

Measurement of the B_s^0 - \bar{B}_s^0 Oscillation Frequency Using Semileptonic Decays

by

Vivek Tiwari

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

at

Carnegie Mellon University
Department of Physics
Pittsburgh, Pennsylvania

Advised by Professor Manfred Paulini

May 30, 2007

Abstract

Measurement of the B_s^0 - \bar{B}_s^0 Oscillation Frequency Using Semileptonic Decays

This thesis reports a time dependent measurement of the B_s^0 - \bar{B}_s^0 oscillation frequency Δm_s using semileptonic decays $B_s^0 \rightarrow D_s^- \ell^+ X$. We use a data sample of 1 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96 \text{ TeV}$ collected with the CDF II detector at the Fermilab Tevatron Collider to reconstruct $\sim 61,500$ semileptonic B_s^0 decays. This analysis of B_s^0 - \bar{B}_s^0 mixing has a sensitivity of 19.4 ps^{-1} and shows an evidence of B_s^0 oscillations at $\Delta m_s \sim 17.75 \text{ ps}^{-1}$ with an amplitude significance of ~ 2 . In combination with the analyses of $\sim 8,700$ hadronic B_s^0 decays at CDF, we have made the first direct observation of time-dependent B_s^0 - \bar{B}_s^0 flavor oscillations measuring $\Delta m_s = 17.77^{+0.09}_{-0.10} \text{ (stat)} \pm 0.07 \text{ (syst)} \text{ ps}^{-1}$. The obtained value of Δm_s agrees with the Standard Model expectation. When combined with the world average values for Δm_d , $m_{\bar{B}^0}$ and $m_{\bar{B}_s^0}$, along with other theoretical input, this result yields the ratio of CKM matrix elements $|V_{td}/V_{ts}| = 0.2060 \pm 0.0007 \text{ (exp)}^{+0.0081}_{-0.0060} \text{ (theor)}$.

Acknowledgments

I would like to express my profound gratitude to my advisor Professor Manfred Paulini for his able guidance during the course of my thesis research. I have learned a lot of things from Manfred about research in general, but most importantly he has taught me how to methodically approach an analysis with persistence and diligence. I am privileged to have worked with him. I am also deeply indebted to him for always supporting me in everything I undertook during my graduate career.

My graduate career has immensely benefited from the advice that I received from Professor James Russ. He has provided invaluable suggestions to approach problems in hand, and has inspired me to always strive for excellence. I am very thankful to him for helping me learn several aspects of the CDF detector and trigger systems.

I would also like to thank Guillermo Gomez-Ceballos, with whom I worked on the semileptonic B_s^0 mixing analysis, for his enormous contributions to the mixing analyses at CDF. His dedication to work is an example for me to follow. I am also indebted for his help during the process of preparation of this thesis report.

I am thankful to my thesis committee members Professors Manfred Paulini (chair), Joseph Boudreau, Roy Briere, Richard Holman, and James Russ for reading the thesis report, and providing many useful comments for its improvement.

I am very grateful to the B_s -mixing group leaders Professors Franco Bedeschi, Matthew Herndon, Joseph Kroll, and Kevin Pitts for their leadership and numerous useful suggestions for improvements in this analysis.

The results discussed in the current dissertation have been made possible by the active collaborative efforts between several members of the B_s -mixing group. I am especially grateful to Professor Christoph Paus, Guillermo Gomez-Ceballos, Aart Heijboer, Ivan Furic, Ilya Kravchenko, Stephanie Menzemer, Konstantin Anikeev, Jonathan Piedra, Jeffrey Miles, Alberto Belloni, Nuno Leonardo, Gian Piero Di Giovanni and Bruno Casal for being wonderful collaborators. I would particularly like to thank Christoph for giving me an opportunity to work closely with him. His leadership qualities and decisiveness helped immensely in the timely completion of the projects that I undertook with him. I am also grateful to Aart Heijboer for accompanying me to Indian lunches.

I have learned about many aspects of electron identification at CDF from Barry Wicklund, who has been a constant source of interesting ideas. I am very thankful to Barry for his help in improving the results discussed in this thesis.

I would like to thank Sandro De Cecco and Masa Tanaka for their contributions to the mixing analysis. They have provided many useful suggestions and ideas used in the current measurement.

My life has immensely improved due to close association with Karen Gibson, Gavril Giurgiu and Ryuji Kusumoto. I thank them for their friendship. They have provided enormous help and support during the course of my graduate career, physics or otherwise. I will long cherish the memories of the time that we all spent together at the CMU house at Fermilab. I am also indebted to Karen and Ryuji for their help during my job search.

I would also like to thank the Fermilab staff and members of the CDF Collaboration in keeping the Tevatron and CDF experiment running. None of the results discussed in this thesis would have been possible without the tireless efforts from several collaborators at CDF and members of the Accelerator Division at Fermilab. I am extremely grateful to the United States Department of Energy for funding and supporting the particle physics research I have conducted at Fermilab.

Lastly, I thank my parents and sisters for their unconditional love and support.

Contents

1	Introduction	1
1.1	The CKM Matrix	2
1.1.1	The CKM Matrix and Wolfenstein Parameterization	3
1.1.2	Constraining the CKM Matrix	5
1.2	B Mixing Phenomenology	7
1.3	B Mixing in the Standard Model	11
1.4	Status of B Mixing Measurements	12
1.4.1	B_s^0 Mixing: Amplitude Technique	13
1.5	Overview of the Measurement	16
2	Experimental Apparatus	21
2.1	Accelerator Complex	21
2.2	The CDF-II Detector	24
2.3	CDF Coordinate System	25
2.4	Tracking Systems	28
2.4.1	Silicon Tracking Detectors	29
2.4.2	Central Outer Tracker	30
2.4.3	Pattern Recognition	32
2.5	Time of Flight	33
2.6	Calorimeters	35
2.7	Muon Systems	37
2.8	The CDF-II Trigger System	38
2.8.1	Level 1 Trigger	40
2.8.2	Level 2 Trigger	41
2.8.3	Level 3 Trigger	43
2.9	Monte Carlo Simulation	43
3	Particle Identification	45
3.1	Kaon and Pion Identification	45
3.1.1	Particle Identification Using dE/dx	45
3.1.2	Particle Identification Using TOF	48
3.1.3	Combination of dE/dx and TOF	50

3.2	Electron Reconstruction and Identification	52
3.2.1	Data Samples for Studying Electron Identification	53
3.2.2	Electron Identification Quantities	56
3.2.3	Electron Identification through Likelihood Method	66
3.3	Muon Identification	71
3.4	Sources of Leptons and Utilization of Lepton Likelihood	73
4	Signal Selection	76
4.1	Semileptonic Decays of the B Mesons	76
4.2	Online Selection: Trigger Paths	78
4.2.1	Two-Track Trigger	79
4.2.2	ℓ +SVT Trigger	83
4.3	Preparation of Charged Particle Tracks	84
4.4	Reconstruction of B Mesons	85
4.5	Signal Optimization	86
4.6	Signal Selection Requirements	87
5	Characterization of B Meson Candidates	91
5.1	Sources of Backgrounds in Semileptonic B Decays	91
5.2	Further Classification of B Meson Samples	92
5.3	Characterization of the B_s^0 Samples and Fitting Technique	94
5.3.1	Characterization in D_s^- Mass	94
5.3.2	Determination of Physics Background Fractions	99
5.3.3	Characterization in $m_{\ell+D_s^-}$	99
5.3.4	Summary of Characterization in D_s^- Mass and $m_{\ell+D_s^-}$	107
5.3.5	Decay Time Measurements in Semileptonic B_s^0 Decays	113
5.3.6	Characterization in Pseudo-Proper Decay Length ct^*	122
5.3.7	Summary of Characterization in D_s^- Mass, $m_{\ell+D_s^-}$, and ct^* : B_s^0 Lifetime Measurements	128
5.4	Characterization of the B^+ and B^0 Samples	133
5.4.1	Sample Composition of the ℓD Samples	133
5.4.2	Characterization in D Mass, $m_{\ell D}$ and ct^*	136
6	Initial State Flavor Tagging	150
6.1	Opposite Side Flavor Tags	153
6.1.1	Track-Based Jet Reconstruction	154
6.1.2	Selection of Opposite Side Tag Candidates	155
6.1.3	Electron Tagging	156
6.1.4	Muon Tagging	162
6.1.5	Jet Charge Tagging	163
6.1.6	Opposite Side Kaon Tagging	164
6.1.7	Combination of Opposite Side Tags	165
6.2	Same Side Flavor Tagging	167

6.2.1	Selection of the Same Side Tag Candidates	168
6.2.2	Same Side Tagging Algorithms	169
6.2.3	Performance of the Same Side Tagging Algorithms	171
6.2.4	Systematic Uncertainties in the Same Side Tag Dilution	172
6.2.5	Summary of Same Side Kaon Tag Performance	173
6.3	Tagging Application	175
7	Calibration of Opposite Side Flavor Tagging and Measurement of B^0-\bar{B}^0 Oscillations	176
7.1	Time Evolution PDF for B^0 - \bar{B}^0 Mixing	177
7.2	Time Evolution PDF for B^+	178
7.3	Time Evolution PDF for Backgrounds	179
7.3.1	Oscillating Backgrounds	179
7.3.2	Non-oscillating Backgrounds	180
7.4	Determination of Inputs for the OST Calibration and B^0 - \bar{B}^0 Mixing	182
7.5	Results	183
7.6	Systematic Uncertainties	186
7.6.1	Fake Lepton Background	187
7.6.2	Sample Composition of ℓD Samples	188
7.6.3	Physics Background Fractions	189
7.6.4	$\xi(t^*)$ Determination from Monte Carlo	189
7.6.5	Combinatorial Background Characterization	190
7.6.6	Fixing $\tau(B)$ to the World Average Value	190
7.6.7	σ_{t^*} Calibration Uncertainty	190
7.6.8	Uncertainty in Proper Time Scale	191
7.6.9	Summary of Systematic Uncertainties	191
8	B_s^0-\bar{B}_s^0 Oscillations	193
8.1	Time Evolution PDF for B_s^0 - \bar{B}_s^0 Mixing	193
8.2	Time Evolution PDF for Backgrounds	194
8.2.1	Oscillating Backgrounds	195
8.2.2	Non-oscillating Backgrounds	196
8.2.3	D^- Reflection Background	198
8.3	B_s^0 - \bar{B}_s^0 Mixing Analysis	199
8.4	Amplitude Method in Toy Monte Carlo	208
8.5	Systematic Uncertainties	209
8.5.1	Tagging Dilution Scale Factors	211
8.5.2	Same Side Tagging Dilution for Physics Background	211
8.5.3	D^- Reflection Background	211
8.5.4	Fake Lepton Background	212
8.5.5	Physics Background Fractions	212
8.5.6	$\xi(t^*)$ Determination from Monte Carlo	212

8.5.7	Combinatorial Background Characterization	213
8.5.8	Fixing $\tau(B_s^0)$ to the World Average Value	213
8.5.9	σ_{t^*} Calibration Uncertainty	213
8.5.10	Composition of $B_s^0 \rightarrow \ell^+ D_s^- X$ Decays	213
8.5.11	Non-Gaussian Detector Resolution	214
8.5.12	Non-vanishing Value of $\Delta\Gamma_s/\Gamma_s$	214
8.5.13	Summary of the Systematic Uncertainties	214
8.6	Amplitude Scan in Blinded Data	217
8.7	Results	218
9	Summary and Conclusions	224
9.1	Combination with Hadronic Modes	224
9.2	Conclusions	228
A	Corrections to CES $q\Delta X$ and ΔZ	231
A.1	Corrections to CES $q\Delta X$	231
A.2	Correction to CES ΔZ	232
B	Electron Identification after the First 360 pb⁻¹ of Data	237
B.1	Electron Identification in the Later Data	237
C	Amplitude Scan for Various Subsets of Data	241

List of Tables

2.1	Accelerator parameters for the Run II configuration of the Tevatron Collider as of February 2006.	24
2.2	Relevant parameters for the layout of the sensors of different SVX-II layers.	29
2.3	Pseudorapidity coverage, energy resolution and thickness for the different calorimeter subdetectors.	36
2.4	Pseudorapidity coverage, pion interaction lengths and minimum muon p_T for detection for the various muon detector devices at CDF.	37
3.1	Electron likelihood performance on isolated electrons and pions. . . .	70
3.2	Electron likelihood performance on non-isolated electrons and pions. .	71
3.3	Efficiencies of muon likelihood requirements for real and fake muons.	74
4.1	Optimized selection criteria for the $\ell^+ D_s^-$ samples.	90
4.2	Optimized selection criteria for the ℓD samples.	90
5.1	D_s^- mass fitting range for the three $\ell^+ D_s^-$ modes.	95
5.2	Fractions of combinatorial background in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates. .	97
5.3	Fractions of D^- reflection background relative to the sum of signal, physics, prompt and D^- reflection backgrounds in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates.	98
5.4	Mass ranges for defining the D_s^- sideband regions in the three $\ell^+ D_s^-$ modes.	99
5.5	Mass ranges for defining the D_s^- signal regions in the three $\ell^+ D_s^-$ modes.	99
5.6	Expected fraction of each physics background with respect to the signal and physics backgrounds in the $\ell^+ D_s^-$ modes.	100
5.7	Fractions of fake lepton background relative to the sum of signal, physics and prompt backgrounds in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates.	106
5.8	Number of true semileptonic $B_s^0 \rightarrow \ell^+ D_s^- X$ signal events in the three decay modes, after taking into account the contributions from all the backgrounds.	108
5.9	B_s^0 lifetime fit results in signal Monte Carlo samples.	122
5.10	B_s^0 lifetime fit results in various modes and subsamples for different periods of data-taking.	129

5.11	B^+ and B^0 sample composition parameters used in this measurement.	134
5.12	Table of the various B^+ and B^0 decay sequences and their contributions to the sample composition of the three general categories of decay signatures.	135
5.13	Relative efficiencies of the 24 B signal Monte Carlo samples with respect to three reference modes.	137
5.14	Fractions of combinatorial background in $B \rightarrow \ell DX$ candidates. . . .	138
5.15	Expected fraction of each physics background with respect to the signal and physics backgrounds in the ℓD modes.	139
5.16	Fractions of fake lepton background relative to the sum of signal, physics and prompt backgrounds for $B \rightarrow \ell DX$ candidates.	140
5.17	Number of true $B^{+,0} \rightarrow \ell DX$ signal events in the three decay modes.	140
6.1	Tagging performances of the various opposite side tagging algorithms in the ℓ +SVT samples.	166
6.2	Performance of the maximum p_L^{rel} algorithm in data and Monte Carlo.	173
6.3	Performance of the maximum $CLL(K)$ algorithm in data and Monte Carlo.	173
6.4	Performance of the NN combination algorithm in data and Monte Carlo.	173
7.1	Combined OST efficiencies for the combinatorial background in $B \rightarrow \ell DX$ candidates.	185
7.2	Combined OST dilutions for the combinatorial background in $B \rightarrow \ell DX$ candidates.	186
7.3	Combined opposite side tagging performance for the three data-taking periods.	187
7.4	Comparison of the combined OST performance using event by event dilution and average dilution for the three data-taking periods.	188
7.5	Systematic uncertainties for the combined opposite side tag dilution scaling factor S_D^{OST} and Δm_d	192
8.1	Dilution scale factors in different data-taking periods for the same side kaon tagging.	200
8.2	Combined OST efficiencies for combinatorial background in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates.	203
8.3	Combined OST dilutions for combinatorial background in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates.	204
8.4	SSKT efficiencies for combinatorial background in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates.	205
8.5	SSKT dilutions for combinatorial background in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates.	206
8.6	Efficiencies and dilutions for the SSKT as measured in $B^{+,0} \rightarrow \ell^+ D^- X$ samples.	207

8.7	SSKT and OST tagging efficiencies on data for B_s^0 signal candidates.	207
8.8	Summary of the systematic uncertainties on the amplitude \mathcal{A} at Δm_s = 0, 5, 10 and 15 ps^{-1}	215
8.9	Summary of the systematic uncertainties on the amplitude \mathcal{A} at Δm_s = 17, 20, 25 and 30 ps^{-1}	215

List of Figures

1.1	The first order Feynman diagram for $K^0 \rightarrow \mu^+\mu^-$ decays in Cabibbo's model prior to the GIM mechanism.	2
1.2	Additional diagram contributing to $K^0 \rightarrow \mu^+\mu^-$ decays due to the GIM mechanism.	3
1.3	The unitarity triangle in the complex plane.	5
1.4	Regions of 68% and 95% probability for the fit result of the unitarity triangle parameters $(\bar{\rho}, \bar{\eta})$, overlaid on experimental constraints, as of early 2006.	6
1.5	Lowest order Feynman diagrams contributing to B_s^0 and B^0 mixing. .	11
1.6	Status of measurements of Δm_d from various experiments.	13
1.7	Amplitude measurements and sensitivities for various experiments at $\Delta m_s = 17.5 \text{ ps}^{-1}$ as of early 2006.	14
1.8	A combination of Δm_s amplitude measurements from all the experiments as of early 2006.	15
1.9	Schematic representation of a B decay.	17
2.1	Layout of the Fermilab accelerator complex.	22
2.2	Peak luminosities for stores collided between April 2001 and May 2007, and delivered and recorded integrated luminosities at CDF to date. .	25
2.3	Conceptual depiction of particle detection at CDF.	26
2.4	The CDF-II detector	27
2.5	A diagram of the CDF tracker layout showing the different subdetector systems.	28
2.6	Silicon tracking detectors	30
2.7	Layout of wire planes on a COT endplate.	31
2.8	Layout of wires in a COT supercell.	32
2.9	Location of the Time of Flight system inside the CDF detector. . . .	34
2.10	CDF muon chamber coverage.	38
2.11	Diagram of the CDF-II trigger system.	40
2.12	Diagram of the different trigger paths at Level 1 and Level 2.	41
2.13	SVT impact parameter resolution.	42
3.1	dE/dx in the 8.5 atm $Ar - CH_4$	46

3.2	$\ln\langle dE/dx \rangle$ versus particle $\beta\gamma$ and dE/dx resolution measured for muons.	48
3.3	Measured performance of the dE/dx particle identification.	49
3.4	Expected performance of the TOF system.	50
3.5	TOF instrumental resolution	51
3.6	Separation power of the combined dE/dx and TOF particle identification.	52
3.7	$\Delta \cot \theta$ distribution for photon conversion candidates.	55
3.8	K_S^0 mass distribution for pions faking electrons.	56
3.9	Local isolation I for conversion electrons and K_S^0 pions.	57
3.10	HadE/EmE distributions for real and fake electrons.	58
3.11	EmE/ p distributions for real and fake electrons.	60
3.12	CES χ_x^2 distributions for real and fake electrons.	62
3.13	CES χ_z^2 distributions for real and fake electrons.	63
3.14	Parameterization of the width σ_x of the corrected $q\Delta X$ in the CES vs p_T .	64
3.15	CES $q\Delta X/\sigma_x$ distributions for real and fake electrons.	65
3.16	Parameterization of the width σ_z of the corrected ΔZ in the CES vs p_T .	66
3.17	CES $\Delta Z/\sigma_z$ distributions for real and fake electrons.	67
3.18	E_{CES}/p^* distributions for real and fake electrons.	68
3.19	Q_{CPR} distributions for real and fake electrons.	69
3.20	dE/dx pull distributions for electron hypothesis in pure samples of electrons and pions.	70
3.21	Empirical parameterizations of EmE/ p for locally isolated electrons and pions in the $2.0 < p_T < 4.0$ GeV/ c range.	71
3.22	Likelihood distributions for locally isolated and non-isolated electrons and pions.	72
3.23	Transverse distance between the extrapolated track and muon stub, and the hadronic energy distributions for fake and real CMU muons in different momentum ranges.	73
3.24	Muon likelihood distributions \mathcal{L}_μ for real and fake muons.	74
4.1	Schematic representation of a semileptonic $B_s^0 \rightarrow D_s^- \ell^+ X$ decay.	77
4.2	The impact parameter resolution for tracks with and without attaching L00 hits.	84
4.3	An example of signal optimization using $L_{xy}/\sigma L_{xy}(PV \rightarrow D)$ in the reconstructed ℓD^0 modes.	88
5.1	D_s^- mass distribution for the B triggers in $B \rightarrow \mu^+ D^- X$ Monte Carlo.	96
5.2	$m_{e^+ D_s^-}$ distribution for B triggers in $B_s^0 \rightarrow e^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ signal Monte Carlo events.	102
5.3	$m_{\ell^+ D_s^-}$ distributions for combinatorial and D^- reflection backgrounds in B triggers.	103

5.4	$m_{\ell^+ D_s^-}$ distribution for various sources of physics background in B triggers using Monte Carlo simulated events.	104
5.5	$m_{\ell^+ D_s^-}$ distribution for fake lepton background obtained in the $B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ candidates.	105
5.6	Fits of the D_s^- mass and $m_{\ell^+ D_s^-}$ distributions for $B \rightarrow \ell D_s(\phi\pi)$ in B triggers.	109
5.7	Fits of the D_s^- mass and $m_{\ell^+ D_s^-}$ distributions for $B \rightarrow \ell D_s(\phi\pi)$ candidates in D triggers.	109
5.8	Fits of the D_s^- mass and $m_{\ell^+ D_s^-}$ distributions for $B \rightarrow \ell D_s(\phi\pi)$ candidates in ℓ +SVT triggers.	110
5.9	Fits of the D_s^- mass and $m_{\ell^+ D_s^-}$ distributions for $B \rightarrow \ell D_s(K^* K)$ candidates in B triggers.	110
5.10	Fits of the D_s^- mass and $m_{\ell^+ D_s^-}$ distributions for $B \rightarrow \ell D_s(K^* K)$ candidates in D triggers.	111
5.11	Fits of the D_s^- mass and $m_{\ell^+ D_s^-}$ distributions for $B \rightarrow \ell D_s(K^* K)$ candidates in ℓ +SVT triggers.	111
5.12	Fits of the D_s^- mass and $m_{\ell^+ D_s^-}$ distributions for $B \rightarrow \ell D_s(3\pi)$ candidates in B triggers.	112
5.13	Fits of the D_s^- mass and $m_{\ell^+ D_s^-}$ distributions for $B \rightarrow \ell D_s(3\pi)$ candidates in D triggers.	112
5.14	Fits of the D_s^- mass and $m_{\ell^+ D_s^-}$ distributions for $B \rightarrow \ell D_s(3\pi)$ candidates in ℓ +SVT triggers.	113
5.15	Proper time distribution measured in the D^- +track calibration sample and variation of the resolution scale factor with the z -coordinate of the reconstructed D^- +track vertex.	115
5.16	κ -factor distributions for $B_s^0 \rightarrow \ell^+ D_s^- X$ decays obtained from B_s^0 signal Monte Carlo samples.	118
5.17	k -factor distributions and the corresponding proper decay time uncertainties in several $m_{\ell^+ D_s^-}$ ranges for the $B_s^0 \rightarrow \ell^+ D_s^-$, $D_s^- \rightarrow \phi \pi^-$ decays.	119
5.18	$\xi(t^*)$ distributions for $B_s^0 \rightarrow \ell^+ D_s^- X$ decays obtained from Monte Carlo.	121
5.19	Fits in the ct^* space for $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ signal Monte Carlo events.	123
5.20	Pseudo-proper decay time distributions for combinatorial and D^- reflection backgrounds in B triggers.	124
5.21	Pseudo-proper decay time distribution for fake lepton background in the $B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ candidates.	125
5.22	Mean and RMS values of the ct^* resolution as a function of $m_{\ell D}$ in ℓD modes for background subtracted signal and sidebands.	127
5.23	σ_{ct^*} distributions for signal and combinatorial background in B triggers in $\ell^+ D_s^-$ modes.	128
5.24	Fits to the ct^* distributions for $B \rightarrow \ell D_s(\phi\pi)$ candidates.	130
5.25	Fits to the ct^* distributions for $B \rightarrow \ell D_s(K^* K)$ candidates.	131

5.26	Fits to the ct^* distributions for $B \rightarrow \ell D_s(3\pi)$ candidates.	132
5.27	A schematic representation of all possible $B \rightarrow \ell D^{(*)} X$ transitions resulting in a ℓD final state.	134
5.28	D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ \bar{D}^0 X$ candidates in the B triggers.	141
5.29	D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ \bar{D}^0 X$ candidates in the D triggers.	142
5.30	D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ \bar{D}^0 X$ candidates in the $l + SVT$ triggers.	143
5.31	D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^{*-} X$ candidates in B triggers.	144
5.32	D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^{*-} X$ candidates in D triggers.	145
5.33	D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^{*-} X$ candidates in $l + SVT$ triggers.	146
5.34	D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^- X$ candidates in B triggers.	147
5.35	D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^- X$ candidates in D triggers.	148
5.36	D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^- X$ candidates in $l + SVT$ triggers.	149
6.1	Sketch of typical $b\bar{b}$ event indicating several B flavor tagging techniques.	152
6.2	The definition of signed impact parameter of the SVT track and invariant mass distribution of the $\ell + SVT$ pairs after background subtraction using the signed SVT impact parameter.	154
6.3	Cartoons depicting signed impact parameters for electrons originating from conversions and B decays.	158
6.4	Distribution of d_0^{sign} for identified conversions.	159
6.5	Tagging efficiencies and dilutions in likelihood bins separately for $d_0^{sign} > 0$ and $d_0^{sign} < 0$ for the opposite side electron tags.	160
6.6	Variation of dilution of the electron tags with p_T^{rel} for $\mathcal{L}_e > 0.85$ in the $e + SVT$ data.	161
6.7	Variation of muon tag dilution with muon likelihood and p_T^{rel}	162
6.8	Jet charge Q_{jet} for Class 1 jets for events tagged as B and \bar{B} , and the dependence of dilution on $ Q_{jet} \cdot \mathcal{P}_{nn}$ for Class 1 jets.	164
6.9	The dependence of OSKT tagging dilution on the fraction of jet momentum carried by the SECVTX jet.	165
6.10	Linear dependence of tagging dilution from the combined OST versus output of the neural network.	166
6.11	Charge correlation of B mesons with fragmentation tracks.	167
6.12	Illustration of the construction of the kinematic quantities p_T^{rel} and p_L^{rel} .	169

6.13	Data and Monte Carlo comparison for several SST related quantities for fully reconstructed \bar{B}_s^0 decays.	170
6.14	Dilution dependence for the three SSKT algorithms.	172
6.15	Comparisons between SSKT candidates in data and Monte Carlo for the tagging track multiplicity around B_s^0 mesons, and $CLL(K)$ values for tagging tracks.	174
7.1	The combined OST \mathcal{D} normalized distributions for signal and combinatorial background in the B and D triggers in $\ell^+\bar{D}^0$ modes.	184
7.2	Time-dependent mixing asymmetry using the combined OST algorithm as applied to semileptonic B^0 and B^+ samples.	187
8.1	Dilution PDFs for the SSKT and combined OST algorithms for B triggers in the ℓD_s channels.	202
8.2	Example of an amplitude scan on a toy Monte Carlo simulating the data generated with $\Delta m_s = 17.75 \text{ ps}^{-1}$	209
8.3	Parameterization of the overall systematic uncertainty versus Δm_s	216
8.4	Blinded amplitude scan in data for the $B_s^0 \rightarrow \ell^+ D_s^- X$ decays.	217
8.5	Unblinded amplitude scan in data for the $B_s^0 \rightarrow \ell^+ D_s^- X$ decays.	219
8.6	Combined overall error including the systematic and statistical errors versus Δm_s	220
8.7	Likelihood profile as a function of Δm_s from the unblinded data for $B_s^0 \rightarrow \ell^+ D_s^- X$ decays.	222
8.8	Error distributions for 1300 toy Monte Carlo experiments simulating the real data generated with $\Delta m_s = 17.75 \text{ ps}^{-1}$	223
9.1	The combined amplitude scan for the hadronic modes.	226
9.2	The combination of the semileptonic and hadronic amplitude scans.	226
9.3	Λ as a function of Δm_s in the range 0-35 ps^{-1} , and in the region near the minimum.	227
9.4	Distribution of the minimal likelihood ratio Λ in scans with randomized tag decisions.	227
9.5	Status of constraints on the unitarity triangle in (ρ', η') space as of early 2006, and after including the CDF measurement of Δm_s as of late 2006.	229
A.1	$\langle q\Delta X \rangle$ with respect to p_T and η	231
A.2	Parameterization of η dependence of $\langle q\Delta X \rangle$ in three bins of p_T	233
A.3	$\langle q\Delta X \rangle$ with respect to p_T and η after corrections.	234
A.4	$\langle \Delta Z \rangle$ with respect to p_T and η before corrections.	234
A.5	Parameterization of η dependence of $\langle \Delta Z \rangle$ in three p_T bins.	235
A.6	$\langle \Delta Z \rangle$ with respect to p_T and η after corrections.	236
B.1	$\langle Q_{CPR} \rangle$ for conversion electrons in CPR2 before applying corrections.	238

B.2	$\langle Q_{CPR} \rangle$ for conversion electrons in CPR2 after applying corrections. .	239
B.3	Q_{CPR} distributions for electrons and pions after applying corrections.	239
B.4	$\langle q\Delta X \rangle$ for conversion electrons for the old and the later data.	240
B.5	$\langle \Delta Z \rangle$ for conversion electrons for the old and the later data.	240
C.1	Unblinded amplitude scan in Period 1 data for the $B_s \rightarrow \ell^+ D_s^- X$ decays.	241
C.2	Unblinded amplitude scan in Period 2 data for the $B_s \rightarrow \ell^+ D_s^- X$ decays.	242
C.3	Unblinded amplitude scan in Period 3 data for the $B_s \rightarrow \ell^+ D_s^- X$ decays.	242
C.4	Unblinded amplitude scan in data for the $B_s \rightarrow \ell^+ D_s^- X$ decays using same side tags alone.	243
C.5	Unblinded amplitude scan in data for the $B_s \rightarrow \ell^+ D_s^- X$ decays using opposite side tags alone.	243

Chapter 1

Introduction

B mesons are bound states of an anti- b quark and a lighter quark $q = u, d, s, c$. They are usually denoted by B_q or as $|\bar{b}q\rangle$. Neutral B mesons (consisting of a bottom antiquark and either a down or a strange quark) oscillate between their particle and antiparticle state due to flavor-changing interactions. In the framework of the Standard Model, the properties of flavor-changing decays are described by the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix [1]. Since the first observation of particle-antiparticle transformations in neutral B mesons in 1987 [2, 3], the determination of the B_s^0 - \bar{B}_s^0 oscillation frequency Δm_s via a time-dependent measurement of B_s^0 - \bar{B}_s^0 oscillations has been a major objective of experimental particle physics. This frequency can be used to extract the magnitude of one of the nine elements of the CKM matrix.

It is not always straightforward to interpret experimental measurements in terms of the CKM matrix elements. B decays provide access to five out of the nine CKM matrix elements. Currently, there are four measurements which provide clean inputs into the determination of the CKM matrix elements: the relative decay rate of $B \rightarrow \pi l \nu$ to $B \rightarrow D l \nu$, CP violation in the kaon and the B^0 meson systems, and the B^0 oscillation frequency Δm_d . Due to large theoretical uncertainties from non-perturbative QCD calculations, precise constraints on the CKM matrix elements from Δm_d are currently not possible. However, a measurement of the B_s^0 mixing frequency Δm_s together with Δm_d can be used to construct a ratio of frequencies, thereby canceling common theoretical uncertainties. Thus, together with existing constraints, the ratio of oscillation frequencies can be used to over-constrain the CKM matrix.

This thesis presents a measurement of the B_s^0 - \bar{B}_s^0 oscillation frequency Δm_s using semileptonic decays of the B_s^0 mesons in $B_s^0 \rightarrow \ell^+ D_s^- X$ modes. A measurement of the B^0 - \bar{B}^0 mixing frequency Δm_d using $B \rightarrow \ell D X$ decays is also accomplished in an effort to calibrate inputs for the Δm_s measurement. In this chapter, we will review the development of the CKM matrix, constraining the CKM matrix parameters using input from experiments, discuss a quantum mechanical treatment of neutral B - \bar{B} oscillations, and neutral B - \bar{B} mixing in the Standard Model together with a

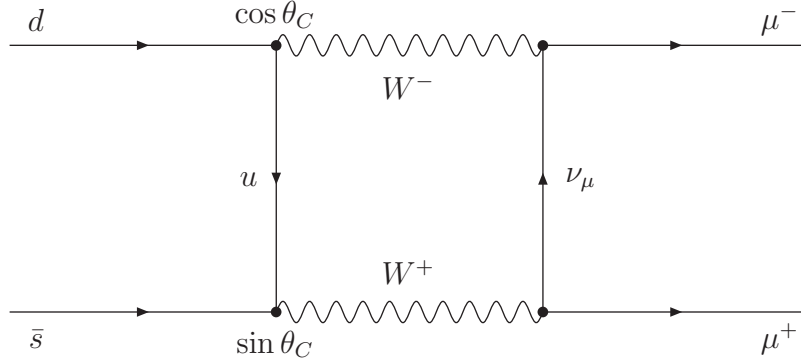


Figure 1.1: The first order Feynman diagram for $K^0 \rightarrow \mu^+ \mu^-$ decays in Cabibbo's model prior to the GIM mechanism.

history of B - \bar{B} mixing measurements. Finally, we conclude with an overview of this measurement.

1.1 The CKM Matrix

Flavor-changing decays in the Standard Model are described via weak currents. Before the quarks were postulated and discovered, Cabibbo [4] examined weak hadronic currents in $\Delta S = 0$ (decays like $n \rightarrow pe\bar{\nu}_e$) and $\Delta S = 1$ (decays such as $K^+ \rightarrow \mu^+ \nu_\mu$) transitions. To explain the experimental data, he proposed a rotation between the $\Delta S = 0$ current and the $\Delta S = 1$ current, by an angle θ_c . In Cabibbo's model, the u quark is weakly coupled to a superposition of d and s quarks, $|d'\rangle = |d\rangle \cos \theta_C + |s\rangle \sin \theta_C$. Thus, the decay rates of $s \rightarrow u$ transitions could now be expressed in terms of $\sin \theta_C$, while $d \rightarrow u$ decay widths can be described by $\cos \theta_C$. However, not all predictions were correct in Cabibbo's model; the $K^0 \rightarrow \mu^+ \mu^-$ decay rate was substantially overestimated as compared to the experimental measurement. Figure 1.1 shows the Feynman diagram that contributed to this decay in the Cabibbo model.

In 1970, Glashow, Iliopoulos, and Maiani [5] came to the rescue of Cabibbo's model. They proposed an extension to it by postulating the existence of a fourth quark, the charm (c) quark. The c quark was proposed to couple to a superposition of s and d quarks. A "mixing matrix" was introduced that would rotate the mass eigenstates d, s into the weak eigenstates d', s' which coupled to the u and c quarks:

$$\begin{bmatrix} d' \\ s' \end{bmatrix} = \begin{bmatrix} \cos \theta_C & \sin \theta_C \\ -\sin \theta_C & \cos \theta_C \end{bmatrix} \begin{bmatrix} d \\ s \end{bmatrix}. \quad (1.1)$$

This led to the introduction of a second Feynman diagram for the $K^0 \rightarrow \mu^+ \mu^-$ decay, which is shown in Figure 1.2. The diagrams would cancel exactly if the c and u quarks had the same mass. Because of the difference in mass of c and u quarks, there is not a perfect cancellation. However, the new diagram still suppresses the decay through what is called the GIM mechanism.

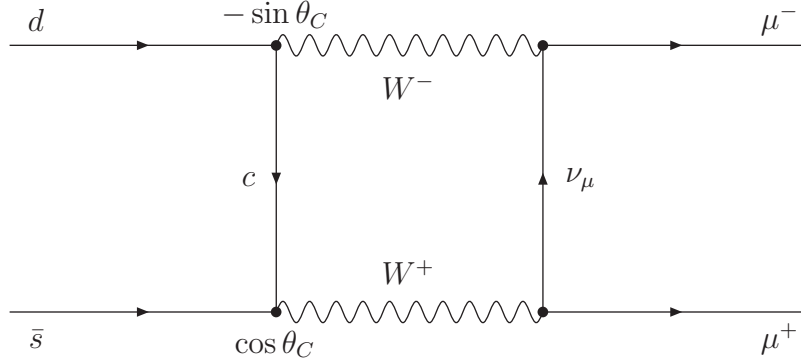


Figure 1.2: Additional diagram contributing to $K^0 \rightarrow \mu^+ \mu^-$ decays due to the GIM mechanism.

The GIM mechanism was confirmed by the discovery of J/ψ meson, a $c\bar{c}$ state, in 1974 [6, 7]. Before that, in 1973, Kobayashi and Maskawa [1] proposed a further extension of the quark model from four to six quarks by adding a third generation of quarks (the top and bottom quarks). They generalized the GIM mixing matrix to be the most general unitary transformation from the flavor eigenstates of the down-type quarks (d, s, b) to their weak eigenstates (d', s', b') . The motivation was to account for the CP violation observed by Cronin and Fitch in 1964, in the decays of the K^0 meson [8]. With three generations, Kobayashi and Maskawa demonstrated the existence of a complex phase in the unitary 3×3 mixing matrix for the six quark model, thus accommodating CP violation. Subsequently, at Fermilab, the bottom quark was discovered in 1977 [9] and the top quark was discovered in 1993 [10] confirming the existence of three quark generations.

1.1.1 The CKM Matrix and Wolfenstein Parameterization

The unitary matrix which rotates the electroweak eigenstates into the mass eigenstates is known as the CKM matrix, shown in Equation (1.2).

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (1.2)$$

with

$$VV^\dagger = 1, \quad (1.3)$$

or equivalently

$$\sum_j V_{ij} V_{kj}^* = \delta_{ik}. \quad (1.4)$$

The CKM matrix can be parameterized by three mixing angles θ_{12} , θ_{23} , and θ_{13} , and a CP -violating phase δ . A standard choice of the parameterization is [11]

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (1.5)$$

where $s_{ij} = \sin \theta_{ij}$, $c_{ij} = \cos \theta_{ij}$. Several experimental and phenomenological results have demonstrated that s_{13} and s_{23} are small and of $\mathcal{O}(10^{-3})$ and $\mathcal{O}(10^{-2})$, respectively. Consequently, $c_{23} \simeq c_{13} \simeq 1$ is a good approximation, and therefore, the four independent parameters are given by

$$s_{12} = |V_{us}|, \quad s_{23} = |V_{cb}|, \quad s_{13} = |V_{ub}|, \quad \text{and} \quad \delta. \quad (1.6)$$

The Wolfenstein parameterization [12] is another approximation of the CKM matrix, in which each element of the matrix is expanded as a power series in $\lambda = \sin \theta_C = s_{12} = |V_{us}| \approx 0.22$ [13],

$$V = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (1.7)$$

The four independent parameters that define the matrix are now given by

$$\lambda = \sin \theta_C, \quad A, \quad \rho, \quad \eta. \quad (1.8)$$

This parameterization is accurate to third order in λ , and illustrates the observed hierarchy in the magnitude of the CKM matrix elements. The diagonal terms are of order unity, while the off-diagonal elements are smaller by factors of λ . Thus, we observe from this parameterization that cross-generational weak decays are CKM suppressed by factors of λ .

In the unitarity condition in Equation (1.4), the off-diagonal conditions ($i \neq k$) can be represented as triangles in the complex plane. Each of the three terms is a vector in the complex plane, and the sum of these vectors is zero. If the CKM matrix elements had all been real numbers, this triangle would lie entirely on the real axis. The greater the CP violating effects, the larger is the area of each triangle.

It is easier to measure the area of a triangle when all of the angles are roughly the same size. The triangle defined by the unitarity condition

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \quad (1.9)$$

which can be written to leading order in λ as

$$A\lambda^3(\rho + i\eta) - A\lambda^3 + A\lambda^3(1 - \rho - i\eta) = 0, \quad (1.10)$$

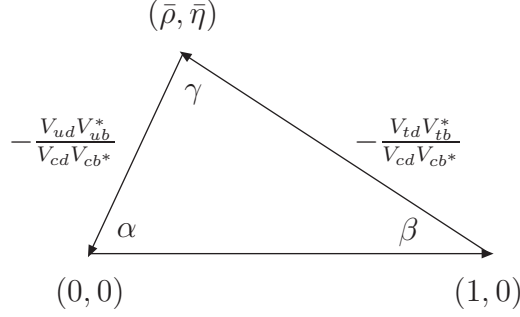


Figure 1.3: The unitarity triangle in the complex plane.

is roughly equilateral, with each side proportional to λ^3 . The triangle condition expressed in Equation (1.9) is called the unitarity triangle. Equation (1.9) is divided by $V_{cd}V_{cb}^*$, so that the length of one side of the triangle becomes unity, and is aligned with the real axis as displayed in Figure 1.3. The coordinates of the triangle in the Wolfenstein parameterization are $(0,0)$, $(1,0)$ and $(\bar{\rho}, \bar{\eta}) = (1 - \lambda^2/2)(\rho, \eta)$. The three angles are given by

$$\alpha = \arg \left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right], \quad \beta = \arg \left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right], \quad \gamma = \arg \left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right]. \quad (1.11)$$

A quantity of special interest in the current measurement is the length of the side between the vertices at $(1,0)$ and $(\bar{\rho}, \bar{\eta})$, expressed as

$$R_t \equiv \left| \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} \right| = \sqrt{(1 - \bar{\rho})^2 + \bar{\eta}^2} = \frac{1}{\lambda} \left| \frac{V_{td}}{V_{ts}} \right|. \quad (1.12)$$

As described in Section 1.3, the ratio $\left| \frac{V_{td}}{V_{ts}} \right|$ can be determined from the measurements of the B^0 and B_s^0 oscillation frequencies.

If there had been more generations of quarks, Equation (1.9) would have more terms (one for each new generation), and the associated figure in the complex plane would be a polygon with the number of sides equal to the number of generations. If it is experimentally found that the unitarity triangle is not closed, this would be evidence for physics beyond the Standard Model. Thus, one of the priorities of experimental particle physics is to over-constrain the unitarity triangle by determining all of its sides and angles, to deduce if the triangle is closed or not.

1.1.2 Constraining the CKM Matrix

As mentioned before, not all flavor changing meson decays contribute to the determination of CKM matrix elements. The reason for this is that in many decays, there are considerable theoretical uncertainties involved in extracting CKM matrix elements from the measured partial decay widths.

Figure 1.4 from Ref. [14] shows a projection of current constraints on the CKM matrix parameters and regions of 68% and 95% probability for the fit result, as of

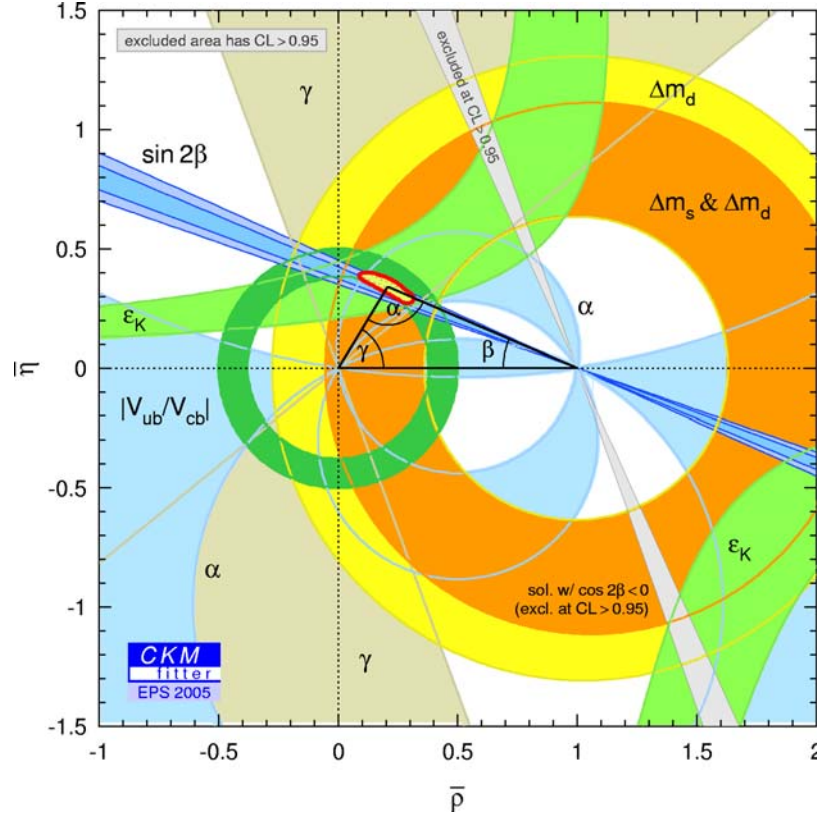


Figure 1.4: Regions of 68% and 95% probability for the fit result of the unitarity triangle parameters $(\bar{\rho}, \bar{\eta})$, overlaid on experimental constraints, as of early 2006.

the beginning of the year 2006. The following five measurements furnish the cleanest constraints on the values of the parameters ρ, η of the unitarity triangle, and are used in the fit. In Figure 1.4, these are displayed as the $|V_{ub}/V_{cb}|$, ϵ_K , $\sin(2\beta)$, Δm_d , and the $\Delta m_d/\Delta m_s$ constraints, described as follows:

1. The measurement of the relative rate of $b \rightarrow ul\nu$ to $b \rightarrow cl\nu$ decays determines the $|V_{ub}/V_{cb}|$ constraint:

$$\sqrt{\bar{\rho}^2 + \bar{\eta}^2} = \left(1 - \frac{\lambda^2}{2}\right) \frac{1}{\lambda} \left| \frac{V_{ub}}{V_{cb}} \right|. \quad (1.13)$$

The ratio constrains the length of the side between $(0,0)$ and $(\bar{\rho}, \bar{\eta})$. The best results for this measurement are obtained from the CLEO [15], BaBar [16] and Belle [17] experiments.

2. The variable ϵ_K is defined as:

$$\epsilon_K = \frac{A(K_L \rightarrow (\pi\pi)_{I=0})}{A(K_S \rightarrow (\pi\pi)_{I=0})}, \quad (1.14)$$

where I denotes the isospin. It is a measure of the CP violation asymmetry in the kaon system, and translates roughly into a hyperbolic constraint on the unitarity triangle [14].

3. The measurement of $\sin(2\beta)$ constrains the angle β of the unitarity triangle. These measurements are dominated by the B -factories [18, 19].
4. The length of the side between $(0,0)$ and $(\bar{\rho}, \bar{\eta})$ corresponds to constraints imposed by Δm_d and limits on Δm_s . The B^0 oscillation frequency Δm_d is measured precisely by the B -factories [13], but currently only a lower limit exists on $\Delta m_s > 14.4 \text{ ps}^{-1}$ [13].

1.2 B Mixing Phenomenology

To study a neutral B meson system, let the neutral B meson state be denoted by $|B_q^0\rangle$, where $q = s, d$. We will omit the q index since the following discussion is applicable to both B^0 and B_s^0 mesons. b quarks are produced via strong or electromagnetic interactions, and hence the neutral B meson states originate in one of the two flavor eigenstates, $|B^0\rangle$ and $|\bar{B}^0\rangle$. Due to mixing, an initially pure $|B^0\rangle$ or $|\bar{B}^0\rangle$ state evolves as a mixture of the two flavor eigenstates. The flavor of the state is observed at a later time t usually through its decay. If $|B^0\rangle(t)$ ($|\bar{B}^0\rangle(t)$) denotes the state vector of $|B^0\rangle$ ($|\bar{B}^0\rangle$) at t , then the time evolution of these states is governed by a Schrödinger equation:

$$i \frac{d}{dt} \begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix} = \left(\mathbf{M} - i \frac{\mathbf{\Gamma}}{2} \right) \begin{pmatrix} |B^0(t)\rangle \\ |\bar{B}^0(t)\rangle \end{pmatrix}, \quad (1.15)$$

where the “mass matrix” \mathbf{M} and the “decay matrix” $\mathbf{\Gamma}$ are t -independent, Hermitian 2×2 matrices. CPT invariance guarantees equality of mass and lifetime of particles and antiparticles, and therefore

$$M_{11} = M_{22} \equiv m, \quad \Gamma_{11} = \Gamma_{22} \equiv \Gamma. \quad (1.16)$$

The off-diagonal terms in $\mathbf{M} - i \mathbf{\Gamma}/2$ can induce transitions between the charge-conjugate states. The mass eigenstates are defined as the eigenvectors of $\mathbf{M} - i \mathbf{\Gamma}/2$, and can be expressed in terms of the flavor eigenstates as

$$|B_{L,H}^0\rangle = p |B^0\rangle \pm q |\bar{B}^0\rangle, \quad (1.17)$$

with $|p|^2 + |q|^2 = 1$, and L and H refer to the “light” and “heavy” mass eigenstates, respectively. The corresponding eigenvalues are given by

$$\lambda_{L,H} = \left(m - \frac{i}{2} \Gamma \right) \pm \frac{q}{p} \left(M_{12} - \frac{i}{2} \Gamma_{12} \right), \quad (1.18)$$

together with

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \quad (1.19)$$

Expressing the two eigenvalues as $M_L - i\Gamma_L/2$ and $M_H - i\Gamma_H/2$, we obtain

$$M_{L,H} = \text{Re}(\lambda_{L,H}), \quad \text{and} \quad \Gamma_{L,H} = -2\text{Im}(\lambda_{L,H}). \quad (1.20)$$

Therefore, the time evolution of the mass eigenstates is given by

$$|B_{L,H}^0(t)\rangle = e^{-(iM_{L,H} + \frac{1}{2}\Gamma_{L,H})t} |B_{L,H}^0\rangle = e^{-i\lambda_{L,H}t} |B_{L,H}^0\rangle. \quad (1.21)$$

It is thus concluded that the average mass of the eigenstates is m and the average width is Γ , i.e.

$$\frac{M_H + M_L}{2} = m, \quad \text{and} \quad \frac{\Gamma_L + \Gamma_H}{2} = \Gamma. \quad (1.22)$$

We further define two more observables, the mass and width differences of the mass eigenstates as Δm and $\Delta\Gamma$:

$$\Delta m \equiv M_H - M_L, \quad \text{and} \quad \Delta\Gamma \equiv \Gamma_L - \Gamma_H. \quad (1.23)$$

Inverting Equation (1.17) to express the flavor eigenstates $|B^0\rangle$ and $|\bar{B}^0\rangle$ in terms of the mass eigenstates, and using Equation (1.21), the time evolution of the flavor eigenstates produced at $t = 0$ is given by:

$$\begin{aligned} |B^0(t)\rangle &= \frac{1}{2p} \left[e^{-(iM_L + \frac{\Gamma_L}{2})t} |B_L^0\rangle + e^{-(iM_H + \frac{\Gamma_H}{2})t} |B_H^0\rangle \right], \\ |\bar{B}^0(t)\rangle &= \frac{1}{2q} \left[e^{-(iM_L + \frac{\Gamma_L}{2})t} |B_L^0\rangle - e^{-(iM_H + \frac{\Gamma_H}{2})t} |B_H^0\rangle \right]. \end{aligned} \quad (1.24)$$

Using Equation (1.17), we finally obtain the following expression for the time evolution of the pure flavor eigenstates:

$$\begin{aligned} |B^0(t)\rangle &= g_+(t) |B^0\rangle + \frac{q}{p} g_-(t) |\bar{B}^0\rangle, \\ |\bar{B}^0(t)\rangle &= \frac{p}{q} g_-(t) |B^0\rangle + g_+(t) |\bar{B}^0\rangle, \end{aligned} \quad (1.25)$$

where

$$g_{\pm}(t) = \frac{1}{2} (e^{-i\lambda_L t} \pm e^{-i\lambda_H t}). \quad (1.26)$$

The normalizations of these states are given by,

$$\begin{aligned} \eta^2 &= \int_0^\infty \langle B^0(t) | B^0(t) \rangle dt = \frac{\Gamma}{2} \left[\frac{1 + |q/p|^2}{\Gamma^2 - \Delta\Gamma^2/4} + \frac{1 - |q/p|^2}{\Gamma^2 + \Delta m^2} \right], \\ \bar{\eta}^2 &= \int_0^\infty \langle \bar{B}^0(t) | \bar{B}^0(t) \rangle dt = \frac{\Gamma}{2} \left| \frac{p}{q} \right|^2 \left[\frac{1 + |q/p|^2}{\Gamma^2 - \Delta\Gamma^2/4} - \frac{1 - |q/p|^2}{\Gamma^2 + \Delta m^2} \right]. \end{aligned} \quad (1.27)$$

Using the definitions of $g_{\pm}(t)$ from Equation (1.26), we have

$$|g_{\pm}(t)|^2 = \frac{1}{2}e^{-\Gamma t} \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) \pm \cos(\Delta m t) \right]. \quad (1.28)$$

The probabilities of observing a flavor change at time t are thus given by:

$$\begin{aligned} \mathcal{P}_{\text{mix}}(t)(B^0 \rightarrow \bar{B}^0) &= \mathcal{P}_{\text{mix}}^{B^0}(t) = \frac{1}{\bar{\eta}^2} |\langle B^0 | \bar{B}^0(t) \rangle|^2 = \frac{1}{\bar{\eta}^2} \left| \frac{p}{q} \right|^2 |g_{-}(t)|^2 \\ &= \frac{2}{\Gamma} \left[\frac{1 + |q/p|^2}{\Gamma^2 - \Delta\Gamma^2/4} - \frac{1 - |q/p|^2}{\Gamma^2 + \Delta m^2} \right]^{-1} \\ &\quad \times \frac{1}{2} e^{-\Gamma t} \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) - \cos(\Delta m t) \right], \end{aligned} \quad (1.29)$$

$$\begin{aligned} \mathcal{P}_{\text{mix}}^{\bar{B}^0}(t)(\bar{B}^0 \rightarrow B^0) &= \mathcal{P}_{\text{mix}}^{\bar{B}^0}(t) = \frac{1}{\eta^2} |\langle \bar{B}^0 | B^0(t) \rangle|^2 = \frac{1}{\eta^2} \left| \frac{q}{p} \right|^2 |g_{-}(t)|^2 \\ &= \frac{2}{\Gamma} \left[\frac{1 + |q/p|^2}{\Gamma^2 - \Delta\Gamma^2/4} + \frac{1 - |q/p|^2}{\Gamma^2 + \Delta m^2} \right]^{-1} \left| \frac{q}{p} \right|^2 \\ &\quad \times \frac{1}{2} e^{-\Gamma t} \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) - \cos(\Delta m t) \right], \end{aligned} \quad (1.30)$$

while the probabilities that the initial states stay unchanged at time t are similarly obtained as:

$$\begin{aligned} \mathcal{P}_{\text{unmix}}(t)(B^0 \rightarrow B^0) &= \mathcal{P}_{\text{unmix}}^{B^0}(t) = \frac{1}{\eta^2} |\langle B^0 | B^0(t) \rangle|^2 = \frac{1}{\eta^2} |g_{+}(t)|^2 \\ &= \frac{2}{\Gamma} \left[\frac{1 + |q/p|^2}{\Gamma^2 - \Delta\Gamma^2/4} + \frac{1 - |q/p|^2}{\Gamma^2 + \Delta m^2} \right]^{-1} \\ &\quad \times \frac{1}{2} e^{-\Gamma t} \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + \cos(\Delta m t) \right], \end{aligned} \quad (1.31)$$

$$\begin{aligned} \mathcal{P}_{\text{unmix}}(t)(\bar{B}^0 \rightarrow \bar{B}^0) &= \mathcal{P}_{\text{unmix}}^{\bar{B}^0}(t) = \frac{1}{\bar{\eta}^2} |\langle \bar{B}^0 | \bar{B}^0(t) \rangle|^2 = \frac{1}{\bar{\eta}^2} |g_{+}(t)|^2 \\ &= \frac{2}{\Gamma} \left[\frac{1 + |q/p|^2}{\Gamma^2 - \Delta\Gamma^2/4} - \frac{1 - |q/p|^2}{\Gamma^2 + \Delta m^2} \right]^{-1} \left| \frac{q}{p} \right|^2 \\ &\quad \times \frac{1}{2} e^{-\Gamma t} \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + \cos(\Delta m t) \right]. \end{aligned} \quad (1.32)$$

It is observed that the obtained expressions are not symmetric between the $B^0(t)$ and $\bar{B}^0(t)$ states.

Two interesting limiting cases of the mixing and no-mixing probabilities obtained in Equations (1.29) and (1.31) can be found by neglecting CP violation in the mixing, or neglecting the lifetime difference between the light and heavy states $\Delta\Gamma$. We discuss these two cases in the following.

The heavy and light states in Equation (1.17) can be re-expressed as:

$$\begin{aligned} |B_{L,H}^0\rangle &= \frac{p+q}{2} \left[(|B^0\rangle + |\bar{B}^0\rangle) \pm \frac{1-q/p}{1+q/p} (|B^0\rangle - |\bar{B}^0\rangle) \right] \\ &= \frac{p+q}{2} \left[(|B^0\rangle + |\bar{B}^0\rangle) \pm \epsilon_B (|B^0\rangle - |\bar{B}^0\rangle) \right], \end{aligned} \quad (1.33)$$

where $\epsilon_B = \frac{1-q/p}{1+q/p}$ corresponds to the amount by which the $|B_L^0\rangle$ and $|B_H^0\rangle$ differ from the CP eigenstates. ϵ_B is expected to be very small in the Standard Model, $\mathcal{O}(10^{-3})$ for the B^0 - \bar{B}^0 and B_s^0 - \bar{B}_s^0 systems [13]. The limit of no CP violation in mixing is thus

$$\frac{q}{p} = 1. \quad (1.34)$$

In this limit, the mixing and no-mixing probabilities are given by

$$\mathcal{P}_{\text{mix}}^{B^0}(t) = \mathcal{P}_{\text{mix}}^{\bar{B}^0}(t) = \frac{\Gamma}{2} e^{-\Gamma t} \left(1 - \frac{\Delta\Gamma^2}{4\Gamma^2} \right) \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) - \cos(\Delta m t) \right], \quad (1.35)$$

$$\mathcal{P}_{\text{unmix}}^{B^0}(t) = \mathcal{P}_{\text{unmix}}^{\bar{B}^0}(t) = \frac{\Gamma}{2} e^{-\Gamma t} \left(1 - \frac{\Delta\Gamma^2}{4\Gamma^2} \right) \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + \cos(\Delta m t) \right]. \quad (1.36)$$

Thus, the symmetry between the time evolution of B^0 and \bar{B}^0 is restored in the case of no CP violation. This formulation is suitable for B_s^0 mesons, which are not expected to be affected by large CP violation effects.

The other limiting case arises by considering $\Delta\Gamma$ to be negligible. The $\Delta\Gamma$ for B^0 mesons is expected to be very small in the Standard Model $\Delta\Gamma_d/\Gamma_d < 1\%$, while it could be large for the B_s^0 system $\Delta\Gamma_s/\Gamma_s \sim 10\%$ [13]. First, we express the off-diagonal matrix elements M_{12} and Γ_{12} in terms of Δm and $\Delta\Gamma$

$$(\Delta m)^2 - \frac{1}{4}(\Delta\Gamma)^2 = 4|M_{12}|^2 - |\Gamma_{12}|^2, \quad (1.37)$$

$$\Delta m \Delta\Gamma = -4 \text{Re}(M_{12}\Gamma_{12}^*), \quad (1.38)$$

Therefore, from $\Delta\Gamma = 0 \Rightarrow \Gamma_{12} = 0$, and we obtain using Equation (1.19)

$$\frac{q}{p} = \sqrt{\frac{M_{12}^*}{M_{12}}} = e^{i\phi}. \quad (1.39)$$

Thus, $|q/p| = 1$ in the case of setting $\Delta\Gamma = 0$, and we finally obtain the mixed and unmixed probabilities at time t

$$\mathcal{P}_{\text{mix}}^{B^0}(t) = \mathcal{P}_{\text{mix}}^{\bar{B}^0}(t) = \frac{\Gamma}{2} e^{-\Gamma t} [1 - \cos(\Delta m t)], \quad (1.40)$$

$$\mathcal{P}_{\text{unmix}}^{B^0}(t) = \mathcal{P}_{\text{unmix}}^{\bar{B}^0}(t) = \frac{\Gamma}{2} e^{-\Gamma t} [1 + \cos(\Delta m t)]. \quad (1.41)$$

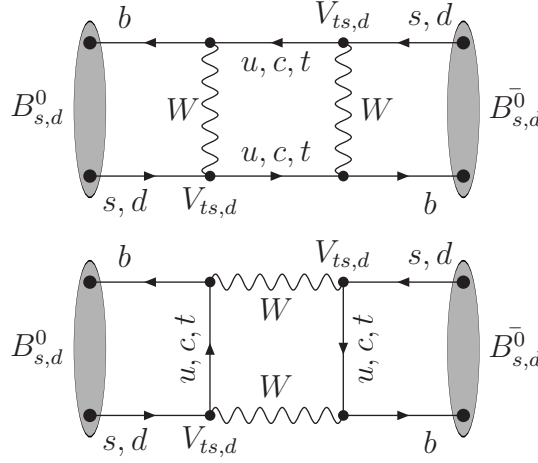


Figure 1.5: Lowest order Feynman diagrams contributing to B_s^0 and B^0 mixing.

The mixed and unmixed probabilities of the B^0 and \bar{B}^0 states are equal. This compact formulation is particularly suitable for the B^0 meson case, where a large CP violating phase ϕ is possible but does not affect the mixed and unmixed probabilities.

1.3 B Mixing in the Standard Model

In the Standard Model, $B_q^0 - \bar{B}_q^0$ transitions between the two flavor eigenstates $|B_q^0\rangle$ and $|\bar{B}_q^0\rangle$ occur by the fourth order flavor-changing weak interactions described, at the lowest order, by the box diagrams in Figure 1.5, involving two W bosons and two up-type quarks. The up-type quarks could be u , c or t quarks. However, diagrams involving exchange of top quarks dominate the interactions since the interaction amplitude is proportional to m_q^2/m_W^2 , where m_q is the quark mass and m_W^2 is the mass of the W boson. In the approximation that only the diagrams involving the top quarks contribute, a theoretical calculation [20] of the dispersive (mass) part of the box diagrams yields the following predictions for the off-diagonal elements M_{12} and Γ_{12} :

$$|M_{12}| = \frac{G_F^2 m_W^2 m_{B_q} f_{B_q}^2 B_{B_q} \eta_B}{12\pi^2} S_0\left(\frac{m_t^2}{m_W^2}\right) |V_{tq}^* V_{tb}|^2, \quad (1.42)$$

where G_F is the Fermi constant. The quantities m_{B_q} , f_{B_q} and B_{B_q} are the mass, weak decay constant and the bag parameter of the B_q^0 meson, respectively. The parameter η_B is a QCD correction factor of order unity [21]. The quantities f_{B_q} and B_{B_q} are calculated in lattice QCD calculations [22]. The known function $S_0(x_t)$ can be well approximated by $0.784x_t^{0.76}$ [23].

The relationship between M_{12} and Δm is given by Equations (1.37) and (1.38). Therefore, in principle, we also need the knowledge of the matrix element Γ_{12} . However, $\Gamma_{12}/M_{12} \approx m_b^2/m_t^2 \ll 1$. Under this approximation, we obtain Δm and $\Delta\Gamma$

from Equations (1.37) and (1.38) as follows

$$\Delta m = 2|M_{12}|, \quad (1.43)$$

$$\Delta \Gamma = 2 \frac{\text{Re}(M_{12}\Gamma_{12}^*)}{|M_{12}|}. \quad (1.44)$$

Thus, the mixing frequencies for the B^0 - \bar{B}^0 and B_s^0 - \bar{B}_s^0 systems can be expressed as

$$\Delta m_d = \frac{G_F^2 m_W^2 m_{B^0} f_{B^0}^2 B_{B^0} \eta_B}{6\pi^2} S_0\left(\frac{m_t^2}{m_W^2}\right) |V_{td}^* V_{tb}|^2, \quad (1.45)$$

$$\Delta m_s = \frac{G_F^2 m_W^2 m_{B_s^0} f_{B_s^0}^2 B_{B_s^0} \eta_B}{6\pi^2} S_0\left(\frac{m_t^2}{m_W^2}\right) |V_{ts}^* V_{tb}|^2, \quad (1.46)$$

$$(1.47)$$

Thus, measurements of Δm_d and Δm_s can be utilized to determine the CKM matrix elements V_{td} and V_{ts} , respectively. In practice, however, the theoretical uncertainties on the weak decay constants f_{B_q} and B_{B_q} limit the precision in the determination of V_{td} and V_{ts} . Therefore, we construct the ratio between Δm_d and Δm_s to obtain

$$\frac{\Delta m_d}{\Delta m_s} = \frac{m_{B^0}}{m_{B_s^0}} \frac{1}{\xi^2} \left| \frac{V_{td}}{V_{ts}} \right|^2, \quad (1.48)$$

where the parameter ξ is calculated from lattice QCD $\xi = 1.21^{+0.047}_{-0.035}$ [24]. Hence, several uncertainties common in the determination of f_{B^0} (B_{B^0}) and $f_{B_s^0}$ ($B_{B_s^0}$) cancel in the ratio. The ratio $|V_{td}/V_{ts}|$ corresponds to one side of the unitarity triangle (see Section 1.1.1). Together with the precise determination of $\Delta m_d = 0.507 \pm 0.005 \text{ ps}^{-1}$ [13], a measurement of the B_s^0 - \bar{B}_s^0 oscillation frequency Δm_s would therefore help infer the ratio $|V_{td}/V_{ts}|$ with a significantly smaller uncertainty, and contribute to a stringent test of the unitarity of the CKM matrix.

1.4 Status of B Mixing Measurements

The first evidence for neutral B meson oscillations was reported in 1987 by the UA1 Collaboration with a study of like-sign muon pairs produced in $\bar{p}p$ collisions in a time-integrated measurement [2]. Shortly after, the ARGUS Collaboration reported [3] the first observation of time-integrated mixing in a like-sign dilepton study for the B^0 - \bar{B}^0 system. This observation was confirmed in 1989 by the CLEO experiment [25]. Since then, time-dependent measurements of B^0 - \bar{B}^0 oscillations have yielded precise measurements of the oscillation frequency, $\Delta m_d = 0.507 \pm 0.005 \text{ ps}^{-1}$ [26, 13]. A summary of Δm_d results is presented in Figure 1.6.

The UA1 measurement, when combined with the time-integrated mixing results from the B^0 system, indicated that oscillations occur in the B_s^0 - \bar{B}_s^0 system, although the oscillation frequency was too high to resolve. In the following, we describe a procedure to search for B_s^0 oscillations in data.

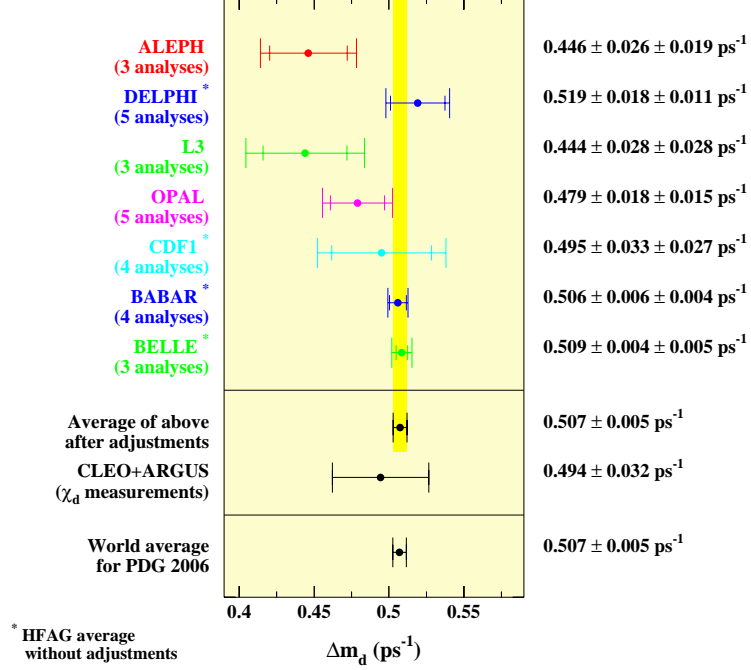


Figure 1.6: Status of measurements of Δm_d from various experiments.

1.4.1 B_s^0 Mixing: Amplitude Technique

To study B_s^0 - \bar{B}_s^0 oscillations, a method mentioned in Ref. [27] is utilized. The so called amplitude technique is a Fourier type of approach for discerning a resonance at the frequency of flavor oscillations present in data. In case of a measurement of B^0 - \bar{B}^0 mixing, the corresponding oscillation frequency Δm_d is directly measured as a parameter in the fit. In contrast, since the B_s^0 oscillation frequency is unknown and has been outside the reach of experiments so far, the amplitude method involves floating a parameter called the amplitude (\mathcal{A}) which multiplies the oscillating term describing the B_s^0 signal as follows:

$$[1 \pm \cos(\Delta m_s t)] \rightarrow [1 \pm \mathcal{A} \cdot \cos(\Delta m_s t)].$$

In the amplitude method, a range of fixed Δm_s frequencies are investigated in the B_s^0 data, and the corresponding fit parameter \mathcal{A} is determined for each value of Δm_s . It is expected that $\mathcal{A} \sim 1$ in case the probed frequency coincides with the true oscillation frequency, and $\mathcal{A} \sim 0$ for the remaining investigated values of the Δm_s spectrum. As a result of the procedure, the amplitude $\mathcal{A}(\Delta m_s)$ and the error on the amplitude $\sigma_{\mathcal{A}}(\Delta m_s)$ are obtained for several probed values of Δm_s .

The error on the amplitude $\sigma_{\mathcal{A}}(\Delta m_s)$ is Gaussian distributed. Therefore, a given Δm_s value can be excluded at 95% confidence level if the following condition is

satisfied:

$$\mathcal{A}(\Delta m_s) + 1.645 \cdot \sigma_{\mathcal{A}}(\Delta m_s) \leq 1. \quad (1.49)$$

For a given $\sigma_{\mathcal{A}}(\Delta m_s)$, statistical fluctuations in the amplitude $\mathcal{A}(\Delta m_s)$ can result in more or less favorable exclusion regions of Δm_s . Hence, a quantity called “sensitivity” is defined to represent the largest value of Δm_s ($= \Delta m_s^{\text{sens}}$) excluded if $\mathcal{A}(\Delta m_s) = 0$ for all $\Delta m_s \leq \Delta m_s^{\text{sens}}$, i.e. the value of Δm_s for which

$$1.645 \cdot \sigma_{\mathcal{A}}(\Delta m_s^{\text{sens}}) = 1. \quad (1.50)$$

The amplitude method is very advantageous, because it allows for a simple combination of independent amplitude measurements from one or more experiments. Such a combination can be achieved by averaging over the results of the corresponding amplitude values at each Δm_s , a feature originating from the linear dependence of the mixing term in \mathcal{A} .

A list of Δm_s amplitude measurements from LEP, SLD and CDF is compiled in Figure 1.7 for a probed value of $\Delta m_s = 17.5 \text{ ps}^{-1}$ [28]. The combined amplitude scan from the various experiments is displayed in Figure 1.8. The exclusion limit obtained from this combination is $\Delta m_s > 14.4 \text{ ps}^{-1}$ at the 95% confidence level.

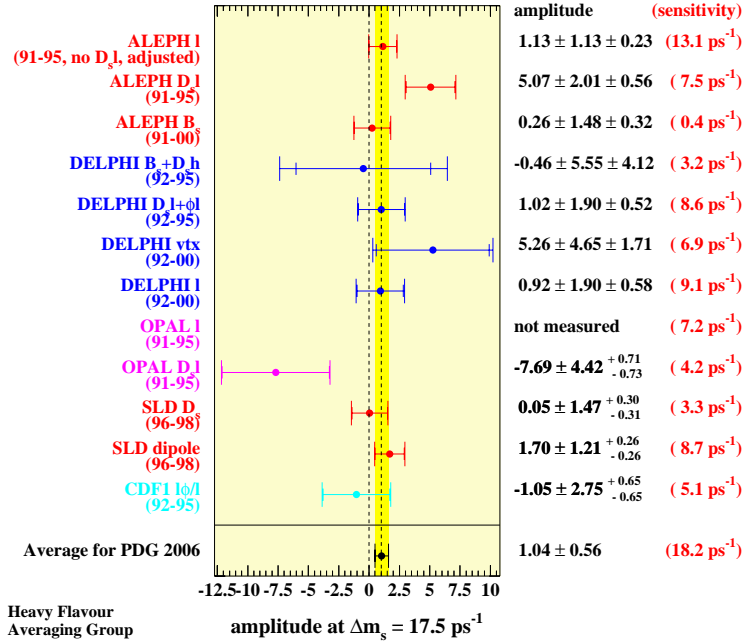


Figure 1.7: Amplitude measurements and sensitivities for various experiments at $\Delta m_s = 17.5 \text{ ps}^{-1}$ as of early 2006.

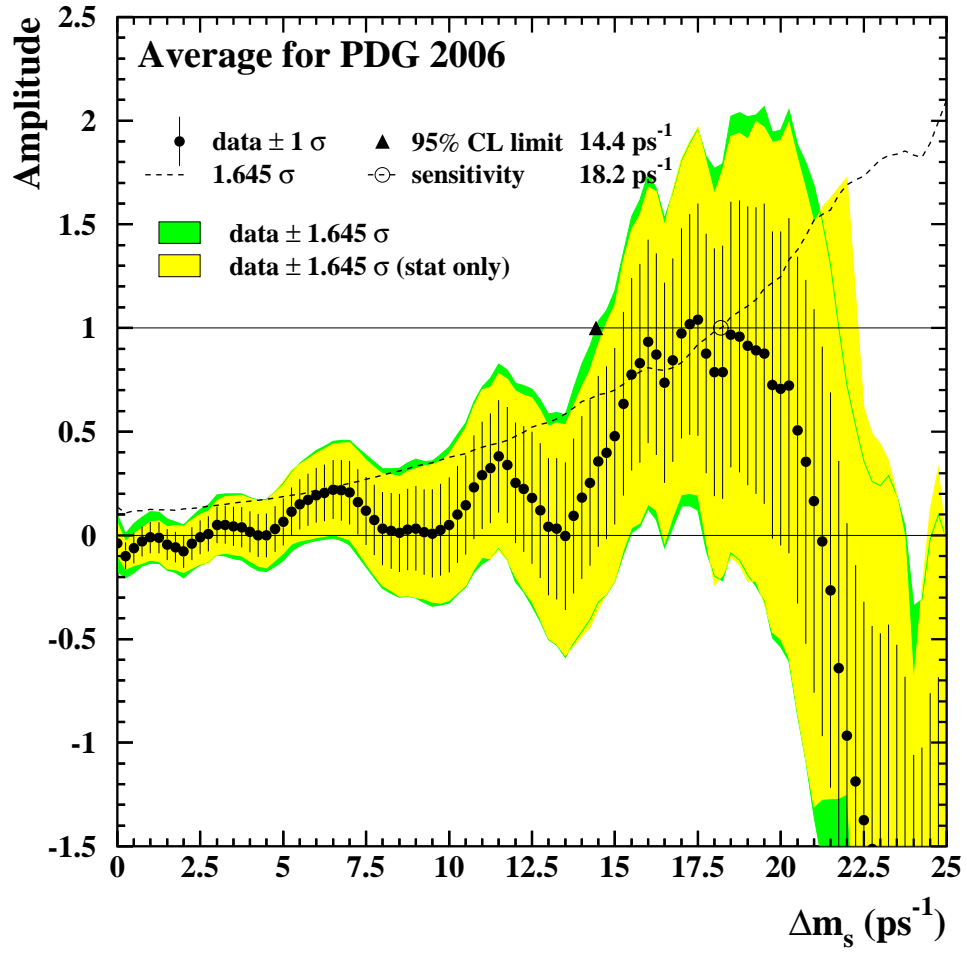


Figure 1.8: A combination of Δm_s amplitude measurements from all the experiments as of early 2006.

1.5 Overview of the Measurement

The goal of this dissertation is to measure the B_s^0 - \bar{B}_s^0 oscillation frequency Δm_s through a time-dependent analysis of semileptonic $B_s^0 \rightarrow \ell^+ D_s^- X$ decays. The basic steps involved in a time-dependent analysis of B meson decays are as follows:

1. Accumulate large samples of B meson decays. This involves a dedicated effort to isolate clean $B \rightarrow \ell D X$ signal decays produced in the $p\bar{p}$ interactions. Further processing of the collected samples is performed in order to reduce backgrounds.
2. For each of the selected B candidates from item 1, determine the production point (primary vertex), and the decay point (secondary vertex). From these two measurements, calculate the decay length of the B candidate. The decay length is used to calculate the decay time, when combined with a B momentum measurement.
3. For each B candidate, determine (tag) the flavor at production, and at decay. In addition, determine the probability that the flavor tag is correct.

The statistical significance of a time-dependent mixing measurement at an oscillation frequency Δm_s is obtained from Ref. [27] as

$$\mathcal{S}(\Delta m_s) = \sqrt{\frac{S}{S+B}} \sqrt{\frac{S\varepsilon\mathcal{D}^2}{2}} \exp\left(\frac{-\sigma_t^2 \Delta m_s^2}{2}\right), \quad (1.51)$$

where S and B denote the number of signal candidates and the number of background candidates, respectively. Both are associated with item 1 described above. The quantity $\varepsilon\mathcal{D}^2$ corresponds to the available flavor tagging power and is related to item 3, while σ_t refers to the proper time resolution and corresponds to errors in the determination of the decay length in item 2. The quantity $\mathcal{S}(\Delta m_s)$ corresponds to the inverse of the measured amplitude error $\sigma_{\mathcal{A}}(\Delta m_s)$, at a given value of Δm_s (see Section 1.4.1).

The exponential dependence implies that the proper time resolution is an important element in this measurement owing to the large expected value of Δm_s . The flavor tagging is challenging in a hadron collider environment, and thus the available resources should be maximally utilized.

In the following, we briefly describe the roadmap of the current measurement.

Collection of Data Samples

At the Tevatron, B mesons originate from the hadronization of b quarks produced in the $p\bar{p}$ interactions. The primary mechanism of b production at the Tevatron is through the creation of $b\bar{b}$ pairs. Upon production, the b quarks undergo a fragmentation process forming mesons and baryons such as B^+ , B^0 , B_s^0 , Λ_b^0 , B_c etc. Thus, all the species of B mesons are produced at the Tevatron, making it a unique place for

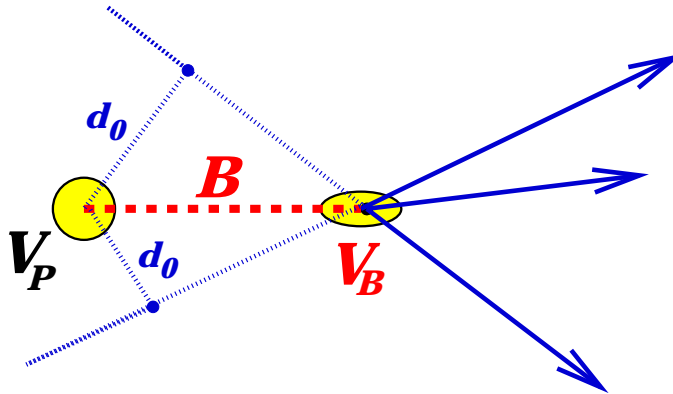


Figure 1.9: Schematic representation of a B decay. B decay daughters typically have large d_0 with respect to the primary vertex V_P .

studying the properties of B_s^0 and B_c mesons as well as Λ_b^0 baryons, until the LHC turn-on.

The b production cross section at the Tevatron is $\mathcal{O}(0.1\%)$ of the total $p\bar{p}$ cross-section of ~ 75 mb. This poses a challenge in collecting samples of B decays. The signatures which allow the selection of large samples of B mesons are their distinctive long lifetimes and relatively higher momentum decay daughters, as compared to background. The interaction point where a B meson is produced is called the primary vertex. B mesons can typically travel a few mm in the CDF-II detector before decaying, and thus the B decay products tend to have a large impact parameter d_0 with respect to the primary vertex. This situation is schematically displayed in Figure 1.9.

The accumulation of B physics events is accomplished via a mixture of hardware and software devices called “triggers”. The stable B decay daughters, that are charged, leave traces in the tracking detectors, which are utilized to reconstruct trajectories called “tracks”. First, momentum measurements are accomplished in the drift chamber. Then, information from the silicon detectors is utilized to make precise impact parameter measurements on the tracks. A given track is then defined to be “displaced”, if the impact parameter is significantly large after taking the measurement errors into account. This unique capability of making real-time decisions on displaced tracks has provided CDF-II with a rich B physics program. A displaced track is then combined with another displaced track or a lepton, and additional kinematical requirements are imposed to reduce backgrounds. The B decay candidates thus collected “online” are stored for further processing.

A description of the CDF-II detector and trigger systems is detailed in Chapter 2.

Further Treatment of Data to Reduce Backgrounds

The obtained B decay candidates are an admixture of the signal $B \rightarrow \ell DX$ decays and a variety of backgrounds. Therefore, the candidates are subjected to “offline” requirements to reduce contamination from backgrounds. For this purpose, dedicated charged particle identification algorithms are utilized to identify the B decay daughters, and signal Monte Carlo samples are generated to study the properties of B decays. Based on the true B decay characteristics, several kinematical and topological criteria are devised, and imposed on the B candidates.

The charged particle identification is described in Chapter 3, followed by a description of the B meson candidate selection procedure in Chapter 4.

Characterization of Data

Despite imposing several selection requirements, backgrounds are still present in our B meson samples. Therefore, a few discriminating variables between signal and background are identified, and a detailed understanding of the nature of these backgrounds is achieved with the help of these variables. The discriminating variables available at our disposal identify various sources of backgrounds differently. Thus, a methodical approach is formulated to characterize backgrounds and determine their contributions in our samples.

The characterization of the B meson samples is detailed in Chapter 5.

Reconstruction of Decay Time

Since we intend to study the time-dependent oscillations of neutral B mesons, the reconstruction of the proper decay time is of crucial importance. The proper decay time of the B mesons can be calculated using the decay length L , which is the distance from the B meson production point to the B meson decay point. The decay length L is related to the proper decay time t in the B rest frame by the Lorentz boost $\beta\gamma$ as $L = \beta\gamma ct$. Typically, L and p are better measured in the plane transverse to the $p\bar{p}$ beam axis, and hence the proper decay time is reconstructed using the variant $L_T = (\beta\gamma)_T ct$. The boost value for the semileptonic decays is inferred using average correction factors from Monte Carlo samples, since the momentum carried away by the missing neutrino and other neutral particles is unknown on a candidate by candidate basis. Corrections are also applied to take into account the finite precision of the L_T measurement, and biases in the proper time acceptance of candidates induced by the trigger requirements.

Precise proper decay time determination is key to observing the rapid B_s^0 - \bar{B}_s^0 oscillations, since large errors in its measurement are likely to obscure the mixing signal. Semileptonic decays used in the current dissertation have an inherent shortcoming: the incomplete reconstruction of such decays leads to a momentum uncertainty, which contributes to a proper time uncertainty that scales with the decay length. Thus,

semileptonic decays with small momentum uncertainty or short decay time are, in general, most sensitive to B_s^0 - \bar{B}_s^0 oscillations. Therefore, we utilize a variable to quantify the missing momentum, and also split our samples depending on the nature of trigger bias on the proper decay time.

The proper time resolution arising from the decay length determination is calibrated using “ B -like” combinations of D mesons and tracks, that are consistent with originating from the primary vertex. Such combinations are thus expected to originate at $t = 0$, and therefore a measurement of their time resolution directly corresponds to the actual proper time resolution.

The reconstruction of decay time, and its resolution, together with its usage as a discriminating variable between signal and backgrounds is contained in Chapter 5. We measure the B_s^0 meson lifetimes as a final validation of our characterization of the semileptonic B_s^0 samples.

Flavor Tagging

Another critical component of this study is B flavor tagging. To measure the B_s^0 - \bar{B}_s^0 oscillation frequency, we need to know the flavor of the B mesons both at production and decay. The task of flavor tagging refers to the determination of the b or \bar{b} quark content of the B mesons. The flavor content of the B mesons in semileptonic decays $B_s^0 \rightarrow \ell^+ D_s^- X$ is determined unambiguously by the charge of the lepton or the D_s^- meson.

Production flavor tagging attempts to determine the b or \bar{b} quark content of a given B meson at its production point i.e., the primary vertex. Broadly speaking, there are two different strategies for tagging the production flavor: same side tagging and opposite side tagging. The same side tagging method exploits correlations between the particles produced in the fragmentation process of the b quark and the resulting B meson. Thus, the same side tagging performance is expected to depend on the B meson species of interest. In the case of B_s^0 mesons, the accompanying particle is expected to be a kaon, and hence particle identification capabilities at CDF-II are likely to play a major role in the same side tag performance. The same side kaon tagging for B_s^0 decays is studied on Monte Carlo samples, and the results thus obtained are applied on data.

Since b quarks are predominantly produced in $b\bar{b}$ pairs at the Tevatron, opposite side tagging attempts to utilize information from the \bar{b} quark in the event. The opposite side tagging algorithms look for leptons from semileptonic decays of the opposite side B meson, or a charged kaon from the $\bar{b} \rightarrow \bar{c} \rightarrow \bar{s}$ transition. Another algorithm is based on identifying jets from the opposite side \bar{b} quark. The opposite side tagging algorithms are studied and optimized on data.

The flavor tagging methods are characterized by a tagging efficiency ε and tagging dilution \mathcal{D} , where $\mathcal{D} = 1 - 2p_W$, with p_W being the probability of a wrong flavor prediction (mistag probability). The figure of merit for the flavor tagging algorithms is

given by $\varepsilon\mathcal{D}^2$, which represents the effective statistical power of the tagging algorithm. The flavor tagging algorithms are described in Chapter 6.

Calibration of Flavor Tagging and B^0 - \bar{B}^0 Mixing

As a necessary step towards studying B_s^0 - \bar{B}_s^0 oscillations, we perform an opposite side tagging calibration, and a measurement of the B^0 - \bar{B}^0 oscillation frequency. The data used for this purpose are $B^+/B^0 \rightarrow \ell DX$ samples collected via the same triggers, and reconstructed using very similar requirements as the $B_s^0 \rightarrow \ell^+ D_s^- X$ signal. The relevant time evolution probability distributions functions (PDFs) for flavor oscillations in the B^0 - \bar{B}^0 systems have been derived in Section 1.2 (Equations (1.40) and (1.41)). The effects of tagging, missing momentum, time resolution, and biases induced by the trigger and reconstruction criteria are then included in the PDFs describing the B^0 and B^+ candidates. For calibration of the opposite side tagging, a simple dilution scaling factor is introduced which multiplies the dilution provided by the opposite side tag. Finally, a simultaneous fit including the B^0 and B^+ mesons is performed, and the values of Δm_d and the dilution scaling factor are extracted.

A noteworthy point here is that only opposite side tagging is calibrated through measurements on B^0 and B^+ mesons; the opposite side tagging performance is expected to be independent of the B meson species in question. Thus, results obtained on B^0 and B^+ mesons for the opposite side tagging methods are equally valid on B_s^0 decays. The same side tagging algorithm, on the other hand, inherently exploits the production properties of B mesons. Thus, same side tagging cannot be calibrated using B^0 and B^+ mesons.

The flavor tagging calibration and Δm_d measurement are detailed in Chapter 7.

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

The final step in the current measurement is the determination of the amplitude \mathcal{A} (see Section 1.4.1) with respect to probing Δm_s values. For the analysis of B_s^0 - \bar{B}_s^0 oscillations, we include the same side kaon tagging in the fit framework. Therefore, we can possibly have upto two tagging decisions for each B_s^0 candidate. Modifications to the time evolution PDFs (with respect to the description for the B^0 - \bar{B}^0 system) for signal and backgrounds are accordingly achieved. The details about the B_s^0 oscillations analysis are presented in Chapter 8.

Finally, the results of the B_s^0 mixing analysis using semileptonic decays are combined with the analyses utilizing hadronic B_s^0 decays at CDF. The combined results are contained in Chapter 9.

Chapter 2

Experimental Apparatus

Fermilab's Tevatron Collider represents the high energy frontier in particle physics. It is currently the source of the highest energy proton - antiproton ($p\bar{p}$) collisions. The collisions occur at two points in an underground ring, which has a radius of about 1 km. Located at these collision points are two detectors: the Collider Detector at Fermilab (CDF-II) and D0. This dissertation uses data collected by the CDF-II experiment.

Between 1997 and 2001, both the accelerator complex and the collider detectors underwent major upgrades, mainly aimed at increasing the luminosity of the accelerator, and gathering data samples of 6 fb^{-1} or more. This is an enormous increase in CDF data sample size as compared to $\sim 120 \text{ pb}^{-1}$ accumulated during the entire Run I. The upgraded machine accelerates 36 bunches of protons and antiprotons, whereas the previous version of the accelerator operated with 6 bunches of protons and antiprotons. Consequently, the time between bunch crossings has been decreased from $3.5 \mu\text{s}$ to 396 ns for the current collider.

The new configuration required detector upgrades for CDF-II to fully utilize the shorter time between beam crossings. In the following pages, we describe how the proton and antiproton beams are produced, accelerated to their final center of mass energy of 1.96 TeV, and collided. We then describe the components used to identify and measure properties of the particles produced in the collision.

2.1 Accelerator Complex

To create the world's most powerful particle beams, Fermilab uses a series of accelerators. The diagram in Figure 2.1 shows the paths taken by protons and antiprotons from initial acceleration to collision in the Tevatron.

The proton acceleration chain begins with hydrogen gas in the Cockcroft-Walton [29] pre-accelerator. Hydrogen gas is ionized inside this device to create H^+ ions and accelerated to 750 keV. The H^+ ions are then sent to a linear accelerator (Linac) [30], approximately 500 feet long, which accelerates the ions to 400 MeV. The H^+ ions

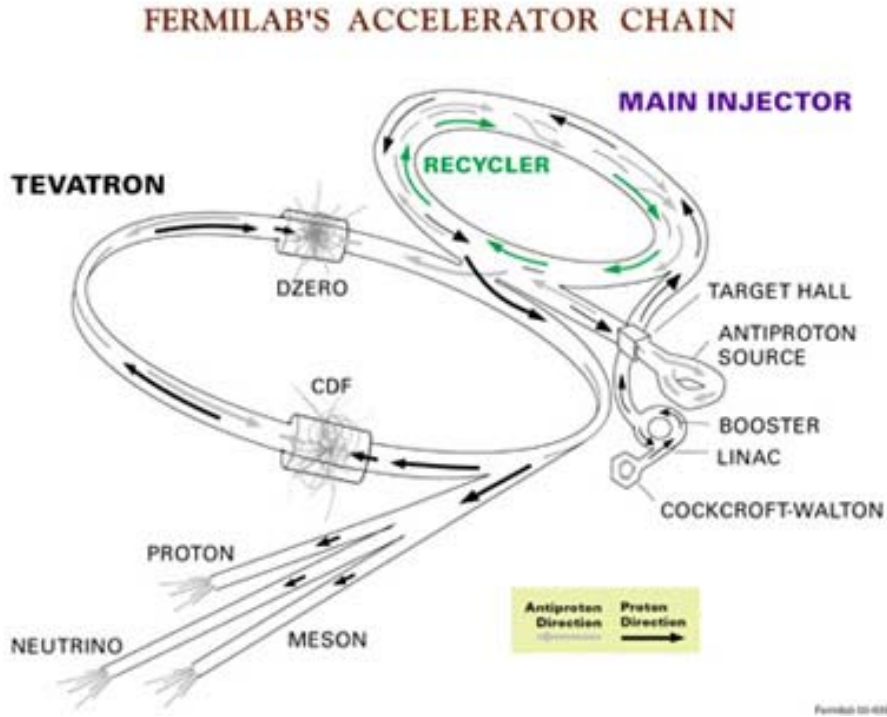


Figure 2.1: Layout of the Fermilab accelerator complex.

are then sent through a carbon foil, stripping off the electrons from the ions, leaving just the bare protons being injected into the Booster [30]. The Booster is a circular synchrotron [30] 74.5 m in diameter. Inside the Booster, the protons are accelerated from 400 MeV to 8 GeV, before sending them to the Main Injector [31].

The Main Injector has four functions. It accelerates protons from 8 GeV to 150 GeV before injection into the Tevatron, it produces 120 GeV protons, which are used for antiproton production, it receives antiprotons from the Antiproton Source and accelerates them to 150 GeV for injection into the Tevatron, and finally, it injects protons and antiprotons into the Tevatron.

To produce antiprotons, 120 GeV protons from the Main Injector are collided into a nickel target in the Target Hall. In these collisions, about 20 antiprotons are produced per one million protons, with a mean kinetic energy of 8 GeV, and collected in the Antiproton Source. A lithium lens is used to focus these particles toward a pulsed magnet, separating the antiprotons from the other particles produced in the collision. Before the antiprotons can be used in narrow beams required for the collider, the difference in kinetic energy between the antiprotons need to be reduced. Since this process reduces the spread of the kinetic energy spectrum of the beam, it is referred to as “cooling” the beam. New batches of antiprotons are initially cooled in the Debuncher synchrotron, collected and further cooled using stochastic cooling [32] in the 8 GeV Accumulator synchrotron. It takes between 10 to 20 hours to build up

a “stack” of antiprotons which is then used for collisions in the Tevatron. Antiproton availability is the most limiting factor in attaining $p\bar{p}$ collision rates.

Since antiprotons are difficult to accumulate, it is desirable to efficiently use the ones that are created. The Main Injector tunnel also houses the Antiproton Recycler. The Recycler was originally built to reuse the remaining antiprotons from a given Tevatron store, cool them and re-integrate them into the stack, so that they can be used in the next store. Currently, the Recycler is operated as an antiproton storage ring that receives batches of antiprotons from the Accumulator and maintains a “stash” of antiprotons. The scheme purpose of this is two-fold. First, it enables the Accumulator to accrue antiprotons at higher currents as the stacking rate in the Accumulator is reduced at high antiproton intensities. Second, the overall antiproton accumulation capacity is dramatically increased with the Recycler’s ability to successfully store stack sizes of more than 400 mA.

Roughly once a day, the stacked antiprotons (36 bunches, each containing 6×10^{10} antiprotons) are injected back into the Main Injector. They are accelerated to 150 GeV together with 36 bunches of roughly 2.5×10^{11} protons each. Both the protons and antiprotons are transferred to the Tevatron.

The Tevatron is the final component of the Fermilab’s accelerator chain. It receives 150 GeV protons and antiprotons from the Main Injector and accelerates them to 980 GeV. The protons and antiprotons circle the Tevatron in opposite directions. The beams are brought to collision at two “collision points”, B0 and D0, using focusing (quadrupole) magnets. The two collider detectors, the CDF-II and D0 are built around the respective collision points. This signifies the beginning of a “store”.

The “instantaneous luminosity”, proportional to the rate of $p\bar{p}$ collisions, can be expressed as:

$$L = \frac{f N_B N_p N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} \mathcal{H} \left(\frac{\sigma_l}{\beta^*} \right), \quad (2.1)$$

where f is the revolution frequency, N_B is the number of bunches, $N_{p/\bar{p}}$ are the number of protons/antiprotons per bunch, and $\sigma_{p/\bar{p}}$ are the beam sizes at the interaction point. The quantity \mathcal{H} is an “hourglass factor” which corrects for the longitudinal bunch shape and depends on the ratio of σ_l (longitudinal bunch length) to β^* (beta function), at the interaction point. The beta function is a measure of the beam width, and is proportional to the beam’s x and y extent in phase space. In Equation (2.1), an ideal case is assumed where the proton and antiproton beams collide in the absence of a crossing angle.

The instantaneous luminosity degrades over time due to beam-beam interactions and losses of particles, particularly the reduction in antiproton current. At the Tevatron, the luminosity has a lifetime of ~ 7 hours at the start of a store. The “integrated luminosity” is defined as $\mathcal{L} = \int L dt$ and measured in units of “inverse femtobarns” ($1\text{fb}^{-1} = 10^{39}\text{cm}^{-2}$). Table 2.1 shows the Run II [31] accelerator parameters as of February 2006. Figure 2.2 shows peak luminosities for stores used in this dissertation as well as the delivered and integrated luminosities at the CDF-II experiment.

Parameter	Run II
number of bunches (N_B)	36
bunch length [cm]	45
bunch spacing [ns]	396
protons/bunch (N_p)	26.0×10^{10}
antiprotons/bunch ($N_{\bar{p}}$)	6.0×10^{10}
$\beta^*(cm)$	28
\mathcal{H}	0.6-0.7
typical luminosity [$\text{cm}^{-2}\text{s}^{-1}$]	1.5×10^{32}
integrated luminosity [fb^{-1}]	1.0
record luminosity [fb^{-1}]	1.7×10^{32}

Table 2.1: Accelerator parameters for the Run II configuration of the Tevatron Collider as of February 2006.

2.2 The CDF-II Detector

The CDF-II detector [33] is a substantial upgrade of the original CDF detector [34], geared towards acquiring data in Run II of the Tevatron. It is located at the B0 collision point of the Tevatron Collider. The detector is designed to measure properties of particles emanating from $p\bar{p}$ collisions. The design of the detector is not geared toward one particular physics measurement, but rather optimized toward extracting different properties of all particle species created in the $p\bar{p}$ collision. Particle detectors such as CDF-II are often called multi-purpose detectors. The first data with the CDF Run II detector was recorded in June 2001. After spending several months commissioning the various sub-detector systems, the physics quality data-taking began in March 2002. We utilize the data taken between March 2002 and February 2006 in this measurement.

Basic concepts of particle detection at CDF-II is shown in Figure 2.3. Neutral particles pass through the tracking chambers without leaving any trace for detection. Electrons and photons deposit their energy in the electromagnetic calorimeter, while hadrons deposit their energy in the hadronic calorimeter. Muons pass through the entire detector, ionizing the calorimeters minimally before being detected in the muon chambers. Utilizing these basic features to discriminate different species of particles, the CDF-II detector is made up of three fundamental sections: tracking chambers, calorimeters, and muon detectors. These three systems are composed of many subsystems that are discussed in the following.

A diagram of the CDF-II detector is shown in Figure 2.4. One half of the detector is cut out to display the different subdetectors. These subsystems can be grouped as follows. The innermost devices constitute the integrated tracking system. It is barrel-shaped and consists of cylindrical subsystems which are concentric with the beam.

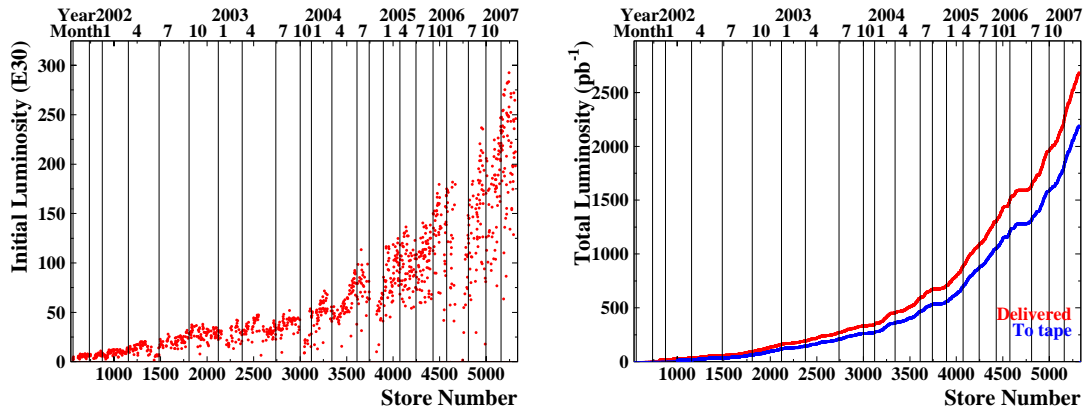


Figure 2.2: Peak luminosities for stores collided between April 2001 and May 2007 (left). Delivered and recorded integrated luminosities at CDF to date (right). We use the data collected between March 2002 and February 2006 in this measurement.

It is designed to detect charged particles, measure their momenta and displacements from the point of collision (primary interaction vertex). The tracking system is surrounded by the Time of Flight (TOF) system, designed to provide particle identification for low-momentum charged particles. Both the tracking and Time of Flight systems are placed inside a superconducting coil, which generates a 1.4 T solenoidal magnetic field. The coil is surrounded by calorimetry systems, which measure the energy of particles that shower when interacting with matter. The calorimetry systems are surrounded by muon detector systems. When interacting with matter, muons act as “minimally ionizing particles” - they only deposit small amounts of ionization energy in the material. Therefore, they are able to penetrate both the tracking and calorimeter systems. The integrated material of the tracking system, TOF, solenoid and calorimetry systems serves as a particle absorber. Particles which penetrate through all that material are mostly muons, and they are detected by leaving tracks in the muon detection system, located outside of the calorimeter.

The most important parts of the CDF-II detector relevant to this dissertation are described in the following sub-sections. A detailed description can be found in the Technical Design Reports of the CDF-I [34] and CDF-II [33] detectors. From here onwards, we will refer to the “CDF-II detector” simply as the “CDF detector”, since we will unambiguously speak of the upgraded CDF detector for Run II.

2.3 CDF Coordinate System

The origin of the CDF detector is located at the beam interaction point. In Cartesian coordinates, the z -axis is defined to be the nominal direction of the proton beam while the x -axis points radially outwards from the center of the Tevatron. The y -axis points upwards. The plane perpendicular to the beam (x - y plane) is referred to as

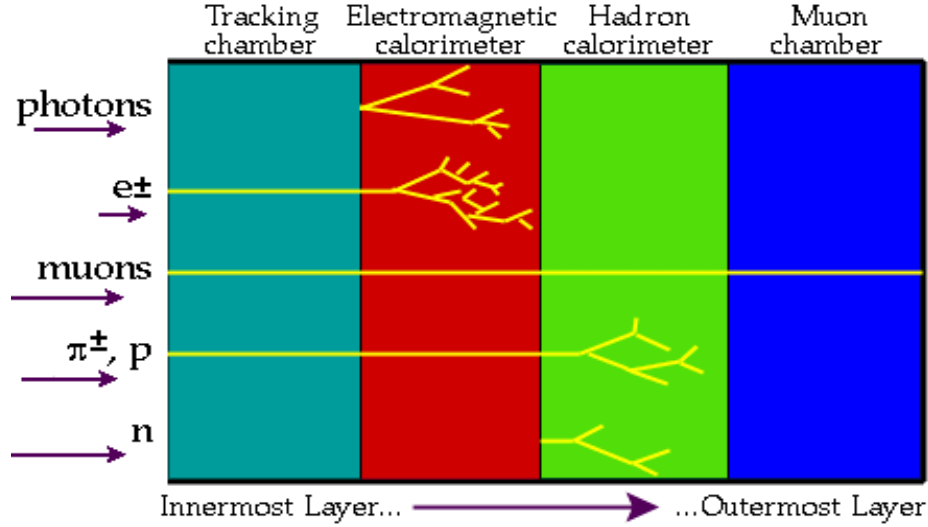


Figure 2.3: Conceptual depiction of particle detection at CDF. The layers represent various detectors at CDF.

the “transverse plane”, and the transverse momentum of the track is referred to as p_T . The CDF detector is a cylindrically and forward-backward symmetric apparatus designed to study $p\bar{p}$ collisions at the Tevatron. This makes it convenient to use a cylindrical (r, φ, z) or polar (r, φ, θ) coordinate system. The quantity r is defined as the radial distance from the center of the detector and φ being the azimuthal angle in the transverse $(x-y)$ plane, counted in the counter-clockwise direction while looking along positive z . Polar angle θ is counted from the positive z -axis.

As opposed to e^+e^- collisions, in $p\bar{p}$ collisions, all of the center of mass energy of the $p\bar{p}$ system is not absorbed in the collision. The colliding partons inside the proton carry only a fraction of the kinetic energy of the proton. As a result, the center of mass system of the parton collisions is boosted along the beam direction (the “longitudinal” direction) by an unknown amount. Quantities defined in the transverse plane are conserved in the collisions. For instance, the sum of all transverse momenta of particles in the collisions is zero ($\sum \vec{p}_T = 0$).

Particles moving through a homogenic solenoidal magnetic field follow helical trajectories. Reconstructed particle trajectories are referred to as “tracks”. To uniquely parameterize a helix in three dimensions, five parameters are needed. The CDF coordinate system chooses three of these parameters to describe a position, and two more to describe the momentum vector at that position. The three parameters which describe a position describe the point of closest approach of the helix to the beam line. These parameters are d_0 , φ_0 , and z_0 , which are the r, φ and z cylindrical coordinates of the point of closest approach of the helix to the beam. The momentum vector is described by the track curvature (C) and the angle of the momentum in the r - z plane ($\cot \theta$). From the track curvature we can calculate the transverse momentum

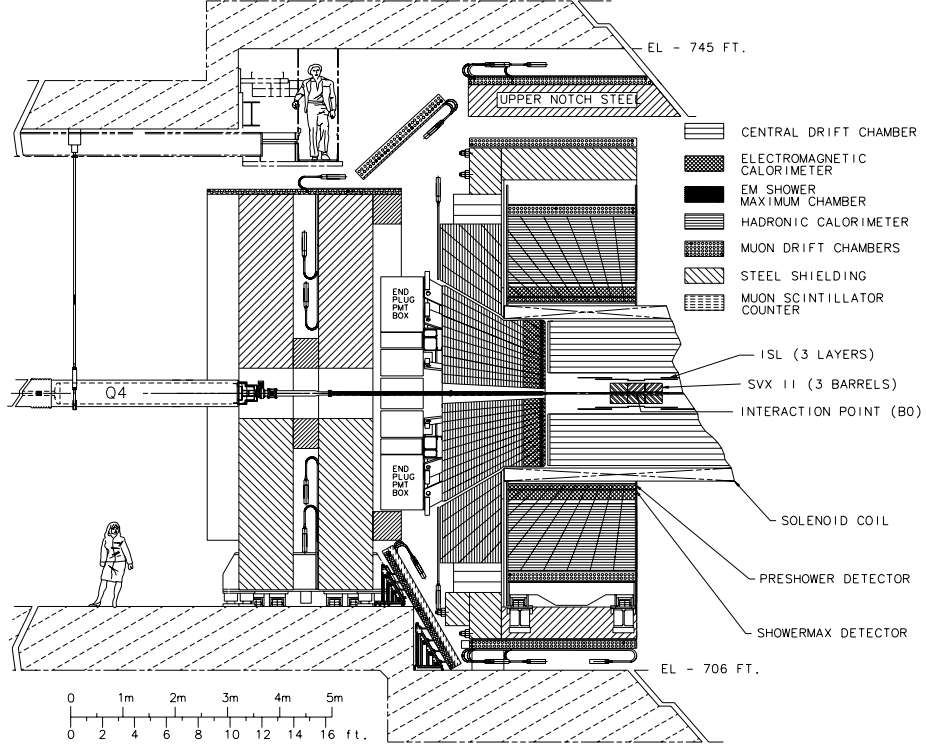


Figure 2.4: Elevation view of the CDF-II detector with one half cut to display the different subdetectors.

p_T . The curvature is signed so that the charge of the particle matches the sign of the curvature. From $\cot \theta$, we can calculate $p_z = p_T \times \cot \theta$. At any given point of the helix, the track momentum is a tangent to the helix. This basically means that the angle φ_0 implicitly defines the direction of the transverse momentum vector p_T at the point of closest approach.

The impact parameter d_0 of a track is another signed variable; its absolute value corresponds to the distance of closest approach of the track to the beamline. The sign of d_0 is taken to be that of $\hat{p} \times \hat{d} \cdot \hat{z}$, where \hat{p} , \hat{d} and \hat{z} are unit vectors in the direction of p_T , d_0 and z , respectively. An alternate variable that describes the angle between the z -axis and the momentum of the particle is the pseudorapidity η defined as:

$$\eta \equiv -\ln \tan(\theta/2). \quad (2.2)$$

The quantity η is the ultrarelativistic/massless limit of the particle's rapidity (y) defined as follows:

$$y \equiv \frac{1}{2} \ln \frac{E + p_z}{E - p_z}. \quad (2.3)$$

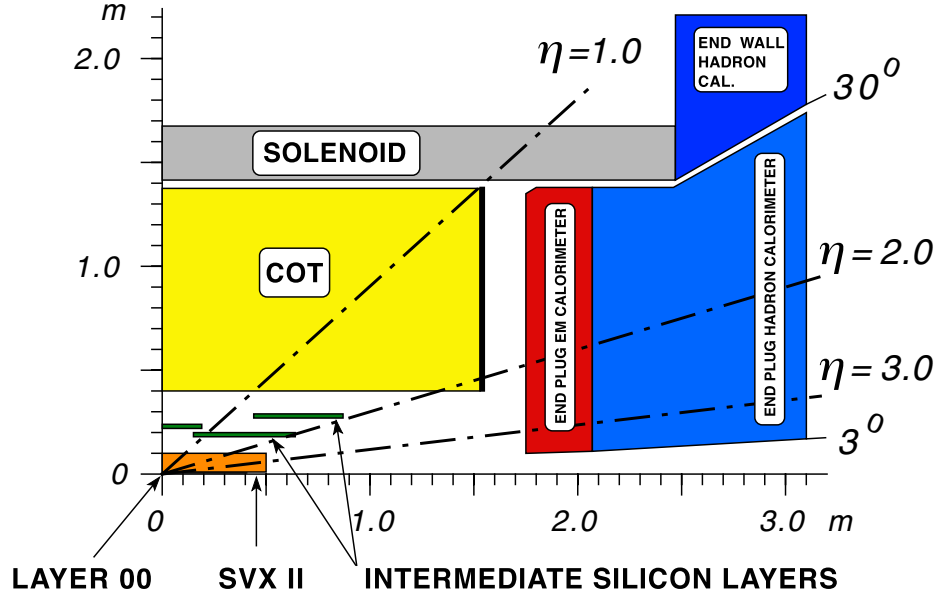


Figure 2.5: A diagram of the CDF tracker layout showing the different subdetector systems.

2.4 Tracking Systems

Charged particles cause ionization as they pass through matter. Typically, this ionization is localized near the trajectory of the particle in small clusters, called hits. Once detected, hits can be used to reconstruct particle's trajectory in the process known as tracking.

The CDF detector has a cylindrical tracking system immersed in a 1.4 T solenoidal magnetic field (directed along the z -axis) for the measurement of charged-particle momenta. Figure 2.5 displays a layout of the CDF tracking system. We will describe this system starting from the devices closest to the beam and moving outwards. The innermost tracking device is a silicon strip vertex detector system, which consists of three subdetectors. A layer of silicon sensors, called Layer 00 (L00) [35], is installed directly onto the beryllium vacuum beam pipe, at a radius of 1.7 cm from the beam. The beam pipe is made of beryllium because this metal has the best mechanical qualities, yet lowest nuclear interaction cross section of all materials.

The layer of silicon on the beam pipe is followed by five concentric layers of silicon sensors (SVXII) [36] located at radii between 2.5 and 10.6 cm. The Intermediate Silicon Layers (ISL) [37] are the outermost silicon subdetector systems, consisting of one layer at a radius of 22 cm in the central region and layers at radii 20 and 28 cm in the forward regions. Surrounding the silicon detector is the Central Outer Tracker (COT) [38], a 3.1-m-long cylindrical open-cell drift chamber covering radii from 40 to 137 cm.

Property	Layer 0	Layer 1	Layer 2	Layer 3	Layer 4
number of φ strips	256	384	640	768	869
number of Z strips	256	576	640	512	869
stereo angle	90°	90 °	+1.2°	90°	-1.2°
φ strip pitch	60 μm	62 μm	60 μm	60 μm	65 μm
Z strip pitch	141 μm	125.5 μm	60 μm	141 μm	65 μm
active width (mm)	15.30	23.75	38.34	46.02	58.18
active length (mm)	72.43	72.43	72.38	72.43	72.43

Table 2.2: Relevant parameters for the layout of the sensors of different SVX-II layers.

2.4.1 Silicon Tracking Detectors

Silicon tracking detectors are used to obtain precise position measurements of the path of a charged particle. A silicon tracking detector is fundamentally a reverse-biased p-n junction. When a charged particle passes through the detector material, it causes ionization. In the case of a semi-conductor material, this means that e^- -hole pairs will be produced. Electrons drift towards the anode, and holes drift toward the cathode, where the charge is gathered. The amount of charge is, to first order, proportional to the path length traversed in the detector material by the charged particle.

By segmenting the p or n side of the junction into “strips” and reading out the charge deposition separately on every strip, we obtain sensitivity to the position of the charged particle. All the CDF silicon tracking detectors are implemented as micro-strip detectors. The typical distance between two strips is about 60 μm . Charge deposition from a single particle passing through the silicon sensor will be read out on one or more strips. This charge deposition is called a “cluster”. There are single and double-sided microstrip detectors. In single-sided detectors, only one (p) side of the junction is segmented into strips. Double-sided detectors have both sides of the junction segmented into strips. The benefit of double-sided detectors is that while one (p) side has strips parallel to the z -direction, providing r - φ position measurements, the other (n) side can have strips at an angle (stereo angle) with respect to the z -direction, which will give z -position information.

The innermost layer, L00, is made of single-sided silicon sensors which provide r - φ position measurements only. The SVX-II and ISL are made of double-sided silicon sensors. As shown in Table 2.2, the SVX-II layers have different stereo angles. Two layers have a small (1.2°) stereo angle and three have a 90° stereo angle. The ISL detector provides small-angle (1.2°) stereo information.

Four silicon sensors are stacked length-wise into a “ladder” structure which is 29 cm long. The readout electronics are mounted onto the ends of the ladders. The ladders are organized in an approximately cylindrical configuration, creating “barrels”. A SVX-II barrel is segmented into 12 wedges, each covering approximately

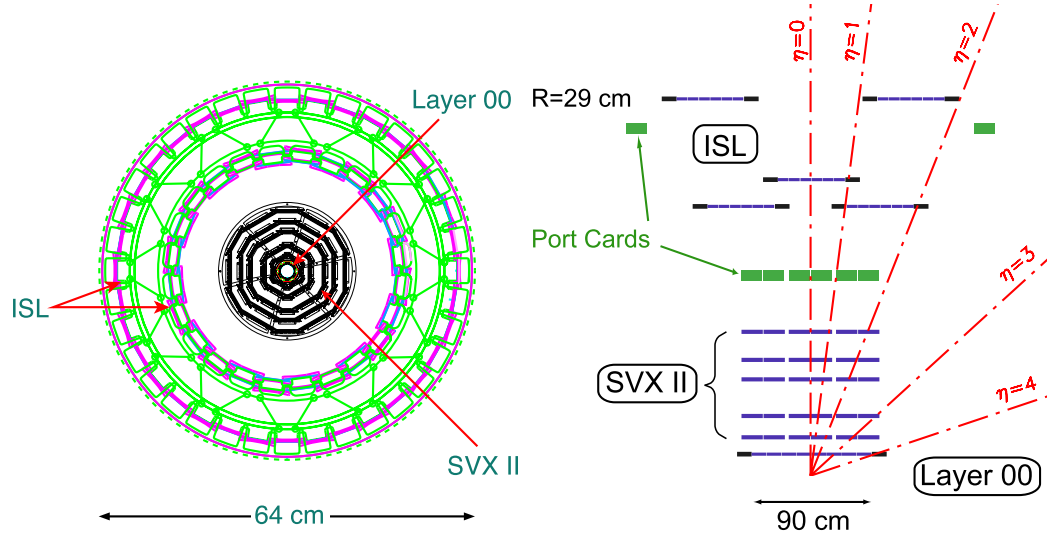


Figure 2.6: Silicon tracking detectors projected on the r - ϕ plane (left) and r - z plane (right), not drawn to scale.

30° in ϕ with a small overlap at the edges, allowing for many silicon hits per track. There are three SVX-II barrels, adjacent to each other along the z -axis, covering the nominal interaction point in the center of the CDF detector. The coverage of the silicon detector subsystems is shown in Figure 2.6.

2.4.2 Central Outer Tracker

The COT drift chamber provides accurate tracking information in the r - ϕ plane for the measurement of transverse momentum, and substantially less accurate information in the r - z plane for the measurement of the z -component of the momentum, p_z . The COT contains 96 sense wire layers, which are radially grouped into eight “superlayers”, as inferred from the end plate section shown in Figure 2.7. Each superlayer is divided in ϕ into “supercells”, and each supercell has 12 sense wires and a maximum drift distance that is approximately the same for all superlayers. Therefore, the number of supercells in a given superlayer scales approximately with the radius of the superlayer. The entire COT contains 30,240 sense wires. Approximately half of the the wires run along the z -direction (“axial”). The other half are strung at a small stereo angle (2°) with respect to the z -direction.

The active volume of the COT begins at a radius of 43 cm from the nominal beamline and extends out to a radius of 133 cm. The chamber is 310 cm long. Particles originating from the interaction point which have $|\eta| < 1$ pass through all 8 superlayers of the COT. Particles which have $|\eta| < 1.3$ pass through 4 or more superlayers.

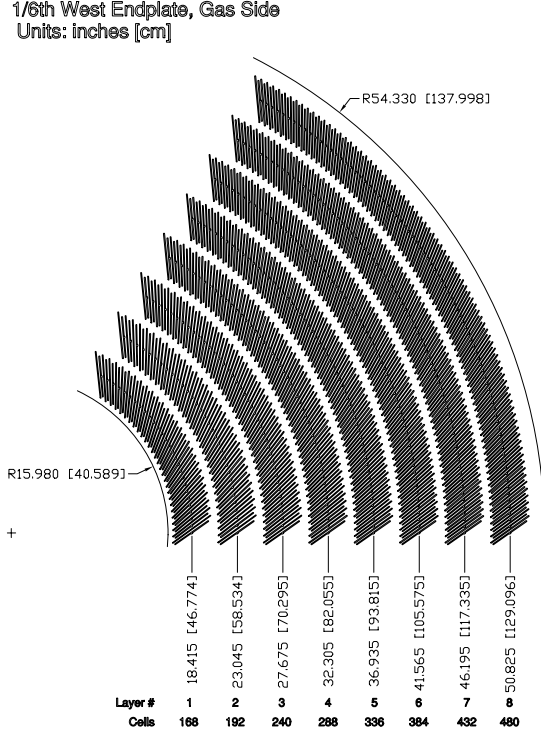


Figure 2.7: Layout of wire planes on a COT endplate.

The supercell layout, shown in Figure 2.8 for superlayer 2, consists of a wire plane containing sense and potential (for field shaping) wires and a field (or cathode) sheet on either side. Both the sense and potential wires are $40\text{ }\mu\text{m}$ diameter gold plated Tungsten. The field sheet is $6.35\text{ }\mu\text{m}$ thick Mylar with vapor-deposited gold on both sides. Each field sheet is shared with the neighboring supercell.

The COT is filled with a 50:50 admixture of Argon-Ethane gas. The mixture is chosen to have a constant drift velocity across the cell width. When a charged particle passes through, the gas is ionized. Electrons drift towards the sense wires. The electric field in a cylindrical system grows exponentially with decreasing radius. As a result, the electric field very close to the sense wire is large, resulting in an avalanche discharge when the charge drifts close to the wire surface. This effect provides a gain of $\sim 10^4$. The maximum electron drift time is approximately 100 ns. Due to the magnetic field in which the COT is immersed, electrons drift at a Lorentz angle of $\sim 35^\circ$. The supercell is therefore tilted by 35° with respect to the radial direction to compensate for this effect.

Signals on the sense wires are processed by the ASDQ (Amplifier, Shaper, Discriminator with charge encoding) chip, which provides input protection, amplification, pulse shaping, baseline restoration, discrimination and charge measurement. The charge measurement is encoded in the width of the discriminator output pulse,

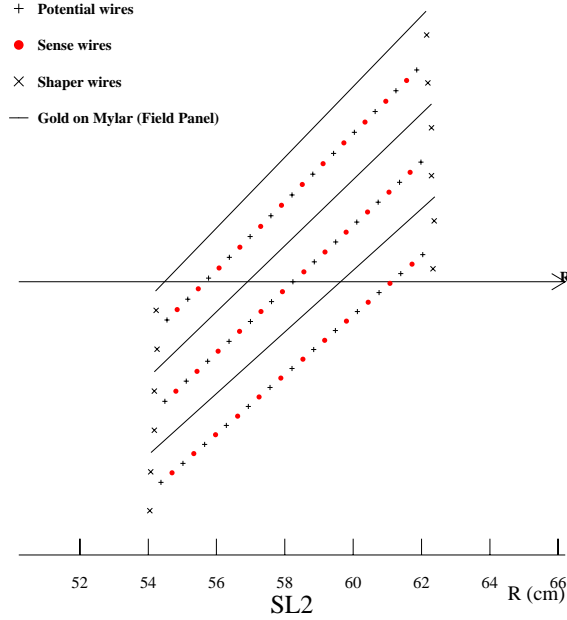


Figure 2.8: Layout of wires in a COT supercell.

and is used for particle identification by measuring the ionization along the trail of the charged particle (dE/dx). The pulse is sent through ~ 35 ft of micro-coaxial cable, via repeater cards to Time to Digital Converter (TDC) boards in the collision hall. The recorded hit times translated into position points are later processed by pattern recognition (tracking) software to form helical tracks.

2.4.3 Pattern Recognition

As explained in the previous sections, charged particles leave small charge depositions as they pass through the tracking system called hits. By following, or “tracking” these hits, pattern recognition algorithms can reconstruct the charged particle track.

There are several pattern recognition algorithms used to reconstruct tracks in the CDF tracking system. Most of the tracks are reconstructed using “Outside-In” algorithms which we will describe here. The name of this group of algorithms suggest that the track is followed from the outside of the tracking system inward.

The track is first reconstructed using only COT information. The COT electronics reports hit time and integrated charge for every wire in an event. The hit time corresponds to the time that an avalanche occurred at a sense wire. The hit time can be interpreted as the drift time of the charge in the gas, but it has to be corrected for time of flight first. The hit timing resolution is of the order of a few ns; this roughly corresponds to the average spread in collision times. It is assumed that the collision times always happen at the same time in a cycle during a store. An average of collision

times is determined for many previous events and used as the event collision time. Hit times corrected for the collision time are interpreted as drift times and used in the pattern recognition algorithm. To perform the final track fit, an additional time of flight correction is performed assuming massless particles.

The helical track, when projected onto the two dimensional r - φ plane, is a circle. This simplifies pattern recognition. The first step of pattern recognition in the COT looks for circular paths in radial superlayers of the COT. Super-cells in the radial superlayers are searched for sets of 4 or more hits that can be fit to a straight line. These sets are called “segments”. The straight-line fit for a segment gives sufficient information to extrapolate rough measurements of curvature and φ_0 . Once segments are found, there are two approaches to finding tracks. One approach is to link together segments for which the measurements of curvature and φ_0 are consistent. The other approach is to improve the curvature and φ_0 measurement of a segment reconstructed in superlayer 8 by constraining its circular fit to the beamline, and then adding hits which are consistent with this path. Once a circular path is found in the r - φ plane, segments and hits in the stereo superlayers are added by their proximity to the circular fit. This results in a three-dimensional track fit. Typically, if one algorithm fails to reconstruct a track, the other algorithm will not. This results in a high track reconstruction efficiency ($\sim 95\%$) in the COT for tracks which pass through all 8 superlayers ($p_T \geq 400 \text{ MeV}/c^2$). The track reconstruction efficiency mostly depends on how many tracks there are to be reconstructed in the event. If there are many tracks present close to each other, hits from one track can shadow hits from the other track, resulting in efficiency loss.

Once a track is reconstructed in the COT, it is extrapolated into the SVX-II. Based on the estimated errors on the track parameters, a three-dimensional “road” is formed around the extrapolated track. Starting from the outermost layer, working inward, silicon clusters found inside the road are added to the track. As a cluster gets added, the road gets narrowed according to the knowledge of the updated track parameters. Reducing the width of the road reduces the chance of adding a wrong hit to the track, and also reduces computation time. In the first pass of this algorithm, only r - φ clusters are added. In the second pass, clusters with stereo information are also added to the track.

2.5 Time of Flight

Outside the tracking system, but still inside the superconducting magnetic coil, the CDF detector has a Time of Flight (TOF) [39] system as shown in Figure 2.9. The TOF system is designed to distinguish low momentum pions, kaons and protons by measuring the time it takes these particles to travel from the primary vertex of the $p\bar{p}$ collision to the TOF system. The system consists of 216 bars of scintillating material, roughly 300 cm in length with a cross-section of $4 \times 4 \text{ cm}$. The bars are arranged into a barrel geometry around the COT outer cylinder. They are surrounded by the

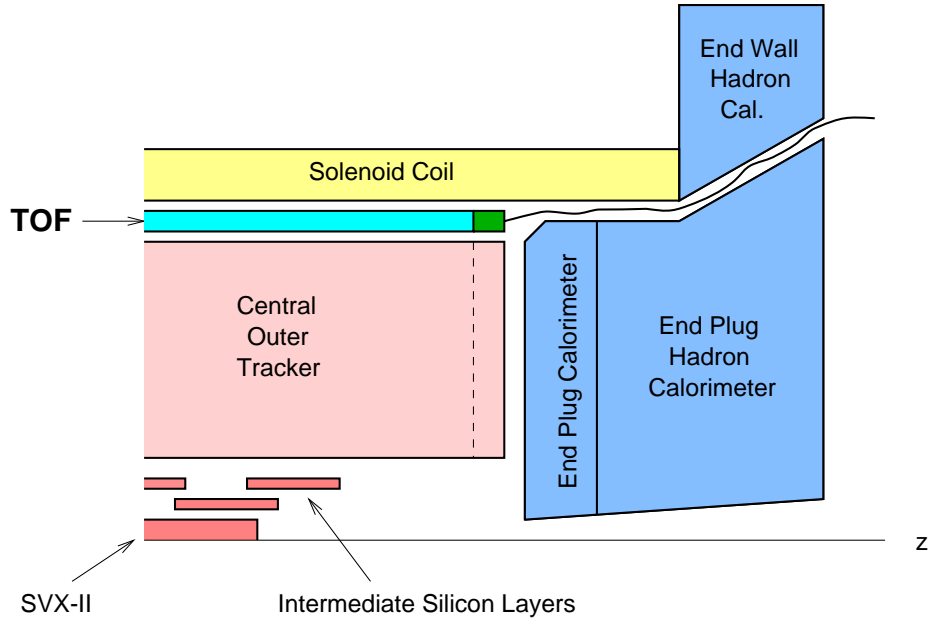


Figure 2.9: Location of the Time of Flight system inside the CDF detector.

superconducting solenoid on the outside.

Particles passing through the scintillator material of the bars result in the emission of photons. Photomultiplier tubes (PMTs), attached to both ends of each bar, collect the emitted photons and provide time and pulse height measurements. By comparing the two timing measurements, the longitudinal z -position of the hit is determined. The z -position, along with the φ location of the bar, is then used to associate tracks in the COT to the TOF information.

The signal from the PMTs is processed by a pre-amplifier circuit mounted directly onto the phototube. The amplified signal is sent via a twisted pair to the readout electronics in the collision hall. The readout electronics performs both time and amplitude digitization of the signal. The TDC information is a digitization of the time when the signal pulse reaches a fixed discriminator threshold. This time depends on the amplitude of the pulse, since a large pulse crosses the threshold earlier (time walk). The digitization of the pulse amplitude is needed to correct for this effect. The timing resolution also varies with the displacement from the photomultiplier tube. Large pulses give better timing resolution, and light attenuates while traveling through the scintillator material. Therefore, particles passing through the bar near the photomultiplier tube have better timing resolution than those which are far away. After correcting for time walk effects, the timing resolution of the TOF system is currently about 110 ps for particles crossing the bar exactly in front of one of the photomultiplier tubes.

2.6 Calorimeters

Segmented electromagnetic and hadronic scintillator sampling calorimeters surround the tracking system and measure the energy flow of interacting particles in the pseudorapidity range $|\eta| < 3.64$. The coverage of the calorimetry system is 2π in azimuthal angle φ . In general, sampling calorimeters consist of an absorber material (e.g. lead, steel) interspersed with an active material (e.g. scintillators). As a particle traverses through a layer of absorber material, it interacts with the material, and its energy is reduced thereby producing a “shower”. The active material then samples the energy of the shower and the total amount of energy collected by all the sampling layers is proportional to the energy of the incident particle. The collected energy is then converted into light in the visible range with the help of wavelength shifting fibers, transformed into an electronic signal via the PMTs, and finally digitized with the help of electronics.

There are two kinds of showers depending on the type of incident particle. Electromagnetic showers develop when an electron or photon interacts with the calorimeter material via electromagnetic interactions, while hadronic showers are produced when a hadron interacts with the material via strong interactions. As electrons or photons traverse the material, electrons radiate photons (Bremsstrahlung radiation) and photons convert into electron-positron pairs (pair production/photon conversions). Eventually there is not enough energy to form any more pairs, and a shower maximum is reached. Thereafter, the electrons mostly lose their residual energy through ionization, while photons undergo Compton scattering. The showering properties of hadrons are inherently different from electrons and photons due to the difference in underlying interactions; hadrons typically interact with the material’s nuclei via strong interactions.

The calorimeter system is divided into two regions: central and plug. The central calorimeter consists of the central electromagnetic (CEM) [40], central hadronic (CHA) [41] and end-wall hadronic (WHA) [41] calorimeters. The plug region contains the plug electromagnetic (PEM) and plug hadronic (PHA) calorimeters. The pseudorapidity coverage, energy resolutions and thickness of the different electromagnetic and hadron calorimeters is given in Table 2.3. The radiation length for the electromagnetic calorimeters (X_0) corresponds to the distance traveled by an electron before losing $1/e$ of its original energy. The amount of material in the hadronic calorimeter is expressed in terms of pion interaction lengths (λ), the distance for which a pion has $1/e$ probability of not interacting with a nuclei.

The CDF calorimeter has a “projective tower” geometry, meaning that it is segmented in η and φ “towers” that point back to the interaction region. Each calorimeter tower consists of an electromagnetic shower counter followed by a hadron calorimeter. This allows for comparison of the electromagnetic and hadronic energies deposited in each tower, and therefore enables the separation of electrons and photons from hadrons. The electrons and photons are expected to shower mostly in

System	η coverage	Energy Resolution	Thickness
CEM	$ \eta < 1.1$	$13.5\%/\sqrt{E_T} \oplus 2\%$	$18 X_0$
PEM	$1.1 < \eta < 3.64$	$16\%/\sqrt{E_T} \oplus 1\%$	$21 X_0$
CHA	$ \eta < 0.9$	$75\%/\sqrt{E_T} \oplus 3\%$	$4.7 \lambda_0$
WHA	$0.7 < \eta < 1.3$	$75\%/\sqrt{E_T} \oplus 4\%$	$4.5 \lambda_0$
PHA	$1.3 < \eta < 3.64$	$80\%/\sqrt{E_T} \oplus 5\%$	$7.0 \lambda_0$

Table 2.3: Pseudorapidity coverage, energy resolution and thickness for the different calorimeter subdetectors of the CDF experiment. The \oplus symbol means that the constant term is added in quadrature to the resolution. λ_0 signifies interaction lengths and X_0 radiation lengths.

the electromagnetic calorimeter while hadrons are expected to deposit most of their energy in the hadronic calorimeter. The electromagnetic and hadronic calorimeters are symmetric in the azimuth and are segmented into 15° wedges in φ , except in the plug calorimeter between $1.10 < |\eta| < 2.11$ where the wedges are segmented into 7.5° wedges in the azimuth. The segmentation in η is 0.11. CEM and PEM use lead sheets, CHA and WHA use steel while iron is used in the PHA as the absorber material, alternated by the active medium made up of polystyrene based scintillators.

Inside the electromagnetic calorimeters at a depth of $\sim 6X_0$, the shower maximum detectors are located, approximately coincident with the region of maximum shower intensity for electrons. These detectors (CES in the central region and PES in the plug) make precise shower position and shower profile measurements within the electromagnetic calorimeters. The increased shower position resolution provides additional selection criteria for electron candidates based on track-shower matching. The CES is a gas proportional chamber (mixture of Ar and CO_2) and measures charge depositions along 128 cathode strips and 64 anode wires running perpendicular to each other in each of the 24 wedges. The wires measure position and shower profile along the transverse (r - φ) direction, while the strips provide the same measurements along the z -axis. The position resolution along both the directions is about 2 mm for electrons. The PEM also has a shower maximum detector (PES) [42] and is composed of eight 45° sectors. Each sector contains two layers (called U and V) of 5 mm wide scintillator strips, measuring shower positions with resolutions of ~ 2 mm.

Located between the CEM and the solenoid coil, there is another set of gas multiwire proportional chambers. These are called the central preshower detector (CPR) [43]. They sample the early development of electromagnetic showers in the solenoid coil material ($1.8 X_0$) in front of them, providing enhanced electron identification. The CPR detector provides shower information only in the r - φ plane. The CPR has been upgraded in 2004 and is replaced by a new detector (CPR2) [44], which has better segmentation. Approximately 64% of the data used in this dissertation is acquired using the CPR2 detector.

System	η coverage	λ	Minimum Muon p_T (in GeV/ c)
CMU	$ \eta < 0.6$	5.5	1.4
CMP	$ \eta < 0.6$	7.8	2.2
CMX	$0.6 < \eta < 1.0$	6.2-10.0	1.4-2.0
IMU	$1.0 < \eta < 1.5$	6.2-20.0	1.4-3.0

Table 2.4: Pseudorapidity coverage, pion interaction lengths and minimum muon p_T for detection listed for the various muon detector devices at CDF.

2.7 Muon Systems

Muons are particles which mainly interact with matter by ionization. For energies relevant to this experiment, they do not cause showers in the electromagnetic or hadronic calorimeters. As a result, if a muon is created in the collision and has enough momentum, it will pass through the calorimeter with minimal interaction with the calorimeter material. Therefore, the calorimeter system can be considered as a “filter” that mostly allows only muons to pass through. Muon detection systems are therefore placed radially outside the calorimeters. Additional steel absorbers are placed between the calorimeters and the muon chambers to further reduce the chance that particles other than muons reach the muon chambers.

The muon detectors are composed of drift chambers and scintillation counters covering an $|\eta| < 1.5$ range. Figure 2.10 shows the coverage of various muon detectors. The CDF detector has four muon systems: the Central Muon Detector (CMU) [45], the Central Muon Upgrade Detector (CMP) [46], the Central Muon Extension Detector (CMX) [46], and the Intermediate Muon Detector (IMU) [33]. The coverage of the muon detectors in η , the amount of absorber in front of them in terms of pion interaction lengths (λ), and the minimum muon p_T needed to reach them is displayed in Table 2.4.

The CMU is located right outside the CHA at a radius of 347 cm from the beam axis and is made up of muon drift cells with seven wires each. It has the same segmentation as the CEM and CHA (15°) in φ . However, there is a 2.4° gap between drift cell arrays leading to an overall φ coverage of 84%. Each wedge is further segmented in the r - φ plane into three 4.2° modules. Each module consists of four layers of four rectangular drift cells. The sense wires in alternating layers are offset by 2 mm for ambiguity resolution. The smallest unit in the CMU, called a “stack”, covers about 1.2° and includes four drift cells, one from each layer. Using the timing information from the drift cells, track segments (called “stubs”) are reconstructed. A second set of muon drift chambers, called the CMP, is located behind an additional 60 cm of absorber material in the form of steel (2.3 pion interaction lengths). The chambers are 640 cm long and arranged axially to form a box around the central detector. The CMP chambers are comprised of rectangular ($2.5 \text{ cm} \times 15 \text{ cm}$), single-

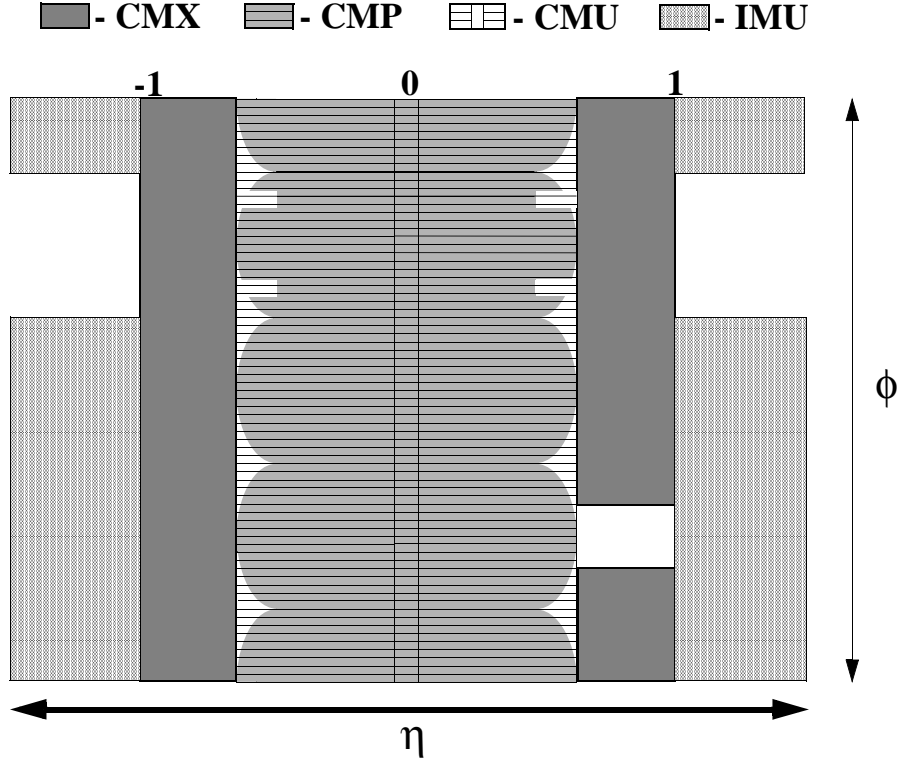


Figure 2.10: CDF muon chamber coverage.

wire drift tubes configured in four layers with alternate half-cell staggering. Muons which register a stub in both the CMU and CMP are called CMUP muons, and are considered the purest types of muons at CDF.

The CMX detector, composed of conical sections drift chambers and scintillation counters, extends the muon coverage from $|\eta| \sim 0.6$ to 1.0. Each 15° wedge is made up of eight layers of drift tubes with six tubes in each layer. The CMX drift tubes have the same rectangular cross section as the CMP drift tubes and only differ in length; they are 180 cm long. The IMU detector further extends the muon $|\eta|$ coverage from 1.0 to 1.5. It consists of a barrel of drift chambers and scintillator counters around two steel toroids with additional counters between the toroids on either side of CDF. The IMU chambers and counters are very similar to those of the central muon systems.

2.8 The CDF-II Trigger System

A trigger system at CDF is necessary because it is not physically possible to store information about every single $p\bar{p}$ collision. Collisions happen roughly at a rate of 2.5 MHz, and the readout of the full detector produces an event roughly the size of 250 kB. There is currently no medium available which is capable of recording data

this quickly, nor would it be practical to analyze all this data later on. The trigger system is thus a “pre-filter”, which reduces data rates and volumes to manageable levels.

The CDF-II triggering system is designed based on three conditions. The first condition is that the trigger operates deadtimeless. This means that the trigger system has to be fast enough to make a decision for every single event, before the next event occurs. The second condition is imposed by the Tevatron upgrade for Run II, and it is the expected time between collisions, 396 ns. The last condition is that the data logging system can write up to about 100 events per second to tape, because of limited resources. In short, the trigger has to be fast enough to analyze every collision, and it has to figure out which 100 of 2.5 million events it should save in a given second. This is achieved by staging trigger decisions in three levels, as shown in Figure 2.11.

Each level of the trigger has a certain maximum amount of time to reach a decision about accepting or rejecting an event. By increasing the time allowed for triggering at different levels of the trigger, the complexity of reconstruction tasks can be increased at every level. Each level uses more restrictive event selection criteria, reducing the event rate such that it can be handled by the next level. At the first level of the trigger, only very rough and quick pattern recognition and filtering algorithms are used. In order to do this in time, the Level 1 and Level 2 triggering mechanisms are based on hardware with custom electronics. The third level of the trigger is implemented as a software trigger using a PC farm with about 500 CPU’s. The delay necessary to make a trigger decision is achieved by storing detector readout information in a storage pipeline. At Level 1, for every Tevatron clock cycle, the event is moved up one slot in the pipeline. By the time it reaches the end of the pipeline, the trigger will have reached a decision whether to accept or reject this event. If the event is accepted, its information will be sent to the higher level of the trigger. Otherwise, the event is simply ignored.

A set of requirements that an event has to fulfill at Level 1, Level 2 and Level 3 constitutes a “trigger path”. In the candidate selection for the B meson decays utilized in this measurement, we require that an event was accepted through a well defined trigger path. This eliminates so called “volunteer events”. A volunteer event is an event which passed a higher level (Level 2 or Level 3) trigger requirement but did not pass the preceding lower level (Level 1 or Level 2) trigger requirement. The CDF trigger system implements about 100 degenerate trigger paths. An event will be accepted if it passes the requirements of any one of these paths. The trigger paths used in this dissertation are the “two-track” and “lepton+SVT” (ℓ +SVT) trigger paths, which are described in detail in Chapter 4. We discuss the general features of the CDF trigger system in the following with an emphasis on trigger components relevant to this dissertation. A block diagram of the different trigger paths at Level 1 and Level 2 is shown in Figure 2.12.

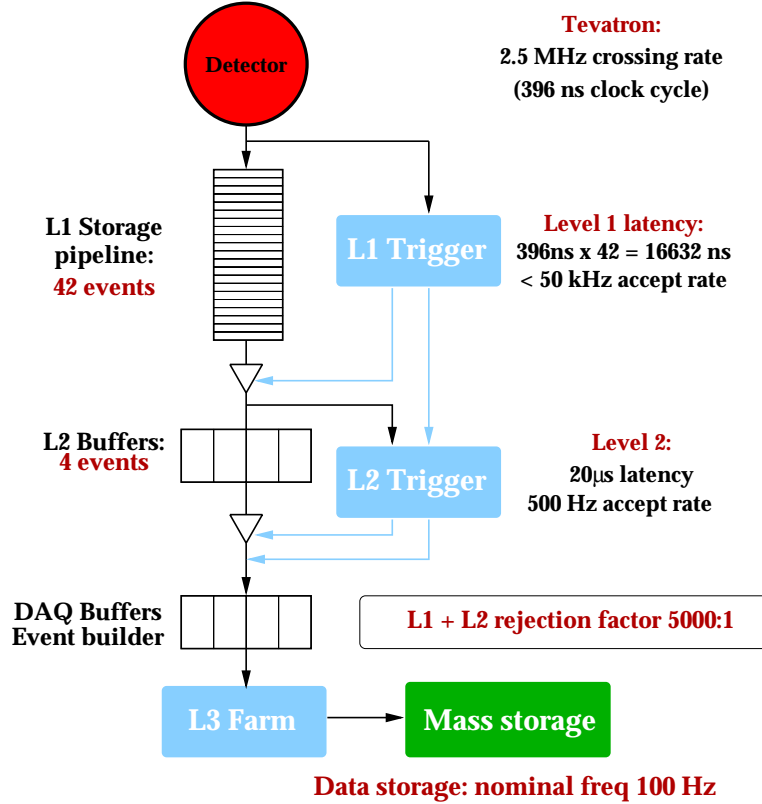


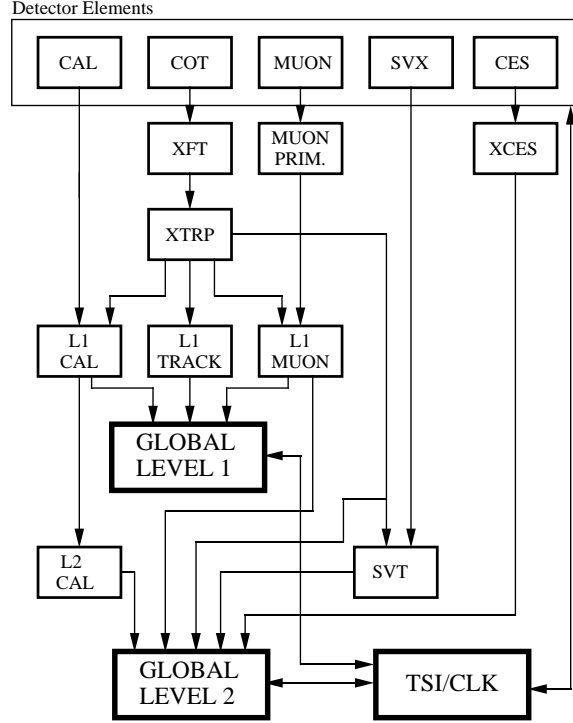
Figure 2.11: Diagram of the CDF-II trigger system.

2.8.1 Level 1 Trigger

The Level 1 trigger operates on every beam crossing and uses custom designed hardware to find physics objects based on a subset of the detector information that includes charged tracks, calorimeter and muon information. At Level 1, track reconstruction is done by the eXtremely Fast Tracker (XFT) [47]. This device examines the COT hits from the four axial superlayers and provides r - φ tracking information. It reports the measurement of the track p_T and φ_6 , the angle of the transverse momentum at the sixth superlayer of the COT, which is located 106 cm radially from the beamline. Based on pre-defined patterns of COT hits, the XFT is capable of recognizing track segments for tracks with $p_T > 1.5$ GeV/ c . It subdivides the COT into azimuthal sections of 1.25° each and places a track into a given section based on its φ_6 angle. The XFT system has high efficiency ($> 90\%$), good transverse momentum resolution, $\delta p_T/p_T = 0.016 p_T$, and pointing resolution, $\delta \varphi_0 = 0.005$ radians, where φ_0 is the azimuthal angle of the track measured at the beamline ($r = 0$).

Within the Level 1 trigger system, tracks found by the XFT are then passed to the Extrapolation System (XTRP) which processes and distributes the tracking information to other trigger elements. Tracks are matched with calorimeter clusters

RUN II TRIGGER SYSTEM



PJW 9/23/96

Figure 2.12: Diagram of the different trigger paths at Level 1 and Level 2.

to identify electrons and with stubs found in the muon detectors to identify muons. Events are also selected at Level 1 based upon charged tracking information only. All the Level 1 objects are sent to the “Global Level 1” system for processing, where they are combined with logical AND and OR gates to form Level 1 triggers. The Global Level 1 has a capacity to accommodate 64 different Level 1 triggers.

2.8.2 Level 2 Trigger

The Level 2 trigger also operates using custom made hardware and reduces the ~ 25 kHz Level 1 accept rate to ~ 500 Hz. At Level 2, information from the silicon detector and the calorimeter shower maximum detectors supplement the information available from the Level 1 trigger. Coarse tracking information from the XFT at Level 1 is combined with high precision SVX-II cluster information by the Silicon Vertex Trigger (SVT) [48]. The goal of the second level of the trigger is to obtain a precise measurement of the track d_0 , and improved measurements of p_T and φ_0 .

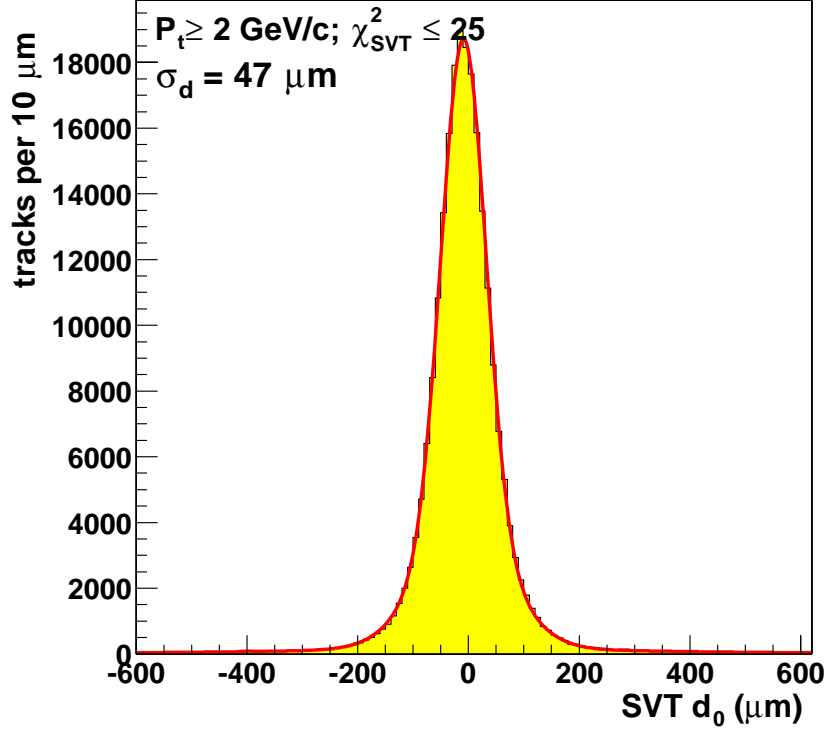


Figure 2.13: SVT impact parameter resolution.

As mentioned in Section 2.4.1, the SVX-II is segmented into 12 wedges in φ and three mechanical barrels along the z -axis. The SVT makes use of this symmetry and performs tracking separately for each wedge and barrel. An SVT track starts with a two dimensional XFT “seed”. The XFT measurement is extrapolated into the SVX-II, forming a “road”. Clusters of charge on the inner four r - φ layers of the given wedge have to be found inside this road. The silicon cluster information and the XFT segment information are fed into a linearized fitter which returns the measurements of p_t , φ_0 and d_0 for the track. The availability of the track impact parameter (d_0) made possible by SVT enables selection of events based upon displaced tracks arising from the decay of weakly decaying particles. Displaced tracks are those which have an impact parameter with respect to the primary $p\bar{p}$ vertex that is inconsistent with having originated from the primary vertex. The impact parameter resolution of the SVT, shown in Figure 2.13, is approximately $\sigma_d \sim 50 \mu\text{m}$. This is a combination of the intrinsic impact parameter resolution of the SVT measurement, and the transverse intensity profile of the $p\bar{p}$ interaction region. The region profile is roughly circular in the transverse plane and can be approximated by a Gaussian distribution with $\sigma \sim 32 \mu\text{m}$. The intrinsic SVT resolution is obtained by subtracting in quadrature the beamline width from the resolution of the SVT d_0 distribution.

To improve electron identification in the Level 2 trigger, information from the elec-

tromagnetic shower maximum detectors (CES) are incorporated in order to perform precise matching between the electron track and calorimeter cluster. This matching drastically reduces the electron fake rate.

2.8.3 Level 3 Trigger

The third level of the trigger system is implemented as a PC farm [49]. Every CPU in the farm provides a processing slot for one event. With roughly 500 CPU's, and a input rate of roughly 500 Hz, this allocates approximately 1 second for event reconstruction and to reach a trigger decision. Events passing the Level 2 trigger are processed by Level 3 running a speed-optimized version of the full event reconstruction, including COT and silicon tracking. The Level 3 selection criteria are similar to the Level 2 criteria. With the additional information and improved resolution provided by the full event reconstruction, the Level 3 confirmation of the Level 1 and Level 2 selection quantities significantly reduces background. Level 3 accepts events from Level 2 at ~ 500 Hz. Events which satisfy the Level 3 selection criteria are written to mass storage at a rate of ~ 100 Hz for subsequent analysis.

2.9 Monte Carlo Simulation

The Monte Carlo simulation enables us to study the behavior of signal and background events in the detector and helps us understand the effects of various biases introduced due to triggering, reconstruction, and analysis requirements. It plays a very important role in this analysis. The Monte Carlo simulation is used to optimize the signal selection requirements, study shapes for signal and backgrounds in various observables, obtain trigger and reconstruction efficiencies for the signal and also to develop and evaluate various flavor tagging algorithms. Same side kaon tagging (see Section 6.2), in particular, utilizes the predicted flavor tag effectiveness based on Monte Carlo simulation.

We use the program **BGenerator** [50] to generate single B mesons in the cases where we are only interested in studying the decay properties of the B hadrons. The production mechanism and fragmentation processes are not relevant in this case as long as they do not influence the properties of the B mesons that we seek to study and the B meson momentum spectrum is correctly reproduced by the simulation. We use the inclusive B meson momentum spectrum measured by CDF [51] as an input to generate B mesons. B meson decays are simulated with **EvtGen** [52], a package which has been extensively tuned by the BaBar and Belle experiments. It contains information about the decay properties of various B and D mesons. We study the accuracy of the Monte Carlo by comparing a large number of kinematic distributions and we find good level of agreement between data and Monte Carlo.

In the case of flavor tagging studies, we use the **PYTHIA** program [53]. **PYTHIA** simulates a complete $p\bar{p}$ interaction: the $b\bar{b}$ pair, the hadronization products, and

the remaining beam fragments from the $p\bar{p}$ scattering referred to as the underlying event. For the simulation of the underlying event, we use “tune A” [54]. Leading and next-to-leading $b\bar{b}$ production mechanisms, namely flavor creation, flavor excitation and gluon splitting have been included [55, 56]. A fraction of 20% of all B^- and \bar{B}^0 mesons originate from B^{**} decays in these PYTHIA Monte Carlo samples, relevant to the same side tagging algorithms discussed in Section 6.2.

After generation and subsequent decay of the simulated particles, the response of the detector is modeled via the CDF-II detector simulation [57]. At this step, the detector geometry and behavior of detector material is simulated using version 3 of the GEANT package [58]. The trigger logic is also emulated and applied to the simulated events. The output of the simulation mimics the structure of the real collision data. Simulated events are processed with the same reconstruction program as the data. After trigger and reconstruction requirements, the Monte Carlo samples typically contain several thousands of reconstructed signal decays, which is several times more than the signal candidates found in data.

Chapter 3

Particle Identification

Each collision inside the CDF detector produces numerous particles that upon production traverse through the various subdetector systems simultaneously. Depending on each particle's type, mass, energy, and lifetime, it could be detected by one or more of the subdetectors. Therefore, measurements from multiple subdetectors are combined for better identification of the particles. In this chapter, we describe the techniques utilized to identify different species of particles. We are mostly concerned with identifying kaons (for better signal selection and initial state flavor tagging) and leptons. We first describe the method used to separate kaons and pions, followed by a description of our lepton (e , μ) identification techniques.

3.1 Kaon and Pion Identification

Identification of charged kaons is an essential ingredient in both B_s^0 meson reconstruction (see Section 4.6) and initial state flavor tagging (see Section 6.2). Charged particles can be distinguished by combining the measurement of the specific ionization per unit track length in the COT (dE/dx) and their time-of-flight measured with the TOF system.

3.1.1 Particle Identification Using dE/dx

When a charged particle traverses the gas volume of a drift chamber, it leaves a trail of ionization along its flight path. Its energy loss in the chamber is proportional to the amount of ionization. The average total energy loss per unit length of a particle with charge q and velocity βc can be described by the Bethe-Bloch formula [13]:

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi N e^4}{m c^2 \beta^2} q^2 \left[\ln \frac{2m c^2 \beta^2 \gamma^2}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right], \quad (3.1)$$

where N is the number density of electrons in the medium, m is the mass of the electron, e is the charge of the electron, I is the mean excitation energy of the atoms

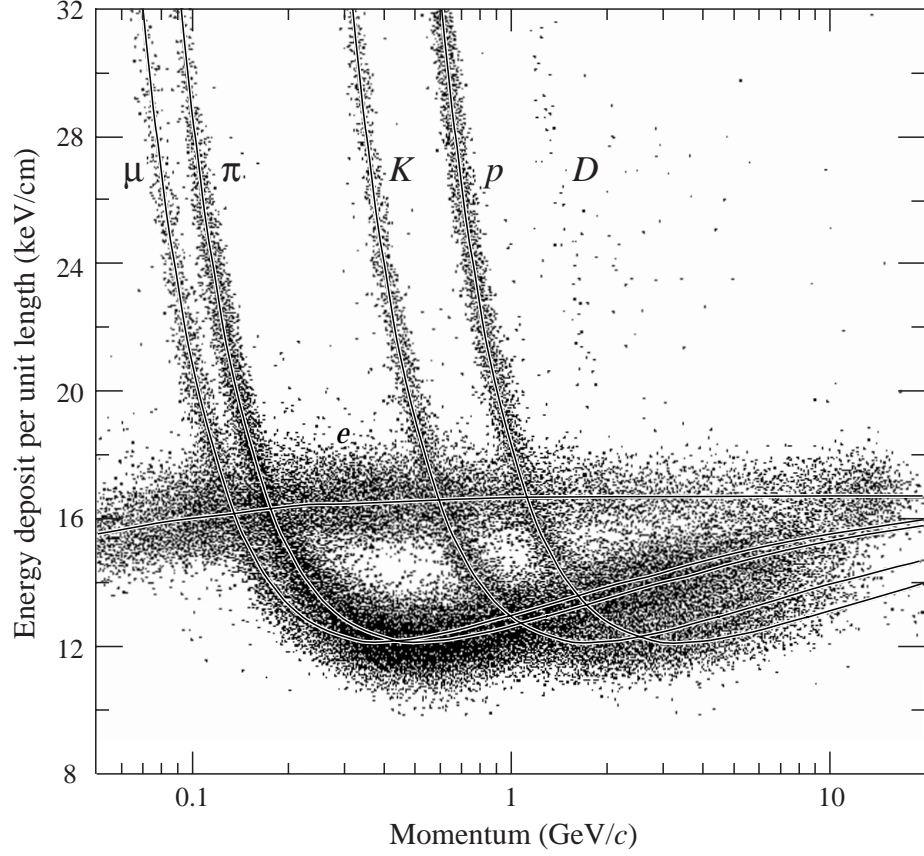


Figure 3.1: dE/dx in the 8.5 atm $Ar - CH_4$ as a function of the track momentum for several particles [13].

in the medium, $\delta(\beta\gamma)$ is the relativistic correction of the density effect at high $\beta\gamma$. Equation (3.1) shows that for a drift chamber with fixed gas properties, dE/dx only depends on the velocity $\beta\gamma$ for a certain particle. Therefore, when combined with a measurement of particle momentum, dE/dx can be used to distinguish particles of different mass. As shown in Figure 3.1, stable particles like electrons, muons, pions, kaons and protons exhibit different dE/dx depending on their momenta.

In the case of CDF, the COT was originally designed for fast and accurate tracking as is evident in the small size of the drift cells in the COT. This reduces the number of charge clusters which can be collected by each wire and leads to an increase in statistical fluctuations. Also, the COT is kept at a pressure of 1 atm, much lower than the drift chambers optimized for dE/dx measurements. Despite these limitations, a meticulous calibration of the dE/dx response in the COT has been carried out which results into excellent separation between various particle species produced in the $p\bar{p}$ collisions.

The COT sense wires collect the charge ionized by the incident charged particle and are read out by the ASDQ front-end electronics. The input analog signal is digitized by ASDQ and the output is a digital pulse. The leading edge gives the arrival time information and the pulse width is related to the amount of charge collected by the wire. The 96 sense wires in the COT provide upto 96 samples of energy loss measurements. Because of the large Landau tail in the distribution of the energy loss measurements, the highest 20% of the measured charge values are discarded and an average is calculated from the remaining sample. This “80% truncated mean” is then taken as the best estimator of the track dE/dx . In the case of the COT, the following empirical variant of the Bethe-Bloch Equation (3.1) better models the average energy loss

$$\left\langle \frac{dE}{dx} \right\rangle = \frac{1}{\beta^2} \left(c_1 \ln \frac{\beta\gamma}{\beta\gamma + b} + c_0 \right) + a_1(\beta - 1) + a_2(\beta - 1)^2 + b. \quad (3.2)$$

In Equation (3.2), describing the so called “universal curve”, the constants c_0 , c_1 , a_1 , a_2 and b are derived from data by plotting $\langle dE/dx \rangle$ as a function of $\beta\gamma$ using different particle species. At CDF, the dE/dx response and its resolution functions have been studied in various $\beta\gamma$ ranges on high statistics and pure samples of electrons, muons, pions, kaons and protons. The calibration samples are obtained from $D^{*+} \rightarrow D^0\pi^+$, $D^0 \rightarrow K^-\pi^+$, $\Lambda^0 \rightarrow p\pi^-$, $J/\psi \rightarrow l^+l^-$ decays and conversion electrons. The measured $\langle dE/dx \rangle$ is displayed as a function of the particle $\beta\gamma$ (universal curve) in Figure 3.2 (left).

For each particle track under consideration, we define the following variable:

$$Z(i) = \ln \left(\frac{dE/dx_{\text{meas}}}{dE/dx_{\text{exp}}(i)} \right), \quad i = e, \mu, \pi, K, p.$$

Here dE/dx_{meas} is the measured dE/dx , and $dE/dx_{\text{exp}}(i)$ represents the expected $\langle dE/dx \rangle$ for the assumed electron, muon, pion, kaon and proton hypotheses extracted from the universal curve. For pure samples of particles, the variable Z is described by a single Gaussian distribution, and shows a resolution of typically 4 – 5%. Figure 3.2 (right) shows the Z distribution for muons with $p_T > 1.5 \text{ GeV}/c$. Using a Gaussian fit to describe the distribution, we extract the resolution σ_Z of Z . σ_Z has been calibrated as a function of the number of sense wire measurements used in calculating the 80% truncated mean for the track and as a function of the track momentum and its direction. Calibrations are derived for separate data-taking periods taking into account observed variations in the operational conditions of the COT [59].

The dE/dx performance in terms of discriminating power between various particle species is shown in Figure 3.3. The dE/dx measurement provides a separation equivalent to 1.4σ for $p > 2 \text{ GeV}/c$ between pions and kaons, and equivalent to 3.0σ at $p = 1.5 \text{ GeV}/c$ (decreasing to $\sim 1\sigma$ at $p \sim 10 \text{ GeV}/c$) between pions and electrons. The efficiency for associating dE/dx information with a charged track reconstructed in the COT in the pseudorapidity range $-1 < \eta < 1$ and with transverse momentum $p_T > 400 \text{ MeV}/c$ is almost 100%. The variables $Z(i)$ are normalized to obtain the

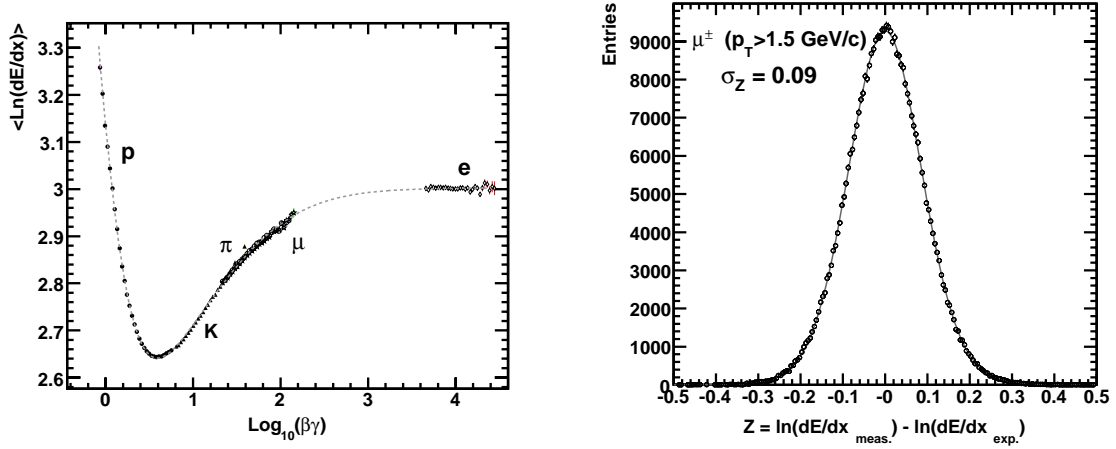


Figure 3.2: Average $\ln(dE/dx)$ versus particle $\beta\gamma$ as measured for various particle species at CDF (left). dE/dx resolution measured for muons with $p > 1.5 \text{ GeV}/c$ (right). The measured width of the $Z(\mu) = \ln(dE/dx_{\text{meas}}) - \ln(dE/dx_{\text{exp}}(\mu))$ distribution corresponds to a resolution of about 4%.

dE/dx probability density distributions $\mathcal{P}_{dE/dx}(i)$ ($i = e, \mu, \pi, K, p$) for all particle species.

3.1.2 Particle Identification Using TOF

For momenta smaller than $p < 1.5 \text{ GeV}/c$, excellent K/π separation is provided by the TOF system, complimentary to the identification with dE/dx for higher momentum, as displayed in Figure 3.4. The TOF system provides better than 2σ K - π separation for $p < 1.4 \text{ GeV}/c$. Particle identification utilizing the TOF detector, described in Section 3.1.2, is performed by measuring the arrival time t of a particle at a TOF scintillator bar with respect to the $p\bar{p}$ interaction time, t_0 . The mass of the incident particle m can then be determined from the particle momentum p and the path length L as measured in the COT, using the following relation [39]:

$$m = \frac{p}{c} \sqrt{\frac{c^2 t^2}{L^2} - 1}, \quad (3.3)$$

where t is referred to as the “time-of-flight” of the incident particle. Two different sources of uncertainty contribute to the time-of-flight resolution: the instrumental (detector and calibration) resolution and the determination of the event production time t_0 .

The operation of a Time of Flight detector in the CDF environment is unique in that it is not possible to synchronize an accelerator timing signal with the respective $p\bar{p}$ interaction time. This is a consequence of the longitudinal spread of the interaction

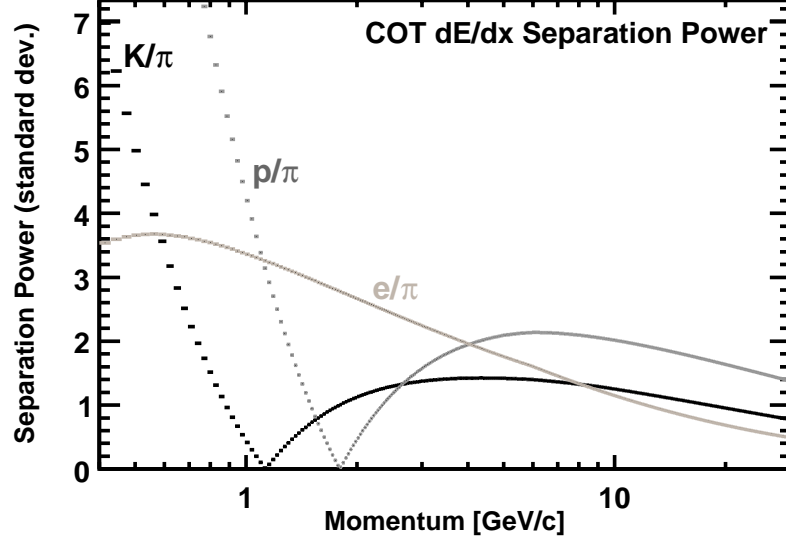


Figure 3.3: Measured performance of the dE/dx particle identification. The vertical axis shows separation power in units of standard deviations versus particle momentum. The K - π separation is shown in black, the p - π separation is displayed in gray, and the e - π separation is depicted in light gray.

region of ~ 30 cm, which yields interactions spread out in time by a few nanoseconds. However, by combining the time-of-flight measurements for all the particles in an event, the value of t_0 can be estimated by assuming that their composition, in terms of stable particles, is known. Using this procedure, CDF obtains a typical t_0 resolution of 50 ps for $b\bar{b}$ events.

To determine the TOF instrumental time resolution, a large sample of $J/\psi \rightarrow \mu^+\mu^-$ decays is studied and the difference (Δ_{TOF}) between the time-of-flight measured for the two muons is investigated. This is particularly advantageous as the event t_0 cancels in the difference. Figure 3.5 shows the distribution of Δ_{TOF} between the two muons. The instrumental resolution can be described by a double Gaussian distribution, with a narrow Gaussian that has a width that varies from 140 ps to 170 ps, depending on the scintillator bar and data-taking period, accounting for $\sim 85\%$ of the area of the resolution function. The broad Gaussian is due to the tracks associated with incorrect TOF information. It has a width of several hundred picoseconds. Using the width of the narrow Gaussian distribution, the instrumental TOF time resolution is estimated to be $\sigma_{\text{TOF}} = \sigma_{\Delta_{\text{TOF}}}/\sqrt{2} = 110$ ps.

In a manner similar to the dE/dx particle identification, the difference of the measured and expected time-of-flight for a given particle hypothesis is used to compute the TOF probability distributions $\mathcal{P}_{\text{TOF}}(i)$ ($i = e, \mu, \pi, K, p$), for a given track.

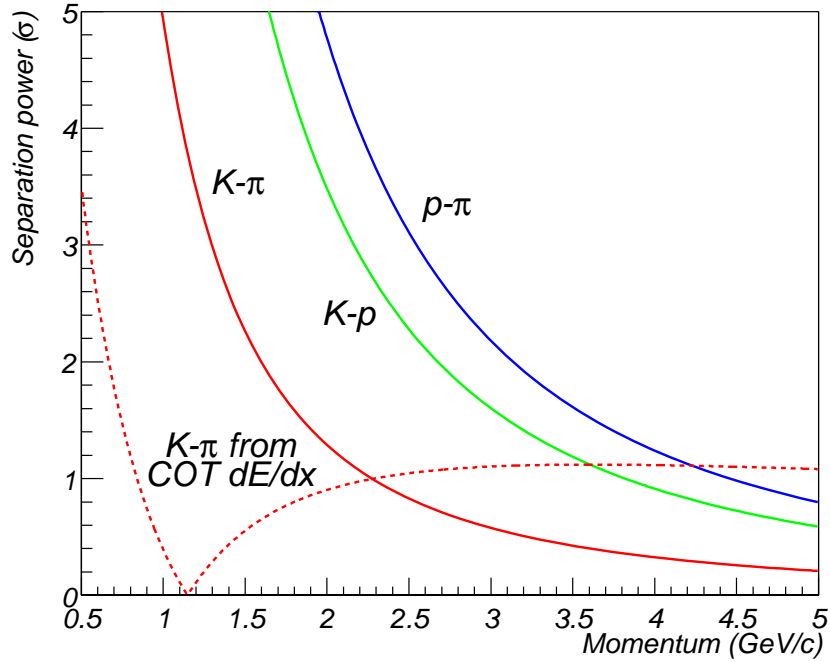


Figure 3.4: Expected performance of the TOF system. The vertical axis shows separation power in units of standard deviations as a function of the particle momentum.

The time-of-flight resolution function has been determined from large samples of high purity pions, kaons and muons obtained from $D^{*+} \rightarrow D^0\pi^+$, $K_S^0 \rightarrow \pi^+\pi^-$ and $J/\psi \rightarrow \mu^+\mu^-$ decays. Multiple particles hitting the same TOF bar tend to distort the time-of-flight measurement, leading to a significant inefficiency in associating the time-of-flight information with a given track. The efficiency for associating TOF information with charged particles in typical $b\bar{b}$ events varies from 50% to 65% and decreases with increasing instantaneous luminosity due to increased track occupancy in the detector.

3.1.3 Combination of dE/dx and TOF

The CDF detector provides two independent sources of charged kaon and pion identification: dE/dx as measured in the COT and time-of-flight measured by the TOF system. The two measurements are complementary. The TOF system is most effective for low momentum tracks ($p < 2$ GeV/c) while dE/dx provides good K - π separation in the intermediate transverse momenta ($p > 2$ GeV/c). By combining information in a single optimized quantity, a K - π separation at least 1.5σ is obtained for track momenta up to 5 GeV/c. At lower momenta, the separation increases.

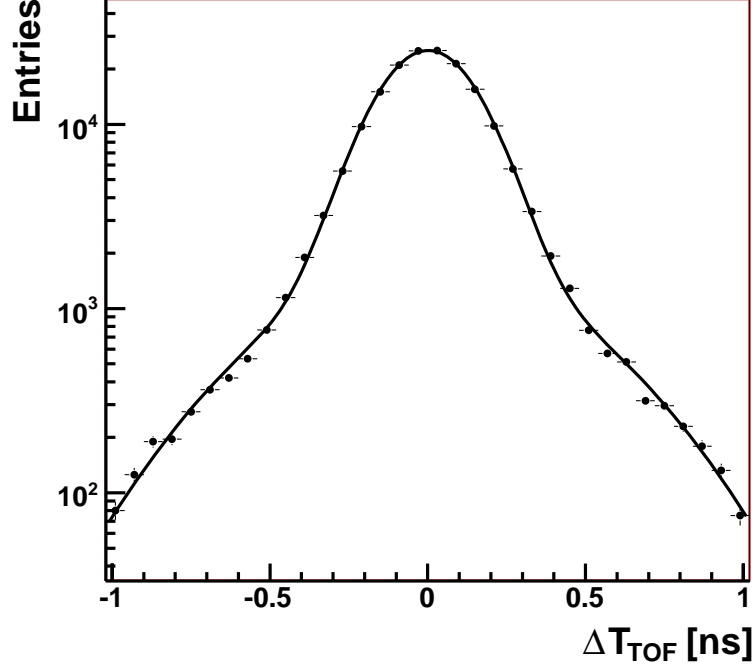


Figure 3.5: Distribution of the time-of-flight difference Δ_{TOF} between the two muons in $J/\psi \rightarrow \mu^+\mu^-$ decays. The instrumental TOF time resolution can be derived from the the width of this distribution using $\sigma_{\text{TOF}} = \sigma_{\Delta_{\text{TOF}}}/\sqrt{2} = 110$ ps.

We combine the dE/dx and time-of-flight measurements for kaon identification in a “combined log likelihood” (CLL) variable, which is defined as:

$$CLL(K) = \log \left(\frac{\mathcal{P}(K)}{f_p \mathcal{P}(p) + f_\pi \mathcal{P}(\pi)} \right), \quad \text{with} \quad \mathcal{P}(i) = \mathcal{P}_{\text{TOF}}(i) \cdot \mathcal{P}_{dE/dx}(i), \quad (3.4)$$

where we choose $f_p = 0.1$ and $f_\pi = 0.9$ as the apriori probabilities for the background composition. $\mathcal{P}_{\text{TOF}}(K, p, \pi)$ and $\mathcal{P}_{dE/dx}(K, p, \pi)$ refer to the probability density distributions for the time-of-flight and dE/dx measurements, as described above. In case the time-of-flight or the dE/dx information is not available for a given track, the corresponding probability is set to $\mathcal{P} = 1$. In Figure 3.6, the K - π separation is shown for the combined particle identification variable as a function of particle momentum as measured in a sample of kaons and pions from $D^{*+} \rightarrow D^0\pi^+$, $D^0 \rightarrow K^-\pi^+$ decays. The achieved separation is compared to the separation power obtained by using only the dE/dx or the time-of-flight measurement. Using the combined kaon identification variable, the K - π separation power is improved by 10% to 25%, depending on the particle momentum, with respect to the separation achieved by using either the TOF or just the dE/dx measurement.

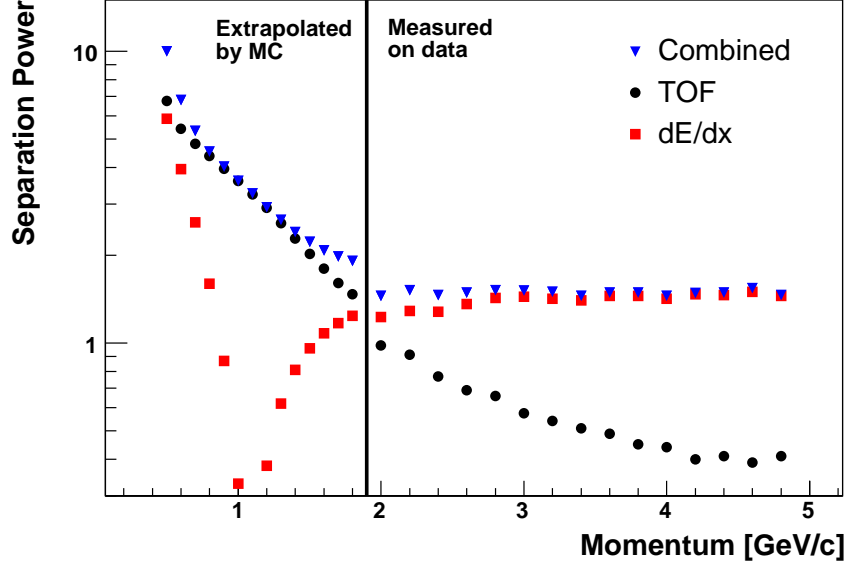


Figure 3.6: Separation power in units of standard deviations versus particle momentum between kaons and pions in $D^{*+} \rightarrow D^0\pi^+$, $D^0 \rightarrow K^-\pi^+$ decays, obtained from TOF (black dots), dE/dx (red dots) and the combined particle identification variable $CLL(K)$ (blue dots).

3.2 Electron Reconstruction and Identification

Electrons are used in this analysis for the reconstruction of semileptonic decays of B mesons and in addition, for opposite side flavor tagging. The reconstruction of electron candidates begins with extrapolating tracks measured in the COT and silicon detectors to the calorimeter. Quantities from each subdetector system are used to separate real electrons from hadrons faking the electron signature, referred to as “fake electrons”. Instead of applying selection requirements on a set of individual electron identification variables, we develop a multivariate approach to combine various electron identification quantities into a global likelihood for higher efficiency and lower mis-identification. To differentiate real electrons from hadrons mimicking the electron signature, we first study the behavior of the discriminating variables for real electrons using electrons from photon conversions $\gamma \rightarrow e^+e^-$. We also select pions from $K_S^0 \rightarrow \pi^+\pi^-$ decays that pass the electron candidate selection criteria to study the behavior of fake electron background from hadrons. The samples of real electrons and background hadrons chosen for this study are selected in an unbiased manner in the electron identification variables.

We reconstruct electron candidates by extrapolating tracks detected in the COT to match shower clusters in the CEM, CES, and CPR (see Section 2.6). First, the

location of the CEM tower that a given track points to, is determined. This tower is referred to as the “seed tower”. Then, a two-tower CEM cluster is formed by adding the nearest neighboring tower in the z direction in the same calorimeter wedge to the seed tower, provided the neighbor’s transverse energy E_T exceeds 100 MeV. Tracks with $p_T > 1.5$ GeV/ c associated with a CEM cluster with a minimum transverse energy E_T of 0.8 GeV are considered as initial electron candidates. Candidates are further required to have a leakage energy from the CEM into the CHA of less than 50%. Finally, the electron candidates have to be fiducial in the CES to ensure that the electromagnetic showers lie within the well-understood regions of the calorimeter. Along the projected trajectory of the tracks, shower clusters in the CES wires and strips, and in the CPR, are also associated. Electron candidates that are not associated with wire or strip clusters in the CES are discarded.

Electron identification utilizes a variety of variables including the energies deposited in the electromagnetic and hadronic calorimeters as well as shower energies, shapes and position measurements in the CES. To further improve the discrimination of electrons from hadron background, we utilize the pulse height in the CPR and include the specific energy loss dE/dx information for electron identification. The particle dE/dx measured in the COT provides good separation between electrons and pions at lower momentum equivalent to 3σ at $p = 1.5$ GeV/ c as described in Section 3.1.1. The electron identification quantities are described in more detail in Section 3.2.2.

3.2.1 Data Samples for Studying Electron Identification

Hadrons can also satisfy the electron candidate requirements above, resulting in fake electron candidates. To improve rejection of fake electrons, we study the behavior of electron identification variables on samples of pure electrons and pure hadrons. We obtain a pure electron sample by reconstructing electrons from photon conversions ($\gamma \rightarrow e^+e^-$). We select pions from $K_S^0 \rightarrow \pi^+\pi^-$ decays to study hadrons faking the electron signature.

Electrons from Conversions

Conversion electrons are reconstructed using ~ 200 pb $^{-1}$ of the 4 GeV and 8 GeV single electron data sets. The recorded events in these data sets are enhanced in electrons, triggered by requiring a minimum transverse momentum of 4 GeV/ c (8 GeV/ c) and at least 4 GeV (8 GeV) of transverse electromagnetic energy in the CEM. The electrons are further required to pass tight electron identification criteria in the trigger, resulting in high purity electron samples.

Photons interact with the material in the CDF detector and convert to e^+e^- pairs. Since photons are massless, tracks originating from a photon conversion are expected to have a very small opening angle at the point of conversion. Furthermore, the trajectory of the two tracks from the photon conversion are parallel to each other

at their distance of closest approach, following the direction of the parent photon. We utilize this property for the reconstruction of conversion candidates by forming pairs of the trigger electron track with other oppositely charged tracks in the event satisfying:

- $|\Delta \cot \theta| < 0.05$ and
- $Sep_{xy} < 0.3$ cm.

Here $|\Delta \cot \theta|$ is the difference in cotangents of the track θ , and Sep_{xy} is defined as the r - φ separation between the track helices at the point of tangency. To ensure good quality tracks, the track pair is required to have at least 10 hits in the axial superlayers, and 10 hits in the stereo superlayers in the COT. The track belonging to a conversion pair with higher (lower) transverse momentum is referred to as the “harder” (“softer”) leg of conversion. The harder leg of each conversion candidate is required to meet the trigger electron criteria for various electron identification quantities, rendering the softer leg unbiased in calorimeter, CES and CPR variables. To avoid biases in the electron identification variables arising from the harder leg being in close vicinity of the softer leg, the conversion electron pair is required to extrapolate to different calorimeter wedges. The softer leg is further required to have at least 3 r - φ hits in the silicon detectors to mimic the track quality criteria imposed on electron candidates for B_s^0 signal reconstruction (see Section 4.3) and opposite side flavor tagging (see Section 6.1.3). This makes the softer leg a suitable source for understanding the behavior of electron identification variables for low p_T electrons.

We select events in which the softer, unbiased leg of conversion has $p_T > 1.5$ GeV/ c . The $\Delta \cot \theta$ distribution for a subset of such conversion candidates ($p_T > 2.0$ GeV/ c) is shown in Figure 3.7. We have $\sim 160,000$ unbiased electrons with $p_T > 2.0$ GeV/ c and a total of $\sim 250,000$ electrons with $p_T > 1.5$ GeV/ c . The $\Delta \cot \theta$ distribution is used for performing background subtraction to extract the distributions for pure electron quantities. For this purpose, the signal region (composed of pure signal and background) is selected with $|\Delta \cot \theta| < 0.01$ and the background regions are chosen between $|\Delta \cot \theta| > 0.02$ and $|\Delta \cot \theta| < 0.03$.

Pions from K_S^0 decays

We use pions from $K_S^0 \rightarrow \pi^+ \pi^-$ decays to study hadrons faking the electron signature. We select events in which at least one of the pion tracks originating from the K_S^0 decay passes the electron identification criteria, as described above. These tracks are thus referred to as fake electrons. The fake electron tracks are required to satisfy $p_T > 1.5$ GeV/ c , while the other K_S^0 daughter pion is allowed to have a transverse momentum as low as 0.4 GeV/ c . The K_S^0 decays are reconstructed using ~ 200 pb $^{-1}$ of data collected via the two-track trigger, described in Section 4.2.1. The two-track trigger does not impose requirements on electron identification variables. Therefore,

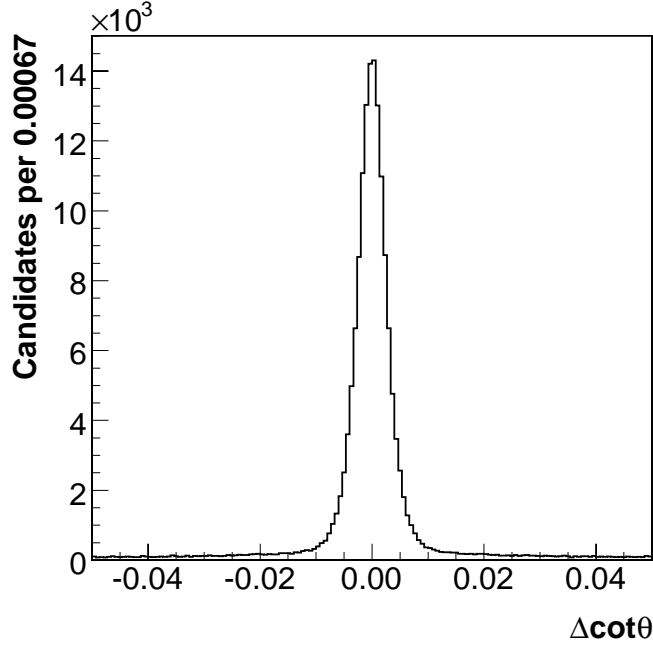


Figure 3.7: $\Delta \cot \theta$ distribution for photon conversion candidates.

pions from $K_S^0 \rightarrow \pi^+\pi^-$ decays reconstructed in the two-track trigger data provide an unbiased source of fake electrons.

The track quality requirements imposed on the pion tracks are the same as for the electron tracks from photon conversions. When reconstructing K_S^0 candidates, we apply a CTVMFT fit to the two pion tracks. To further clean up the K_S^0 signal, we use the following kinematic requirements:

- $L_{xy}/\sigma_{L_{xy}} > 5$, and
- $|d_0(K_S^0)| < 0.01$ cm,

where L_{xy} is the distance in the xy -plane between the primary vertex and the candidate vertex projected along the candidate transverse momentum, $\sigma_{L_{xy}}$ is the error on L_{xy} and $d_0(K_S^0)$ is the impact parameter of the reconstructed K_S^0 .

We obtain a sample of ~ 4 million $K_S^0 \rightarrow \pi^+\pi^-$ decays in which there is at least one pion track with $p_T > 1.5$ GeV/ c satisfying the electron candidate criteria. The K_S^0 signal for fake electrons with $p_T > 2$ GeV/ c is shown in Figure 3.8. The K_S^0 mass distribution is used for performing background subtraction in order to obtain the distributions of electron identification variables for fake electrons. In the K_S^0 mass distribution, the signal region (composed of pure signal and background) is chosen from 0.489 to 0.505 GeV/ c^2 and the background regions are defined from 0.47 to 0.478 GeV/ c^2 and 0.516 to 0.524 GeV/ c^2 .

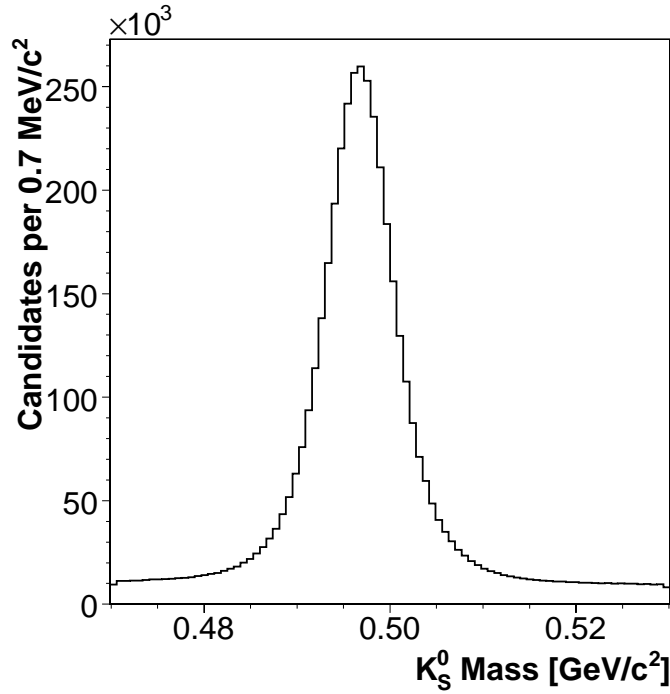


Figure 3.8: K_S^0 mass distribution for pions faking electrons.

3.2.2 Electron Identification Quantities

After selecting real electrons from photon conversions and pions from K_S^0 decays as fake electrons, we use these samples to study the response of the calorimeter, CES, and CPR detectors towards electrons and hadrons. Due to the coarse segmentation in the central calorimeter, the electron identification variables for a given track could be affected by the presence of other tracks in its vicinity. However, not all such tracks reach the same physical location in the calorimeter as the given track due to different p_T and charge. We, therefore, attempt to quantify the “isolation” of an electron candidate track in the calorimeter by defining a local calorimeter isolation variable.

Local Isolation

To define the local isolation of a given track with transverse momentum p_T^{trk} , we begin with an η - φ cone of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2} = 0.7$ around the track. We then extrapolate all the tracks in this cone to the two-tower calorimeter cluster formed by the track. A scalar sum of the p_T of these tracks is performed to obtain $\sum_{0.7} p_T$ and the local isolation is defined as:

$$I = \frac{\sum_{0.7} p_T}{p_T^{trk}}. \quad (3.5)$$

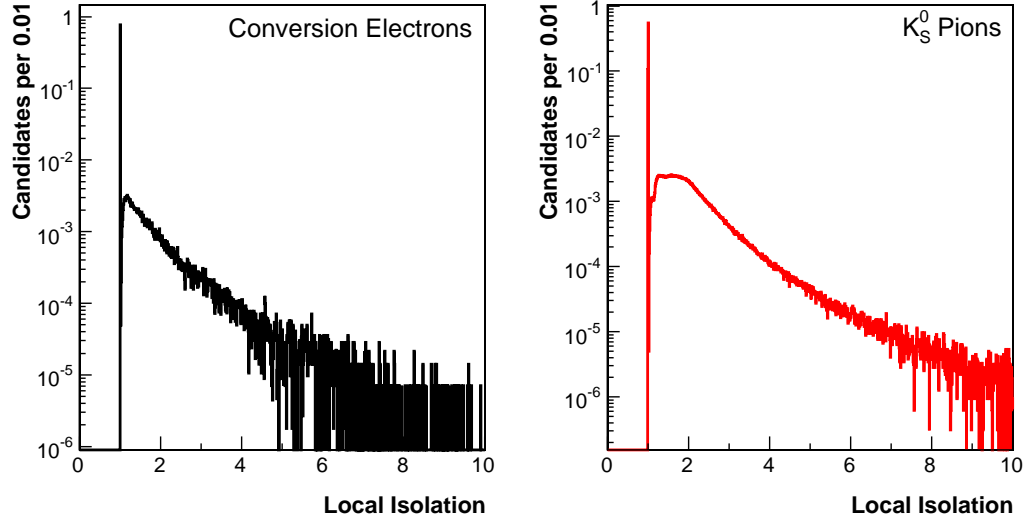


Figure 3.9: Local isolation I for conversion electrons (left) and K_S^0 pions (right).

We define a locally isolated track as the one with $I = 1$, i.e. it is the only track extrapolating to the two-tower cluster. A locally non-isolated track is defined as a track with $I > 1$, i.e. there are other tracks extrapolating to the two-tower cluster. The isolation for electrons from conversions and pions from K_S^0 decays is shown in Figure 3.9. The spike at 1 indicates the locally isolated candidates followed by a long tail consisting of non-isolated candidates. Approximately, 77% of our conversion electrons and 65% of K_S^0 pions are locally isolated.

Some of the electron identification variables are expected to depend on the local calorimeter isolation. Therefore, we split the electron sample from photon conversions and the fake electron sample from K_S^0 decays into isolated and non-isolated subsamples. We use background subtraction in $\Delta \cot \theta$ to obtain a pure sample of electrons, and in the K_S^0 mass distribution to obtain a pure sample of fake electrons. The separation between electrons and hadrons is also expected to improve with increasing p_T of the candidate track, due to better resolution of calorimeter, CES and CPR detectors at higher p_T . To exploit the improvement in discrimination power of the identification variables, we split the electron and pion samples in three p_T bins: 1.5-2.0 GeV/ c , 2.0-4.0 GeV/ c , and > 4.0 GeV/ c . We study the electron identification quantities in the three p_T ranges, separately for the isolated and non-isolated cases. We describe the electron identification variables in the following subsections.

HadE/EmE

HadE/EmE is the ratio of the two-tower energy deposition in the CHA and WHA to the two-tower electromagnetic energy deposited in the CEM. Since the CEM has a

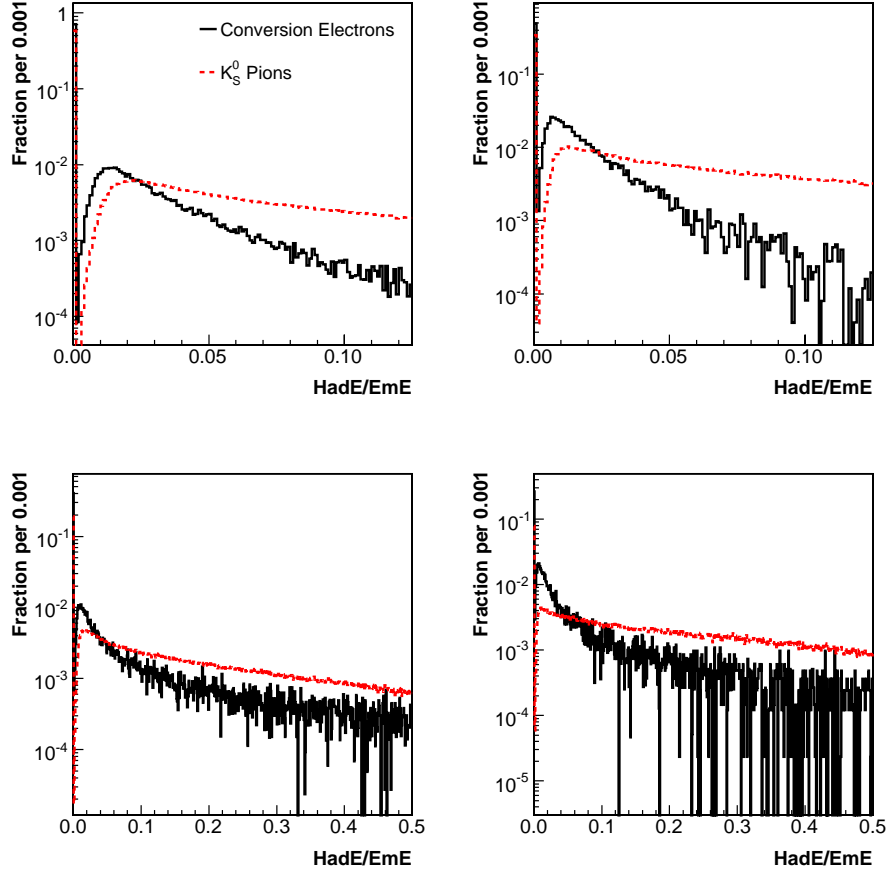


Figure 3.10: HadE/EmE for electrons (black) and pions (red) for isolated (top) and non-isolated (bottom) candidates. The $2.0 < p_T < 4.0$ GeV/ c range is shown on the left and the $p_T > 4.0$ GeV/ c range is displayed on the right.

transverse depth of $18X_0$, the energy leakage for an electron into the CHA and WHA is expected to be less than 1%. Electrons are thus expected to deposit a very small energy fraction in the hadronic calorimeter. The energy in the hadronic calorimeter for real electrons is usually dominated by the energy of the underlying event, and is thus expected to increase with increasing instantaneous luminosity. Hadrons are expected to behave in an opposite manner to the electrons. They deposit a large fraction of their energy in the CHA and WHA.

The HadE/EmE distribution for the locally isolated candidates in the $2.0 < p_T < 4.0$ GeV/ c and $p_T > 4.0$ GeV/ c range is shown in the top-left and top-right of Figure 3.10, respectively. Conversion electrons are displayed in black and pions in red. For non-isolated candidates in the 2.0 to 4.0 GeV/ c range, the distributions of

HadE/EmE are shown in the lower-left plot while for the 4.0 GeV/ c and above range, the corresponding distributions can be found in the lower-right of Figure 3.10. For isolated electrons, we see a big spike around zero in HadE/EmE. For pions, the spike around zero is smaller. The discrimination between electrons and pions is better in the range $p_T > 4.0$ GeV/ c as compared to the 2.0-4.0 GeV/ c range. For non-isolated electrons, HadE/EmE is less discriminating than in the case of isolated candidates due to the presence of additional particles in the vicinity of the electrons. Note, that the horizontal scale extends up to 0.125 for isolated and up to 0.5 for locally non-isolated candidates.

EmE/ p

The EmE/ p is defined as the ratio of the two-tower electromagnetic energy deposited in the CEM to the track momentum. The distributions are shown in Figure 3.11 and follow the same convention as before: isolated candidates on the top and non-isolated ones at the bottom. The 2.0-4.0 GeV/ c range is on the left and $p_T > 4.0$ GeV/ c range on the right. The EmE/ p distributions for pure hadrons are expected to have a minimum ionizing peak around zero. However, such peaks are absent due to the $E_T > 0.8$ GeV/ c and the Had/Em < 0.5 requirements imposed on the electron candidates during the pre-selection.

Since real electrons are expected to deposit their entire energy in the CEM, the ratio EmE/ p is expected to be ~ 1 . However, while traversing the detector material and the magnetic field, electrons lose energy by photon bremsstrahlung. Since the bremsstrahlung photons are emitted in the same direction as the electron, such photons usually end up in the same calorimeter wedge as the electron. Thus, the electromagnetic energy reconstructed in the CEM via the two-tower cluster represents the original electron energy fairly well. On the other hand, the track momentum measured in the COT is mostly after the bremsstrahlung photon emission and hence corresponds to a lower momentum. This causes a positive shift in the value of EmE/ p , as can be seen in Figure 3.11. For isolated electrons EmE/ p is a very strong discriminator against hadrons. The distribution of EmE/ p gets narrower with increasing p_T due to the better CEM resolution at higher energies. Like the HadE/EmE distribution, EmE/ p is also influenced by the presence of other particles in the vicinity of the electrons as well as the energy from the underlying event. For non-isolated electrons, the EmE/ p distribution is much broader and much less discriminating than in the isolated case. This is due to the contribution of energy from other tracks close to the electrons.

CES χ_x^2 and CES χ_z^2

CES χ_x^2 describes the consistency of the measured shower profiles in the CES wires (r - φ view) with the same profile extracted from an electron test beam. Similarly, CES χ_z^2 is a measure of consistency of the shower profile in the CES strips (z view).

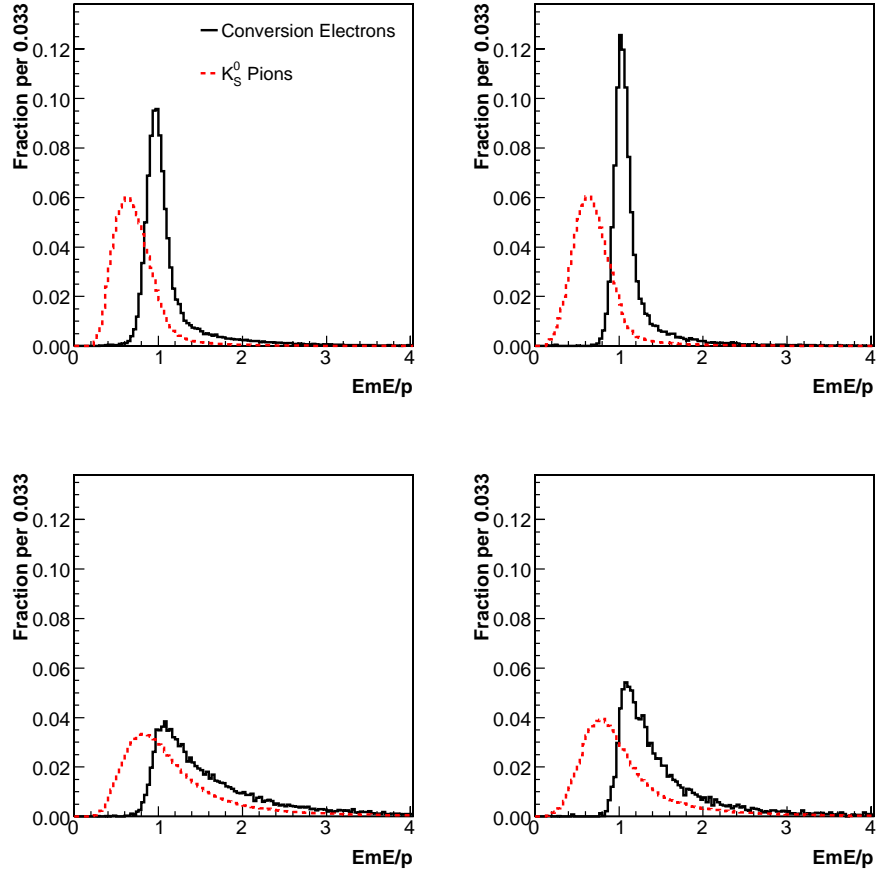


Figure 3.11: EmE/p for electrons (black) and pions (red) for isolated (top) and non-isolated (bottom) candidates. The $2.0 < p_T < 4.0 \text{ GeV}/c$ range is shown on the left and the $p_T > 4.0 \text{ GeV}/c$ range is displayed on the right.

Thus, real electrons are expected to cluster around smaller values of CES χ_x^2 and χ_z^2 . To allow for a comparison that is valid for energies lower than the energy of test beam electrons (10.0 GeV), the measured χ_x^2 and χ_z^2 values are scaled by $(p/10.0)^\alpha$ where p is the track momentum and α is given by

$$\alpha = 0.85 + 0.15 \exp(-p/15.0) - p/1000. \quad (3.6)$$

The distributions for χ_x^2 and χ_z^2 are shown in Figures 3.12 and 3.13, respectively. Both variables are not as strongly discriminating as EmE/p and HadE/EmE . The χ^2 distributions in both wire and strip views look similar for electrons, where most electrons have a low value of χ^2 as expected. On the other hand, pions have a higher value of χ_x^2 and χ_z^2 on average. A notable feature of the plots is that the distributions for non-isolated candidates look very similar to the isolated ones. This is due to the finer segmentation in the CES.

CES $q\Delta X/\sigma_x$

CES ΔX is the distance between the extrapolated COT track position in the CES and the measured CES cluster centroid position in the r - φ plane. This distance is multiplied by the charge of the track to account for the asymmetric tails originating from the bremsstrahlung radiation of electrons. The quantity $q\Delta X$ is further corrected for p_T and η dependent offsets in average $q\Delta X$ (denoted by $\langle q\Delta X \rangle$) for pure electrons, as detailed in Appendix A.1.

Applying the p_T and η dependent corrections, we plot the resolution of the $q\Delta X$ distribution σ_x for pure electrons with respect to p_T (in centimeters). This is shown in Figure 3.14. As expected, the resolution improves with increase in p_T . The dependence of σ_x on p_T is parameterized by a function of the form: $\sigma_x(p_T) = p_0 + p_1/p_T$.

After obtaining the scaling function $\sigma_x(p_T)$ describing the width of the corrected $q\Delta X$ with respect to p_T , we form the ratio $q\Delta X/\sigma_x$ to obtain distributions with constant resolution for pure electrons. These are displayed for electrons and pions in Figure 3.15. For electrons, the $2.0 < p_T < 4.0$ GeV/ c range is combined with the $p_T > 4.0$ GeV/ c range. Separate distributions are prepared for locally isolated and non-isolated electrons. The separation power of $q\Delta X/\sigma_x$ between real electrons and fake electrons is very good, and improves with increasing p_T . In addition, the separation power of $q\Delta X/\sigma_x$ between non-isolated electrons and pions is very similar to the isolated candidates.

CES $\Delta Z/\sigma_z$

The quantity ΔZ in the CES is the distance between the extrapolated COT track position in the CES and the measured cluster centroid position in the CES r - z plane. Unlike the ΔX distribution, the ΔZ distribution is not affected by the charge asymmetry. However, ΔZ is further corrected for p_T and η dependent offsets in average ΔZ (denoted by $\langle \Delta Z \rangle$) for pure electrons, as detailed in Appendix A.2.

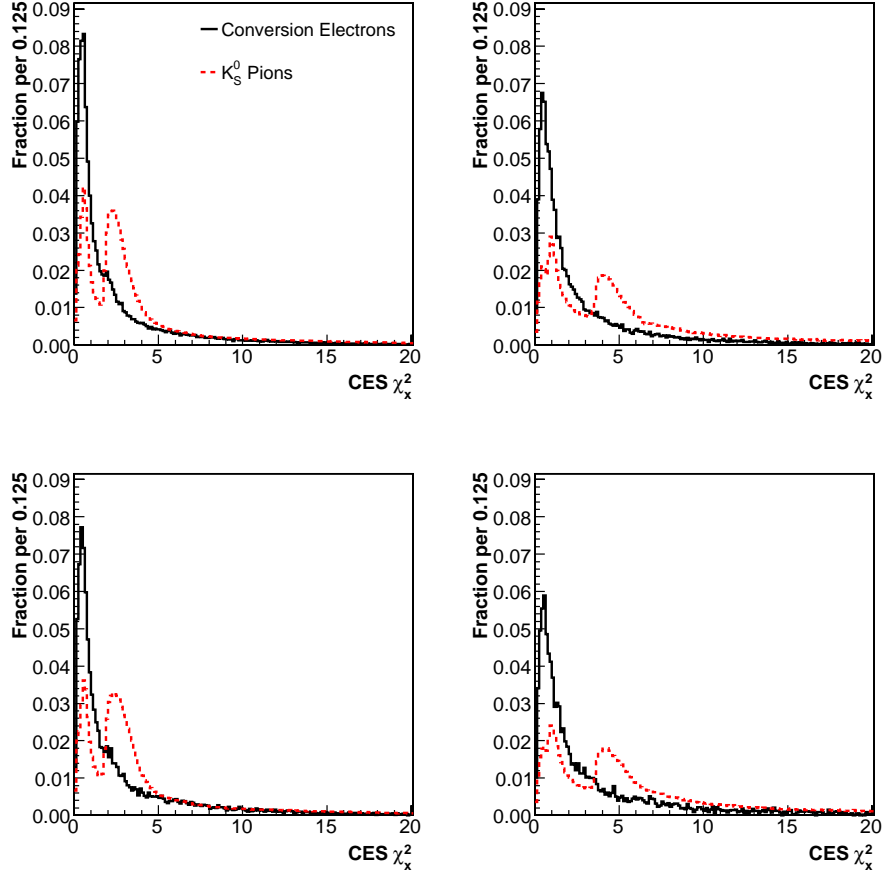


Figure 3.12: CES χ_x^2 for electrons (black) and pions (red) for isolated (top) and non-isolated (bottom) candidates. The $2.0 < p_T < 4.0$ GeV/ c range is shown on the left and the $p_T > 4.0$ GeV/ c range is displayed on the right.

Using the p_T and η dependent corrections, we first plot the width σ_z of the ΔZ distribution for pure electrons with respect to p_T (in centimeters). This is displayed in Figure 3.16. The dependence of σ_z on p_T is parameterized by a function of the form: $\sigma_z(p_T) = p_0 + p_1/p_T$.

After obtaining the scaling function $\sigma_z(p_T)$, we plot $\Delta Z/\sigma_z$ for electrons and pions as shown in Figure 3.17. For electrons, the $2.0 < p_T < 4.0$ GeV/ c range is combined with the $p_T > 4.0$ GeV/ c range. Separate distributions are produced for locally isolated and non-isolated electrons. The discrimination power of $\Delta Z/\sigma_z$ between real and fake electrons is excellent and improves with increasing track p_T , just as for $q\Delta X/\sigma_x$.

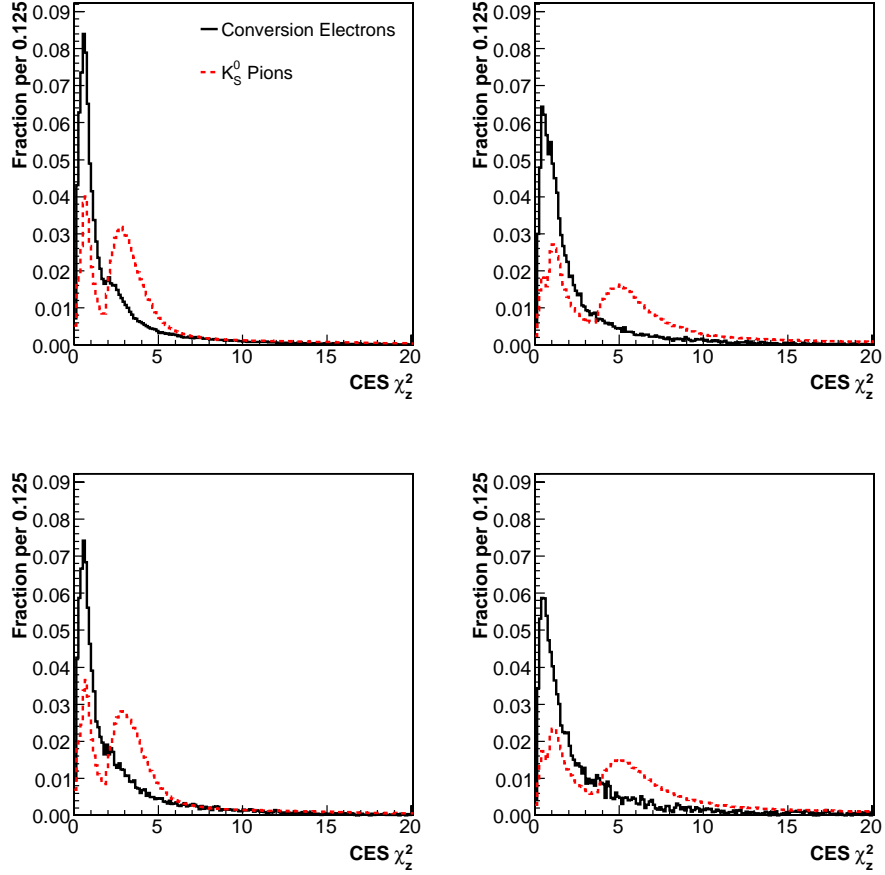


Figure 3.13: CES χ_z^2 for electrons (black) and pions (red) for isolated (top) and non-isolated (bottom) candidates. The $2.0 < p_T < 4.0$ GeV/ c range is shown on the left and the $p_T > 4.0$ GeV/ c range is displayed on the right.

E_{CES}/p^*

E_{CES}/p^* is the wire cluster pulse height measured in the CES, corrected for chamber geometry and cell boundaries, and scaled by $p^* = 10(p/10)^\alpha$, where p is the track momentum and α is defined in Equation (3.6). The distributions for electrons and pions are shown in Figure 3.18. Despite the requirements on the CEM E_T and Had/Em for the electron candidate pre-selection, a large fraction of pions are minimum ionizing in the CES wires. This results in an excellent discrimination power of E_{CES}/p^* between real and fake electrons.

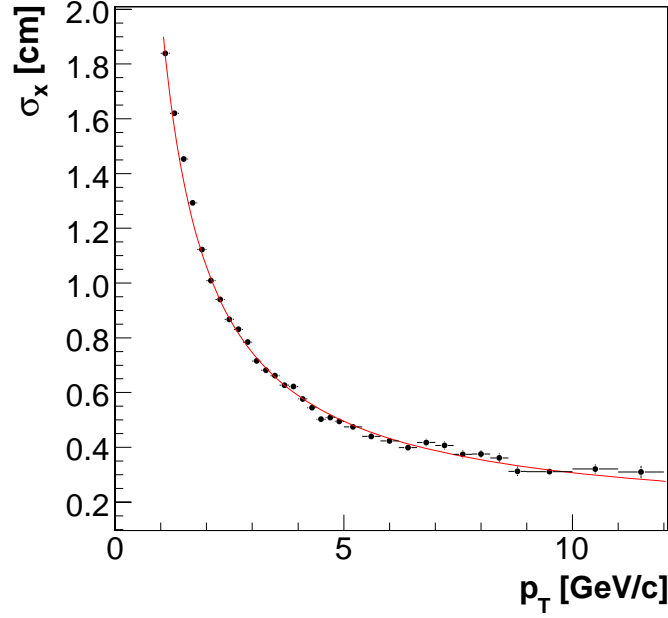


Figure 3.14: Parameterization of the width σ_x of the corrected $q\Delta X$ in the CES vs p_T .

Q_{CPR}

Q_{CPR} is the pulse height in the CPR associated with a track, and corrected for its dependence on track parameters like $\sin \theta$ and p_T . Electron tracks at higher η traverse a longer path length through the solenoid coil as compared to electrons with $\eta \approx 0$ resulting in a θ dependence of the CPR response. Tracks at small $\sin \theta$ interact with more material and produce more secondary particles, leading to higher pulse heights registered in the CPR. In addition, the pulse height is proportional to the transverse momentum of the incident track: tracks at higher transverse momentum tend to deposit more energy in the CPR. These dependences are corrected for and the distributions of corrected Q_{CPR} for electrons and pions are shown in Figure 3.19. Q_{CPR} is a strong discriminator between electrons and pions. A large fraction of the incident pions are minimum ionizing in CPR.

dE/dx

The specific ionization dE/dx is the energy loss measured in COT, as described in Section 3.1.1. To study the dE/dx information for electron identification, the samples of electrons from photon conversions and pions from K_S^0 decays are split into several subsamples based on the momentum and charge of the electron candidate. In each momentum bin, we plot the quantity $Z(e)$ from Equation (3.1.1) normalized to its

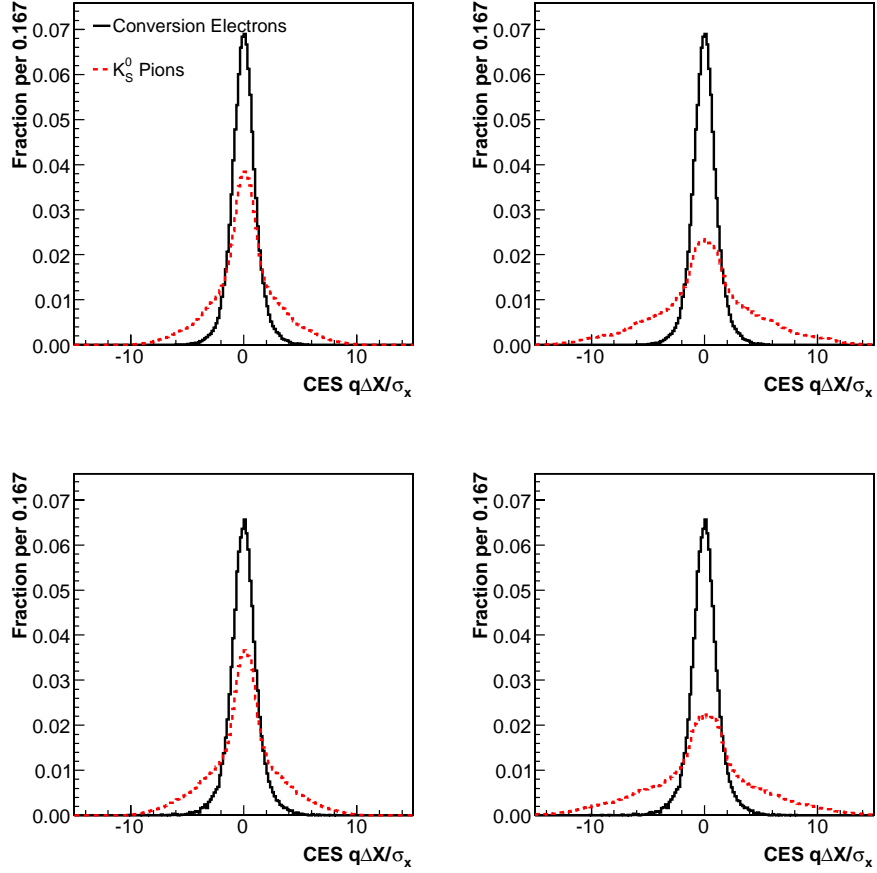


Figure 3.15: CES $q\Delta X/\sigma_x$ for electrons (black) and pions (red) for isolated (top) and non-isolated (bottom) candidates. The $2.0 < p_T < 4.0$ GeV/ c range is shown on the left and the $p_T > 4.0$ GeV/ c range is displayed on the right.

resolution σ_Z , i.e. Z/σ_Z , for the electron hypothesis. An example of the $Z(e)/\sigma_Z$ distribution is shown in Figure 3.20 for e^+ and π^+ candidates in the momentum range 2.3-2.4 GeV/ c . The distributions are described well by a single Gaussian, as detailed in Section 3.1.1.

dE/dx provides a separation of $\sim 3\sigma$ at $p = 1.5$ GeV/ c which diminishes to $\sim 1\sigma$ at $p = 10.0$ GeV/ c , as displayed in Figure 3.3. dE/dx provides strong discrimination between electrons and pions. In particular, its power lies in the fact that its separation gets larger with decreasing momentum, unlike calorimeter, CES and CPR based variables whose separation power degrades with decreasing p_T .

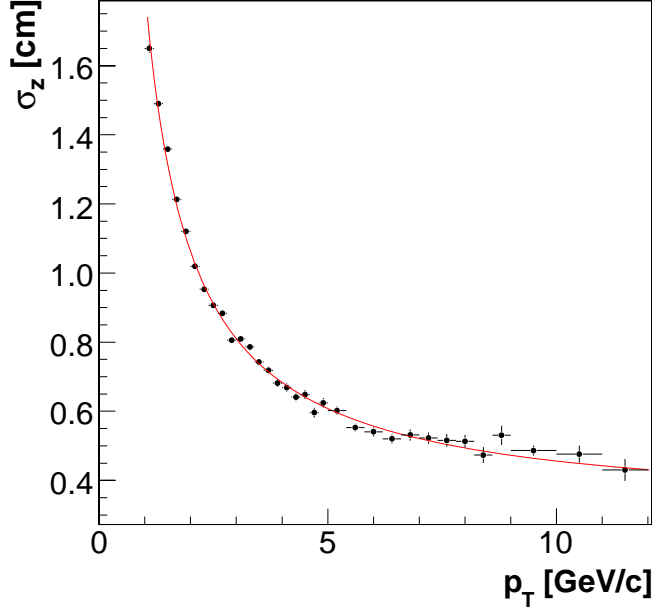


Figure 3.16: Parameterization of the width σ_z of the corrected ΔZ in the CES vs p_T .

3.2.3 Electron Identification through Likelihood Method

To identify real electrons efficiently, we combine the various electron identification variables described above into a global likelihood to maintain good acceptance for real electrons and low efficiency for fakes. The electron likelihood estimates the probability that an electron candidate is indeed a real electron. We describe the likelihood method and its performance in the following.

Description of Likelihood Method

We use the nine variables described in Section 3.2.2 to discriminate real electrons from fake electrons. Using these samples of real electrons and pions faking the electron signature described above, we parameterize the distributions of discriminating variables for both electrons and fakes, using empirical functional forms. We treat the fitted empirical functions as probability distribution functions (PDFs). As an example, Figure 3.21 shows the parameterizations of EmE/p for locally isolated electrons (left) and pions (right) in the $2.0 < p_T < 4.0$ GeV/ c range. Separate templates are used for locally isolated and non-isolated candidates in the three p_T ranges for electrons and pions. When using the dE/dx information, Gaussian PDFs describe the $Z(e)/\sigma_Z$ distributions for electrons and pions in a large number of bins in track momentum. This parameterization is done separately for positive and negative charges.

With this procedure, we obtain nine “signal” PDFs: $S_{\text{EmE}/p}$, $S_{\text{HadE}/\text{EmE}}$, $S_{\chi^2_x}$,

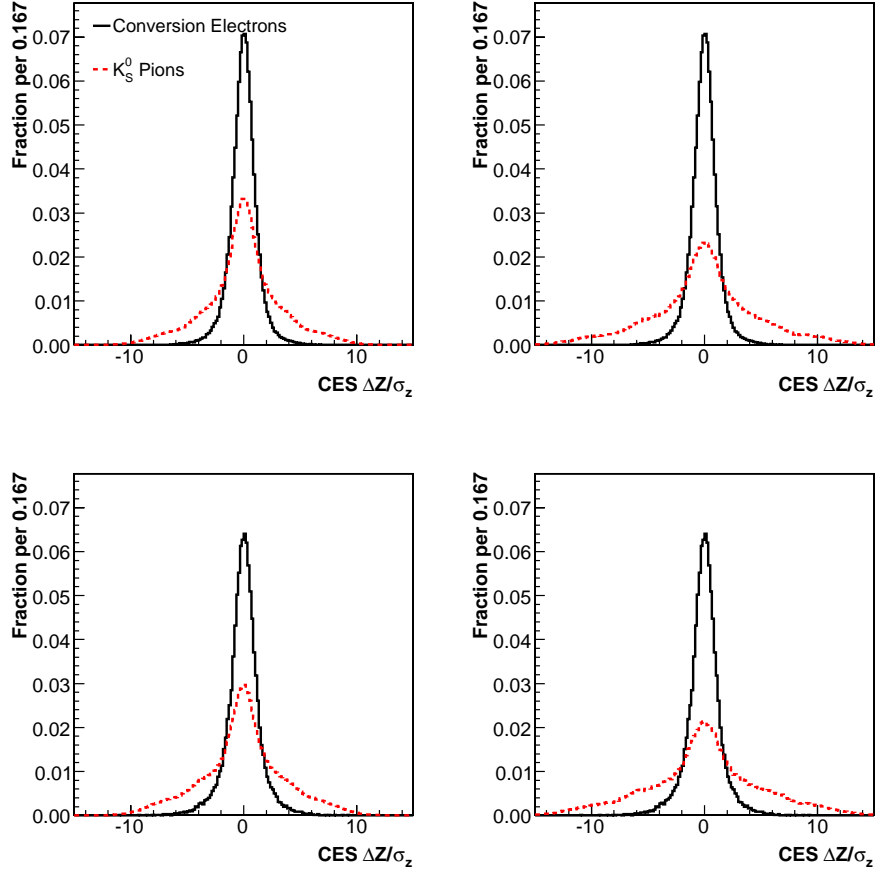


Figure 3.17: CES $\Delta Z/\sigma_z$ for electrons (black) and pions (red) for isolated (top) and non-isolated (bottom) candidates. The $2.0 < p_T < 4.0$ GeV/ c range is shown on the left and the $p_T > 4.0$ GeV/ c range is displayed on the right.

$S_{\chi_z^2}$, $S_{q\Delta X/\sigma_x}$, $S_{\Delta Z/\sigma_z}$, S_{E_{CES}/p^*} , $S_{Q_{CPR}}$, $S_{dE/dx}$ and nine corresponding “background” PDFs: $B_{EmE/p}$, $B_{HadE/EmE}$, $B_{\chi_x^2}$, $B_{\chi_z^2}$, $B_{q\Delta X/\sigma_x}$, $B_{\Delta Z/\sigma_z}$, B_{E_{CES}/p^*} , $B_{Q_{CPR}}$, $B_{dE/dx}$. Assuming all the variables are uncorrelated, the likelihood that an electron candidate is indeed a real electron can be written as

$$S = \prod_{i=1}^9 S_i, \quad (3.7)$$

and the likelihood that the electron candidate is a fake electron is given by

$$B = \prod_{i=1}^9 B_i. \quad (3.8)$$

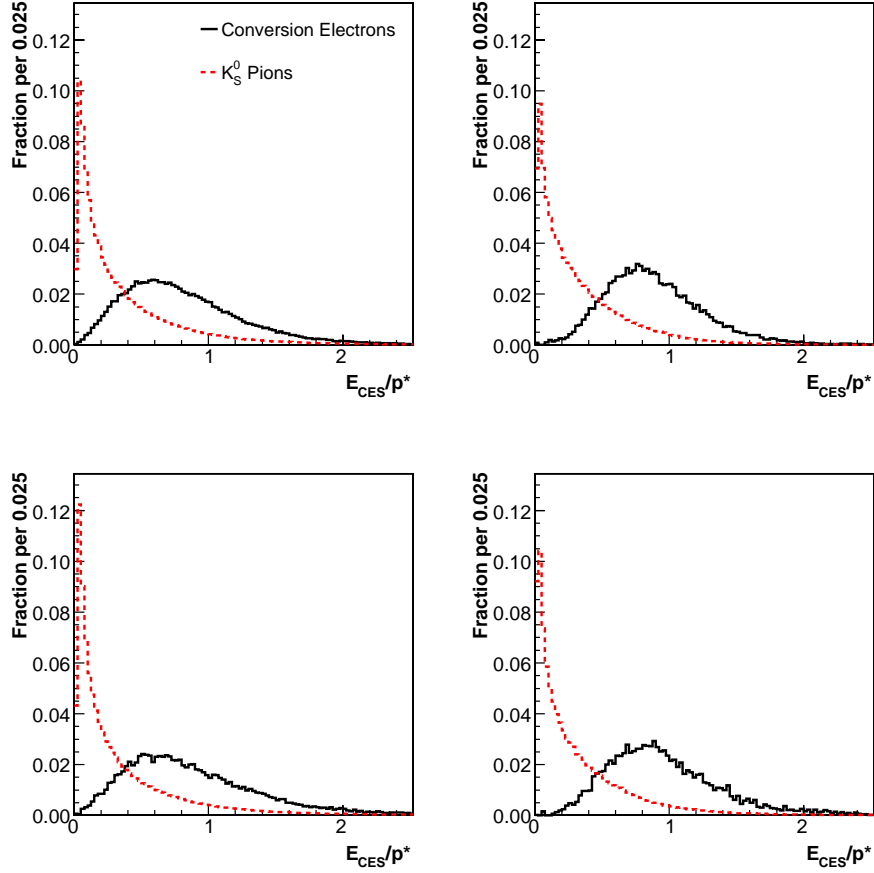


Figure 3.18: E_{CES}/p^* for electrons (black) and pions (red) for isolated (top) and non-isolated (bottom) candidates. The $2.0 < p_T < 4.0$ GeV/ c range is shown on the left and the $p_T > 4.0$ GeV/ c range is displayed on the right.

The global likelihood estimator for a given electron candidate is then defined as the ratio of its joint probability to be an electron to the sum of its joint probabilities to be a real electron or a hadron:

$$\mathcal{L}_e = \frac{S}{S + B}. \quad (3.9)$$

The values for \mathcal{L}_e are bound between zero and one with real electrons occupying the high likelihood region close to unity while fake electrons predominantly populate the low \mathcal{L}_e region close to zero. For defining the electron likelihood, we require the candidate electron to have calorimeter information, be fiducial in the CES detector and be associated with shower clusters in both CES wires and strips. However, the electron candidates are not required to have a cluster in the CPR detector, due to a lower reconstruction efficiency in the CPR detector of $\sim 70\%$. Thus, the electron

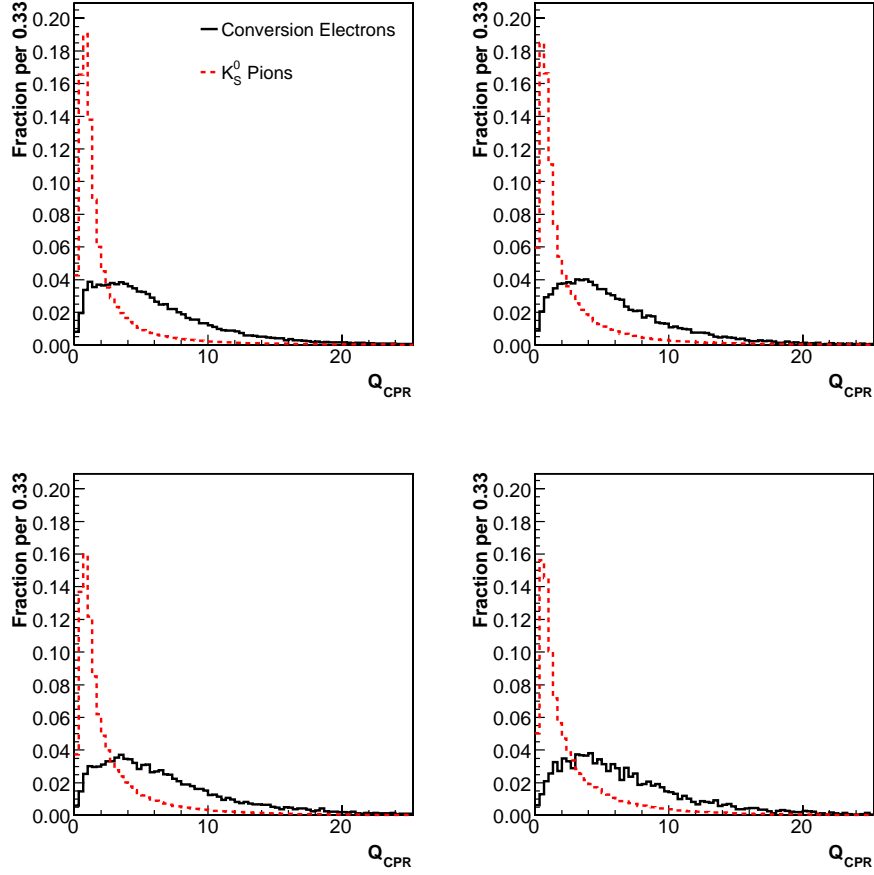


Figure 3.19: Corrected Q_{CPR} for electrons (black) and pions (red) for isolated (top) and non-isolated (bottom) candidates. The $2.0 < p_T < 4.0$ GeV/ c range is shown on the left and the $p_T > 4.0$ GeV/ c range is displayed on the right.

likelihood utilizes nine or eight discriminating variables, depending on whether the CPR information is available or not.

Performance of the Likelihood Estimator

An initial step to test the likelihood algorithm is to study its performance on samples that were used to generate the PDF templates: electrons from photon conversions and pions from K_S^0 . The separation of the likelihood method can qualitatively be characterized by how close pions accumulate near zero and how close electrons cluster towards one.

The left-hand side of Figure 3.22 shows the normalized distributions of \mathcal{L}_e for locally isolated electrons (black) and pions (red). The same is displayed for and

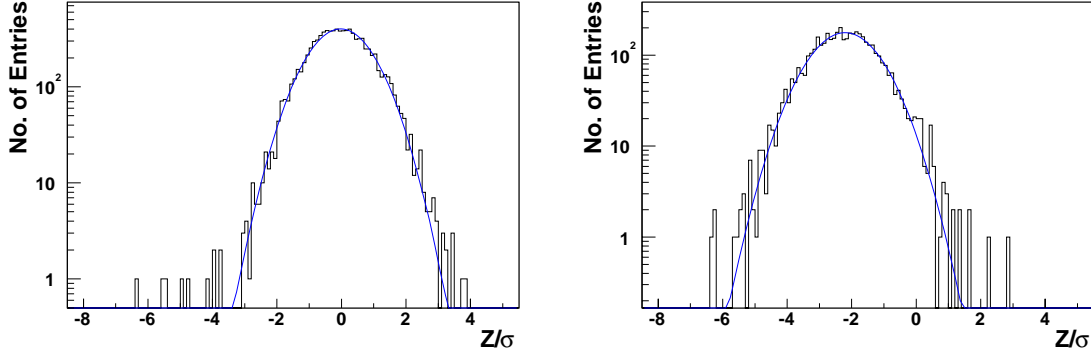


Figure 3.20: dE/dx pull distribution for positron hypothesis for e^+ (left) and π^+ (right) in the momentum range $2.3 \leq p \leq 2.4$ GeV/ c .

$\mathcal{L}_e >$	Efficiency for pions [%]	Efficiency for electrons [%]
0.01	21.0	100.0
0.05	14.5	99.9
0.10	11.1	99.6
0.20	8.5	99.1
0.50	5.0	97.6
0.90	1.9	91.0

Table 3.1: Efficiencies for isolated pure pions and electrons for different requirements on electron likelihood \mathcal{L}_e , relative to the candidates passing the initial electron candidate criteria.

locally non-isolated electrons and pions on the right-hand side of Figure 3.22. As expected, electrons show a large spike at ~ 1 and pions exhibit a large peak at zero. The distribution for isolated and non-isolated candidates look similar.

Tables 3.1 and 3.2 show the efficiencies for pure pions and electrons applying different requirements on \mathcal{L}_e for isolated and non-isolated candidates, respectively. High efficiencies for electrons are achieved while maintaining a large rejection for fakes. The quoted efficiencies are with respect to the candidates passing the initial electron candidate criteria, and thus, the pion efficiencies do not reflect the actual fake rate, which is much lower. The performance of the likelihood is only slightly worse for non-isolated cases, even though EmE/ p and HadE/EmE suffer significant degradation in separation power.

Updates of the Electron Likelihood

Appendix B describes changes to the likelihood templates, which were obtained for the first 360 pb $^{-1}$ of data, taking into account the changes in the detector configuration

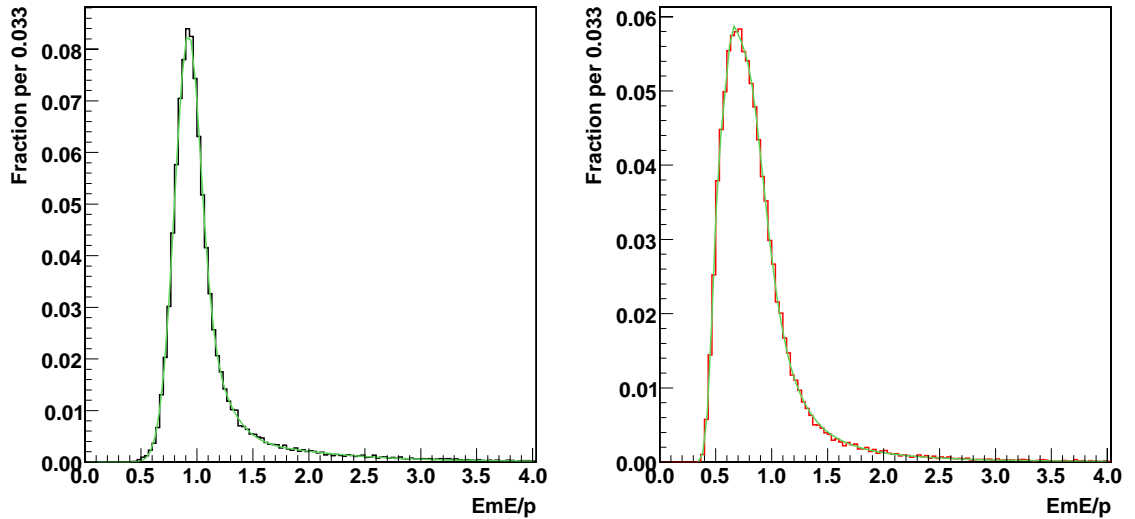


Figure 3.21: Empirical parameterizations of EmE/p for locally isolated electrons (left) and pions (right) in the $2.0 < p_T < 4.0$ GeV/c range.

and electron identification variables in the later ~ 640 pb^{-1} of data utilized in this measurement.

3.3 Muon Identification

Muon candidates are reconstructed by extrapolating tracks measured in the COT and silicon detector to the muon systems, where they are matched to track segments (stubs) reconstructed in the various muon chambers (see Section 2.7). The track-to-stub matching is based on the following observables. The distances in the r - φ (ΔX)

$\mathcal{L}_e >$	Efficiency for pions [%]	Efficiency for electrons [%]
0.01	27.0	100.0
0.05	18.3	99.9
0.10	14.8	99.2
0.20	11.6	98.5
0.50	7.0	96.9
0.90	2.6	90.1

Table 3.2: Efficiencies for non-isolated pions and electrons for different requirements on electron likelihood \mathcal{L}_e , relative to the candidates passing the initial electron candidate criteria.

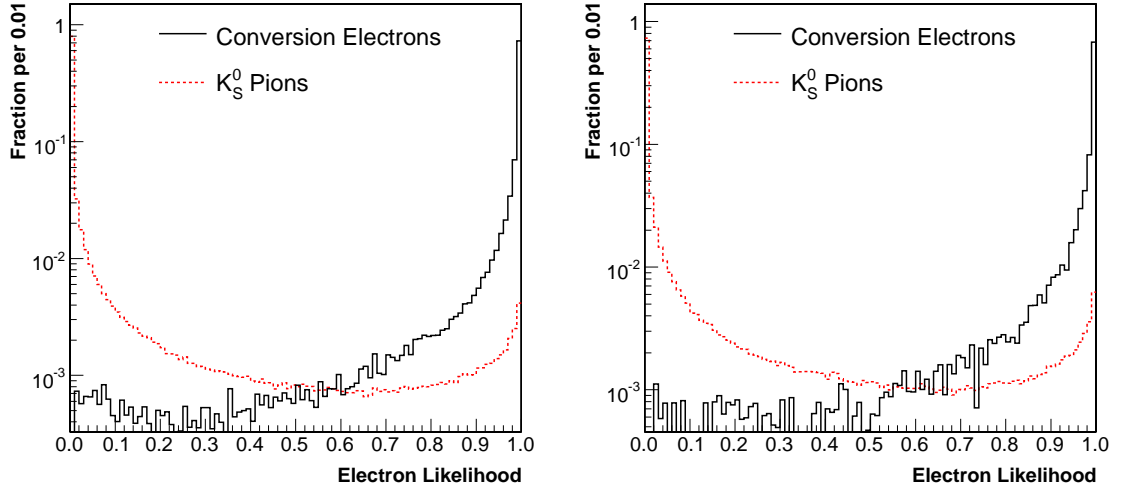


Figure 3.22: Likelihood distributions for locally isolated (left) and locally non-isolated (right) electrons (in black) and pions (in red).

and the longitudinal planes (ΔZ) between the positions of the extrapolated track and the measured stub are calculated. Note, the CMP detectors do not provide a measurement of the longitudinal coordinate. Additionally, the angle $\Delta\varphi$ between the directions of the extrapolated track and the reconstructed stub is determined. The observed deviation between the measured track segment in the muon chambers and the computed extrapolation of the COT track is dominated by multiple scattering in the intermediate detector material which is taken into account when calculating the track to stub matching distance. In addition to these matching variables, the energy deposits in the electromagnetic and hadronic calorimeter towers close to the muon trajectory are used for muon identification. The energies deposited by muons are expected to be typical of a minimum ionizing particle.

Observing a stub in the muon chambers does not guarantee a real muon, because hadrons interacting late in the calorimeters could produce secondary particles that may register hits in the muon chambers. To suppress such hadron backgrounds and efficiently identify real muons, a muon likelihood (\mathcal{L}_μ) is constructed from the calorimeter energies and the muon matching quantities in a manner similar to the one described above. For more details on determining the muon likelihood, see Ref. [60]. To obtain distributions of these variables for real muons, muons are selected from $J/\psi \rightarrow \mu^+\mu^-$ decays. Pions from $K_S^0 \rightarrow \pi^+\pi^-$, kaons from $D^0 \rightarrow K^-\pi^+$ and protons from $\Lambda^0 \rightarrow p\pi^-$ decays with associated stubs in the muon detectors are used to obtain distributions for studying the hadron background. The distributions of the discriminating variables for both muons and backgrounds are parameterized using empirical functions. These parameterizations are again treated as PDFs. Examples of such distributions are given in Figure 3.23. The transverse distance between the

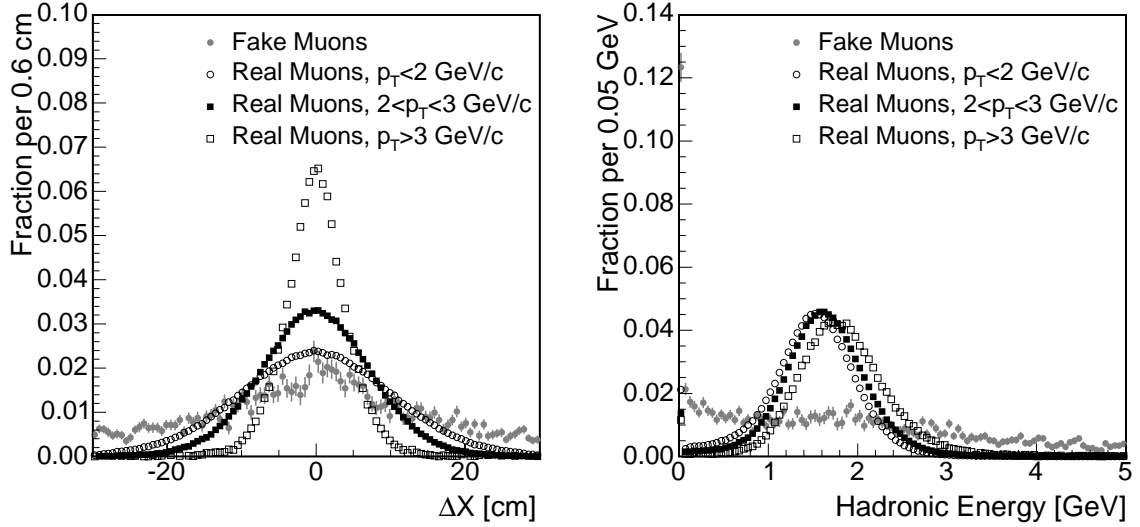


Figure 3.23: Transverse distance between the extrapolated track and muon stub (left) and the hadronic energy (right) for fake and real CMU muons in different momentum ranges.

extrapolated track and muon stub is shown on the left-hand side for fake muons and real muons in different momentum ranges while the hadronic energy is displayed on the right-hand side of Figure 3.23. Calculating $\mathcal{S} = \prod_i S_i$ and $\mathcal{B} = \prod_i B_i$, the muon likelihood is defined as $\mathcal{L}_\mu = \mathcal{S}/(\mathcal{S} + \mathcal{B})$. Figure 3.24 shows such muon likelihood distributions \mathcal{L}_μ for real and fake muons. We find real muons at high likelihood values close to unity while fake muons from hadrons populate the region of small \mathcal{L}_μ values.

The efficiency and background rejection of the muon likelihood depends on the muon detector subsystem. CMUP muon candidates which have stubs in both the CMU and CMP detectors have very low hadron background and consequently, typical muon likelihood requirements imposed on these candidates can be less stringent. The muon likelihood requirements used to select semileptonic B decays are listed in Table 3.3 for various muon subsystems together with the corresponding efficiencies for real and fake muons. The quoted efficiencies are relative to the muon candidates having stubs in the muon chambers, and therefore the efficiencies for fake muons do not reflect the actual fake rate, which is much lower.

3.4 Sources of Leptons and Utilization of Lepton Likelihood

Electron candidates consist of real electrons from heavy flavor decays, electrons from photon conversions ($\gamma \rightarrow e^+e^-$) in the detector material and background from hadrons

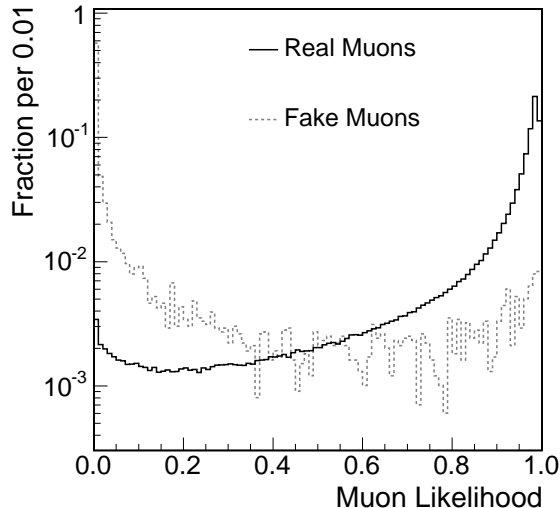


Figure 3.24: Muon likelihood distributions \mathcal{L}_μ for real and fake muons.

Muon System	$\mathcal{L}_\mu >$	Efficiency for real muons [%]	Efficiency for fake muons [%]
CMU	0.5	92.0	13.5
CMP	0.5	88.2	27.1
CMUP	0.05	98.8	55.0
CMX	0.5	91.8	22.2
IMU	0.7	78.8	9.6

Table 3.3: Efficiencies of muon likelihood requirements for real and fake muons satisfying the muon candidate requirements, compiled for different muon detector systems. The efficiencies are relative to the muon candidates having stubs in the muon detectors, and thus do not reflect the true fake rates which are much lower.

faking the electron signature. We identify photon conversions by pairing the electron candidate with another oppositely charged track in the event and require that the pair meet the criteria listed in Section 3.2.1). In case of the signal selection for semileptonic B decays (see Section 4.6) as well as for the opposite side electron tagging (see Section 6.1.3), we reject electron candidates that satisfy our conversion electron criteria. We utilize the electron likelihood \mathcal{L}_e to discriminate against hadrons passing the electron candidate criteria.

Similarly, muon candidates are composed of real muons from B decays, hadrons traveling through the calorimeter material and steel absorber reaching the muon chambers (“punch-through”), and real muons from the decay of pions or kaons before reaching the muon chambers (“decay in flight”). The muon likelihood \mathcal{L}_μ is designed to discriminate against the punch-through hadron background although, some dis-

crimination against decays in flight is also possible.

The main purpose for the development of a global lepton likelihood is to more efficiently identify real leptons from semileptonic B decays together with a lower misidentification rate for hadrons faking the lepton signature. To select the semileptonic B signals (see Section 4.6), we impose the following requirements on the lepton likelihood:

- $\mathcal{L}_e \geq 0.9$ for electrons,
- $\mathcal{L}_\mu \geq 0.05$ for CMUP muons,
- $\mathcal{L}_\mu \geq 0.50$ for CMU, CMP, and CMX muons,
- $\mathcal{L}_\mu \geq 0.70$ for IMU muons.

For opposite side flavor tagging using leptons (see Section 6.1), we utilize the lepton likelihood to characterize the purity of the opposite side lepton tag.

We also use the lepton likelihood to obtain samples enriched in fake leptons. For this purpose, we require $\mathcal{L}_{e,\mu} < 0.05$ while keeping all other requirements the same as imposed for the semileptonic B signal selection. Such an “anti-selection” allows us to acquire large fake lepton samples that are kinematically similar to the semileptonic B signal samples. These fake lepton samples are used to determine the fraction and properties of residual fake leptons found in the semileptonic B samples (see Section 5.1).

Chapter 4

Signal Selection

For measuring the frequency of B_s^0 - \bar{B}_s^0 oscillations, we need to separate events of interest (“signal”) from a large number of uninteresting events (“background”). In this chapter, we describe the selection of signal events for use in the oscillations measurement. Apart from the samples of semileptonic B_s^0 decays, we also reconstruct B^0 and B^+ control samples for understanding the mixing fit framework and calibrating the opposite side flavor tags. The samples used in this measurement have been collected using triggers exploiting the long lifetime of the B meson. B mesons can travel a few millimeters in the CDF detector before weakly decaying into lepton and charm daughters. This implies that the decay products of B mesons likely have a large impact parameter with respect to the primary $p\bar{p}$ interaction vertex. The SVT (see Section 2.8.2) enables online selection of displaced tracks which are inconsistent with the primary production at $\sim 2.5\sigma$ level.

4.1 Semileptonic Decays of the B Mesons

Figure 4.1 shows an illustration of a typical semileptonic B_s^0 decay, $B_s^0 \rightarrow D_s^- \ell^+ X$ in the CDF detector. A B_s^0 meson is produced at the primary $p\bar{p}$ interaction vertex V_P , and decays at the secondary vertex V_{B_s} into a lepton, a D_s^- meson and a neutrino. The D_s^- meson further travels some distance and decays into multiple daughters at the tertiary vertex V_{D_s} . In this dissertation, we attempt to fully reconstruct the hadronic two and three-body decays of the D_s^- meson into kaons and pions together with the lepton from the B_s^0 meson in order to access the semileptonic B_s^0 decays. A useful quantity related to the lifetime of the parent B_s^0 meson is the transverse decay length $L_{xy}^{B_s^0}$. $L_{xy}^{B_s^0}$ is defined as the displacement of the secondary vertex V_{B_s} from the primary vertex V_P , measured in the plane transverse to the beam axis and projected onto the combined transverse momentum vector of the lepton and the D_s^- meson.

We distinguish $L_{xy}^{B_s^0}$ from the true transverse decay length $D_{xy}^{B_s^0}$, which is defined as the decay distance of the B_s^0 meson in the transverse plane. We express the proper

- $B^+/B^0 \rightarrow \ell^+ \bar{D}^0 X, \quad \bar{D}^0 \rightarrow K^+ \pi^-;$
- $B^+/B^0 \rightarrow \ell^+ D^- X, \quad D^- \rightarrow K^+ \pi^- \pi^-;$
- $B^+/B^0 \rightarrow \ell^+ D^{*-} X, \quad D^{*-} \rightarrow \bar{D}^0 \pi_*^-, \bar{D}^0 \rightarrow K^+ \pi^-,$

where the charge conjugated states are also implied. The pion from a D^{*-} decay usually has a low p_T , and is consequently called a “soft pion” (denoted by π_*^+). The reconstruction of the six semileptonic decay modes starts with the selection of good quality tracks detected in the central outer tracker and the silicon detectors. The COT together with the silicon detectors provides excellent momentum resolution which further translates into excellent invariant mass resolution. The superior mass resolution helps in discriminating against backgrounds that inevitably pass the signal selection criteria. In addition to superior tracking, the excellent vertexing resolution from the silicon detectors helps reduce short-lived backgrounds originating from the primary vertex. Finally, the charged particle identification capabilities of the CDF detector for kaons and leptons (see Chapter 3) are utilized to further suppress backgrounds.

In the following, the online selection criteria, i.e. the triggers used to collect the data are described. Then, the procedure for the treatment of tracks and the offline reconstruction of the B meson signal candidates is explained. Finally, we describe the process of signal optimization minimizing the statistical uncertainties on the quantities of interest, in order to arrive at the signal selection criteria.

4.2 Online Selection: Trigger Paths

As illustrated in Figure 4.1, semileptonic B decays result into a lepton and multiple daughters originating from the decay of the D meson. The lepton and the D daughters will often have a large impact parameter with respect to the primary vertex V_P . This signature can be used for online selection of the semileptonic B decays, made possible at CDF by the presence of the SVT system. In this measurement, we utilize two trigger paths:

- The two-track trigger attempts to identify two displaced tracks from B decays based on p_T , impact parameter d_0 , opening angle in the transverse plane $\Delta\varphi$ and transverse decay length L_{xy} of the two tracks. The two trigger tracks can be the lepton and a D daughter, or both from the decay of the D meson. The two-track trigger paths were later appended with an additional path requiring a muon to enhance the online selection of $B\bar{B}$ events, where one B decays semileptonically or hadronically while the other decays semi-muonically.
- The ℓ +SVT trigger is specifically designed to identify the lepton from the semileptonic B decay using electron and muon identification in the trigger, in combination with a displaced track originating from the decay of the D meson. The lepton track has no impact parameter requirement.

We further require that the data used in this dissertation be taken with optimal detector configuration. This implies that the key detector components like COT, silicon detectors, central calorimeters, muon chambers as well as the three levels of triggers be fully functional for accumulating the data used in this measurement.

4.2.1 Two-Track Trigger

There are four different “scenarios” of the two-track trigger used in this dissertation. The nominal two-track trigger path corresponds to the “Scenario A” configuration detailed below. Over the course of a single Tevatron store, the available trigger bandwidth changes because trigger rates decrease as luminosity falls. The trigger rates are higher at higher luminosities, largely related to increased rate of physics events, and also to the increased number of track combinations arising from multiple interactions. To maximize the optimal usage of the trigger bandwidth, CDF uses a “prescaling” system whereby certain triggers using very high fraction of bandwidth are prescaled in real time. A trigger prescaled with a factor N will only pass every N^{th} event which satisfies the specified trigger criteria.

“Scenario C” trigger path was introduced to retain candidates with higher purity at higher luminosities when the prescale on the Scenario A path is high. Scenario C uses stricter requirements on the participating tracks, and consequently has a comparatively lower trigger rate. To utilize remaining bandwidth at lower luminosities, another variant of the Scenario A trigger path was introduced called “Scenario Low”. This trigger path attempts to collect large samples of B decays by relaxing the opposite charge requirement on the two displaced tracks.

To retain the events precious to the $B_s^0\text{-}\bar{B}_s^0$ oscillation measurement, Scenario Low was later supplemented by an additional requirement of an opposite side muon candidate in the trigger. The “Scenario Low plus a Muon” trigger path was designed to keep B meson decays with a muon flavor tag information. This trigger path benefits from the excellent muon triggering capability of CDF at Level 1, which results into lower trigger rates for this path even at higher luminosities. This allows taking data with no or small prescale at high luminosities. For the initial period of data taking with this trigger path, the opposite side requirement was not imposed on the muon, thereby allowing the accumulation of large samples of semi-muonic B meson decays.

In the following, we list the criteria for the various two-track triggers. Some of the notations used are:

- p_T^{XFT} and p_T^{SVT} : The transverse momentum measured by the XFT and SVT, respectively.
- $\sum p_T^{\text{XFT}}$ and $\sum p_T^{\text{SVT}}$: The scalar sum of the two-track transverse momenta measured in the XFT and SVT systems, respectively.
- $\Delta\varphi_6^{\text{XFT}}$: The separation between a pair of tracks in φ as measured by the XFT at the radius of superlayer 6 in the COT.

- $\Delta\varphi_0^{\text{SVT}}$: The difference in the SVT measured track parameter φ_0 between a pair of tracks.
- Δz_0 : Distance between a pair of tracks along the beam axis.

Scenario A

1. Level 1

- Two XFT tracks with opposite charge;
- $p_T^{\text{XFT}} > 2.0 \text{ GeV}/c$ for each XFT track;
- $\sum p_T^{\text{XFT}} > 5.5 \text{ GeV}/c$;
- $0^\circ < \Delta\varphi_6^{\text{XFT}} < 135^\circ$.

2. Level 2

- Both XFT tracks are each matched to an SVT track;
- $p_T^{\text{SVT}} > 2.0 \text{ GeV}/c$ for each SVT track;
- $\sum p_T^{\text{SVT}} > 5.5 \text{ GeV}/c$;
- $2^\circ < \Delta\varphi_0^{\text{SVT}} < 90^\circ$;
- $120 \mu\text{m} \leq |d_0^{\text{SVT}}| \leq 1 \text{ mm}$ for each SVT track;
- $L_{xy}^{\text{SVT}} > 200 \mu\text{m}$ for the two-track vertex.

3. Level 3

- SVT tracks matched to COT+silicon tracks;
- $p_T > 2.0 \text{ GeV}/c$ for each COT+silicon track;
- $\sum p_T > 5.5 \text{ GeV}/c$;
- $2^\circ < \Delta\varphi_0 < 90^\circ$;
- $80 \mu\text{m} \leq |d_0| \leq 1 \text{ mm}$ for each COT+silicon track;
- $|\Delta z_0| < 5 \text{ cm}$;
- $L_{xy} > 200 \mu\text{m}$ for the two-track vertex.

Scenario C

1. Level 1

- Two XFT tracks with opposite charge;
- $p_T^{\text{XFT}} > 2.5 \text{ GeV}/c$ for each XFT track;
- $\sum p_T^{\text{XFT}} > 6.5 \text{ GeV}/c$;

- $0^\circ < \Delta\varphi_6^{\text{XFT}} < 120^\circ$.

2. Level 2

- Both XFT tracks are each matched to an SVT track;
- $p_T^{\text{SVT}} > 2.5 \text{ GeV}/c$ for each SVT track;
- $\sum p_T^{\text{SVT}} > 6.5 \text{ GeV}/c$;
- $2^\circ < \Delta\varphi_0^{\text{SVT}} < 90^\circ$;
- $120 \mu\text{m} \leq |d_0^{\text{SVT}}| \leq 1 \text{ mm}$ for each SVT track;
- $L_{xy}^{\text{SVT}} > 200 \mu\text{m}$ for the two-track vertex.

3. Level 3

- SVT tracks matched to COT+silicon tracks;
- $p_T > 2.5 \text{ GeV}/c$ for each COT+silicon track;
- $\sum p_T > 6.5 \text{ GeV}/c$;
- $2^\circ < \Delta\varphi_0 < 90^\circ$;
- $80 \mu\text{m} \leq |d_0| \leq 1 \text{ mm}$ for each COT+silicon track;
- $|\Delta z_0| < 5 \text{ cm}$;
- $L_{xy} > 200 \mu\text{m}$ for the two-track vertex.

Scenario Low

1. Level 1

- Two XFT tracks without opposite charge requirement;
- $p_T^{\text{XFT}} > 2.0 \text{ GeV}/c$ for each XFT track;
- $0^\circ < \Delta\varphi_6^{\text{XFT}} < 90^\circ$.

2. Level 2

- Both the XFT tracks are each matched to an SVT track;
- $p_T^{\text{SVT}} > 2.0 \text{ GeV}/c$ for each SVT track;
- $2^\circ < \Delta\varphi_0^{\text{SVT}} < 90^\circ$;
- $120 \mu\text{m} \leq |d_0^{\text{SVT}}| \leq 1 \text{ mm}$ for each SVT track;
- $L_{xy}^{\text{SVT}} > 200 \mu\text{m}$ for the two-track vertex.

3. Level 3

- SVT tracks matched to COT+silicon tracks;

- $p_T > 2.0 \text{ GeV}/c$ for each COT+silicon track;
- $2^\circ < \Delta\varphi_0 < 90^\circ$;
- $80 \mu\text{m} \leq |d_0| \leq 1 \text{ mm}$ for each COT+silicon track;
- $|\Delta z_0| < 5 \text{ cm}$;
- $L_{xy} > 200 \mu\text{m}$ for the two-track vertex.

Scenario Low Plus a Muon

1. Level 1

- Two XFT tracks without opposite charge requirement, coupled with a muon candidate detected in CMU and CMX detectors matched to an XFT track;
- $p_T^{\text{XFT}} > 2.0 \text{ GeV}/c$ for two XFT tracks and muon candidate $p_T^{\text{XFT}} > 1.5 \text{ (2.0) GeV}/c$ for CMU(CMX);
- Two XFT tracks with $0^\circ < \Delta\varphi_6^{\text{XFT}} < 90^\circ$.

2. Level 2

- Two of the XFT tracks are each matched to an SVT track;
- $p_T^{\text{SVT}} > 2.0 \text{ GeV}/c$ for each SVT track;
- $2^\circ < \Delta\varphi_0^{\text{SVT}} < 90^\circ$;
- $120 \mu\text{m} \leq |d_0^{\text{SVT}}| \leq 1 \text{ mm}$ for each SVT track;
- $L_{xy}^{\text{SVT}} > 200 \mu\text{m}$ for the two-track vertex;
- $90^\circ < \Delta\varphi_6^{\text{XFT}} < 180^\circ$ between the muon candidate and each of the SVT tracks. This requirement was not imposed for some part of the data collected by this trigger path, leading to large semi-muonic B_s^0 decays accumulated by this trigger.

3. Level 3

- SVT tracks matched to COT+silicon tracks;
- $p_T > 2.0 \text{ GeV}/c$ for each track;
- $2^\circ < \Delta\varphi_0 < 90^\circ$;
- $80 \mu\text{m} \leq |d_0| \leq 1 \text{ mm}$ for each track;
- $|\Delta z_0| < 5 \text{ cm}$;
- $L_{xy} > 200 \mu\text{m}$ for the two-track vertex.

4.2.2 ℓ +SVT Trigger

Additional trigger paths utilized in this dissertation result from dedicated triggers that attempt to recognize and collect semileptonic B decays. These ℓ +SVT triggers, where $\ell = e, \mu$, identify leptons with $p_T > 4 \text{ GeV}/c$ along with a SVT displaced track with $p_T > 2 \text{ GeV}/c$. The ℓ +SVT trigger was traditionally regarded as the main trigger for the semileptonic B decays. However, the two-track trigger allows access to B decays with lower p_T leptons ($p_T > 1.5 \text{ GeV}/c$), and consequently dominates our semileptonic B decay samples (see Section 5.2). The ℓ +SVT sample is collected with an increased p_T requirement on the lepton, but without an impact parameter requirement. This allows accumulation of semileptonic B_s^0 decays at low proper decay time, and consequently better decay time resolution, as described in Section 5.3.5. We list the criteria for ℓ +SVT triggers in the following.

1. Level 1

- An XFT track with $p_T^{\text{XFT}} > 4.0 \text{ GeV}/c$ matched to a CEM cluster ($E_T > 4.0 \text{ GeV}$ and $\text{HadE}/\text{EmE} < 0.125$) or stubs in both CMU and CMP;
- An additional XFT track with $p_T^{\text{XFT}} > 2.0 \text{ GeV}/c$;
- $0^\circ < \Delta\varphi_6^{\text{XFT}} < 100^\circ$.

2. Level 2

- The XFT track with $p_T^{\text{XFT}} > 2.0 \text{ GeV}/c$ at Level 1 is matched to an SVT track;
- Electron candidate track matched with a CES wire cluster with pulse height $> 2.0 \text{ GeV}$;
- $p_T^{\text{SVT}} > 2.0 \text{ GeV}/c$ for the SVT track with $120 \mu\text{m} \leq |d_0^{\text{SVT}}| \leq 1 \text{ mm}$;
- $2^\circ < \Delta\varphi_6^{\text{XFT}} < 90^\circ$.

3. Level 3

- Lepton and SVT tracks matched to COT+silicon tracks;
- $p_T > 4.0 \text{ GeV}/c$ for the lepton;
- Lepton identification requirements based on offline calorimeter and CES quantities in the case of electron candidates, and CMU and CMP in the case of muons;
- $p_T > 2.0 \text{ GeV}/c$ for the SVT track with $120 \mu\text{m} \leq |d_0| \leq 1 \text{ mm}$;
- ℓ +SVT invariant mass less than $5.0 \text{ GeV}/c^2$;
- $2^\circ < \Delta\varphi_0 < 90^\circ$;
- $|\Delta z_0| < 5 \text{ cm}$.

4.3 Preparation of Charged Particle Tracks

We use COT+SVX tracks to represent charged particles for the reconstruction of B mesons. Since we are interested in momentum and lifetime related quantities, we take special care to ensure that the tracks used for reconstructing B candidates are of the best available quality. We therefore apply additional track quality requirements, and perform additional preparation of the tracks to improve the track parameters and their error estimates.

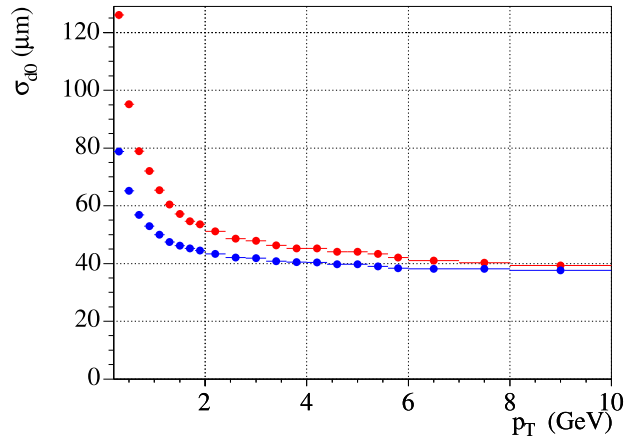


Figure 4.2: The impact parameter resolution for tracks with L00 hits (blue points) and without L00 hits (red points.) Typical tracks originating from B meson decays are below 2 GeV/ c in transverse momentum, where the resolution is significantly improved by adding L00 hits.

For rejecting fake and mis-measured tracks in the COT, we require that the tracks used in this dissertation have at least a total of 10 hits in the axial superlayers, and 10 hits in the stereo superlayers of the COT. To ensure good vertexing and impact parameter resolution, we further require that the tracks have at least 3 r - φ hits in the silicon detectors. We further impose the requirement on the transverse momentum of the tracks, $p_T > 0.4$ GeV/ c . This limits the asymmetry in the detection of positively and negatively charged tracks observed in the COT. The error matrix of the track fit in COT underestimates the measurement errors. To improve this issue, we refit tracks after rescaling the covariance matrix of the COT parent with known scale factors. The rescaled COT track is used as a starting point for the refit of the combined COT+silicon track. The refit of the track takes into account energy loss corrections for kaon, pion or lepton hypothesis according to the requested species. Silicon hits from the L00 detector are also added, if available. Addition of L00 hits significantly improves the impact parameter of low momentum tracks, as illustrated in Figure 4.2.

The efficiency of adding a L00 hit is $\sim 65\%$.

4.4 Reconstruction of B Mesons

B meson candidates are reconstructed in reverse order of their decay sequence. We first reconstruct the D meson by taking all combinations of eligible tracks (tracks passing the quality criteria detailed in Section 4.3) with a hypothesized mass and requiring that they pass through a common point of origin (the D vertex) using the CTVMFT package [61]. If such a vertex is found, the invariant mass of the D candidate daughters is calculated and required to be within reasonable bounds of the known D rest mass. The direction of flight of the D candidate is then computed using the vector sum of the track momenta of the D daughters. The D candidate flight direction is then intersected with a lepton candidate trajectory. In the case of two-track trigger data, the lepton is required to satisfy the initial identification criteria described in Sections 3.2 and 3.3, while we use the trigger lepton in the case of ℓ +SVT triggers. The lepton and the D daughters are required to pass the trigger requirements in any combination, based on the type of trigger (two-track/ ℓ +SVT). The confirmation of the trigger requirements is performed in order to remove volunteer events where the trigger decision is not based on the daughters of B decays (see Section 2.8). This allows the emulation of the trigger behavior on Monte Carlo events possible, which are used to derive various useful quantities (see Section 2.9).

As an illustrative example of a B meson candidate reconstruction, we take the reconstruction of the decay mode $B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$. We start by first constructing $\phi \rightarrow K^+ K^-$ candidates. We assign a kaon mass to two tracks with opposite charge and form a ϕ candidate using the CTVMFT package. The invariant mass of the ϕ candidate is then required to be within a reasonable range of the nominal Particle Data Group (PDG) mass value [13]. The next step is to reconstruct $D_s^- \rightarrow \phi \pi^-$ candidates using ϕ candidates from the previous step. To achieve this, we assign the pion mass to a set of tracks and require them to pass through the spatial location of the ϕ candidate within vertexing errors. The three-track vertex thus formed is interpreted as a D_s^- candidate. The invariant mass of the D_s^- candidate is required to be within acceptable bounds around the PDG mass value of the D_s^- rest mass. Finally, the direction of flight of the D_s^- meson is intersected with the trajectory of a lepton candidate to obtain a B meson candidate. In the case of data collected by the two-track trigger, we require that either the lepton and a D_s^- daughter satisfy the trigger criteria listed in Section 4.2.1, or two of the three D_s^- daughters meet the trigger requirements. In case of the ℓ +SVT data, we demand that the lepton and one of the D_s^- daughters satisfy the ℓ +SVT trigger requirements listed in Section 4.2.2.

The obtained B meson candidate is then stored and analyzed later with a stricter set of requirements imposed for reduction in backgrounds. These requirements are based on the kinematics and decay characteristics of the B meson candidates and its

daughters, and are detailed in Section 4.6.

4.5 Signal Optimization

We would like to maximize the statistical significance of our signal candidates in order to minimize the uncertainties in the measurement of quantities of interest. We use the following figure of merit for determining signal selection criteria: $S/\sqrt{S+B}$, where S denotes the number of signal candidates and B represents the number of background candidates. To avoid enhancing the signal content of our sample artificially, we use Monte Carlo simulated events (see Section 2.9) to describe the signal characteristics of interest. A subsample of the data, containing no signal, is extracted from the sidebands of the reconstructed D meson mass distribution. It is used to study the behavior of the selection requirements on background candidates. Such a sample of background candidates is beneficial because the characteristics of the background from sidebands is kinematically similar to the background underneath the D signal.

First, a set of variables are chosen which can help discriminate signal and background. The goal is to determine the optimal set of requirements on these variables by maximizing the figure of merit $S/\sqrt{S+B}$. The optimization process starts with relaxed requirements on these variables deduced from signal Monte Carlo distributions. First, each variable is varied holding others constant to maximize the figure of merit. This process is iterated until all the variables are at their optimal values.

The distinguishing variables between signal and background are based on kinematic features, quality of vertex fit, and decay characteristics of the reconstructed B meson candidates and its daughters. We describe the discriminating variables in the following:

- B Vertex Probability $\mathcal{P}(B_{vertex})$: Vertex fit probability of the B meson vertex. Background candidates typically have poor fit quality.
- $\chi^2_{xy}(D)$: D vertex fit quality in r - φ plane expressed in terms of χ^2 of the vertex fit.
- $L_{xy}/\sigma L_{xy}(PV \rightarrow D)$: The significance of the transverse decay distance of the D candidate with respect to the primary vertex. Background candidates likely populate lower values of $L_{xy}/\sigma L_{xy}(PV \rightarrow D)$.
- $ct(D)$: Proper decay time of the D meson. It is obtained from the transverse decay distance L_{xy} using $ct(D) = L_{xy}(D) \cdot \frac{m_D^{\text{PDG}}}{p_T(\ell D)}$. Here m_D^{PDG} is the world average rest mass value for the D meson [13]. Background candidates likely occupy lower values of $ct(D)$.
- $L_{xy}/\sigma L_{xy}(PV \rightarrow B)$: The significance of the transverse decay distance of the B meson candidate with respect to the primary vertex. Background candidates have lower values of $L_{xy}/\sigma L_{xy}(PV \rightarrow B)$.

- Track p_T : Background candidates are composed of tracks that prefer lower transverse momentum.
- $m_{\ell D}$: Invariant mass of the lepton and D meson system. Background candidates preferentially populate low $m_{\ell D}$.
- $\sigma_{ct^*}(B)$: Error on B pseudo-proper decay time calculated using $\sigma_{ct^*}(B) = \sigma_{L_{xy}(PV \rightarrow B)} \cdot \frac{m_B^{\text{PDG}}}{p_T(\ell D)}$, where m_B^{PDG} is the nominal value of the mass of the B meson [13].
- $|\cos \psi|$: Cosine of the helicity angle for vector particles like ϕ and K^{*0} . The helicity angle ψ is defined as the angle between the D_s and K^- momentum in the ϕ rest frame. Backgrounds have no preference in this variable, and are distributed flat.
- $CLL(K)$: Combined particle identification variable for kaons defined in Section 3.1.3, which is useful to identify kaons from D decays. Backgrounds are typically composed of pions which are more abundantly produced in $p\bar{p}$ collisions.

Figure 4.3 shows an example of how the optimization is performed on a discriminating quantity, the $L_{xy}/\sigma_{L_{xy}}(PV \rightarrow D)$ in the reconstructed ℓD^0 sample. The requirements on $L_{xy}/\sigma_{L_{xy}}(PV \rightarrow D)$ are varied holding other quantities constant. The two variables that are monitored while optimizing the selection criterion on $L_{xy}/\sigma_{L_{xy}}(PV \rightarrow D)$ are the figure of merit $S/\sqrt{S+B}$, shown in the bottom plot and the number of signal candidates S , shown in the middle plot. The signal Monte Carlo is normalized to the number of signal candidates observed in data for a certain requirement on $L_{xy}/\sigma_{L_{xy}}(PV \rightarrow D)$ ($= 6$ in the middle plot of Figure 4.3). When a set of requirements have a similar figure of merit, the requirement that retains the highest amount of signal candidates is chosen. The optimization process is iterative; the selection requirements are changed to new values at the beginning of every iteration. The bottom plot in Figure 4.3, shows the chosen optimized selection requirement on $L_{xy}/\sigma_{L_{xy}}(PV \rightarrow D)$ to be greater than 6.

4.6 Signal Selection Requirements

To select B meson candidates, we impose requirements on the decay topologies and kinematical variables of the reconstructed candidates to suppress background contamination through a process detailed in the previous section. The requirements on the lepton candidates, crucial for lepton identification in the two-track triggered sample, are listed in Section 3.4. Additionally, we reject electron candidates that appear to originate from photon conversions. We impose the same lepton identification requirements on the 4 GeV/ c lepton candidates collected by the ℓ +SVT trigger.

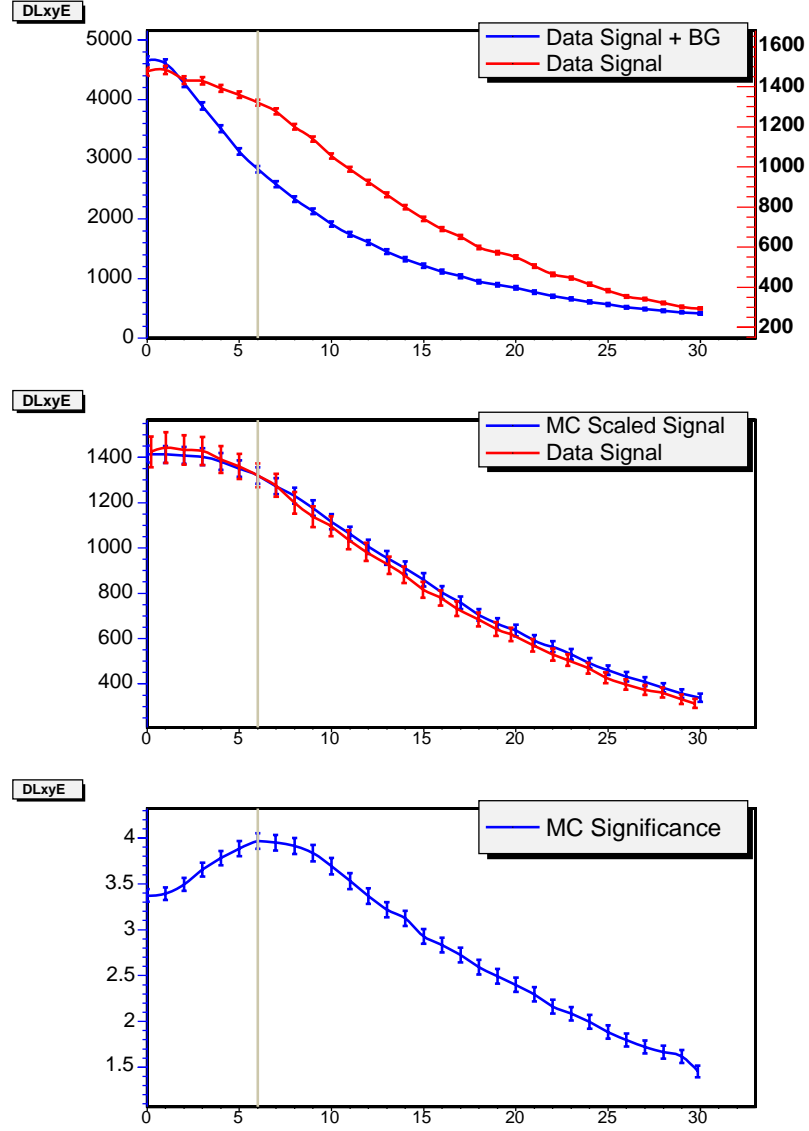


Figure 4.3: Optimization of $L_{xy}/\sigma L_{xy}(PV \rightarrow D)$ in the reconstructed ℓD^0 sample. The middle plot shows the number of signal candidates remaining after imposing requirements on $L_{xy}/\sigma L_{xy}(PV \rightarrow D)$ and the bottom plot shows the corresponding value of the figure of merit $S/\sqrt{S+B}$.

We include charged kaon identification, as detailed in Section 3.1, in the selection of the B_s^0 modes which have a kaon in the final state. This refers to the $D_s^- \rightarrow \phi\pi^-$, $\phi \rightarrow K^+K^-$ and $D_s^- \rightarrow K^{*0}K^-$, $K^{*0} \rightarrow K^+\pi^-$ modes. Kaon identification in these modes serves two purposes:

1. Backgrounds in both decay modes are dominated by pions. Hence, particle identification improves their suppression.
2. The $D_s^- \rightarrow K^{*0}K^-$, $K^{*0} \rightarrow K^+\pi^-$ decay suffers additionally from large backgrounds originating from a reflection of $D^- \rightarrow K^{*0}\pi^-$ and $D^- \rightarrow K^+\pi^-\pi^-$ (non-resonant) decays. Such backgrounds are caused by a mis-assignment of the pions from the D^- decays as kaons. Thus, kaon particle identification is likely to reduce the D^- reflection into D_s^- candidates.

In addition, the narrow invariant mass resonance provided by the ϕ meson while reconstructing the $D_s^- \rightarrow \phi\pi^-$, $\phi \rightarrow K^+K^-$ mode greatly reduces background in this mode. We require the reconstructed ϕ mass to be in the interval $|m_\phi - m_\phi^{\text{PDG}}| < 12 \text{ MeV}/c^2$ around the nominal mass value [13]. In the case of the $D_s^- \rightarrow K^{*0}K^-$, $K^{*0} \rightarrow K^+\pi^-$ decays, we demand the reconstructed K^{*0} mass to obey $|m_{K^{*0}} - m_{K^{*0}}^{\text{PDG}}| < 50 \text{ MeV}/c^2$. After applying these requirements on the two modes, the optimized selection criteria based on the variables discussed in Section 4.5 are obtained for the three ℓD_s modes. In the case of the $D_s^- \rightarrow K^{*0}K^-$, $K^{*0} \rightarrow K^+\pi^-$ mode, we include the D^- reflection shape from a Monte Carlo sample in the background description while performing optimization. The optimal selection criteria are summarized in Table 4.1. The quantity $m_{\ell D}$ is a powerful discriminator against backgrounds, and is hence used directly in the fit framework (see Section 5.3.3).

A signal optimization is performed on similar discriminating variables for the ℓD modes, and the optimized selection criteria is listed in Table 4.2. As mentioned before, the $B^+/B^0 \rightarrow D\ell\nu_\ell X$ modes are used for the calibration of opposite side flavor tagging and checking the fit framework (see Chapter 7). Particle identification is not used for these modes as the number of signal candidates in ℓD modes is large enough for accomplishing the flavor tagging calibration, together with low contamination from backgrounds. The selection criteria for the ℓD^* mode are deliberately kept similar to the ℓD^0 mode in order to simplify the analysis. The ℓD^{*+} candidates are obtained from the sample of ℓD^0 candidates, where a soft pion candidate passing the track requirements detailed in Section 4.3 is vertexed with the D^0 candidate using the CTVMFT package. To isolate ℓD^{*+} candidates, a requirement on the invariant mass difference $0.1435 \text{ GeV}/c^2 \leq |m_{D^{*+}} - m_{D^0}| \leq 0.1475 \text{ GeV}/c^2$ is imposed.

Selection Requirements	$D_s^- \rightarrow \phi\pi^-$	$D_s^- \rightarrow K^{*0}K^-$	$D_s^- \rightarrow \pi^+\pi^-\pi^-$
$\mathcal{P}(B_{vertex}) >$	10^{-7}	10^{-7}	10^{-5}
$\chi_{xy}^2(D_s^-) <$	20	20	20
$L_{xy}/\sigma_{L_{xy}}(D_s^-) >$	5	7	11
$L_{xy}/\sigma_{L_{xy}}(B_s^0) >$	—	2	2
$ \cos\psi >$	0.3	0.3	—
$p_T(\text{tracks}) [\text{GeV}/c] >$	0.4	0.4	0.7
$\sigma_{ct}(B_s^0) [\text{cm}] <$	0.04	0.04	0.04
$m_{\ell^+D_s^-} [\text{GeV}/c^2] >$	2.0	2.0	2.0
$m_{\ell^+D_s^-} [\text{GeV}/c^2] <$	5.5	5.5	5.5
$ct(D) [\text{cm}] >$	-0.0050	-0.0050	+0.0025
$ct(D) [\text{cm}] <$	0.100	0.100	0.100
$CLL(K1) >$	-2.5	-2.25	—
$CLL(K2) >$	-2.5	-1.1	—

Table 4.1: Optimized selection criteria for the $\ell^+D_s^-$ samples.

Selection Requirements	$D^0 \rightarrow K^+\pi^-$	$D^- \rightarrow K^+\pi^-\pi^-$
$\mathcal{P}(B_{vertex}) >$	10^{-6}	10^{-5}
$\chi_{xy}^2(D) <$	20	20
$L_{xy}/\sigma_{L_{xy}}(D) >$	6	11
$L_{xy}/\sigma_{L_{xy}}(B) >$	2	2
$p_t(K) [\text{GeV}/c] >$	0.5	0.7
$\sigma_{ct}(B) [\text{cm}] <$	0.04	0.04
$m_{\ell D} [\text{GeV}/c^2] >$	2.0	2.0
$m_{\ell D} [\text{GeV}/c^2] <$	5.5	5.5
$ct(D) [\text{cm}] >$	-0.0050	-0.0050
$ct(D) [\text{cm}] <$	0.100	0.200

Table 4.2: Optimized selection criteria for the ℓD samples; the selection requirements for the ℓD^{*+} sample are identical to those of the ℓD^0 sample.

Chapter 5

Characterization of B Meson Candidates

In this chapter, we determine the composition of the B_s^0 samples used in the measurement of B_s^0 - \bar{B}_s^0 oscillation frequency. The B_s^0 candidates are reconstructed first with a relaxed set of selection criteria. Stricter requirements based on decay topology and kinematical quantities are imposed later, after a thorough optimization procedure against backgrounds (see Chapter 4) to minimize the statistical uncertainties on the quantities that we are interested in this measurement. However, the ‘incomplete’ or ‘partial’ reconstruction of the semileptonic decays of B mesons leads to contamination from a variety of backgrounds. Our goal in this chapter is to identify and characterize these backgrounds in discriminating variables with respect to the signal, and hence determine the true B_s^0 content of our samples. Such a characterization also yields the shapes of the backgrounds in the discriminating variables. These shapes are later used as PDFs in a maximum likelihood fit to perform the measurement of B_s^0 - \bar{B}_s^0 oscillation frequency.

5.1 Sources of Backgrounds in Semileptonic B Decays

As described in Section 4.1, in semileptonic B decays, the B meson momentum is incompletely reconstructed due to the missing neutrino and sometimes other neutral decay products that we cannot observe in the CDF detector. Thus we are unable to reconstruct the complete invariant mass of the B mesons, and hence cannot definitively infer that the lepton and the D candidate originate from the same parent B meson by just relying on the invariant mass of the ℓD system $m_{\ell D}$. This leads to a variety of backgrounds that contaminate the B samples by yielding a lepton and a D candidate with correct charge correlation passing our B signal selection criteria. These backgrounds are largely classified into four categories:

1. **Combinatorial Background:** This is the class of background candidates that do not originate from true D mesons, but instead result from accidental combinations of tracks that happen to have an invariant mass in the vicinity of the D mass. The contribution from combinatorial background is assessed using the D meson mass distribution.
2. **Physics Background:** This kind of background originates from decays of real B mesons, where a real lepton and a real D meson is present in the final state. Specifically for the $\ell^+ D_s^-$ samples, B^0 , B^+ and B_s^0 mesons sometimes decay into a D_s^- and another charm meson along with other decay products ($B \rightarrow D_s^- D X$ where D can be a D_s^- , D^- or D^0). The D can further decay semileptonically and produce a real lepton. These kind of decays produce $\ell^+ D_s^-$ signatures that feed into our signal B_s^0 candidates and are generally called physics background.
3. **Fake Lepton Background:** Sometimes a real D meson produced in the primary interaction (“prompt” D) or from a B decay (“secondary” D) is combined with a fake lepton track, and meets our signal selection criteria.
4. **D^- Reflection Background:** In the reconstruction of $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$ candidates, pions from $D^- \rightarrow K^{*0} \pi^-$ and $D^- \rightarrow K^+ \pi^- \pi^-$ (non-resonant) decays are mis-assigned a kaon mass hypothesis (see Section 4.6), and populate the same invariant mass region as the correctly reconstructed D_s^- decays. This background is referred to as the D^- reflection background. We use kaon particle identification on the kaon candidate track that significantly reduces this background, but does not eliminate it completely.

In the following, we attempt to determine the contribution from the various sources of backgrounds, and also obtain the PDFs of these backgrounds in three discriminating variables: invariant mass of D , $m_{\ell D}$, and pseudo-proper decay length ct^* .

5.2 Further Classification of B Meson Samples

Before proceeding to understand the composition of the B samples in terms of signal and backgrounds, we would like to sub-divide our samples based on trigger paths and topologies. The semileptonic decays that we are interested in produce at least three daughter tracks. Thus, it is possible to trigger on such decays using different combinations of the tracks. A majority ($\sim 90\%$) of semileptonic decays analyzed in this dissertation are collected via the two-track trigger (see Section 4.2.1). We recognize two different trigger configurations in the data accumulated by the two-track trigger based on whether the lepton is part of the two-track trigger or not. Such a distinction is likely going to affect, for example, the contamination from fake lepton background.

There is a significant overlap between the events triggered by the two-track and ℓ +SVT trigger paths. About 60% of the semileptonic B decays collected via the ℓ +SVT samples are already contained in the two-track trigger data. After removing this overlap, the ℓ +SVT datasets contribute $\sim 10\%$ of semileptonic B meson decays to the total sample.

Based on the trigger topologies, we sub-divide our B meson samples into the following categories:

- B Triggers: The B meson candidates are collected by the two-track trigger. The lepton and a D daughter track form the two-track trigger pair.
- D Triggers: Candidates in the two-track trigger data not belonging to the category of B triggers. Consequently, the two-track trigger pair forms two daughters of the D meson.
- ℓ +SVT Triggers: Exclusive B meson candidates collected by the ℓ +SVT trigger path, but not selected by either B or D triggers.

The relative contribution of the three trigger topologies for signal B_s^0 decays is in the ratio $B : D : \ell$ +SVT $\sim 70 : 20 : 10$. The backgrounds are expected to be different between the three categories of the B meson samples. We expect to benefit from this distinction by localizing backgrounds into certain categories of these trigger classes. To limit contamination from fake lepton backgrounds, we further impose the following requirements on the pseudo-proper decay length of the B mesons:

- B Triggers: $ct^* \in [0.005, \infty)$ cm;
- D Triggers: $ct^* \in [0.010, \infty)$ cm;
- ℓ +SVT Triggers: $ct^* \in [0.010, \infty)$ cm.

Finally, the B meson data is divided into three data-taking periods. The detector configuration changes, particularly with respect to the COT, TOF and CPR detectors, have forced this discrimination. We use particle identification on leptons and kaons in our signal selection, hence the performance of these detectors affects our signal candidates. The luminosities corresponding to the three periods are:

- Period 1: $\sim 355\text{pb}^{-1}$,
- Period 2: $\sim 410\text{pb}^{-1}$,
- Period 3: $\sim 230\text{pb}^{-1}$.

5.3 Characterization of the B_s^0 Samples and Fitting Technique

To extract parameters of interest characterizing signal and backgrounds, we use the unbinned maximum likelihood fitting method. To construct the overall likelihood \mathcal{L} describing a given sample, the likelihood for each candidate is first constructed by forming a joint probability distribution describing the discriminating variables in the various components of the sample. If \mathbf{x} denotes the vector of all the observables and discriminating variables, the likelihood for the n^{th} candidate can be written as

$$L_n(\mathbf{x}) = \sum_i f_i \cdot \mathcal{P}^i(\mathbf{x}), \quad (5.1)$$

where the index i runs over different components of a given sample, and f_i is the fraction of the component with joint PDF \mathcal{P}^i . The fractions f_i satisfy

$$\sum_i f_i = 1. \quad (5.2)$$

The overall likelihood is then expressed in terms of the candidate likelihoods as

$$\mathcal{L}(\mathbf{x}) = \prod_n L_n(\mathbf{x}) = \prod_n \sum_i f_i \cdot \mathcal{P}_n^i(\mathbf{x}), \quad (5.3)$$

where the index n runs over all the candidates in the sample. The PDFs and \mathcal{L} values depend on the parameters of interest. Various values of \mathcal{L} are probed by varying the parameters, and the values of parameters maximizing \mathcal{L} are determined. The maximization procedure is performed using the MINUIT package [62]. It conveniently minimizes the quantity

$$-2 \log \mathcal{L} \quad (5.4)$$

to determine the optimal values of the unknown parameters. In case of multiple samples that are used to determine the unknown parameters, the minimization is performed using

$$-2 \sum_m \log \mathcal{L}_m, \quad (5.5)$$

where the index m runs over all the participating samples.

5.3.1 Characterization in D_s^- Mass

The D_s^- invariant mass distribution provides an opportunity to understand the composition of $\ell^+ D_s^-$ samples in terms of real and incorrectly reconstructed D_s^- mesons. This refers to ascertaining the fractions and mass PDFs of combinatorial and D^- reflection backgrounds, as well as determining the mass PDFs for signal together with physics and fake lepton backgrounds (see Section 5.1).

D_s^- Mass PDF for Signal, Physics and Fake Lepton Backgrounds

The width of the reconstructed D_s^- mass distribution for real D_s^- mesons is dominated by detector resolution. The mass PDF for the components of $\ell^+ D_s^-$ samples containing real D_s^- mesons (signal, physics background and fake lepton background) is described by a normalized sum of two Gaussian distributions with the same mean, and expressed as:

$$\mathcal{P}^{\text{signal}}(m) = \frac{1}{N} [f_m \cdot G(m; M, \sigma_1) + (1 - f_m) \cdot G(m; M, \sigma_2)], \quad (5.6)$$

where $G(x; \mu, \sigma)$ is a Gaussian with mean μ and width σ , f_m is the fraction of the Gaussian with width σ_1 and N is the normalization which depends on the mass fitting range (M_{\min}, M_{\max}). The mass fitting range for the three $\ell^+ D_s^-$ modes is summarized in Table 5.1. The parameters of this PDF are determined directly from data by performing a fit in the D_s^- invariant mass.

Mode	M_{\min} [GeV/ c^2]	M_{\max} [GeV/ c^2]
$D_s^- \rightarrow \phi \pi^-$	1.92	2.04
$D_s^- \rightarrow K^{*0} K^-$	1.90	2.05
$D_s^- \rightarrow \pi^+ \pi^- \pi^-$	1.92	2.04

Table 5.1: D_s^- mass fitting range for the three $\ell^+ D_s^-$ modes.

D_s^- Mass PDF for Combinatorial Background

The combinatorial background is described by a linear parameterization, described by

$$\mathcal{P}^{\text{comb}}(m) = \frac{1}{N} [a \cdot m + b], \quad (5.7)$$

where the normalization N again depends on the mass fitting range. The parameters a and b are determined directly from data.

D_s^- Mass PDF for D^- Reflection Background

The D^- reflection background (only present in the $B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$ mode) is expressed by an exponential distribution convoluted with a Gaussian added to a linear term:

$$\mathcal{P}^{\text{refl}}(m) = \frac{1}{N} [E(m; M_0, \tau) \otimes G(m; M_0, \sigma) + a \cdot m + b], \quad (5.8)$$

where

$$E(m; M_0, \tau) = \frac{1}{\tau} e^{-(m-M_0)/\tau} \theta(m - M_0),$$

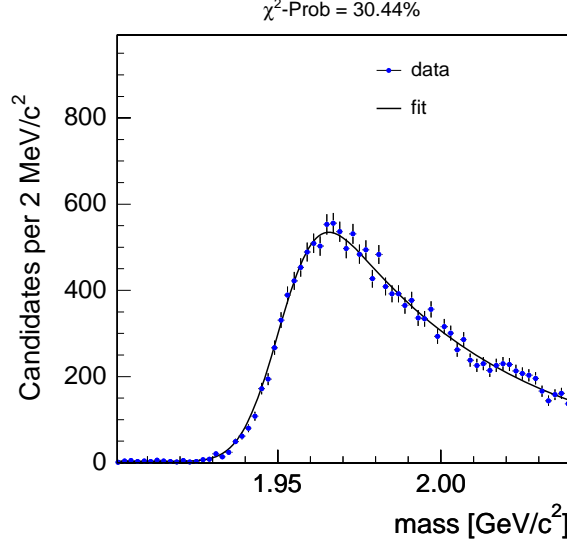


Figure 5.1: D_s^- mass distribution for B triggers in $B \rightarrow \mu^+ D^- X$ Monte Carlo. These decays are reconstructed as $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$.

with $\theta(m - M_0)$ being the step function. The normalization factor depends on the fitting range. The shape of the D^- reflection in the D_s^- invariant mass is obtained from Monte Carlo simulated events and shown in Figure 5.1. However, while performing a fit in D_s^- invariant mass variable on data, the parameter M_0 is left free to accommodate shifts in the invariant mass distribution of this background with respect to Monte Carlo.

The parameterization obtained by convoluting an exponential distribution with a Gaussian is very commonly used in the current analysis to describe the shapes of various backgrounds in D_s^- mass, $m_{\ell^+ D_s^-}$, and ct^* spaces. From here onwards, we refer to this parameterization as the “tailed Gaussian” distribution for simplicity.

D_s^- Mass Fit in Data

Using the above mentioned models for signal and backgrounds in the D_s^- mass, our goal is to determine the PDFs for all the components in D_s^- mass and obtain fractions of combinatorial and D^- reflection backgrounds. The candidate likelihood characterizing the D_s^- mass distribution is given by

$$L(m) = f_{\text{comb}} \cdot \mathcal{P}^{\text{comb}}(m) + f_{\text{phys}} \cdot \mathcal{P}^{\text{phys}}(m) + f_{\text{fake}} \cdot \mathcal{P}^{\text{fake}}(m) + f_{\text{refl}} \cdot \mathcal{P}^{\text{refl}}(m) \\ + (1 - f_{\text{comb}} - f_{\text{phys}} - f_{\text{fake}} - f_{\text{refl}}) \cdot \mathcal{P}^{\text{signal}}(m), \quad (5.9)$$

where f_i are the fractions of the various signal and background components of the B_s^0 samples.

Decay Samples	$\phi\pi^-$ [%]	$K^{*0}K^-$ [%]	$\pi^+\pi^-\pi^-$ [%]
	Period 1		
$\mu D_s, B$ trigger	40.9	61.4	91.0
$e D_s, B$ trigger	44.1	61.9	90.1
$\mu D_s, D$ trigger	27.9	58.0	88.2
$e D_s, D$ trigger	35.9	58.6	89.5
$\mu D_s, \ell + SVT$ trigger	45.2	72.9	88.6
$e D_s, \ell + SVT$ trigger	41.7	67.9	89.3
	Period 2		
$\mu D_s, B$ trigger	43.1	66.3	90.0
$e D_s, B$ trigger	43.4	65.4	89.9
$\mu D_s, D$ trigger	29.3	58.4	91.2
$e D_s, D$ trigger	29.0	59.7	91.4
$\mu D_s, \ell + SVT$ trigger	48.4	75.2	90.3
$e D_s, \ell + SVT$ trigger	42.7	73.6	90.1
	Period 3		
$\mu D_s, B$ trigger	42.5	67.7	90.2
$e D_s, B$ trigger	45.5	62.7	92.2
$\mu D_s, D$ trigger	35.0	60.4	88.3
$e D_s, D$ trigger	36.1	59.7	89.5
$\mu D_s, \ell + SVT$ trigger	46.2	73.1	91.4
$e D_s, \ell + SVT$ trigger	39.4	69.6	90.9

Table 5.2: Fractions of combinatorial background in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates in the three data-taking periods. The absolute uncertainties in these fractions are $\sim 2\%$ each.

Since $\mathcal{P}^{\text{phys}}(m) = \mathcal{P}^{\text{fake}}(m) = \mathcal{P}^{\text{signal}}(m)$, we can write Equation (5.9) as

$$L(m) = f_{\text{comb}} \cdot \mathcal{P}^{\text{comb}}(m) + f_{\text{refl}} \cdot \mathcal{P}^{\text{refl}}(m) + (1 - f_{\text{comb}} - f_{\text{refl}}) \cdot \mathcal{P}^{\text{signal}}(m). \quad (5.10)$$

Using this model for the candidate likelihood, we fit the D_s^- mass distributions in the 18 subsamples for each B_s^0 decay mode (corresponding to the three trigger topologies \times three data-taking periods \times two lepton species = e, μ). The shape of D^- reflection is derived from Monte Carlo as described above, however the offset (parameter M_0) of the tailed Gaussian shape is floated to accommodate shifts in the invariant mass while performing the fit. The other parameters are fixed to their determined values, except the fractions of combinatorial and reflection backgrounds. The fractions of combinatorial background are summarized in Table 5.2. The fractions of D^- reflection background are displayed in Table 5.3. The D_s^- mass PDFs of the five components of the B_s^0 candidates are also obtained in this process.

Decay Samples	$\phi\pi^-$ [%]	$K^{*0}K^-$ [%]	$\pi^+\pi^-\pi^-$ [%]
Period 1			
$\mu D_s, B$ trigger	—	56.3	—
$e D_s, B$ trigger	—	55.2	—
$\mu D_s, D$ trigger	—	55.4	—
$e D_s, D$ trigger	—	55.2	—
$\mu D_s, \ell + SVT$ trigger	—	59.7	—
$e D_s, \ell + SVT$ trigger	—	58.3	—
Period 2			
$\mu D_s, B$ trigger	—	50.7	—
$e D_s, B$ trigger	—	49.3	—
$\mu D_s, D$ trigger	—	58.4	—
$e D_s, D$ trigger	—	51.1	—
$\mu D_s, \ell + SVT$ trigger	—	55.4	—
$e D_s, \ell + SVT$ trigger	—	55.2	—
Period 3			
$\mu D_s, B$ trigger	—	45.0	—
$e D_s, B$ trigger	—	50.9	—
$\mu D_s, D$ trigger	—	53.4	—
$e D_s, D$ trigger	—	56.2	—
$\mu D_s, \ell + SVT$ trigger	—	55.4	—
$e D_s, \ell + SVT$ trigger	—	55.2	—

Table 5.3: Fractions of D^- reflection background relative to the sum of signal, physics, prompt and D^- reflection backgrounds in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates, for the three data-taking periods. The absolute uncertainties in these fractions are $\sim 2\%$ each.

Signal and Sideband Regions in D_s^- Mass

As mentioned before, the D_s^- mass can be used to understand the composition of the $B_s^0 \rightarrow D_s^- \ell^+ X$ samples in terms of real and incorrectly reconstructed D_s^- mesons. In particular, the reconstructed D_s^- mass regions sufficiently far away from the D_s^- peak can be utilized to study the behavior of combinatorial background in the variables of interest. These regions are referred to as the D_s^- mass sidebands, and summarized for the three decay modes in Table 5.4. For the $D_s^- \rightarrow K^{*0} K^-$ mode, we only use the lower mass sideband due to the presence of D^- reflection background in the upper mass sideband. The D_s^- signal regions for the three decay modes are listed in Table 5.5 corresponding to $\sim \pm 3\sigma$ intervals around the nominal D_s^- mass values from Ref. [13]. These signal regions are used in conjugation with the sidebands in order to extract the background subtracted signal distributions for various interesting quantities.

Decay Mode	Lower sideband		Upper sideband	
	M_{min} [GeV/ c^2]	M_{max} [GeV/ c^2]	M_{min} [GeV/ c^2]	M_{max} [GeV/ c^2]
$D_s^- \rightarrow \phi\pi^-$	1.92	1.94	2.00	2.02
$D_s^- \rightarrow K^{*0}K^-$	1.90	1.94	—	—
$D_s^- \rightarrow \pi^+\pi^-\pi^-$	1.92	1.94	2.00	2.02

Table 5.4: Mass ranges for defining the D_s^- sideband regions in the three $\ell^+ D_s^-$ modes. The candidates belonging to these regions are utilized for obtaining properties of the combinatorial background.

Mode	M_{min} [GeV/ c^2]	M_{max} [GeV/ c^2]
$D_s^- \rightarrow \phi\pi^-$	1.95	1.99
$D_s^- \rightarrow K^{*0}K^-$	1.95	1.99
$D_s^- \rightarrow \pi^+\pi^-\pi^-$	1.95	1.99

Table 5.5: Mass ranges for defining the D_s^- signal regions in the three $\ell^+ D_s^-$ modes.

5.3.2 Determination of Physics Background Fractions

After determining the fractions of combinatorial and D^- reflection backgrounds, the next step is to evaluate the contribution from physics backgrounds. As discussed in Section 5.1, physics backgrounds typically arise from B meson decays to a D_s^- and another charm meson, which then decays semileptonically providing us with a signature that mimics our signal. The contributions from the physics backgrounds are assessed using Monte Carlo samples, and weighted by their branching ratios before applying reconstruction efficiencies. In the case of $\ell^+ D_s^-$ modes, sources of such backgrounds and their respective contribution are summarized in Table 5.6 for the three trigger topologies. Physics backgrounds strongly depend on the p_T of the participating lepton; leptons from D decays are, in general, softer than leptons originating directly from a B meson decay. This is evident from Table 5.6 as the D triggers (which are populated by softer leptons as compared to the B and ℓ +SVT triggers) show significantly higher fractions of physics backgrounds.

5.3.3 Characterization in $m_{\ell^+ D_s^-}$

After determining the fractions of combinatorial, physics, and D^- reflection backgrounds, the next step is to determine the fractions of fake lepton background together with the behavior of the signal and various background components in the $m_{\ell^+ D_s^-}$ space. Subsequently, we would know the fraction of each of the five components of our B_s^0 samples. As detailed in Section 5.1, fake lepton background arises from a combination of a prompt or secondary real D meson with a hadron faking the

<i>B</i> Triggers				
Decay	$\phi\pi^-$ [%]	$K^{*0}K^-$ [%]	$\pi^+\pi^-\pi^-$ [%]	
$B^+ \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow \mu^+Y$	7.8 ± 3.1	7.9 ± 3.1	7.6 ± 3.0	
$B^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow \mu^+Y$	7.8 ± 3.1	8.0 ± 3.2	7.7 ± 3.0	
$B_s^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow \mu^+Y$	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	
$B_s^0 \rightarrow D_s^{(*)+}D_s^{(*)-}X, \quad D_s^{(*)} \rightarrow \mu^+Y$	2.9 ± 0.9	2.8 ± 0.9	2.8 ± 0.8	
total	18.6 ± 6.5	18.8 ± 6.6	18.2 ± 6.4	
Decay	$\phi\pi^-$ [%]	$K^{*0}K^-$ [%]	$\pi^+\pi^-\pi^-$ [%]	
$B^+ \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow e^+Y$	6.7 ± 2.7	6.7 ± 2.7	7.1 ± 2.8	
$B^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow e^+Y$	6.7 ± 2.7	6.8 ± 2.7	7.1 ± 2.8	
$B_s^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow e^+Y$	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	
$B_s^0 \rightarrow D_s^{(*)+}D_s^{(*)-}X, \quad D_s^{(*)} \rightarrow e^+Y$	2.0 ± 0.6	2.0 ± 0.6	1.9 ± 0.6	
total	15.5 ± 5.4	15.6 ± 5.5	16.2 ± 6.0	
<i>D</i> Triggers				
Decay	$\phi\pi^-$ [%]	$K^{*0}K^-$ [%]	$\pi^+\pi^-\pi^-$ [%]	
$B^+ \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow \mu^+Y$	11.9 ± 4.8	11.9 ± 4.8	11.3 ± 4.5	
$B^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow \mu^+Y$	11.9 ± 4.8	11.8 ± 4.8	11.3 ± 4.5	
$B_s^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow \mu^+Y$	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	
$B_s^0 \rightarrow D_s^{(*)+}D_s^{(*)-}X, \quad D_s^{(*)} \rightarrow \mu^+Y$	3.7 ± 1.1	3.6 ± 1.0	3.7 ± 1.0	
total	27.6 ± 9.7	27.4 ± 9.6	26.4 ± 9.5	
Decay	$\phi\pi^-$ [%]	$K^{*0}K^-$ [%]	$\pi^+\pi^-\pi^-$ [%]	
$B^+ \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow e^+Y$	11.0 ± 4.3	11.2 ± 4.4	9.5 ± 3.8	
$B^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow e^+Y$	11.1 ± 4.4	11.2 ± 4.4	9.5 ± 3.8	
$B_s^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow e^+Y$	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	
$B_s^0 \rightarrow D_s^{(*)+}D_s^{(*)-}X, \quad D_s^{(*)} \rightarrow e^+Y$	2.9 ± 0.9	3.0 ± 0.8	2.7 ± 0.8	
total	25.1 ± 8.8	25.5 ± 8.9	21.8 ± 7.6	
ℓ +SVT Triggers				
Decay	$\phi\pi^-$ [%]	$K^{*0}K^-$ [%]	$\pi^+\pi^-\pi^-$ [%]	
$B^+ \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow \mu^+Y$	3.8 ± 1.5	3.9 ± 1.6	3.8 ± 1.5	
$B^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow \mu^+Y$	3.7 ± 1.5	3.8 ± 1.5	3.8 ± 1.5	
$B_s^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow \mu^+Y$	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	
$B_s^0 \rightarrow D_s^{(*)+}D_s^{(*)-}X, \quad D_s^{(*)} \rightarrow \mu^+Y$	1.4 ± 0.4	1.6 ± 0.4	1.4 ± 0.4	
total	9.0 ± 3.1	9.4 ± 3.3	9.1 ± 3.1	
Decay	$\phi\pi^-$ [%]	$K^{*0}K^-$ [%]	$\pi^+\pi^-\pi^-$ [%]	
$B^+ \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow e^+Y$	3.4 ± 1.3	3.1 ± 1.2	3.0 ± 1.2	
$B^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow e^+Y$	3.3 ± 1.3	3.1 ± 1.2	3.0 ± 1.2	
$B_s^0 \rightarrow D_s^{(*)}D^{(*)}X, \quad D^{(*)} \rightarrow e^+Y$	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	
$B_s^0 \rightarrow D_s^{(*)+}D_s^{(*)-}X, \quad D_s^{(*)} \rightarrow e^+Y$	1.1 ± 0.3	1.2 ± 0.4	1.2 ± 0.4	
total	7.9 ± 2.8	7.4 ± 2.6	7.3 ± 2.6	

Table 5.6: Expected fraction of each physics background with respect to the signal and physics backgrounds in the $\ell^+D_s^-$ modes for the *B*, *D* and ℓ +SVT triggers.

lepton signature. The category of fake lepton candidates also includes cases when the lepton is real but originates from prompt sources. One such example is a $c\bar{c}$ event, where one charm hadronizes into a D_s^- meson and the other produces a lepton. Another example is the case when an electron from an undetected photon conversion (see Section 6.1.3) is paired with a real D_s^- meson. In both of these cases, we obtain the same signature as our signal B_s^0 decays.

For determining the fraction of the fake lepton background, we utilize the fake sample obtained using the anti-selection on lepton likelihood (see Section 3.4). This sample is highly enriched in fake leptons and is kinematically similar to the B_s^0 samples. We first derive the shape of $m_{\ell+D_s^-}$ for the signal, combinatorial background, D^- reflection background, and physics background, followed by the shape for fake leptons. After deriving the preliminary PDFs in the $m_{\ell+D_s^-}$ space for all the five components of the B_s^0 samples, we perform a simultaneous fit in the D_s^- mass and $m_{\ell+D_s^-}$ spaces for the B_s^0 candidates. This provides additional separation power between the different components of the sample, and thus helps in determining the fake fractions and final PDFs for the various components more precisely. The D_s^- mass PDF parameters previously determined are fixed in this fit. The fake lepton fraction, apart from a few of the $m_{\ell+D_s^-}$ PDF parameters are left free in the fit.

$m_{\ell+D_s^-}$ PDF for Signal

The signal shapes for $m_{\ell+D_s^-}$ are derived using Monte Carlo samples. They are expressed by a sum of four Gaussian distributions. An example of a signal distribution is illustrated in Figure 5.2 for $B_s^0 \rightarrow e^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ decays in the B triggered sample. The four Gaussian fit is overlaid.

$m_{\ell+D_s^-}$ PDF for Combinatorial Background

The $m_{\ell+D_s^-}$ template for combinatorial background is determined by fitting the $m_{\ell+D_s^-}$ distribution for the candidates obtained from the D_s^- mass sidebands, as defined in Table 5.4. The parameterization used in this case is obtained by a sum of two tailed Gaussians together with another Gaussian. This is illustrated in Figure 5.3 (left) for the combinatorial background candidates obtained from B triggers in the $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ mode. The combinatorial background candidates preferentially populate low $m_{\ell+D_s^-}$ values.

Some of the parameters obtained above for the PDFs describing the combinatorial background are left free in the final D_s^- mass and $m_{\ell+D_s^-}$ simultaneous fit to the B_s^0 candidates. This is done in order to accommodate changes in the behavior of the combinatorial background underneath the D_s^- mass peak with respect to the background behavior in the sidebands. The floating parameters include the most probable values of the two tailed Gaussians along with the mean of the additional Gaussian.

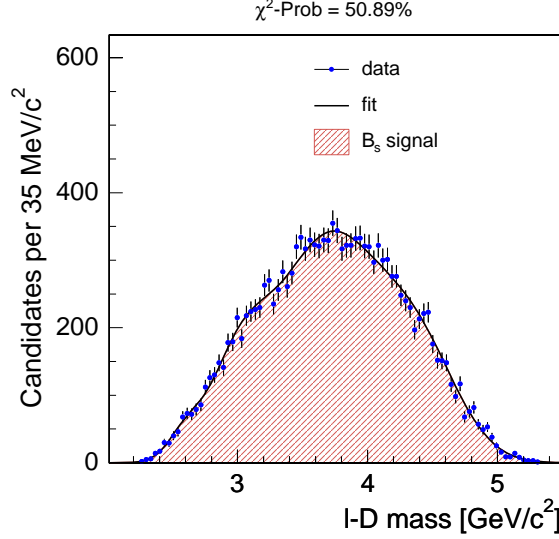


Figure 5.2: $m_{e^+D_s^-}$ distribution for B triggers in $B_s^0 \rightarrow e^+D_s^-X$, $D_s^- \rightarrow \phi\pi^-$, $\phi \rightarrow K^+K^-$ signal Monte Carlo events.

$m_{\ell^+D_s^-}$ PDF for D^- Reflection Background

The $m_{\ell^+D_s^-}$ templates for the D^- reflection background are obtained from Monte Carlo simulated events. The D^- reflection background originates from real B meson decays, and consequently is expected to populate regions very similar to the $B_s^0 \rightarrow D_s^- \ell^+ X$ signal in the invariant mass distribution of the lepton and the D meson system. We express the D^- reflection background shape in $m_{\ell^+D_s^-}$ by a sum of four Gaussian distributions (same parameterization as the signal). Figure 5.3 (right) shows an example of the $m_{\ell^+D_s^-}$ distribution for the D^- reflection background obtained in B triggers from $B \rightarrow \mu^+ D^- X$ Monte Carlo, and reconstructed as $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$ decays using our signal selection criteria, where the four Gaussian fit is also overlaid.

$m_{\ell^+D_s^-}$ PDF for Physics Background

The $m_{\ell^+D_s^-}$ shapes for the various sources of physics background (see Section 5.3.2) are obtained from the corresponding Monte Carlo samples. The shapes for these backgrounds are described by a sum of three tailed Gaussians. Figure 5.4 shows an example of a $m_{\ell^+D_s^-}$ distribution obtained from Monte Carlo for the different sources of physics background reconstructed as B triggers in $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$ candidates. The fit using the three tailed Gaussian combination is overlaid. The different sources of physics background have one common feature: they mostly occupy the low $m_{\ell^+D_s^-}$ region in contrast to the signal B_s^0 events.

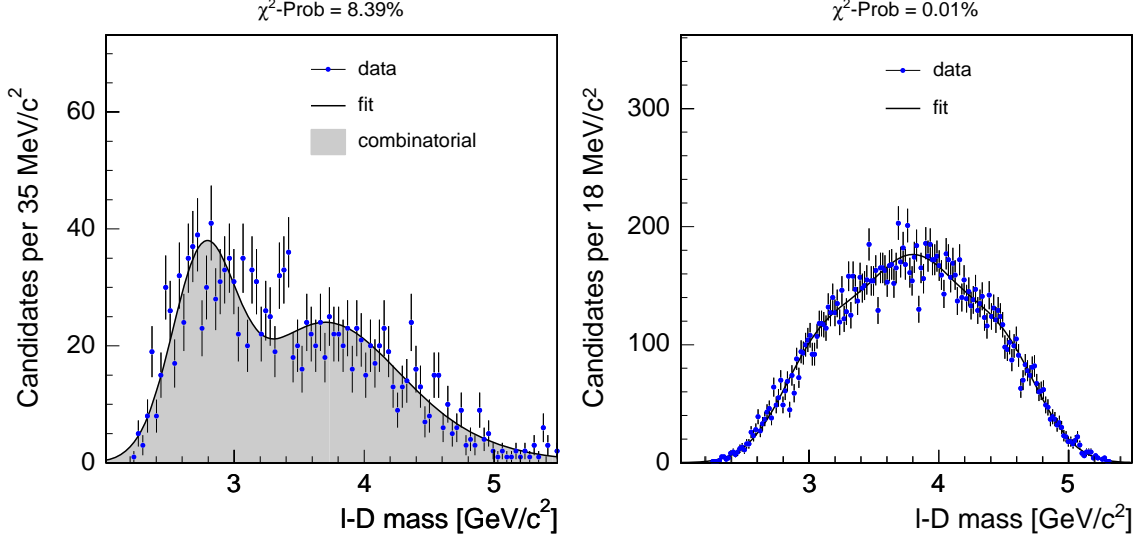


Figure 5.3: $m_{\ell^+ D_s^-}$ distribution for combinatorial background in B triggers in $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ candidates using D_s^- mass sideband candidates (left). $m_{\ell^+ D_s^-}$ distribution for the D^- reflection background in B triggers in the $B \rightarrow \mu^+ D^- X$ Monte Carlo. These decays are reconstructed as $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$ (right).

$m_{\ell^+ D_s^-}$ PDF for Fake Lepton Background

The $m_{\ell^+ D_s^-}$ template for the fake lepton background is obtained from the fake lepton samples (see Section 3.4) by applying an anti-selection on the lepton likelihood ($\mathcal{L}_{e,\mu} < 0.05$). Such a sample is expected to describe both fake leptons and prompt real leptons paired with a real D_s^- meson, as described above. The fit model we employ for describing this class of background is a sum of two tailed Gaussians together with another Gaussian distribution.

Figure 5.5 illustrates the determination of the fake lepton shape in $m_{\ell^+ D_s^-}$. First, the combinatorial background shape is extracted from the D_s^- sidebands in the fake lepton samples. Then, a fit to the entire D_s^- mass region is performed, obtaining the parameters describing the fake lepton background behavior in $m_{\ell^+ D_s^-}$. The plot shown in Figure 5.5 is for the $B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ mode, combining B , D and ℓ +SVT triggers.

Fake lepton background has a strong preference for low $m_{\ell^+ D_s^-}$; a feature that should enable us to discriminate against them by utilizing the $m_{\ell^+ D_s^-}$ distribution. After deriving the shape of the fake lepton background, we are ready to fit the $m_{\ell^+ D_s^-}$ distribution of B_s^0 candidates in data.

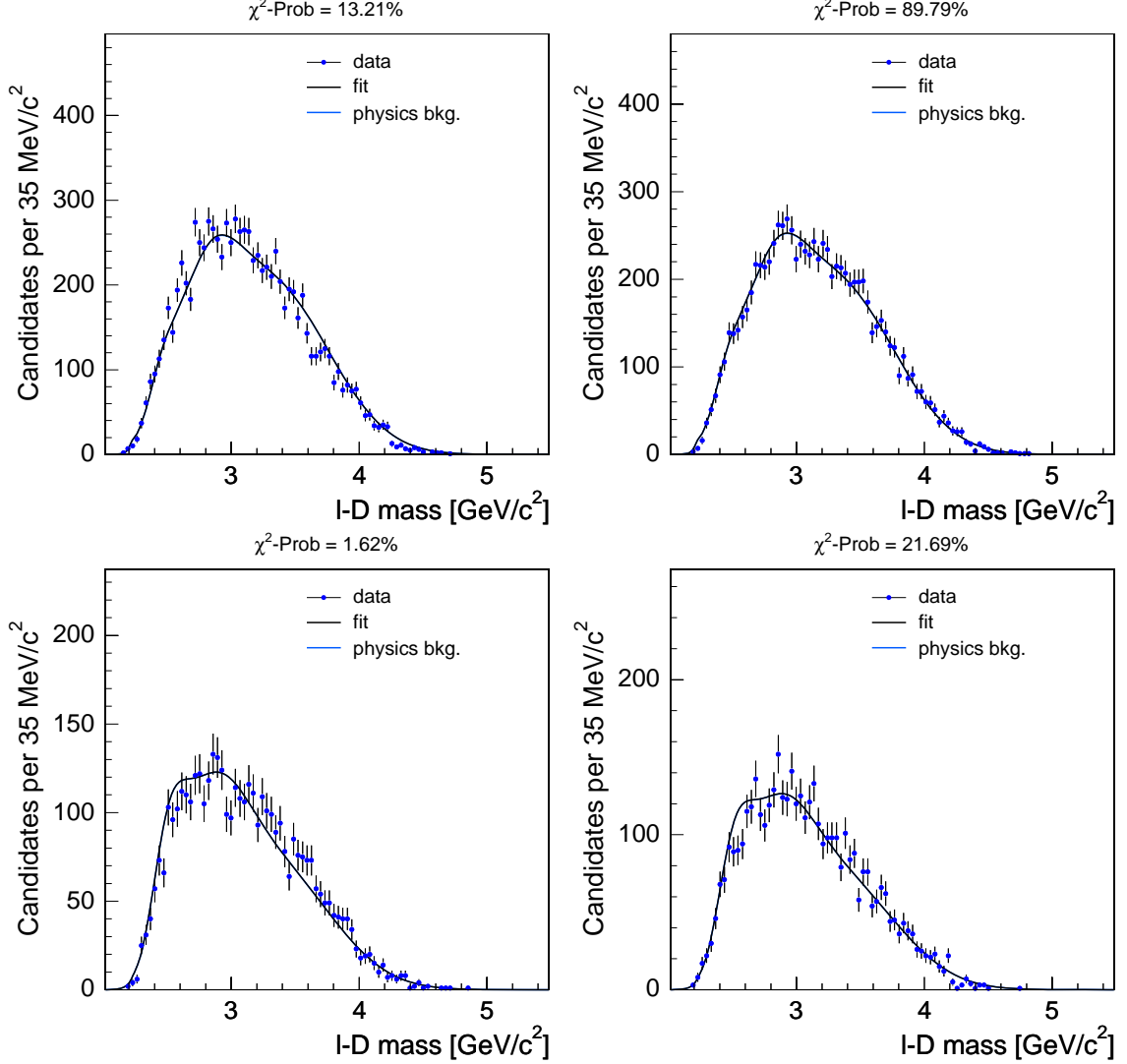


Figure 5.4: $m_{\ell^+ D_s^-}$ distribution for various sources of physics background in B triggers using Monte Carlo simulated events, reconstructed as $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$. The top-left plot is for physics backgrounds from $B^+ \rightarrow D_s^{(*)} D^{(*)} X$, $D^{(*)} \rightarrow \mu^+ Y$ decays, the top-right plot is for $B^0 \rightarrow D_s^{(*)} D^{(*)} X$, $D^{(*)} \rightarrow \mu^+ Y$ decays, the bottom-left plot is for $B_s^0 \rightarrow D_s^{(*)} D^{(*)} X$, $D^{(*)} \rightarrow \mu^+ Y$ decays, and the bottom-right plot is for the physics background shape for $B_s^0 \rightarrow D_s^{(*)} D_s^{(*)} X$, $D_s^{(*)} \rightarrow \mu^+ Y$ decays feeding into our selected B_s^0 candidates.

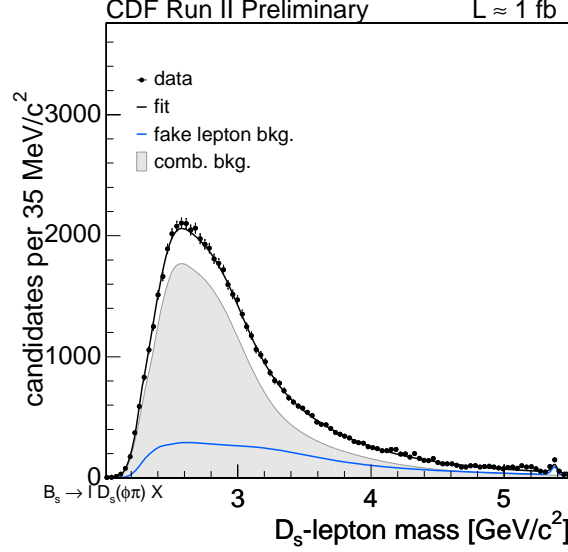


Figure 5.5: $m_{\ell+D_s^-}$ distribution for fake lepton background obtained in the $B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ candidates.

$m_{\ell+D_s^-}$ Fit in Data

We have already determined the contributions of combinatorial, physics, and D^- reflection backgrounds in our B_s^0 samples. We have also obtained parameterizations for the description of the true B_s^0 signal along with various kinds of backgrounds in the two pivotal discriminating variables: D_s^- invariant mass and $m_{\ell+D_s^-}$. The next step is to determine the fraction of fake lepton background and the final descriptions of the various component PDFs in the $m_{\ell+D_s^-}$ space. The candidate likelihood in $m_{\ell+D_s^-}$ space for the B_s^0 samples can be written as

$$\begin{aligned}
 L(m_{\ell+D_s^-}) &= f_{\text{comb}} \cdot \mathcal{P}^{\text{comb}}(m_{\ell+D_s^-}) + f_{\text{phys}} \cdot \mathcal{P}^{\text{phys}}(m_{\ell+D_s^-}) \\
 &\quad + f_{\text{fake}} \cdot \mathcal{P}^{\text{fake}}(m_{\ell+D_s^-}) + f_{\text{refl}} \cdot \mathcal{P}^{\text{refl}}(m_{\ell+D_s^-}) \\
 &\quad + (1 - f_{\text{comb}} - f_{\text{phys}} - f_{\text{fake}} - f_{\text{refl}}) \cdot \mathcal{P}^{\text{signal}}(m_{\ell+D_s^-}), \quad (5.11)
 \end{aligned}$$

where f_i are the fractions of the various signal and background components of the B_s^0 samples, and $\mathcal{P}^i(m_{\ell+D_s^-})$ are the corresponding PDFs. The overall candidate likelihood in the D_s^- mass and $m_{\ell+D_s^-}$ spaces is expressed as

$$\begin{aligned}
 L(m, m_{\ell+D_s^-}) &= \sum_i f_i \cdot \mathcal{P}^i(m, m_{\ell+D_s^-}) \\
 &= \sum_i f_i \cdot \mathcal{P}^i(m) \cdot \mathcal{P}^i(m_{\ell+D_s^-}), \quad (5.12)
 \end{aligned}$$

where $\mathcal{P}^i(m)$ and $\mathcal{P}^i(m_{\ell+D_s^-})$ are the individual PDFs for the i^{th} component in the D_s^- mass and $m_{\ell+D_s^-}$ spaces, respectively. Using this model for the likelihood, we perform

a simultaneous fit in the D_s^- mass and $m_{\ell+D_s^-}$ variables in the different subsamples to obtain the fake lepton background contributions. While fitting for the fraction of the fake lepton background, some of the parameters obtained for the PDFs describing the combinatorial background in the $m_{\ell+D_s^-}$ space are allowed to vary as mentioned above. These include the most probable values of the two tailed Gaussians along with the mean of the additional Gaussian. All the parameterizations in the D_s^- mass space are fixed to their determined values, as detailed in Section 5.3.1. The obtained fractions of fake lepton backgrounds are summarized in Table 5.7. The final PDFs for the signal along with the four kinds of backgrounds in $m_{\ell+D_s^-}$ space are hence determined for usage in the next step of analysis of the B_s^0 candidates.

Decay Samples	$\phi\pi^-$ [%]	$K^{*0}K^-$ [%]	$\pi^+\pi^-\pi^-$ [%]
	Period 1		
$\mu D_s, B$ trigger	10.0	1.9	2.5
$e D_s, B$ trigger	4.9	0.0	2.0
$\mu D_s, D$ trigger	11.3	3.4	15.2
$e D_s, D$ trigger	14.4	2.7	7.4
$\mu D_s, \ell + SVT$ trigger	0.6	0.3	1.1
$e D_s, \ell + SVT$ trigger	1.0	0.9	0.9
	Period 2		
$\mu D_s, B$ trigger	5.4	0.2	2.5
$e D_s, B$ trigger	2.5	0.0	2.0
$\mu D_s, D$ trigger	6.3	2.8	15.2
$e D_s, D$ trigger	14.4	2.4	7.4
$\mu D_s, \ell + SVT$ trigger	0.4	0.4	1.1
$e D_s, \ell + SVT$ trigger	0.9	0.3	0.6
	Period 3		
$\mu D_s, B$ trigger	2.9	0.0	2.5
$e D_s, B$ trigger	0.2	0.0	2.0
$\mu D_s, D$ trigger	5.6	3.1	15.2
$e D_s, D$ trigger	12.3	5.9	7.4
$\mu D_s, \ell + SVT$ trigger	0.4	0.3	1.1
$e D_s, \ell + SVT$ trigger	1.0	0.4	0.8

Table 5.7: Fractions of fake lepton background relative to the sum of signal, physics and prompt backgrounds in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates in the three data-taking periods. The relative uncertainties in these fractions are $\sim 50\%$ each.

5.3.4 Summary of Characterizations in D_s^- Mass and $m_{\ell+D_s^-}$

We have completed determining the contributions of all five components of the semileptonic B_s^0 samples using simultaneous fits in the D_s^- mass and $m_{\ell+D_s^-}$ variables. The reconstructed D_s^- invariant mass distribution for the B_s^0 candidates enables us to distinguish against combinatorial and D^- reflection backgrounds. The quantity $m_{\ell+D_s^-}$ discriminates B_s^0 signal primarily from combinatorial, physics and fake lepton backgrounds. The fits in the D_s^- mass yield measurements of the fractions of combinatorial and D^- reflection backgrounds, while simultaneously fitting in the D_s^- mass and $m_{\ell+D_s^-}$ spaces provides a measurement of the contribution from fake lepton background. The fraction of physics background is estimated using Monte Carlo samples. In the process, we have also extracted the PDFs describing the signal, and combinatorial, physics, D^- reflection, and fake lepton backgrounds in the D_s^- mass and $m_{\ell+D_s^-}$ spaces.

Having determined the fractions of all the components of our B_s^0 samples, we can now quote the amount of true semileptonic $B_s^0 \rightarrow D_s^- \ell^+ X$ signal in our data. Table 5.8 summarizes the signal content of all the B_s^0 samples for different topologies and data-taking periods. We have $\sim 29,600$ $D_s^- \rightarrow \phi\pi^-$, $\sim 22,000$ $D_s^- \rightarrow K^{*0}K^-$, and $\sim 9,900$ $D_s^- \rightarrow \pi^+\pi^-\pi^-$ decays available for analyzing B_s^0 - \bar{B}_s^0 oscillations. The fractions of the three trigger topologies are $B : D : \ell+\text{SVT} \sim 0.7 : 0.2 : 0.1$, while the relative contribution of the two lepton species in $B_s^0 \rightarrow \ell^+ D_s^- X$ decays is $e : \mu \sim 40 : 60$. Thus, our B_s^0 samples are dominated by the B triggers where the lepton is a displaced SVT track. Figures 5.6 to 5.14 show the distributions of the D_s^- mass and $m_{\ell+D_s^-}$ in the three B_s^0 modes for the three trigger topologies, with fits of the signal and different background components overlaid.

Decay	$\phi\pi^-$	$K^{*0}K^-$	$\pi^+\pi^-\pi^-$
Samples	Period 1		
$\mu D_s, B$ trigger	4225	3361	1307
$e D_s, B$ trigger	2468	1979	823
$\mu D_s, D$ trigger	964	788	511
$e D_s, D$ trigger	500	493	322
$\mu D_s, \ell + SVT$ trigger	1092	423	454
$e D_s, \ell + SVT$ trigger	807	370	258
	Period 2		
$\mu D_s, B$ trigger	5706	4705	1827
$e D_s, B$ trigger	3353	2623	937
$\mu D_s, D$ trigger	1166	922	436
$e D_s, D$ trigger	584	599	282
$\mu D_s, \ell + SVT$ trigger	1477	631	530
$e D_s, \ell + SVT$ trigger	879	404	258
	Period 3		
$\mu D_s, B$ trigger	2519	2089	735
$e D_s, B$ trigger	1530	1283	343
$\mu D_s, D$ trigger	448	414	240
$e D_s, D$ trigger	255	235	155
$\mu D_s, \ell + SVT$ trigger	1038	448	329
$e D_s, \ell + SVT$ trigger	563	270	152

Table 5.8: Number of true semileptonic $B_s^0 \rightarrow \ell^+ D_s^- X$ signal events in the three decay modes, after taking into account the contributions from all the backgrounds.

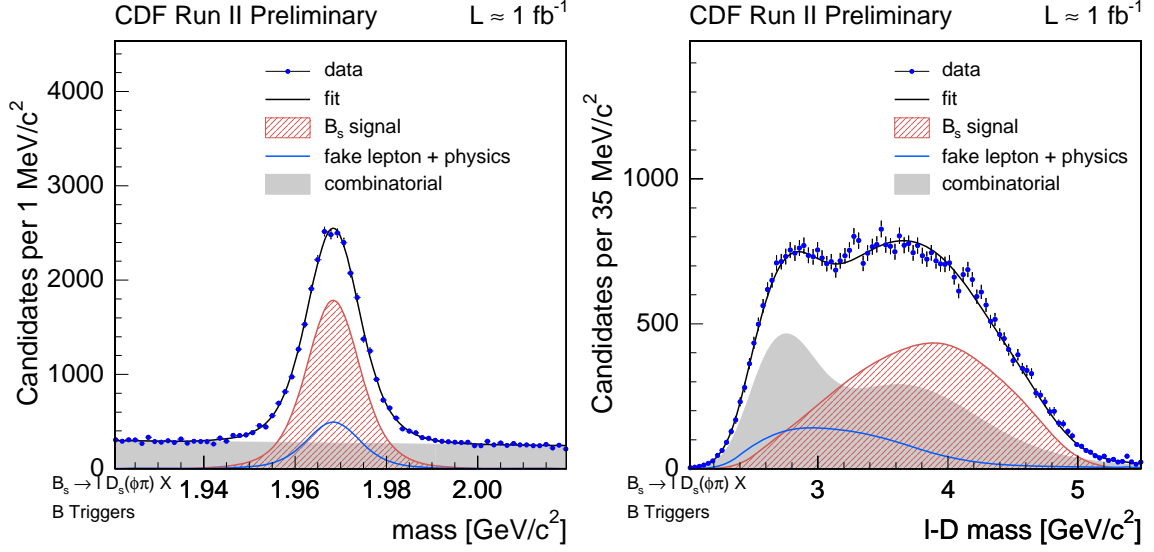


Figure 5.6: Fits of the D_s^- mass and $m_{\ell+D_s^-}$ distributions for $B \rightarrow \ell D_s(\phi\pi)$ in B triggers.

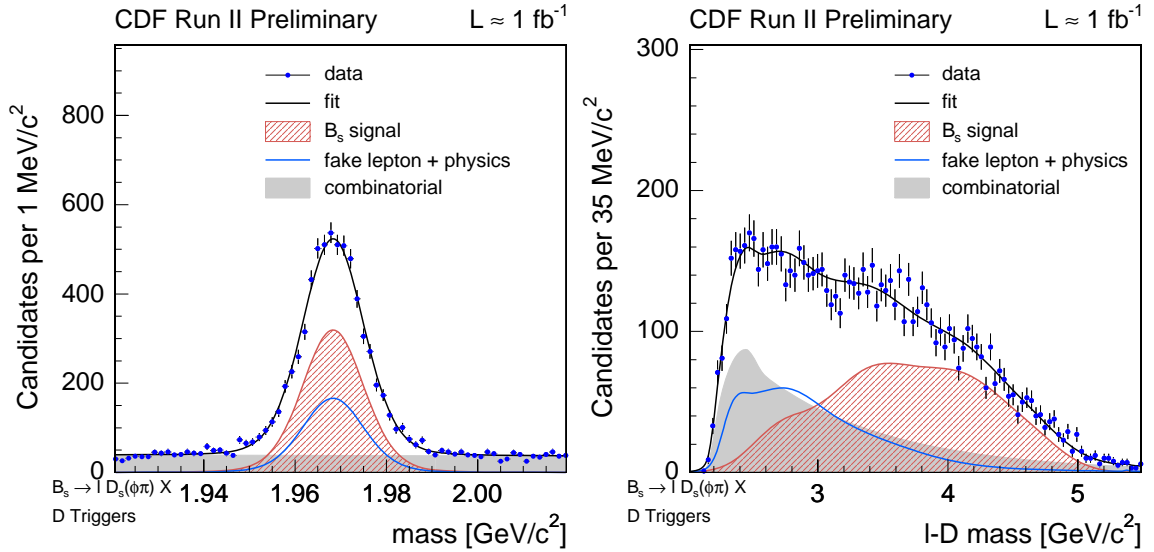


Figure 5.7: Fits of the D_s^- mass and $m_{\ell+D_s^-}$ distributions for $B \rightarrow \ell D_s(\phi\pi)$ candidates in D triggers.

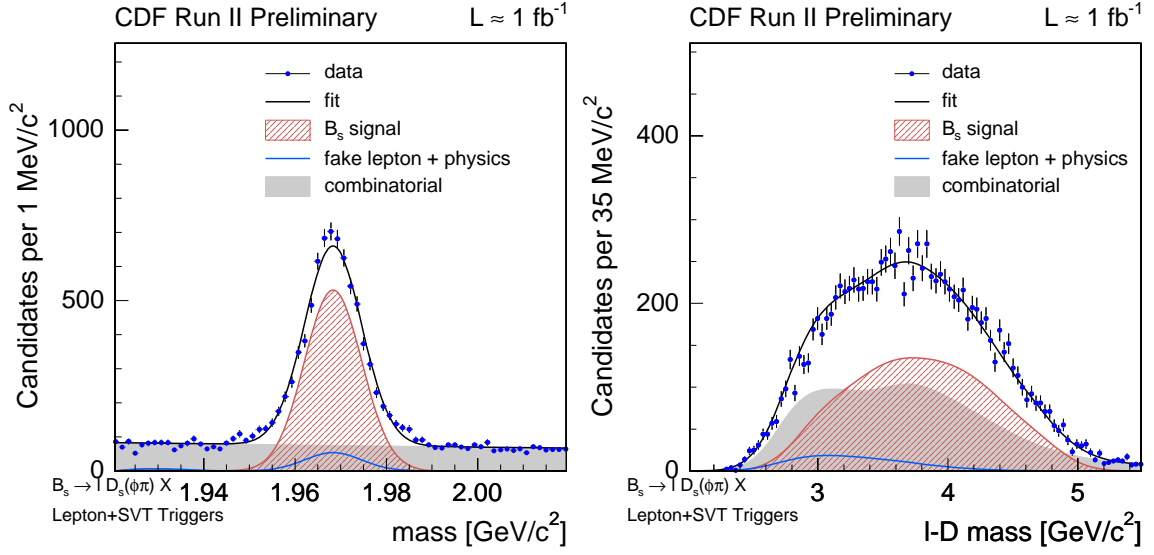


Figure 5.8: Fits of the D_s^- mass and $m_{\ell+D_s^-}$ distributions for $B \rightarrow \ell D_s(\phi\pi)$ candidates in ℓ +SVT triggers.

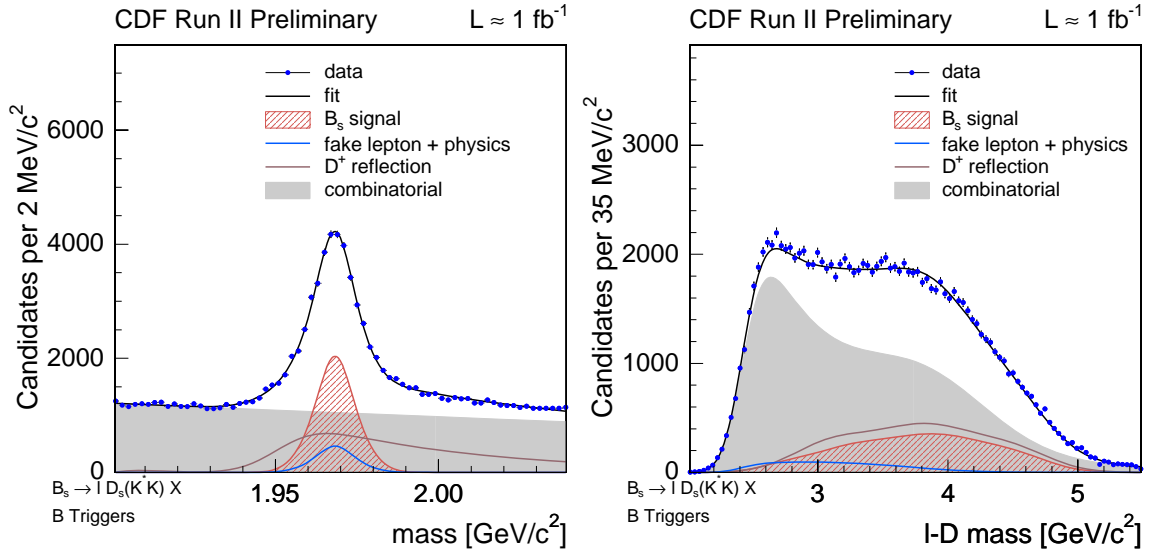


Figure 5.9: Fits of the D_s^- mass and $m_{\ell+D_s^-}$ distributions for $B \rightarrow \ell D_s(K^*K)$ candidates in B triggers.

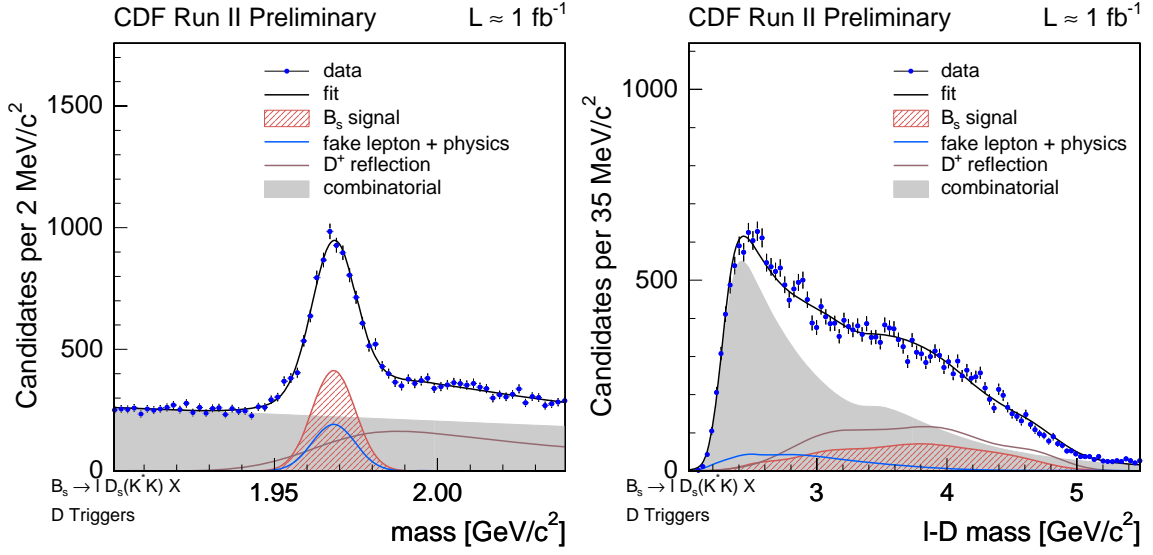


Figure 5.10: Fits of the D_s^- mass and $m_{\ell+D_s^-}$ distributions for $B \rightarrow \ell D_s(K^* K)$ candidates in D triggers.

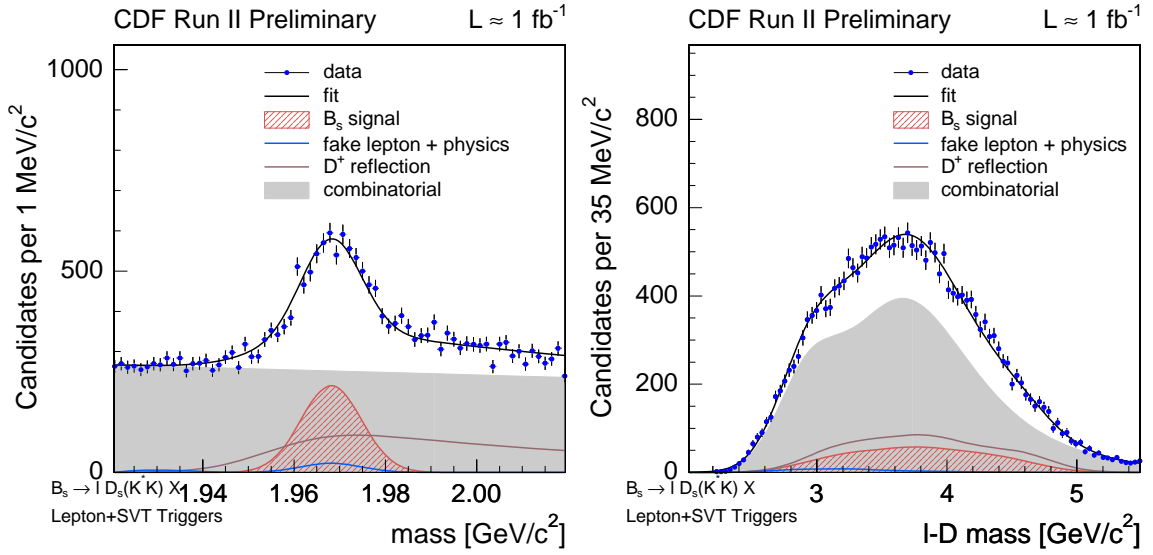


Figure 5.11: Fits of the D_s^- mass and $m_{\ell+D_s^-}$ distributions for $B \rightarrow \ell D_s(K^* K)$ candidates in ℓ +SVT triggers.

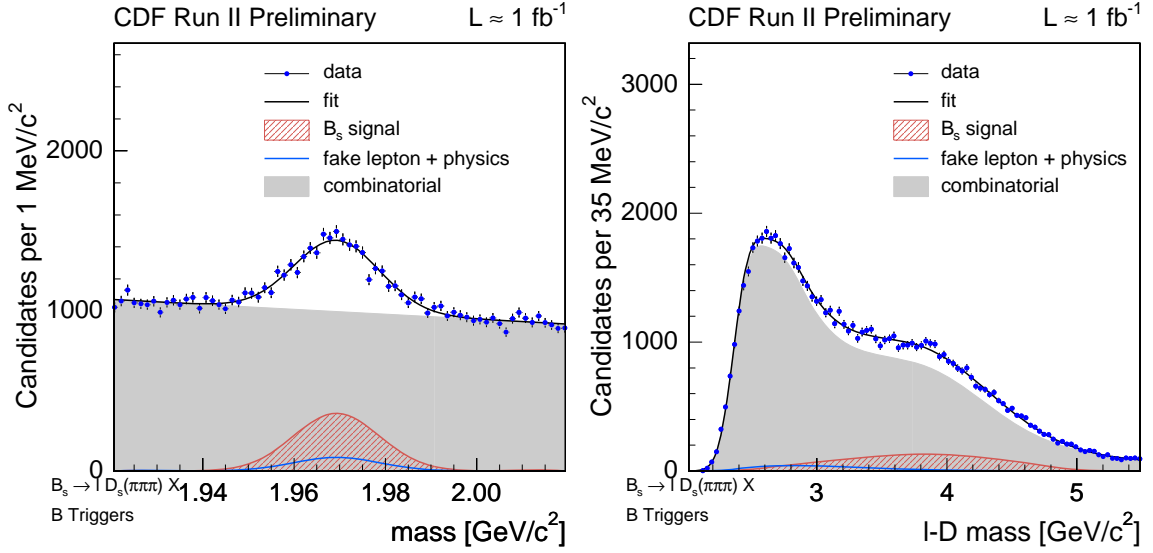


Figure 5.12: Fits of the D_s^- mass and $m_{\ell+D_s^-}$ distributions for $B \rightarrow \ell D_s(3\pi)$ candidates in B triggers.

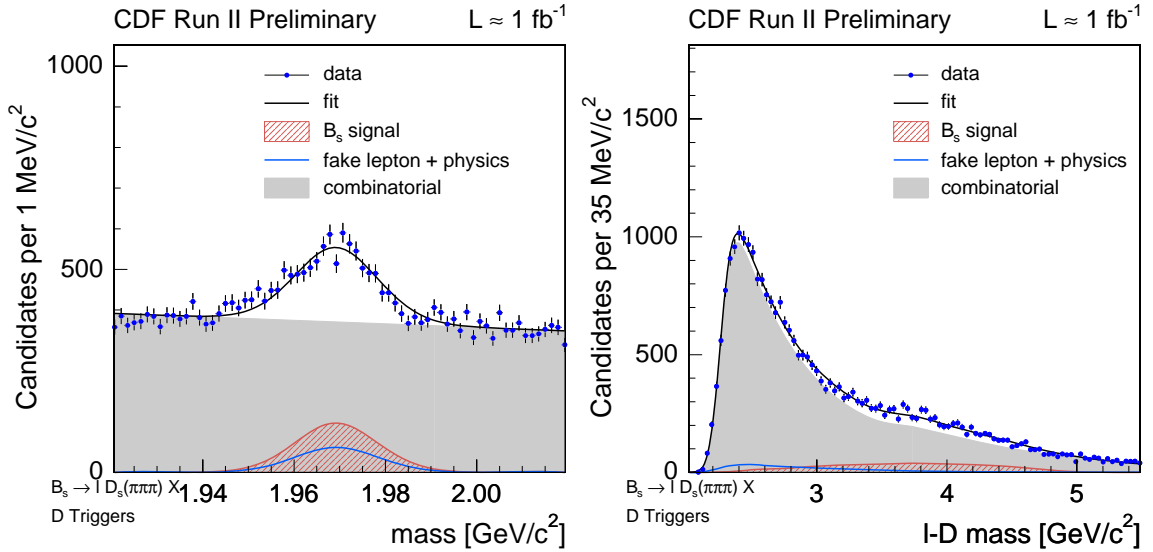


Figure 5.13: Fits of the D_s^- mass and $m_{\ell+D_s^-}$ distributions for $B \rightarrow \ell D_s(3\pi)$ candidates in D triggers.

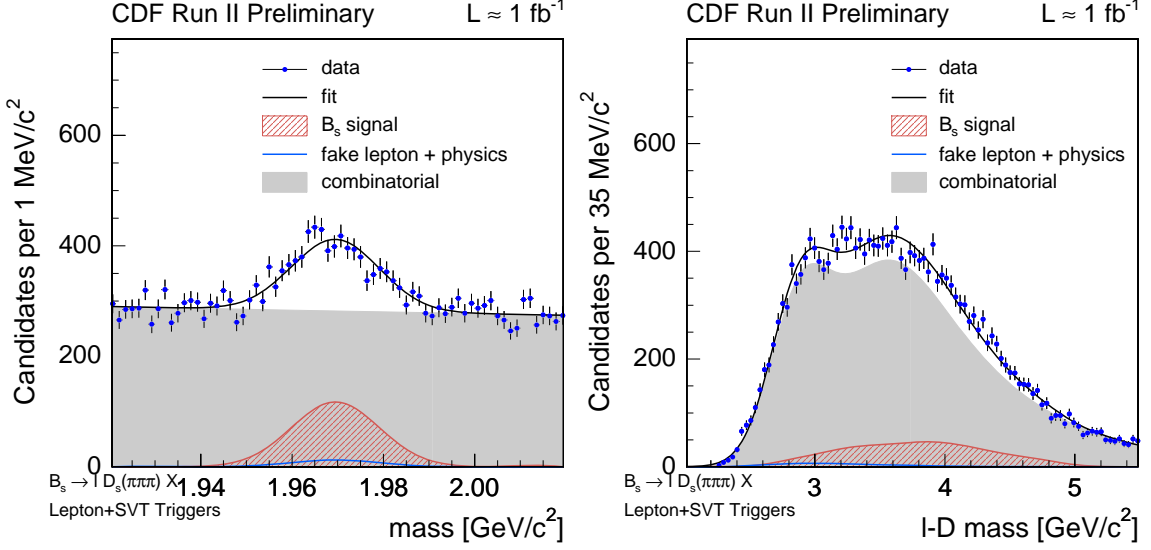


Figure 5.14: Fits of the D_s^- mass and $m_{\ell+D_s^-}$ distributions for $B \rightarrow \ell D_s(3\pi)$ candidates in ℓ +SVT triggers.

5.3.5 Decay Time Measurements in Semileptonic B_s^0 Decays

A quantity of central importance in performing time-dependent measurements of B_s^0 meson properties is the decay time distribution of the B_s^0 mesons. As described in Section 4.1, the loss of the neutrino in the reconstruction of semileptonic B_s^0 meson decays complicates the measurement of the proper decay time. The “visible” time in semileptonic decays is called the pseudo-proper decay time, and defined in Equation (4.2) as

$$t^* = L_{xy}^{B_s^0} \frac{M^{B_s^0}}{p_T^{\ell^+ D_s^-}}.$$

The proper decay time t defined in Equation (4.1) is related to t^* by a correction factor κ defined as follows:

$$\kappa = \frac{D_{xy}^{B_s^0} \cdot p_T^{\ell^+ D_s^-}}{L_{xy}^{B_s^0} \cdot p_T^{B_s^0}}, \quad (5.13)$$

such that $t = t^* \cdot \kappa$. The kinematic correction factor κ is obtained in form of a distribution $\mathcal{H}(\kappa)$ derived from the Monte Carlo simulation.

The mass of the B_s^0 meson used in the calculation of t^* is the nominal mass obtained from Ref. [13]. A necessary component of t^* determination is the transverse decay length $L_{xy}^{B_s^0}$. The quantity $L_{xy}^{B_s^0}$ is a signed variable which can be negative for the configuration, where a particle seems to decay before the point of its production. To determine the transverse coordinates of the primary vertex (V_P in Figure 4.1) and the transverse coordinates of the secondary vertex (V_{B_s} in Figure 4.1) need to be

known. The uncertainties in the determination of these two positions contribute to the uncertainties in t^* . The error on the transverse decay length is returned by the CTVMFT vertex fit and is denoted by $\sigma_{L_{xy}^{B_s^0}}$. The resolution on the pseudo-proper decay time is denoted by σ_{t^*} , and is deduced from the error $\sigma_{L_{xy}^{B_s^0}}$ on $L_{xy}^{B_s^0}$ as follows:

$$\sigma_{t^*} = \sigma_{L_{xy}^{B_s^0}} \frac{M^{B_s^0}}{p_T^{\ell^+ D_s^-}}. \quad (5.14)$$

The error on the transverse decay length determination depends on the understanding of the track parameter errors, which in turn depend on the comprehension of hit resolutions in the tracking detectors. This results in $\sigma_{L_{xy}^{B_s^0}}$ being underestimated as returned by the CTVMFT fit. Therefore, a scaling factor is applied to $\sigma_{L_{xy}^{B_s^0}}$ for each candidate to properly describe the decay length resolution. As explained above, the quantity $\sigma_{L_{xy}^{B_s^0}}$ has contributions from the determination of the primary and secondary vertices. We describe the determination of the primary vertex and its uncertainties, followed by a description of the deduction of the secondary vertex uncertainties.

Uncertainties in Primary Vertex Determination

The coordinates of the primary vertex are determined for each event using a CTVMFT vertex fit to the prompt tracks from the underlying event that surround the B candidate. Tracks that seem to originate from multiple interactions or other secondary vertices are removed from the fit. To obtain a measurement of uncertainty from the primary vertex determination, the tracks used in the vertex fit are divided into two subsamples and a primary vertex position is fit for each subsample. A comparison is performed between the consistency of primary vertex positions returned by the two fits, and the fit errors returned by CTVMFT. This procedure unveils that the primary vertex errors returned by CTVMFT are systematically underestimated by a factor of 1.38. The uncertainties on the primary vertex position used in the CTVMFT fit are thus scaled by this factor.

Uncertainties in Secondary Vertex Determination

To determine whether the uncertainties on the secondary vertex position are correctly determined by CTVMFT, prompt D^- mesons are combined with another randomly selected track in the event using the CTVMFT vertex fit. The selection requirements imposed on such D^- +track combinations mimic the signal selection criteria used in this measurement. The distribution of the proper decay time measured in this sample is shown in Figure 5.15 (left), and is dominated by a prompt component around $t = 0$. Thus, the true error on the proper time measurement can be determined from the width of the prompt component and compared to the vertex position errors returned

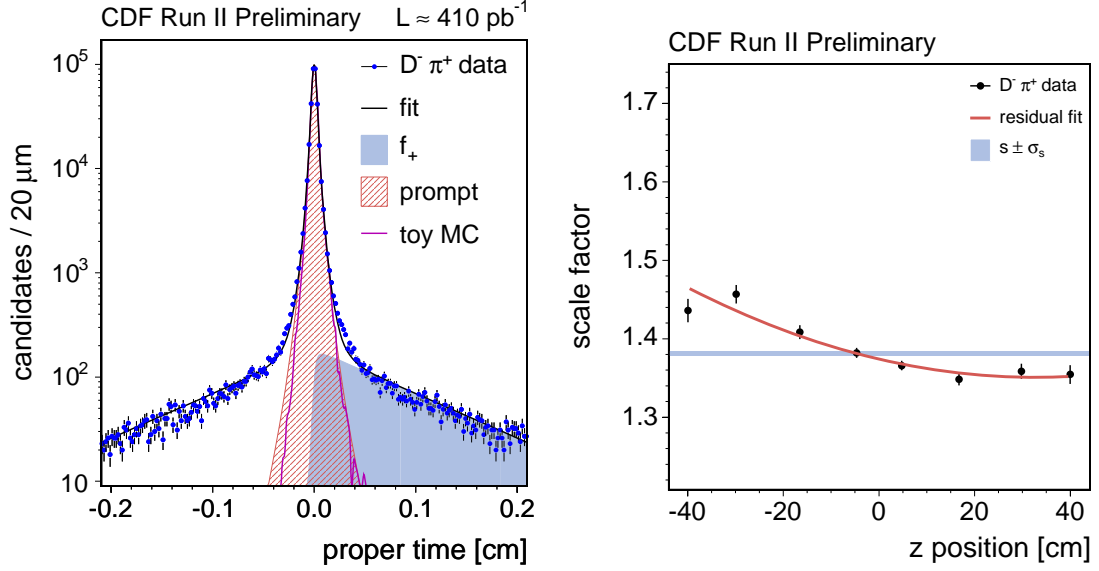


Figure 5.15: Proper time distribution measured in the D^- +track calibration sample (left). Variation of the resolution scale factor with the z -coordinate of the reconstructed D^- +track vertex (right).

from the CTVMFT fit which includes the primary vertex contribution. This comparison exhibits that the decay length uncertainties returned by the CTVMFT fit require to be scaled by a factor that depends on kinematic properties of the decay. The scale factor is therefore parameterized as a function of various kinematical variables like $\Delta R(D, \ell)$, η , z , χ_{xy}^2 and isolation (the ratio between the p_T of the B_s^0 candidate and the p_T of all tracks in a cone $\Delta R < 0.7$ around the B_s^0 candidate direction). An example of this parameterization with respect to the z -position of the D^- +track is displayed in Figure 5.15 (right). This parameterization is used to determine the $\sigma_{L_{xy}^B}$ scaling factor for each B_s^0 meson candidate. These scale factors typically vary between 1.35 and 1.65.

From here on, unless specified otherwise, the quantities $\sigma_{L_{xy}^B}$ and σ_{t^*} refer to the scaled quantities and correspond to the true errors on the transverse decay length and pseudo-proper decay time, respectively.

In the following, we describe the procedure to extract signal PDFs in t^* for the semileptonic decays of B_s^0 mesons.

Pseudo-Proper Decay Time PDF for Signal

In the absence of measurement errors, reconstruction biases, and missing particles, the proper decay time of a B meson is governed by an exponential distribution:

$$E(t; \tau) = \frac{1}{\tau} e^{-\frac{t}{\tau}} \theta(t) \quad (5.15)$$

However, we do not fully reconstruct the semileptonic decays of the B mesons, and we are limited by the finite detector resolution. Furthermore, our samples are biased by trigger and reconstruction requirements. Therefore, we need to take these effects into account in order to understand the behavior of real B mesons in the pseudo-proper decay time variable. First, we include the effect of incomplete reconstruction, i.e. the missing neutrino and other neutral particles, via a distribution ($\mathcal{H}(\kappa)$) of the kinematic correction factor κ defined in Equation (5.13). The description of the proper decay time in Equation (5.15) now becomes (using $t = \kappa \cdot t^*$):

$$\begin{aligned} F(t^*; \tau) &= E(\kappa t^*; \tau) \otimes_{\kappa} \mathcal{H}(\kappa) \\ &= \int \frac{\kappa}{\tau} e^{-\frac{\kappa t^*}{\tau}} \theta(\kappa t^*) \mathcal{H}(\kappa) d\kappa, \end{aligned} \quad (5.16)$$

where the operation \otimes_{κ} represents a convolution in κ space. Next, we model the effect arising from the finite detector resolution via a Gaussian distribution and include in the description of the signal PDF:

$$\begin{aligned} I(t^*; \sigma_{t^*}, \tau) &= E(\kappa t^*; \tau) \otimes_{\kappa} \mathcal{H}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \\ &= E(\kappa t^*; \tau) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \otimes_{\kappa} \mathcal{H}(\kappa) \\ &= \left[\int_{-\infty}^{+\infty} \frac{\kappa}{\tau} e^{-\frac{\kappa t'}{\tau}} \theta(\kappa t') \frac{1}{\sqrt{2\pi}\sigma_{t^*}} e^{-\frac{1}{2}\left(\frac{t' - t^*}{\sigma_{t^*}}\right)^2} dt' \right] \otimes_{\kappa} \mathcal{H}(\kappa) \\ &= \left[\frac{\kappa}{2\tau} e^{-\frac{\kappa}{\tau}\left(t^* - \frac{\kappa\sigma_{t^*}^2}{2\tau}\right)} \text{Erfc}\left(\frac{\kappa\sigma_{t^*}^2 - t^*\tau}{\sqrt{2}\sigma_{t^*}\tau}\right) \right] \otimes_{\kappa} \mathcal{H}(\kappa) \\ &= \int \frac{\kappa}{2\tau} e^{-\frac{\kappa}{\tau}\left(t^* - \frac{\kappa\sigma_{t^*}^2}{2\tau}\right)} \text{Erfc}\left(\frac{\kappa\sigma_{t^*}^2 - t^*\tau}{\sqrt{2}\sigma_{t^*}\tau}\right) \mathcal{H}(\kappa) d\kappa \end{aligned} \quad (5.17)$$

where $\text{Erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp(-u^2) du$ is the complimentary error function. Finally, after including the effects from missing particles and detector resolution, we include the sculpting of the decay time by trigger and reconstruction requirements. This is modeled through an efficiency function $\xi(t^*)$. We define the signal PDF in the pseudo-proper decay time space as:

$$\begin{aligned} \mathcal{P}(t^*; \sigma_{t^*}, \tau) &= \frac{1}{N} I(t^*; \sigma_{t^*}, \tau) \cdot \xi(t^*) \\ &= \frac{1}{N} \int \frac{\kappa}{2\tau} e^{-\frac{\kappa}{\tau}\left(t^* - \frac{\kappa\sigma_{t^*}^2}{2\tau}\right)} \text{Erfc}\left(\frac{\kappa\sigma_{t^*}^2 - t^*\tau}{\sqrt{2}\sigma_{t^*}\tau}\right) \mathcal{H}(\kappa) d\kappa \cdot \xi(t^*), \end{aligned} \quad (5.18)$$

where N is the normalization given by:

$$N(\sigma_{t^*}, \tau) = \int \frac{\kappa}{2\tau} \exp\left(\frac{\kappa^2 \sigma_{t^*}^2}{2\tau^2}\right) \left[\int_{-\infty}^{+\infty} e^{-\frac{\kappa}{\tau} t^*} \text{Erfc}\left(\frac{\kappa \sigma_{t^*}^2 - t^* \tau}{\sqrt{2} \sigma_{t^*} \tau}\right) \xi(t^*) dt^* \right] \mathcal{H}(\kappa) d\kappa. \quad (5.19)$$

Having formulated the signal PDF in the pseudo-proper decay time, we now discuss its two important ingredients: the κ factor distribution $\mathcal{H}(\kappa)$ and the efficiency function $\xi(t^*)$ both of which are obtained Monte Carlo.

Signal PDF: Correction for Missing Momentum

As mentioned earlier, the reconstruction of semileptonic decays of the B_s^0 mesons is incomplete because of missing neutral particles that are not reconstructed in the CDF detector. This implies that the missing momentum needs to be corrected for by using a distribution of the κ -factor defined in Equation (5.13). The distributions of κ -factors $\mathcal{H}(\kappa)$ obtained from $B_s^0 \rightarrow \mu^+ D_s^- X$ Monte Carlo samples are shown in Figure 5.16. The distributions are arranged in the following order: $D_s^- \rightarrow \phi \pi^-$ (top), $D_s^- \rightarrow K^{*0} K^-$ (middle), and $D_s^- \rightarrow \pi^+ \pi^- \pi^-$ (bottom) with B triggers on the left, D triggers in the middle, and ℓ +SVT triggers on the right.

The value of the κ -factor for a given $B_s^0 \rightarrow D_s^- \ell^+ X$ signal candidate is not known in the data. Therefore, we apply the missing momentum correction via a distribution of all possible κ -factors. However, the κ -factor distribution is correlated with $m_{\ell^+ D_s^-}$; the closer $m_{\ell^+ D_s^-}$ is to the B_s^0 mass, the smaller is the missing momentum, and consequently, the higher is the value of the average κ -factor. Furthermore, the phase space available for the missing neutral particles is reduced with higher $m_{\ell^+ D_s^-}$, thereby reducing the width of the κ -factor distribution. The uncertainty in the reconstructed proper decay time can be expressed as:

$$\sigma_t = \sigma_t^* \oplus \frac{\sigma_{p_T}^{B_s^0}}{p_T^{\ell^+ D_s^-}} \times t^*. \quad (5.20)$$

The first term is defined in Equation (5.14). It is related to the resolution of the transverse decay length, $\sigma_{L_{xy}}^B$ which is constant with respect to the decay time. The second term grows with increasing decay time and is directly proportional to the momentum uncertainty of the B_s^0 meson. The momentum uncertainty, in turn, is directly proportional to the width of the κ -factor distribution. Thus, semileptonic B_s^0 candidates with higher $m_{\ell^+ D_s^-}$ have a narrower κ -factor distribution, and consequently, a lower proper time uncertainty. Furthermore, the growth in the proper time uncertainty with decay time is also less pronounced for higher $m_{\ell^+ D_s^-}$ candidates. Since the sensitivity to B_s^0 - \bar{B}_s^0 oscillations degrades exponentially with the square of the proper time resolution (see Section 1.5), this implies that B_s^0 candidates with high $m_{\ell^+ D_s^-}$ or small decay time are most sensitive to rapid B_s^0 - \bar{B}_s^0 oscillations.

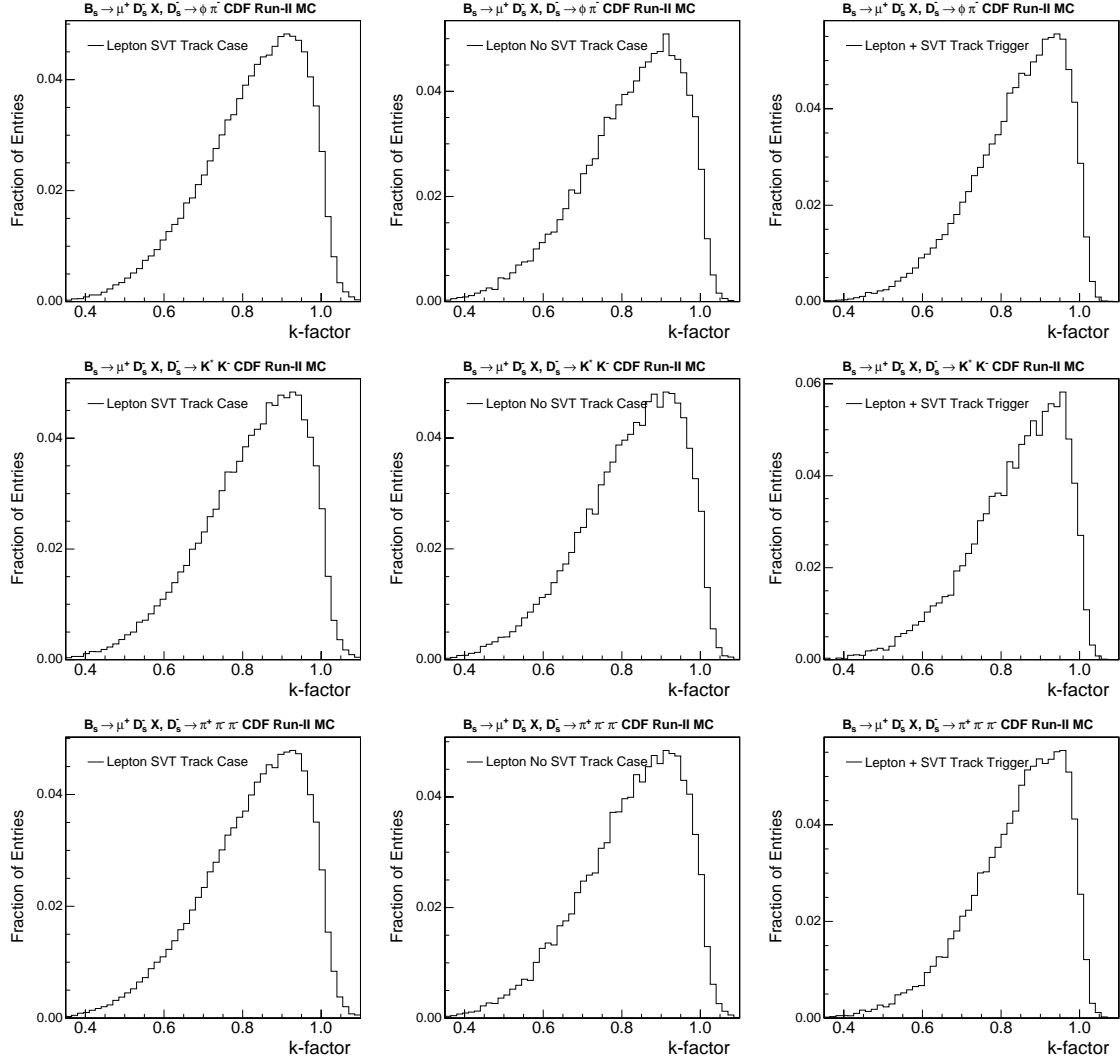


Figure 5.16: κ -factor distributions for $B_s^0 \rightarrow \ell^+ D_s^- X$ decays obtained from B_s^0 signal Monte Carlo samples.

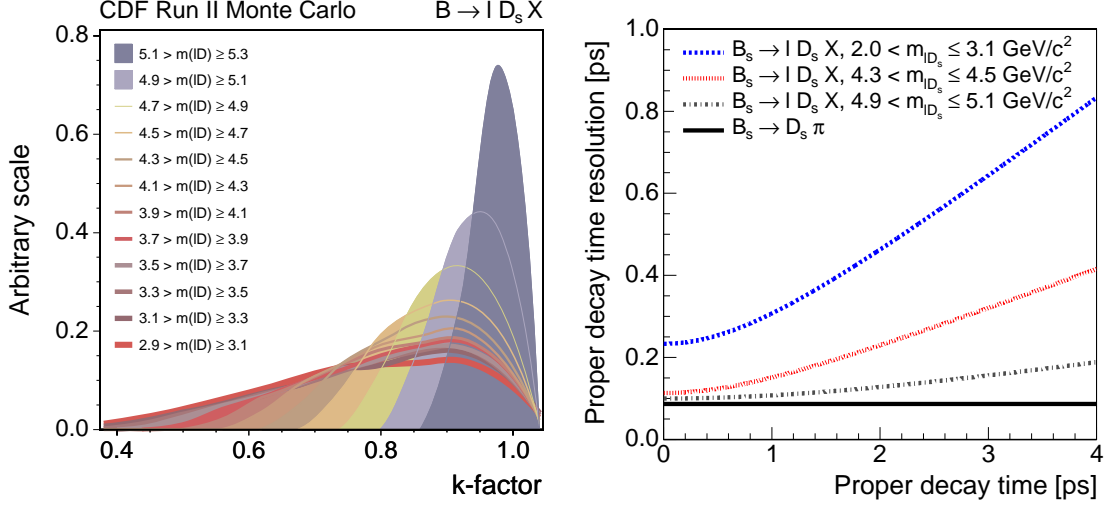


Figure 5.17: k -factor distributions in several $m_{\ell^+ D_s^-}$ ranges for the $B_s^0 \rightarrow \ell^+ D_s^-, D_s^- \rightarrow \phi \pi^-$ decays (left). The corresponding proper decay time uncertainties for B_s^0 candidates in the specified $m_{\ell^+ D_s^-}$ ranges with respect to proper decay time. For comparison, the proper decay time uncertainty for hadronic decays is shown in black (right).

To maximally exploit the dependence of the κ -factor on $m_{\ell^+ D_s^-}$, we derive separate κ -factor distributions in the following 12 ranges in $m_{\ell^+ D_s^-}$:

$$m_{\ell^+ D_s^-} \in (2.0, 3.1], (3.1, 3.3], (3.3, 3.5], (3.5, 3.7], (3.7, 3.9], (3.9, 4.1], \\ (4.1, 4.3], (4.3, 4.5], (4.5, 4.7], (4.7, 4.9], (4.9, 5.1], (5.1, 5.5) \text{ GeV}/c^2.$$

The $\mathcal{H}(\kappa)$ distributions for the various several $m_{\ell^+ D_s^-}$ ranges for $B_s^0 \rightarrow \ell^+ D_s^-, D_s^- \rightarrow \phi \pi^-$ decays are displayed in Figure 5.17 (left). Figure 5.17 (right) illustrates the corresponding proper decay time uncertainties for the B_s^0 semileptonic decays in the specified $m_{\ell^+ D_s^-}$ ranges with respect to proper decay time. For comparison, the proper decay time uncertainty for hadronic $B_s^0 \rightarrow D_s^- \pi^+$ decays is shown in black. The general behavior of B_s^0 mesons with low and high $m_{\ell^+ D_s^-}$, in terms of their proper time uncertainties, are in agreement with the description above.

Signal PDF: Corrections for Trigger and Reconstruction Biases

The effect of trigger and reconstruction selection criteria on B_s^0 candidates is studied via a Monte Carlo simulation of B_s^0 semileptonic decays. The Monte Carlo samples are obtained after complete detector and trigger simulation (see Section 2.9). The selection requirements on the B_s^0 candidates imposed offline are also imposed on the Monte Carlo samples. The corrections for the trigger and reconstruction requirements resulting in decay time biases in the B_s^0 candidates are implemented via the

efficiency function, $\xi(t^*)$. The efficiency function is obtained using the ratio of the pseudo-proper time distribution obtained from B_s^0 Monte Carlo samples after trigger and reconstruction requirements, to the analytically computed distribution of the pseudo-proper time of B_s^0 events obtained in an unbiased manner without any selection bias. For each event passing the selection criteria in the Monte Carlo, the expected t^* distribution without selection bias is obtained by convolving the B_s^0 lifetime exponential with the κ -factor distribution followed by smearing with a Gaussian resolution function. The width of the Gaussian resolution function is obtained from the σ_{t^*} obtained from the Monte Carlo event. Thus, the efficiency function $\xi(t^*)$ is defined as

$$\xi(t^*) \equiv \frac{\text{reconstructed } t^* \text{ after trigger + selection}}{\sum_{\sigma_{t^*}} \frac{1}{\tau} \exp\left(-\frac{t^*}{\tau}\right) \otimes_{t'} G(t' - t^*; \sigma_{t^*}) \otimes_k \mathcal{H}(k)}, \quad (5.21)$$

where τ is the lifetime of the B meson used in Monte Carlo production, and the sum in the denominator is over all possible value of σ_{t^*} obtained from the Monte Carlo events. The efficiency functions $\xi(t^*)$ obtained from the $B_s^0 \rightarrow \mu^+ D_s^- X$ Monte Carlo samples are shown in Figure 5.18. The plots are arranged in the following order: $D_s^- \rightarrow \phi \pi^-$ (top), $D_s^- \rightarrow K^{*0} K^-$ (middle), and $D_s^- \rightarrow \pi^+ \pi^- \pi^-$ (bottom) with the B triggers on the left, D triggers in the middle, and ℓ +SVT triggers on the right. The efficiency function, as expected, has a different behavior between B and D triggers. B triggers are selected by the lepton and a D daughter passing the SVT based requirements, and consequently, are significantly depleted in the low t^* region. D and ℓ +SVT triggers, on the other hand, retain high efficiency for low decay time candidates. The efficiency function is parameterized by the following formulation:

$$\xi(t^*) = (p_0 + p_1 \cdot t^* + p_2 \cdot (t^*)^2)(f \cdot \exp(-t^*/\lambda_1) + (1 - f) \cdot \exp(-t^*/\lambda_2)). \quad (5.22)$$

The choice of the $\xi(t^*)$ parameterization is motivated by the possibility of an analytical computation of the normalization of the signal PDF in Equation (5.19) in the t^* space, which greatly reduces the computation time in evaluating the signal likelihood terms.

After obtaining the full description of the time dependence of semileptonic B_s^0 decay candidates from Monte Carlo samples, in terms of $\mathcal{H}(\kappa)$ and $\xi(t^*)$, the signal PDF described in Equation (5.18) can be computed for each candidate. The integral in κ -factor space for the signal PDF and its normalization in Equation (5.19) is approximated by a finite sum

$$\int \mathcal{H}(\kappa) d\kappa \rightarrow \sum_i \mathcal{H}(\kappa_i) \Delta\kappa,$$

where the sum is taken over bins in the histogrammed $\mathcal{H}(\kappa)$ distribution with bin width $\Delta\kappa$. In addition, the normalization of the signal PDF in Equation (5.19) is analytically calculated in the t^* space using the convenient parameterization of $\xi(t^*)$ from Equation (5.22).

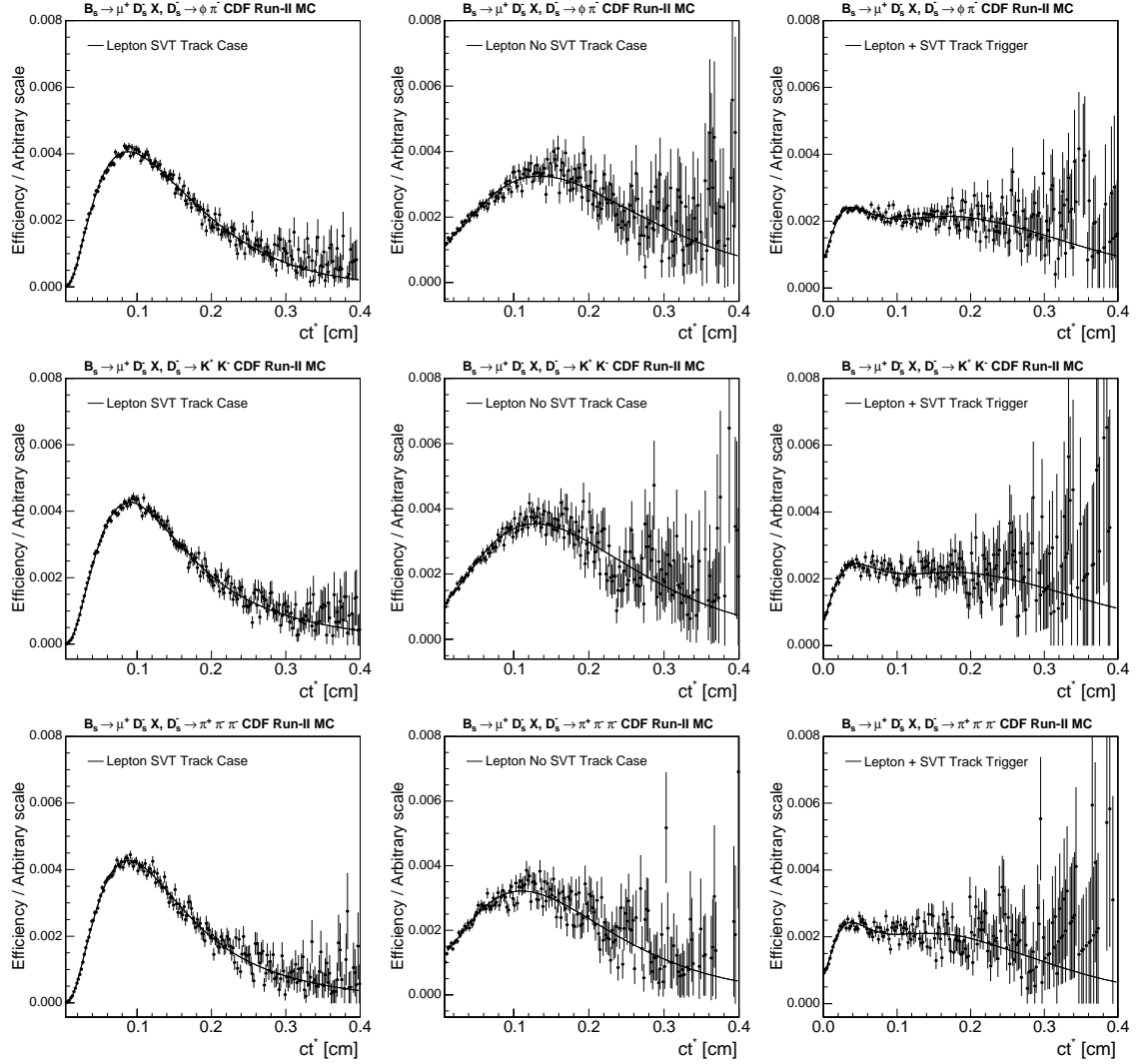


Figure 5.18: $\xi(t^*)$ distributions for $B_s^0 \rightarrow \ell^+ D_s^- X$ decays obtained from Monte Carlo. The normalization scale is in arbitrary units.

Decay Mode	B_s^0 Lifetime [μm] in μ modes	B_s^0 Lifetime [μm] in e modes
$D_s^- \rightarrow \phi\pi^- \quad B$	438 ± 2	437 ± 7
$D_s^- \rightarrow \phi\pi^- \quad D$	440 ± 3	441 ± 11
$D_s^- \rightarrow \phi\pi^- \quad \ell + SVT$	437 ± 4	438 ± 11
$D_s^- \rightarrow K^{*0}K^- \quad B$	437 ± 3	446 ± 7
$D_s^- \rightarrow K^{*0}K^- \quad D$	438 ± 4	441 ± 10
$D_s^- \rightarrow K^{*0}K^- \quad \ell + SVT$	441 ± 5	436 ± 11
$D_s^- \rightarrow \pi^+\pi^-\pi^- \quad B$	440 ± 3	436 ± 5
$D_s^- \rightarrow \pi^+\pi^-\pi^- \quad D$	441 ± 4	441 ± 10
$D_s^- \rightarrow \pi^+\pi^-\pi^- \quad \ell + SVT$	435 ± 5	438 ± 11

Table 5.9: B_s^0 lifetime fit results in signal Monte Carlo samples. The input B_s^0 lifetime value in the Monte Carlo samples is $c\tau(B_s^0) = 438 \mu\text{m}$.

Signal PDF Validation: B_s^0 Lifetime Fits in Monte Carlo

To verify the implementation of the signal PDF in Equation (5.18), including the $\mathcal{H}(\kappa)$ distributions and $\xi(t^*)$ functions, we fit for the B_s^0 lifetime in Monte Carlo samples. The Monte Carlo samples are the same as the ones used for the derivation of $\mathcal{H}(\kappa)$ and $\xi(t^*)$. Figure 5.19 shows the fits to the pseudo-proper decay length (ct^*) in the $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow \phi\pi^-$, $\phi \rightarrow K^+ K^-$ signal Monte Carlo samples for the various trigger topologies. The fit results of the B_s^0 lifetimes in all the B_s^0 signal Monte Carlo samples are summarized in Table 5.9, and are compatible with the input B_s^0 lifetime value $c\tau(B_s^0) = 438 \mu\text{m}$ within the measurement errors.

5.3.6 Characterization in Pseudo-Proper Decay Length ct^*

After gaining confidence in description of the $B_s^0 \rightarrow D_s^- \ell^+ X$ signal in the pseudo-proper decay time via B_s^0 lifetime fits in Monte Carlo, we now describe the characterization of the B_s^0 samples in terms of the combined signal and background components in the pseudo-proper decay length. This characterization is achieved in conjugation with the description in the D_s^- mass and $m_{\ell^+ D_s^-}$ spaces. As a cross-check of understanding the composition of the B_s^0 samples, we determine the B_s^0 meson lifetime to validate the obtained fractions and PDFs describing the signal and backgrounds in the D_s^- mass, $m_{\ell^+ D_s^-}$, and ct^* spaces. First, we describe the procedure to obtain PDFs for the various backgrounds in the following.

ct^* PDF for Combinatorial Background

The ct^* PDF for combinatorial background is again determined from candidates in the D_s^- mass sidebands obtained using the definition in Table 5.4. The parameterization used for describing the combinatorial background shape in ct^* is a sum of two tailed

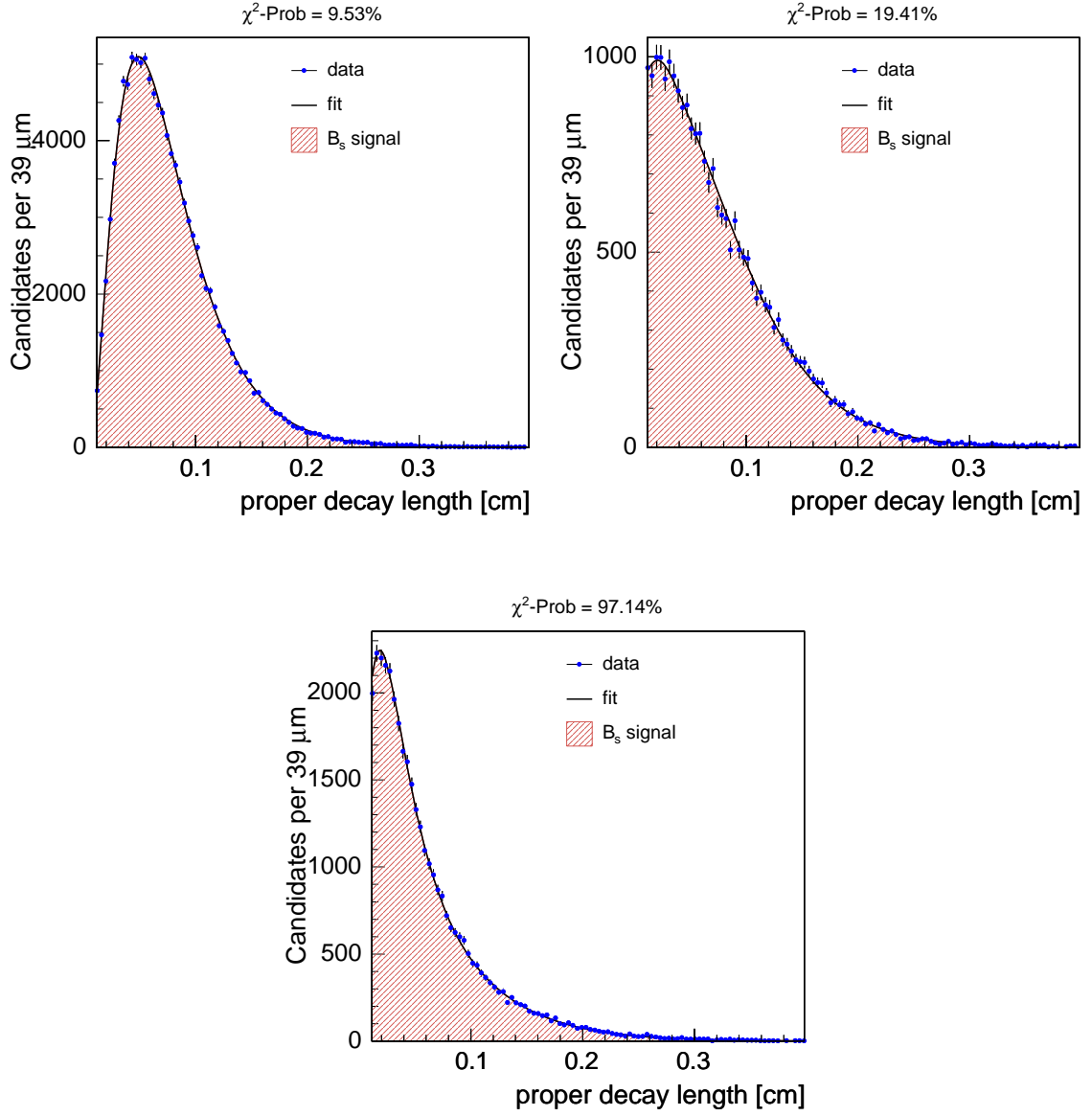


Figure 5.19: Fits in the ct^* space for $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ signal Monte Carlo events in the B (top-left), D (top-right), and ℓ +SVT (bottom) triggers.

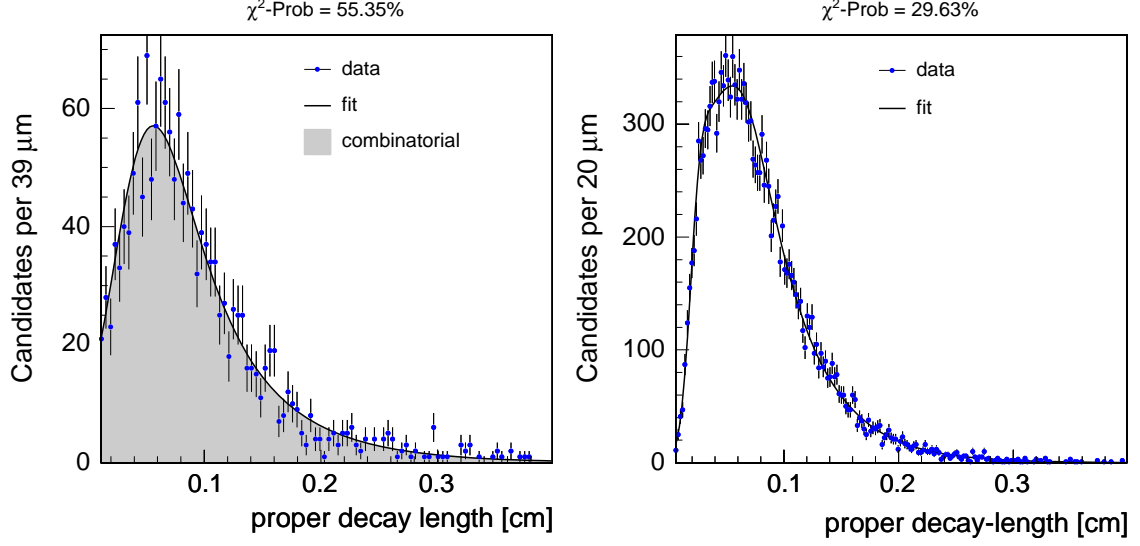


Figure 5.20: Pseudo-proper decay time distribution for combinatorial background in B triggers in the $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ candidates using sideband candidates (left). Pseudo-proper decay time distribution for the D^- reflection background in B triggers in the $B \rightarrow \mu^+ D^- X$ Monte Carlo. These decays are reconstructed as $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$ (right).

Gaussians together with another Gaussian. This is displayed in Figure 5.20 (left) for the combinatorial background candidates obtained from B triggers in the $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ mode.

Some of the parameters obtained for the PDFs describing the combinatorial background in ct^* space are left free in the final D_s^- mass, $m_{\ell^+ D_s^-}$ and ct^* simultaneous fit to the B_s^0 candidates, to again accommodate alterations in the background behavior underneath the D_s^- mass peak with respect to the sidebands. The floating parameters include the most probable values of the two tailed Gaussians along with the mean of the additional Gaussian.

ct^* PDF for D^- Reflection Background

The description of D^- reflection background in ct^* is obtained in the same fashion as the signal utilizing Equation (5.18). The $\mathcal{H}(\kappa)$ distributions and $\xi(t^*)$ functions are derived from Monte Carlo utilizing $B \rightarrow \ell^+ D^- X$ decays reconstructed as $B_s^0 \rightarrow \mu^+ D_s^- X$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$ candidates after passing the signal selection criteria. The behavior of the D^- reflection background in ct^* space is expected to be very similar to the true B_s^0 signal component because it also originates from real B decays. Figure 5.20 (right) illustrates the ct^* description for the D^- reflection background obtained in B triggers from the $B \rightarrow \mu^+ D^- X$ Monte Carlo sample. The

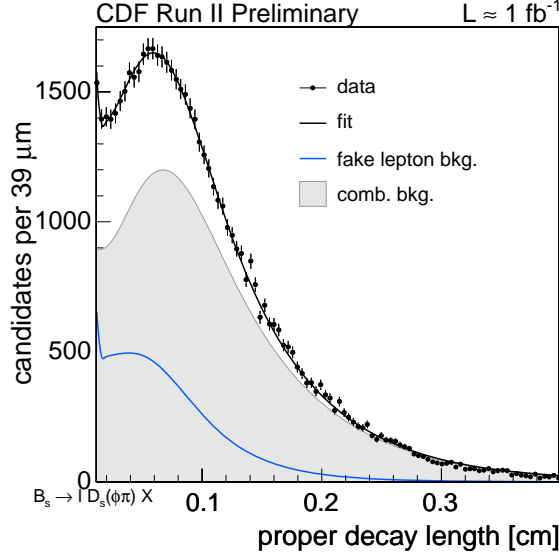


Figure 5.21: Pseudo-proper decay time distribution for fake lepton background in the $B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ candidates.

likelihood fit projection is also overlaid, and confirms that the $\mathcal{H}(\kappa)$ distributions and $\xi(t^*)$ were obtained properly from Monte Carlo. For deriving $\mathcal{H}(\kappa)$ distributions and $\xi(t^*)$ functions from $B \rightarrow \ell^+ D^- X$ decays in Monte Carlo, the sample composition in terms of B^+ and B^0 contributions (see Section 5.4.1) is constrained to the description obtained from the ℓD data (see Section 5.4.2).

ct^* PDF for Physics Background

Since physics background originates from real B decays, we model its PDF using the same description as the signal in Equation (5.18). For this purpose, separate $\mathcal{H}(\kappa)$ distributions and $\xi(t^*)$ functions are derived from the Monte Carlo samples representing the physics background.

ct^* PDF for Fake Lepton Background

The ct^* PDF for the fake lepton background is obtained from the fake lepton samples using the same procedure as for the derivation of the $m_{\ell^+ D_s^-}$ template described in Section 5.3.3. The model utilized for describing the ct^* shape for the fake lepton background is a sum of two tailed Gaussians together with another Gaussian distribution. Figure 5.21 shows the fake lepton distribution in ct^* for the $B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$ candidates, combining B , D and ℓ +SVT triggers. Fake lepton background, in general, has a visible prompt peak in ct^* , although its contribution is reduced with the application of the minimum ct^* requirements detailed in Section 5.2. Also, the classification of the samples into B , D , and

ℓ +SVT triggers further isolates the fake lepton background into certain subsets of the B_s^0 samples, namely the D triggered samples.

ct^* Fit in Data

Using fits in the D_s^- mass and $m_{\ell+D_s^-}$ spaces, we have determined the fractions of signal and background components, together with final PDFs for the various sample components in the D_s^- mass and $m_{\ell+D_s^-}$ spaces. We now include the discrimination provided by the ct^* information in describing the sample. The purpose of doing this is two fold: to validate the sample description in terms of its components by extracting its B_s^0 lifetime values, and to determine the final PDFs in ct^* space for use in the B_s^0 - \bar{B}_s^0 oscillations measurement.

The candidate likelihood in the D_s^- , $m_{\ell+D_s^-}$ and ct^* spaces is described by

$$\begin{aligned} L(m, m_{\ell+D_s^-}, ct^*, \sigma_{ct^*}) &= \sum_i f_i \cdot \mathcal{P}^i(m, m_{\ell+D_s^-}, ct^*, \sigma_{ct^*}) \\ &= \sum_i f_i \cdot \mathcal{P}^i(m) \cdot \mathcal{P}^i(m_{\ell+D_s^-}, ct^*, \sigma_{ct^*}). \end{aligned} \quad (5.23)$$

The contribution to the joint PDFs from the D_s^- mass factorizes out because it is independent of the other variables. However, the other variables used in the fit are correlated to each other. In particular, $m_{\ell+D_s^-}$ and σ_{ct^*} are correlated to the opening angle ΔR between the lepton and the D_s^- meson; higher $m_{\ell+D_s^-}$ values, in general, occur at higher ΔR , where the vertex resolution improves. Therefore, σ_{ct^*} is lower at higher values of $m_{\ell+D_s^-}$. We account for this correlation by binning the sample in the same $m_{\ell+D_s^-}$ ranges as is done for obtaining the $\mathcal{H}(\kappa)$ distributions. The variation of the mean and RMS of the σ_{ct^*} distribution in the $m_{\ell+D_s^-}$ bins is illustrated in Figure 5.22 for the signal and combinatorial backgrounds. The σ_{ct^*} distributions for the combinatorial background are derived using the candidates from the D_s^- mass sidebands, as defined in Table 5.4. The signal distributions for the σ_{ct^*} are obtained from the D_s^- mass signal region defined in Table 5.5, after performing sideband subtraction of background candidates. After taking the dependence of σ_{ct^*} on $m_{\ell+D_s^-}$ into account, we can rewrite the candidate likelihood as

$$L(m, m_{\ell+D_s^-}, ct^*, \sigma_{ct^*}) = \sum_i f_i \cdot \mathcal{P}^i(m) \cdot \mathcal{P}^i(m_{\ell+D_s^-}) \cdot \mathcal{P}^i(ct^*, \sigma_{ct^*}), \quad (5.24)$$

where binning in $m_{\ell+D_s^-}$ is implicit. We can further factorize Equation (5.24) as

$$L(m, m_{\ell+D_s^-}, ct^*, \sigma_{ct^*}) = \sum_i f_i \cdot \mathcal{P}^i(m) \cdot \mathcal{P}^i(m_{\ell+D_s^-}) \cdot \mathcal{P}^i(ct^* | \sigma_{ct^*}) \cdot \mathcal{P}^i(\sigma_{ct^*}). \quad (5.25)$$

$\mathcal{P}^i(\sigma_{ct^*})$ are unconditional PDFs for the σ_{ct^*} and are derived for the signal and combinatorial background components as described above. The normalized σ_{ct^*} PDFs for

the sideband-subtracted signal and combinatorial background for the combined $\ell^+ D_s^-$ modes are shown in Figure 5.23. The $\mathcal{P}^i(\sigma_{ct^*})$ PDFs for the physics, D^- reflection, and fake lepton backgrounds are the same as the one for the signal.

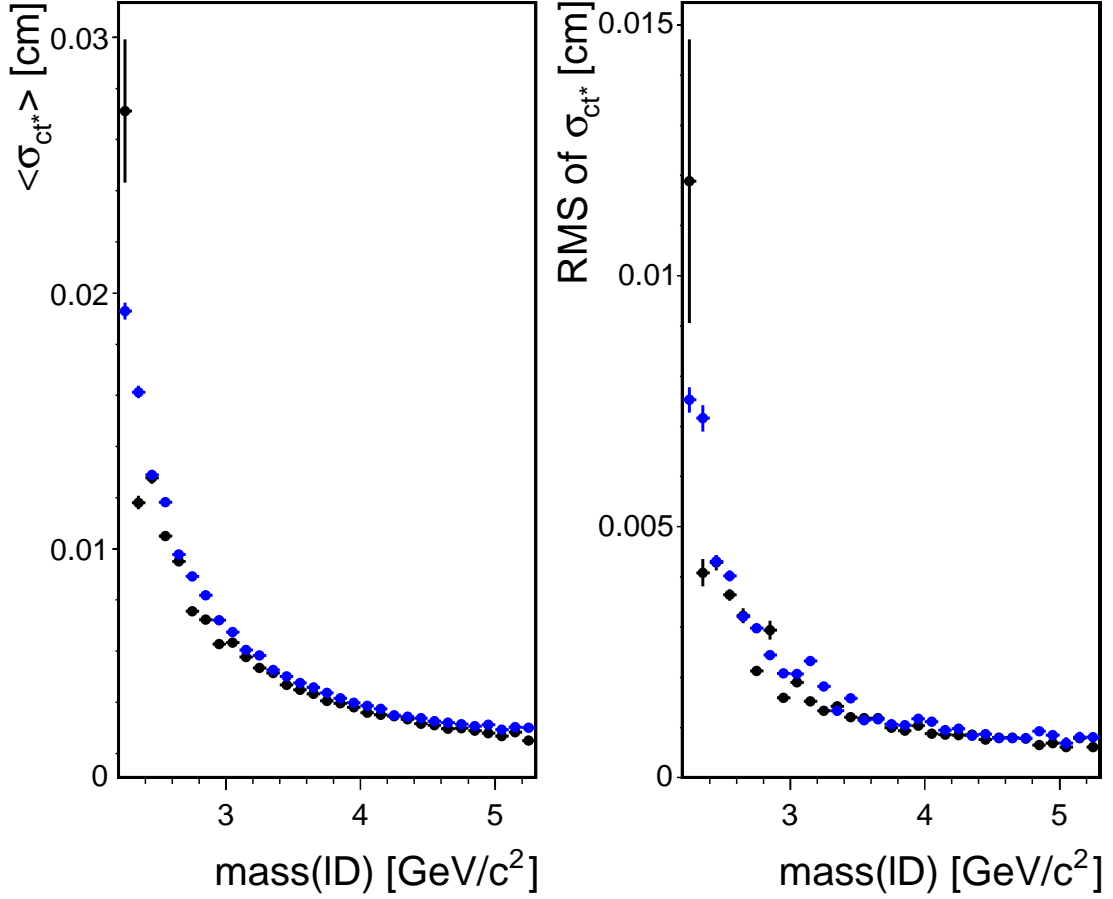


Figure 5.22: Mean (left) and RMS (right) values of the ct^* resolution as a function of $m_{\ell D}$ in ℓD modes for background subtracted signal (black) and sidebands (blue).

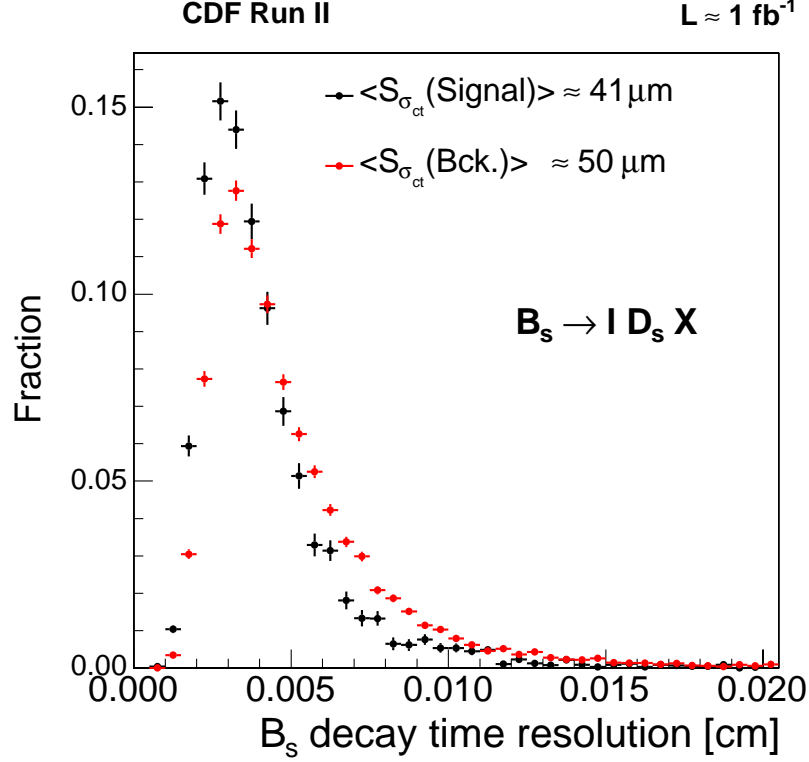


Figure 5.23: σ_{ct^*} distributions for signal and combinatorial background in B triggers in $\ell^+ D_s^-$ modes.

5.3.7 Summary of Characterizations in D_s^- Mass, $m_{\ell^+ D_s^-}$, and ct^* : B_s^0 Lifetime Measurements

Having obtained the descriptions of the signal and different background components in ct^* space, we perform a fit to the B_s^0 candidates simultaneously in the D_s^- mass, $m_{\ell^+ D_s^-}$, and ct^* spaces as described by Equation (5.25). The B_s^0 lifetime is left free in the fit together with some parameters describing the combinatorial background ct^* , as detailed above. The B_s^0 lifetime values obtained from the fit are summarized in Table 5.10. The quoted errors include statistical errors only. The combined B_s^0 lifetimes in the various data-taking periods and trigger topologies are in good agreement with the world average B_s^0 meson lifetime $c\tau(B_s^0) = 438 \pm 18 \mu\text{m}$ [13]. Figures 5.24 to 5.26 display the distributions of ct^* in the three B_s^0 decay modes for the different trigger topologies (B , D , and ℓ +SVT triggers), with fits of the signal and various background components overlaid.

Decay Samples	$\phi\pi^-$ [μm]	$K^{*0}K^-$ [μm]	$\pi^+\pi^-\pi^-$ [μm]
	Period 1		
$\mu D_s, B$ trigger	406.7 ± 10.2	431.1 ± 15.5	430.6 ± 28.9
$e D_s, B$ trigger	421.4 ± 13.4	429.2 ± 18.4	459.7 ± 37.6
$\mu D_s, D$ trigger	439.4 ± 09.6	433.3 ± 13.2	425.1 ± 24.2
$e D_s, D$ trigger	418.6 ± 16.2	441.8 ± 20.2	392.7 ± 30.0
$\mu D_s, \ell + SVT$ trigger	449.2 ± 12.3	435.2 ± 23.9	526.4 ± 66.6
$e D_s, \ell + SVT$ trigger	433.3 ± 14.0	421.1 ± 24.8	404.3 ± 34.1
	Period 2		
$\mu D_s, B$ trigger	423.5 ± 17.3	452.1 ± 17.9	437.0 ± 51.0
$e D_s, B$ trigger	422.3 ± 15.5	468.0 ± 38.2	444.6 ± 37.1
$\mu D_s, D$ trigger	415.7 ± 25.5	416.7 ± 30.6	482.5 ± 39.2
$e D_s, D$ trigger	454.7 ± 23.6	381.9 ± 43.0	484.4 ± 54.6
$\mu D_s, \ell + SVT$ trigger	378.8 ± 30.3	382.0 ± 21.3	483.3 ± 54.4
$e D_s, \ell + SVT$ trigger	379.4 ± 21.0	360.7 ± 37.7	414.1 ± 32.7
	Period 3		
$\mu D_s, B$ trigger	410.9 ± 16.3	412.7 ± 46.1	344.7 ± 31.1
$e D_s, B$ trigger	437.5 ± 14.4	409.4 ± 34.3	471.9 ± 35.2
$\mu D_s, D$ trigger	455.5 ± 17.7	469.8 ± 31.6	391.4 ± 37.1
$e D_s, D$ trigger	470.6 ± 20.8	456.3 ± 38.4	431.7 ± 46.8
$\mu D_s, \ell + SVT$ trigger	451.4 ± 23.6	468.6 ± 54.9	453.1 ± 39.7
$e D_s, \ell + SVT$ trigger	448.3 ± 19.1	456.1 ± 42.5	408.9 ± 61.4
Combined	429.4 ± 3.6	430.1 ± 5.7	427.5 ± 8.7

Table 5.10: B_s^0 lifetime fit results in various modes and subsamples for different periods of data-taking. The errors quoted are statistical only.

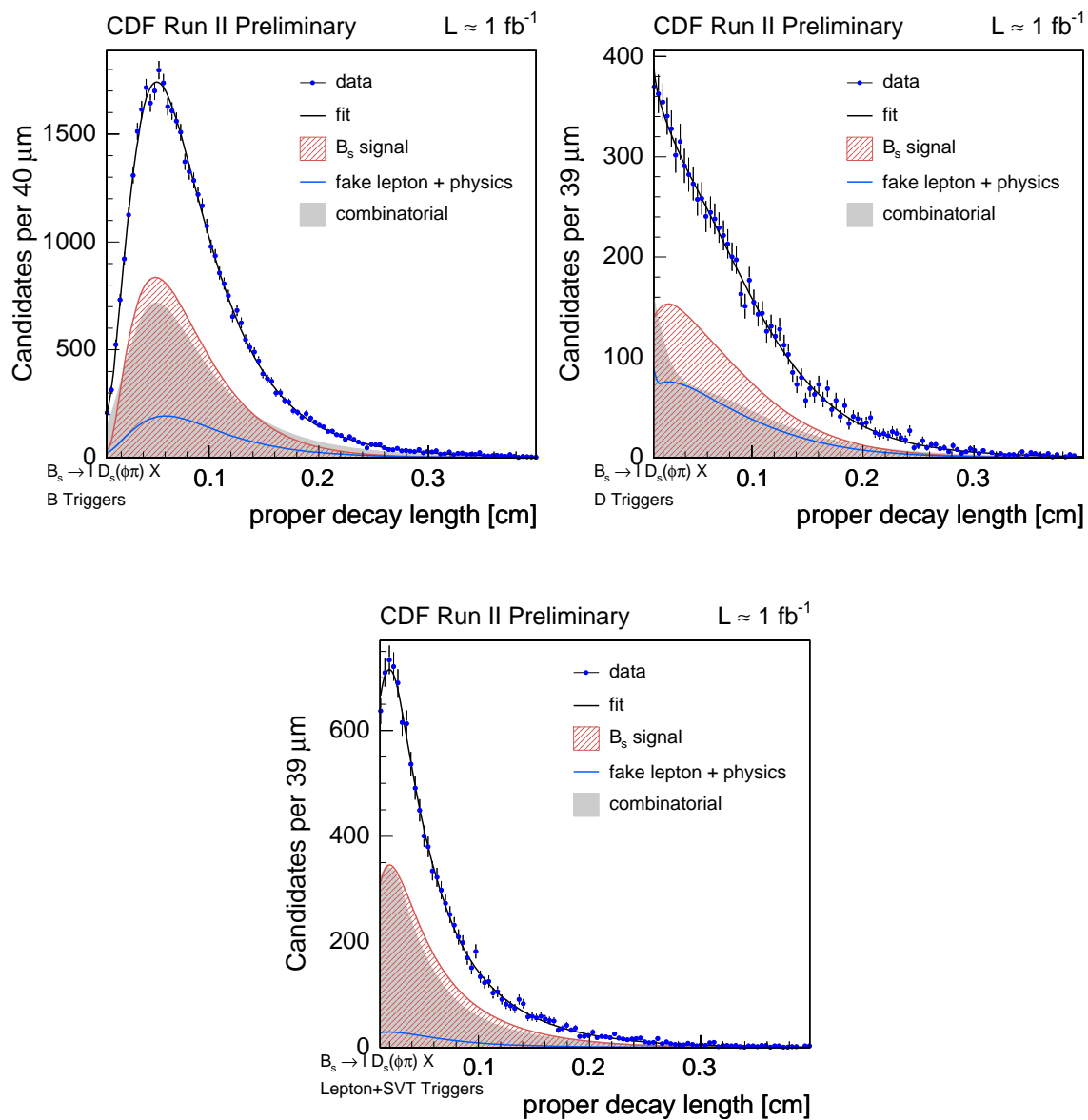


Figure 5.24: Fits to the ct^* distributions for $B \rightarrow \ell D_s(\phi\pi)$ candidates in B (top-left), D (top-right), and ℓ +SVT (bottom) triggers.

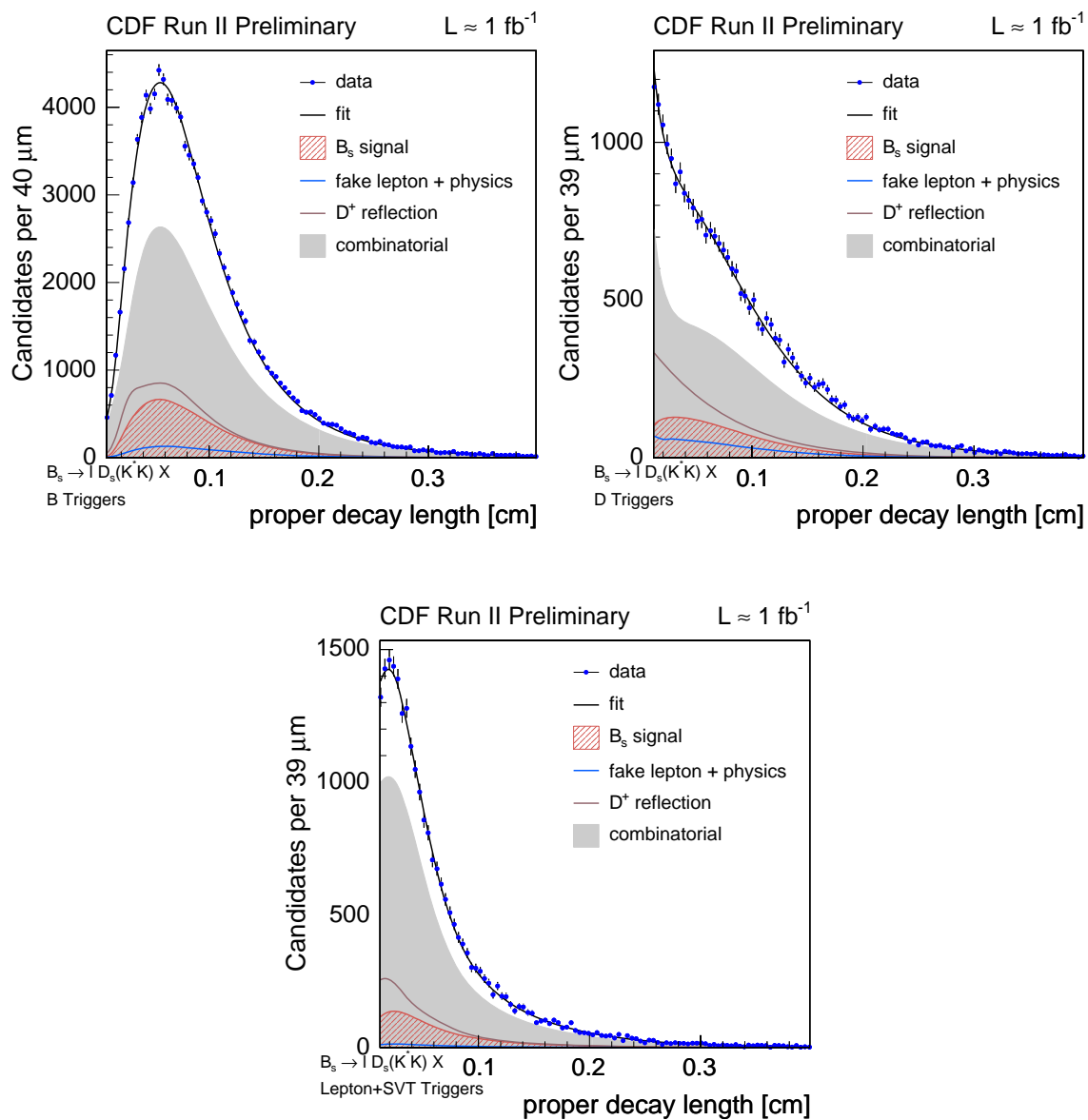


Figure 5.25: Fits to the ct^* distributions for $B \rightarrow \ell D_s(K^*K)$ candidates in B (top-left), D (top-right), and ℓ +SVT (bottom) triggers.

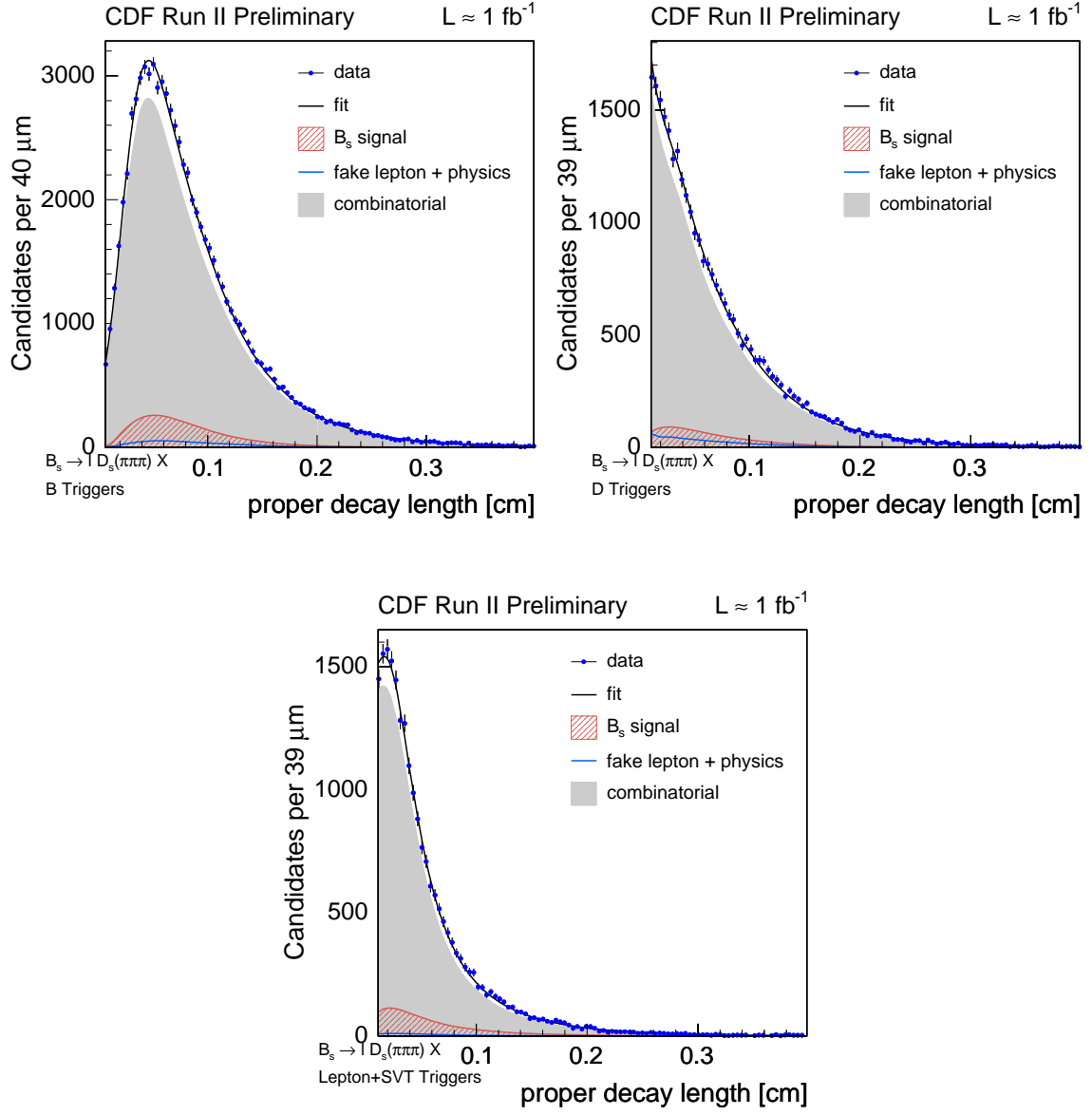


Figure 5.26: Fits to the ct^* distributions for $B \rightarrow \ell D_s(3\pi)$ candidates in B (top-left), D (top-right), and ℓ +SVT (bottom) triggers.

5.4 Characterization of the B^+ and B^0 Samples

We reconstruct the B^+ and B^0 mesons in the following three final states: $\ell^+\bar{D}^0X$, ℓ^+D^-X , and $\ell^+D^{*-}X$. The methodology used for determining the fractions and PDFs of signal and background components for B_s^0 samples is replicated for the case of B^+ and B^0 samples. However, we note two important differences with respect to the analysis of the B_s^0 decay modes:

- None of the three ℓD decay modes are composed of pure B^+ or B^0 contribution. Instead, they are admixtures of the decays of the B^+ and B^0 mesons. This issue complicates the analysis as described below.
- All background types discussed in Section 5.1 are also present in these samples, except for the background from the D^- reflection.

We first discuss the issue of disentangling the contributions from B^+ and B^0 mesons into a specific ℓD final state.

5.4.1 Sample Composition of the ℓD Samples

A necessary complication of analyzing the three ℓD final states is unraveling the various types of cross-talks between the B^+ and B^0 mesons. The incomplete reconstruction in semileptonic B^+ and B^0 decays to excited charm states (D^* and D^{**}) results in cross-talk between the B^+ and B^0 mesons, producing ℓD final states that are admixtures of the two mesons. The ℓD modes reconstructed in this measurement are thus composed of 24 different decay sequences that are schematically represented in Figure 5.27. The contribution of B^+ and B^0 to each of the three ℓD final states is calculated using the branching ratios, as well as trigger and reconstruction efficiencies for each of the 24 decay sequences.

We study the sample composition of the ℓD modes by analyzing the impact of relative branching ratios at each decay step, invoking the following parameterization of the various decay sequences [63]. The ratio of semileptonic B decays into the various charm mesons can conveniently be re-expressed using ratios relative to the inclusive branching fraction to the lowest-lying D state, including decays via intermediate D^* and D^{**} states, $\mathcal{B}(B \rightarrow \nu \ell D X)$:

$$f_{0,+} \equiv \frac{\mathcal{B}(B^{0,+} \rightarrow \nu \ell D)}{\mathcal{B}(B^{0,+} \rightarrow \nu \ell D X)}, \quad (5.26)$$

$$f_{0,+}^* \equiv \frac{\mathcal{B}(B^{0,+} \rightarrow \nu \ell D^*)}{\mathcal{B}(B^{0,+} \rightarrow \nu \ell D X)}, \quad (5.27)$$

$$f_{0,+}^{**} \equiv \frac{\mathcal{B}(B^{0,+} \rightarrow \nu \ell D^{**})}{\mathcal{B}(B^{0,+} \rightarrow \nu \ell D X)}. \quad (5.28)$$

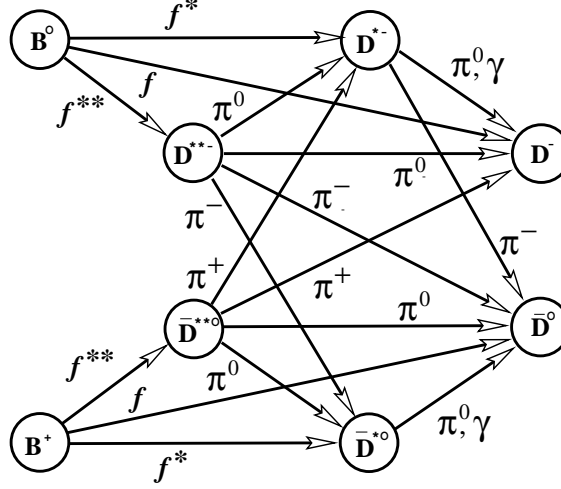


Figure 5.27: A schematic representation of all possible $B \rightarrow \ell D^{(*)} X$ transitions resulting in a ℓD final state.

Ref. [63] assumes that all the charged and neutral ratios are equal, for example, $f \equiv f_0 = f_+$. Thus, in the schematic diagram shown in Figure 5.27, f , f^* , and f^{**} denote the fraction of B decays to D^+ or D^0 , D^{*+} or D^{*0} , and D^{*-} or D^{*0} , respectively. Since the D^* and D^{**} decay strongly they all ultimately result in a DX signature, and thus $f + f^* + f^{**} \equiv 1$. The fraction of D^{**} decays to D^* states with respect to the fraction of D^{**} decays to D states is represented by another parameter P_V . The ratio f^*/f is denoted by R_f . The efficiency of reconstructing the D^{*+} in the $D^{*+} \rightarrow D^0 \pi^+$ decay is symbolized by ϵ^* . Thus, the sample composition of the ℓD modes before application of trigger and reconstruction efficiencies is completely represented by the four independent parameters: f^{**} , R_f , P_V , and ϵ^* . The sample composition parameters used in this measurement are summarized in Table 5.11. The rates of the 24 decay sequences in terms of the sample composition parameters are summarized in Table 5.12. The branching ratio of $D^* \rightarrow X$ is denoted by $\mathcal{B}^*(X)$, and is obtained from Ref. [13].

Parameter	Value
R_f	2.14 ± 0.14
f^{**}	0.31 ± 0.05
P_V	0.627 ± 0.26

Table 5.11: B^+ and B^0 sample composition parameters used in this measurement.

After evaluating the contribution of the various branching ratios to the sample composition of the ℓD samples, the trigger and reconstruction efficiencies are applied

B^0 Decay Sequences	Decay Signatures		
	ℓD^{*-}	ℓD^-	ℓD^0
$\rightarrow D^{*-}\ell^+\nu$			
$D^{*-} \rightarrow \bar{D}^{*0}\pi_{**}^-$			
$\bar{D}^{*0} \rightarrow \bar{D}^0\pi_*^0$	—	—	$f^{**}P_V\frac{2}{3}\mathcal{B}^*(D^-\pi_*^0)$
$\bar{D}^{*0} \rightarrow \bar{D}^0\gamma$	—	—	$f^{**}P_V\frac{2}{3}\mathcal{B}^*(\bar{D}^0\gamma)$
$D^{*-} \rightarrow D^{*-}\pi_{**}^0$	$f^{**}P_V\frac{1}{3}\mathcal{B}^*(\bar{D}^0\pi_*^-\epsilon(\pi_*))$	—	$f^{**}P_V\frac{1}{3}\mathcal{B}^*(\bar{D}^0\pi_*^-)(1-\epsilon(\pi_*))$
$D^{*-} \rightarrow D^-\pi_*^0$	—	$f^{**}P_V\frac{1}{3}\mathcal{B}^*(D^-\pi_*^0)$	—
$D^{*-} \rightarrow D^-\gamma$	—	$f^{**}P_V\frac{1}{3}\mathcal{B}^*(D^-\gamma)$	—
$D^{*-} \rightarrow \bar{D}^0\pi_{**}^-$	—	—	$f^{**}(1-P_V)\frac{2}{3}$
$D^{*-} \rightarrow D^-\pi_{**}^0$	—	$f^{**}(1-P_V)\frac{1}{3}$	—
$\rightarrow D^{*-}\ell^+\nu$			
$D^{*-} \rightarrow \bar{D}^0\pi_*^-$	$f^*\mathcal{B}^*(\bar{D}^0\pi_*^-\epsilon(\pi_*))$	—	$f^*\mathcal{B}^*(\bar{D}^0\pi_*^-)(1-\epsilon(\pi_*))$
$D^{*-} \rightarrow D^-\pi_*^0$	—	$f^*\mathcal{B}^*(D^-\pi_*^0)$	—
$D^{*-} \rightarrow D^-\gamma$	—	$f^*\mathcal{B}^*(D^-\gamma)$	—
$\rightarrow D^-\ell^+\nu$	—	f	—
B^+ Decay Sequences			
$\rightarrow \bar{D}^{*0}\ell^+\nu$			
$\bar{D}^{*0} \rightarrow D^{*-}\pi_{**}^+$			
$D^{*-} \rightarrow \bar{D}^0\pi_*^-$	$f^{**}P_V\frac{2}{3}\mathcal{B}^*(\bar{D}^0\pi_*^-\epsilon(\pi_*))$	—	$f^{**}P_V\frac{2}{3}\mathcal{B}^*(\bar{D}^0\pi_*^-)(1-\epsilon(\pi_*))$
$D^{*-} \rightarrow D^-\pi_*^0$	—	$f^{**}P_V\frac{2}{3}\mathcal{B}^*(D^-\pi_*^0)$	—
$D^{*-} \rightarrow \bar{D}^0\gamma$	—	$f^{**}P_V\frac{2}{3}\mathcal{B}^*(D^-\gamma)$	—
$\bar{D}^{*0} \rightarrow \bar{D}^{*0}\pi_{**}^0$			
$\bar{D}^{*0} \rightarrow \bar{D}^0\pi_*^0$	—	—	$f^{**}P_V\frac{1}{3}\mathcal{B}^*(\bar{D}^0\pi_*^0)$
$\bar{D}^{*0} \rightarrow \bar{D}^0\gamma$	—	—	$f^{**}P_V\frac{1}{3}\mathcal{B}^*(\bar{D}^0\gamma)$
$\bar{D}^{*0} \rightarrow D^-\pi_{**}^+$	—	$f^{**}(1-P_V)\frac{2}{3}$	—
$\bar{D}^{*0} \rightarrow \bar{D}^0\pi_{**}^0$	—	—	$f^{**}(1-P_V)\frac{1}{3}$
$\rightarrow D^{*0}\ell^+\nu$			
$\bar{D}^{*0} \rightarrow \bar{D}^0\pi_*^0$	—	—	$f^*\mathcal{B}^*(\bar{D}^0\pi_*^0)$
$\bar{D}^{*0} \rightarrow \bar{D}^0\gamma$	—	—	$f^*\mathcal{B}^*(\bar{D}^0\gamma)$
$\rightarrow D^0\ell^+\nu$	—	—	f

Table 5.12: Table of the various B^+ and B^0 decay sequences and their contributions to the sample composition of the three general categories of decay signatures ($\ell^+ D^{*-}$, $\ell^+ D^-$, and $\ell^+ \bar{D}^0$) [63]. The total contribution of a particular B meson to a ℓD final state is the sum over the entries in the corresponding vertical column.

to obtain the actual contributions in data. The different decay sequences are likely different in their kinematical behavior. This results in different efficiencies for the trigger selection and application of reconstruction requirements. Thus, the relative contribution of the decay sequences to a given ℓD final state is altered. Since we are only interested in relative efficiencies in this measurement, the efficiencies for the 24 decay sequences are computed with respect to the efficiencies of three reference decay sequences:

1. $B^0 \rightarrow \ell^+ D^{*-} \nu_\ell$, $D^{*-} \rightarrow \bar{D}^0 \pi_*^-$, $\bar{D}^0 \rightarrow K^+ \pi^-$,
2. $B^0 \rightarrow \ell^+ D^- \nu_\ell$, $D^- \rightarrow K^+ \pi^- \pi^-$,
3. $B^+ \rightarrow \ell^+ \bar{D}^0 \nu_\ell$, $\bar{D}^0 \rightarrow K^+ \pi^-$.

The relative efficiencies for triggering and reconstruction of the remaining 21 decay sequences are summarized in Table 5.13, as obtained from Monte Carlo samples. Using the information from Tables 5.11 to 5.13, we determine that the fraction of B^0 in our data is $\sim 28\%$ for the $B \rightarrow \ell^+ \bar{D}^0 X$ mode, $\sim 84\%$ for $B \rightarrow \ell^+ D^- X$, and $\sim 90\%$ for the $B \rightarrow \ell^+ D^{*-} X$ mode.

5.4.2 Characterizations in D Mass, $m_{\ell D}$ and ct^*

The determination of signal and background components in the ℓD modes in the corresponding D mass, $m_{\ell D}$ and ct^* spaces follows closely the procedure described in the Section 5.3 for the B_s^0 candidates. As noted above, the ℓD samples also have contamination from combinatorial, physics and fake lepton backgrounds but the D^- reflection background is absent. The fraction of combinatorial background is obtained in a manner similar to the B_s^0 modes (see Section 5.3.1), by fitting the D invariant mass distribution. These fractions are summarized in Table 5.14 for the various subsamples and different data-taking periods. The physics background contamination in the ℓD samples is determined from Monte Carlo and summarized in Table 5.15. The fake lepton contribution is determined via a simultaneous fit to the D mass and $m_{\ell D}$ variables and is displayed in Table 5.16. The number of true B^+ and B^0 candidates in our ℓD samples is summarized in Table 5.17.

The final PDFs for the description of signal and background components are also derived in the simultaneous fits in the three discriminating variables: D mass, $m_{\ell D}$ and ct^* . The lifetime values for B^0 and B^+ mesons are extracted and are in agreement within uncertainties with their corresponding world average values [13]. Figures 5.28 to 5.36 display the fits to the D mass, $m_{\ell D}$ and ct^* spaces in the three B semileptonic decay modes for the different trigger topologies (B , D , and ℓ +SVT triggers), with fits of the signal and various background components overlaid.

	Decay channel	Relative Efficiencies	
		Trigger	Reconstruction
ℓD^{*-}	$B^0 \rightarrow D^{*-} \rightarrow \pi_*^- \bar{D}^0 \rightarrow K\pi$	1	1
	$B^0 \rightarrow D^{*-} \rightarrow \pi_{**}^0 D^{*-} \rightarrow \pi_*^- \bar{D}^0 \rightarrow K\pi$	0.68 ± 0.01	1.05 ± 0.02
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^+ D^{*-} \rightarrow \pi_*^- \bar{D}^0 \rightarrow K\pi$	0.66 ± 0.01	0.99 ± 0.02
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^+ D^{*-} \rightarrow \pi_*^- \bar{D}^0 \rightarrow K\pi$	0.66 ± 0.01	0.99 ± 0.02
ℓD^-	$B^0 \rightarrow D^{*-} \rightarrow \pi_*^- \bar{D}^0 \rightarrow K\pi\pi$	1	1
	$B^0 \rightarrow D^{*-} \rightarrow \gamma D^- \rightarrow K\pi\pi$	1.20 ± 0.02	0.96 ± 0.01
	$B^0 \rightarrow D^{*-} \rightarrow \pi_*^0 D^- \rightarrow K\pi\pi$	1.19 ± 0.02	0.96 ± 0.01
	$B^0 \rightarrow D^{*-} \rightarrow \pi_{**}^0 D^- \rightarrow K\pi\pi$	0.85 ± 0.01	0.95 ± 0.01
	$B^0 \rightarrow D^{*-} \rightarrow \pi_{**}^0 D^{*-} \rightarrow \gamma D^- \rightarrow K\pi\pi$	0.78 ± 0.01	0.97 ± 0.01
	$B^0 \rightarrow D^{*-} \rightarrow \pi_{**}^0 D^{*-} \rightarrow \pi_*^0 D^- \rightarrow K\pi\pi$	0.78 ± 0.01	0.95 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^+ D^- \rightarrow K\pi\pi$	1.06 ± 0.02	0.77 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^+ D^{*-} \rightarrow \gamma D^- \rightarrow K\pi\pi$	0.91 ± 0.01	0.82 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^+ D^{*-} \rightarrow \pi_*^0 D^- \rightarrow K\pi\pi$	0.91 ± 0.01	0.81 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^+ D^{*-} \rightarrow \pi_*^0 D^- \rightarrow K\pi\pi$	0.91 ± 0.01	0.81 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^+ D^{*-} \rightarrow \pi_*^0 D^- \rightarrow K\pi\pi$	0.91 ± 0.01	0.81 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^+ D^{*-} \rightarrow \pi_*^0 D^- \rightarrow K\pi\pi$	0.91 ± 0.01	0.81 ± 0.01
ℓD^0	$B^0 \rightarrow D^{*-} \rightarrow \pi_*^- \bar{D}^0 \rightarrow K\pi$	1.22 ± 0.02	0.90 ± 0.01
	$B^0 \rightarrow D^{*-} \rightarrow \pi_{**}^0 D^{*-} \rightarrow \pi_*^- \bar{D}^0 \rightarrow K\pi$	0.85 ± 0.01	0.92 ± 0.01
	$B^0 \rightarrow D^{*-} \rightarrow \pi_{**}^0 \bar{D}^0 \rightarrow K\pi$	0.95 ± 0.01	0.91 ± 0.01
	$B^0 \rightarrow D^{*-} \rightarrow \pi_{**}^- D^{*0} \rightarrow \gamma \bar{D}^0 \rightarrow K\pi$	0.83 ± 0.01	0.92 ± 0.01
	$B^0 \rightarrow D^{*-} \rightarrow \pi_{**}^- D^{*0} \rightarrow \pi_*^0 \bar{D}^0 \rightarrow K\pi$	0.84 ± 0.01	0.91 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^+ \bar{D}^0 \rightarrow K\pi$	1	1
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \gamma \bar{D}^0 \rightarrow K\pi$	1.32 ± 0.02	0.93 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^0 \bar{D}^0 \rightarrow K\pi$	1.32 ± 0.02	0.93 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^0 \bar{D}^0 \rightarrow K\pi$	0.83 ± 0.01	0.94 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^0 \bar{D}^0 \rightarrow \gamma \bar{D}^0 \rightarrow K\pi$	0.94 ± 0.01	0.95 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^0 \bar{D}^0 \rightarrow \pi_*^0 \bar{D}^0 \rightarrow K\pi$	0.84 ± 0.01	0.93 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^0 \bar{D}^0 \rightarrow \pi_*^0 \bar{D}^0 \rightarrow K\pi$	0.83 ± 0.01	0.91 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^0 \bar{D}^0 \rightarrow \pi_*^0 \bar{D}^0 \rightarrow K\pi$	0.83 ± 0.01	0.91 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^0 \bar{D}^0 \rightarrow \pi_*^0 \bar{D}^0 \rightarrow K\pi$	0.83 ± 0.01	0.91 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^0 \bar{D}^0 \rightarrow \pi_*^0 \bar{D}^0 \rightarrow K\pi$	0.83 ± 0.01	0.91 ± 0.01
	$B^+ \rightarrow \bar{D}^{*0} \rightarrow \pi_{**}^0 \bar{D}^0 \rightarrow \pi_*^0 \bar{D}^0 \rightarrow K\pi$	0.83 ± 0.01	0.91 ± 0.01

Table 5.13: Relative efficiencies of the 24 B signal Monte Carlo samples with respect to three reference modes.

Decay Samples	$\ell^+ D^0$ [%]	$\ell^+ D^{*-}$ [%]	$\ell^+ D^-$ [%]
	Period 1		
μD , B trigger	38.2	0.7	68.3
$e D$, B trigger	38.7	0.7	67.5
μD , D trigger	29.8	0.2	64.5
$e D$, D trigger	30.4	0.3	65.2
μD , $\ell + SVT$ trigger	40.8	7.2	64.1
$e D$, $\ell + SVT$ trigger	33.7	6.7	57.9
Period 2			
μD , B trigger	38.2	0.7	68.8
$e D$, B trigger	37.8	0.6	66.8
μD , D trigger	28.6	0.4	64.4
$e D$, D trigger	28.4	0.2	65.8
μD , $\ell + SVT$ trigger	40.3	7.3	63.0
$e D$, $\ell + SVT$ trigger	32.1	6.3	58.3
Period 3			
μD , B trigger	38.2	1.0	69.1
$e D$, B trigger	37.8	0.4	67.0
μD , D trigger	28.6	0.7	64.4
$e D$, D trigger	28.4	0.0	65.1
μD , $\ell + SVT$ trigger	40.3	7.8	63.8
$e D$, $\ell + SVT$ trigger	32.1	6.6	59.5

Table 5.14: Fractions of combinatorial background in $B \rightarrow \ell D X$ candidates for the three data-taking periods. The absolute uncertainties in these fractions are $\sim 2\%$ each.

<i>B</i> Triggers			
Decay	D^0 [%]	D^* [%]	D^+ [%]
$B \rightarrow \nu\tau D^{(*)}X, \tau \rightarrow \mu\nu\nu$	2.1 ± 0.5	2.0 ± 0.5	2.1 ± 0.5
$B \rightarrow D_s^{(*)}D^{(*)}X, D_s^{(*)} \rightarrow \mu X$	2.3 ± 0.6	1.6 ± 0.4	3.6 ± 0.9
$B \rightarrow D^{(*)}D^{(*)}K, D^{(*)} \rightarrow \mu X$	1.1 ± 0.3	1.2 ± 0.3	2.1 ± 0.5
$B_s \rightarrow \nu\mu D_s^{**}X, D_s^{**} \rightarrow DK$	1.2 ± 0.3	0.8 ± 0.2	1.5 ± 0.4
total	6.7 ± 1.6	5.6 ± 1.4	9.3 ± 2.3
Decay	D^0 [%]	D^* [%]	D^+ [%]
$B \rightarrow \nu\tau D^{(*)}X, \tau \rightarrow e\nu\nu$	1.9 ± 0.4	1.8 ± 0.4	2.0 ± 0.5
$B \rightarrow D_s^{(*)}D^{(*)}X, D_s^{(*)} \rightarrow eX$	2.0 ± 0.5	1.3 ± 0.3	3.2 ± 0.8
$B \rightarrow D^{(*)}D^{(*)}K, D^{(*)} \rightarrow eX$	0.7 ± 0.2	0.7 ± 0.2	1.7 ± 0.4
$B_s \rightarrow \nu e D_s^{**}X, D_s^{**} \rightarrow DK$	1.0 ± 0.3	0.7 ± 0.2	1.4 ± 0.4
total	5.6 ± 1.4	4.5 ± 1.1	8.4 ± 2.1
<i>D</i> Triggers			
Decay	D^0 [%]	D^* [%]	D^+ [%]
$B \rightarrow \nu\tau D^{(*)}X, \tau \rightarrow \mu\nu\nu$	2.6 ± 0.6	2.4 ± 0.6	2.3 ± 0.6
$B \rightarrow D_s^{(*)}D^{(*)}X, D_s^{(*)} \rightarrow \mu X$	3.4 ± 0.9	2.1 ± 0.5	4.9 ± 1.2
$B \rightarrow D^{(*)}D^{(*)}K, D^{(*)} \rightarrow \mu X$	2.1 ± 0.5	1.9 ± 0.5	3.0 ± 0.8
$B_s \rightarrow \nu\mu D_s^{**}X, D_s^{**} \rightarrow DK$	1.2 ± 0.3	0.8 ± 0.2	1.4 ± 0.4
total	9.3 ± 2.3	7.2 ± 1.8	11.6 ± 2.9
Decay	D^0 [%]	D^* [%]	D^+ [%]
$B \rightarrow \nu\tau D^{(*)}X, \tau \rightarrow e\nu\nu$	2.6 ± 0.6	2.4 ± 0.6	2.4 ± 0.6
$B \rightarrow D_s^{(*)}D^{(*)}X, D_s^{(*)} \rightarrow eX$	3.1 ± 0.8	2.0 ± 0.5	4.3 ± 1.1
$B \rightarrow D^{(*)}D^{(*)}K, D^{(*)} \rightarrow eX$	1.5 ± 0.4	1.3 ± 0.3	3.0 ± 0.8
$B_s \rightarrow \nu e D_s^{**}X, D_s^{**} \rightarrow DK$	1.0 ± 0.2	0.7 ± 0.2	1.3 ± 0.4
total	8.2 ± 2.3	6.4 ± 1.6	11.0 ± 2.7
ℓ +SVT Triggers			
Decay	D^0 [%]	D^* [%]	D^+ [%]
$B \rightarrow \nu\tau D^{(*)}X, \tau \rightarrow \mu\nu\nu$	1.4 ± 0.4	1.4 ± 0.4	1.6 ± 0.4
$B \rightarrow D_s^{(*)}D^{(*)}X, D_s^{(*)} \rightarrow \mu X$	1.3 ± 0.3	1.0 ± 0.2	2.5 ± 0.6
$B \rightarrow D^{(*)}D^{(*)}K, D^{(*)} \rightarrow \mu X$	0.5 ± 0.1	0.4 ± 0.1	1.5 ± 0.4
$B_s \rightarrow \nu\mu D_s^{**}X, D_s^{**} \rightarrow DK$	1.2 ± 0.3	0.7 ± 0.2	1.6 ± 0.6
total	4.4 ± 1.1	3.5 ± 0.9	7.2 ± 1.8
Decay	D^0 [%]	D^* [%]	D^+ [%]
$B \rightarrow \nu\tau D^{(*)}X, \tau \rightarrow e\nu\nu$	1.3 ± 0.4	1.2 ± 0.3	1.5 ± 0.4
$B \rightarrow D_s^{(*)}D^{(*)}X, D_s^{(*)} \rightarrow eX$	1.1 ± 0.3	0.8 ± 0.2	1.7 ± 0.4
$B \rightarrow D^{(*)}D^{(*)}K, D^{(*)} \rightarrow eX$	0.2 ± 0.1	0.2 ± 0.1	0.7 ± 0.2
$B_s \rightarrow \nu e D_s^{**}X, D_s^{**} \rightarrow DK$	1.2 ± 0.3	0.7 ± 0.2	1.6 ± 0.4
total	3.8 ± 1.0	2.9 ± 0.9	7.2 ± 1.8

Table 5.15: Expected fraction of each physics background with respect to the signal and physics backgrounds in the ℓD modes for the B , D and ℓ +SVT triggers.

Decay Samples	$\ell^+ D^0$ [%]	$\ell^+ D^{*-}$ [%]	$\ell^+ D^-$ [%]
Period 1			
$\mu D, B$ trigger	3.0	1.5	4.1
$e D, B$ trigger	2.4	3.2	0.1
$\mu D, D$ trigger	11.8	9.7	11.6
$e D, D$ trigger	12.1	8.8	12.8
$\mu D, \ell + SVT$ trigger	0.9	3.0	1.7
$e D, \ell + SVT$ trigger	0.7	1.4	1.3
Period 2			
$\mu D, B$ trigger	1.6	0.5	4.3
$e D, B$ trigger	2.2	2.7	4.1
$\mu D, D$ trigger	7.0	10.0	12.6
$e D, D$ trigger	13.4	9.9	7.1
$\mu D, \ell + SVT$ trigger	1.0	1.9	1.7
$e D, \ell + SVT$ trigger	1.6	0.5	1.0
Period 3			
$\mu D, B$ trigger	1.1	1.9	4.2
$e D, B$ trigger	3.6	4.6	2.5
$\mu D, D$ trigger	12.6	12.6	8.6
$e D, D$ trigger	13.0	10.7	11.9
$\mu D, \ell + SVT$ trigger	1.0	2.3	1.7
$e D, \ell + SVT$ trigger	1.7	2.7	1.3

Table 5.16: Fractions of fake lepton background relative to the sum of signal, physics and prompt backgrounds for $B \rightarrow \ell D X$ candidates for the three data-taking periods. The relative uncertainties on these fractions are $\sim 50\%$ each.

Decay Mode	$B^+ + B^0$ Yield	B^0 Fraction [%]
$B^{+,0} \rightarrow \ell^+ D^0 X$	674271 ± 7345	~ 28
$B^{+,0} \rightarrow \ell^+ D^- X$	322710 ± 6482	~ 84
$B^{+,0} \rightarrow \ell^+ D^{*-} X$	77556 ± 673	~ 90

Table 5.17: Number of true $B^{+,0} \rightarrow \ell D X$ signal events in the three decay modes with their corresponding B^0 fractions, after taking into account the contributions from all the backgrounds.

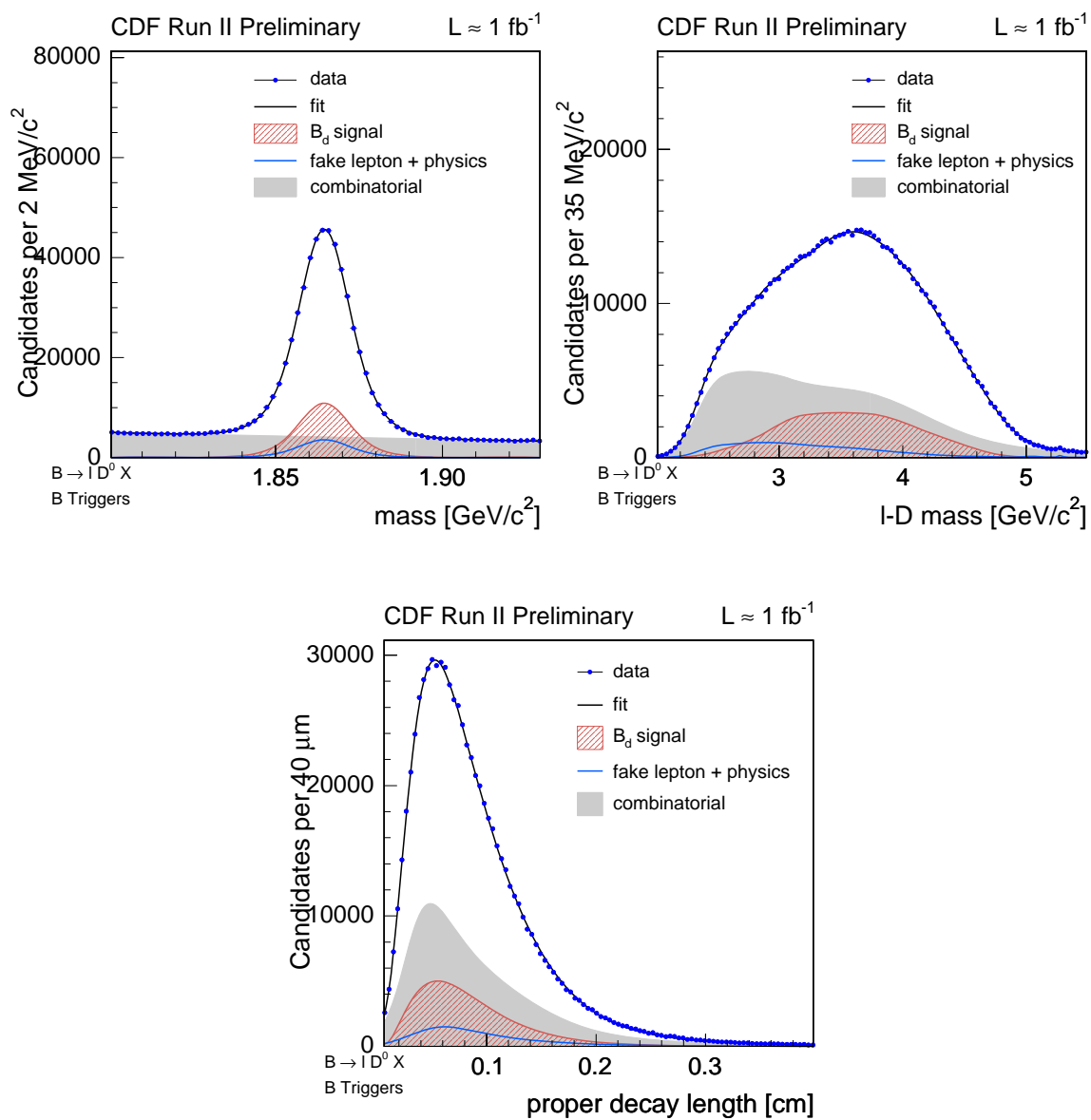


Figure 5.28: D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ \bar{D}^0 X$ candidates in the B triggers.

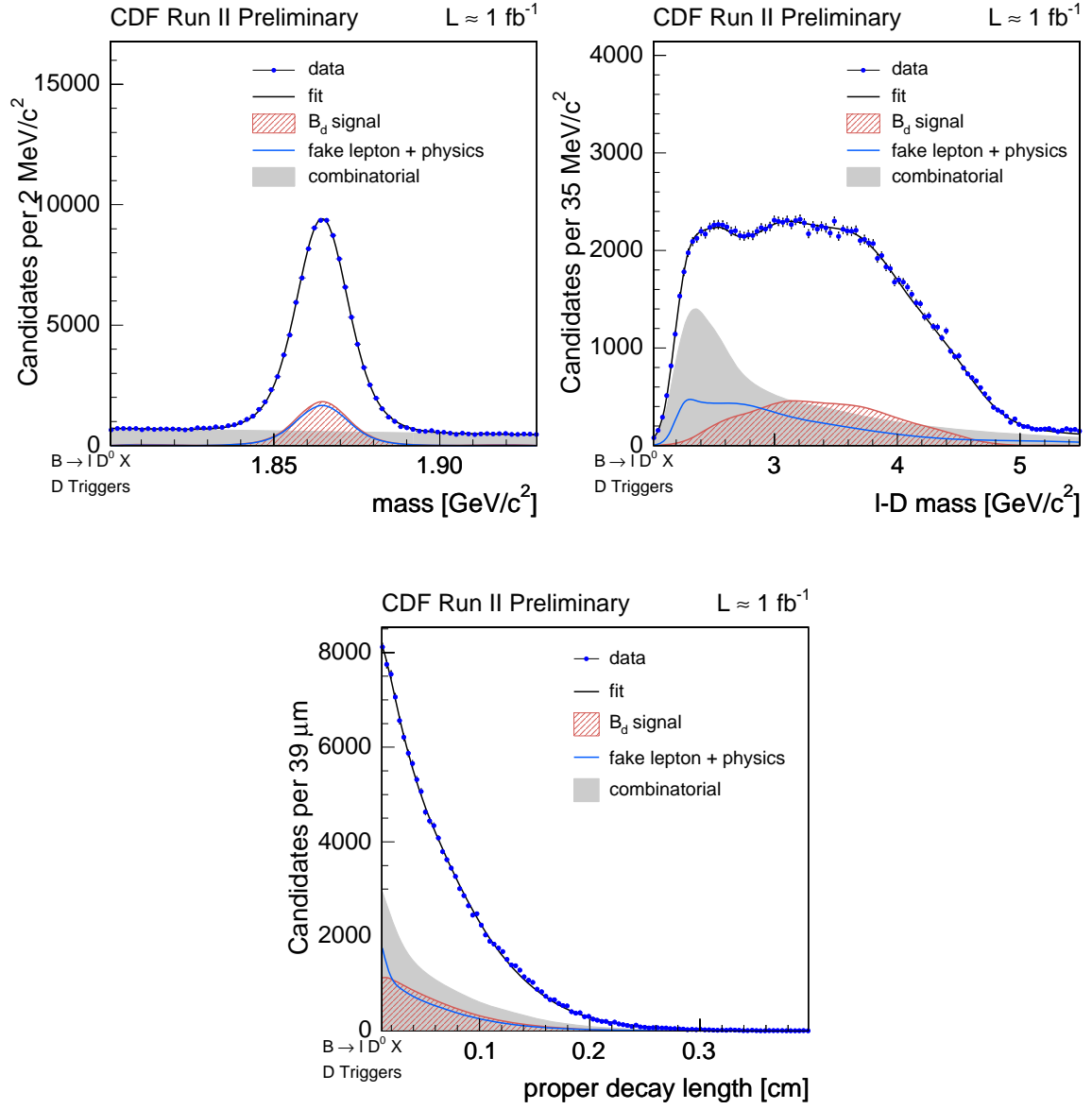


Figure 5.29: D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ \bar{D}^0 X$ candidates in the D triggers.

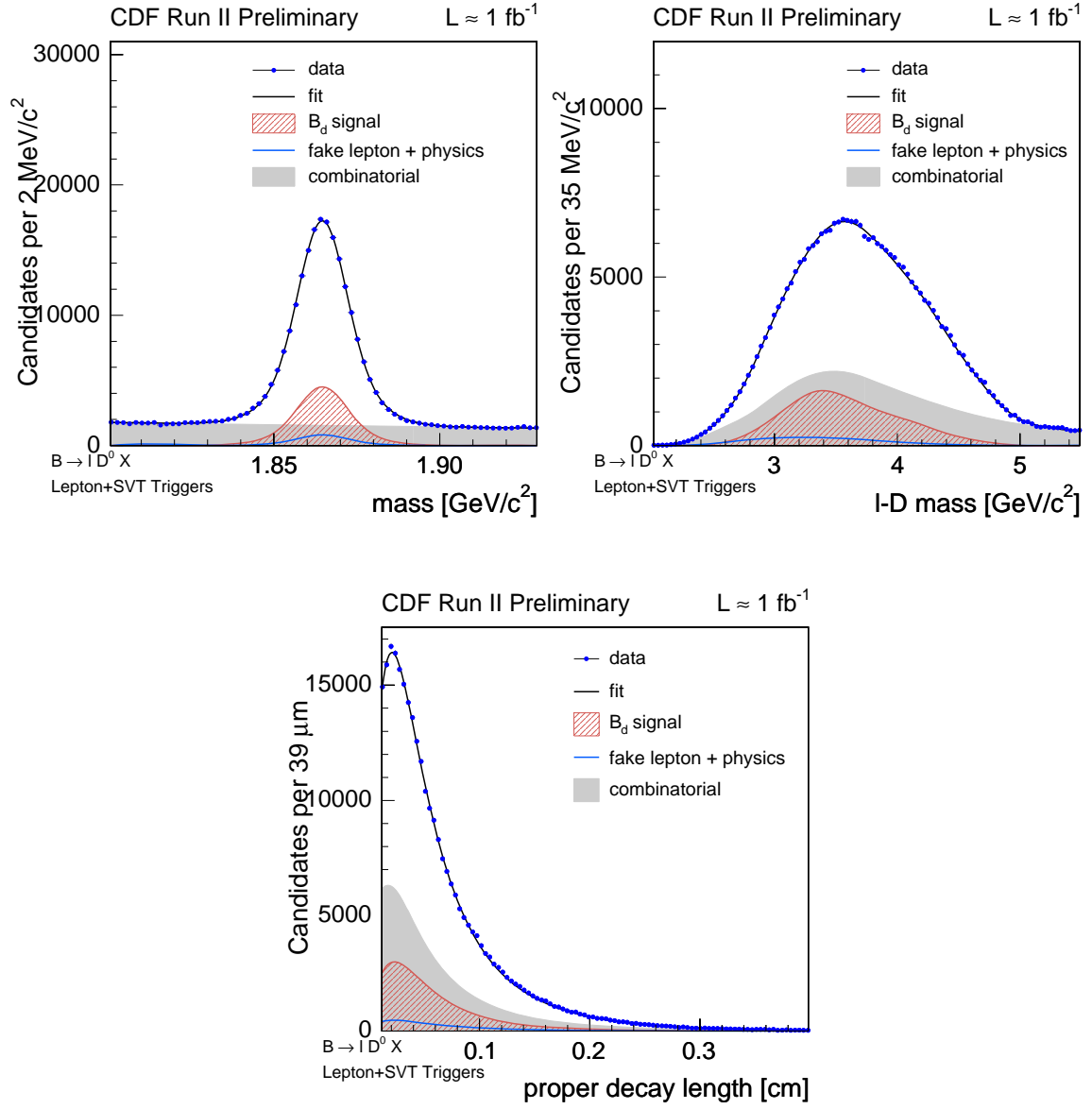


Figure 5.30: D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ \bar{D}^0 X$ candidates in the $l + SVT$ triggers.

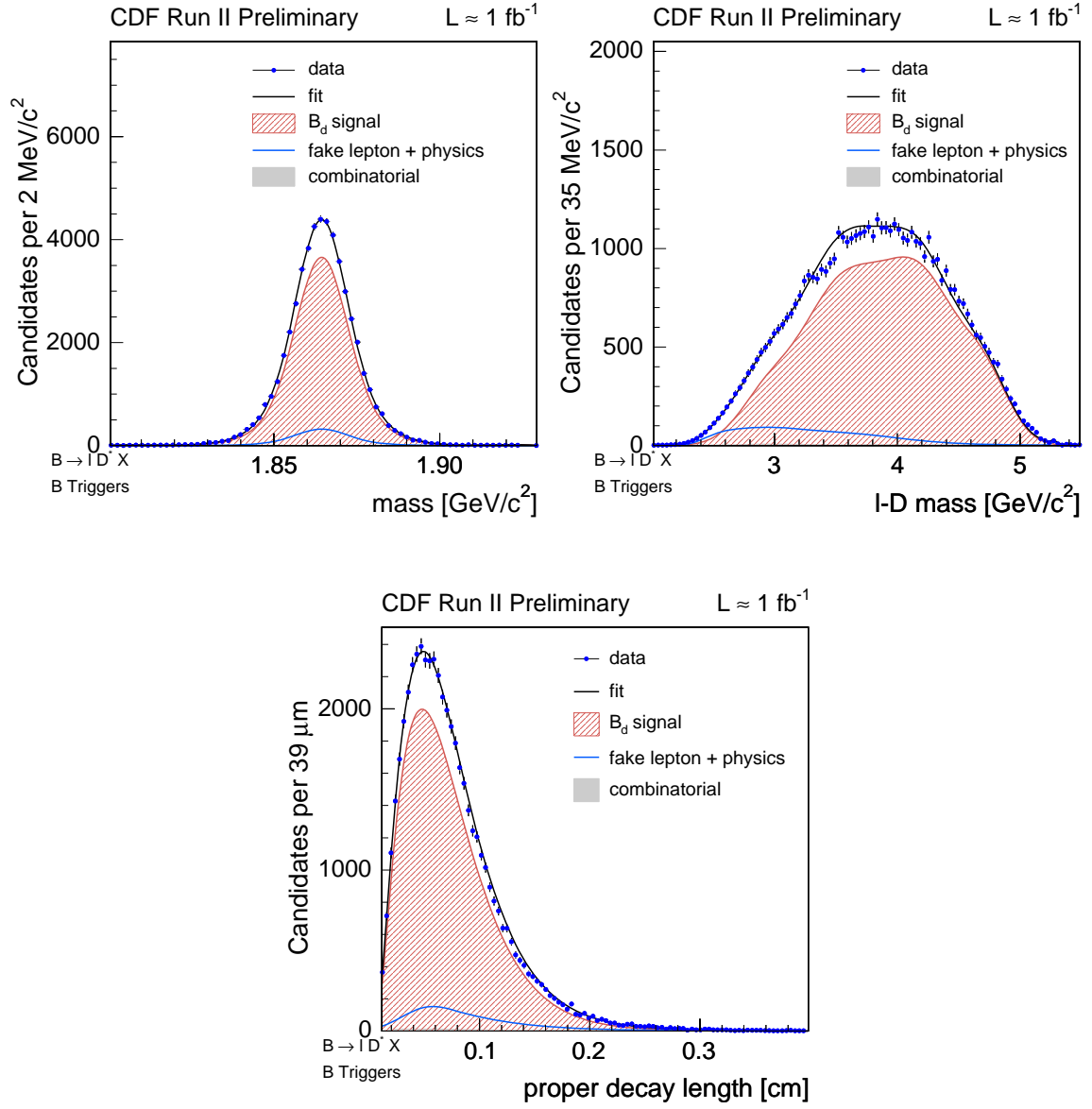


Figure 5.31: D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^{*-} X$ candidates in B triggers.

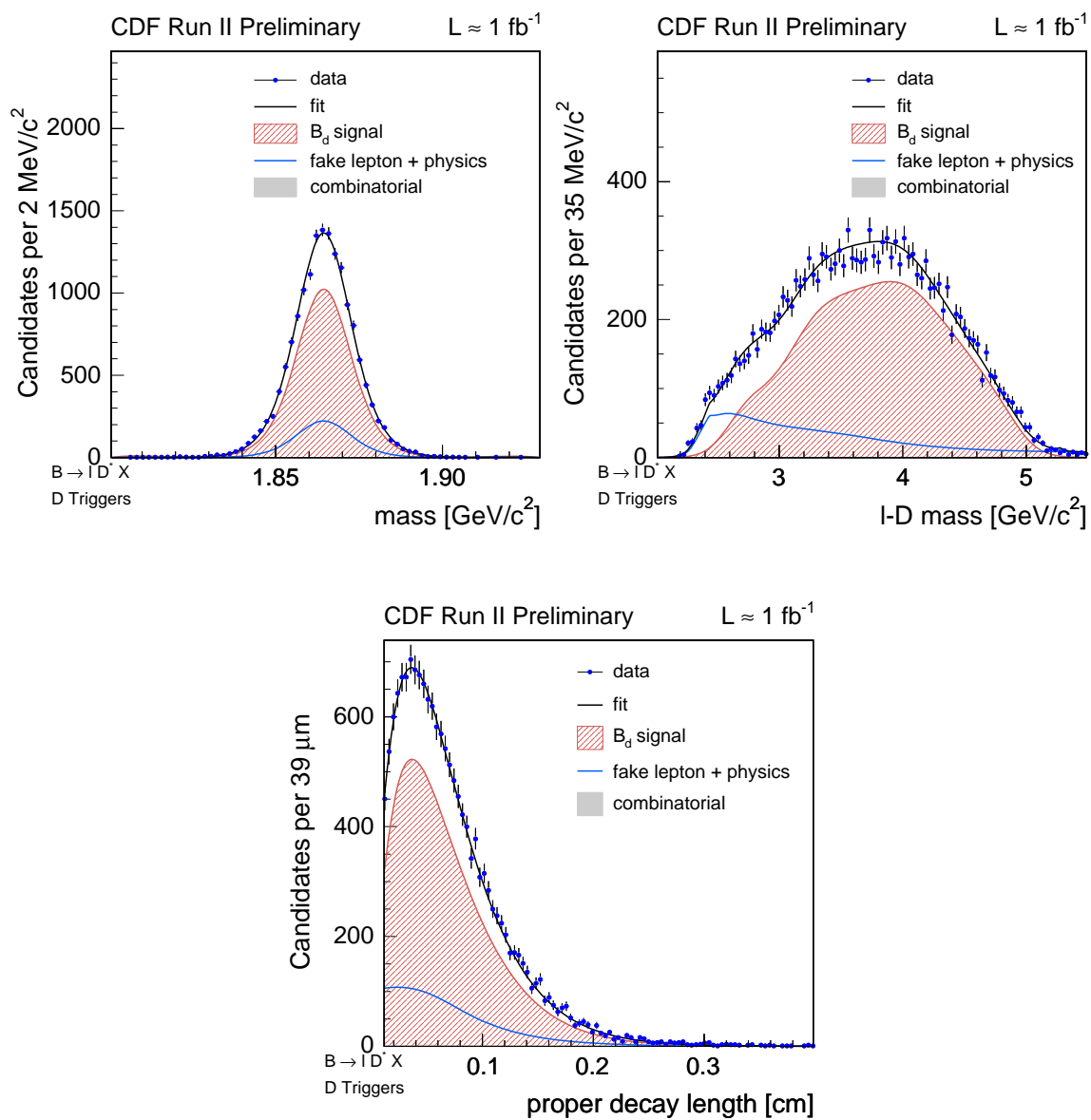


Figure 5.32: D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^{*-} X$ candidates in D triggers.

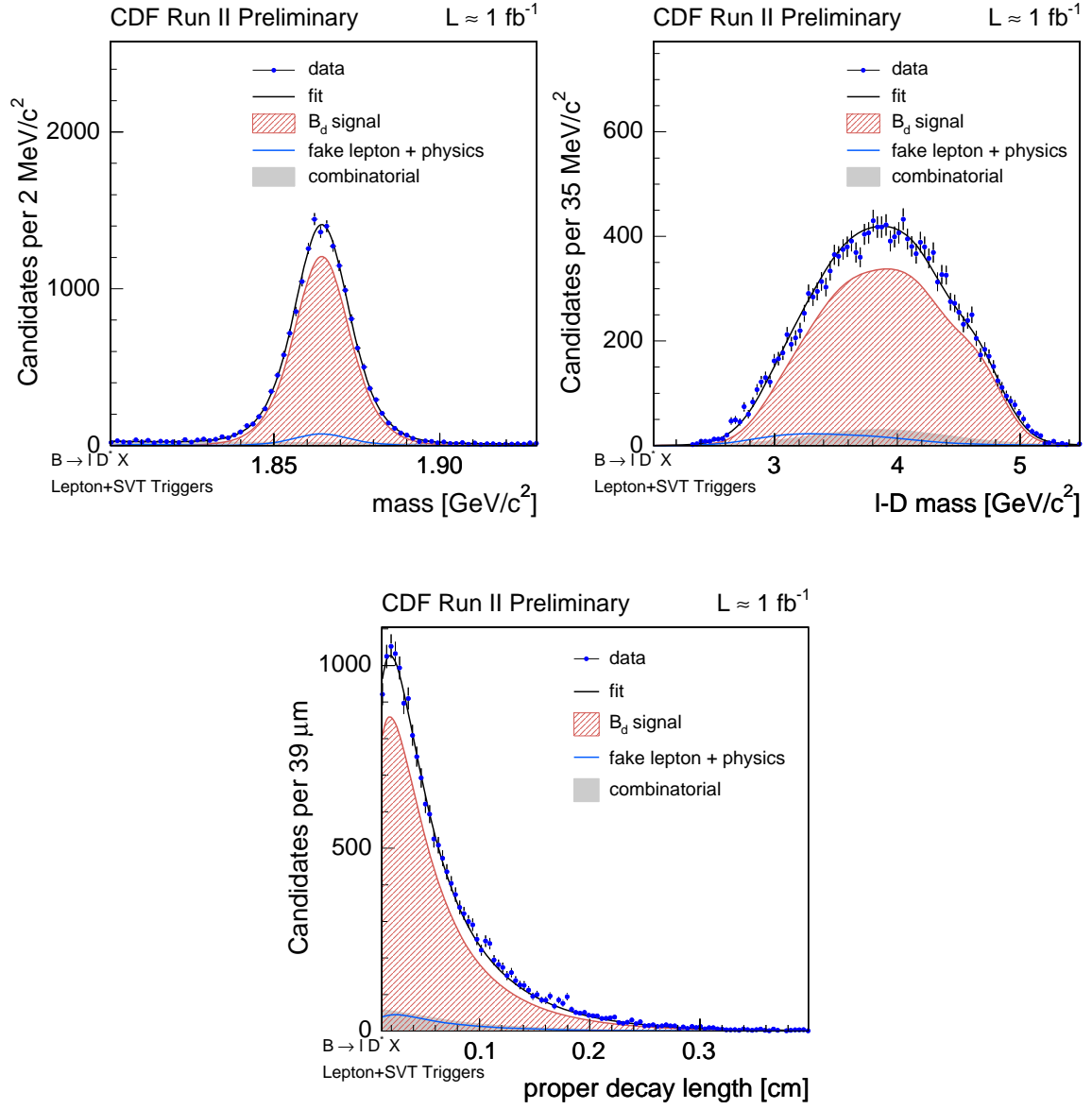


Figure 5.33: D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^{*-} X$ candidates in $l + SVT$ triggers.

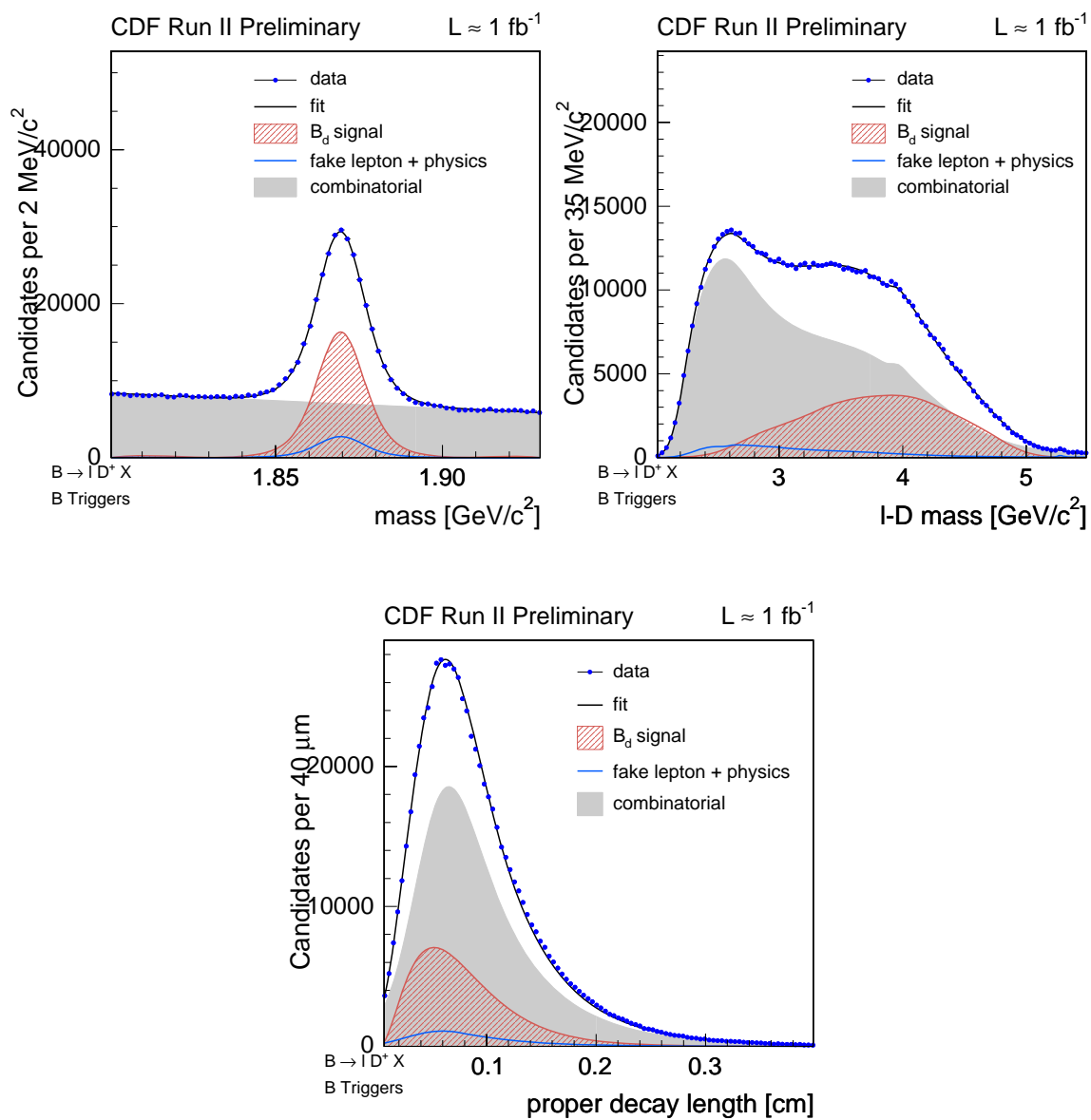


Figure 5.34: D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^- X$ candidates in B triggers.

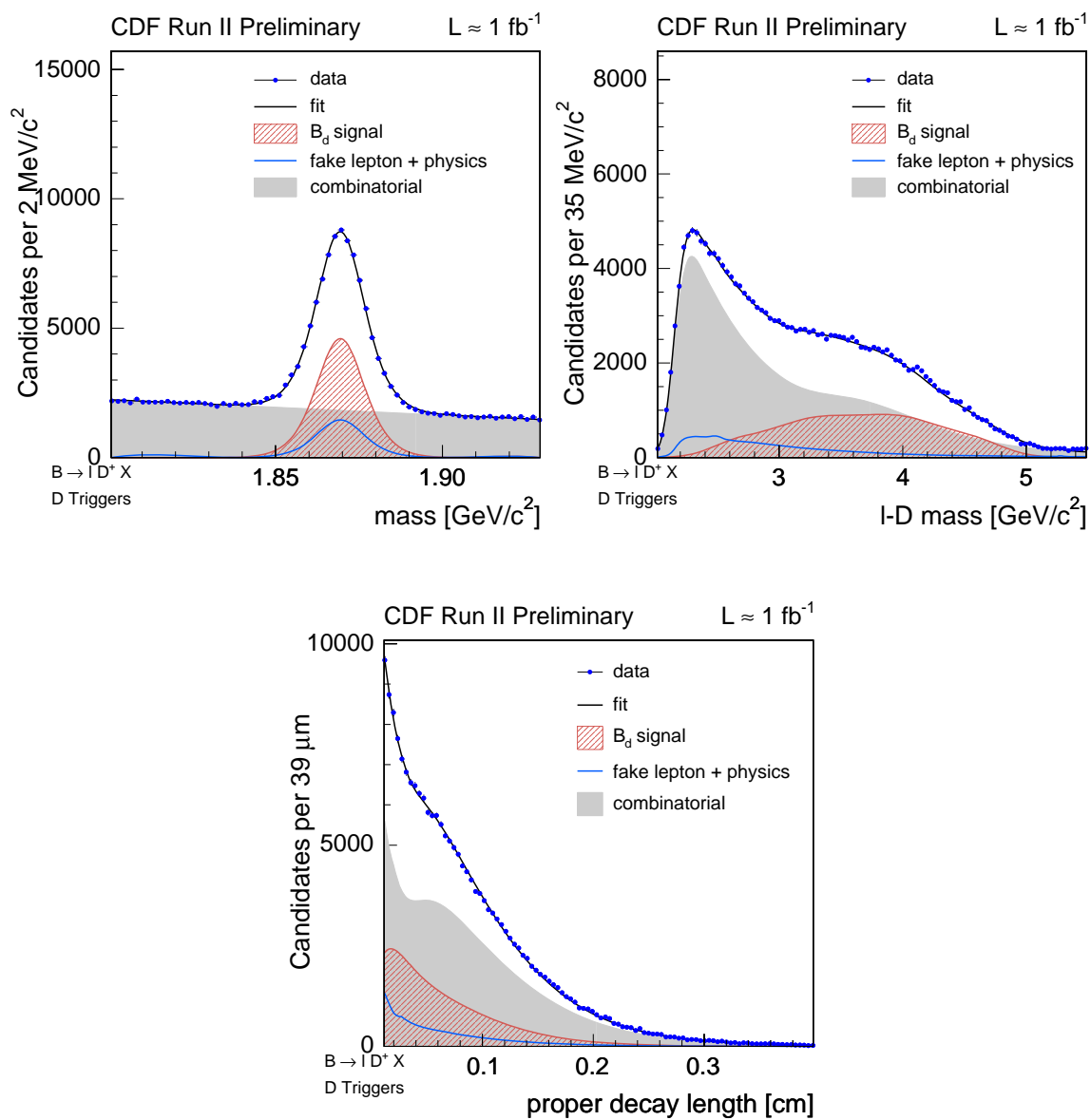


Figure 5.35: D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^- X$ candidates in D triggers.

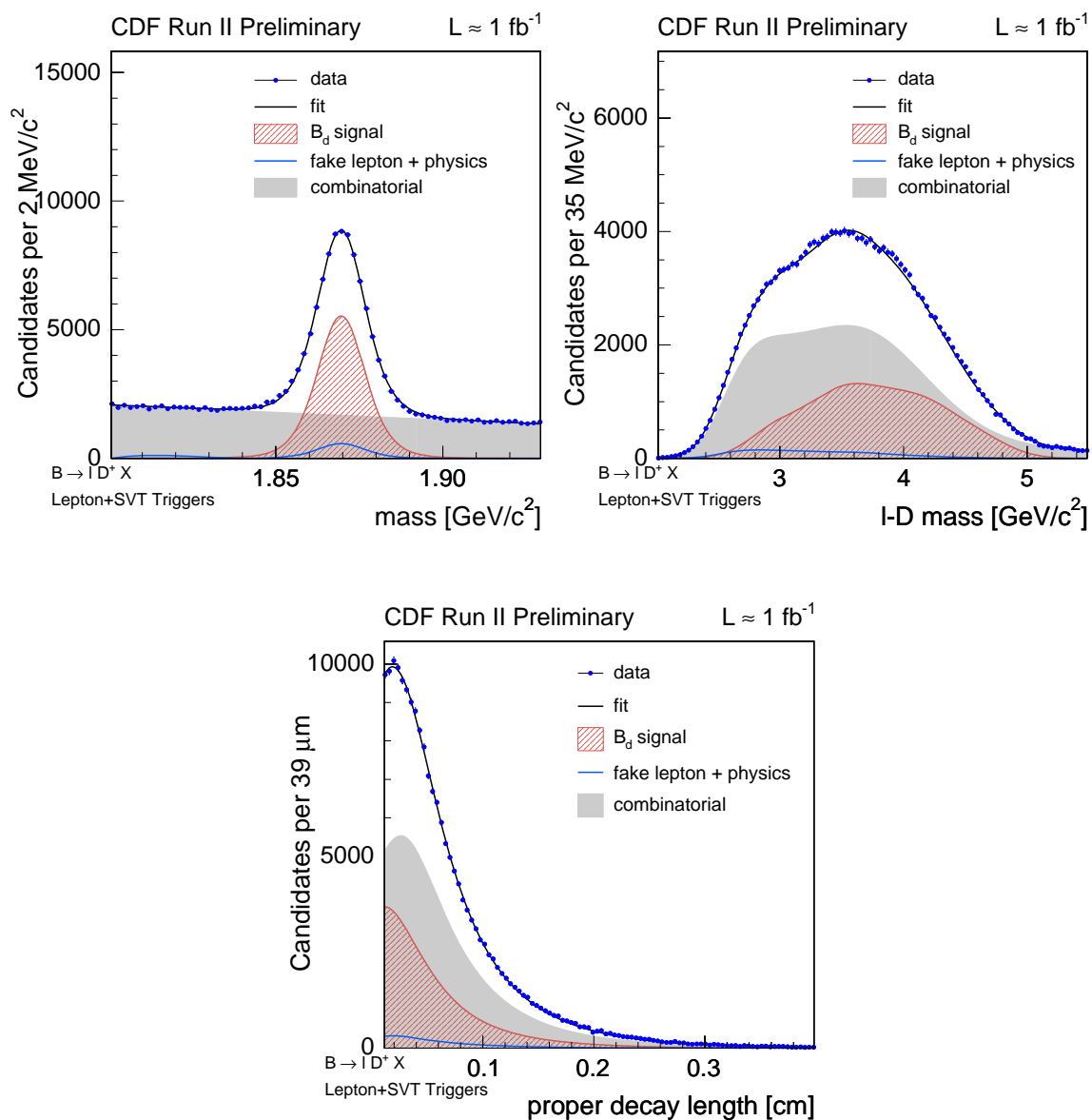


Figure 5.36: D mass, $m_{\ell D}$ and ct^* distributions for the $B^{+,0} \rightarrow \ell^+ D^- X$ candidates in $l + SVT$ triggers.

Chapter 6

Initial State Flavor Tagging

One of the components of measuring neutral B meson flavor oscillations is identifying the flavor of the B meson (containing a \bar{b} antiquark) or the \bar{B} meson (containing a b quark) at both production and decay in order to determine whether the neutral B meson has oscillated. We refer to this method of identifying the B meson flavor as “ B flavor tagging”. Events are classified on the basis of the sign of the production and decay tagging decision as mixed or unmixed. The figure of merit to compare different flavor tagging techniques is the so-called effective tagging power $\varepsilon \mathcal{D}^2 = \varepsilon(1 - 2p_W)^2$, where the efficiency ε represents the fraction of events for which a flavor tag exists and p_W is the mistag probability indicating the fraction of events with a wrong flavor tag. The mistag probability is related to the dilution, another quantity used to express the power of a flavor tag: $\mathcal{D} = 1 - 2p_W$. The dilution \mathcal{D} itself is defined as the number of correctly tagged events N_R minus the number of incorrectly identified events N_W divided by the sum of both:

$$\mathcal{D} = \frac{N_R - N_W}{N_R + N_W}. \quad (6.1)$$

Thus, a flavor tag which always returns the correct tag has a dilution of 1, while a random tag yielding the correct flavor 50% of the time has a dilution of zero. Finally, the efficiency ε of flavor tagging is expressed as:

$$\varepsilon = \frac{N_R + N_W}{N}, \quad (6.2)$$

where N is the total number of candidates that a flavor tag is applied to.

In this measurement, the b quark flavor at decay time is identified by the charge of the lepton in the \bar{B}_s^0 decay ($\bar{b} \leftrightarrow D_s^-$ and $b \leftrightarrow D_s^+$). Several methods to tag the production b quark flavor exist. The production flavor tags can be divided into two groups, those that tag the initial charge of the b quark contained in the \bar{B} candidate itself (same side tag - SST) and those that tag the initial charge of the other quark (\bar{b}) produced in the same event (opposite side tag - OST). The dominant production

mechanisms of b quarks in hadronic $p\bar{p}$ collisions at the Tevatron produce $b\bar{b}$ pairs. Opposite-side flavor tags exploit this fact by identifying the flavor of the second bottom meson in the event and inferring the production flavor of the first B meson of interest.

Figure 6.1 is a sketch of a $b\bar{b}$ event showing the B and \bar{B} mesons originating from the primary $p\bar{p}$ interaction vertex. The B meson decays at a secondary vertex indicating possible flavor tags on the decay vertex side (SST) as well as opposite side tags. Three methods of opposite side flavor tagging are employed in this analysis. One method, called “lepton tagging”, looks for an electron or muon from the semileptonic decay of the opposite side B meson in the event. The charge of this lepton is correlated with the flavor of the B meson: an ℓ^- comes from a $b \rightarrow c \ell^- \bar{\nu} X$ transition, while an ℓ^+ originates from a \bar{b} quark. Since the semileptonic B branching fraction is small, $\mathcal{B}(B \rightarrow \ell X) \sim 20\%$, lepton tags are expected to have low efficiency but high dilution because of the high purity of lepton identification (see Sections 3.2 and 3.3). Second, the charge of a K^\pm from the subsequent charm decay $c \rightarrow s X$ is also correlated with the B flavor: a K^- results from the decay chain $b \rightarrow c \rightarrow s$ while a K^+ signals a \bar{b} flavor. Searching for a charged kaon from the opposite side B meson decay is referred to as “kaon tagging”. This method is expected to have high efficiency but low dilution since the challenge is to first identify kaons among a vast background of pions, and secondly to find the kaon candidate from the B decay among all other kaons present in the hadronic collision event. Third, a method called “jet charge tagging” exploits the fact that the sign of the momentum weighted sum of the particle charges of the opposite side b jet is correlated to the charge of the b quark producing this jet. Jet charge tags are expected to have high efficiency but lower dilution.

Additionally, the flavor of a B meson can also be tagged on the same side as the B meson of interest by exploiting correlations of the B flavor with the charge of particles produced in association with the B meson. Such correlations are expected to arise from b quark hadronization and from B^{**} decays. In the case of a B^- or \bar{B}^0 mesons, the fragmentation particles are mainly pions while \bar{B}_s^0 mesons are primarily accompanied by fragmentation kaons. In the \bar{B}_s^0 meson case we thus refer to this method as “same side kaon tagging” (SSKT).

As mentioned above (see also Section 1.5), the statistical power of a given data sample scales with the effective tagging power $\varepsilon\mathcal{D}^2$ of the flavor tagging method. At the Tevatron, the typical flavor tagging power of a single tagging algorithm is $\mathcal{O}(1\%)$. Limitations in opposite side tagging algorithms arise because the second bottom meson is inside the detector acceptance in less than 40% of the time or it is possible that the second B meson is a neutral B meson that mixed into its antiparticle. For example, the low efficiency of an opposite side lepton tag of $\sim 20\%$ from the semileptonic B meson branching fraction together with an average dilution of $\sim 30\%$ results in an estimated $\varepsilon\mathcal{D}^2 \sim 0.4 \times 0.2 \times 0.3^2 \sim 0.01$.

In applying these tagging methods, we quantify the tagging performance on an event by event basis, in order to weight events in the mixing fit based upon our level

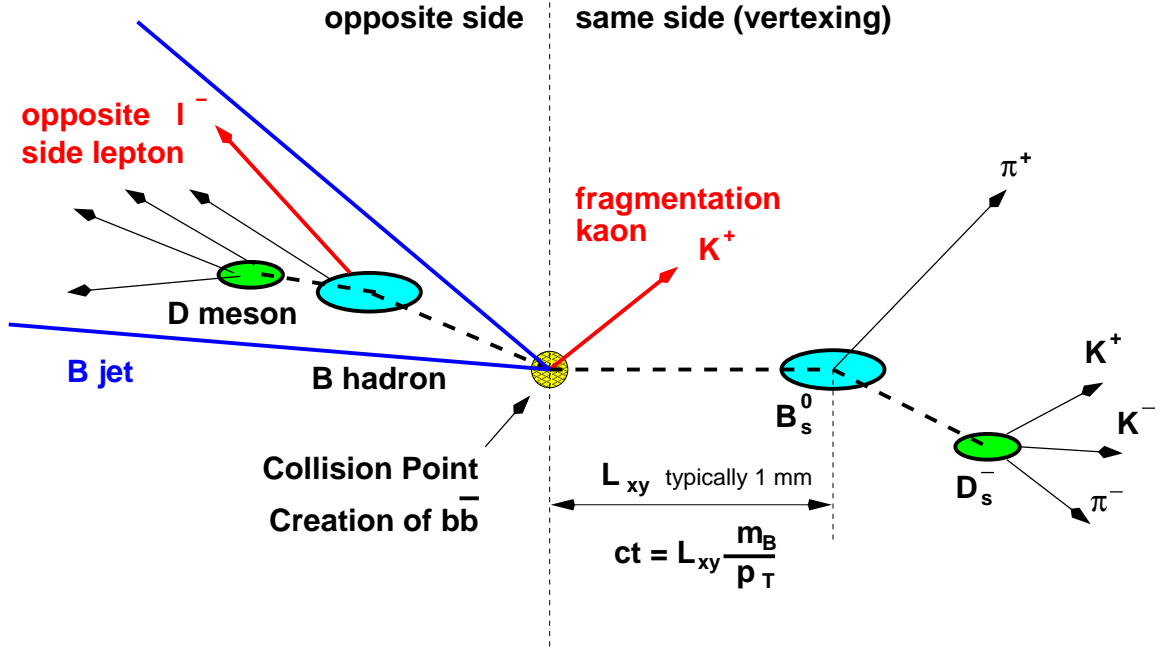


Figure 6.1: Sketch of typical $b\bar{b}$ event indicating several B flavor tagging techniques.

of confidence that the B flavor tag is correct. In addition, we combine the decision of each of the three opposite side tagging algorithms using an artificial neural network (ANN) [67] to improve the tagging power by taking into account the correlations among the tagging algorithms. Finally, we combine the decision of the same side and opposite side tags as described in Section 6.3.

In the case of opposite side flavor tags, the dilution is expected to be independent of the type of B meson that produces the hadronic or semileptonic decay. We first study the performance of a particular OST on a large sample of semileptonic B meson decays collected with the ℓ +SVT trigger (see Section 2.8). Lepton tags have low efficiency, and consequently require a large sample for calibrating the tag dilution. This large sample is essential to determining the dependence of the tagging dilution as a function of kinematic variables such as p_T^{rel} , the lepton momentum transverse to the bottom jet direction (see Sections 6.1.3 and 6.1.4), with small statistical uncertainties. The obtained parameterization of the dilution against a kinematic quantity is used to predict the dilution of a particular opposite side tag in the B_s^0 mixing analysis. Rather than just predicting the sign of a flavor tag, this procedure maximizes the effectiveness of a tagging method on an event by event basis (see Chapter 7).

6.1 Opposite Side Flavor Tags

As outlined above, we utilize three opposite side tagging techniques in this measurement: lepton tagging, kaon tagging and jet charge tagging. The performance of the opposite side tagging algorithms is studied in high statistics samples of semileptonic B meson decays collected with the ℓ +SVT trigger described in Section 4.2.2. The charge of the trigger lepton in the ℓ +SVT sample provides an estimate of the trigger side B meson flavor. Hence, the opposite side algorithms can be applied on an event by event basis to the unbiased away side to determine whether the flavor tag has the same or opposite sign as the trigger lepton charge. However, the ℓ +SVT data is not a pure sample of B decays. In addition to events from semileptonic B decays, it also contains semileptonic charm decays, hadrons that fake the trigger lepton and other backgrounds which do not originate from bottom mesons. Therefore, prior to using the ℓ +SVT sample for opposite side tagging studies, the sample composition of the ℓ +SVT data needs to be understood. A background subtraction procedure is utilized to determine the B purity of the sample. This allows us to study the performance that an opposite side flavor tag would achieve in a pure sample of bottom mesons and to quantify the tag performance in terms of its dilution \mathcal{D} and $\varepsilon\mathcal{D}^2$.

The background subtraction procedure is based on the assumption that all sources of background are symmetric in the signed impact parameter of the SVT track which is defined as

$$\delta(\text{SVT}) = |\vec{d}_0| \text{sign}[\vec{d}_0 \cdot \vec{p}(\ell + \text{SVT})]. \quad (6.3)$$

where \vec{d}_0 is the vector pointing from the primary vertex to the point of closest approach of the SVT track with respect to the primary $p\bar{p}$ interaction vertex and $\vec{p}(\ell + \text{SVT})$ is the vector direction sum momentum of the trigger lepton plus SVT track. The convention of signing the impact parameter of the SVT track with respect to the ℓ +SVT direction can be understood with the sketch on the left-hand side of Figure 6.2. If the SVT track intersects the reference vector along the positive (negative) ℓ +SVT-axis with respect to the position of the primary $p\bar{p}$ vertex, $\delta(\text{SVT})$ is defined to be positive (negative) as shown in the figure. Backgrounds arising from uncorrelated cases where the lepton and SVT track do not originate from the same parent particle, for example, when prompt tracks are identified as the trigger leptons, are expected to populate positively and negatively signed impact parameters equally. Thus, this technique of signing the impact parameter of the SVT tracks can be used to statistically subtract such backgrounds. To obtain a distribution that is characteristic for a pure bottom meson sample, the distribution with negative $\delta(\text{SVT})$ is subtracted from the corresponding distribution with positive $\delta(\text{SVT})$.

After performing the background subtraction using the $\delta(\text{SVT})$ distribution, Figure 6.2 shows on the right-hand side the invariant mass distribution of the lepton and SVT track system assuming the pion mass for the SVT track. The shapes of the ℓ +SVT invariant mass distribution are overlaid for the case when both the lep-

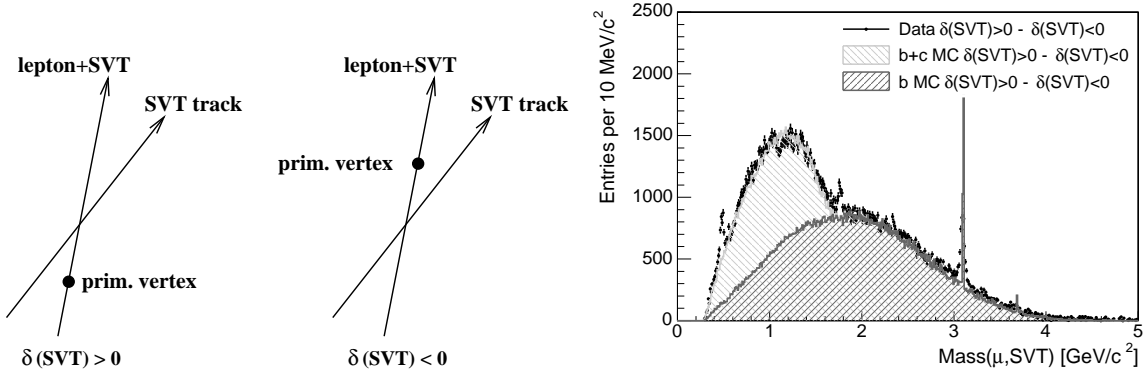


Figure 6.2: Defining signed impact parameter of the SVT track (left). Invariant mass distribution of the ℓ +SVT pairs after background subtraction using the signed SVT impact parameter (right).

ton and the SVT track originate from a bottom (charm) meson as obtained from a large Monte Carlo sample of inclusive bottom (charm) meson decays. To reject ℓ +SVT events from charm production, only event candidates are selected with $2 \text{ GeV}/c^2 < m_{\ell+\text{SVT}} < 4 \text{ GeV}/c^2$ for opposite side tagging studies, where $m_{\ell+\text{SVT}}$ is the invariant mass of the lepton and SVT track system. After removing charm decays and other background with the signed impact parameter subtraction method, a sample containing pure bottom meson decays on the trigger side is obtained. In order to correctly use the charge of the trigger side lepton as indicator of the B flavor on the trigger side, a correction has to be taken into account for the cases where the trigger side lepton originates from a double-sequential leptonic decay ($b \rightarrow c \rightarrow s\ell$) or the trigger side B meson is neutral and oscillated to the flavor opposite its production flavor. These effects are corrected on average using a large Monte Carlo sample of inclusive bottom meson decays. This average correction is called the trigger side dilution $\mathcal{D}_{\text{trig}}$. We find

$$\mathcal{D}_{\text{trig}} = \begin{cases} 0.6412 \pm 0.0015 \text{ (stat)}^{+0.0141}_{-0.0226} \text{ (syst)} & \text{for } \mu+\text{SVT} \text{ and,} \\ 0.6412 \pm 0.0015 \text{ (stat)}^{+0.0215}_{-0.0367} \text{ (syst)} & \text{for } e+\text{SVT}, \end{cases} \quad (6.4)$$

where the systematic errors are dominated by the uncertainty in the trigger side fake lepton contribution. The true dilution of the opposite side flavor tag is then obtained from its measured dilution \mathcal{D}_{raw} as $\mathcal{D}_{\text{true}} = \mathcal{D}_{\text{raw}}/\mathcal{D}_{\text{trig}}$.

6.1.1 Track-Based Jet Reconstruction

The opposite side flavor tagging algorithms described below require the reconstruction of jets that originate from the production of b quarks. We refer to such jets as “ b jets”. Due to the steeply falling b quark production cross section at the Tevatron and the trigger conditions, in most of the recorded events the b jet has low momentum.

Thus its energy is usually not very precisely measured in the calorimeter. Jets are therefore formed from COT tracks using a cone-clustering algorithm. Tracks with $p_T > 1.0 \text{ GeV}/c$ are used as jet seeds. If two seeds are within a cone with radius $\Delta R < 0.7$, the momenta of the seeds are added together to form a new seed. After all possible seed merging, lower momentum tracks ($0.4 < p_T < 1.0 \text{ GeV}/c$) that are within $\Delta R < 0.7$ of a seed are added in to form the final jets. A jet can, in principle, consist of a single track with $p_T > 1 \text{ GeV}/c$.

In order to increase the probability that such a track-based jet originates from a b quark, the reconstruction of the decay point of the B meson (referred to as the secondary vertex) within the jet is attempted. We refer to jets with an identified secondary vertex as **SECVTX** jets [64]. The search for a secondary vertex in the jet is a two stage process. In both stages, tracks in the jet are selected for reconstruction of a secondary vertex based on the significance of their impact parameter with respect to the primary vertex, d_0/σ_{d_0} , where σ_{d_0} is the estimate of the uncertainty on d_0 . The uncertainty σ_{d_0} includes contributions from both the primary vertex and the track parameters. The first stage requires at least three candidate tracks for the reconstruction of the secondary vertex. Tracks consistent with coming from the decay $K_S^0 \rightarrow \pi^+\pi^-$ or $\Lambda^0 \rightarrow p\pi^-$ are not used as candidate tracks. Two candidate tracks are constrained to pass through the same space point to form a seed vertex. If at least one additional candidate track is consistent with intersecting this seed vertex, then the seed vertex is used as the secondary vertex. If the first stage is not successful in finding a secondary vertex, the second stage is attempted. More stringent track requirements (on d_0/σ_{d_0} and p_T , for example) are imposed on the candidate tracks. All candidate tracks satisfying these stricter criteria are constrained to pass through the same space point to form a seed vertex. This vertex has an associated χ^2 . Candidate tracks that contribute too much to the χ^2 are removed, and a new seed vertex is formed. This procedure is iterated until a seed vertex remains that has at least two associated tracks and an acceptable value of χ^2 .

6.1.2 Selection of Opposite Side Tag Candidates

Given a sample of B meson decay candidates, we define the same side or the “trigger side” as an η - φ cone of $\Delta R < 0.7$ around the B meson direction. All the tracks in a given event within this cone are considered to be originating from the same side b quark. Therefore, to search for opposite side tag candidates, we look outside the same side cone. That is, tracks satisfying $\Delta R > 0.7$ with respect to the candidate B meson are considered for use in the opposite side tagging algorithms. With this criteria for the selection of the opposite side tag candidates, we describe the various flavor tagging algorithms used in this measurement in the following.

6.1.3 Electron Tagging

The goal of the electron tagging algorithm is to find an electron arising from a semileptonic decay of the opposite side B meson. To this end, we first select opposite side tracks based on the criteria described in Section 6.1.2 and impose the following additional requirements:

- Tracks with transverse momentum $p_T > 2.0$ passing the track quality criteria described in Section 4.3, where L00 hits are not used on the tag candidates.
- Remove tag candidates consistent with originating from photon conversions, as discussed in Section 3.4.
- $|\Delta z| < 5.0$ cm, the distance the between the tagging track and the B candidate along the beam axis.

The tracks satisfying the above criteria are further required to meet the initial electron candidate criteria, as detailed in Section 3.2. At this point, an electron likelihood \mathcal{L}_e for the electron tag candidate is calculated using information from various discriminating variables in the calorimeter, COT, CES and CPR detectors. The electron likelihood is described in Section 3.2.3.

The eligible electron candidates are now associated with track-based jets reconstructed in the event as detailed above in Section 6.1.1. Sometimes, an electron tag candidate is found to be “isolated” and is not associated with a jet. Such candidates are called “globally isolated” to distinguish them from the local isolation defined in the calorimeter. Next, we calculate the p_T^{rel} of the candidate tracks. The quantity p_T^{rel} is defined as the magnitude of the component of the candidate track momentum that is perpendicular to the axis of the jet associated with the lepton tag. The electron candidate is removed from the jet to which it was assigned, and the jet axis is recalculated to determine p_T^{rel} .

If there are two or more electron candidates in an event which satisfy the above selection criteria, we first consider all the tag candidates that pass a minimum likelihood requirement of $\mathcal{L}_e > 0.05$. Among the candidates with $\mathcal{L}_e > 0.05$, we find a globally isolated candidate first as it is more likely to originate from the semileptonic decay of the opposite side B meson. Otherwise, the candidate with the highest p_T^{rel} is selected among all the candidates passing the minimum likelihood requirement. If none of the multiple candidates pass the minimum likelihood cut $\mathcal{L}_e > 0.05$ requirement, we select the globally isolated candidate first (if one exists), otherwise we choose the electron tag candidate with the maximum p_T^{rel} .

Backgrounds for Electron Tagging and Discriminating Variables

As described in Section 3.4, there are various sources of real and fake electrons. We are interested in identifying real electrons originating from the semileptonic decays of the

opposite side B mesons. The main sources of backgrounds for opposite side electron tagging are hadrons faking the electron signature, electrons from sequential $b \rightarrow c \rightarrow \ell$ decays, and electrons from unidentified photon conversions. The unidentified photon conversions include the cases where the conversion pair partner to the electron candidate track has a very low transverse momentum and is hence not detected in the COT. Additionally, there are photon conversions that are not identified because of the inefficiency of the conversion identification criteria described in Section 3.4.

We analyze a variety of variables that can enable us understand various backgrounds for opposite side electron tags. The electron likelihood \mathcal{L}_e is a strong discriminator between electrons and hadrons, the most prominent source of background. However, we cannot use \mathcal{L}_e to discriminate against electrons from background sources. p_T^{rel} is expected to discriminate between electrons from direct semileptonic B decays and electrons from sequential decays, fake electrons as well as electrons from unidentified photon conversions. p_T^{rel} is, in general, larger for an electron originating from a semileptonic B decay as compared to the background sources. In the following, we discuss a discriminating variable which can be used to distinguish electrons from massive decays and electrons from photon conversions.

We define the signed impact parameter for a given track (d_0^{sign}) as the product of impact parameter d_0 signed with the charge q of the track, $d_0^{\text{sign}} = d_0 \times q$. In the absence of resolution effects in tracking, electrons emanating from conversions, $\gamma \rightarrow e^+e^-$, always have a positively signed impact parameter [65]. This is due to the fact that photons are massless, and usually point back to the primary interaction vertex because they are predominantly produced in $\pi^0 \rightarrow \gamma\gamma$ decays, where the π^0 is directly produced in the $p\bar{p}$ interaction. This is depicted in the cartoon on the left-side of Figure 6.3. A photon from a π^0 decay travels a finite distance, interacts with the detector material and converts into an e^+e^- pair. The impact parameter of the resulting e^+e^- pair is itself a signed quantity. For example, the positron e^+ in the cartoon, which goes to the left, has a positive impact parameter, while e^- which goes to the right has a negative impact parameter. Therefore, signing the impact parameter with the charge of the track, the signed impact parameter becomes positive for both e^+ and e^- . Electrons produced in a semileptonic B decay or hadrons faking the electron signature are expected to be symmetric in the signed impact parameter d_0^{sign} , because they originate from massive decays. To check whether this is indeed true for electrons from identified conversions (see Section 3.2.1), we plot the signed impact parameter of the unbiased leg of the conversion. This is displayed in Figure 6.4. As expected, a large excess is seen on the positive values of d_0^{sign} as compared to the negative values. Therefore, electrons from photon conversions tend to have a preference for $d_0^{\text{sign}} > 0$.

The purity of the opposite side electron tags is likely to depend on the three discriminating variables described above: p_T^{rel} of the tagging track, likelihood \mathcal{L}_e , and the signed impact parameter of the track d_0^{sign} .

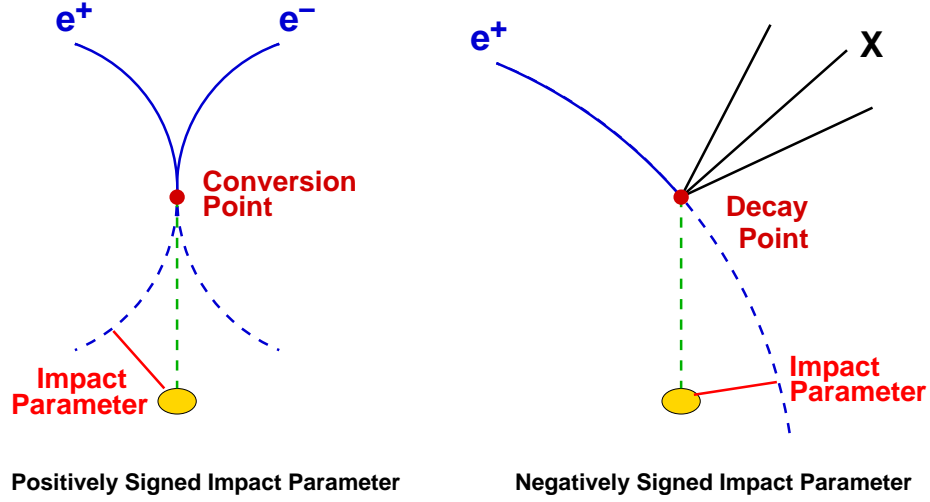


Figure 6.3: Cartoons depicting signed impact parameters for electrons originating from conversions (left) and electrons from B decays (right).

Electron Tagging Performance

The performance of the opposite side electron tagging algorithm is determined in the high statistics samples of semileptonic B meson decays collected with the ℓ +SVT trigger as described above, and in Section 4.2.2. We utilize $\sim 180 \text{ pb}^{-1}$ of the e +SVT and μ +SVT datasets. Based on the discussion about the backgrounds for opposite side electrons in the previous section, we expect the tagging dilution to depend on the three discriminating variables. Ultimately, we would like to separate the opposite side electron candidates into samples according to the various discriminating variables and make use of the dependence of dilution on them. We first split the opposite side electron samples in the ℓ +SVT data sets into two bins in d_0^{sign} : $d_0^{sign} > 0$ and $d_0^{sign} < 0$. We then investigate the dependence of dilution on \mathcal{L} and p_T^{rel} in each of the bins of d_0^{sign} .

High quality electrons have \mathcal{L}_e values close to one while low quality electrons have \mathcal{L}_e values close to zero. There are also electron candidates that have intermediate likelihood values, populated with an admixture of fake and real electrons. Since the dilution of fake electrons should be close to zero, we expect that the higher the value of \mathcal{L}_e , the higher the dilution would be. To investigate the dependence of dilution on the value of \mathcal{L}_e , we subdivide the $d_0^{sign} > 0$ and $d_0^{sign} < 0$ subsamples into six bins of \mathcal{L}_e : $0.0 \leq \mathcal{L} \leq 0.001$, $0.001 < \mathcal{L} \leq 0.05$, $0.05 < \mathcal{L} \leq 0.65$, $0.65 < \mathcal{L} \leq 0.99$, $0.99 \leq \mathcal{L} \leq 0.999$ and $0.999 \leq \mathcal{L} \leq 1.0$. The efficiencies and dilutions for these subsamples binned in \mathcal{L}_e are shown in Figure 6.5 for the μ +SVT data. For each bin in \mathcal{L}_e , dilution is calculated using Equation (6.1). The cases where the electron tagging algorithm finds an opposite side tag with a charge opposite to the trigger lepton are

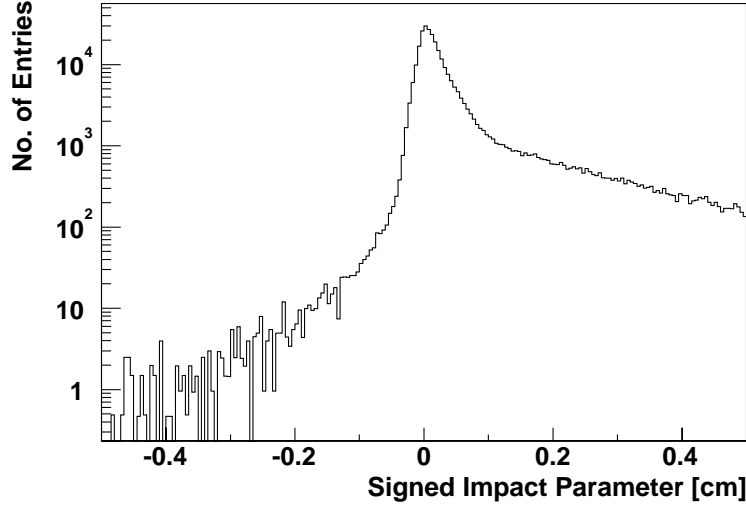


Figure 6.4: Distribution of d_0^{sign} for identified conversions.

labeled as correct tags, while the cases where the tag electron and the trigger lepton are of the same charge are flagged as incorrectly tagged. A large fraction of tagging efficiency is achieved in the first two bins of \mathcal{L}_e . The efficiencies for $d_0^{sign} > 0$ and $d_0^{sign} < 0$ are consistent in these two bins supporting our argument that hadrons are symmetric in d_0^{sign} . The dilution in those two bins is ~ 0 indicating that these candidates are almost completely dominated by hadrons, which do not provide flavor information of the opposite side B meson. In the next two bins in \mathcal{L}_e , we begin to notice an asymmetry in efficiency for $d_0^{sign} > 0$ and $d_0^{sign} < 0$. The background from unidentified conversions is expected to preferentially populate $d_0^{sign} > 0$, and therefore we see a significantly higher efficiency as compared to $d_0^{sign} < 0$. The dilutions, though, seem to be consistent $\mathcal{D} \sim 5\text{-}10\%$ between the $d_0^{sign} > 0$ and $d_0^{sign} < 0$ samples. These two bins are occupied by a mix of electrons and hadrons, and hence exhibit small dilution. The last two bins in \mathcal{L}_e also show a larger asymmetry in efficiency for $d_0^{sign} > 0$ and $d_0^{sign} < 0$, caused by the presence of unidentified conversion electrons in the tag candidates. The dilution for the $d_0^{sign} > 0$ candidates is significantly worse than the $d_0^{sign} < 0$ candidates. This indicates that the background from unidentified conversions is significant, and the discriminating variable d_0^{sign} is effective in isolating higher purity electrons from semileptonic B meson decays.

To study the dependence of dilution on p_T^{rel} , the opposite side electron candidates in each likelihood bin are further split into a few bins in p_T^{rel} , in the $d_0^{sign} > 0$ and $d_0^{sign} < 0$ subsamples. The p_T^{rel} dependence of dilution for high likelihood candidates ($\mathcal{L}_e > 0.85$) for the $e\text{+SVT}$ data is shown in Figure 6.6. The dilution increases with increase in p_T^{rel} of the electron tag candidates, as the low p_T^{rel} region is dominated by electrons from sequential $b \rightarrow c \rightarrow \ell$ decays and fake electrons. Globally isolated

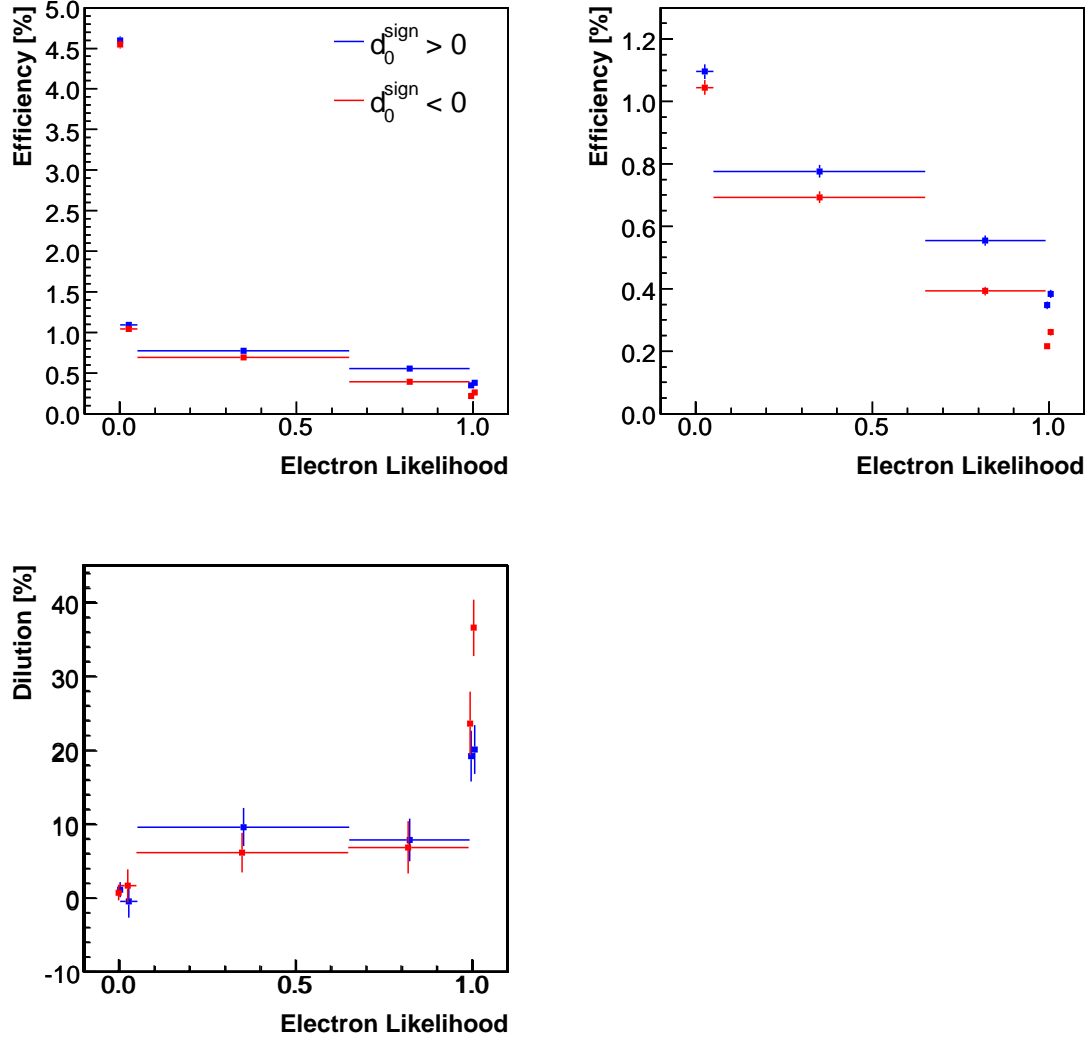


Figure 6.5: Tagging efficiencies (top) and dilutions (bottom) in likelihood bins separately for $d_0^{sign} > 0$ (blue) and $d_0^{sign} < 0$ (red) for the electron tag in the μ +SVT sample. The top-right plot is a zoomed-in version of the top-left plot. The bottom-left plot shows dilution in bins of \mathcal{L}_e .

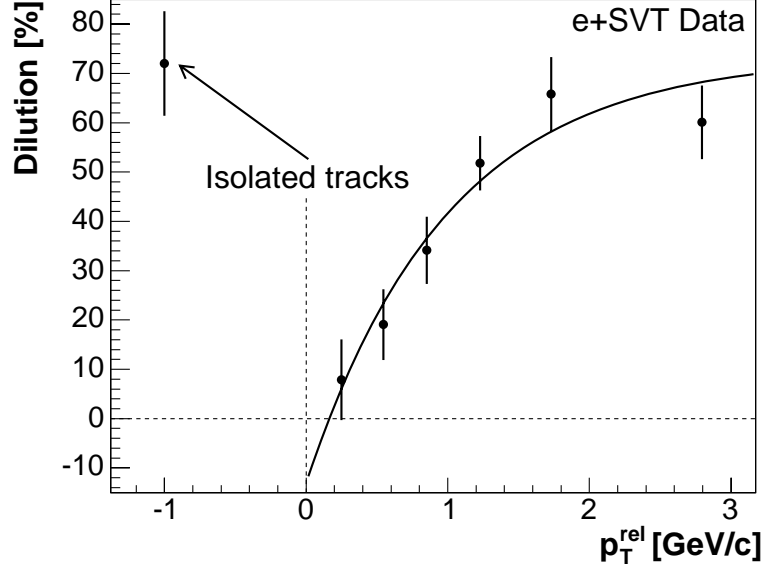


Figure 6.6: Variation of dilution of the electron tags with p_T^{rel} for $\mathcal{L}_e > 0.85$ in the e +SVT data.

electrons, defined above, are displayed at $p_T^{\text{rel}} = -1$ and exhibit very high dilution. The p_T^{rel} dependence of \mathcal{D} is described by the following parameterization:

$$\mathcal{D}(p_T^{\text{rel}}) = A \cdot \left(1 - e^{-p_T^{\text{rel}} + B}\right). \quad (6.5)$$

The form of Equation (6.5) is empirical and is found to describe the shape of \mathcal{D} as a function of p_T^{rel} well in Monte Carlo. The parameters A and B are determined for each of the six likelihood bins by fitting the p_T^{rel} distribution with the parameterization in Equation (6.5), separately in the $d_0^{\text{sign}} > 0$ and $d_0^{\text{sign}} < 0$ subsamples. In the fit describing neutral B mixing, these three pieces of information are used to predict more accurately the tag dilution on an event by event basis, hence maximizing the effectiveness of the electron tagging.

Summary

As established above, the dilution of the opposite side electron tagging algorithm depends strongly on three variables: d_0^{sign} , \mathcal{L}_e and p_T^{rel} . The signed impact parameter d_0^{sign} discriminates electron candidates from massive decays against electrons from unidentified conversions. The electron likelihood \mathcal{L}_e gives a probability estimate that an electron candidate is a real electron, while p_T^{rel} is related to the chance that the electron originates from a B meson. We exploit the dependence of dilution on these

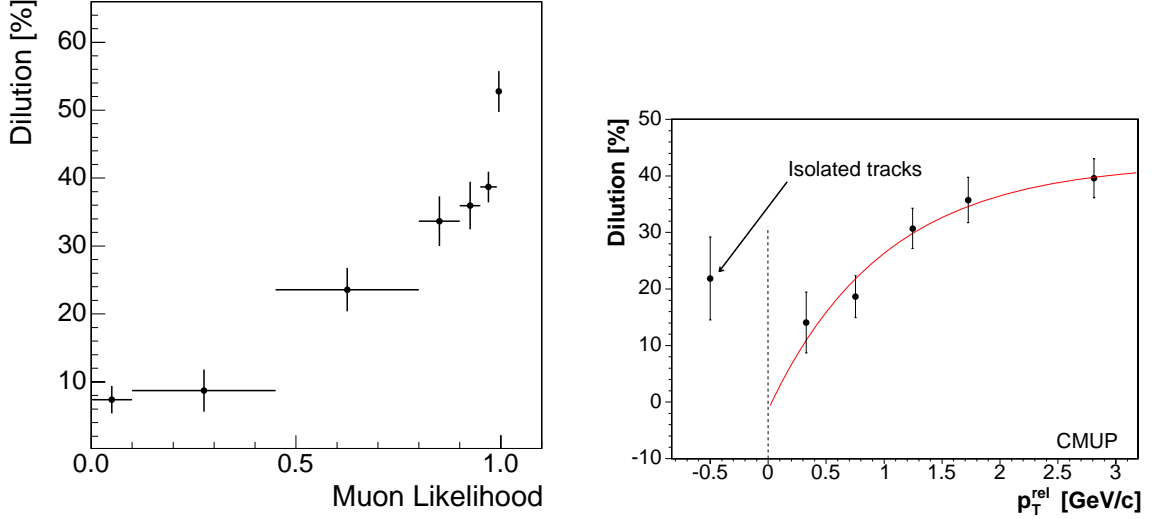


Figure 6.7: Variation of dilution of the muon tags with muon likelihood (left). Variation of raw dilution of the CMUP muon tags with respect to p_T^{rel} (right).

discriminating variables to improve the performance of the opposite side electron tagging algorithm.

6.1.4 Muon Tagging

To find a muon from a semileptonic decay of the opposite side B meson, we utilize the muon likelihood \mathcal{L}_μ described in Section 3.3. The opposite side muon tagging algorithm works in a manner similar to the electron tagging algorithm. Muons [60] are required to have $p_T > 1.5$ GeV/ c . Muon tags are distinguished using the muon likelihood \mathcal{L}_μ , p_T^{rel} and the muon subsystem (CMU, CMP, CMX, or IMU) that recorded them. The variation of the dilution for muon tags as a function of likelihood \mathcal{L}_μ is displayed on the left-hand side of Figure 6.7. Background from hadrons misidentified as leptons is expected to populate the low likelihood region (see Section 3.3), showing an increase of the dilution for larger lepton likelihood values.

The muon tags are then grouped into three likelihood bins: $0.0 \leq \mathcal{L}_\mu < 0.8$, $0.8 \leq \mathcal{L}_\mu < 0.95$, and $0.95 \leq \mathcal{L}_\mu \leq 1.0$. The variation of the dilution as a function of p_T^{rel} is exploited as can be seen on the right-hand side of Figure 6.7 for the CMUP muon tags. The behavior of dilution is similar to the electron tagging case; the dilution is lower for low p_T^{rel} because fake muons and muons from sequential semileptonic decays ($b \rightarrow c \rightarrow \mu^+$) tend to have relatively low p_T^{rel} values. Muons with large p_T^{rel} are more likely to come from primary B meson decays ($b \rightarrow \mu^-$). The p_T^{rel} dependence of \mathcal{D} is parameterized using the same functional form in Equation (6.5) for the electron tagging algorithm. The parameters A and B are measured separately for the three

likelihood bins and the different muon systems.

6.1.5 Jet Charge Tagging

The jet charge tagging algorithm is based on the momentum weighted sum of the charges of tracks associated with a jet. Jets on the opposite side are reconstructed using a cone-clustering algorithm (see Section 6.1.1) with a fixed cone radius $\Delta R = 1.5$. We use an ANN to enhance the probability of identifying jets from opposite side B mesons [66, 67]. In order to train the ANN, we generate a large PYTHIA Monte Carlo, as described in Section 2.9.

The ANN used to determine a probability for a jet to be a b jet is introduced in two stages. First, using a set of track based input variables, a track probability neural network (TrackNet) is constructed to estimate the likelihood of a given track to originate from a B meson decay. The input tracks to the TrackNet are reconstructed by attaching L00 hits, which significantly improves the resolution on the track impact parameter (see Section 4.3). The output of the TrackNet is the probability \mathcal{P}_{trk} that the track is a B decay product. The TrackNet output for each track in a jet together with other jet related kinematic input variables are then fed into a jet probability neural network (JetNet). The jet with the highest probability (\mathcal{P}_{nn}) is selected as the tagging jet.

The opposite side flavor is given by the sign of the jet charge Q_{jet} defined as

$$Q_{\text{jet}} = \frac{\sum_i Q_i \cdot p_T^i \cdot (1 + \mathcal{P}_{\text{trk}}^i)}{\sum_i p_T^i \cdot (1 + \mathcal{P}_{\text{trk}}^i)}, \quad (6.6)$$

where the index i runs over all the tracks in the selected jet, Q_i is the track charge, p_T^i is the track transverse momentum and $\mathcal{P}_{\text{trk}}^i$ the track probability. To better utilize the statistical power of the jet charge tagging algorithm, the tagged jets are further split into three types based on their quality and the predicted tag dilution is calculated separately for these types. Class 1 tags contain jets with a secondary vertex (SECVTX jets) that has a decay length significance, $L_{xy}/\sigma(L_{xy}) > 3$. These jets have the highest purity and best tagging dilution. Class 2 includes all jets not contained in Class 1 but with at least one track in the jet that has a track probability $\mathcal{P}_{\text{trk}} > 50\%$. Finally, Class 3 jets are the remainder of tagged jets not contained in Class 1 but with $\mathcal{P}_{\text{trk}} < 50\%$ for all tracks in the jet. The tagging classes are mutually exclusive. The left-hand side of Figure 6.8 shows the distribution of jet charge Q_{jet} for Class 1 jets in the e +SVT sample. The tags are shown separately for events tagged as B (open) and \bar{B} (solid). The separation of both distributions is a measure for the tagging power of the neural network based jet charge tagging algorithm to distinguish opposite side b jets versus \bar{b} jets.

To again maximize the effectiveness of the tagging method on an event by event basis, the dilution of the jet charge tags is expressed as a linear function in the quantity $|Q_{\text{jet}}| \cdot \mathcal{P}_{\text{nn}}$ as displayed in Figure 6.8 for Class 1 jets found in the e +SVT data. The

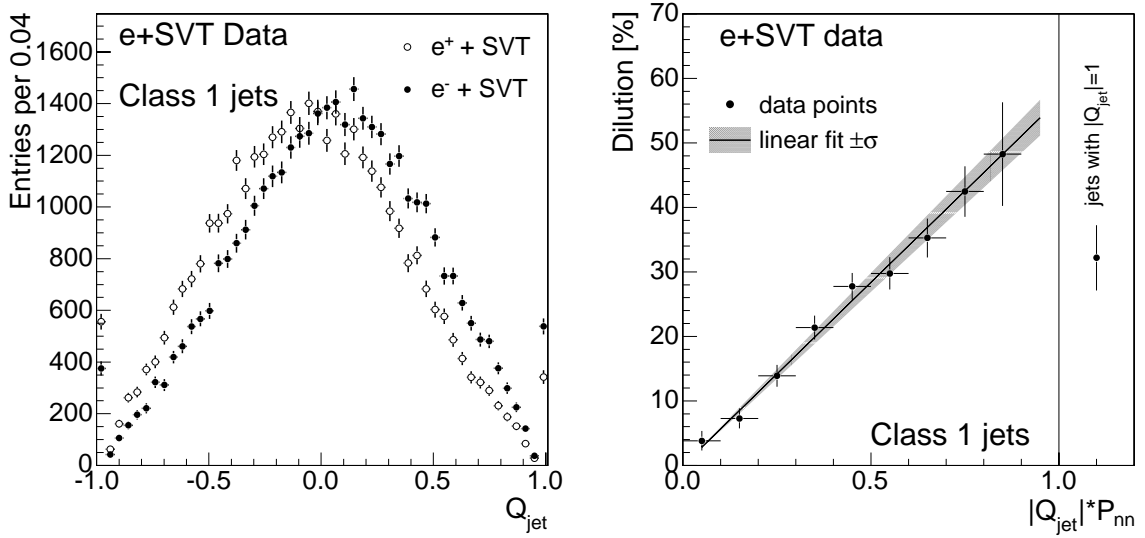


Figure 6.8: (left) Jet charge Q_{jet} for Class 1 jets for the $e+SVT$ sample shown separately for events tagged as B (open) and \bar{B} (solid). (right) Dilution dependence on $|Q_{jet}| \cdot P_{nn}$ for Class 1 tagged jets in the $e+SVT$ data.

dilution corresponding to jets with $Q_{jet} = 1$ is not included in the linear model but recorded separately. In the case of jet charge tagging, flavor misidentification can occur because the momentum weighted charge of the jet does not always reflect the true charge of the original b quark. In addition, the selected tagging jet may contain only a few or no tracks from the opposite side B meson decay.

6.1.6 Opposite Side Kaon Tagging

The opposite side kaon tagging algorithm (OSKT) examines opposite side jets to search for kaons from the decay sequence $b \rightarrow c \rightarrow s$ using particle identification. To discriminate against large backgrounds from pions, we combine information from both the TOF detector and dE/dx measured in the COT to construct a likelihood ratio for a track to agree with the kaon hypothesis. For a given track, the likelihood ratio $\mathcal{LR}(K)$ is constructed as

$$\mathcal{LR}(K) = \log \left(\frac{\mathcal{P}(K)}{0.2 \mathcal{P}(K) + 0.7 \mathcal{P}(\pi) + 0.1 \mathcal{P}(p)} \right), \quad (6.7)$$

where $\mathcal{P}(i) = \mathcal{P}_{TOF}(i) \cdot \mathcal{P}_{dE/dx}(i)$ for the particle hypotheses $i = K, \pi, p$.

We require $\mathcal{LR}(K) > -0.3$ to identify kaons candidates. After that, we must distinguish kaons from B meson decays from kaons originating from the primary $p\bar{p}$ interactions. We utilize the impact parameter significance ($|d_0|/\sigma(d_0)$) of the kaon tracks with respect to the primary vertex to find displaced kaon candidates. To reduce

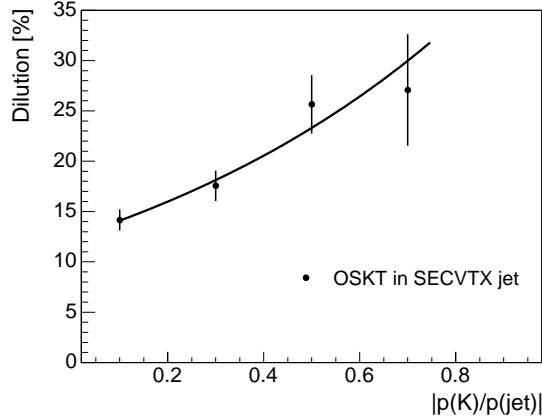


Figure 6.9: The dependence of OSKT tagging dilution on the fraction of jet momentum carried by the SECVTX jet. The exponential fit parameterization is overlaid.

the uncertainty on the impact parameter measurement, we again attach L00 hits to the candidate kaon tracks whenever possible.

In order to improve the effectiveness of the OSKT, we split the kaon tags into two categories. The first group contains tags where the identified kaon candidate is part of a SECVTX jet and the second category includes all remaining kaon tags. To specify the dilution of the opposite side kaon tags on an event by event basis, we parameterize the dependence of the dilution on a variable x with a functional form $\mathcal{D}(x) = e^{A+Bx}$, where A and B are parameters derived from the data. For the first category of kaon tags, the dilution is parameterized as a function of $p(K)/p(\text{jet})$, the fraction of jet momentum carried by the kaon candidate. This dependence is displayed in Figure 6.9 with the exponential fit parameterization overlaid. For the second category of kaon tags, where no secondary vertex is found in the jet, we parameterize the dilution as a function of the impact parameter significance $|d_0|/\sigma(d_0)$ of the kaon track. Within this second category, jets containing a single track (isolated tracks) have a higher purity. We therefore group them into a separate sub-category and parameterize their predicted dilution separately.

6.1.7 Combination of Opposite Side Tags

The performance of the various opposite side flavor tagging algorithms discussed above is studied with a pure sample of B mesons obtained from the ℓ +SVT data after background subtraction using the signed impact parameter of the SVT track (see Section 6.1). The obtained efficiencies ε , effective dilutions $\langle \mathcal{D} \rangle$ and effective tagging powers $\varepsilon \mathcal{D}^2$ are shown in Table 6.1 for individual tagging algorithms. The effective dilution is determined from $\langle \mathcal{D} \rangle \sim \sqrt{\varepsilon \mathcal{D}^2 / \varepsilon}$.

These opposite side tags are not mutually exclusive and multiple tagging decisions

Opposite Side Tag	Efficiency ε [%]	Effect. dilution $\langle \mathcal{D} \rangle$ [%]	Tagging Power $\varepsilon \mathcal{D}^2$ [%]
Muon	4.6 ± 0.0	34.7 ± 3.5	0.58 ± 0.02
Electron	3.2 ± 0.0	30.3 ± 0.7	0.29 ± 0.01
Jet Charge	95.5 ± 0.1	9.7 ± 0.2	0.90 ± 0.03
Kaon	18.1 ± 0.1	11.1 ± 0.9	0.23 ± 0.02
Combined NN	95.8 ± 0.1	12.7 ± 0.2	1.55 ± 0.04

Table 6.1: Tagging performances of the various opposite side tagging algorithms in the ℓ +SVT samples. The performance of a neural network based combined opposite side tagging technique is also shown. All errors given are statistical.

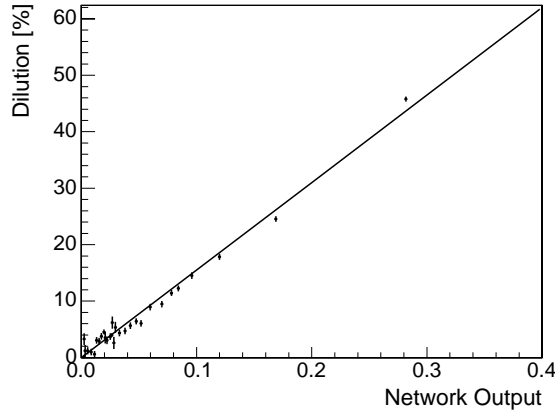


Figure 6.10: Linear dependence of tagging dilution from the combined OST versus output of the neural network.

are possible for a given event. We combine the information from all the opposite side tagging algorithms using an artificial neural network. Apart from the simplicity in dealing with a single opposite side tagging decision for each event, the purpose of combining the various opposite side tags into one tagging variable is to exploit correlations among the opposite side tags and improve the purity of a given tag. The neural network is again trained on the ℓ +SVT data. The set of input variables to the ANN includes the tagging decisions, predicted tag dilutions and various kinematic quantities characteristic for each individual opposite side tag. The dilution of the combined tag has a linear dependence on the output of the neural network as displayed in Figure 6.10. The performance of the combined opposite side tagging algorithm is also included in Table 6.1 as obtained from the ℓ +SVT data.

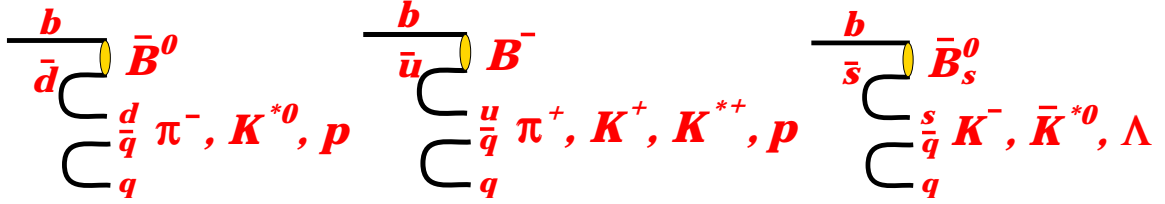


Figure 6.11: Charge correlation of B mesons with fragmentation tracks.

6.2 Same Side Flavor Tagging

Same side flavor tagging attempts to identify the production flavor of a B meson using information from the tracks close by. It has been suggested that the electric charge of particles produced in the vicinity of a B meson can be used to determine its initial flavor [68]. This can be visualized from a simple illustration shown in Figure 6.11. If a b quark hadronizes to form a \bar{B}^0 meson, the remaining d quark may combine with a \bar{u} quark to produce a π^- . Similarly, we can expect that a π^+ or a K^+ will be produced in conjugation with a B^- , and a K^- will be created in association with a \bar{B}_s^0 meson. Another source of correlated particles is the decays of the orbitally excited P wave B mesons (B^{**}) [68]. The goal of the same side tagging algorithm is to exploit the correlation between particles produced in the fragmentation of b quarks to B mesons.

Same side tagging benefits from several features in comparison with the opposite side tagging algorithms. The leading fragmentation track is likely to be within the η - φ space spanned by the same side B decay products. Thus, the detector acceptance for the fragmentation tracks close to the reconstructed B meson is expected to be large. Additionally, same side tagging does not suffer from limitations due to branching ratios of B mesons, like in the case of opposite side lepton tags. Additionally, the same side tagging algorithm searches for tagging tracks in close vicinity of the reconstructed B meson, and is thus robust against background tracks originating from the underlying event or multiple interactions. Finally, opposite side tagging methods suffer from the inevitable degradation of tagging power arising from opposite side neutral B mixing. Thus, the same side tagging algorithm is expected to perform significantly better than the opposite side tagging methods described in Section 6.1. Despite these advantages, the same side tagging has a limitation imposed by the very nature of its tagging approach. Since the same side tagging inherently depends on the fragmentation process of the reconstructed B mesons, its performance is dependent on the B mesons of interest. Therefore, we cannot calibrate the same side tagging dilution based on the decays of B^- and \bar{B}^0 mesons. The performance of the same side tagging algorithm can only be evaluated using \bar{B}_s^0 decays. However, this necessitates a precision measurement of the yet unknown B_s^0 - \bar{B}_s^0 oscillation frequency. Therefore, we utilize a PYTHIA Monte Carlo sample (see Section 2.9) to study the performance of the same side tagging algorithms for \bar{B}_s^0 mesons. The fragmentation mechanism

of the B mesons in the Monte Carlo sample is validated by comparing the variables sensitive to same side tagging. For this comparison, we utilize high statistics samples of B^- and \bar{B}^0 decays in data and Monte Carlo.

6.2.1 Selection of the Same Side Tag Candidates

The same side tagging algorithm first identifies potential tagging tracks by imposing the following preselection requirements:

- Tracks detected in the central detector region with $|\eta| < 1.0$ satisfying quality tracking criteria ensuring good momentum and impact parameter resolution; imposing the same COT and silicon hit requirements discussed in Section 4.3 and additionally, requiring track $p_T \geq 0.45$ GeV/ c to limit the asymmetry in the detection of positively and negatively charged tracks observed in the COT.
- Track should be close to the reconstructed B meson candidate. The η - φ separation between the B meson and track direction is required to be $\Delta R(B\text{-track}) \leq 0.7$. This also avoids overlap between the same side and the opposite side tagging, as the tracks chosen for all the opposite side tagging algorithms satisfy $\Delta R(B\text{-track}) > 0.7$ (see Section 6.1.2).
- Track impact parameter significance $|d_0|/\sigma_{d_0} < 4.0$ to select tracks originating from the primary $p\bar{p}$ interaction.
- Longitudinal distance between the track and the B candidate along the beam axis $|\Delta z(B\text{-track})| < 1.2$ cm, to reject particles from multiple interactions in the same event.
- Identified leptons using requirements on $\mathcal{L}_{e,\mu}$ are rejected (see Section 3.4).
- Track should not originate from photon conversions according to the criteria listed in Section 3.4.

Applying these requirements to the potential tagging tracks, we investigate the agreement between data and Monte Carlo in several kinematical and particle identification variables. Some examples of such variables are: the transverse momentum of the B meson, the tagging candidate multiplicity, ΔR between the tag candidate and the B meson, and the combined kaon identification variable $CLL(K)$ of tag candidates (see Section 3.1.3). Among the other variables of interest investigated for the same side tagging are the tag candidate kinematical variables like p_T , p_L^{rel} , and p_T^{rel} illustrated in Figure 6.12. In general, good agreement is observed between data and Monte Carlo in these variables for the fully reconstructed \bar{B}_s^0 candidates (see Section 9.1), and for the large control samples of the fully reconstructed \bar{B}^0 and B^- decays. Some examples of these comparisons are illustrated in Figure 6.13.

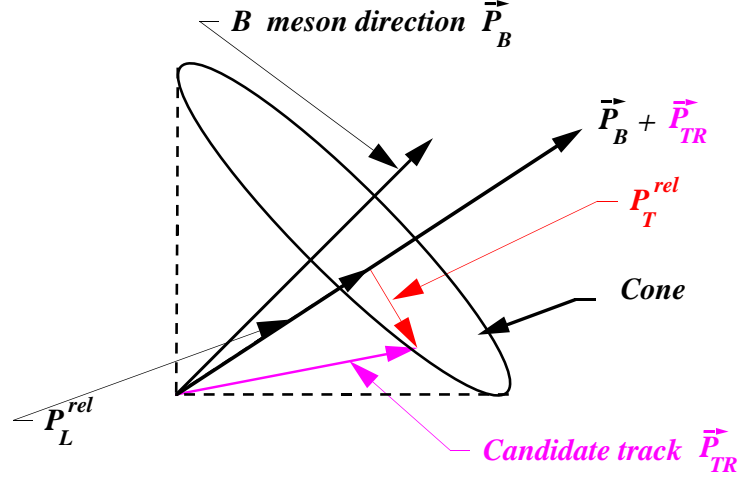


Figure 6.12: Illustration of the construction of the kinematic quantities p_T^{rel} and p_L^{rel} .

6.2.2 Same Side Tagging Algorithms

The tag candidates passing the preselection criteria described above, sometimes result in multiple eligible tag candidates differing in charge for the same \bar{B}_s^0 candidate. This occurs in $\sim 40\%$ of the tagged \bar{B}_s^0 candidates, with $\sim 65\%$ of those cases leading to tagging candidates with disagreeing tag decisions. To resolve this ambiguity and provide better tag decisions based on kinematical and particle identification characteristics of tag candidates, the following same side tagging algorithms are studied:

1. Among possible tag candidates, the maximum p_L^{rel} algorithm attempts to predict flavor information based on the charge of the tag candidate with the highest p_L^{rel} with respect to the B meson.
2. The maximum $CLL(K)$ algorithm selects tagging tracks based on the highest value of $CLL(K)$ among the tag candidates.
3. The neural net (NN) combination algorithm combines information from several kinematic variables of the tag candidates such as p_T^{rel} , p_L^{rel} , ΔR , p_T and $CLL(K)$. The NN is trained on \bar{B}_s^0 Monte Carlo samples assigning a “signal” hypothesis to the kaons with the correct charge correlation, and a “background” hypothesis to kaons, pions and protons with incorrect charge correlation. In this algorithm, the tagging track that maximizes the NN output is chosen as the same side tag.

In a manner similar to the opposite side tagging algorithms, the dependence of tagging dilution with respect to kinematical quantities is investigated to obtain a suitable measure of the same side tag purity. As described in Section 6.1 above, this method

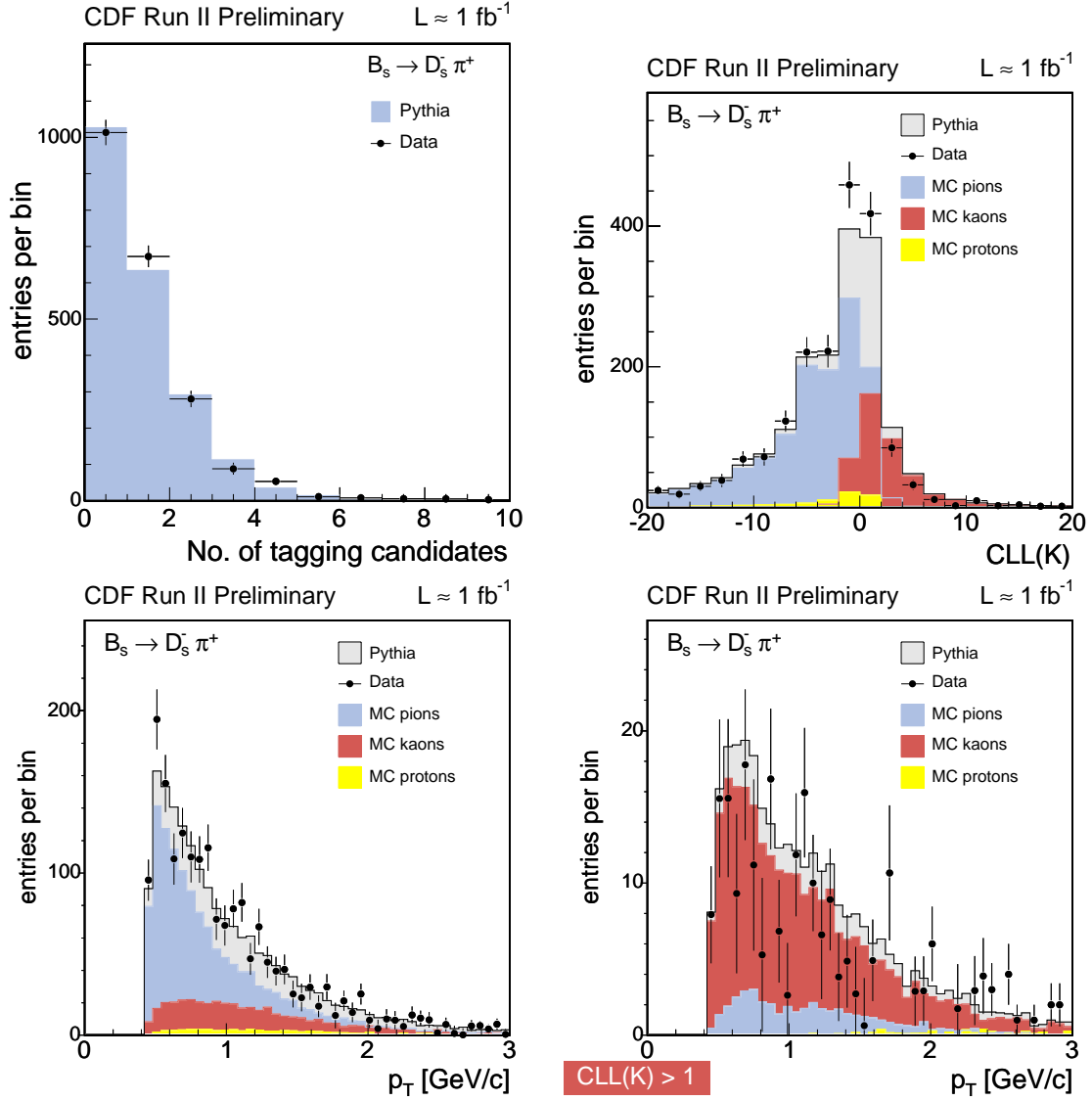


Figure 6.13: Data and Monte Carlo comparison for several SST related quantities for fully reconstructed \bar{B}_s^0 decays: Tagging track multiplicity (top-left), distribution of $CLL(K)$ for tag candidates (top-right), p_T of tag candidates (bottom-left) and p_T of tag candidates that are more likely to be kaons with $CLL(K) > 1$ (bottom-right). The Monte Carlo distributions are normalized to the same number of \bar{B}_s^0 mesons observed in data.

maximizes the effectiveness of the same side tagging algorithms on an event by event basis. The dilution of the maximum p_L^{rel} algorithm is parameterized as a function of the p_T of the tagging track. The dilution of the maximum $CLL(K)$ algorithm is parameterized with respect to the value of $CLL(K)$ of the tag and the dilution of the NN combination algorithm is parameterized as a function of the NN output variable nn with the following functional form:

$$\mathcal{D}(nn) = a + b \cdot nn + c \cdot nn^2 + d \cdot nn^3, \quad (6.8)$$

where a , b , c , and d are constants determined from data. The same side tags are divided into two categories as follows:

- Agreeing Cases: The tagged \bar{B}_s^0 candidates for the case when there is a unique tagging track or when all the tagging candidates have the same charge.
- Disagreeing Cases: The tagged \bar{B}_s^0 candidates when there are more than one tagging tracks and the charges of some of the tagging tracks are different.

The dilution for the agreeing and disagreeing subsamples is parameterized separately, as the disagreeing cases in general have a lower dilution. The dependence of dilution for the three algorithms and their parameterizations are displayed in Figure 6.14, for the cases when there is a unique tag candidate. The dilution of the same side tag increases as a function of tag p_T for the maximum p_L^{rel} algorithm. The dilution for the NN combination of the same side tagging algorithm also increases with the increase in the value of the NN output for the tag. The dilution for the maximum $CLL(K)$ algorithm first increases, and then decreases at high values of $CLL(K)$. High values of $CLL(K)$ are achieved at low track momentum for real kaons, when both dE/dx and TOF provide large separation between kaons and pions. Thus, the tag candidates occupying the high $CLL(K)$ region are real low momentum kaons, and consequently their purity suffers due to contamination of (background) kaons from the underlying event.

6.2.3 Performance of the Same Side Tagging Algorithms

The performance of the three same side tagging algorithms has been evaluated for the various decay modes in B^- , \bar{B}^0 and \bar{B}_s^0 decays. The performances are summarized in Tables 6.2 to 6.4. The agreement between the Monte Carlo predictions and data is good, and demonstrates that the same side tagging is reproduced well by the Monte Carlo simulation across all the B meson species. Since we rely on the Monte Carlo predictions for the same side tag dilution in the case of \bar{B}_s^0 mesons, systematic uncertainties on the Monte Carlo dilution predictions need to be evaluated before applying the predictions on data.

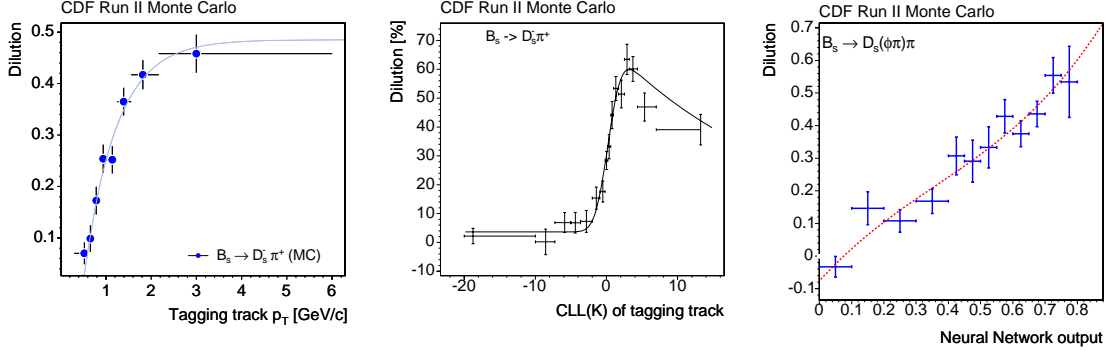


Figure 6.14: Dilution of the maximum p_L^{rel} algorithm as a function of tagging track p_T (left). Dilution for the maximum $CLL(K)$ algorithm as a function of the kaon identification variable $CLL(K)$ (middle). Dilution of the NN combination algorithm as a function of NN output variable (right). \bar{B}_s^0 meson candidates with a unique same side tag are displayed.

6.2.4 Systematic Uncertainties in the Same Side Tag Dilution

The sources of systematic errors on the dilution of the same side tag are as follows:

Level of Data-Monte Carlo Agreement in B^- and \bar{B}^0 decays

The data and Monte Carlo agree well in terms of the tagging dilution of the same side tag within the uncertainties for the B^- and \bar{B}^0 decay modes. However, the precision of the comparison is limited by the uncertainties in the measured dilutions for data and Monte Carlo. Therefore, the combined uncertainties from the B^- and \bar{B}^0 modes have been taken as a systematic error on the Monte Carlo predictions.

Particle Content Around the B Mesons

An independent measurement of the fraction of various particle species around B mesons in semileptonic decays has been accomplished in the ℓ +SVT data [69]. This study verifies that the kaon fractions around the B mesons are reproduced well in the Monte Carlo samples for B^- and \bar{B}^0 decays, but indicates a slightly lower kaon fraction in \bar{B}_s^0 data as compared to the Monte Carlo. The Monte Carlo sample is re-weighted to accommodate the lower kaon fraction around the \bar{B}_s^0 meson and the variation in the dilution is taken as a systematic error.

Other Sources of Systematic Uncertainties

Apart from the largest sources of systematic errors described above, additional sources have been considered. These include the ability to accurately model the B meson frag-

	[%]	$B^+ \rightarrow J/\psi K^+$	$B^0 \rightarrow J/\psi K^{*0}$	$B_s^0 \rightarrow J/\psi \phi$
MC	ϵ	55.9 ± 0.1	56.6 ± 0.1	52.1 ± 0.3
	$\langle \mathcal{D} \rangle$	29.0 ± 0.3	17.2 ± 0.4	22.8 ± 0.7
data (1 fb ⁻¹)	ϵ	58.2 ± 0.3	57.1 ± 0.3	49.2 ± 1.3
	$\langle \mathcal{D} \rangle$	29.7 ± 0.7	17.8 ± 1.6	—

Table 6.2: Performance of the maximum p_L^{rel} algorithm in data and Monte Carlo (statistical uncertainties only).

	[%]	$B^+ \rightarrow J/\psi K^+$	$B^0 \rightarrow J/\psi K^{*0}$	$B_s^0 \rightarrow J/\psi \phi$
MC	ϵ	55.9 ± 0.1	56.6 ± 0.1	52.1 ± 0.3
	$\langle \mathcal{D} \rangle$	27.3 ± 0.2	17.1 ± 0.4	27.9 ± 0.7
data (1 fb ⁻¹)	ϵ	58.4 ± 0.3	57.3 ± 0.3	49.4 ± 1.3
	$\langle \mathcal{D} \rangle$	27.4 ± 0.8	16.9 ± 1.7	—

Table 6.3: Performance of the maximum $CLL(K)$ algorithm in data and Monte Carlo (statistical uncertainties only).

	[%]	$B^+ \rightarrow J/\psi K^+$	$B^0 \rightarrow J/\psi K^{*0}$	$B_s^0 \rightarrow J/\psi \phi$
MC	ϵ	55.9 ± 0.1	56.6 ± 0.1	52.1 ± 0.3
	$\langle \mathcal{D} \rangle$	26.8 ± 0.2	16.1 ± 0.6	29.2 ± 0.7
data (1 fb ⁻¹)	ϵ	58.2 ± 0.3	57.2 ± 0.3	49.3 ± 1.3
	$\langle \mathcal{D} \rangle$	26.4 ± 0.8	15.2 ± 1.7	—

Table 6.4: Performance of the NN combination algorithm in data and Monte Carlo (statistical uncertainties only).

mentation, multiple interactions in the same event, B meson production mechanisms, the B^{**} rate and the description of the particle identification quantities (dE/dx and TOF) in the Monte Carlo simulation. The effect of those variations are small as compared to the systematic errors from the data-Monte Carlo agreement and the particle content around the B_s^0 mesons.

The agreement between data and Monte Carlo is displayed in Figure 6.15, after including all the systematic uncertainties. They agree well within the variations introduced by the systematic errors.

6.2.5 Summary of Same Side Kaon Tag Performance

The same side tagging performance evaluated on a sample of fully reconstructed $B_s^0 \rightarrow D_s^- \pi^+$ data, using the dilution parameterization derived from Monte Carlo, is

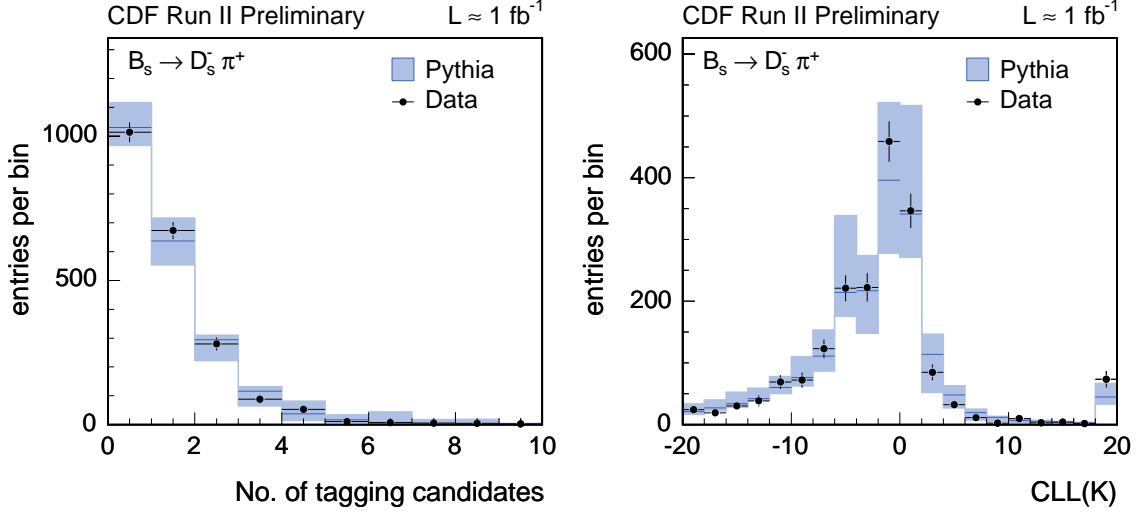


Figure 6.15: Comparisons between data and Monte Carlo for the tagging track multiplicity around B_s^0 mesons (left) and $CLL(K)$ values for tagging tracks (right) for same side tag candidates. Blue bands indicate the range of variations due to systematic uncertainties assigned.

as follows:

- Maximum p_L^{rel} algorithm: $\varepsilon\mathcal{D}^2 = 2.8_{-0.8}^{+0.6} \%$;
- Maximum $CLL(K)$ algorithm: $\varepsilon\mathcal{D}^2 = 2.9 - 4.0_{-1.2}^{+0.9} \%$;
- NN combination algorithm: $\varepsilon\mathcal{D}^2 = 3.1 - 4.3_{-1.4}^{+1.0} \%$.

The variations in the performance of the maximum $CLL(K)$ and NN combined algorithms arise from the changes in the performances of the Time of Flight detector and the dE/dx in the three data-taking periods (see Section 5.2). The performance of the NN combined algorithm is the best among the three algorithms, and consequently we utilize it for this measurement.

The dilution prediction from Monte Carlo is based on hadronic $B_s^0 \rightarrow D_s^- \pi^+$ decays. The semileptonic B decays collected via the two-track and the ℓ +SVT triggers, in general, have a higher parent B transverse momentum as compared to the hadronic decays. This improves the dilution of the same side kaon tagging due to slightly harder fragmentation tracks. Therefore, the event by event dilution prediction for the NN same side kaon tag is scaled by a correction factor of 1.06 for use in the current measurement utilizing semileptonic B_s^0 meson decays.

6.3 Tagging Application

We permit each B meson candidate to be tagged by as many as two flavor tags: one opposite side and one same side [70]. For candidates that have both the same side and the combined opposite side tags, the tagging decisions from both the tags are combined assuming that they are independent as described in Section 8.1.

Chapter 7

Calibration of Opposite Side Flavor Tagging and Measurement of B^0 - \bar{B}^0 Oscillations

The initial calibration of the event by event dilution for the combined opposite side tag (see Section 6.1.7) is performed on inclusive semileptonic B meson events reconstructed in the ℓ +SVT trigger data. However, the kinematical requirements imposed in the ℓ +SVT triggers are different from the ones required by the two-track trigger for online selection of events (see Section 4.2). In addition, the average B meson momentum for semileptonic B decays reconstructed in the ℓ +SVT trigger data is higher than for semileptonic B candidates found in the two-track trigger data. These differences can influence the dilution prediction of the combined OST in the two-track trigger data. Furthermore, the B_s^0 decays collected in the ℓ +SVT data, that are utilized for the measurement of B_s^0 - \bar{B}_s^0 oscillations, are a small subset of the sample used for OST studies. Therefore, we re-calibrate the combined opposite side tagging algorithm with the help of the observed asymmetry in the B^+ decays and a measurement of the B^0 oscillation frequency Δm_d . The B meson candidates used for this purpose are collected by the same trigger paths as the B_s^0 sample and are, hence, kinematically similar to the B_s^0 decays used for the current measurement.

In the current chapter, we discuss the inclusion of flavor tagging information in the description of the B^+ and B^0 meson samples. In Chapter 5, we have derived the fractions and PDFs describing the signal and backgrounds in D mass, $m_{\ell D}$, and ct^* spaces. As described in the following, this arms us to introduce tagging information in the overall likelihood describing the B samples, and allows the simultaneous extraction of Δm_d and tagging calibration parameters.

7.1 Time Evolution PDF for B^0 - \bar{B}^0 Mixing

As described in Section 1.2, the difference in mass Δm_d between the two mass eigenstates of the B^0 meson leads to a phase difference between their wave functions. Utilizing Equation (1.40) for the B^0 - \bar{B}^0 system under the approximation $\Delta\Gamma_d = 0$, the probability of flavor change of an initially pure B^0 state at decay time t is given by

$$\mathcal{P}_{\text{mix}}^{B^0}(t) = \mathcal{P}_{B^0 \rightarrow \bar{B}^0}(t) = \frac{\Gamma_d e^{-\Gamma_d t}}{2} \theta(t) [1 - \cos(\Delta m_d t)] , \quad (7.1)$$

while the probability that it decays as B^0 is given by Equation (1.41)

$$\begin{aligned} \mathcal{P}_{\text{unmix}}^{B^0}(t) = \mathcal{P}_{B^0 \rightarrow B^0}(t) &= 1 - \mathcal{P}_{\text{mix}}(t) \\ &= \frac{\Gamma_d e^{-\Gamma_d t}}{2} \theta(t) [1 + \cos(\Delta m_d t)] , \end{aligned} \quad (7.2)$$

where $\theta(t)$ is the step function. In performing a study of oscillations using semileptonic decays $B \rightarrow \ell DX$, the flavor of the B^0 meson at decay is inferred from the charge of the lepton, while the flavor at production is determined by the combined OST (see Section 6.1.7). The combined OST provides a tagging decision T , with $T = +1$ signaling that the charge of the lepton from the B decay and the tag decision (corresponding to the B flavor at production) are the same, while $T = -1$ indicating vice-versa. If the combined OST always furnishes a correct decision ($\mathcal{D} = 1$) with 100% efficiency (see Chapter 6), then the time evolution of the B^0 system can be written as:

$$\mathcal{P}_T^{B^0}(t) = \frac{\Gamma_d e^{-\Gamma_d t}}{2} \theta(t) [1 + T \cos(\Delta m_d t)] , \quad (7.3)$$

where $T = -1$ signals flavor change or mixing, while $T = +1$ indicates no mixing. The effect of an imperfect combined OST with mistagging probability p_W can be taken into account as

$$\begin{aligned} \mathcal{P}_T^{B^0}(t) &= \frac{\Gamma_d e^{-\Gamma_d t}}{2} \theta(t) [(1 - p_W) \cdot \{1 + T \cos(\Delta m_d t)\} + p_W \cdot \{1 - T \cos(\Delta m_d t)\}] \\ &= \frac{\Gamma_d e^{-\Gamma_d t}}{2} \theta(t) [1 + (1 - 2p_W) \cdot T \cos(\Delta m_d t)] . \end{aligned} \quad (7.4)$$

Using the definition of dilution $\mathcal{D} = 1 - 2p_W$ (see Chapter 6), we obtain

$$\mathcal{P}_T^{B^0}(t|\mathcal{D}) = \frac{\Gamma_d e^{-\Gamma_d t}}{2} \theta(t) [1 + \mathcal{D} \cdot T \cos(\Delta m_d t)] . \quad (7.5)$$

Apart from mistagging, the combined OST may not always furnish a tagging decision ($T = 0$). If ε is the tagging efficiency, we can express Equation (7.5) as

$$\mathcal{P}_T^{B^0}(t|\mathcal{D}) = \begin{cases} \frac{\varepsilon}{2} \cdot \Gamma_d e^{-\Gamma_d t} \theta(t) [1 + \mathcal{D} \cdot T \cos(\Delta m_d t)] & \text{for } T = \pm 1 , \\ (1 - \varepsilon) \cdot \Gamma_d e^{-\Gamma_d t} \theta(t) & \text{for } T = 0 . \end{cases} \quad (7.6)$$

The decay time t is reconstructed using the formalism developed in Section 5.3.5. The incomplete reconstruction of semileptonic decays is corrected for utilizing the κ -factor distribution $\mathcal{H}(\kappa)$, the reconstruction and trigger effects are taken into account using $\xi(t^*)$, and the pseudo-proper decay time resolution is included via a Gaussian distribution $G(t' - t^*, \sigma_{t^*})$. Therefore, the PDF for the time evolution describing the B^0 meson system in Equation (7.6) can be re-written for semileptonic decays in terms of the observable time t^* as follows:

tagged ($T = \pm 1$):

$$\begin{aligned} \mathcal{P}_T^{B^0}(t^* | \mathcal{D}, \sigma_{t^*}) &= \frac{\varepsilon}{N^{B^0}} \frac{\Gamma_d}{2} e^{-\Gamma_d \kappa t^*} \theta(\kappa t^*) [1 + \mathcal{D}T \cos(\Delta m_d \kappa t^*)] \\ &\quad \otimes_{\kappa} \mathcal{H}^{B^0}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^0}(t^*), \end{aligned} \quad (7.7)$$

untagged ($T = 0$):

$$\mathcal{P}_T^{B^0}(t^* | \mathcal{D}, \sigma_{t^*}) = \frac{1 - \varepsilon}{N^{B^0}} \Gamma_d e^{-\Gamma_d \kappa t^*} \theta(\kappa t^*) \otimes_{\kappa} \mathcal{H}^{B^0}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^0}(t^*). \quad (7.8)$$

Here, N^{B^0} is the normalization given by Equation (5.19). Thus, the non-tagged component is simply described by a lifetime PDF as obtained in Equation (5.18). Another quantity of interest in the mixing measurements is the asymmetry $\mathcal{A}(t)$ defined as

$$\mathcal{A}(t) = \frac{N_+(t) - N_-(t)}{N_+(t) + N_-(t)}, \quad (7.9)$$

where $N_+(t)$ ($N_-(t)$) is the number of candidates which are flagged as unmixed (mixed) as a function of proper decay time t .

7.2 Time Evolution PDF for B^+

The B^+ meson does not undergo flavor oscillations. Hence, its description is obtained by setting $\Delta m_d \rightarrow \Delta m_u = 0$ in Equations (7.6) and (7.7). Therefore, the time evolution of the B^+ mesons is described by

tagged ($T = \pm 1$):

$$\begin{aligned} \mathcal{P}_T^{B^+}(t^* | \mathcal{D}, \sigma_{t^*}) &= \frac{\varepsilon}{N^{B^+}} \frac{\Gamma_u}{2} e^{-\Gamma_u \kappa t^*} \theta(\kappa t^*) [1 + \mathcal{D}T] \otimes_{\kappa} \mathcal{H}^{B^+}(\kappa) \\ &\quad \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^+}(t^*), \end{aligned} \quad (7.10)$$

untagged ($T = 0$):

$$\mathcal{P}_T^{B^+}(t^*|\mathcal{D}, \sigma_{t^*}) = \frac{1-\varepsilon}{N^{B^+}} \Gamma_u e^{-\Gamma_u \kappa t^*} \theta(\kappa t^*) \otimes_{\kappa} \mathcal{H}^{B^+}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^+}(t^*). \quad (7.11)$$

Separate κ -factor distributions $\mathcal{H}^{B^+}(\kappa)$ and efficiency functions $\xi^{B^+}(t^*)$ are obtained from B^+ Monte Carlo simulations. The efficiencies and dilutions for the combined OST are the same for signal B^+ decays as in the case of signal B^0 decays, since the performance of an opposite side tag does not distinguish between the various B meson species.

7.3 Time Evolution PDF for Backgrounds

There are three types of backgrounds (see Section 5.4) for the asymmetry analysis in B^+ decays and for the mixing measurement in the case of B^0 decays. These are classified into two types based on whether they oscillate or not:

1. Oscillating backgrounds: Physics background originating from B^0 and B_s^0 decays belong to this category.
2. Non-oscillating backgrounds: Physics background arising from B^+ decays, as well as combinatorial and fake lepton backgrounds are associated with this category.

We briefly discuss the treatment of these backgrounds in the following.

7.3.1 Oscillating Backgrounds

Physics Backgrounds from B^0 Decays

The time evolution PDF for describing the physics background from B^0 decays follows the same description as obtained for the B^0 - \bar{B}^0 mixing term discussed in Section 7.1. Therefore, the PDF is given by **tagged** ($T = \pm 1$):

$$\begin{aligned} \mathcal{P}_T^{B^0 \text{ phys}}(t^*|\mathcal{D}, \sigma_{t^*}) &= \frac{\varepsilon}{N^{B^0 \text{ phys}}} \frac{\Gamma_d}{2} e^{-\Gamma_d \kappa t^*} \theta(\kappa t^*) [1 + \mathcal{D}T \cos(\Delta m_d \kappa t^*)] \\ &\otimes_{\kappa} \mathcal{H}^{B^0 \text{ phys}}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^0 \text{ phys}}(t^*), \end{aligned} \quad (7.12)$$

untagged ($T = 0$) :

$$\begin{aligned} \mathcal{P}_T^{B^0 \text{ phys}}(t^*|\mathcal{D}, \sigma_{t^*}) &= \frac{1-\varepsilon}{N^{B^0 \text{ phys}}} \Gamma_d e^{-\Gamma_d \kappa t^*} \theta(\kappa t^*) \otimes_{\kappa} \mathcal{H}^{B^0 \text{ phys}}(\kappa) \\ &\otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^0 \text{ phys}}(t^*), \end{aligned} \quad (7.13)$$

where we have emphasized that the κ -factor distribution $\mathcal{H}^{B^0 \text{ phys}}(\kappa)$, and the efficiency function $\xi^{B^0 \text{ phys}}(t^*)$ are derived from the B^0 physics background Monte Carlo. The efficiencies and dilutions for the combined OST for the physics background are taken to be the same as the signal B^+ and B^0 decays.

Physics Backgrounds from B_s^0 Decays

For the physics background originating from B_s^0 decays, the fast oscillations result into a fully mixed sample i.e. we assume that at a given time $t > 0$, an initially pure B_s^0 state consists of an equal admixture of B_s^0 and \bar{B}_s^0 states. Therefore,

$$\mathcal{P}_{\text{mix}}^{B_s^0}(t) = \mathcal{P}_{\text{unmix}}^{B_s^0}(t) = \frac{\Gamma_s e^{-\Gamma_s t}}{2} \theta(t). \quad (7.14)$$

Thus, the PDF describing the time evolution of physics background from B_s^0 decays is given by

tagged and untagged ($T = \pm 1$ and $T = 0$):

$$\begin{aligned} \mathcal{P}_T^{B_s^0 \text{ phys}}(t^* | \mathcal{D}, \sigma_{t^*}) &= \frac{1}{N^{B_s^0 \text{ phys}}} \frac{\Gamma_s}{2} e^{-\Gamma_s \kappa t^*} \theta(\kappa t^*) \otimes_{\kappa} \mathcal{H}^{B_s^0 \text{ phys}}(\kappa) \\ &\quad \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B_s^0 \text{ phys}}(t^*). \end{aligned} \quad (7.15)$$

The time evolution of physics backgrounds from B_s^0 decays is thus independent of the flavor tag.

7.3.2 Non-oscillating Backgrounds

Physics Backgrounds from B^+ Decays

The physics background originating from B^+ decays are treated the same way as the B^+ signal. Therefore, we have

tagged ($T = \pm 1$):

$$\begin{aligned} \mathcal{P}_T^{B^+ \text{ phys}}(t^* | \mathcal{D}, \sigma_{t^*}) &= \frac{\varepsilon}{N} \frac{\Gamma_u}{2} e^{-\Gamma_u \kappa t^*} \theta(\kappa t^*) [1 + \mathcal{D}T] \otimes_{\kappa} \mathcal{H}^{B^+ \text{ phys}}(\kappa) \\ &\quad \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^+ \text{ phys}}(t^*), \end{aligned} \quad (7.16)$$

untagged ($T = 0$):

$$\begin{aligned} \mathcal{P}_T^{B^+}(t^* | \mathcal{D}, \sigma_{t^*}) &= \frac{1 - \varepsilon}{N} \Gamma_u e^{-\Gamma_u \kappa t^*} \theta(\kappa t^*) \otimes_{\kappa} \mathcal{H}^{B^+ \text{ phys}}(\kappa) \\ &\quad \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^+ \text{ phys}}(t^*), \end{aligned} \quad (7.17)$$

where we have again emphasized that $\mathcal{H}^{B^0 \text{ phys}}(\kappa)$, and $\xi^{B^0 \text{ phys}}(t^*)$ are derived using Monte Carlo appropriate for the B^+ physics background.

Combinatorial Background

The combinatorial background is assumed to be non-oscillating, and is hence assigned a time-averaged asymmetry $\mathcal{D}_{avg}^{\text{comb}}$ and the corresponding combined OST efficiency $\varepsilon^{\text{comb}}$. The time evolution behavior for the combinatorial background is derived by utilizing the fact that if the OST decision signals a mixed (unmixed) state, it does so with the probability p_W^{comb} ($1 - p_W^{\text{comb}}$). Thus, we obtain

tagged ($T = \pm 1$):

$$\mathcal{P}_T^{\text{comb}}(t^* | \mathcal{D}, \sigma_{t^*}) = \varepsilon^{\text{comb}} \frac{(1 + T \cdot \mathcal{D}_{avg}^{\text{comb}})}{2} \cdot \mathcal{P}^{\text{comb}}(t^*, \sigma_{t^*}), \quad (7.18)$$

untagged ($T = 0$):

$$\mathcal{P}_T^{\text{comb}}(t^* | \mathcal{D}, \sigma_{t^*}) = (1 - \varepsilon^{\text{comb}}) \cdot \mathcal{P}^{\text{comb}}(t^*, \sigma_{t^*}), \quad (7.19)$$

where $\mathcal{P}^{\text{comb}}(t^*, \sigma_{t^*})$ is the joint PDF describing the combinatorial background in the t^* and σ_{t^*} spaces, obtained in Section 5.3.6. The efficiency and dilution for the combinatorial background are derived using the candidates from the D mass sidebands. The obtained OST efficiency for the combinatorial background is very similar to that for the signal.

Fake Lepton Background

The fake lepton background is assumed to be non-oscillating, with an OST tagging efficiency which is the same as the B^0 signal. The time-averaged asymmetry of the fake lepton background is assumed to be zero ($\mathcal{D}_{avg}^{\text{fake}} = 0$). These assumptions have been cross-checked with the fake lepton sample, and found to yield consistent results. However, we have studied systematic errors as a consequence of this assumption in Section 7.6.1. The time evolution for the fake lepton background is given by,

tagged ($T = \pm 1$):

$$\mathcal{P}_T^{\text{fake}}(t^* | \mathcal{D}, \sigma_{t^*}) = \frac{\varepsilon}{2} \cdot \mathcal{P}^{\text{fake}}(t^*, \sigma_{t^*}), \quad (7.20)$$

untagged ($T = 0$):

$$\mathcal{P}_T^{\text{fake}}(t^* | \mathcal{D}, \sigma_{t^*}) = (1 - \varepsilon) \cdot \mathcal{P}^{\text{fake}}(t^*, \sigma_{t^*}), \quad (7.21)$$

where $\mathcal{P}^{\text{fake}}(t^*, \sigma_{t^*})$ is the joint PDF describing the fake lepton background in the t^* and σ_{t^*} spaces as determined in Section 5.3.6.

7.4 Determination of Inputs for the OST Calibration and B^0 - \bar{B}^0 Mixing

As described in Section 5.3, we utilize an unbinned maximum likelihood fit to extract parameters of interest from the data. In this fit, decay candidates enter the likelihood with different weights, based on their characterization in D mass, $m_{\ell D}$, ct^* , σ_{ct^*} , and the tagging space characterized by the dilution \mathcal{D} . Consequently, candidates populating regions in D mass, $m_{\ell D}$, and ct^* spaces with higher probabilities to be signal candidates, together with well measured σ_{ct^*} and high \mathcal{D} affect the output of the fit more significantly as compared to candidates lacking these features.

The overall likelihood is given by Equation (5.3), and written as

$$\mathcal{L} = \prod_n L_n = \prod_n \sum_i f_i \cdot \mathcal{P}^i(m, m_{\ell D}, ct^*, \sigma_{ct^*}, T, \mathcal{D}), \quad (7.22)$$

where the index i (n) runs over the various components (candidates) of a given sample, L_n is the candidate likelihood for the n^{th} candidate, and \mathcal{P}^i is the joint PDF in D mass, $m_{\ell D}$, ct^* , σ_{ct^*} and tagging spaces for the i^{th} component. The candidate likelihood L is given by,

$$L(m, m_{\ell D}, ct^*, \sigma_{ct^*}, T, \mathcal{D}) = \sum_i f_i \cdot \mathcal{P}^i(m, m_{\ell D}, ct^*, \sigma_{ct^*}, T, \mathcal{D}). \quad (7.23)$$

Using Equation (5.25), the joint PDF \mathcal{P}^i can be written as

$$\begin{aligned} \mathcal{P}^i(m, m_{\ell D}, ct^*, \sigma_{ct^*}, T, \mathcal{D}) &= \mathcal{P}^i(m) \cdot \mathcal{P}^i(m_{\ell D}) \cdot \mathcal{P}_T^i(ct^* | \mathcal{D}, \sigma_{ct^*}) \\ &\quad \cdot \mathcal{P}^i(\mathcal{D}) \cdot \mathcal{P}^i(\sigma_{ct^*}), \end{aligned} \quad (7.24)$$

where $\mathcal{P}^i(\mathcal{D})$ and $\mathcal{P}^i(\sigma_{ct^*})$ are the unconditional PDFs for σ_{ct^*} and \mathcal{D} of the i^{th} component. In terms of the notation that we have been using, therefore, the candidate likelihood can be written as

$$\begin{aligned} L(m, m_{\ell D}, ct^*, \sigma_{ct^*}, T, \mathcal{D}) &= \sum_i f_i \mathcal{P}^i(m) \cdot \mathcal{P}^i(m_{\ell D}) \cdot \mathcal{P}_T^i(ct^* | \mathcal{D}, \sigma_{ct^*}) \\ &\quad \cdot \mathcal{P}^i(\mathcal{D}) \cdot \mathcal{P}^i(\sigma_{ct^*}). \end{aligned} \quad (7.25)$$

We are interested in a measurement of Δm_d together with a calibration of the combined OST. To characterize the tagging performance, we introduce a scaling factor S_D^{OST} that multiplies the event by event tagging dilution, as predicted by the combined OST. The scaling factor S_D^{OST} is assumed to be the same for B^+ and B^0 decays. Therefore,

$$\mathcal{D} \rightarrow S_D^{OST} \cdot \mathcal{D}. \quad (7.26)$$

A value of $S_D^{OST} < 1$ thus indicates an overestimation of the predicted dilution of the tag, while $S_D^{OST} > 1$ indicates an undervaluation of the predicted OST dilution.

A simultaneous minimization is performed to extract the values of Δm_d and S_D^{OST} . The contribution to the Δm_d measurement is due to the argument of the cosine modulation term in the description of the time evolution of B^0 mesons, while both B^+ and B^0 decays contribute to the determination of S_D^{OST} . The signal part of the joint PDF in Equation (5.25) can thus be written as:

$$\begin{aligned} f_{\text{signal}} \cdot \mathcal{P}_T^{\text{signal}}(ct^* | S_D^{OST} \cdot \mathcal{D}, \sigma_{ct^*}, \Delta m_d) &= f_{B^0} \cdot \mathcal{P}_T^{B^0}(ct^* | S_D^{OST} \cdot \mathcal{D}, \sigma_{ct^*}, \Delta m_d) \\ &+ f_{B^+} \cdot \mathcal{P}_T^{B^+}(ct^* | S_D^{OST} \cdot \mathcal{D}, \sigma_{ct^*}), \end{aligned} \quad (7.27)$$

where the PDF $\mathcal{P}_T^{B^0}$ is given by Equations (7.7) and (7.8), while $\mathcal{P}_T^{B^+}$ is given by Equations (7.10) and (7.11).

We have obtained the PDFs describing signal and backgrounds in $\mathcal{P}^i(m)$, $\mathcal{P}^i(m_{\ell D})$, and $\mathcal{P}^i(ct^* | \sigma_{ct^*})$ in Section 5.4.2. With the introduction of tagging information, the ct^* PDFs are now coupled with the tagging space, and the modified PDFs $\mathcal{P}_T^i(ct^* | \mathcal{D}, \sigma_{ct^*})$ are derived in Sections 7.1, 7.2, and 7.3, respectively. The unconditional PDFs in the σ_{ct^*} space are obtained in the same manner as described in Section 5.3.6 for B_s^0 modes. The unconditional PDFs describing the OST dilutions $\mathcal{P}^i(\mathcal{D})$ are also obtained using the same procedure as for σ_{ct^*} . The distributions of \mathcal{D} for the combinatorial background are derived using the candidates from the D mass sidebands. The signal distributions for the $\mathcal{P}^i(\mathcal{D})$ are obtained from the D mass signal region after performing sideband subtraction of background candidates. The $\mathcal{P}^i(\mathcal{D})$ PDFs for the physics and fake lepton backgrounds are the same as the ones for the signal. The $\mathcal{P}^i(\mathcal{D})$ PDFs for the signal and combinatorial background in the $\ell^+ \bar{D}^0$ modes are shown for the B (left) and D (right) triggers in Figure 7.1.

After obtaining the PDFs describing the signal and different sources of backgrounds, we next determine the tagging efficiencies and dilutions for these components. The tagging efficiency $\varepsilon^{\text{comb}}$ and the average dilution $\mathcal{D}_{avg}^{\text{comb}}$ for the combinatorial background are summarized in Tables 7.1 and 7.2, respectively. Next, the efficiency for the signal candidates is determined. We again emphasize that the efficiencies and dilutions for the physics background are constrained to be the same as the signal, since the performance of the combined OST is independent of the B meson species. The efficiency of the fake lepton background is assumed to be the same as the signal, with zero dilution (see Section 7.3.2).

7.5 Results

After the determination of PDFs describing signal and backgrounds in the different spaces, and the corresponding tagging efficiencies and dilutions, we simultaneously fit for the OST dilution scaling factor S_D^{OST} , and the B^0 - \bar{B}^0 oscillation frequency Δm_d in the combined $\ell^+ D^-$, $\ell^+ \bar{D}^0$ and $\ell^+ D^{*-}$ samples. The likelihood fit projections for the time-dependent asymmetry defined in Equation (7.9) is shown in Figure 7.2. The

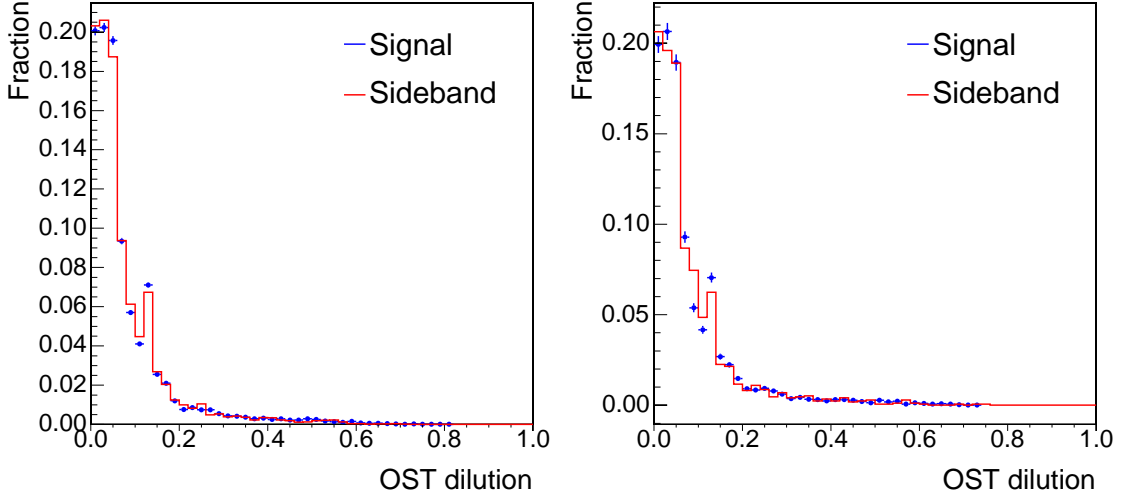


Figure 7.1: The combined OST \mathcal{D} normalized distributions for signal and combinatorial background in the B (left) and D (right) triggers in $\ell^+ \bar{D}^0$ modes.

data points represent the asymmetry calculated in each bin in ct , after subtracting backgrounds using their corresponding PDF descriptions. Since this sample is an admixture of charged and neutral B mesons, the effect on the time-dependent mixing asymmetry from B^+ (straight line) and B^0 (cosine) is indicated separately in Figure 7.2. The amplitude of the time-dependent mixing asymmetry at proper time $t = 0$ is a measure of the effective dilution of the tagging method yielding a direct calibration of the tagging dilution. The results for the dilution scaling factor S_D^{OST} and the signal tagging efficiencies ε in the three data-taking periods are presented in Table 7.3, together with the combined measurement of S_D^{OST} for 1 fb^{-1} of data. The overall OST dilution scaling factor is

$$S_D^{OST} = 107.9 \pm 1.0 \text{ (stat.)} \pm 3.3 \text{ (syst.)} \text{ \%}.$$

The overall systematic uncertainties are also displayed in Table 7.3, and described in Section 7.6. The measured B^0 oscillation frequency is given by,

$$\Delta m_d = 0.510 \pm 0.009 \text{ (stat.)} \pm 0.016 \text{ (syst.)} \text{ ps}^{-1}.$$

The value of Δm_d agrees with the world average value [13] within errors. The uncertainty in the Δm_d measurement is dominated by the systematic uncertainty, as discussed in Section 7.6.

A quantity of significant importance in the current measurement is the overall tagging effectiveness, $\varepsilon \mathcal{D}^2$. It is readily calculated from the knowledge of the dilution scaling factor S_D^{OST} and the event by event predicted OST dilution as follows:

$$\varepsilon [S_D^{OST}]^2 \langle \mathcal{D}^2 \rangle,$$

Decay Samples	$\ell^+ D^0$	$\ell^+ D^{*-}$	$\ell^+ D^-$
	Period 1		
$\mu D, B$ trigger	0.9614	0.9615	0.9626
$e D, B$ trigger	0.9621	0.9427	0.9643
$\mu D, D$ trigger	0.9584	0.9530	0.9654
$e D, D$ trigger	0.9639	0.9530	0.9605
$\mu D, \ell + SVT$ trigger	0.9777	0.9730	0.9756
$e D, \ell + SVT$ trigger	0.9748	0.9785	0.9730
	Period 2		
$\mu D, B$ trigger	0.9668	0.9712	0.9661
$e D, B$ trigger	0.9675	0.9532	0.9680
$\mu D, D$ trigger	0.9604	0.9605	0.9660
$e D, D$ trigger	0.9661	0.9605	0.9665
$\mu D, \ell + SVT$ trigger	0.9819	0.9732	0.9803
$e D, \ell + SVT$ trigger	0.9820	0.9505	0.9773
	Period 3		
$\mu D, B$ trigger	0.9703	0.9575	0.9694
$e D, B$ trigger	0.9712	0.9375	0.9698
$\mu D, D$ trigger	0.9661	0.9632	0.9720
$e D, D$ trigger	0.9624	0.9632	0.9703
$\mu D, \ell + SVT$ trigger	0.9850	0.9712	0.9790
$e D, \ell + SVT$ trigger	0.9831	0.9895	0.9811

Table 7.1: Combined OST efficiencies for the combinatorial background in $B \rightarrow \ell DX$ candidates in the three data-taking periods.

where the quantity $\langle \mathcal{D}^2 \rangle$ is calculated by averaging over the event by event dilution, and applying background subtraction using the D mass sidebands. The combined OST performance in the various data-taking periods is shown in Table 7.3 and the overall tagging performance is obtained as

$$\epsilon \mathcal{D}^2 = 1.82 \pm 0.04 \text{ (stat.)} \pm 0.11 \text{ (syst.) } \% \quad .$$

Table 7.4 shows a comparison between the observed dilution when using the predicted dilution as furnished by the combined opposite side tag (based on quantities like p_T^{rel} , $\mathcal{L}_{e,\mu}$, jet charge etc.), and using an average dilution assigned to tags irrespective of their quality. Utilization of the event by event predicted dilution improves the overall dilution by $\sim 60\%$.

Decay	$\ell^+ D^0$	$\ell^+ D^{*-}$	$\ell^+ D^-$
Samples	Period 1		
$\mu D, B$ trigger	+0.0333	+0.1200	+0.0059
$e D, B$ trigger	+0.0226	-0.0556	+0.0201
$\mu D, D$ trigger	+0.0121	+0.0000	+0.0207
$e D, D$ trigger	+0.0325	+0.0000	+0.0111
$\mu D, \ell + SVT$ trigger	+0.0190	-0.1300	+0.0068
$e D, \ell + SVT$ trigger	+0.0344	+0.2308	+0.0519
	Period 2		
$\mu D, B$ trigger	+0.0269	+0.0693	+0.0108
$e D, B$ trigger	+0.0326	-0.0612	+0.0045
$\mu D, D$ trigger	+0.0330	+0.0000	+0.0151
$e D, D$ trigger	+0.0129	+0.0000	+0.0051
$\mu D, \ell + SVT$ trigger	+0.0257	+0.1035	+0.0043
$e D, \ell + SVT$ trigger	+0.0328	+0.0000	+0.0234
	Period 3		
$\mu D, B$ trigger	+0.0244	-0.0222	+0.0189
$e D, B$ trigger	+0.0170	-0.0667	+0.0156
$\mu D, D$ trigger	+0.0445	+0.0000	-0.0070
$e D, D$ trigger	-0.0045	+0.0000	+0.0035
$\mu D, \ell + SVT$ trigger	+0.0242	+0.0693	+0.0213
$e D, \ell + SVT$ trigger	+0.0283	+0.0346	+0.0087

Table 7.2: Combined OST dilutions for the combinatorial background in $B \rightarrow \ell DX$ candidates in the three data-taking periods.

7.6 Systematic Uncertainties

We evaluate the systematic uncertainties on the two fit parameters of the combined fit of the three ℓD modes: the dilution scaling factor for the combined opposite side tag algorithm S_D^{OST} and Δm_d . Systematic uncertainties originate from several sources in this measurement. All the systematic errors have been evaluated directly on data, except in the case of systematic errors due to the proper time scale uncertainty. Toy Monte Carlo implementation for studying systematic errors is not required for the B^0 and B^+ samples, as the statistical power of the samples is large enough so that any systematic variation in the central values is always statistically significant. The systematic errors are determined by varying the sources of error within reasonable ranges, repeating the fit of the two parameters, and assigning the difference in fit values between the default and the altered fit as the systematic error.

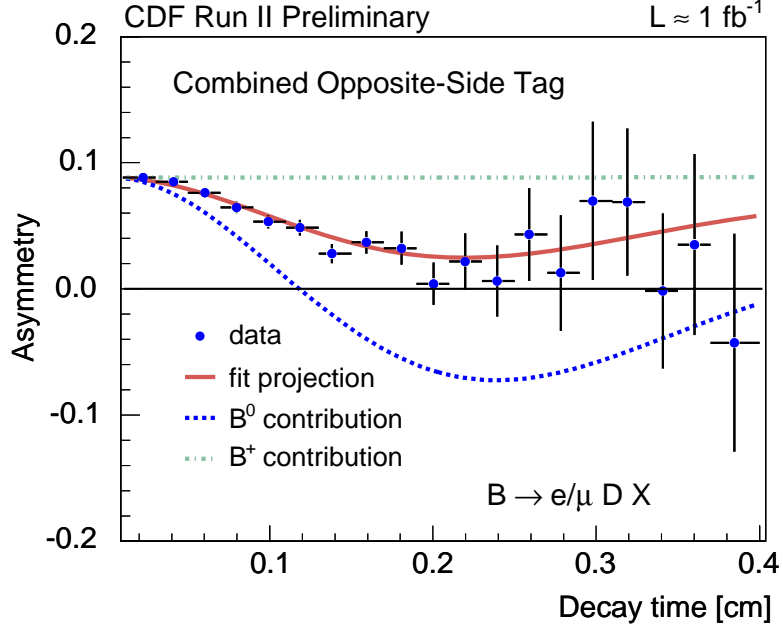


Figure 7.2: Time-dependent mixing asymmetry using the combined OST algorithm as applied to semileptonic B^0 and B^+ samples.

7.6.1 Fake Lepton Background

The dominant source of systematic errors for the S_D^{OST} determination is the contamination of ℓD samples from fake lepton background. The three uncertainties arising from the fake lepton background are related to its dilution, fraction, and ct^* shape.

In the final fit for S_D^{OST} and Δm_d , the dilution of the fake lepton background is fixed to zero (see Section 7.3). The fake lepton dilution has been cross-checked utilizing the fake lepton sample obtained from the anti-likelihood selection (see Section 3.4), and found to be consistent with the zero dilution hypothesis, albeit with large statistical errors. For studying the systematic uncertainties arising from the zero dilution hypothesis, we have undertaken a conservative approach. We first assume

Dataset	$\sqrt{\langle \mathcal{D}_{pred}^2 \rangle}$ [%]	ε [%]	S_D^{OST} [%]	$\varepsilon \mathcal{D}^2$ [%]
Period 1	12.49	95.29 ± 0.04	$103.4 \pm 1.8 \pm 3.3$	$1.60 \pm 0.06 \pm 0.11$
Period 2	13.00	96.05 ± 0.03	$109.8 \pm 1.6 \pm 3.3$	$1.96 \pm 0.06 \pm 0.12$
Period 3	13.09	96.32 ± 0.04	$110.3 \pm 2.2 \pm 3.3$	$2.01 \pm 0.09 \pm 0.13$
All	12.77	95.88 ± 0.02	$107.9 \pm 1.0 \pm 3.3$	$1.82 \pm 0.04 \pm 0.11$

Table 7.3: Combined opposite side tagging performance for the three data-taking periods.

Dataset	$S_D^{OST} \sqrt{\langle \mathcal{D}_{\text{pred}}^2 \rangle} [\%]$	$\langle \mathcal{D} \rangle [\%]$
Period 1	12.91 ± 0.22	7.91 ± 0.24
Period 2	14.27 ± 0.21	9.11 ± 0.22
Period 3	14.44 ± 0.29	9.27 ± 0.31
All	13.79 ± 0.13	8.71 ± 0.14

Table 7.4: Comparison of the combined OST performance using event by event dilution and average dilution for the three data-taking periods.

that the fake lepton background has a positive dilution equal in magnitude to the dilution predicted by the combined opposite side tag for the signal candidates, and repeat the fit for S_D^{OST} and Δm_d . Next, we assume that the fake lepton background candidates have a dilution equal in magnitude to the signal candidates with negative sign, and again perform the fit. The largest variation in the fit values of S_D^{OST} and Δm_d in the two cases is then taken as an estimate of the systematic error from the uncertainty in fake lepton dilution.

The fractions of the fake lepton contribution to the ℓD samples is determined via a simultaneous fit to the D mass and $m_{\ell D}$ distributions (see Sections 5.4.2 and 5.3.4). The statistical uncertainties in the fractions of the fake lepton background are used to compute the systematic error. The fractions are varied within the range of statistical errors, and the largest variation in the values of S_D^{OST} and Δm_d is then inferred as the systematic error.

The third source of systematic error from fake lepton background originates from the uncertainty in the ct^* shape, as derived using the fake lepton sample. The fake lepton background shape is described by two tailed Gaussians together with another Gaussian distribution (see Section 5.3.6). The “slope” of the lifetime parameters describing the fake lepton background in the ct^* space is altered by a relative amount of 15%, corresponding to the statistical error in its determination. The associated variation in the S_D^{OST} and Δm_d fit values is utilized as the systematic error from the uncertainty in the fake lepton ct^* shape.

7.6.2 Sample Composition of ℓD Samples

The sample composition of ℓD samples in terms of the B^+ and B^0 meson fractions constitutes the largest source of systematic uncertainty in the Δm_d determination. The three ℓD modes result from a multitude of decay sequences of the B^+ and B^0 mesons (see Section 5.4.1). We have utilized the world average values of the sample composition parameters [13]. To assess the systematic errors arising from the uncertainties in the sample composition parameters as summarized in Table 5.11, the fit to S_D^{OST} and Δm_d is repeated with the parameters varied within the quoted errors of

the respective world average values. Changing the values of the sample composition parameters not only changes the relative contributions of B^+ and B^0 mesons in the ℓD samples, but also leads to alterations in the derived κ -factor distributions and $\xi(t^*)$ functions. Therefore, the systematic error is computed after taking into account all these modifications.

The uncertainties in the sample composition of the ℓD samples also result in slightly different shapes of the B meson signal in the $m_{\ell D}$ distributions for the ℓD modes. As a consequence, some of the fits in the $m_{\ell D}$ space for the ℓD candidates are of poorer quality. To estimate the systematic uncertainty, we let the relative contributions of B^+ and B^0 mesons free in the fit to $m_{\ell D}$, and determine the fractions in the ℓD data. The difference in B^+ and B^0 contributions thus obtained is utilized for assigning the associated systematic error.

7.6.3 Physics Background Fractions

The ℓD samples are also contaminated with a variety of sources of physics backgrounds (see Section 5.4.2 and Table 5.15). The contribution from the various sources of physics backgrounds is evaluated using Monte Carlo samples. However, this involves knowledge of branching ratios for the various contributing decay modes, some of which have large uncertainties. Consequently, the estimated contamination from these sources is poorly known, and we assign a 25% relative uncertainty in the fractions of the physics backgrounds. The $S_{\mathcal{D}}^{\text{OST}}$ and Δm_d fit is repeated with the altered physics background fractions to extract the corresponding systematic uncertainty.

7.6.4 $\xi(t^*)$ Determination from Monte Carlo

The efficiency functions $\xi(t^*)$ describe the effect of trigger and reconstruction biases on the pseudo-proper decay time of the signal B mesons. These biases are derived using Monte Carlo samples that include a full detector and trigger simulation. There are two sources of systematic errors in determining the $\xi(t^*)$ functions. The $\xi(t^*)$ shapes are derived utilizing the signal Monte Carlo samples that use the world average value [13] of the B meson lifetimes. Therefore, we vary the lifetime values used in the Monte Carlo simulation within the errors of the respective world average lifetime values. The fit to $S_{\mathcal{D}}^{\text{OST}}$ and Δm_d is repeated with the new $\xi(t^*)$ functions, and the corresponding systematic error is evaluated.

The second source of systematic error in the $\xi(t^*)$ determination originates from the description of correlation between the online and offline impact parameters in Monte Carlo as compared to data. The difference in the online (as measured by the SVT) and offline (as measured by the SVX) impact parameters in Monte Carlo, when compared to data, warrant an additional smearing of $\sim 12 \mu\text{m}$ in the Monte Carlo before the application of SVT trigger confirmation requirements. We repeat the fit to $S_{\mathcal{D}}^{\text{OST}}$ and Δm_d , and assign a systematic error equal to the difference in the measured

values.

7.6.5 Combinatorial Background Characterization

Two sources of systematic errors are addressed to account for the uncertainties in the characterization of the combinatorial background. The first source of systematic error arises from the uncertainty in the fraction of the combinatorial background, as determined by a fit to the D invariant mass of the ℓD candidates (see Section 5.3.1). We vary the contributions of the combinatorial background within their statistical uncertainties, and assess the systematic errors from the observed deviations in the default fit values for $S_{\mathcal{D}}^{\text{OST}}$ and Δm_d .

The second source of systematic error originates from the description of the combinatorial background in the $m_{\ell D}$ and ct^* variables. The combinatorial background shapes in the $m_{\ell D}$ and ct^* spaces are obtained by fitting the background candidates obtained from the D mass sidebands. However, the shapes thus obtained may not adequately describe the combinatorial background underneath the D mass peak. We address this issue by letting certain background parameters float while performing fits in the $m_{\ell D}$ and ct^* variables for the ℓD modes (see Sections 5.3.3 and 5.3.6). To further account for the systematic errors introduced in the characterization of the combinatorial background, the choice of sidebands is shifted $\pm 50 \text{ MeV}/c^2$ from their nominal values. New shapes for the combinatorial background description in the $m_{\ell D}$ and ct^* spaces is obtained, and we repeat the fit utilizing the newly obtained shapes. The systematic uncertainty is estimated using the differences from the default fit values for $S_{\mathcal{D}}^{\text{OST}}$ and Δm_d .

7.6.6 Fixing $\tau(B)$ to the World Average Value

While performing a fit determining $S_{\mathcal{D}}^{\text{OST}}$ and Δm_d , the lifetimes of the B^+ and B^0 mesons are fixed to their corresponding world average values [13]. As a source of systematic error, we vary the lifetimes within the quoted errors on these world average values. The variations in the values of the two fit parameters is taken as the associated systematic uncertainties.

7.6.7 σ_{t^*} Calibration Uncertainty

As described in Section 5.3.5, the value of the error on the pseudo-proper decay length σ_{t^*} as returned by the vertex fit underestimates the true error. A calibration of σ_{t^*} is performed on a sample of prompt D^- +track combinations to obtain a scaling correction factor for each ℓD candidate, and properly describe the pseudo-proper decay length resolution. For estimating the systematic error due to this σ_{t^*} determination, we alter the values of the scale factors using the uncertainties quoted

from the $D^- + \text{track}$ study. The scale factors are changed by ± 0.06 , and we assess the systematic error by repeating the fit.

7.6.8 Uncertainty in Proper Time Scale

Uncertainties in the proper time scale result into systematic errors in the determination of Δm_d . Three sources of uncertainties in the time scale have been identified. The first source of uncertainty in the absolute time scale is caused by the imperfect knowledge of the wafer positions in the silicon detectors (silicon alignment) [71]. To ascertain the impact of the silicon alignment on the measurement of B meson lifetimes, distortions in the form of radial movements and bowing of silicon wafers are introduced in Monte Carlo simulation. Then, a fit for the B meson lifetimes is performed with the nominal alignment table. The maximum variation in measured lifetimes due to the distortions is found to be $\pm 1 \mu\text{m}$, which corresponds to a 0.2% systematic uncertainty on the absolute time scale.

The second source of systematic error arises from biases introduced in the primary vertex determination (see Section 5.3.5). The same side B daughter tracks are ruled out of the primary vertex calculation, and the primary vertex is determined by vertexing the underlying event tracks. However, inclusion of the tracks from the opposite side B , that form another secondary vertex in the event, can bias the measurement of the primary vertex. To estimate the extent of this bias, fully reconstructed samples of B^+ and B^0 decays are analyzed, and the primary vertex positions from the average beam position and the primary vertex finder are compared. To enhance the effect of the bias, only candidates with opposite side tags are included in the study. The maximum variation found is $\pm 1 \mu\text{m}$, and we take 50% of this value as the uncertainty in the primary vertex determination.

The third source of systematic error originates from biases in the track helix fit procedure. Inaccurate measurements of the track curvature feed into the transverse decay length determination via the vertex fit, and also into proper decay time calculation through the mis-measured B transverse momentum. Thus, the measured lifetimes of the B mesons are affected. The effect of the biases have been studied in a Monte Carlo simulation of the tracking detectors and determined to be $\pm 1.3 \mu\text{m}$, corresponding to 0.3% systematic error on the proper time scale of the CDF detector.

7.6.9 Summary of Systematic Uncertainties

Table 7.5 summarizes the systematic uncertainties from the various sources. The dominant sources of systematic uncertainties for the dilution scale factor of the combined opposite side tagging algorithm S_D^{OST} are the dilution and the fraction of the fake lepton background. The dominant source of uncertainty for the Δm_d measurement is the sample composition of the ℓD samples, which limits the determination of Δm_d in semileptonic B decays.

Source of Systematic Error	S_D^{OST} [%]	Δm_d [ps ⁻¹]
Dilution of fake background	2.50	0.0015
Fraction of fake background	1.78	0.0018
Shape of fake background	0.15	0.0009
Sample composition	0.98	0.0147
$m_{\ell D}$ description	0.03	0.0007
Fraction of physics bkg.	0.32	0.0033
ct^* efficiency	0.14	0.0011
Combinatorial background	0.06	0.0009
Fixed $\tau_{B_{d,u}}$	0.04	0.0005
σ_{ct}	0.00	0.0005
ct scale	0.00	0.0020
Total systematic error	3.32	0.0155

Table 7.5: Systematic uncertainties for the combined opposite side tag dilution scaling factor S_D^{OST} and Δm_d .

Chapter 8

B_s^0 - \bar{B}_s^0 Oscillations

We have achieved a good understanding of the B_s^0 samples in terms of their signal and background contributions as described in Sections 5.3.4 and 5.3.7. The tools necessary to analyze B_s^0 - \bar{B}_s^0 oscillations have also been developed in the previous chapters. The proper time resolution of the B_s^0 decays has been understood using a calibration sample of prompt D^+ plus track combinations. The opposite side flavor tagging algorithms have been studied in large semileptonic samples collected via the ℓ +SVT triggers, and further calibrated utilizing the observed dilution in B^+ and B^0 samples. The likelihood fit framework has been extensively tested with the measurement of B^0 - \bar{B}^0 oscillations. The same side kaon tagging, developed using Monte Carlo samples, boosts the flavor tagging power provided by the OST further by a factor of 2-3. Thus, to analyze B_s^0 - \bar{B}_s^0 oscillations, we intend to incorporate the SSKT into the fit. This implies that for a given B_s^0 candidate, we can have up to two tagging decisions. The fit framework has to be modified in order to handle candidates that are tagged by both algorithms. In addition, a new parameter is introduced in the fit to search for a peak in the power spectrum of the data as a function of frequency. This parameter, called the amplitude \mathcal{A} , has been described in Section 1.4.1.

8.1 Time Evolution PDF for B_s^0 - \bar{B}_s^0 Mixing

The PDF describing the B_s^0 - \bar{B}_s^0 system has already been derived in Section 1.2. With the introduction of the amplitude \mathcal{A} , and the same side kaon tagging, we derive the time evolution PDFs for the B_s^0 - \bar{B}_s^0 system using Equations (1.35) and (1.36), neglecting CP violation in the B_s^0 system and setting $\Delta\Gamma_s$ to zero. $\Delta\Gamma_s$ is expected to be large $\sim 0.1 \times \Gamma_s$ for the B_s^0 - \bar{B}_s^0 system, however neglecting it in the time evolution PDFs results into little or no systematic uncertainties (see Section 8.5.12). We follow the same steps as described in Section 7.1 for B^0 - \bar{B}^0 mixing to obtain the time evolution PDFs for the B_s^0 - \bar{B}_s^0 system. If T_{OST} and T_{SST} denote the tagging decisions of the opposite side and same side tagging algorithms, respectively, then we can

express the mixing PDF as follows:
untagged ($T_{\text{OST}} = 0$ and $T_{\text{SST}} = 0$):

$$\begin{aligned} \mathcal{P}_T^{B_s^0}(t^* | \mathcal{D}, \sigma_{t^*}) &= \frac{1}{N^{B_s^0}} (1 - \varepsilon_{\text{OST}}) (1 - \varepsilon_{\text{SST}}) \Gamma_s e^{-\Gamma_s \kappa t^*} \theta(\kappa t^*) \\ &\quad \otimes_{\kappa} \mathcal{H}(\kappa)^{B_s^0} \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B_s^0}(t^*), \end{aligned} \quad (8.1)$$

OST tagged only ($T_{\text{OST}} = \pm 1$ and $T_{\text{SST}} = 0$):

$$\begin{aligned} \mathcal{P}_T^{B_s^0}(t^* | \mathcal{D}, \sigma_{t^*}) &= \varepsilon_{\text{OST}} (1 - \varepsilon_{\text{SST}}) \frac{1}{N^{B_s^0}} \frac{\Gamma_s}{2} e^{-\Gamma_s \kappa t^*} \theta(\kappa t^*) [1 + \mathcal{A} \mathcal{D}_{\text{OST}} T_{\text{OST}} \cos(\Delta m_s \kappa t^*)] \\ &\quad \otimes_{\kappa} \mathcal{H}^{B_s^0}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B_s^0}(t^*), \end{aligned} \quad (8.2)$$

SST tagged only ($T_{\text{SST}} = \pm 1$ and $T_{\text{OST}} = 0$):

$$\begin{aligned} \mathcal{P}_T^{B_s^0}(t^* | \mathcal{D}, \sigma_{t^*}) &= \varepsilon_{\text{SST}} (1 - \varepsilon_{\text{OST}}) \frac{1}{N^{B_s^0}} \frac{\Gamma_s}{2} e^{-\Gamma_s \kappa t^*} \theta(\kappa t^*) [1 + \mathcal{A} \mathcal{D}_{\text{SST}} T_{\text{SST}} \cos(\Delta m_s \kappa t^*)] \\ &\quad \otimes_{\kappa} \mathcal{H}^{B_s^0}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B_s^0}(t^*), \end{aligned} \quad (8.3)$$

OST & SST tagged ($T_{\text{OST}} = \pm 1$ and $T_{\text{SST}} = \pm 1$):

$$\begin{aligned} \mathcal{P}_T^{B_s^0}(t^* | \mathcal{D}, \sigma_{t^*}) &= \frac{\varepsilon_{\text{OST}} \varepsilon_{\text{SST}}}{N^{B_s^0}} \frac{\Gamma_s}{2} e^{-\Gamma_s \kappa t^*} \theta(\kappa t^*) \cdot \frac{1}{2} \cdot [(1 + T_{\text{OST}} \mathcal{D}_{\text{OST}} T_{\text{SST}} \mathcal{D}_{\text{SST}}) \\ &\quad + \mathcal{A} (T_{\text{OST}} \mathcal{D}_{\text{OST}} + T_{\text{SST}} \mathcal{D}_{\text{SST}}) \cos(\Delta m_s \kappa t^*)] \\ &\quad \otimes_{\kappa} \mathcal{H}^{B_s^0}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B_s^0}(t^*), \end{aligned} \quad (8.4)$$

where the combination of the OST and the SST has been carried out assuming that the two algorithms provide independent tagging decisions [70]. The same side tagging is only applied to B_s^0 candidates with high $m_{\ell^+ D_s^-}$, as the behavior of the same side kaon tag is not easily understood on the physics background from $B \rightarrow DDX$ type of decays. In order to reduce this contamination, B_s^0 candidates with $m_{\ell^+ D_s^-} > 3.5 \text{ GeV}/c^2$ are selected for the same side tagging application, where the contamination from $B \rightarrow DDX$ types of decays is reduced to $< 2\%$ of its overall fraction. This issue is further addressed in the evaluation of systematic errors in Section 8.5.2.

8.2 Time Evolution PDF for Backgrounds

In the case of the B_s^0 samples, there are four types of backgrounds (see Section 5.1). We group the background components into three types :

1. Oscillating backgrounds: Physics background originating from B^0 and B_s^0 decays belong to this category.
2. Non-oscillating backgrounds: Physics background arising from B^+ decays as well as combinatorial and fake lepton backgrounds are associated with this category.
3. D^- reflection background: Since the D^- reflection background originates from real B decays, the OST behavior for the D^- reflection background in the $D_s^- \rightarrow K^{*0} K^-$ mode is modeled in the same manner as the signal. For the same side kaon tagging, the D^- reflection background is assigned an average asymmetry obtained directly from measurements on $B \rightarrow \ell^+ D^- X$ data.

The background description in the case of B_s^0 decays closely follows the treatment in the case of B^0 and B^+ decays in Section 7.3. The efficiencies and dilutions for the different sources of physics background are assumed to be the same for the OST and the SSKT (after the $m_{\ell^+ D_s^-} > 3.5 \text{ GeV}/c^2$ requirement) as the signal B_s^0 decays. The efficiency for the fake lepton background is also assumed to be the same as the signal B_s^0 decays, while it is hypothesized to have a null dilution.

8.2.1 Oscillating Backgrounds

Physics Backgrounds from B^0 Decays

The time evolution PDF to describe the physics background from B^0 decays follows the same description as obtained for describing B_s^0 - \bar{B}_s^0 mixing discussed in Section 8.1, with the following modifications in Equations (8.1) to (8.4):

$$\Gamma_s \rightarrow \Gamma_d, \quad \mathcal{A} = 1, \quad \Delta m_s \rightarrow \Delta m_d, \quad \mathcal{H}^{B_s^0}(\kappa) \rightarrow \mathcal{H}^{B^0 \text{ phys}}(\kappa), \quad \xi^{B_s^0}(t^*) \rightarrow \xi^{B^0 \text{ phys}}(t^*).$$

The efficiency and dilution for the physics background from B^0 decays are taken to be the same as for the signal B_s^0 decays, with the modification that the OST tagging decision for the $B \rightarrow DD\bar{X}$ type of decays is flipped to account for the wrong charge of the lepton with respect to the flavor of the B meson. The κ -factor distribution $\mathcal{H}^{B^0 \text{ phys}}(\kappa)$ and the efficiency function $\xi^{B^0 \text{ phys}}(t^*)$ specific to the B^0 physics background are derived from Monte Carlo.

Physics Backgrounds from B_s^0 Decays

The physics background due to the B_s^0 decays arise due to contributions from:

1. $B_s^0 \rightarrow D_s^{(*)} D^{(*)} X, D^{(*)} \rightarrow \ell^+ Y,$
2. $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-} X, D_s^{(*)} \rightarrow \ell^+ Y.$

These decays of B_s^0 mesons are of the type $B_s^0 \rightarrow D_1^+ D_2^- X$, where contributions from $B_s^0 \rightarrow D_1^+ D_2^- X$ and $B_s^0 \rightarrow D_1^- D_2^+ X$ are equally likely. Consequently, the physics background from B_s^0 decays produces an equal admixture of negatively and positively charged leptons. Hence, the behavior of physics background originating from B_s^0 decays is described by setting the flavor tag dilution $\mathcal{D} = 0$ in Equations (8.1) to (8.4). Thus, the obtained PDF for the physics background from B_s^0 decays is the same as in Equation (7.15).

8.2.2 Non-oscillating Backgrounds

This category includes backgrounds that do not oscillate such as the physics background from B^+ decays, or backgrounds that are not explicitly described by a time dependent asymmetry like the D^- reflection background.

Physics Backgrounds from B^+ Decays

The physics background arising from B^+ decays are described by implementing the following modifications to the PDF for B_s^0 decays in Equations (8.1) to (8.4):

$$\Gamma_s \rightarrow \Gamma_u, \quad \mathcal{A} = 1, \quad \Delta m_s = 0, \quad \mathcal{H}^{B_s^0}(\kappa) \rightarrow \mathcal{H}^{B^+ \text{ phys}}(\kappa), \quad \xi^{B_s^0}(t^*) \rightarrow \xi^{B^+ \text{ phys}}(t^*).$$

The efficiency and dilution for the B^+ physics background are assumed to be the same as for the signal, with the same modification as in the case of B^0 : the OST tagging decision for the $B \rightarrow DDX$ type of decays is flipped to account for the wrong charge of the lepton with respect to the flavor of the B meson.

Combinatorial Background

The combinatorial background is hence assigned time-averaged asymmetries $\mathcal{D}_{\text{OST}}^{\text{comb}}$ and $\mathcal{D}_{\text{SST}}^{\text{comb}}$, and the corresponding tagging efficiencies $\varepsilon_{\text{OST}}^{\text{comb}}$ and $\varepsilon_{\text{SST}}^{\text{comb}}$. Therefore, the PDFs describing the combinatorial background are given by:

untagged ($T_{\text{OST}} = 0$ and $T_{\text{SST}} = 0$):

$$\mathcal{P}_T^{\text{comb}}(t^* | \mathcal{D}, \sigma_{t^*}) = (1 - \varepsilon_{\text{OST}}^{\text{comb}}) (1 - \varepsilon_{\text{SST}}^{\text{comb}}) \cdot \mathcal{P}^{\text{comb}}(t^*, \sigma_{t^*}), \quad (8.5)$$

OST tagged only ($T_{\text{OST}} = \pm 1$ and $T_{\text{SST}} = 0$):

$$\mathcal{P}_T^{\text{comb}}(t^* | \mathcal{D}, \sigma_{t^*}) = \varepsilon_{\text{OST}}^{\text{comb}} (1 - \varepsilon_{\text{SST}}^{\text{comb}}) \frac{1}{2} \cdot (1 + T_{\text{OST}} \cdot \mathcal{D}_{\text{OST}}^{\text{comb}}) \cdot \mathcal{P}^{\text{comb}}(t^*, \sigma_{t^*}), \quad (8.6)$$

SST tagged only ($T_{\text{SST}} = \pm 1$ and $T_{\text{OST}} = 0$):

$$\mathcal{P}_T^{\text{comb}}(t^* | \mathcal{D}, \sigma_{t^*}) = \varepsilon_{\text{SST}}^{\text{comb}} (1 - \varepsilon_{\text{OST}}^{\text{comb}}) \frac{1}{2} \cdot (1 + T_{\text{SST}} \cdot \mathcal{D}_{\text{SST}}^{\text{comb}}) \cdot \mathcal{P}^{\text{comb}}(t^*, \sigma_{t^*}), \quad (8.7)$$

OST & SST tagged ($T_{\text{OST}} = \pm 1$ and $T_{\text{SST}} = \pm 1$):

$$\begin{aligned} \mathcal{P}_T^{\text{comb}}(t^* | \mathcal{D}, \sigma_{t^*}) = & \frac{1}{2} \cdot \varepsilon_{\text{OST}}^{\text{comb}} \varepsilon_{\text{SST}}^{\text{comb}} \left[\frac{1}{2} \cdot (1 + T_{\text{OST}} \cdot \mathcal{D}_{\text{OST}}^{\text{comb}}) (1 + T_{\text{SST}} \cdot \mathcal{D}_{\text{SST}}^{\text{comb}}) \right] \\ & \cdot \mathcal{P}^{\text{comb}}(t^*, \sigma_{t^*}), \end{aligned} \quad (8.8)$$

where $\mathcal{P}^{\text{comb}}(t^*, \sigma_{t^*})$ is the joint PDF describing the combinatorial background in the t^* and σ_{t^*} spaces, obtained in Section 5.3.6. The efficiencies for the combinatorial background for the OST and the SST tags are assumed to be different as compared to those for the signal, and determined using the behavior of tagging on the candidates belonging to the D_s^- mass sidebands defined in Table 5.4.

Fake Lepton Background

The fake lepton background is assumed to be non-oscillating, with the OST and the SST tagging efficiency being the same as the signal. Furthermore, we assume that the time-averaged asymmetry of the fake lepton background is zero ($\mathcal{D}_{\text{avg}}^{\text{fake}} = 0$). These assumptions have been cross-checked with the fake lepton sample, which yield consistent results with large uncertainties on the asymmetry $\mathcal{D}_{\text{avg}}^{\text{fake}} = 0$. The systematic uncertainty arising from this hypothesis has been studied and is summarized in Section 7.6.1 for the ℓD samples and in Section 8.5.4 for the $D_s^- \ell^+$ data. The time evolution for the fake lepton background can be obtained from the description for the combinatorial background in Equations (8.5) to (8.8), with the following modifications:

$$\begin{aligned} \mathcal{D}_{\text{OST}}^{\text{comb}} &\rightarrow 0, \quad \mathcal{D}_{\text{SST}}^{\text{comb}} \rightarrow 0, \quad \varepsilon_{\text{OST}}^{\text{comb}} \rightarrow \varepsilon_{\text{OST}}, \quad \varepsilon_{\text{SST}}^{\text{comb}} \rightarrow \varepsilon_{\text{SST}}, \\ \mathcal{P}^{\text{comb}}(t^*, \sigma_{t^*}) &\rightarrow \mathcal{P}^{\text{fake}}(t^*, \sigma_{t^*}), \end{aligned}$$

where the efficiency is set to be the same as the B_s^0 signal, and the dilution is set to zero as mentioned earlier. This yields the following PDFs for describing the fake lepton background:

untagged ($T_{\text{OST}} = 0$ and $T_{\text{SST}} = 0$):

$$\mathcal{P}_T^{\text{fake}}(t^* | \mathcal{D}, \sigma_{t^*}) = (1 - \varepsilon_{\text{OST}}) (1 - \varepsilon_{\text{SST}}) \cdot \mathcal{P}^{\text{fake}}(t^*, \sigma_{t^*}), \quad (8.9)$$

OST tagged only ($T_{\text{OST}} = \pm 1$ and $T_{\text{SST}} = 0$):

$$\mathcal{P}_T^{\text{fake}}(t^* | \mathcal{D}, \sigma_{t^*}) = \frac{1}{2} \cdot \varepsilon_{\text{OST}} (1 - \varepsilon_{\text{SST}}) \cdot \mathcal{P}^{\text{fake}}(t^*, \sigma_{t^*}), \quad (8.10)$$

SST tagged only ($T_{\text{OST}} = \pm 1$ and $T_{\text{SST}} = 0$):

$$\mathcal{P}_T^{\text{fake}}(t^* | \mathcal{D}, \sigma_{t^*}) = \frac{1}{2} \cdot \varepsilon_{\text{SST}} (1 - \varepsilon_{\text{OST}}) \cdot \mathcal{P}^{\text{fake}}(t^*, \sigma_{t^*}), \quad (8.11)$$

OST & SST tagged ($T_{\text{OST}} = \pm 1$ and $T_{\text{SST}} = \pm 1$):

$$\mathcal{P}_T^{\text{fake}}(t^* | \mathcal{D}, \sigma_{t^*}) = \frac{1}{4} \cdot \varepsilon_{\text{OST}} \varepsilon_{\text{SST}} \cdot \mathcal{P}^{\text{fake}}(t^*, \sigma_{t^*}), \quad (8.12)$$

where $\mathcal{P}^{\text{fake}}(t^*, \sigma_{t^*})$ is the joint PDF describing the fake lepton background in the t^* and σ_{t^*} spaces as determined in Section 5.3.6.

8.2.3 D^- Reflection Background

The D^- reflection in the $B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$ mode arises from real B decays (see Section 5.1). Therefore, the OST behavior for the D^- reflection background in the $D_s^- \rightarrow K^{*0} K^-$ mode is modeled in the same fashion as the signal with dedicated $\mathcal{H}(\kappa)$ distributions and $\xi(t^*)$ functions obtained from the $B \rightarrow \ell^+ D^- X$ Monte Carlo (see Section 5.3.6). The Monte Carlo contains the same admixture of B^+ and B^0 mesons as determined by the $\ell^+ D^-$ modes reconstructed for studying B^0 - \bar{B}^0 oscillations and flavor tagging. The OST efficiency and dilution for the D^- reflection background is expected to be the same as the signal. The D^- reflection behavior for the SST is modeled by assigning an average dilution ($\mathcal{D}_{\text{SST}}^{\text{refl}}$) together with the corresponding efficiency ($\varepsilon_{\text{SST}}^{\text{refl}}$), and is measured on the $B \rightarrow \ell^+ D^- X$ data. The PDFs for the D^- reflection are obtained as follows:

untagged ($T_{\text{OST}} = 0$ and $T_{\text{SST}} = 0$):

$$\begin{aligned} \mathcal{P}_T^{D^+ \text{ refl}}(t^* | \mathcal{D}, \sigma_{t^*}) &= (1 - \varepsilon_{\text{OST}}) (1 - \varepsilon_{\text{SST}}^{\text{refl}}) [f_{B^+} \cdot \left\{ \frac{1}{N_{B^+}} \Gamma_u e^{-\Gamma_u \kappa t^*} \theta(\kappa t^*) \otimes_{\kappa} \mathcal{H}^{B^+}(\kappa) \right. \\ &\quad \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^+}(t^*) \} + f_{B^0} \cdot \left\{ \frac{1}{N_{B^0}} \Gamma_d e^{-\Gamma_d \kappa t^*} \theta(\kappa t^*) \right. \\ &\quad \left. \otimes_{\kappa} \mathcal{H}^{B^0}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^0}(t^*) \} \right], \end{aligned} \quad (8.13)$$

OST tagged only ($T_{\text{OST}} = \pm 1$ and $T_{\text{SST}} = 0$):

$$\begin{aligned}
\mathcal{P}_T^{D^+ \text{ refl}}(t^* | \mathcal{D}, \sigma_{t^*}) &= \frac{1}{2} \varepsilon_{\text{OST}} (1 - \varepsilon_{\text{SST}}^{\text{refl}}) \{ f_{B^+} \cdot [1 + \mathcal{D}_{\text{OST}} T_{\text{OST}}] \cdot [\frac{1}{N^{B^+}} \Gamma_u e^{-\Gamma_u \kappa t^*} \\
&\quad \theta(\kappa t^*) \otimes_{\kappa} \mathcal{H}^{B^+}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^+}(t^*)] \\
&\quad + f_{B^0} \cdot [1 + \mathcal{D}_{\text{OST}} T_{\text{OST}} \cos(\Delta m_d \kappa t^*)] \cdot [\frac{1}{N^{B^0}} \Gamma_d e^{-\Gamma_d \kappa t^*} \theta(\kappa t^*) \\
&\quad \otimes_{\kappa} \mathcal{H}^{B^0}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^0}(t^*)] \}, \tag{8.14}
\end{aligned}$$

SST tagged only ($T_{\text{SST}} = \pm 1$ and $T_{\text{OST}} = 0$):

$$\begin{aligned}
\mathcal{P}_T^{D^+ \text{ refl}}(t^* | \mathcal{D}, \sigma_{t^*}) &= \frac{1}{2} \varepsilon_{\text{SST}}^{\text{refl}} (1 - \varepsilon_{\text{OST}}) \{ f_{B^+} \cdot [1 + \mathcal{D}_{\text{SST}}^{\text{refl}} T_{\text{SST}}] \cdot [\frac{1}{N^{B^+}} \Gamma_u e^{-\Gamma_u \kappa t^*} \\
&\quad \theta(\kappa t^*) \otimes_{\kappa} \mathcal{H}^{B^+}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^+}(t^*)] \\
&\quad + f_{B^0} \cdot [1 + \mathcal{D}_{\text{SST}}^{\text{refl}} T_{\text{SST}} \cos(\Delta m_d \kappa t^*)] \cdot [\frac{1}{N^{B^0}} \Gamma_d e^{-\Gamma_d \kappa t^*} \theta(\kappa t^*) \\
&\quad \otimes_{\kappa} \mathcal{H}^{B^0}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^0}(t^*)] \}, \tag{8.15}
\end{aligned}$$

OST & SST tagged ($T_{\text{OST}} = \pm 1$ and $T_{\text{SST}} = \pm 1$):

$$\begin{aligned}
\mathcal{P}_T^{D^+ \text{ refl}}(t^* | \mathcal{D}, \sigma_{t^*}) &= \frac{1}{2} \varepsilon_{\text{OST}} \varepsilon_{\text{SST}}^{\text{refl}} \cdot \frac{1}{2} \{ f_{B^+} \cdot [(1 + \mathcal{D}_{\text{OST}} T_{\text{OST}} \mathcal{D}_{\text{SST}}^{\text{refl}} T_{\text{SST}}) + \mathcal{D}_{\text{OST}} T_{\text{OST}} \\
&\quad + \mathcal{D}_{\text{SST}}^{\text{refl}} T_{\text{SST}}] \cdot [\frac{1}{N^{B^+}} \Gamma_u e^{-\Gamma_u \kappa t^*} \theta(\kappa t^*) \otimes_{\kappa} \mathcal{H}^{B^+}(\kappa) \\
&\quad \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^+}(t^*)] + f_{B^0} \cdot [(1 + \mathcal{D}_{\text{OST}} T_{\text{OST}} \mathcal{D}_{\text{SST}}^{\text{refl}} T_{\text{SST}}) \\
&\quad (\mathcal{D}_{\text{OST}} T_{\text{OST}} + \mathcal{D}_{\text{SST}}^{\text{refl}} T_{\text{SST}}) \cos(\Delta m_d \kappa t^*)] \cdot [\frac{1}{N^{B^0}} \Gamma_d e^{-\Gamma_d \kappa t^*} \theta(\kappa t^*) \\
&\quad \otimes_{\kappa} \mathcal{H}^{B^0}(\kappa) \otimes_{t'} G(t' - t^*, \sigma_{t^*}) \cdot \xi^{B^0}(t^*)] \}, \tag{8.16}
\end{aligned}$$

where the fractions of B^+ and B^0 contributions in the D^+ reflection background are f_{B^0} and f_{B^+} , respectively.

8.3 B_s^0 - \bar{B}_s^0 Mixing Analysis

The joint PDFs describing signal and backgrounds in the tagging and ct^* space have been derived in Sections 8.1 and 8.2. We utilize an unbinned maximum likelihood fit to extract parameters of interest. The likelihood description for the B meson samples has been explained in Section 7.4. This likelihood framework is adopted for studying B_s^0 - \bar{B}_s^0 oscillations with some modifications:

- Same side tagging is introduced in the fit, and in case a candidate is tagged by both OST and SST, a combination of tagging is implemented as described

Parameter	Period 1	Period 2	Period 3
S_D^{SST} (%)	$105.3^{+10.7}_{-14.3}$	$102.2^{+10.8}_{-14.4}$	$101.3^{+10.8}_{-14.4}$

Table 8.1: Dilution scale factors in different data-taking periods for the same side kaon tagging.

in Sections 8.1 and 8.2 [70]. The implementation assumes that the two kinds of flavor tagging algorithms are mutually independent (see Section 6.2.1 for selection requirements on SST tags that ensures independence of the OST and SST).

- As detailed in Sections 1.4.1 and 8.1, we introduce a Fourier coefficient called the amplitude \mathcal{A} , which multiplies the cosine term describing the oscillations.

$$[1 \pm \cos(\Delta m_s t)] \rightarrow [1 \pm \mathcal{A} \cdot \cos(\Delta m_s t)].$$

\mathcal{A} is the only free parameter in the fit, and is determined at various probe values of Δm_s . Unlike the Δm_d measurement, letting \mathcal{A} free in the fit does not render a measurement of the dilution. A measurement of \mathcal{A} requires that the flavor tags are properly calibrated i.e., their tagging purity is known in advance. This makes the calibration of flavor tagging a necessity, until the B_s^0 - \bar{B}_s^0 oscillations are observed.

- Exploiting the independence of the OST performance with the respect to the trigger side B meson species, its calibration has already been achieved using high statistics samples of B^+ and B^0 decays in Chapter 7. As a consequence, scaling factors (S_D^{OST}) for the predicted event by event OST dilution have been obtained in Table 7.3. The performance of the same side tags has been studied using Monte Carlo samples (see Section 6.2). The dilution scale factors for the same side tagging S_D^{SST} have been evaluated for semileptonic B_s^0 decays (see Section 6.2.5) and summarized in Table 8.1 for the three data-taking periods. The predicted event by event dilutions are scaled by these values before being applied to the B_s^0 candidates i.e.,

$$\begin{aligned} \mathcal{D}^{OST} &\rightarrow S_D^{OST} \cdot \mathcal{D}^{OST} \\ \mathcal{D}^{SST} &\rightarrow S_D^{SST} \cdot \mathcal{D}^{SST} \end{aligned} \quad (8.17)$$

The candidate likelihood from Equation (7.25) is given by

$$\begin{aligned} L(m, m_{m_{\ell^+ D_s^-}}, ct^*, \sigma_{ct^*}, T, \mathcal{D}) &= \sum_i f_i \mathcal{P}^i(m) \cdot \mathcal{P}^i(m_{\ell^+ D_s^-}) \cdot \mathcal{P}_T^i(ct^*, |\mathcal{D}, \sigma_{ct^*}) \\ &\quad \cdot \mathcal{P}^i(\mathcal{D}) \cdot \mathcal{P}^i(\sigma_{ct^*}), \end{aligned} \quad (8.18)$$

The PDFs describing the signal and backgrounds in the $\mathcal{P}^i(m)$, $\mathcal{P}^i(m_{\ell^+ D_s^-})$, and $\mathcal{P}^i(ct^* | \sigma_{ct^*})$ spaces have been obtained in Sections 5.3.4 and 5.3.7, together with their corresponding contributions. The modifications to the PDFs in order to incorporate the tagging information are described in Sections 8.1 and 8.2. The unconditional PDFs in the σ_{ct^*} space are derived as described in Section 5.3.6. The unconditional PDFs for the SST and OST dilutions are obtained using the same procedure as detailed in Section 7.4. The \mathcal{D} distributions for the SSKT behavior in case of the D^- reflection background are obtained from the $B \rightarrow \ell^+ D^- X$ samples. The $\mathcal{P}^i(\mathcal{D})$ PDFs for the signal and combinatorial background in the $\ell^+ D_s^-$ modes are shown in Figure 8.1 for the SSKT (top) and the OST (bottom) in B triggers.

We now have access to all the PDFs needed for the likelihood computation in Equation (8.18). We next determine the tagging efficiencies and dilutions for the various components of the B_s^0 samples. As mentioned above, the SSKT is only applied for candidates satisfying $m_{\ell^+ D_s^-} > 3.5 \text{ GeV}/c^2$. First, the tagging efficiency ϵ^{comb} and the average dilution $\mathcal{D}_{\text{avg}}^{\text{comb}}$ for the combinatorial background are measured from the D_s^- mass sideband candidates, separately for OST and SSKT. These are summarized in Tables 8.2 and 8.3 for OST, and Tables 8.4 and 8.5 for SSKT. The SSKT efficiency and time-averaged dilution for the D^- reflection background are determined from $B \rightarrow \ell^+ D^- X$ data, and shown in Table 8.6. The SSKT performance for the physics background candidates surviving the $m_{\ell^+ D_s^-} > 3.5 \text{ GeV}/c^2$ requirement is hypothesized to be the same as the B_s^0 signal. The systematic uncertainty arising from this assumption is addressed in Section 8.5.2. The tagging efficiencies for the B_s^0 signal are computed for the SSKT and the OST, and displayed in Table 8.7. The SSKT efficiencies for the D^- reflection are higher as compared to those observed for the signal B_s^0 candidates. This is mainly due to the presence of additional tracks around $B^{+,0}$ mesons from excited B^{**} decays.

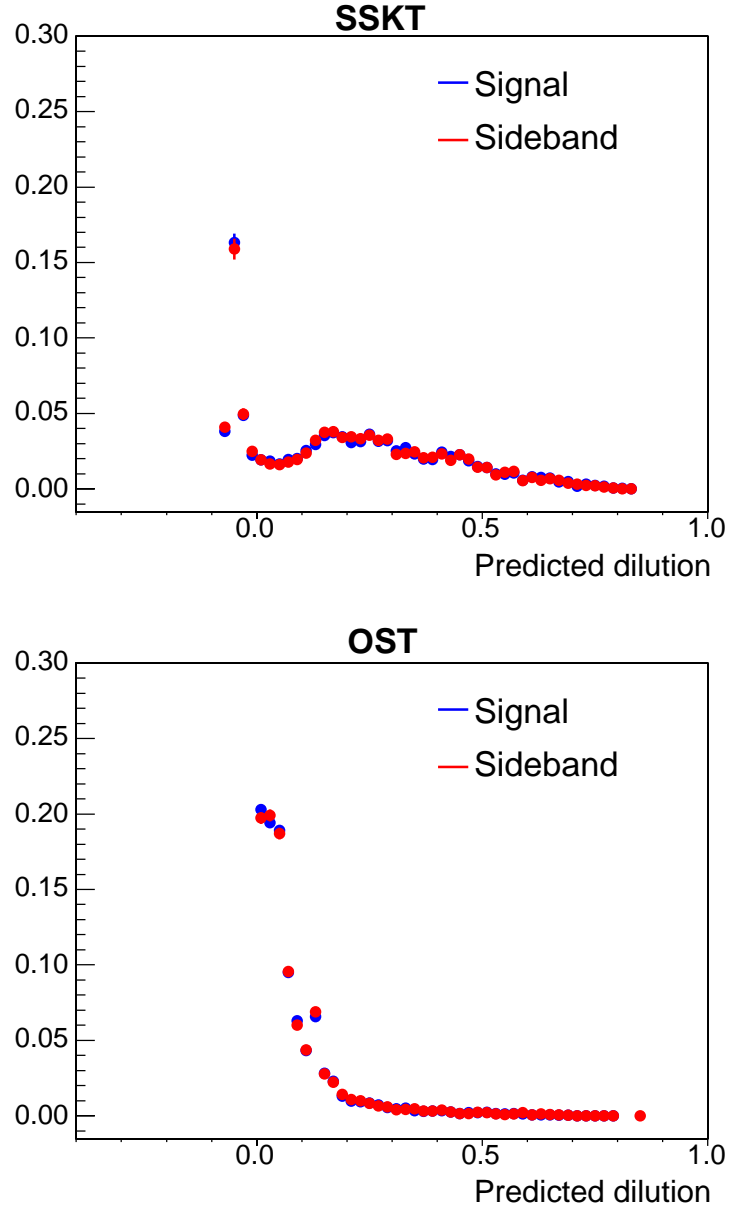


Figure 8.1: Dilution PDFs for the SSKT (top) and combined OST (bottom) algorithms for B triggers in the ℓD_s channels in the Period 1 data.

Decay	$\phi\pi^-$	$K^{*0}K^-$	$\pi^+\pi^-\pi^-$
Samples	Period 1		
$\mu D_s, B$ trigger	0.959	0.958	0.965
$e D_s, B$ trigger	0.953	0.963	0.964
$\mu D_s, D$ trigger	0.946	0.959	0.970
$e D_s, D$ trigger	0.936	0.962	0.966
$\mu D_s, \ell + SVT$ trigger	0.986	0.976	0.983
$e D_s, \ell + SVT$ trigger	0.968	0.960	0.971
	Period 2		
$\mu D_s, B$ trigger	0.963	0.963	0.968
$e D_s, B$ trigger	0.964	0.970	0.969
$\mu D_s, D$ trigger	0.969	0.978	0.965
$e D_s, D$ trigger	0.970	0.968	0.971
$\mu D_s, \ell + SVT$ trigger	0.986	0.979	0.980
$e D_s, \ell + SVT$ trigger	0.968	0.978	0.986
	Period 3		
$\mu D_s, B$ trigger	0.965	0.961	0.974
$e D_s, B$ trigger	0.979	0.968	0.976
$\mu D_s, D$ trigger	0.943	0.967	0.958
$e D_s, D$ trigger	0.977	0.970	0.966
$\mu D_s, \ell + SVT$ trigger	0.989	0.983	0.982
$e D_s, \ell + SVT$ trigger	0.954	0.993	0.971

Table 8.2: Combined OST efficiencies for combinatorial background in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates in the three data-taking periods.

Decay Samples	$\phi\pi^-$	$K^{*0}K^-$	$\pi^+\pi^-\pi^-$
	Period 1		
$\mu D_s, B$ trigger	+0.017	+0.005	+0.025
$e D_s, B$ trigger	+0.044	+0.024	+0.032
$\mu D_s, D$ trigger	+0.000	-0.026	+0.027
$e D_s, D$ trigger	+0.094	-0.037	-0.019
$\mu D_s, \ell + SVT$ trigger	-0.060	-0.037	+0.042
$e D_s, \ell + SVT$ trigger	+0.077	-0.046	+0.007
	Period 2		
$\mu D_s, B$ trigger	+0.016	+0.017	+0.014
$e D_s, B$ trigger	+0.004	+0.011	+0.006
$\mu D_s, D$ trigger	-0.106	+0.067	+0.020
$e D_s, D$ trigger	-0.108	-0.031	-0.023
$\mu D_s, \ell + SVT$ trigger	+0.024	+0.010	+0.024
$e D_s, \ell + SVT$ trigger	+0.076	+0.129	+0.056
	Period 3		
$\mu D_s, B$ trigger	+0.063	+0.032	+0.046
$e D_s, B$ trigger	+0.003	-0.019	+0.005
$\mu D_s, D$ trigger	-0.074	+0.067	+0.077
$e D_s, D$ trigger	-0.126	+0.015	+0.000
$\mu D_s, \ell + SVT$ trigger	-0.019	+0.049	+0.006
$e D_s, \ell + SVT$ trigger	+0.088	+0.100	+0.059

Table 8.3: Combined OST dilutions for combinatorial background in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates in the three data-taking periods.

Decay Samples	$\phi\pi^-$	$K^{*0}K^-$	$\pi^+\pi^-\pi^-$
	Period 1		
$\mu D_s, B$ trigger	0.712	0.715	0.718
$e D_s, B$ trigger	0.678	0.716	0.723
$\mu D_s, D$ trigger	0.671	0.745	0.764
$e D_s, D$ trigger	0.775	0.761	0.762
$\mu D_s, \ell + SVT$ trigger	0.800	0.737	0.828
$e D_s, \ell + SVT$ trigger	0.692	0.660	0.782
	Period 2		
$\mu D_s, B$ trigger	0.728	0.716	0.746
$e D_s, B$ trigger	0.733	0.729	0.736
$\mu D_s, D$ trigger	0.736	0.768	0.771
$e D_s, D$ trigger	0.695	0.767	0.807
$\mu D_s, \ell + SVT$ trigger	0.742	0.844	0.796
$e D_s, \ell + SVT$ trigger	0.774	0.746	0.772
	Period 3		
$\mu D_s, B$ trigger	0.667	0.707	0.734
$e D_s, B$ trigger	0.685	0.756	0.759
$\mu D_s, D$ trigger	0.600	0.787	0.745
$e D_s, D$ trigger	0.642	0.696	0.757
$\mu D_s, \ell + SVT$ trigger	0.791	0.760	0.789
$e D_s, \ell + SVT$ trigger	0.770	0.762	0.780

Table 8.4: SSKT efficiencies for combinatorial background in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates in the three data-taking periods.

Decay	$\phi\pi^-$	$K^{*0}K^-$	$\pi^+\pi^-\pi^-$
Samples	Period 1		
$\mu D_s, B$ trigger	+0.082	+0.004	+0.001
$e D_s, B$ trigger	+0.026	+0.090	+0.021
$\mu D_s, D$ trigger	+0.209	-0.047	+0.010
$e D_s, D$ trigger	+0.032	+0.125	+0.011
$\mu D_s, \ell + SVT$ trigger	+0.012	+0.196	+0.084
$e D_s, \ell + SVT$ trigger	-0.154	-0.131	+0.058
	Period 2		
$\mu D_s, B$ trigger	+0.021	+0.020	-0.000
$e D_s, B$ trigger	+0.078	-0.029	+0.022
$\mu D_s, D$ trigger	-0.014	-0.045	+0.067
$e D_s, D$ trigger	-0.125	+0.000	+0.113
$\mu D_s, \ell + SVT$ trigger	+0.004	-0.032	+0.019
$e D_s, \ell + SVT$ trigger	+0.094	+0.113	+0.110
	Period 3		
$\mu D_s, B$ trigger	+0.025	-0.049	+0.017
$e D_s, B$ trigger	+0.024	+0.009	-0.047
$\mu D_s, D$ trigger	+0.142	+0.096	-0.079
$e D_s, D$ trigger	-0.001	+0.130	+0.115
$\mu D_s, \ell + SVT$ trigger	-0.069	-0.035	+0.021
$e D_s, \ell + SVT$ trigger	+0.065	-0.081	-0.039

Table 8.5: SSKT dilutions for combinatorial background in $B_s^0 \rightarrow \ell^+ D_s^- X$ candidates in the three data-taking periods.

Decay	ϵ [%]	\mathcal{D} [%]
Samples	Period 1	
$\mu D_s, B$ trigger	62.2 ± 0.3	5.0 ± 0.8
$e D_s, B$ trigger	62.8 ± 0.4	4.4 ± 1.0
$\mu D_s, D$ trigger	62.9 ± 0.5	1.5 ± 1.4
$e D_s, D$ trigger	66.2 ± 0.7	6.1 ± 1.9
$\mu D_s, \ell + SVT$ trigger	68.1 ± 0.5	4.2 ± 1.4
$e D_s, \ell + SVT$ trigger	63.2 ± 0.6	3.6 ± 1.6
	Period 2	
$\mu D_s, B$ trigger	59.3 ± 0.3	3.4 ± 1.0
$e D_s, B$ trigger	59.4 ± 0.5	2.5 ± 1.3
$\mu D_s, D$ trigger	60.8 ± 0.7	3.7 ± 1.8
$e D_s, D$ trigger	62.7 ± 0.9	4.8 ± 2.4
$\mu D_s, \ell + SVT$ trigger	64.6 ± 0.6	3.5 ± 1.7
$e D_s, \ell + SVT$ trigger	61.3 ± 0.8	8.3 ± 2.1
	Period 3	
$\mu D_s, B$ trigger	60.8 ± 0.5	1.6 ± 1.5
$e D_s, B$ trigger	60.8 ± 0.7	2.5 ± 1.9
$\mu D_s, D$ trigger	63.4 ± 1.0	5.0 ± 2.6
$e D_s, D$ trigger	64.6 ± 1.3	8.6 ± 3.6
$\mu D_s, \ell + SVT$ trigger	65.3 ± 0.7	6.4 ± 2.0
$e D_s, \ell + SVT$ trigger	61.1 ± 1.0	6.6 ± 2.7

Table 8.6: Efficiencies and dilutions for the SSKT as measured in $B^{+,0} \rightarrow \ell^+ D^- X$ samples. This represents the efficiencies and dilutions for the SSKT for the D^- reflection background in the $B_s^0 \rightarrow \ell^+ D_s^- X$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$ mode.

Parameter	Fit value (0d)	Fit value (0h)	Fit value (0i)
$\epsilon_{\text{SSKT}}[\%]$	53.83 ± 0.50	53.26 ± 0.44	55.03 ± 0.62
$\epsilon_{\text{OST}}[\%]$	95.94 ± 0.12	96.30 ± 0.11	96.58 ± 0.15

Table 8.7: SSKT and OST tagging efficiencies on data for B_s^0 signal candidates. Tagging efficiencies for the SSKT case are computed from candidates with $m_{\ell^+ D_s^-} > 3.5 \text{ GeV}/c^2$.

8.4 Amplitude Method in Toy Monte Carlo

Once the PDFs describing the signal and backgrounds in the D_s^- mass, $m_{\ell+D_s^-}$, ct^* , σ_{ct^*} and tagging spaces have been determined, we can model the B_s^0 samples using toy Monte Carlo. Such an implementation enables simulation of a realistic description of the data, and allows validation checks of the fit framework. High statistics toy Monte Carlo samples are particularly useful to study the systematic errors associated with the current analysis (see Section 8.5), as they allow access to various mixing frequencies not present in the data, together with providing an approach to decouple systematic errors from statistical fluctuations.

We construct toy Monte Carlo samples describing the B_s^0 data in various discriminating variables, tagging efficiencies and dilutions, as well as measurement resolutions. The fit framework designed for the measurement of \mathcal{A} in data is then applied to the toy Monte Carlo samples. As a result of the amplitude procedure, a measurement of the amplitude $\mathcal{A}(\Delta m_s)$, together with the error associated with its measurement $\sigma_{\mathcal{A}}(\Delta m_s)$, is obtained. As an example, the amplitude method is applied to an ensemble of toy Monte Carlo experiments generated with $\Delta m_s = 17.75 \text{ ps}^{-1}$ with the same statistical power as the data. The frequency domain is discretized in increments of 0.25 ps^{-1} , and the pairs $[\mathcal{A}(\Delta m_s), \sigma_{\mathcal{A}}(\Delta m_s)]$ obtained for each Δm_s hypothesis in the range $[0, 30] \text{ ps}^{-1}$ are displayed for each value of the investigated frequency in an “amplitude scan” as shown in Figure 8.2. Each point in the amplitude scan represents the measured value of $\mathcal{A}(\Delta m_s)$, while the error bars correspond to the statistical error $\sigma_{\mathcal{A}}(\Delta m_s)$ at the hypothesized Δm_s value. The yellow band signifies the $1.645 \cdot \sigma_{\mathcal{A}}(\Delta m_s)$ deviation from the mean amplitude value $\mathcal{A}(\Delta m_s)$. Thus, utilizing Equation (1.49), the frequencies that are below the point where the yellow band intersects the $\mathcal{A} = 1$ line are excluded at 95% confidence level. The dotted curve indicates $1.645 \cdot \sigma_{\mathcal{A}}(\Delta m_s)$ computed at each Δm_s value. Therefore, the point in the amplitude scan Δm_s^{sens} where the dotted curve intersects the $\mathcal{A} = 1$ line indicates the statistical sensitivity of the amplitude scan (see Equation (1.50)). The amplitude values below Δm_s^{sens} ($= 19.9 \text{ ps}^{-1}$ in this toy Monte Carlo example) are more precisely measured as compared to the Δm_s values above Δm_s^{sens} . Thus, the precision of the amplitude measurement diminishes with increasing values of Δm_s , as expected from Equation (1.51) defining the significance of the measurement. The amplitude values above the sensitivity are, in general, so imprecisely measured that the distinction between $\mathcal{A} = 0$ and $\mathcal{A} = 1$ is no longer noticeable. As described in Section 1.4.1, statistical fluctuations in the measured amplitude $\mathcal{A}(\Delta m_s)$ can result in more or less favorable exclusion regions of Δm_s . The exclusion conditions thus depend on the measured amplitude values, and consequently the true oscillation frequency. Therefore, the 95% confidence level exclusion limit obtained from toy Monte Carlo is meaningless. However, the derived toy Monte Carlo sensitivity, which corresponds to the largest Δm_s that can be excluded given the statistical significance of the measurement, is a meaningful quantity. Thus, the toy Monte Carlo example

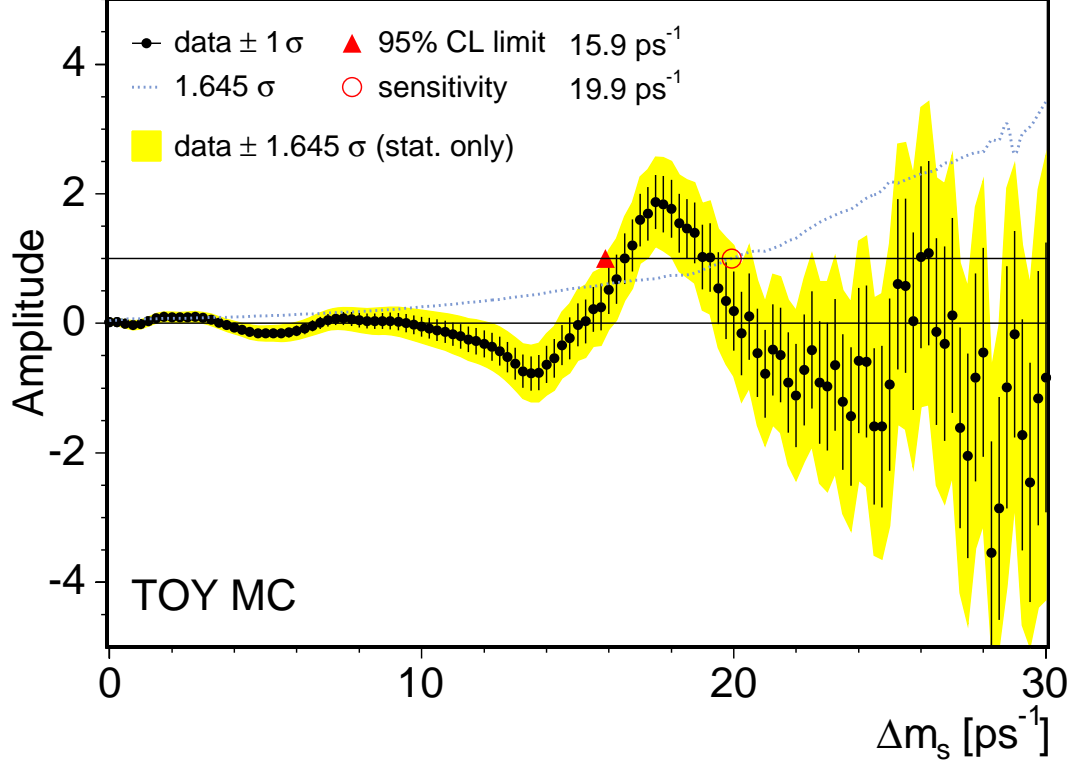


Figure 8.2: Example of an amplitude scan on a toy Monte Carlo simulating the data generated with $\Delta m_s = 17.75\text{ ps}^{-1}$. Only statistical errors are shown here.

indicates that the estimated statistical sensitivity of the semileptonic B_s^0 samples for the B_s^0 - \bar{B}_s^0 oscillations analysis is $\sim 20\text{ ps}^{-1}$.

We utilize a high statistics toy Monte Carlo implementation of the amplitude technique to study the systematic errors associated with the measurement of \mathcal{A} . This is described in the next section.

8.5 Systematic Uncertainties

The procedure for the determination of systematic uncertainties in the amplitude method is discussed in Ref. [27]. An important point to note is that at a given Δm_s value in the amplitude scan, as a result of the systematic variation, the deviations in the amplitude and its statistical error are correlated. Hence, in order to correctly

evaluate the systematic uncertainty, the following prescription is employed [27]:

$$\sigma_{\mathcal{A}}^{syst} = \Delta\mathcal{A} + (1 - \mathcal{A}) \cdot \frac{\Delta\sigma_{\mathcal{A}}}{\sigma_{\mathcal{A}}}, \quad (8.19)$$

where $\Delta\mathcal{A}$ and $\Delta\sigma_{\mathcal{A}}$ respectively denote the differences observed in the values of the amplitude and the values of its systematic uncertainty as a consequence of the variation of the systematic source.

For the purpose of evaluating systematic errors, we utilize a toy Monte Carlo strategy. The motivation for using a toy Monte Carlo approach is two fold. First, the generation of a large ensemble of toy Monte Carlo samples allows disentangling systematic contributions from statistical effects. Second, the evaluation of systematic errors utilizing Equation (8.19) assumes that the values of $\Delta\mathcal{A}$ and $\Delta\sigma_{\mathcal{A}}$ are extracted at a particular value of Δm_s with the oscillating signal at the same frequency. Therefore, the data cannot be used for this purpose.

We perform a construction of toy Monte Carlo models describing data in terms of observed characteristics of signal, backgrounds, tagging, and measurement errors. To evaluate the contribution from systematic uncertainties in the amplitude scan at a given Δm_s value, we generate a large ensemble of toy Monte Carlo samples using the default configuration of the data at that Δm_s value. The toy Monte Carlo samples are then fit using the nominal settings to obtain the values of the amplitude ($\mathcal{A}^{\text{default}}$) and its statistical error ($\sigma_{\mathcal{A}}^{\text{default}}$). Then, for each source of systematic error under investigation, the fit parameters are modified to take into account the corresponding systematic deviation. The samples are again analyzed with the updated fit configuration to obtain the new values of the amplitude (\mathcal{A}^{new}) and its statistical error ($\sigma_{\mathcal{A}}^{\text{new}}$). Using Equation (8.19), we construct the systematic error for each sample i as follows:

$$\sigma_{\mathcal{A}}^{syst}|_i = (\mathcal{A}^{\text{new}} - \mathcal{A}^{\text{default}}) + (1 - \mathcal{A}^{\text{default}}) \cdot \frac{\sigma_{\mathcal{A}}^{\text{new}} - \sigma_{\mathcal{A}}^{\text{default}}}{\sigma_{\mathcal{A}}^{\text{default}}}. \quad (8.20)$$

The results $\sigma_{\mathcal{A}}^{syst}|_i$ are then plotted to obtain a distribution of systematic errors for each source. The associated systematic uncertainties are then obtained from the mean values of the corresponding distributions.

To understand systematic uncertainties arising from the non-Gaussian detector resolution and a non-zero value of $\Delta\Gamma_s/\Gamma_s$, we perform toy Monte Carlo studies in a different manner as compared to the ones described above. The values of $\mathcal{A}^{\text{default}}$ and its error $\sigma_{\mathcal{A}}^{\text{default}}$ for the default data and fit configuration are obtained in the same manner as outlined above. However, for the two studies in question, we generate a separate set of toy Monte Carlo samples with the systematic effect introduced in the configuration. The fits to the samples are repeated with the nominal settings to obtain the updated values of the amplitude (\mathcal{A}^{new}) and its statistical error ($\sigma_{\mathcal{A}}^{\text{new}}$). The associated systematic error is then calculated using Equation (8.20).

The systematic uncertainties are expected to have a smooth functional dependences on Δm_s across the range of scanned frequencies. Therefore, we evaluate the

systematic errors at some of the Δm_s values in the probed region, and interpolate to other frequencies in the amplitude scan using a functional dependence. In the following, we describe the various sources of systematic uncertainties affecting the amplitude measurements. The sources of systematic error listed in the case of the Δm_d measurement (see Section 7.6) are investigated together with additional sources relevant for the Δm_s amplitude measurement.

8.5.1 Tagging Dilution Scale Factors

The dilution scale factors for the same side and the combined opposite side tags are fixed in the nominal fit. The uncertainties on the measured amplitude related to the dilution scale factors are evaluated by modifying them in the fit taking into account the errors on the calibrated values (see Tables 7.3 and 8.1). The systematic error is obtained by taking the largest variation in the obtained amplitude value induced by coherent increase or decrease of the tagging dilution scale factors.

8.5.2 Same Side Tagging Dilution for Physics Background

The presence of physics background in the B_s^0 samples can alter the same side tagging performance. This results from the additional B daughters produced in physics background events, apart from the lepton and the D_s^- meson, which can be mistaken as a tagging track by the same side kaon tagging algorithm. For this purpose, we apply the SSKT algorithm only on the candidates satisfying $m_{\ell^+ D_s^-} > 3.5 \text{ GeV}/c^2$ where the physics background originating from $B \rightarrow DDX$ decays is reduced to a minimal level ($< 2\%$ of its fraction in the $2.0 < m_{\ell^+ D_s^-} < 5.5$ region). The remaining physics background behaves similar to the signal in the same side tag performance, and is hence assigned the same dilution. For estimating the corresponding systematic uncertainty, we have varied the dilution for the physics backgrounds by an absolute amount of $\pm 4\%$ about the default values.

8.5.3 D^- Reflection Background

One of the three B_s^0 decay modes analyzed in this measurement is $B_s^0 \rightarrow \ell^+ D_s^- X$, with $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$. This mode suffers from the D^- reflection background (see Sections 5.1 and 5.3.1). Two sources of systematic errors can arise from this background. The contribution of the D^- reflection in the above mentioned sample is determined via a fit to the D_s^- invariant mass. The statistical errors in the determination of this fraction are then used to evaluate the corresponding systematic error. Another source of systematic error from the D^- reflection background is its dilution used for the same side kaon tag. We alter the SSKT dilutions measured on the $B^{+,0} \rightarrow \ell^+ D^- X$ samples within their statistical uncertainties, and assign the associated systematic error.

8.5.4 Fake Lepton Background

Three sources of systematic errors arise from fake lepton background, as described in Section 7.6.1. In summary, the fake lepton dilution, fraction, and ct^* shape are considered for systematic uncertainty evaluation. The dilution of the fake lepton background in the default fit is fixed to zero, and systematic errors are evaluated by varying its dilution between \mathcal{D}_{signal} and $-\mathcal{D}_{signal}$, where \mathcal{D}_{signal} is the predicted dilution for the signal as given by the OST and the SST algorithms. The fractions of the fake lepton contributions to the $\ell^+ D_s^-$ samples are determined as a fit parameter in the simultaneous fit to the D_s^- mass and $m_{\ell^+ D_s^-}$ spaces (see Sections 5.3.3 and 5.3.4), and the statistical uncertainties in the fractions are used to compute the corresponding systematic error. The ct^* shape for the fake lepton background is obtained using the fake lepton sample. The systematic uncertainty due to the ct^* description of the fake lepton background is estimated using the statistical uncertainties of the parameters describing its shape.

8.5.5 Physics Background Fractions

The fractions of the various sources of physics backgrounds are estimated using Monte Carlo samples (see Section 5.3.2). Poorly measured branching ratios involved in the determination of these backgrounds lead to large uncertainties in the contributions from these sources. Therefore, we assign a 30-40% fractional uncertainty in the contributions from the physics backgrounds, corresponding to the uncertainties in their determination from Monte Carlo. The systematic uncertainty is evaluated using toy Monte Carlo samples with the fractions of the physics background altered to reflect this uncertainty.

8.5.6 $\xi(t^*)$ Determination from Monte Carlo

The efficiency functions $\xi(t^*)$ describe the effect of trigger and reconstruction biases on the pseudo-proper decay time for the signal B_s^0 mesons (see Section 5.3.5). The related sources of systematic errors originating from the derivation of the efficiency functions is detailed in Section 7.6.4. Two sources of systematic errors arise from the $\xi(t^*)$ extraction from signal Monte Carlo samples. The $\xi(t^*)$ shapes are obtained assuming the world average value of the B_s^0 meson lifetime [13] in the Monte Carlo simulations. We alter the input B_s^0 lifetime value in the Monte Carlo samples, and derive the updated $\xi(t^*)$ function for the systematic error evaluation. The second source of systematic error due to $\xi(t^*)$ arises from the difference in data and Monte Carlo simulation in terms of reproducing the correlation between the online and offline impact parameters. An additional smearing of $\sim 12 \mu\text{m}$ is applied in the Monte Carlo samples before the SVT trigger confirmation, and the re-derived efficiency functions are used to estimate the systematic error.

8.5.7 Combinatorial Background Characterization

The fractions and shapes of the combinatorial background in the $m_{\ell D}$ and ct^* spaces are investigated as sources of systematic errors, as described in Section 7.6.5. The fractions of the combinatorial background are varied within their statistical uncertainties, and the associated systematic errors are assessed. The descriptions of the combinatorial background in the $m_{\ell D}$ and ct^* variables are obtained from the candidates in the D_s^- mass sideband. We shift the sideband window for the selection of background candidates by ± 50 MeV/ c^2 in the D_s^- mass, and evaluate the systematic error utilizing the updated templates for the background description.

8.5.8 Fixing $\tau(B_s^0)$ to the World Average Value

The lifetime of the B_s^0 meson is fixed to its world average value [13] in the nominal fit. We alter the lifetime within the errors in the world average value, and generate toy Monte Carlo samples with the updated lifetimes. The toy Monte Carlo samples are then fit with the nominal fit configuration to estimate the systematic error.

8.5.9 σ_{t^*} Calibration Uncertainty

The error on σ_{t^*} as returned by the CTVMFT vertex fit underestimates the true error, and a scaling correction factor is applied to σ_{t^*} for each $\ell^+ D_s^-$ candidate (see Section 5.3.5). The systematic error is estimated by altering the scale factors by ± 0.06 corresponding to the uncertainty in the scale factor determination from the $D^- + \text{track}$ studies. A set of toy Monte Carlo samples are generated with the altered values, and then fit with nominal settings to obtain the associated systematic error.

8.5.10 Composition of $B_s^0 \rightarrow \ell^+ D_s^- X$ Decays

This source of systematic error is analogous to the uncertainty due to imperfect knowledge of the sample composition of the ℓD sample. In the B_s^0 samples, the branching ratios of $B_s^0 \rightarrow \ell^+ D_s^- X$ and $B_s^0 \rightarrow \ell D_s^* X$, $D_s^* \rightarrow D_s \gamma / \pi^0$ are not well known [13]. The relative contributions between the two sources can be constrained using the fully reconstructed modes $B_s^0 \rightarrow D_s^- \pi^+$ and $B_s^0 \rightarrow D_s^* \pi$, where the latter is present in the fully reconstructed $B_s^0 \rightarrow D_s^- \pi^+$ sample located in the mass region just below the nominal B_s^0 mass. The agreement between the expectation in the signal Monte Carlo and the relative fraction obtained in the data is at the level of 10-20%. Hence, the relative contributions have been varied by 20%, and the associated κ -factor and $\xi(t^*)$ distribution for the signal description have been re-derived. The systematic error is estimated by finding the difference between this result and the default configuration.

8.5.11 Non-Gaussian Detector Resolution

The nominal fit model assumes that the detector resolution for the pseudo-proper decay time is a single Gaussian (see Section 5.3.5). The width of the Gaussian distribution is obtained from the ct^* error returned for each candidate by the CTVMFT fit, after multiplication by a scaling factor obtained via the proper time calibration study (see Section 5.3.5). Studies on the high statistics proper time calibration samples indicate that a double Gaussian plus a symmetric exponential provide a better description of the proper time resolution. The associated systematic uncertainty is obtained by generating toy Monte Carlo samples using the proper time shape obtained from calibration samples, and fitting with the nominal settings.

8.5.12 Non-vanishing Value of $\Delta\Gamma_s/\Gamma_s$

The signal PDF used in the B_s^0 mixing measurement assumes that the lifetime difference between the heavy and light B_s^0 mass eigenstates is zero. Thus, the signal PDF is characterized by a simple exponential modulated by a fast oscillating term. In the case of a non-negligible value of $\Delta\Gamma_s/\Gamma_s$, the time evolution of the B_s^0 - \bar{B}_s^0 system gets altered as seen in Equations (1.35) and (1.36):

$$\begin{aligned} \frac{e^{-\Gamma_s t}}{2} [1 \pm \cos(\Delta m_s t)] &\rightarrow \frac{e^{-\Gamma_s t}}{2} \left(1 - \frac{\Delta\Gamma_s^2}{4\Gamma_s^2}\right) \left[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) \pm \cos(\Delta m_s t) \right] \\ &\rightarrow \frac{e^{-\Gamma_s t}}{2} \left[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) \pm \cos(\Delta m_s t) \right], \end{aligned}$$

where we have neglected the contribution from the $\left[\frac{\Delta\Gamma_s}{\Gamma_s}\right]^2$ term. Effectively, the signal description is thus composed of two components with different τ values, and the oscillations get damped in time. We study the effect of having a non-zero $\Delta\Gamma$ between the two B_s^0 states utilizing a toy Monte Carlo approach. We use a conservative value of $\Delta\Gamma_s/\Gamma_s = 0.2$ for estimating the systematic error.

8.5.13 Summary of the Systematic Uncertainties

The evaluation of systematic errors is repeated at several values of Δm_s . The magnitudes of the systematic uncertainties $\sigma_A^{\text{syst}}(\Delta m_s)$ from different sources are added in quadrature for a given oscillation frequency, and the total systematic uncertainty at each value of Δm_s is obtained. The resulting values are then interpolated using a smooth functional dependence to obtain the contribution from systematic uncertainties at all probed frequencies. Tables 8.8 and 8.9 display the contributions from various sources of systematic uncertainties. Figure 8.3 shows the variation of the various systematic uncertainties together with the total systematic uncertainty with respect to Δm_s . Systematic error originating from the uncertainties in the tagging

dilution scale factors dominates the total systematic uncertainty, where a large contribution comes from the dilution scale factor for the same side kaon tag.

Source of Systematic Error	$\sigma_A^{\text{syst}}(\Delta m_s)$ for $\Delta m_s =$			
	0 ps ⁻¹	5 ps ⁻¹	10 ps ⁻¹	15 ps ⁻¹
Dilution scale factors	0.121	0.121	0.122	0.122
σ_{t^*} calibration	0.000	0.015	0.038	0.055
Combinatorial background	0.020	0.023	0.027	0.038
Fraction of physics background	0.047	0.037	0.034	0.032
Dilution of physics background	0.000	0.007	0.000	0.000
Fake Lepton background dilution	0.029	0.008	0.008	0.002
Fake Lepton background fraction	0.025	0.025	0.024	0.022
Fake Lepton background ct^* shape	0.001	0.004	0.005	0.003
Sample composition	0.003	0.005	0.007	0.009
t^* efficiency	0.001	0.001	0.001	0.001
$\tau(B_s^0)$	0.000	0.006	0.008	0.008
$\Delta\Gamma_s/\Gamma_s$	0.000	0.001	0.001	0.000
Non-Gaussian detector resolution	0.010	0.024	0.037	0.034
D^+ reflection	0.003	0.001	0.000	0.000
Total systematic uncertainty	0.137	0.134	0.143	0.148
Statistical uncertainty	0.023	0.068	0.154	0.340

Table 8.8: Summary of the systematic uncertainties on the amplitude \mathcal{A} at $\Delta m_s = 0, 5, 10$ and 15 ps^{-1} .

Source of Systematic Error	$\sigma_A^{\text{syst}}(\Delta m_s)$ for $\Delta m_s =$			
	17 ps ⁻¹	20 ps ⁻¹	25 ps ⁻¹	30 ps ⁻¹
Dilution scale factors	0.121	0.121	0.119	0.119
σ_{t^*} calibration	0.058	0.061	0.064	0.066
Combinatorial background	0.039	0.047	0.063	0.072
Fraction of physics background	0.031	0.032	0.031	0.030
Dilution of physics background	0.000	0.000	0.000	0.000
Fake Lepton background dilution	0.004	0.003	0.011	0.014
Fake Lepton background fraction	0.021	0.019	0.014	0.010
Fake Lepton background ct^* shape	0.003	0.002	0.001	0.004
Sample composition	0.010	0.010	0.015	0.016
t^* efficiency	0.002	0.002	0.004	0.004
$\tau(B_s^0)$	0.008	0.007	0.004	0.004
$\Delta\Gamma_s/\Gamma_s$	0.002	0.003	0.003	0.007
Non-Gaussian detector resolution	0.030	0.023	0.008	0.003
D^+ reflection	0.000	0.000	0.000	0.000
Total systematic uncertainty	0.149	0.150	0.154	0.158
Statistical uncertainty	0.439	0.610	1.368	2.012

Table 8.9: Summary of the systematic uncertainties on the amplitude \mathcal{A} at $\Delta m_s = 17, 20, 25$ and 30 ps^{-1} .

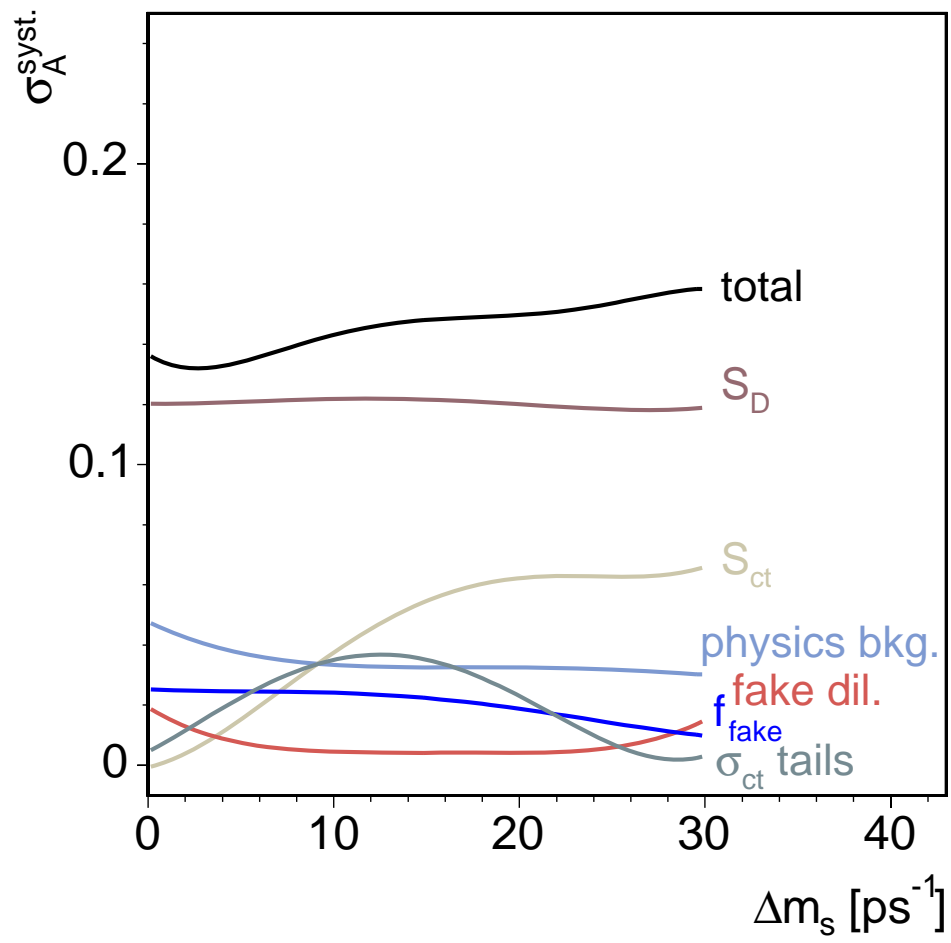


Figure 8.3: Parameterization of the overall systematic uncertainty versus Δm_s .

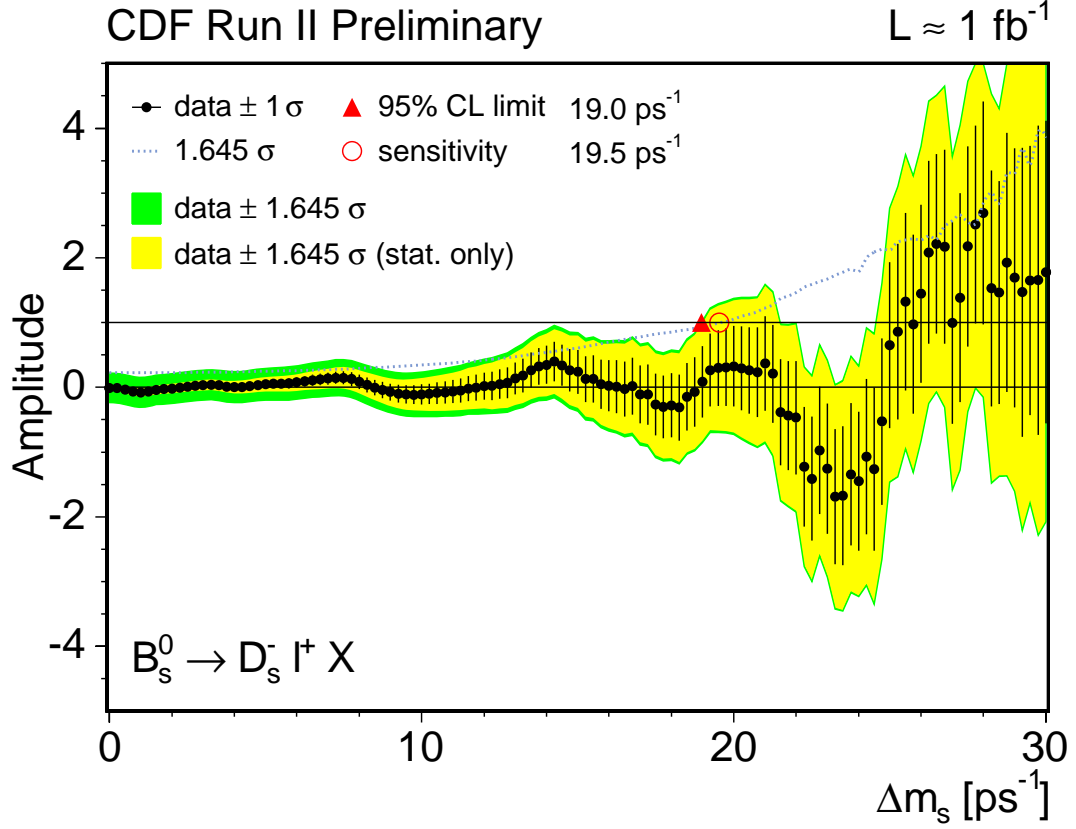


Figure 8.4: Blinded amplitude scan in data for the $B_s^0 \rightarrow \ell^+ D_s^- X$ decays. Statistical and systematic errors are evaluated and displayed.

8.6 Amplitude Scan in Blinded Data

We have by now established the systematic uncertainties involved in the measurement of $\mathcal{A}(\Delta m_s)$. As a first step to applying the amplitude measurement technique on the data, we perform a “blind” analysis by randomizing the flavor tag information. The flavor tagging decisions for the B_s^0 candidates are multiplied by a factor $(-1)^n$, where n is an integer corresponding to the order of the candidates in the B_s^0 samples. The scrambling of flavor tagging information thus allows full characterization of the data except for the values of the measured amplitude. This method further facilitates the validation of the fit framework in an unbiased manner. The relevant information in the amplitude scan, obtained from the blinded data in this fashion, is the measured sensitivity derived from the uncertainty on the measured amplitude $\sigma_{\mathcal{A}}(\Delta m_s)$ at each probed value of Δm_s . The blinded amplitude scan for the data is shown in Figure 8.4. The yellow band corresponds to the $1.645 \cdot \sigma_{\mathcal{A}}^{\text{stat}}(\Delta m_s)$ deviation from the central

amplitude value $\mathcal{A}(\Delta m_s)$, while the green band includes the total systematic error obtained in Section 8.5 added in quadrature to the statistical uncertainty at each value of Δm_s i.e.,

$$\sigma_{\mathcal{A}}^{\text{total}}(\Delta m_s) = \sqrt{[\sigma_{\mathcal{A}}^{\text{stat}}(\Delta m_s)]^2 + [\sigma_{\mathcal{A}}^{\text{syst}}(\Delta m_s)]^2}. \quad (8.21)$$

Thus, the green band is associated with the $1.645 \cdot \sigma_{\mathcal{A}}^{\text{total}}(\Delta m_s)$ variation around the mean measured amplitude \mathcal{A} . The obtained sensitivity of the semileptonic B_s^0 samples is derived from the total uncertainty on the amplitude and the obtained value is $\sim 19.5 \text{ ps}^{-1}$. The sensitivity using only the statistical errors in the amplitude is 19.9 ps^{-1} , similar to the prediction from the toy Monte Carlo example in Section 8.4. The measured amplitude values, and hence the 95% confidence level exclusion limit, are unimportant in this figure.

8.7 Results

We have thus far evaluated the systematic errors associated with the measurement of \mathcal{A} , together with the assessment of the measurement sensitivity using the blinded amplitude scan. The final step in the analysis of $B_s^0\text{-}\bar{B}_s^0$ oscillations in the semileptonic B_s^0 decays is to “unblind” the data by removing the tag randomization, and repeating the amplitude measurements using the correct flavor tag information. The results of the unblinded amplitude scan are displayed in Figure 8.5. The overall sensitivity of the measurement is confirmed in the unblinded data with a sensitivity of $\sim 19.4 \text{ ps}^{-1}$. The Δm_s frequencies below 16.5 ps^{-1} are excluded at 95% or more confidence level. An evidence of $B_s^0\text{-}\bar{B}_s^0$ oscillations is seen at $\sim \Delta m_s = 17.75 \text{ ps}^{-1}$ with an amplitude significance $\mathcal{A}/\sigma_{\mathcal{A}} \sim 2$. The amplitude scans in the three data-taking periods are included in Appendix C. Amplitude scans prepared using only the SSKT or the OST are also shown in Appendix C.

Tables 8.8 and 8.9 summarize the contributions from various sources of systematic uncertainties, and a comparison of the total systematic error with the statistical uncertainty observed at a given value of Δm_s . Figure 8.6 compares the total systematic uncertainty with the statistical error at each probed value of Δm_s . We draw two important conclusions about the uncertainties involved in the amplitude scan determination:

- The systematic error originating from the uncertainty in the tagging dilution scale factors dominates the total systematic uncertainty.
- The evaluated total systematic uncertainties are considerably smaller than the corresponding statistical errors, particularly at the higher probed values of Δm_s .

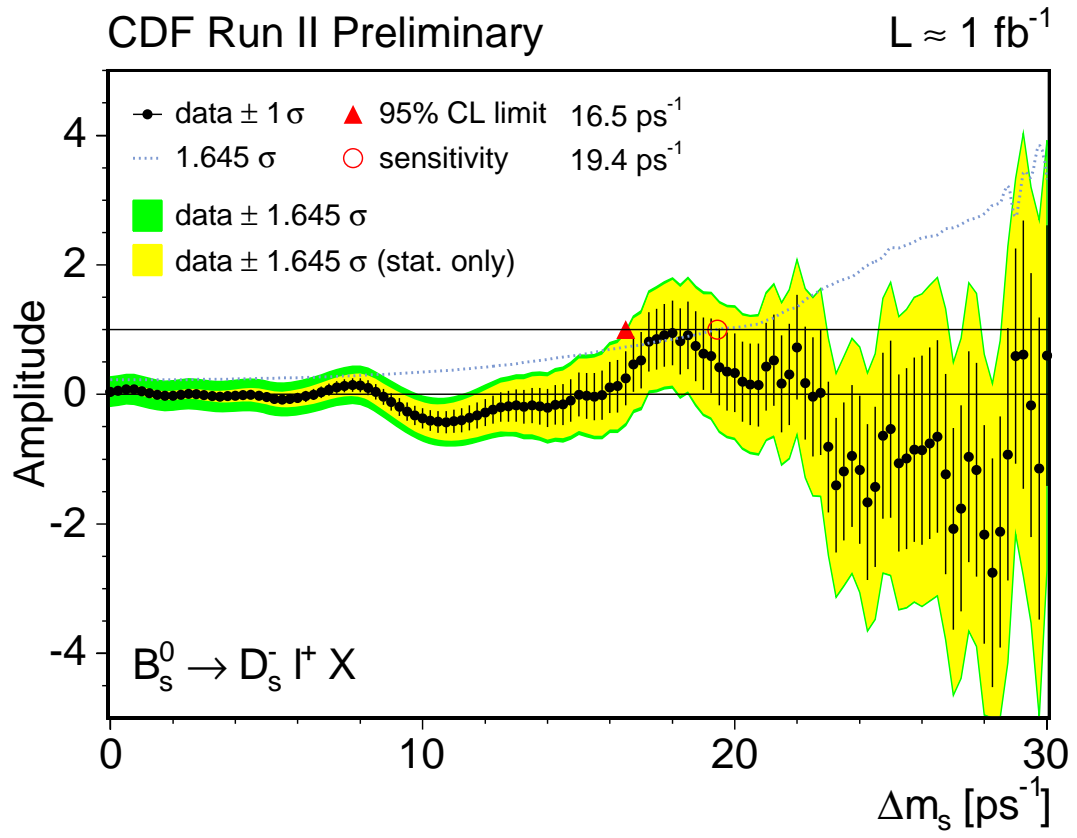


Figure 8.5: Unblinded amplitude scan in data for the $B_s^0 \rightarrow \ell^+ D_s^- X$ decays. The errors displayed include statistical and systematic errors.

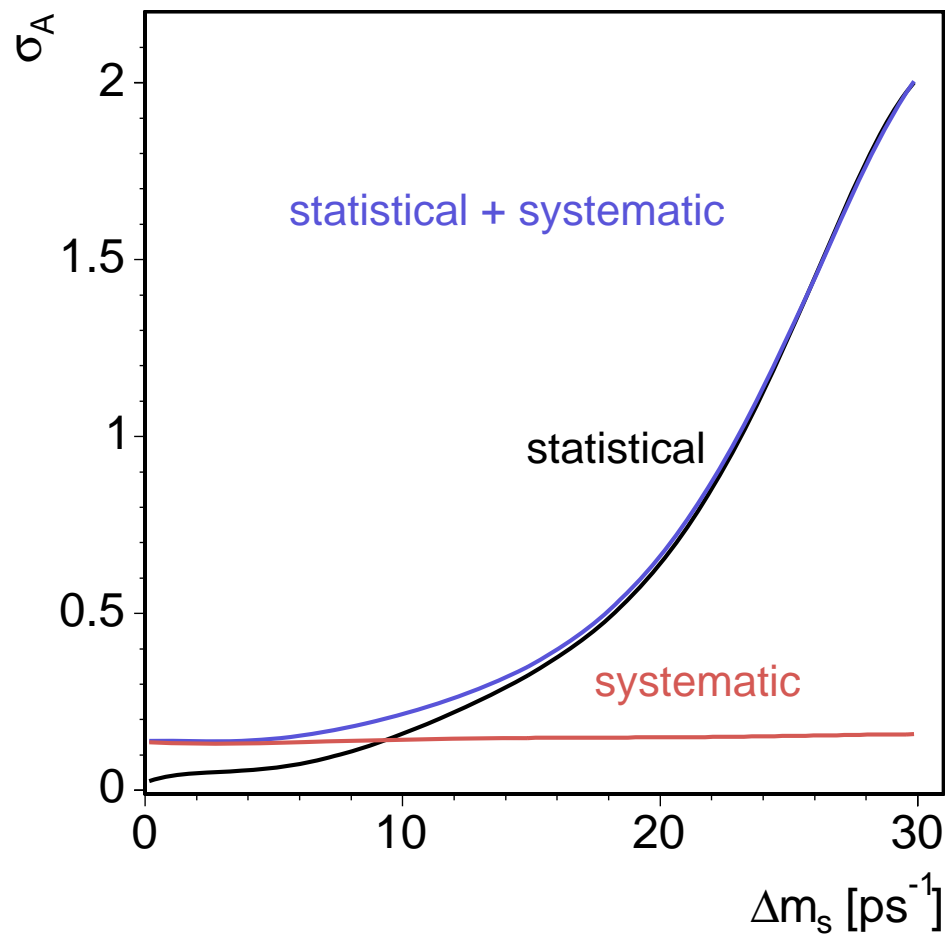


Figure 8.6: Combined overall error including the systematic and statistical errors versus Δm_s .

To assess the profile of the likelihood in the region of evidence around $\Delta m_s \sim 18 \text{ ps}^{-1}$, we construct the logarithm of the ratio of likelihoods:

$$\Lambda = \log \frac{\mathcal{L}^{\mathcal{A}=0}}{\mathcal{L}^{\mathcal{A}=1}(\Delta m_s)}, \quad (8.22)$$

where $\mathcal{L}^{\mathcal{A}=1}(\Delta m_s)$ is the likelihood of the data under the hypothesis that Δm_s is the true oscillation frequency. The likelihood $\mathcal{L}^{\mathcal{A}=0}$ is independent of Δm_s and represents the likelihood for $\mathcal{A} = 0$, which is equivalent to oscillations with $\Delta m_s = \infty$. Figure 8.7 shows the variation of Λ as a function of Δm_s in the range $\Delta m_s \in [15, 20] \text{ ps}^{-1}$. The variable Λ has a minimum around $\sim \Delta m_s = 17.75 \text{ ps}^{-1}$ and some “wiggles” close by. However, if we assign a strict Gaussian interpretation to the minimum, the $\pm 1\sigma$ bounds on \mathcal{L} can be calculated using:

$$\Delta \mathcal{L} = \frac{1}{2} \sigma, \quad (8.23)$$

The thin red lines in Figure 8.7 indicate the 1σ and 2σ regions. We observe that the $\pm 1\sigma$ bounds correspond to an error of 0.2-0.4 ps^{-1} . We have confirmed that this behavior is expected for a true oscillation signal by generating 1300 toy Monte Carlo experiments at $\Delta m_s = 17.75 \text{ ps}^{-1}$ with the same statistical power as the data. Figure 8.8 shows the negative and positive error distributions. We notice that uncertainties between 0.2 and 0.4 ps^{-1} are the most probable ones.

Our goal is to combine the analysis of B_s^0 - \bar{B}_s^0 oscillations using the semileptonic B_s^0 decays with similar analyses carried out in the hadronic decays of B_s^0 mesons at CDF. This combination is discussed in the next chapter.

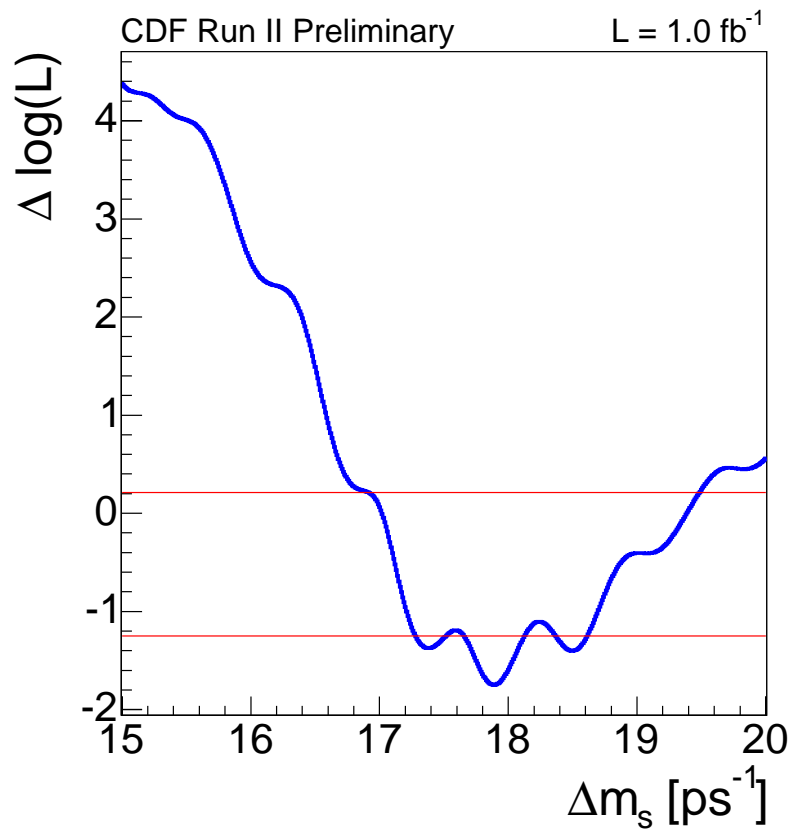


Figure 8.7: Likelihood profile as a function of Δm_s from the unblinded data for $B_s^0 \rightarrow \ell^+ D_s^- X$ decays. The thin red lines indicate the 1σ and 2σ regions.

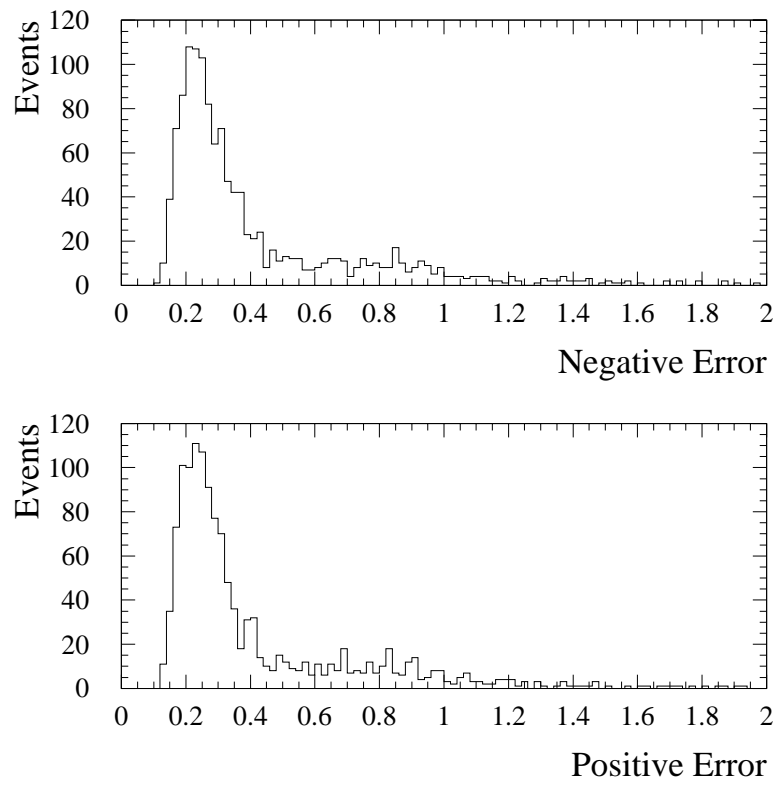


Figure 8.8: Negative (top) and positive (bottom) error distributions from 1300 toy Monte Carlo experiments simulating the real data with B_s^0 - \bar{B}_s^0 oscillations at $\Delta m_s = 17.75 \text{ ps}^{-1}$.

Chapter 9

Summary and Conclusions

9.1 Combination with Hadronic Modes

In conjugation with the analysis of semileptonic B_s^0 decays searching for B_s^0 - \bar{B}_s^0 oscillations, hadronic decay modes of the B_s^0 mesons have also been analyzed [72, 73]. These include the fully reconstructed modes:

- $B_s^0 \rightarrow D_s^- \pi^+ (D_s^- \pi^+ \pi^- \pi^+)$, $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$;
- $B_s^0 \rightarrow D_s^- \pi^+ (D_s^- \pi^+ \pi^- \pi^+)$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$;
- $B_s^0 \rightarrow D_s^- \pi^+ (D_s^- \pi^+ \pi^- \pi^+)$, $D_s^- \rightarrow \pi^+ \pi^- \pi^-$,

and the partially reconstructed hadronic modes $B_s^0 \rightarrow D_s^{*-} \pi^+$, $D_s^{*-} \rightarrow D_s^- \gamma / \pi^0$ and $B_s^0 \rightarrow D_s^- \rho^+$, $\rho^+ \rightarrow \pi^+ \pi^0$, with $D_s^- \rightarrow \phi \pi^-$, $\phi \rightarrow K^+ K^-$. The main advantage of using the fully reconstructed hadronic decays over the semileptonic decays is that their momentum uncertainty is negligible. Thus, they have no variation in their proper time resolution with respect to the proper decay time. This is displayed on the right-hand side in Figure 5.17 showing a comparison of decay time resolutions between fully reconstructed hadronic and semileptonic decays. The partially reconstructed hadronic decays miss soft neutral particles such as a π^0 or γ in their reconstruction, and consequently have a narrow distribution of the corresponding κ -factor. An artificial neural network is used for candidate selection in the hadronic decays, utilizing kinematical as well as particle identification quantities. Other aspects of the analysis in the hadronic decays, for example flavor tagging, are the same as in the case of semileptonic decays. The B_s^0 - \bar{B}_s^0 mixing analyses in the case of hadronic decays utilize $\sim 5,600$ fully reconstructed, and $\sim 3,100$ partially reconstructed decays.

The unblinded amplitude scan for the hadronic decays, including the fully and partially reconstructed samples, is shown in Figure 9.1. In this sample, the statistical uncertainty dominates for values of $\Delta m_s > 5 \text{ ps}^{-1}$. The sensitivity is 30.7 ps^{-1} , and the 95% confidence level exclusion limit is $\Delta m_s > 17.1 \text{ ps}^{-1}$. The amplitude scan for the combination of the semileptonic (shown in Figure 8.5) and hadronic samples

is derived from the combined likelihood. The result is shown in Figure 9.2. The dominant systematic uncertainties on the amplitude are fully correlated between the hadronic and semileptonic amplitude scans. The systematic uncertainty assigned to the combined amplitude is the maximum of the two.

The statistical uncertainty on the combined amplitude scan exceeds the systematic error for all Δm_s values of interest. The sensitivity of the scan is 31.3 ps^{-1} , and the limit obtained from the combined semileptonic and hadronic analysis is $\Delta m_s > 17.2 \text{ ps}^{-1}$ at 95% confidence level. This is significantly lower than the expected limit because the amplitude shows a value consistent with unity at $\Delta m_s = 17.75 \text{ ps}^{-1}$. The value of the amplitude at this value of Δm_s is $\mathcal{A} = 1.21 \pm 0.20$ (stat.), indicating that the data are compatible with B_s^0 - \bar{B}_s^0 oscillations with that frequency, while the amplitude is inconsistent with zero: amplitude significance $\mathcal{A}/\sigma_{\mathcal{A}} = 6.05$, where $\sigma_{\mathcal{A}}$ is the statistical uncertainty on \mathcal{A} (the ratio has negligible systematic uncertainties). The small uncertainty on \mathcal{A} at $\Delta m_s = 17.75 \text{ ps}^{-1}$ is a consequence of the superior decay time resolution of the hadronic decay modes, as mentioned above.

Figure 9.3 shows the value of Λ (see Equation (8.22)) as a function of Δm_s . The minimum of $\Lambda = -17.26$ is observed corresponding to $\Delta m_s = 17.77 \text{ ps}^{-1}$. The significance of the signal is the probability that randomly tagged data would produce a value of Λ lower than -17.26 at any value of Δm_s . To evaluate this probability, the likelihood scans are repeated on the data with random flavor tagging decisions. The distribution of the minimum value of Λ found in the randomized scans is shown in Figure 9.4. Out of 3.5×10^8 trials, 28 have $\Lambda < -17.26$. This means the probability for random scans to produce a signal as significant as the one we see in data is 8×10^{-8} (or 5.4σ). This is well below the 5σ threshold of 5.7×10^{-7} .

A measurement of Δm_s is extracted from the shape of the Λ function near the minimum, as shown on the right-hand side of Figure 9.3. The central value of Δm_s is taken to be the minimum of Λ . Confidence intervals are given by the region where the likelihood deviates less than $1/2$, 1.64 and 1.90 for 1σ , 90% and 95% CL, respectively. The measured value of Δm_s is:

$$\Delta m_s = 17.77^{+0.09}_{-0.10} \text{ (stat)} \pm 0.07 \text{ (syst)} \text{ ps}^{-1}.$$

The systematic uncertainty on Δm_s is from the uncertainty on the absolute scale of the decay time measurement (see Section 7.6.8). Contributions to this uncertainty include biases in the primary vertex reconstruction due to the presence of the opposite side b -hadron, uncertainties in the silicon detector alignment, and biases in track fitting. Additional source of systematic uncertainty arises from κ -factors for partially reconstructed decays. This only contributes to the measurement using the partially reconstructed sample, and has negligible effect on the overall systematic error.

In addition, we determine the following confidence intervals:

$$17.56 < \Delta m_s < 17.96 \text{ ps}^{-1} \text{ at } 90\% \text{ CL},$$

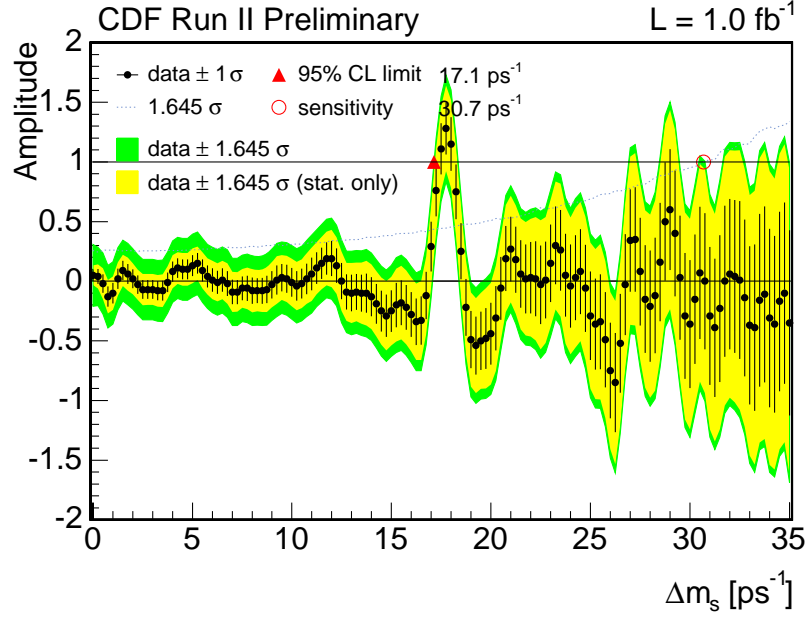


Figure 9.1: The combined amplitude scan for the hadronic modes.

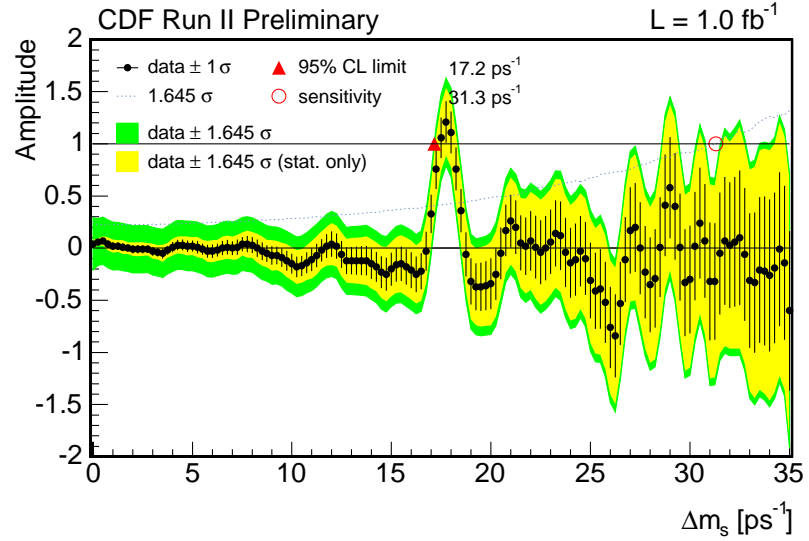


Figure 9.2: The combination of the semileptonic and hadronic amplitude scans.

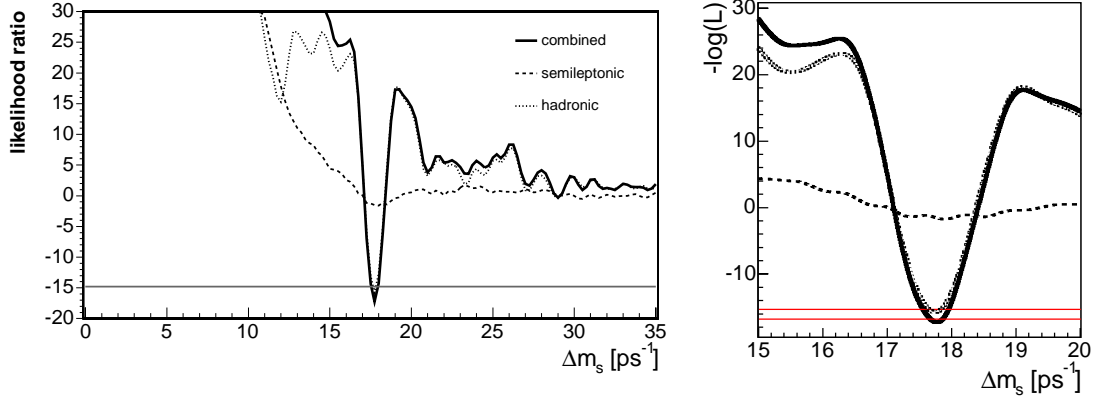


Figure 9.3: (left) Λ as a function of Δm_s in the range 0-35 ps⁻¹. The horizontal line indicates the value of Λ that corresponds to a deviation from randomly tagged data with 5 σ significance. (right) Λ in the region near the minimum. The horizontal lines indicate values of $\Delta\Lambda$ corresponding to 1 σ and 95% CL intervals.

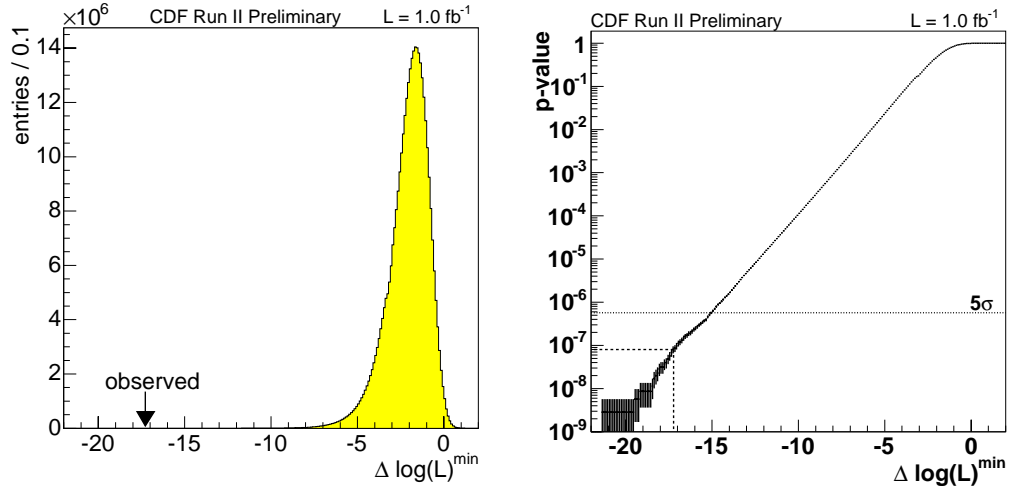


Figure 9.4: (left) Distribution of the minimal likelihood ratio Λ in scans with randomized tag decisions. The arrow indicates the value observed in the measurement of -17.26. (right) Probability for randomly tagged data as function of the minimal observed likelihood ratio.

and

$$16.51 < \Delta m_s < 18.00 \text{ ps}^{-1} \text{ at 95\% CL},$$

which includes both statistical and systematic uncertainties.

The measured B_s^0 - \bar{B}_s^0 oscillation frequency is used to derive the ratio $|V_{td}/V_{ts}|$ using Equation (1.48). As inputs, we use $m_{B^0}/m_{B_s^0} = 0.98390$ [74] with negligible uncertainty, $\Delta m_d = 0.507 \pm 0.005 \text{ ps}^{-1}$ [13] and $\xi = 1.21^{+0.047}_{-0.035}$ [24]. We extract:

$$|V_{td}/V_{ts}| = 0.2060 \pm 0.0007 \text{ (exp)}^{+0.0081}_{-0.0060} \text{ (theor)},$$

where the first uncertainty is due to the CDF measurement of Δm_s and the second uncertainty is the combination in quadrature of the uncertainties from the measurement of Δm_d and the calculation of ξ , which is the dominant contribution. The impact of the Δm_s measurement on the unitarity triangle is shown in Figure 9.5, where the top plot is from early 2006 without the Δm_s measurement, while the bottom plot is after the inclusion of the observed Δm_s frequency. The uncertainty ellipse in $|V_{td}/V_{ts}|$ shrinks considerably after the measurement, and the measurement agrees with the Standard Model expectation.

9.2 Conclusions

In the current dissertation, we have analyzed B_s^0 - \bar{B}_s^0 oscillations in semileptonic B_s^0 decays using 1 fb^{-1} of data from the CDF detector at Fermilab's Tevatron Collider. We have shown that it is possible to identify clean semileptonic B decays in the two-track trigger sample. The inclusion of the two-track sample in the current dissertation greatly increases the statistical power of the semileptonic mixing analysis. An electron identification technique via an electron likelihood has been developed and is used for opposite side flavor tagging as well as isolating semileptonic B decays on the trigger side. The electron likelihood is further utilized by introducing the anti-likelihood method for obtaining samples of fake lepton background. The fractions of fake lepton background in the B_s^0 samples are determined with a novel $m_{\ell^+ D_s^-}$ fitting technique, whereby fake lepton backgrounds are clearly separated with respect to the signal B_s^0 decays. Furthermore, the usage of $m_{\ell^+ D_s^-}$ directly in the unbinned likelihood fit framework enables further discrimination against various sources of backgrounds in the B_s^0 samples. We thus significantly improve the sensitivity of the semileptonic mixing analysis by isolating the most sensitive candidates in $m_{\ell^+ D_s^-}$. Particle identification is introduced in the candidate selection in an effort to further minimize contamination from combinatorial and D^- reflection backgrounds. We have also accomplished a flavor tagging calibration for the application of opposite side tagging in B_s^0 candidates, together with a measurement of the B^0 - \bar{B}^0 oscillation frequency Δm_d .

The semileptonic analysis of B_s^0 - \bar{B}_s^0 mixing has a sensitivity of 19.4 ps^{-1} and shows an evidence of B_s^0 oscillations at $\Delta m_s \sim 17.75 \text{ ps}^{-1}$ with an amplitude significance of ~ 2 . In combination with the analyses of hadronic B_s^0 decays at CDF, we have

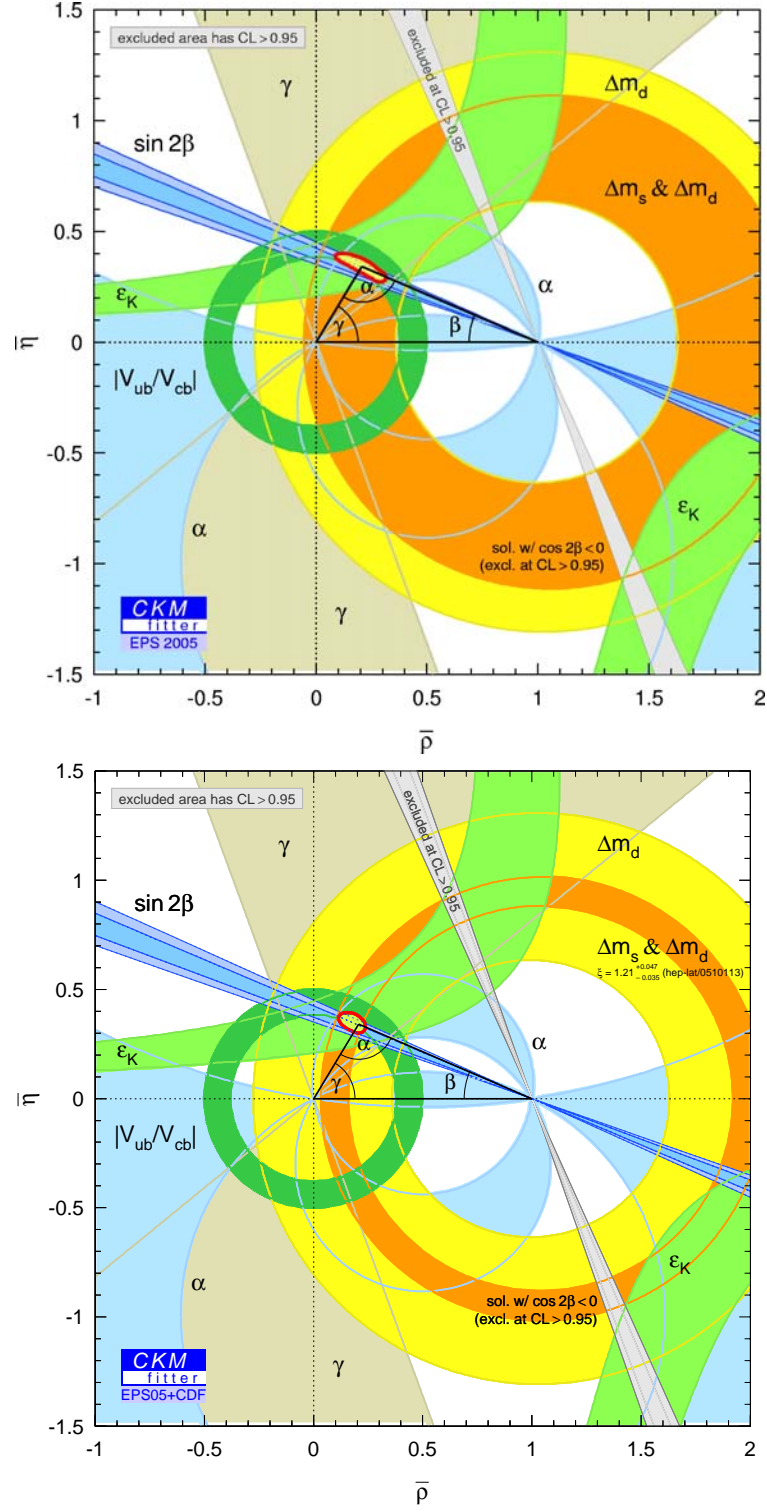


Figure 9.5: (top) Status of constraints on the unitarity triangle in (ρ', η') space as of early 2006. (bottom) After including the CDF measurement of Δm_s as of late 2006.

made the first direct observation of time-dependent B_s^0 - \bar{B}_s^0 flavor oscillations. From this observation, we measure

$$\Delta m_s = 17.77^{+0.09}_{-0.10} \text{ (stat)} \pm 0.07 \text{ (syst)} \text{ ps}^{-1}.$$

When combined with the world average values for Δm_d , $m_{\bar{B}^0}$ and $m_{\bar{B}_s^0}$, along with other theoretical input, this result yields

$$|V_{td}/V_{ts}| = 0.2060 \pm 0.0007 \text{ (exp)}^{+0.0081}_{-0.0060} \text{ (theor)}.$$

Appendix A

Corrections to CES $q\Delta X$ and ΔZ

In this appendix, we describe the corrections applied to the measurement of $q\Delta X$ and ΔZ in the CES. The dependences arise because of electrons traverse different amount of material in the CDF detector, depending on their transverse momentum p_T and pseudorapidity η . Additionally, uncorrected offsets between the COT and the CES detectors result into shifts in these distributions. The methodology applied to obtain these corrections is the same in the case of both $q\Delta X$ and ΔZ : obtain η dependent corrections in bins of track p_T .

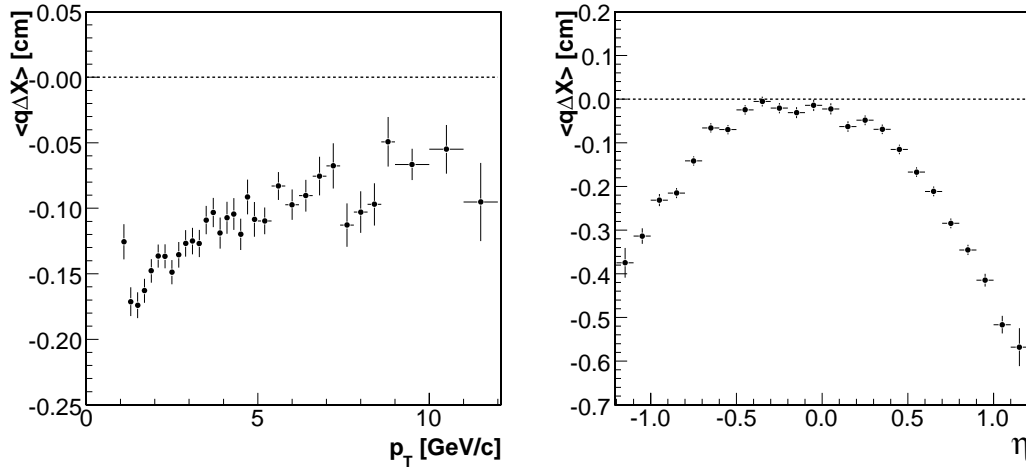


Figure A.1: $\langle q\Delta X \rangle$ with respect to p_T (left) and η (right) before corrections.

A.1 Corrections to CES $q\Delta X$

Figure A.1 shows the dependence of the mean of $q\Delta X$ with respect to p_T (left) and η (right). In order to better understand the behavior of this variable for low p_T

electrons, we have added electrons from conversions with $1.0 \leq p_T \leq 1.5 \text{ GeV}/c$ to the sample of $p_T > 1.5 \text{ GeV}/c$ electrons under study. In Figure A.1, we observe that the extrapolated position of the electron tracks is increasingly underestimated with decreasing p_T . This is due to the bremsstrahlung radiation of the electrons while interacting with the material in the detector. An electron at lower p_T is curved more in the detector, thus traveling through more detector material on average. This leads to an increased level of underestimation of the extrapolated CES position. A similar effect is visible in $\langle q\Delta X \rangle$ with respect to η , where tracks at higher η , in general, pass through more material in the detector.

The dependences of $\langle q\Delta X \rangle$ on p_T and η are correlated. Therefore, we split the conversion electron sample into three p_T bins: $1.0 < p_T < 2.5 \text{ GeV}/c$, $2.5 < p_T < 5.0 \text{ GeV}/c$, and $p_T > 5.0 \text{ GeV}/c$. We then plot $\langle q\Delta X \rangle$ with respect to η in each of the three p_T bins, and parameterize its dependence using a second order polynomial in each p_T bin. This is shown in Figure A.2. The top-left plot is for the 1.0-2.5 GeV/ c range, the top-right plot is for the 2.5-5.0 GeV/ c range, and bottom-left plot is for $p_T > 5.0 \text{ GeV}/c$. As expected, the dependence is more dramatic for lower p_T electrons. The parameterization is used to apply additive η dependent corrections in each p_T range such that $\langle q\Delta X \rangle$ is close to zero. After applying these corrections, $\langle q\Delta X \rangle \sim 0.0$ in p_T and η , as shown in Figure A.3.

A.2 Correction to CES ΔZ

Figure A.4 shows the dependence of the mean of ΔZ with respect to p_T (left) and η (right). Although the $\langle \Delta Z \rangle$ distribution does not show dependence with respect to p_T , we obtain the offset corrections in a manner similar to the $q\Delta X$ case. An η dependence is clearly visible in Figure A.4. Electrons at higher η on average go through more material in the detector leading to more bremsstrahlung radiation and consequently, an increased under-estimation of the extrapolated CES position. In addition, we note that the sign of $\langle \Delta Z \rangle$ changes as we go from $\eta < 0$ to $\eta > 0$. This is possibly due to the alignment effects between the COT and the CES detectors.

Following the same procedure as before when correcting $q\Delta X$, we split the sample of conversion electrons into three p_T bins: $1.0 < p_T < 2.5 \text{ GeV}/c$, $2.5 < p_T < 5.0 \text{ GeV}/c$, and $p_T > 5.0 \text{ GeV}/c$. We then plot $\langle \Delta Z \rangle$ with respect to η in each of the three p_T bins and parameterize the dependence using the following functional form for $1.0 < p_T < 2.5 \text{ GeV}/c$:

$$\langle \Delta Z \rangle(\eta) = p_0 \tanh(p_1 \eta + p_2) + p_3 \sinh(p_4(\eta - p_5)). \quad (\text{A.1})$$

For the $2.5 < p_T < 5.0 \text{ GeV}/c$ and $p_T > 5.0 \text{ GeV}/c$ ranges, we use the following functional form:

$$\langle \Delta Z \rangle(\eta) = p_0 \tanh(p_1 \eta + p_2) + p_3. \quad (\text{A.2})$$

The result of the parameterization is shown in Figure A.5. The top-left plot is for the 1.0-2.5 GeV/ c range, the top-right plot is for the 2.5-5.0 GeV/ c range, and

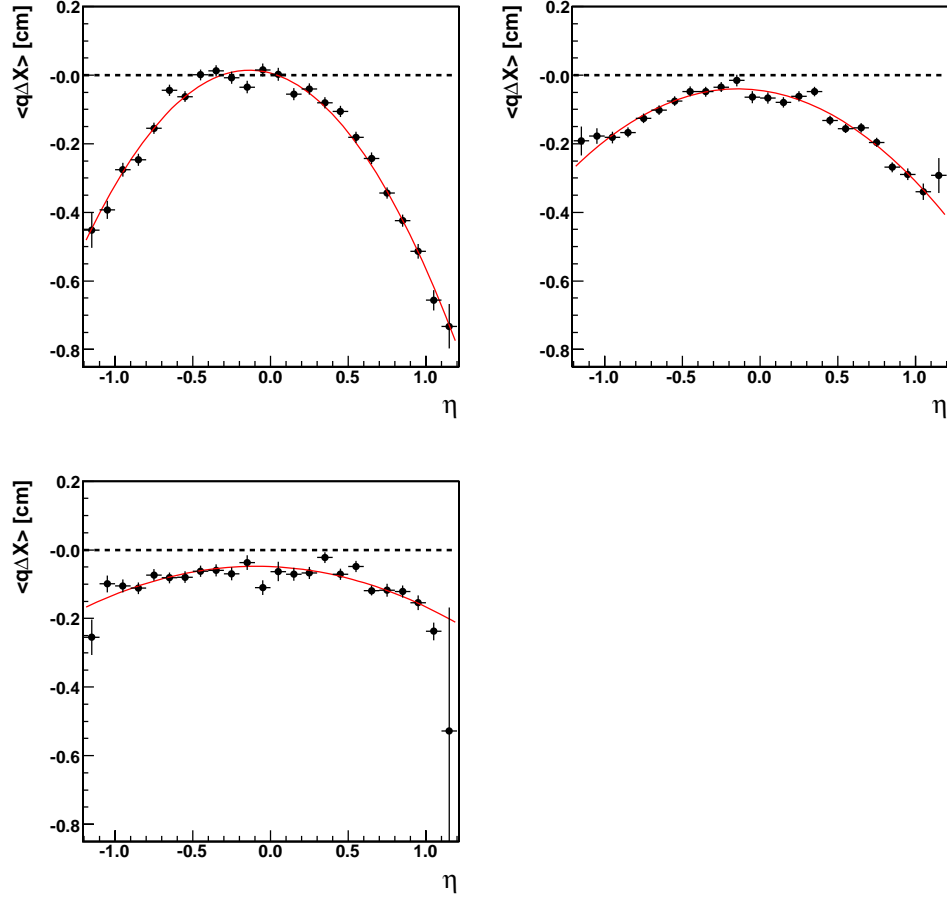


Figure A.2: Parameterization of η dependence of $\langle q\Delta X \rangle$ in bins of p_T : $1.0 < p_T < 2.5$ GeV/c (top-left), $2.5 < p_T < 5.0$ GeV/c (top-right), and $p_T > 5.0$ GeV/c (bottom-left).

bottom-left plot is for $p_T > 5.0$ GeV/c. We do not observe significant differences between the different p_T ranges. The parameterization is used to apply additive η dependent corrections in each p_T range such that $\langle \Delta Z \rangle$ is close to zero after correction. The variations in $\langle \Delta Z \rangle$ with respect to η are greatly reduced after applying these corrections, and $\langle \Delta Z \rangle \sim 0.0$ as shown in Figure A.6. The p_T dependence of CES $\langle \Delta Z \rangle$ was flat before corrections and remains flat after corrections.

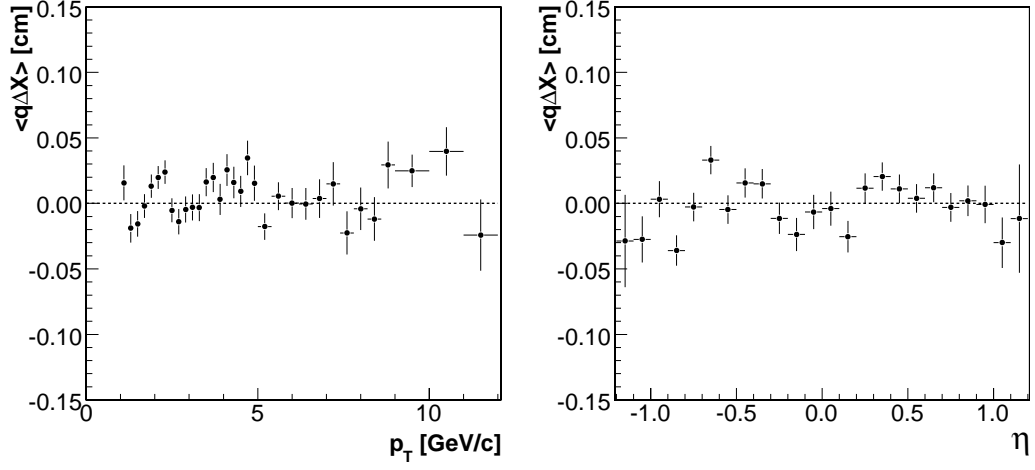


Figure A.3: $\langle q\Delta X \rangle$ with respect to p_T (left) and η (right) after corrections.

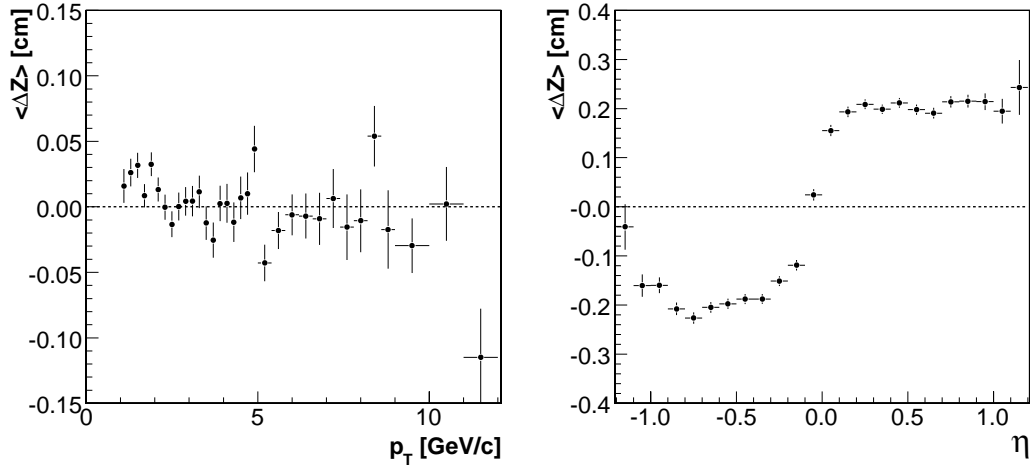


Figure A.4: $\langle \Delta Z \rangle$ with respect to p_T (left) and η (right) before corrections.

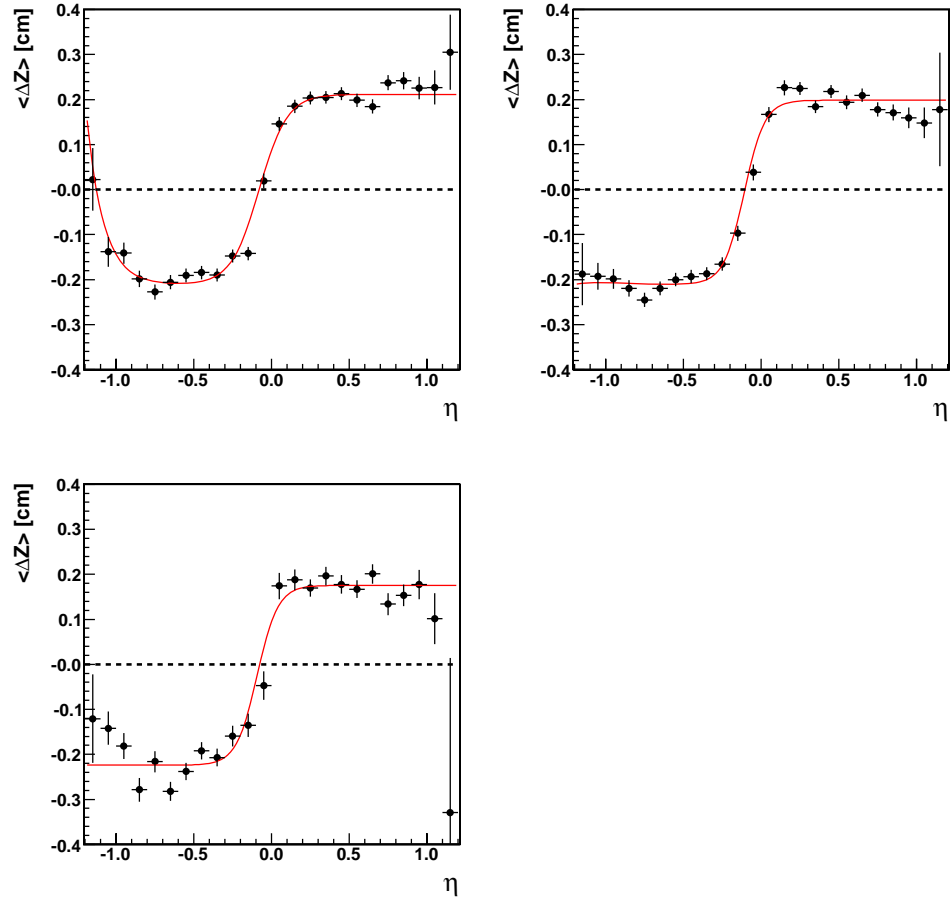


Figure A.5: Parameterization of η dependence of $\langle \Delta Z \rangle$ in three p_T bins: $1.0 < p_T < 2.5$ GeV/ c (top-left), $2.5 < p_T < 5.0$ GeV/ c (top-right), and $p_T > 5.0$ GeV/ c (bottom-left).

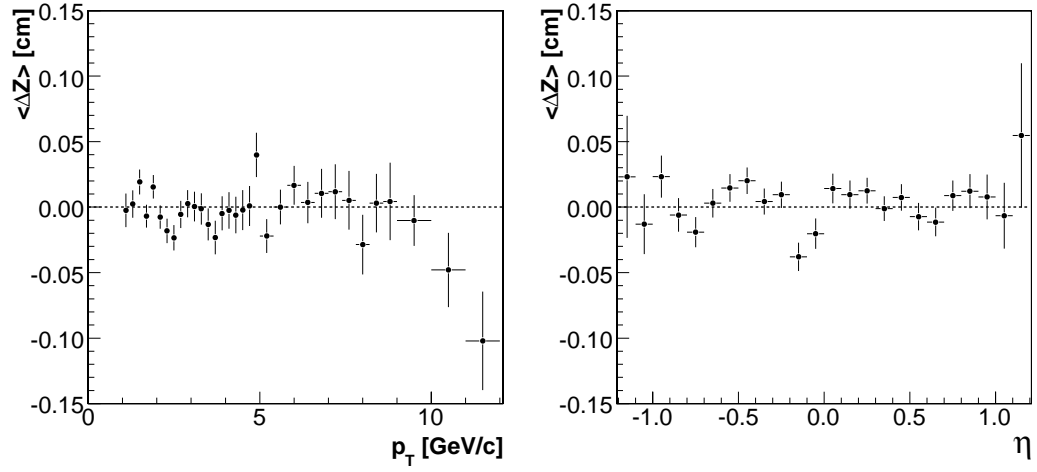


Figure A.6: $\langle \Delta Z \rangle$ with respect to p_T (left) and η (right) after corrections.

Appendix B

Electron Identification after the First 360 pb⁻¹ of Data

Electron identification for the first 360 pb⁻¹ data (“old data”) is described in detail in Section 3.2. Information from various electron identification quantities from the calorimeter, CPR, CES and dE/dx are combined into a global likelihood to distinguish real electrons from fakes. Templates for constructing the electron likelihood for the electron identification are obtained from conversion electrons ($\gamma \rightarrow e^+e^-$) and K_S^0 pions ($K_S^0 \rightarrow \pi^+\pi^-$). The templates are separated for isolated and non-isolated candidates as well as in three ranges in p_T : $1.5 < p_T < 2.0$ GeV/ c , $2.0 < p_T < 4.0$ GeV/ c , and $p_T > 4.0$ GeV/ c . In this appendix, we briefly describe changes to the likelihood templates obtained for the first 360 pb⁻¹. This is done in order to account for the changes in the detector configuration and electron identification quantities in the later ~ 640 pb⁻¹ data (“later data”) utilized in this dissertation.

B.1 Electron Identification in the Later Data

A new CPR detector (CPR2) was installed in the CDF detector during the accelerator shutdown period in 2004. About 640 pb⁻¹ of data used in this analysis is accumulated using CPR2. In terms of reconstruction of electron identification quantities in the CPR2 detector, the main difference with respect to using the old CPR based pulse height Q_{CPR} is that the response of the CPR2 detector for real electrons requires corrections to account for the known p_T and $\sin \theta$ dependences. We utilize ~ 410 pb⁻¹ of the later data to study these dependences, and establish a calibrated Q_{CPR} for the data collected with the CPR2 detector.

These dependences are displayed in Figure B.1 for the electrons obtained from photon conversions. In the plot on the left, the mean CPR2 pulse height $\langle Q_{CPR} \rangle$ in multiples of energy deposited by minimum ionizing particles Q_{MIP} is plotted with respect to track $\sin \theta$, and in the plot on the right, $\langle Q_{CPR} \rangle$ is plotted with respect to track p_T . Empirical fits to these dependences are also overlaid. Using these fits,

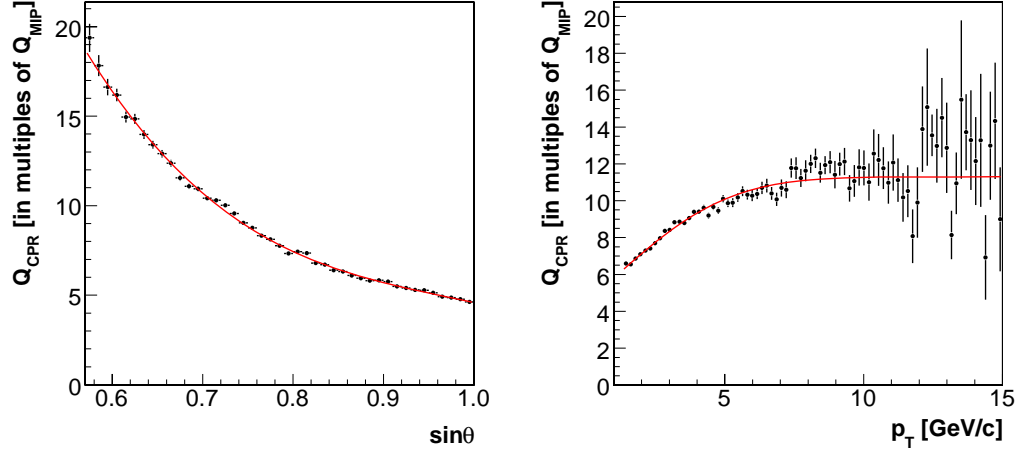


Figure B.1: $\langle Q_{CPR} \rangle$ for conversion electrons in CPR2 (in multiples of pulse height deposited by minimum ionizing particles Q_{MIP} with respect to $\sin \theta$ (left) and p_T (right) before applying corrections.

$\langle Q_{CPR} \rangle$ for real electrons with $p_T > 1.5$ GeV/ c is corrected one-by-one in $\sin \theta$ and p_T such that the resulting $\langle Q_{CPR} \rangle$ is flat in these two variables, as shown in Figure B.2. Applying these corrections, we obtain templates for real electrons from photon conversions and pions from K_S^0 decays. These templates for the $2.0 < p_T < 4.0$ GeV/ c range are shown in Figure B.3 for isolated and non-isolated electrons and pions.

We also compare $\langle q\Delta X \rangle$ and $\langle \Delta Z \rangle$ with respect to track η between the later and the old data (see Appendix A) for electrons from photon conversions. These comparisons are shown in Figures B.4 and B.5 for $2.5 < p_T < 5.0$ GeV/ c electrons. The quantity $\langle q\Delta X \rangle$ has similar offsets in the later data as the old data, hence we retain the old corrections. This is not true for offsets in $\langle \Delta Z \rangle$, and hence we use these offsets to correct ΔZ in the new data by using the same procedure as described in Section A.2. Using these corrections, we derive new templates for electrons and pions for the new data.

The behavior of the other discriminating variables used in the electron likelihood was checked in the later data and found to be similar to the old data. An electron likelihood for the later data was formed using the same procedure as described in Section 3.2.3.

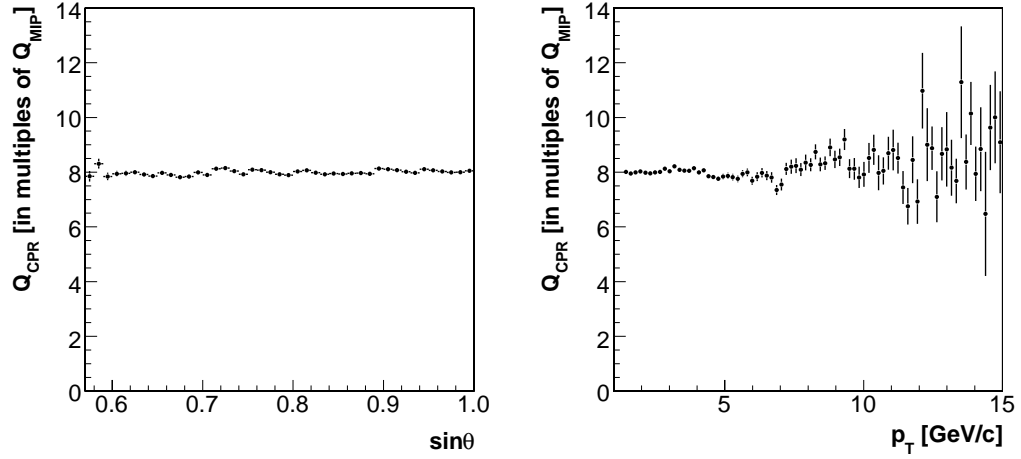


Figure B.2: $\langle Q_{CPR} \rangle$ for conversion electrons in CPR2 (in multiples of Q_{MIP}) with respect to $\sin\theta$ (left) and p_T (right) after applying corrections.

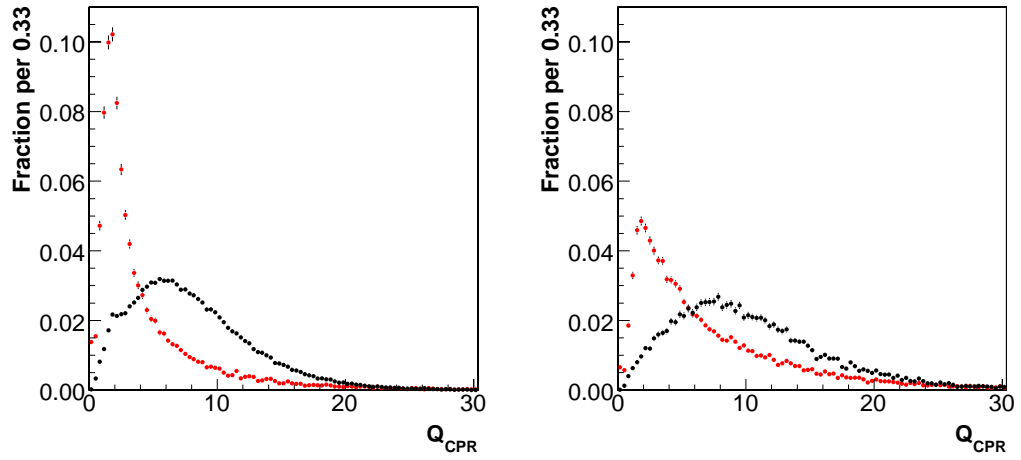


Figure B.3: Q_{CPR} distributions (in multiples of Q_{MIP}) for isolated (left) and non-isolated (right) electrons and pions in the $2.0 < p_T < 4.0$ GeV/c range after applying corrections.

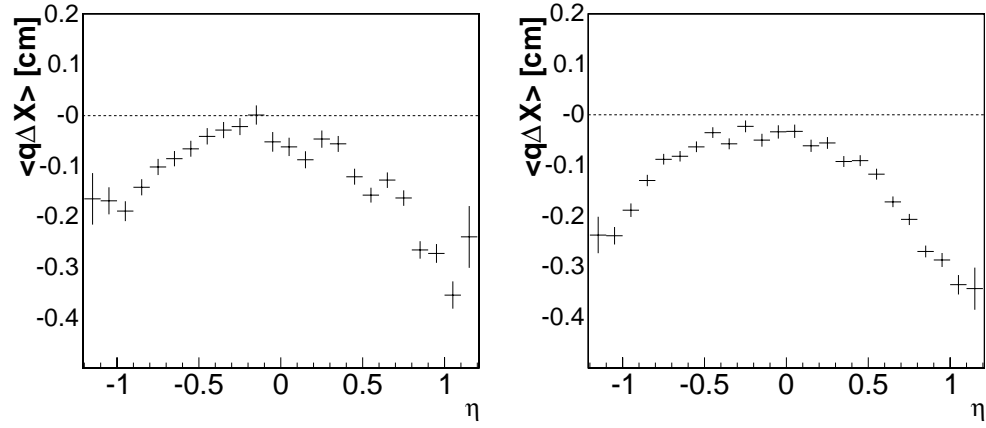


Figure B.4: $\langle q\Delta X \rangle$ for conversion electrons with respect to η for the old (left) and the later data (right).

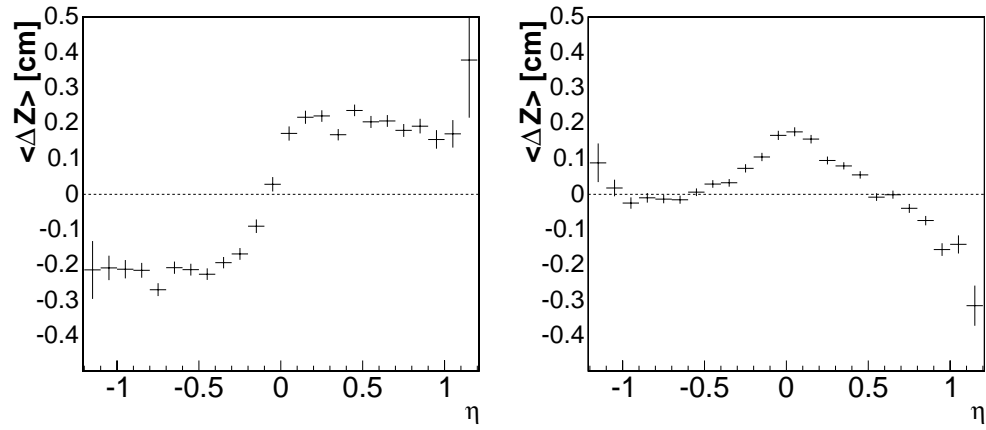


Figure B.5: $\langle \Delta Z \rangle$ for conversion electrons with respect to η for the old (left) and the later data (right).

Appendix C

Amplitude Scan for Various Subsets of Data

The unblinded amplitude scan distributions for the three data-taking periods are shown in Figures C.1, C.2 and C.3, respectively. Separate amplitude scans using the SSKT algorithm only or the OST only are presented in Figures C.4 and C.5, respectively.

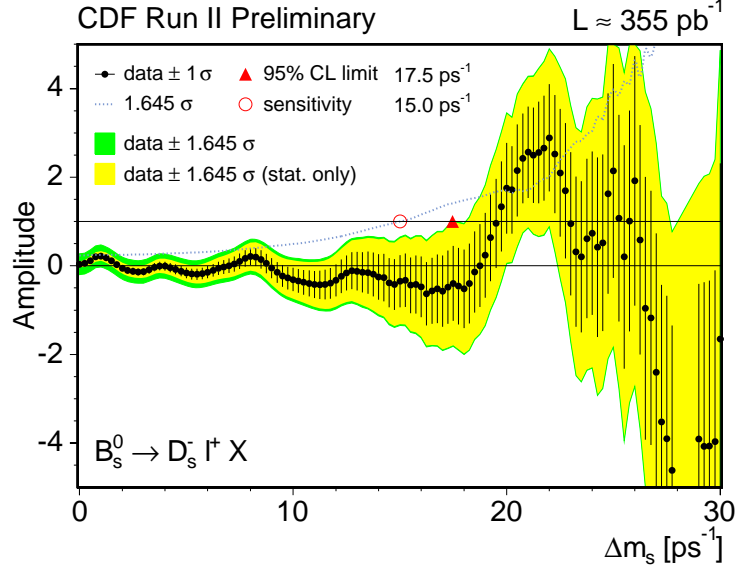


Figure C.1: Unblinded amplitude scan in Period 1 data for the $B_s \rightarrow \ell^+ D_s^- X$ decays. The errors displayed include statistical and systematic errors.

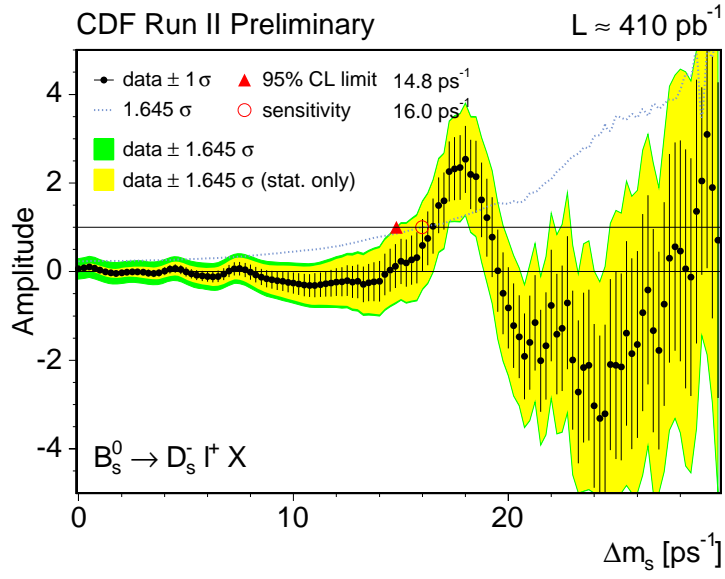


Figure C.2: Unblinded amplitude scan in Period 2 data for the $B_s \rightarrow \ell^+ D_s^- X$ decays. The errors shown include statistical and systematic errors.

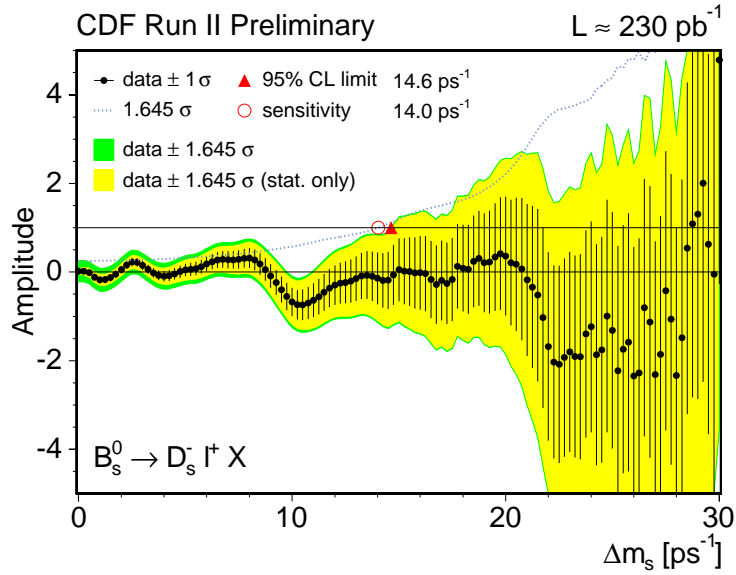


Figure C.3: Unblinded amplitude scan in Period 3 data for the $B_s \rightarrow \ell^+ D_s^- X$ decays. The errors displayed include statistical and systematic errors.

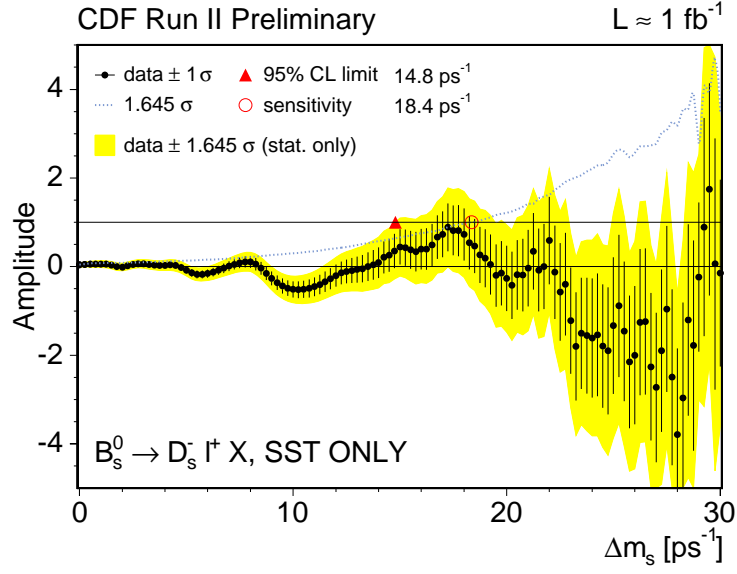


Figure C.4: Unblinded amplitude scan in data for the $B_s \rightarrow \ell^+ D_s^- X$ decays using same side tags alone. Only statistical errors are shown here.

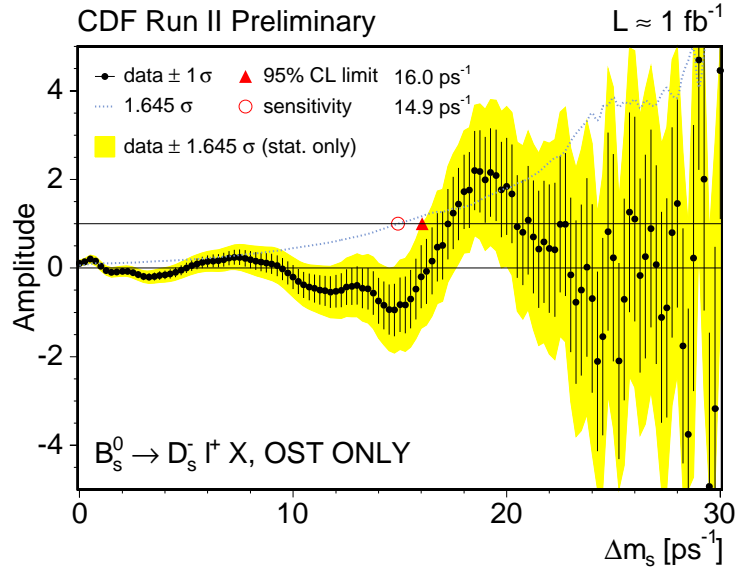


Figure C.5: Unblinded amplitude scan in data for the $B_s \rightarrow \ell^+ D_s^- X$ decays using opposite side tags alone. Only statistical errors are displayed here.

Bibliography

- [1] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
- [2] C. Albajar *et al.* [UA1 Collaboration], Phys. Lett. B **186**, 247 (1987)
[Erratum-ibid. **197B**, 565 (1987)].
- [3] H. Albrecht *et al.* [ARGUS Collaboration], Phys. Lett. B **192**, 245 (1987).
- [4] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).
- [5] S. L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D **2**, 1285 (1970).
- [6] J. J. Aubert *et al.* [E598 Collaboration], Phys. Rev. Lett. **33**, 1404 (1974).
- [7] J. E. Augustin *et al.* [SLAC-SP-017 Collaboration], Phys. Rev. Lett. **33**, 1406 (1974).
- [8] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).
- [9] S. W. Herb *et al.*, Phys. Rev. Lett. **39**, 252 (1977).
- [10] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **74**, 2626 (1995)
[arXiv:hep-ex/9503002].
- [11] L. L. Chau and W. Y. Keung, Phys. Rev. Lett. **53**, 1802 (1984).
- [12] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).
- [13] W. M. Yao *et al.* [Particle Data Group], J. Phys. G **33**, 1 (2006).
- [14] J. Charles *et al.* [CKMfitter Group], Eur. Phys. J. C **41**, 1 (2005)
[arXiv:hep-ph/0406184]. See <http://ckmfitter.in2p3.fr/>
- [15] A. Bornheim *et al.* [CLEO Collaboration], Phys. Rev. Lett. **88**, 231803 (2002)
[arXiv:hep-ex/0202019].
- [16] B. Aubert *et al.* [BaBar Collaboration], arXiv:hep-ex/0408075.

- [17] A. Limosani *et al.* [Belle Collaboration], Phys. Lett. B **621**, 28 (2005) [arXiv:hep-ex/0504046].
- [18] B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. Lett. **94**, 161803 (2005) [arXiv:hep-ex/0408127].
- [19] K. Abe *et al.* [Belle Collaboration], arXiv:hep-ex/0507037.
- [20] A. J. Buras, W. Slominski and H. Steger, Nucl. Phys. B **245**, 369 (1984).
- [21] A. J. Buras, M. Jamin and P. H. Weisz, Nucl. Phys. B **347**, 491 (1990).
- [22] C. Gay, Ann. Rev. Nucl. Part. Sci. **50**, 577 (2000) [arXiv:hep-ex/0103016].
- [23] T. Inami and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981) [Erratum-ibid. **65**, 1772 (1981)].
- [24] A. Gray *et al.* [HPQCD Collaboration], Phys. Rev. Lett. **95**, 212001 (2005) [arXiv:hep-lat/0507015];
S. Aoki *et al.* [JLQCD Collaboration], Phys. Rev. Lett. **91**, 212001 (2003) [arXiv:hep-ph/0307039];
M. Okamoto, PoS **LAT2005**, 013 (2006) [arXiv:hep-lat/0510113].
- [25] M. Artuso *et al.*, Phys. Rev. Lett. **62**, 2233 (1989).
- [26] K. Abe *et al.* [BELLE Collaboration], Phys. Rev. D **71**, 072003 (2005); **71**, 079903(E) (2005);
N. C. Hastings *et al.* [BELLE Collaboration], Phys. Rev. D **67**, 052004 (2003);
B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. Lett. **88**, 221803 (2002).
- [27] H. G. Moser and A. Roussarie, Nucl. Instrum. Meth. A **384**, 491 (1997).
- [28] J. Abdallah *et al.* [DELPHI Collaboration], Eur. Phys. J. C **35**, 35 (2004) [arXiv:hep-ex/0404013];
K. Abe *et al.* [SLD Collaboration], Phys. Rev. D **67**, 012006 (2003);
A. Heister *et al.* [ALEPH Collaboration], Eur. Phys. J. C **29**, 143 (2003).
- [29] G. Aubrecht *et al.*, “A Teachers Guide to the Nuclear Science Wall Chart”, Contemporary Physics Education Project, 2003.
<http://www.lbl.gov/abc/wallchart/teachersguide/pdf/Chap11.pdf>.
- [30] C. W. Schmidt, “The Fermilab 400-MeV Linac upgrade”
FERMILAB-CONF-93-111 (1993).
- [31] Fermilab Beams Division, “Run II Handbook”,
<http://www-bd.fnal.gov/runII/index.html>.

- [32] J. Marriner, “Stochastic Cooling Overview”,
FERMILAB-CONF-03-158(2003).
- [33] R. Blair *et al.*, “The CDF-II detector: Technical design report”,
FERMILAB-PUB-96/390-E (1996).
- [34] F. Abe *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **271**, 387 (1988).
- [35] C. S. Hill [On behalf of the CDF Collaboration], Nucl. Instrum. Meth. A **530**, 1 (2004).
- [36] P. Azzi [CDFII Collaboration], Nucl. Instrum. Meth. A **419**, 532 (1998);
A. Sill [CDF Collaboration], Nucl. Instrum. Meth. A **447**, 1 (2000).
- [37] A. A. Affolder *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **453**, 84 (2000).
- [38] A. A. Affolder *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **526**, 249 (2004).
- [39] S. Cabrera *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **494**, 416 (2002).
- [40] L. Balka *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **267**, 272 (1988).
- [41] S. Bertolucci *et al.* [CDF Collaboration], Nucl. Instrum. Meth. A **267**, 301 (1988).
- [42] G. Apollinari, K. Goulianos, P. Melese and M. Lindgren, Nucl. Instrum. Meth. A **412**, 515 (1998).
- [43] F. Abe *et al.* [CDF Collaboration], Phys. Rev. Lett. **73**, 2662 (1994)
[Erratum-ibid. **74**, 1891 (1995)].
- [44] S. Kuhlmann *et al.*, Nucl. Instrum. Meth. A **518**, 39 (2004)
[arXiv:physics/0310155].
- [45] G. Ascoli *et al.*, Nucl. Instrum. Meth. A **268**, 33 (1988).
- [46] T. Dorigo [CDF Collaboration], Nucl. Instrum. Meth. A **461**, 560 (2001).
- [47] E. J. Thomson *et al.*, IEEE Trans. Nucl. Sci. **49**, 1063 (2002).
- [48] A. Bardi *et al.*, Nucl. Instrum. Meth. A **485**, 178 (2002);
B. Ashmanskas *et al.* [CDF-II Collaboration], Nucl. Instrum. Meth. A **518**, 532 (2004)
[arXiv:physics/0306169].
- [49] G. Gomez-Ceballos *et al.*, Nucl. Instrum. Meth. A **518** (2004) 522.

- [50] P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. B **303**, 607 (1988); P. Nason, S. Dawson and R. K. Ellis, Nucl. Phys. B **327**, 49 (1989) [Erratum-ibid. B **335**, 260 (1990)].
- [51] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. D **71**, 032001 (2005) [arXiv:hep-ex/0412071].
- [52] D. J. Lange, Nucl. Instrum. Meth. A **462**, 152 (2001).
- [53] T. Sjostrand, L. Lonnblad and S. Mrenna, arXiv:hep-ph/0108264. We use version 6.216.
- [54] A. A. Affolder *et al.* [CDF Collaboration], Phys. Rev. D **65**, 092002 (2002).
- [55] R. D. Field, Phys. Rev. D **65**, 094006 (2002) [arXiv:hep-ph/0201112].
- [56] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. D **71**, 092001 (2005) [arXiv:hep-ex/0412006].
- [57] E. Gerchtein and M. Paulini, *In the Proceedings of 2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Jolla, California, 24-28 Mar 2003, pp TUMT005* [arXiv:physics/0306031].
- [58] R. Brun, R. Hagelberg, M. Hansroul and J. C. Lassalle, CERN-DD-78-2-REV.
- [59] D. Tonelli, Ph. D. thesis, Scuola Normale Superiore, 2006, FERMILAB-THESIS-2006-23.
- [60] G. Giurciu, Ph. D. thesis, Carnegie Mellon University, 2005, FERMILAB-THESIS-2005-41.
- [61] J. Marriner, "Secondary Vertex Fit with Mass and Pointing Constraints (CTVMFT)", CDF Note 1996 (1993).
- [62] F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).
- [63] F. Abe *et al.* [CDF Collaboration], Phys. Rev. D **59**, 032001 (1999) [arXiv:hep-ex/9806026].
- [64] A. Abulencia *et al.* [CDF Collaboration], Phys. Rev. Lett. **97**, 082004 (2006) [arXiv:hep-ex/0606017].
- [65] A. A. Affolder, Ph. D. thesis, University of California at Berkeley, 2002, FERMILAB-THESIS-2002-18.
- [66] C. Lecci, Ph. D. thesis, University of Karlsruhe, 2005, IEKP-KA/2005-13.

- [67] M. Feindt and U. Kerzel, Nucl. Instrum. Meth. A **559**, 190 (2006).
- [68] A. Ali and F. Barreiro, Z. Phys. C **30**, 635 (1986);
M. Gronau, A. Nippe, J. L. Rosner, Phys. Rev. D **47**, 1988 (1993);
M. Gronau and J. L. Rosner, Phys. Rev. D **49**, 254 (1994).
- [69] D. Usynin, Ph. D. thesis, University of Pennsylvania, 2005,
FERMILAB-THESIS-2005-68.
- [70] A. A. Affolder *et al.* [CDF Collaboration], Phys. Rev. D **61**, 072005 (2000)
[arXiv:hep-ex/9909003].
- [71] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. Lett. **94**, 101803 (2005)
[arXiv:hep-ex/0412057].
- [72] A. Belloni, Ph. D. thesis, Massachusetts Institute of Technology, thesis in preparation.
- [73] J. Miles, Ph. D. thesis, Massachusetts Institute of Technology, thesis in preparation.
- [74] D. Acosta *et al.* [CDF Collaboration], Phys. Rev. Lett. **96**, 202001 (2006)
[arXiv:hep-ex/0508022].