basic conceptual principle of the L2STT. L1CTT sends a list of tracks to L2STT for every event. In the L2STT, a ± 2 mm road is defined around each L1CTT track and the SMT hits within that road are associated with the track [37]. The L2STT uses axial hits in the silicon ladders, hits in the innermost and outermost layers of the CFT, and hits in at least three of the four layer of the SMT to fit track parameters. The results of the track fit are sent to L2CTT as well as Level 3.

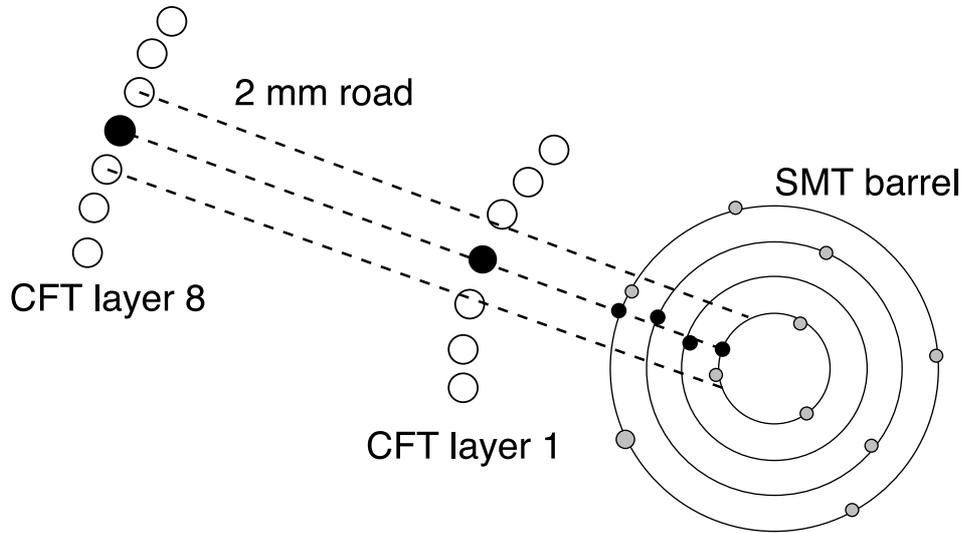


Figure 4.27: The definition of roads based on L1 tracks and SMT hits in L2STT [35]

The SMT ladders are arranged in twelve 30° azimuthal sectors. There is a slight overlap in sectors such that greater than 98% of tracks are contained within a single sector. Each 30° sector is therefore treated independently by L2STT with negligible loss.

L2CTT

The L2CTT preprocessor receives inputs from L1CTT and L2STT, as mentioned above. The system works in two different modes of operation: *a)* with track inputs directly from L1CTT and *b)* with track inputs from L2STT. In the first mode, lists of L1CTT tracks from different ϕ sections are sent to L2STT which then combines them into a single list ordered by track p_T . The p_T measurement is then improved by using additional hit and tracking information that is unavailable at L1. Furthermore, the value of the azimuthal angle with respect to the beamline ϕ_0 , the azimuthal angle with respect to the third layer of the EM calorimeter ϕ_{em3} , and the isolation are

calculated in the L2CTT. The p_T -sorted list is then reported to L2Global.

In the second mode of operation, lists of L2STT tracks as well as improved L2 track p_T 's are sent as inputs. ϕ_0 , ϕ_{em3} , and the isolation are then calculated for these tracks. Two sorted lists are then passed to L2Global, one ordered by p_T and the other by impact parameter.

L2PS

At Level 2 data from the CPS and FPS are processed independently. However, the principles of the L2CPS and L2FPS are similar. Stereo clusters are sent directly to L2 while axial clusters are combined into azimuthal quadrant before being transferred to the L2PS. Cluster centroids are compared to produce η and ϕ coordinates for clusters that match in all three layers. The presence or absence of CFT trigger tracks is also provided at this level of processing. This enables the flagging of output clusters as electrons (associated to a track) or photons (no track). The η and ϕ coordinates are binned to correspond to the calorimeter trigger tower geometry of $\eta \times \phi = 0.2 \times 0.2$ to facilitate the matching of preshower hits to calorimeter objects. This information is then passed to L2Global.

L2Muon

L2Muon improves muon identification by using calibration and more precise timing information. Track segments are searched for in small regions of the detector independently so that the total execution time of the algorithm is independent of the number of hits. Integrated muon candidates constructed from these small track segments are formed by L2Muon and sent to L2Global.

L2CAL

The L2CAL preprocessor performs jet and electron/photon identification and calculates global event \cancel{E}_T . The jet algorithm clusters 5×5 groups

of calorimeter trigger towers centered on seed towers, which themselves are E_T -ordered with $E_T > 2$ GeV. The electron/photon algorithm creates an E_T -ordered list of EM towers with $E_T > 1$ GeV. For each seed tower, the neighboring tower with the largest E_T is combined with the seed to create an EM cluster. The L2CAL \cancel{E}_T algorithm calculates the vector sum \cancel{E}_T from individual trigger tower E_T received from L1. A list of jet and electron/photon candidates as well as \cancel{E}_T is then sent to L2Global.

4.3.3 Level 3 Trigger and Data Acquisition

The Level 3 trigger is a fully programmable software trigger composed which performs a fast reconstruction of events. The final trigger decision is made based on complete physics objects (such as electrons, muons, and jets) as well as the relationship between objects (such as the azimuthal angle separating objects or their combined invariant mass).

When an L2 accept is issued, data is transferred out of readout crates by a single board computer (SBC) situated in each crate. All SBC's transfer their data via a single Cisco 6509 gigabit ethernet switch [45] to farm nodes specified by routing instructions received from the routing master (RM) process. The RM is a dedicated routing process which executes on an SBC in a special VME crate containing an interface to the TFW. The TFW provides trigger information to the RM and allows it to asynchronously disable the firing of L1 triggers if the L3 farm cannot keep up with the rate of incoming data.

The farm nodes run two different programs: an event builder (EVB) process and an event filter process. The EVB builds a complete event from the event fragments received by the SBC's. For each event, the RM sends a list of expected crates to the EVB of a particular node so that the EVB will know when an event is complete. If some SBC fails to deliver a fragment that the EVB expects, the event is discarded. Completed events are placed into shared memory buffers for processing by the event filters. The EVB routinely reports the number of free buffers it has to the RM so the RM will

be able to efficiently route events to nodes that have the capacity to process them.

The second program on each node runs the fast reconstruction and event filtering processes. *Physics tool* algorithms create the physics objects and relations between objects by unpacking the raw data, applying calibrations, locating hits and clusters, and reconstructing the objects. Calls to the tools are made by *filter scripts* that define selection criteria used by the tools or imposed on their results. These filter scripts specify a set of reference parameters (referred to as a *refset*) to be used by each tool. These refsets define physics objects precisely for use by tools; the jet refset, for example, specifies the size of the jet cone as well as other jet parameters. An event passes an L3 trigger if all the filters for any of the filter scripts pass. Accepted events are then sent to the online system to be stored on tape drives.

A supervisor process running on a separate SBC interfaces between the DØ run control (COOR) and the Level 3 data acquisition system (L3DAQ). When a run is configured, the supervisor passes run and trigger information to the RM and the L3 filter configuration to the EVB processes on the relevant farm nodes. Figure 4.28 shows the flow of data through the L3DAQ system. The system is designed for a bandwidth of 250 MB/s, corresponding to an average event size of ~ 200 kB at an L2 accept rate of 1 kHz.

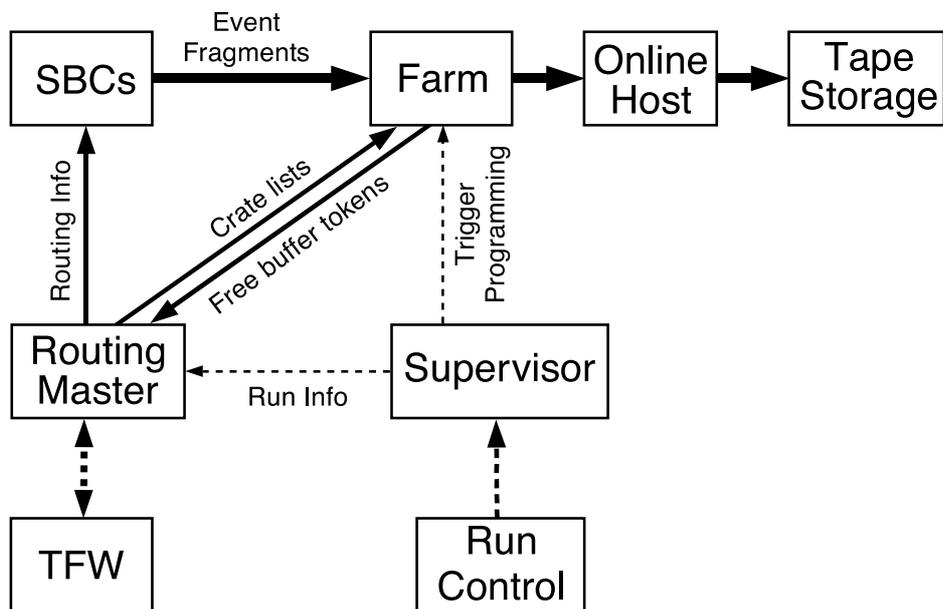


Figure 4.28: Schematic illustration of the data flow through the L3DAQ system [35].

Chapter 5

Multidimensional Opposite-Side Flavor Tagging

A crucial component of any mixing analysis is the initial state flavor tagger. The first analysis presented here is a measurement of Δm_d in the semileptonic decays $B \rightarrow \mu^+ \bar{D}^0 X$ and $B \rightarrow \mu^+ D^{*-} X$. B^+ decays give the main contribution in the first sample, and B^0 decays dominate in the second sample.

The measurement of Δm_d allows us to determine the performance of the initial state flavor tagger. Initial state flavor tagging is the determination of the flavor of the b quark at the time of the production of a reconstructed B meson. It is a crucial component of any B_s^0 mixing analysis, as described in Ch. 3. We use properties of the b quark opposite to the one from which the reconstructed meson is created, called opposite-side flavor tagging. This type of flavor tagger is constructed to be independent of the species of reconstructed B meson (i.e., B^+ , B^0 , B_s^0 ...), which is why we determine its performance on B^0 and B^+ mesons before using it in the B_s^0 analysis.

The flavor tagging dilution is determined independently for reconstructed B^+ and B^0 events. This technique allows us to verify the assumption of independence of the the opposite-side flavor tagging on the type of reconstructed

B meson.

Two separate flavor taggers are employed for the Δm_d analysis: a multidimensional likelihood flavor tagger based on Monte Carlo simulations and a combined one-dimensional likelihood flavor tagger based on data. The use of two taggers formed from different sources enables us to cross-check their results against each other and to try to gain improvement in tagging. In this chapter, we describe the principles and construction of the Monte Carlo flavor tagger in detail. The data-based tagger performs marginally better, so it was used in the Δm_s analysis in Ch. 6 where it is briefly described. The Monte Carlo tagger, however, has the advantage of taking correlations between discriminants into account. We discuss the benefits of using discriminant correlations in flavor tagging in Sec. 5.2.

5.1 Likelihood Based Flavor Tagging

In this section we use a likelihood approach to derive formulas similar to Eq. 3.18. We make a slight change of notation here for the sake of simplicity: the probability for a meson to not mix is written \mathcal{P}_- and the probability to mix is written \mathcal{P}_+ . Let $P(B^0)$ be the probability that the tagged meson is a B^0 at $t = 0$ and $P(\bar{B}^0)$ be the probability the tagged meson is a \bar{B}^0 at $t = 0$. We can then write the observed oscillation probability as

$$\mathcal{P}_{\pm}^{obs}(\xi_{B^0}, t) = P(\bar{B}^0)\mathcal{P}_{\pm\xi_{B^0}}(t) + P(B^0)\mathcal{P}_{\mp\xi_{B^0}}(t) \quad (5.1)$$

where ξ_{B^0} denotes the flavor of the meson at $t = 0$. We choose the convention $\xi_{B^0} = \pm 1$ when the meson flavor at production is $\bar{B}^0(B^0)$. Expressing the probabilities as ratios of likelihoods,

$$P(B^0) = \frac{\mathcal{L}(B^0)}{\mathcal{L}(B^0) + \mathcal{L}(\bar{B}^0)}, \quad P(\bar{B}^0) = \frac{\mathcal{L}(\bar{B}^0)}{\mathcal{L}(B^0) + \mathcal{L}(\bar{B}^0)} \quad (5.2)$$

we can rewrite Eq. 5.1 as

$$\mathcal{P}_{\pm}^{obs}(\xi_{B^0}, t) = \frac{1}{2}\Gamma e^{-\Gamma t} [1 \pm d\xi_{B^0} \cos(\Delta m dt)], \quad (5.3)$$

where the *tagger output*, d , is defined as,

$$d = \frac{\mathcal{L}(\bar{B}^0) - \mathcal{L}(B^0)}{\mathcal{L}(\bar{B}^0) + \mathcal{L}(B^0)}. \quad (5.4)$$

The statistical error for a given single event, with index i , due to imperfect tagging will then be given by,

$$\sigma_i^2 \propto \frac{1}{\epsilon \langle d^2 \rangle}, \quad (5.5)$$

where ϵ is the efficiency of the tagger, $N_{\text{tagged}}/N_{\text{reconstructed}}$. The average variance for N tagged events is then,

$$\langle \sigma_i^2 \rangle \propto \frac{1}{N\epsilon \langle d^2 \rangle}. \quad (5.6)$$

We define the *effective efficiency*, ϵ_{eff} , as

$$\epsilon_{\text{eff}} = \epsilon \langle d^2 \rangle, \quad (5.7)$$

and can therefore identify the analogue of the dilution, \mathcal{D} , for likelihood-based flavor tagging as

$$\mathcal{D} \approx \sqrt{\langle d^2 \rangle}. \quad (5.8)$$

5.2 Multidimensional Likelihoods

The flavor likelihoods, $\mathcal{L}(B^0)$ and $\mathcal{L}(\bar{B}^0)$, in Eq. 5.2 can, in principle, be functions of more than one variable. If they are, we refer to them as *multidimensional* and write them as $\mathcal{L}(B^0, \vec{x})$ and $\mathcal{L}(\bar{B}^0, \vec{x})$, where \vec{x} is a vec-

tor of discriminating variables. Multidimensional likelihoods are superior to one-dimensional likelihoods in the following respects: tagging performance is improved because correlations between discriminants are taken into account, and the tagger output is more linear (i.e., $\mathcal{D}(d)$, where \mathcal{D} is the measured dilution, is more linear). Here we show how correlations between discriminants can be used to improve tagging performance. The linearity of the flavor tagging output is discussed in Sec. 5.7.6 where the results of the Δm_d oscillation analysis are presented.

Figure 5.1(a) shows distributions of qp_T for opposite-side muons from Monte Carlo simulations of $B^+ \rightarrow J/\psi K^+$, and Fig. 5.1(b) shows distributions of the impact parameter significance, $b/\sigma(b)$, for the same Monte Carlo sample. Muons arising from b quark hadronization have an asymmetric distribution in qp_T , while muons arising from other source have a symmetric distribution. This is expected because $b \rightarrow c \rightarrow \mu$ transitions will cause the muon to have the same charge sign as the b quark and have a softer momentum spectrum than muons coming directly from b 's, $b \rightarrow \mu$. At the same time, muons coming from b hadronization have longer tails in their impact parameter significance distributions than those from other sources. Therefore, by using the correlation between these variables, we can improve the modelling of the likelihood of opposite-side muons and thereby increase the performance of the flavor tagging algorithm. We further explore correlations between discriminants in Sec. 5.5.

5.3 Monte Carlo Samples and Selections

The likelihoods of flavor at creation in Eq. 5.4 can in general be functions of kinematic variables \vec{x} which discriminate between B^0 and \bar{B}^0 . We therefore write the likelihoods as $\mathcal{L}(B^0, \vec{x})$ and $\mathcal{L}(\bar{B}^0, \vec{x})$.

We obtain our likelihoods from Monte Carlo samples of $B^\pm \rightarrow J/\psi K^\pm$ with the J/ψ decaying to $\mu^+\mu^-$. This final state does not oscillate and is

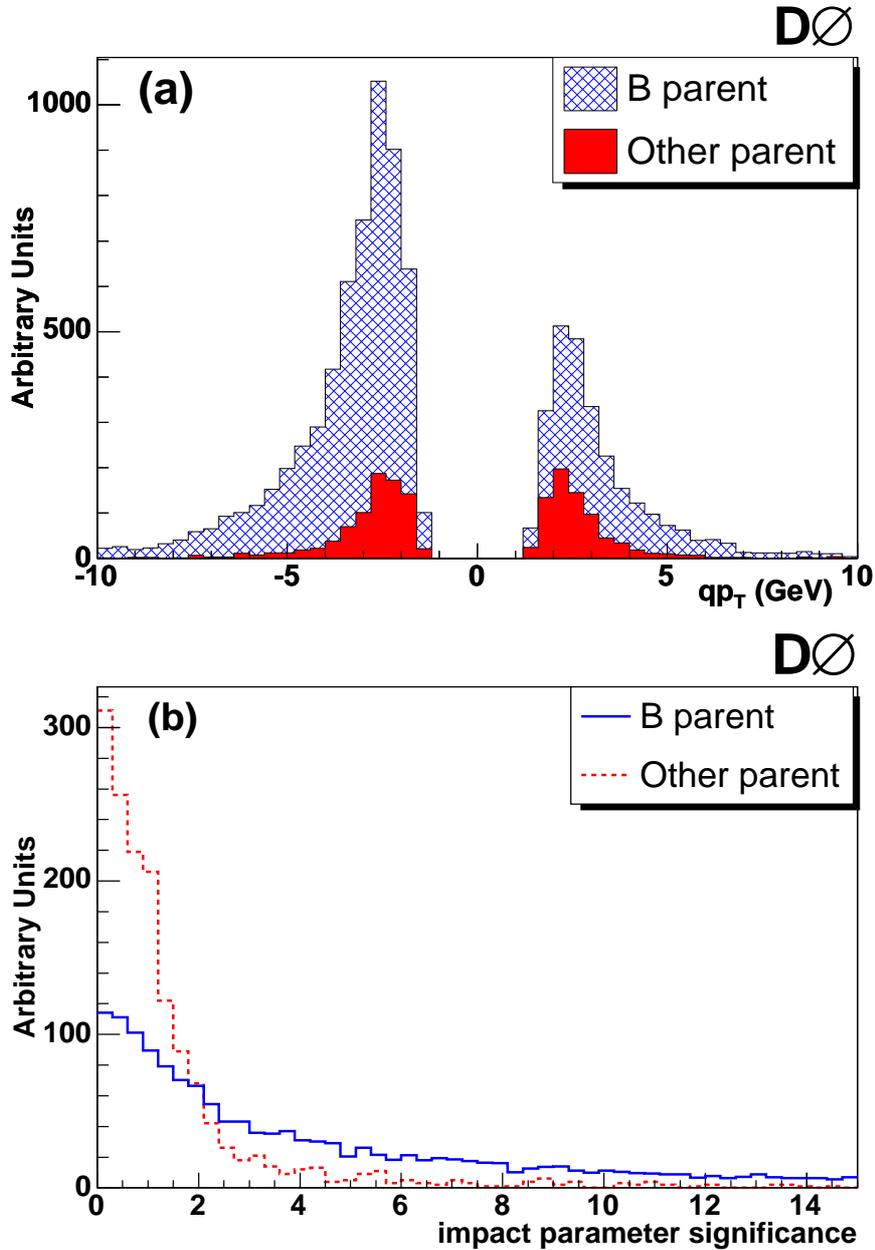


Figure 5.1: (a) qp_T distributions and (b) impact parameter significance, $b/\sigma(b)$, distributions of opposite-side muons from Monte Carlo simulations of $B^+ \rightarrow J/\psi K^+$. Muons are classified in two categories depending on whether or not they are the products of b quark hadronization. The qp_T distribution is asymmetric for muons coming from b quarks and symmetric otherwise. The impact parameter distribution of muons coming from b quarks has a longer tail than that of muons from other sources.

therefore flavor pure. The Monte Carlo sample $B^+ \rightarrow J/\psi K^+$ is used to create $\mathcal{L}(B^0)$, and $B^- \rightarrow J/\psi K^-$ is used to create $\mathcal{L}(\bar{B}^0)$. The selections for reconstruction are the same as in [46] and are as follows:

- K^\pm
 - Number of SMT hits > 1 .
 - Tracks are rejected if they can be identified as products of K_S^0 decays, Λ^0 decays, or photon conversions.
 - $p_T(K^\pm) > 0.5$ GeV.
 - $|\vec{p}| > 0.7$ GeV.
 - If $p_T(K^\pm) < 1.0$ GeV the kaon is required to be in the same jet as the J/ψ . Jets are constructed using the Durham algorithm [31].
 - The 3-D impact parameter significance of the kaon relative to the primary vertex > 3 .
- J/ψ
 - $p_T(J/\psi) > 5$ GeV.
 - $2.80 < m(J/\psi) < 3.34$ GeV.
 - The candidate mass is constrained to the J/ψ nominal value 3.09687 GeV [13].
- B^\pm
 - 2 of the 3 tracks must have at least 2 hits in the SMT
 - χ^2 of the three track vertex < 16 ; if $p_T(K) < 1.0$ GeV/c then this $\chi^2 < 9$.
 - The transverse decay length significance, $L_{xy}/\sigma_L > 4.5$; if $p_T(K) < 1.0$ GeV/c then this is tightened to 5.5.

- $\cos(\alpha) > 0.9$ where α is the angle between the B momentum and the vector from the primary to the B vertex.

In practice, the likelihoods are histograms that have one dimension per discriminant whose bin contents have been normalized to the total number of events in the sample. For a given event, the tagger output d is obtained by substituting the appropriate normalized bin contents into Eq. 5.4. We discuss the discriminating variables used to construct the multidimensional likelihoods, $\mathcal{L}(B^0, \vec{x})$ and $\mathcal{L}(\bar{B}^0, \vec{x})$, from the samples of $B^\pm \rightarrow J\psi K^\pm$ in the following sections.

5.4 Discriminating Variables

In each analyzed event, we search for an additional muon. This muon was required to be classified as “loose” by the standard DØ muon identification algorithm [47], to have at least one hit in the muon chambers, and to have $\cos\phi(\mathbf{p}_\mu, \mathbf{p}_B) < 0.8$, where \mathbf{p}_B is the three-momentum of the reconstructed B meson. If more than one muon was found, the muon with the highest number of hits in the muon chambers was used. If more than one muon with the same number of hits in the muon chambers was found, the muon with the highest transverse momentum p_T was used. For this muon, a *muon jet charge* Q_J^μ was constructed as:

$$Q_J^\mu = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}. \quad (5.9)$$

The sum was taken over all charged particles, including the muon, satisfying the condition $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$. $\Delta\phi$ and $\Delta\eta$ were computed with respect to the muon direction. Daughters of the reconstructed B meson were explicitly excluded from the sum. In addition, any charged particle with $\cos\phi(\mathbf{p}, \mathbf{p}_B) > 0.8$ was excluded. We use this p_T -weighted cone charge to estimate the charge associated with the hadronization of a postulated b

or \bar{b} quark. This is helpful to account for processes such as $b \rightarrow c \rightarrow \mu^+ \bar{\nu}$ where the muon is the “wrong” charge and is also soft. This soft muon would be weighted less in the muon jet charge than harder tracks that hadronized directly from the b quark.

We use two other kinematic properties of opposite-side muons: the momentum projected along the direction perpendicular to the jet axis momentum, $p_T^{rel} = |\vec{p}| \sin \theta_{jet}$, and the $R\phi$ impact parameter significance, $S_{R\phi} = d_{R\phi}/\sigma_{R\phi}$, using the notation of Sec. 3.1.1. Muons coming from B hadron decays are expected to have larger momentum transverse to the B momentum, approximated by the jet axis momentum vector. The usefulness of the transverse ($R\phi$) impact parameter significance as a correlated flavor discriminant is described in Sec. 5.2.

An additional identified electron [48] was used for the flavor tagging if $\cos \phi(\mathbf{p}_e, \mathbf{p}_B) < 0.8$. For this electron, an *electron jet charge* Q_J^e was constructed as:

$$Q_J^e = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}. \quad (5.10)$$

The sum was taken over all charged particles, including the electron, with $\Delta R < 0.5$, as above.

A secondary vertex corresponding to the decay of B mesons was searched for using all charged particles in the event. The secondary vertex was required to contain at least 2 particles with axial impact parameter significance greater than 3. The distance l_{xy} from the primary to the secondary vertex should satisfy the condition: $l_{xy} > 4\sigma(l_{xy})$. The details of the secondary vertex search are give in Sec. 3.1.2.

The momentum of the secondary vertex \mathbf{p}_{SV} was defined as the sum of all momenta of particles included in the secondary vertex. Secondary vertices used for flavor tagging were required to satisfy $\cos \phi(\mathbf{p}_{SV}, \mathbf{p}_B) < 0.8$.

A *secondary vertex charge* Q_{SV} was defined as the third discriminating

variable:

$$Q_{SV} = \frac{\sum_i q^i (p_L^i)^\kappa}{\sum_i (p_L^i)^\kappa}. \quad (5.11)$$

where the sum was taken over all particles included in the secondary vertex. Daughters of the reconstructed B meson were explicitly excluded from the sum. In addition, any charged particle with $\cos \phi(\mathbf{p}, \mathbf{p}_B) > 0.8$ was excluded. p_L^i is the longitudinal momentum of a given particle with respect to the direction of the secondary vertex momentum. Secondary vertex charge works as a flavor discriminant because of the long lifetime of B hadrons. Tracks coming from a true separated B vertex should be boosted in the direction of the B momentum, approximated by the vector sum of the track momenta associated with the given secondary vertex. Thus, we use the track momenta projected along the direction of the secondary vertex as a weight factor to reduce the contribution from spurious tracks wrongly associated to the secondary vertex. The exponential factor, κ , is an ad-hoc parameter obtained from studying the dependence of dilution on Q_{SV} , as seen in Fig. 5.2. Values of κ less than 1 give more weight to softer tracks than when $\kappa = 1$. Figure 5.2 indicates that these softer tracks contain flavor information that is useful to improve the tagging performance. We therefore use $\kappa = 0.6$ in the construction of the secondary vertex charge.

The *transverse momentum of the secondary vertex* p_T^{SV} was also used as the discriminating variable. Events with fake vertices are not sensitive to the charge of the B meson on the reconstructed side. Their contribution decreases the tagging purity. Usually, they are constructed from the low momentum tracks and their p_T^{SV} is softer. Furthermore, p_T^{SV} can discriminate between charm and bottom secondary vertices because the larger mass of B hadrons causes the SV to have larger p_T .

Finally, the *event charge* Q_{EV} was constructed as:

$$Q_{EV} = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}. \quad (5.12)$$

The sum was taken over all charged particles with $0.5 < p_T < 50 \text{ GeV}/c$ and having $\cos\phi(\mathbf{p}, \mathbf{p}_B) < 0.8$. Daughters of the reconstructed B meson were explicitly excluded from the sum.

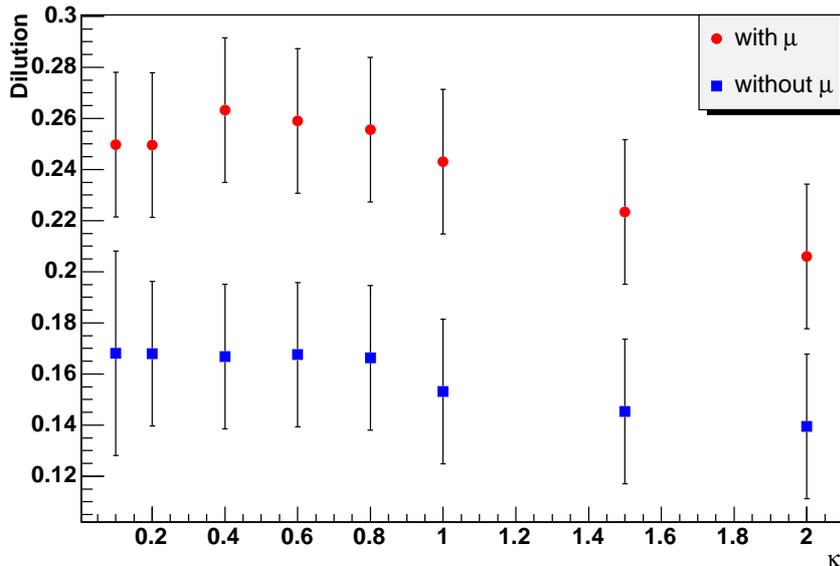


Figure 5.2: Dilution of events tagged by Q_{SV} versus the coefficient κ . Statistical errors shown are correlated.

5.5 Likelihoods, Binnings, and Tagging Logic

In principle, one could combine all of the discriminants mentioned in Sec. 5.4 into a single multidimensional likelihood and use that as a flavor tagger. However, one must remember that these are binned likelihoods and that in order to achieve a reasonable resolution in any given discriminant, the binning must be fine enough to resolve its useful features. On the other hand, binning too finely results in not enough entries per bin, or even bins

with zero entries. This means that one must choose discriminants wisely when attempting to make a combination.

We therefore divide events into three categories based on their opposite side contents:

1. μ candidate and a secondary vertex.
2. μ candidate without a secondary vertex.
3. Secondary vertex without a μ candidate.

To aid in determining an appropriate grouping of discriminants, correlations between all discriminants were studied in a 550 pb^{-1} data sample of $B^\pm \rightarrow J/\psi K^\pm$ reconstructed with the same selections as in Sec. 5.3. The signal region was chosen to be $5.14 < m(J/\psi K^\pm) < 5.40 \text{ GeV}$ and an opposite-side μ was required according to the criteria in Sec. 5.4. The resulting sample contains 451 events of which 83% are estimated to be signal events by fitting to a Gaussian plus exponential background. The resulting correlation matrix is shown in Table 5.1.

	p_T	p_T^{rel}	Imp. Sig.	Q_{SV}	Q_J^μ	Q_{EV}	p_T^{SV}
p_T	1	0.353	0.060	0.136	0.159	0.137	0.071
p_T^{rel}		1	0.103	0.318	0.152	0.100	-0.214
Imp. Sig.			1	0.080	0.009	0.002	0.116
Q_{SV}				1	0.442	0.310	0.229
Q_J^μ					1	0.414	0.131
Q_{EV}						1	0.0001
p_T^{SV}							1

Table 5.1: Correlation matrix for all tagging discriminants as measured on a 550 pb^{-1} $B^\pm \rightarrow J/\psi K^\pm$ opposite-side μ -tagged data sample.

Corresponding to the three categories above and being mindful of the correlations, we choose the following sets of discriminants:

1. $\text{Tag}(\mu+\text{SV})=\{Q_J^\mu; p_T^{\text{rel}}(\mu); Q_{\text{SV}}\}$
2. $\text{Tag}(\mu \text{ without SV})=\{Q_J^\mu; p_T^{\text{rel}}(\mu); p_T(\mu); S_{R\phi}(\mu)\}$
3. $\text{Tag}(\text{SV without } \mu)=\{Q_{\text{EV}}; Q_{\text{SV}}; p_T^{\text{SV}}\}$

The primary reason for the above grouping is that when Monte Carlo statistics are limited we want to use the best discriminants available. Thus, for the case where both a μ and secondary vertex exist on the opposite side, we want to be sure to use both the μ jet charge as well as the secondary vertex charge because we know that these two variables are the strongest discriminants. For the case where there is a μ but no secondary vertex, we have more freedom to add in discriminants. Finally, for the last case, we simply do not have many discriminants available.

Distributions of the tagging variable d for the above three taggers are shown in Fig. 5.3. The distributions shown in this figure are made by applying the taggers to the Monte Carlo $B^\pm \rightarrow J/\psi K^\pm$ samples from which they are created. Figure 5.3 clearly shows separation between the B^+ and B^- sample and the separation increases with increasing $|d|$.

The binnings used for the discriminants in each tagger are as follows:

- $N_{\text{bins}}(Q_J^\mu) = 6 : \{-1, -0.5, -0.25, 0, 0.25, 0.5, 1\}$
- $N_{\text{bins}}(Q_{\text{SV}}) = 6 : \{-1, -0.5, -0.25, 0, 0.25, 0.5, 1\}$
- $N_{\text{bins}}(p_T^{\text{rel}}) = 3 : \{0, 1.5, 3.5, > 3.5\} \text{ GeV}$
- $N_{\text{bins}}(p_T) = 3 : \{0, 5, 10, > 10\} \text{ GeV}$
- $N_{\text{bins}}(\text{imp. sig.}) = 2 : \{0, 2, > 2\}$
- $N_{\text{bins}}(Q_{\text{EV}}) = 6 : \{-1, -0.5, -0.25, 0, 0.25, 0.5, 1\}$
- $N_{\text{bins}}(p_T^{\text{SV}}) = 3 : \{0, 5, 10, > 10\} \text{ GeV.}$

The final multidimensional tagger employed the following logic to decide which of its sub-taggers to use:

1. If the opposite side contains a μ + secondary vertex, use Tag(μ +SV).
2. If the opposite side contains a μ and no secondary vertex, use Tag(μ without SV).
3. If the opposite side contains an electron, use the electron tagger described in Sec. 5.4. Note that this tagger is not multidimensional and is not derived from Monte Carlo.
4. If the opposite side contains a secondary vertex, use Tag(SV without μ).

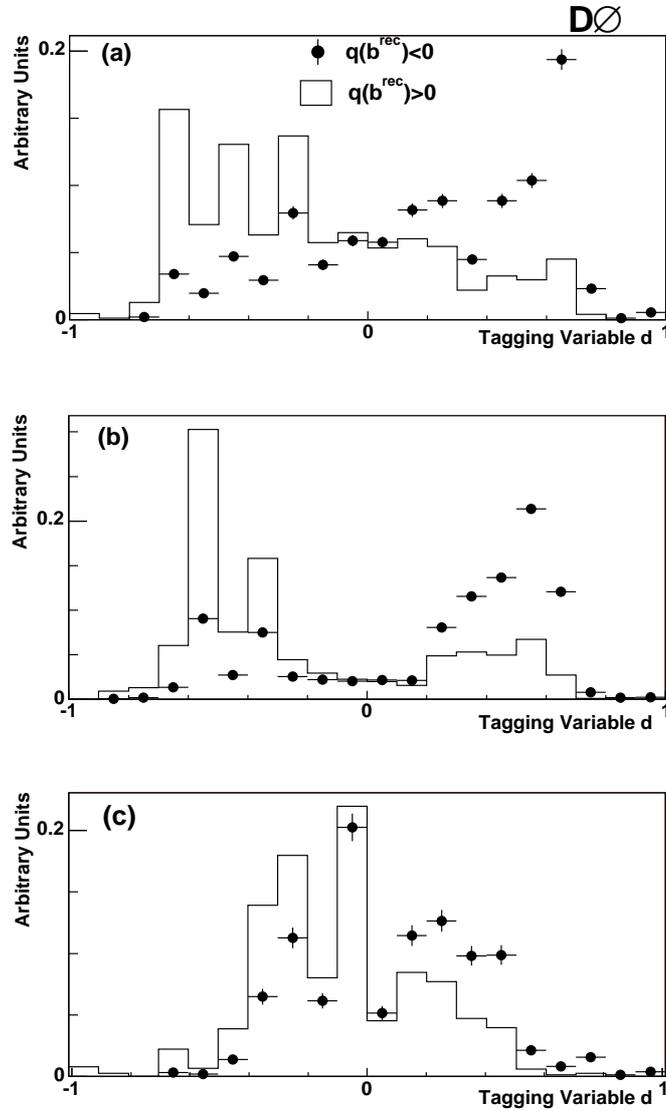


Figure 5.3: Normalized distributions of the combined tagging variable for the 3 multidimensional taggers on the Monte Carlo samples $B^\pm \rightarrow J/\psi K^\pm$ from which they are created. $q(b^{rec})$ is the charge of the b quark from the reconstructed side. a) Distribution of d for $\text{Tag}(\mu + \text{SV})$; b) for $\text{Tag}(\mu \text{ without SV})$; c) and for $\text{Tag}(\text{SV without } \mu)$.

5.6 Binned Likelihood Reweighting

Two types of reweightings are applied to the binned likelihoods to compensate for the limited statistics of the Monte Carlo samples from which they are created: adaptive binning for bins which have no events in either $\mathcal{L}(B^0)$ or $\mathcal{L}(\bar{B}^0)$, and Poisson reweighting for bins which have less than 20 but greater than 0 events in either $\mathcal{L}(B^0)$ or $\mathcal{L}(\bar{B}^0)$.

5.6.1 Adaptive Binning

Suppose, as in Sec. 5.3, we have a set of discriminating variables $\vec{x} = \{x_1, x_2 \dots x_n\}$. Each variable corresponds to a discriminant axis in the multidimensional likelihoods $\mathcal{L}(B^0, \vec{x})$ and $\mathcal{L}(\bar{B}^0, \vec{x})$ such that the vector \vec{x} is associated with the vector of bin numbers $\vec{b} = \{b_1, b_2 \dots b_n\}$. We can therefore write the likelihoods as $\mathcal{L}(B^0, \vec{b})$ and $\mathcal{L}(\bar{B}^0, \vec{b})$. In the case where $\mathcal{L}(B^0, \vec{b}) = 0$ and $\mathcal{L}(\bar{B}^0, \vec{b}) = 0$, we define new likelihoods by adding more bins to the original likelihoods. First we find the discriminant that has the largest number of bins and expand the binning in that direction,

$$\mathcal{L}'_{B^0, \bar{B}^0} = \sum_i^{\mathcal{L}' \neq 0} \mathcal{L}_{B^0, \bar{B}^0}(b_1 + i, b_2 \dots b_n) + \mathcal{L}_{B^0, \bar{B}^0}(b_1 - i, b_2 \dots b_n), \quad (5.13)$$

where b_1 corresponds to the discriminant with the largest number of bins. The sum continues until $\mathcal{L}' \neq 0$. If no events are found after including all the bins of b_1 , we move onto the discriminant with the second largest number of bins,

$$\mathcal{L}''_{B^0, \bar{B}^0} = \mathcal{L}'_{B^0, \bar{B}^0} + \sum_i^{\mathcal{L}'' \neq 0} \mathcal{L}_{B^0, \bar{B}^0}(b_1, b_2 + i \dots b_n) + \mathcal{L}_{B^0, \bar{B}^0}(b_1, b_2 - i \dots b_n). \quad (5.14)$$

The adaptive binning process continues thusly until a non-zero likelihood is obtained for either $\mathcal{L}(B^0)$ or $\mathcal{L}(\bar{B}^0)$.

5.6.2 Poisson Reweighting

The second type of weighting applied to the likelihoods is Poisson weighting to account for fluctuations in low statistics bins. Assume that the number of observed \bar{B}^0 and B^0 to be n and m respectively for a particular bin of the likelihoods. Let $N_{B^0(\bar{B}^0)}$ be the total number of $B^0(\bar{B}^0)$ events in the likelihoods. The probability to observe i events for \bar{B}^0 and j events for B^0 from other samples, $p(i, j)$, can then be written as a product of Poisson probabilities,

$$\begin{aligned} p(i, j) &= f(i; n)f(j; m) \\ f(r; \mu) &= \frac{\mu^r}{r!}e^{-\mu}; \\ &\{f(r; 0) = 1(r = 0), f(r; 0) = 0(r > 0)\} \end{aligned} \quad (5.15)$$

Therefore, the expected value of d from another sample can be calculated as,

$$\begin{aligned} \langle d \rangle &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} d(i, j)f(i; n)f(j; m) \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{i/N_{\bar{B}^0} - j/N_{B^0}}{i/N_{\bar{B}^0} + j/N_{B^0}} f(i; n)f(j; m). \end{aligned} \quad (5.16)$$

When n or m is larger than 20, the Poisson probability becomes effectively Gaussian. d is therefore calculated as,

$$\begin{aligned} d &= \sum_{i=0}^{50} \sum_{j=0}^{50} \frac{i/N_{\bar{B}^0} - j/N_{B^0}}{i/N_{\bar{B}^0} + j/N_{B^0}} f(i; n)f(j; m) \quad (n < 20 \text{ and } m < 20) \\ &= \frac{n/N_{\bar{B}^0} - m/N_{B^0}}{n/N_{\bar{B}^0} + m/N_{B^0}} \quad (n \geq 20 \text{ or } m \geq 20) \end{aligned} \quad (5.17)$$

5.7 A Measurement of Δm_d

5.7.1 Data Sample and Event Selection for Δm_d Measurement

This measurement exploits the large semileptonic data sample corresponding to approximately 1 fb^{-1} of integrated luminosity, accumulated by the DØ detector between April 2002 and October 2005.

B mesons were selected using their semileptonic decays ¹ $B \rightarrow \mu^+ \nu \bar{D}^0 X$ and were divided into two exclusive groups: the D^* sample, containing all events with reconstructed $D^{*-} \rightarrow \bar{D}^0 \pi^-$ decays, and the D^0 sample, containing all remaining events. Experimental results show that the D^* sample is dominated by $B_d^0 \rightarrow \mu^+ \nu D^{*-} X$ decays, while the D^0 sample is dominated by $B^+ \rightarrow \mu^+ \nu \bar{D}^0 X$ decays. The purity of the flavor tagging as well as Δm_d are measured in the D^* sample as well as the D^0 sample to test the independence of the flavor tagger on the reconstructed side meson.

Muons for this analysis were required to be identified as “loose” by the standard DØ identification tools [47], have hits in more than one muon chamber ($n_{\text{seg}} > 1$), an associated track in the central tracking system with hits in both SMT and CFT present, transverse momentum $p_T^\mu > 2 \text{ GeV}$ as measured in the central tracker, pseudo-rapidity $|\eta^\mu| < 2$ and total momentum $p^\mu > 3 \text{ GeV}$.

All charged particles in a given event were clustered into jets using the DURHAM clustering algorithm, as noted in Sec. 3.1.2. Events with more than one identified muon in the same jet were rejected, as well as events with identified $J/\psi \rightarrow \mu^+ \mu^-$ decays.

The D^0 candidate was constructed from two particles of opposite charge belonging to the same jet as the reconstructed muon. Both particles were required to have transverse momentum $p_T > 0.7 \text{ GeV}$, and pseudo-rapidity $|\eta| < 2$. They were required to form a common D -vertex with the fit $\chi^2 <$

¹Charge-conjugate states are implied throughout

9. The combined significance $\sqrt{(d_{R\phi}/\sigma_{R\phi}^2) + (\epsilon_{Rz}/\sigma_{Rz})^2}$ was required to be greater than 2. The distance d_T^D between the primary and D vertices in the axial plane was required to exceed 4 standard deviations: $d_T^D/\sigma(d_T^D) > 4$. The accuracy of the distance d_T^D determination was required to be better than 500 μm . The angle α_T^D between the D^0 momentum and the direction from the primary to the D^0 vertex in the axial plane was required to satisfy the condition: $\cos(\alpha_T^D) > 0.9$. The tracks of muon and D^0 candidate were required to form a common B vertex with the fit $\chi^2 < 9$. The momentum of the B candidate was computed as the sum of the momenta of the μ and D^0 . The mass of the $(\mu^+ \bar{D}^0)$ system was required to fall within $2.3 < M(\mu^+ \bar{D}^0) < 5.2$ GeV. Note that the nominal mass of the B^0 is $M(B^0) = 5.279$ GeV [13]. The masses of kaon and pion were assigned to the particles according to the charge of the muon, requiring the $\mu^+ K^+ \pi^-$ final system or its charge conjugate.

If the distance d_T^B between the primary and B vertices in the axial plane exceeded $4\sigma(d_T^B)$, the angle α_T^B between the B momentum and the direction from the primary to the B -vertex in the axial plane was demanded to satisfy the condition $\cos(\alpha_T^B) > 0.95$. The distance d_T^B was allowed to be greater than d_T^D , provided that the distance between the B and D vertices d_T^{BD} was less than $3\sigma(d_T^{BD})$. The error $\sigma(d_T^B)$ was required to be less than 500 μm . In addition, the cut $p_T(\bar{D}^0) > 5$ GeV/ c^2 was applied.

For the $\mu^+ D^{*-}$ candidates, we searched for an additional pion with $p_T > 0.18$ GeV and charge opposite to the charge of the muon. The mass difference $\Delta M = M(\bar{D}^0 \pi) - M(\bar{D}^0)$ for D^* candidates having $1.75 < M(\bar{D}^0) < 1.95$ GeV is shown in Fig. 5.4. The peak corresponding to $M(D^*) - M(D^0)$ is clearly seen.

All events with $0.1425 < \Delta M < 0.1490$ GeV/ c^2 were included in the D^* sample. All remaining events were included in D^0 sample. The $K\pi$ mass distribution for these two samples together with the results of the fit is shown in Fig. 5.5. The procedure to fit these mass spectra is described

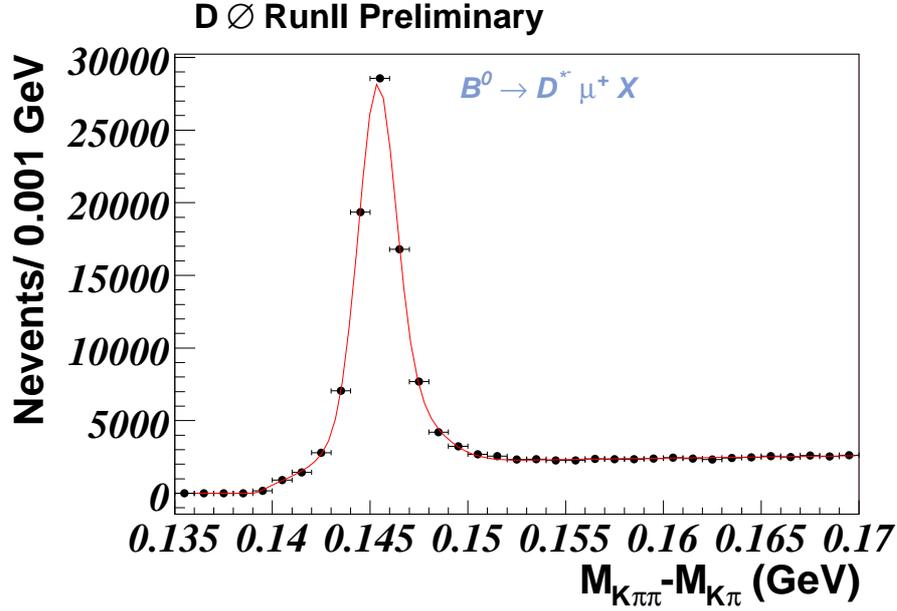


Figure 5.4: The $M(K\pi\pi) - M(K\pi)$ invariant mass for selected μD^* candidates. The curve shows the result of the fit described in Sec. 5.7.2.

on section 5.7.2. In total, $230551 \pm 1627(\text{stat.})$ $B \rightarrow \mu^+ \nu \bar{D}^0$ decays and $73532 \pm 304(\text{stat.})$ $B \rightarrow \mu^+ \nu \bar{D}^{*}$ decays were reconstructed.

5.7.2 Fitting Procedure for Δm_d Measurement

The performance of the flavor tagging and the B_d^0 mixing parameter Δm_d were obtained from the study of evolution of the flavor asymmetry as a function of the B meson decay length.

The flavor asymmetry A is defined as:

$$A = \frac{N^{\text{nos}} - N^{\text{osc}}}{N^{\text{nos}} + N^{\text{osc}}}, \quad (5.18)$$

Here N^{nos} is the number of non-oscillating B decays and N^{osc} is the number of oscillating B decays. An event $B \rightarrow \mu^+ \nu \bar{D}^0 X$ with $q(\mu) \cdot \text{sign}(d) <$

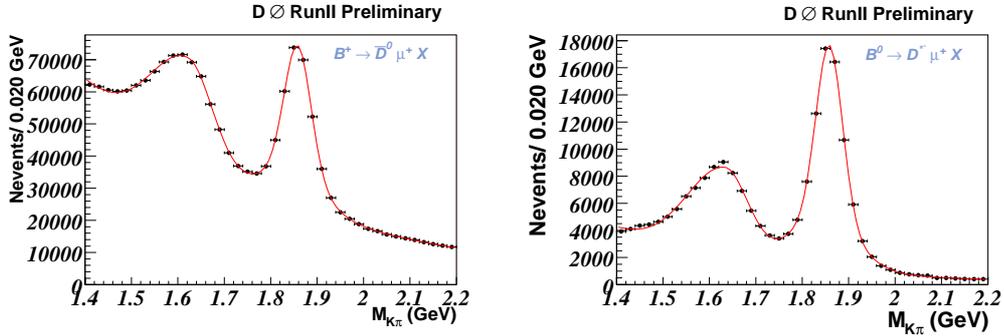


Figure 5.5: The $K\pi$ invariant mass with the result of the fit described in Sec. 5.7.2 overlaid. The μD^0 sample is shown on the left and the μD^* sample is shown on the right.

0 where d is the flavor tagging variable was tagged as non-oscillating, and an event with $q(\mu) \cdot \text{sign}(d) > 0$ was tagged as oscillating.

All events in the D^0 and D^* samples were divided into 7 bins according to the measured VPDL as defined in Eq. 3.11. The VPDL intervals are given in Table 5.5. The number of oscillating N_i^{osc} and non-oscillating N_i^{nos} signal events in each interval i was determined from a fit of the D^0 signal in the $K\pi$ invariant mass distribution for both samples.

5.7.3 Mass Fit

The fitting function was chosen to give the best χ^2 fit to the $K\pi$ mass spectrum of the entire sample of $B \rightarrow \mu^+ \bar{D}^0 X$ events shown in Fig. 5.5. The signal peak corresponding to the decay $D^0 \rightarrow K^- \pi^+$ peaks can be seen at a mass of 1.857 GeV. The background to the right of the signal region is adequately described by an exponential function:

$$f_1^{\text{bkg}} = a_0 \cdot \exp(x/b_0). \quad (5.19)$$

The peak in the background to the left of the signal is due to events in which D^0 decays to $K\pi X$ where X is not reconstructed, e.g. $X = \pi^0$. We

model this shape with a bifurcated Gaussian:

$$\begin{aligned} f_2^{bkg} &= N_0 \cdot \exp\left(-\frac{(x - \mu_0)^2}{2\sigma_R^2}\right) \quad \text{for } (x - \mu_0) > 0.0 \\ &= N_0 \cdot \exp\left(-\frac{(x - \mu_0)^2}{2\sigma_L^2}\right) \quad \text{for } (x - \mu_0) < 0.0. \end{aligned} \quad (5.20)$$

Here N_0 is the normalization of the background gaussian, μ_0 is the mean of the gaussian and σ_L and σ_R are the two widths of the bifurcated gaussian.

The signal peak is modelled by the sum of two Gaussians:

$$\begin{aligned} f^{sig} &= 0.5A(1 + R) \exp\left(-\frac{(x - \mu_1)^2}{2\sigma_1^2}\right) + \\ &\quad 0.5A(1 - R) \exp\left(-\frac{(x - \mu_2)^2}{2\sigma_2^2}\right), \\ N^{sig} &= \frac{\sqrt{2\pi}}{2} A((1 + R)\sigma_1 + (1 - R)\sigma_2), \end{aligned} \quad (5.21)$$

where N^{sig} is the number of signal events, μ_1 and μ_2 are the means of the Gaussians, σ_1 and σ_2 are the widths of the Gaussians, and R controls the relative contribution of the two Gaussians.

The two Gaussians were not constrained to have the same mean. Table 5.2 shows the improvement in the χ^2 value for this choice over the cases where a single Gaussian is used, or the case when the means are constrained to the same value.

The total fitting function, which in general has 12 free parameters, is:

$$f = f^{sig} + f_1^{bkg} + f_2^{bkg}. \quad (5.22)$$

The low statistics in tagged VPDL bins, which have as few as ten events, do not permit a free fit to this function and some parameters should be constrained or fixed. In order to do this it was necessary to show that the

Fit function	$N^{sig}/1000$	μ (GeV)	χ^2	d.o.f.
Single Gaussian	210.5 ± 0.8	1.8571 ± 0.0001	769	31
Double Gaussian, means constrained	226.0 ± 1.0	1.8569 ± 0.0001	117	29
Double Gaussian, means free	230.5 ± 0.9	1.8676 ± 0.0017 1.8487 ± 0.0017	57	28

Table 5.2: Table showing the χ^2 of different fits to the entire $B^+ \rightarrow \mu^+ \nu \bar{D}^0 X$ sample.

constraints on the parameters are valid over all VPDL bins for tagged events. Unconstrained fits were performed to several high statistic samples. The set of all events was fitted as a reference fit. Events divided into VPDL bins were fitted to investigate changes with VPDL. Three samples were made to test whether the presence of a flavor tag changes the mass spectrum: all tagged events over the entire VPDL range; all events in the short VPDL range (0,0.05) tagged unmixed (i.e. enriched with correct tags); all events in VPDL range (0,0.05) tagged mixed, (i.e. enriched with wrong tags).

Tables 5.3 and 5.4 show the results of the unconstrained fits. The first column shows parameter values from the sample of all events. The fitted parameters were chosen to optimize the fit procedure, by enabling an initial fit with “difference” terms such as $(\mu_1 - \mu_2)$ constrained to zero, followed by a full fit with these terms unconstrained. Subsequent columns show the significance of the deviation from the fitted value, $(X_{all} - X_{sample}) / \sqrt{\sigma^2(X_{all}) + \sigma^2(X_{sample})}$.

The parameter b_0 shows significant variation across the samples. This is consistent with our expectation that the shape of the background depends on VPDL. The other parameters that determine the shape and position of the signal and background peaks in each sample are broadly consistent with the fit to all events.

Therefore when fitting the plots, the parameters describing the width and position of the Gaussians were fixed using a free fit to the total D^0 or

Parameter		VPDL							
		(-0.025, 0.000)	(0.000, 0.025)	(0.025, 0.050)	(0.050, 0.075)	(0.075, 0.100)	(0.100, 0.125)	(0.125, 0.250)	
b_0	0.464 ± 0.001 GeV	-28	-21	-2.5	13	24	26	35	
μ_0	1.620 ± 0.001 GeV	-0.5	-1.2	-1.8	0.3	1.4	2.6	1.4	
$\frac{\sigma_R + \sigma_L}{2}$	0.0669 ± 0.0005 GeV	1.4	0.1	0.0	-1.7	-1.0	0.8	-1.7	
$\frac{\sigma_R - \sigma_L}{\sigma_R + \sigma_L}$	-0.17 ± 0.01	0.9	-0.3	-1.8	-0.3	1.3	1.8	0.8	
μ_1	1.8570 ± 0.0001 GeV	0.5	-2.9	-0.1	1.4	-0.5	1.7	3.0	
$\frac{\sigma_1 + \sigma_2}{2}$	0.042 ± 0.001 GeV	0.2	0.0	-0.7	-0.3	-0.6	-0.5	-0.6	
R	0.47 ± 0.04	1.2	-0.1	-0.7	0.1	-0.6	-0.4	0.1	
$\frac{\sigma_1 - \sigma_2}{\sigma_1 + \sigma_2}$	-0.34 ± 0.01	1.2	-0.1	-0.7	0.1	-0.6	-0.4	0.1	

Table 5.3: The results of a free fit to $B^+ \rightarrow \mu^+ \nu \bar{D}^0 X$ samples. The first column contains parameter values from the fit to all events. Subsequent columns show the significance of the deviation from this value, for the fits to other samples. The background slope parameter b_0 shows significant variation across VPDL bins, consistent with our expectation that the shape of the background depends on VPDL.

D^* sample. This left four free parameters: the number of events in the signal peak, background peak, and exponential background, and the slope constant of the exponential background. The number of (non-)oscillating signal events in each VPDL interval and for different values of the flavor tagging variable is given in the Tables (5.5-5.8). The fit to the $K\pi$ mass spectrum for events with $|d| > 0.37$ selected by the multidimensional tagger is shown in Fig.5.6. This cut on the tagging variable d was selected to maximize tagging performance for the binned oscillation asymmetry fit. We remove this cut when using the flavor tagger in the unbinned Δm_s analysis shown in Ch. 6.

Parameter		VPDL						
		(-0.025, 0.000)	(0.000, 0.025)	(0.025, 0.050)	(0.050, 0.075)	(0.075, 0.100)	(0.100, 0.125)	(0.125, 0.250)
b_0	0.337 ± 0.003 GeV	-5.9	-6.2	1.6	1.8	6.2	6.5	9.0
μ_0	1.628 ± 0.001 GeV	0.4	-1.0	-0.4	-0.1	0.9	0.2	0.9
$\frac{\sigma_R + \sigma_L}{2}$	0.069 ± 0.001 GeV	-0.4	0.2	1.3	-1.1	-0.6	-0.2	-0.7
$\frac{\sigma_R - \sigma_L}{\sigma_R + \sigma_L}$	-0.12 ± 0.02	-0.4	0.3	-0.2	-0.2	0.5	-0.1	-0.1
μ_1	1.8565 ± 0.0002 GeV	-0.7	-3.9	0.7	0.4	2.0	2.5	1.5
$\frac{\sigma_1 + \sigma_2}{2}$	0.042 ± 0.001 GeV	-1.2	0.4	0.0	0.8	0.1	-0.7	-1.2
R	0.39 ± 0.04	2.3	0.0	0.2	0.7	0.3	-0.5	-1.0
$\frac{\sigma_1 - \sigma_2}{\sigma_1 + \sigma_2}$	-0.36 ± 0.01	-0.6	-0.4	-1.2	0.3	0.0	-0.7	0.2

Table 5.4: The results of a free fit to $B_d^0 \rightarrow \mu^+ \nu D^{*-} X$ samples. The first column contains parameter values from the fit to all events. Subsequent columns show the significance of the deviation from this value, for the fits to other samples. The background slope parameter b_0 shows significant variation across VPDL bins, consistent with our expectation that the shape of the background depends on VPDL.

5.7.4 Expected Flavor Asymmetry

For a given type of B_q meson ($q = u, d, s$) the distribution of the visible proper decay length x is given by:

$$n_u^{nos}(x, K) = \frac{K}{c\tau_{B^+}} \exp\left(-\frac{Kx}{c\tau_{B^+}}\right) \frac{1 + \mathcal{D}_u}{2}, \quad (5.23)$$

$$n_u^{osc}(x, K) = \frac{K}{c\tau_{B^+}} \exp\left(-\frac{Kx}{c\tau_{B^+}}\right) \frac{1 - \mathcal{D}_u}{2}, \quad (5.24)$$

$$n_d^{nos}(x, K) = \frac{K}{c\tau_{B^0}} \exp\left(-\frac{Kx}{c\tau_{B^0}}\right) \cdot 0.5 \cdot (1 + \mathcal{D}_d \cos(\Delta m_d K x / c)), \quad (5.25)$$

$$n_d^{osc}(x, K) = \frac{K}{c\tau_{B^0}} \exp\left(-\frac{Kx}{c\tau_{B^0}}\right) \cdot 0.5 \cdot (1 - \mathcal{D}_d \cos(\Delta m_d K x / c)), \quad (5.26)$$

$$n_s^{nos}(x, K) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5, \quad (5.27)$$

$$n_s^{osc}(x, K) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5. \quad (5.28)$$

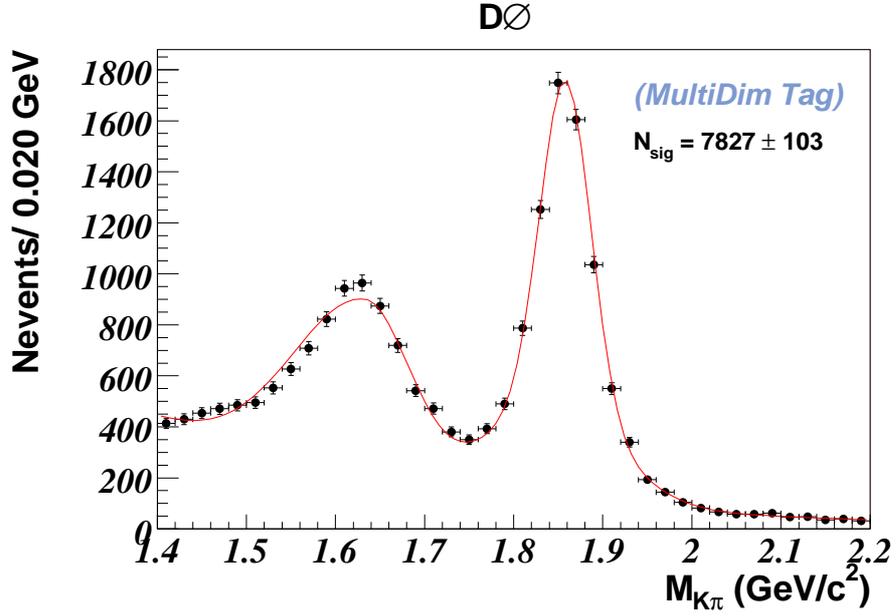


Figure 5.6: The tagged $M(K\pi)$ distribution for events in the D^* sample with $|d| > 0.37$.

Here τ is the lifetime of B meson, Δm_d is the mixing parameter of B^0 , and the K -factor is as defined in Sec. 3.2. The B^+ meson does not oscillate and it is assumed that the B_s meson oscillates with infinite frequency. The flavor tagging dilution is given by \mathcal{D} . In general case it can be different for B^0 and B^+ . In our study we verify the assumption that $\mathcal{D}_d = \mathcal{D}_u$ for our opposite-side flavor tagger.

The transition from the true x to the experimentally measured visible proper decay length, x^M , is achieved by integrating over the K -factor distribution and convoluting with the resolution function:

$$N_{q,j}^{nos/osc}(x^M) = \int dx R_j(x - x^M) \varepsilon_j(x) \theta(x) \int dK D_j(K) n_{q,j}^{nos/osc}(x, K) \quad (5.29)$$

Here $R_j(x - x^M)$ is the detector resolution in VPDL and $\varepsilon_j(x)$ is the reconstruction efficiency for a given B meson decay channel j . The step function $\theta(x)$ ensures that x is positive in the integration. x^M can have negative values due to resolution effects. The function $D_j(K)$ is a normalised distribution of the K -factor in a given channel j .

In addition to the main decay channels $B \rightarrow \mu^+ \nu \bar{D}^0 X$, the process $c\bar{c} \rightarrow \mu^+ \nu \bar{D}^0 X$ gives a contribution in the selected final state. A dedicated analysis was developed to study this process, both in data and in simulation. It shows that the pseudo-decay length, constructed from the crossing of the μ and \bar{D}^0 trajectories, is distributed around zero with $\sigma \sim 150\mu\text{m}$. The distribution $N^{c\bar{c}}(x^M)$ of VPDL for this process was taken from simulation. It was assumed that the production ratio $(c \rightarrow D^*)/(c \rightarrow D^0)$ is the same as in the semileptonic B decays and that the flavor tagging for the $c\bar{c}$ events gives the same rate of oscillated and non-oscillated events. The fraction $f_{c\bar{c}}$ of $c\bar{c}$ events was obtained from the fit.

Taking into account all mentioned contributions, the expected number of (non-) oscillated events in the i -th bin of VPDL is:

$$N_i^{e,nos/osc} = \int_i dx^M \left((1 - f_{c\bar{c}}) \left(\sum_{q=u,d,s} \sum_j (Br_j \cdot N_{q,j}^{nos/osc}(x^M)) \right) + f_{c\bar{c}} N_{c\bar{c}}(x^M) \right) \quad (5.30)$$

Here the integration $\int_i dx^M$ is taken over a given interval i , the sum \sum_j is taken over all decay channels $B \rightarrow \mu^+ \nu \bar{D}^0 X$, contributing to the selected sample and the Br_j is the branching rate of a given channel j .

Finally, the expected value A_i^e for each interval i of the measured VPDL is given by:

$$A_i^e(\Delta m, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) = \frac{N_i^{e,nos} - N_i^{e,osc}}{N_i^{e,nos} + N_i^{e,osc}} \quad (5.31)$$

The expected asymmetry can be computed both for D^* and D^0 samples. The only difference between them is the contributing decay channels of B

mesons.

For the computation of A_i^c , the B meson lifetimes and the branching rates Br_j were derived from the PDG [13]. They are discussed in the following section. The functions $D_j(K)$, $R_j(x)$ and $\varepsilon_j(x)$ were found using Monte Carlo simulations.

5.7.5 Sample Composition

There is cross-contamination between $B^0 \rightarrow \mu^+\nu\bar{D}^0X$, $B_s^0 \rightarrow \mu^+\nu\bar{D}^0X$ and $B^+ \rightarrow \mu^+\nu\bar{D}^0X$ samples because B mesons in semileptonic decays are not fully reconstructed. To determine the composition of selected samples we list all possible decay chains for B^0 , B_s^0 and B^+ with corresponding branching ratios from which we estimate the sample composition of the D^* and D^0 samples.

The following decay channels of B mesons (B^0, B^\pm, B_s^0) were considered for the D^* sample:

- $B^0 \rightarrow \mu^+\nu D^{*-}$;
- $B^0 \rightarrow \mu^+\nu D^{*-} \rightarrow \mu^+\nu D^{*-} X$;
- $B^+ \rightarrow \mu^+\nu \bar{D}^{*0} \rightarrow \mu^+\nu D^{*-} X$;
- $B_s^0 \rightarrow \mu^+\nu D^{*-} X$.

and for the D^0 sample:

- $B^+ \rightarrow \mu^+\nu \bar{D}^0$;
- $B^+ \rightarrow \mu^+\nu \bar{D}^{*0}$;
- $B^+ \rightarrow \mu^+\nu \bar{D}^{*0} \rightarrow \mu^+\nu \bar{D}^0 X$;
- $B^+ \rightarrow \mu^+\nu \bar{D}^{*0} \rightarrow \mu^+\nu \bar{D}^{*0} X$;
- $B^0 \rightarrow \mu^+\nu D^{*-} \rightarrow \mu^+\nu \bar{D}^0 X$;

- $B^0 \rightarrow \mu^+ \nu D^{*-} \rightarrow \mu^+ \nu \bar{D}^{*0} X$;
- $B_s^0 \rightarrow \mu^+ \nu \bar{D}^0 X$;
- $B_s^0 \rightarrow \mu^+ \nu \bar{D}^{*0} X$.

Here and in the following the symbol “ D^{**} ” denotes both narrow and wide D^{**} resonances, together with non-resonant $D\pi$ and $D^*\pi$ production. The contribution of $D\pi\pi$ final states was neglected.

In addition, various decay chains are affected differently by the B -meson selection cuts, so additional reconstruction efficiencies are introduced to correct for this effect. The reconstruction efficiencies were determined from Monte Carlo simulations for each of the corresponding channels.

The latest PDG values [13] were used to determine the branching fractions of decays contributing to the D^0 and D^* samples:

- $\text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^0) = (2.15 \pm 0.22)\%$;
- $\text{Br}(B^0 \rightarrow \mu^+ \nu D^-) = (2.14 \pm 0.20)\%$;
- $\text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^{*0}) = (6.5 \pm 0.5)\%$;
- $\text{Br}(B^0 \rightarrow \mu^+ \nu D^{*-}) = (5.44 \pm 0.23)\%$.

$\text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^{**0})$ was estimated using the following inputs:

- $\text{Br}(B^+ \rightarrow \mu^+ \nu X) = (10.73 \pm 0.28)\%$;
- $\text{Br}(B^0 \rightarrow \mu^+ \nu X) = \tau^0/\tau^+ \cdot \text{Br}(B^+ \rightarrow \mu^+ \nu X)$;
- $\text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^{**0}) = \text{Br}(B^+ \rightarrow \mu^+ \nu X) - \text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^0) - \text{Br}(B^+ \rightarrow \mu^+ \nu D^{*-})$;
- $\text{Br}(B^0 \rightarrow \mu^+ \nu D^{*-}) = \text{Br}(B^0 \rightarrow \mu^+ \nu X) - \text{Br}(B^0 \rightarrow \mu^+ \nu D^-) - \text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^{*0})$;
- $\text{Br}(B^0 \rightarrow \mu^+ \nu D^{**}) = \tau^0/\tau^+ \cdot \text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^{**0})$;

to obtain the following value:

$$\text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^{**0}) = (2.70 \pm 0.47)\%. \quad (5.32)$$

$\text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^{**0} \rightarrow \mu^+ \nu D^{*-} X)$ was estimated from the following inputs:

- $\text{Br}(\bar{b} \rightarrow \ell^+ \nu D^{*-} \pi^+ X) = (4.73 \pm 0.77 \pm 0.55) \cdot 10^{-3}$ [49];
- $\text{Br}(\bar{b} \rightarrow \ell^+ \nu D^{*-} \pi^+ X) = (4.80 \pm 0.9 \pm 0.5) \cdot 10^{-3}$ [50];
- $\text{Br}(\bar{b} \rightarrow \ell^+ \nu D^{*-} \pi^- X) = (0.6 \pm 0.7 \pm 0.2) \cdot 10^{-3}$ [50];

and assuming $\text{Br}(b \rightarrow B^+) = 0.397 \pm 0.010$ [13]. The usual practice in estimating this decay rate is to neglect the contributions of decays $D^{**} \rightarrow D^* \pi \pi$. However, the data listed above allow us to also take into account these decays. Neglecting the decays $D^{**} \rightarrow D^* \pi \pi$, which give $\sim 1\%$ contribution to D^{**} decays according to simulation, the available measurements can be expressed as:

$$\begin{aligned} \text{Br}(\bar{B} \rightarrow \ell^+ \nu D^{*-} \pi^+ X) &= \text{Br}(B^+ \rightarrow \ell^+ \nu D^{*-} \pi^+ X^0) + \text{Br}(B^0 \rightarrow \ell^+ \nu D^{*-} \pi^+ \pi^-), \\ \text{Br}(\bar{B} \rightarrow \ell^+ \nu D^{*-} \pi^- X) &= \text{Br}(B^0 \rightarrow \ell^+ \nu D^{*-} \pi^+ \pi^-). \end{aligned}$$

From these relations and the measurements listed above, we obtain:

$$\text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^{**0} \rightarrow \ell^+ \nu D^{*-} X) = (1.06 \pm 0.24)\%. \quad (5.33)$$

All other rates $\text{Br}(B \rightarrow \mu^+ \nu \bar{D}^{**} \rightarrow \mu^+ \nu \bar{D}^* X)$ were obtained using the following relations:

- $\text{Br}(B^0 \rightarrow \mu^+ \nu D^{**0} (D^* \pi)) = \tau^0 / \tau^+ \cdot \text{Br}(B^+ \rightarrow \mu^+ \nu \bar{D}^{**0} (D^* \pi) X)$;
- $\text{Br}(B \rightarrow \mu^+ \nu \bar{D}^{**} (\bar{D}^* \pi^+) X) = 2 \cdot \text{Br}(B \rightarrow \mu^+ \nu \bar{D}^{**} (\bar{D}^* \pi^0) X)$ (isospin invariance).

The following inputs were used to estimate $\text{Br}(B \rightarrow \mu^+ \nu \bar{D}^{**} \rightarrow \mu^+ \nu \bar{D} X)$:

- $\text{Br}(B \rightarrow \mu^+ \nu \bar{D}^{**}(\bar{D}\pi^+)X) = 2 \cdot \text{Br}(B \rightarrow \mu^+ \nu \bar{D}^{**}(\bar{D}\pi^0)X)$ (isospin invariance);
- $\text{Br}(B \rightarrow \mu^+ \nu \bar{D}^{**}(\bar{D}\pi)X) = \text{Br}(B \rightarrow \mu^+ \nu \bar{D}^{**}) - \text{Br}(B \rightarrow \mu^+ \nu \bar{D}^{**}(\bar{D}^*\pi)X)$.

To estimate branching rates of B_s^0 decays, the following inputs were used:

- $\text{Br}(B_s^0 \rightarrow \mu^+ \nu X) = \tau^s / \tau^d \cdot \text{Br}(B^0 \rightarrow \mu^+ \nu X)$;
- $\text{Br}(B_s^0 \rightarrow \mu^+ \nu D_s^- X) = (7.9 \pm 2.4)\%$ [13];
- $\text{Br}(B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu D^{*-} X) = \text{Br}(B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu \bar{D}^{*0} X)$ (isospin invariance).

In addition, it was assumed that:

$$R_s^{**} = \frac{\text{Br}(B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu D^* X)}{\text{Br}(B_s^0 \rightarrow \mu^+ \nu D_s^{**})} = 0.35. \quad (5.34)$$

There is no experimental measurement of this rate yet. This rate was varied between 0 and 1 to estimate the systematic uncertainty on the contribution to the samples.

According to the above, the D^* sample consists of 89% B^0 and 10% B^+ decays. The D^0 sample correspondingly has a 85% contribution from B^+ and 15% contribution from B^0 . These numbers do not take into account the reconstruction efficiencies in different channels and the B_s^0 contribution.

The sample composition can also be extracted from the MC simulation. The decay rates used for the DØ generation of B hadron decays give an 87% contribution from B^0 and a 13% contribution from B^+ in the D^* sample. The D^0 sample has an 83% contribution from B^+ and a 17% contribution from B^0 . These numbers are in good agreement with the experimental measurements. Taking into account the reconstruction efficiencies, B meson lifetimes and

the B_s^0 contribution, the D^* sample contains 89% B^0 , 10% B^+ and 1% B_s^0 , while the D^0 sample contains 83% B^+ , 16% B^0 and 1% B_s^0 .

Since the D^* sample is selected by a cut on the mass difference $\Delta M = M(D^0\pi) - M(D^0)$, there is a small additional contribution of $B \rightarrow \mu^+\nu\bar{D}^0$ events in the D^* sample when a D^0 is randomly combined with a pion from the combinatorial background. The fraction of this contribution was estimated using $\mu^+\bar{D}^0\pi^+$ events. These events were selected applying all criteria for the D^* sample, described in Sec. 5.7.1, except that the wrong charge correlation of muon and pion was required. The number of D^0 events was determined using the same fitting procedure as for D^* sample and the additional fraction of $B \rightarrow \mu^+\nu\bar{D}^0$ events in the D^* sample was estimated to be $(4.00 \pm 0.85)\%$.

5.7.6 Results

For any sample of tagged events, the observed and expected asymmetries were determined using Eq. 5.18 and Eq. 5.31 in each VPDL bin and the values of Δm_d , $f_{c\bar{c}}$, \mathcal{D}_d and \mathcal{D}_u were obtained from a simultaneous χ^2 fit of:

$$\begin{aligned} \chi^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) &= \chi_{D^*}^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) + \chi_{D^0}^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) \\ \chi_{D^*}^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) &= \sum_i \frac{(A_{i,D^*} - A_{i,D^*}^e(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u))^2}{\sigma^2(A_{i,D^*})} \\ \chi_{D^0}^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) &= \sum_i \frac{(A_{i,D^0} - A_{i,D^0}^e(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u))^2}{\sigma^2(A_{i,D^0})}. \end{aligned} \quad (5.35)$$

Here \sum_i is the sum over all VPDL bins. The measured and expected flavor asymmetry for samples with different value of the combined tagging variable are given in Tables 5.5–5.8. Examples of the fits of the $(K\pi)$ mass distribution used to determine the flavor asymmetry are shown in Figs. 5.7–5.9. Examples of the fit of the flavor asymmetry using Eq. 5.35 are shown in Figs. 5.10–5.12.

The top plot in Fig. 5.10 shows clear oscillations in the D^* sample as

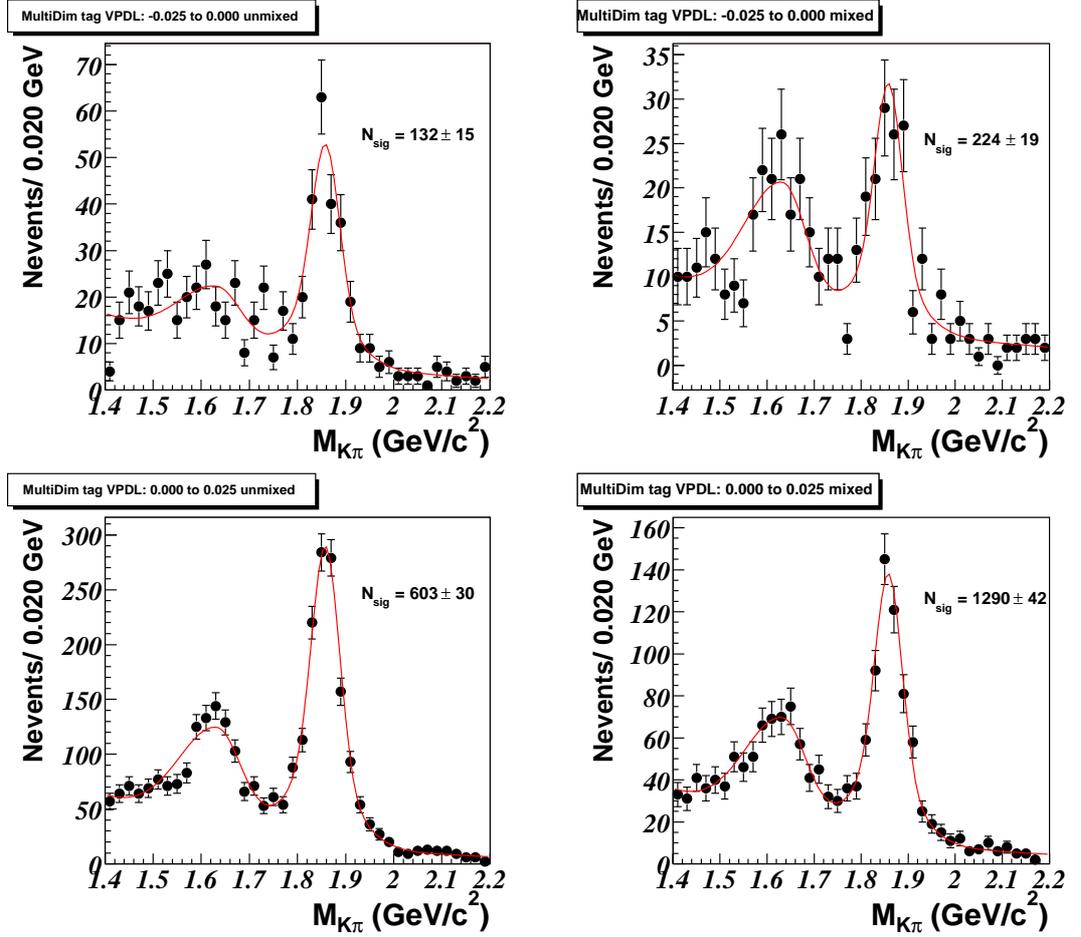


Figure 5.7: The fit to $M(K\pi)$ mass for non-oscillating (left) and oscillating (right) for μ^+D^{*-} events tagged by the multidimensional tagger with $|d| > 0.37$ for the VPDL ranges $\{-0.025, 0.0\}$ and $\{0.0 - 0.025\}$ (cm).

Bin No.	VPDL range, cm	N_i^{osc}	N_i^{nosc}	A_i	A_i^e
D^* Sample					
0	-0.025-0.000	97 ± 13	72 ± 11	-0.147 ± 0.094	0.071
1	0.000-0.025	391 ± 23	494 ± 26	0.117 ± 0.038	0.109
2	0.025-0.050	393 ± 23	505 ± 27	0.125 ± 0.038	0.104
3	0.050-0.075	272 ± 20	321 ± 22	0.084 ± 0.048	0.086
4	0.075-0.100	210 ± 17	240 ± 18	0.067 ± 0.054	0.061
5	0.100-0.125	120 ± 13	135 ± 14	0.056 ± 0.073	0.033
6	0.125-0.250	229 ± 18	233 ± 18	0.008 ± 0.054	-0.015
D^0 Sample					
0	-0.025-0.000	162 ± 28	258 ± 28	0.228 ± 0.095	0.107
1	0.000-0.025	1105 ± 49	1378 ± 54	0.110 ± 0.029	0.168
2	0.025-0.050	1036 ± 47	1434 ± 54	0.161 ± 0.029	0.175
3	0.050-0.075	666 ± 40	1116 ± 47	0.253 ± 0.034	0.170
4	0.075-0.100	519 ± 35	729 ± 38	0.169 ± 0.041	0.164
5	0.100-0.125	402 ± 31	489 ± 32	0.098 ± 0.049	0.157
6	0.125-0.250	712 ± 42	1017 ± 47	0.177 ± 0.036	0.144

Table 5.5: For each VPDL bin, the measured number of D^* and D^0 events with opposite sign and same sign in the tagging variable range ($0.2 < |d| < 0.35$) N_i^{nos} , N_i^{osc} , the measured asymmetry A_i and expected asymmetry A_i^e corresponding to the fitted value $\Delta m_d = 0.486 \text{ ps}^{-1}$ are given.

function of VPDL. This sample is dominated by B^0 , so we see larger oscillations than in the D^0 sample (the bottom plot of Fig. 5.10). This is expected as the D^0 sample is dominated by B^+ , which does not oscillate. The residual oscillation is due to the $\sim 15\%$ B^0 cross-contamination in the D^0 sample. Also note that in Figs. 5.11 and 5.12 that as the value of $|d|$ increases, the amplitude of the oscillation in the D^* sample increases.

Results are given in Tables 5.9–5.11. All errors are statistical and do not include systematic uncertainties. The tagging efficiencies shown in Tables 5.9 and 5.10 were computed using events with the VPDL in the range $\{0.025, 0.250\}$. This selection reduces the contribution from $c\bar{c} \rightarrow \mu^+ \nu D^0 X$

Bin No.	VPDL range, cm	N_i^{osc}	N_i^{nos}	A_i	A_i^e
D^* Sample					
0	-0.025-0.000	79 ± 12	131 ± 15	0.247 ± 0.085	0.178
1	0.000-0.025	431 ± 25	737 ± 32	0.263 ± 0.033	0.272
2	0.025-0.050	365 ± 22	662 ± 30	0.290 ± 0.034	0.258
3	0.050-0.075	306 ± 21	435 ± 25	0.175 ± 0.042	0.206
4	0.075-0.100	219 ± 18	276 ± 20	0.116 ± 0.052	0.136
5	0.100-0.125	165 ± 15	165 ± 16	0.000 ± 0.064	0.060
6	0.125-0.250	328 ± 21	248 ± 19	-0.139 ± 0.048	-0.075
D^0 Sample					
0	-0.025-0.000	259 ± 32	240 ± 32	-0.038 ± 0.090	0.163
1	0.000-0.025	1090 ± 53	1810 ± 62	0.249 ± 0.028	0.255
2	0.025-0.050	1060 ± 51	1878 ± 61	0.278 ± 0.026	0.266
3	0.050-0.075	806 ± 44	1325 ± 53	0.243 ± 0.032	0.260
4	0.075-0.100	583 ± 37	1013 ± 45	0.270 ± 0.036	0.250
5	0.100-0.125	377 ± 31	712 ± 38	0.308 ± 0.044	0.240
6	0.125-0.250	788 ± 45	1172 ± 51	0.196 ± 0.034	0.220

Table 5.6: For each VPDL bin, the measured number of D^* and D^0 events with opposite sign and same sign of the tagging variable range ($0.35 < |d| < 0.45$) N_i^{nos} , N_i^{osc} , the measured asymmetry A_i and expected asymmetry A_i^e corresponding to the fitted value $\Delta m_d = 0.486 \text{ ps}^{-1}$ are given.

events, since they have a VPDL distribution with zero mean and $\sigma \sim 150 \mu\text{m}$ according to our studies.

The following results were obtained:

$$\varepsilon \mathcal{D}_d^2 = (1.71 \pm 0.19)\%; \quad (5.36)$$

$$\Delta m_d = 0.502 \pm 0.025 \text{ ps}^{-1};$$

$$f_{c\bar{c}} = (3.1 \pm 1.4)\%.$$

One of the goals of this measurement is to verify the assumption of independence of the opposite-side flavor tagging on the type of the reconstructed

Bin No.	VPDL range, cm	N_i^{osc}	N_i^{nos}	A_i	A_i^e
D^* Sample					
0	-0.025-0.000	34 ± 8	67 ± 10	0.335 ± 0.118	0.264
1	0.000-0.025	152 ± 15	415 ± 24	0.465 ± 0.044	0.403
2	0.025-0.050	207 ± 16	410 ± 23	0.331 ± 0.042	0.383
3	0.050-0.075	118 ± 13	249 ± 19	0.359 ± 0.057	0.307
4	0.075-0.100	120 ± 13	174 ± 15	0.184 ± 0.065	0.205
5	0.100-0.125	77 ± 11	104 ± 12	0.152 ± 0.086	0.092
6	0.125-0.250	185 ± 16	170 ± 15	-0.043 ± 0.060	-0.104
D^0 Sample					
0	-0.025-0.000	66 ± 19	199 ± 24	0.507 ± 0.115	0.260
1	0.000-0.025	482 ± 35	1154 ± 47	0.411 ± 0.034	0.407
2	0.025-0.050	503 ± 34	1285 ± 48	0.437 ± 0.031	0.424
3	0.050-0.075	385 ± 30	893 ± 41	0.398 ± 0.038	0.414
4	0.075-0.100	263 ± 26	642 ± 35	0.420 ± 0.045	0.399
5	0.100-0.125	195 ± 22	475 ± 30	0.419 ± 0.052	0.382
6	0.125-0.250	356 ± 31	707 ± 38	0.330 ± 0.045	0.350

Table 5.7: For each VPDL bin, the measured number of D^* and D^0 events with opposite sign and same sign of the tagging variable range ($0.45 < |d| < 0.6$) N_i^{nos} , N_i^{osc} , the measured asymmetry A_i and expected asymmetry A_i^e corresponding to the fitted value $\Delta m_d = 0.486 \text{ ps}^{-1}$ are given.

B meson. It can be seen from Tables 5.9 and 5.10 that the measured flavor tagging performance for B^0 events is slightly better than for B^+ events, both for individual and combined taggers. This difference can be explained by a better selection of $\mu^+ \nu D^{*-}$ events due to an additional requirement of the charge correlation between muon and pion from $D^{*-} \rightarrow D^0 \pi^-$ decay. The D^0 sample can contain events with a wrongly selected muon. Since the charge of the muon determines the flavor asymmetry, such a background can reduce the measured B^+ dilution. The charge correlation between the muon and the pion can suppress this background and result in a better measurement of the tagging performance.

Bin No.	VPDL range, cm	N_i^{osc}	N_i^{nosc}	A_i	A_i^e
D^* Sample					
0	-0.025-0.000	29 ± 7	38 ± 8	0.133 ± 0.150	0.298
1	0.000-0.025	89 ± 12	248 ± 19	0.472 ± 0.058	0.455
2	0.025-0.050	121 ± 14	286 ± 20	0.407 ± 0.054	0.432
3	0.050-0.075	83 ± 11	178 ± 15	0.366 ± 0.065	0.344
4	0.075-0.100	73 ± 11	123 ± 13	0.256 ± 0.080	0.227
5	0.100-0.125	47 ± 9	61 ± 9	0.130 ± 0.117	0.097
6	0.125-0.250	111 ± 13	79 ± 11	-0.165 ± 0.085	-0.129
D^0 Sample					
0	-0.025-0.000	75 ± 16	100 ± 19	0.143 ± 0.138	0.261
1	0.000-0.025	300 ± 28	725 ± 38	0.415 ± 0.043	0.408
2	0.025-0.050	273 ± 27	729 ± 38	0.456 ± 0.044	0.426
3	0.050-0.075	234 ± 24	551 ± 32	0.404 ± 0.049	0.416
4	0.075-0.100	182 ± 21	366 ± 28	0.337 ± 0.060	0.401
5	0.100-0.125	130 ± 18	288 ± 24	0.378 ± 0.069	0.384
6	0.125-0.250	232 ± 26	512 ± 33	0.377 ± 0.054	0.351

Table 5.8: For each VPDL bin, the measured number of D^* and D^0 events with opposite sign and same sign of the tagging variable range ($|d| > 0.6$) N_i^{nosc} , N_i^{osc} , the measured asymmetry A_i and expected asymmetry A_i^e corresponding to the fitted value $\Delta m_d = 0.486 \text{ ps}^{-1}$ are given.

To test this hypothesis, a special sample of events satisfying all the conditions of the D^* sample, except for the requirement of the charge correlation between the muon and the pion, was selected. The dilution \mathcal{D}'_d for this sample is shown in Table 5.10. It can be seen that \mathcal{D}'_d is statistically compatible with \mathcal{D}_u for all samples and all taggers since the $\chi^2/\text{d.o.f}$ for the difference in dilutions is found to be 1.06. We can compare this $\chi^2/\text{d.o.f}$ to 1.27 in the case where only right-sign events are considered for the D^* sample.

This result confirms the assumption of the same performance of the opposite-side flavor tagging for B^+ and B^0 events. It also shows that unaccounted contribution of background in the D^0 sample reduces the measured

Sample	$\varepsilon(\%)$	\mathcal{D}_d	$\varepsilon\mathcal{D}_d^2(\%)$
$ d > 0.37$	10.98 ± 0.15	0.395 ± 0.022	1.71 ± 0.19
$0.20 < d < 0.35$	5.3 ± 0.1	0.110 ± 0.031	0.063 ± 0.035
$0.35 < d < 0.45$	6.2 ± 0.1	0.300 ± 0.064	0.56 ± 0.24
$0.45 < d < 0.60$	3.6 ± 0.1	0.440 ± 0.097	0.70 ± 0.31
$0.60 < d < 1.00$	2.3 ± 0.1	0.506 ± 0.044	0.58 ± 0.10

Table 5.9: Tagging performance for events with a reconstructed B^0 for different subsamples.

Sample	$\varepsilon(\%)$	\mathcal{D}_u	$\varepsilon\mathcal{D}_u^2(\%)$	\mathcal{D}'_d
$ d > 0.37$	11.67 ± 0.11	0.363 ± 0.012	1.540 ± 0.106	0.414 ± 0.025
$0.20 < d < 0.35$	5.46 ± 0.08	0.201 ± 0.020	0.220 ± 0.044	0.127 ± 0.034
$0.35 < d < 0.45$	6.49 ± 0.09	0.286 ± 0.018	0.531 ± 0.066	0.300 ± 0.015
$0.45 < d < 0.60$	3.83 ± 0.07	0.439 ± 0.020	0.740 ± 0.067	0.440 ± 0.030
$0.60 < d < 1.00$	2.39 ± 0.05	0.443 ± 0.027	0.469 ± 0.059	0.534 ± 0.048

Table 5.10: Tagging performance for events with reconstructed B^+ for different taggers and subsamples. For comparison, the dilution \mathcal{D}'_d measured in the D^* sample with addition of wrong sign $\mu^+\nu\bar{D}^0\pi^+$ events is also shown.

Sample	$\Delta m_d \text{ ps}^{-1}$	$f_{c\bar{c}}$
$ d > 0.37$	0.502 ± 0.025	0.031 ± 0.014
$0.20 < d < 0.35$	0.448 ± 0.112	0.219 ± 0.145
$0.35 < d < 0.45$	0.561 ± 0.041	0.076 ± 0.042
$0.45 < d < 0.60$	0.448 ± 0.036	0.000 ± 0.057
$0.60 < d < 1.00$	0.488 ± 0.041	0.054 ± 0.031

Table 5.11: Measured value of Δm_d and $f_{c\bar{c}}$ for different subsamples.

dilution for B^+ events. This background is suppressed by the requirement of the charge correlation between the muon and the pion. Thus, the dilution measured in the D^* sample can be used for the B_s mixing measurement, where a similar charge correlation between the muon and D_s is required.

The dilution for any event should strongly depend on the magnitude of the variable d . This property becomes important in the B_s mixing measurement, since in this case the dilution of each event can be estimated using the value of d and can be included in the likelihood function, improving the sensitivity of the measurement. To calibrate the dependence of the dilution on d , all tagged events were divided into subsamples with $0.2 < |d| < 0.35$, $0.35 < |d| < 0.45$, $0.45 < |d| < 0.6$, and $|d| > 0.6$. The overall efficiency of this sample is $(17.4 \pm 0.2)\%$. The dilutions obtained are shown in Table 5.9. Their strong dependence on the value of the tagging variable is clearly seen. The overall tagging power, computed as the sum of tagging powers of all subsamples is:

$$\varepsilon \mathcal{D}_d^2 = (1.90 \pm 0.41)\%. \quad (5.37)$$

The measured oscillation parameter Δm_d for all considered taggers and subsamples is given in Table 5.11. It is compatible with the world average value $\Delta m_d = 0.509 \pm 0.004 \text{ ps}^{-1}$. This indicates that the estimate of the tagging power, $\varepsilon \mathcal{D}_d^2$ is reliable.

Finally, the mixing parameter Δm_d was obtained from the simultaneous fit of the flavor asymmetry in the subsamples defined above. The fraction $f_{c\bar{c}}$ was constrained to be the same for all subsamples. The result obtained is:

$$\begin{aligned} \Delta m_d &= 0.486 \pm 0.021 \text{ ps}^{-1}; \\ f_{c\bar{c}} &= (2.5 \pm 1.1)\%. \end{aligned} \quad (5.38)$$

The statistical precision of Δm_d from the simultaneous fit is about 10% better than that from the fit of events with $|d| > 0.3$. This improvement is directly related with a better overall tagging power (5.37) for the sum of

subsamples as compared to the result (5.36) for the sample with $|d| > 0.3$.

Dilution Calibration

We calibrate the multidimensional flavor tagging by fitting the dependence of the measured dilution in the $D^*\mu X$ channel, \mathcal{D}_d , on the absolute value of the tagger output variable, d , to a functional form. We can then use this function to obtain an event-by-event dilution for the Δm_s analysis. In Fig. 5.13, we see that a linear function describes the calibration well, with a χ^2 of 0.9375 for 2 degrees of freedom. Linear dependence is expected because the multidimensional likelihoods take correlations between discriminants into account. Ignoring these correlations can introduce non-linearities into the tagger output, with a corresponding increase in the systematic error assigned to the dilution calibration.

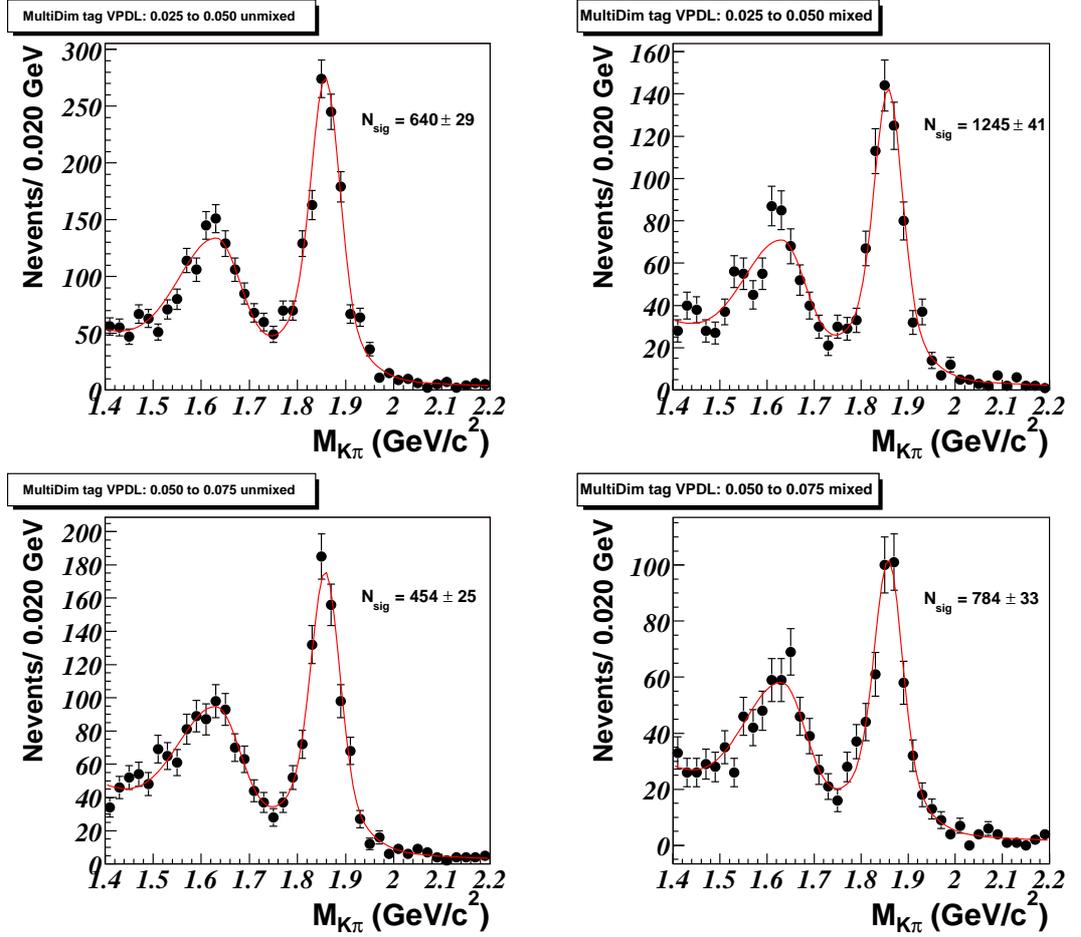


Figure 5.8: The fit to $M(K\pi)$ mass for non-oscillating (left) and oscillating (right) for μ^+D^{*-} events tagged by the multidimensional tagger with $|d| > 0.37$ for the VPDL ranges $\{0.025 - 0.050\}$ and $\{0.050 - 0.075\}$ (cm).

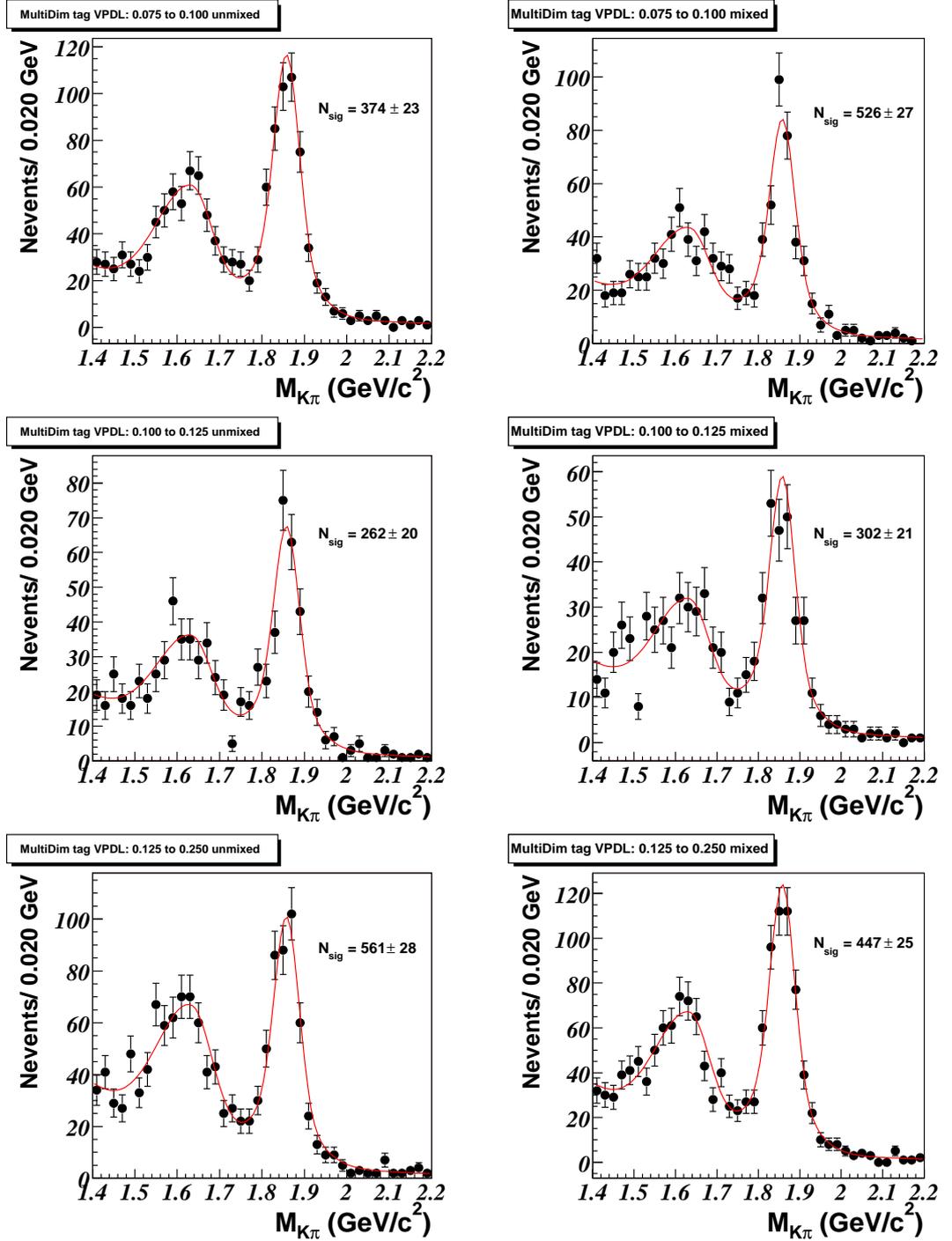


Figure 5.9: The fit to $M(K\pi)$ mass for non-oscillating (left) and oscillating (right) for μ^+D^{*-} events tagged by the multidimensional tagger with $|d| > 0.37$ for the VPDL ranges $\{0.075-0.100\}$, $\{0.100-0.125\}$, and $\{0.125-0.250\}$ (cm).

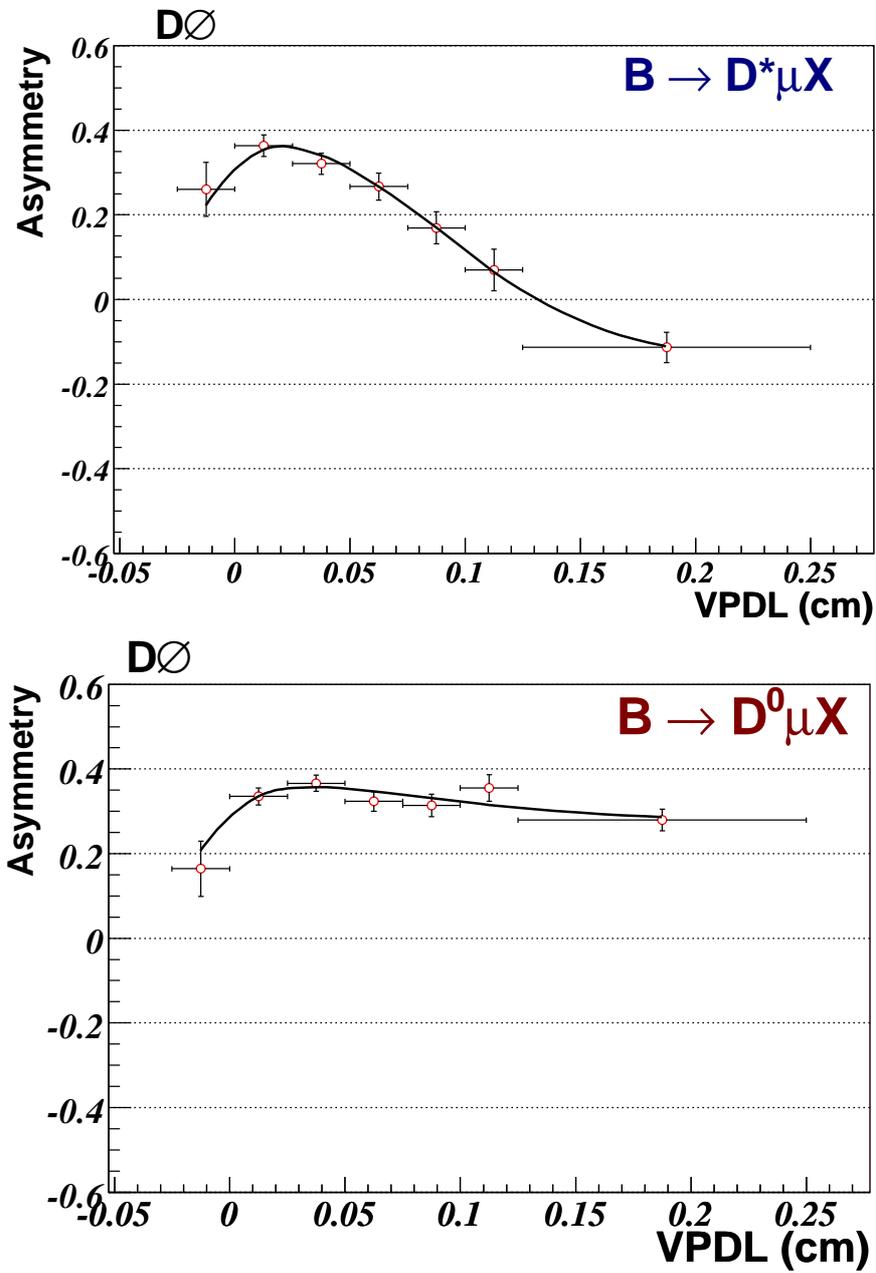


Figure 5.10: The asymmetries obtained in the D^* (top plot) and D^0 (bottom plot) samples with the result of the fit superimposed for the $|d| > 0.37$ sample.

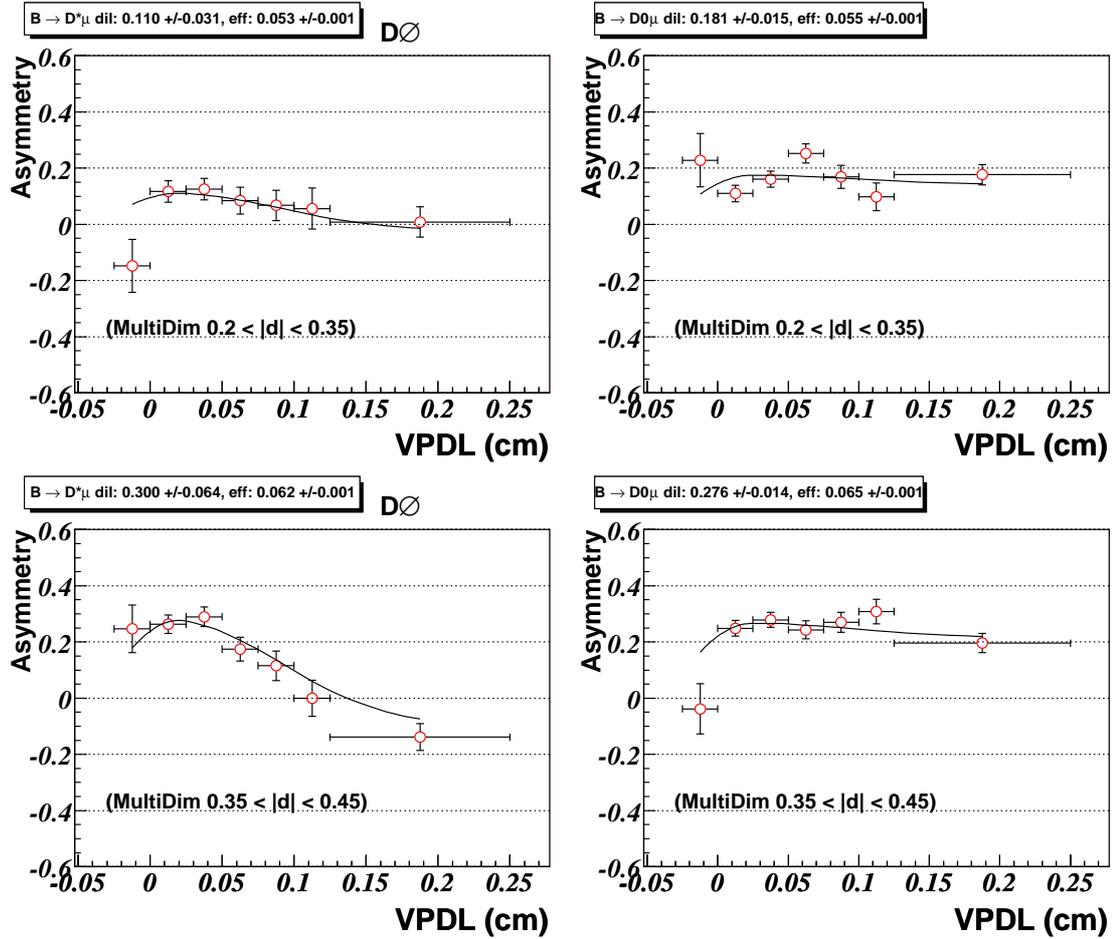


Figure 5.11: The asymmetries obtained in the D^* and D^0 samples with the result of the fit superimposed for the $0.2 < |d| < 0.35$ and $0.35 < |d| < 0.45$ samples.

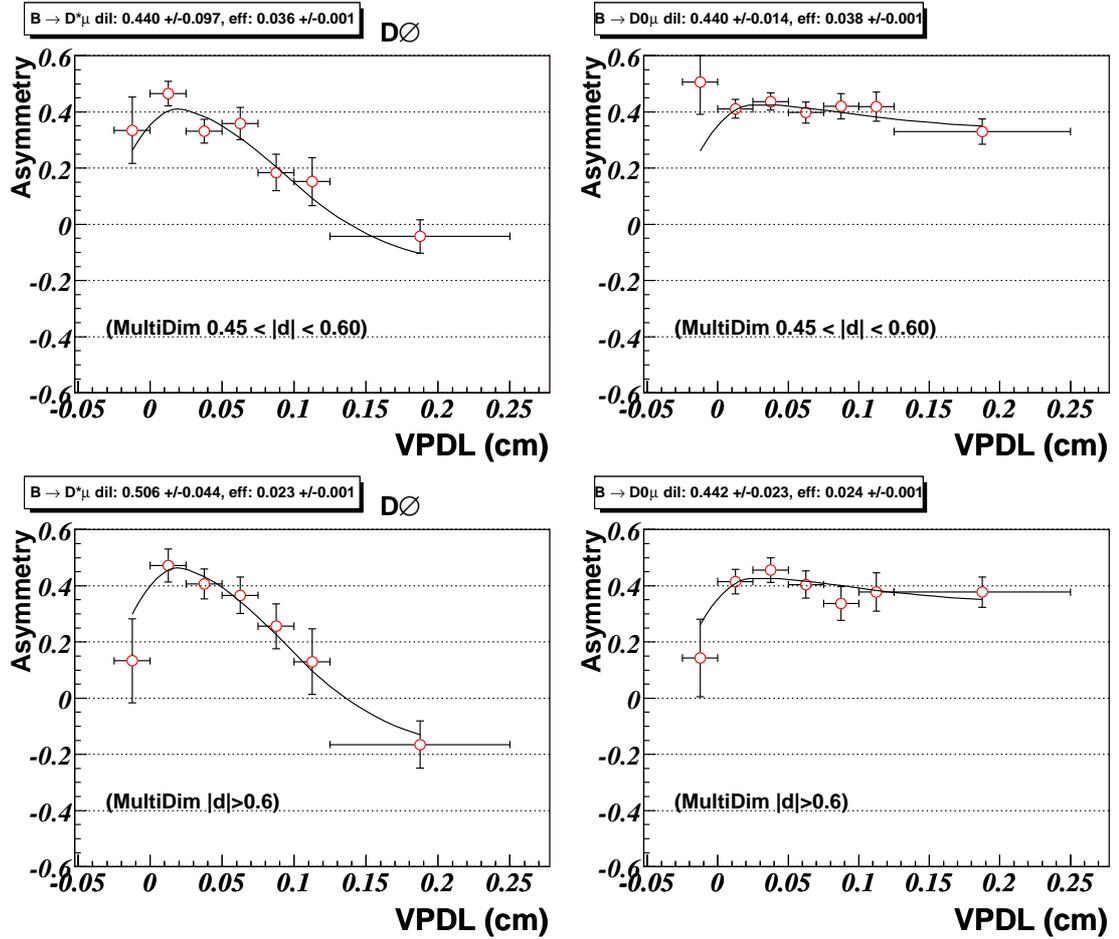


Figure 5.12: The asymmetries obtained in the D^* and D^0 samples with the result of the fit superimposed for the $0.45 < |d| < 0.60$ and $|d| > 0.60$ samples.

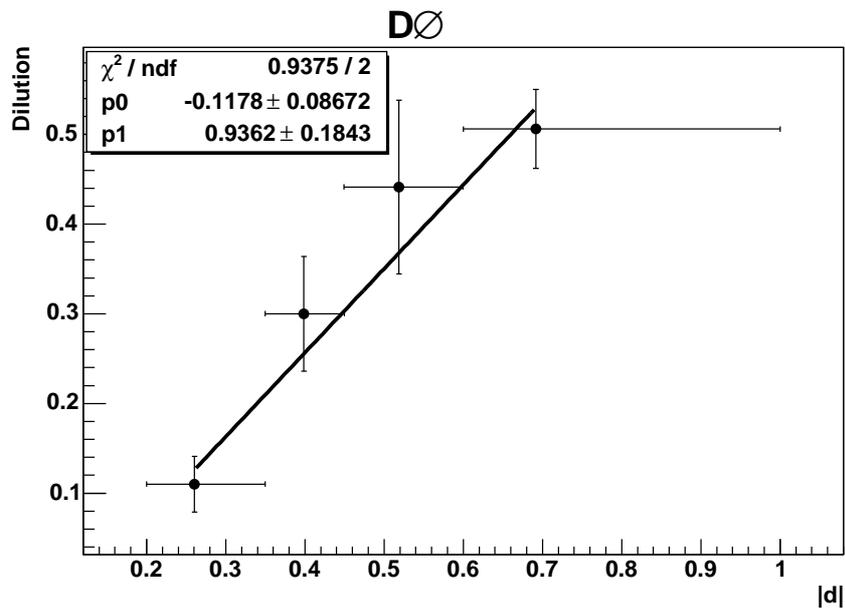


Figure 5.13: Measured dilution of the multidimensional flavor tagger in the $D^*\mu X$ sample (referred to as \mathcal{D}_d in the text) versus the absolute value of the tagger output variable d . A linear fit is superimposed.

Chapter 6

B_s^0 Oscillation Analysis

We search for B_s^0 oscillations in the decay channel $B_s^0 \rightarrow D_s^- \mu^+ X$, $D_s^- \rightarrow K_S^0 K^-$. The branching ratio for this decay is 30% smaller [13] than the $D_s \rightarrow \phi \pi$, $\phi \rightarrow K^+ K^-$ channel analyzed in [16]. Furthermore, due to the presence of the long-lived ($c\tau = 2.7$ cm) K_S^0 , the reconstruction efficiency for this channel is lower than in Ref. [16]. We therefore expect to find fewer events in this channel and consequently a smaller significance (see Eq. 3.20) than in Ref. [16]. Although the expected sensitivity will therefore be low, the result is relevant because it can be combined with the other B_s^0 modes: $\phi \pi \mu$, $\phi \pi e$, and $K^* K \mu$ for greater sensitivity [51].

To reconstruct this channel, we will first search for a K_S^0 candidate decaying to $\pi^+ \pi^-$ and then combine this candidate with a third track, assumed to be kaon, to form a D_s candidate. The D_s is then combined with a muon to form the B_s candidate. We then flavor tag the candidate as described in Sec. 6.3 to obtain the b flavor at the time of production. The flavor at decay is given by the charge of the muon in the semileptonic decay; by comparing the production and decay flavor of the candidate, we are able to classify it as oscillated or non-oscillated. We next perform an unbinned likelihood framework to measure the B_s^0 lifetime and Δm_d as cross-checks. Finally, we use the unbinned likelihood framework to perform an amplitude scan (see

Sec. 6.8) to search for B_s^0 oscillations.

6.1 Data Sample and Event Selections

This measurement uses the large semileptonic sample corresponding to approximately 1.2 fb^{-1} of integrated luminosity, accumulated by the DØ detector during the period from April 2002 to February 2006, also referred to as Run IIa. B_s^0 hadrons were selected in the semileptonic channel $B_s^0 \rightarrow D_s^- \mu^+ \nu X$ where $D_s^- \rightarrow K_S^0 K^-$. Due to the long K_S^0 lifetime, it is possible to obtain a pure sample of K_S^0 's by cutting on their decay length. When this K_S^0 candidate is combined with another track, assumed to be a kaon, and then combined with a muon, we obtain a B_s^0 sample with a good signal to background ratio.

The reconstruction begins with the identification of a muon candidate through the standard DØ algorithm [47]. The requirements on the muon are as follows:

- $p_T > 2 \text{ GeV}/c$;
- $|\vec{p}| > 3 \text{ GeV}/c$;
- hits in both the CFT and SMT; and
- at least two measurements in the muon chambers.

Next, we search for a K_S^0 decaying to $\pi^+ \pi^-$. We consider all tracks that share the same primary vertex as the muon. The cuts on the two tracks composing the K_S^0 are as follows:

- each track must have at least 4 two-dimensional hits, at least 2 of which must be CFT hits;
- the tracks must have opposite charge;

- for the two-pion mass hypothesis, $460 < m(\pi_1, \pi_2) < 525 \text{ MeV}/c^2$;
- to reduce the contribution from photon conversions we require candidates to have a two-photon mass hypothesis $m(\gamma_1, \gamma_2) > 25 \text{ MeV}/c^2$;
- the combined significance of the transverse and the longitudinal impact parameter projection with respect to the primary vertex, $\epsilon_{sig} = \sqrt{(d_{R\phi}/\sigma(d_{R\phi}))^2 + (d_{Rz}/\sigma(d_{Rz}))^2}$ was required to be greater than 3 for each track;
- events that had $\epsilon_{sig} < 4$ for either track were rejected;
- $p_T(K_S^0) > 650 \text{ MeV}/c$;
- the transverse decay length of the K_S^0 is required to be greater than 0.3 cm;
- the angle $\alpha_T^{K_S^0}$ between the K_S^0 momentum and direction from the primary to the K_S^0 vertex in the transverse plane was required to satisfy $\cos(\alpha_T^{K_S^0}) > 0.8$;
- the transverse decay length significance of the K_S^0 candidate was required to be $d_T^{K_S^0}/\sigma(d_T^{K_S^0}) > 2$; and
- the K_S^0 is constrained to its nominal mass [13].

The invariant mass of the K_S^0 candidates, with all of the above applied cuts, is shown in Fig. 6.1.

The K_S^0 is then combined with a third track, the kaon candidate, to form a D_s^- . The cuts on the kaon candidate and D_s^- are as follows:

- $p_T(K) > 1.5 \text{ GeV}/c$;
- $\epsilon_{sig}(K) > 2$;
- the χ^2 of the vertex fit was required to be $\chi^2 < 16$;

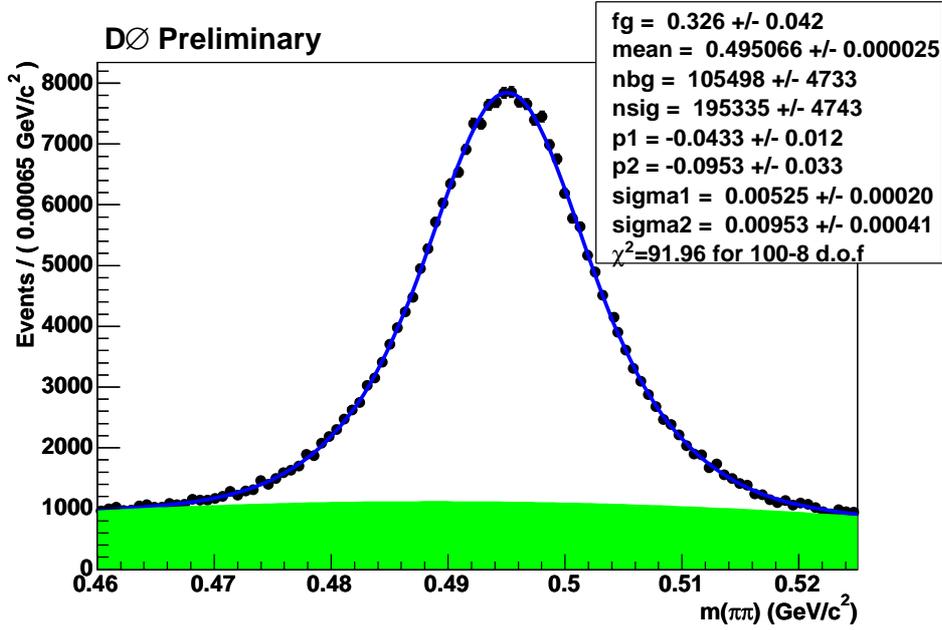


Figure 6.1: Invariant mass of the K_S^0 candidates, $m(\pi^+\pi^-)$. The fit indicates that there are 195335 events in the K_S^0 signal. Note that this spectrum is made after all preselections in Sec. 6.1, but before application of the likelihood ratio selections described in Sec. 6.1.1.

- the transverse decay length significance of the D_s^- candidate was required to be, $d_T^D/\sigma(d_T^D) > 4$; and
- the angle α_T^D between the momentum of the D_s^- candidate and the direction from the primary to the D_s^- vertex in the transverse plane was required to be $\cos(\alpha_T^D) > 0.9$.

Finally, the D_s^- is combined with the μ^+ to form the B_s^0 candidate. The cuts on the $\mu^+ + D_s^-$ vertex are as follows:

- $q(\mu) \cdot q(K) < 0$;
- the mass of the $\mu^+ + D_s^-$ system is required to be in the range $2.6 < m(\mu D_s) < 5.4 \text{ GeV}/c^2$, i.e., close to the nominal B_s^0 mass or less;

- the χ^2 of the vertex fit is required to be $\chi^2 < 9$;
- if the transverse distance d_T^B between the primary and B vertex exceeds $4\sigma(d_T^B)$, the angle α_T^B between the B momentum and direction is required to satisfy $\cos(\alpha_T^B) > 0.95$;
- if $d_T^B > d_T^D$, the transverse distance significance between the B and D is required to be $d_T^{BD} < 2$;
- the isolation, defined as $\text{Iso} = p(\mu D_s) / (p(\mu D_s) + \sum p_i)$ where $\sum p_i$ is taken over all charged particles in the cone $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$ and $\Delta\eta(\phi)$ is the pseudorapidity (azimuthal angle) with respect to the μD_s direction, is required to be $\text{Iso} > 0.3$.

Candidates passing these cuts were then passed through a likelihood ratio selection as described in the next section.

6.1.1 Likelihood Ratio Method and Selection

To further increase the signal-to-noise ratio, we construct a multivariate selection. We choose a set of discriminating variables x_1, \dots, x_n for each event and construct probability density functions for signal, $f^s(x_i)$, and background, $f^b(x_i)$. We then define a combined selection variable y as,

$$y = \prod_{i=1}^n y_i; \quad y_i = \frac{f_i^b(x_i)}{f_i^s(x_i)}. \quad (6.1)$$

In case a variable x_i cannot be constructed for a particular event, we set the corresponding y_i to 1. We select signal events by applying a cut on the combined variable, $y < y_0$ [52].

The following discriminating variables were used in the construction of the likelihood ratio probability density functions (*pdf*'s):

- $p_T(K)$;

- $p_T(K_S^0)$;
- The transverse decay length of the K_S^0 ;
- $m(\pi_1, \pi_2)$ of the K_S^0 candidate;
- The χ^2 of the D_s^- vertex fit;
- The isolation of the B as defined in Sec. 6.1;
- $m(\mu D_s)$.

The probability density functions were constructed using data events. We define signal (S) and background (B) regions as:

$$\begin{aligned}
 B & : 1.90 < M(D_s^-) < 2.02 \text{ GeV}, q_\mu \cdot q_K < 0 \text{ (right-sign)}; \\
 S & : 1.90 < M(D_s^-) < 2.02 \text{ GeV}, q_\mu \cdot q_K > 0 \text{ (wrong-sign)}.
 \end{aligned}$$

Note that these band definitions arise from an estimate of $\pm 3\sigma$ of the D_s mass peak, where σ is the width of the peak. The signal probability density function was constructed by subtracting the distributions of events in region B from the distribution of events in region S . In Fig. 6.2 we see the right- and wrong-sign combinations of events passing the pre-selections in Sec. 6.1 before applying likelihood ratio selections. We use these events to extract discriminants from which we calculate likelihood ratios. The normalized distributions for all discriminants are shown in Fig. 6.3.

Figure 6.4 shows an estimate of the signal significance versus $\log_{10} y$. The significance is estimated as $N_{sig}/\sqrt{N_{sig} + N_{bg}}$ with the signal (background) sample defined as the right-sign (wrong-sign) sample. The number of events is obtained by counting events in the signal region and scaling the background sample to the right-sideband region ($2.2 < M(K_S^0 K) < 2.4$) GeV). We use this figure to determine our final cut on the combined variable, $\log_{10} y < -0.08$.

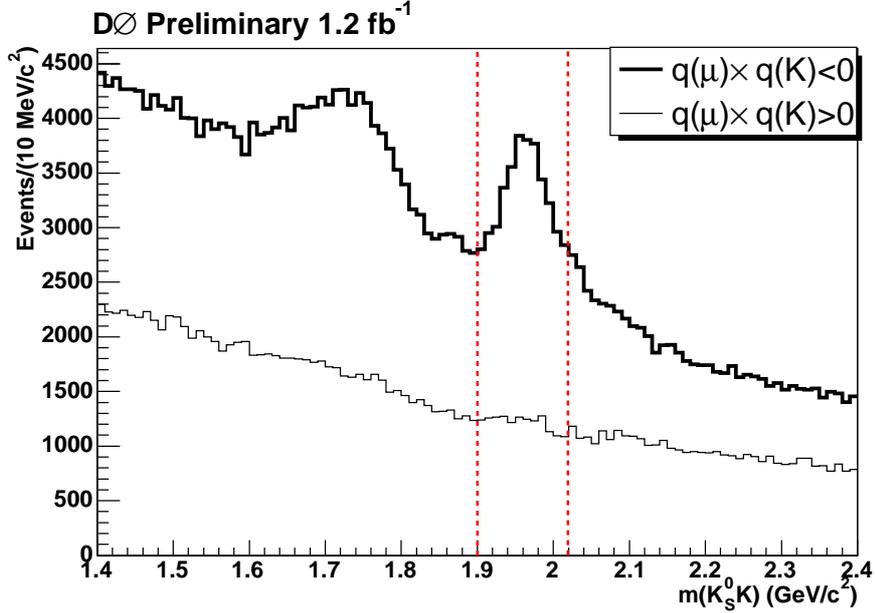


Figure 6.2: Distributions of $m(K_S^0 K)$ for right-sign ($q_\mu \cdot q_K < 0$) and wrong-sign ($q_\mu \cdot q_K > 0$) candidates after the pre-selections described in Sec. 6.1 and before likelihood ratio selections. The mass band used to define the signal and background regions for the likelihood ratio technique is indicated by dashed lines.

Figure 6.5 shows distributions of $\log_{10} y$ for signal and background regions as defined below,

$$\begin{aligned}
 B_L & : 1.400 < M(D_s^-) < 1.434 \text{ GeV}/c^2 \\
 S & : 1.89 < M(D_s^-) < 2.05 \text{ GeV}/c^2 \\
 B_R & : 2.22 < M(D_s^-) < 2.40 \text{ GeV}/c^2.
 \end{aligned} \tag{6.2}$$

In Fig. 6.6 the mass spectrum of $m(K_S^0 K)$ after applying final likelihood ratio selections is shown. We can see that the signal to background ratio improved after the application of likelihood ratio selections by comparing

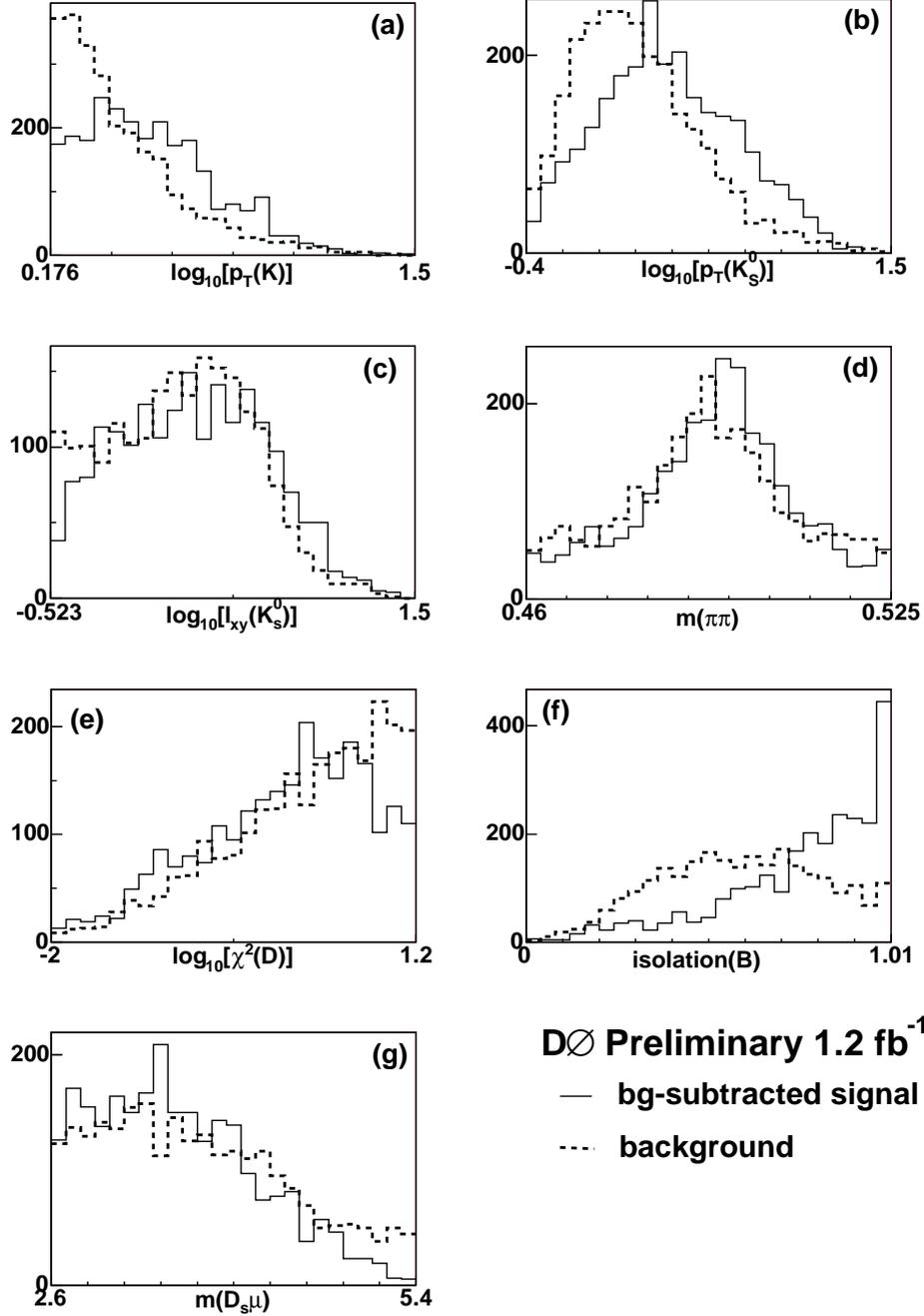


Figure 6.3: $B_s^0 \rightarrow D_s^- \mu^+ X$, $D_s^- \rightarrow K_s^0 K^-$ signal and background for the discriminants used in the likelihood ratio selections. (a) $\log_{10}[p_T(K)]$, (b) $\log_{10}[p_T(K_s^0)]$, (c) $l_{xy}(K_s^0)$, (d) $m(\pi_1, \pi_2)$, (e) $\log_{10}[\chi^2(D \text{ vertex})]$, (f) isolation, (g) $m(\mu + D_s)$. Note that the scale on all plots is arbitrary.

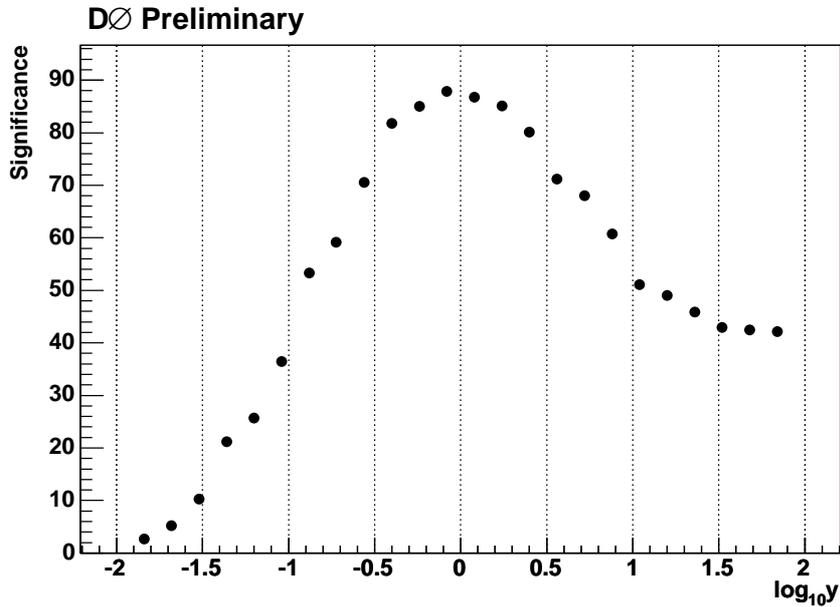


Figure 6.4: Signal significance versus likelihood ratio selection variable, $\log_{10} y$. The significance here is estimated by $N_{sig}/\sqrt{N_{sig} + N_{bg}}$. The number of signal events is obtained by counting the number of events in the right-sign sample ($q_\mu \cdot q_K < 0$) and the number of background events is similarly taken from the wrong-sign sample ($q_\mu \cdot q_K > 0$). The number of background events is scaled so that the number of events in the right-sideband ($2.2 < M(K_S^0)K < 2.4$ GeV) is the same for both samples.

Fig. 6.6 to Fig. 6.2. In Fig. 6.7 we overlay the mass spectrum before and after likelihood ratio selections are applied. Using techniques described in Sec. 6.2, we estimate that the signal to background ratio in the $\pm 3\sigma$ region around the D_s signal mass peak goes from $S/B = 0.06$ to $S/B = 0.17$ after the application of likelihood ratio selections.

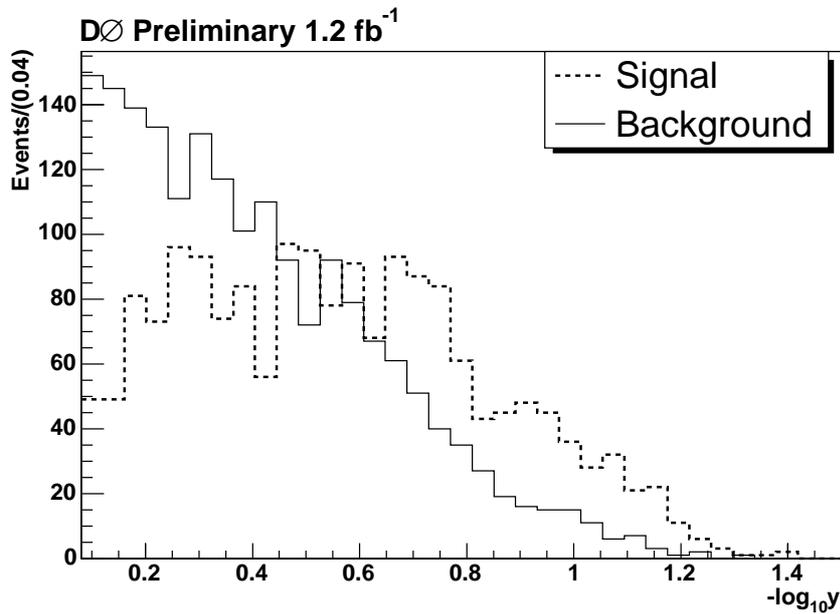


Figure 6.5: The distribution of the combined selection variable, $-\log_{10} y$ for signal and background.

6.2 Mass Fitting Procedure

$D\bar{0}$ does not have a system that can be used for particle identification. When we select the $D_s \rightarrow K_S^0 K$, we therefore must assume that a track is a kaon when it may not be. The mass spectrum shown in Fig. 6.6 contains contributions from modes other than the $D_s^- \rightarrow K_S^0 K^-$ channel, some of which are due to the misassignment of tracks as kaons. The main contributions to the invariant mass spectrum are listed below:

1. $D_s \rightarrow K_S^0 K$, signal;
2. $D^+ \rightarrow K_S^0 \pi$;
3. $D^+ \rightarrow K_S^0 K$, which we refer to as the *Cabbibo-suppressed* mode;
4. $\Lambda_c \rightarrow K_S^0 p$;

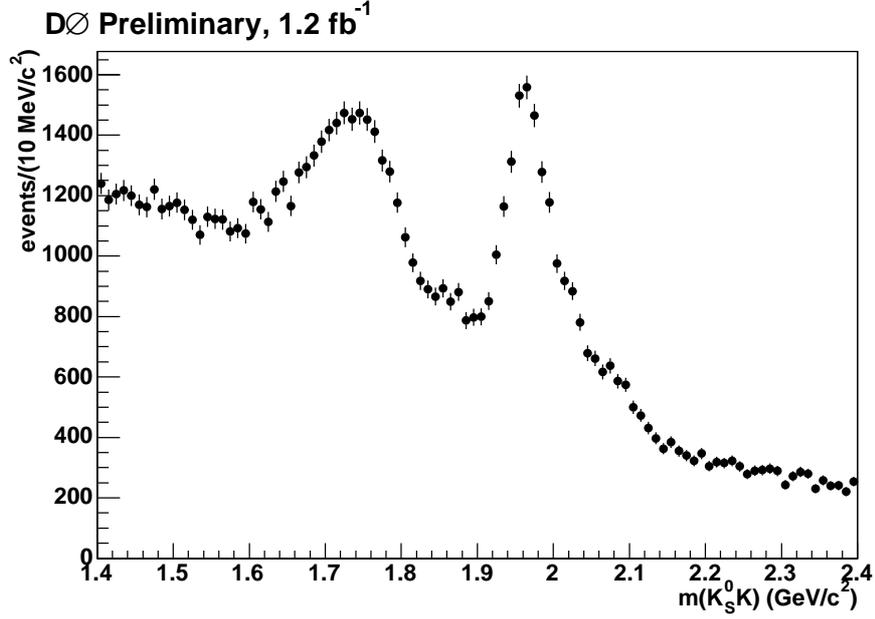


Figure 6.6: Mass spectrum for $m(K_S^0 K)$ after all selections, including the likelihood ratio selection..

5. $D^{(0,+,*)} \rightarrow K_S^0 \pi X$, which we refer to as the *low-mass* mode.

We have developed an unbinned likelihood technique to separate these kinematic reflections which we describe below.

6.2.1 Description of Technique

Consider a decay $X \rightarrow K_S^0 + \text{track}$, where the X can be D_s , D^+ , or Λ_c . Let the track be identified as a K . For the $D^+(\Lambda_c)$ system this would be a misassignment of the $\pi(p)$. The mass of the $K_S^0 + \text{track}$ system is then:

$$M_{mis}^2 = M_{K_S}^2 + M_K^2 + 2E_{K_S}E_{mis} - 2\vec{p}_{K_S} \cdot \vec{p}_{trk} \quad (6.3)$$

$$= M^2(X) + M_K^2 - M_{trk}^2 + 2E_{K_S}E_K - 2E_{K_S}E_{trk}. \quad (6.4)$$

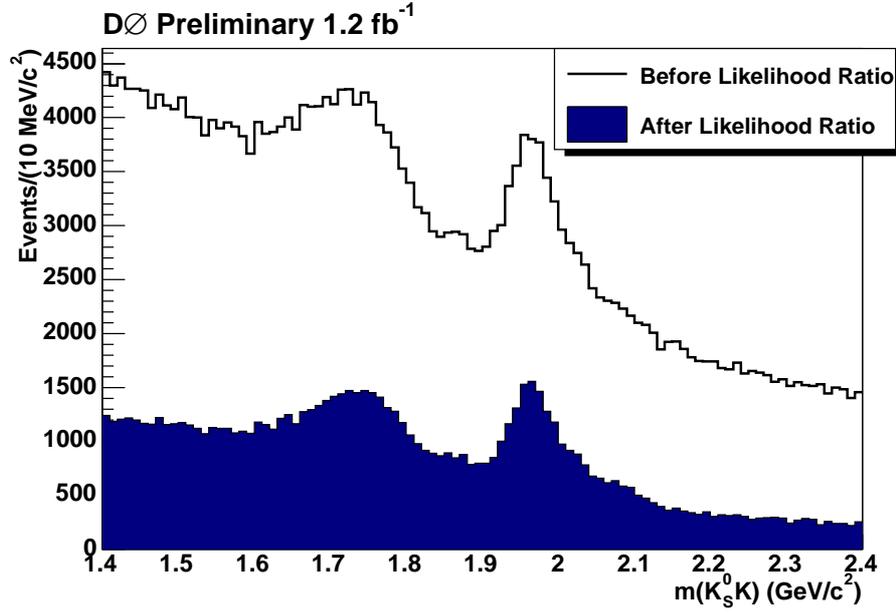


Figure 6.7: Mass spectrum for $m(K_S^0 K)$ before and after likelihood ratio selections.

We can Taylor expand the relativistic energy in M/p as

$$E = p \left[1 + \frac{1}{2} \left(\frac{M}{p} \right)^2 + \dots \right], \quad (6.5)$$

so that we can write:

$$\begin{aligned} 2E_{K_S} (E_K - E_{trk}) &= 2p_{K_S} \left[1 + \frac{1}{2} \left(\frac{M_{K_S}}{p_{K_S}} \right)^2 \right] \left(p_{trk} + \frac{M_K^2}{2p_{trk}} - p_{trk} - \frac{M_{trk}^2}{2p_{trk}} \right) \\ &= \frac{p_{K_S}}{p_{trk}} \left[1 + \frac{1}{2} \left(\frac{M_{K_S}}{p_{K_S}} \right)^2 \right] (M_K^2 - M_{trk}^2). \end{aligned} \quad (6.6)$$

We can then write Eq. 6.4 as:

$$\mathcal{M}^2 = M^2(X) + (M_K^2 - M_{trk}^2) \left[1 + \frac{p_{K_S}}{p_{trk}} \left(1 + \frac{M_{K_S}^2}{2p_{K_S}^2} \right) \right]. \quad (6.7)$$

If $M_{K_S} \ll p_{K_S}$ we then have:

$$\mathcal{M}^2(\lambda) = M_X^2 + \left(\frac{2}{1-\lambda} \right) (M_K^2 - M_{trk}^2), \quad (6.8)$$

where $\lambda = (p_{K_S} - p_{trk}) / (p_{K_S} + p_{trk})$ is the momentum asymmetry.

We can then insert a kinematic term in the likelihood as follows:

$$\mathcal{L}_i^{mass} = P_i(\lambda) \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{M_{measured}(K_S K) - \mathcal{M}_i(\lambda)}{\sigma} \right)^2 \right], \quad (6.9)$$

where the i represents the various modes, $P_i(\lambda)$ are distributions obtained from Monte Carlo, and $\mathcal{M}_i(\lambda)$ is Eq. 6.8 for the D^+ and Λ_c modes. The mass of the D_s and Cabbibo-suppressed mode do not depend on λ and therefore the likelihoods for them are simply taken as double Gaussians in mass. We use a bifurcated Gaussian in mass to model the low-mass peak.

The functions $P_i(\lambda)$ are formed by fitting polynomials to Monte Carlo distributions of λ . The function for the low mass peak, $P_{low}(\lambda)$, is constructed using information from $D^+ \rightarrow K_S^0 \pi^+ \pi^0$ Monte Carlo.

6.2.2 Tests of the Technique

We use $D^+ \rightarrow K_S^0 \pi^+$ Monte Carlo to form a profile histogram of $m(K_S^0 K)$ vs. λ which we then fit to Eqn. 6.8. We fit with the mass of the D^+ allowed to float and obtain the same mass as when we fit the distribution of $m(K_S^0 \pi)$. The result is shown in Fig. 6.8.

Because the mass of the proton from $\Lambda_c \rightarrow K_S^0 p$ decays is large compared to the pion mass, the Taylor expansion in Eqn. 6.6 does not work as well as

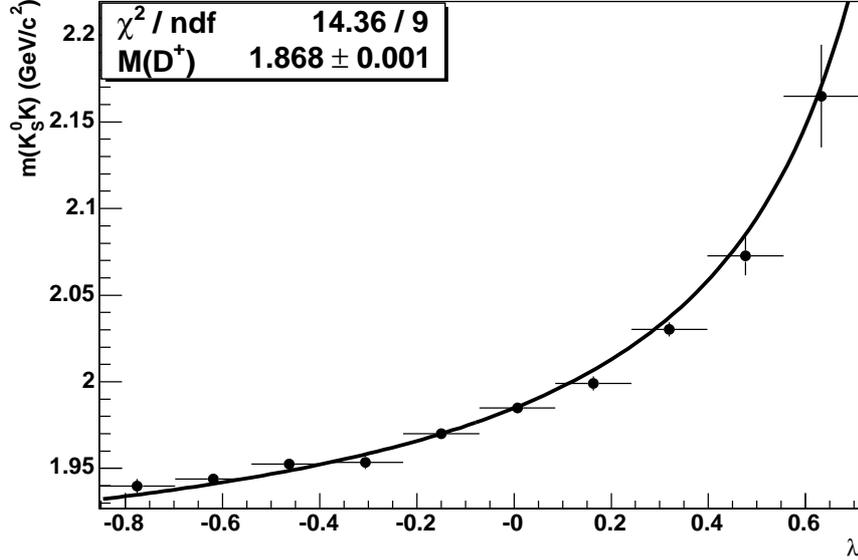


Figure 6.8: The fit of $D^+ \rightarrow K_S^0 \pi$ Monte Carlo to Eqn. 6.8. The mass of the D^+ is left allowed to vary in the fit.

it does for the $D^+ \rightarrow K_S^0 \pi^+$. We find that we have to modify Eqn. 6.8 by adding two ad-hoc correction terms:

$$\mathcal{M}^2(\lambda) = M_X^2 + \left(\frac{2}{1-\lambda} \right) (M_K^2 - M_{trk}^2) + a \cdot \frac{1-\lambda}{1+\lambda} + b \quad (6.10)$$

for the Λ_c mode. We fit for these two terms in Λ_c Monte Carlo, fixing $m(\Lambda_c)$ to the value obtained from fitting the $m(K_S^0 p)$ spectrum.

To test for possible biases, we perform an ensemble test of 1000 simulated experiments. The component yields are modelled as Gaussians centered on the yields found in the full tagged sample with the fitted uncertainty in the data used as the width of the Gaussians. An example experiment is shown in Fig. 6.10.

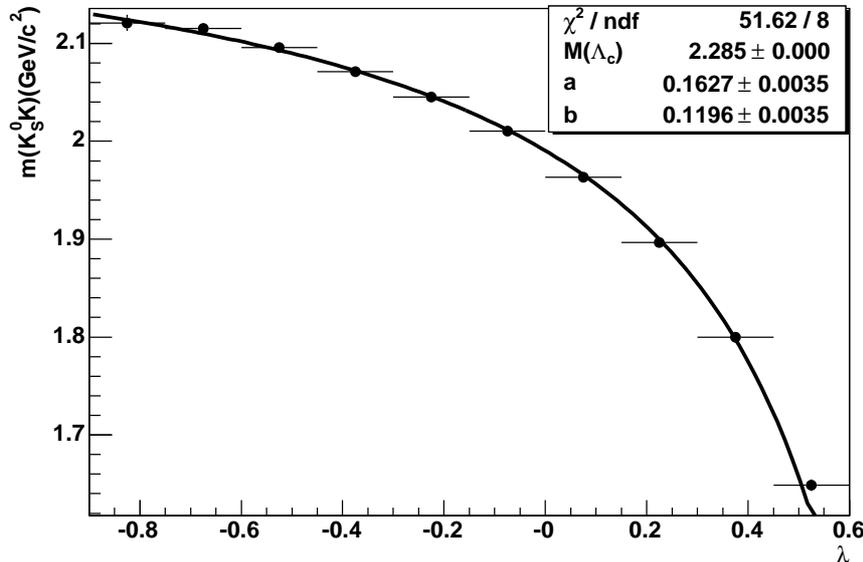


Figure 6.9: The fit of $\Lambda_c \rightarrow K_S^0 p$ Monte Carlo to Eqn. 6.10. The ad-hoc terms a and b are left floating in the fit.

The pull of the D_s signal yield shows no bias, as seen in Fig. 6.11. However, the width of the pulls in the ensemble test indicates that yield errors are underestimated by 15%. We account for this effect as well as other effects of variations in yields in the systematic uncertainty of the mixing measurement.

6.2.3 Application to Data

Figure 6.12 shows a fit to the entire untagged data sample. We found 2603 ± 110 signal events in the sample. We modelled the combinatorial background with a third-order Chebyshev polynomial. It should be noted that when we fit the data, we constrain the yield of the Cabbibo-suppressed mode to $0.13 \times N(D^+ \rightarrow K_S^0 \pi)$. The multiplicative constraint 0.13 is obtained by comparing branching ratios and Monte Carlo reconstruction efficiencies

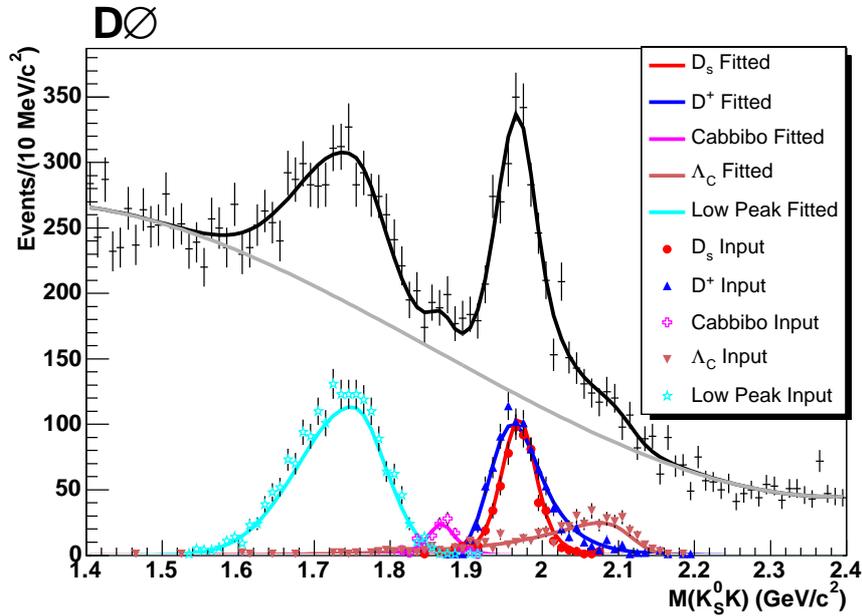


Figure 6.10: An example experiment from an ensemble test of the asymmetry mass fitting technique. Simulated input data is represented as markers while projections of the fitted yields are drawn as functions.

for the two modes. We varied this multiplicative constraint as a systematic uncertainty.

6.3 Initial State Flavor Tagging

A short description of the data-based flavor tagger mentioned in Ch. 5 is presented here. The discriminants used in this tagger are the same as those used for the Monte Carlo based tagger of Ch. 5. The differences between the tagger described there and the one described here are:

- the samples from which the taggers are created;
- the way in which tagging information is combined.

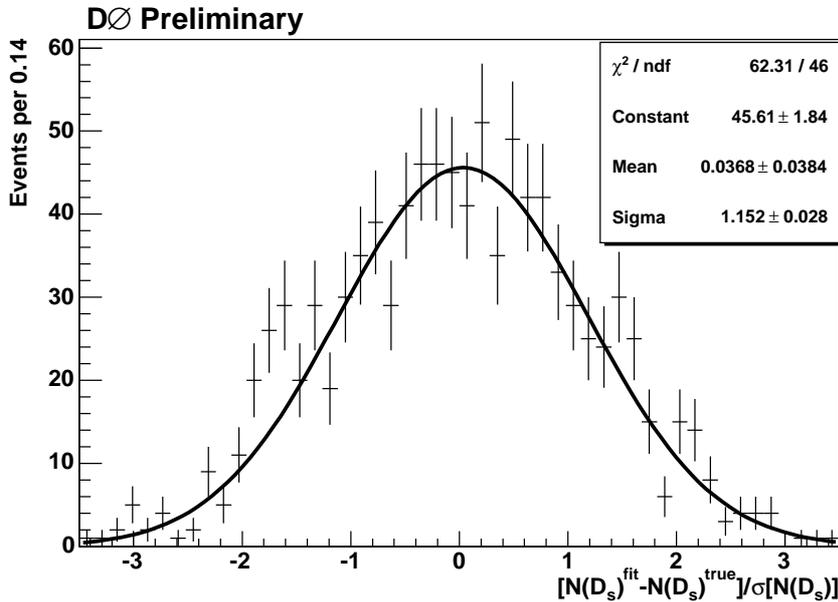


Figure 6.11: The pull of the signal yield, $(N(D_s)^{\text{fitted}} - N(D_s)^{\text{true}}) / \sigma(D_s)$, from 1000 simulated experiments fitted to a Gaussian. No signal yield bias is observed, but the fitted width indicates that errors are underestimated by 15%.

The data-based tagger uses one-dimensional probability density functions. The probability density function for each discriminating variable was constructed using events from the D^0 sample as defined in Sec. 5.7.1 with a cut applied to the visible proper decay length of the B candidate of $0 < x^M < 500 \mu\text{m}$. The decays $B^+ \rightarrow \mu^+ \nu \bar{D}^0$ compose the majority of this sample (see Sec. 5.7.5), while $B_d^0 \rightarrow \mu^+ \nu D^{*+}$ events make up 16% of the sample. Due to the cut on the visible proper decay length, the B_d^0 contribution is dominated by non-oscillated decays. Therefore, the initial flavor of a b -quark is determined by the charge of the muon. According to the MC estimates, the purity of such identification of the initial flavor in the selected sample is 0.98 ± 0.01 , where the error reflects the uncertainty in branching ratios of B

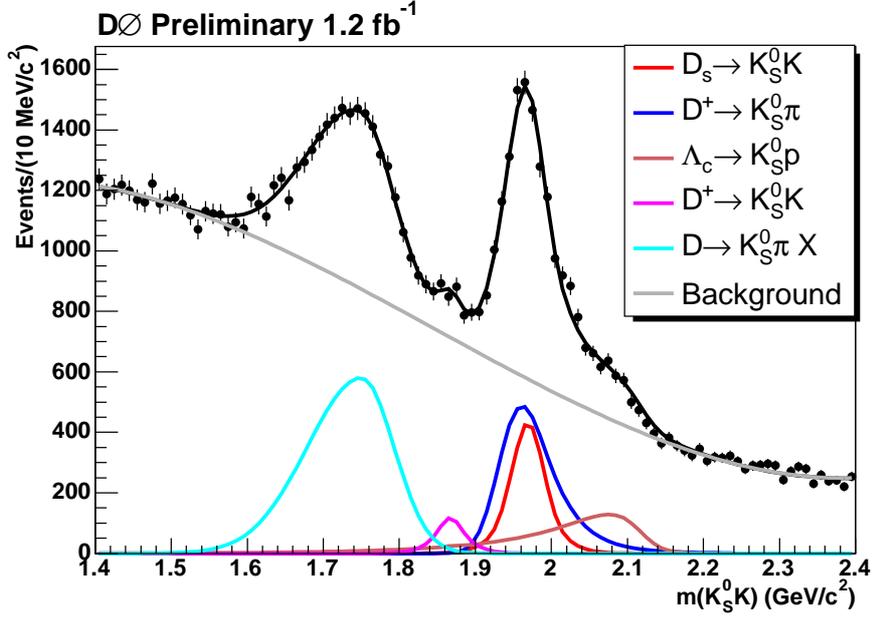


Figure 6.12: A fit to the entire untagged data sample.

decays.

For each discriminating variable, the signal band containing all events with $1.80 < M(K\pi) < 1.92$ GeV, and the background sideband containing all events with $1.94 < M(K\pi) < 2.2$ GeV were defined. The PDF distribution was constructed as the difference of distribution for the signal band and for the background band multiplied by 0.74. The coefficient 0.74 was chosen so that the number of events in the background band corresponds to the estimated number of background events events in the signal band.

The combination of tagging information proceeds through the construction of a combined flavor tagging variable y :

$$y = \prod_{i=1}^n y_i; \quad y_i = \frac{f_i^{\bar{b}}(x_i)}{f_i^b(x_i)}, \quad (6.11)$$

where the $f_i^b(x_i)$ are the probability. We convert this variable to the tagging variable d , defined in Eq. 5.4, via the transformation,

$$d = \frac{1 - y}{1 + y}. \quad (6.12)$$

As in Sec. 5.1, an event with $d > 0$ is tagged as a b quark and one with $d < 0$ as a \bar{b} quark, with larger $|d|$ values corresponding to higher tagging purities.

The flavor tagging algorithm was calibrated in data by applying it to the events containing B^0 and B^+ decays using the same techniques as described in Ch. 5. Fits to the asymmetry distribution, in various ranges of $|d|$ for these events show clear B_d oscillations with Δm_d values consistent with the world average value [13].

6.3.1 Tagged Sample Yields

We fit the tagged sample using the technique in Sec. 6.2 to obtain yields of signal and reflections. For the binned likelihood fit for Δm_s , we remove events having low tagger output dilution, $|d| < 0.3$. For the unbinned likelihood, we use the event-by-event dilution and therefore do not cut out these events. Figure 6.13 shows the fitted distribution without this cut and Fig. 6.14 shows the fitted distribution with this cut.

In Table 6.1 we list the yields and fractions for the components enumerated in Sec. 6.2 for the untagged, full tagged, and tagged with $|d| > 0.3$ samples, corresponding to Figs. 6.12, 6.13, and 6.14. We do not expect flavor tagging to affect the component fractions $f(D_s)$, $f(D^+)$, and $f(\Lambda_c)$. We verified that the component fractions in Fig. 6.12, Fig. 6.13, and Fig. 6.14 agree within errors, as seen in Table 6.1.

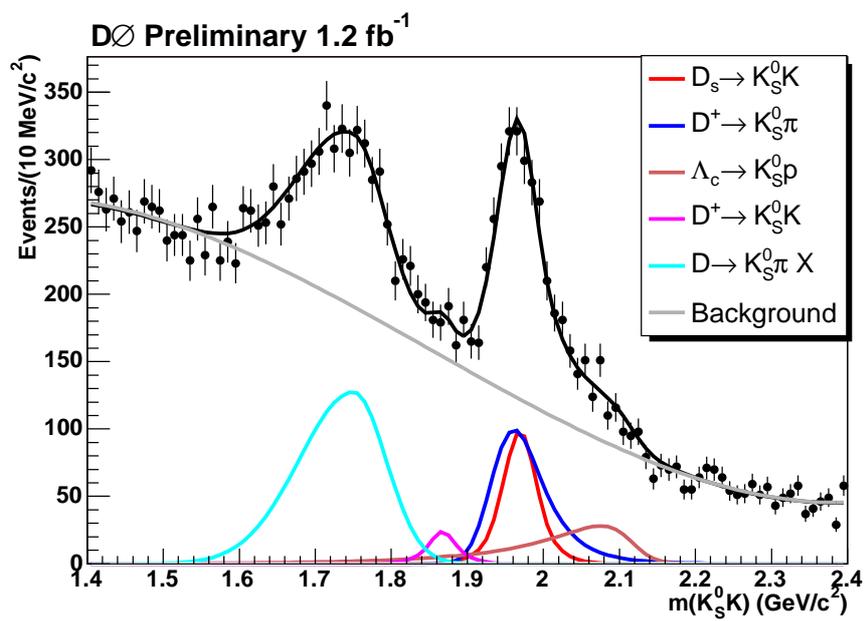


Figure 6.13: Mass fit to the entire tagged sample.

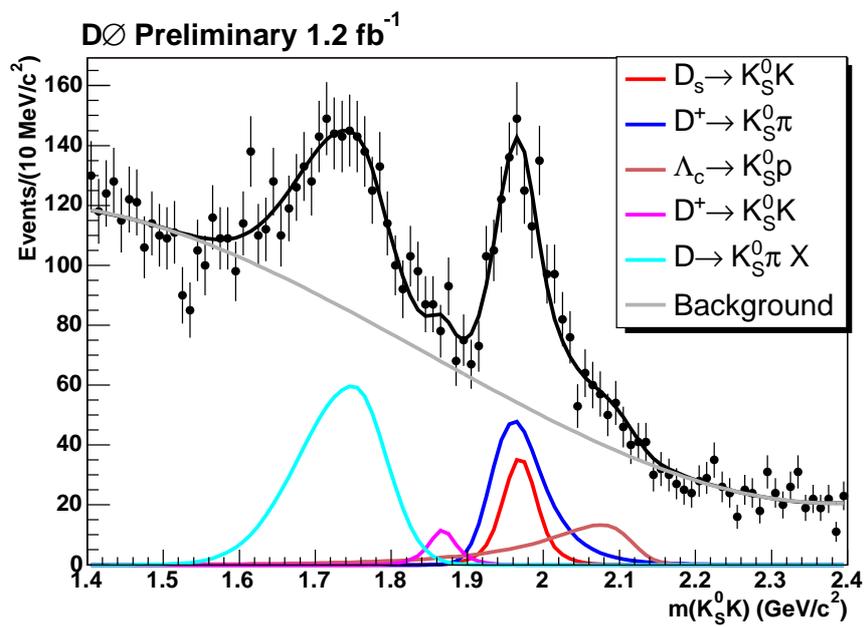


Figure 6.14: Mass fit to the tagged sample having $|d| > 0.3$.

Component	Variable	Untagged	Tagged	$ d > 0.3$
$D_s \rightarrow K_S^0 K$	N	2603 ± 110	593 ± 67	215 ± 34
	f	0.030 ± 0.001	0.032 ± 0.004	0.026 ± 0.004
$D^+ \rightarrow K_S^0 \pi$	N	4481 ± 106	914 ± 64	441 ± 32
	f	0.051 ± 0.001	0.049 ± 0.003	0.053 ± 0.004
$\Lambda_c \rightarrow K_S^0 p$	N	2244 ± 86	490 ± 42	232 ± 27
	f	0.026 ± 0.001	0.026 ± 0.002	0.028 ± 0.003
$D^{(0,+,*)} \rightarrow K_S^0 \pi X$	N	8314 ± 187	1828 ± 112	855 ± 59
	f	0.095 ± 0.002	0.098 ± 0.006	0.103 ± 0.007

Table 6.1: Component yields and fractions for the untagged, tagged, and tagged with $|d| > 0.3$ samples. N is the number of events and f is the component fraction.

6.4 Unbinned Likelihood Fit Method

The likelihood for an event to arise from a specific source in the sample depends on the visible proper decay length x^M , its error (σ_{x^M}), the mass of the D_s^- meson candidate (m) and the predicted dilution (d). All these quantities are known on an event-by-event basis. The *pdf* for each source can be expressed by the following formula:

$$\mathcal{P}_i = P_i^{x^M}(x^M, \sigma_{x^M}, d) P_i^m(m, \lambda) P_i^{\sigma_{x^M}} P_i^d P_i^y. \quad (6.13)$$

The function $P_i^{\sigma_{x^M}}$ is the *pdf* for the VPDL uncertainty, $P_i^m(m, \lambda)$ is the mass *pdf* as in Sec. 6.2, P_i^d is the *pdf* for the dilution and P_i^y is the *pdf* for the selection variable y . The function $P_i^{x^M}(x^M, \sigma_{x^M}, d)$ will be defined later.

The sources considered for the entire $K_S^0 K$ mass region ($1.4 < m(K_S^0 K) < 2.4 \text{ GeV}/c^2$) are the same as those enumerated in the beginning of Sec. 6.2.

The total pdf for the j^{th} B candidate has the form:

$$\mathcal{P}_j = \mathcal{F}_{D_s} \mathcal{P}_{D_s} + \mathcal{F}_{D^+} \mathcal{P}_{D^+} + 0.13 \cdot \mathcal{F}_{D^+} \mathcal{P}_{Cabbibo} + \mathcal{F}_\Lambda \mathcal{P}_\Lambda + \mathcal{F}_{low} \mathcal{P}_{low} + (1 - \mathcal{F}_{D_s} - 1.13 \cdot \mathcal{F}_{D^+} - \mathcal{F}_\Lambda - \mathcal{F}_{low}) \mathcal{P}_{bg}. \quad (6.14)$$

The fractions \mathcal{F}_{D_s} , \mathcal{F}_{D^+} , \mathcal{F}_Λ , \mathcal{F}_{low} are determined from a fit to the total tagged sample (see Fig. 6.13).

The shape of the combinatorial background is VPDL-dependent. We therefore bin the tagged sample in 11 bins of VPDL and fit for the background shape parameters in each bin. The shape of the background in these bins is seen in Fig. 6.15. Note that this shape is most important in the signal region ($\approx 1.85 < m(K_s^0 K) < 2.1 \text{ GeV}/c^2$).

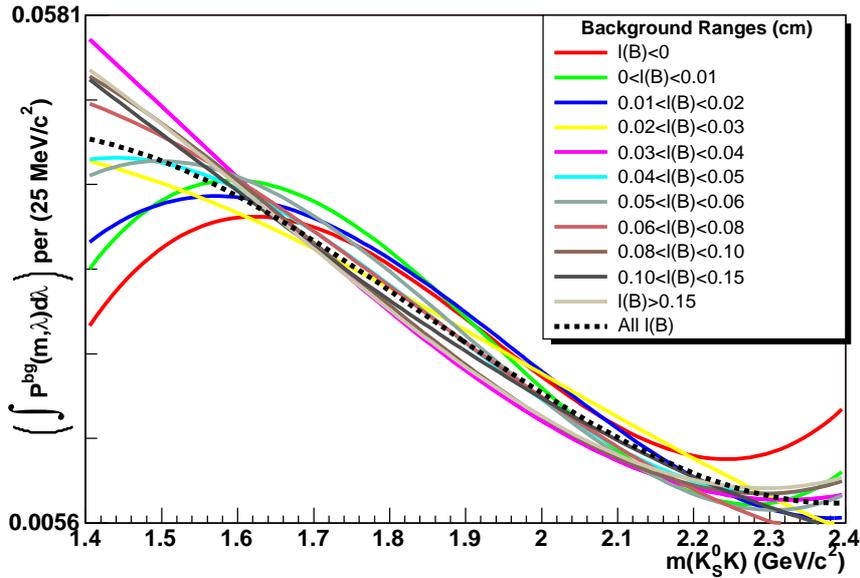


Figure 6.15: Fitted values of the 3^{rd} order Chebyshev polynomial background in 11 bins of VPDL obtained from the full tagged sample. Also overlaid is the shape of the background from the full tagged sample.

We perform a log-likelihood minimization of

$$\mathcal{L} = -2 \sum_{j=1}^{N_{events}} \ln \mathcal{P}_j \quad (6.15)$$

using MINUIT [55].

The signal distribution for VPDL error (Fig. 6.16), predicted dilution (Fig. 6.17), and combined selection variable (Fig. 6.5) are obtained by background-subtracting right and left sideband regions from the signal region. The definitions of signal and sideband region for this background subtraction are given in Eqn. 6.2. These signal distributions are used for all components except the combinatorial background. Note that the distribution of the predicted dilution shown in Fig. 6.17 has spikes due to Q_J^μ , Q_J^e , Q_{SV} , and Q_{EV} taking on discrete values when all tracks in the jet have the same charge (see Sec. 5.4).

As can be seen in Eq. 3.20, the significance of a mixing measurement is a function of the proper time resolution. In our measurement, the analogue of the proper time resolution is the visible proper decay length resolution, σ_{VPDL} . We expect that the presence of a long-lived K_S^0 in the decay chain of the channel we are analyzing causes a degradation of σ_{VPDL} . We compare the signal distribution of $\sigma(VPDL)$ in Fig. 6.16 to the same distribution in the $\phi\pi\mu$ mode [16] to get an estimate of this degradation. The means of the distributions are $\langle \sigma(VPDL)_{\phi\pi\mu} \rangle = 72 \mu m$ and $\langle \sigma(VPDL)_{K_S^0 K \mu} \rangle = 86 \mu m$. The most probable values are $\sigma(VPDL)_{\phi\pi\mu}^{max} = 32 \mu m$ and $\sigma(VPDL)_{K_S^0 K \mu}^{max} = 45 \mu m$. We therefore see a 20% degradation in the mean of the resolution and a 40% degradation in its most probable value.

6.4.1 pdf for the μD_s Signal

The μD_s sample is composed mostly of B_s^0 mesons with some contributions from B_u and B_d mesons. Different species of B mesons behave differ-

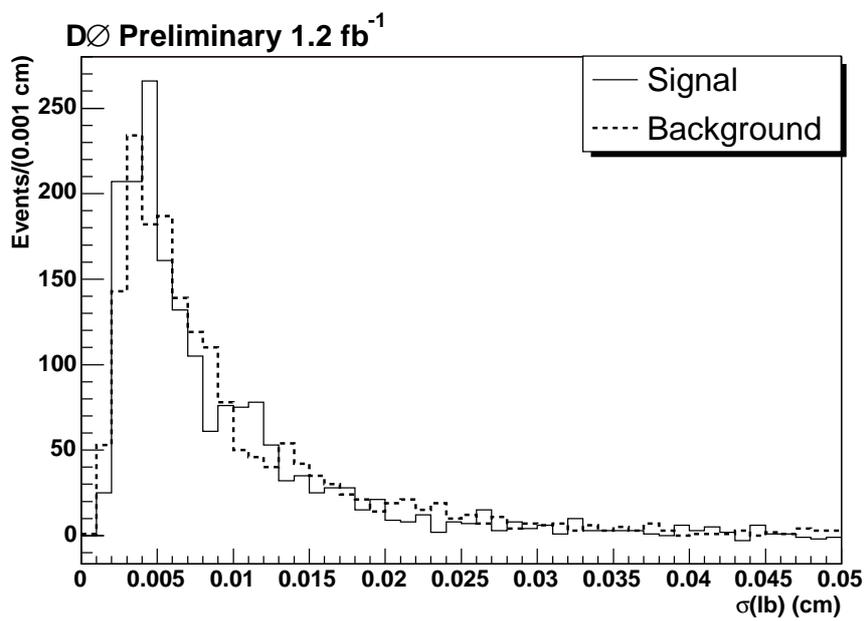


Figure 6.16: The distribution of VPD errors for signal and background.

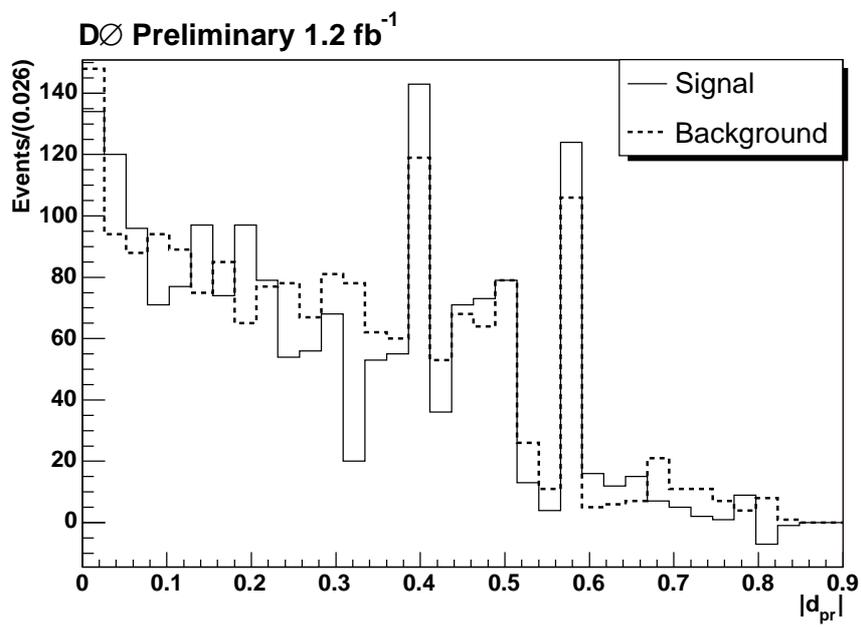


Figure 6.17: The distribution of predicted dilution for signal and background.

ently with respect to oscillations. Neutral B_d and B_s mesons do oscillate (with different frequencies) while charged B_u (i.e., B^+) mesons do not.

For a given type of B hadron (i.e., d , u , s) or b -baryon, the distribution of the visible proper decay length x is given by:

$$p_s^{nos}(x, K, d) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5 \cdot (1 + \mathcal{D}(d) \cos(\Delta m_s \cdot Kx/c)), \quad (6.16)$$

$$p_s^{osc}(x, K, d) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5 \cdot (1 - \mathcal{D}(d) \cos(\Delta m_s \cdot Kx/c)), \quad (6.17)$$

$$p_{D_s D_s}^{osc}(x, K) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5, \quad (6.18)$$

$$p_{D_s D_s}^{nos}(x, K) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5, \quad (6.19)$$

$$p_u^{nos}(x, K, d) = \frac{K}{c\tau_{B_u}} \exp\left(-\frac{Kx}{c\tau_{B_u}}\right) \cdot 0.5 \cdot (1 - \mathcal{D}(d)), \quad (6.20)$$

$$p_u^{osc}(x, K, d) = \frac{K}{c\tau_{B_u}} \exp\left(-\frac{Kx}{c\tau_{B_u}}\right) \cdot 0.5 \cdot (1 + \mathcal{D}(d)), \quad (6.21)$$

$$p_d^{nos}(x, K, d) = \frac{K}{c\tau_{B_d}} \exp\left(-\frac{Kx}{c\tau_{B_d}}\right) \cdot 0.5 \cdot (1 - \mathcal{D}(d) \cos(\Delta m_d \cdot Kx/c)), \quad (6.22)$$

$$p_d^{osc}(x, K, d) = \frac{K}{c\tau_{B_d}} \exp\left(-\frac{Kx}{c\tau_{B_d}}\right) \cdot 0.5 \cdot (1 + \mathcal{D}(d) \cos(\Delta m_d \cdot Kx/c)), \quad (6.23)$$

$$\text{where } K = P_T^{\mu D_s^-} / P_T^B. \quad (6.24)$$

Here τ is the lifetime of the B hadron and or b baryon and K is the K factor (see Sec. 3.2). Note that there is a sign swap in Eqs. 6.20–6.23 with respect to Eq. 6.16 and Eq. 6.17 due to the anti-correlation of muon charge for $B \rightarrow DD_s$; $D \rightarrow \mu X$ processes.

$\mathcal{D}(|d|)$ in Eqs. 6.16–6.23 is the dilution calibration of the data-based fla-

vor tagger (see Sec. 5.7.6). Figure 6.18 shows the results of a fit to a 3rd order polynomial to the distribution of measured dilution, obtained from the Δm_d analysis in the same manner as for the Monte Carlo based flavor tagger, versus $|d|$. The resulting calibration curve is parameterized,

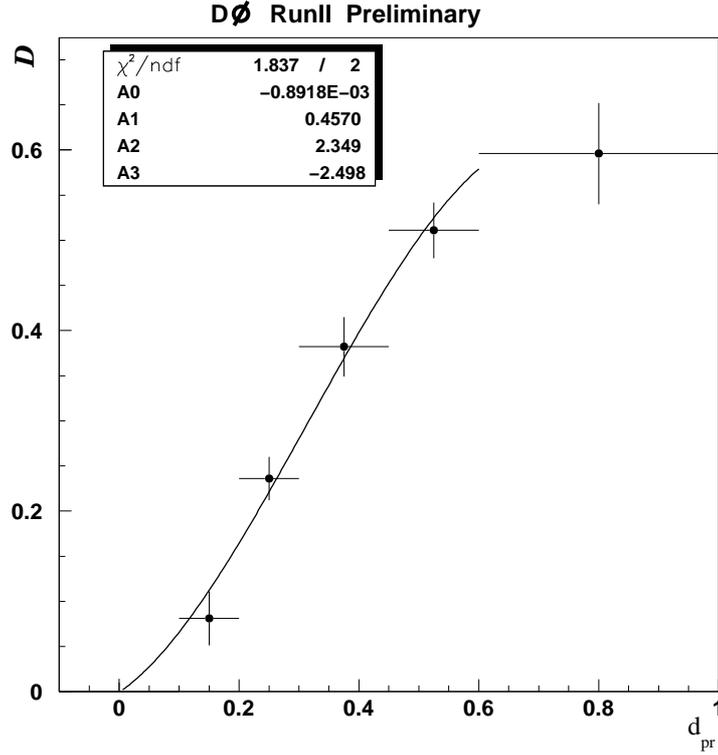


Figure 6.18: Dilution calibration, taken from DØ Note 4991.

$$\mathcal{D}(d) = \begin{cases} 0.457 \cdot |d| + 2.349 \cdot |d|^2 - 2.498 \cdot |d|^3, & d < 0.6 \\ 0.6; & d \geq 0.6. \end{cases}$$

The translation to the measured VPDL, x^M is achieved by a convolution

of the K factors and resolution functions as specified below.

$$P_{(d,u,s,\Lambda),j}^{osc,nos}(x^M, \sigma_{x^M}, d) = \quad (6.25)$$

$$\int_{K_{min}}^{K_{max}} dK D_j(K) \cdot \frac{Eff_j(x^M)}{N_j(K, \sigma_{x^M}, d)} \int_0^\infty dx G(x - x^M, \sigma_{x^M}) \quad (6.26)$$

$$\cdot p_{(d,u,s,\Lambda),j}^{osc,nos}(x, K, d).$$

$$\text{Here } G(x - x^M, \sigma_{x^M}) = \frac{1}{\sqrt{2\pi}\sigma_{x^M}} \exp\left(-\frac{(x - x^M)^2}{2\sigma_{x^M}^2}\right) \quad (6.27)$$

is the detector resolution of the VPDL and σ_{x^M} is given by

$$\sigma_{x^M} = (f_1 \cdot SF_1 \cdot \sigma_{x^M} + (1 - f_1) \cdot SF_2 \cdot \sigma_{x^M}),$$

where f_1 , SF_1 , and SF_2 are the resolution scale factors components as discussed in Sec. 6.6. $Eff_j(x)$ is the reconstruction efficiency for a given decay channel j of this type of B meson as a function of VPDL. The function $D_j(K)$ gives the normalized distribution of the K factor in a given channel j . The normalization factor N_j is calculated by integration over the entire VPDL region:

$$N_j(K, \sigma_{x^M}, d) = \frac{1}{2} \int_{-\infty}^{\infty} dx^M Eff_j(x^M) \int_0^\infty dx G(x - x^M, \sigma_{x^M}) \quad (6.28)$$

$$\cdot (p_{(d,u,s,\Lambda),j}^{osc}(x, K, d) + p_{(d,u,s,\Lambda),j}^{nos}(x, K, d)).$$

The total VPDL *pdf* for the μD_s signal is a sum of all the contributions

which give the D_s mass peak:

$$P_{\mu D_s}^{osc, nos}(x^M, \sigma_{x^M}, d) = \quad (6.29)$$

$$\begin{aligned} & \left(\sum_j SC_j \cdot P_{d, j}^{osc, nos}(x^M, \sigma_{x^M}, d) \right. \\ & + \sum_j SC_j \cdot P_{u, j}^{osc, nos}(x^M, \sigma_{x^M}, d) \\ & \left. + \sum_j SC_j \cdot P_{s, j}^{osc, nos}(x^M, \sigma_{x^M}, d) \right) \\ & \times (1 - \mathcal{F}_{c\bar{c}}) + \mathcal{F}_{c\bar{c}} \cdot P_{c\bar{c}}^{osc, nos}(x^M). \end{aligned} \quad (6.30)$$

The sum \sum_j is taken over all decay channels $B \rightarrow \mu^+ \nu D_s^- X$ and SC_j is the sample composition for a given channel j as described in Sec. 6.5.1. The functions $D_j(K)$ were taken from Monte Carlo simulation and are input into the fit in the form of histograms. Uncertainties in all of these inputs will contribute to the systematic uncertainties.

We have found that the cut on the K_S^0 decay length does not bias the VPDL of the B candidate. We therefore use the same *pdf* for the $c\bar{c}$ contribution, $P_{c\bar{c}}(x^M)$ as in Ref. [54].

The B meson lifetimes and efficiencies $Eff_j(x)$ are highly correlated. The efficiencies determined using Monte Carlo do not take into account the trigger selection and therefore measurements of the B meson lifetimes with such efficiencies give biased results. It is necessary to mention that the lifetime does not directly influence the measurements of the B_s^0 oscillation frequency though the error on oscillation frequency can be sensitive to the modeling of the background. Therefore, the B_s^0 lifetime was determined from data using the efficiencies measured in Monte Carlo. The results of the lifetime fit are discussed in Sec. 6.7.

6.4.2 *pdf* for μD^+ Components

As noted in Sec. 6.2 there are two μD^+ components present in the $m(K_S^0 K)$ spectrum, $D^+ \rightarrow K_S^0 \pi$ and $D^+ \rightarrow K_S^0 K$. We use Eqs. 6.23 and 6.22 to model the B_d components and Eqs. 6.21 and 6.22 to model the B_u components of these decays.

Monte Carlo studies indicate that the $K_S^0 \pi$ mode is composed of 85% B_d and 15% B_u . For the low mass peak, $K_S^0 \pi X$, we use $D^+ \rightarrow K_S^0 \pi \pi^0$ to estimate the sample composition as 65% B_d and 35% B_u . We model the Cabbibo-suppressed peak as pure B_d .

6.4.3 *pdf* for the $\mu \Lambda_C$ Component

We use Eqs. 6.31 and 6.32 to model the Λ_c component:

$$p_{\Lambda}^{nos}(x, K, d) = \frac{K}{c\tau_{\Lambda}} \exp\left(-\frac{Kx}{c\tau_{\Lambda}}\right) \cdot 0.5 \cdot (1 - \mathcal{D}(d)), \quad (6.31)$$

$$p_{\Lambda}^{osc}(x, K, d) = \frac{K}{c\tau_{\Lambda}} \exp\left(-\frac{Kx}{c\tau_{\Lambda}}\right) \cdot 0.5 \cdot (1 + \mathcal{D}(d)), \quad (6.32)$$

where τ_{Λ} is the lifetime of the Λ_c baryon [13].

6.4.4 *pdf* for the Combinatorial Background

The following contributions to the combinatorial background were considered:

1. Quasi-vertices distributed around the primary vertex - described as a Gaussian with width $\sigma_{peak.bg}$; fraction in the background: $\mathcal{F}_{peak.bg}$.
2. A negative exponential to account for outliers in the negative x^M tail - fraction in the background: $(1 - \mathcal{F}_{peak.bg}) \cdot (1 - \mathcal{F}_{mix}) \cdot \mathcal{F}_{neg}$.
3. A long-lived background insensitive to tagging - described as an exponential with decay length $c\tau_{bg}$ convoluted with the resolution containing

a background scale factor s_{bg} ; fraction in the background $(1 - \mathcal{F}_{peak.bg}) \cdot (1 - \mathcal{F}_{mix}) \cdot (1 - \mathcal{F}_{neg})$.

4. A non-oscillating long-lived background sensitive to tagging - described similarly to the insensitive long-lived background except for the multiplication of the dilution factor $1 \pm \mathcal{D}$; fraction in the background $(1 - \mathcal{F}_{peak.bg}) \cdot \mathcal{F}_{mix} \cdot (1 - \mathcal{F}_{Bd})$.
5. An long-lived background sensitive to tagging and oscillating at the frequency Δm_d - described similarly to the non-oscillating tag-sensitive background except for the multiplication of $\cos(\Delta m_d x/c)$; fraction in the background $(1 - \mathcal{F}_{peak.bg}) \cdot \mathcal{F}_{mix} \cdot \mathcal{F}_{Bd}$.

The fractions of these contributions and their parameters were determined from fitting the lifetime distribution in the data sample. The background *pdf*

is expressed in the following form:

$$P_{bg}(x^M, \sigma_{x^M}, d) = \mathcal{F}_{peak_bg} G(0 - x^M, \sigma_{peak_bg}) + (1 - \mathcal{F}_{peak_bg}) P_{bg}^{res}(x^M, \sigma_{x^M}, d), \quad (6.33)$$

$$P_{bg}^{res}(x^M, \sigma_{x^M}, d) = (1 - \mathcal{F}_{mix}) P_{\pm} + \mathcal{F}_{mix} (\mathcal{F}_{Bd} P_{Bd} + (1 - \mathcal{F}_{Bd}) P_{Bu}), \quad (6.34)$$

$$P_{\pm}(x^M, \sigma_{x^M}, d) = \mathcal{F}_{neg} \cdot \frac{-1}{c\tau_{neg}} \exp\left(-\frac{x^M}{c\tau_{neg}}\right) + (1 - \mathcal{F}_{neg}) \frac{\epsilon(x^M)}{N} \int_0^{\infty} dx \frac{1}{c\tau_{bg}} \exp\left(-\frac{x}{c\tau_{bg}}\right) \cdot G(x - x^M, s_{bg}\sigma_{x^M}), \quad (6.35)$$

$$P_{Bu}^{osc,nonosc}(x^M, \sigma_{x^M}, d) = \frac{\epsilon(x^M)}{N} \int_0^{\infty} dx \frac{1}{c\tau_{bg}} \exp\left(-\frac{x}{c\tau_{bg}}\right) (1 \pm \mathcal{D}) \cdot G(x - x^M, s_{bg}\sigma_{x^M}), \quad (6.36)$$

$$P_{Bd}^{osc,nonosc}(x^M, \sigma_{x^M}, d) = \frac{\epsilon(x^M)}{N} \int_0^{\infty} dx \frac{1}{c\tau_{bg}} \exp\left(-\frac{x}{c\tau_{bg}}\right) (1 \pm \mathcal{D} \cos(\Delta m_d x/c)) \cdot G(x - x^M, s_{bg}\sigma_{x^M}), \quad (6.37)$$

where N is the normalization constant and the fit parameters are \mathcal{F}_{peak_bg} , σ_{peak_bg} , \mathcal{F}_{mix} , \mathcal{F}_{Bd} , \mathcal{F}_{Bu} , τ_{bg} , τ_{neg} , and s_{bg} . The efficiency for the $B_d^0 \rightarrow D^- \mu^+ \nu X$ channel was used for $\epsilon(x^M)$.

6.5 Inputs to the Fit

6.5.1 Sample Composition

To determine the composition of the selected μD_s sample, the following decay channels of B mesons were considered :

- $B_s^0 \rightarrow \mu^+ \nu D_s^-$;
- $B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu D_s^-$;
- $B_s^0 \rightarrow \mu^+ \nu D_{s0}^{*-} \rightarrow \mu^+ \nu D_s^-$;
- $B_s^0 \rightarrow \mu^+ \nu D_{s1}^{\prime-} \rightarrow \mu^+ \nu D_s^-$;
- $B_s^0 \rightarrow \tau^+ \nu D_s^- X, \tau \rightarrow \mu \nu \nu$;
- $B_s^0 \rightarrow D_s^+ D_s^- X; D_s^- \rightarrow \mu \nu X$;
- $B_s^0 \rightarrow D_s D X; D \rightarrow \mu \nu X$;
- $B^+ \rightarrow D D_s^- X; D \rightarrow \mu \nu X$;
- $B^0 \rightarrow D D_s^- X; D \rightarrow \mu \nu X$.

The latest PDG values [13] were used to determine the branching fractions of decays contributing to the D_s^- sample. EvtGen [56] inputs were used for the branching fractions which are not given in the PDG.

- $\text{Br}(B_s^0 \rightarrow \mu^+ \nu D_s^- X) = (7.9 \pm 2.4)\%$, total semileptonic Br was taken from PDG, fractions of exclusive channels were taken from EvtGen:
 - $\text{Br}(B_s^0 \rightarrow \mu^+ \nu D_s^-) = 2.0\%$;
 - $\text{Br}(B_s^0 \rightarrow \mu^+ \nu D_s^{*-}) = 5.3\%$;
 - $\text{Br}(B_s^0 \rightarrow \mu^+ \nu D_{s0}^{*-}) = 0.19\%$;
 - $\text{Br}(B_s^0 \rightarrow \mu^+ \nu D_{s1}^{\prime-}) = 0.35\%$;

- $\text{Br}(B_s^0 \rightarrow \tau^+ \nu D_s^- X) = 2.9\%$, from Evtgen;
- $\text{Br}(\tau^+ \rightarrow \mu \nu \nu) = (17.36 \pm 0.06)\%$, from PDG;
- $\text{Br}(B_s^0 \rightarrow D_s^+ D_s^- X) = (23_{-13}^{+21})\%$, from PDG;
- $\text{Br}(B_s^0 \rightarrow DD_s^- X) = 15.4\%$, from EvtGen;
- $\text{Br}(B^+ \rightarrow DD_s^- X) = (10.5 \pm 2.6)\%$, from PDG;
- $\text{Br}(B^0 \rightarrow DD_s^- X) = (10.5 \pm 2.6)\%$, from PDG;
- $\text{Br}(D_s^{-*} \rightarrow D_s^- X) = 100\%$;
- $\text{Br}(D_{s0}^{*-} \rightarrow D_s^- X) = 100\%$;
- $\text{Br}(D_{s1}^{\prime-} \rightarrow D_s^- X) = 100\%$;
- $\text{Br}(D_s^- \rightarrow \mu \nu X) = (6.3 \pm 0.8)\%$, from PDG, assuming the same partial width as for D^0 and D^+ ;
- $\text{Br}(D^0 \rightarrow \mu \nu X) = (6.5 \pm 0.8)\%$, from PDG;
- $\text{Br}(D^+ \rightarrow \mu \nu X) = (17.2 \pm 1.9)\%$, from PDG;
- $\text{Br}(\bar{b} \rightarrow B^0) = (39.7 \pm 1.0)\%$, from PDG;
- $\text{Br}(\bar{b} \rightarrow B^+) = (39.7 \pm 1.0)\%$, from PDG;
- $\text{Br}(\bar{b} \rightarrow B_s^0) = (10.7 \pm 1.1)\%$, from PDG;

The reconstruction efficiency used to determine the sample composition does not include lifetime cuts. Lifetime dependent efficiencies are handled separately.

Process	$p_T(\mu) > 2 \text{ GeV}/c$	$p_T(\mu) > 6 \text{ GeV}/c$
$B_s^0 \rightarrow \mu^+ \nu D_s^-$	20.87%	21.98%
$B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu D_s^- X$	57.65%	60.68%
$B_s^0 \rightarrow \mu^+ \nu D_{s0}^{*-} \rightarrow \mu^+ \nu D_s^- X$	1.46%	1.54%
$B_s^0 \rightarrow \mu^+ \nu D_{s1}^{\prime-} \rightarrow \mu^+ \nu D_s^- X$	3.32%	3.49%
$B_s^0 \rightarrow \tau^+ \nu D_s^- \rightarrow \mu^+ \nu D_s^- X$	1.99%	2.09%
$B_s^0 \rightarrow D_s^+ D_s^- X; D_s^- \rightarrow \mu \nu X$	0.99%	0.79%
$B_s^0 \rightarrow DD_s^- X; D \rightarrow \mu \nu X$	1.60%	1.14%
$B^+ \rightarrow DD_s^- X; D \rightarrow \mu \nu X$	5.78%	3.79%
$B^0 \rightarrow DD_s^- X; D \rightarrow \mu \nu X$	6.36%	4.51%

Table 6.2: Sample composition calculated before the application of lifetime-biasing cuts. The top group of processes are “useful” for mixing because they are B_s^0 decays. To be conservative, we do not consider the $B_s^0 \rightarrow \tau X$ channel as a signal process. The overall usable sample fraction for $p_T(\mu) > 2 \text{ GeV}$ is 83.3%.

6.5.2 K Factors

As described earlier, semileptonic B decays necessarily have an undetected neutrino present in the decay chain, making a precise determination of the B meson kinematics difficult. In addition, other neutral or non-reconstructed charged particles can be present in the decay chain of the B meson. This leads to a bias towards smaller values of the momentum of the B meson calculated using reconstructed particles only. We correct for this bias by scaling the measured momentum by a K -factors as defined in Eq. 3.8. These K -factors were obtained from Monte Carlo simulations and were calculated before lifetime cuts were applied. For the computation of p_T , generator level information was used. We use the reconstructed level information as a systematic uncertainty.

Fig. 6.19 shows the distributions of the K factors for the semileptonic decays $B_s^0 \rightarrow \mu^+ \nu D_s^-$ and $B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu D_s^-$. As expected, the K factor for D_s^{*-} decays had a lower mean value because more decay products

are missing. Note that since the K factors in Eq. 3.8 were defined as the ratio of transverse momenta, they can exceed 1.

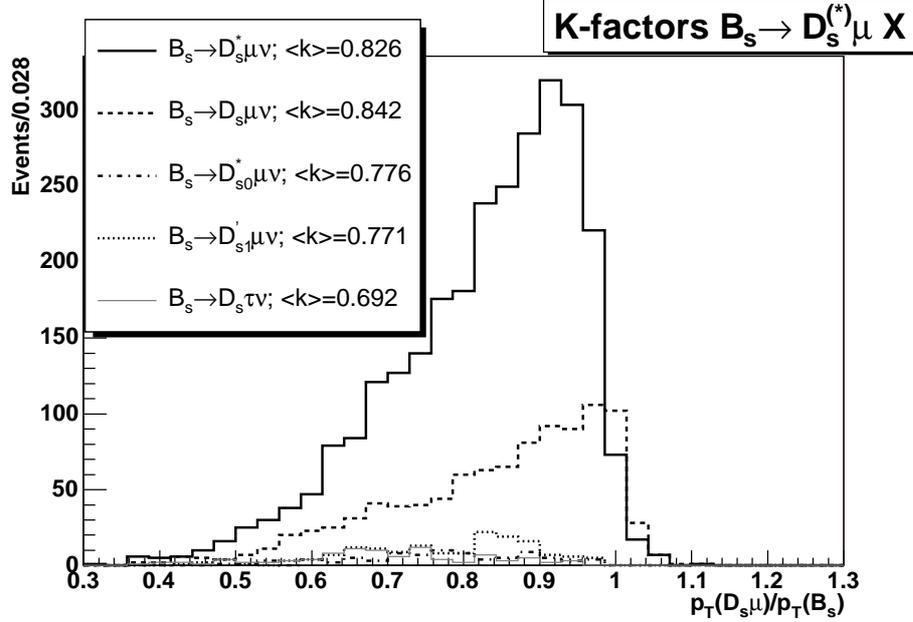


Figure 6.19: K factors for $B_s^0 \rightarrow D_s^- \mu^+ X$ and $B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu D_s^-$ processes.

The K factor distributions are divided into four $m(D_s^- \mu)$ bins for the likelihood fit. The mass bins are: $m(D_s^- \mu) < 3.5 \text{ GeV}/c^2$, $3.5 < m(D_s^- \mu) < 4.0 \text{ GeV}/c^2$, $4.0 < m(D_s^- \mu) < 4.5 \text{ GeV}/c^2$, and $m(D_s^- \mu) < 4.5 \text{ GeV}/c^2$. The K factor distributions for the four $m(D_s^- \mu)$ bins for $B_s^0 \rightarrow \mu^+ \nu D_s^-$ decays are shown in fig. 6.20. Fig. 6.21 shows the distributions for $B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu D_s^-$ decays.

Fig. 6.22 shows the K factor distribution for $B \rightarrow \mu \nu D^+ \rightarrow K_S^0 \pi$.

Fig. 6.23 shows the K factor distribution for $B \rightarrow \mu \nu D^+ \rightarrow K_S^0 \pi \pi^0$.

Fig. 6.24 shows the K factor distribution for $B^0 \rightarrow D_s D$.

Fig. 6.25 shows the K factor distribution for $B^- \rightarrow D_s D$.

Fig. 6.26 shows the K factor distribution for $B_s^0 \rightarrow D_s D_s$.

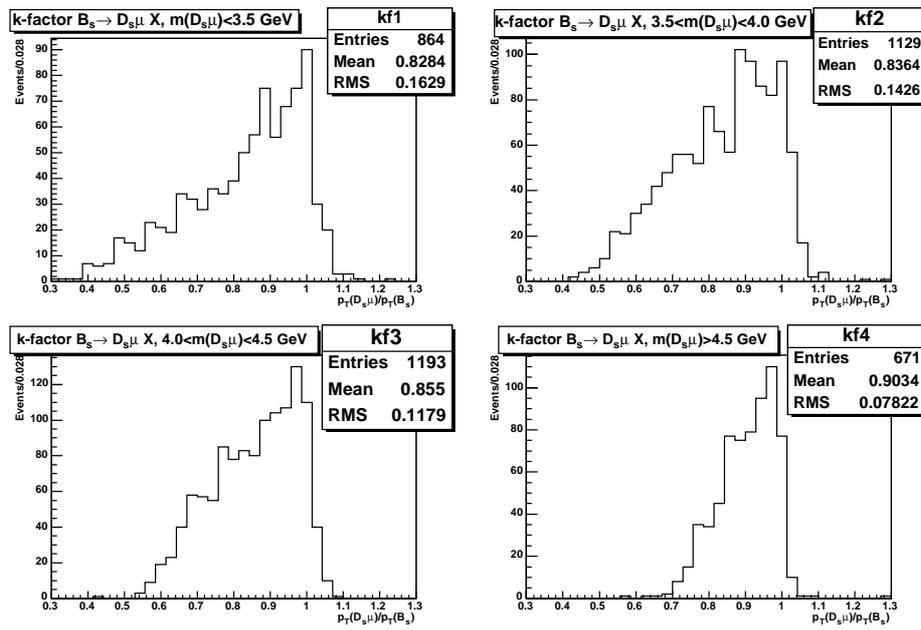


Figure 6.20: K factors for $B_s^0 \rightarrow D_s^- \mu^+ X$ divided into 4 bins of $m(D_s\mu)$.

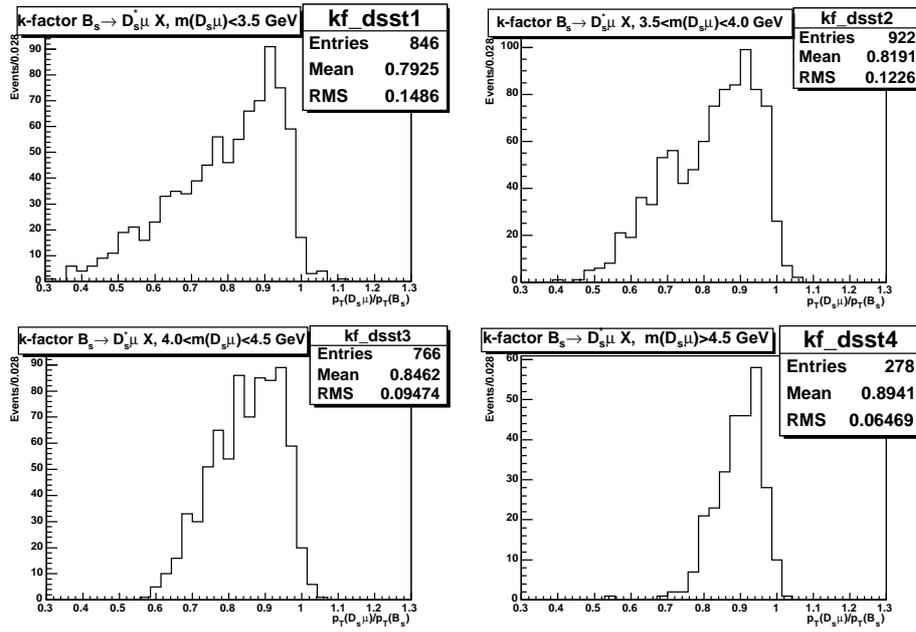


Figure 6.21: K factors for $B_s^0 \rightarrow \mu^+ \nu D_s^{*-} \rightarrow \mu^+ \nu D_s^-$ divided into 4 bins of $m(D_s \mu)$.

Fig. 6.27 shows the K factor distribution for $B_s^0 \rightarrow D_s D$.

Fig. 6.28 shows the K factor distribution for $\Lambda_b \rightarrow \mu\nu\Lambda_c \rightarrow K_S^0 p$.

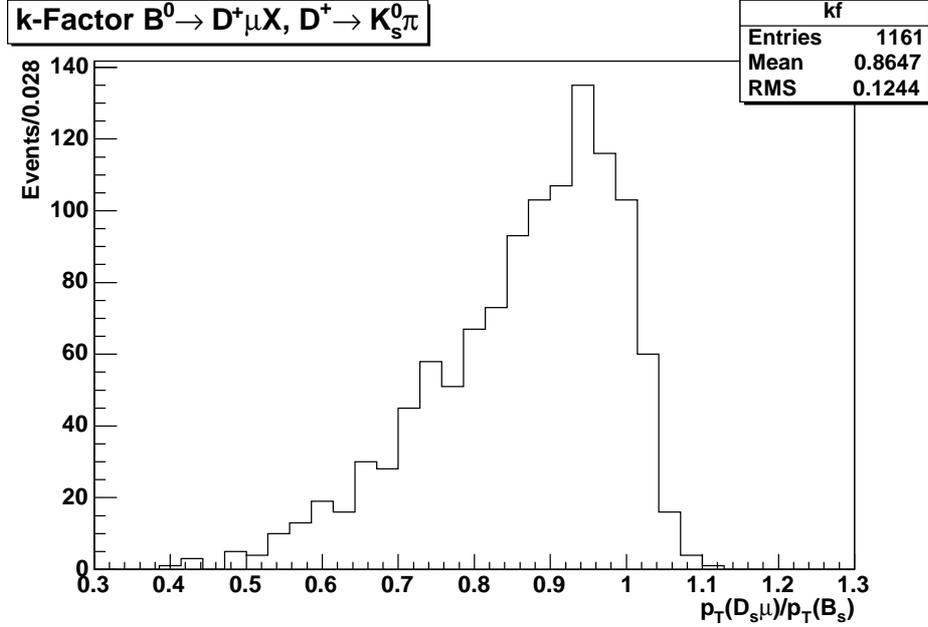
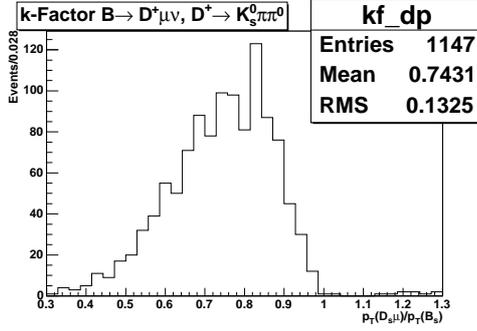
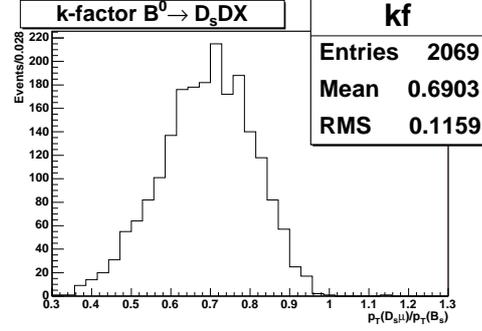
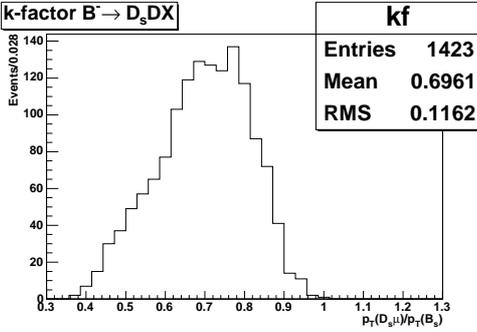
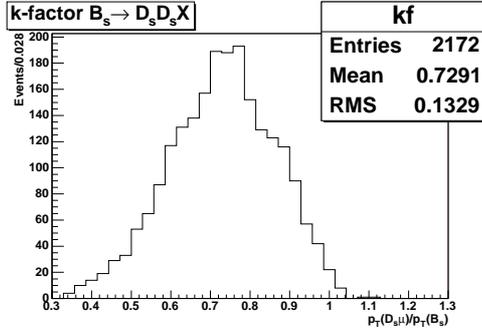
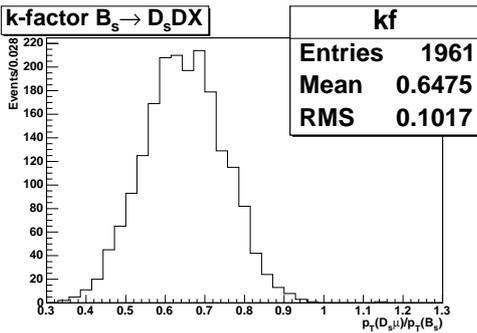
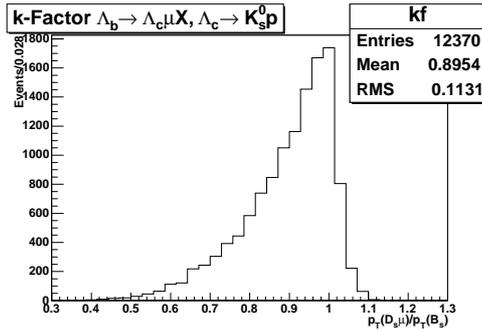


Figure 6.22: K factor for $B \rightarrow \mu\nu D^+ \rightarrow K_S^0 \pi$.

6.5.3 Reconstruction Efficiencies

The reconstruction efficiency of different modes contributing to the B_s was determined using the Monte Carlo simulations. The efficiency strongly depends on the decay length due to the lifetime biased selections for the sample. We determined the efficiency as a function of the reconstructed VPDL. The fit function is

$$Eff(x^M) = p_0 \cdot (1 - (p_2 + p_3 \cdot x^M + p_4 \cdot (x^M)^2 + p_5 \cdot (x^M)^3) \cdot \exp(-(x^M)^2/p_1)). \quad (6.38)$$

Figure 6.23: K factor for $B \rightarrow \mu \nu D^+ \rightarrow K_S^0 \pi \pi^0$.Figure 6.24: K factor for $B^0 \rightarrow D_s D$.Figure 6.25: K factor for $B^- \rightarrow D_s D$.Figure 6.26: K factor for $B^0 \rightarrow D_s D_s$.Figure 6.27: K factor for $B_s \rightarrow D_s D$.Figure 6.28: K factor for $\Lambda_b \rightarrow \mu \nu \Lambda_c \rightarrow K_S^0 sp$.

This functional form was selected because it allows for analytical calculation of the normalization integral (see Eq. 6.29).

Fig. 6.29 shows the efficiency as a function of VPDL for $B_s^0 \rightarrow \mu^+ \nu D_s^- \rightarrow K_S^0 K^-$.

Fig. 6.30 shows the efficiency as a function of VPDL for $B \rightarrow \mu \nu D^+ \rightarrow K_S^0 \pi$.

Fig. 6.31 shows the efficiency as a function of VPDL for $B^0 \rightarrow D_s D$.

Fig. 6.32 shows the efficiency as a function of VPDL for $B^- \rightarrow D_s D$.

Fig. 6.33 shows the efficiency as a function of VPDL for $B_s^0 \rightarrow D_s D_s$.

Fig. 6.34 shows the efficiency as a function of VPDL for $B_s \rightarrow D_s D$.

Fig. 6.35 shows the efficiency as a function of VPDL for $\Lambda_b \rightarrow \mu \nu \Lambda_c \rightarrow K_S^0 p$.

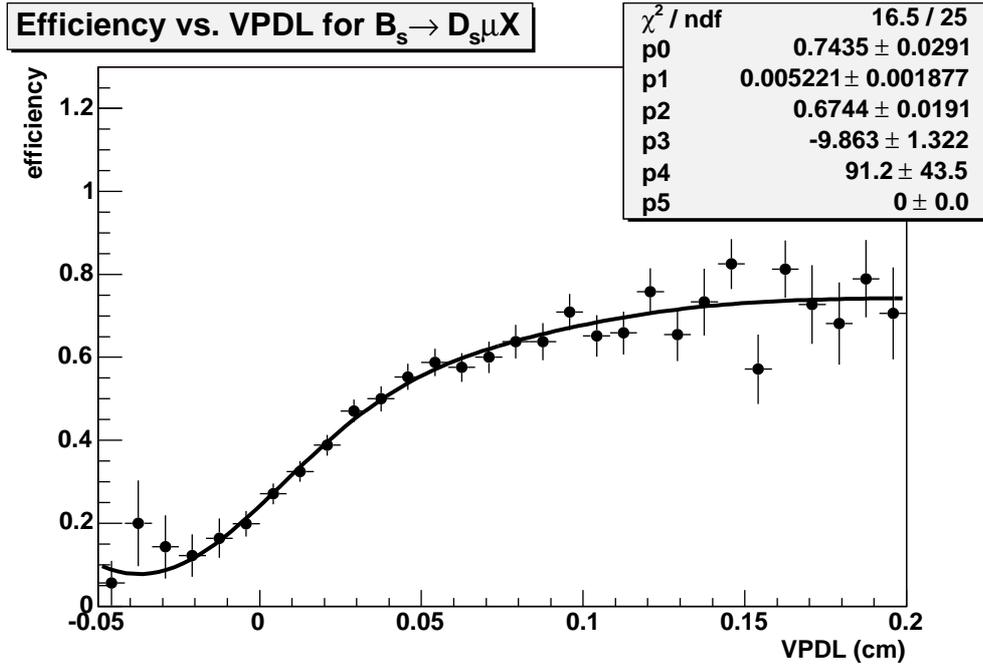


Figure 6.29: Efficiency as a function of VPDL for $B_s^0 \rightarrow D_s^- \mu^+ X$.

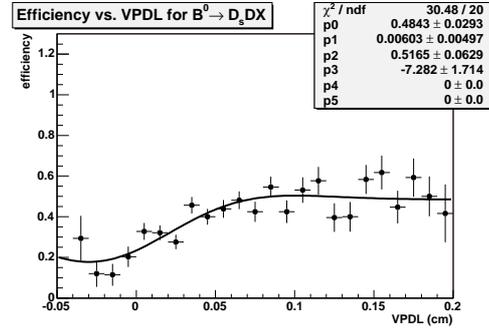
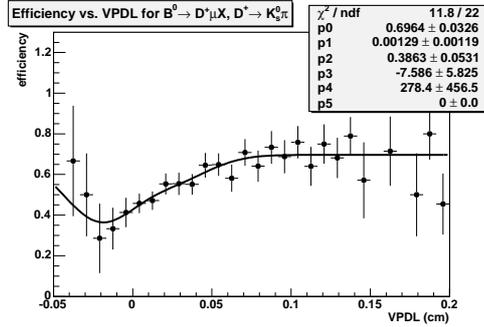


Figure 6.30: Efficiency as a function of VPDL for $B \rightarrow \mu\nu D^+ \rightarrow K_S^0\pi$.

Figure 6.31: Efficiency as a function of VPDL for $B^0 \rightarrow D_s D$.

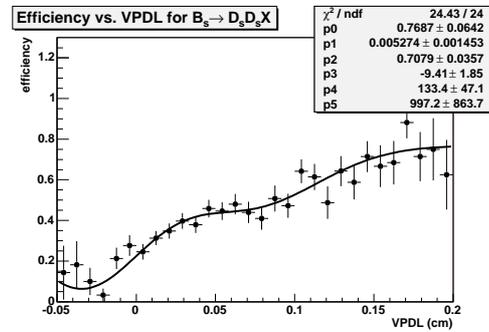
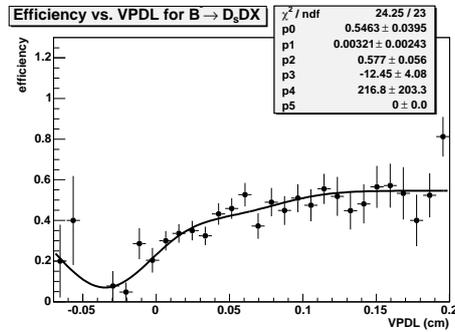


Figure 6.32: Efficiency as a function of VPDL for $B^- \rightarrow D_s D$.

Figure 6.33: Efficiency as a function of VPDL for $B_s \rightarrow D_s D_s$.

6.6 Resolution Scale Factor

The vertexing algorithm we use provides an estimate of σ_{VPDL} due to estimated errors on track parameters. Since not all detector systematics are included in the track parameter estimate, these are often underestimated. The *resolution scale factor* allows the level of this misestimation to be determined by the data itself.

The resolution scale factor for the signal component has been determined previously by examining the proper decay length resolution of prompt J/ψ events [16]. To briefly summarize the idea of this study, the negative tail of

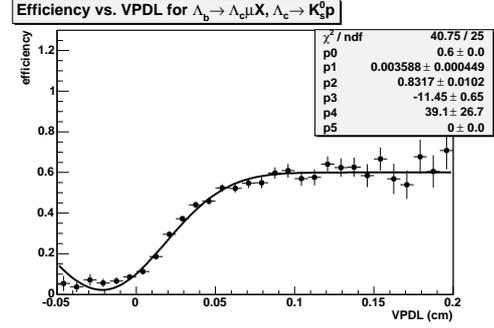
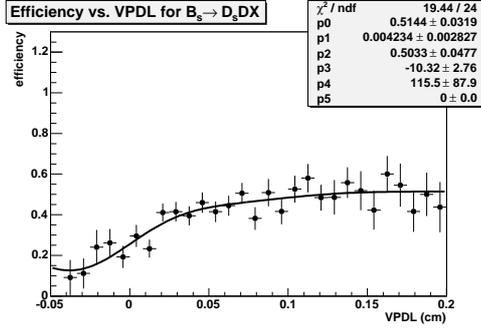


Figure 6.34: Efficiency as a function of VPDL for $B_s \rightarrow D_s D$.

Figure 6.35: Efficiency as a function of VPDL for $\Lambda_b \rightarrow \mu \nu \Lambda_c \rightarrow K_S^0 p$.

the pull of the J/ψ vertex distribution with respect to the primary vertex should be a Gaussian with a sigma of 1 if errors assigned to the vertex coordinates were correct. This is because events with negative decay length are dominated by reconstruction errors while positive decay length events may be daughters of long-lived particles (such as D or B hadrons). Therefore, the measured width of the negative side of this distribution can be used to scale the decay length errors, i.e. it provides an estimate of the resolution scale factor.

It is unclear from the above study, however, if the presence of a long-lived K_S^0 ($c\tau = 2.7$ cm) in the decay has an effect on the scale factor. We therefore studied the proper decay length pull distribution of a separate prompt decay, $D^{*+} \rightarrow D^0 \pi^+$ where $D^0 \rightarrow K_S^0 \pi^- \mu^+ X$. This decay has the advantage of having a simple way of estimating the combinatorial background through the charge correlation of the two pions and being close in topology to our signal channel.

We begin the reconstruction of this channel by searching for a muon and K_S^0 having the same properties as in Sec. 6.1. We then search for a pion coming from the D^0 . The pion selection cuts are as follows:

- the track must have at least two hits in the SMT and CFT;

- the track must come from the same primary vertex as the muon;
- the track must be in the same jet as the muon;
- $q(\pi) \cdot q(\mu) < 0$;
- $p_T(\pi) > 750 \text{ MeV}/c^2$.

We then form a D^0 vertex from these three particles. We require:

- $0.9 < m(\mu + \pi + K_S^0) < 2.0 \text{ GeV}/c^2$;
- $\chi^2(D^0 \text{ vertex fit}) < 25$.

Finally, we search for a slow pion, π_{slow} , to add to the D^0 to make a D^* . We require:

- the track must have hits in both the CFT and SMT;
- the track must have the same primary vertex as the D^0 ;
- the track must be in the same jet as the D^0 ;
- $\chi^2(D^* \text{ vertex fit}) < 16$.

The resulting mass distributions for both charge correlations of $q(\pi) \times q(\pi_{slow})$ are shown in Fig. 6.36.

An excess of events corresponding to D^* events is seen at low ΔM . We make signal pull distributions of $\text{PDL}(D^*)/\sigma_{\text{PDL}}(D^*)$ by taking all events having $\Delta M < 175 \text{ MeV}/c^2$ and subtracting distributions with $q(\pi) \times q(\pi_{slow}) > 0$ from distributions with $q(\pi) \times q(\pi_{slow}) < 0$. The negative side of these pull distributions are fit to double Gaussians as shown in Fig. 6.37. We also apply a cut of $p_T(\pi_{slow}) > 1.0 \text{ GeV}/c$ to bring all the tracks into the same kinematic range as the $K_S^0 K$ data sample; the fit on the right side of Fig. 6.37 is done on the sample with the p_T cut on the slow pion.

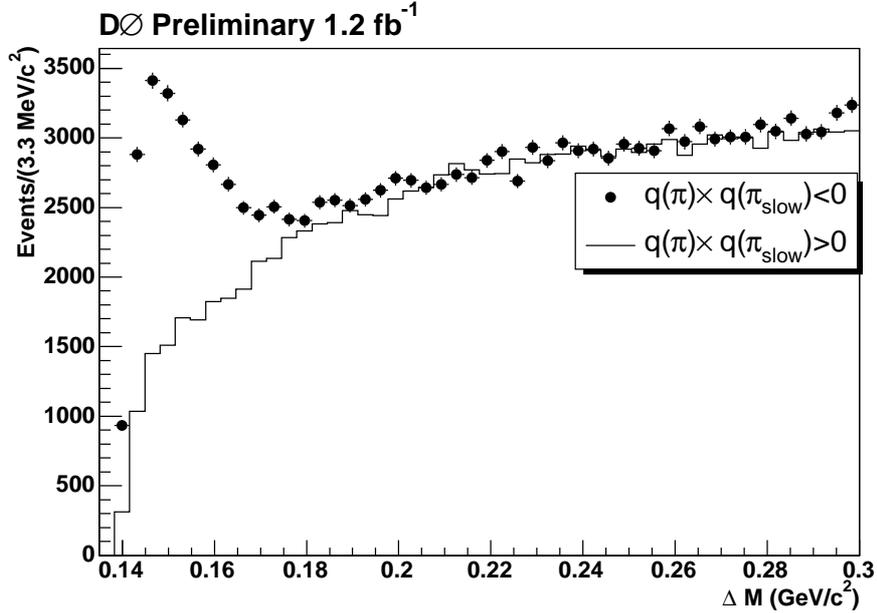


Figure 6.36: Distributions of $\Delta M = m(D^*) - m(D^0)$ for $D^{*+} \rightarrow D^0\pi^+$, $D^0 \rightarrow K_S^0\pi^-\mu^+X$ events. Both charge correlations of $q(\pi) \times q(\pi_{slow})$, where π refers to the pion from the D^0 and π_{slow} refers to the pion from the D^* , are shown.

For the $p_T(\pi_{slow}) > 1.0$ GeV/ c sample, we have 85% of the events having a scale factor of 0.966 and 15% having a scale factor of 2.48. We use these scale factors as the default signal resolution scale factors for this analysis.

We then bin the sample in three bins of $l_{xy}(K_S^0)$ and again fit double Gaussians to the pull distribution to test for a possible dependence of the scale factor on the K_S^0 decay length. Figure 6.38 shows the result of this study. No clear dependence on $l_{xy}(K_S^0)$ is seen and we therefore assume no dependence of the scale factor on $l_{xy}(K_S^0)$ for the remainder of this note.

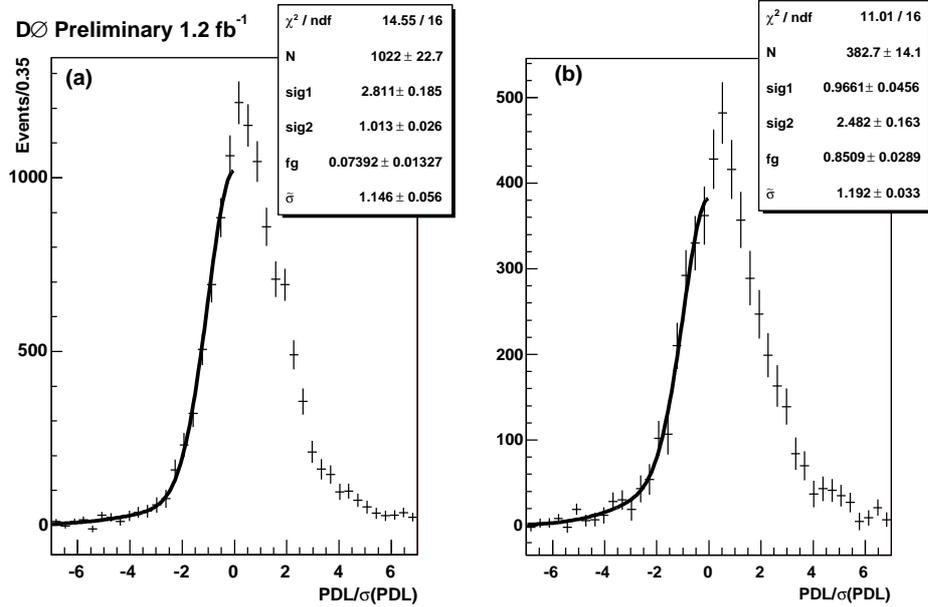


Figure 6.37: Double Gaussian fits to the negative part of background subtracted $PDL(D^*)/\sigma_{PDL}(D^*)$ distributions. a) No cut on $p_T(\pi_{slow})$. b) $p_T(\pi_{slow}) > 1.0$ GeV/c.

6.7 Results of the Lifetime Fit

We first employ the unbinned framework described in Sec. 6.4 to fit for the lifetime of the B_s^0 . The total tagged sample in the entire mass range $1.4 < m(K_s^0 K) < 2.4$ GeV/ c^2 was used to determine the parameters as seen in Table 6.3.

The lifetime we obtain, $c\tau_{B_s} = 498 \pm 39 \mu m$, is 1.5 standard deviations from the PDG value $c\tau_{B_s} = 438 \mu m$ [13]. We fix the lifetime at the PDG value and rescan the amplitude as an estimate of the systematic uncertainty due to the effect of the fitted lifetime.

Fig. 6.39 shows the VPD distribution of events in the entire mass range $1.4 < m(K_s^0 K) < 2.4$ GeV/ c^2 with the lifetime fit projected using the above

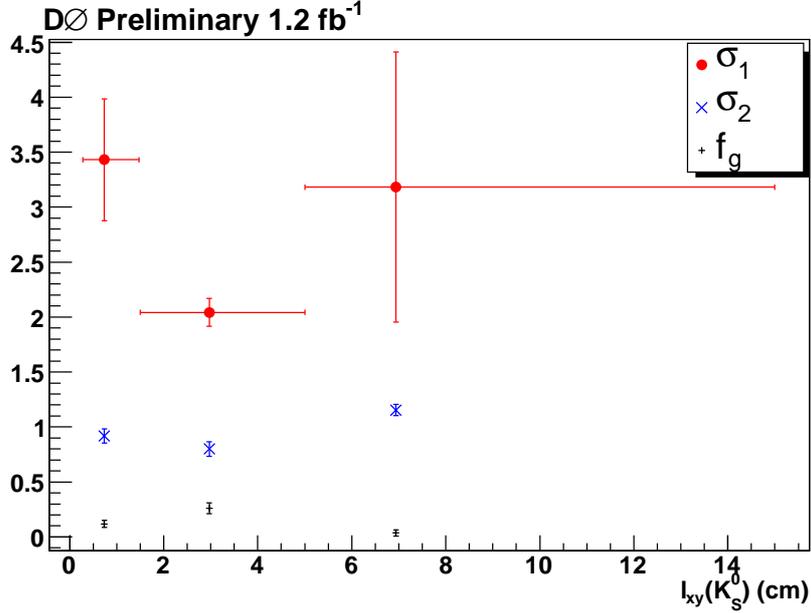


Figure 6.38: Widths and fractions of double Gaussian fits to pull distributions versus $l_{xy}(K_S^0)$.

optimal fit parameters.

Fig. 6.40 shows the VPDL distribution of events in the 3σ signal mass range $1.89 < m(K_S^0 K) < 2.05 \text{ GeV}/c^2$ with the lifetime fit projected using the above optimal fit parameters.

Fig. 6.41 shows the VPDL distribution of events in the right ($2.22 < m(K_S^0 K) < 2.4 \text{ GeV}/c^2$) and left ($1.4 < m(K_S^0 K) < 1.434 \text{ GeV}/c^2$) side-band mass ranges with the lifetime fit projected using the above optimal fit parameters. Only x^M and σ_{x^M} pdf's were used to produce these plots.

6.8 Fitting Procedure for the Δm_s Limit

We use the *amplitude fit method* [29] to scan for Δm_s and to set a limit on B_s^0 oscillations.

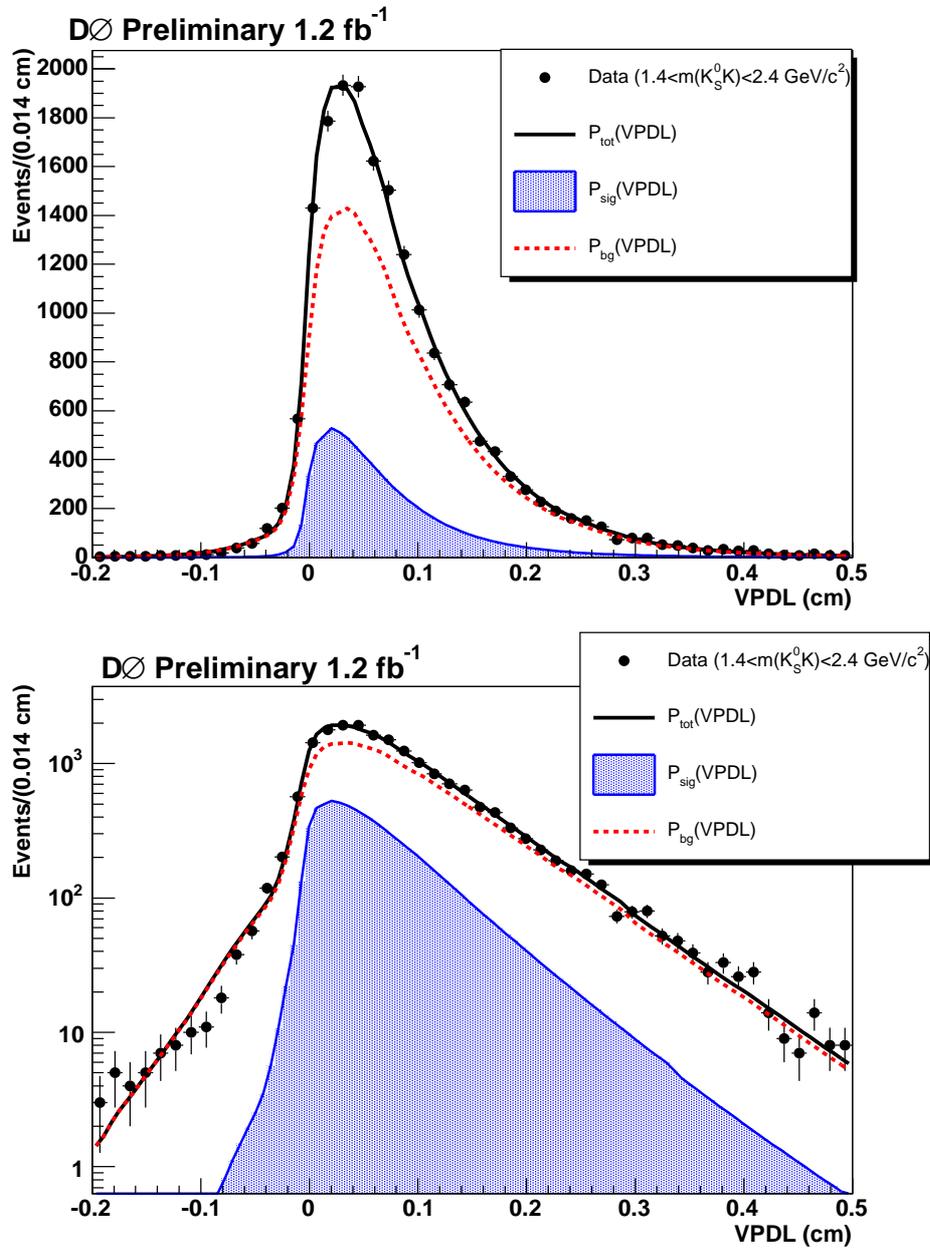


Figure 6.39: Lifetime fit projected on the entire mass range, $1.4 < m(K_S^0 K) < 2.4 \text{ GeV}/c^2$ in linear scale on the top and logarithmic scale on the bottom.

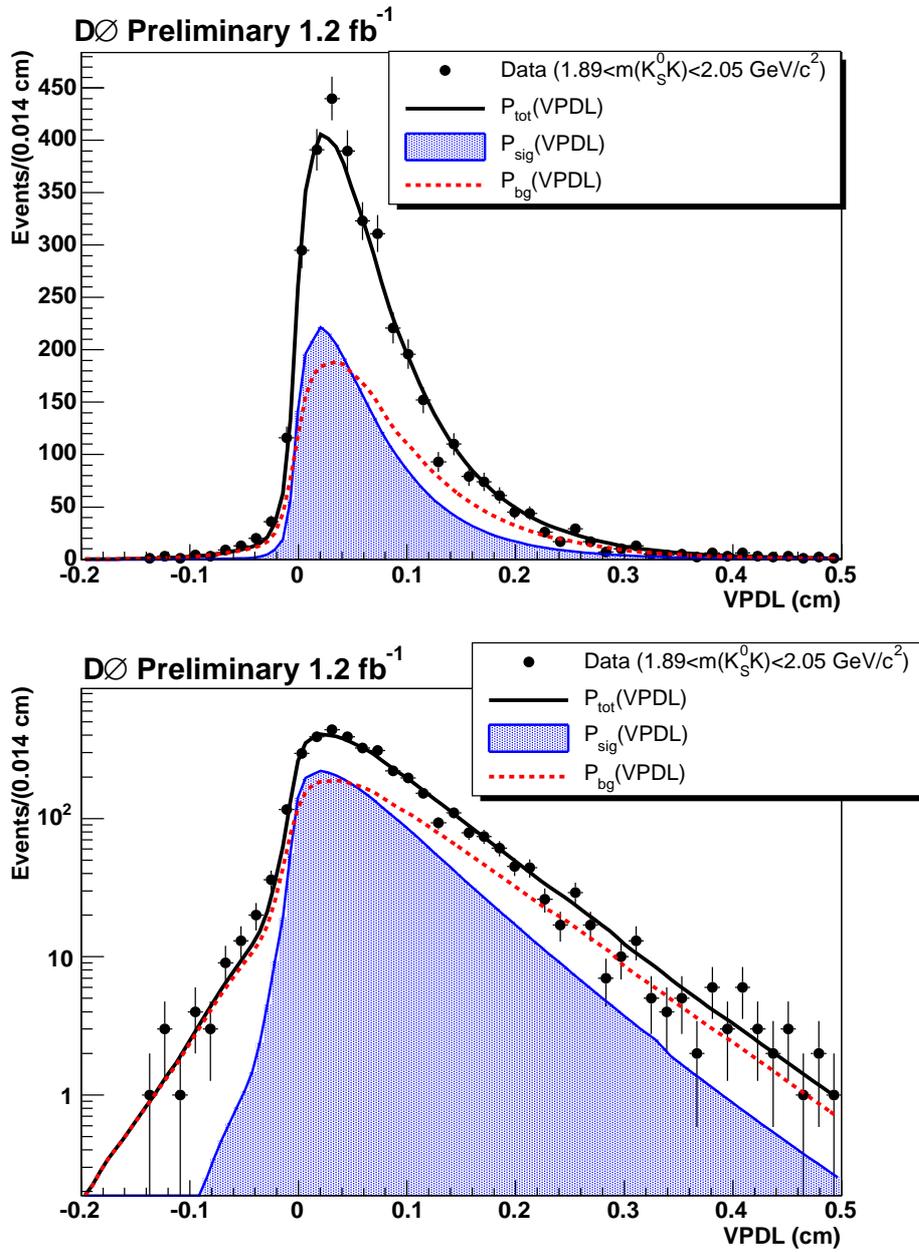


Figure 6.40: Lifetime fit projected on the 3σ signal region mass range, $1.89 < m(K_S^0 K) < 2.05 \text{ GeV}/c^2$ in linear scale on the top and logarithmic scale on the bottom.

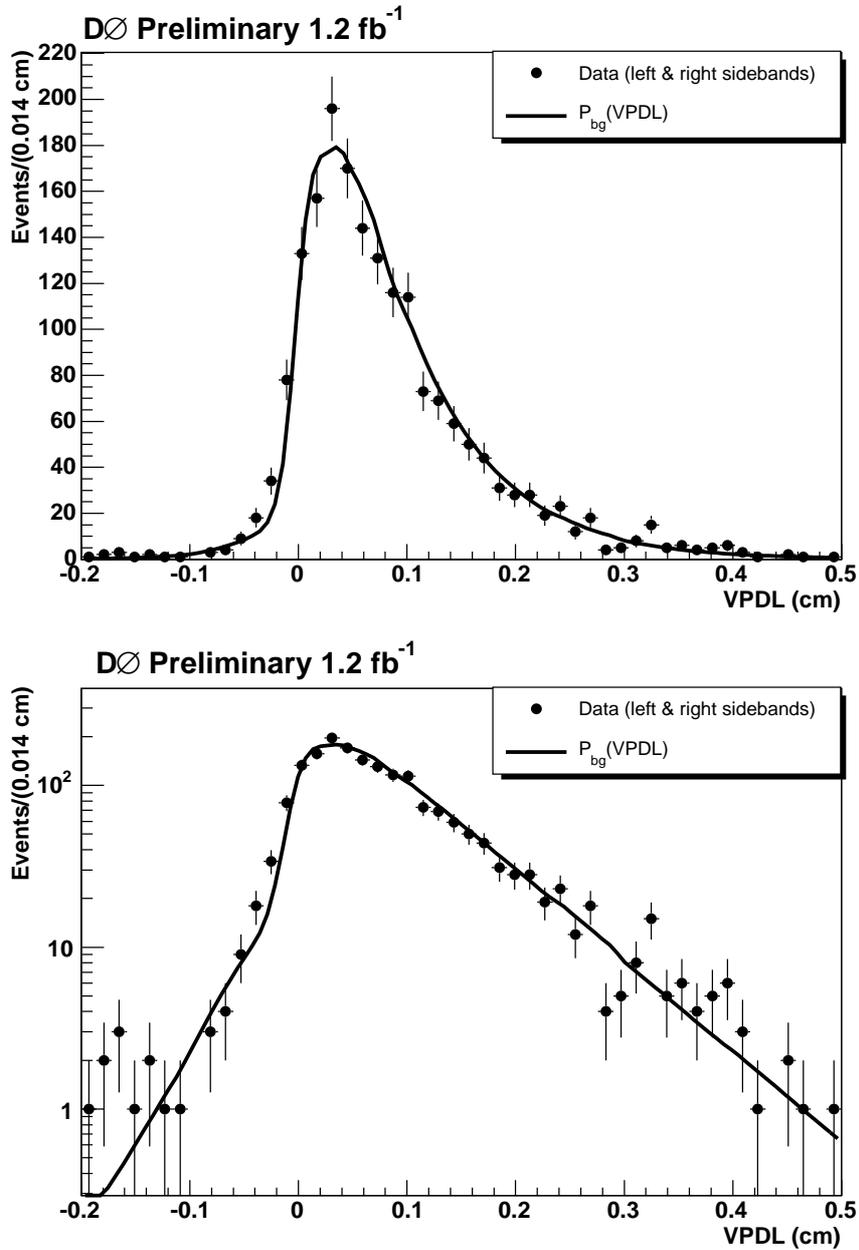


Figure 6.41: Lifetime fit projected on the right ($2.22 < m(K_S^0 K) < 2.40$ GeV/ c^2) and left ($1.400 < m(K_S^0 K) < 1.434$ GeV/ c^2) sidebands in linear scale on the top and logarithmic scale on the bottom.

Fit Parameter	Meaning	Value
$\mathcal{F}_{\text{peak.bg}}$	Peaking background fraction	0.022 ± 0.007
\mathcal{F}_{mix}	Long-lived background fraction	0.609 ± 0.0455
\mathcal{F}_{B_d}	Oscillating background fraction	0.462 ± 0.057
\mathcal{F}_{neg}	Negative exponential fraction	0.0022 ± 0.005
s_{bg}	Background scale factor	2.51 ± 0.06
$c\tau_{\text{neg}}$	Negative exponential lifetime	$72 \pm 20 \mu\text{m}$
$c\tau_{bg}$	Background lifetime	$771 \pm 8 \mu\text{m}$
$c\tau_{B_s}$	B_s^0 lifetime	$490 \pm 39 \mu\text{m}$

Table 6.3: Parameters determined from a lifetime fit to the full tagged sample.

For a given type of B hadron (i.e., d , u , s), the distribution of the VPDL is modified from equation 6.16 and 6.17 to,

$$p_s^{\text{nos}}(x) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5 \cdot (1 + \mathcal{A} \cdot \mathcal{D} \cos(\Delta m_s \cdot Kx/c)), \quad (6.39)$$

$$p_s^{\text{osc}}(x) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5 \cdot (1 - \mathcal{A} \cdot \mathcal{D} \cos(\Delta m_s \cdot Kx/c)). \quad (6.40)$$

where τ is the lifetime of B hadron, K is the K factor and \mathcal{A} is a fit parameter. Different values of Δm_s are fixed as input parameters and a fitted value of \mathcal{A} is returned. By plotting the fitted value of \mathcal{A} as a function of the input value of Δm_s , one searches for a peak of $\mathcal{A}=1$ to obtain a measurement of Δm_s . For any value of Δm_s not equal to the “true” value of B_s oscillation frequency, the amplitude \mathcal{A} should be zero. If no peak is found, limits can easily be set on Δm_s using this method. The sensitivity of a measurement is determined by calculating the probability that at a non-“true” value of Δm_s , the amplitude could fluctuate to $\mathcal{A}=1$. This occurs at the lowest value of Δm_s for which $1.645 \sigma_{\Delta m_s} = 1$ for a 95% CL, where $\sigma_{\Delta m_s}$ is the uncertainty on the value of \mathcal{A} at the point Δm_s . The limit is determined by calculating the probability that a fitted value of \mathcal{A} could fluctuate to $\mathcal{A} = 1$. This occurs

at the lowest value of Δm_s for which $\mathcal{A}_{\Delta m_s} + 1.645\sigma_{\Delta m_s} = 1$.

6.9 Results

Fig. 6.42 shows the dependence of the parameter \mathcal{A} and its error on Δm_s . A 95% confidence level limit on the oscillation frequency $\Delta m_s > 1.10 \text{ ps}^{-1}$ and sensitivity of 1.90 ps^{-1} were obtained.

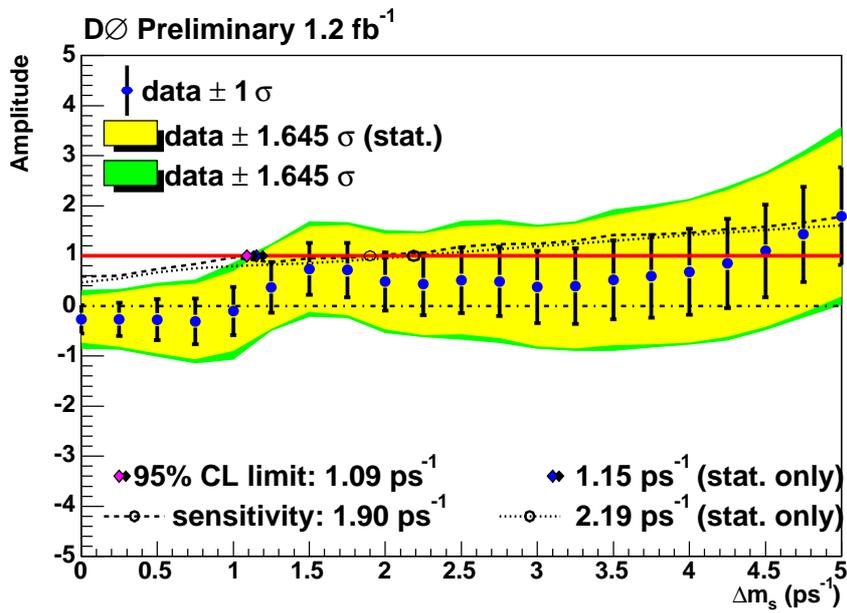


Figure 6.42: B_s^0 oscillation amplitude with statistical and systematic errors.

6.10 Cross Checks and Systematic Errors

We expect the following to contribute to the systematic uncertainty of the limit:

- Dilution;
- Mass fitting procedure;
- Resolution scale factor;
- Sample composition;
- K factors;
- Randomized flavor tagging;
- Variations of likelihood fit parameters.

The contribution to the systematic error from each variation can be estimated using the formula [29]:

$$\sigma_{\mathcal{A}}^{sys} = \Delta\mathcal{A} + (1 - \mathcal{A})\frac{\Delta\sigma_{\mathcal{A}}}{\sigma_{\mathcal{A}}} \quad (6.41)$$

Typically one of the inputs is varied or an alternate form is used and the entire analysis is repeated to calculate $\Delta\mathcal{A}$ and $\Delta\sigma_{\mathcal{A}}$ for a given value of Δm_s . The values of $\Delta\mathcal{A}$, $\Delta\sigma_{\mathcal{A}}$, and $\sigma_{\mathcal{A}}^{sys}$ are collected in Table 6.4.

6.10.1 Dilution

There are three components in the sample which oscillate at Δm_d :

1. $D^+ \rightarrow K_S^0\pi$;
2. $D^+ \rightarrow K_S^0K$;
3. A long-lived background component.

We scan components 1) and 3) separately and then scan all three together for Δm_d .

Figure 6.43 shows the amplitude scan for component 1.

Figure 6.44 shows the amplitude scan for component 3.

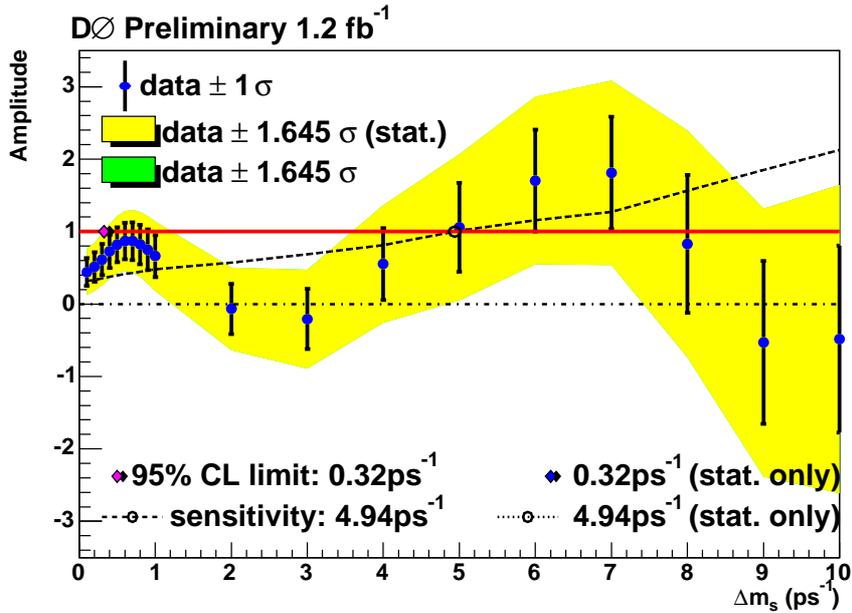


Figure 6.43: The $B_d^0 - \bar{B}_d^0$ oscillation amplitude scanning the $D^+ \rightarrow K_S^0 \pi$ component.

Figure 6.45 shows the amplitude scan for all three components.

The amplitude peak at $\Delta m_d \approx 0.5 \text{ ps}^{-1}$ is in agreement with 1 for all three scans, confirming both that the dilution calibration was performed correctly and that there is a combinatorial background component oscillating at Δm_d . After transforming the scan in Fig. 6.45 to a likelihood referenced to infinity as described in Ref. [29], we obtain $\Delta m_d = 0.50 \pm 0.13$, in agreement with the world average [57].

This cross-check also shows the ability of the method to detect an oscillation signal and the ability of the asymmetry mass fitting procedure to

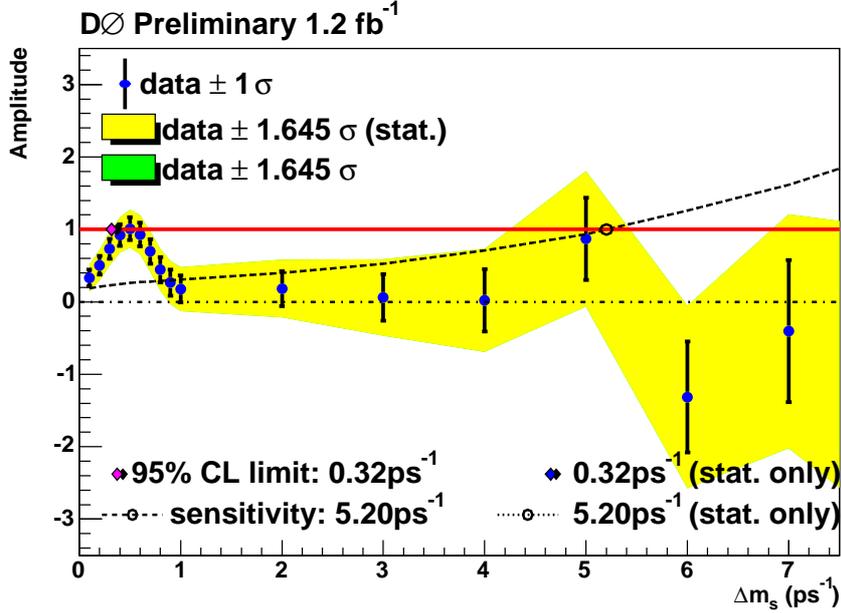


Figure 6.44: The $B_d^0 - \bar{B}_d^0$ oscillation amplitude scanning the long-lived oscillating background component.

accurately separate reflection components.

We also perform a systematic variation of the dilution calibration by using an alternate calibration function as in Eq. 6.42:

$$\mathcal{D}(d_{pr}) = \frac{0.6}{1 + \exp\left(-\frac{d_{pr} - 0.312}{0.108}\right)}. \quad (6.42)$$

6.10.2 Mass fitting procedure

We refit the tagged mass spectrum fixing $N(D_s)$ at $\pm 1.15\sigma$ and obtaining the other yields as a systematic variation. The multiplicative factor 1.15 is taken from the width of the signal yield pull obtained from toy Monte Carlo studies and shown in Fig. 6.11. We also do fits fixing $N(D^+)$ at $\pm 1\sigma$ as another systematic variation. We vary the multiplicative constraint on

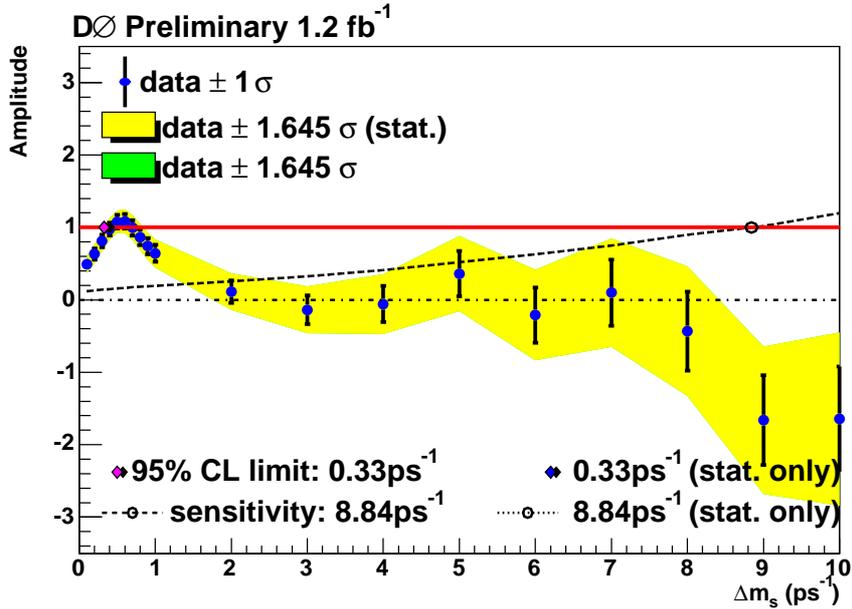


Figure 6.45: The $B_d^0 - \bar{B}_d^0$ oscillation amplitude scanning all three Δm_d components.

the Cabbibo-suppressed mode $N(D^+ \rightarrow K_S^0 K)/N(D^+ \rightarrow K_S^0 \pi) = 0.13$ by ± 0.035 , taking into account statistical and systematic uncertainties in the Monte Carlo samples used to calculate the constraint. Lastly, we use the shape of the background mass spectrum from the entire tagged sample (see the dashed line in Fig. 6.15) as an alternate background VPDL shape-dependent parameterization.

6.10.3 Resolution scale factor

We apply a cut of $p_T(\mu) > 6 \text{ GeV}/c$ to the procedure in Sec. 6.6 to obtain a systematic variation on the resolution scale factor. Without this cut, the scale factor values are (see Sec. 6.6): $\sigma_1 = 2.48 \pm 0.16$, $\sigma_2 = 0.966 \pm 0.46$, and $f = 0.15 \pm 0.03$. The scale factor values obtained after applying the $p_T(\mu)$

cut are: $\sigma_1 = 2.7 \pm 0.4$, $\sigma_2 = 1.05 \pm 0.06$, and $f = 0.09 \pm 0.04$. The average scale factor for this cut is then $\tilde{\sigma} = 1.19 \pm 0.05$.

6.10.4 Sample Composition

The $B_s \rightarrow D_s D_s$ branching was changed from 10% to 4.7% (EvtGen value) and 23% (PDG value). The branching ratio $B_s \rightarrow D_s \mu X$ was also changed from 7.9% to 5.5% (PDG uncertainty). The variation of the branching ratio $B_s \rightarrow D_s \mu X$ gives the largest change in the signal fraction.

The sample composition was also determined with the muon p_T cut of greater than 6 GeV (see Sec. 6.5.1).

6.10.5 K factors

Four additional sets of K factor distributions were generated to estimate contributions to the systematic error.

We vary the K factors by $\pm 2\%$ because that is the maximum variation we observe in the means of the K factor distributions when we apply the cut $p_T(\mu) > 6$ GeV/ c . In one set, the K factor defined in Eq. 3.8 was scaled up by 2% (i.e., multiplied by 1.02). In the second set of K -factor distributions, the K factor was scaled down by 2% (i.e., multiplied by 0.98).

In the third set, the distributions were smoothed using the ROOT function “Smooth” (with argument 1, which applies the smoothing algorithm once).

A final set of histograms was generated using the definition

$$K = p_T^{reco}(\mu D_s^-) / p_T^{MC}(B). \quad (6.43)$$

The resulting systematic errors were obtained using Eqn. 6.41 and summed in quadrature. The result is shown in Fig. 6.42 and in Table 6.4.

6.10.6 Randomized flavor tagging

As a cross-check, we simulate $\Delta m_s = \infty$ by randomizing the sign of d_{pr} and scan for B_s^0 oscillations. We obtain a sensitivity of 2.13 ps^{-1} using statistical errors only, similar to the unblinded sensitivity of 2.19 ps^{-1} .

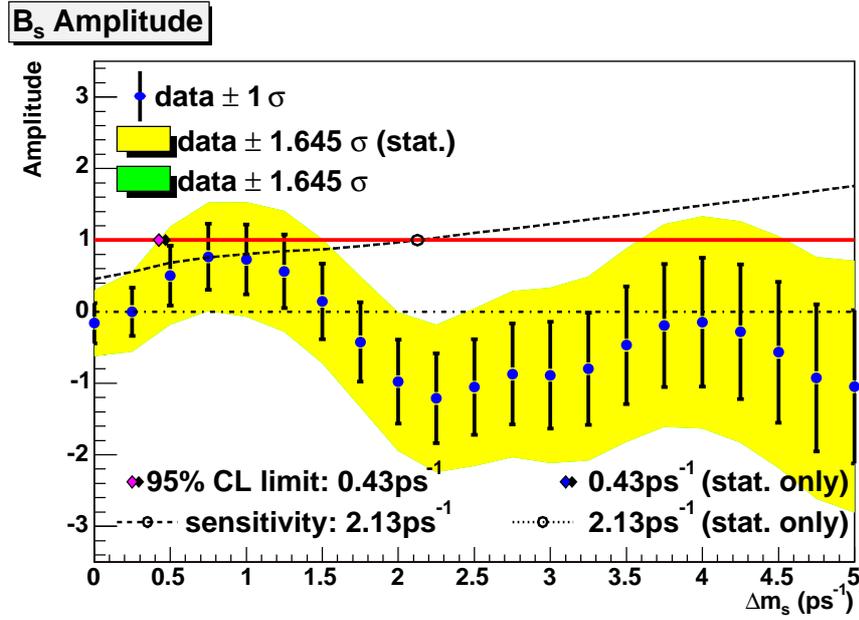


Figure 6.46: B_s^0 oscillation amplitude with randomized flavor tagging. Only statistical errors are shown.

6.10.7 Variations of fit parameters

The following systematic variations of fit parameters were considered:

- $Fr_{c\bar{c}} + 1\sigma$,
- $s_{bg} = 2.0$,
- $\mathcal{F}_{mix} \pm 1\sigma$,

- $\mathcal{F}_{Bd} \pm 1\sigma$,
- $\mathcal{F}_{neg} \pm 1\sigma$,
- $\mathcal{F}_{peak.bg} \pm 1\sigma$,
- $c\tau_{B_s} = 438 \mu m$, the world average.

Table 6.4: Systematic uncertainties on the amplitude. The shifts of both the measured amplitude, $\Delta\mathcal{A}$, and its statistical uncertainty, $\Delta\sigma$, are listed

Osc. frequency (ps^{-1})		0.00	0.50	1.00	1.50	2.00	2.50	3.00	3.50	4.00	4.50	5.00
\mathcal{A}		-0.284	-0.282	-0.188	0.771	0.432	0.539	0.358	0.543	0.668	1.130	1.934
Stat. uncertainty		0.281	0.420	0.489	0.528	0.595	0.678	0.754	0.829	0.908	0.980	1.036
Dilution	$\Delta\mathcal{A}$	-0.013	-0.034	-0.006	-0.022	+0.002	+0.002	-0.003	+0.008	+0.049	+0.069	+0.085
	$\Delta\sigma$	-0.002	-0.003	-0.004	-0.003	-0.003	-0.004	-0.005	-0.008	-0.009	-0.009	-0.010
Scale Factor	$\Delta\mathcal{A}$	-0.002	+0.009	+0.027	+0.024	+0.007	-0.014	-0.029	-0.038	-0.038	-0.030	-0.018
	$\Delta\sigma$	-0.000	+0.000	-0.000	-0.000	-0.000	-0.001	-0.002	-0.004	-0.003	-0.002	-0.000
$Br(B_s \rightarrow D_s \mu \nu) = 5.5\%$	$\Delta\mathcal{A}$	+0.014	+0.018	-0.002	+0.020	+0.011	+0.014	+0.009	+0.017	+0.021	+0.039	+0.069
	$\Delta\sigma$	+0.010	+0.015	+0.019	+0.020	+0.023	+0.026	+0.030	+0.033	+0.036	+0.039	+0.041
$Br(B_s \rightarrow D_s D_s) = 5.5\%$	$\Delta\mathcal{A}$	+0.012	+0.016	-0.003	+0.024	+0.014	+0.017	+0.011	+0.020	+0.025	+0.044	+0.079
	$\Delta\sigma$	+0.012	+0.018	+0.021	+0.023	+0.026	+0.030	+0.033	+0.037	+0.040	+0.044	+0.047
$Br(B_s \rightarrow D_s D_s) = 23\%$	$\Delta\mathcal{A}$	-0.008	-0.009	-0.003	+0.014	+0.010	+0.010	+0.007	+0.009	+0.012	+0.020	+0.034
	$\Delta\sigma$	+0.005	+0.008	+0.009	+0.009	+0.010	+0.012	+0.013	+0.014	+0.016	+0.017	+0.018
$c\bar{c}$: 4.62%	$\Delta\mathcal{A}$	-0.002	+0.001	+0.005	+0.018	+0.010	+0.007	+0.002	+0.002	+0.002	+0.007	+0.019
	$\Delta\sigma$	+0.003	+0.005	+0.006	+0.006	+0.007	+0.008	+0.009	+0.010	+0.012	+0.013	+0.014
$c\tau_{B_s} = 438\mu m$	$\Delta\mathcal{A}$	-0.016	-0.006	-0.079	+0.049	-0.049	+0.031	-0.024	+0.024	-0.014	+0.034	+0.150
	$\Delta\sigma$	-0.002	+0.014	+0.009	+0.013	+0.020	+0.028	+0.036	+0.046	+0.055	+0.065	+0.069
$p_{T\mu} > 6 \text{ GeV}/c$	$\Delta\mathcal{A}$	-0.035	-0.032	-0.082	+0.011	-0.070	+0.011	-0.034	+0.004	-0.038	-0.013	+0.058
	$\Delta\sigma$	-0.016	-0.008	-0.017	-0.014	-0.012	-0.008	-0.005	+0.000	+0.005	+0.010	+0.011
D_s yield $\pm 1.15\sigma$	$\Delta\mathcal{A}$	+0.071	+0.121	-0.035	+0.136	+0.019	+0.115	+0.066	+0.140	+0.079	+0.151	+0.363
	$\Delta\sigma$	-0.028	-0.026	-0.037	+0.067	+0.082	+0.098	+0.117	+0.134	+0.149	+0.166	+0.176
D^+ yield $\pm 1\sigma$	$\Delta\mathcal{A}$	+0.038	+0.079	-0.043	+0.082	-0.095	+0.083	+0.027	+0.084	+0.032	+0.097	+0.267
	$\Delta\sigma$	-0.018	-0.012	-0.021	+0.040	-0.016	+0.063	+0.076	+0.089	+0.102	+0.116	+0.123
k-factor $\pm 2\%$	$\Delta\mathcal{A}$	-0.027	-0.006	-0.108	+0.038	-0.087	+0.030	+0.005	+0.046	+0.050	+0.149	+0.258
	$\Delta\sigma$	-0.003	+0.012	+0.005	+0.011	+0.020	+0.028	+0.026	+0.047	+0.055	+0.063	+0.071
k-factor smoothed	$\Delta\mathcal{A}$	-0.017	-0.009	-0.086	+0.034	-0.055	+0.032	-0.022	+0.028	-0.010	+0.046	+0.153
	$\Delta\sigma$	-0.003	+0.011	+0.006	+0.010	+0.017	+0.025	+0.032	+0.042	+0.050	+0.060	+0.064
Reco k-factor	$\Delta\mathcal{A}$	-0.018	-0.009	-0.089	+0.040	-0.031	+0.059	-0.014	+0.048	+0.045	+0.109	+0.166
	$\Delta\sigma$	-0.003	+0.012	+0.009	+0.015	+0.024	+0.034	+0.044	+0.057	+0.068	+0.079	+0.086
BG Scale Factor = 2.0	$\Delta\mathcal{A}$	-0.018	-0.021	-0.109	+0.030	-0.062	+0.026	-0.028	+0.018	-0.021	+0.027	+0.139
	$\Delta\sigma$	-0.003	+0.015	+0.008	+0.012	+0.018	+0.026	+0.034	+0.044	+0.052	+0.061	+0.065
$frNeg + 1\sigma$	$\Delta\mathcal{A}$	-0.017	-0.010	-0.086	+0.035	-0.056	+0.028	-0.022	+0.026	-0.012	+0.035	+0.143
	$\Delta\sigma$	-0.004	+0.011	+0.006	+0.010	+0.016	+0.023	+0.031	+0.040	+0.048	+0.057	+0.061
$f_{cc}(bg) \pm 1\sigma$	$\Delta\mathcal{A}$	-0.021	-0.003	-0.090	+0.040	-0.063	+0.034	-0.019	+0.027	-0.012	+0.035	+0.144
	$\Delta\sigma$	-0.003	+0.011	+0.006	+0.010	+0.016	+0.023	+0.031	+0.040	+0.048	+0.057	+0.061
frBd $\pm 1\sigma$	$\Delta\mathcal{A}$	-0.060	-0.007	-0.115	+0.045	-0.060	+0.029	-0.022	+0.026	-0.011	+0.034	+0.142
	$\Delta\sigma$	-0.003	+0.012	+0.006	+0.009	+0.016	+0.023	+0.031	+0.041	+0.048	+0.057	+0.060
frMix $\pm 1\sigma$	$\Delta\mathcal{A}$	-0.067	+0.059	-0.111	+0.041	-0.076	+0.048	-0.006	+0.038	-0.001	+0.044	+0.150
	$\Delta\sigma$	-0.003	+0.011	+0.005	+0.009	+0.017	+0.022	+0.029	+0.038	+0.046	+0.055	+0.059
Background Mass Shape	$\Delta\mathcal{A}$	-0.034	-0.023	-0.064	+0.030	-0.047	+0.016	-0.021	+0.053	+0.020	+0.030	+0.100
	$\Delta\sigma$	+0.001	+0.015	+0.009	+0.012	+0.017	+0.031	+0.041	+0.048	+0.049	+0.050	+0.050
$N(D^+ \rightarrow K_S^0 K)/N(D^+ \rightarrow K_S^0 \pi) \pm 0.035$	$\Delta\mathcal{A}$	+0.005	+0.022	-0.076	+0.057	-0.074	+0.044	-0.009	+0.042	-0.002	+0.049	+0.178
	$\Delta\sigma$	-0.010	+0.002	-0.005	+0.021	+0.003	+0.037	+0.047	+0.058	+0.067	+0.078	+0.083
Total syst.	σ_{tot}^{sys}	0.216	0.186	0.338	0.260	0.236	0.300	0.211	0.343	0.208	0.270	0.467
Total	σ_{tot}	0.357	0.449	0.589	0.580	0.626	0.720	0.753	0.860	0.885	0.962	1.081

Chapter 7

Discussion and Conclusions

7.1 Combination and Conclusions

We use the amplitude method to combine the results presented in Chapter 6 with the other three analyzed results at $D\bar{D}$: $\mu\phi\pi$ [16], $\mu\phi e$ [58], and μK^*K [59]. The advantage of the amplitude method is that the result of the analysis in each channel is a single amplitude value, with statistical and systematic uncertainties, at each scanned value of Δm_s . The combination is performed by taking the weighted average of amplitudes for a combined amplitude scan.

As in Ref. [51], we use the COMBOS program [60] developed at LEP to combine results, taking into account correlated errors properly. We combine the $\mu\phi\pi$, $\mu\phi e$, μK^*K , and $\mu K_S^0 K$ modes taking the following uncertainties as 100% correlated:

- $Br(B_s \rightarrow X\mu D_s)$;
- $Br(B_s \rightarrow XD_s D_s)$;
- Signal decay length resolution for all semi-muonic modes;
- $\Delta\Gamma/\Gamma$.

Figure 7.1 shows the result of the combined amplitude scan. A limit of $\Delta m_s > 14.9 \text{ ps}^{-1}$ at 95% C.L. is obtained, with an expected limit of 16.5 ps^{-1} . However, an excess is observed, i.e., a signal consistent with amplitude ≈ 1 at $\Delta m_s \sim 19 \text{ ps}^{-1}$. The value of the amplitude at $\Delta m_s = 19 \text{ ps}^{-1}$ is $\mathcal{A} = 1.05 \pm 0.76$ (*stat.*) corresponds to a 16.8% fluctuation away from the null hypothesis, $\mathcal{A}(\Delta m_s = 19 \text{ ps}^{-1}) = 0$. The probability that given $\Delta m_s = 19 \text{ ps}^{-1}$ we would see a fluctuation to the observed value of $\mathcal{A}(\Delta m_s = 19 \text{ ps}^{-1})$ is 94.7%.

Once a possible signal is observed, it is more appropriate to examine it with a likelihood scan to be able to assign errors and/or confidence regions to the measurement. We convert the amplitude scan to a log likelihood scan by using the following formulae [29],

$$\begin{aligned}\mathcal{L} &\equiv -\Delta \log(L) = \frac{1}{2} \left(\frac{1 - 2\mathcal{A}}{\sigma_{\mathcal{A}}^2} \right)^2, \\ \sigma_{\mathcal{L}} &= \frac{1}{\sigma_{\mathcal{A}}}.\end{aligned}\tag{7.1}$$

Figure 7.2 shows the resulting log likelihood scan. The preferred value of the oscillation frequency is $\Delta m_s = 19 \text{ ps}^{-1}$ with a 90% confidence level interval ($\mathcal{L} = 1.355$) of $17 < \Delta m_s < 21 \text{ ps}^{-1}$, assuming Gaussian uncertainties. In the original DØ publication of the $\mu\phi\pi$ channel [16], this represented the first ever two-sided bound on the oscillation frequency for B_s . We note that the Δm_s range observed above is consistent with the Standard Model prediction obtained from CKM fits where no experimental information on Δm_s is used, $\Delta m_s^{\text{indirect}} = 18.4 \pm 2.4 \text{ ps}^{-1}$ [61].

In the previous analysis using only the $B_s^- \rightarrow D_s^- \mu^+ X$ ($D_s^- \rightarrow \phi\pi^-$) decay mode [16], the probability of background fluctuations to give a minimum of equal or greater depth in this interval was determined to be 5% using ensemble tests. Comparing the change in likelihood at $\Delta m_s = 19 \text{ ps}^{-1}$ and the likelihood at $\Delta m_s = \infty$ [62] also yields a 5% probability for a background fluctuation. For the combined results shown in Figs. 7.1 and 7.2, a compari-

son of the change in likelihood between $\Delta m_s = 19 \text{ ps}^{-1}$ and $\Delta m_s = \infty$ yields an 8% probability for a background fluctuation.

Figures 7.3 and 7.4 show the constraints on the unitarity triangle in the $\bar{\rho}-\bar{\eta}$ plane before and after the inclusion of the combined DØ Δm_s constraint presented in this work [23]. The value of the right side of the unitarity triangle defined as,

$$R_t \equiv \sqrt{(1 - \bar{\rho})^2 + \bar{\eta}^2}, \quad (7.2)$$

goes from $R_t = 0.863_{-0.041}^{+0.047}$ before the DØ result to $R_t = 0.849_{-0.025}^{+0.073}$ with the inclusion of the results presented in Figs. 7.1 and 7.2. When the subsequent CDF measurement of $\Delta m_s = 17.77 \pm 0.10(\text{stat.}) \pm 0.07(\text{syst.})$ [17] is included $R_t = 0.868_{-0.025}^{+0.060}$. Note that the lower bound on R_t is significantly improved just through the inclusion of the DØ two-sided bound on Δm_s .

If we interpret the DØ combined result as a measurement of Δm_s , we can use it to extract $|V_{td}/V_{ts}|$ from Eq. 2.66. As inputs, we use the lattice QCD result given in Eq. 2.67, $m(B^0)/m(B_s^0) = 0.98390$ [63] with negligible uncertainty, and $\Delta m_d = 0.507 \pm 0.005$ [13]. We obtain,

$$\left| \frac{V_{td}}{V_{ts}} \right| = 0.199 \pm 0.003(\text{exp.}) \begin{matrix} +0.008 \\ -0.006 \end{matrix}(\text{lat.}). \quad (7.3)$$

As can be seen above in Eq. 7.3, the dominant uncertainties on $|V_{td}/V_{ts}|$ after the inclusion of the combined DØ result are theoretical.

We use the CDF and DØ Δm_s results to put constraints on new phenomena in a model-independent manner. Following the method of Ref. [61] we define,

$$C_{B_q} e^{2i\phi_{B_q}} = \frac{\langle B_q | H_{\text{eff}}^{\text{full}} | \bar{B}_q \rangle}{\langle B_q | H_{\text{eff}}^{\text{SM}} | \bar{B}_q \rangle}, \quad (7.4)$$

so that the shift induced in the $B_q-\bar{B}_q$ mixing frequency by new phenomena effects is parameterized by C_{B_q} and the corresponding change in the phase

is parameterized by ϕ_{B_q} , i.e.,

$$\begin{aligned}\Delta m_s &= C_{B_s} \cdot \Delta m_s^{\text{SM}} \quad \text{and} \\ \beta_s^{\text{exp}} &= \beta_s^{\text{SM}} - \phi_{B_s}.\end{aligned}\tag{7.5}$$

In the absence of new phenomena, we expect $C_{B_q} = 1$ and $\phi_{B_q} = 0$.

Figure 7.6 shows the bounds on the ϕ_{B_s} vs. C_{B_s} plane using all available data on the B_s^0 system. The constraints are $C_{B_s} = 1.13 \pm 0.35$ and $\phi_{B_s} = (-3 \pm 19) \cup (94 \pm 19)^\circ$ [61]. Both values are consistent with SM expectations. Note that the measurements of Δm_s from CDF and DØ strongly constrain C_{B_s} so that it is already known better than C_{B_d} whose current value is $C_{B_d} = 1.25 \pm 0.43$.

In summary, using a signal of $593 B_s^0 \rightarrow D_s^- \mu^+ X$ where $D_s^- \rightarrow K_S^0 K^-$ and an opposite-side flavor tagging algorithm, we performed a search for $B_s^0 - \bar{B}_s^0$ oscillations. We obtain a 95% confidence level limit on the oscillation frequency $\Delta m_s > 1.10 \text{ ps}^{-1}$ and an expected limit of 1.92 ps^{-1} . Results are presented when this new channel is combined with other decay channels from the DØ Collaboration. The combined result provides powerful constraints on the CKM unitarity triangle. Results in all cases are consistent with Standard Model expectations.

7.2 Outlook

As Eq. 3.20 indicates, there are four main ways in which the sensitivity to oscillations may be improved:

- increase the statistics of the sample,
- improve the signal-to-background ratio,
- improve the flavor tagging performance $\epsilon \mathcal{D}^2$, and
- improve the proper time resolution.

DØ is currently making improvements in all three categories. Some details of these improvements are described below.

7.2.1 Increasing the Statistics and Signal to Background

The DØ experiment continues to operate in a stable mode and collect data. Current projections are that the final delivered luminosity will be $\mathcal{L} \approx 8 \text{ fb}^{-1}$. The results presented in this analysis as well as in Refs. [16] and [51] are based on a 1 fb^{-1} dataset. In addition, new modes such as $B_s \rightarrow D_s \pi$ ($D_s \rightarrow \phi \pi$) are being added. To improve the signal to noise ratio, selections involving boosted decision trees are now in development.

7.2.2 Improving the Flavor Tagging

In this analysis, we have only used properties of the decay products of the B hadron opposite to the B_s decay we are analyzing to determine the initial state flavor. DØ is now developing techniques to use objects on the same side of the event as the reconstructing meson for flavor tagging. These techniques will improve the efficiency of the tagging algorithm and thereby better its performance. In addition, studies are underway to use the energy loss in the silicon, dE/dx , to separate pions from kaons. Because the hadronization products associated with a B_s must have strangeness, same-side kaon identification can improve $\epsilon \mathcal{D}^2$.

7.2.3 Improving the Proper Time Resolution

The installation of an inner layer of silicon, called “layer 0”, near the beampipe in the Summer of 2006 allows us to take data with smaller tracking and vertexing errors. This data will therefore have improved proper time resolution. In addition, hadronic events such as the one mentioned above ($B_s \rightarrow D_s \pi$) tend to have better proper time resolution because they are fully

reconstructed. The events will therefore be more sensitive to oscillations and should greatly improve the measurement of Δm_s .

Because of these improvements, the prospects for DØ to observe a significant signal for B_s mixing in 2007 or 2008 are strong.

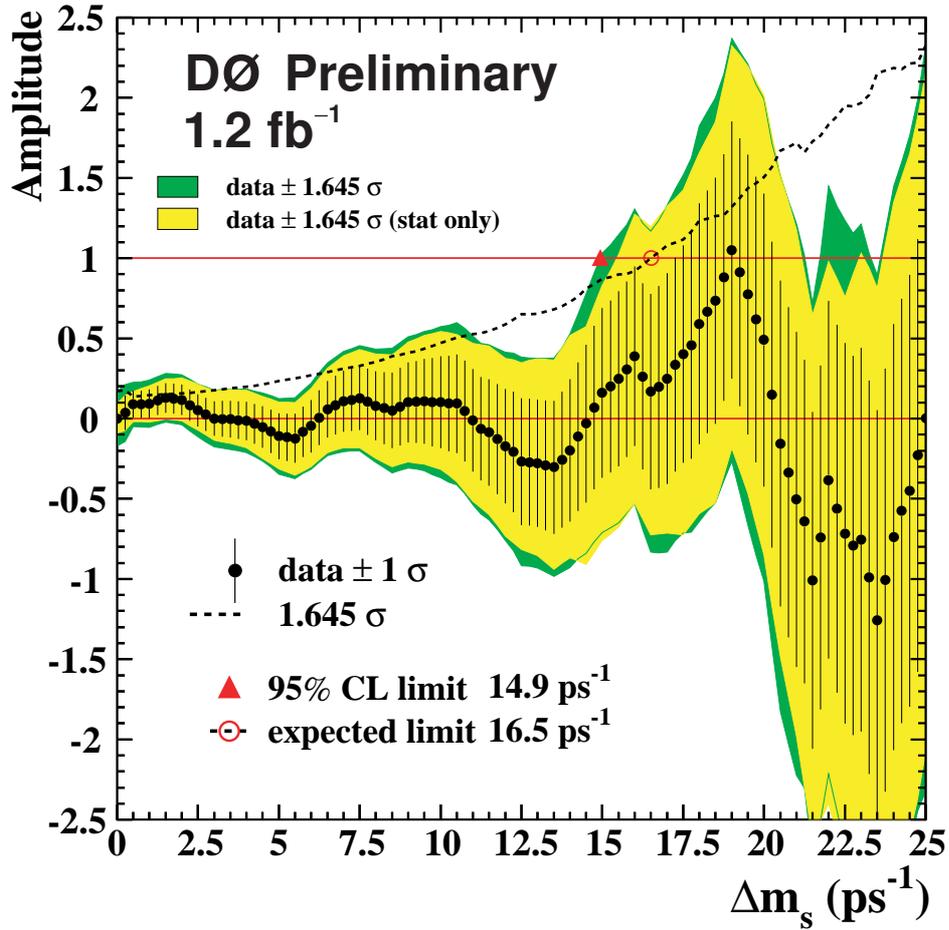


Figure 7.1: B_s^0 oscillation amplitude with statistical and systematic errors for $B_s^0 \rightarrow D_s^- e^+ X$ ($D_s^- \rightarrow \phi \pi^-$) and $B_s^0 \rightarrow D_s^- \mu^+ X$ ($D_s^- \rightarrow \phi \pi^-$, $D_s^- \rightarrow K^* K$ and $D_s^- \rightarrow K_S^0 K^-$).

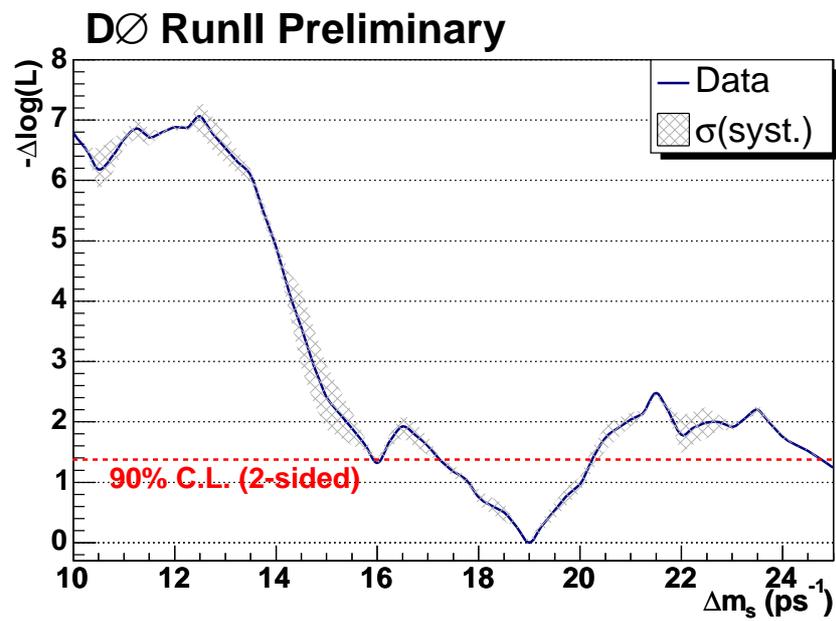


Figure 7.2: Log likelihood scan for the combination of the four decay modes listed in the text obtained from the amplitude scan using the statistical uncertainty with a systematic uncertainty band overlaid.

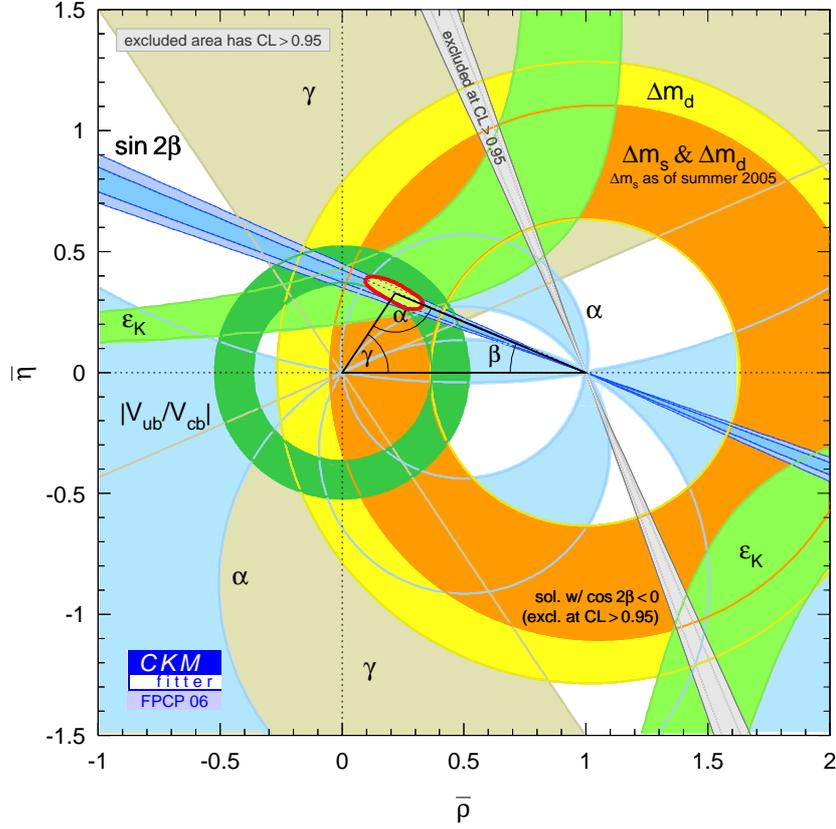


Figure 7.3: Experimental constraints on the $\bar{\rho}\text{-}\bar{\eta}$ plane before the inclusion of the new $D\bar{O}$ combined Δm_s result. The Δm_s data used for this plot was current as of the Summer of 2005. The green circle centered at $(0,0)$ comes from charmless B decays and the hyperbolic curves come from kaon CP -violation results. The area between the blue lines passing through $(1,0)$ come from measurements of $\sin 2\beta$. The yellow circle constraining the right side of the triangle comes from measuring Δm_d , and the orange circle is the extra constraint from Δm_s measurements. The red ellipse indicates the intersection of all constraints at 95% C.L. [23].

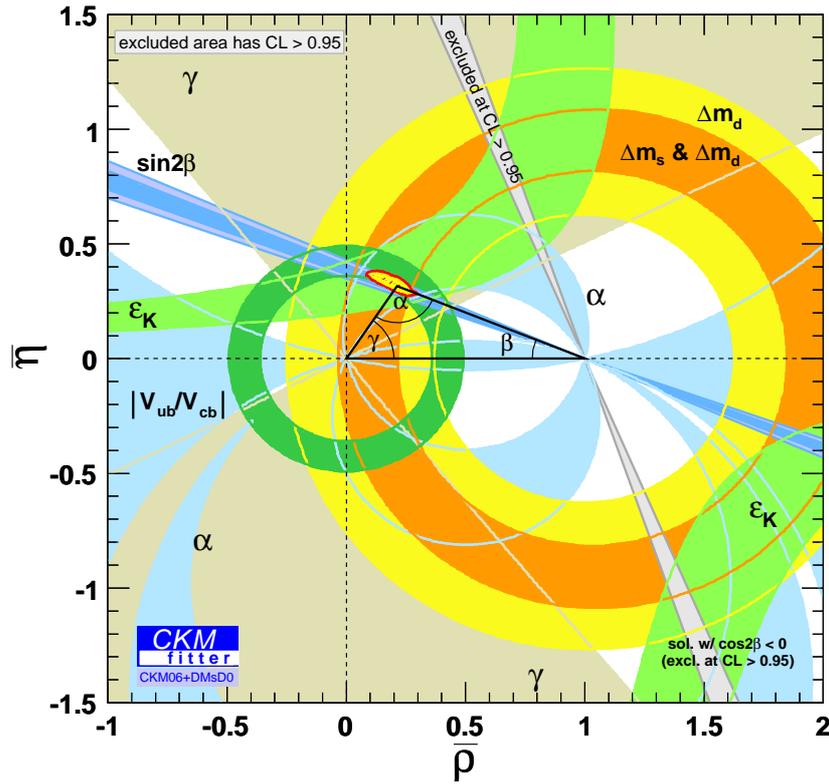


Figure 7.4: Experimental constraints on the $\bar{\rho}$ - $\bar{\eta}$ plane after the inclusion of the new $D\bar{O}$ combined Δm_s result. The colors and curves on the plot represent the same constraints as in Fig. 7.3 [23]. Note the reduction in the area of the orange annulus and constrained region of the triangle vertex when compared to Fig. 7.3.

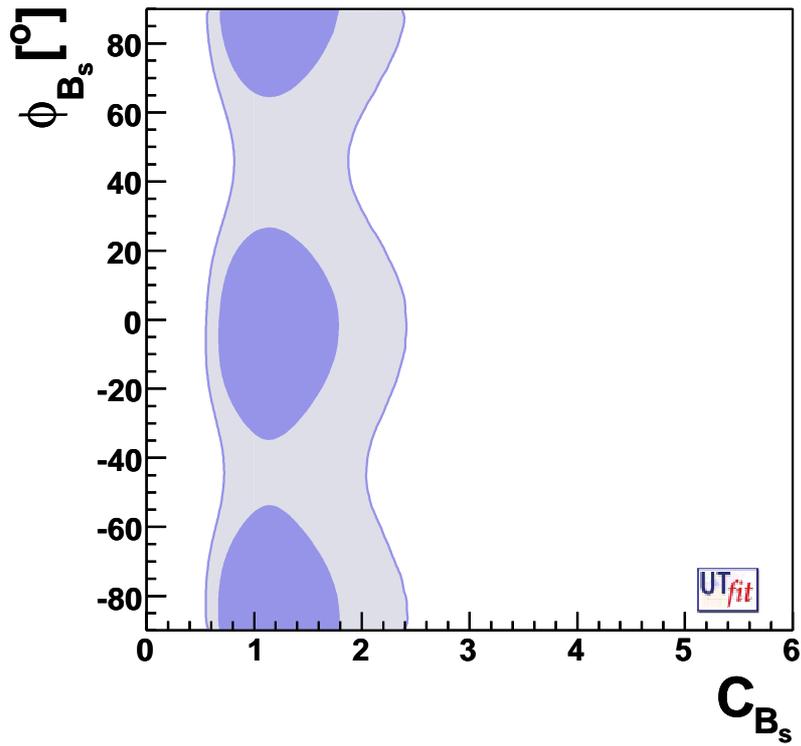


Figure 7.6: Constraints on the model-independent new phenomena parameters C_{B_s} and ϕ_{B_s} (see Eq. 7.4) in the B_s^0 system. The dark (light) area corresponds to the 68% (95%) probability region [61].

Appendix A

Luminosity Monitor Calibration

The author performed significant work on the luminosity monitor that is described in this appendix. As mentioned in Sec. 4.2.5, time and charge signals from the luminosity monitor are digitized in six time to digital converter (TDC) boards. These TDC's must be calibrated to linearize the charge output and time-of-flight information as well as correct for charge-slewing.

Each luminosity monitor TDC board accepts eight photomultiplier signals and a common stop signal via front-panel LEMO connectors. The TDC's precisely measure the arrival time of particles striking the scintillator so that the luminosity can be determined in accord with Eq. 4.4. Time-to-charge conversion is performed on each channel by switching on a current source when the PMT signal crosses a programmable threshold and switching off the current source when the common stop signal is received. The charge from the switched current source is integrated and digitized using CAFÉ daughter cards developed for the CDF calorimeter readout [64]. The PMT signals are also fed into a second CAFÉ card to measure their charge. This charge is used to generate a time-slewing correction to maintain good timing resolution over a wide range of scintillator pulse-heights.

The calibration of these TDC boards proceeds in three stages:

- calibration of individual CAFÉ cards,
- global time calibration using collider data,
- loading lookup tables (LUT's) using the calibration constants determined in the previous two steps.

Lookup tables are stored onboard the CAFÉ cards in 1 Megabit AMD Am29F100 flash memory [65].

A.1 CAFÉ Calibration

CAFÉ daughter cards have a variable current source for calibration. The magnitude of the calibration current is determined by the 16-bit VCAL DAC, which provides a 0–10 V calibration voltage. The calibration circuit produces a 1 mA/V current source which is then integrated for 132 ns. The integrated charge provided by the circuit is given by,

$$\begin{aligned} Q &= (132 \text{ ns}) \times \left(\frac{VCAL}{2^{16}} \times 10 \text{ V} \right) \times (1 \text{ mA/V}) \\ &= (0.0201 \text{ pC}) \times VCAL, \end{aligned} \tag{A.1}$$

where $VCAL$ is the value of the VCAL DAC. By measuring the average ADC output of a CAFÉ card as a function of the VCAL DAC setting, the CAFÉ cards may be calibrated. CAFÉ cards have eight ranges, with approximately a factor of two difference in gain between the ranges, and four separate integration capacitors for each range. There are therefore 32 sets of calibration constants for each CAFÉ card.

For a given range and integration capacitor, the CAFÉ cards are nearly linear over their operating range. A linear fit should be sufficient to calibrate

a given CAFÉ range and integration capacitor,

$$Q = Q_0 + Q_{\text{Slope}} \times ADC. \quad (\text{A.2})$$

The data stored for each CAFÉ calibration are shown in Table A.1.

TDC Module ID
TDC Channel Number (0 – 7)
QFlag (0 = Time, 1 = Charge)
Range (0 – 7)
Capacitor ID (0 – 3)
Q_0 (pC)
Q_{Slope} (pC/ADC count)
RMS of Fit Residual (pC)

Table A.1: CAFÉ calibration constants.

A.2 Global Time Calibration

The global time calibration provides the time-of-flight calibration for each counter. It is performed with a global fit of all luminosity monitor channels using collider data. The best estimate of the particle arrival time at the counters, including a charge slewing correction is,

$$T = T_0 + T_{\text{Slope}} \times Q_T - \frac{K}{\sqrt{Q_P}}, \quad (\text{A.3})$$

where Q_T is the charge measured by the arrival time CAFÉ card, Q_P is the charge measured by the pulse height CAFÉ card, and T_0 , T_{Slope} , and K are calibration constants. The expected arrival time for a particle from a single $p\bar{p}$ interaction depends on the z -coordinate of the interaction vertex, z_v , and

the time the interaction took place, T_I :

$$\begin{aligned}\bar{T}_N &= T_I + \frac{L + z_v}{c}, \\ \bar{T}_S &= T_I + \frac{L - z_v}{c},\end{aligned}\tag{A.4}$$

where $\bar{T}_{N(S)}$ is the expected time for the north (south) counters and L is the distance between the counters and the origin. The time origin, $T = 0$, is arbitrary and is defined below.

The calibration constants T_0 , T_{slope} , and K for each channel are obtained from a global fit of the luminosity monitor time and charge measurements for a sample of collider data. This is accomplished by minimizing the following χ^2 ,

$$\chi^2 = \sum_{Events} \left(\langle T_N \rangle - \langle T_S \rangle - \frac{2z_v}{c} \right)^2 + \frac{\lambda}{N_{Hits}} \sum_{Events} \sum_{Counters} T,\tag{A.5}$$

where $\langle T_{N(S)} \rangle$ is the average time for the north (south) LM counters. The first term in the χ^2 is used to fit the LM vertex position to the tracker vertex position. The second term removes the ambiguity in the definition of $T = 0$ by imposing a Lagrange constraint that $\langle T \rangle = 0$, averaging over all hits in the data sample. The three calibration constants for all 48 LM channels are determined simultaneously by minimizing the above χ^2 with respect to the calibration constants and the Lagrange multiplier λ . Table A.2 shows the calibration constants stored for each counter following the global time calibration.

A.3 Loading Lookup Tables

The lookup table (LUT) operates in two modes depending on the setting of an “external control” bit. In VME mode, the LUT address is determined

TDC Module ID
TDC Channel Number (0 – 7)
T_0 (ns)
T_{Slope} (ns/pC)
K (ns $\sqrt{\text{pC}}$)
RMS of time residual (ns)

Table A.2: Luminosity monitor global time calibration constants.

by the VME address. In readout mode, the address of the LUT depends on the capacitor ID, range, and ADC values as shown in Table A.3. The LUT

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
H/L	Cap ID	Range				ADC									

Table A.3: Lookup table address bitmap. If the 16th bit, H/L , is high, the table operates in “pass-through” mode where the data output is identical to the address.

contains 2^{16} addressable 16-bit data words which have been divided into two halves. The H/L bit determines which half of the LUT is used. If $H/L = 1$, a permanently programmed “pass-through” table is used where the LUT data output is identical to the address input. We use pass-through mode to determine individual CAFÉ card calibration constants as described above in Sec. A.1. If $H/L = 0$, the output of the LUT is determined by the values programmed into the lower half of the lookup tables. The address 0xFFFF is a special case, returning the serial number of the CAFÉ card.

The values programmed into the LUT depend on the calibration constants determined from the CAFÉ and global time calibrations. For a given LUT address, the *Cap ID* and *Range* bits are decoded and used to determine the CAFÉ calibration constants Q_0 and Q_{Slope} . The ADC value associated with

this address is then used to calculate the CAFÉ charge as shown in Eq. A.2.

For arrival time CAFÉ cards, the LUT output is calculated as follows,

$$t = T_{off} + \text{NINT} \left(\frac{T_0 + T_{\text{Slope}} \cdot Q_T}{T_{lsb}} \right), \quad (\text{A.6})$$

where T_{off} defines the time lookup table offset, T_{lsb} is the time binning, and NINT is the nearest integer function. We are using 12-bit time measurements and therefore, a valid time measurement must have $0 < t < 0\text{xFFF}$. The data for the arrival time LUT is loaded as shown in Table A.4. The valid time

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
0	0	0	VT	Time											

Table A.4: Arrival time CAFÉ lookup table data bitmap. The VT bit is high if $T_{min} < t < T_{max}$ and low otherwise.

bit, VT , is set to 1 if t is within the range specified by the global calibration constants T_{min} and T_{max} .

For pulse height CAFÉ cards, the charge slewing correction is given by,

$$\text{Slew} = S_{off} - \text{NINT} \left(\frac{K}{\sqrt{Q_P} \cdot T_{lsb}} \right), \quad (\text{A.7})$$

where S_{off} defines the LUT output for infinite pulse height, and T_{lsb} is the same time binning constant used to calculate the time for arrival time CAFÉ cards. We have 7 bits available to store the slew and therefore valid charge measurements must give rise to a slew in the range, $0 < \text{Slew} < 0\text{x7F}$. A value of $\text{Slew} = 0$ indicates the pulse height Q_P was outside the range specified by the global calibration constants QS_{max} and QS_{min} .

In addition to storing the slew correction, we also store the measured pulse height in the LUT's for the pulse height CAFÉ cards. Since the pulse height has a large dynamic range, it is divided into four ranges and stored in

a piecewise linear fashion. Each range has a minimum and maximum charge, Q_{min} and Q_{max} , associated with it. The appropriate range is determined by finding where $Q_{min} \leq Q_P < Q_{max}$. Nominal values for Q_{min} and Q_{max} are given in Table A.5. The pulse height value is given by,

Q Range	Q_{min}	Q_{max}
0	-2 pC	10.8 pC
1	10.8 pC	36.4 pC
2	36.4 pC	87.6 pC
3	87.6 pC	190 pC

Table A.5: Charge ranges for storing pulse height information in lookup tables on pulse height CAFÉ cards

$$Q_{Data} = 128 \times \text{INT} \left(\frac{Q_P - Q_{min}}{Q_{max} - Q_{min}} \right), \quad (\text{A.8})$$

where the INT function truncates the fractional part of its arguments. If $Q_P < Q_{min}$ ($Range = 0$), then the range and data are set to 0. If $Q_P > Q_{max}$ ($Range = 3$), then the range is set to 3 and the data is set to 0x1FF. The slew and pulse height corrections are loaded into the LUT as shown in Table A.6.

15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
0	0	0	VT	Time											

Table A.6: Pulse height lookup table bitmap.

The TDC board also calculates a “corrected time” by summing Time and Slew. The least significant 8 bits of the corrected time is included in the TDC event data. The global calibration constants defined above are summarized in

Table A.7. A schematic overview of the operations of the luminosity monitor lookup tables is shown in Fig. A.1.

Parameter	Definition	Nominal Value
T_{off}	Time offset	0x800
T_{lsb}	Time and Slew binning	50 ps
T_{min}	Minimum valid Time	0
T_{max}	Maximum valid Time	0x990
S_{off}	Slew offset	0x80
QS_{min}	Minimum valid Q_P	-9999 pC
QS_{max}	Maximum valid Q_P	9999 pC
Q_{min}	Lower Q_P bin edges	(-10, 54, 182, 438) pC
Q_{max}	Upper Q_P bin edges	(54, 182, 438, 950) pC

Table A.7: Global luminosity monitor calibration constants.

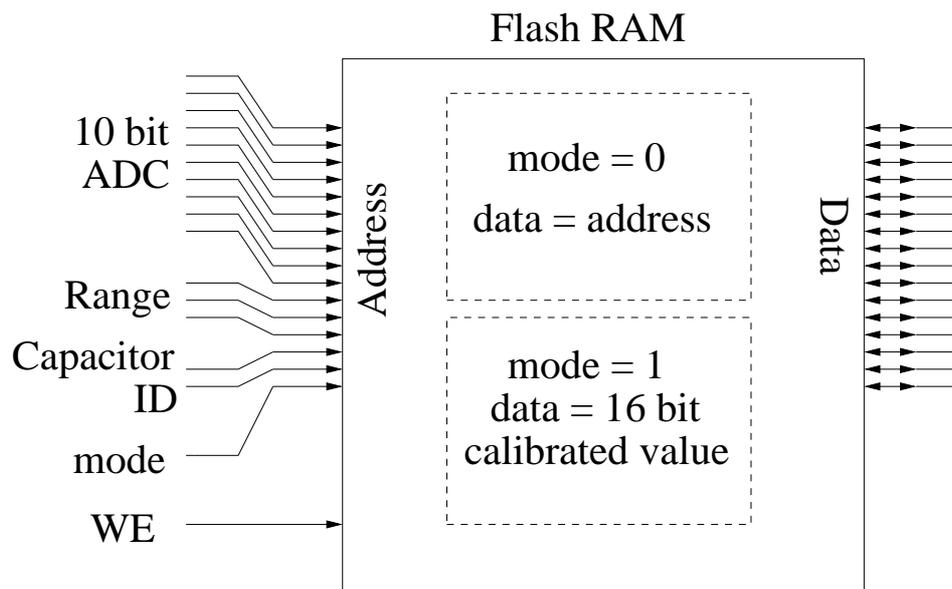


Figure A.1: Schematic drawing of the operation of lookup tables loaded on Am29F100 flash RAM on TDC CAFÉ daughter cards. *WE* is the “write-enable” bit, also referred to as the “external control” bit, which must be set for the VME bus to access the flash RAM during programming. The *mode* bit determines whether or not the table is read in “pass-through” mode where the data output is identical to the address input.

Bibliography

- [1] The LEP Electroweak Working Group, arXiv:hep-ex/0511027 (2005).
- [2] M. Gell-Mann and A. Pais, Phys. Rev. **97**, 1387 (1955).
- [3] K. Lande *et al.*, Phys. Rev. **103**, 1901 (1956).
- [4] N. Cabbibo, Phys. Rev. Lett. **10**, 531 (1963).
- [5] S. Glashow, J. Iliopoulos, L. Maiani, Phys. Rev. D **2**, 1285, (1970).
- [6] J. H. Christenson *et al.*, Phys. Rev. Lett. **13**, 138 (1964).
- [7] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
- [8] S. W. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977).
- [9] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995);
CDF Collaboration, F. Abe *et al.*, *ibid.*, 2626 (1995).
- [10] UA1 Collaboration, C. Albajar *et al.*, Phys. Lett. **B186**, 247 (1987).
- [11] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. **B192**, 245 (1987).
- [12] CLEO Collaboration, M. Artuso *et al.*, Phys. Rev. Lett. **62**, 2233 (1989).
- [13] S. Eidelman *et al.*, Phys. Lett. **B592**, 1 (2004).
- [14] Belle Collaboration, K. Abe *et al.*, Phys. Rev. D **71**, 072003 (2005);
Belle Collaboration, N. C. Hastings *et al.*, Phys. Rev. D **67**, 052004 (2003);
BaBar Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **88**, 221803 (2002).
- [15] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. **B284**, 177 (1992);
ibid., Z. Phys. **C57**, 181 (1994); DELPHI Collaboration, P. Abreu *et al.*,
Phys. Lett. **B332**, 488 (1994).

- [16] DØ Collaboration, V.M. Abazov *et al.*, Phys. Rev. Lett. **97**, 21802 (2006).
- [17] CDF Collaboration, A. Abulencia *et al.*, Phys. Rev. Lett. **97**, 62003 (2006).
- [18] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967); A. Salam, “Elementary Particle Theory”, ed. N. Svartholm, Almquist and Wiksells, Stockholm, p. 367, (1969).
- [19] M. Peskin and D. Schroeder, “An Introduction to Quantum Field Theory”, Perseus Books, Cambridge, Massachusetts, (1995).
- [20] F. Mandl and G. Shaw, “Quantum Field Theory”, John Wiley & Sons, New York, (1993).
- [21] K. Anikeev *et al.*, “B Physics at the Tevatron: Run II and Beyond”, arxiv:hep-ph/0201071, (2001).
- [22] C. Jarlskog, Phys. Rev. Lett. **55**, 1039 (1985).
- [23] CKMfitter Group, J. Charles *et al.*, Eur. Phys. J. **C41**, 1-131 (2005) [hep-ph/0406184]; updated results and plots available at <http://ckmfitter.in2p3.fr>.
- [24] F. J. Gilman *et al.* Phys. Lett. **B592**, 793 (2004).
- [25] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).
- [26] C. Gay, Ann. Rev. Nucl. Part. Sci. **50**, 577(2000).
- [27] T. Imani and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981).
- [28] M. Okamoto, plenary talk at the XXIIIth International Symposium on Lattice Field Theory, Dublin, (2005) [hep-lat/0510113]; these estimates are obtained by combining the results in HPQCD Collaboration, A. Gray *et al.*, Phys. Rev. Lett. **95**, 212001 (2005); JLQCD Collaboration, S. Aoki *et al.*, Phys. Rev. Lett. **91**, 212001 (2003).
- [29] H.G. Moser and A. Roussarie, Nucl. Instrum. Methods Phys. Res. A **384**, 491 (1997).

-
- [30] DELPHI Collaboration, J. Abdallah *et al.*, Eur. Phys. J. **C32**, 185 (2004).
- [31] S. Catani *et al.*, Phys. Lett. **B269**, 432 (1991).
- [32] T. Cole *et al.*, Technical report, Fermilab TM-1909 (1994).
- [33] S. Mishra, FERMILAB-CONF-03-194, *Presented at Particle Accelerator Conference (PAC 03)*, (2003).
- [34] S. van der Meer, CERN/ISR-PO/72-31, (1994).
- [35] DØ Collaboration, V.M. Abazov *et al.*, Nucl. Instrum. Methods Phys. Res. **A565**, 463 (2006).
- [36] A. Peisert, “Instrumentation in High Energy Physics”, ed. F. Sauli, World Scientific, New Jersey, p. 3, (1992).
- [37] T. Bose, “Search for B_s^0 oscillations at DØ”, Ph.D. Thesis, Columbia University, New York (2005).
- [38] R. Yarema *et al.*, Fermilab-TM-1892 (1994, revised 1996).
- [39] E. Segré, “Nuclei and Particles”, Benjamin, New York (1964).
- [40] R. Fernow, “Introduction to Experimental Particle Physics”, Cambridge University Press, New York (1986).
- [41] S. Snyder, “Measurement of the Top Quark Mass at DØ”, Ph.D. Thesis, S.U.N.Y., Stony Brook (1995).
- [42] V.M. Abazov *et al.*, Nucl. Instrum. Methods Phys. Res. **A552**, 372-398 (2005).
- [43] S. Abachi *et al.*, Nucl. Instrum. Methods Phys. Res. **A338**, 185 (1994).
- [44] B. Casey *et al.*, *DØ Note 4958*, “Determination of the Effective Inelastic $p\bar{p}$ Cross-Section for the DØ Luminosity Measurement Using Upgraded Electronics”, (2006).
- [45] Cisco System Inc., <http://www.cisco.com>.

- [46] G. Borrisov *et al.*, *DØ Note 4481*, “Reconstructions of B hadron signals at DØ”, (2004).
- [47] http://www-d0.fnal.gov/computing/algorithms/muon/muon_algo.html.
- [48] F. Beaudette and J.F. Grivaz, *DØ Note 3976*, “The Road Method (an algorithm for identification of electrons in jets)”, (2002).
- [49] ALEPH Collaboration, D. Buskulic *et al.*, *Zeit. Phys.* **C73**, 601 (1997).
- [50] DELPHI Collaboration, P. Abreu *et al.*, *Phys. Lett.* **B475**, 407 (2000).
- [51] G. Borrisov *et al.*, *DØ Note 5207*, Combination of B_s Oscillation Results from DØ, (2006).
- [52] G. Borisov, *Nucl. Instrum. Methods Phys. Res. A* **417**, 384 (1994).
- [53] DØ Collaboration, V.M. Abazov *et al.*, *Phys. Rev. D* **74**, 112002 (2006).
- [54] B. Abbot *et al.*, *DØ Note 4842*, “ B_s mixing semileptonic B_s decays using $D_s \rightarrow \phi\pi$ decay mode.”
- [55] F. James, MINUIT - Function Minimization and Error Analysis, CERN Program Library Long Writeup D506 (1998).
- [56] D.J. Lange, *Nucl. Instrum. Methods Phys. Res. A* **462**, 152 (2001); for details see <http://www.slac.stanford.edu/~lange/EvtGen>.
- [57] Heavy Flavor Averaging Group, <http://www.slac.stanford.edu/xorg/hfag/index.htm>.
- [58] T. Moulik *et al.*, “ B_s Mixing in $B_s^0 \rightarrow D_s^- e^+ \nu_e X$, $D_s^- \rightarrow \phi\pi^-$ Decay Mode”, *DØ Note 5174*, (2006).
- [59] Md. Naimuddin *et al.*, “ B_s Mixing Studies with $B_s^0 \rightarrow D_s^- \mu^+ X$ ($D_s^- \rightarrow K^{*0} K^-$) Decay Using Unbinned Fit”, *DØ Note 5172*, (2006).
- [60] <http://lepbosec.web.cern.ch/LEPBOSC/combos/>.
- [61] UTFit Collaboration, M. Bona *et al.*, hep-ph/0606167, hep-ph/0605213, (2006); updated results at <http://utfit.roma1.infn.it>.
- [62] D. Abbaneo and G. Boix, *J. High Energy Phys.* **08**, 004 (1999).

-
- [63] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. Lett. **96**, 202001 (2006).
- [64] T. Shaw, C. Nelson, and T. Wesson, IEEE Trans. Nucl. Sci. **47**, 1834-1838 (2000).
- [65] Advanced Micro Devices, One AMD Place, P.O. Box 3453, Sunnyvale, CA 94088-3453.