B Flavor Tagging Calibration and Search for B_s^0 Oscillations in Semileptonic Decays with the CDF Detector at Fermilab

by
Gavril A. Giurgiu

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
at
Carnegie Mellon University
Department of Physics
Pittsburgh, Pennsylvania
September 2005

Advisor: Prof. Manfred Paulini

B Flavor Tagging Calibration and Search for B_s^0 Oscillations in Semileptonic Decays with the CDF Detector at Fermilab

Abstract

In this thesis we present a search for oscillations of B_s^0 mesons using semileptonic $B_s^0 \to D_s^- \ell^+ \nu$ decays. Data were collected with the upgraded Collider Detector at Fermilab (CDFII) from events produced in collisions of 980 GeV protons and antiprotons accelerated in the Tevatron ring. The total proton-antiproton center-of-mass energy is 1.96 TeV. The Tevatron is the unique source in the world for B_s^0 mesons, to be joined by the Large Hadron Collider at CERN after 2007.

We establish a lower limit on the B_s^0 oscillation frequency $\Delta m_s > 7.7$ ps^{-1} at 95% Confidence Level. We also present a multivariate tagging algorithm that identifies semileptonic $B \to \mu X$ decays of the other B mesons in the event. Using this muon tagging algorithm as well as opposite side electron and jet charge tagging algorithms, we infer the B_s^0 flavor at production. The tagging algorithms are calibrated using high statistics samples of B^0 and B^+ semileptonic $B^{0/+} \to D\ell\nu$ decays. The oscillation frequency Δm_d in semileptonic $B^0 \to D\ell\nu$ decays is measured to be $\Delta m_d = (0.501 \pm 0.029(stat.) \pm 0.017(syst.))$ ps^{-1}.

Acknowledgments

I thank the CDF collaboration and the Fermilab staff for making this work possible.

I thank my advisor, Prof. Manfred Paulini, for guiding me through the past five years in all the projects I was involved in, beginning with pedagogical simulation projects and ending with the topics described in this thesis. I have learned from him physics and data analysis techniques. I learned to work carefully and study things in detail. I learned to steadily follow my goals and to be self-confident and patient.

I thank Prof. James Russ, from whom I got valuable advice whenever I needed it. Besides a lot of physics, he taught me to think critically and strive to understand things throughly. His insights helped me solve numerous problems that came up in my analyses.

I thank both Prof. Paulini and Prof. Russ for creating a great working environment and for taking good care of me.

I thank Masashi Tanaka, our collaborator in the B^0 and B_s^0 mixing analyses. He also had many helpful suggestions related to the likelihood muon tagging algorithm. I learned many things from him about B physics and analysis techniques. I learned to push my limits and improve myself. His honesty, good thinking, and will power are an example for me to follow.

I thank Sinead Farrington for her contribution to the B_s^0 mixing analysis. Her physics background studies are crucial for setting a correct lower limit on the B_s^0 mixing frequency.

I thank Prof. Joe Boudreau, Prof. Helmut Vogel and Prof. Lincoln Wolfenstein for their comments and suggestions for writing the thesis.

I thank the Mixing Analysis Coordinators, Prof. Joseph Kroll and Prof. Franco Bedeschi, as well as the B Physics conveners, Prof. Stefano Giagu, Prof. Christoph Paus and Prof. Matthew Herndon, for guidance and useful comments and suggestions.

I thank Guillelmo Gomez-Ceballos, Stephanie Menzemer, Nuno Leonardo, Cristoph Paus, and the other members of the Cantabria-Chicago-Karlsruhe-MIT group who did an independent mixing analysis in parallel with us. We performed many cross checks which resulted in a better final result.

I thank Vivek Tiwari, who worked on electron tagging while I was working on muon tagging. We had countless fruitful discussions which led to better lepton tagging algorithms. He also helped me with using the electron tagger in the mixing analyses.

I thank Karen Gibson for all the help she gave me in the first year when my C++ and Root skills were non-existent. She helped me a lot with signal reconstruction and optimization of the $D\ell$ modes. We also had many useful discussions.

I thank both Karen and Vivek for their friendship.

I thank Ilya Kravchenko for numerous useful comments and for teaching me how to use the jet charge tagger, a tool that he developed and which proved crucial in the mixing analyses.

I thank Matthew Jones, Denys Usynin, Joseph Kroll, and Barry Wicklund for helping us understand issues related to the composition of the lepton-SVT data sample

and for developing the initial cut-based muon tagger on which the likelihood muon tagging algorithm was developed.

I thank Lucio Cerrito for many useful suggestions on the likelihood muon tagger.

I thank Prof. Marjorie Shapiro for helping us understand the event-by-event primary vertex algorithm and for many useful comments on our analyses.

I thank Soon Yung Jun for helping me to get started with the CDF software and for many useful comments on the muon tagger and other projects.

I thank Min-Jeong Kim who helped me get started with the muon tagging project.

I thank Jason Galyardt for numerous discussions.

I thank William Wester, Cristian Gingu, Kazu Hanagaki, Len Christofek and Petros Rapidis with whom I worked on testing SVX4 chips for the CDF RunIIB upgrade.

I thank the members of the B Physics group at CDF for many useful suggestions and comments.

I thank my wife for making my life beautiful and my parents for everything they have done for me.

I thank my friend Valentin Necula for many discussions about physics in general.

Contents

1 Introduction **7**
- 1.1 Theoretical Overview . . . 7
  - 1.1.1 The CKM Matrix . . . 8
  - 1.1.2 Oscillations of Neutral B Mesons . . . 10
- 1.2 Experimental Methods . . . 12
  - 1.2.1 Proper Decay Time . . . 12
  - 1.2.2 Flavor Tagging . . . 14
  - 1.2.3 Overview of B Mixing Analysis . . . 16
  - 1.2.4 The Amplitude Method . . . 18

2 Experimental Apparatus **20**
- 2.1 The CDFII Detector . . . 20
  - 2.1.1 Silicon Detector . . . 20
  - 2.1.2 Central Outer Tracker . . . 22
  - 2.1.3 Calorimeters . . . 24
  - 2.1.4 Muon Detectors . . . 26
- 2.2 Trigger . . . 28
  - 2.2.1 The Lepton-SVT Trigger Path . . . 29

3 Likelihood B Flavor Tagging Using Opposite Side Muons **31**
- 3.1 Description of Likelihood Method . . . 32
- 3.2 J/ψ Muon Templates . . . 33
  - 3.2.1 Track-Stub Matching Templates . . . 34
  - 3.2.2 Calorimeter Templates . . . 36
- 3.3 Sources of Fake Muons . . . 38
- 3.4 Fake Muon Templates . . . 41
- 3.5 Likelihood of Muons from $J/\psi \to \mu\mu$ and Protons from $\Lambda \to p\pi$ . . . 44
- 3.6 Tagging Algorithm Performance . . . 46
  - 3.6.1 Trigger Side Selection . . . 49
  - 3.6.2 Opposite Side Selection . . . 50
  - 3.6.3 Likelihood on Opposite Side Muons . . . 51
  - 3.6.4 Dependence of Dilution on Likelihood . . . 51
  - 3.6.5 Dependence of Dilution on p_T^{rel} . . . 53
  - 3.6.6 Evaluation of $\varepsilon\mathcal{D}^2$ . . . 53
- 3.7 Likelihood Tagger in a Mixing Analysis . . . 54

4 Flavor Tagging Re-Calibration and Measurement of $B^0\bar{B}^0$ Oscillation Frequency in Semileptonic B Decays **59**
- 4.1 Data Selection and Signal Reconstruction . . . 59
  - 4.1.1 Data Selection . . . 59
  - 4.1.2 $D\ell$ Signal Reconstruction . . . 60
- 4.2 Flavor Tagging . . . 65
- 4.3 Sample Composition . . . 67

4.4 Decay Time of B Mesons . . . 71
4.5 Biases on the B Lifetime . . . 73
4.6 Maximum Likelihood Fit . . . 74
4.7 Toy Monte Carlo Tests of the Fitting Framework . . . 80
4.8 Fit Results from Data . . . 82
4.9 Systematic Uncertainties . . . 87
4.9.1 Prompt Background . . . 87
4.9.2 Sample Composition . . . 88
4.9.3 Resolution Scale Factor . . . 89
4.9.4 SVT Bias Curve . . . 89
4.9.5 Asymmetry of Prompt Background . . . 89
4.9.6 Fitting Method . . . 90
4.10 Evaluation of the Tagging Performance . . . 90

5 Search for $B_s^0\bar{B}_s^0$ Oscillations in Semileptonic $B_s^0 \to D_s^-\ell^+\nu$ Decays **93**
5.1 Data Selection and Signal Reconstruction . . . 93
5.1.1 $D_s^-\ell^+$ Signal Reconstruction . . . 93
5.2 Flavor Tagging . . . 97
5.3 Decay Time of B_s^0 Mesons . . . 99
5.4 Biases on the B_s^0 Lifetime . . . 104
5.5 Maximum Likelihood Fit . . . 104
5.6 Physics Backgrounds . . . 106
5.7 Tests of the Unbinned Likelihood Fitting Framework . . . 110
5.8 B_s^0 Lifetime Measurement . . . 112
5.8.1 Systematic Uncertainties on the B_s^0 Lifetime Measurement . . . 112
5.9 Amplitude Scan . . . 115
5.10 Evaluation of Systematic Uncertainties . . . 117
5.10.1 Prompt Background Fraction . . . 119
5.10.2 Prompt Background Asymmetry . . . 120
5.10.3 Physics Backgrounds . . . 120
5.10.4 Resolution Scale Factor . . . 121
5.10.5 Dilution Scale Factors . . . 121
5.10.6 Sample Composition . . . 121
5.10.7 SVT Bias . . . 121
5.10.8 B_s^0 Lifetime . . . 121
5.10.9 Event-by-Event Primary Vertex . . . 121
5.10.10 Non-Zero $\Delta\Gamma/\Gamma$ . . . 122
5.10.11 Detector Resolution Function . . . 122
5.10.12 Summary of Systematic Uncertainties . . . 122
5.11 Final Result - The Unblinded Amplitude Scan . . . 125

6 Conclusions **127**

A Appendix **130**
A.1 Improvement of Tagging Power by Using the Dilution Dependence on a Given Quantity . . . 130
A.2 Parameters Describing the p_T Dependence of the Track-Stub Matching Variables . . . 131
A.3 Consistency Checks of Fake Parameterizations . . . 131
A.4 Consistency Checks on Decay-in-Flight and Punch-Through Hypotheses . . . 132
A.5 Consistency Checks of the Likelihood Algorithm . . . 138
A.6 Trigger Side and Opposite Side Dilutions . . . 139
A.7 Detailed Calculation of $\varepsilon\mathcal{D}^2$ . . . 140
A.8 Correlations Between Likelihood Variables . . . 141
A.9 K-Factor Distributions . . . 143
A.10 SVT Trigger Bias Parameterizations . . . 144
A.11 Convolution Integrals . . . 144
A.12 Amplitude Scans in Each D_s^- Decay Mode . . . 145
A.13 Amplitude Scan Result . . . 146

1 Introduction

The neutral $B_s^0\bar{B}_s^0$ system is the subject of intense research, because the measurement of the oscillation frequency of B_s^0 mesons will provide crucial information about the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix [1, 3] and might reveal new physics phenomena to which B_s^0 oscillations are sensitive. Although different experiments have searched for B_s^0 mixing, the effect has not yet been observed. The current combined lower limit on the B_s^0 oscillation frequency Δm_s with contributions from the ALEPH, CDF, DELPHI, OPAL and SLD experiments is $\Delta m_s > 14.4$ ps^{-1} at 95% Confidence Level [1].

The focus of this thesis is a search for B_s^0 oscillations using semileptonic $B_s^0 \to D_s^+\ell^-\nu$ decays with the CDF detector at Fermilab. In Section 1.1 we give a brief theoretical overview of the CKM matrix and of the neutral $B^0/\bar{B}^0$ and $B_s^0/\bar{B}_s^0$ systems while in Section 1.2 we describe the experimental methods involved in a B mixing analysis. Section 2 contains a short description of the CDF detector. A detailed treatment of a multivariate B flavor tagging technique involving semileptonic $B \to \mu X$ decays is presented in Section 3. Here, we describe the calibration of the tagging algorithm using a high statistics inclusive semileptonic B sample. Section 4 describes the re-calibration of all the tagging algorithms used for the B_s^0 mixing analysis: opposite side muon and electron tagging as well as opposite side jet charge tagging. The tagging re-calibration is performed using semileptonic $B^{0/+} \to D\ell X$ decays. A measurement of the B^0 oscillation frequency Δm_d is also performed in order to validate the tagging performance. Finally, in Section 5 we present the search for B_s^0 mixing. Conclusions are discussed in Section 6.

1.1 Theoretical Overview

In the Standard Model of Elementary Particle Physics [2], the building blocks of matter, the quarks and the leptons, are arranged in three generations:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad \begin{pmatrix} t \\ b \end{pmatrix} \tag{1}$$

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}, \quad \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \quad \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \tag{2}$$

where the second and the third generations are heavier replicas of the first one. Interactions between matter particles are mediated by gauge bosons. Electromagnetic interactions between charged particles are mediated by photons, weak interactions take place via $W^\pm$ and Z^0 boson exchange. The strong interaction between quarks is carried by the octet of colored gluons and finally, the still-undiscovered gravitons are the exchange particles mediating the gravitational force that affects all massive particles.

The charged current weak interactions of the quarks are mediated by $W^\pm$ bosons and are described by the following Lagrangian:

$$\mathcal{L} = \frac{g}{2\sqrt{2}}[\,\bar{d}'\gamma^\mu(1-\gamma^5)u + \bar{s}'\gamma^\mu(1-\gamma^5)c + \bar{b}'\gamma^\mu(1-\gamma^5)t\,]\,W_\mu^- + \text{h. c.} \tag{3}$$

where g is a coupling constant related to the electron charge e and the Weinberg angle θ_W by $g = e/\sin\theta_W$. The Dirac spinors describing the quark mass eigenstates are denoted by u, d, c, s, t and b. The matrices γ^μ and $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ are 4×4 Dirac matrices. The operator $\frac{1-\gamma^5}{2}$ ensures that only the left-handed (right-handed) chirality component of the fermion (anti-fermion) wave functions participate in the weak interactions. $W^\pm$ are the charged bosons which along with the neutral Z^0 boson mediate the weak interactions. The weak (gauge) eigenstates d', s' and b' are different than the quark mass eigenstates d, s and b. The two sets of eigenstates are related by linear combinations:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad \text{where} \quad V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{4}$$

is the Cabibbo-Kobayashi-Maskawa (CKM) matrix [3]. By convention, this matrix acts on the down-type quarks, but in principle, it could also act on the up-type quarks. No physical observable would be affected by the chosen convention.

1.1.1 The CKM Matrix

The CKM matrix satisfies the unitarity relation:

$$V_{CKM} V_{CKM}^\dagger = \mathbb{1}_{3\times3} \tag{5}$$

or explicitly:

$$\sum_{j=1}^{3} V_{CKM}^{ij} V_{CKM}^{kj^*} = \delta_{ik}. \tag{6}$$

A 3×3 unitary matrix can be parameterized by three real rotational angles and six complex phases. Five out of the six complex phases can be removed by redefinitions of the quark fields. and thus, the CKM matrix is uniquely determined by four parameters: three angles θ_{12}, θ_{23} and θ_{13} and a complex phase δ_{13}. Using a common parameterization [4] the CKM matrix can be expressed as::

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \tag{7}$$

where $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ with $i, j = 1, 2, 3$. Using the Wolfenstein parameterization [5] the CKM matrix can be re-written as:

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{8}$$

where $\lambda = s_{12}$ is the sine of the Cabibbo angle θ_C which was initially introduced to describe the mixing between the first and the second quark generations. A, ρ and η are real parameters of order unity. This parameterization clearly shows how the

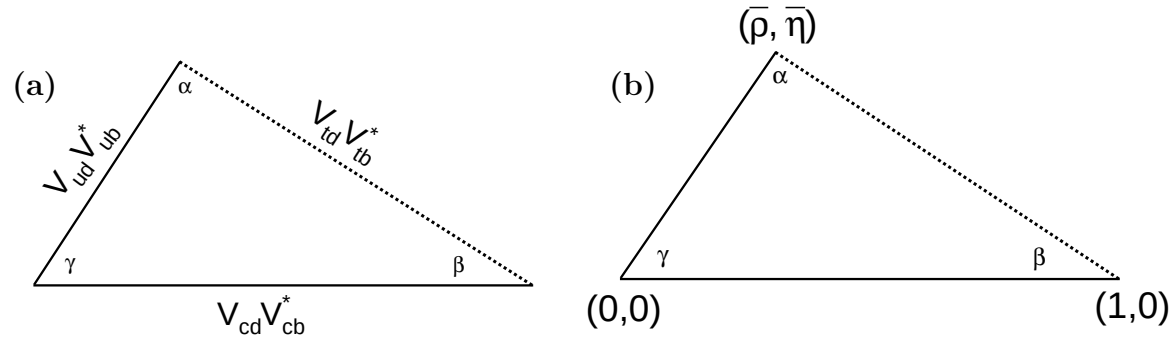

Figure 1: The geometric representation of Equation (9) in the complex plane (a) and a similar representation (b) after rescaling the unitarity triangle as described in the text.

cross talk between different quark families is suppressed by powers of λ. For example, $b \to c$ transitions are suppressed by λ^2 and $b \to u$ transitions are suppressed by λ^3. If we choose $i = 1$ and $k = 3$ in Equation (6) we obtain:

$$V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0. \tag{9}$$

In the complex plane, the above unitarity relation can be represented by a triangle as shown in Figure 1. Other choices of i and k in Equation (6) lead to different unitarity triangles. Using the parameterization from Equation (8), Formula (9) becomes

$$A\lambda^3(\rho + i\eta) - A\lambda^3 + A\lambda^3(1 - \rho - i\eta) = 0, \tag{10}$$

showing that all sides of this unitarity triangle are of order λ^3. After rescaling the triangle so that the side given by $V_{cd}V^*_{cb}$ becomes unity, the triangle vertices are defined by $(0, 0)$, $(1, 0)$ and $(\bar\rho, \bar\eta) = (1 - \lambda^2/2)(\rho, \eta)$ as also shown in Figure 1. The length of the side between $(1, 0)$ and $(\bar\rho, \bar\eta)$ may be approximated as

$$\left|\frac{V_{td}V^*_{tb}}{V_{cd}V^*_{cb}}\right| \;\approx\; \left|\frac{V_{td}V^*_{tb}}{V_{cd}V^*_{ts}}\right| \;\sim\; \left|\frac{V_{td}}{V_{ts}}\right| \tag{11}$$

since $|V_{cb}|$ and $|V_{ts}|$ are equal within very good approximation as seen from Equation (8), $V_{tb} \approx 1$ and $V_{cd} = \sin\theta_C$. As we will see in Section 1.1.2, the ratio of the parameters V_{td} and V_{ts} may be determined by measuring the oscillation frequencies of B^0 and B^0_s mesons.

The size of the CP violating effects due to the CKM mechanism is proportional to the area of the unitarity triangle [6]. If $\eta = 0$, the CKM matrix is real. In this case there are no CP violating effects due to the CKM mechanism and the triangle collapses to a segment on the real axis.

The CKM matrix contains fundamental information about the most basic interactions among elementary particles. Measuring its parameters and verifying the unitarity conditions given by Equation (5) is a high priority project in experimental particle physics. Figure 2 taken from Reference [7] shows the constraints on the apex of the unitarity triangle originating from different physics processes. Measurements of

the angle β from decays like $B^0 \to J/\psi K^0_S$ were recently performed by the B factories BaBar [8] and Belle [9]. The length of the side between $(0,0)$ and $(\bar{\rho},\bar{\eta})$ is given by measurements of $|V_{ub}/V_{cb}|$. Such measurements are obtained by analyzing semileptonic $b \to u\ell\bar{\nu}$ and $b \to c\ell\bar{\nu}$ decays. The best results come from the CLEO collaboration [10] as well as from from the B factories [11, 12]. The CP violating parameter in the neutral K system ε_K provides a hyperbolic constraint on the unitarity triangle apex as discussed in Reference [13]. Finally, the length of the side between $(1,0)$ and $(\bar{\rho},\bar{\eta})$ is constrained by measurements of the B^0 oscillation frequency Δm_d and lower limits on the B^0_s oscillation frequency Δm_s. The frequency Δm_d is precisely measured by many experiments and the world average is $\Delta m_d = (0.502 \pm 0.007)$ ps^{-1} [1]. Since the B^0_s mesons oscillate much faster than B^0 mesons, measuring Δm_s is more difficult. Until now, B^0_s oscillations have not been observed but lower limits on the oscillation frequency have been established by different experiments and the combined lower limit at 95% Confidence Level is $\Delta m_s > 14.4$ ps^{-1} [1]. Note, the mass difference Δm is usually given in ps^{-1}, where 1 ps^{-1} corresponds to 6.58×10^{-4} eV/c^2.

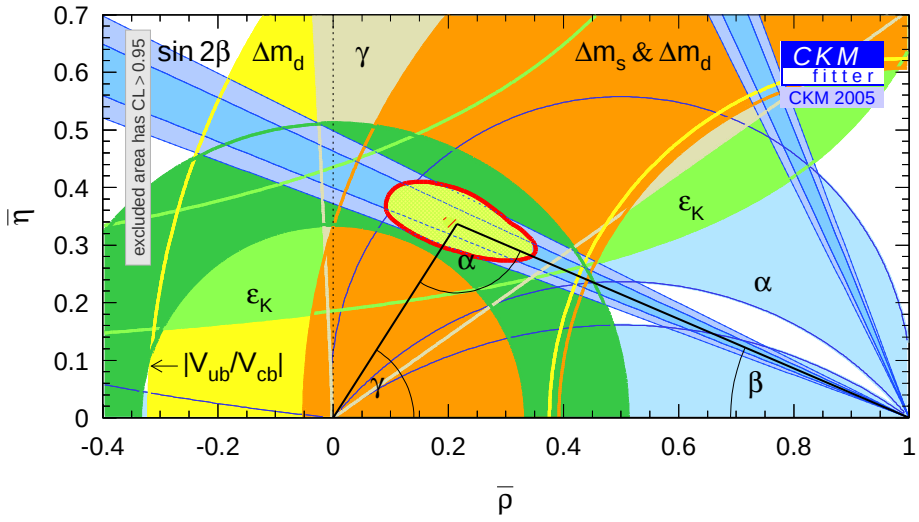

Figure 2: Constraints on the unitarity triangle from different physics processes.

1.1.2 Oscillations of Neutral B Mesons

Neutral B^0 mesons are composed of an $\bar{b}$ anti-quark and a d quark. Similarly, B^0_s mesons are composed of an $\bar{b}$ anti-quark and an s quark:

$$|B^0\rangle = |\bar{b}d\rangle\,, \qquad |\bar{B}^0\rangle = |b\bar{d}\rangle, \tag{12}$$

$$|B^0_s\rangle = |\bar{b}s\rangle\,, \qquad |\bar{B}^0_s\rangle = |b\bar{s}\rangle. \tag{13}$$

The B^0, $\bar{B}^0$ and the B^0_s, $\bar{B}^0_s$ states are superpositions of corresponding light $|B_L\rangle$ and heavy $|B_H\rangle$ mass eigenstates. Neglecting CP violation effects, the mass eigenstates can be written as:

$$|B_L\rangle = \frac{1}{\sqrt{2}}\left(\,|B^0_q\rangle + |\bar{B}^0_q\rangle\,\right) \tag{14}$$

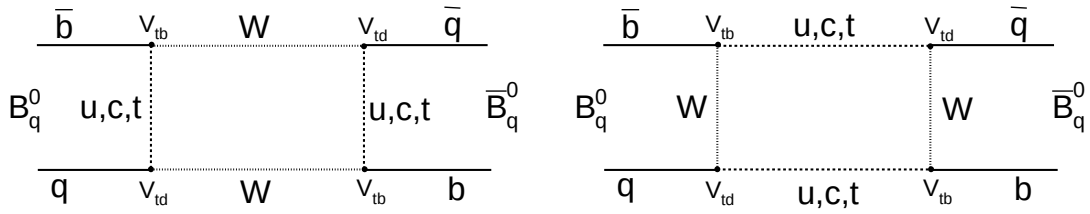

Figure 3: Lowest order Feynman diagrams that induce $B_q^0 \bar{B}_q^0$ oscillations.

$$|B_H\rangle = \frac{1}{\sqrt{2}} \left(|B_q^0\rangle - |\bar{B}_q^0\rangle \right) \tag{15}$$

where $q = d, s$. The mass difference between the heavy and light mass eigenstates is denoted by $\Delta m_q = m_H - m_L$ and the total decay width difference by $\Delta\Gamma_q = \Gamma_H - \Gamma_L$, where m_H, m_L, Γ_H and Γ_L are the masses and the decay widths of the heavy and light states, respectively. The time evolution of the $|B_q^0\rangle$ and $|\bar{B}_q^0\rangle$ states is governed by the following equations:

$$|B_q^0(t)\rangle = g_+(t)\ |B_q^0\rangle + g_-(t)\ |\bar{B}_q^0\rangle \tag{16}$$

$$|\bar{B}_q^0(t)\rangle = g_+(t)\ |\bar{B}_q^0\rangle + g_-(t)\ |B_q^0\rangle \tag{17}$$

where $g_\pm = \frac{1}{2}(e^{-im_H t - \Gamma_H t/2} \pm e^{-im_L t - \Gamma_L t/2})$. The probability that an initial state $|B_q^0\rangle$ remains unchanged after time t is:

$$\mathcal{P}_{\text{unmix}} = |\langle B_q^0 | B_q^0(t)\rangle|^2 = |g_+(t)|^2 = \frac{\Gamma_q}{2} e^{-\Gamma_q t}\ (\cosh\frac{\Delta\Gamma_q t}{2} + \cos\Delta m_q t) \tag{18}$$

while the probability that an initial state $|B_q^0\rangle$ oscillates into a $|\bar{B}_q^0\rangle$ state after time t is:

$$\mathcal{P}_{\text{mix}} = |\langle \bar{B}_q^0 | B_q^0(t)\rangle|^2 = |g_-(t)|^2 = \frac{\Gamma_q}{2} e^{-\Gamma_q t}\ (\cosh\frac{\Delta\Gamma_q t}{2} - \cos\Delta m_q t) \tag{19}$$

where Γ_q is defined as $\Gamma_q = (\Gamma_H + \Gamma_L)/2$. The lifetime and the total decay width of a particle are related by $\tau = 1/\Gamma$.

In the Standard Model $B_q^0 \bar{B}_q^0$ oscillations are described, at lowest order in perturbation theory, by the Feynman diagrams shown in Figure 3. These box diagrams involve exchange of W bosons and top quarks. The oscillation frequency depends on the ratio $\frac{m_q^2}{m_W^2}$, where m_q is the mass of the up-type quark involved in the mixing amplitude and m_W is the mass of the W boson. Because the t quark is much heavier than the u and c quarks, the t quark exchange dominates these transitions. Although c and u quarks may be exchanged, their contributions to the total oscillation amplitude is negligible.

Theoretical calculations of these box diagrams [14] find that the mass difference between the heavy and light mass eigenstates is given by:

$$\Delta m_q = \frac{G_F^2 m_W^2 \eta_B m_{B_q} B_{B_q} f_{B_q}^2}{6\pi^2} S_0(\frac{m_t^2}{m_W^2}) |V_{tq}^* V_{tb}|^2 \tag{20}$$

where G_F is the Fermi constant of the weak interaction, m_W and m_t are the mass of the W boson and the mass of the t quark, respectively. η_B is a QCD correction of order unity which accounts for exchange of gluons running between the quark lines. m_{B_q}, f_{B_q} and B_{B_q} are the mass, bag parameter and the weak decay constant of the B_q^0 meson. The latter two parameters are calculated in lattice QCD [15]. The function S_0 is approximated by $S_0(x) \approx 0.784\, x^{0.76}$ [16].

In principle, using Equation (20), V_{td} and V_{ts} can be inferred individually from measurements of Δm_d and Δm_s. However, the determination of V_{td} from the above equation is affected by large theoretical uncertainties related to the poor knowledge of the B^0 meson weak decay constant f_{B^0} and the bag parameter B_{B^0} which limit the direct extraction of the CKM matrix element from a measurement of Δm_d. In the ratio between Δm_s and Δm_d, several theoretical uncertainties cancel providing a tighter constraint on the side of the unitarity triangle defined by the vertices $(1, 0)$ and $(\bar{\rho}, \bar{\eta})$:

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m_{B_s^0}}{m_{B^0}} \xi^2 \left| \frac{V_{ts}}{V_{td}} \right|^2 \tag{21}$$

where ξ^2 is calculated from lattice QCD calculations and found to be of order unity $\xi^2 = 1.56 \pm 0.26$ [17], resulting in $\xi = 1.25 \pm 0.10$.

Thus, $B^0\bar{B}^0$ and $B_s^0\bar{B}_s^0$ flavor oscillations measure the Cabibbo-Kobayashi-Maskawa matrix elements V_{td} and V_{ts} and the ratio $\Delta m_d/\Delta m_s$ will ultimately determine one of the sides of the unitarity triangle. A measurement of Δm_s will be the next crucial test of the Standard Model probing whether the obtained result will fit to the current constraints on the CKM triangle. It is worth noting that physics with B_s^0 mesons is unique to the Tevatron until the start of the Large Hadron Collider.

1.2 Experimental Methods

In general, a measurement of the time dependence of neutral B meson oscillations requires the knowledge of:

- the proper decay time t of the B meson
- the flavor of the B meson (either B or $\bar{B}$) at production and decay in order to determine whether the B meson has oscillated or not.

1.2.1 Proper Decay Time

At high energy colliders such as the Tevatron, the B meson proper decay time t can be obtained from a measurement of the distance X^B between the B production vertex and the B decay vertex. The proper time t is related to the decay distance X^B by:

$$t = \frac{X^B}{\beta\gamma} = X^B \frac{m(B)}{p\,(B)}, \tag{22}$$

where β is the speed of the B meson, γ is the Lorentz boost factor and $m(B)$ and $p\,(B)$ are the mass and the momentum of the B meson. The above relation is projected in

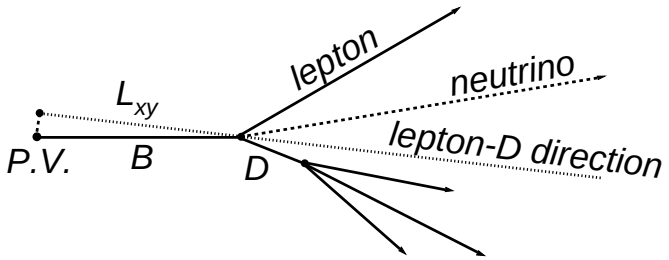

Figure 4: Schematic representation of a semileptonic $B \to D\ell\nu$ decay.

the plane transverse to the beam line:

$$t = X^B_{xy} \frac{m(B)}{p_T(B)}, \tag{23}$$

since at CDF the transverse distance X^B_{xy} and the transverse momentum $p_T(B)$ are measured more accurately than X^B and $p\,(B)$.

In semileptonic $B \to D\ell\nu$ decays, one cannot reconstruct the full momentum of the B meson because the neutrino is not detected. Instead, the combined momentum of the $D\ell$ pair is used to calculate the "pseudo proper decay time":

$$t^* = X^B_{xy} \frac{m(B)}{p_T(D\ell)}. \tag{24}$$

The transverse decay length L^B_{xy} is defined as the displacement $\vec{X}^B_{xy}$ in the transverse plane of the secondary B vertex from the primary event vertex (P.V.) projected onto the $D\ell$ momentum as shown in Figure 4:

$$L^B_{xy} = \frac{\vec{X}^B_{xy} \cdot \vec{p_T}(D\ell)}{|\vec{p_T}(D\ell)|}. \tag{25}$$

The decay length L^B_{xy} is a signed variable which can be negative for the configuration where the particle seems to decay before the point where it is produced. From Monte Carlo simulation studies we have determined that the difference between $|\vec{X}^B_{xy}|$ and L^B_{xy} is so small, that for any practical purpose we can use L^B_{xy} as an excellent approximation of $|\vec{X}^B_{xy}|$. The B meson pseudo proper decay time is then calculated from the measured decay length L^B_{xy} as:

$$t^* = X^B_{xy} \frac{m(B)}{p_T(D\ell)} \approx L^B_{xy} \frac{m(B)}{p_T(D\ell)}. \tag{26}$$

The actual proper decay time is obtained using the following correction:

$$t = t^* \times K \tag{27}$$

where the factor K is the ratio between the observed $D\ell$ momentum the true B momentum:

$$K = \frac{p_T(D\ell)}{p_T(B)}. \tag{28}$$

Although it is impossible to determine the correction factor K on an event-by-event basis, we can still determine the K-factor distribution $\mathcal{F}(K)$ from Monte Carlo simulations (see Section 4.3).

The probability distribution $\mathcal{P}(t)$ of the proper time t is hypothesized to be a decay exponential convoluted with a detector resolution function assumed to be Gaussian:

$$\mathcal{P}(t) = \int_0^\infty \frac{1}{\tau} e^{-t_0/\tau} \frac{1}{\sqrt{2\pi}\sigma} e^{-(t-t_0)^2/2\sigma^2} dt_0, \tag{29}$$

where t_0 is the true decay time and σ is the width of the Gaussian detector resolution function. Given the probability distribution functions $\mathcal{P}(t)$ and $\mathcal{F}(K)$ and using Equation (27), we determine the probability distribution of the pseudo proper decay time:

$$\mathcal{P}^*(t^*) = \int_{-\infty}^{\infty} dt \int_{K_{min}}^{K_{max}} dK\, \mathcal{P}(t)\, \mathcal{F}(K)\, \delta(t^* - \frac{t}{K}) = \int_{K_{min}}^{K_{max}} \mathcal{P}(Kt^*)\, K\, \mathcal{F}(K)\, dK \tag{30}$$

where we have used the fact that the correlation between the K-factor and the proper decay time is negligible and that $K > 0$.

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

$$\sigma_t = \sqrt{\left(\sigma(L_{xy}^B)\frac{m(B)}{p_T(B)}\right)^2 \;+\; t^2 \left(\frac{\sigma(p_T(B))}{p_T(D\ell)}\right)^2}. \tag{31}$$

The proper decay time resolution σ_t depends on the uncertainty $\sigma(L_{xy}^B)$ to infer the decay length from the primary to the B decay vertex and on the $p_T(B)$ momentum resolution $\sigma(p_T(B))$. Note that the latter uncertainty scales with the proper decay time t, while the vertexing resolution is independent of proper time and only adds a constant uncertainty. To resolve the fast B_s^0 oscillations the error on the decay time should be smaller than the oscillation period. In a search for B_s^0 mixing, events with small decay time are thus the most sensitive to oscillations if $\sigma(p_T)$ is significant as is the case in semileptonic decays.

1.2.2 Flavor Tagging

Events containing a neutral B meson are classified as mixed or unmixed based on the sign of the production and decay tagging variables. To determine the flavor of the B meson it is necessary to determine its b quark content (b or $\bar{b}$). Knowing the mixing status of each event along with the proper decay time, Equations (18) and (19) determine the oscillation frequency Δm_q. The flavor of a neutral B meson at decay is given by the charges of the decay products. For example, a B^0 meson decay into $D^-\pi^+$ or $D^-\ell^+X$ while a $\bar{B}^0$ meson decays into $D^+\pi^-$ or $D^+\ell^-X$.

The procedure of determining the B flavor at production is called B flavor tagging. The figure of merit used to compare different flavor tags is the effective tagging efficiency $\varepsilon(1-2\,p_W)^2$, where the efficiency ε represents the fraction of events in which a flavor tag is found and p_W is the mistag probability indicating the fraction of

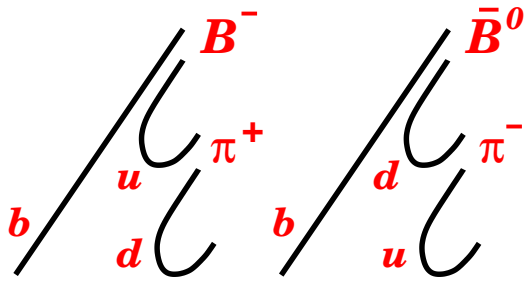

Figure 5: Example showing how in the fragmentation process a b quark hadronizes into either a B^+ or a $\bar{B}^0$ meson.

events with a wrong flavor tag. The mistag probability is related to the dilution $\mathcal{D}$, another quantity used to express the power of a flavor tag:

$$\mathcal{D} = 1 - 2\,p_W. \tag{32}$$

The dilution $\mathcal{D}$ is defined as the number of correctly tagged events N_R minus the number of incorrectly identified events N_W divided by the sum:

$$\mathcal{D} = \frac{N_R - N_W}{N_R + N_W}. \tag{33}$$

It can be misleading to quantify the tagging power with an expression called "dilution". A flavor tagging algorithm which always returns the correct tag has a dilution of one, while a flavor tagging algorithm giving the correct tag 50% of the cases has a dilution of zero. This means that a tagging algorithm with a large dilution is desirable, while a small dilution characterizes a less powerful tagging method.

Several techniques can be used to determine the neutral B meson flavor at production. One technique uses the correlations between the charge of the initial $b/\bar{b}$ quark and the charges of the fragmentation particles around the B meson direction or the decay products of B^{**} states. This method is called "same side tagging" [18] and is exemplified in Figure 5. The closest fragmentation track to the B meson has a charge correlated with the charge of the b quark. Positive pions are associated with B^- and B^0 mesons while negative pions are associated with B^+ and $\bar{B}^0$ mesons. The correlation is lost if the fragmentation track is neutral. Another possibility is to use the fact that the $b/\bar{b}$ quark of interest is always produced together with another $\bar{b}/b$ quark of opposite charge. The second $\bar{b}/b$ is called "the opposite side b". Determining the flavor of the opposite side B hadron in the event provides information on the B meson of interest.

The opposite side B flavor can be identified by the charge of a lepton from a semileptonic B decay, by the charge of the kaon from the subsequent charm meson decay or by analyzing the jet associated with the opposite side b quark. Figure 6 is a schematic view of a $b\bar{b}$ event showing different tagging possibilities: same side tagging and opposite side tagging using leptons, kaons or jets. In this analysis we use only opposite side lepton and jet charge tagging algorithms. In a direct $b \to \ell^-$ transition, the charge of the lepton reflects the b flavor. However, other processes can

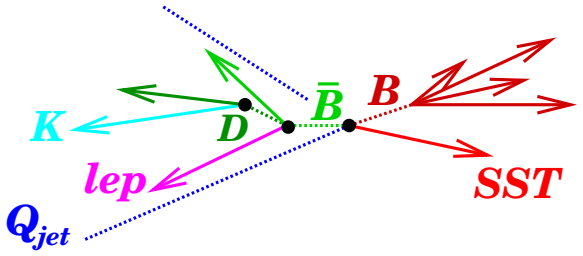

Figure 6: Schematic view of a $b\bar{b}$ event with different ways of inferring the B flavor at production: same side and opposite side (lepton, kaon, jet) flavor tagging.

also give a lepton in the final state such as cascade decays $b \to c \to \ell^+$ resulting in a wrong-sign tag, right-sign cascade decays $b \to W^-$ with $W^- \to \bar{c} \to \ell^-$, semileptonic τ decays $b \to W^- \to \tau^- \to \ell^-$ or $b \to J/\psi X \to \ell^\pm$ decays giving both sign leptons. A misidentified fake lepton would provide a random tag. Another opposite side tagging method determines the flavor of the opposite side b quarks by analyzing the jets associated with these quarks. The sum of the charges of all tracks in a jet weighted by the momenta and the displacements of the tracks is correlated to the charge of the initial $b/\bar{b}$ quark that produced the jet. The methods used in this analysis to tag the production b quark flavor (e.g. opposite side lepton or jet charge tagging) are discussed in more detail in Section 3 and Section 4.2.

1.2.3 Overview of B Mixing Analysis

In a B mixing measurement, the oscillation frequency Δm_q is extracted from the data using a maximum likelihood method. In the following, we discuss the essential steps for a B^0 mixing analysis determining Δm_d. We use the example of an analysis where same-sign (opposite-sign) events describe mixed (unmixed) events as would be the case in an analysis using semileptonic B decays with opposite side lepton tags. We start with a pure sample of B^0 mesons and assume that the lepton tag is always correct. In this case, an event with an opposite-sign lepton pair signals an unmixed event, while a same-sign lepton pair indicates a mixed event. In this case the probabilities for an opposite-sign event $\mathcal{P}_{OS}$ and a same-sign event $\mathcal{P}_{SS}$ are directly related to the mixing probabilities:

$$\mathcal{P}_{\rm SS}(t) = \mathcal{P}_{\rm mix}(t) \qquad \text{and} \qquad \mathcal{P}_{\rm OS}(t) = \mathcal{P}_{\rm unmix}(t). \tag{34}$$

The mixing probabilities are obtained from Equations (18) and (19) by neglecting the width difference between the heavy and light mass eigenstates:

$$\mathcal{P}_{\rm unmix/mix}(t) = \frac{1}{2}\,\Gamma \mathrm{e}^{-\Gamma t}\,(1 \pm \cos\Delta m\, t). \tag{35}$$

The observable asymmetry $\mathcal{A}_{\rm mix}$ is defined as:

$$\mathcal{A}_{\rm mix}(t) = \frac{\mathcal{P}_{\rm unmix}(t) - \mathcal{P}_{\rm mix}(t)}{\mathcal{P}_{\rm unmix}(t) + \mathcal{P}_{\rm mix}(t)} = \cos\Delta m\, t. \tag{36}$$

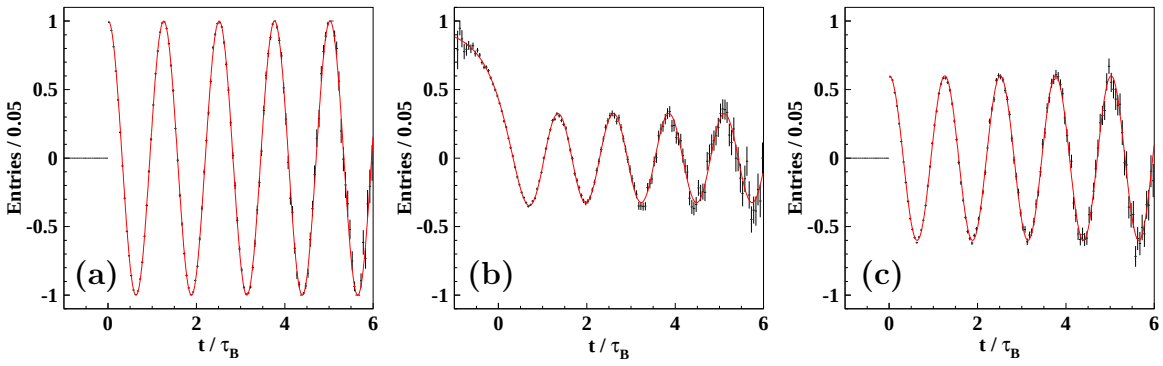

Figure 7: Illustration of various detector and analysis effects on the mixing amplitude $\mathcal{A}_{\rm mix}$: (a) ideal case, (b) finite decay length resolution, (c) mistag probability. The exponential decay terms cancel in the asymmetry ratio, but one can see that the statistical errors increase with decay time.

Next, we introduce several effects that we will have to include in a realistic mixing analysis. First, we consider the mixing asymmetry $\mathcal{A}_{\rm mix}$ as a function of the decay time for $\Delta m = 5$ ps^{-1} assuming an ideal case with perfect tagging, ideal proper time resolution and no background as shown in Figure 7(a). Then, we introduce a vertexing resolution function $\mathcal{G}$ which smears the decay time measurement and effectively reduces the amplitude of the oscillation as shown in Figure 7(b). The oscillation amplitude is also reduced by mistag effects and the resulting amplitude is proportional with the tagging dilution $\mathcal{D}$. The effect of mis-tagging is exemplified in Figure 7(c). The effect of introducing a momentum resolution, for example through a K-factor distribution $\mathcal{F}(K)$ in the case of a partially reconstructed decay, is also considered and shown in Figure 8.

The event selection efficiency might not be constant as a function of the decay time as described in Section 4.5. We account for effects that would bias the lifetime distribution by introducing an acceptance function $\mathcal{E}$. Including all these effects, the opposite/same-sign probability would be modified as (see Section 4.6):

$$\mathcal{P}_{\rm OS/SS}(t^*) \sim \int dK\ K\mathcal{E}(Kt^*) \int \frac{1}{2}\,\Gamma\,{\rm e}^{-\Gamma t}\,(1 \pm \mathcal{D}\,\cos(\Delta m_d\,t))\mathcal{G}(t'-t,\sigma)dt\ |_{t'=Kt^*}. \quad (37)$$

In a real measurement there will be background such as combinatorial background under a charm signal. We define $\mathcal{P}^{\rm sig}_{\rm SS}$ and $\mathcal{P}^{\rm sig}_{\rm OS}$ as $\mathcal{P}_{\rm SS}$ and $\mathcal{P}_{\rm OS}$ in Equation (37) and obtain:

$$\begin{aligned} \mathcal{P}_{\rm SS}(t) &= (1-f_{\rm bg})\,\mathcal{P}^{\rm sig}_{\rm SS}(t) + f_{\rm bg}\,f_{\rm SS}\,\mathcal{P}^{\rm bg}(t), \\ \mathcal{P}_{\rm OS}(t) &= (1-f_{\rm bg})\,\mathcal{P}^{\rm sig}_{\rm OS}(t) + f_{\rm bg}\,(1-f_{SS})\,\mathcal{P}^{\rm bg}(t), \end{aligned} \quad (38)$$

where $f_{\rm bg}$ is the fraction of background in a given sample, $f_{\rm SS}$ is the fraction of same-sign events in the background, while $\mathcal{P}^{\rm bg}$ is the probability function that describes the proper time t distribution of the background events. To extract the value Δm_q of the oscillation frequency, the following likelihood function is minimized:

$$\mathcal{L} = -\sum_{\rm SS} \ln \mathcal{P}^{\rm SS}(t^*) - \sum_{\rm OS} \ln \mathcal{P}^{\rm OS}(t^*). \quad (39)$$

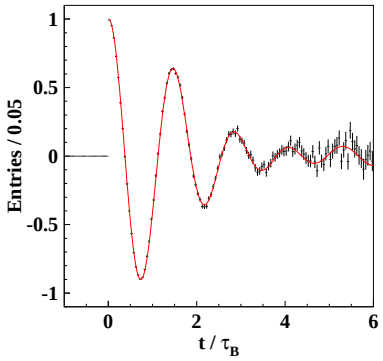

Figure 8: Illustration of decay length and momentum resolution plus mis-tagging and background effects on the mixing asymmetry $\mathcal{A}_{\rm mix}$.

The statistical significance $\mathcal{S}$ of a B^0 or B_s^0 oscillation signal can be approximated as discussed in Reference [19]:

$$\mathcal{S} \approx \sqrt{\frac{N}{2}}\, f_{\rm sig}\, \mathcal{D}\, {\rm e}^{-(\Delta m\, \sigma_t)^2/2} = \sqrt{\frac{\varepsilon \mathcal{D}^2}{2}} \frac{S}{\sqrt{S+B}}\, {\rm e}^{-(\Delta m\, \sigma_t)^2/2} \tag{40}$$

where N and $f_{\rm sig}$ are the number of tagged candidate events and the fraction of signal in the selected sample and σ_t is the resolution on proper decay time. S and B are the numbers of signal and background events. The sensitivity $\mathcal{S}$ decreases rapidly as Δm increases. This dependence is controlled by σ_t which means that excellent proper time resolution is needed to explore high frequency B_s^0 oscillations. The sensitivity also increases with the tagging power $\varepsilon \mathcal{D}^2$, the signal fraction and the total number of events.

1.2.4 The Amplitude Method

In B_s^0 mixing analyses, the amplitude method [19] is used to set limits on Δm_s and combine results from different experiments when no oscillation signal is observed. An amplitude $\mathcal{A}$ is introduced in the expressions describing the mixed and unmixed probabilities:

$$\mathcal{P}_{\rm unmix}^{B_s} = \frac{1}{2}\Gamma {\rm e}^{-\Gamma t}\, (1 + \mathcal{A} \cos \Delta m_s t) \tag{41}$$

and similarly:

$$\mathcal{P}_{\rm mix}^{B_s} = \frac{1}{2}\Gamma {\rm e}^{-\Gamma t}\, (1 - \mathcal{A} \cos \Delta m_s t). \tag{42}$$

The amplitude method works as follows. A B_s^0 oscillation amplitude $\mathcal{A}$ and its error $\sigma_{\mathcal{A}}$ are extracted as a function of a fixed test value of Δm_s using a likelihood method in analogy to Equation (39) based on the physics functions defined in Equation (41) and (42). To a very good approximation, the statistical uncertainty on $\mathcal{A}$ is Gaussian and equal to the inverse of the significance $1/\mathcal{S}$ defined in Equation (40). If Δm_s equals its true value $\Delta m_s^{\rm true}$, the amplitude method expects $\mathcal{A} = 1$ within the total uncertainty $\sigma_{\mathcal{A}}$. If Δm_s is tested far from its true value, a measurement

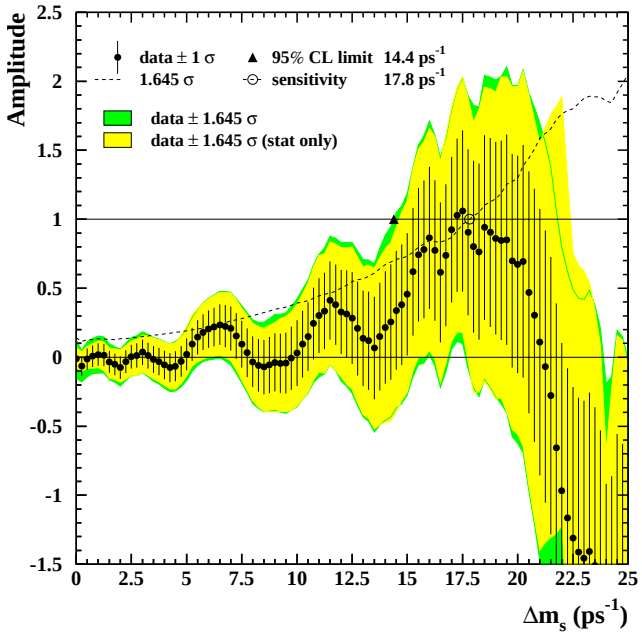

Figure 9: Combined measurements of the B_s^0 oscillation amplitude as a function of Δm_s, including all published result from LEP, SLD and the Tevatron as compiled in Reference [1].

consistent with $\mathcal{A} = 0$ is expected. A value of Δm_s can be excluded at 95% Confidence Level if $\mathcal{A} + 1.645\,\sigma_{\mathcal{A}} \leq 1$. Because of proper time resolution, the quantity $\sigma_{\mathcal{A}}(\Delta m_s)$ is an increasing function of Δm_s. It is therefore expected that individual values of Δm_s can be excluded up to $\Delta m_s^{\rm sens}$, where $\Delta m_s^{\rm sens}$ is called the sensitivity of the analysis defined by $1.645\,\sigma_{\mathcal{A}}(\Delta m_s^{\rm sens}) = 1$. If no signal is observed, a lower limit is set with 95% Confidence Level at the mixing frequency for which $\mathcal{A} + 1.645\,\sigma_{\mathcal{A}} = 1$. An interesting feature of the amplitude methods is that the results from different analyses and experiments can be combined (after accounting for correlations between the systematic errors) by simple averaging of different amplitude spectra.

The combined measurements of the B_s^0 oscillation amplitude as a function of Δm_s, including all published results from LEP, SLD and the Tevatron Run I provide a lower limit on the B_s^0 mixing frequency of $\Delta m_s > 14.4$ ps^{-1} at 95% Confidence Level with a sensitivity of 17.8 ps^{-1} [1] as shown in Figure 9.

2 Experimental Apparatus

In this analysis we use data collected between February 2002 and August 2004 with the upgraded Collider Detector at Fermilab (CDFII). The total integrated luminosity is about 355 pb^{-1}. The CDFII detector records events from proton (p) antiproton ($\bar{p}$) collisions at a center-of-mass energy $\sqrt{s} = 1.96$ TeV. The Fermilab Tevatron, a 1 km radius synchrotron, accelerates the protons and antiprotons up to 980 GeV. Both protons and antiprotons are separated in 36 bunches which circulate the Tevatron ring in opposite directions and collide every 396 ns at two interaction points where the CDFII and D0 detectors are located.

In the CDFII coordinate system, the z-direction is defined by the proton beam line. The plane perpendicular to the beam line is called the transverse or $r\phi$-plane. The momentum of a particle projected onto the transverse plane is called the transverse momentum p_T. The polar angles θ is defined with respect to the z-direction. The azimuthal angle ϕ together with the radius r define positions in the transverse plane. A commonly used quantity in CDF analyses is the pseudorapidity $\eta = -\ln[\tan(\theta/2)]$ with $\eta = 0$ defining the transverse plane.

2.1 The CDFII Detector

The CDFII detector [20, 21] is a multipurpose detector designed to record the interaction products from $p\bar{p}$ collisions. Figure 10 shows a schematic view of the detector. The silicon vertex detector [22] is the innermost tracking device that provides three dimensional track reconstruction. It is used to identify tracks and vertices displaced with respect to the primary interaction vertex. Such displaced tracks and vertices are in general associated with decays of long lived hadrons containing charm (c) and bottom (b) quarks. Surrounding the silicon vertex detector is the Central Outer Tracker (COT) [23], a multi wire drift chamber located inside a superconducting solenoid that generates a 1.4 Tesla magnetic field along the z-axis. Charged particles have helical trajectories in uniform magnetic fields. The transverse momentum of a charged particle is determined by measuring the curvature of the trajectory. Right outside the solenoid, the time-of-flight detector (TOF) [24] provides particle identification for low momentum particles. Surrounding the drift chamber are the electromagnetic and hadronic calorimeters which measure the energy depositions of particles emerging from the $p\bar{p}$ interactions. The drift chambers and scintillator counters for muon identification are the outermost detectors.

2.1.1 Silicon Detector

The silicon detector is composed of a total of seven layers of silicon strip sensors as shown in Figure 11. Figure 12 shows that the silicon layers are concentric around the beam line forming barrel shaped structures. Along the z-axis, the silicon detector is divided in three barrels. The silicon layers may be either single-sided which means that only one side of the silicon layer is segmented into strips or double sided in which case, both sides are segmented. In a double-sided layer, one side has strips parallel to

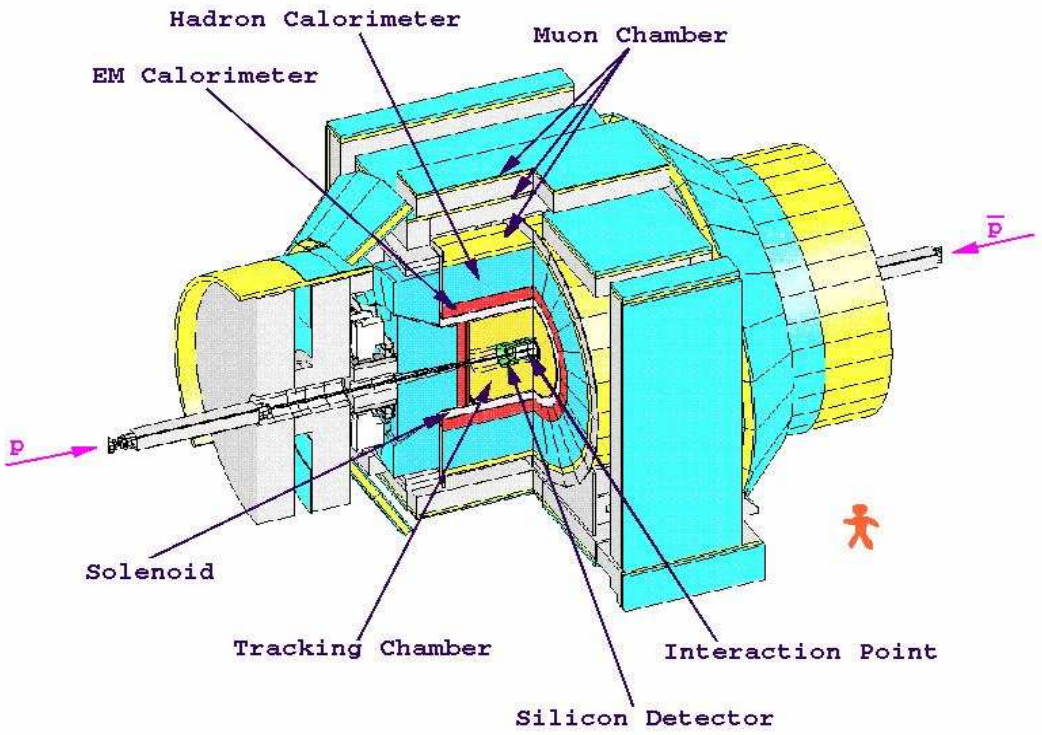

Figure 10: Schematic view of the CDF detector.

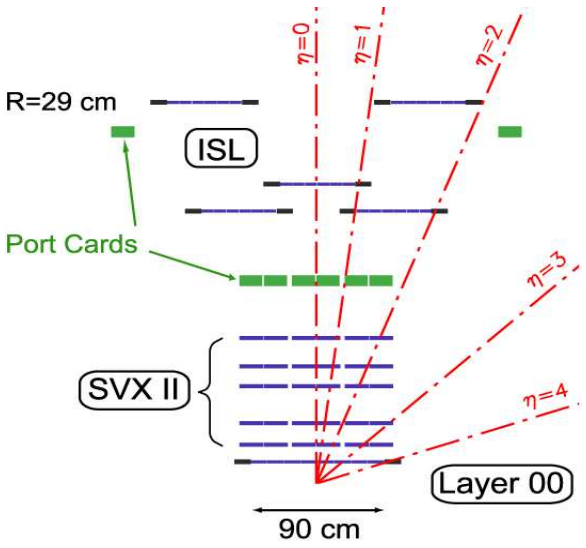

Figure 11: A side view of half of the CDFII silicon system on a scale in which the z coordinate is highly compressed.

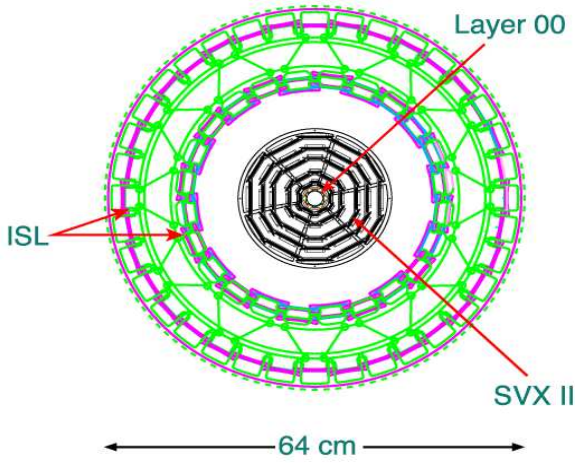

Figure 12: An end view of the CDFII silicon system including the SVX II cooling bulkheads and ISL support structure.

the beam line direction (axial sensors) providing information about the $r\phi$ position and the other side has strips that make either a small angle or a 90° angle with the beam line direction (stereo sensors) and thus, providing information about the z position of the track hits. The innermost silicon layer L00 of the silicon detector is single sided with axial sensors and is located immediately outside the beam pipe at a radius of 1.7 cm away from the beam line. Outside L00 is the SVXII component of the silicon detector which consists of five concentric cylindrical layers of double-sided silicon sensors with radii between 2.5 and 10.6 cm. Finally, the Intermediate Silicon Layers (ISL) consist of double sided stereo sensors located at radii of 22 cm in the central region and 20 and 28 cm in the forward regions. The z coverage of the silicon detector is about 90 cm corresponding to a pseudorapidity coverage of $|\eta| < 2$.

2.1.2 Central Outer Tracker

The COT is a cylindrical drift chamber that provides accurate measurements of the particle momenta. The drift chamber is filled with a mixture of 50% Argon and 50% Ethane. There are more than 32 thousands sense wires inside the COT. The sense wires are grouped in concentric superlayers as shown in Figure 13. Each superlayer is sectioned in ϕ into separate cells as shown in Figure 14. A cell is defined as one sense plane with two adjacent grounded field sheets. The sense plane is composed of 40 μm gold plated tungsten wires. The eight superlayers of the COT alternate between stereo and axial, beginning with superlayer 1 which is a stereo layer. Roughly half of the sense wires are parallel (axial) to the z-direction providing information on the $r\phi$ position, while the other half make 2° angles (stereo) with the z direction providing information on the z-position in addition to $r\phi$ information. The active volume of the drift chamber covers the region $|z| < 155$ cm corresponding to $|\eta| < 1$ extending radially from 40 to 140 cm.

The trajectory of a charged particle in a uniform magnetic field is a helix. A helix is completely determined by five parameters [25]. At CDF, the following quantities

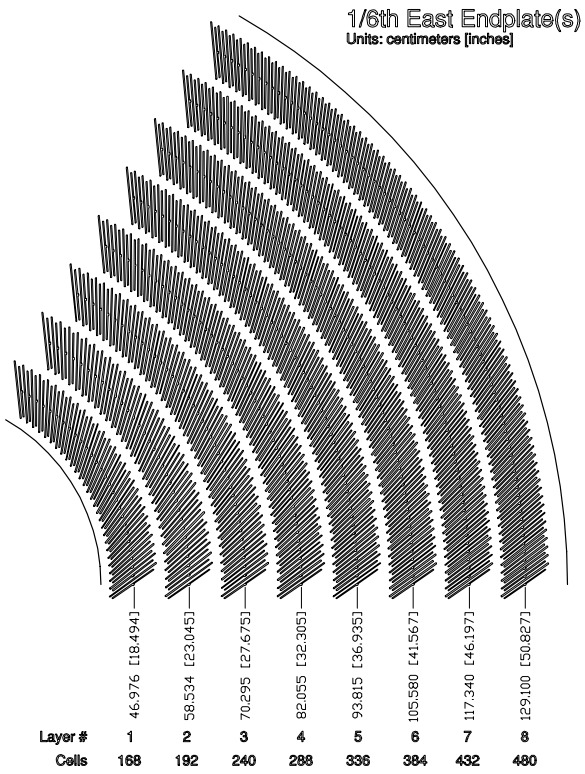

Figure 13: Geometrical arrangement of the eight superlayers of the CDF drift chamber.

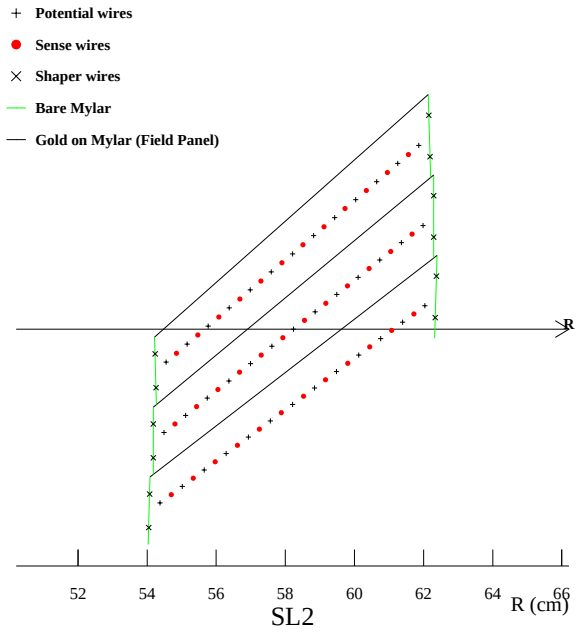

Figure 14: Nominal cell layout for superlayer 2.

are used to parameterize the helical particle trajectories. Three parameters define the position of the point of closest approach: the position along the beam line z, the impact parameter of the track d_0 which is the transverse distance from the primary vertex to the point of closest approach, and the azimuthal angle ϕ. The particle transverse momenta are determined from the curvature C of the reconstructed track and finally, the direction of the momentum vector at the point of closest approach projected into the rz plane is given by the polar angle θ.

Charged particles passing through the SVX and COT cause ionization of the silicon material and of the gas that fills the drift chamber volume. The charge detected by the silicon sensors or the COT wires provide information about the position where a charged particle passed through the detector. These positions are called "hits". The amount of charge deposited on the sense wires is also used for particle identification through specific energy loss (dE/dx). A charged particle produces a series of hits in the detector along its trajectory. The trajectory reconstructed by following the hits of the particle in the drift chamber and in the silicon detector is called a "track". At the Tevatron, a typical event contains several tens of tracks and the track reconstruction involves pattern recognition algorithms that search for helix trajectories. Initially, the pattern recognition algorithm reconstructs line segments from hits in each superlayer. First, the line segments from the axial layers that are tangent to a common circle are linked together to form a track candidate and the hit positions are fit to a circle. Secondly, the line segments in stereo layers are linked to the two dimensional track and a helix fit is performed. In the next step, the track reconstructed in COT is extrapolated into the silicon detector. Hits that are consistent with the track are added successively. A window around the COT track is determined based on the error matrix of the COT track. If a hit in the outer silicon layer is found within the window, it is added to the track. The track is refit and a new error matrix of the five track parameters is determined. Using the new error matrix, a new search window is established which is then used to add hits from the next silicon layer. The procedure is repeated for each silicon layer. If a certain silicon layer has no hits inside the search window, the pattern recognition algorithm proceeds to the next layer. If silicon hits may be attached to the COT track in different combinations corresponding to different tracks, the track with the largest number of silicon hits is selected. A track with both COT and SVXII hits is reconstructed only if at least three $r\phi$ hits in the silicon detector are associated with the original COT track. Tracks formed with hits from the silicon detector only are called silicon stand alone tracks.

2.1.3 Calorimeters

Outside the solenoid are the electromagnetic calorimeters followed by the hadronic calorimeters. Comparison of the energy depositions in electromagnetic and hadronic calorimeters provides separation between electrons and photons (which deposit most of the energy in the electromagnetic calorimeter) and hadrons (which deposit most of the energy in the hadronic calorimeter). Both electromagnetic and hadronic calorimeters are segmented in ϕ and η towers so that each tower points to the interaction region. This arrangement is called a projective tower geometry. The calorime-

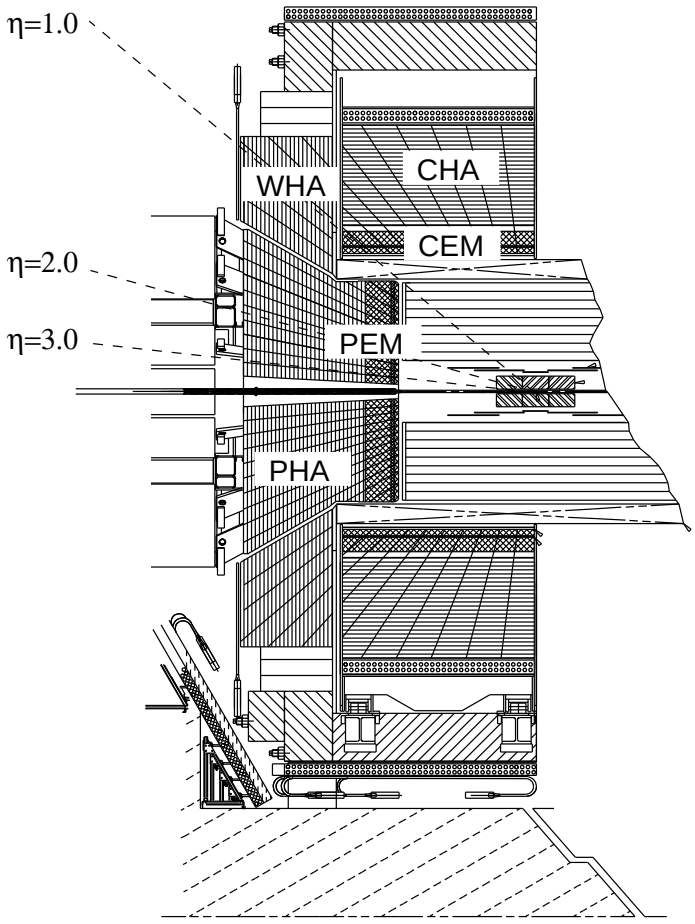

Figure 15: Schematic view of the CDF detector showing different calorimeter subsystems. The central electromagnetic (CEM) and hadronic (CHA) calorimeters extend up to $\eta \sim 1$. The plug electromagnetic (PEM) and hadronic (PHA) systems extend between $\eta \sim 1.1$ and $\eta \sim 3$. The wall hadronic calorimeter (WHA) is located between the central and the forward calorimeters ($\theta \sim 30^\circ$).

ters are divided into central and forward subsystems corresponding to $|\eta| < 1$ and $1.1 < |\eta| < 3.6$ respectively. Figure 15 is a schematic view of one half of the CDF detector displaying the components of the CDF calorimeter.

The central electromagnetic and hadronic calorimeters covers the region $|\eta| < 1$. Each tower is 15 degrees in azimuth and 0.11 in pseudorapidity. The electromagnetic towers consist of alternating layers of lead and scintillator and have a depth of ~ 18 radiation lengths of material. The hadronic towers are composed of alternating layers of iron and scintillator and correspond to ~ 4.7 interaction lengths of material. The wall hadronic calorimeter is situated between the central and the forward calorimeters as seen in Figure 15 and has a similar design as the central hadronic calorimeter.

The plug calorimeters cover the region $1.1 < |\eta| < 3.6$ and is segmented in 7.5° towers for the $\eta < 2.1$ region and 15° towers for the $\eta > 2.1$ region. The plug

electromagnetic calorimeter corresponds to $\sim$ 23 interaction lengths while the plug hadronic calorimeter has a depth of $\sim$ 7 interaction lengths. For each calorimeter subsystem, the scintillation light is directed to photomultiplier tubes that measure the energy depositions.

The shower maximum detector (CES) [26] is a multi-wire proportional chamber located inside the central electromagnetic calorimeter. The CES detector measures the coordinates of showers produced by electrons and photons with 2-3 mm accuracy. It is used for identification of electrons and photons by matching the position of electromagnetic showers with incident tracks. The transverse shower profile is used to separate photons from neutral pions and the pulse height helps to identify electromagnetic showers.

The central pre-radiator detector (CPR) [27] is located between the solenoid coil and the central electromagnetic calorimeter at a radius of $\sim$ 168 cm. It is a single plane multi-wire proportional chamber with 32 sense wires per wedge along the z-direction. The chamber is $\sim$ 116 cm long. It provides separation between electrons and minimum ionizing particles like muons or hadrons.

2.1.4 Muon Detectors

In general, electrons and hadrons deposit most of their energy in the calorimeter material and do not reach the muon detectors. On the other hand, muons which are minimum ionizing, leave only a small amount of their energy in the calorimeters and given enough momentum, pass through calorimeters and reach the muon chambers. The calorimeters act as shielding for all particles except muons. It is possible that a small fraction of hadrons that interact late in the calorimeters will produce secondary particles that reach the muon chambers and are identified as muons. Hadrons detected by the muon detectors are called "fake" muons. A detailed discussion on fake muons is given in Section 3.

The CDF detector contains four muon systems: the central muon detector (CMU), the central muon extension (CMX), the central muon upgrade (CMP) and the intermediate muon upgrade (BMU). Figure 16 shows the $\eta\phi$ coverage of each muon system.

The CMU detector is located outside the central calorimeter at a radius of 347 cm from the beam line. It is composed of drift cells. Each cell covers 12.6° in ϕ and contains seven wires 226 cm long and parallel to the beamline. Between drift cells there are 2.4° gaps which limit the CMU coverage to about 84% in ϕ. Each wedge is segmented azimuthally into three 4.2° modules. Each module consists of four rectangular drift cells. A track segment detected in these drift chambers is called a CMU stub. The pseudorapidity coverage of CMU relative to the interaction region is $|\,\eta\,| < 0.63$. The CMU shielding provided by the calorimeter corresponds to 5.5 pion interaction lengths.

The CMP detector consists of drift chambers located behind an additional 60 cm of iron shield. It covers the region $55° < \theta < 90°$. The CMP chambers are rectangular, single-wire drift tubes configured in four layers. These chambers are 640 cm long and arranged axially to form a box around the central detector. Since the CMP geometry is not cylindrical, the η coverage varies slightly as a function of ϕ. In

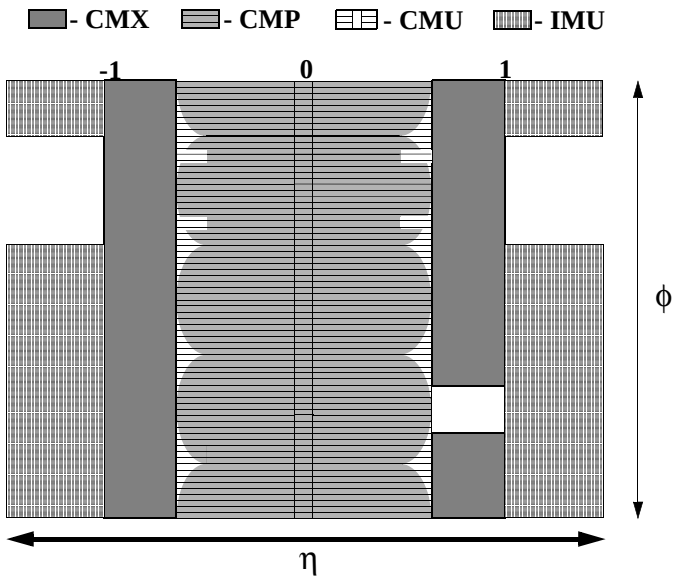

Figure 16: Coverage in $\eta\phi$ of each muon system.

average the pseudorapidity coverage is $|\eta| < 0.6$. The calorimeter and the additional iron shielding correspond to 7.8 hadronic absorption length. Muons which have stubs in both CMU and CMP are called CMUP muons and are, in general, very pure muons.

The CMX detector consists of conical sections of drift tubes located at each end of the central detector and covers the region $42° < \theta < 55°$ which corresponds to a pseudorapidity coverage of $0.6 < |\eta| < 1$. The calorimeter, magnet yoke of the detector and the steel support structure provide shielding of about 6.2 pion interaction lengths. The CMX drift tubes are 180 cm long. They are grouped in four layers of twelve tubes for each 15° ϕ sector.

Finally, the BMU detector consists of drift chambers similar to the CMP ones. The chambers are $\sim$ 30 cm long and each chamber covers 12.5° in azimuth. The BMU covers the rapidity region $1 < |\eta| < 1.5$ and is shielded by material corresponding to $6.2 - 20$ pion interaction lengths.

At CDF, "muon objects" are reconstructed by matching the stub position in the muon detector with a track measured in COT and extrapolated to the muon chambers. To reconstruct the muon stubs, linear segments are formed by using alternate layers that are within 7.5 cm from each other which corresponds to a maximum angle of 65° relative to the radial direction. Then, the pattern recognition algorithm searches in the remaining layers for hits that are within 0.5 cm from the line segment. The procedure is repeated until the optimal set of hits is found. The muon stub is obtained by fitting the hits using a least square method to a line segment. A stub must contain hits in at least three out of four layers. Stubs that are reconstructed in the CMU are matched to tracks with a minimum p_T of 1.3 GeV/c. The tracks are extrapolated to the CMU using a simplified geometry model of the muon candidate's motion in the non-uniform magnetic field of the calorimeter. The distance $\Delta r\phi$ (also called ΔX) in the $r\phi$-plane between the track projected to the muon chambers and

the muon stub is required to satisfy $\Delta r\phi < 30$ cm. Similarly, for CMP, CMX and BMU muons, the minimum muon transverse momenta are 2.2 GeV/c, 1.5 GeV/c and $\sim$ 2.5 GeV/c respectively while the matching in the $r\phi$ plane between the extrapolated track and the muon stub must be less than 60 cm, 50 cm and 90 cm, respectively. The minimum momentum and matching requirements are chosen to maximize the muon identification efficiency while maintaining high muon purity at the same time. The angle $\Delta\Phi$ between the direction of the muon stub and the direction of the extrapolated track projected onto the $r\phi$-plane as well as the longitudinal distance (along z direction) ΔZ between the stub and track are also measured (except for CMP which does not measure ΔZ). These variables could also be used for muon identification as described in Section 3.

2.2 Trigger

At the Tevatron Run II, the $p\bar{p}$ crossing rate is about 2.5 MHz corresponding to a bunch spacing of 396 ns. Each bunch crossing that produces a $p\bar{p}$ collision with interaction products recorded by the detector is called an "event". It is impossible to record every event at this high rate. In fact, the rate at which the CDF experiment can write events to tape is about 100 Hz. The purpose of the trigger is to select only events that are important for specific physics studies. Only the selected events are recorded while the rest are discarded. The CDF triggers achieve a total rate reduction of almost 10^5 by employing a three level architecture in which each level provides a rate reduction which allows processing in the next level with minimal deadtime.

The total interaction cross section at the Tevatron is about one thousand times larger than the $b\bar{b}$ production cross section. The CDF detector has different B triggers designed to select the $b\bar{b}$ events. The "lepton-SVT" trigger is designed to select semileptonic decays of B mesons like $B \to \ell DX$ by triggering on signatures with a high momentum lepton and a displaced track coming from the B daughters. The "two track trigger" is geared toward selecting fully hadronic B decays like $B \to D\pi$. The main requirements of this trigger are two tracks with large impact parameters and transverse momenta larger than 2 GeV/c. Finally, many B events are selected by the "di-muon" trigger which requires a pair of muons (either CMU-CMU or CMU-CMX) with the invariant mass close to the J/ψ mass.

The Level 1 trigger is implemented using custom designed hardware. Its output rate is about 25 kHz. This rate reduction is achieved by analyzing only partial information from different detectors subsystems like calorimeters, tracking chamber and muon systems. Events are selected based on the energy/momentum of electrons, muons or jets, missing energy or simple tracks as well as based on the number of such objects. Events accepted by Level 1 are further processed by Level 2. The Level 2 trigger also uses custom hardware which provides an output rate of about 400 Hz. Events are selected based on more precise information from the same detectors as in the Level 1 case, but also on information from other detector subsystems. Very important for B physics, the Level 2 trigger takes decisions based on track displacement as measured by the silicon vertex tracker (SVT) [28]. Heavy hadrons containing c and b quarks are long lived and travel in the laboratory frame distances of the order of

~ 1 mm before they decay. The decay products of such long lived particles will have in general large impact parameters with respect to the primary interaction point. The SVT system uses information from both the COT and SVXII as well as the position of the beam line to determine online track parameters like d_0, ϕ and p_T with precisions comparable to the offline analysis. The total d_0 resolution as measured by the SVT is about 47 μm for tracks with $p_T > 2$ GeV/c as shown in Figure 17. The error on the beam line position is ~ 33 μm from which a 30 μm intrinsic SVT resolution is inferred.

Finally, the Level 3 trigger reduces the output rate down to ~ 100 Hz by analyzing fully reconstructed events. This trigger is software based and is implemented on a farm with about 300 CPU's.

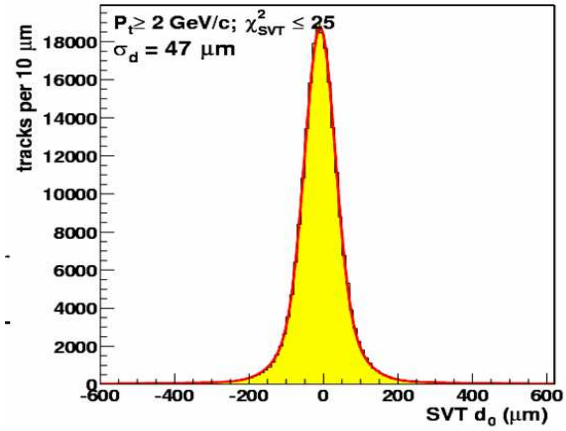

Figure 17: Impact parameter resolution as measured by the silicon vertex trigger (SVT) for tracks with $p_T > 2$ GeV/c.

2.2.1 The Lepton-SVT Trigger Path

A trigger path is a set of trigger requirements that a given event must pass. In this analysis we mainly use a dataset collected using the lepton-SVT trigger path, which we describe in the following.

At Level 1 a high p_T lepton must be identified. The lepton could be either an electron or a muon. The electron must have a transverse momentum larger than 4 GeV/c. The transverse momentum is determined by the extremely fast tracker (XFT) [29], a device that uses COT hit information from the outer four axial superlayers to perform fast track reconstruction. The transverse energy in the central electromagnetic calorimeter associated with the track must also be larger than 4 GeV. The ratio between the hadronic and electromagnetic energy depositions must satisfy $E_{HAD}/E_{EM} < 0.125$. The muon must have stubs in both the CMU and CMP detectors and the transverse momentum must be larger than 4 GeV/c.

At Level 2 the electron requirements are reconfirmed and in addition, the electron must have 2 GeV energy measured by CES. There are no requirements on the muon at Level 2. In addition to the 4 GeV/c lepton, a displaced track with 120 μm $< d_0 <$

1 mm and $p_T > 2$ GeV/c must be found by the SVT. For electrons only, the transverse angle between the lepton momentum and the SVT track momentum must satisfy $2^\circ < \Delta\phi\,(\text{lepton, SVT track}) < 90^\circ$.

At Level 3 the lepton identification criteria are tightened. In the electron case, the transverse distance Δx between the shower centroid as measured by CES and the extrapolated track position is required to be less than 5 cm, while the corresponding longitudinal distance Δz must be less than 3 cm. χ^2 comparisons between the transverse and longitudinal profiles of the showers in the CES and the corresponding shapes obtained from test beam electrons are performed and the differences are expressed by χ^2_x and χ^2_z which must be less than 10 and 15, respectively. Another quantity that is used for electron identification is the lateral shower sharing $Lshr$ [30] which measures the difference between the observed sharing of energy deposition between towers in the CEM and the deposition expected from real electromagnetic showers. To pass the lepton-SVT trigger path, an electron must satisfy $Lshr < 0.2$. In the muon case, the matching ΔX between the CMU stub and the extrapolated track must be less than 15 cm while the same quantity corresponding to CMP stubs must be less than 20 cm. For both electrons and muons, the transverse angle $\Delta\phi$ between the lepton and the SVT track is required to be satisfy $2^\circ < \Delta\phi\,(\text{lepton, SVT track}) < 90^\circ$, while the invariant mass of the lepton-SVT pair must be less than 5 GeV.

3 Likelihood B Flavor Tagging Using Opposite Side Muons

A crucial task in a B mixing analysis is to determine the flavor of the neutral B mesons at production and decay time. The flavor of a neutral B meson is given by its b quark content. A B^0 meson contains a $\bar{b}$ quark, while a $\bar{B}^0$ meson contains a b quark. In a mixed (unmixed) event the flavor of the B meson at production is different than (the same as) the flavor of the B meson at decay time. In $p\bar{p}$ collisions a b quark is always produced together with a $\bar{b}$ anti-quark. Both b quarks hadronize into B mesons (or baryons). The B hadron that produces decay products satisfying the trigger requirements is called the "trigger B" while the other B hadron in the event is called the "opposite side B". We study opposite side lepton tagging, where the lepton is a muon.

The performance of this tagging method is established on an inclusive semileptonic sample selected by the lepton-SVT trigger described in Section 2.2.1. This trigger requires a displaced track as signature of a long lived particle in the event and a 4 GeV/c lepton as signature of a semileptonic decay. The charge of the trigger lepton provides an estimate of the trigger side B flavor at decay time. However, the lepton-SVT data sample 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. A background subtraction procedure has been established [31] to determine the B purity of the sample. This procedure allows one to study the performance that an opposite side flavor tag would achieve in a pure b sample and to quantify the tagging performance in terms of dilution $\mathcal{D}$ and $\varepsilon\mathcal{D}^2$. This method will be described in detail in Section 3.6.

To determine the flavor of the B meson at production, we search for a muon from the semileptonic decay of the opposite side B hadron in the event. The charge of this lepton is correlated with the flavor of the B hadron: a negative charge lepton ℓ^- comes from a $b \to c\,\ell^-\bar{\nu}X$ transition, while a positive charge one ℓ^+ originates from a $\bar{b}$ quark.

The subject of our study is to combine information from track-stub matching quantities and calorimeter energies into a global likelihood to discriminate real muons from fake muons. This type of multivariate analysis aims to identify muons without loss in efficiency from the application of cuts while achieving a better separation between real muons and fakes. The muon likelihood is constructed so that it approaches one for real muons and zero for fake muons. We determine that the tagging dilution increases with muon likelihood. High purity muons have high dilution while the low purity ones have low dilution. Also, as established in Run I [32], the tagging dilution of opposite side leptons increases with the transverse momentum p_T^{rel} of the lepton with respect to the axis of the jet in which the lepton is found. A schematic view of how p_T^{rel} is calculated is shown in Figure 18. Jets produced by heavy b quarks are spread more widely in the plane transverse to the b direction than jets originating from c quarks. Consequently, on average, leptons from b quark decays will have larger p_T^{rel} than leptons from c quark decays.

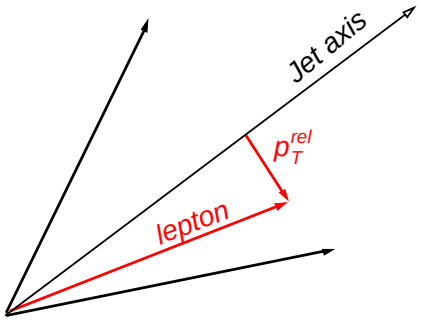

Figure 18: Transverse momentum p_T^{rel} of the opposite side lepton with respect to the axis of the jet in which the lepton is found.

The tagging power is given by $\varepsilon\mathcal{D}^2$, which is proportional to the inverse square of the expected statistical uncertainty in a mixing measurement. To minimize the statistical error, for example on Δm_d, $\varepsilon\mathcal{D}^2$ has to be maximized. As discussed in Appendix A.1, the tagging power given by $\varepsilon\mathcal{D}^2$ can be significantly improved by using the dilution dependence on different quantities. In the case of the muon tagging algorithm we use the dilution dependence on muon likelihood and p_T^{rel} to improve $\varepsilon\mathcal{D}^2$ by $\sim 45\%$.

3.1 Description of Likelihood Method

At CDF, muon objects consist of a stub in a muon chamber associated with a track that extrapolates from the drift chamber to the vicinity of the stub. A muon object does not always correspond to a real muon. For example, high momentum hadrons that interact late in the calorimeter may produce decay particles that reach the muon detectors and are identified as muon objects. These objects are called "fake muons". We use the track-stub matching quantities, ΔX, $\Delta\Phi$ and ΔZ introduced in Section 2.1.4 along with energy depositions in the electromagnetic and hadronic calorimeters, E_{EM} and E_{HAD}, to discriminate real muons from fake muons. We expect that the matching distributions are narrow for muons and wide for fakes. Also, the hadronic energy E_{HAD} would be centered around a minimum ionizing peak for real muons, while fake muons would be characterized by a flat hadronic energy distribution.

To study the distributions of discriminating variables for real muons, we use muons from $J/\psi \to \mu^+\mu^-$ decays. We select pions from $K_S^0 \to \pi^+\pi^-$, kaons and pions from $D^0 \to \pi^+K^-$ and protons from $\Lambda \to p\pi^-$ to study the behavior of fake muons. We study separately the real and fake muon distributions for different muon types. The muon type labels the detector that identified the muon stub. We study five muon types: CMU, CMP, CMUP (muons with stubs in both CMU and CMP), CMX and BMU.

Using these samples of real and fake muons, we study the distributions of separating variables for both muons and fakes. For computational convenience we param-

eterize these distributions with empirical functional forms and we treat the fitting functions as probability distribution functions (PDF). We obtain five signal PDF's: S_X, S_Φ, S_Z, S_{EM} and S_{HAD} and five corresponding background PDF's: B_X, B_Φ, B_Z, B_{EM} and B_{HAD}. The likelihood that a muon object is a real muon can be written as:

$$S = \prod_i S_i \tag{43}$$

while the likelihood that the muon object is a fake muon is:

$$B = \prod_i B_i \tag{44}$$

where $i = X, \Phi, Z, EM, HAD$. The final likelihood estimator is constructed as:

$$\mathcal{L} = \frac{S}{S+B}. \tag{45}$$

We expect that muon objects with values of $\mathcal{L}$ close to unity are real muons while values of $\mathcal{L}$ close to zero indicate most probably fake muons.

In general, variables can be lost from the likelihood estimator when they are measured to be outside the range in which the PDF's are defined. These ranges are chosen large enough so that almost no real muon could fall outside the range. Muon objects that miss one variable can still be real muons, but we found that the ones that miss two or more variables are mostly fakes.

3.2 J/ψ Muon Templates

To study the distributions of discriminating variables for real muons, we use muons from $J/\psi \to \mu^+\mu^-$ decays. We analyze two data samples. In one sample events are selected by a trigger which requires two muons with invariant mass around the J/ψ mass. The two triggered muons are either two CMU muons or a CMU muon and a CMX muon. This sample is rich in CMU, CMUP and CMX muons but depleted of CMP and BMU muons. To enhance our samples of CMP and BMU muons we also reconstruct J/ψ states from the muon-SVT sample. One muon from the $J/\psi \to \mu^+\mu^-$ is the trigger muon, while the second muon could be detected by any muon system. We use a standard CDF software package [33] to reconstruct the J/ψ states. We select tracks with $p_T > 1$ GeV/c and $\eta < 2$ to form the J/ψ candidates. Each track corresponding to a CMU, CMP, CMUP or CMX muon is required to have at least 20 axial and 20 stereo COT hits and at least 3 $r\phi$ silicon hits. Tracks corresponding to BMU muons are required to have at least 10 axial and 10 stereo COT hits and at least 3 $r\phi$ silicon hits. The tracks are refitted accounting for the energy loss corrections for the muon hypothesis. Layer 00 silicon hits are dropped before refitting. The J/ψ daughters are required to be associated with muon objects and the distance along the z-axis between the two tracks must be less than 5 cm. The di-muon invariant mass is required to be between 2.8 and 3.4 GeV/c^2. For the candidates remaining after the mass cuts, a three dimensional vertex fit [34] of the two tracks is required to converge. In the vertex fit, the track parameters of the

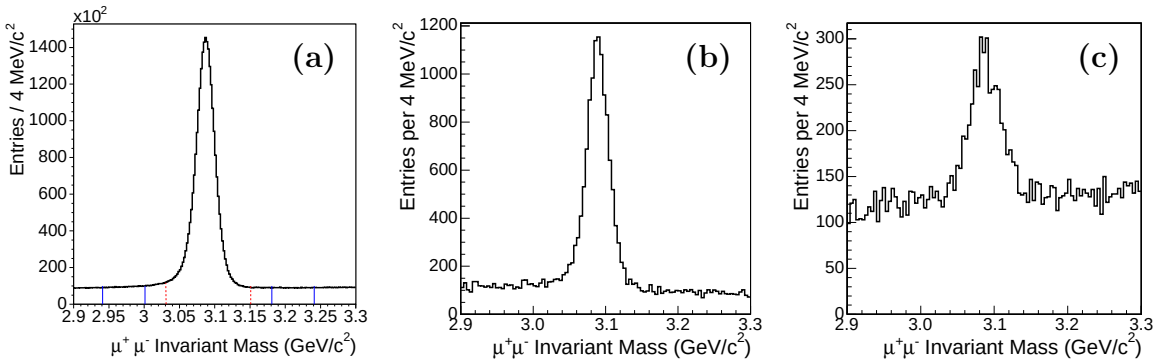

Figure 19: J/ψ mass peaks from (a) the di-muon dataset where the muon types are CMU, CMUP or CMX, and from the muon-SVT dataset where one leg of the J/ψ is (b) a CMP muon or (c) a BMU muon. Dashed vertical lines indicate the signal region and the solid vertical lines indicate the side-band regions.

two muons are allowed to float within their errors and the χ^2 is minimized under the assumption that the two tracks originate from a common (J/ψ) vertex. After the vertex fit, the di-muon invariant mass range is tightened to be between 2.9 and 3.3 GeV/c^2.

The J/ψ mass peaks are shown in Figure 19. We obtain large samples of CMU ($\approx$ 3 millions), CMUP ($\approx$ 1 million) and CMX ($\approx$ 0.5 million) muons but smaller numbers of CMP and BMU muons. We collect a total of about fifteen thousand CMP and two thousand BMU muons as seen in Figure 19. The high background of the J/ψ peak where one leg is a BMU muon reflects the large fraction of fake muons in this muon detector.

3.2.1 Track-Stub Matching Templates

The distributions of the track-stub matching variables ΔX, $\Delta\Phi$ and ΔZ depend on the muon momentum. At large momenta, the Coulomb multiple scattering is small and the extrapolated track matches well the muon stub. As the momentum decreases the multiple scattering effects are larger and the distance between the extrapolated track and the stub increases and the matching is less accurate. The p_T dependence of the matching quantities ΔX, $\Delta\Phi$ and ΔZ has been studied extensively in References [35, 36, 37]. Figure 20 illustrates the p_T dependence of matching variable for CMU muons. The dependences of each matching variable of CMU, CMP and CMX muons is studied in Reference [35] where the width of each matching variable is parameterized as a function of the muon transverse momentum p_T. The functional forms of the parameterizations $\sigma_{\Delta X}(p_T)$, $\sigma_{\Delta\Phi}(p_T)$, and $\sigma_{\Delta Z}(p_T)$ are given by:

$$\sigma_{\Delta X}^{CMU}(p_T) = \frac{a + e^{b+cp_T}}{p_T} + d, \tag{46}$$

$$\sigma_{\Delta\Phi,\Delta Z}^{CMU}(p_T) = a + \frac{b}{p_T} + \frac{c}{p_T^2}, \tag{47}$$

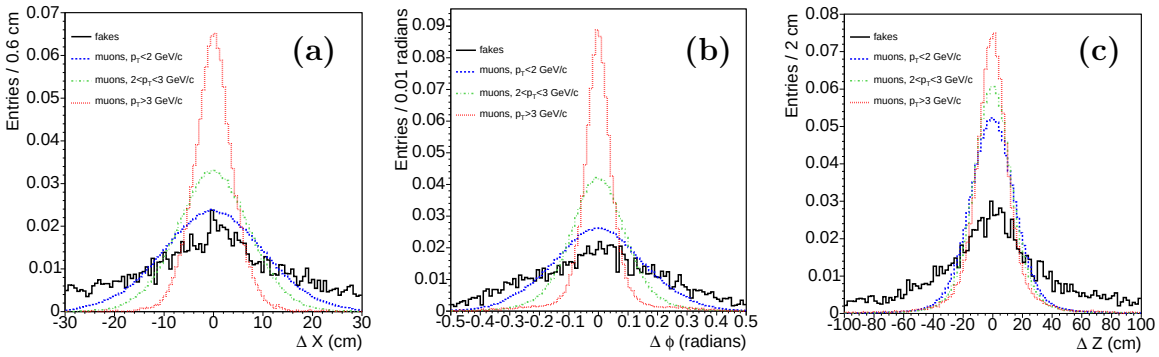

Figure 20: CMU track-stub matching variables: (a) ΔX, (b) $\Delta\Phi$ and (c) ΔZ. Solid lines correspond to proton distributions. Muon distributions are shown for three p_T ranges: $p_T < 2$ GeV/c (dashed), $2 < p_T < 3$ GeV/c (dash-dotted), $p_T > 3$ GeV/c (dotted). The width of the muon matching distributions decreases with increasing transverse momentum.

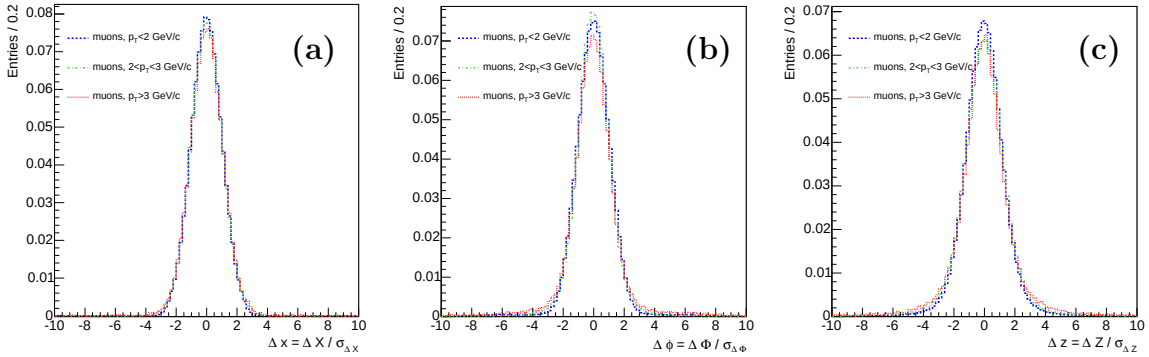

Figure 21: Scaled variables (a) $\Delta x = \Delta X/\sigma_{\Delta X}(p_T)$, (b) $\Delta\phi = \Delta\Phi/\sigma_{\Delta\Phi}(p_T)$ and (c) $\Delta z = \Delta Z/\sigma_{\Delta Z}(p_T)$ for CMU muons in three p_T ranges: $p_T < 2$ GeV/c (dashed), $2 < p_T < 3$ GeV/c (dash-dotted) and $p_T > 3$ GeV/c (dotted).

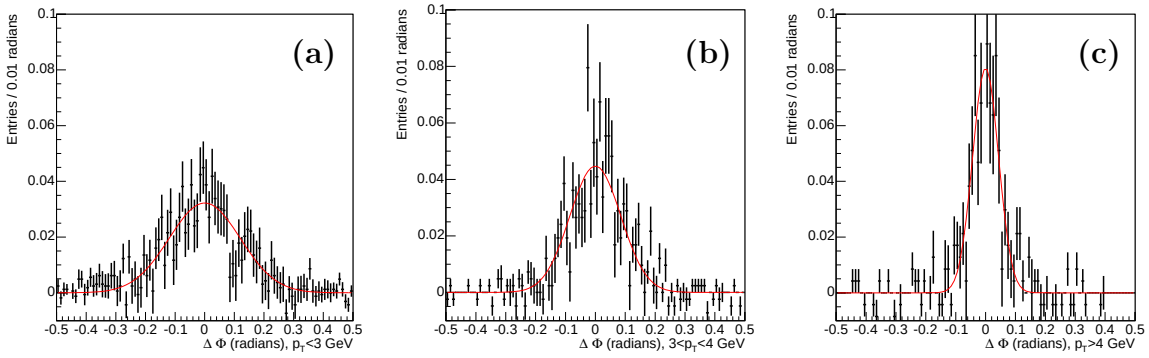

Figure 22: $\Delta\Phi$ of BMU muons in three p_T ranges: (a) $p_T < 3$ GeV/c, (b) $3 < p_T < 4$ GeV/c and (c) $p_T > 4$ GeV/c.

$$\sigma_{\Delta X}^{CMP}(p_T, \phi) = A(p_T) + B(p_T) \times (\sin(4\phi + \pi/2) - 1), \quad (48)$$

where $A(p_T) = \frac{a+e^{b+cp_T}}{p_T}$ and $B(p_T) = \mathrm{Min}(e + fp_T, 0)$,

$$\sigma_{\Delta\Phi,\Delta Z}^{CMP}(p_T) = a + \frac{b}{p_T} + \frac{c}{p_T^2}, \quad (49)$$

and

$$\sigma_{\Delta X,\Delta\Phi,\Delta Z}^{CMX}(p_T) = \mathrm{Min}(a + bp_T, 0) \times (|\eta| - 0.675) + c + \frac{d}{p_T} + \frac{e}{p_T^2}. \quad (50)$$

The parameters corresponding to each muon type and matching variable are given in Appendix A.2. We obtain p_T independent quantities by scaling the matching variables with the p_T dependent widths: $\Delta x = \Delta X/\sigma_{\Delta X}(p_T)$, $\Delta\phi = \Delta\Phi/\sigma_{\Delta\Phi}(p_T)$ and $\Delta z = \Delta Z/\sigma_{\Delta Z}(p_T)$. Figure 21 shows the scaled variables for CMU muons in three p_T ranges: $p_T < 2$ GeV/c, $2 < p_T < 3$ GeV/c and $p_T > 3$ GeV/c . It is clear that the scaled variables are p_T independent. We obtain similar results for CMP, CMUP and CMX muons. For BMU muons there exist no available scaling functions to obtain the p_T independent track-stub matching quantities. The sample of BMU muons used in this analysis is not large enough to derive these scaling functions. We solve this problem by simply parameterizing the matching variables in different p_T ranges. Figure 22 shows, as an example, the matching variable $\Delta\Phi$ in each p_T range for BMU muons.

3.2.2 Calorimeter Templates

In addition to the track-stub matching variables, the muon likelihood also contains information on the electromagnetic and hadronic energy depositions of the muon candidates. We study the dependence of the electromagnetic energy E_{EM} and the hadronic energy E_{HAD} distributions on isolation and transverse momentum p_T of the muon. We define the isolation of a given track as the ratio between the transverse momentum of the track and the sum of transverse momenta of all tracks in an $\eta\phi$ cone of $\Delta R = 0.4$:

$$I = \frac{p_T}{\sum_i p_{T_i}} \quad (51)$$

where i runs over all tracks inside a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ around the track in question. Figure 23 shows the isolation distribution of CMU muons from $J/\psi \to \mu^+\mu^-$ and CMU protons from $\Lambda \to p\pi^-$ (see Sec. 3.4). We find that muons from J/ψ decays are mostly isolated. Other muon types have similar distributions.

The electromagnetic energy has some dependence on isolation. Figure 24 shows the E_{EM} energy distribution for isolated muons ($I > 0.5$) and non-isolated muons ($I < 0.5$). Different templates corresponding to these two isolation ranges are used in the final likelihood. As shown in Figure 24, the dependence of the hadronic energy on isolation is very small and we neglect it. A similar behavior is found for all other muon types.

The dependence of E_{EM} energy distributions on p_T is very small and we neglect it as well. However, E_{HAD} energy has some p_T dependence as shown in Figure 25.

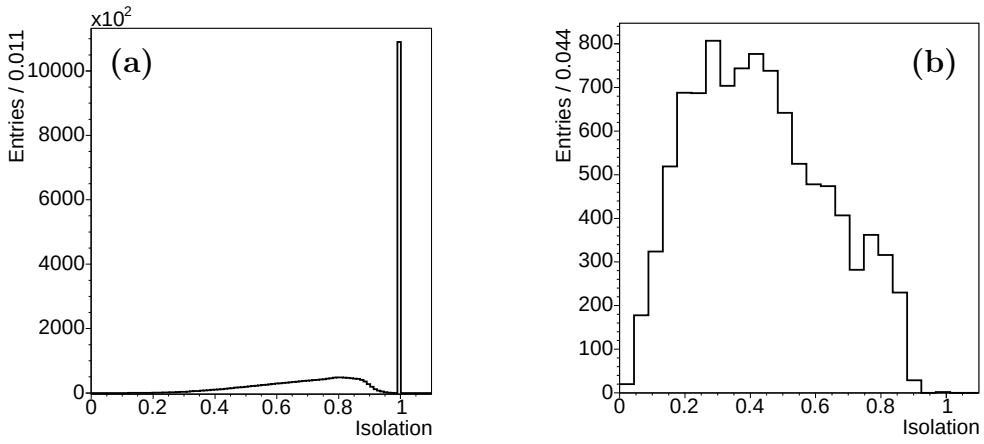

Figure 23: Isolation of (a) CMU muons from $J/\psi \to \mu^+\mu^-$ and (b) isolation of CMU protons from $\Lambda \to p^+\pi^-$. Entries at one correspond to totally isolated muons for which there is no other track in a 0.4 $\eta\phi$ cone around the muon track.

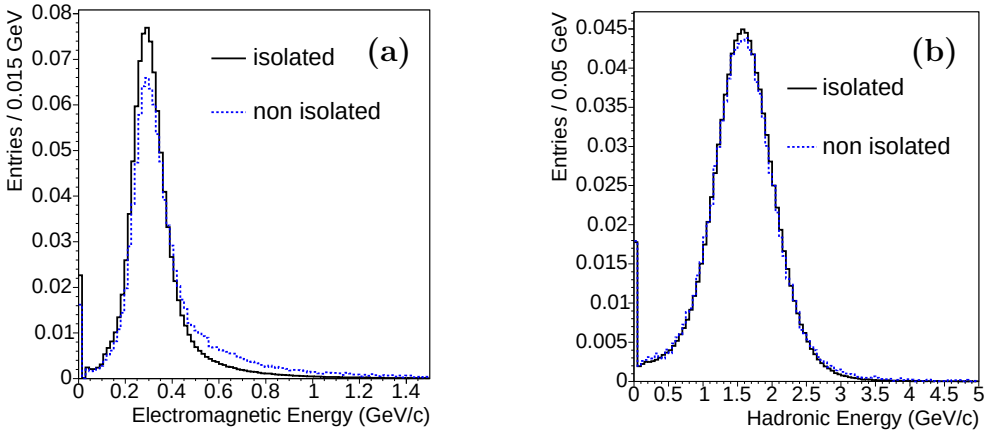

Figure 24: Dependence of the (a) electromagnetic energy E_{EM} and (b) hadronic energy E_{HAD} on isolation for CMU muons. Solid line corresponds to isolated muons ($I > 0.5$) and dashed line corresponds to non-isolated muons ($I < 0.5$). While there is some isolation dependence in E_{EM} distributions, the E_{HAD} distributions show negligible dependence.

Rather than attempting to find p_T independent scaling functions as we did in the case of the track-stub matching variables, we use in the final likelihood different E_{HAD} templates corresponding to three p_T ranges: $p_T < 2$ GeV/c, $2 < p_T < 3$ GeV/c and $p_T > 3$ GeV/c for CMU, $p_T < 3$ GeV/c, $3 < p_T < 4$ GeV/c and $p_T > 4$ GeV/c for CMUP, CMX and BMU and $p_T < 3$ GeV/c, $3 < p_T < 5$ GeV/c and $p_T > 5$ GeV/c for CMP muons. We chose different ranges for different muon types for two reasons. First, we account for the lowest transverse momentum p_T a muon must have in order to reach each muon detector. Secondly, we select the boundaries in each range so that the sample sizes of different ranges are comparable.

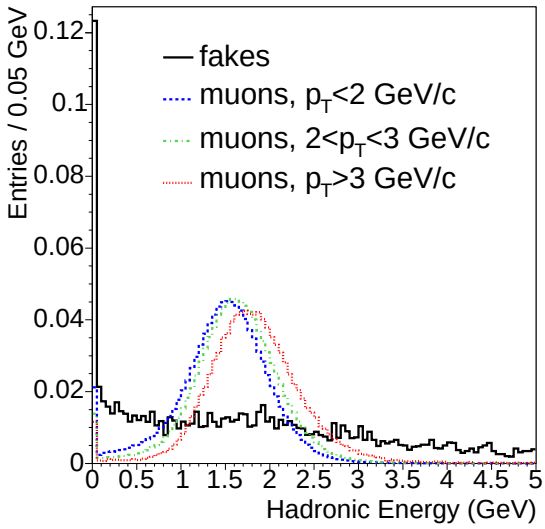

Figure 25: Hadronic energy E_{HAD} of CMU muons in different p_T ranges: $p_T < 2$ GeV/c (dashed), $2 < p_T < 3$ GeV/c (dot-dashed) and $p_T > 3$ GeV/c (dotted) compared to hadronic energy of CMU protons (solid line).

As discussed in more detail in Sec. 3.4, the hadronic energy distributions have the most power in separating muons from fakes. Secondly, the track-stub matching quantities are also very discriminating at high p_T, but in the low p_T regime, the matching distributions for both muons and fakes are wide and would add less separation power in the combined likelihood. The electromagnetic energy distributions contribute the least to the likelihood separation. Figure 26 shows the large overlap between real muon and fake muon E_{EM} distributions.

We fit both real and fake (see Section 3.4) muon distributions with empirical functional forms and use the fitting functions as PDF's in the global likelihood. Figure 26 shows as an example the parameterized electromagnetic energy distributions for isolated and non-isolated muons.

The procedure described to obtain the CMU templates is similarly repeated for all muon types. A total of about one hundred templates are parameterized for use as PDF's in the muon identification likelihood algorithm.

3.3 Sources of Fake Muons

Pions, kaons and protons can be misidentified as muons by punch-through (PT) or decay-in-flight (DIF). We define the fake muons are hadrons that reach the muon chambers by punching-through the calorimeter and steel shielding. Punch-through happens when a hadron interacts late in the calorimeter and a leading interaction product with small transverse momentum with respect to the original track and large momentum fraction exits the calorimeter, producing a muon stub. The punch-through events can be produced by pions, kaons or protons. We use pions from $K^0_S \to \pi^+\pi^-$, kaons from $D^0 \to K^+\pi^-$ and protons from $\Lambda \to p^+\pi^-$ to study how hadrons produce

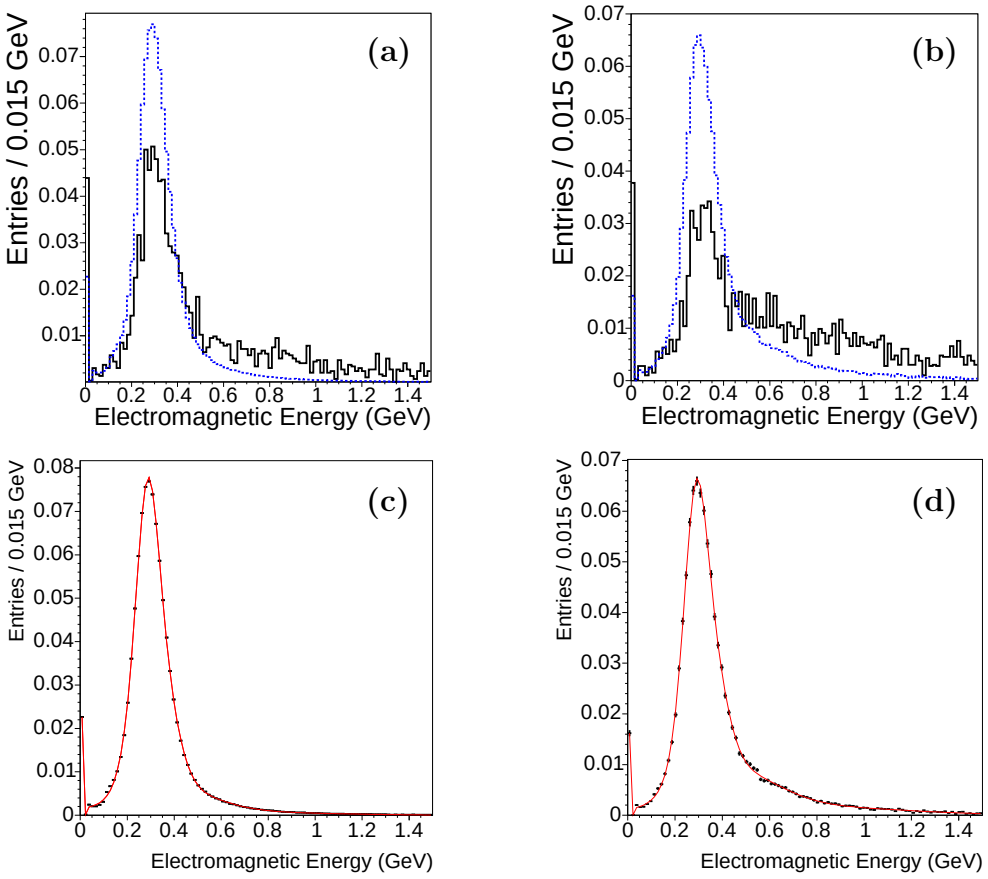

Figure 26: Electromagnetic energy for (a) isolated tracks and (b) non-isolated tracks with solid line corresponding to fake muons (protons) and the dashed line corresponding to real muons. Parameterization of (c) isolated and (d) non-isolated CMU muon E_{EM} distributions .

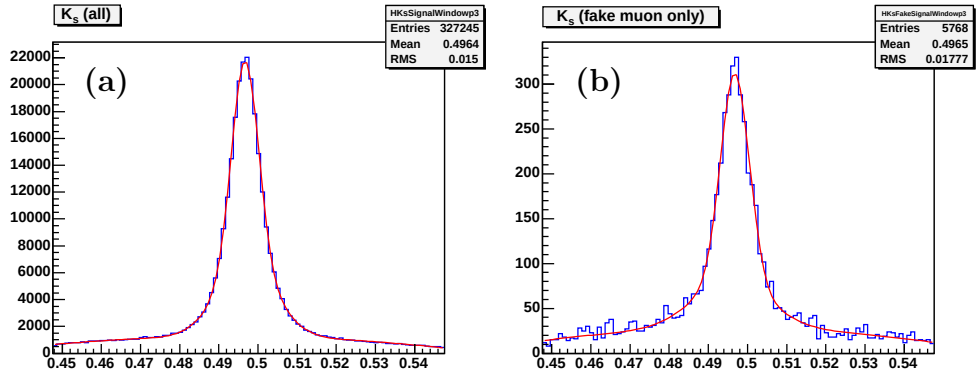

Figure 27: K^0_S mass peak for (a) all events in which the positive daughter pion has $3 < p_T < 4$ GeV/c and (b) K^0_S mass peak for events in which the positive daughter pion with $3 < p_T < 4$ GeV/c fakes a muon object.

p_T (GeV/c)	1.5-2.0	2.0-3.0	3.0-4.0	4.0-5.0	5.0-7.0	>7.0
π^- probability (%)	0.7	1.0	1.5	1.8	2.8	4.2
π^+ probability (%)	0.6	1.1	1.5	2.0	2.5	2.7

Table 1: Probabilities that a pion from K^0_S fakes a muon object as a function of the pion transverse momentum p_T. The statistical uncertainties on these probabilities are less than 0.1%.

p (GeV/c)	1.0	2.0	3.0	4.0	5.0	6.0	7.0
Pion probability (%)	5.2	2.6	1.8	1.3	1.1	0.9	0.8
Kaon probability (%)	32.9	18.1	12.4	9.5	7.7	6.4	5.5

Table 2: Probabilities that a pion or a kaon decays in flight before reaching the muon chambers for various momenta p as obtained from Equation (52).

muon objects.

We estimate the probability that a pion from a $K^0_S \to \pi^+\pi^-$ decay fakes a muon by measuring the fraction of K^0_S events in which one of the daughter pions is misidentified as a muon object. We evaluate the fake fractions as a function of pion transverse momentum p_T. In different p_T ranges, the fake fraction is given by the ratio between the number of signal events that fake a muon to the total number of signal events. As an example, Figure 27 shows the K^0_S peaks for all events in which the positive pion has $3 < p_T < 4$ GeV/c and for events where one pion is identified as a muon.

Table 1 shows the probability that pions fake muons as a function of p_T. For a typical opposite side muon of 3-4 GeV/c, we expect a fake fraction of $\sim$ 1.5%. Considering that in the Tevatron environment a single event may contain several tens of tracks, the chance that one of them would fake a muon object seems to be quite large.

The numbers in Table 1 include both PT and DIF. There appears to be no straightforward way to disentangle these contributions. However, one can calculate the probability that a pion or a kaon decays in flight before it reaches the muon chambers (i.e. it travels a distance of $\sim$ 3 m):

$$\mathcal{P}(p) = \int_0^3 \frac{m}{p\tau} e^{-mx/p\tau} dx, \tag{52}$$

where m, p and τ are the mass, momentum and lifetime of the pion or kaon, respectively. Table 2 shows the probabilities that a pion or a kaon decays in flight before reaching the muon chambers. Kaons have a higher chance to decay inside the tracking volume, but the branching fraction for a kaon to decay into a muon is less than that of a pion, $BR(K^+ \to \mu^+\nu) \approx 63.5\%$ and $BR(\pi^+ \to \mu^+\nu) \approx 100\%$.

We expect that many of the decay-in-flight muons are removed by the track reconstruction algorithms, especially the decays in which the direction of the muon is significantly different from the direction of its parent particle resulting in a "kink" in the trajectory. This effect is probably enhanced in kaon decays because kaons are heavier than pions and more available phase-space would produce a larger "kink" in

the trajectory. Tables 1 and 2 show that DIF are favored at low momenta and PT are favored at high momenta.

3.4 Fake Muon Templates

To identify fake muons, we select events in which the track of the hadron (pion, kaon or proton) from the reconstructed state (K^0_S, D^0 or Λ) is matched with a muon stub. If multiple tracks are attached to the same muon stub, the track that is closest to the stub is chosen. The K^0_S, D^0 and Λ signals are obtained from the two-track-trigger dataset in which the main selection requirements are two displaced tracks with impact parameter $|d_0| > 120$ μm and $p_T >$ 2 GeV/c. Candidates are reconstructed using a software package described in Reference [33]. Tracks with $P_T > 0.4$ GeV/c and $\eta < 2$ are selected to form the candidates. All tracks are required to have at least 20 axial and 20 stereo COT hits and are refitted accounting for energy loss corrections for pion, kaon and proton hypothesis. Layer 00 silicon hits are are not used for refitting the tracks. The candidate invariant mass is required to be between 0.42 and 0.58 GeV/c^2 for K^0_S, 1.65 and 2.55 GeV/c^2 for D^0 and 0.95 and 1.21 GeV/c^2 for Λ. For the candidates remaining after the mass cuts, a three dimensional vertex fit [34] of the two tracks is required to converge and the invariant mass range is tightened to 0.47 and 0.53 GeV/c^2 for K^0_S, 1.70 and 2.05 GeV/c^2 for D^0 and 1.0 and 1.16 GeV/c^2 for Λ. To obtain a reasonable signal to background ratio, we impose the following selection requirements requirements:

- $\Lambda \rightarrow p\pi^-$ selection:
  - $L_{xy}/\sigma_{L_{xy}} > 20$
  - $|d_0(\Lambda)| < 0.02$ cm
  - $|\Delta Z(p,\pi)| < 2$ cm
  - $\chi^2_{vx} < 10$
- $K^0_S \rightarrow \pi^+\pi^-$ selection:
  - $L_{xy}/\sigma_{L_{xy}} > 10$
  - $|d_0(K^0_S)| < 0.02$ cm
- $D^0 \rightarrow K^+\pi^-$ selection:
  - $L_{xy}/\sigma_{L_{xy}} > 5$
  - $\chi^2_{vx} < 20$
  - $p_T(\pi) > 2$ GeV/c
  - $p_T(D^0) > 5.5$ GeV/c
  - $0.12 < d_0(\pi/K) < 1.0$ mm

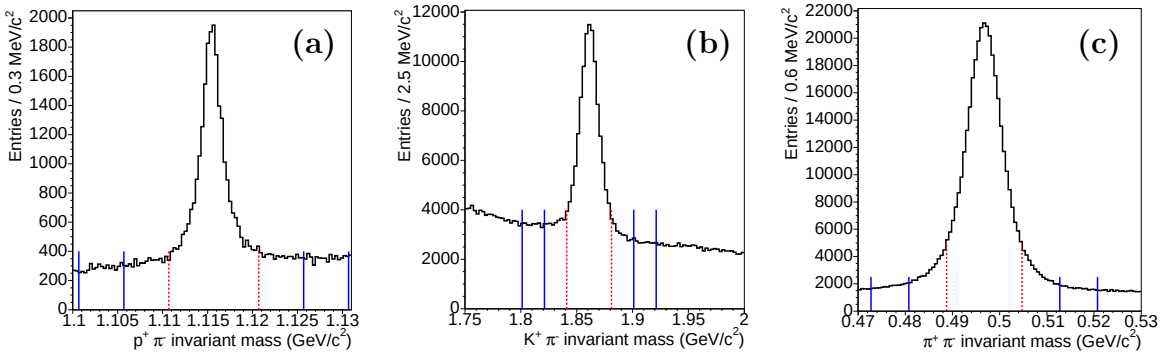

Figure 28: Mass peak of (a) $\Lambda \to p^+\pi^-$, (b) $K^0_S \to \pi^+\pi^-$ and (c) $D^0 \to K^+\pi^-$. Protons from Λ, pions from K^0_S and kaons from D^0 are matched with muon stubs.

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}, d_0 is the impact parameter of the corresponding particle, $|\Delta Z|$ is the distance along the z-axis between the two daughter tracks and χ^2_{vx} refers to the vertex fit. The purpose of the decay length significance $L_{xy}/\sigma_{L_{xy}}$ requirements is to select long lived particles. The impact parameter of the Λ and K^0_S candidates is required to be consistent with prompt production of these particles. Vertex χ^2 requirements ensure that the decay products come from a common vertex. We select samples of about 3×10^5 pions from $K^0_S \to \pi^+\pi^-$ decays, 7.5×10^4 kaons from $D^0 \to K^+\pi^-$ decays and 2×10^4 protons from $\Lambda \to p\pi^+$ decays where the pion, kaon or the proton are associated with muon objects. Signal K^0_S, D^0 and Λ mass peaks are shown in Figure 28.

We investigated the possibility of using the high statistics pion and kaon distributions to obtain the fake templates, but since kaons and pions are unstable and decay mainly into muons, our method of matching hadron tracks to muon objects, selects not only fake muons but also decay-in-flight muons. Since DIF muons are, in general, not usable for flavor tagging, one could imagine that including them in the fake sample would help to separate them from non-DIF muons. However, DIF and non-DIF muons have very similar detector signatures and it appears impossible to separate them on a statistical basis. Also, combining PT and DIF to obtain the fake templates reduces the separation between real muons and PT. We decide to use only protons as representative for all hadrons that produce muon stubs by punch-through. The validity of this assumption is demonstrated in Appendix A.3. Our attempts to further identify DIF using track χ^2 and the track impact parameter are described in Appendix A.4.

As in the case of real muons, we find that the electromagnetic energy of protons has some dependence on isolation, while the hadronic energy shows no significant dependence. This is exemplified for CMU protons in Figure 29. The ΔX, $\Delta\Phi$ and ΔZ distributions for CMU protons are included in Figure 20 and the E_{EM} and E_{HAD} distributions are shown in Figure 26. Proton track-stub variables have very small p_T dependence as shown in Figure 30. The small dependence of the hadronic energy of CMU protons is resolved by parameterizing it in three p_T ranges.

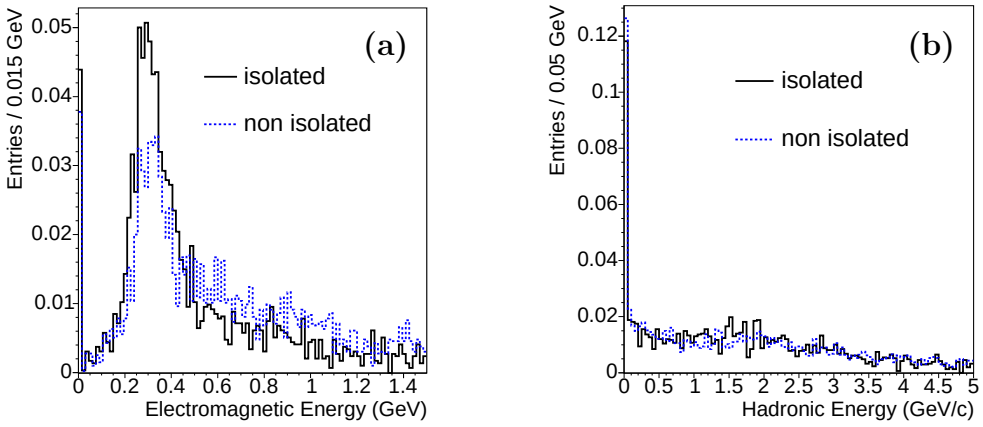

Figure 29: Isolation dependence of (a) electromagnetic energy and (b) hadronic energy distributions for CMU protons. Solid line corresponds to isolated protons ($I > 0.5$) and dashed line corresponds to non-isolated ones ($I < 0.5$).

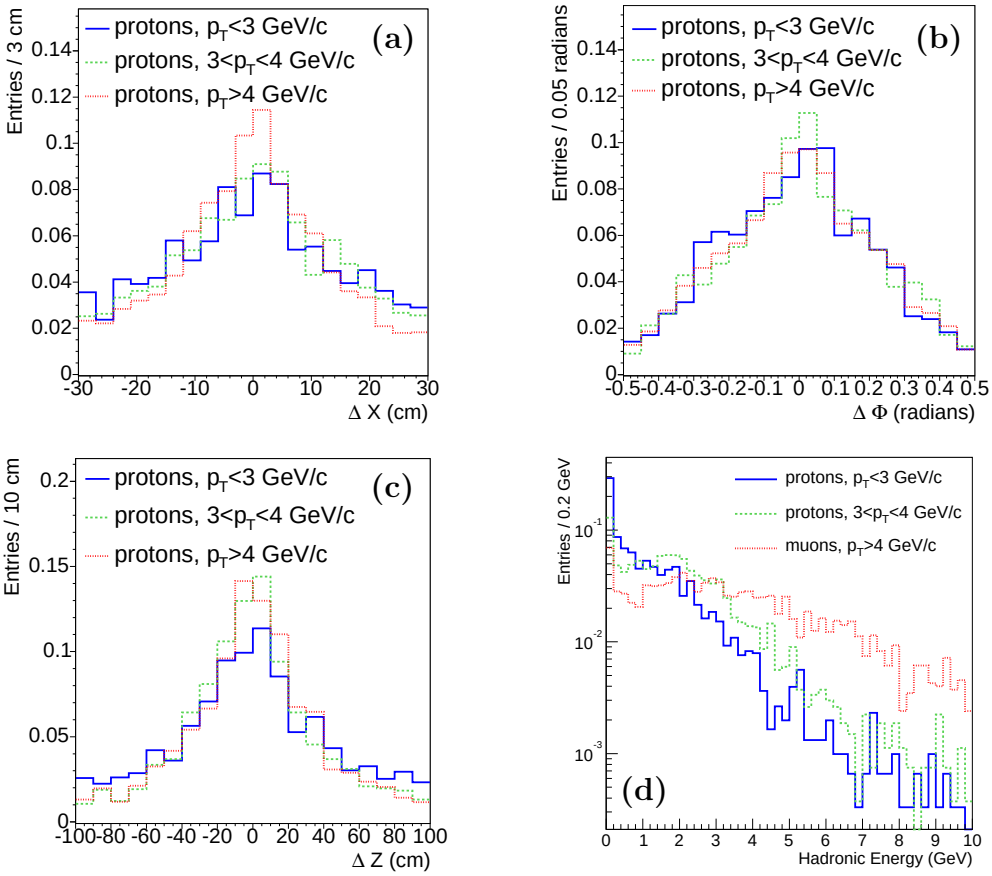

Figure 30: p_T dependence of CMU proton distributions: (a) ΔX, (b) $\Delta\Phi$, (c) ΔZ and (d) hadronic energy. $p_T < 3$ GeV/c (solid), $3 < p_T < 4$ GeV/c (dashed) and $p_T > 4$ GeV/c (dotted).

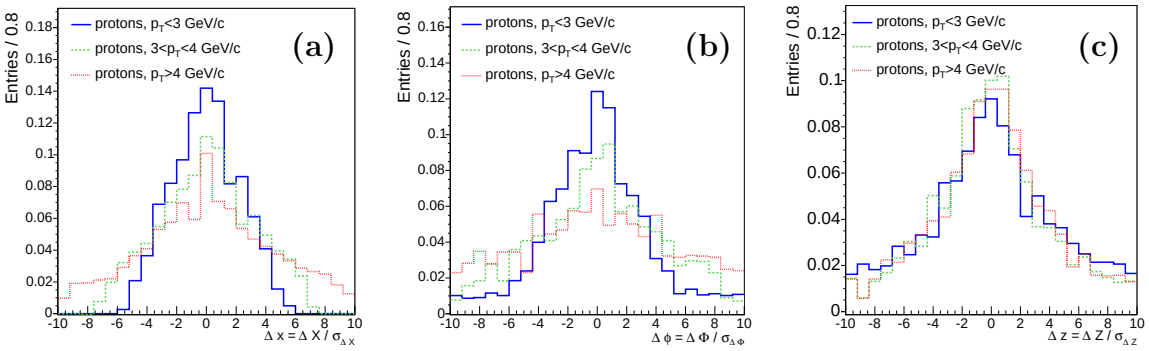

Figure 31: p_T dependence of CMU proton matching distributions after scaling: (a) $\Delta x = \Delta X/\sigma_{\Delta X}(p_T)$, (b) $\Delta\phi = \Delta\Phi/\sigma_{\Delta\Phi}(p_T)$ and (c) $\Delta z = \Delta Z/\sigma_{\Delta Z}(p_T)$; $p_T < 3$ GeV/c (solid), $3 < p_T < 4$ GeV/c (dashed) and $p_T > 4$ GeV/c (dotted).

Before applying the scaling functions that make the matching quantities p_T independent for real muons, the track-stub matching quantities of protons that fake muons have negligible p_T dependence. However, after we apply the p_T scaling procedure described in Section 3.2.1 to protons, we obtain p_T dependent Δx and $\Delta\phi$ distributions. Figure 31 illustrates this effect. We use different parameterizations in three p_T ranges: $p_T < 3$ GeV/c, $3 < p_T < 4$ GeV/c and $p_T > 4$ GeV/c.

3.5 Likelihood of Muons from $J/\psi \to \mu\mu$ and Protons from $\Lambda \to p\pi$

An initial step to test the likelihood algorithm is to study its performance on the samples that were used to generate the template distributions for the discriminating variables: muons from $J/\psi \to \mu^+\mu^-$ and protons from $\Lambda \to p^+\pi^-$. The separation power of the likelihood method can qualitatively be characterized by how close protons accumulate near zero and how close muons cluster toward one. Figure 32 shows the likelihood distributions of real muons for each of the five muon types. Each likelihood distribution is obtained as the difference between likelihood distributions in the signal region and the side-band region of the J/ψ mass peak. All muon types, except CMP, cluster sharply at one. For CMP muons, ΔZ matching between the track and stub is not measured and the likelihood only contains four variables at most.

Figure 32 also shows the likelihood distributions for protons. The same side-band subtraction procedure as for real muons is used to obtain the likelihood distributions for fake muons. Protons cluster very close to zero. The likelihood distribution of CMP protons has a longer tail towards one, but still have a sharp spike at zero. Further consistency checks of the likelihood algorithm are presented in Appendix A.5.

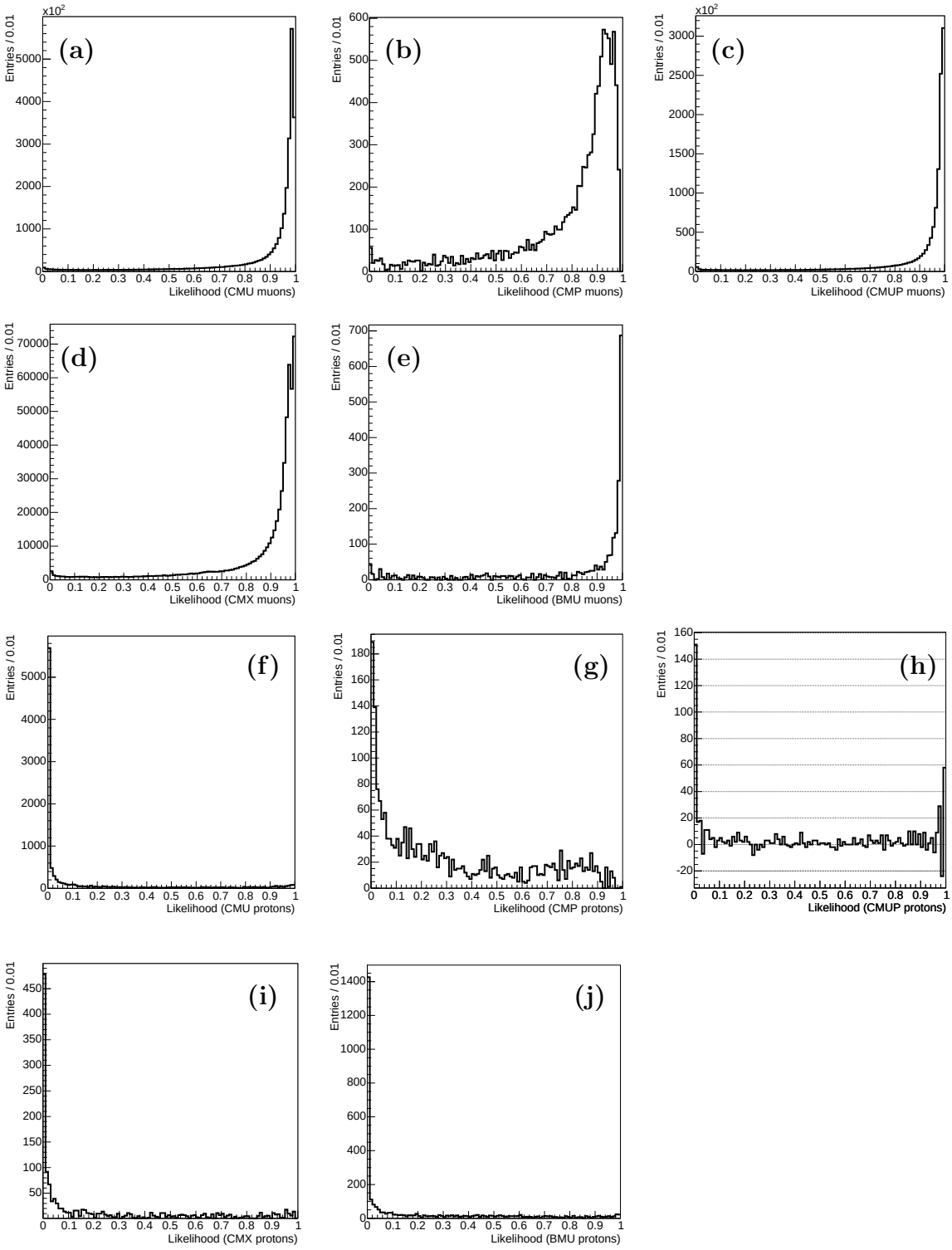

Figure 32: Likelihood distributions for real (a) CMU muons, (b) CMP muons, (c) CMUP muons, (d) CMX muons and (e) BMU muons. Likelihood distributions for (f) CMU protons, (g) CMP protons, (h) CMUP protons, (i) CMX protons and (j) BMU protons.

3.6 Tagging Algorithm Performance

To evaluate the performance of the likelihood tagging algorithm, we follow the procedure described in References [31, 42] . We use the muon-SVT and electron-SVT datasets. As described in Section 2.2.1, these events are triggered by a 4 GeV/c lepton and a displaced track (SVT) with impact parameter 0.12 μm $< |d_0| <$ 1 mm. In addition, the SVT track must have a minimum transverse momentum requirement $p_T > 2$ GeV/c. The role of the displaced track is to select the long lived B mesons. The trigger lepton is the signature of a semileptonic B decay and provides an indication of the flavor of the trigger side B meson at decay. Apart from B events, the trigger lepton may come from prompt charm mesons and other background sources where the hadron is misidentified as either the trigger muon or electron. The prompt charm contribution in this sample is removed by requiring that the invariant mass of the trigger lepton and the SVT track is above 2 GeV/c^2. From Monte Carlo simulation studies we find that the fraction of prompt charm particles that produce a trigger lepton and an SVT track with invariant mass above 2 GeV/c^2 is negligible. We also require that the lepton SVT invariant mass is less than 4 GeV/c^2 to suppress other sources of background like the events in which the trigger lepton is fake.

If we restrict the analysis to events with lepton-SVT mass between 2 and 4 GeV/c^2, this sample will contain b events and background, but no charm events. The method developed in Reference [31] is used to determine the properties of the b subsample by removing the background contribution. The basic assumption is that all sources of background are symmetric in the signed impact parameter of the SVT track which is defined as:

$$\delta(SVT) = |d_0|\, sign(\vec{d_0} \cdot \vec{p}_{lep-SVT}), \tag{53}$$

where $\vec{d_0}$ is the displaced track impact parameter with direction from the primary vertex to the point of closest approach of the SVT track and $\vec{p}_{lep-SVT}$ is the combined momentum of the trigger lepton plus SVT track. Figure 33 shows the topological configurations corresponding to both positive and negative signed impact parameter of the SVT track and Figure 34 shows the invariant mass of the lepton-SVT pair for positive and negative SVT track impact parameter. To obtain any distribution that is characteristic for a pure b sample, we subtract the distribution with negative $\delta(SVT)$ from the corresponding distribution with positive $\delta(SVT)$.

In our case, the two quantities of interest are the efficiency and dilution of the tagging algorithm. In the particular case of the muon tagging algorithm, the efficiency is given by the fraction of events with opposite side muons found:

$$\varepsilon = \frac{N_{OS} + N_{SS}}{N}, \tag{54}$$

where $N_{OS/SS}$ is the number of opposite/same-sign tagged events and N is the total number of events in the sample. Including the background $\delta(SVT)$ subtraction, the efficiency is calculated as:

$$\varepsilon = \frac{(N^+_{OS} - N^-_{OS}) + (N^+_{SS} - N^-_{SS})}{N^+ - N^-}, \tag{55}$$

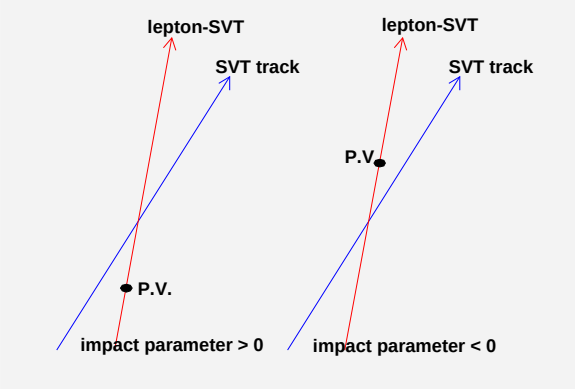

Figure 33: Two configurations corresponding to positive (left) and negative (right) signed impact parameter of the SVT track. If the intersection of the SVT track with the combined lepton-SVT direction is on the positive lepton-SVT axis with respect to the primary vertex, the signed impact parameter of the SVT track is positive. If the intersection is on the negative lepton-SVT axis, the SVT impact parameter is negative.

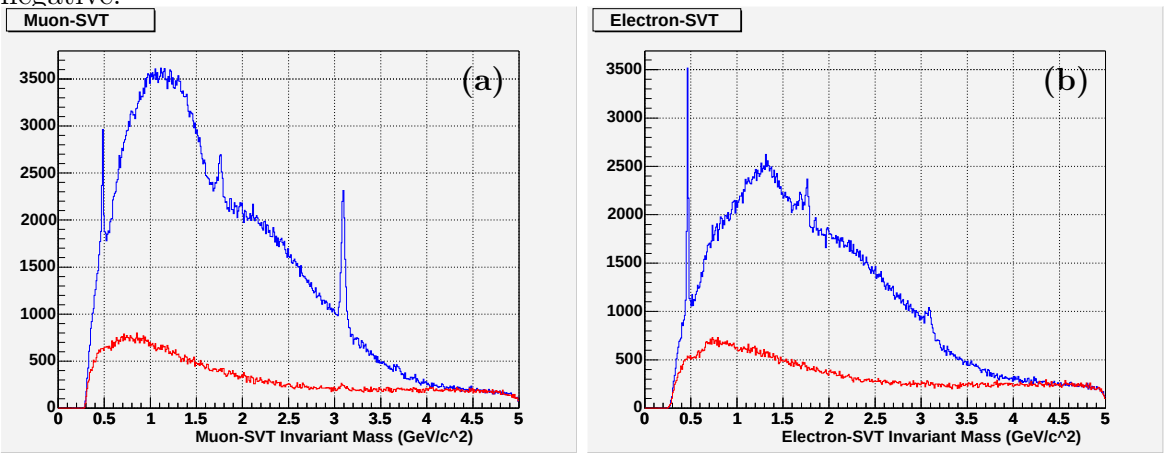

Figure 34: The invariant mass of the SVT track combined with the trigger (a) muon or (b) electron. The blue (red) histogram corresponds to positive (negative) impact parameter of the SVT track. The peaks originate from $K^0_S \to \pi^+\pi^-$ and $D^0 \to \pi^+K^-$ decays where one of the decay mesons is misidentified as a lepton or $J/\psi \to \mu^+\mu^-$ and $J/\psi \to e^+e^-$ decays in which one of the muons or electrons is identified as the trigger lepton and the other one is the SVT track. From Monte Carlo simulation studies, the contribution of charm mesons above 2 GeV/c^2 is found to be negligible.

where $N^{\pm}_{OS/SS}$ is the number of events in which the trigger and the opposite side muons have opposite/same-sign charges and $N^{\pm}$ is the total number of events, tagged and untagged. The $+$ and $-$ superscripts indicate $\delta(SVT) > 0$ and $\delta(SVT) < 0$, respectively.

If the opposite side muon originates from a semileptonic decay of the "opposite side" B meson and the trigger lepton originates from a semileptonic decay of the "trigger side" B meson, the charges of the two leptons can be used to decide whether

the trigger side B meson mixed or not. The accuracy of this determination is altered whenever one of the leptons is fake, when it does not come from a direct semileptonic B decay or when the opposite side B meson is neutral and oscillates into a state different than the production state. If p_W^{trg} and p_W^{opp} are the mistag probabilities on the trigger side and on the opposite side, respectively, we can introduce the trigger side dilution $\mathcal{D}_{trg} = 1 - 2p_W^{trg}$ and the opposite dilution $\mathcal{D}_{opp} = 1 - 2p_W^{trg}$. Since the latter is the quantity of interest, we will call the opposite side dilution simply $\mathcal{D}$. We introduce the raw dilution calculated as:

$$\mathcal{D}_{raw} = \frac{N_{OS} - N_{SS}}{N_{OS} + N_{SS}} \tag{56}$$

After applying the background subtraction procedure, the above equation becomes:

$$\mathcal{D}_{raw} = \frac{(N_{OS}^{+} - N_{OS}^{-}) - (N_{SS}^{+} + N_{SS}^{-})}{(N_{OS}^{+} - N_{OS}^{-}) + (N_{SS}^{+} - N_{SS}^{-})} \tag{57}$$

The trigger and opposite side dilutions are related to the raw dilution, as discussed in Appendix A.6, by the simple relation:

$$\mathcal{D} = \frac{\mathcal{D}_{raw}}{\mathcal{D}_{trg}} \tag{58}$$

To determine the intrinsic dilution of the opposite side muon tagger one must account for the fact that the trigger lepton may be unrelated to the flavor of the trigger side B meson due to sequential decays ($b \to c \to \ell$), mixing of the trigger side B hadron or other processes. The trigger side dilutions in both muon-SVT and electron-SVT data samples are calculated from Monte Carlo simulations of semileptonic B decays followed by detector and trigger simulations. The correlations between the charge of the trigger side b quark and the charge of the trigger lepton have been studied resulting in the following values for the trigger side dilutions:

$$\mathcal{D}_{trg} = \begin{cases} 0.6270 \pm 0.0032(stat), & \text{muon-SVT} \\ \\ 0.6625 \pm 0.0027(stat), & \text{electron-SVT} \end{cases} \tag{59}$$

These results are in good agreement with previous studies [42] that used a simplified detector simulation in the Monte Carlo studies. For consistency reasons we use the results from Reference [42]:

$$\mathcal{D}_{trg} = \begin{cases} 0.6412 \pm 0.0015(stat)^{+0.0141}_{-0.0226}(syst), & \text{muon-SVT} \\ \\ 0.6412 \pm 0.0015(stat)^{+0.0215}_{-0.0367}(syst), & \text{electron-SVT} \end{cases} \tag{60}$$

We treat the trigger side dilution as a correction to the raw dilution due to the trigger leptons being uncorrelated to the flavor of the trigger side B mesons. The corrected raw dilution is then the intrinsic dilution of the opposite side tagging algorithm.

The errors on ε and $\mathcal{D}$ are calculated by simple error propagation of the uncertainties on $N^{+/-}_{OS/SS}$ and $N^{+/-}$ and are obtained as:

$$\sigma^2_\varepsilon = \frac{(N^+_{OS} + N^-_{OS} + N^+_{SS} + N^-_{SS})(N^+ - N^-)^2 + (N^+_{OS} - N^-_{OS} + N^+_{SS} - N^-_{SS})^2(N^+ + N^-)}{(N^+ - N^-)^4} \tag{61}$$

and:

$$\sigma^2_\mathcal{D} = \frac{4[(N^+_{OS} - N^-_{OS})^2(N^+_{SS} + N^-_{SS}) + (N^+_{SS} - N^-_{SS})^2(N^+_{OS} + N^-_{OS})]}{(N^+_{OS} - N^-_{OS} + N^+_{SS} - N^-_{SS})^4}. \tag{62}$$

3.6.1 Trigger Side Selection

We evaluate the performance of the likelihood muon tagger using the muon-SVT and electron-SVT data sets. We reconfirm offline the Level 3 trigger decision by using a software module [38] which emulates the trigger response. The trigger lepton and the SVT tracks were refitted within the appropriate particle hypothesis. We use runs within the range from 150799 to 168889. A good run selection was applied by requiring that the detector components used in this analysis (silicon detector, drift chamber, calorimeters and muon systems) were performing well in each selected run. The total offline integrated luminosity is (168.8 ± 9.9) pb^{-1}.

Events were selected imposing the following requirements to the trigger lepton and SVT track:

- the lepton and SVT tracks have at least 20 axial and 20 stereo hits in the drift chamber to ensure high quality of the tracks
- the distance between the trigger lepton and the SVT track projected along the beam direction must satisfy $|Z(lepton) - Z(SVT)| < 2.5$ cm to avoid tracks originating from different primary interactions
- the lepton has at least 4 $r\phi$ SVXII silicon hits
- some background events originate from trigger electrons from photon conversion into electron pairs as a result of interaction with detector material ($\gamma \to e^+e^-$). In such events, the two electron tracks have a very small angular separation, reflecting the fact that the photon is massless. We partially remove conversions by three requirements:
  - transverse separation between trigger lepton track and conversion partner is less than 1.0 cm
  - the angle between the trigger lepton track and conversion partner is very small: $|\Delta cot\,\theta| < 0.05$
  - a three dimensional vertex fit ensures that the trigger lepton track and the conversion partner originate from a common vertex
- in case of multiple lepton-SVT pairs in an event, we select the pair with highest combined transverse momentum

- the invariant mass between the trigger lepton and the SVT track must be consistent with B decays: $2 < m < 4$ GeV/c^2

The selected events are divided in two subsamples according to the sign of the SVT track impact parameter. In the muon-SVT data sample, we find 376,000 (77,000) events with positive (negative) impact parameter of the SVT track and in the electron-SVT data, we find 340,000 (94,000) events with positive (negative) impact parameter of the SVT track.

The trigger muons have both CMU and CMP stubs and the track-stub matching ΔX is required to be less than 15 cm and 20 cm, respectively. With these tight selection criteria we expect that the trigger muons are mostly pure with small fake muon contamination from punch-through effects. As expected, we find in Appendix A.4 that the fake muon fraction for the trigger muons is $\sim 3\%$.

3.6.2 Opposite Side Selection

Once an event passes the trigger side requirements, we search for another muon called the opposite side or soft muon which must be different from the trigger lepton. To avoid selecting muons from sequential decays in which both the B meson and its charm daughter decay semileptonically, we require that the trigger lepton and the opposite side muons are in different jets. To find the jets in an event we use a standard CDF software package [39] which implements a track based cone clustering algorithm [41]. This algorithm starts by selecting all tracks with $p_T > 1$ GeV/c as the jet seeds. The two highest p_T tracks with $\Delta R < 0.7$ are merged by adding their momenta to form a new jet seed. This process is repeated until no track pair satisfies the merging criteria. Finally, all tracks within a $\Delta R < 0.8$ with respect to the final jet seeds and with $0.4 < p_T < 1$ GeV are added to form the final jets.

The opposite side muon is required to pass the following requirements:

- $p_T > 1.5$ GeV/c for any muon type
- the requirement that the muon impact parameter $|d_0| < 0.3$ cm rejects hadrons that decay into muons
- the opposite side muon candidate and the trigger lepton are part of different jets
- the opposite side muon candidate and the SVT track are different tracks
- the invariant mass of the trigger lepton, SVT track and the opposite side muon is greater than 5 GeV/c. This requirement ensures that the tag muon cannot originate from the same B decay as the trigger lepton and SVT track
- the distance between the trigger lepton and the tag muon projected along the beam line $|\Delta Z(\text{trigger lepton, soft muon})| < 5.0$ cm

If there is more than one muon candidate that passes the above selection criteria, the isolated muon is chosen if found. An isolated muon is defined as a muon which

Muon type	CMU	CMP	CMUP	CMX	BMU
Fake fraction (%)	46	42	12	20	72

Table 3: Fake fractions for each muon type. The errors associated with these fractions are less than $\sim 2\%$.

represents its own jet with no other tracks within $\Delta R = 0.7$ distance from the muon. About 8% of the soft muons are found to be isolated. If none of the muon candidates is isolated, the muon with the highest p_T^{rel} is selected.

3.6.3 Likelihood on Opposite Side Muons

Ultimately, we want to use the likelihood method to evaluate the quality of opposite side muons for the purpose of flavor tagging. Figure 35 shows the likelihood distributions for the opposite side muons separately for each of the five muon types. Each plot shows sharp spikes at zero which correspond to low quality muons and peaks close to one which correspond to pure muons. Directly from the plots, one can read off that the BMU detectors allow the largest fraction of fakes, while CMUP muons are the purest muon type.

We repeat the procedure described in Appendix A.5 to roughly estimate the fake fractions in each of the five muon systems. The resulting fractions are shown in Table 3. As a consistency check, we study the subsample of opposite side muons that have all five (four for CMP) variables inside the chosen PDF ranges. Fake fractions in this case are shown in Table 4. As expected, a clear decrease in the fraction of fakes is observed when all likelihood variables are available.

muon type	CMU	CMP	CMUP	CMX	BMU
fake fraction (%)	31	30	9	17	46

Table 4: Fake fractions for each muon type after requiring all variables to be available in the likelihood calculation. The errors associated with these fractions are less than $\sim 2\%$.

3.6.4 Dependence of Dilution on Likelihood

As seen in previous sections, high quality muons have likelihood values that are close to one and low quality muons have likelihood values close to zero. There are also muon objects that have intermediate likelihood values. This sub-sample is more evenly populated with both fakes and real muons. The dilution of fake muons should be very small, so we expect that the harder we cut on likelihood the higher the dilution would be. Our assumption is based on the basic fact that a more pure sample of muons should reflect the charge of the initial b quarks more accurately.

As a demonstration, Figure 36 shows the efficiency, raw dilution and $\varepsilon\mathcal{D}^2_{raw}$ versus likelihood cut for all muon types combined. After the first cut at $\mathcal{L} > 2\%$, the efficiency drops from $\sim 6\%$ to $\sim 4.5\%$ as most of the low quality muons are removed.

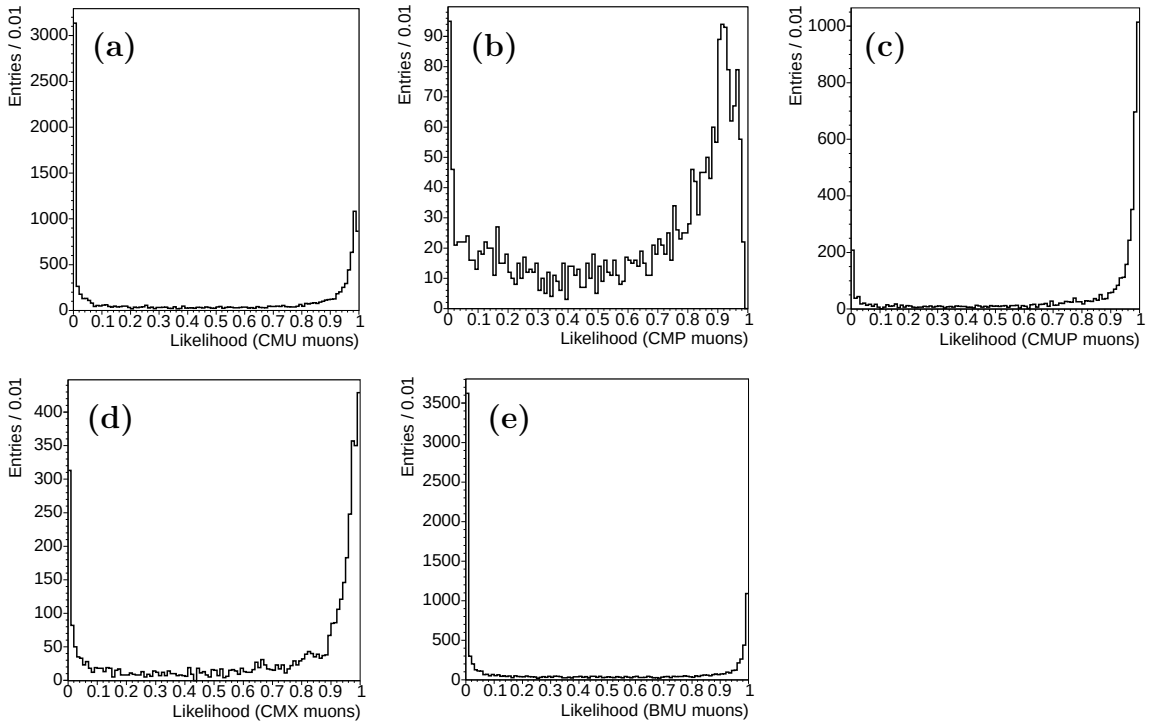

Figure 35: Likelihood distributions for opposite side (a) CMU, (b) CMP, (c) CMUP (d) CMX and (e) BMU muons in the combined electron-SVT and muon-SVT samples.

Muon Type (GeV/c)	CMU	CMP	CMUP	CMX	BMU
Likelihood Cut	0.5	0.5	0	0.5	0.7

Table 5: Likelihood values that maximize the tagging power $\varepsilon\mathcal{D}^2$ for each muon type.

The efficiency drop is accompanied by a sudden rise in dilution from $\sim 16\%$ to $\sim 20\%$. As we cut harder in likelihood, the efficiency drops further, while the dilution increases steadily up to $\sim 30\%$. $\varepsilon\mathcal{D}^2_{raw}$ exhibits a maximum around a 50% likelihood cut. We split the opposite side muons into five subsamples corresponding to the five muon types and find that the general features from Figure 36 are reproduced for each muon type.

Knowing that the dilution has a strong dependence on the likelihood cut, we study the dilution dependence on the likelihood variable itself. To demonstrate this dependence, we calculate the dilution in seven different likelihood bins: $0.0 \leq \mathcal{L} \leq 0.1$, $0.1 < \mathcal{L} \leq 0.4$, $0.4 < \mathcal{L} \leq 0.8$, $0.8 < \mathcal{L} \leq 0.9$, $0.9 < \mathcal{L} \leq 0.95$, $0.95 < \mathcal{L} \leq 0.99$ and $0.99 < \mathcal{L} \leq 1.0$. The first bin contains the lowest quality muon objects, while the next bins are samples of increasing muon purity. The dependence of the raw dilution on likelihood is shown in Figure 37 for all muon systems combined. The absolute dilution of the opposite side muon tagging algorithm is easily obtained after multiplication with the trigger side corrections given by Equation (60).

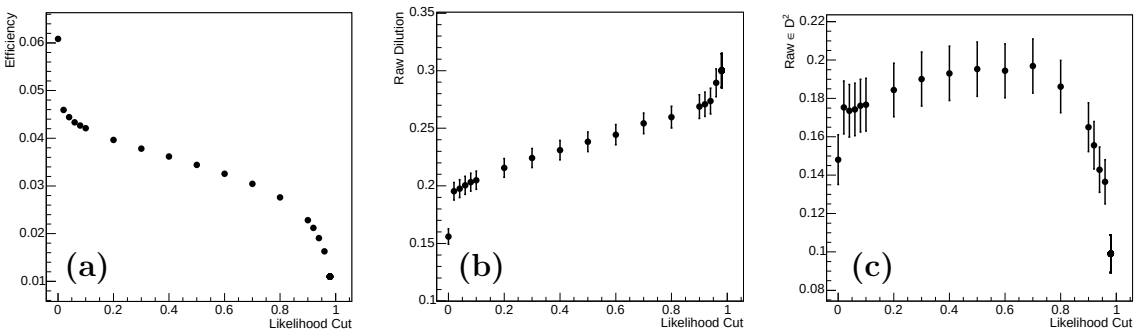

Figure 36: Dependence of (a) efficiency ε, (b) raw dilution $\mathcal{D}_{raw}$ and (c) $\varepsilon\mathcal{D}^2_{raw}$ as functions of likelihood cut for all muon types combined in the lepton-SVT sample.

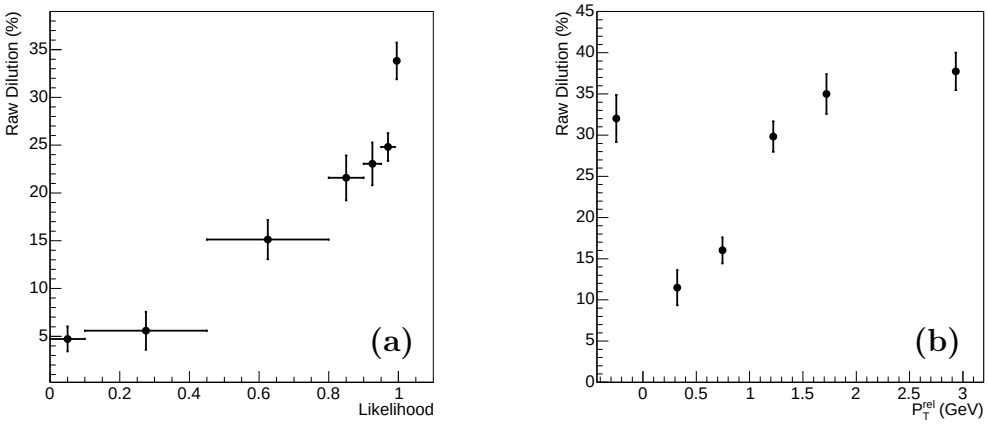

Figure 37: Raw dilution as function of (a) likelihood and (b) $p_T^{\rm rel}$ as obtained from the lepton-SVT sample. Dilution of isolated muons is shown as an entry at negative $p_T^{\rm rel}$.

3.6.5 Dependence of Dilution on p_T^{rel}

As discussed in the beginning of Section 3, the opposite side lepton dilution increases with increasing p_T^{rel}. Muons coming from b quark decays have larger $p_T^{\rm rel}$ than muons from c quark decays due to larger phase space available in a decay of B mesons. Figure 37 shows the dependence of dilution as a function of $p_T^{\rm rel}$ for electron-SVT and muon-SVT datasets combined. This dependence was established with the likelihood cuts given in Table 5 which were found to maximize $\varepsilon\mathcal{D}^2$. Isolated muons have a raw dilution of $\sim 30\%$. The muons in the lowest $p_T^{\rm rel}$ bin have a raw dilution of $\sim 10\%$ while the muons in the highest $p_T^{\rm rel}$ bin have a dilution of $\sim 40\%$.

3.6.6 Evaluation of $\varepsilon\mathcal{D}^2$

A better understanding of dilution can be achieved by analyzing each muon system separately and also by studying the dilution dependence on both likelihood and $p_T^{\rm rel}$ simultaneously. These dependencies are described in detail in Section 3.7. As es-

tablished in Sections 3.6.4 and 3.6.5, the dilution depends strongly on both muon likelihood and $p_T^{\rm rel}$. The likelihood function gives the probability that a muon object is a real muon and $p_T^{\rm rel}$ is related to the chance that the muon comes from a B meson. In a mixing analysis, one can use both these pieces of information to predict more accurately the dilution on a event-by-event basis. This procedure would improve the tagging power given by $\varepsilon\mathcal{D}^2$. We can do an a priori evaluation of the tagger performance by dividing the opposite side muons in six $p_T^{\rm rel}$ bins (1 for isolated muon + 5 $p_T^{\rm rel}$ bins) and three likelihood bins. In the limit of very large number of events, we would split the data in as many bins as possible, but having a limited number of events, we choose to split the data in fewer bins. We then evaluate $\varepsilon\mathcal{D}^2$ separately in each subsample. All opposite side muons that pass the requirements in Section 3.6.2 are selected. There is no cut on likelihood. The likelihood is allowed to be formed with any number of available variables. The total $\varepsilon\mathcal{D}^2$ is obtained by summing over the eighteen subsamples. In Appendix A.7 we report the $\varepsilon\mathcal{D}^2$ values for each muon type for both muon-SVT and electron-SVT data samples. Table 6 gives $\varepsilon\mathcal{D}^2_{raw}$ for each muon type for both the muon-SVT and electron-SVT data samples. Using the trigger side correction from Equation (60), the total corrected $\varepsilon\mathcal{D}^2$ is obtained:

$$\varepsilon\mathcal{D}^2 = (0.688 \pm 0.053(stat)^{+0.051}_{-0.029}(syst))\% \text{ in muon-SVT sample,}$$

$$\varepsilon\mathcal{D}^2 = (0.708 \pm 0.065(stat)^{+0.088}_{-0.046}(syst))\% \text{ in electron-SVT sample.}$$

The quoted systematic errors are based only on the uncertainties from the trigger side correction. These errors are obtained by simple error propagation of the systematic errors from Equation (60). The values of $\varepsilon\mathcal{D}^2$ obtained in electron-SVT and muon-SVT samples are in very good agreement with each other.

The above performance evaluation of the muon tagging algorithm assumed a known dependence of the tagging dilution as a function of the muon likelihood. The likelihood algorithm can also be used by cutting on the muon likelihood instead of parameterizing the dilution as a function of the likelihood. Even though this approach is not optimal for a mixing analysis, it has the advantage of selecting high purity muons. The harder the cut, the more pure selected muons are. As an example, we evaluate $\varepsilon\mathcal{D}^2$ using the same cuts as the ones given in Table 5. Table 7 shows raw dilution and efficiency in muon-SVT and electron-SVT data sets. Finally, with this method, we obtain the total $\varepsilon\mathcal{D}^2$ as:

$$\varepsilon\mathcal{D}^2 = (0.619 \pm 0.051)\% \text{ in muon+SVT sample,}$$

$$\varepsilon\mathcal{D}^2 = (0.607 \pm 0.057)\% \text{ in electron+SVT sample.}$$

3.7 Likelihood Tagger in a Mixing Analysis

We discuss a method of implementing the likelihood information in the measurement of a CP asymmetry or a B mixing analysis. The knowledge of the dilution

$\varepsilon\mathcal{D}^2_{raw}(\%)$	CMU	CMUP	CMX	CMP	BMU
muon-SVT	0.0571 ± 0.0099	0.0807 ± 0.0114	0.0483 ± 0.0090	0.0340 ± 0.0078	0.0628 ± 0.0104
electron-SVT	0.0584 ± 0.0130	0.0817 ± 0.0132	0.0650 ± 0.0120	0.0308 ± 0.0094	0.0555 ± 0.0116

Table 6: $\varepsilon\mathcal{D}^2_{raw}$ for each muon type in muon-SVT and electron-SVT data sets.

	CMU	CMP	CMUP	CMX	BMU
muon-SVT					
$\mathcal{D}_{raw}(\%)$	16.7 ± 2.0	29.4 ± 3.9	27.0 ± 2.3	25.1 ± 2.7	28.0 ± 2.6
$\varepsilon(\%)$	1.145 ± 0.024	0.303 ± 0.012	0.833 ± 0.020	0.556 ± 0.016	0.659 ± 0.018
electron-SVT					
$\mathcal{D}_{raw}(\%)$	17.6 ± 2.4	25.8 ± 4.6	29.3 ± 2.6	27.7 ± 3.1	27.1 ± 3.0
$\varepsilon(\%)$	1.131 ± 0.027	0.283 ± 0.013	0.809 ± 0.022	0.558 ± 0.018	0.635 ± 0.019

Table 7: Efficiency and raw dilution in electron-SVT and muon-SVT data sets using likelihood cuts.

dependence on likelihood and $p_T^{\rm rel}$ as well as on muon type can be used to predict the dilution on an event-by-event basis.

The study described in this section uses about 365 pb^{-1} of data [44]. We evaluate $\varepsilon\mathcal{D}^2$ with this enhanced data sample containing about 2.3 times more events than the initial sample. The updated tagging power is consistent with the initial study:

$$\varepsilon\mathcal{D}^2 = (0.677 \pm 0.035(stat))\% \text{ in muon-SVT sample}$$

$$\varepsilon\mathcal{D}^2 = (0.679 \pm 0.039(stat))\% \text{ in electron-SVT sample}$$

Figure 38 shows the raw dilution as a function of $p_T^{\rm rel}$ in the three likelihood ranges, $\mathcal{L}1 = [0, 0.80)$, $\mathcal{L}2 = [0.8, 0.95)$ and $\mathcal{L}3 = [0.95, 1.0]$ for all muon types combined. For a more accurate prediction, we study this dependence for each muon type individually. Figures 39 up to 43 show the dilution as a function of $p_T^{\rm rel}$ in three likelihood bins for each muon type. To fit the data, the following parameterization was used:

$$\mathcal{D}(p_T^{\rm rel}) = P_0 \cdot (1 - e^{-p_T^{\rm rel} + P_1}) \tag{63}$$

where P_1 represents the plateau at which the raw dilution reaches its maxim value as a function of $p_T^{\rm rel}$. Table 8 gives the values of P_1 as obtained from the fits and the dilution for isolated muons in different likelihood bins and for each muon type. Clearly, for each muon type, the dilution increases with increasing likelihood.

In a mixing analysis, for a given event, the likelihood that the opposite side muon is a real muon can be calculated. With this information and knowing the $p_T^{\rm rel}$ of the tag muon, the dilution can be extracted from the corresponding functional form of dilution versus $p_T^{\rm rel}$. This tagging method is applied in the mixing analyses described in Sections 4 and 5 but the dilution dependencies on likelihood and p_T^{rel} are obtained using the initial 168 pb^{-1} data sample.

A potential improvement would be to linearly interpolate between different likelihood bins. In this way, a more accurate event-by-event dilution could be predicted. Figure 44 shows the two dimensional dependence of absolute dilution of CMU muons (corrected by trigger side dilution) on likelihood and $p_T^{\rm rel}$. Similar shapes are obtained for other muon types.

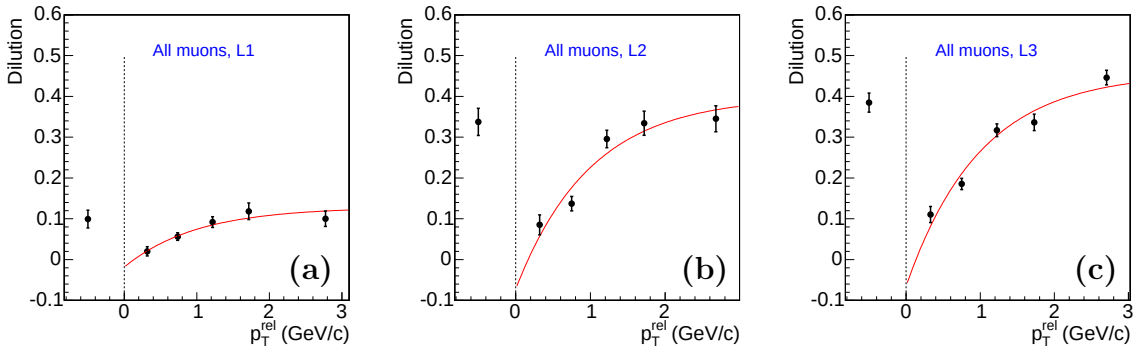

Figure 38: Raw dilution as a function of $p_T^{\rm rel}$ in the three likelihood ranges, (a) $\mathcal{L}1$, (b) $\mathcal{L}2$ and (c) $\mathcal{L}3$, for all muon types combined. Entries at $p_T^{rel} = -0.5$ GeV/c correspond to isolated muons.

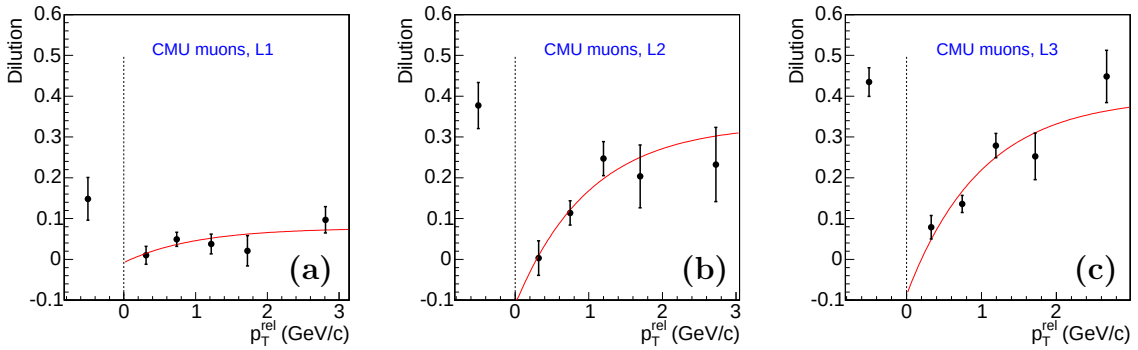

Figure 39: Raw dilution as a function of $p_T^{\rm rel}$ in three likelihood ranges, (a) $\mathcal{L}1$, (b) $\mathcal{L}2$ and (c) $\mathcal{L}3$, for CMU muons. Entries at $p_T^{rel} = -0.5$ GeV/c correspond to isolated muons.

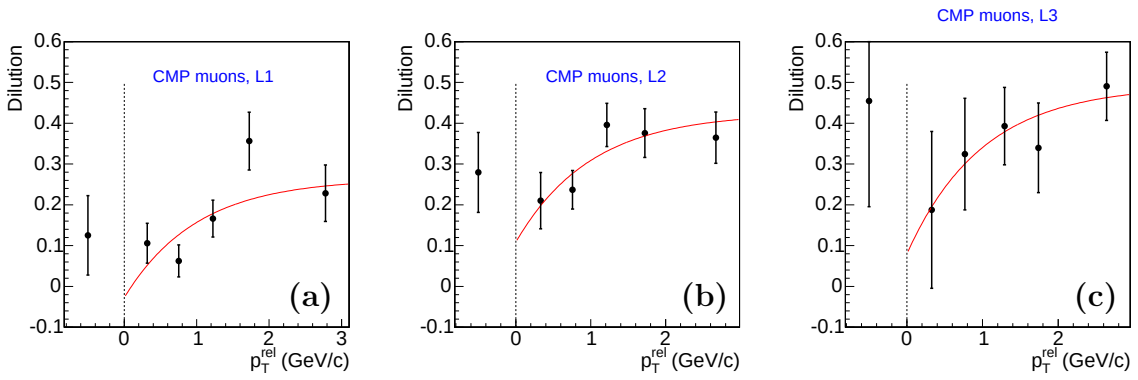

Figure 40: Raw dilution as a function of $p_T^{\rm rel}$ in three likelihood ranges, (a) $\mathcal{L}1$, (b) $\mathcal{L}2$ and (c) $\mathcal{L}3$, for CMP muons. Entries at $p_T^{rel} = -0.5$ GeV/c correspond to isolated muons.

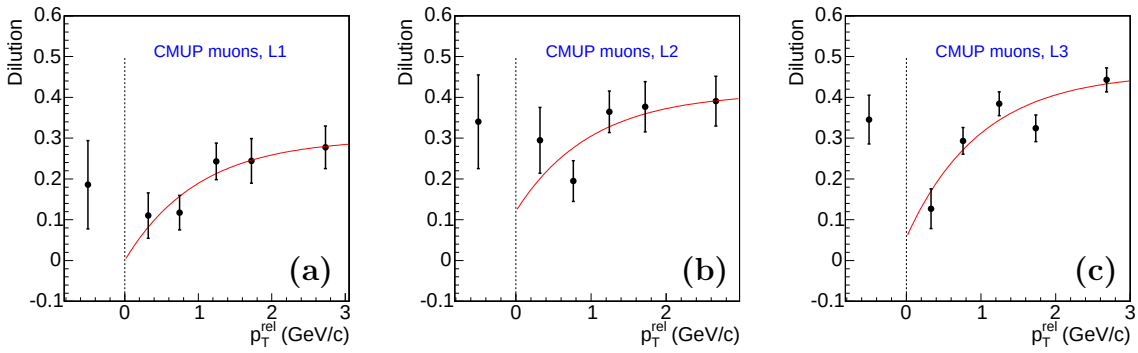

Figure 41: Raw dilution as a function of $p_T^{\rm rel}$ in three likelihood ranges, (a) $\mathcal{L}1$, (b) $\mathcal{L}2$ and (c) $\mathcal{L}3$, for CMUP muons. Entries at $p_T^{rel} = -0.5$ GeV/c correspond to isolated muons.

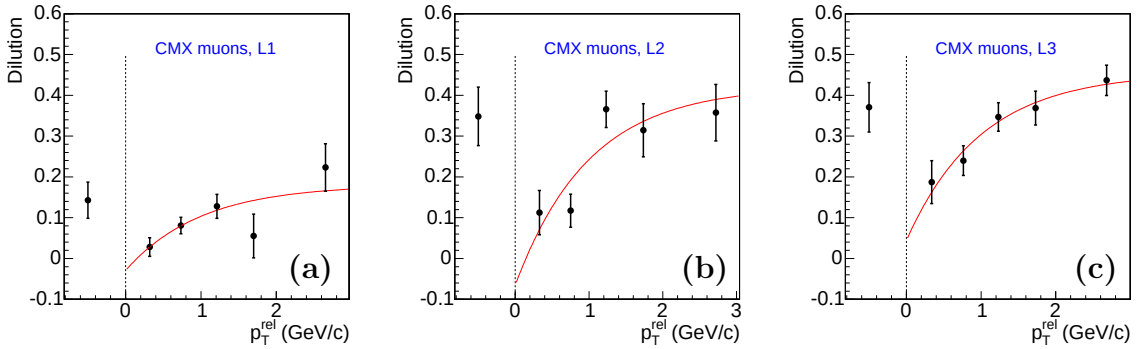

Figure 42: Raw dilution as a function of $p_T^{\rm rel}$ in three likelihood ranges, (a) $\mathcal{L}1$, (b) $\mathcal{L}2$ and (c) $\mathcal{L}3$, for CMX muons. Entries at $p_T^{rel} = -0.5$ GeV/c correspond to isolated muons.

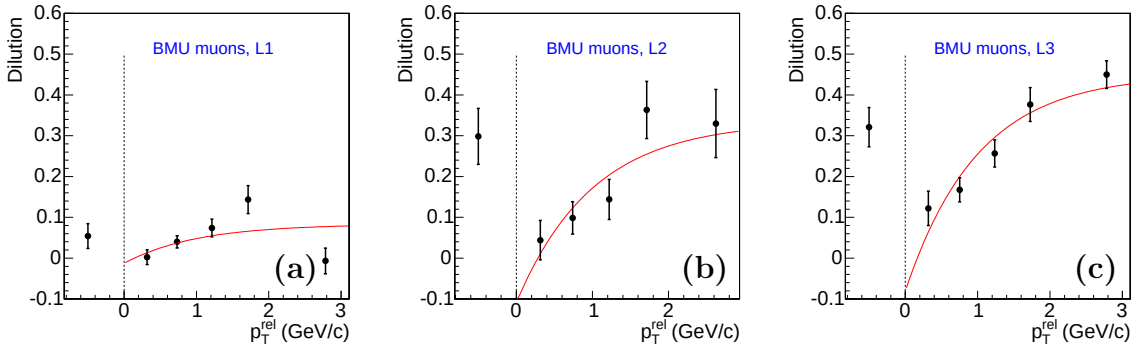

Figure 43: Raw dilution as a function of $p_T^{\rm rel}$ in three likelihood ranges, (a) $\mathcal{L}1$, (b) $\mathcal{L}2$ and (c) $\mathcal{L}3$, for BMU muons. Entries at $p_T^{rel} = -0.5$ GeV/c correspond to isolated muons.

	CMU	CMP	CMUP	CMX	BMU
$0 < \mathcal{L} < 0.8$					
$\mathcal{D}_{max}(\%)$	8 ± 2	26 ± 5	30 ± 4	18 ± 3	8 ± 2
$\mathcal{D}_{iso}(\%)$	15 ± 5	13 ± 10	19 ± 11	14 ± 4	5 ± 3
$0.8 < \mathcal{L} < 0.95$					
$\mathcal{D}_{max}(\%)$	33 ± 5	43 ± 5	41 ± 5	42 ± 5	33 ± 5
$\mathcal{D}_{iso}(\%)$	38 ± 6	28 ± 10	34 ± 11	35 ± 7	30 ± 7
$0.95 < \mathcal{L} < 1.0$					
$\mathcal{D}_{max}(\%)$	40 ± 4	49 ± 8	46 ± 3	46 ± 3	45 ± 3
$\mathcal{D}_{iso}(\%)$	43 ± 4	45 ± 25	35 ± 6	37 ± 6	32 ± 5

Table 8: Maximum raw dilution $\mathcal{D}_{max}$ and raw dilution of isolated muons $\mathcal{D}_{iso}$ for each muon type in different likelihood bins. The dilution clearly increases with increasing likelihood. The absolute dilution can be obtained from raw dilution by division with the trigger side dilution given in Equation 60.

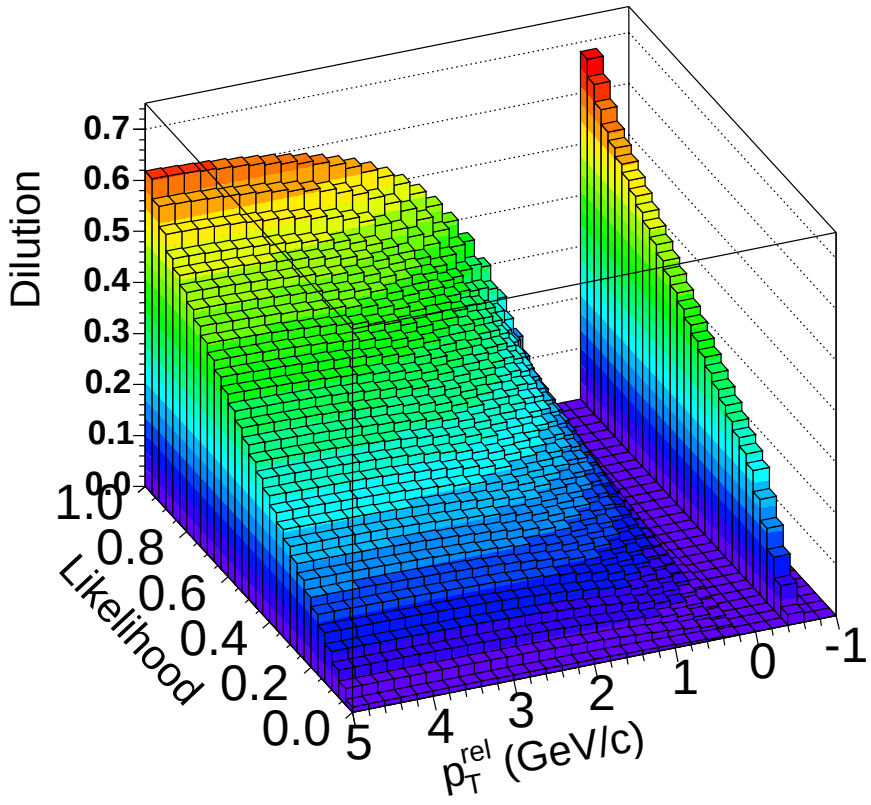

Figure 44: Absolute dilution as a function of $p_T^{\rm rel}$ and likelihood for CMU muons. Entries at $p_T^{rel} = -0.5$ GeV/c correspond to isolated muons. The dilution is corrected by: $\mathcal{D} = \mathcal{D}_{raw}/\mathcal{D}_{trg}$.

4 Flavor Tagging Re-Calibration and Measurement of $B^0\bar{B}^0$ Oscillation Frequency in Semileptonic B Decays

As already discussed in Section 3, several methods of same side and opposite side B flavor tagging exist. Opposite side tagging methods identify the flavor of the other B hadron produced in the initial collision along with the B meson of interest. Using large samples of inclusive lepton-SVT events, the performance of several opposite side flavor taggers has been studied at CDF in Run II: jet charge tagging (JQT) [41], opposite side muon tagging (SMT) [42, 43, 44] (discussed in detail in Section 3) and opposite side electron tagging (SET) [45].

In this section we describe a study of the above three opposite side tagging methods as well as a measurement of the $B^0\bar{B}^0$ mixing frequency Δm_d with semileptonic $B \to D\ell\nu$ decays. Using events from CDF's lepton-SVT trigger, the three decay channels $B \to D^+\ell\nu X$ with $D^+ \to K^-\pi^+\pi^+$, $B \to D^0\ell\nu X$ with $D^0 \to K^-\pi^+$ and $B \to D^{*+}\ell\nu X$ with $D^{*+} \to D^0\pi^+$ are reconstructed, providing high statistics B samples. The $D^+\ell$ and $D^{*+}\ell$ samples are dominated by neutral B^0 decays while the $D^0\ell$ signature corresponds to a sample enriched in B^+ decays. These samples can be used at CDF for precision measurements of B lifetimes τ_{B^0} and τ_{B^+} as well as the B^0 oscillation frequency Δm_d. On the other hand, since B^0 mixing is a well established phenomenon and Δm_d has been precisely measured [1], the semileptonic B samples can be used to verify and re-calibrate the predicted dilution of the three opposite side flavor tagging algorithms. While the initial calibration was performed using the inclusive lepton-SVT dataset, the re-calibration on the $D^{+/0}\ell$ samples provides a dilution measurement in an environment very similar to the $D_s\ell$ sample on which the search for B_s^0 oscillations is performed. The first Run II CDF result on B_s^0 oscillations (see Section 5) is not an observation but rather a lower limit on Δm_s. In this case, the dilutions of the flavor tags need to be known and used as input in the B_s^0 mixing fit. Thus they must be estimated from other $B \to D\ell\nu$ samples.

There is another significant difference between B^0 and B_s^0 mixing. Since the B^0 oscillation is relatively slow, a χ^2 fit in bins of the B proper decay time was sufficient in the past to measure Δm_d [46]. In order to detect the rapid B_s^0 oscillations, an unbinned likelihood fitting method is needed to exploit the proper time information as well as predicted tagging dilutions on an event-by-event basis. This study also serves as a validation of the unbinned likelihood fit procedure on real data.

4.1 Data Selection and Signal Reconstruction

4.1.1 Data Selection

The data sample used for this analysis was recorded from February 2002 until August 2004 with the CDF detector. The events were selected by the lepton-SVT trigger described in Section 2.2.1. We confirm offline the trigger requirements using a package that emulates the Level 3 trigger decision [38]. The integrated luminosity

of the dataset used in this analysis is ~ 355 pb^{-1}, twice as large as the data sample used for the study of the opposite side muon tagger discussed in Section 3. We require a good-run selection list consisting of 1348 runs as recommended by the B Physics Analysis Kernel (BPAK) group [47]. This selection ensures that all detector components were properly functioning while data used in this analysis were recorded. All tracks are refit using the description of passive material settings and COT covariance scaling as outlined by the BPAK group [47]. If available, Layer 00 silicon hits are included in the track refit procedure. In the flavor tagging analysis described in Section 3, the Layer 00 information was not used. It became available during the analysis development, so we include the Layer 00 hits for the mixing analysis. With an extra constraint on the particle tracks we expect that the vertexing resolution and consequently the decay length resolution will improve significantly. We find that, indeed, the decay distance resolution improves by $\sim 10\%$ after Layer 00 information is used. Basic track quality cuts on all tracks used in this analysis require two or more axial and stereo COT superlayers with at least 5 hits each and at least 3 silicon $r\phi$ hits on a track.

4.1.2 $D\ell$ Signal Reconstruction

The lepton-SVT data set is used to search for semileptonic decays $B \to D\ell\nu$ decays. A schematic representation of such a semileptonic B decay is displayed in Figure 4. First, a D candidate is searched by requiring the SVT track to be one of the tracks from the D meson decay. The D candidates are reconstructed in three decay modes $D^0 \to K^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$ and $D^{*+} \to D^0\pi^+$ ($D^0 \to K^-\pi^+$). Using a secondary vertex fitting software package [34], the tracks forming the D candidate are refit with a common vertex constraint, referred to as the tertiary vertex. The secondary vertex, where the B decays to a lepton and a D meson, is obtained by simultaneously intersecting the trajectory of the lepton track with the flight path of the D candidate.

An optimization of the three $D\ell$ signal yields was performed with the purpose of maximizing $S/\sqrt{S+B}$ which, as seen in Equation (40) is one of the factors determining the sensitivity to neutral B oscillations. Here, S and B represent the number of signal and background events, respectively. For each $D\ell$ mode, the signal region around the D meson mass is in the range $(m_D - 3\sigma, m_D + 3\sigma)$ with $\sigma \approx 8$ MeV/c^2. The signal events for each $D\ell$ mode are obtained from an inclusive $B \to D\ell\nu$ Monte Carlo sample while the background events are selected from the side-bands of the D meson invariant mass distributions from real data. The number of signal events from the Monte Carlo sample is comparable to the number of signal events observed in real data. The original selection requirements from this optimization procedure are summarized in Tables 9, 10 and 11. In order to facilitate better cross checking with an independent analysis proceeding in parallel [49], we use a slightly modified set of cuts that would constitute the same basic event sample to be used by both analyses. The selection requirement from this combined proposal are also listed in Tables 9 through 11. The meaning of some of the variables listed in Tables 9, 10 and 11 is given below:

- $p_T(K)$: Transverse momentum of the kaon from D meson decays
- Probability (B Vertex): Vertex probability for the B vertex fit
- 2-dimensional $\chi^2_{xy}(D)$: Two-dimensional χ^2 for the D vertex fit
- $L_{xy}(PV \rightarrow D)/\sigma(L_{xy})$: Significance of reconstructed L_{xy} from primary vertex to the D vertex
- $L_{xy}(PV \rightarrow D)$: Straight cut on reconstructed L_{xy} from primary vertex to the D meson vertex
- $\sigma_{ct^*}(PV \rightarrow B)$: Error on the B meson pseudo proper decay time
- $ct(D)$: Proper decay time of the D meson as obtained from the decay length between the B meson and D meson decay points
- $m(D\ell)$: Invariant mass of the $D\ell$ pair
- $m(D^{*+} - m(D^0))$: Mass difference between the D^{*+} and D^0 states
- $p_T(\pi_*)$: Transverse momentum of the pion from $D^{*+} \rightarrow D^0\pi^+_*$ decays

Removing the $D^{*+}\ell$ candidates from the $D^0\ell$ sample increases the charged B content of the $D^0\ell$ data. Care was taken in this analysis to treat the D^{*+} and D^0 reconstruction as similarly as possible. The D^{*+} candidates were formed by vertexing the $K\pi$ combinations just like in the D^0 case and pointing the D^0 candidate back to the lepton to obtain the B vertex. The soft pion π^+_* from the decay $D^{*+} \rightarrow D^0\pi^+_*$ was not used in the vertexing procedure but only used to determine the mass difference $m(K\pi\pi_*) - m(K\pi)$ and to decide whether the D^0 combination originates from a D^{*+} candidate. To increase the D^{*+} detection efficiency, we allow the soft pion track to be formed with silicon hits only and also we require a low p_T cut on the soft pion $p_T(\pi_*) > 0.2$ GeV/c. This increases the D^{*+} identification efficiency and thus increases the B^+ purity of the $D^0\ell$ sample.

The corresponding D mass distributions obtained with the combined proposal of selection requirements indicate large charm signals and are shown in Figure 45 for the right-sign lepton-D combinations: $D^+\ell^-$, $D^{*+}\ell^-$ and $D^0\ell^-$ with and without the $D^{*+}\ell^-$ candidates removed. The D mass distributions in Figure 45 are fit to a double Gaussian signal function plus a linear background function. The natural width of the D mass is described by a Breit-Wigner function with a width of the order of ~ 10 meV. The mass resolution of the CDF detector is of the order of ~ 10 MeV, so the observed mass distribution is dominated by the detector resolution. This resolution is expected to be Gaussian with the width increasing with decreasing momentum of the tracks involved in the mass reconstruction (low momentum tracks have larger curvature leading to better momentum measurement). Effectively, a double Gaussian function accurately describes the D meson mass distributions. The second Gaussian function is introduced to handle events with larger mass resolution, typically because of one or more mis-measured tracks. The fit results for each decay signature are shown

	Original optimization	Combined proposal
$p_T(K)$ (GeV/c)	> 0.6	> 0.7
Probability (B Vertex)	$> 10^{-5}$	$> 10^{-5}$
2-dimensional $\chi^2_{xy}(D)$	< 7	< 10
$L_{xy}(PV \to D)/\sigma(L_{xy})$	> 12	> 11
$L_{xy}(PV \to D)$ (cm)	–	> 0.04
$\sigma_{ct^*}(PV \to B)$ (cm)	< 0.04	< 0.04
$ct(D)$ (cm)	$-0.05 < ct < 0.2$	$-0.015 < ct < 0.2$
$m(D^+\ell)$ (GeV/c^2)	$2.3 < m < 5.3$	$2.3 < m < 5.0$

Table 9: Summary of selection requirements as a result of the original optimization procedure and the combined set of cuts for the $D^+\ell$ sample.

	Original optimization	Combined proposal
$p_T(K)$ (GeV/c)	> 0.4	> 0.4
Probability (B Vertex)	$> 10^{-6}$	$> 10^{-6}$
2-dimensional $\chi^2_{xy}(D)$	< 10	< 10
$L_{xy}(PV \to D)/\sigma(L_{xy})$	> 6.5	> 6
$L_{xy}(PV \to D)$ (cm)	–	> 0.02
$\sigma_{ct^*}(PV \to B)$ (cm)	< 0.04	< 0.04
$ct(D)$ (cm)	$-0.05 < ct < 0.1$	$-0.015 < ct < 0.1$
$m(D^{*+}\ell)$ (GeV/c^2)	$2.3 < m < 5.3$	$2.3 < m < 5.0$
$m(D^{*+} - m(D^0))$ (GeV/c^2)	$0.1440 < m < 0.1475$	$0.1435 < m < 0.1475$
$p_T(\pi_*)$ (GeV/c)	> 0.2	> 0.2

Table 10: Summary of selection requirements as a result of the original optimization procedure and the combined set of cuts for the $D^{*+}\ell$ sample.

	Original optimization	Combined proposal
$p_T(K)$ (GeV/c)	> 0.4	> 0.4
Probability(B Vertex)	$> 10^{-6}$	$> 10^{-6}$
2-dimensional $\chi^2_{xy}(D)$	< 10	< 10
$L_{xy}(PV \to D)/\sigma(L_{xy})$	> 6.5	> 6
$L_{xy}(PV \to D)$ (cm)	–	> 0.02
$\sigma_{ct^*}(PV \to B)$ (cm)	< 0.04	< 0.04
$ct(D)$ (cm)	$-0.05 < ct < 0.1$	$-0.015 < ct < 0.1$
$m(D^0\ell)$ (GeV/c^2)	$2.3 < m < 5.3$	$2.3 < m < 5.0$

Table 11: Summary of selection requirements as a result of the original optimization procedure and the combined set of cuts for the $D^0\ell$ sample.

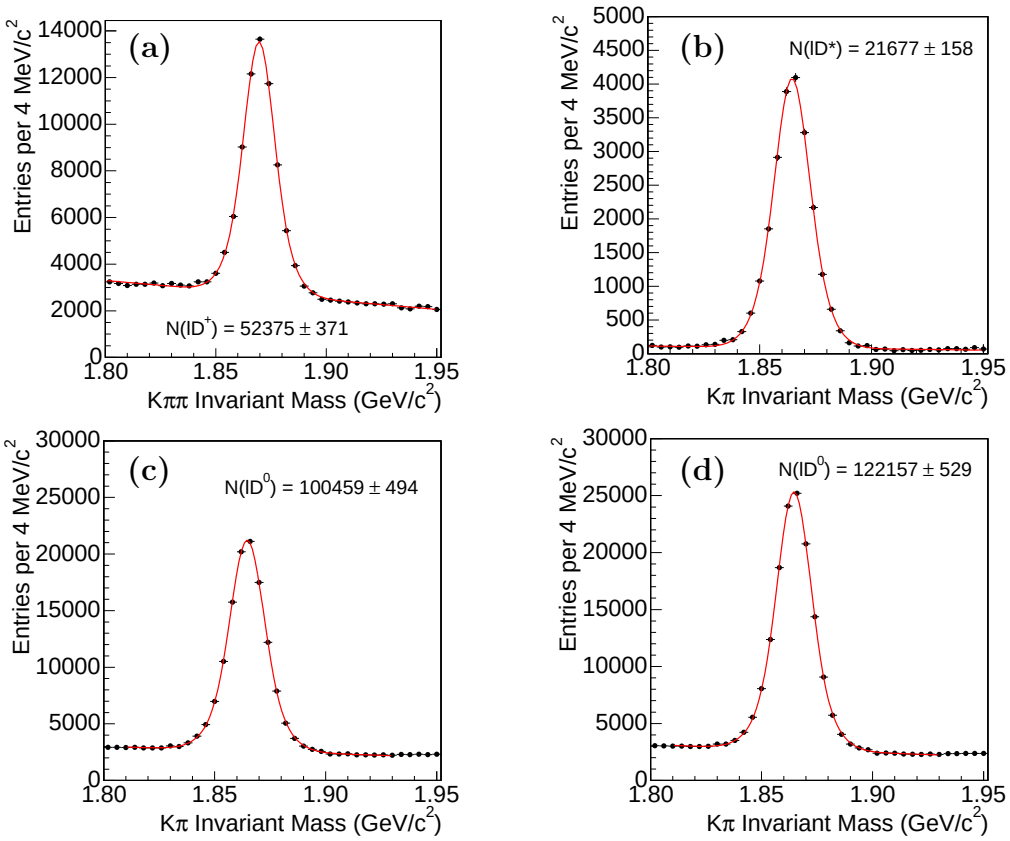

Figure 45: D meson mass distribution for the right-sign lepton-D combinations: (a) $D^+\ell^-$, (b) $D^{*+}\ell^-$, (c) $D^0\ell^+$ with the D^{*+} candidates removed and (d) $D^0\ell^+$ where no D^{*+} removal was performed. The D mass distributions are fit with a double Gaussian signal function plus a linear background.

Parameter	$D^+\ell$	$D^{*+}\ell$	$D^0\ell$ (no D^{*+})	$D^0\ell$ (including D^{*+})
$m(D)$ (MeV/c^2)	1869.62 ± 0.05	1864.47 ± 0.07	1864.67 ± 0.04	1864.65 ± 0.03
f_1	0.504 ± 0.101	0.508 ± 0.127	0.567 ± 0.052	0.556 ± 0.049
σ_1 (MeV/c^2)	6.12 ± 0.34	6.96 ± 0.50	7.02 ± 0.19	7.00 ± 0.18
σ_2 (MeV/c^2)	10.4 ± 0.68	11.4 ± 0.75	12.1 ± 0.50	12.0 ± 0.44
c_1	18200 ± 368	855 ± 73	14500 ± 516	15700 ± 523
c_2	-8270 ± 194	-412 ± 38	-6370 ± 273	-6990 ± 277

Table 12: Mass fit results for each of the three decay signatures. The mass peaks are fit with two Gaussian functions for signal distributions and a linear function for background distributions. The two Gaussian functions have the same mean $m(D)$. The weight of the first Gaussian is given by f_1. The widths are σ_1 and σ_2. The parameters c_1 and c_2 describe the offset and the slope of the linear background function.

	N	S/B	$S/\sqrt{S+B}$
$D^+\ell$	$52,375 \pm 371$	2.0	183.3
$D^{*+}\ell$	$21,677 \pm 158$	23.1	141.7
$D^0\ell$ (no D^{*+})	$100,458 \pm 494$	3.5	274.5
$D^0\ell$ (incl. D^{*+})	$122,157 \pm 529$	4.2	308.0

Table 13: Summary of signal yield N, signal to background ratio S/B and $S/\sqrt{S+B}$ for the lepton-charm modes. We list separately $D^0\ell^+$ yields with the D^{*+} candidates removed (no D^{*+}) and with no D^{*+} removal (incl. D^{*+}).

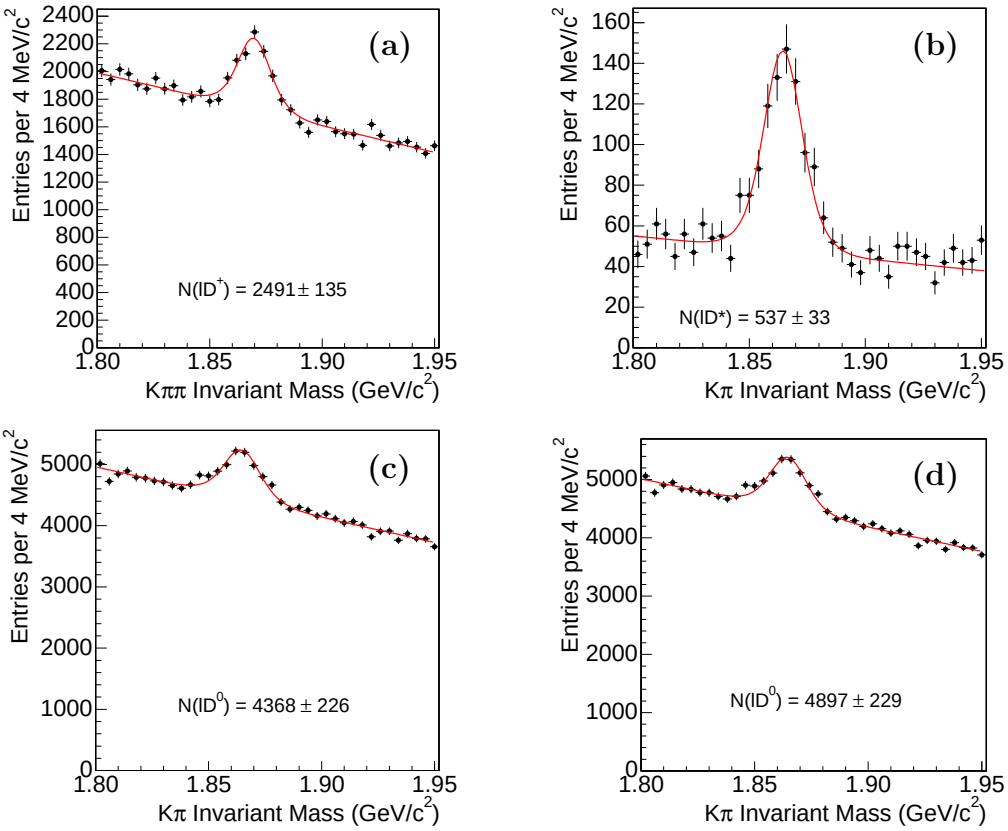

Figure 46: D mass distribution for the wrong-sign lepton-D combinations: (a) $D^+\ell^+$, (b) $D^{*+}\ell^+$, (c) $D^0\ell^+$ with the D^{*+} candidates removed and (d) $D^0\ell^+$ where no D^{*+} removal was performed. The D mass distributions are fit to a double Gaussian signal function plus a linear background. The double Gaussian shape is fixed from the right-sign fit.

in Table 12. We define the signal region as $\pm 3\,\sigma$ around the fitted D meson mass and the side-band regions as $\pm$ (4-8) σ away from the fitted D mass peak. Here σ is the width of the narrow Gaussian. The chosen side-band regions are wider that the signal region to obtain a more accurate description of the side-band distributions. The signal yields are summarized in Table 13.

Figure 46 shows the D mass distribution for the wrong-sign lepton-D combination:

	N	S/B	$S/\sqrt{S+B}$	$f_{WS}(\%)$
$D^+\ell$	2491 ± 135	0.2	18.0	4.8
$D^{*+}\ell$	537 ± 33	1.0	16.1	2.5
$D^0\ell$ (no D^{*+})	4368 ± 226	0.1	19.5	4.4
$D^0\ell$ (incl. D^{*+})	4897 ± 229	0.1	20.7	4.0

Table 14: Summary of signal yield N, S/B, $S/\sqrt{S+B}$ for the wrong-sign lepton-charm modes. The fraction of wrong-sign signal yield with respect to the right-sign signal yield f_{WS} is also shown. We again list separately the $D^0\ell^+$ yields with the D^{*+} candidates removed (no D^{*+}) and with no D^{*+} removal (incl. D^{*+}).

$D^+\ell^+$, $D^{*+}\ell^+$, $D^0\ell^+$ with the D^{*+} candidates removed and $D^0\ell^+$ where no D^{*+} removal was performed. The D mass distributions are again fit to a double Gaussian signal function plus a linear background. The parameters describing the signal distribution, m_D, f_1, σ_1 and σ_2 are fixed from the right-sign fit result. Clear D mass signals are observed, corresponding to a yield of the order of a few percent of the corresponding right-sign yields: $f_{WS}(D^+) = 4.8\%$, $f_{WS}(D^{*+}) = 2.5\%$, $f_{WS}(D^0) = 4.4\%$ with D^{*+} removal and $f_{WS}(D^0) = 4.0\%$ with no D^{*+} removal.

A summary of the $D\ell$ signal yields, S/B and $S/\sqrt{S+B}$ for the wrong-sign lepton-charm modes are listed in Table 14. We again provide separately the $D^0\ell^+$ yields with the D^{*+} candidates removed and with no D^{*+} removal.

4.2 Flavor Tagging

This analysis employs three opposite side flavor tagging algorithms: opposite side electron and muon tagging as well as jet charge tagging. The tagging algorithms provide a predicted dilution based on event-by-event kinematic information such as $p_T^{\rm rel}$ or the value of the lepton likelihood indicating the purity of the lepton identification. The dilution predictions for each tagger have been studied on a large inclusive B sample using the lepton-SVT trigger data.

The opposite side muon tagging algorithm (SMT) was described in detail in Section 3. The dilution for the muon tagging method is predicted as a function of the muon type, the muon likelihood and $p_T^{\rm rel}$. For isolated muons, only the muon type and the muon likelihood is used to evaluate the dilution.

The opposite side electron tagging algorithm (SET) [45] combines various electron identification variables such as hadronic and electromagnetic energies, electron momentum, information from the central pre-radiator detector (CPR), χ^2 matching variables in the maximum shower detector (CES) as well as track energy loss in the drift chamber $\mathrm{d}E/\mathrm{d}x$ into a likelihood function. This variable represents the probability that a track extrapolated to the calorimeter is a real electron. A strong dependence of the dilution on likelihood, $p_T^{\rm rel}$ and the signed impact parameter d_0^{sign} is found. The quantity d_0^{sign} is used to detect background from unseen conversions. The unseen conversions are events in which a photon interacting with the electromagnetic field of a nucleus converts into a e^+e^- pair and the conversion removal algorithm

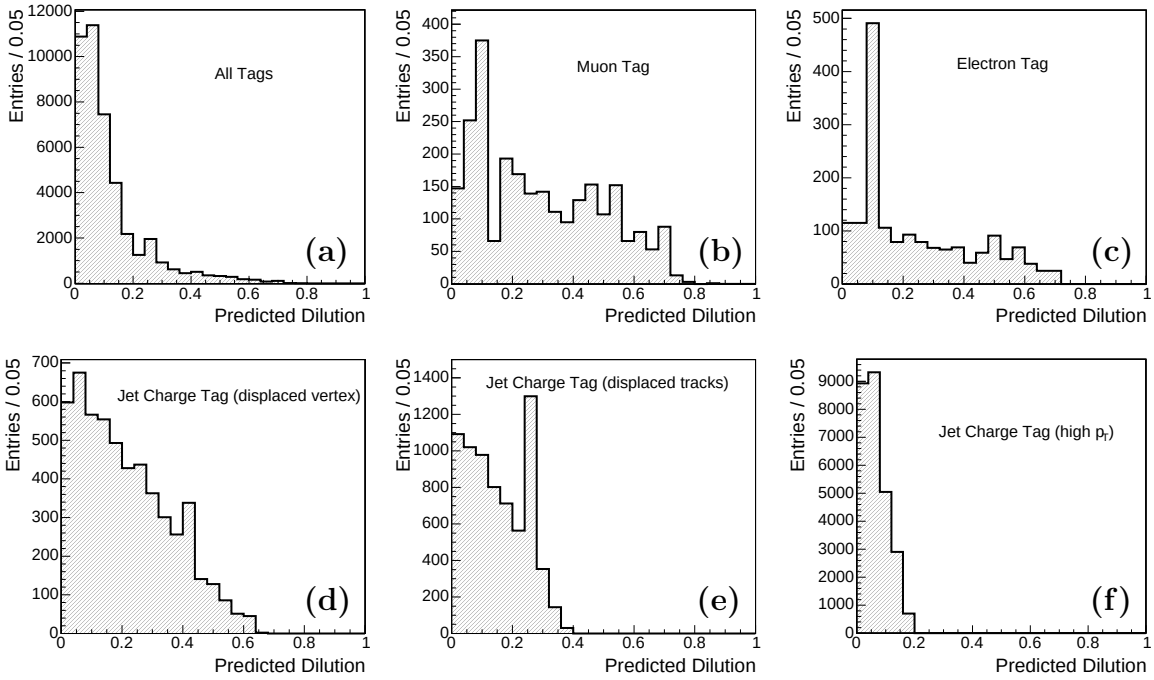

Figure 47: Predicted dilution distribution for lepton-D^+ candidates. The six plots correspond to (a) all tags combined, (b) muon tag, (c) electron tag, jet charge tag with (d) displaced vertex, (e) displaced tracks and (e) high p_T

does not identify these events. The signed impact parameter d_0^{sign} is defined as the product of the beam-spot corrected impact parameter d_0 signed by the charge q of the electron candidate, $d_0^{sign} = d_0 \times q$. Conversions originating from photon conversions have, up to resolution effects, positive signed impact parameter. This is explained by the fact that the photons are massless and mainly originate from the primary vertex, from $\pi^0 \to \gamma\gamma$ decays. On the other hand, photons from massive particle decays have symmetric signed impact parameter distributions. The opposite side tag electron are divided in two subsamples with positive and negative signed impact parameter. It has been found [45] that the tag electrons with positive signed impact parameter have lower dilution as this sample contains the unseen conversion events. The dilution predictions for electron tagging are determined by binning the data in electron likelihood and parameterizing the dilution versus track $p_T^{\rm rel}$. Different parameterizations are available for d_0^{sign} being positive (potential unseen conversion background) or negative.

In general, jet charge tagging exploits the fact that the sign of the momentum weighted sum of the particle charges of an opposite side b jet is the same as the charge of the b quark producing this jet. The dilution prediction for the jet charge tagging algorithm (JQT) [41] is parameterized as a function of the calculated value of jet charge. Different parameterizations are calculated for subsamples of different tag categories. Jets containing a displaced vertex with respect to the primary vertex constitute the highest quality jets very likely coming from heavy flavor decays. Sec-

ondly, jets containing displaced tracks based on a likelihood that a tracks comes from a long-lived particle are used. Finally, all remaining jets are utilized, but the jet with the highest transverse momentum p_T is chosen in case multiple jets are found in a certain event.

For a given event more than one opposite side tag can be available. Only one tag is chosen using the following decision hierarchy: if an SMT tag is available it is chosen first, while SET is next. Both muon and electron tags must have a lepton likelihood greater than 5%. If no lepton tag is available, the jet charge tag is used with displaced vertex first, displaced tracks next and high-momentum jet charge tag last. If not tag is found in an event, it is considered a no-tag event and the predicted dilution is set to zero.

Figure 47 shows, as an example, the predicted dilution distributions for the lepton-D^+ sample. The six plots correspond to all tags combined and then separately: muon tag, electron tag, jet charge tag with displaced vertex, displaced tracks and high-momentum respectively. The spikes in the lepton tag dilution distributions correspond to low likelihood leptons for which the dilution versus p_T^{rel} dependence is almost flat resulting in a narrow range of the predicted dilution. The spikes in the jet charge tag dilution distributions correspond to jets with charge equal to unity.

4.3 Sample Composition

None of the three decay final states $D^+\ell$, $D^{*+}\ell$ and $D^0\ell$ represent pure samples of either B^0 or B^+ decays. Each $D\ell$ channel corresponds to a certain mixture of B^0 and B^+ mesons in the initial state. Figure 48 is a schematic illustration of the cross talk and feed-down contributing to the sample composition of the various lepton-D final states. The sample composition problem is to determine the fractions of the B^0/B^+ components in each of the three decay signatures as well as the K-factor distributions and trigger efficiencies (see Section 4.5) corresponding the B^0 and B^+ components.

The sample composition can be determined in various ways. We parameterize the sample composition using the method outlined in Reference [46]. In this approach, four parameters f^{**}, R_f, P_V and ϵ^* are used to fully characterize the sample composition of each decay signature. The first parameter f^{**} represents the fraction that a B meson decays into a D^{**} state, where by D^{**} we understand any excited D state other than D^*. R_f is defined as $R_f = f^*/f$, where f^* (f) represents the fraction how often a B meson decays into a D^* (D) state. The probability that a D^{**} state decays into a D^* meson is given by P_V, while $1 - P_V$ gives the probability that a D^{**} decays directly into a D meson. To conserve the total decay probability, the three parameters f^{**}, f^* and f are related by $f^{**}+f^*+f = 1$. The reconstruction efficiency of the soft pion π_*^+ from $D^{*+} \to D^0\pi_*^+$ is given by ϵ^*. Table 15 shows the various B decay chains and their contributions to the sample composition of the three decay channels $D^+\ell$, $D^{*+}\ell$ and $D^0\ell$ including the relevant branching fractions in terms of the sample composition parameters. The actual values of the branching fractions [1] involved are given in Table 16.

In addition to the branching fractions of the various decay modes involved [1], differences in the trigger and reconstruction efficiencies due to different kinematics of the

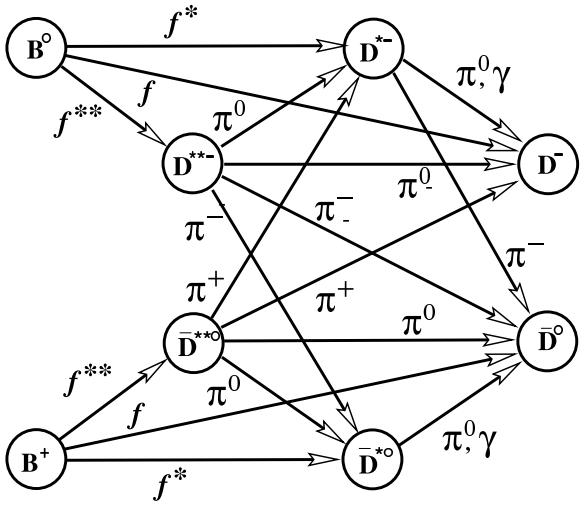

Figure 48: A schematic view illustrating the sample composition of the various $B \to D\ell\nu$ transitions.

i	Mode		Decay Channel	Sample Composition Fraction (f_{SC})
0	ℓD^+	$\bar{B}^0$	$\to D^+ \to K\pi\pi$	$fBR(K\pi\pi)$
1	ℓD^+	$\bar{B}^0$	$\to D^{*+} \to \gamma D^+ \to K\pi\pi$	$f^*BR(\gamma D^+)BR(K\pi\pi)$
2	ℓD^+	$\bar{B}^0$	$\to D^{*+} \to \pi^0 D^+ \to K\pi\pi$	$f^*BR(\pi^0 D^+)BR(K\pi\pi)$
3	ℓD^+	$\bar{B}^0$	$\to D^{**+} \to \pi^0 D^+ \to K\pi\pi$	$f^{**}(1-P_V)(1/3)BR(K\pi\pi)$
4	ℓD^+	$\bar{B}^0$	$\to D^{**+} \to \pi^0 D^{*+} \to \gamma D^+ \to K\pi\pi$	$f^{**}P_V(1/3)BR(\gamma D^+)BR(K\pi\pi)$
5	ℓD^+	$\bar{B}^0$	$\to D^{**+} \to \pi^0 D^{*+} \to \pi^0 D^+ \to K\pi\pi$	$f^{**}P_V(1/3)BR(\pi^0 D^+)BR(K\pi\pi)$
6	ℓD^+	B^-	$\to D^{**0} \to \pi^- D^+ \to K\pi\pi$	$f^{**}(1-P_V)(2/3)BR(K\pi\pi)$
7	ℓD^+	B^-	$\to D^{**0} \to \pi^- D^{*+} \to \gamma D^+ \to K\pi\pi$	$f^{**}P_V(2/3)BR(\gamma D^+)BR(K\pi\pi)$
8	ℓD^+	B^-	$\to D^{**0} \to \pi^- D^{*+} \to \pi^0 D^+ \to K\pi\pi$	$f^{**}P_V(2/3)BR(\pi^0 D^+)BR(K\pi\pi)$
9	ℓD^0	$\bar{B}^0$	$\to D^{*+} \to \pi^+ D^0 \to K\pi$	$f^*BR(\pi^+ D^0)BR(K\pi)(1-\epsilon^*)$
10	ℓD^0	$\bar{B}^0$	$\to D^{**+} \to \pi^0 D^{*+} \to \pi^+ D^0 \to K\pi$	$f^{**}P_V(1/3)BR(\pi^+ D^0)BR(K\pi)(1-\epsilon^*)$
11	ℓD^0	$\bar{B}^0$	$\to D^{*+} \to \pi^+ D^0 \to K\pi$	$f^{**}(1-P_V)(2/3)BR(K\pi)$
12	ℓD^0	$\bar{B}^0$	$\to D^{**+} \to \pi^+ D^{*0} \to \gamma D^0 \to K\pi$	$f^{**}P_V(2/3)BR(\gamma D^0)BR(K\pi)$
13	ℓD^0	$\bar{B}^0$	$\to D^{**+} \to \pi^+ D^{*0} \to \pi^0 D^0 \to K\pi$	$f^{**}P_V(2/3)BR(\pi^0 D^0)BR(K\pi)$
14	ℓD^0	B^-	$\to D^0 \to K\pi$	$fBR(K\pi)$
15	ℓD^0	B^-	$\to D^{*0} \to \gamma D^0 \to K\pi$	$f^*BR(\gamma D^0)BR(K\pi)$
16	ℓD^0	B^-	$\to D^{*0} \to \pi^0 D^0 \to K\pi$	$f^*BR(\pi^0 D^0)BR(K\pi)$
17	ℓD^0	B^-	$\to D^{**0} \to \pi^0 D^0 \to K\pi$	$f^{**}(1-P_V()1/3)BR(K\pi)$
18	ℓD^0	B^-	$\to D^{**0} \to \pi^0 D^{*0} \to \gamma D^0 \to K\pi$	$f^{**}P_V(1/3)BR(\gamma D^0)BR(K\pi)$
19	ℓD^0	B^-	$\to D^{**0} \to \pi^0 D^{*0} \to \pi^0 D^0 \to K\pi$	$f^{**}P_V(1/3)BR(\pi^0 D^0)BR(K\pi)$
20	ℓD^0	B^-	$\to D^{**0} \to \pi^- D^{*+} \to \pi^+ D^0 \to K\pi$	$f^{**}P_V(2/3)BR(\pi^0 D^0)BR(K\pi)(1-\epsilon^*)$
21	ℓD^{*+}	$\bar{B}^0$	$\to D^{*+} \to \pi^+ D^0 \to K\pi$	$f^*BR(\pi^+ D^0)BR(K\pi)\epsilon^*$
22	ℓD^{*+}	$\bar{B}^0$	$\to D^{**+} \to \pi^0 D^{*+} \to \pi^+ D^0 \to K\pi$	$f^{**}P_V(1/3)BR(\pi^+ D^0)BR(K\pi)\epsilon^*$
23	ℓD^{*+}	B^-	$\to D^{**0} \to \pi^- D^{*+} \to \pi^+ D^0 \to K\pi$	$f^{**}P_V(2/3)BR(\pi^+ D^0)BR(K\pi)\epsilon^*$

Table 15: Various B decay chains and their contributions to the sample composition of the three decay signatures $D^+\ell$, $D^{*+}\ell$ and $D^0\ell$ including the relevant branching fractions.

various decay chains also change the sample composition. These relative efficiencies are estimated using realistic Monte Carlo simulations of individual $B \to D\ell\nu$ decay

Decay	Branching Fraction (%)
$D^0 \to K^-\pi^+$	3.8
$D^+ \to K^-\pi^+\pi^+$	9.2
$D^{*+} \to D^+\gamma$	1.6
$D^{*+} \to D^+\pi^0$	30.7
$D^{*+} \to D^+\pi^0$	30.7
$D^0 \to D^0\gamma$	38.1
$D^0 \to D^0\pi^0$	61.9
$D^{*+} \to D^0\pi^+$	67.7

Table 16: Branching fractions from Reference [1] used to determine the sample composition.

modes. The events were generated using the software package `HeavyQuarkGen` [50]. This program uses theoretical calculations [51] of the b production differential cross section with respect to the b transverse momentum and pseudorapidity to generate single b quarks which are hadronized using the model described in Reference [52]. The B hadrons are then decayed using the `EvtGen` package [53] developed by the BaBar collaboration and are then passed through the full CDF detector and trigger simulation [54]. Each possible decay chain from Figure 48 was generated and simulated individually. Relevant D^{**} decay modes were then combined to match the 24 decay channels used in Table 15. The trigger and analysis reconstruction relative efficiencies, ϵ^i_{trig} and ϵ^i_{rec}, respectively, for the 24 decay channels are obtained from these Monte Carlo samples and shown in Table 17.

Using the information in Tables 15 and 17, we determine the fractions of B^0/B^+ in each of the three samples $D^+\ell$, $D^{*+}\ell$, $D^0\ell$:

$$f_{B^0}^{D^+\ell} = \sum_{i=0}^{5} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}, \qquad f_{B^+}^{D^+\ell} = \sum_{i=6}^{8} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}, \tag{64}$$

$$f_{B^0}^{D^0\ell} = \sum_{i=9}^{13} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}, \qquad f_{B^+}^{D^0\ell} = \sum_{i=14}^{20} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}, \tag{65}$$

$$f_{B^0}^{D^{*+}\ell} = \sum_{i=21}^{22} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}, \qquad f_{B^+}^{D^{*+}\ell} = \sum_{i=23}^{23} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}, \tag{66}$$

where $f^i_{SC} = f^i_{SC}(f^{**}, R_f, P_V, \epsilon^*)$ are given in Table 15.

The K-factor distributions are obtained in a similar way, by adding the distributions corresponding to individual decay modes weighted by the respective efficiencies:

$$\mathcal{F}_{B^0}^{D^+\ell}(K) = \sum_{i=0}^{5} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}\mathcal{F}_i(K), \qquad \mathcal{F}_{B^+}^{D^+\ell}(K) = \sum_{i=6}^{8} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}\mathcal{F}_i(K), \tag{67}$$

$$\mathcal{F}_{B^0}^{D^0\ell(K)} = \sum_{i=9}^{13} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}\mathcal{F}_i(K), \qquad \mathcal{F}_{B^+}^{D^0\ell}(K) = \sum_{i=14}^{20} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}\mathcal{F}_i(K), \tag{68}$$

$$\mathcal{F}_{B^0}^{D^{*+}\ell(K)} = \sum_{i=21}^{22} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}\mathcal{F}_i(K), \qquad \mathcal{F}_{B^+}^{D^{*+}\ell}(K) = \sum_{i=23}^{23} f^i_{SC}\epsilon^i_{trig}\epsilon^i_{rec}\mathcal{F}_i(K), \tag{69}$$

where the distributions $\mathcal{F}_i(K)$, $i = 0, 23$ correspond to the 24 decay modes listed in Table 15.

In appendix A.9, we show the K-factor distributions for all the 24 decay chains defined above. The histograms have 50 bins each from 0.2 to 1.2 and are used for numerical integration in the likelihood function (see Sec. 4.6). As seen in Table 17, the mean of each of these distributions is around 0.80 which means that the correction due to missing neutrino is about 20%. In the likelihood fit to the $D\ell$ final state distributions, the set of six K-factor distributions obtained after combining the initial 24 distribution with the appropriate sample composition weights are needed to describe the B^+ and B^0 contributions to the three final states D^+, D^0 and D^{*+}. These contributions depend on the four sample composition parameters described above.

Channel	ϵ_{trig}	ϵ_{rec}	$< K >$	RMS(K)
0	1	1	0.86	0.13
1	1.14	0.96	0.84	0.11
2	1.14	0.96	0.84	0.11
3	0.74	1.01	0.78	0.13
4	0.74	0.98	0.76	0.11
5	0.74	0.98	0.76	0.11
6	0.88	0.86	0.77	0.12
7	0.80	0.91	0.76	0.11
8	0.80	0.91	0.76	0.11
9	1.18	1.00	0.84	0.12
10	0.75	1.02	0.77	0.11
11	0.81	0.94	0.78	0.12
12	0.78	0.96	0.77	0.11
13	0.78	0.97	0.77	0.11
14	1	1	0.85	0.13
15	1.21	1.01	0.84	0.12
16	1.21	1.01	0.84	0.12
17	0.73	1.06	0.78	0.11
18	0.76	1.02	0.77	0.11
19	0.76	1.02	0.77	0.11
20	0.80	0.96	0.77	0.11
21	1	1	0.84	0.12
22	0.63	1.04	0.77	0.11
23	0.68	0.96	0.77	0.11

Table 17: Trigger (ϵ_{trig}) and reconstruction (ϵ_{rec}) efficiencies and the mean and RMS of the K-factor distributions for the various semileptonic decay channels contributing to the $D^+\ell$, $D^{*+}\ell$ and $D^0\ell$ final states as determined from Monte Carlo simulations. Given the large Monte Carlo simulation samples used to obtain these efficiencies, the statistical errors are negligible.

4.4 Decay Time of B Mesons

In this analysis, for each candidate in an event, a primary interaction vertex [47, 48] is reconstructed by considering all tracks in the event which are consistent with originating from the primary interaction point. Tracks that have large impact parameter or the position along the z-axis is not consistent with the primary vertex are removed. Also, tracks that form the B candidate are removed. All remaining tracks are fitted to a common primary vertex. The tracks used in this process are refit including Layer 00 information. The event-by-event primary vertex coordinates plus calculated vertex uncertainties are used together with the B vertex and the B vertex errors for the determination of decay lengths and decay length errors in a given event. Adding the Layer 00 information together with the event-by-event primary vertex determination improves the decay time resolution of B mesons in semileptonic decays by $\sim 14\%$. The main contribution to this improvement comes from using Layer 00 hit information.

The pseudo-proper decay length ct^* distributions for the right-sign lepton-D combinations are shown in Figure 49. Indications of prompt backgrounds around $ct^* \sim 0$ are seen in the $D^+\ell$ distribution. These prompt backgrounds are clearly visible in the corresponding ct^* distributions for the wrong-sign lepton-D combination as shown in Figure 50.

As we will see in Section 4.9, the prompt background observed in the wrong-sign $D\ell$ combinations is a source of large systematic uncertainties. Prompt background events may be produced by different mechanisms. A real D meson can be associated with either a fake lepton or a real lepton which is not a B daughter from semileptonic $B \to D\ell X$ decays. When the lepton is fake, the events would populate evenly the right and wrong-sign $D\ell$ samples. Another source of prompt background are $b\bar{b}$ events in which the D meson originates from the b quark and the lepton comes from the semileptonic decay of the $\bar{b}$ quark. Such events will have wrong $D\ell$ charge combination unless one of the B mesons oscillates or it comes from a sequential $b \to c \to \ell$ decay which flips the sign of the lepton charge with respect to the initial b quark. The third source of prompt background are $c\bar{c}$ events in which one c quark hadronizes into a D meson that we reconstruct and the other c quark decays semileptonically. These events will have the right $D\ell$ charge correlation. Using Monte Carlo studies [46] it was shown the $c\bar{c}$ fraction in the lepton-SVT dataset is very small.

In a B mixing measurement it is crucial that the decay time resolution is small compared to the oscillation period. In the case of B^0 mesons, the oscillation period is $\sim 2.0 \times 10^{-12}$ seconds. In Figure 51 we show distributions of the vertex error contribution to the ct^* error:

$$\sigma_{ct^*}(\text{vertex}) = \sigma_{L_{xy}} \frac{m(B)}{p_T(D\ell)}, \tag{70}$$

for the right-sign lepton-D combinations: $D^+\ell$, $D^{*+}\ell$ and $D^0\ell$ with again the D^* candidates excluded. The means of the side-band subtracted distributions are ~ 42.7 μm, ~ 50.5 μm and ~ 47.6 μm, respectively. These resolutions translate into typical pseudo proper decay time errors of less than $\sim 0.2 \times 10^{-12}$ seconds which are about an order of magnitude less than the B^0 oscillation period.

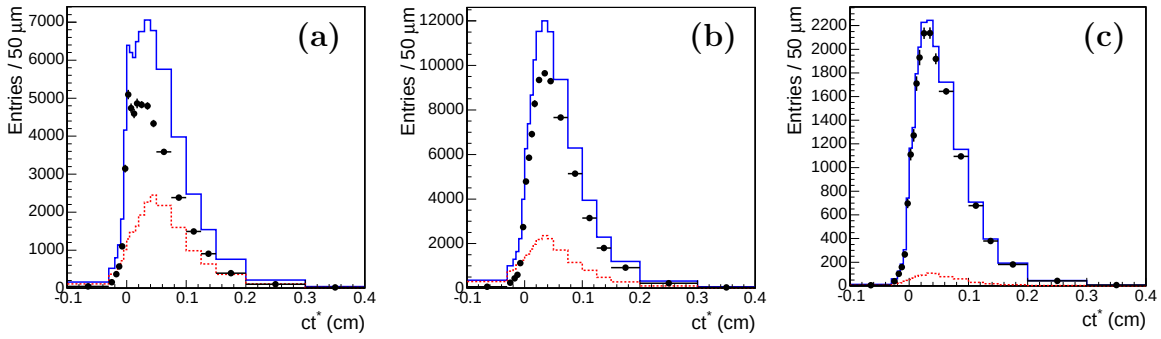

Figure 49: Distributions of the pseudo proper decay time ct^* for the right-sign lepton-D combinations: (a) $D^+\ell$, (b) $D^0\ell$ and (c) $D^{*+}\ell$. The lepton-D^* candidates are excluded from the $D^0\ell$ candidates. The solid (dashed) histograms are the distributions in the signal (side-band) region, while the points with error bars represent the side-band subtracted signal distribution.

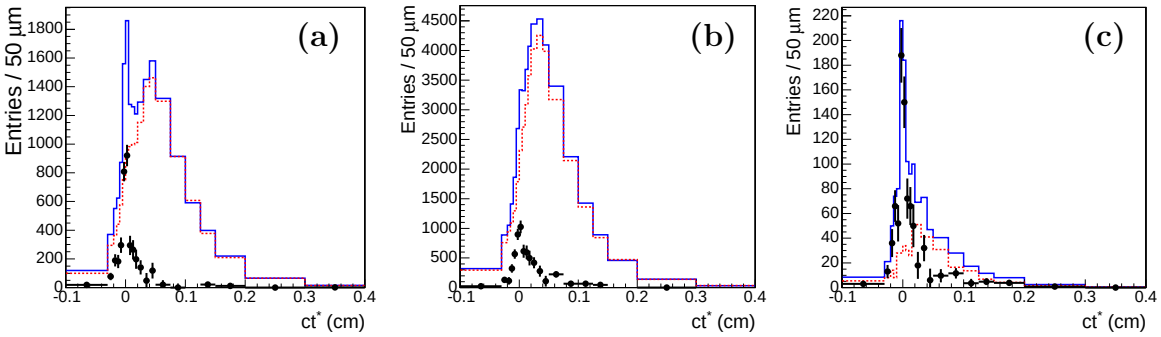

Figure 50: Distributions of ct^* for the wrong-sign lepton-D combinations: (a) $D^+\ell$, (b) $D^0\ell$ and (c) $D^{*+}\ell$.

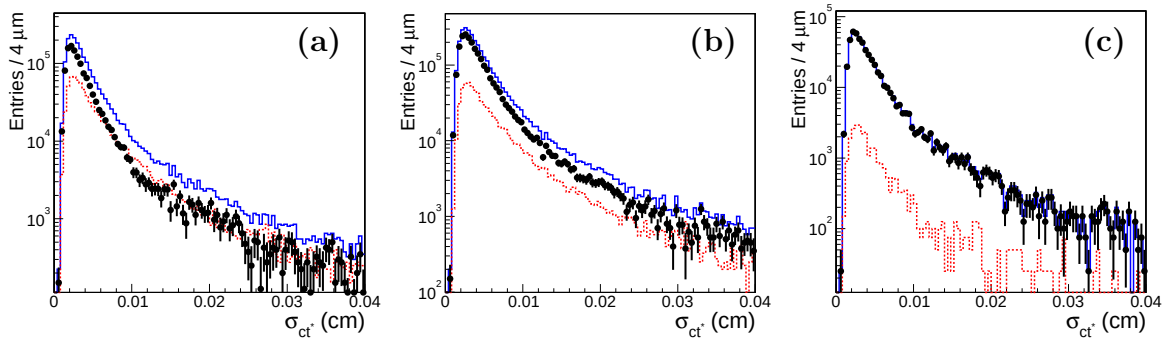

Figure 51: Distributions of estimated errors on the pseudo proper decay length σ_{ct^*} for the right-sign lepton-D combinations: (a) $D^+\ell$, (b) $D^0\ell$ and (c) $D^{*+}\ell$.

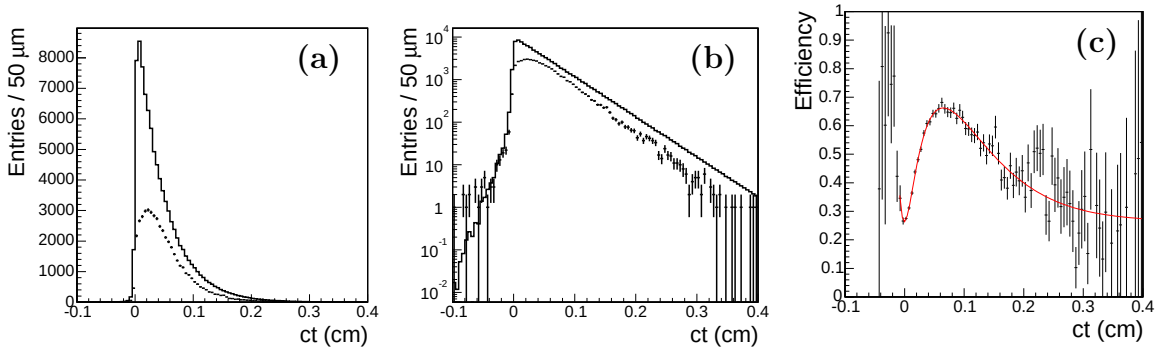

Figure 52: Proper time ct distributions before (solid) and after (error bars) the trigger and offline selection cuts for $B^0 \to D^-\ell^+$ (channel 0 in Table 15) on (a) linear and (b) logarithmic scale as obtained from the realistic Monte Carlo simulation. The efficiency $\mathcal{E}(ct)$ bias curve with fit parameterization superimposed is also shown (c).

4.5 Biases on the B Lifetime

Due to the SVT track requirement in the lepton-SVT trigger data and the analysis requirements summarized in Tables 9 through 11, the reconstructed proper decay time distribution t does not follow a pure exponential (modulo resolution effects) but is biased at both small and large lifetimes. This bias expressed as an efficiency curve $\mathcal{E}(t)$ is obtained using realistic Monte Carlo simulations. In general, we refer to this bias curve as the trigger efficiency or SVT bias. A pure exponential with the respective B^0/B^+ lifetime was convoluted with a detector resolution function for each decay mode to represent the reconstructed decay time distribution t before the trigger requirement. The detector resolution function was assumed to be Gaussian. The width of the Gaussian was set on an event-by-event basis as obtained from Equation (70). The decay time t distribution of these events defines the denominator for $\mathcal{E}(t)$. The numerator was obtained from the corresponding t distribution after SVT-trigger requirement and offline event selection as determined from the Monte Carlo simulation samples. Figure 52 shows, as an example, the ct distributions before and after the trigger and offline selection requirements for $B^0 \to D^-\ell^+$ (channel 0 in Table 15). The ratio of both distributions corresponds to the efficiency bias curve $\mathcal{E}(ct)$. These bias curves are parameterized with a common functional form:

$$\mathcal{E}(ct) = (p_0 + p_1\, ct + p_2\, ct^2) \times (f_1\, e^{-\lambda_1 ct} + f_2\, e^{-\lambda_2 ct}) + p_{\text{offset}}. \tag{71}$$

This functional form is used for analytical integration in the maximum likelihood fit. The parameters λ_1 and λ_2 are constrained for different decay chains to be the same for each final state (D^+, D^0, and D^*). This facilitates an easier combination of the bias curves in the likelihood function. Table 47 in Appendix A.10 summarizes the values of these parameters for each of the 24 decay channels in Table 15.

The six sets of parameters describing the trigger efficiencies for the B^+ and B^0 subsamples in each of the $D^+\ell$, $D^0\ell$ and $D^{*+}\ell$ modes are obtained in the same way as the K-factor distributions (see Equation (69)) by replacing $\mathcal{F}(K)$ with each of the

parameters p_0, p_1, p_2, f_1, f_2 and p_{offset} from Equation (71).

4.6 Maximum Likelihood Fit

To measure the $B^0\bar{B}^0$ oscillation frequency Δm_d from the data, an unbinned maximum likelihood method is used. Neglecting effects from CP violation as well as a possible lifetime difference between the heavy and light mass eigenstates of the B^0 meson, the probability $\mathcal{P}$ for a B^0 meson produced at time $t_0 = 0$ to decay as B^0 at proper time $t > 0$ is given by:

$$\mathcal{P}^{\text{unmix}}_{B^0}(t) = \mathcal{P}_{B^0 \to B^0}(t) \;=\; \frac{1}{2}\,\Gamma\, e^{-\Gamma t}\,(1 + \cos(\Delta m_d t)). \tag{72}$$

Similarly, the probability for the B^0 to decay as $\bar{B}^0$ is given as:

$$\mathcal{P}^{\text{mix}}_{B^0}(t) = \mathcal{P}_{B^0 \to \bar{B}^0}(t) \;=\; \frac{1}{2}\,\Gamma\, e^{-\Gamma t}(1 - \cos(\Delta m_d t)). \tag{73}$$

In practice we need to determine the flavor of the B^0 mesons at decay time and at production time. In semileptonic decays $B \to D\ell X$, the charge of the lepton gives the flavor of the B meson at decay. The flavor at production is obtained using opposite side flavor taggers. For simplicity of the following example, let us assume that we only use one opposite side lepton tagger. The lepton from the semileptonic decay of the opposite side B meson is referred to as the "tag lepton". The correlation between the charge of the trigger lepton and the charge of the tag lepton provides information about the mixing status in a given event.

In the ideal case in which we have a pure sample of B^0 and $\bar{B}^0$ mesons and the flavor tag is always correct, an event with an opposite-sign lepton pair signals an unmixed event, while a same-sign lepton pair indicates a mixed event. In this case the probabilities for an opposite-sign event $\mathcal{P}_{OS}$ and a same-sign event $\mathcal{P}_{SS}$ are directly related to the mixing probabilities:

$$\mathcal{P}_{\text{SS}}(t) = \mathcal{P}_{\text{mix}}(t), \qquad\qquad \mathcal{P}_{\text{OS}}(t) = \mathcal{P}_{\text{unmix}}(t). \tag{74}$$

If the lepton tag in our example does not always tag correctly but with a mistag probability p_W, we measure the following same-sign and opposite-sign fractions:

$$\mathcal{P}_{\text{SS}}(t) = (1 - p_W)\,\mathcal{P}_{\text{mix}}(t) + p_W\,\mathcal{P}_{\text{unmix}}(t), \tag{75}$$

$$\mathcal{P}_{\text{OS}}(t) = p_W\,\mathcal{P}_{\text{mix}}(t) + (1 - p_W)\,\mathcal{P}_{\text{unmix}}(t). \tag{76}$$

After substitution of $\mathcal{P}_{\text{mix}}$ and $\mathcal{P}_{\text{unmix}}$ from Equations (72) and (73) in Equation (75) and (76), we obtain:

$$\mathcal{P}_{\text{SS}}(t) = \frac{1}{2}\,\Gamma\, e^{-\Gamma t}\,(1 - \mathcal{D}\cos(\Delta m_d t)), \tag{77}$$

$$\mathcal{P}_{\text{OS}}(t) = \frac{1}{2}\,\Gamma\, e^{-\Gamma t}\,(1 + \mathcal{D}\cos(\Delta m_d t)), \tag{78}$$

where $\mathcal{D} = 1 - 2p_W$ is the tagging dilution which is determined for each event by one of the opposite side flavor taggers described in Section 4.2. Since the taggers were calibrated on the inclusive lepton-SVT sample and Δm_d is measured on a subsample we allow for possible differences in predicted dilutions by introducing a dilution scale factor $\mathcal{S}_\mathcal{D}$ for each of the five taggers (opposite side muon and electron, jet charge with displaced vertex, displaced track and high momentum). We can then float the dilution scale factors in the likelihood fit to re-calibrate the predicted tagging dilution. If tagging calibrations from the inclusive lepton-SVT sample were perfect and if the sample on which we measure Δm_d and the lepton-SVT sample are very similar, then we expect to obtain dilution scale factors close to one. We also introduce the amplitude $\mathcal{A}$, which is a Fourier coefficient multiplying the oscillatory term as described in Section 1.2.4:

$$(1 \pm \mathcal{D} \cos \Delta m_d \, t) \quad \rightarrow \quad (1 \pm \mathcal{A} \; \mathcal{S}_\mathcal{D} \; \mathcal{D} \cos \Delta m_d \, t) \tag{79}$$

By fixing the oscillation frequency Δm_d and floating the amplitude in the likelihood fit we expect to obtain $\mathcal{A} \sim 1$ when Δm_d is close to its true value and $A \sim 0$ when Δm_d is far from the true value. This method is described in detail in Reference [19] and is particularly suitable for $B_s^0 \bar{B}_s^0$ mixing searches. As we will use this method for our Δm_s analysis in Section 5, we first test the amplitude technique on $B^0 \bar{B}^0$ oscillations.

Including a Gaussian vertex resolution $\mathcal{G}(t' - t, s_\sigma \sigma)$ and an efficiency function $\mathcal{E}(t)$ (Section 4.5), the same-sign probability from Equation (77) is modified to be:

$$\mathcal{P}'_{\rm SS}(t') = \mathcal{N} \int_0^\infty \frac{1}{2} \Gamma \, \mathrm{e}^{-\Gamma t} \, (1 - \mathcal{A} \, \mathcal{S}_\mathcal{D} \, \mathcal{D} \cos \Delta m_d \, t) \mathcal{G}(t' - t, s_\sigma \sigma) dt \times \mathcal{E}(t'). \tag{80}$$

Here, σ is the uncertainty on the decay time t given by:

$$\sigma = \sigma_{L_{xy}} \frac{m_B}{p_T(B)} \tag{81}$$

and s_σ is a resolution scale factor which accounts for underestimation of the decay vertex errors calculated by the vertex fitter. This resolution scale factor is determined from the reconstruction of J/ψ states. A subsample of J/ψ consistent with originating from the primary vertex mesons is used to determine the resolution scale factor and a value of 1.35 is obtained [55]. Instead of using an average dilution for each tagger, the dilution $\mathcal{D}$ used in the probability density function of the likelihood fit is an event-by-event predicted dilution as described in Sec. 4.2. The normalization $\mathcal{N}$ in Equation (80) is determined from:

$$1/\mathcal{N} = \int_{-\infty}^{\infty} dt' \; \int_0^\infty dt \; \frac{1}{2} \; \Gamma \; \mathrm{e}^{-\Gamma t} \; \mathcal{G}(t' - t, s_\sigma \sigma) \times \mathcal{E}(t'). \tag{82}$$

Note that we need to normalize the sum $\mathcal{P}_{\rm SS} + \mathcal{P}_{\rm OS}$ and not each term $\mathcal{P}_{\rm SS}$ and $\mathcal{P}_{\rm OS}$ individually.

After including the momentum uncertainty expressed through a normalized K-factor distribution $\mathcal{F}(K)$, as given in Equation (30) , the same-sign probability is

modified as:

$$\mathcal{P}_{\rm SS}(t^*) = \int_{K_{min}}^{K^{max}} dK\ \mathcal{P}'_{\rm SS}(Kt^*)K\,\mathcal{F}(K). \tag{83}$$

A formula similar to Equation (83) is obtained for opposite-sign probability by changing the sign of the oscillatory term from "−" to "+" in Equation (80). The calculation of the integrals in Equation (83) is described in Appendix A.11.

For the evaluation of the likelihood function, four input quantities are used on an event-by-event basis: the pseudo proper decay time t^* of the B meson, its error σ_{t^*}, the D meson mass m and the dilution $\mathcal{D}$. Here, we refer to the dilution as the dilution predicted from the tagger $\mathcal{D}_{pred}$ times the tag decision s_{tag}, which is $+1$ if the tag is same-sign or -1 if the tag is opposite-sign with respect to the trigger lepton, $\mathcal{D} = \mathcal{D}_{pred} \times s_{tag}$. With this notation, the likelihood function for signal events is:

$$\begin{aligned} L^{t^*}_{sig}(t^*, \sigma_{t^*}, \mathcal{D}) &= L_{\rm SS}(t^*, \sigma_{t^*}, \mathcal{D}) \ \text{ if } \ s_{tag} = +1 \\ L^{t^*}_{sig}(t^*, \sigma_{t^*}, \mathcal{D}) &= L_{\rm OS}(t^*, \sigma_{t^*}, \mathcal{D}) \ \text{ if } \ s_{tag} = -1 \end{aligned} \tag{84}$$

The complete likelihood function $\mathcal{L}$ is described by the following equation:

$$\begin{aligned} \mathcal{L} = f_{sig}(1-f_{pmt})L^{t^*}_{sig}(t^*, \sigma_{t^*}, \mathcal{D})L^{m}_{sig}(m)L^{\sigma_{t^*}}_{sig}(\sigma_{t^*})L^{\mathcal{D}}_{sig}(\mathcal{D}_{pred}) \\ + f_{sig}f_{pmt}L^{t^*}_{pmt}(t^*, \mathcal{D}_{pmt})L^{m}_{sig}(m)L^{\sigma_{t^*}}_{sig}(\sigma_{t^*})L^{\mathcal{D}}_{sig}(\mathcal{D}_{pred}) \\ + (1-f_{sig})L^{t^*}_{bkg}(t^*, \mathcal{D}_{bkg})L^{m}_{bkg}(m)L^{\sigma_{t^*}}_{bkg}(\sigma_{t^*})L^{\mathcal{D}}_{bkg}(\mathcal{D}_{pred}), \end{aligned} \tag{85}$$

where f_{sig} is the fraction of real D meson signal containing real semileptonic $B \to D\ell\nu$ events as well as some prompt background contributing to the $D\ell$ final state. These signal fractions are calculated individually for each of the three decay signatures and separately for each tagged sample corresponding to the five taggers. The fraction of prompt background f_{pmt} is estimated using the wrong-sign lepton-D sample. This fraction is determined as the ratio between the yield of wrong-sign lepton-D candidates to the yield of right-sign lepton-D candidates and are summarized in Table 14. For the case of $D^0\ell$ and $D^{*+}\ell$ we directly use the wrong-sign yield assuming $N_{WS} \approx N_{RS}$ while for $D^+\ell$ the number of wrong-sign events is scaled by a factor of 1.6 to achieve the same level of combinatorial background in both side-band distributions.

$L^m_{sig}(m)$ and $L^m_{bkg}(m)$ represent the probability distribution functions for the signal and combinatorial background D meson mass distribution expressed by a double Gaussian for the signal and by a linear function for background. $L^{\mathcal{D}}_{sig/bkg}$ and $L^{\sigma_{t^*}}_{sig/bkg}$ describe the predicted dilution and proper time resolution distributions for both signal and background events. The dilution and the decay time error enter the likelihood function on an event-by-event basis. The distributions of these quantities for signal and background events are different and this effect must be accounted for in the likelihood function. It has been shown [56] that without using separate distributions for signal and background the likelihood fit produces biased results. The term $L^{t^*}_{bkg}(t^*, \mathcal{D}_{bkg})$ corresponds to the probability distribution function of the combinatorial background on t^*. This dependence is obtained from the D meson mass side-bands

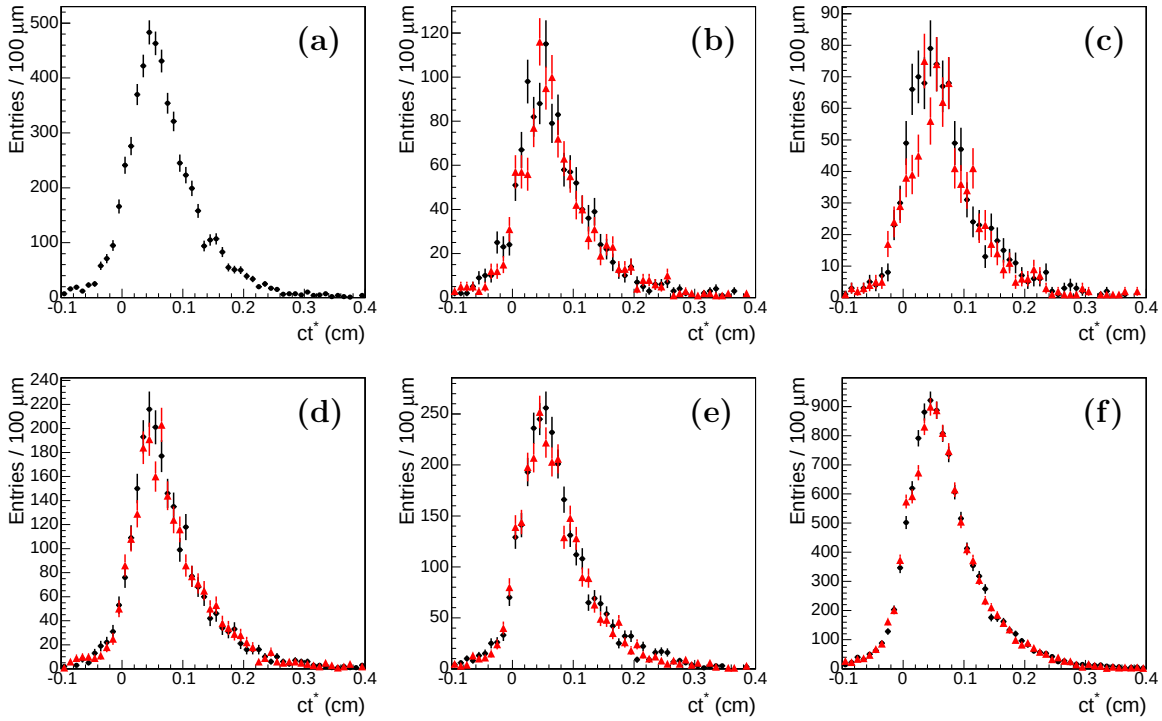

Figure 53: The distributions of the pseudo proper decay time ct^* for combinatorial background events from D meson side-bands for (a) no tag, (b) SMT, (c) SET, jet charge (d) displaced vertex, (e) displaced track and (f) high p_T tagged $D^+\ell$ events are shown. Triangles (circles) correspond to same (opposite) sign events.

and is parameterized as:

$$L^{t^*}_{bkg}(t^*, \mathcal{D}_{bkg}) = \frac{1-\mathcal{D}_{bkg}}{2} N \left[f \int_0^\infty \frac{1}{\tau} e^{-t/\tau} \mathcal{G}(t-t^*, \sigma) dt + (1-f)\mathcal{G}(t^*, \sigma) \right] \quad (86)$$

where $\mathcal{D}_{bkg}$ is the asymmetry of the combinatorial background multiplied by the tag decision s_{tag}:

$$\mathcal{D}_{bkg} = \frac{N_{OS} - N_{SS}}{N_{OS} + N_{SS}} \times s_{tag}. \quad (87)$$

With this definition we observe that for $s_{tag} = -1$ $(+1)$, $\frac{1-\mathcal{D}_{bkg}}{2}$ represents the fraction of opposite (same) sign events in the combinatorial background sample. Figure 53 shows the ct^* distributions of $D^+\ell$ events for each tagger, separated in same-sign and opposite-sign. Although same-sign and opposite-sign distributions are very similar, we account for small differences by fitting them separately with the functional form given in Equation (86). The resulting fitting functions are used in the likelihood fit as templates for the ct^* distributions of combinatorial background events. We apply the same procedure to $D^0\ell$ and $D^{*+}\ell$ events.

The term $L^{t^*}_{pmt}(t^*, \mathcal{D}_{pmt})$ represents the corresponding probability distribution function of t^* for the prompt background events and is parameterized with the same

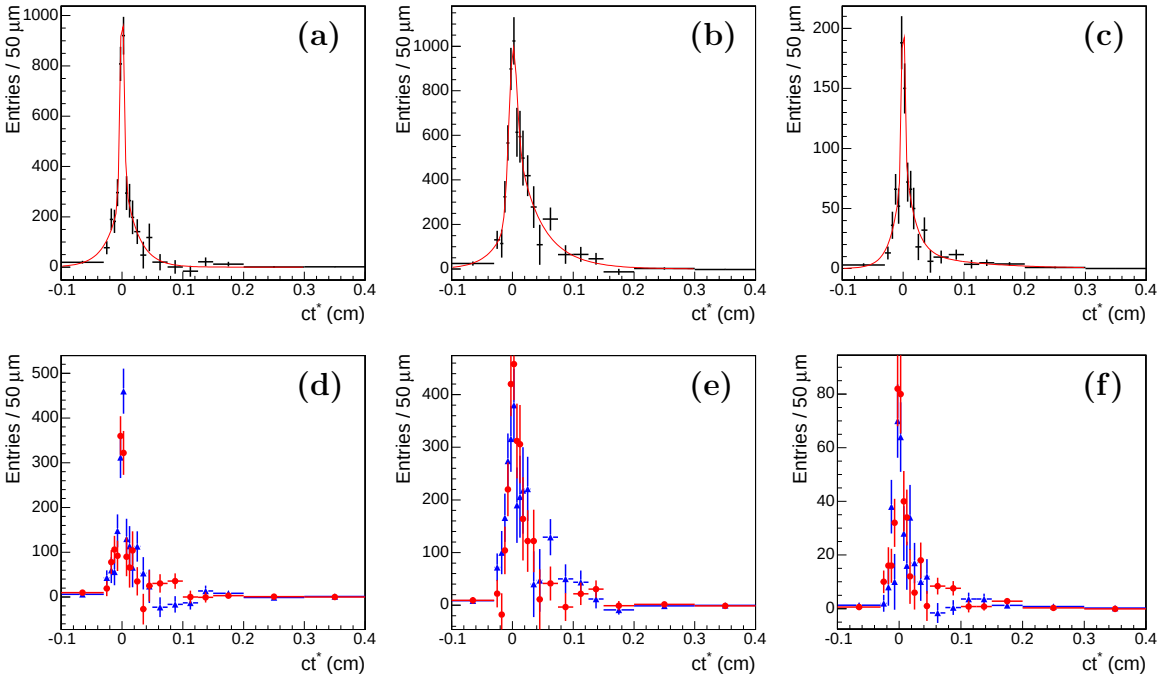

Figure 54: Distributions of the pseudo proper decay time $ct^*(B)$ for wrong sign (a) $D^+\ell^+$, (b) $D^0\ell^+$ and (c) $D^{*+}\ell^+$ combinations after side-band subtraction with fit results superimposed. The same plots are split in same-sign (circles) and opposite-sign (triangles) between the trigger lepton charge and the tag sign: (d) $D^+\ell^+$, (e) $D^0\ell^+$ and (f) $D^{*+}\ell^+$.

functional form as the combinatorial background:

$$L^{t^*}_{pmt}(t^*, \mathcal{D}_{pmt}) = \frac{1-\mathcal{D}_{pmt}}{2} N\,[\, f \int_0^\infty \frac{1}{\tau} e^{-t/\tau} \mathcal{G}(t-t^*, \sigma) dt + (1-f)\mathcal{G}(t^*, \sigma)\,]. \quad (88)$$

Similar to the case of the combinatorial background, $\mathcal{D}_{pmt}$ represents the asymmetry of the prompt background and for $s_{tag} = -1$ $(+1)$, $\frac{1-D_{pmt}}{2}$ represents the fraction of opposite (same) sign events in the prompt background samples. For a better understanding of the prompt background, we study the side-band subtracted ct^* distributions from the wrong-sign $D\ell$ samples. Figure 54 shows these distributions plus the result of their empirical parameterizations. These parameterizations are used in the likelihood fit as templates for the prompt background distributions of ct^*. Figure 54 also shows the same ct^* distribution for the wrong-sign lepton-D combinations now divided into events with opposite-sign and same-sign flavor tags. Although these distributions are statistically limited, we do not observe a significant asymmetry. Thus for the default analysis, we assume the asymmetry of the prompt background $\mathcal{D}_{pmt} = 0$. The mass, the pseudo proper decay time uncertainty and predicted dilution terms for the prompt background are taken to be the same as the signal term distributions in the likelihood. Since our knowledge of the prompt background is limited by the statistics of the $D\ell$ wrong-sign candidates, the assumptions above are valid with limited certainty and contribute to the systematic error.

Finally, the probability distribution functions for the proper time t^* of each of the three $B \to D\ell\nu$ decay signatures, $L_{sig}^{t^*}$, is given by:

$$L_{sig}^{t^*} = f_0 L_0 + (1 - f_0) L_+ \tag{89}$$

where f_0 is the fraction of B^0 mesons and $1 - f_0$ is the fraction of B^+ mesons in the respective data sample. L_0 and L_+ are the likelihood terms for neutral and charged B mesons given, according to Equation (83), as:

$$L_0(t^*, \sigma_{t^*}, \mathcal{D}) =$$

$$= \mathcal{N} \int dK K\, \mathcal{E}(Kt^*) \int_0^\infty \frac{\Gamma_0}{2} \mathrm{e}^{-\Gamma_0 t} (1 \pm \mathcal{A}\ \mathcal{S}_\mathcal{D}\ \mathcal{D} \cos(\Delta m_d\, t)) \mathcal{G}(t' - t, s_\sigma \sigma) dt|_{t'=Kt^*} \tag{90}$$

and

$$L_+(t^*, \sigma_{t^*}, \mathcal{D}) = \mathcal{N} \int dK K\, \mathcal{E}(Kt^*) \int_0^\infty \frac{\Gamma_+}{2} \mathrm{e}^{-\Gamma_+ t} (1 \pm \mathcal{S}_\mathcal{D}\ \mathcal{D}) \mathcal{G}(t' - t, s_\sigma \sigma) dt|_{t'=Kt^*}. \tag{91}$$

The total likelihood function is a product over all decay signatures ($D^+\ell$, $D^0\ell$ and $D^{*+}\ell$) and all events in each decay signature:

$$\mathcal{L}_{TOT} = \prod_i \prod_j \mathcal{L}_{ij} \tag{92}$$

where i runs over the $D^+\ell$, $D^0\ell$ and $D^{*+}\ell$ signatures and j runs over all events in each decay sample with $\mathcal{L}_{ij}$ given by Equation (89). The function to be minimized is defined as:

$$-2 \ln \mathcal{L}_{TOT}. \tag{93}$$

With this convention, the deviation in each parameter that produces a variation of the likelihood function from the minimum of one unit, is defined as the 1σ Gaussian error on that parameter. The fitting framework is structured such that we can float the four sample composition parameters f^{**}, R_f, P_V and ϵ^*. The fraction f^{**} and the ratio R_f are obtained from Reference [46]:

$$f^{**} = 0.31 \pm 0.05 \tag{94}$$

and

$$R_f = 2.18 \pm 0.14. \tag{95}$$

We use these values together with their errors to constrain f^{**} and R_f with χ^2-terms added to the total likelihood function. We obtain further constraints on the sample composition parameters from the ratios of D^* and D^+ states to the total number of D^0 states (including D^*):

$$R^* = \frac{N(D^*)}{N(D^0) + N(D^*)} \tag{96}$$

$$R^+ = \frac{N(D^+)}{N(D^0) + N(D^*)} \tag{97}$$

where $N(D^*)$, $N(D^0)$ and $N(D^+)$ are the numbers of D^*, D^0 and D^+ states after subtracting the fractions of prompt background: $N(D^*) = N_m(D^*)(1 - f_{pmt}(D^*))$, $N(D^0) = N_m(D^0)(1 - f_{pmt}(D^0))$ and $N(D^+) = N_m(D^+)(1 - f_{pmt}(D^+))$ with $N_m(D^*)$, $N_m(D^0)$ and $N_m(D^+)$ representing the measured yields given in Table 13. $N_m(D^0)$ refers to the number of D^0 states from which the D^{*+} states are removed. The fractions of prompt background are given in Table 19. Finally, we obtain:

$$R^* = 0.18 \pm 0.01 \tag{98}$$

$$R^+ = 0.41 \pm 0.05 \tag{99}$$

At the same time, the above ratios can be written in terms of the sample composition parameters:

$$R^* = \frac{f_{B^+}^{D^{*+}\ell} + f_{B^0}^{D^{*+}\ell}}{f_{B^+}^{D^{*+}\ell} + f_{B^0}^{D^{*+}\ell} + f_{B^0}^{D^0\ell} + f_{B^+}^{D^0\ell}} \tag{100}$$

$$R^+ = \frac{f_{B^0}^{D^+\ell} + f_{B^+}^{D^+\ell}}{f_{B^+}^{D^{*+}\ell} + f_{B^0}^{D^{*+}\ell} + f_{B^0}^{D^0\ell} + f_{B^+}^{D^0\ell}} \tag{101}$$

where $f_{B^+/B^0}^{D^{*+}\ell/D^0\ell/D^+\ell}$ depend on the fractions $f_{SC}^i(f^{**}, R_f, P_V, \epsilon^*)$ as shown in Equation (66). In turn, the dependence of these fractions on the four sample composition parameters $f_{SC}^i(f^{**}, R_f, P_V, \epsilon^*)$ is given in Table 15.

We add four χ^2 terms to the likelihood function corresponding to f^{**}, R_f, R^* and R^+, where the last two ratios depend on all four sample composition parameters:

$$-2 \ln \mathcal{L}_{TOT} \;\rightarrow\; -2 \ln \mathcal{L}_{TOT} \;+\; \chi^2(f^{**}) \;+\; \chi^2(R_f) \;+\; \chi^2(R^*) \;+\; \chi^2(R^+) \tag{102}$$

The χ^2-terms are defined as $\chi^2(x) = (\frac{x - x_0}{\sigma_x})^2$, where x_0 is the central value and σ_x the constraining error on x. By floating the sample composition parameters, the fractions of B^0/B^+ as well as the shapes of the K-factor and trigger efficiency distributions will change accordingly as indicated in Equation (69).

4.7 Toy Monte Carlo Tests of the Fitting Framework

In order to test the integrity of the likelihood fitting framework, we perform toy Monte Carlo experiments (pseudo-experiment). In such experiments, a data sample is simulated first. Events are generated according to the 24 decay channels in Table 15 with the corresponding fractions given by the sample composition parameters and relative selection efficiencies. For each event, a random decay time t is generated according to an exponential decay with either B^0 or B^+ lifetime. Detector resolution effects are accounted for by adding a Gaussian distributed error to the generated lifetime: $t' = t + \delta t$. The pseudo proper decay time is then obtained as: $t^* = t'/K$ where the K-factor accounting for the missing neutrino is randomly generated according to the K-factor distributions obtained from realistic simulation of B decays followed by detector and trigger simulation. The event is randomly identified as mixed

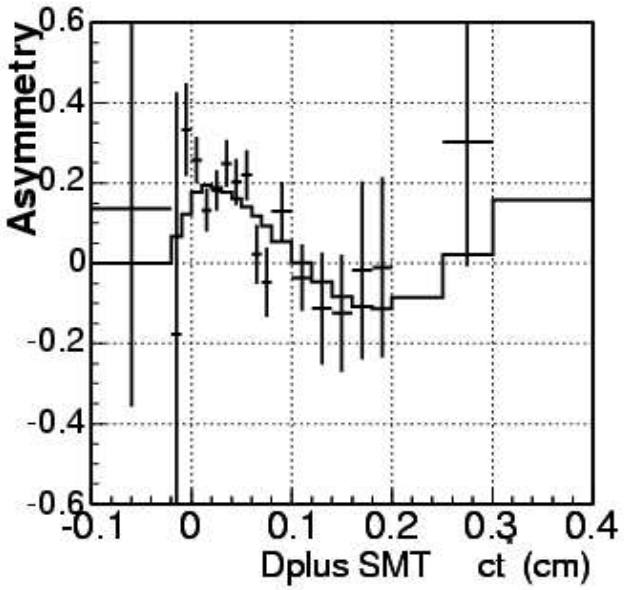

Figure 55: Likelihood projection of the oscillation asymmetry for a toy Monte Carlo experiment with lepton-D^+ events with SMT tags only.

or unmixed according to the oscillation frequency Δm_d. The decision of the tagging algorithms is also generated according to the tagging efficiency (only a fraction of events are tagged) and dilution (events are tagged correctly with probability $\mathcal{P} = \frac{1+\mathcal{D}}{2}$). The output of the generated sample is a set of events, where each event contains values for the pseudo-proper decay time t^*, the error σ_{t^*}, the tagging decision (mixed or unmixed event) along with the predicted dilution $\mathcal{D}_{pred}$ and the mass of the D meson m_D. Once the toy Monte Carlo data-sample is generated, the pseudo-experiment is performed by fitting the simulated sample. The input values for the oscillation frequency Δm_d and the lifetimes of B^0 and B^+ mesons used to generate the sample are expected to be returned by the fit within statistical errors. The asymmetry distribution with the fit projection of one particular toy Monte Carlo experiment is shown in Figure 55 for $D^+\ell$ events with opposite side muon tags and the same number of tagged events and the same dilution as in real data.

To further evaluate the fitter integrity, 500 toy Monte Carlo experiments are run and the pull distribution $(x_{fit} - x_{input})/\sigma(x_{fit})$ for various fit parameters are examined. Figure 56 shows the pull distributions for the lifetime $c\tau_{B^0}$, the lifetime ratio τ_{B^+}/τ_{B^0}, the oscillation frequency Δm_d and the amplitude $\mathcal{A}$. In each pseudo-experiment we generate a total of $\sim 500,000$ events, approximately about the number of $D\ell$ events in real data. Table 18 summarizes the results of the 500 pseudo-experiments. The fit input values compared to the average result of the lifetime $c\tau_{B^0}$, lifetime ratio τ_{B^+}/τ_{B^0}, oscillation frequency Δm_d and amplitude $\mathcal{A}$ are shown. The mean of the pull distributions for $c\tau_{B^0}$ and amplitude $\mathcal{A}$ agree with zero as expected, while for τ_{B^+}/τ_{B^0} and Δm_d we observe a possible small bias of 0.3 and 0.4 σ, respectively. These small effects will be treated as systematic uncertainties in Section 4.9. The

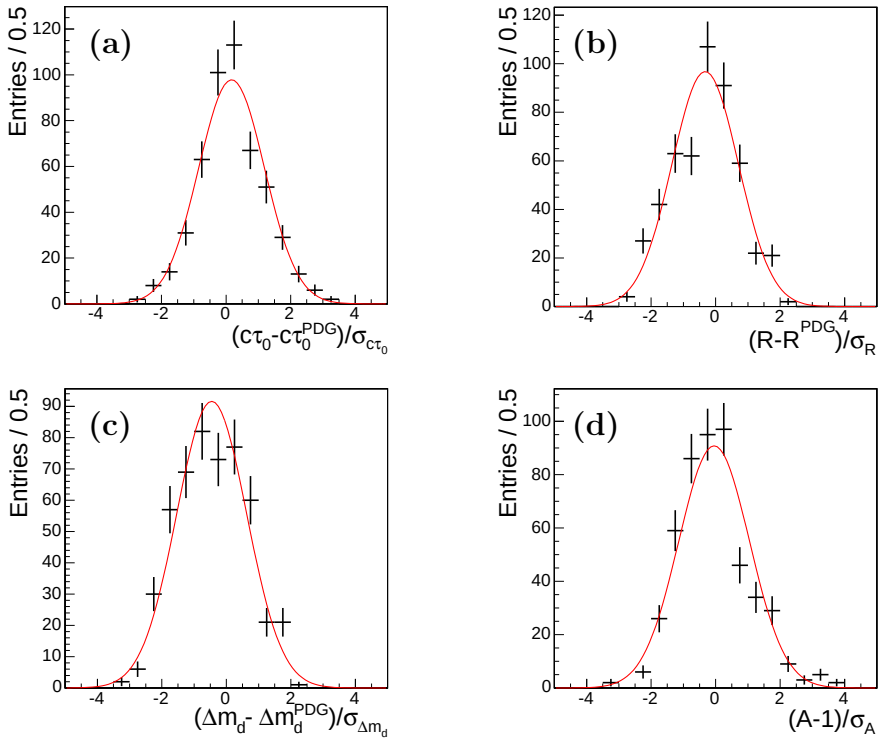

Figure 56: Pull distributions for (a) $c\tau_{B^0}$, (b) $R = \tau_{B^+}/\tau_{B^0}$ (c) Δm_d and (d) the amplitude $\mathcal{A}$ for 500 toy Monte Carlo experiments.

	Input	Average	Average Error	Pull	σ(Pull)
$c\tau_{B^0}$	460.8 μm	461.2 μm	2.5 μm	0.17 ± 0.05	1.02 ± 0.03
τ_{B^+}/τ_{B^0}	1.086	1.083	0.011	-0.32 ± 0.05	1.03 ± 0.03
Δm_d	0.502 ps^{-1}	0.492 ps^{-1}	0.021 ps^{-1}	-0.45 ± 0.05	1.08 ± 0.04
Amplitude $\mathcal{A}$	1.0	0.998	0.023	-0.05 ± 0.05	1.10 ± 0.04

Table 18: Summary of 500 toy Monte Carlo experiments. Input values, average fit results, average fit errors as well as means ans widths of the pull distributions are shown for $c\tau_{B^0}$, τ_{B^+}/τ_{B^0}, Δm_d and amplitude $\mathcal{A}$.

widths of all pull distributions agree with unity as expected.

4.8 Fit Results from Data

We use the likelihood fit framework described in Sec. 4.6 to measure various physics quantities such as the B^0/B^+ lifetimes, the B^0 oscillation frequency Δm_d and dilution scale factors from the data. All events with 1.80 GeV/c < m(D) < 1.93 GeV/c enter the sum in the total likelihood function. Since the full fit runs over about 0.5 million events (signal and side-band events for right-sign and wrong-sign $D\ell$ events) and could potentially include hundreds of fit parameters (signal and background fractions, D mass parameterizations, ct^* parameterizations for signal, side-bands, right-sign and

Parameter / Tag	No Tag	SMT	SET	JQT1	JQT2	JQT3
$B \to D^+\ell^-$						
Signal Fraction $f_{sig}(\%)$	39.9	35.7	32.9	37.4	38.4	37.7
Comb. Bkg. Asymmetry $\mathcal{D}_{bkg}(\%)$	–	2.7	5.3	1.2	1.6	0.4
Prompt Bkg. Fraction $f_{pmt}(\%)$	4.8	4.8	4.8	4.8	4.8	4.8
$B \to D^0\ell^-$						
Signal Fraction $f_{sig}(\%)$	58.7	56.8	51.8	58.3	59.8	56.6
Comb. Bkg. Asymmetry $\mathcal{D}_{bkg}(\%)$	–	11.0	9.5	3.7	2.3	1.9
Prompt Bkg. Fraction $f_{pmt}(\%)$	4.3	4.3	4.3	4.3	4.3	4.3
$B \to D^{*+}\ell^-$						
Signal Fraction $f_{sig}(\%)$	89.9	87.9	88.4	89.4	89.8	89.9
Comb. Bkg. Asymmetry $\mathcal{D}_{bkg}(\%)$	–	10.8	19.1	−5.2	0.0	−2.9
Prompt Bkg. Fraction $f_{pmt}(\%)$	2.5	2.5	2.5	2.5	2.5	2.5

Table 19: Fractions of signal, asymmetries of the combinatorial background and fractions of prompt background in each decay signature and each tagged samples. JQT1, JQT2 and JQT3 refer to the jet charge tagging algorithm with displaced vertex, displaced track and high p_T jets, respectively.

wrong-sign $D\ell$, dilutions, sample composition parameters, lifetimes, oscillation frequency, etc), it is technically challenging to run the maximization fitting package `MINUIT` [57] with so many events and free fit parameters. However, most of the fit parameters can be fixed by distributions from the side-bands and wrong-sign events. We therefore perform a two step fit procedure. We first determine the fit parameters such as lifetime shapes of the combinatorial and prompt background from the side-band and wrong-sign events (Figures 53 and 54) and fix them when fitting for the physics quantities of interest. We also determine the signal fractions, the asymmetries of the combinatorial background and the fractions of prompt background in each decay signature for each of the tagged samples. Table 19 summarizes these parameters.

The signal fraction f_{sig} in each subsample is determined simply as the number of signal events divided by the total number of events in that sample. The asymmetry of the combinatorial background is calculated as in Equation (87). The fraction of prompt background is calculated as the ratio between the number of signal D events in the wrong-sign $D\ell$ sample and the number of D events in the right-sign $D\ell$ sample. It is assumed to be independent of the tag type.

The next step is to determine the sample composition of each of the three $D\ell$ decay signatures. We fix the sample composition parameters R_f and f^{**} to the values given in Equations (94) and (95). To measure the two remaining sample composition parameters, we fix every parameter in the likelihood fit except for P_V and ϵ^*. The B lifetimes and oscillation frequency are fixed to the world average value [1]. The resulting sample composition parameters are shown in Table 20. These values are then fixed when fitting for lifetimes, Δm_d or dilution scale factors.

We run two types of fit configurations. In the "Amplitude Fit" configuration, we

Parameter	Fit result	PDG
$c\tau_{B^0}$	445.5 ± 2.9 μm	460.8 ± 4.2 μm
τ_{B^+}/τ_{B^0}	1.096 ± 0.013	1.086 ± 0.017
Δm_d	0.501 ± 0.029 ps^{-1}	0.502 ± 0.007 ps^{-1}
$\mathcal{A}$	0.958 ± 0.028	–
R_f	2.18 (fixed)	2.18 ± 0.14
f^{**}	0.31 (fixed)	0.31 ± 0.05
P_V	0.85 (fixed)	–
ϵ^*	0.54 (fixed)	–

Table 20: Fit result of Amplitude Fit configuration on real data compared to the world average [1].

	$c\tau_{B^0}$	τ_{B^+}/τ_{B^0}	Δm_d	A
$c\tau_{B^0}$	1.0	-0.85	-0.46	-0.01
τ_{B^+}/τ_{B^0}	-0.85	1.0	0.05	0.02
Δm_d	-0.46	0.05	1.0	0.46
$\mathcal{A}$	-0.01	0.02	0.46	1.0

Table 21: Fit parameter correlations for the Amplitude Fit to the data.

fit for the B^0 lifetime $c\tau_{B^0}$, the ratio between the B^0 and B^+ lifetimes τ_{B^+}/τ_{B^0}, the oscillation frequency Δm_d and the amplitude $\mathcal{A}$ with all other parameters including the sample composition fixed, using the dilution predictions from the flavor taggers. The value of the amplitude returned by the fit is expected to be consistent with one. Since one of the objectives of this analysis is to cross check the dilution predictions of the flavor tags, we also run the fit in a "Dilution Fit" configuration, where we fit for five global dilution scale factors $\mathcal{S}_\mathcal{D}$ for the five flavor tags: SMT, SET, JQT (displaced vertex, displaced track and high p_T).

The result of the Amplitude Fit to real data is shown in Table 20. The sizes of the statistical uncertainties are consistent with the expectations from the Toy Monte Carlo. The B^0 lifetime is short by $\sim 3\%$ compared to the world average [1]. This is a long standing issue at CDFII which has not yet been solved. We believe that the short lifetime problem originates from an insufficient understanding of the prompt background. We study the systematic uncertainties associated with the prompt background in Section 4.9.1. The measured Δm_d and τ_{B^+}/τ_{B^0} values are:

$$\Delta m_d \ = 0.501 \pm 0.029 \text{ ps}^{-1} \tag{103}$$

and

$$\tau_{B^+}/\tau_{B^0} = 1.096 \pm 0.013 \tag{104}$$

which are both consistent with the world average expectations [1]. The fitted amplitude $\mathcal{A}$ is also consistent with one as expected. The correlation matrix of the fit parameters for the Amplitude Fit to the data is compiled in Table 21.

The result of the Dilution Fit to real data is shown in Table 22. The scale factors for all taggers are consistent with one within 1-2 standard deviations. The correlation

Parameter	Fit result	PDG
$c\tau_{B^0}$	461 μm (fixed)	460.8 ± 4.2 μm
τ_{B^+}/τ_{B^0}	1.086 (fixed)	1.086 ± 0.017
Δm_d	0.502 ps^{-1} (fixed)	0.502 ± 0.007 ps^{-1}
D_{SMT}	0.949 ± 0.037	–
D_{SET}	1.003 ± 0.056	–
$D_{JQT(vertex)}$	0.950 ± 0.061	–
$D_{JQT(track)}$	0.856 ± 0.079	–
$D_{JQT(high\ p_T)}$	1.079 ± 0.092	–
R_f	2.18 (fixed)	2.18 ± 0.14
f^{**}	0.31 (fixed)	0.31 ± 0.05
P_V	0.85 (fixed)	–
ϵ^*	0.54 (fixed)	–

Table 22: Fit result of Dilution Fit configuration for the dilution scale factor on real data.

	D_{SMT}	D_{SET}	D_{SecVtx}	D_{Jpb}	D_{Hpt}
D_{SMT}	1.0	0.0	0.0	0.0	0.0
D_{SET}	0.0	1.0	0.0	0.0	0.0
$D_{JQT(vertex)}$	0.0	0.0	1.0	0.0	0.0
$D_{JQT(track)}$	0.0	0.0	0.0	1.0	0.0
$D_{JQT(high\ p_T)}$	0.0	0.0	0.0	0.0	1.0

Table 23: Parameter correlations of the Dilution Fit for the dilution scale factors. Since the data samples corresponding to the five taggers are independent, the correlation matrix is diagonal.

matrix of the fit parameter correlations for the Dilution Fit to the data is shown in Table 23, indicating that all dilution scale factors are uncorrelated. This is expected since the different tags are mutually exclusive.

Likelihood projections of the asymmetry from the Amplitude Fit for each of the three $D^+\ell$, $D^0\ell$ and $D^{*+}\ell$ final states and each of the five flavor tags (SMT, SET, JQT-vertex, JQT-track and JQT-High-p_T) are shown in Figures 57 and 58. We observe clear oscillations in the $D^+\ell$ and $D^{*+}\ell$ samples which are dominated by B^0 decays, while in the $D^0\ell$ sample, dominated by B^+ decays, the asymmetry is almost flat. With the sample composition parameters given in Table 20, the fractions of B^0 mesons are about 87% in $D^+\ell$, 28% in $D^0\ell$ and 82% in $D^{*+}\ell$ sample. The amplitude of the oscillation asymmetries gives the average dilution for each tagger. As expected from Figure 47, we observe that the SMT tags have the highest dilution followed by SET, JQT(vertex), JQT(track) and finally JQT(high p_T).

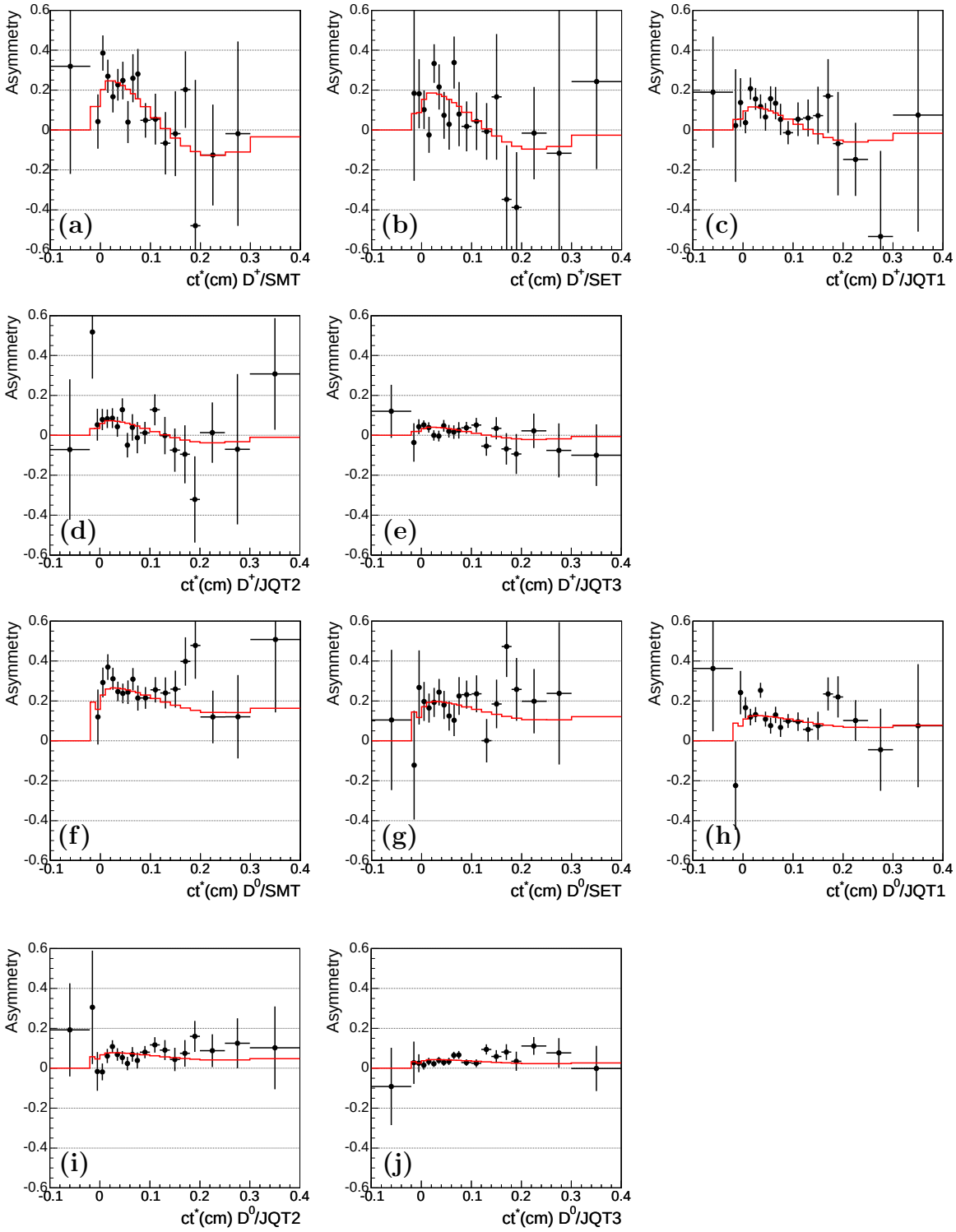

Figure 57: Likelihood projections of asymmetry for lepton-D^+/lepton-D^0 events with each tagger: (a)/(f) SMT, (b)/(g) SET, (c)/(h) JQT (vertex), (d)/(i) JQT (track) and (e)/(j) JQT (high p_T).

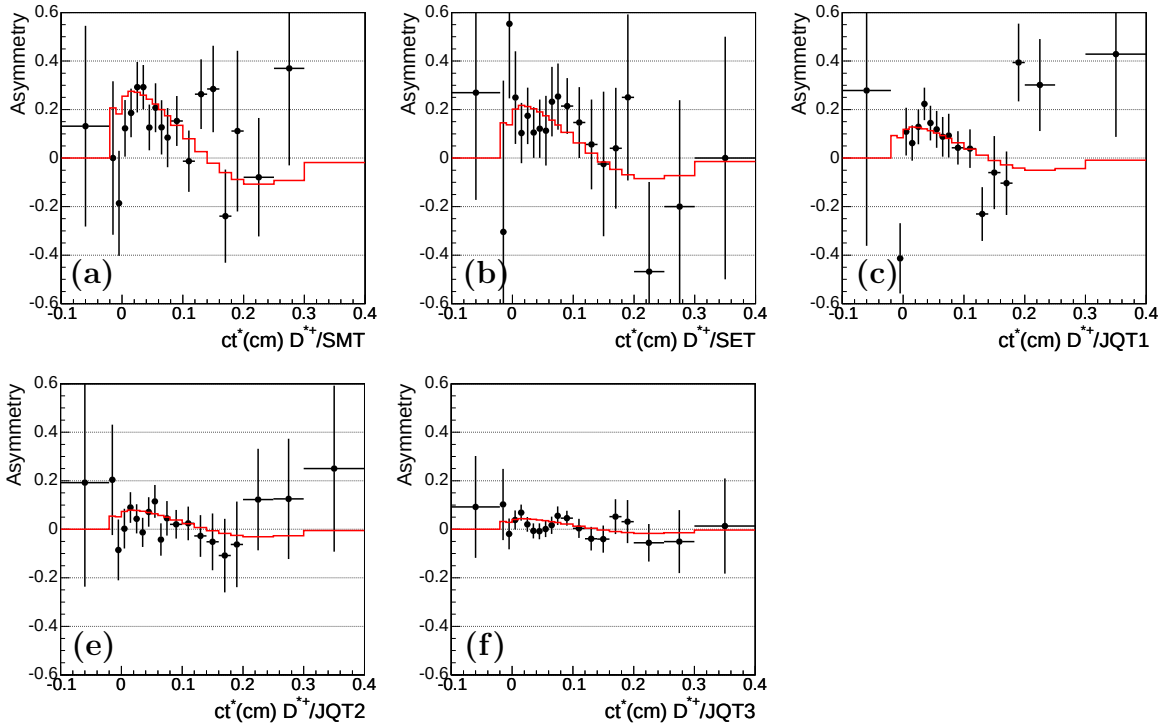

Figure 58: Likelihood projections of asymmetry for lepton-D^{*+} events with each tagger: (a) SMT, (b) SET, (c) JQT (vertex), (d) JQT (track) and (e) JQT (high p_T).

4.9 Systematic Uncertainties

In this section we discuss the various sources of systematic uncertainties on the B^0 lifetime, the B^0 to B^+ lifetime ratio, the oscillation frequency Δm_d, the amplitude $\mathcal{A}$ and the dilution scale factors $\mathcal{S}_\mathcal{D}$.

4.9.1 Prompt Background

The prompt background is one of the main sources of systematic uncertainties. In the default fit configuration, we assume that the fractions and ct^* distribution shapes of the prompt background are the same in both-right and wrong-sign $D\ell$ samples. With this assumption, we determined the fractions of prompt background as the ratios between the number of signal $D^{+/0}$ events in the wrong-sign sample and the corresponding numbers of events in the right-sign sample. For $D^+\ell$ events, we also scale the fraction by the ratio between the combinatorial background levels in right and wrong-sign samples: $f_{pmt} \rightarrow 1.6 \times f_{pmt}$. The shapes of the ct^* distributions obtained from the wrong-sign sample were used as templates in the right-sign sample weighted by the corresponding fractions. Although plausible, the above assumptions may not be absolutely true. To estimate the systematic uncertainties associated with the prompt background, which is expected to be the main reason for the short B lifetimes observed in $B \rightarrow D\ell X$ semileptonic decays, we run the Amplitude and Dilution Fit configuration without including the contribution from wrong-sign $D^+\ell^+$

Parameter	No Prompt Background
$c\tau_{B^0}$	411.7 ± 2.6 μm
τ_{B^+}/τ_{B^0}	1.177 ± 0.014
Δm_d	0.496 ± 0.030 ps^{-1}
A	0.904 ± 0.027

Parameter	No prompt Background
D_{SMT}	0.901 ± 0.036
D_{SET}	0.957 ± 0.054
$D_{JQT(vertex)}$	0.906 ± 0.058
$D_{JQT(track)}$	0.803 ± 0.074
$D_{JQT(high\ p_T)}$	1.014 ± 0.087

Table 24: Fit results without the prompt background included in the likelihood fit.

Parameter	Scaled Prompt Background	Unscaled prompt Background
$c\tau_{B^0}$	448.1 ± 2.9 μm	434.1 ± 2.8 μm
τ_{B^+}/τ_{B^0}	1.071 ± 0.013	1.136 ± 0.013
Δm_d	0.495 ± 0.030 ps^{-1}	0.504 ± 0.029 ps^{-1}
A	0.949 ± 0.028	0.952 ± 0.028
D_{SMT}	0.944 ± 0.037	0.944 ± 0.037
D_{SET}	0.997 ± 0.056	0.997 ± 0.056
$D_{JQT(vertex)}$	0.941 ± 0.060	0.941 ± 0.060
$D_{JQT(track)}$	0.846 ± 0.078	0.846 ± 0.078
$D_{JQT(high\ p_T)}$	1.069 ± 0.091	1.069 ± 0.091

Table 25: Fit results with and without the prompt background scaled by combinatorial background levels in all three decay signatures.

events in the fit. The outcome of these fits is summarized in Table 24. We have also repeated the study in two other configurations. First, we scale all prompt background fractions by the ratio between the combinatorial background levels. The scaling factors are 1.6 for $D^+\ell$, 0.6 for $D^0\ell$ and 1.8 for $D^{*+}\ell$. In the second configuration we removed all scaling factors. As shown in Table 25, we find that in both configurations, the effect on the dilution scale factors and lifetimes is much smaller than by totally removing the prompt background. The effect on Δm_d is comparable with the case of allowing no prompt background. We estimate the systematic uncertainties associated with the prompt background as 50% of the deviations of the measured quantities for the case in which the prompt background is totally removed with respect to the default case. The systematic errors associated with this effect are shown in Table 27.

4.9.2 Sample Composition

To evaluate the systematic uncertainties associated with the sample composition, we repeat the fit with the sample composition parameters floating. The result of those fits are summarized in Table 26. We find that in general the systematic errors are smaller than those corresponding to the prompt background except for Δm_d, for which the sample composition is the main source of systematic uncertainties.

Parameter	Dilution Fit	Amplitude Fit
$c\tau_{B^0}$	461 μm (fixed)	438.1 ± 4.4 μm
τ_{B^+}/τ_{B^0}	1.086 (fixed)	1.125 ± 0.015
Δm_d	0.502 ps^{-1} (fixed)	0.512 ± 0.032 ps^{-1}
A	1.0 (fixed)	0.956 ± 0.029
D_{SMT}	0.932 ± 0.037	1.0 (fixed)
D_{SET}	0.985 ± 0.056	1.0 (fixed)
$D_{JQT(vertex)}$	0.927 ± 0.060	1.0 (fixed)
$D_{JQT(track)}$	0.846 ± 0.077	1.0 (fixed)
$D_{JQT(high\ p_T)}$	1.070 ± 0.091	1.0 (fixed)
R_f	2.228 ± 0.132	2.247 ± 0.133
f^{**}	0.120 ± 0.027	0.328 ± 0.042
P_V	0.210 ± 0.252	0.547 ± 0.093
ϵ^*	0.596 ± 0.035	0.604 ± 0.036

Table 26: Fit result with the sample composition parameters floating.

4.9.3 Resolution Scale Factor

In the likelihood fit, we use a global ct^* resolution scale factor on the event-by-event lifetime uncertainty σ_{ct^*} which is obtained from Reference [55] and fixed to be 1.35. We study the systematic effect due to the limited knowledge on this scale factor by varying it up to 1.4 and down to 1.3. The minor effects on all fit parameters are summarized in Table 27.

4.9.4 SVT Bias Curve

Another source of systematic uncertainties comes from the SVT trigger bias curve. The difference between the SVT impact parameter and the SVX impact parameter Δd_0, as seen in data, can be different from Δd_0 from Monte Carlo simulations. In Reference [60] the difference between $\Delta d_0^{\rm data}$ and $\Delta d_0^{\rm MC}$ was fit to a double Gaussian where the narrow component dominates ($\approx 80\%$) indicating a width of (26.2 ± 0.2) μm in Monte Carlo and (28.8 ± 0.1) μm in data. We introduce an extra smearing of ≈ 12 μm to the SVT track impact parameter before the trigger confirmation to account for this difference and re-derive the trigger efficiency curves. With the new efficiency curve we repeat the fit and assign a systematic error equal to the difference between the new result and the default fit.

4.9.5 Asymmetry of Prompt Background

When including the prompt background from the wrong-sign $D\ell$ events in the analysis, we assume the asymmetry of the prompt background to be zero as discussed in Section 4.6 and supported by Figure 54. We explicitly calculate the prompt background asymmetry between wrong-sign and right-sign flavor tags. We obtain asymmetries of $(1 \pm 6)\%$, $(6 \pm 5)\%$ and $(-10 \pm 7)\%$ in the $D^+\ell$, $D^0\ell$ and $D^{*+}\ell$ samples, respectively. Conservatively, we assign a $\pm 10\%$ systematic uncertainty on our

knowledge of the prompt background asymmetry. In an additional cross check, we count the number of right-sign and wrong-sign tag events but apply a weight according to the predicted dilution. We again do not see an asymmetry larger than 10%. Thus we repeat the fits with ±10% asymmetry for the prompt background and note the excursions of the fit results. In our default fit we assume that the dilution and resolution distributions for the prompt background are the same as for signal. We change this assumption by using the combinatorial background distributions instead. The results are almost identical to the case in which the signal distributions are used and the systematic errors are negligible. The effects of the prompt background asymmetry on all fit parameters are summarized in Table 27.

4.9.6 Fitting Method

We study the systematic uncertainties related to possible biases in our fitting method. Using the values in Table 18 we assign systematic errors on the B^0 lifetime $c\tau_{B^0}$, lifetime ratio τ_{B^+}/τ_{B^0}, oscillation frequency Δm_d and amplitude $\mathcal{A}$ as the offset of the pull distributions multiplied by the average errors on the fit results as shown in Table 27.

Parameter	SC(%)	PB(%)	SF(%)	SVT(%)	PBA(%)	Fit(%)	Tot(%)
$c\tau_{B^0}$	1.69	3.80	0.06	0.68	0.02	0.09	4.22
τ_{B^+}/τ_{B^0}	2.65	3.45	0.03	0.43	0.03	0.28	4.38
Δm_d	2.20	0.50	0.04	0.01	1.41	1.95	3.30
$\mathcal{A}$	0.21	2.80	0.02	0.01	2.70	0.10	3.90
D_{SMT}	1.79	2.55	0.04	0.04	1.17	–	3.33
D_{SET}	1.79	2.15	0.02	0.04	1.15	–	3.03
$D_{JQT(vertex)}$	2.42	2.30	0.02	0.06	2.65	–	4.26
$D_{JQT(track)}$	1.14	3.30	0.02	0.13	4.80	–	5.94
$D_{JQT(high\ p_T)}$	0.83	3.00	0.01	0.12	8.87	–	9.40

Table 27: Systematic relative errors due to sample composition (SC), prompt background (PB), scale factor (SF), SVT bias (SVT), prompt background asymmetry (PBA) and fitting method (Fit) as well as total systematic errors.

The total systematic error for each parameter in Table 27 is obtained by adding in quadrature the systematic uncertainties from different sources. The final result on the oscillation frequency Δm_d including systematic uncertainties is:

$$\Delta m_d \ = (0.501 \pm 0.029(stat.) \pm 0.017(syst.))\ \mathrm{ps}^{-1}. \tag{105}$$

4.10 Evaluation of the Tagging Performance

One of the main purposes of the $B^0\bar{B}^0$ mixing analysis is to re-calibrate the dilutions of the flavor taggers. The dependence of the dilutions as function of different parameters has been established using the inclusive lepton-SVT data sample. In the Δm_d measurement, we assume the same dependence but allow the absolute dilution,

Tagger	Efficiency (%)	Dilution Scale Factor
SMT	5.015 ± 0.074	$0.949 \pm 0.037 \pm 0.032$
SET	3.339 ± 0.064	$1.003 \pm 0.056 \pm 0.030$
JQT (vertex)	10.169 ± 0.101	$0.950 \pm 0.061 \pm 0.040$
JQT (track)	13.870 ± 0.114	$0.856 \pm 0.079 \pm 0.051$
JQT (high p_T)	52.293 ± 0.166	$1.079 \pm 0.092 \pm 0.101$
No Tag	15.313 ± 0.117	–
Total	100.0 ± 0.245	–

Table 28: Efficiency and dilution scale factors for each tagger. The first error of the dilution scale factors is statistical and the second is systematic.

Tagger	$< \mathcal{D} >$ (%)	$\sqrt{< \mathcal{D}^2 >}$ (%)	$S_{\mathcal{D}} < \mathcal{D} >$ (%)	$S_{\mathcal{D}}\sqrt{< \mathcal{D}^2 >}$ (%)
SMT	29.8 ± 0.3	36.2 ± 0.3	28.3 ± 1.5	34.3 ± 1.7
SET	21.7 ± 0.3	29.5 ± 0.3	21.8 ± 1.4	29.6 ± 1.9
JQT (vertex)	14.0 ± 0.1	17.0 ± 0.1	13.3 ± 1.0	16.2 ± 1.2
JQT (track)	9.6 ± 0.1	11.4 ± 0.1	8.2 ± 0.9	9.8 ± 1.1
JQT (high p_T)	4.3 ± 0.0	5.1 ± 0.0	4.6 ± 0.1	5.5 ± 0.1

Table 29: Average dilution with and without utilizing the p_T^{rel}, lepton likelihood and jet charge dependence before and after dilution scaling.

measured by the dilution scale factors, to float in the maximum likelihood fit. We find that the central values of all dilution scale factors are consistent with one within 2σ as shown in Table 28. For each tagger we calculate the efficiency as:

$$\varepsilon_{Tag} = \frac{\sum_{i \in Sig\&Tag} 1 - S_{Bkg} \sum_{i \in Bkg\&Tag} 1}{\sum_{i \in Sig} 1 - S_{Bkg} \sum_{i \in Bkg} 1} \tag{106}$$

where, for example, $\sum_{i \in Sig\&Tag}$ means sum over all events in signal region tagged by a certain tagger "Tag" and S_{Bkg} is the ratio between the number of background events in the signal region and the number of background events in the side-bands. Similarly, we obtain the tagging dilution without using the known dependences of dilution on parameters like p_T^{rel}, lepton likelihood or jet charge. This is simply the average dilution calculated as:

$$< \mathcal{D}_{Tag} >= \frac{\sum_{i \in Sig\&Tag} \mathcal{D}_{pred,i} - S_{Bkg} \sum_{i \in Bkg\&Tag} \mathcal{D}_{pred,i}}{\sum_{i \in Sig} 1 - S_{Bkg} \sum_{i \in Bkg} 1}. \tag{107}$$

However, in this analysis we use the dependence of the tagging dilution on p_T^{rel}, lepton likelihood and jet charge on an event-by-event basis. Instead of assigning the same average dilution to all events, we assign higher dilutions to events with high p_T^{rel}, lepton likelihood or jet charge. As shown in Reference [40], this procedure enhances the effective dilution due to the quadratic dependence of the tagging power $\sim \varepsilon \mathcal{D}^2$:

$$\sqrt{< \mathcal{D}_{Tag}^2 >} = \sqrt{\frac{\sum_{i \in Sig\&Tag} \mathcal{D}_{pred,i}^2 - S_{Bkg} \sum_{i \in Bkg\&Tag} \mathcal{D}_{pred,i}^2}{\sum_{i \in Sig} 1 - S_{Bkg} \sum_{i \in Bkg} 1}}. \tag{108}$$

Tagger	$\varepsilon < \mathcal{D} >^2$ (%)	$\varepsilon < \mathcal{D}^2 >$ (%)	$\varepsilon S_{\mathcal{D}}^2 < \mathcal{D} >^2$ (%)	$\varepsilon S_{\mathcal{D}}^2 < \mathcal{D}^2 >$ (%)
SMT	0.425 ± 0.011	0.657 ± 0.014	0.401 ± 0.024	0.592 ± 0.035
SET	0.158 ± 0.006	0.291 ± 0.009	0.159 ± 0.012	0.292 ± 0.020
JQT (vertex)	0.198 ± 0.004	0.295 ± 0.005	0.179 ± 0.015	0.266 ± 0.022
JQT (track)	0.129 ± 0.002	0.181 ± 0.002	0.094 ± 0.012	0.133 ± 0.017
JQT (high p_T)	0.095 ± 0.001	0.136 ± 0.001	0.110 ± 0.013	0.158 ± 0.019
Total	1.025 ± 0.013	1.559 ± 0.017	0.943 ± 0.035	1.441 ± 0.052

Table 30: Tagging power with and without utilizing the $p_T^{\rm rel}$, lepton likelihood and jet charge dependence before and after dilution scaling.

Combining the five opposite side tagging algorithms, the total tagging power corrected with the dilution scale factors $S_{\mathcal{D},Tag}$ is given by:

$$\varepsilon\mathcal{D}^2 \;=\; \sum_{Tag} \varepsilon_{Tag}\; S^2_{\mathcal{D},Tag}\; < \mathcal{D}^2_{Tag} > \;=\; (\; 1.441 \pm 0.052 \;)\;\% \tag{109}$$

As seen in Tables 29 and 30, using the dilution dependence on $p_T^{\rm rel}$, lepton likelihood and jet charge increases the dilution by $\sim 20\%$ and the total tagging power by $\sim 50\%$. The uncertainty for the total $\varepsilon\mathcal{D}^2$ is calculated assuming no correlations between the uncertainties of the different taggers.

5 Search for $B_s^0\bar{B}_s^0$ Oscillations in Semileptonic $B_s^0 \to D_s^-\ell^+\nu$ Decays

In this section we describe the search for particle-antiparticle oscillations in the neutral $B_s^0\bar{B}_s^0$ system. The B_s^0 meson candidates are partially reconstructed using semileptonic $B_s^0 \to D_s^-\ell^+\nu$ decays. Opposite side lepton and jet charge tagging algorithms are used to determine the B flavor at production. The tagging dilutions measured along with Δm_d in semileptonic $B^{0/+} \to D\ell\nu$ decays (see Section 4) are used as input for the B_s^0 mixing search. Using the lepton likelihood, p_T^{rel} or the jet charge, the dilution is predicted on an event-by-event basis by the opposite side lepton or jet charge tagging algorithms. To establish a lower limit on the oscillation frequency Δm_s, we employ the amplitude method introduced in Section 1.2.4. The analysis is initially performed using randomized tagging decisions. With the "blind analysis" we determine the sensitivity of our measurement and only after the sensitivity is obtained, the analysis is "un-blinded" and the lower limit on Δm_s established.

5.1 Data Selection and Signal Reconstruction

We use the same data set described in Section 4.1.1 to search for semileptonic decays $B_s^0 \to D_s^-\ell^+\nu$. In a similar fashion, first the D_s^- candidate is searched by requiring the SVT track to be one of the tracks from the D_s^- decay. The D_s^- candidates are reconstructed in three decay modes

- $D_s^- \to \phi\pi^-$, $\phi \to K^+K^-$
- $D_s^- \to K^{*0}K^-$, $K^{*0} \to K^+\pi^-$
- $D_s^- \to \pi^+\pi^-\pi^-$.

To partially reconstruct the B_s^0 candidates we apply a procedure similar to the one described in Section 4.1.2 for $B^{0/+}$ reconstruction. The same secondary vertex fitting package discussed in Reference [34] is utilized to determine the D_s and B_s^0 vertices. The tracks forming the D_s^- candidate are refit with a common vertex constraint referred to as the tertiary vertex. The secondary vertex, where the B_s^0 decays to a lepton and a D_s^- meson, is obtained by intersecting the trajectories corresponding to the lepton and D_s flight paths.

5.1.1 $D_s^-\ell^+$ Signal Reconstruction

An optimization procedure of the $D_s^- \to \phi\pi^-$ signal yield was performed using $S/\sqrt{S+B}$ as figure of merit using the procedure already outlined in Section 4.1.2 for D^0/D^+ signal optimization. The signal events for the D_s^- decay are obtained from an inclusive $B_s^0 \to D_s^-\ell^+\nu$ Monte Carlo simulation while the background events are selected from the side-bands of the invariant mass distribution of D_s^- from data. The selection requirements from this optimization procedure are summarized in Table 31. In order to facilitate better cross checking with an independent analysis described in

	Original optimized	Combined proposal
$p_T(K)$ (GeV/c)	> 0.6	> 0.6
Probability (B Vertex)	$> 10^{-8}$	$> 10^{-7}$
2-dimensional $\chi^2_{xy}(D_s^-)$	< 10	< 10
$L_{xy}(PV \to D_s^-)/\sigma(L_{xy})$	> 5.5	> 5
$L_{xy}(PV \to D_s^-)$ (cm)	–	> 0.02
$\sigma_{ct^*}(PV \to B_s)$ (cm)	< 0.04	< 0.04
$ct(D_s^-)$ (cm)	$-0.05 < ct < 0.1$	$-0.015 < ct < 0.1$
$m(D_s^-\ell^+)$ (GeV/c^2)	$2.3 < m < 5.3$	$2.3 < m < 5.3$
$\lvert m(KK) - m(\phi)\rvert$ (GeV/c^2)	< 0.0075	< 0.0075
Helicity $\lvert\cos\Psi\rvert$	> 0.2	> 0.2

Table 31: Summary of selection requirements as a result of the original optimization procedure and the combined proposal of cuts for the $D_s^- \to \phi\pi^-$ sample.

	$D_s^- \to K^{*0}K^-$	$D_s^- \to \pi^+\pi^-\pi^-$
$p_T(K)$ (GeV/c)	> 0.7	> 0.7
Probability (B Vertex)	$> 10^{-3}$	$> 10^{-3}$
$L_{xy}(PV \to D_s^-)/\sigma(L_{xy})$	> 8	> 11
$L_{xy}(B \to D_s^-)/\sigma(L_{xy})$	> -1	> 2
$L_{xy}(PV \to D_s^-)$ (cm)	> 0.04	> 0.04
$\sigma_{ct^*}(PV \to B_s)$ (cm)	< 0.04	< 0.04
$ct(D_s^-)$ (cm)	$-0.015 < ct < 0.1$	$-0.015 < ct < 0.1$
$m(D_s^-\ell^+)$ (GeV/c^2)	$2.3 < m < 5.3$	$2.3 < m < 5.3$
$\lvert m(K\pi) - m(K^*)\rvert$ (GeV/c^2)	< 0.060	–
Helicity $\lvert\cos\Psi\rvert$	> 0.2	–

Table 32: Summary of selection requirements for the $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^+\pi^-\pi^-$ sample.

Reference [59], we adopt a common set of selection requirements that will provide the same event sample for both analyses. The selection requirement from this combined proposal is also listed in Table 31. For the other two decay modes, $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^+\pi^-\pi^-$, the requirements from the combined proposal are listed in Table 32.

The meaning of the variables listed in Tables 31 and 32 is the same as described in Section 4.1.2. A variable used for D_s signal optimization that was not used for $D^{0/+}$ optimization is the helicity angle Ψ. Since the ϕ meson has spin one and both the D_s^- and π^- have spin zero, the helicity angle Ψ, which is the angle between the K^- and D_s^- directions in the ϕ rest frame, exhibits a distribution $dN/d(\cos\Psi) \sim \cos^2\Psi$. A requirement that $|\cos\Psi| > 0.2$ is applied to suppress combinatorial background, which is found to be flat in the $\cos\Psi$ distribution.

In the case of the $D_s^- \to K^{*0}K^-$ channel, it is possible that $D^- \to K^+\pi^-\pi^-$ decays, in which one of the pions is misidentified as a kaon, are reconstructed as $D_s^- \to K^{*0}K^-$ decays. We avoid this possibility by removing the $K\pi\pi$ combinations

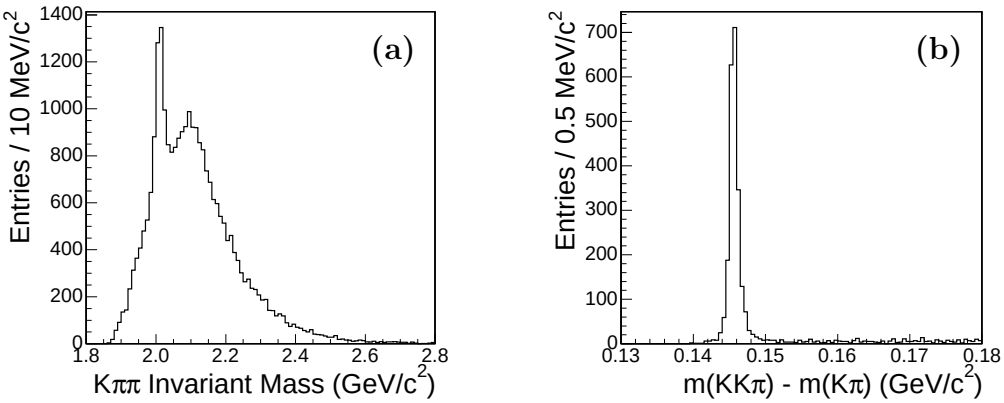

Figure 59: Invariant mass distribution of (a) $K\pi\pi$ and (b) the difference between the $K\pi\pi$ and the $K\pi$ invariant masses.

with invariant mass around the D^- mass: $|m(K^+\pi^-\pi^-) - m(D^+)| > 20$ MeV/c^2, where the negative kaon from the D_s decays is now assigned the pion mass. We note that the positive kaon (from $K^{*0} \to K^+\pi^-$) could not be misidentified as a pion from D^- decays because both pions in the latter decay are negatively charged. In the case of the $D_s^- \to \pi^+\pi^-\pi^-$ channel, we remove candidates that come from $D^{*+} \to K\pi\pi$ decays by requiring $|m(K\pi\pi) - m(K\pi)| > 150$ MeV/c^2. Figure 59 shows the $K\pi\pi$ invariant mass distribution and also the mass difference $m(K\pi\pi) - m(K\pi)$.

The D_s^- mass distributions corresponding to the above selection requirements are shown in Figure 60 for both right and wrong-sign $D_s\ell$ combinations. Again, we refer to $D_s^-\ell^+$ ($D_s^-\ell^-$) as right-sign (wrong-sign) $D\ell$ combinations. Besides the main D_s^- mass peak at ~ 1.968 GeV/c^2, the right-sign distributions also contain a second mass peak at ~ 1.869 GeV/c^2 corresponding to the Cabibbo suppressed $D^- \to K^+K^-\pi^-$ decays. The D_s^- mass distributions in Figure 60 are fit to a double Gaussian signal function plus a linear background for $D_s^- \to \phi\pi^-$, a single Gaussian plus linear background for $D_s^- \to K^{*0}K^-$ and a single Gaussian plus a quadratic background for $D_s^- \to \pi^+\pi^-\pi^-$. The Cabibbo suppressed D^+ mass peak is fit with a single Gaussian function. Table 33 summarizes the fit parameters for the three decay modes.

We define the signal region as $\pm 3\sigma$ of the fitted signal width around the nominal D_s^- mass value. In the case of $D_s^- \to \phi\pi^-$, σ is the width of the narrow Gaussian. The obtained signal yields are summarized in Table 34 together with the signal-to-background ratio S/B, the significance $S/\sqrt{S+B}$ and the fit probability. The fit probabilities for each decay mode are: 63.9%, 3.4% and 32.7%, respectively.

Figure 61 shows the D_s^- mass distribution for the wrong-sign $D_s^-\ell^-$ combinations. The D_s^- mass distribution for $D_s^- \to \phi\pi^-$ is again fit to a double Gaussian signal function plus a linear background. The parameters describing the signal distribution are fixed from the right-sign fit result. A small D_s^- mass signal corresponding to a yield of $(3.2 \pm 0.8)\%$ of the corresponding right-sign yield is observed. There is no evidence for a wrong-sign mass peak in the other two $D_s^-\ell^-$ mass distributions which are fit to a linear and quadratic background function, respectively.

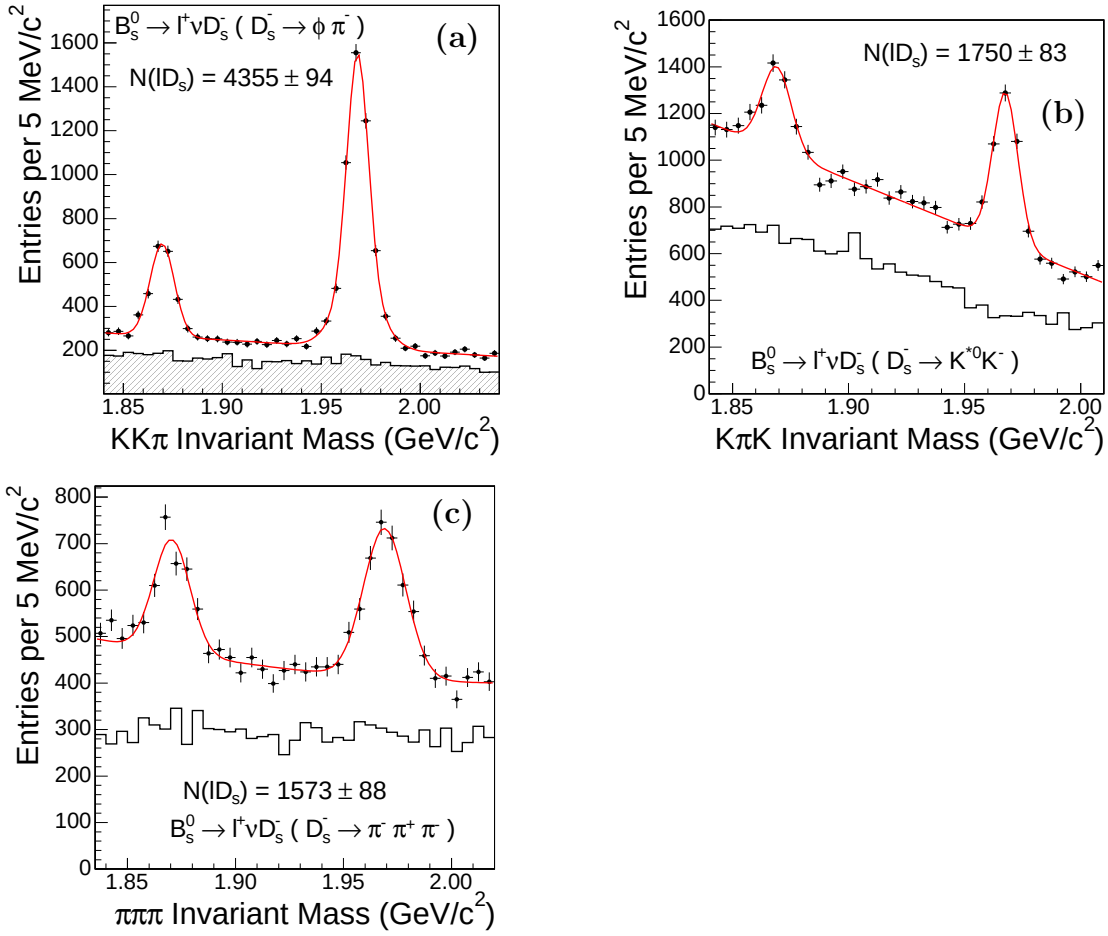

Figure 60: D_s^- mass distributions for the right-sign (points with error bars) and wrong-sign (solid) $D\ell$ combinations: (a) $D_s^- \to \phi\pi^-$, (b) $D_s^- \to K^{*0}K^-$, and (c) $D_s^- \to \pi^+\pi^-\pi^-$. The fit to D_s^- right-sign mass distribution is superimposed.

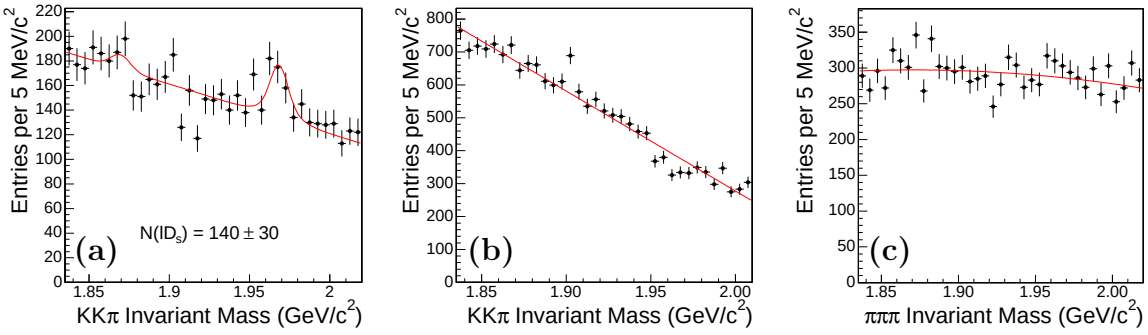

Figure 61: D_s^- mass distribution for wrong-sign $D_s^-\ell^-$ combinations: (a) $D_s^- \to \phi\pi^-$, (b) $D_s^- \to K^{*0}K^-$and (c) $D_s^- \to \pi^+\pi^-\pi^-$.

Parameter	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to \pi^+\pi^-\pi^-$
c_1	1265.16 ± 97.28	8582.02 ± 244.27	9641.03 ± 2655.02
c_2	-535.75 ± 49.72	-4033.82 ± 125.62	-9042.19 ± 2757.68
c_3	–	–	2211.63 ± 715.01
$N(D^+)$	1290.56 ± 65.53	1131.66 ± 120.11	1027.02 ± 86.11
$m(D^+)$ (MeV/c^2)	1869.59 ± 0.30	1869.01 ± 0.60	1870.47 ± 0.67
$\sigma(D^+)$ (MeV/c^2)	6.02 ± 0.31	6.25 ± 0.68	8.35 ± 0.75
$N(D_s)$	4355.90 ± 94.39	1750.42 ± 83.21	1573.98 ± 88.80
$m(D_s)$ (MeV/c^2)	1968.38 ± 0.13	1967.56 ± 0.27	1969.08 ± 0.57
$\sigma_1(D_s)$ (MeV/c^2)	10.79 ± 1.47	5.35 ± 0.26	9.71 ± 0.58
f_1	0.35 ± 0.10	1.0 (fixed)	1.0 (fixed)
$\sigma_2(D_s)$ (MeV/c^2)	5.23 ± 0.35	–	–

Table 33: Mass fit results for the three D_s^- decay modes described in the text. The mass peaks are fit with either a single or a double Gaussian function for signal distributions and either a linear or quadratic function for background distributions.

	N	S/B	$S/\sqrt{S+B}$	Fit Probability (%)
$D_s^- \to \phi\pi^-$	4355 ± 94	3.12	55.9	63.9
$D_s^- \to K^{*0}K^-$	1750 ± 83	0.42	22.8	3.4
$D_s^- \to \pi^+\pi^-\pi^-$	1573 ± 88	0.32	19.4	32.7

Table 34: Summary of signal yield N, signal to background ratio S/B, significance $S/\sqrt{S+B}$ and fit probability for the three reconstructed D_s^- decay modes.

5.2 Flavor Tagging

For the $B_s^0\bar{B}_s^0$ mixing search we use the same B flavor tagging algorithms that were utilized for the Δm_d measurement discussed in Section 4: opposite side electron and muon tagging as well as jet charge tagging. The tagging algorithms provide the predicted dilution as a function of the opposite side lepton likelihood, lepton p_T^{rel} or the jet charge. The dilution is determined on an event-by-event basis. The dilution predictions have initially been studied on an inclusive B sample based on the lepton-SVT trigger data. Using this data set, the dependence of the lepton tag dilution on $p_T^{\rm rel}$ and likelihood as well as the dependence of jet charge tag dilution on the charge of the jet were determined. The dilutions in a sample of semileptonic $B \to D\ell\nu$ decays may vary. This effect has been studied in the measurement of the B^0 mixing frequency Δm_d using semileptonic $B^{0/+} \to D\ell\nu$ decays as described in Section 4. Since we observe the B^0 oscillations, we can also measure the tag dilutions. In fact, we determine the dilution scale factors for each of the five taggers. These correction factors multiply the predicted dilutions and the adjusted values are used as inputs in the Δm_s analysis. To facilitate the comparison with an independent analysis described in Reference [59], we use the dilution scale factors from Reference [49]. Table 35 summarizes the dilution scale factors for each tagger and shows very good agreement with the corresponding scale factors listed in Table 28.

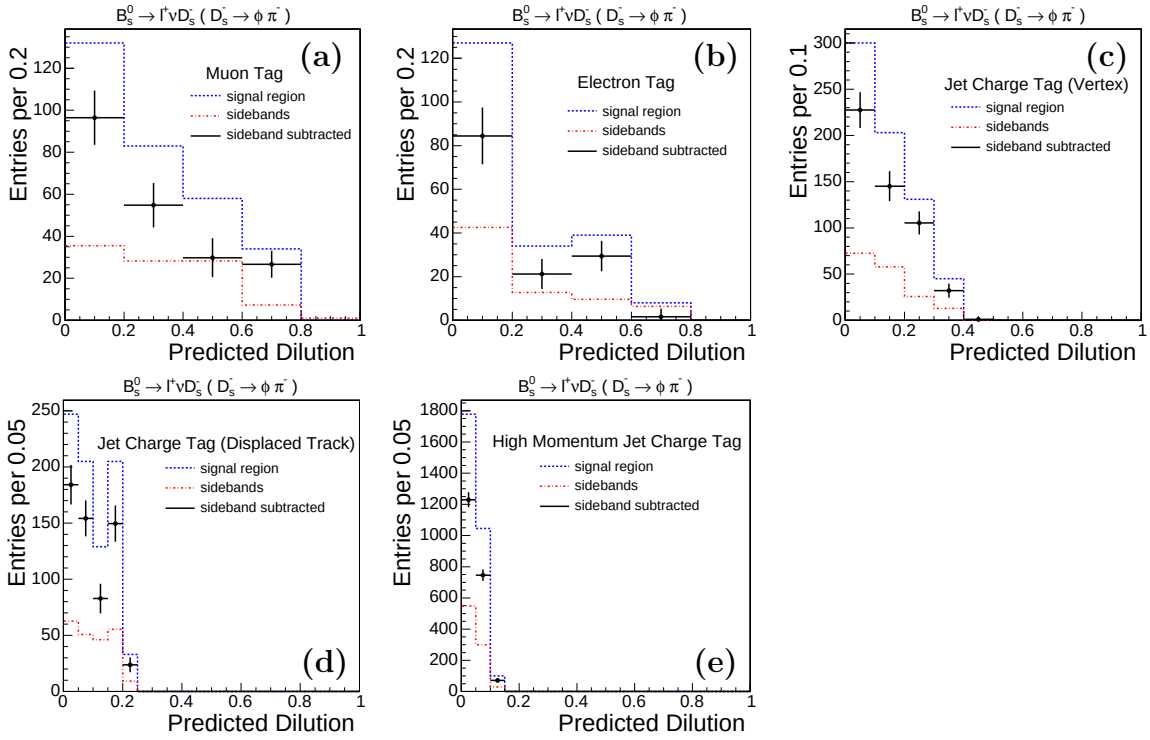

Figure 62: Predicted dilution distributions for the $D_s^- \to \phi\pi^-$ sample. The points with error bars in each diagram are the distributions for the signal after side-band subtraction. The dashed histograms correspond to the signal region and the dot-dashed ones correspond to the side-band events. The five plots correspond to (a) muon tag, (b) electron tag, (c) jet charge displaced vertex tag, (d) jet charge displaced track tag and (e) jet charge high p_T tag.

Tagger	Dilution Scale Factor
SMT	0.926 ± 0.048
SET	0.980 ± 0.063
JQT (vertex)	0.971 ± 0.074
JQT (track)	0.903 ± 0.091
JQT (high p_T)	1.082 ± 0.127

Table 35: Dilution scale factors used to correct the dilution predicted by each of the five taggers.

Using the same procedure as for the Δm_d measurement, only one tag is chosen in a given event using the following decision hierarchy: an SMT tag is chosen first if available, while SET is next. The tag lepton is required to have a likelihood larger that 5%. If no lepton tag is available, the JQT tagger is used starting with a displaced vertex tag first, displaced track tag next and high p_T jet charge tag last. If no tag is found in an event, it is considered a no-tag event and the predicted dilution is set to zero. Figure 62 shows the predicted dilution distributions applying the above tagging

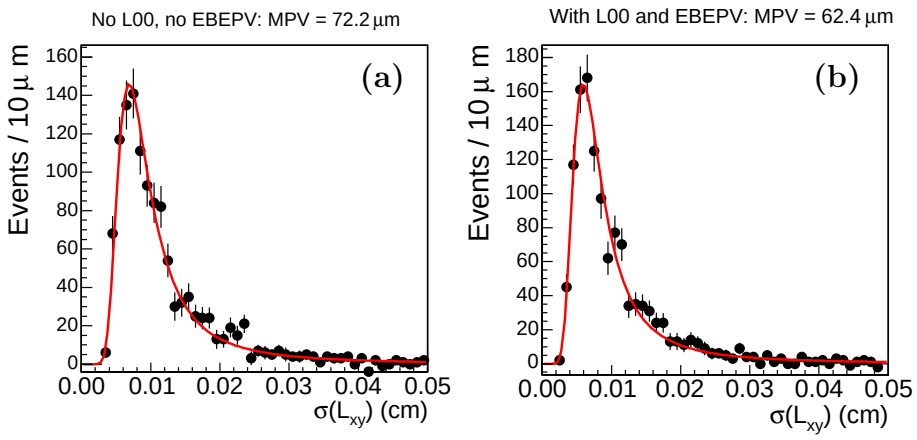

Figure 63: Distributions of the decay length resolution (a) without Layer 00 and event-by-event primary vertex and (b) with both Layer 00 and event-by-event primary vertex. Both distributions are fit with Landau functions and the most probable value in each distribution is interpreted as a measure of the decay length resolution. Adding Layer 00 and event-by-event primary vertex improves the decay length resolution by $\sim 14\%$.

algorithms to the $D_s^- \ell^+$ data.

5.3 Decay Time of B_s^0 Mesons

The measurement of the B_s^0 decay time follows the procedure outlined in Sections 1.2.1 and 4.4. Figure 63 shows that the improvement of the decay length uncertainty after inclusion of Layer 00 hits and use of event-by-event primary vertices is $\sim 14\%$.

The distributions of the pseudo-proper decay time ct^* and the resolution σ_{ct^*} for the right-sign lepton-D_s^- combinations are shown in Figures 64. Because the vertex resolutions are in general underestimated [55] by the vertex fitter we correct the decay time resolution with a scaling function which depends on various kinematic quantities such as the momentum of the lepton $p_T(\ell)$, the opening angle between the lepton and the D_s^- candidate $\Delta R(D_s\ell)$, the isolation $I(D_s\ell)$ and the pseudorapidity $\eta(D_s\ell)$ of the $D_s\ell$ pair. For this correction procedure, the isolation is defined as the ratio between the transverse momentum of the $D_s\ell$ pair and the transverse momentum of all tracks in a cone of $\Delta R \leq 0.8$ around the $D_s\ell$ momentum direction. The functional form of the scaling function is given by:

$$f(I,\eta,\Delta R,p_T) = (a_1 + a_2\, I + a_3\, I^2) \times (b_1 + b_2\,\eta + b_3\,\eta^2) \times (c_1 + c_2\,\Delta R + c_3\,\Delta R^2) \times (d_1 + d_2\, p_T + d_3\, p_T^2), \quad (110)$$

where the parameters a_i, b_i, c_i and d_i as well as a detailed description of the scaling procedure can be found in Reference [61]. Figure 65 shows the distributions of the isolation, pseudorapidity and opening angle for the $D_s\ell$ pair. With this procedure we obtain the resolution scale factor distributions shown in Figure 66. Table 36

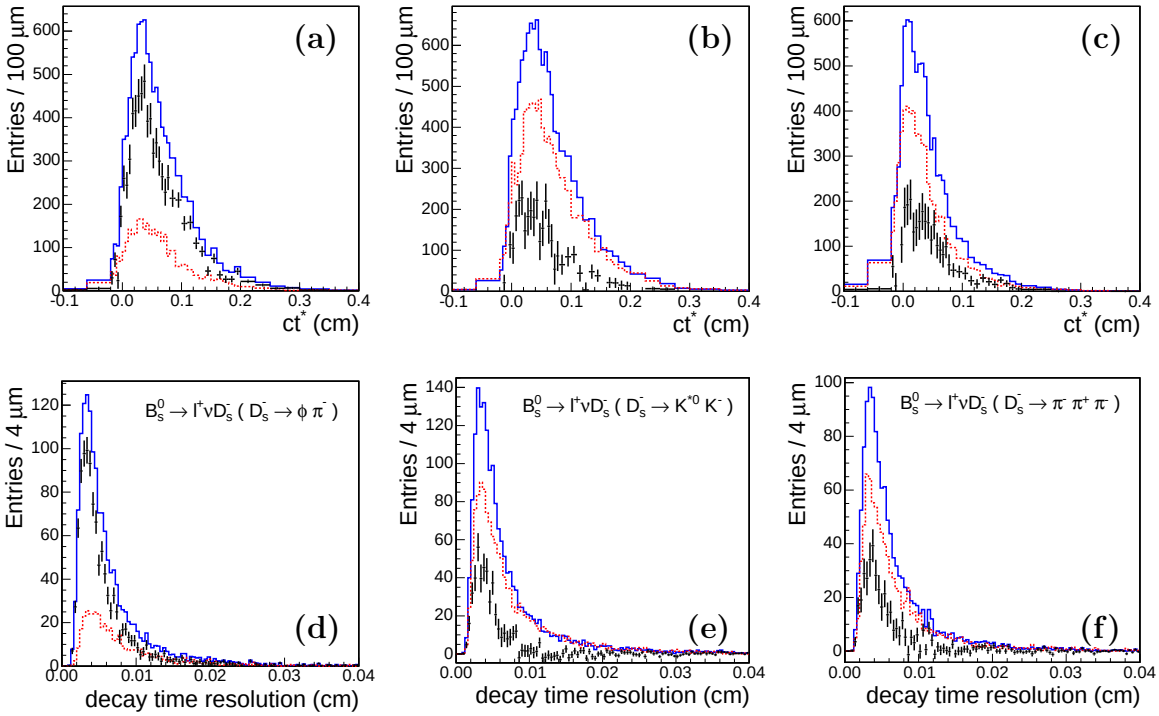

Figure 64: Distributions of the pseudo-proper decay time ct^* for right-sign $D_s^-\ell^+$ combinations with: (a) $D_s^- \to \phi\pi^-$, (b) $D_s^- \to K^{*0}K^-$and (c) $D_s^- \to \pi^+\pi^-\pi^-$. Distributions of the scaled pseudo proper decay time uncertainty σ_{ct^*} for right-sign $D_s^-\ell^+$ combinations: (d) $D_s^- \to \phi\pi^-$, (e) $D_s^- \to K^{*0}K^-$and (f) $D_s^- \to \pi^+\pi^-\pi^-$. The solid (dashed) histograms correspond to the signal (side-band) region, while the points with error bars represent the side-band subtracted signal distribution.

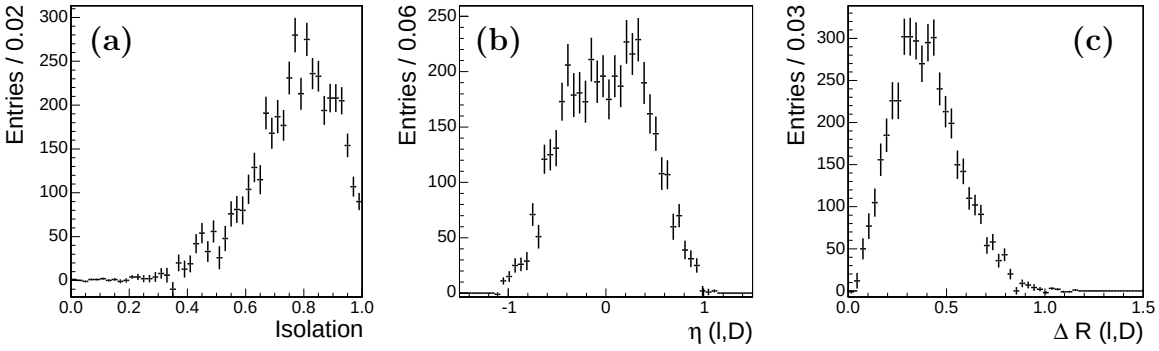

Figure 65: Distributions of (a) isolation, (b) pseudorapidity η and (c) ΔR for the $D_s^-\ell$ pair in $D_s^- \to \phi\pi^-$ decay mode. These are the variables which are used to determine the parameterization of the decay time resolution scale factor. The distributions corresponding to $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^+\pi^-\pi^-$ decay modes are similar.

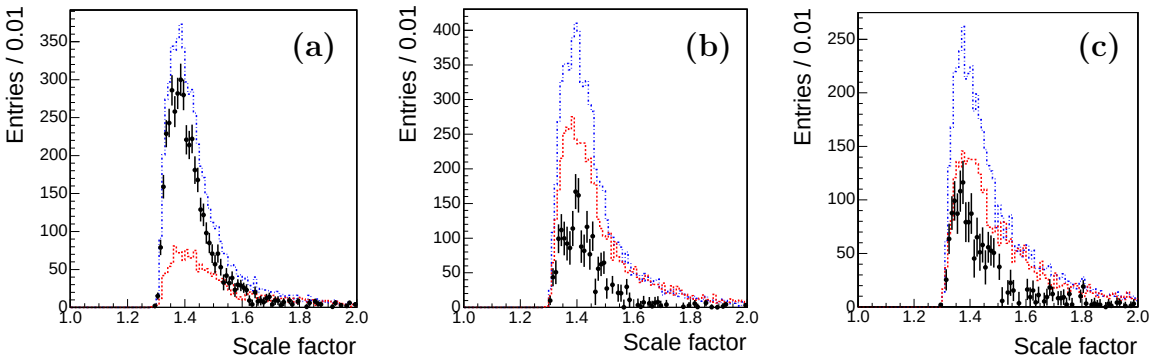

Figure 66: Distributions of the resolution scale factors for (a) $D_s^- \to \phi\pi^-$, (b) $D_s^- \to K^{*0}K^-$and (c) $D_s^- \to \pi^+\pi^-\pi^-$decay modes. The dot-dashed (dashed) histograms correspond to the signal (side-band) region, while the points with error bars correspond to the side-band subtracted distributions.

	Signal Region	Side-band Region	Side-band Subtracted
$D_s^- \to \phi\pi^-$	1.45	1.50	1.43
$D_s^- \to K^{*0}K^-$	1.45	1.47	1.44
$D_s^- \to \pi^+\pi^-\pi^-$	1.48	1.50	1.45

Table 36: Mean values of the resolution scale factor distributions.

summarizes the mean values of these distributions for signal and side-band regions as well as the mean values of the side-band subtracted distributions.

We also analyze the pseudo-proper decay time distributions of the wrong-sign $D_s^-\ell^-$ candidates and find evidence of prompt background contamination in the $D_s^- \to \phi\pi^-$ mode as seen in Figure 67. The prompt background distribution is fit with a single Gaussian with a width of ~ 95 μm. The fitting function will later be used as a template to describe the prompt background of the right-sign $D_s^-\ell^+$ combinations in the likelihood fitting procedure (see Section 5.5). No evidence of prompt background is found in the $D_s^- \to K^{*0}K^-$ or $D_s^- \to \pi^+\pi^-\pi^-$ modes. This is consistent with the large displacement requirement of the D_s vertex in these two decay channels.

As discussed in Section 1.2.1, to relate the pseudo proper decay time to the actual proper decay time, we introduce a correction factor which is the ratio between the observed transverse momentum of the $D_s^-\ell^+$ pair and the full transverse momentum of the B_s^0 meson:

$$K = \frac{p_T(D_s^-\ell^+)}{p_T(B_s)}. \tag{111}$$

The K-factor distribution $\mathcal{F}(K)$ is obtained from a realistic Monte Carlo simulation of the combined $B_s^0 \to D_s^{(*/**)-}\ell^+\nu$. The mixture of semileptonic B_s^0 decays into D_s^-, D_s^{*-} and D_s^{**-} states is taken from the `EvtGen` decay table [53] corresponding to the PDG values [1]. Separate MC samples for the three decay modes $\phi\pi^-$, $K^{*0}K^-$ and $\pi^+\pi^-\pi^-$ have been generated and used accordingly.

Figure 68 shows the K-factor distributions for the $D_s^- \to \phi\pi^-$ mode. The upper

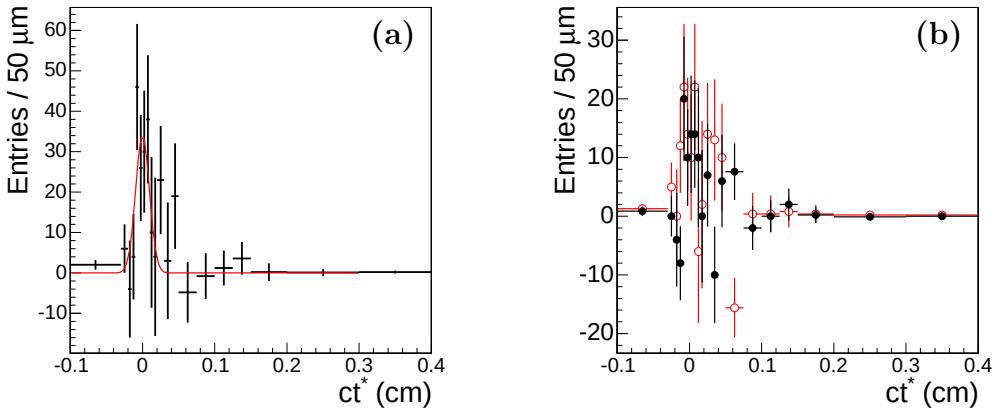

Figure 67: Distribution of (a) $ct^*(B_s)$ for wrong-sign lepton-D_s^- combinations in the $D_s^- \to \phi\pi^-$ mode after side-band subtraction plus the result of a fit described in the text and (b) the $ct^*(B_s)$ distribution divided into same-sign (full circles) and opposite-sign (open circles) tagged events.

$m(D_s^-\ell^+)$	$D_s^- \to \phi\pi^-$		$D_s^- \to K^{*0}K^-$		$D_s^- \to \pi^+\pi^-\pi^-$	
(GeV/c^2)	Mean	RMS	Mean	RMS	Mean	RMS
$2.3-5.3$	0.83	0.12	0.83	0.12	0.84	0.12
$2.3-3.2$	0.78	0.15	0.80	0.15	0.80	0.14
$3.2-3.7$	0.80	0.13	0.82	0.13	0.82	0.13
$3.7-4.2$	0.83	0.11	0.83	0.11	0.84	0.11
$4.2-4.7$	0.87	0.08	0.87	0.08	0.87	0.08
$4.7-5.3$	0.92	0.05	0.92	0.05	0.92	0.05

Table 37: Mean values and RMS of K-factor distributions in bins of $m(D_s^-\ell^+)$ for $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^+\pi^-\pi^-$.

left plot corresponds to all events with invariant mass of the $D_s\ell$ pair between 2.3 and 5.3 GeV/c^2 while the remaining five histograms show the K-factor distribution in different ranges of $D_s\ell$ invariant mass. The mean values and root-mean-squares of the K-factor distributions in bins of $m(D_s^-\ell^+)$ are shown in Table 37. The K-factor distributions for the $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^+\pi^-\pi^-$ modes are very similar to $D_s^- \to \phi\pi^-$as also shown in Table 37. It is found that the RMS/mean of the K-factor distribution decreases/increases with increasing $m(D_s^-\ell^+)$. This means that events with high $m(D_s^-\ell^+)$ are less affected by the missing neutrino momentum and are thus more valuable in a Δm_s measurement. The K-factor distributions for the $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^+\pi^-\pi^-$ decay modes are shown in Figure 69. The K-factor histograms have 50 bins each from 0.2 to 1.2 and are used for numerical integration in the likelihood function.

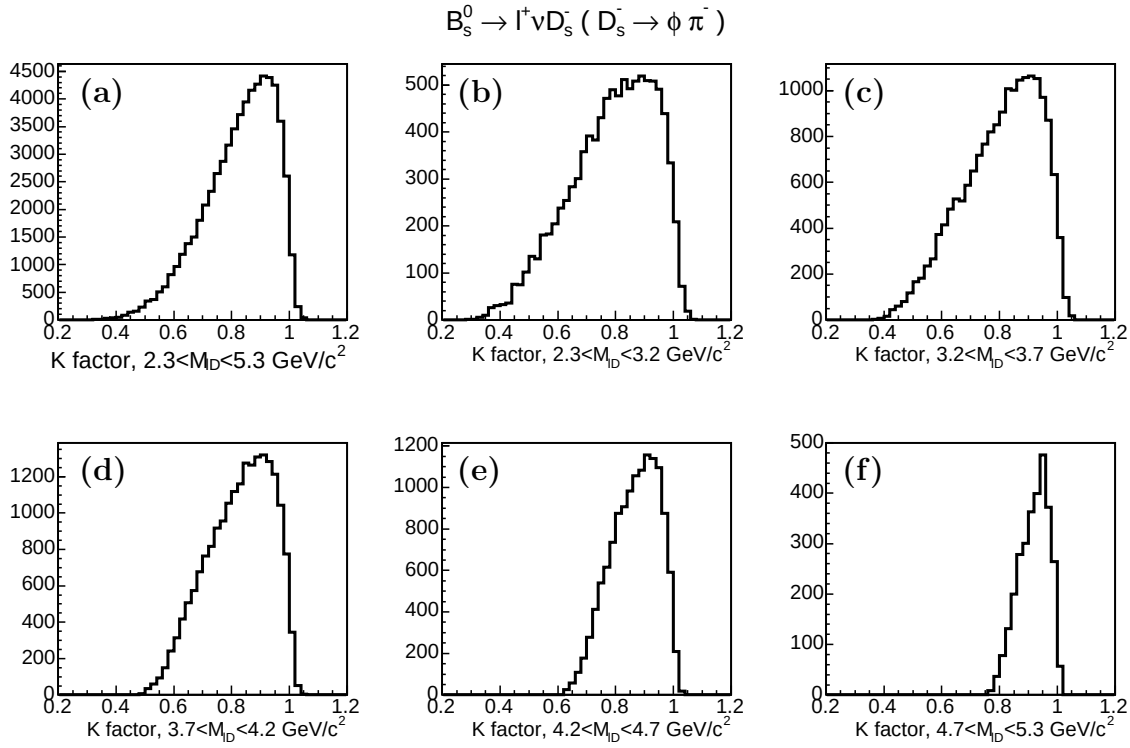

Figure 68: K-factor distributions in different $m(D_s^-\ell^+)$ ranges for the $D_s^- \to \phi\pi^-$ decay mode obtained from a realistic Monte Carlo $D_s^- \to \phi\pi^-$ sample: (a) $2.3 < m < 5.3$ GeV/c^2, (b) $2.3 < m(D_s^-\ell^+) < 3.1$ GeV/c^2, (c) $3.2 < m(D_s^-\ell^+) < 3.7$ GeV/c^2, (d) $3.7 < m(D_s^-\ell^+) < 4.2$ GeV/c^2, (e) $4.2 < m(D_s^-\ell^+) < 4.7$ GeV/c^2, (f) $4.7 < m(D_s^-\ell^+) < 5.3$ GeV/c^2.

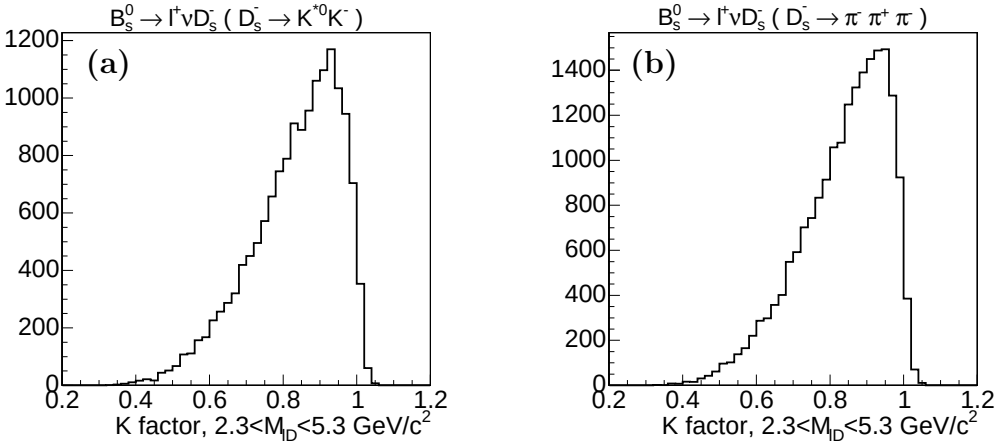

Figure 69: K-factor distributions for (a) the $D_s^- \to K^{*0}K^-$ and (b) the $D_s^- \to \pi^+\pi^-\pi^-$ decay modes obtained from realistic Monte Carlo samples.

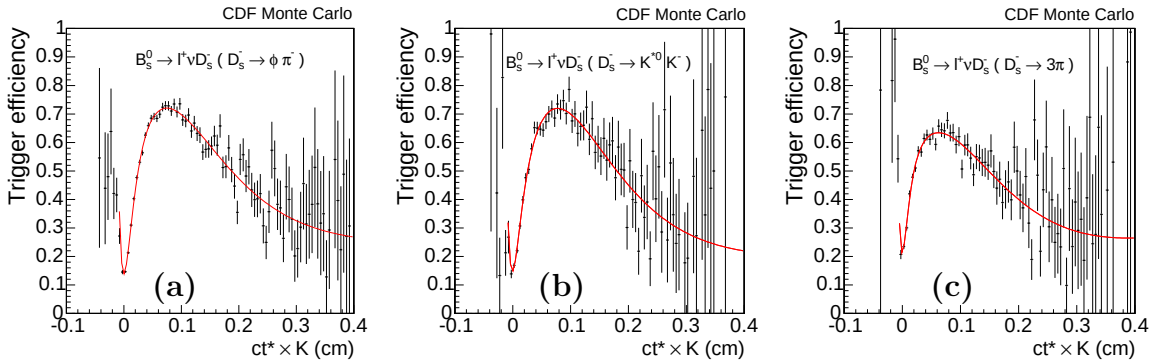

Figure 70: SVT efficiency bias curve $\epsilon(Kt^*)$ with fit parameterization for (a) $D_s^- \to \phi\pi^-$ (b) $D_s^- \to K^{*0}K^-$ and (c) $D_s^- \to \pi^+\pi^-\pi^-$ decay channels.

5.4 Biases on the B_s^0 Lifetime

As discussed in Section 4.5 in the $D^{0/+}\ell$ case, due to the track displacement requirement of the SVT trigger in the lepton-SVT data set and also due to various analysis displacement requirements described in Tables 31 and 32, the reconstructed proper decay time distribution of B_s^0 mesons is biased. This bias expressed as an efficiency curve $\mathcal{E}(t)$ is obtained as explained in Section 4.5. except that in this case, the exponential corresponds to the PDG B_s^0 lifetime of 438 μm. Again, the bias curves are fit to the same functional form:

$$\mathcal{E}(t) = (p_0 + p_1 t + p_2 t^2) \times (f_1 e^{-\lambda_1 t} + f_2 e^{-\lambda_2 t}) + p_{\text{offset}}. \tag{112}$$

The SVT bias curves $\mathcal{E}(t)$ for the $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^+\pi^-\pi^-$ decay modes are shown in Figure 70.

5.5 Maximum Likelihood Fit

To search for the $B_s^0\bar{B}_s^0$ oscillation frequency Δm_s we use an unbinned maximum likelihood method similar to the one described in Section 4.6. The likelihood function $\mathcal{L}$ is very similar to the one presented in Section 4.6. The difference is that in this case we do not have the sample composition problem but, on the other hand we have to deal with the physics backgrounds discussed in detail in Section 5.6. The likelihood function is described by the following equation:

$$\begin{aligned}
\mathcal{L} = \; & f_{\text{sig}} \quad (1 - f_{\text{pmt}} - f_{B_s}^{\text{bkg}} - f_{B^+}^{\text{bkg}} - f_{B^0}^{\text{bkg}}) L_{\text{sig}}^{t^*}(t^*, \sigma_{t^*}, \mathcal{D}) L_{\text{sig}}^{m}(m) L_{\text{sig}}^{\sigma_{t^*}}(\sigma_{t^*}) L_{\text{sig}}^{\mathcal{D}}(\mathcal{D}_{\text{pred}}) \\
& + \quad f_{\text{sig}}\, f_{\text{pmt}} L_{\text{pmt}}^{t^*}(t^*, \mathcal{D}_{\text{pmt}}) L_{\text{sig}}^{m}(m) L_{\text{sig}}^{\sigma_{t^*}}(\sigma_{t^*}) L_{\text{sig}}^{\mathcal{D}}(\mathcal{D}_{\text{pred}}) \\
& + \quad f_{\text{sig}}\, f_{B_s}^{\text{bkg}} L_{B_s}^{t^*}(t^*, \mathcal{D}) L_{\text{sig}}^{m}(m) L_{\text{sig}}^{\sigma_{t^*}}(\sigma_{t^*}) L_{\text{sig}}^{\mathcal{D}}(\mathcal{D}_{\text{pred}}) \\
& + \quad f_{\text{sig}}\, f_{B^+}^{\text{bkg}} L_{B^+}^{t^*}(t^*, \mathcal{D}) L_{\text{sig}}^{m}(m) L_{\text{sig}}^{\sigma_{t^*}}(\sigma_{t^*}) L_{\text{sig}}^{\mathcal{D}}(\mathcal{D}_{\text{pred}}) \\
& + \quad f_{\text{sig}}\, f_{B^0}^{\text{bkg}} L_{B^0}^{t^*}(t^*, \mathcal{D}) L_{\text{sig}}^{m}(m) L_{\text{sig}}^{\sigma_{t^*}}(\sigma_{t^*}) L_{\text{sig}}^{\mathcal{D}}(\mathcal{D}_{\text{pred}}) \\
& + \quad (1 - f_{\text{sig}})\, L_{\text{bkg}}^{t^*}(t^*, \mathcal{D}_{\text{bkg}}) L_{\text{bkg}}^{m}(m) L_{\text{bkg}}^{\sigma_{t^*}}(\sigma_{t^*}) L_{\text{bkg}}^{\mathcal{D}}(\mathcal{D}_{\text{pred}}),
\end{aligned} \tag{113}$$

where $f_{\rm sig}$ is the fraction of the D_s^- signal containing the real semileptonic $B_s^0 \to D_s^-\ell^+\nu$ decays as well as prompt and physics backgrounds contributing to the $D_s^-\ell^+$ final state. $f_{B_s}^{\rm bkg}$, $f_{B^+}^{\rm bkg}$, and $f_{B^0}^{\rm bkg}$ are the fractions of physics backgrounds from the various decays, $B_s^0/B^{0/+} \to D_s^- D_{(s)}$ and $B_s^0 \to D_s^-\tau X$, that produce the same signature as the signal decay $B_s^0 \to D_s^-\ell^+$. The determination of these fractions and the corresponding probability distribution functions of the pseudo proper decay time t^* as templates T of these backgrounds are discussed in detail in Section 5.6. The parameter $f_{\rm pmt}$ is the fraction of the prompt background. The likelihood function for each of the three decay modes considered in this analysis contains a prompt background template parameterized by a single Gaussian with a width of ~ 95 μm. In the $D_s^- \to \phi\pi^-$ mode, we observe a small mass peak in the wrong-sign $D_s^-\ell^-$ sample as shown in Figure 61. Assuming that the prompt background events populate evenly the right and wrong-sign $D_s^-\ell^+$ samples, we assign a prompt background fraction of 3.2%. We scale this number up by a factor of 1.5 which is the ratio between the combinatorial background levels in right- and wrong-sign $D_s\ell$ samples. No wrong-sign mass peak is seen in the $D_s^- \to K^{*0}K^-$ and $D_s^- \to \pi^-\pi^+\pi^-$ decay modes. However, we still see slight evidence of the presence of the prompt background in the corresponding ct^* distributions. In our default fitter, we assume that the prompt background fraction in each decay mode is the same as in $D_s^- \to \phi\pi^-$. The effect of the prompt background on the B_s^0 lifetime measurement and the amplitude scan is discussed in the evaluation of systematic uncertainties in Section 5.10.

The term $L_{sig}^{t^*}(t^*, \sigma_{t^*}, \mathcal{D})$ is the probability distribution function of the pseudo proper decay time t^* describing the signal events. It is similar to the corresponding Δm_d likelihood function in Equation (80):

$$L_{sig}^{t^*}(t^*, \sigma_{t^*}, \mathcal{D}) = \int_{K_{min}}^{K^{max}} dK\ \mathcal{P}'_{\rm SS}(Kt^*)K\,\mathcal{F}(K), \tag{114}$$

where

$$\mathcal{P}'_{\rm SS}(t') = \mathcal{N}\int_0^\infty \frac{1}{2}\,\Gamma\,{\rm e}^{-\Gamma t}\,(1 - \mathcal{A}\,\mathcal{S}_{\mathcal{D}}\,\mathcal{D}\cos\Delta m_s\,t)\mathcal{G}(t'-t, s_\sigma\sigma)dt \times \mathcal{E}(t'). \tag{115}$$

with $\mathcal{N}$ given by Equation (82).

The terms $L_{\rm sig}^m(m)$ and $L_{\rm bkg}^m(m)$ represent the probability distribution functions for the signal and background D_s^- meson mass distributions expressed by a double Gaussian for the signal and by a linear function for the background. $L_{sig/bkg}^{\mathcal{D}}$ and $L_{sig/bkg}^{\sigma_{t^*}}$ are the probability distribution functions for the predicted dilution and proper time resolution for signal/background events. These probability functions are obtained from the normalized distribution shown in Figures 62 and 64. As shown in Reference [56], quantities that enter the likelihood function on an event-by-event basis exhibiting different distributions for signal and background events may introduce biases in the maximum likelihood procedure. To avoid such biases, we use separate templates for signal and background distributions.

The term $L_{\rm bkg}^{t^*}(t^*, \mathcal{D}_{\rm bkg})$ corresponds to the probability distribution function of the pseudo proper decay time t^* for combinatorial background events from the D_s^- mass side-bands. It is parameterized as in Equation (86).

The term $L^{t^*}_{pmt}(t^*, \mathcal{D}_{pmt})$ represents the corresponding probability distribution function of t^* for the prompt background events and has a parameterization similar to the combinatorial background. The right-hand side of Figure 67 shows the same $ct^*(B_s)$ distribution for the wrong-sign $D_s^-\ell^-$ combinations separated into events with same-sign and opposite-sign tags. Within the limited statistics of these distributions, we do not observe a significant asymmetry. Thus, for the default analysis, we assume that the prompt background has no asymmetry $\mathcal{D}_{\rm pmt} = 0$. The mass, dilution and decay time resolution templates for the prompt background are taken to be the same as the signal. Since our knowledge of the prompt background is limited by the statistics of the $D_s\ell$ wrong-sign sample, the above assumptions are valid with limited certainty and evaluated as systematic uncertainties in Section 5.10.

The mass, dilution and pseudo proper decay time resolution terms for the physics background events are described by the same probability distribution functions as the signal events. The t^* templates are given by:

$$L^{t^*}_{B_s}(t^*, \mathcal{D}) = \frac{1}{2} T_{B_s}(t^*), \qquad L^{t^*}_{B^+}(t^*, \mathcal{D}) = \frac{1-\mathcal{D}}{2} T_{B^+}(t^*) \tag{116}$$

$$L^{t^*}_{B^0}(t^*, \mathcal{D}) = \frac{1-\mathcal{D}}{2} T^{\rm unmix}_{B^0}(t^*) + \frac{1+\mathcal{D}}{2} T^{\rm mix}_{B^0}(t^*). \tag{117}$$

The templates T_{B_s}, T_{B^+} and T_{B^0} are described in Section 5.6. It is clear that B^+ and B^0 events that produce a $D_s^-\ell^+$ signature must be treated as background events. In the case of B_s^0 events where $B_s^0 \to D_s\ \tau/D/D_s$ and $\tau/D/D_s \to e/\mu\ X$, we treat them as background although they are real B_s^0 events. Since the electron or muon does not come directly from the B_s^0 decay, but from a long lived B_s^0 daughter, the intersection between the reconstructed D_s and the electron/muon trajectories will be displaced with respect to the real B_s^0 decay point and would give the wrong proper decay time. In principle, one could still treat these events as signal and add extra displacement/smearing to the B_s^0 decay point in the likelihood function. However, since such events represent only about 5% of the total B_s^0 signal and the gain in sensitivity would be minute due to degraded decay time resolution, we choose to conservatively treat these events as background.

5.6 Physics Backgrounds

The maximum likelihood fitting method employed in this analysis takes into account combinatorial backgrounds by fitting the side-band regions and fixing the fit parameters as the background contribution to the signal region. However, there are additional backgrounds which result in a real D_s^- meson and a real lepton ℓ^+ but do not come from the semileptonic $B_s^0 \to D_s^-\ell^+\nu$ decay. These backgrounds contribute to the signal, but not to the side-band sample. Their contributions to the lifetime distribution must be determined and incorporated in the maximum likelihood fit. Simulation samples of these background processes allow the determination of their ct^* distributions and their reconstruction efficiencies relative to that of the $B_s^0 \to D_s^-\ell^+\nu$ decays. The background fractions expected in the data are calculated using branching

ratios from the PDG [1] unless stated otherwise. There are five such background components.

The first background contribution is from B^- meson decays to $D_s^{(*)-}D^{(*)0/+}X$, with the $D^{(*)0/+}$ decaying semileptonically with a positively charged lepton ℓ^+ in the final state. The D_s^{*-} meson always decays to a D_s^- meson. This background therefore results in a real D_s^- meson and a real lepton of opposite charge ℓ^+. Similarly, the second source of backgrounds are decays of the $\bar{B}^0$ meson to $D_s^{(*)-}D^{(*)+/0}X$ with semileptonic $D^{(*)+/0}$ decays result in the same signature. The efficiency for these decays to pass the trigger is expected to be lower than for $B_s^0 \to D_s^-\ell^+\nu$ events since the momentum spectrum of the lepton in the $D^{(*)0/+}$ decay is softer. A realistic Monte Carlo simulation sample of B^- and $\bar{B}^0$ decays is generated using the program `HeavyQuarkGen`, decayed using the program `EvtGen` and are then passed through the full detector and lepton-SVT trigger simulation. The $D_s^- \to \phi\pi^-$ and $\phi \to K^+K^-$ decays are forced. The decays of D^0 and D^+ are open. The events which pass the trigger simulation are then passed through the $B_s^0 \to D_s^-\ell^+\nu$ analysis chain. The efficiency to reconstruct a $D_s^-\ell^+$ combination in the B^- events relative to the efficiency for reconstructing a $D_s^-\ell^+$ in $B_s^0 \to D_s^-\ell^+\nu$ decays is found to be:

$$\epsilon^{\rm rel}_{B^-\to D_sD} = \frac{\epsilon_{D_sD}(B^- \to {D_s}^{(*)-}D^{(*)0/+}X,\ D^{(*)0/+} \to \ell^+X)}{\epsilon_{D_s\ell}(B_s^0 \to D_s^-\ell^+\nu)} = (1.17 \pm 0.02)\ \%. \quad (118)$$

The corresponding efficiency to reconstruct a $D_s^-\ell^+$ combination in $\bar{B}^0$ events relative to the efficiency for reconstructing $D_s^-\ell^+$ in $B_s^0 \to D_s^-\ell^+\nu$ decays is:

$$\epsilon^{\rm rel}_{\bar{B}^0\to D_sD} = \frac{\epsilon_{D_sD}(\bar{B}^0 \to {D_s}^{(*)-}D^{(*)+/0}X,\ D^{(*)0/+} \to \ell^+X)}{\epsilon_{D_s\ell}(B_s^0 \to D_s^-\ell^+\nu)} = (1.39 \pm 0.02)\ \%. \quad (119)$$

The fraction of false signal events expected from this irreducible background relative to true signal events is given by:

$$f_{B^{-/0}\to D_sD} = \epsilon^{\rm rel}_{B^{-/0}\to D_sD}\frac{f_{u/d}}{f_s}\frac{BR(B \to D_s^-X)}{BR(B_s^0 \to D_s^-\ell^+\nu)}, \quad (120)$$

where $f_u = f_d = (39.7 \pm 1.0)\%$ and $f_s = (10.7 \pm 1.1)\%$ are the B fragmentation fractions which account for the relative production fractions of the B^-, B^0 and B_s^0 mesons [1]. The world average measurement of $BR(B \to D_s^-X) = (10.5\pm2.6)\%$ is for the B^0/B^+ admixture and the branching ratio for $B_s^0 \to D_s^-\ell^+\nu$ is $(7.9\pm2.4)\%$ [1]. The resulting contributing fraction from B^- is:

$$f_{B^-\to D_sD} = (5.77 \pm 2.34)\% \quad (121)$$

and for B^0 it is:

$$f_{B^0\to D_sD} = (6.85 \pm 2.78)\%. \quad (122)$$

The third background contribution is from B_s^0 meson decays to ${D_s}^{(*)-}{D_s}^{(*)+}X$ where one D_s decays semileptonically and the other decays to $\phi\pi$. Similarly to the B^+/B^0 case, simulation samples of $B_s^0 \to D_s^{(*)-}D_s^{(*)+}X$ events are generated. The

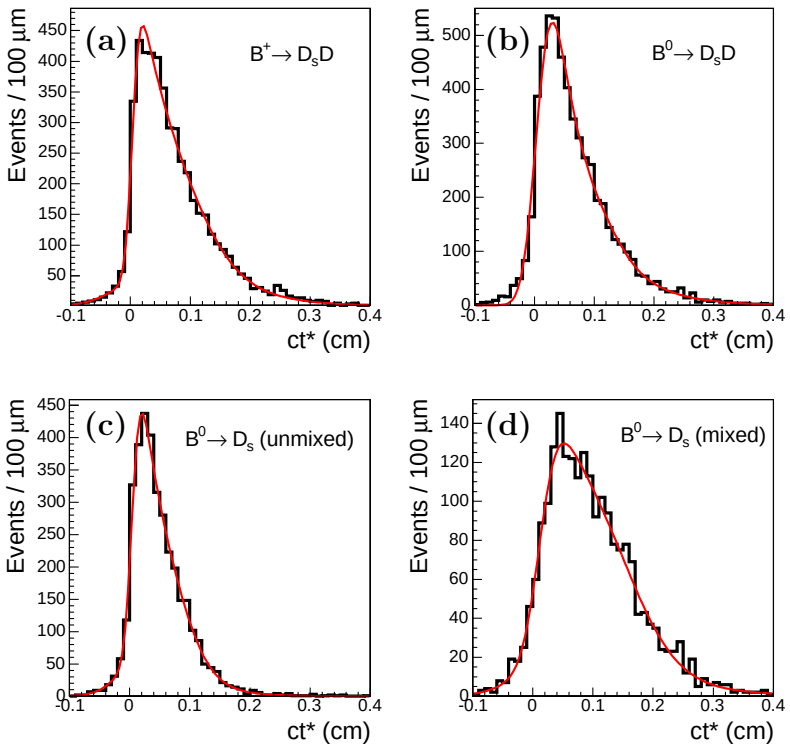

Figure 71: Lifetime templates of (a) $B^0 \to D^- D_s^+$ and (b) $B^+ \to D_s^+ \bar{D}^0$ physics backgrounds to B_s^0 semileptonic decays. The B^0 template is split up in (c) an unmixed and (d) a mixed component.

efficiency for reconstructing D_s^- mesons in this sample relative to that in the $B_s^0 \to D_s^- \ell^+ \nu$ data is:

$$\epsilon^{\rm rel}_{D_s D_s} = \frac{\epsilon_{D_s D_s}(B_s \to D_s{}^{(*)-} D_s{}^{(*)+} X)}{\epsilon_{D_s \ell}(B_s^0 \to D_s^- \ell^+ \nu)} = (27.86 \pm 0.47)\ \%. \tag{123}$$

The fraction of false signal events which are expected from this irreducible background to true signal events is given by:

$$f_{D_s D_s} = \epsilon^{\rm rel}_{D_s D_s} \frac{BR(B_s \to D_s{}^{(*)-} D_s{}^{(*)+} X) \times BR(D_s \to \ell X) \times 2}{BR(B_s^0 \to D_s^- \ell^+ \nu)} \tag{124}$$

where the factor of two accounts for the two possible combinations $D_s^+ D_s^- \to \ell^+ \phi \pi^-(X)$ $/ \phi \pi^+ \ell^-(X)$. The world average measurement of the branching ratio for $B_s \to D_s{}^{(*)-} D_s{}^{(*)+} X$ is quite inaccurately measured and thus the `EvtGen` model prediction of 4.72% is used. Similarly, the measured semileptonic branching ratio of $D_s^- \to \ell^- X$ is measured to be $8^{+6}_{-5}\%$. Therefore a prediction of 6.32% is made in `EvtGen` which is compatible with the assumtion that the partial widths of the D^0 and D^+ decays to semileptonic final states are the same as compared to D_s^- decays as predicted by the spectator model. This results in a contribution to the sample of:

$$f_{B_s^0 \to D_s D_s} = (2.10 \pm 0.64)\ \%. \tag{125}$$

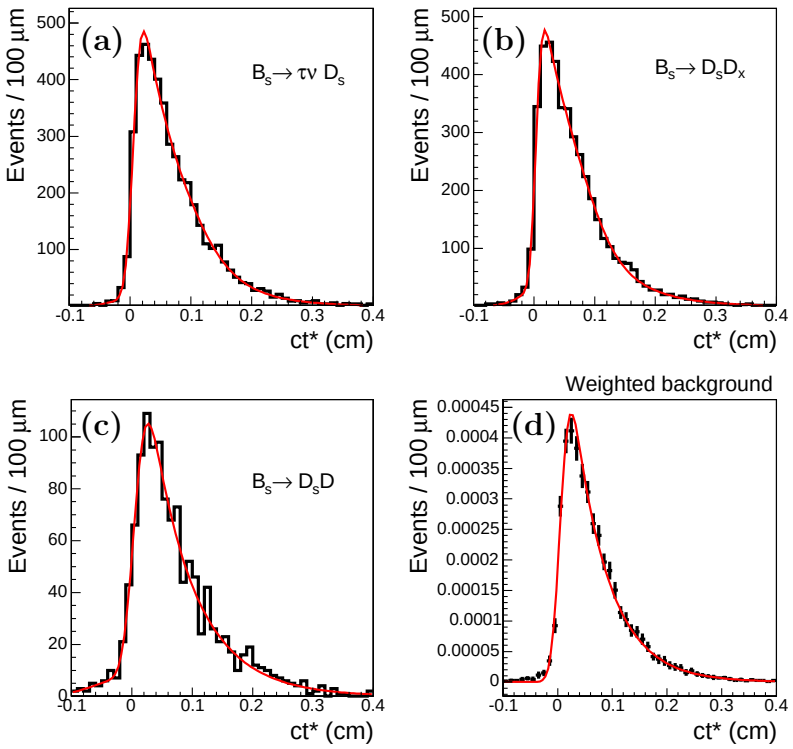

Figure 72: Lifetime templates of (a) $B_s^0 \to D_s^- \tau^+ \nu$, (b) $B_s^0 \to D_s^+ D_s^-$ and (c) $B_s^0 \to D_s^+ D^-$ physics backgrounds. The lifetime template corresponding to the weighted B_s^0 backgrounds (d) is obtained by summing the templates (a), (b) and (c) with the corresponding contributing fractions.

The fourth background contribution is also a decay of the B_s^0 meson, in this case $B_s^0 \to {D_s}^{(*)-} D^{(*)+/0} X$. Simulation samples are generated for this decay mode as for the previous cases. The D^0 and D^+ are not forced to decay semileptonically. The relative number of reconstructed D_s^- mesons in this sample relative to the $B_s^0 \to D_s^- \ell^+ \nu$ simulation sample is:

$$\epsilon^{\rm rel}_{B_s^0 \to D_s D} = \frac{\epsilon_{B_s^0 \to D_s D}(B_s^0 \to {D_s}^{(*)-} D^{(*)+/0} X)}{\epsilon_{D_s \ell}(B_s^0 \to D_s^- \ell^+ \nu)} = (0.76 \pm 0.02)\ \%. \tag{126}$$

The fraction of these events expected in the data sample is given by:

$$f_{B_s^0 \to D_s D} = \epsilon^{\rm rel}_{B_s^0 \to D_s D} \frac{BR(B_s^0 \to {D_s}^{(*)-} D^{(*)+/0} X)}{BR(B_s^0 \to D_s^- \ell^+ \nu)}. \tag{127}$$

Using the `EvtGen` prediction $BR(B_s^0 \to D_s D) = 15.4\%$, the resulting fraction is:

$$f_{B_s^0 \to D_s D} = (1.48 \pm 0.45)\ \%. \tag{128}$$

The fifth and final background is from semileptonic decays of the B_s^0 which include a τ lepton. For this decay the relative efficiency is.

$$\epsilon^{\rm rel}_{B_s \to D_s \tau} = (29.29 \pm 0.48)\ \%. \tag{129}$$

The fraction of false signal originating from this source of physics background is:

$$f_{B_s \to D_s^- \tau^+ \nu} = \epsilon^{\rm rel}_{B_s \to D_s \tau} \frac{BR(B_s^0 \to D_s^- \tau^+ X) \times BR(\tau \to \ell X)}{BR(B_s^0 \to D_s^- \ell^+ \nu)}. \quad (130)$$

Using the `EvtGen` prediction $BR(B_s^0 \to D_s^- \tau^+ X) = 2.92\%$ and $BR(\tau \to \ell) = 17.36 \pm 0.06\%$ we obtain:

$$f_{B_s \to D_s^- \tau^+ \nu} = (1.88 \pm 0.57)\ \%. \quad (131)$$

The contributions from these backgrounds are included in the likelihood fit (see Equation (113)) by adding extra terms to the signal likelihood function. These extra terms consist of a relative fraction, discussed above, and a lifetime template T for each irreducible background contribution. These are obtained from fits to the ct^* distributions given by the simulation samples. These ct^* distributions are shown in Figures 71 and 72.

The contributions of the physics backgrounds to the $K^{*0}K^-$ and $\pi^+\pi^-\pi^-$ channels are estimated in the same way as for $\phi\pi^-$ channel. A summary of these fractions for each of the three decay modes is given in Table 38.

	D_s^- Decay Channel		
Background Mode	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to \pi^+\pi^-\pi^-$
$B^0 \to D_s D$ (%)	6.85 ± 2.78	6.01 ± 2.44	5.86 ± 2.31
$B^+ \to D_s D$ (%)	5.77 ± 2.34	5.28 ± 2.15	5.18 ± 2.11
$B_s \to D_s D$ (%)	1.48 ± 0.45	1.15 ± 0.07	1.03 ± 0.06
$B_s \to D_s D_s$ (%)	2.10 ± 0.64	1.97 ± 0.06	2.01 ± 0.07
$B_s \to D_s \tau$ (%)	1.88 ± 0.57	1.66 ± 0.04	1.64 ± 0.05

Table 38: Fractions of the physics backgrounds contributing to the $B_s^0 \to D_s^- \ell^+ \nu$ signal.

5.7 Tests of the Unbinned Likelihood Fitting Framework

To test the unbinned maximum likelihood fitting framework, we first perform a fit for the B_s^0 lifetime using a Monte Carlo sample of $B_s^0 \to D_s^- \ell^+ \nu$, $D_s \to \phi\pi^-$ decays. This sample, of $\sim 60,000$ events, is also used to generate the K-factor distributions and SVT-bias curves. The fit result is $c\tau(B_s^0) = (439.1 \pm 1.9)\ \mu$m in good agreement with the world average input value of 438 μm used to generate this Monte Carlo sample. The fit probability is 9.2% corresponding to $\chi^2/dof = 106.0/88$. Figure 73 shows the lifetime fit projection. We repeat the lifetime fit using K-factor distributions binned in $m(D_s^- \ell^+)$, we find $c\tau(B_s) = (438.5 \pm 1.9)\mu$m.

In addition, we study the maximum likelihood fitter using toy Monte Carlo samples. The study is done for the $D_s^- \to \phi\pi^-$ channel. Each sample corresponds to one pseudo-experiment. We generate 1000 pseudo-experiments with a total of 50,000 events each. These samples contain the expected signal to background fractions

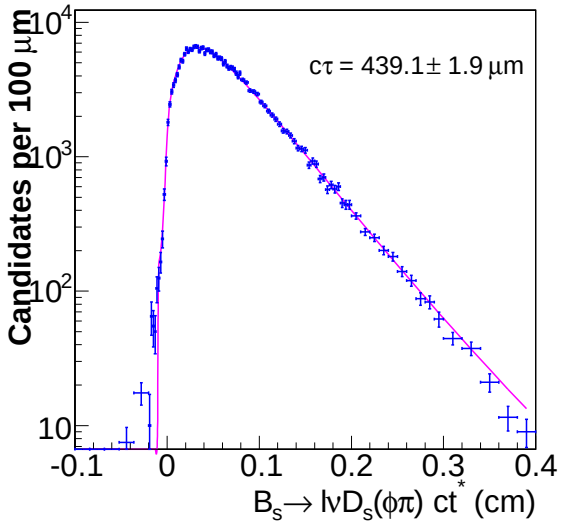

Figure 73: B_s^0 lifetime fit projection using a Monte Carlo sample of $B_s^0 \to D_s^- \ell^+ \nu$, $D_s^- \to \phi\pi^-$ decays.

measured in data. Prompt and physics backgrounds are also included with the corresponding fractions and ct^* distributions observed in data or realistic Monte Carlo simulations. The lifetime distribution for signal events is obtained in a few steps. First we generate the decay times according to a decaying exponential corresponding to a lifetime $c\tau(B_s^0) = 438$ μm. Then, a Gaussian resolution function is used to smear the decay time. The width of the Gaussian has the distribution from Figure 64. The events are then selected according to the SVT trigger efficiencies discussed in Section 5.4 and shown in Figure 70. The remaining events are finally smeared using the K-factor distributions shown in Figure 68. For background events we use the lifetime distribution of the D_s side-band events from data. To generate the decay time uncertainties for signal and background events, we use separate distributions as shown in Figure 64. The mass distributions for signal and background events are generated according to a Gaussian and a linear background function, respectively. The B_s^0 oscillation frequency is fixed at $\Delta m_s^{input} = 5$ ps^{-1}. We employ realistic tagging with a combined $\varepsilon\mathcal{D}^2$ adding up to $\sim 1.5\%$. To achieve this we use tagging efficiencies similar to the ones measured in data: $\epsilon_{SMT} \sim 4\%$, $\epsilon_{SET} \sim 3\%$, $\epsilon_{JQT(vertex)} \sim 9\%$, $\epsilon_{JQT(track)} \sim 10\%$ and $\epsilon_{JQT(highp_T)} \sim 48\%$. The predicted dilutions are generated separately for signal and background events using the distributions shown in Figure 62. For each event, we generate four quantities: the pseudo proper decay time and the decay time error of the B_s^0 meson, the mass of the D_s meson and the predicted dilution. The result of the toy Monte Carlo studies in which we fit for Δm_s are shown in Figure 74 which shows the pull distribution for Δm_s given by $(\Delta m_s - \Delta m_s^{input})/\sigma_{\Delta m_s}$. The pull distribution has a mean of 0.05 ± 0.03 and a width of $\sigma = 1.09 \pm 0.03$. Figure 74 also shows the corresponding results of the toy Monte Carlo studies when fitting for the B_s^0 lifetime and the amplitude $\mathcal{A}$ with the B_s^0 oscillation frequency fixed at $\Delta m_s = 5$ ps^{-1}. The mean and width of the pull distributions are 0.13 ± 0.03 and

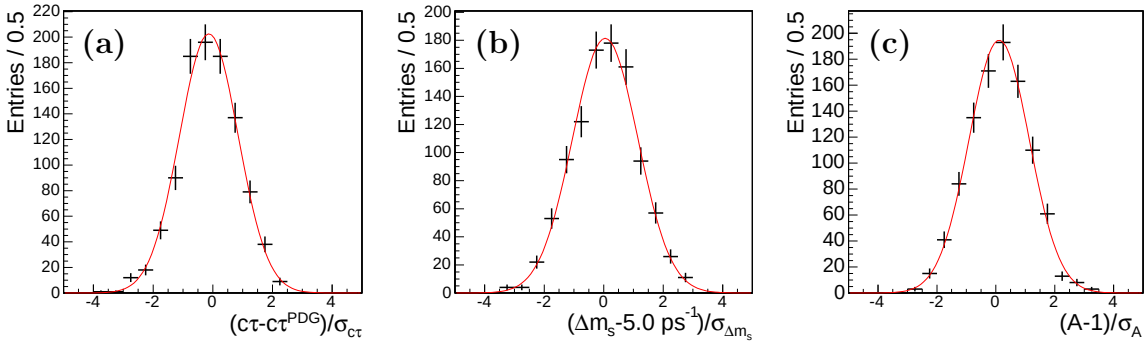

Figure 74: Results of the toy Monte Carlo pseudo-experiments. Pull distributions for (a) B_s^0 lifetime, (b) B_s^0 oscillation frequency fixed at $\Delta m_s = 5$ ps^{-1}, and (c) amplitude $\mathcal{A}$.

Mode	Fit result $c\tau(B_s^0)$
$D_s^- \to \phi\pi^-$	455.0 ± 11.9 μm
$D_s^- \to K^{*0}K^-$	413.8 ± 20.1 μm
$D_s^- \to \pi^+\pi^-\pi^-$	422.6 ± 25.7 μm
Combined	442.7 ± 9.5 μm
PDG	438 ± 17 μm

Table 39: B_s^0 lifetime fit results.

$\sigma = 0.99 \pm 0.02$ for the B_s^0 lifetime fits and 0.11 ± 0.03 and $\sigma = 1.02 \pm 0.03$ for the amplitude fits. The pull distributions are in good agreement with expectations and indicate that the maximum likelihood fitter is unbiased.

5.8 B_s^0 Lifetime Measurement

As an additional check of the maximum likelihood fitting procedure, we measure the lifetime of the B_s^0 meson. This is also an important step toward a time dependent Δm_s measurement. In the likelihood fit, all parameters are fixed except the B_s^0 lifetime. Table 39 shows the result of the B_s^0 lifetime fits for the three individual D_s^- decay channels and for all three decay modes combined. Figure 75 shows the B_s^0 lifetime fit projection for each of the three decay modes.

5.8.1 Systematic Uncertainties on the B_s^0 Lifetime Measurement

We study the main sources of systematic uncertainties on the B_s^0 lifetime measurement. They are listed below and summarized in Table 40.

Prompt Background Fraction:
The largest systematic uncertainty on the B_s^0 lifetime measurement originates from the limited knowledge of the prompt background fraction discussed in Section 5.3. To understand the effect of the prompt background on the B_s^0 lifetime result, we totally remove the prompt background component from the fit. We assign a systematic error

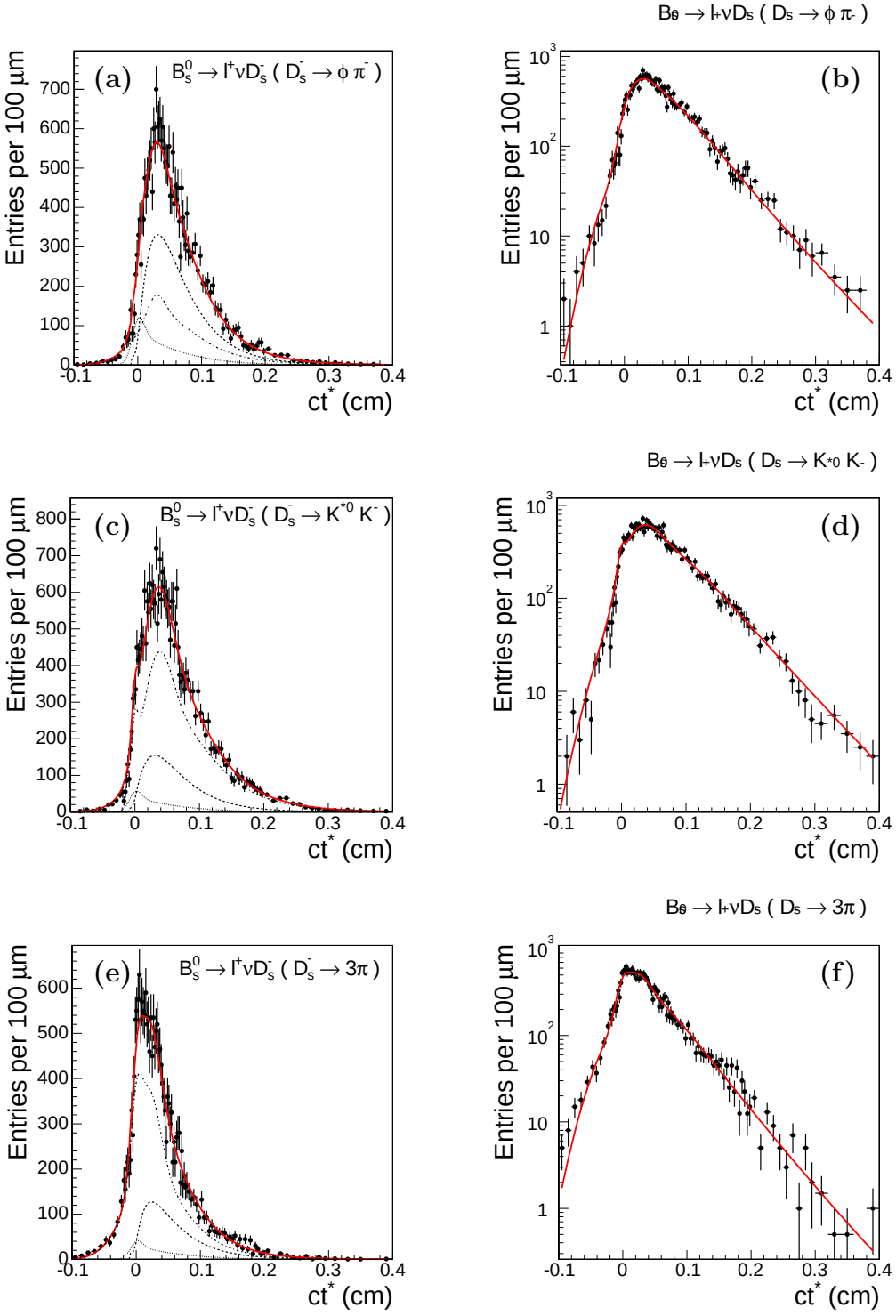

Figure 75: B_s^0 lifetime fit projection for $D_s^- \to \phi\pi^-$ mode on (a) linear and (b) logarithmic scales, for $D_s^- \to K^{*0}K^-$ mode on (c) linear and (d) logarithmic scales and for $D_s^- \to \pi^+\pi^-\pi^-$ mode on (e) linear and (f) logarithmic scales. The points with error bars represent data and the solid lines are the fit projections. The signal (dashed), combinatorial background (dot-dashed) and the sum of the prompt and physics backgrounds (dotted) are also shown on the linear scale plots.

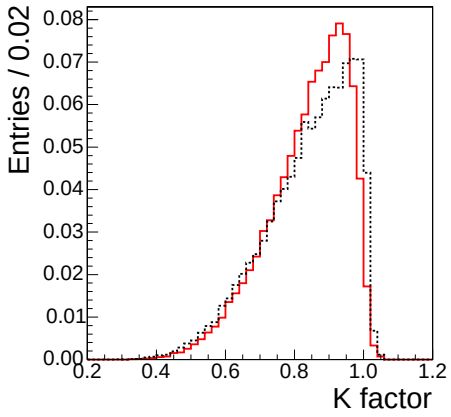

Figure 76: K-factor distributions for semileptonic B_s^0 decays directly to $D_s^- \ell^+$ (dotted) and into D_s^{*-} and D_s^{**-} states only (solid).

of 50% of the difference between the new lifetime value and the one from the original fit.

Physics Backgrounds:
The second largest systematic error originates from the various physics backgrounds discussed in Sec. 5.6. We evaluate the systematic errors by changing all fractions of physics backgrounds at the same time by $+1\sigma$ and -1σ from the central values shown in Table 38. We assign the systematic error as the change in the lifetime with respect to the default fit value.

SVT Trigger Efficiency:
Another source of systematic uncertainties originates from the SVT trigger bias curve. To estimate these uncertainties we use the method discussed in Section 4.9.4.

Sample Composition:
The fractions of semileptonic B_s^0 decays into D_s^-, D_s^{*-} and D_s^{**-} states are not measured. The fractions used in this analysis are the default values from the `EvtGen` program and are based on theoretical estimations. To study the effect of the $D_s^{(*/**)-}$ mixture on the B_s^0 lifetime, we repeat the fit with two different K-factor distributions and two different SVT bias curves: one in which the B_s^0 decays directly to D_s^- ($B_s^0 \to D_s^- \ell^+ \nu$) and another one in which the B_s^0 decays semileptonically into D_s^{*-} and D_s^{**-} states only. Figure 76 shows the K-factor distributions for B_s^0 decays directly to D_s^- and into D_s^{*-} and D_s^{**-} states only. We find the difference between the two K-factor distributions to be small.

Resolution Scale Factor:
We vary the decay time resolution scale factor by $\pm 5\%$ and find a small systematic error associated with this effect.

Event-by-Event Primary Vertex:
Finally, we study possible systematic effects associated with the event-by-event primary vertex determination by smearing the B_s^0 decay length in the Monte Carlo simulation with a Gaussian of 21 μm width as suggested by studies described in Reference [48]. The SVT efficiency curves are re-evaluated and used to assign systematic

uncertainties. We evaluated a third set of SVT efficiency curves by smearing both the B_s^0 decay length and the distance from the D_s^- decay point to the primary vertex by the same 21 μm width Gaussian. We find that the SVT efficiency curves are negligibly affected by this smearing.

To summarize the result of the B_s^0 lifetime study from the three D_s decay modes, we obtain a combined B_s^0 lifetime:

$$c\tau(B_s^0) = (\ 442.7\ \pm\ 9.5\ (stat.)\ \pm\ 12.7\ (syst.)\ )\ \mu m \tag{132}$$

This result is in good agreement with the PDG value of (438 ± 17) μm. The total systematic errors is obtained by adding in quadrature the systematic uncertainties listed in Table 40. We note that the statistical uncertainty is about half of the world average error. However, once the systematic error is added in quadrature, the total uncertainty on the B_s^0 lifetime is just slightly better than the PDG uncertainty. With better understanding of the prompt background, which is the main source of systematic uncertainty on the B_s^0 lifetime, the CDF experiment will perform one of the most competitive B_s^0 lifetime measurements. However, for the moment, we only use this measurement as a cross check of our mixing analysis framework.

Prompt background	10.8 μm
Physics background	4.1 μm
SVT efficiency	3.9 μm
Sample composition	3.4 μm
Event-by-event vertex	1.5 μm
Scale factor	0.2 μm
Total	12.7 μm

Table 40: Summary of main systematic effects on the B_s lifetime measurement.

5.9 Amplitude Scan

Using the data from the three $B_s^0 \to D_s^- \ell^+ \nu$ samples ($D_s \to \phi\pi^-$, $D_s \to K^{*0}K^-$ and $D_s \to \pi^+\pi^-\pi^-$), we search for the B_s^0 oscillation frequency Δm_s using the amplitude method introduced in Section 1.2.4. This method is similar to a Fourier analysis for determining an oscillation frequency. In the case of B^0 mesons where we observe the actual oscillation, it is natural to measure the oscillation frequency Δm_d by using a fit in which Δm_d is a free floating parameter and the value for which the fit finds the minimum of the likelihood function will be the result of our measurement. In the case of B_s^0 mesons, the oscillation frequency Δm_s is at least ~ 30 times larger than Δm_d. This means that Δm_s may be beyond our experiment's ability to observe the oscillation. In such a case, a fit in which Δm_s is a free floating parameter will not converge. The amplitude method was specially designed for such cases to establish a lower limit on the oscillation frequency. It is also very convenient for averaging lower limits from different experiments. As discussed in Section 1.2.4, in the amplitude

method, instead of allowing Δm_s to float in the maximum likelihood fit, we fix it at different test values in the range of interest. In particular, our test values start from $\Delta m_s = 0$ ps^{-1} and go up to $\Delta m_s = 20$ ps^{-1} in steps of 0.25 ps^{-1}. We also introduce the Fourier coefficient called amplitude $\mathcal{A}$ which multiplies the oscillation term:

$$\cos(\Delta m_s t) \quad \rightarrow \quad \mathcal{A} \times \cos(\Delta m_s t). \tag{133}$$

When the test value of Δm_s is close to the true value, the fit will return a value of $\mathcal{A}$ close to one. When Δm_s is far from its true value, the returned amplitude value will be consistent with zero.

Before the actual search for Δm_s, we perform a "blind" analysis in which the tag decision is randomized by multiplying it with $(-1)^{(event\ number)}$ as suggested in Reference [62]. This procedure allows us to study the sensitivity to the B_s^0 oscillation frequency Δm_s and also to perform systematic uncertainty studies without knowing the actual lower limit on Δm_s. We find that, at this time, the sensitivity of our experiment is below the world average lower limit, so we do not expect to observe the Δm_s signal yet, but rather to set a lower limit. Once the sensitivity is established, the systematic errors are evaluated and in general, the analysis procedure is complete, we remove the randomization of the B_s^0 flavor tags to measure the actual lower limit on Δm_s.

The fit configuration which we define as the "default" setting is described as follows:

- The amplitude $\mathcal{A}$ is the only free parameter in this fit.
- The dilution scale factors are fixed to their values in Table 35.
- Different K-factor distributions are used in different ranges of the $D_s\ell$ invariant mass $m(D_s^-\ell^+)$. We observe that, as expected, this procedure enhances slightly our sensitivity to Δm_s because we treat separately events with smaller missing momentum from unseen neutrinos.
- The B_s^0 lifetime is fixed to its world average value $c\tau = 438$ μm.
- The resolution scale factor for each candidate is calculated as a function of the $D_s\ell$ isolation, pseudorapidity η, the opening ΔR between the lepton and the D_s^- candidate and the transverse momentum p_T of the lepton, following the procedure described in Section 5.3.
- The physics and prompt backgrounds are included in the default fit.

The result of the blinded amplitude scan is shown in Figure 77. The points with error bars represent the fitted amplitude $\mathcal{A}$ with statistical error σ for each value of the assumed oscillation frequency Δm_s. The yellow band indicates 1.645 σ deviation from the central value of $\mathcal{A}$. Every value of Δm_s for which the yellow band is below one is excluded at more than 95% Confidence Level. The statistical error σ multiplied by 1.645 is plotted as a dotted line. The value of Δm_s for which this curve reaches

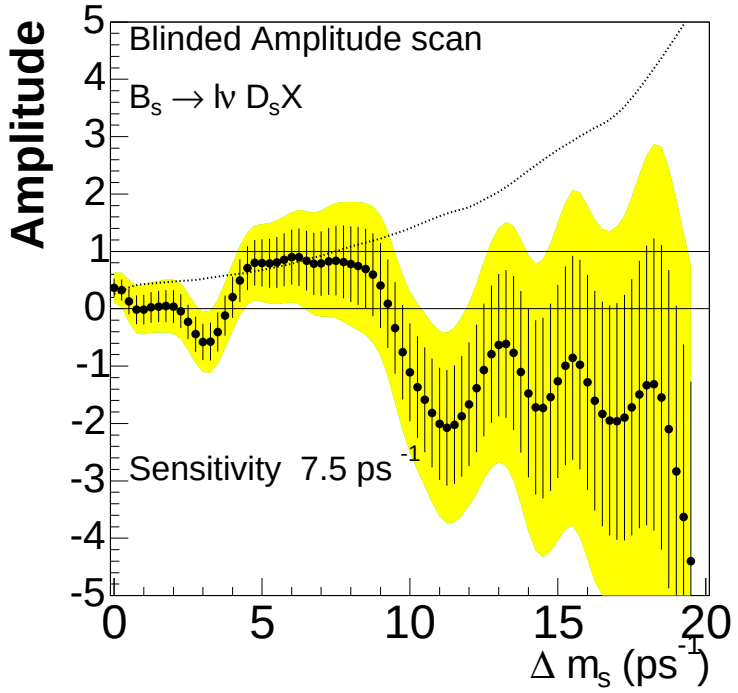

Figure 77: The blinded amplitude scan for the default fitting configuration including only the statistical errors indicates a sensitivity of 7.5 ps^{-1}.

the horizontal line at one gives the sensitivity of our measurement. For values of Δm_s above this sensitivity, the amplitude errors are so large that the amplitude is consistent with both zero and one. The value of the sensitivity depends on how statistically strong we want our final result to be. A general convention is to express the results with 95% Confidence Level. From this distribution, we find a sensitivity of 7.5 ps^{-1}. This would approximate the mean of the Δm_s lower limit distribution from an ensemble of statistically independent experiments. For one single measurement the actual lower limit on Δm_s will fluctuate around the sensitivity value.

5.10 Evaluation of Systematic Uncertainties

Reference [19] details a prescription for setting a lower limit on the B_s^0 oscillation frequency Δm_s using the amplitude scan method and also discusses a procedure for evaluating the systematic uncertainties on amplitude. The systematic uncertainty which is to be added in quadrature to the statistical error on the amplitude is given as:

$$\sigma_{sys} = \Delta\mathcal{A} + (1 - \mathcal{A})\frac{\Delta\sigma_{\mathcal{A}}}{\sigma(\mathcal{A})}. \tag{134}$$

Here $\Delta\mathcal{A}$ is obtained as the difference between the value of the amplitude from the default fit and the value of the amplitude from the fit including the corresponding systematic excursion. $\Delta\sigma(\mathcal{A})$ is the difference between the statistical errors on the

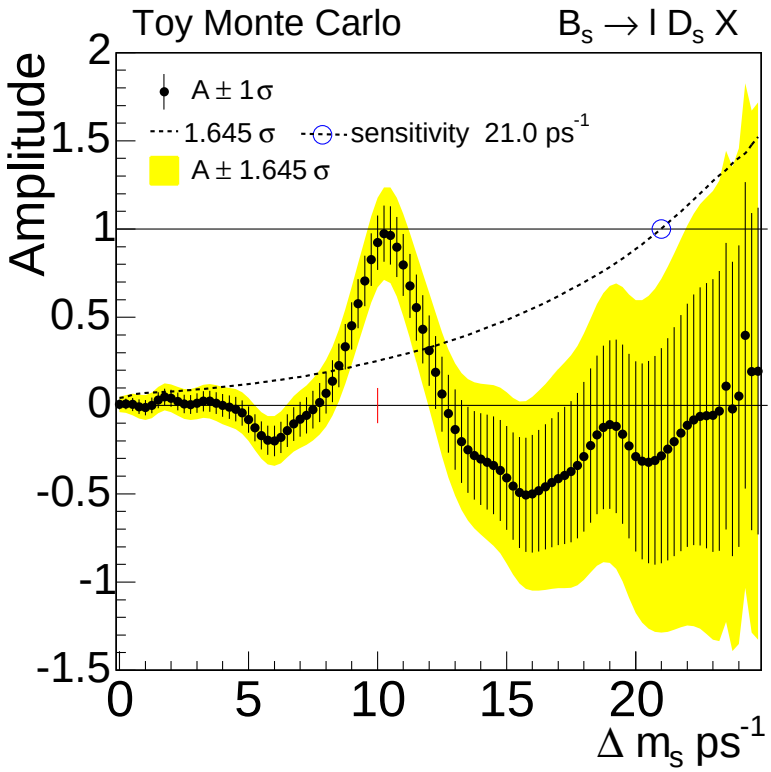

Figure 78: Result of the amplitude scan on a toy Monte Carlo sample with ∼10 times more statistics than the data. Events are generated with Δm_s=10 ps^{-1}.

amplitude value in the default fit and the fit including the systematic deviation. $\mathcal{A}$ and $\sigma(\mathcal{A})$ are defined as the amplitude value and its error from the default fit.

We use toy Monte Carlo samples to evaluate the systematic uncertainty on the amplitude scan at several Δm_s values: 0, 5, 10, 15, 20 ps^{-1}. It is necessary to use toy Monte Carlo and not the blinded data, since the blinded data contains no information on the actual B_s^0 oscillations.

A large set of toy Monte Carlo samples is required for each systematic effect considered at each value of Δm_s. We fit each Monte Carlo sample twice. First we fit the sample with the default fit settings and then we repeat the fit with the systematic effect under consideration varied in the fit configuration. A set contain about $\sim$ 500 statistically independent toy Monte Carlo samples. For each sample we calculate the systematic error:

$$\sigma^i_{sys} = (\mathcal{A}^1 - \mathcal{A}^0) + (1 - \mathcal{A}^0)\frac{\sigma^1 - \sigma^0}{\sigma^0}, \tag{135}$$

where $\mathcal{A}^0$ is the amplitude resulting from fitting the toy Monte Carlo sample in the default fit configuration and $\mathcal{A}^1$ is the amplitude obtained from the fit in which the configuration is varied according to the systematic effect considered. σ^0 and σ^1 are the corresponding statistical errors on $\mathcal{A}^0$ and $\mathcal{A}^1$. The assigned systematic uncertainty is the mean of the σ^i_{sys} distribution.

The results of this procedure are systematic uncertainties as a function of the Δm_s

value $\sigma_{\mathcal{A}}^{sys}(\Delta m_s)$. To reduce the computation time, the systematic uncertainties are evaluated for the values 0, 5, 10, 15, 20 ps^{-1} and interpolated for all other points in the scan. The systematic uncertainties are then combined by adding them in quadrature for a given Δm_s value. The combined systematic uncertainty is then folded into the exclusion limit by requiring:

$$\mathcal{A} + 1.645\sqrt{\sigma_{\mathcal{A}}^{stat}(\Delta m_s)^2 + \sigma_{\mathcal{A}}^{sys}(\Delta m_s)^2} < 1. \tag{136}$$

For this study, it is crucial to have toy Monte Carlo samples which reproduce the data very well. The toy Monte Carlo sample is fine tuned so that for similar statistics it gives similar sensitivity to what we observe in the data. In each pseudo-experiment we generate the same number of signal and background events as in data. Prompt and physics backgrounds are included. We simulate all three decay modes of the D_s meson. Events are generated as described in Section 5.7.

Figure 78 shows, as an example, the amplitude scan result of a toy Monte Carlo experiment where the oscillation frequency is set to $\Delta m_s = 10$ ps^{-1} and the number of events is, for clarity, about ten times larger than in real data.

Figure 79 shows an example of the evaluation of the systematic uncertainty on the amplitude $\mathcal{A}$ due to the uncertainty on the fraction of physics backgrounds. We run 500 pseudo-experiments. The fractions of physics backgrounds are increased by one standard deviation with respect to the default fit configuration. We use the fractions of physics backgrounds and the errors from Table 38. The amplitude and the amplitude errors are shown for the default and for the increased physics background configuration. The distribution of the systematic error σ_{sys}^i is shown as well. The assigned systematic error on amplitude is the mean of this distribution $< \sigma_{sys} >= 0.134$. The average of σ_{sys}^i is taken as the systematic uncertainty while the statistical error on the systematic uncertainty is defined as root-mean-square/$\sqrt{N}$, where N is the number of toy Monte Carlo samples. This procedure is repeated for all sources of systematic uncertainties considered and for different values of Δm_s. A summary of all systematic errors for all value of Δm_s considered is shown in Table 41.

5.10.1 Prompt Background Fraction

The prompt background seen in the wrong-sign $D_s^-\ell^-$ combinations constitutes one of the largest systematics uncertainties on the amplitude. Similar to the discussion in Section 4.4, prompt background events have the correct signature $D_s^+\ell^-$, but the lepton is either fake, or it does not come from the semileptonic decay of the B_s^0 meson.

In the default fitter, the fraction of prompt background is the same for all three decay modes. The actual fraction is $4.8\% = 1.5 \times 3.2\%$, where 3.2% is the $D_s^-\ell^-$ yield in wrong-sign lepton $D_s^- \to \phi\pi^-$ sample and 1.5 is the ratio between the combinatorial background levels in right-sign and wrong-sign samples. We evaluate the systematic uncertainties in three different cases. First, we completely remove the prompt background component from the maximum likelihood description. This is an extreme case since we clearly observe the prompt background signal. It is just that we cannot exactly evaluate its fraction. In the second case, we increase the prompt

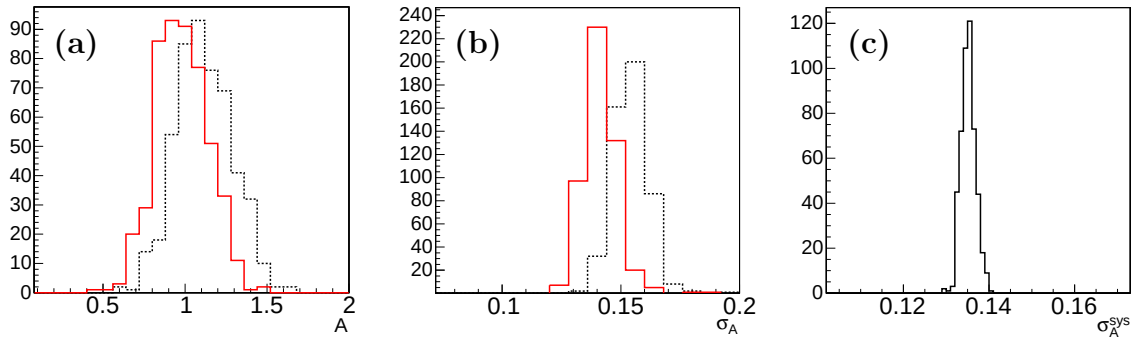

Figure 79: Example of systematic uncertainty evaluation of physics backgrounds using toy Monte Carlo experiments at $\Delta m_s = 0$ ps^{-1}. The amplitude (a) and the amplitude errors (b) are shown for the default (solid) and for the increased physics background configuration (dotted). Also the distribution of the systematic error σ^i_{sys} is shown (c).

background fraction by one standard deviation in each decay mode. By one standard deviation we understand the uncertainty of the D_s^- mass yield in the wrong sign samples. These uncertainties are 0.8% in the $\phi\pi^-$ mode, 3% in the $K^{*0}K^-$ and 6% in the $\pi^+\pi^-\pi^-$ mode. Finally, in the third case, we change the shape of the prompt background to an exponential convoluted with a Gaussian function. We compare the fit results with the default condition. To evaluate the final systematic error we add in quadrature the systematic errors coming from the first and the third scenarios. We do not use the second case since it is correlated to the first one and its effect is much smaller than the errors found in the first case.

5.10.2 Prompt Background Asymmetry

Figure 67 shows no significant difference between right-sign and wrong-sign tagged events for the prompt background which suggested to set the dilution of the prompt background to zero in the default fit. We study possible systematic uncertainties coming from the non-zero asymmetry of the prompt background. We repeat the fit setting the prompt background asymmetry to +10% and to −10%, the same as in the Δm_d analysis, and note the deviations from the default fit results.

5.10.3 Physics Backgrounds

In the default fit, the fractions of the physics backgrounds are set to their central values as shown in Table 38. The errors on the physics background fractions come from uncertainties on fragmentation fractions, decay branching ratios of B mesons as well as selection efficiency uncertainties. To assign the systematic error, the fractions are coherently varied by $+1\sigma$ and -1σ.

5.10.4 Resolution Scale Factor

In our maximum likelihood fit, we use an event-by-event decay time error obtained from the primary and secondary vertex errors returned by the vertex fitter. We know that the decay time errors are underestimated in this procedure and correct for this effect. We use a correction scale factor which depends on different kinematic variables as described in Section 5.8. To obtain the systematic uncertainty associated with the scale factor we change the the value of the scale factor by $\pm 5\%$ in the fit and compare the results with the default.

5.10.5 Dilution Scale Factors

We fix the dilution scale factors in the default fit. To determine the systematics associated with this choice, we take the deviation from the fit coherently changing the dilution scale factors for all five tag categories to $\pm 1\sigma$ with respect to their default values.

5.10.6 Sample Composition

To evaluate the uncertainties on amplitude associated with the the fraction of semileptonic B_s^0 decays into D_s^-, D_s^{*-} and D_s^{**-} states, we apply the procedure discussed in Section 4.9. We evaluate two sets of K-factor distributions and SVT bias curves. One set corresponds to direct decays of B_s^0 mesons to D_s^- states and another set corresponds to decays of B_s^0 mesons to D_s^{*-} and D_s^{**-} states. We repeat the amplitude scan with the new K-factor distributions and SVT bias curves and assign the systematic uncertainty as the variation of the amplitude from the default fit configuration.

5.10.7 SVT Bias

To evaluate the systematic uncertainties from the SVT bias curves, we apply the same procedure as in Sections 4.9.4 and 5.8. We smear the offline SVT impact parameter of the SVT tracks with a Gaussian with 12 μm width and then re-apply the trigger requirement on the SVT impact parameter $|d_0| > 120$ μm. With the new efficiency curve we repeat the fit.

5.10.8 B_s^0 Lifetime

As a default, the lifetime of the B_s^0 meson which has been used in generating the simulation samples is the central value of the world average measurement $c\tau(B_s^0) = 438$ μm. This is varied by one standard deviation (± 17 μm) to obtain a systematic error from the experimental uncertainty of the B_s^0 lifetime.

5.10.9 Event-by-Event Primary Vertex

The same procedure used for establishing the systematic uncertainty on the B_s^0 lifetime due to event-by-event vertexing procedure (see Section 5.8) is applied to

determine the systematic uncertainty on the amplitude.

5.10.10 Non-Zero $\Delta\Gamma/\Gamma$

In the default configuration of the maximum amplitude fit we assume that both heavy and light B_s^0 mass eigenstates have the same lifetime by setting $\Delta\Gamma/\Gamma = 0$. If $\Delta\Gamma/\Gamma$ is not zero the probability distribution function of the proper decay time is given by:

$$\frac{e^{-\Gamma t}}{2}(\cosh\frac{\Delta\Gamma t}{2} \pm \cos(\Delta m t)). \tag{137}$$

We study the effect of a non zero value of $\Delta\Gamma/\Gamma$ by modifying the mixing probability with the above formula. The effect of $\Delta\Gamma/\Gamma = 0.2$ on the amplitude is studied and found to be of order 1%.

5.10.11 Detector Resolution Function

The detector resolution on the B_s^0 proper decay time is assumed to be Gaussian in the default fit configuration. We take into account possible deviations of the detector resolution function from this Gaussian shape by adding exponential tails:

$$G(\Delta t) = f\frac{1}{\sqrt{2\pi}\sigma}e^{-\Delta t^2/2\sigma^2} + (1-f)\frac{1}{2\lambda}e^{-|\Delta t|/\lambda} \tag{138}$$

where Δt is the lifetime error. The fraction of non-Gaussian resolution events $(1-f)$ is set to 8% and $\lambda = 100$ μm which are typical values observed in $B \to J/\psi K$ decays [63].

5.10.12 Summary of Systematic Uncertainties

We summarize the systematic uncertainties on the amplitude from various systematic sources discussed above in Table 41. Systematic uncertainties are calculated at fixed values of $\Delta m_s = 0,\ 5,\ 10,\ 15,\ 20$ ps^{-1}. Each of the systematics uncertainties are estimated continuously in the range from 0 to 20 ps^{-1} by fitting the five points with polynomial functions. In the case of the prompt background asymmetry and the physics backgrounds, for a better understanding of the non-monotonous dependence of the systematic errors with Δm_s, the errors are evaluated in steps of 1 ps^{-1}. The total systematic uncertainty is obtained by adding in quadrature the systematic errors from all sources. Figure 80 shows the profile of each systematic uncertainty contribution as a function of Δm_s with the polynomial fitting function superimposed. Figure 81 shows the interpolated individual and total systematic error contributions.

We repeat the blinded amplitude scan including the above systematic errors added in quadrature. The result of this amplitude scan is shown in Figure 82 indicating a sensitivity of 7.3 ps^{-1}. Clearly, the measurement is dominated by statistical errors.

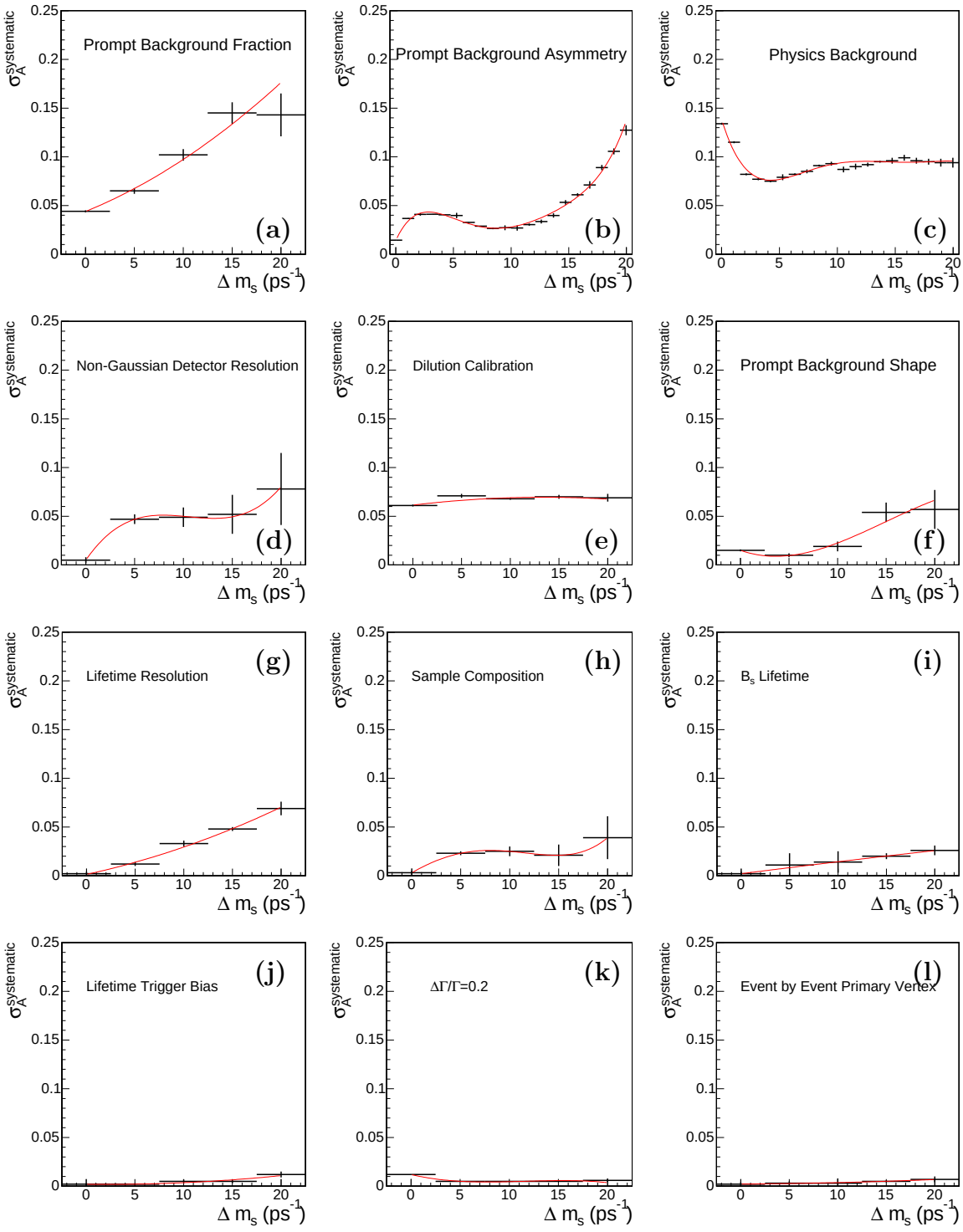

Figure 80: Systematic uncertainties as a function of Δm_s: (a) prompt background fraction, (b) prompt background asymmetry, (c) physics backgrounds, (d) non-Gaussian detector resolution, (e) dilution calibration factors, (f) prompt background shape, (g) decay time resolution, (h) sample composition, (i) B_s^0 lifetime, (j) SVT trigger bias, (k) non-zero $\Delta\Gamma/\Gamma$, (l) event-by-event primary vertex.

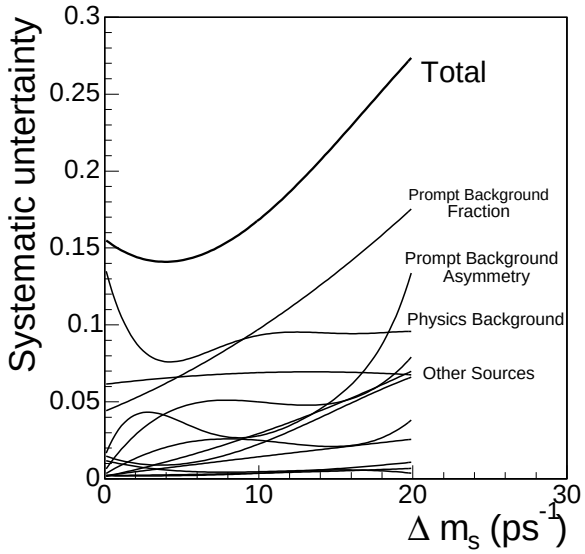

Figure 81: Systematic uncertainties as a function of the Δm_s. All contributions are shown as interpolated shapes and also the total systematic uncertainty is shown.

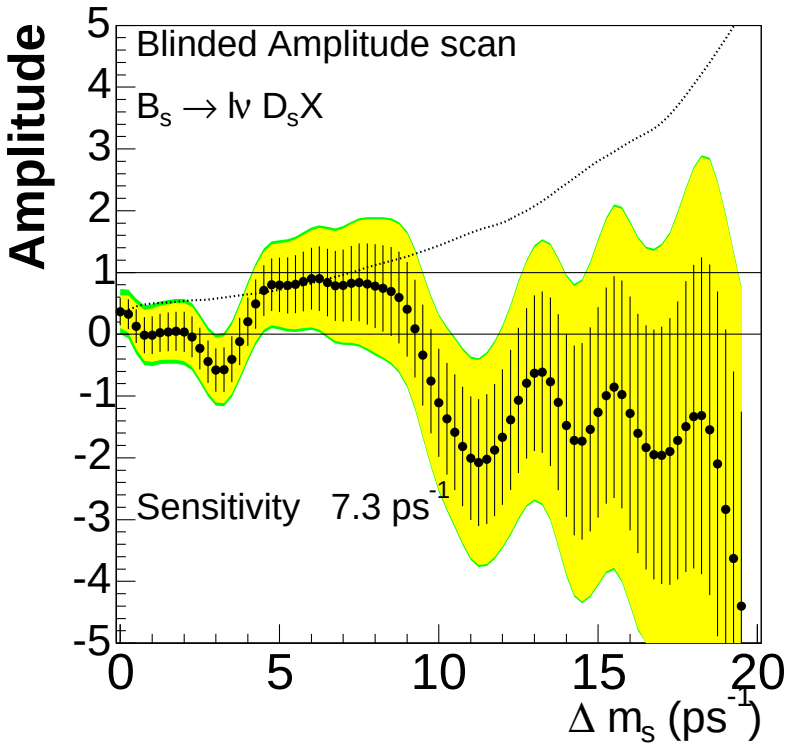

Figure 82: Result of blinded amplitude scan including systematic uncertainties. The yellow band corresponds to the statistical error only and the green band contains both the statistical and the systematic errors added in quadrature.

Source of Systematics	Selected Δm_s values				
	0 ps^{-1}	5 ps^{-1}	10 ps^{-1}	15 ps^{-1}	20 ps^{-1}
Prompt background fraction	0.044	0.065	0.102	0.145	0.143
Prompt background asymmetry	0.014	0.040	0.027	0.062	0.157
Prompt background shape	0.015	0.010	0.019	0.054	0.057
Physics background fraction	0.134	0.078	0.093	0.096	0.103
Resolution scale factor	0.002	0.012	0.033	0.047	0.065
Dilution scale factors	0.061	0.071	0.068	0.070	0.069
Sample composition	0.002	0.015	0.022	0.021	0.039
SVT efficiency bias curve	0.002	0.001	0.005	0.005	0.012
B_s^0 lifetime	0.001	0.011	0.014	0.020	0.026
Event-by-event primary vertex	0.002	0.003	0.003	0.005	0.007
Detector resolution function	0.005	0.047	0.049	0.052	0.078
$\Delta\Gamma/\Gamma$	0.012	0.005	0.005	0.005	0.006
Total Systematics	0.155	0.134	0.160	0.213	0.262
Statistical Error	0.159	0.406	0.856	1.654	3.364

Table 41: Summary of systematic uncertainties for $\Delta m_s = 0,\ 5,\ 10,\ 15,\ 20$ ps^{-1}.

5.11 Final Result - The Unblinded Amplitude Scan

Since we have now established the sensitivity of our measurement and evaluated the systematic uncertainties on the blinded sample, we can proceed to the final step of the analysis, setting a lower limit on the B_s^0 oscillation frequency Δm_s using semileptonic $B_s^0 \to D_s^- \ell^+ X$ decays. After we remove the randomization procedure of the tag decision, we repeat the amplitude scan. The result is shown in Figure 83. From this distribution, a lower limit on the B_s^0 mixing frequency of $\Delta m_s > 7.7$ ps^{-1} is derived at the 95% Confidence Level with a sensitivity of 7.4 ps^{-1}. Figure (96) from Appendix A.12 shows the amplitude scans corresponding to each of the three D_s decay modes. Clearly, the dominant mode is $D_s^- \to \phi\pi^-$, while the $D_s^- \to K^{*0}K^-$and $D_s^- \to \pi^+\pi^-\pi^-$modes by themselves have very low sensitivity to Δm_s. The numerical values of the amplitude at each value of Δm_s together the with statistical and systematic errors are given in Appendix A.13.

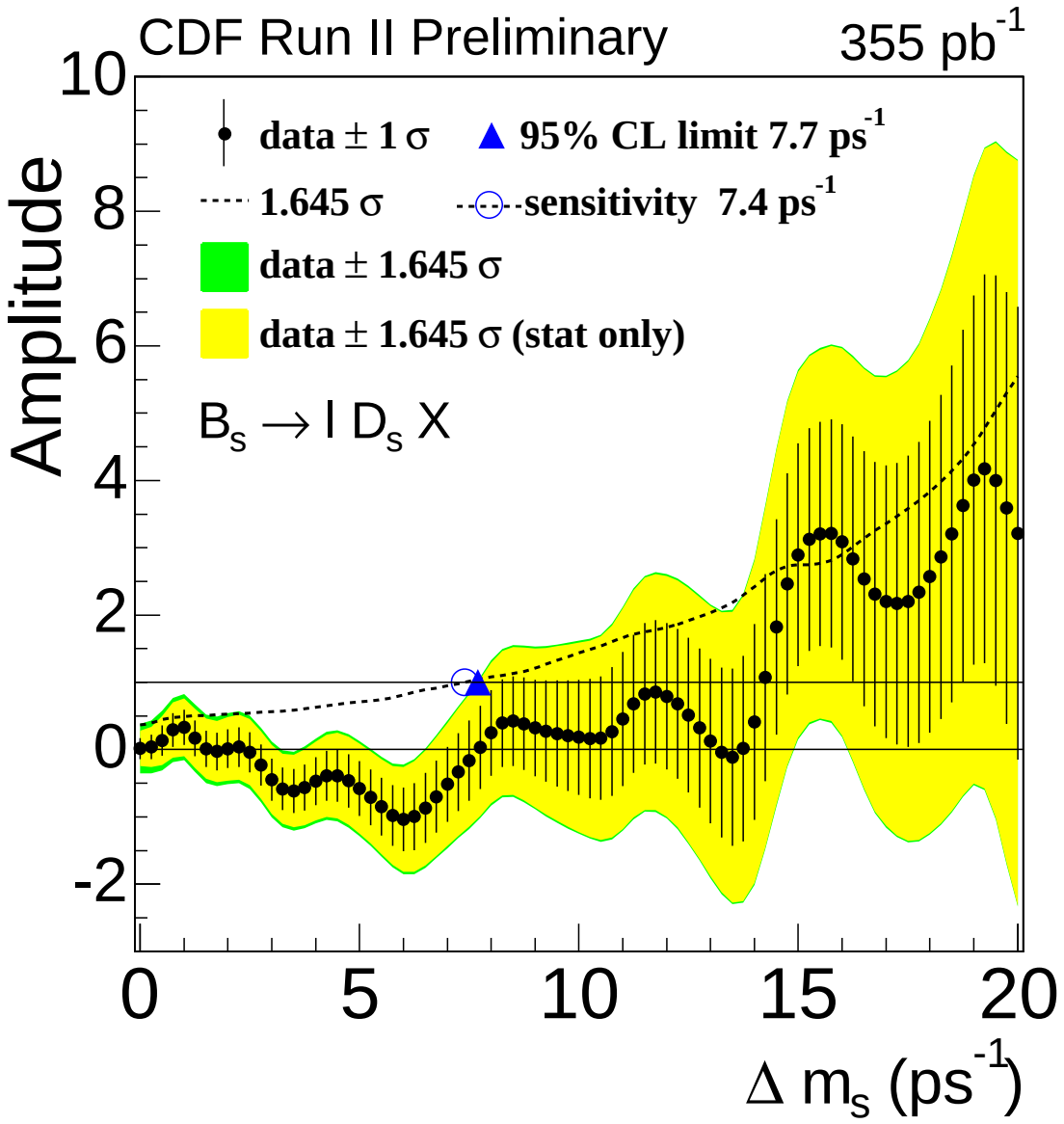

Figure 83: The unblinded amplitude scan reveals a lower limit on the B_s^0 oscillation frequency Δm_s of 7.7 ps^{-1} at 95% Confidence Level.

6 Conclusions

In this thesis we discuss three major topics: an opposite side B flavor tagging algorithm, a measurement of the B^0 oscillation frequency Δm_d in semileptonic B decays plus a calibration of the tagging dilutions of opposite side lepton and jet charge taggers as well as a search for B_s^0 mixing in semileptonic $B_s^0 \to D_s^- \ell^+ X$ decays.

The likelihood based B flavor tagging algorithm uses semileptonic $B \to \mu X$ decays of the opposite side b quark in a $b\bar{b}$ event to determine the production flavor of the trigger side neutral B meson. To perform this study we use an inclusive sample enriched in semileptonic B decays. The events in this sample are selected by the lepton-SVT trigger which requires a 4 GeV/c lepton, a displaced SVT track with impact parameter satisfying 120 μm $< d_0 <$ 1 mm. The likelihood algorithm combines information from the electromagnetic and hadronic calorimeters with track-stub matching quantities to form a global likelihood function used for muon identification. We determine that the B tagging dilution increases as a function of the muon likelihood. We use the dependence of the dilution on the muon likelihood as well as the dependence on $p_T^{\rm rel}$ to improve the overall tagging power of the opposite side muon tagger.

In the measurement of the B^0 oscillation frequency Δm_d in semileptonic $B \to D\ell X$ decays, the charm mesons are fully reconstructed from the lepton-SVT sample. For B flavor tagging, we use opposite side muon and electron tagging algorithms as well as jet charge tagging. In this analysis we utilize the dependences of the tagging dilution on lepton likelihood, $p_T^{\rm rel}$ or jet charge as determined from the inclusive lepton-SVT sample. However, the value of this predicted tagging dilution is multiplied by a freely floating parameter which is determined from the unbinned maximum likelihood fit that measures Δm_d. These parameters are found to be consistent with one for each tagging algorithm employed in this analysis. The absolute values of the tagging dilutions as determined with this calibration method are then used in our search for B_s^0 oscillations.

The search for B_s^0 mixing is performed using semileptonic $B_s^0 \to D_s^- \ell^+ \nu$ decays, where the D_s^- meson is fully reconstructed from three hadronic decay modes $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$and $D_s^- \to \pi^+\pi^-\pi^-$. The analysis is initially performed "blindly" with randomized tagging decisions. After the sensitivity and the systematic uncertainties are determined, we repeat the analysis with correct tagging information and determine a lower limit on the B_s^0 oscillation frequency $\Delta m_s > 7.7$ ps^{-1} at 95% Confidence Level.

In parallel with the search for B_s^0 mixing in semileptonic decays, another analysis employing fully hadronic $B_s^0 \to D_s^- \pi^+$ decays was performed at CDF [64]. The advantage of using hadronic decays is that the B momentum can be fully reconstructed as all final state particles are observed. With full B momentum reconstruction, the uncertainty on the proper decay time is smaller than in semileptonic decays. On the

other hand, the data sample accumulated in the hadronic analysis is about an order of magnitude smaller than the semileptonic sample. For this reason, the limit set in the hadronic analysis is null, $\Delta m_s > 0.0$ ps^{-1} at 95% Confidence Level with a sensitivity of 0.4 ps^{-1}. We expect that in the future, with more accrued data, the hadronic analysis will start to dominate at large values of Δm_s as the statistical error on the amplitude in the hadronic modes increases slower than in semileptonic modes due to better proper decay time resolution. Already, we find that after combining the amplitude scans from the semileptonic and hadronic analyses, the total sensitivity increases to 8.4 ps^{-1} with a lower limit $\Delta m_s > 7.9$ ps^{-1} [65] as shown in Figure 84. The current PDG world average lower limit is $\Delta m_s > 14.4$ ps^{-1}, not including the Tevatron Run II results. The best single-experiment limit is $\Delta m_s > 10.9$ ps^{-1} at 95% Confidence Level measured by the ALEPH experiment [66].

It is estimated [67] that the CDF experiment will cover the Standard Model favored Δm_s region by reaching a sensitivity of ~ 20 ps^{-1} with 8 fb^{-1} after including other tagging techniques currently under development [68]. Figure 85 indicates the sensitivity to Δm_s for the CDF experiment [67] using both semileptonic and hadronic samples. The shown predictions consider both 95% Confidence Level exclusion and 5σ observation. In each case three scenarios are studied: the current analyses with no improvements, conservative improvements and optimistic improvements. The conservative (optimistic) improvement scenario assumes that the tagging power given by $\varepsilon\mathcal{D}^2$ increases by 1% (3%) and the decay time resolution in hadronic modes improve by 10% (30%). It is clear that a sensitivity of at least 20 ps^{-1} can be reached with ~ 5 times more data assuming the conservative improvement scenario. However, to make a 5σ observation at $\Delta m_s = 20$ ps^{-1}, we would need ~ 15 times more data assuming the optimistic scenario. Under any assumption, the CDF experiment will at least establish the best lower limit on the B_s^0 oscillation frequency Δm_s in the near future.

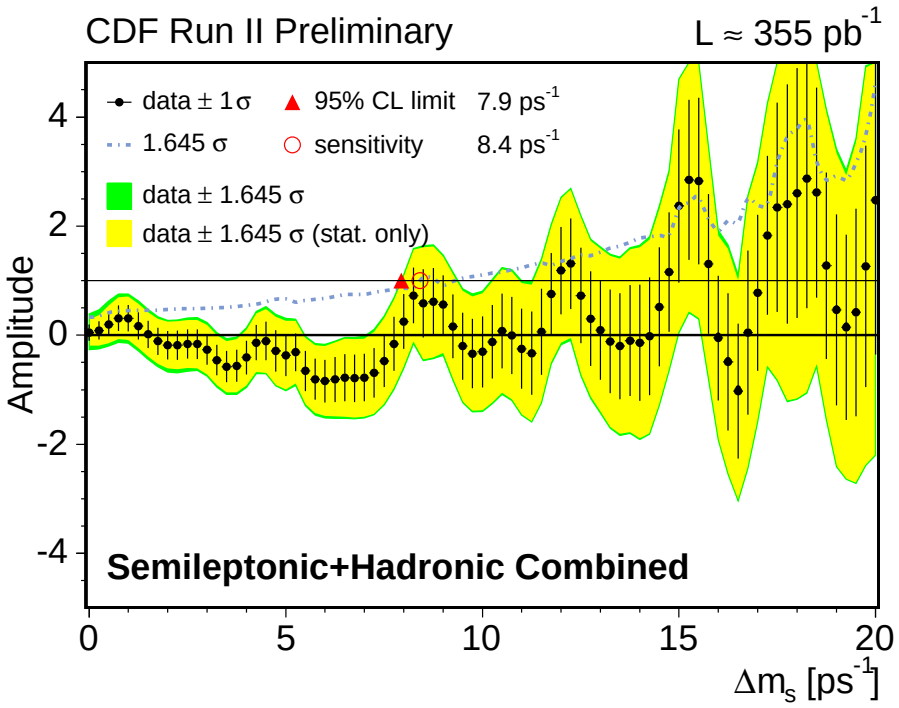

Figure 84: Amplitude scan obtained after combining the semileptonic and hadronic amplitude scans.

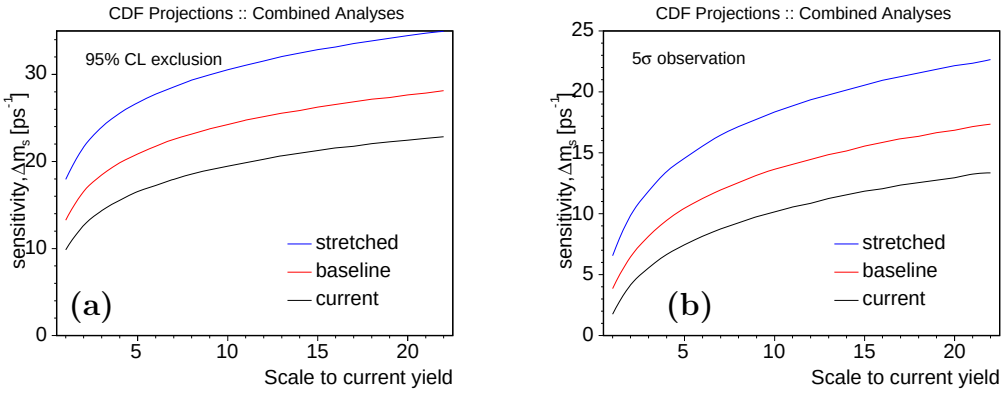

Figure 85: Sensitivity of the combined semileptonic and hadronic analyses as a function of the scale to the current B_s^0 yield for (a) 95% Confidence Level exclusion and (b) 5σ observation.

A Appendix

A.1 Improvement of Tagging Power by Using the Dilution Dependence on a Given Quantity

Let us consider a sample of B^0 decays with N events on which we apply a flavor tagging algorithm with efficiency ε and dilution $\mathcal{D}$. The error on the oscillation frequency $\sigma_{\Delta m}$ scales with the effective number of events:

$$\frac{1}{\sigma^2_{\Delta m}} \sim N\varepsilon\mathcal{D}^2. \tag{139}$$

Suppose that the dilution $\mathcal{D}$ depends on some variable q: $\mathcal{D} = \mathcal{D}(q)$. We can separate the data sample in n subsamples corresponding to n ranges of q. In each range the average dilution is $\mathcal{D}_i$, $i = 1, n$. We measure the oscillation frequency Δm in each q range and find the result Δm_i with statistical error $\sigma^i_{\Delta m}$. The best estimate of the true value of the oscillation frequency is given by the weighted average of the independent measurements Δm_i:

$$\Delta m = \frac{\sum_i^n \Delta m_i/\sigma^{i^2}_{\Delta m}}{\sum_i^n 1/\sigma^{i^2}_{\Delta m}}, \tag{140}$$

and the error on Δm is given by:

$$\frac{1}{\sigma^2_{\Delta m}} \sim \sum_i^n \frac{1}{\sigma^2_{\Delta m_i}} = N\sum_i^n \varepsilon_i\mathcal{D}_i^2. \tag{141}$$

For simplicity, we assume that each subsample i contains $\frac{N}{n}$ events and the tagging efficiency is the same in each subsample $\varepsilon_i = \frac{\varepsilon}{n}$. We also note that the dilution $\mathcal{D}$ used in Equation (139) is the average of dilutions $\mathcal{D}_i$ in samples i:

$$\mathcal{D} = \frac{1}{n}\sum_i^n \mathcal{D}_i \tag{142}$$

Using the inequality:

$$(\sum_i^n \mathcal{D}_i)^2 < n\sum_i^n \mathcal{D}_i^2, \tag{143}$$

we infer that:

$$N\varepsilon\mathcal{D}^2 < N\sum_i^n \varepsilon_i\mathcal{D}_i^2. \tag{144}$$

The above relation implies that the error calculated in Equation (141) is always smaller than the error from Equation (139). This means that knowing the dilution dependence on any variable would improve the tagging power of the flavor tagging algorithm characterized by $\varepsilon\mathcal{D}^2$.

A.2 Parameters Describing the p_T Dependence of the Track-Stub Matching Variables

Parameter	a	b	c	d	e	f
CMU ΔX	14.65	2.555	−0.890	0.048	–	–
CMU ΔZ	8.33	−5.25	22.96	–	–	–
CMU $\Delta\phi$	0.0120	0.0330	0.352	–	–	–
CMP ΔX	39.38	4.55	−0.782	0.102	2.8	−0.27
CMP $\Delta\phi$	0.04022	−0.1853	1.43	–	–	–
CMX ΔX	−43.58	6.50	0.675	0.366	27.92	77.71
CMX ΔZ	−33.01	4.996	0.675	6.552	−3.1	87.91
CMX $\Delta\phi$	−0.2228	0.04	0.675	0.0074	0.0695	0.655

Table 42: Parameters describing the p_T dependence of the track-stub matching variables for CMU, CMP and CMX muons.

A.3 Consistency Checks of Fake Parameterizations

Hadrons that are associated with muon stubs interact late in the shielding material which means that they must be minimum ionizing in the calorimeter. We expect that the minimum ionizing energy distributions for pions, kaons and protons are similar. As a first step in convincing ourselves that different hadron types faking muons behave similarly in the calorimeter, we compare the energy depositions of protons and anti-protons that fake muons. We find that they are almost identical. Figure 86 shows the CMU proton and anti-proton electromagnetic energy and hadronic energy as well as the CMU proton and anti-proton distributions of ΔX, $\Delta\Phi$, ΔZ (before scaling).

Furthermore, we find that pions and kaons can be well described by the proton distributions. Pions from $K^0_S \to \pi^+\pi^-$ and kaons from $D^0 \to K^+\pi^-$ that are associated with muon objects, are either punch-through hadrons or decay-in-flight muons. This means that their distributions can be treated as superpositions of punch-through and decay-in-flight. distributions. If protons correctly describe generic punch-through hadrons and distributions from decay-in-flight muons are similar to real muons not from decay-in-flight, then the pion and kaon distributions from K^0_S and D^0 decays can be described as superpositions of proton and real muon distributions:

$$\text{Pion Distribution} = \alpha \times \text{Muon Template} + (1-\alpha) \times \text{Proton Template} \quad (145)$$

$$\text{Kaon Distribution} = \beta \times \text{Muon Template} + (1-\beta) \times \text{Proton Template} \quad (146)$$

where α and β are the DIF fractions in the pion and kaon samples.

Figure 87 shows that the pion and kaon distributions are almost identical and indeed bound between the proton and real muon distributions. The "hump" centered at ≈ 1.5 GeV in the pion and kaon hadronic energy distributions is an indication that some of the pions and kaons decay into muons before they reach the hadronic

calorimeter and the corresponding energy deposition has a muon-like minimum ionizing peak shape. We find that for $\alpha \approx 35\%$ and $\beta \approx 25\%$ the CMU pion and kaon distributions match the combination of protons and muons reasonably well as can be seen in Figures 88 and 89.

A.4 Consistency Checks on Decay-in-Flight and Punch-Through Hypotheses

As a consistency check, we also study the likelihood algorithm on CMU, CMUP and CMX pions from K^0_S. The corresponding distributions are shown in Figure 90. As mentioned in Section 3.3, pions from K^0_S that match a muon stub can either be PT or DIF. Consistent with this hypothesis, the likelihood distributions for pions have spikes at zero that correspond to PT but also peaks at one which correspond to DIF. Similar effects are seen for CMP and BMU pions and also for kaons from D^0 that fake muons. CMUP pions have the highest rate of DIF. Hadrons are unlikely to produce CMUP muons by PT which is consistent with the very low number of CMUP protons (about 400 events) that we collected from the entire two-track-trigger data set. Using the likelihood method, we can separate the PT pions from the DIF pions. For example, Figure 90 shows the likelihood distributions for CMU, CMUP and CMX pions. The spikes at zero corresponding to PT and the peaks at one corresponding to DIF are quite obvious. By selecting pions from $K^0_S \to \pi^+\pi^-$ that have a muon object associated and high muon likelihood ($\mathcal{L} > 0.8$), we obtain a sample enriched in DIF pions. Similarly, by selecting pions from K^0_S decays that are associated with a muon object and have low likelihood ($\mathcal{L} < 0.2$), we obtain a sample enriched in PT. We note that the trajectory of a pion or kaon that decays in flight into a muon is not a simple helix. Because the muon momentum has in general a direction different than the initial direction of the meson momentum, the combined trajectory will have a discontinuous derivative. Also the curvature of the muon in the magnetic field is larger than that of the meson since the muon momentum is smaller. The track associated with a meson that decays into a muon is fit to a helix. Since the actual trajectory is described by two helices, we expect that the quality of the fit will be worse especially when the muon momentum makes a large angle ("a kink") with the momentum of the initial meson. Figure 91 compares the χ^2 divided by the number of degrees of freedom $\chi^2/ndof$ of the track fit and the impact parameters of PT and DIF pions. The $\chi^2/ndof$ of DIF pions is slightly but systematically shifted with respect to the $\chi^2/ndof$ of PT pions. This observation is consistent with the expectation of a "kink" in the DIF trajectory which worsens the quality of the track fit. Similar, shifts are observed when only silicon or only COT hits are included in the $\chi^2/ndof$. Also, the impact parameter distribution of high likelihood pions is wider than that of low likelihood ones. This supports our assumption that high likelihood pions are dominated by DIF while low likelihood ones are dominated by PT. From the sharp cuts at 120 μm and 1 mm, it can be seen that most of the pions are SVT tracks. The tails in the impact parameter distribution come from non-trigger tracks.

We check that the hadronic energy distribution of the low likelihood pions has no

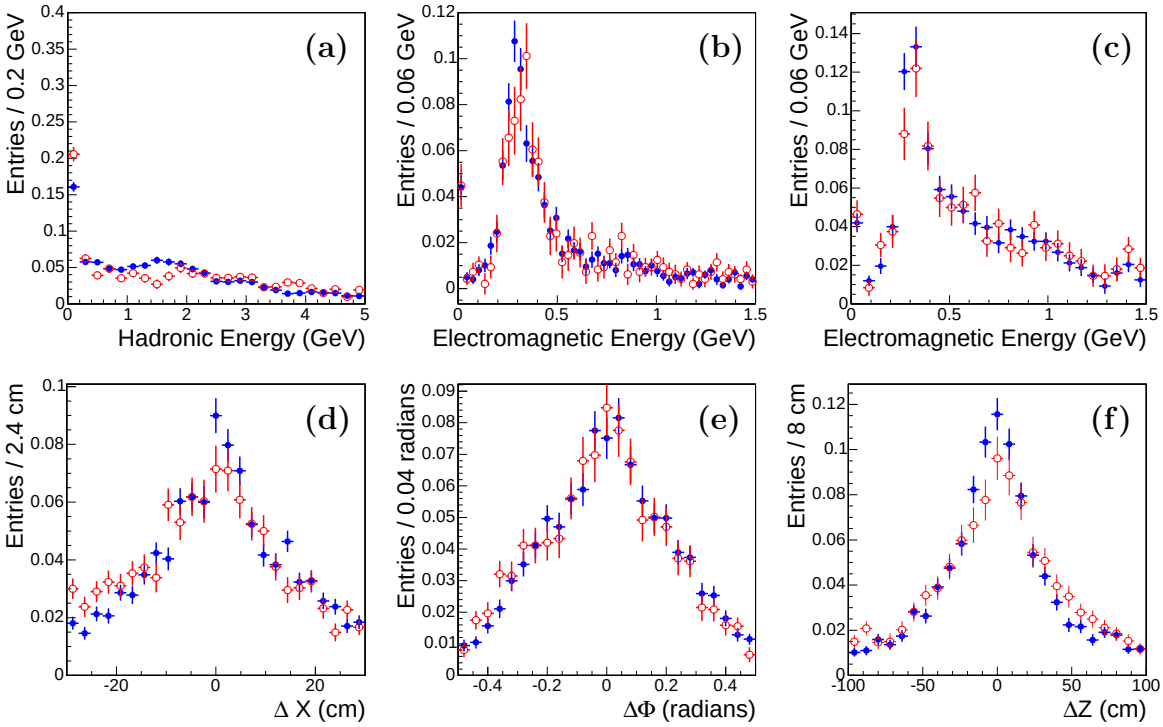

Figure 86: Proton (full circles) and anti-proton (open circles) distributions of (a) hadronic energy, (b) electromagnetic energy of isolated tracks, (c) electromagnetic energy of non-isolated tracks, (d) ΔX, (e) $\Delta\Phi$ and (f) ΔZ.

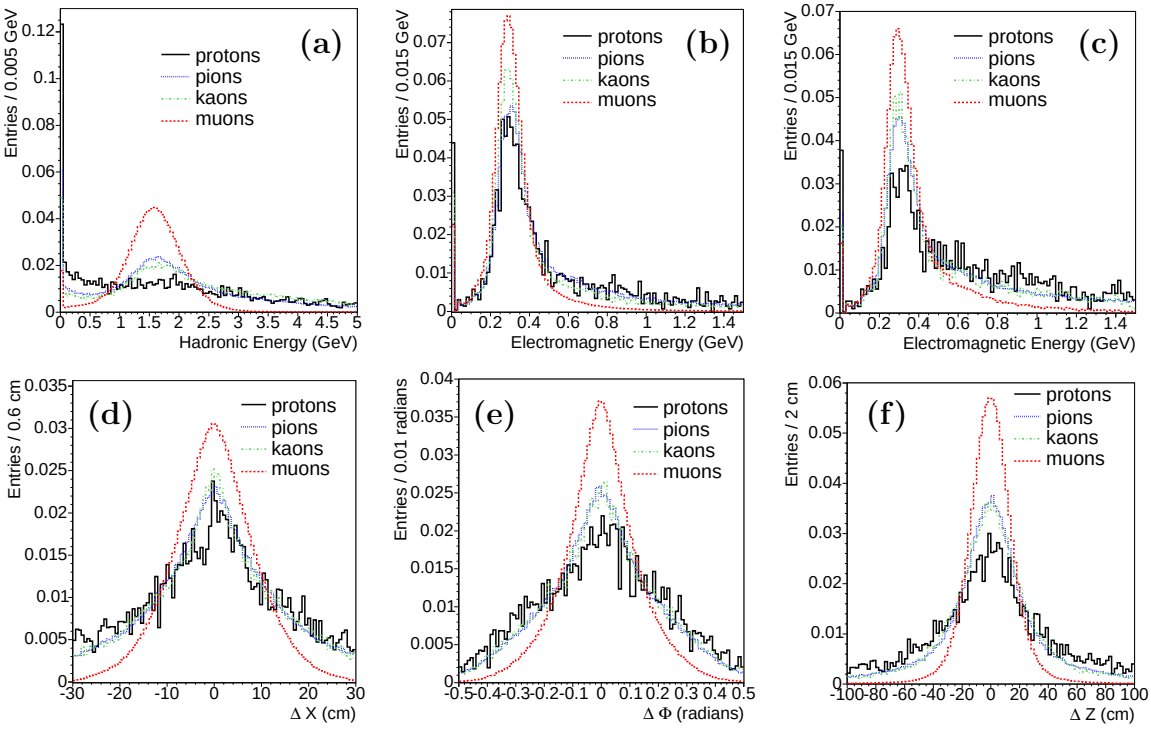

Figure 87: Proton (solid), pion (dotted), kaon (dot-dashed) and muon (dashed) distributions of (a) hadronic energy, (b) electromagnetic energy of isolated tracks, (c) electromagnetic energy of non-isolated tracks, (d) ΔX, (e) $\Delta\Phi$ and (f) ΔZ.

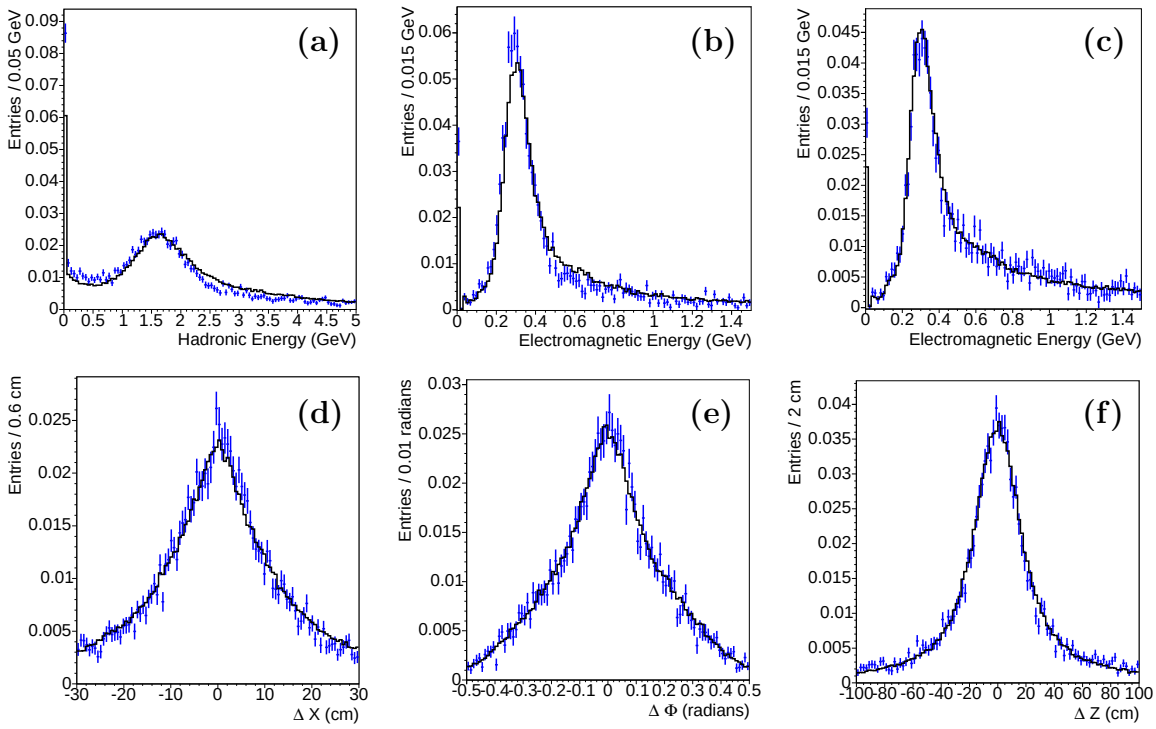

Figure 88: Superpositions of muons and protons are shown with error bars, while the pion distribution are shown as solid histograms: (a) hadronic energy, (b) electromagnetic energy (isolated), (c) electromagnetic energy (non-isolated), (d) ΔX, (e) $\Delta\Phi$, (f) ΔZ.

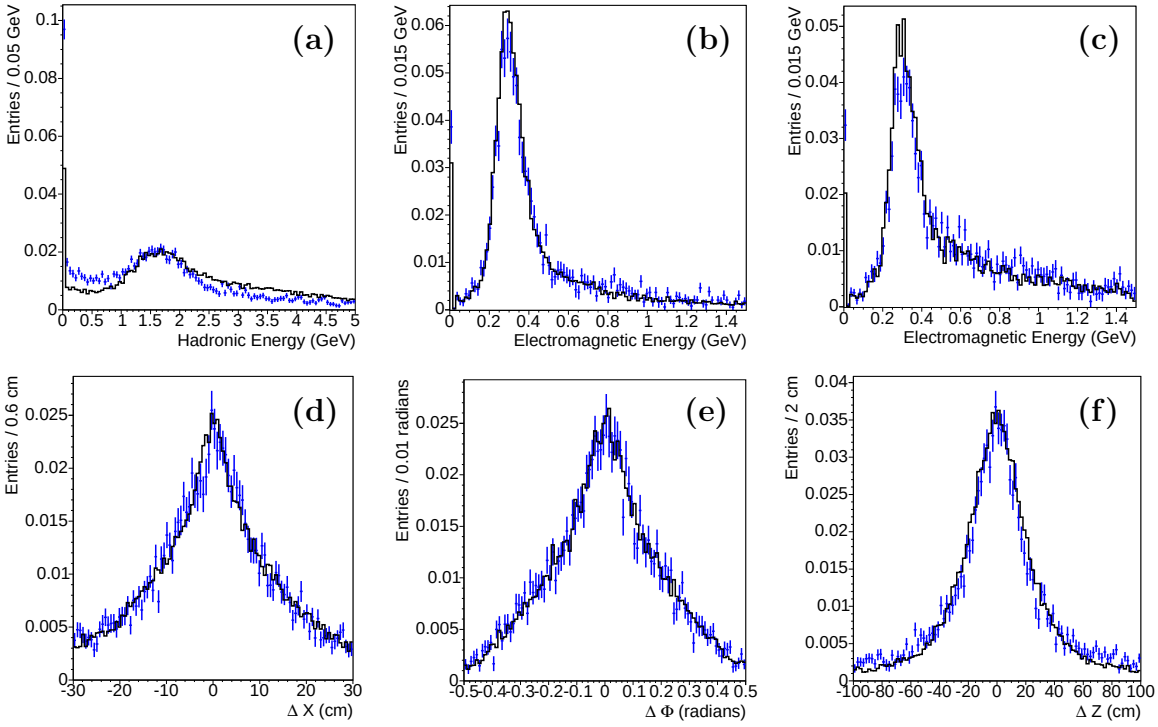

Figure 89: Superpositions of muons and protons are shown with error bars, while the kaons distribution are shown as solid histograms (a) hadronic energy, (b) electromagnetic energy (isolated), (c) electromagnetic energy (non-isolated), (d) ΔX, (e) $\Delta\Phi$, (f) ΔZ.

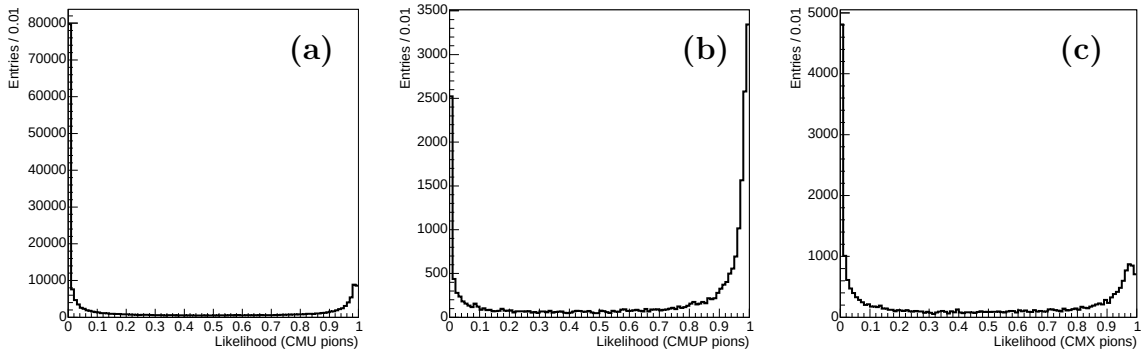

Figure 90: Likelihood distributions for (a) CMU, (b) CMUP and (c) CMX pions from K^0_S. As expected, besides the spikes at zero corresponding to PT pions, we also see peaks at one corresponding to DIF pions.

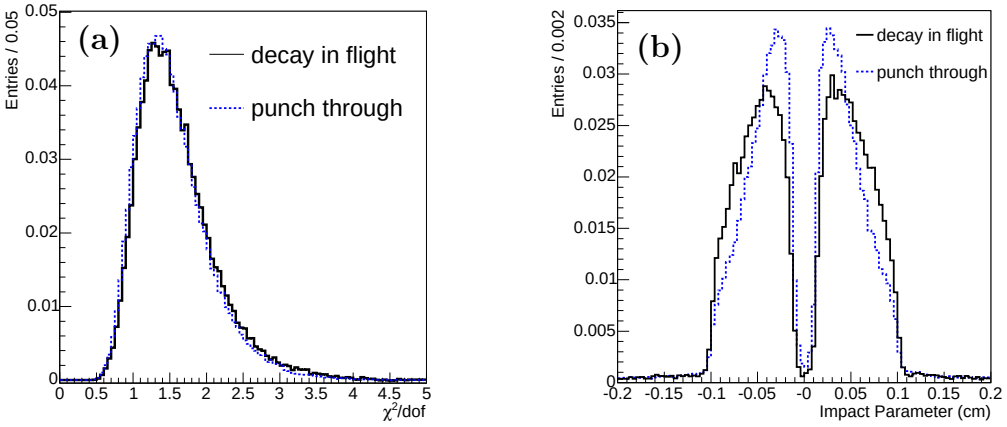

Figure 91: (a) $\chi^2/ndof$ of PT pions (dashed) and DIF pions (solid) and (b) impact parameter of PT pions (dashed) and and DIF pions (solid).

"hump" at ≈ 1.5 GeV, which means that this subsample of pions does not have a DIF component. As an example, Figure 92 shows that the hadronic energy distributions of CMU and CMX pions with $\mathcal{L} < 0.2$ are almost identical with those of CMU and CMX protons.

$\mathcal{L}$ cut (%)	2	10	75
CMU (%)	0.6	1.9	15.6
CMP (%)	0.7	2.1	27.0
CMUP (%)	0.7	1.9	14.1
CMX (%)	0.5	1.8	17.5
BMU (%)	2.6	5.6	22.9

2	10	75
62.1	73.8	91.6
14.9	32.8	85.7
39.8	50.0	72.5
32.8	52.6	88.4
51.3	63.4	92.1

Table 43: Fractions of rejected real muons (left) and rejected fake muons (right) for different likelihood lower cuts.

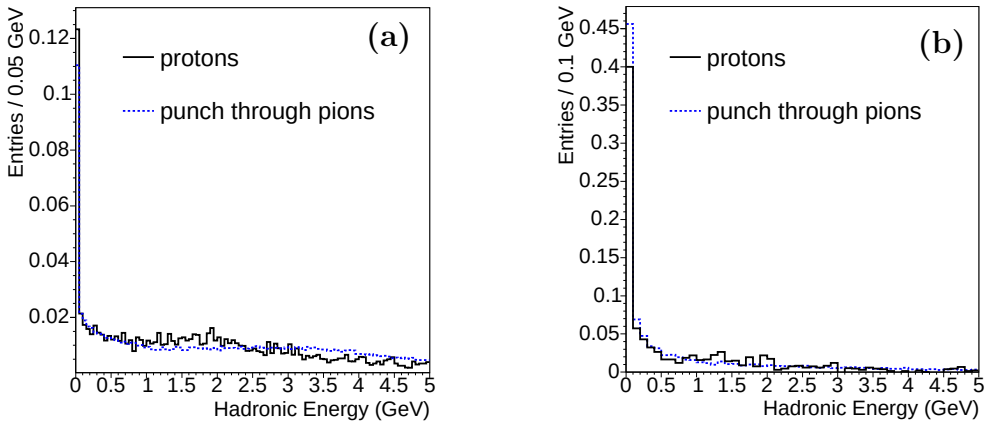

Figure 92: Hadronic energy distributions of low likelihood ($\mathcal{L} < 0.2$) pions (dashed) and hadronic energy distribution of protons (solid) for (a) CMU muons and (b) CMX muons.

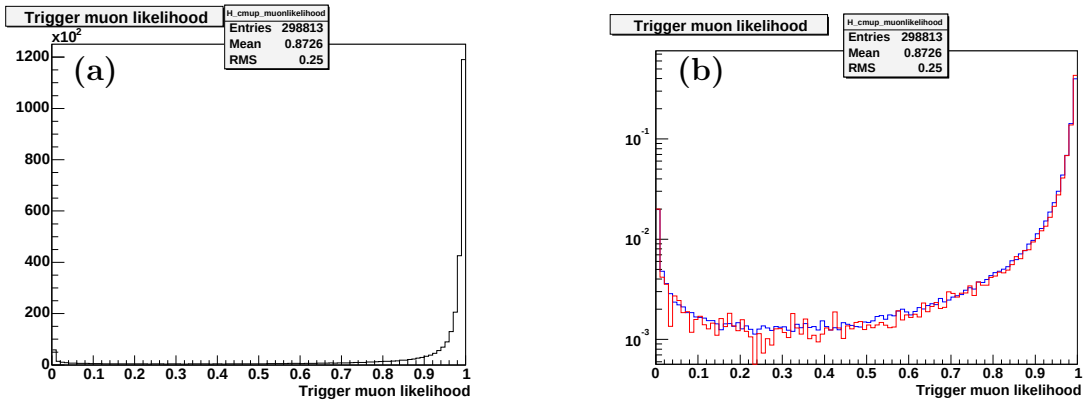

Figure 93: Likelihood distribution of (a) trigger muons on a linear scale and (b) logarithmic scale. The linear combination of real muons and protons is superimposed (red histogram).

A.5 Consistency Checks of the Likelihood Algorithm

A simple way to check the validity of the likelihood method is to test it on the high purity trigger muons from the muon-SVT data sample. These muons have $p_T > 4$ GeV/c, have stubs in both CMU and CMP muon detectors and have tight matching cuts: $\Delta X < 15$ cm for the CMU stub and $\Delta X < 20$ cm for the CMP stub. The likelihood distribution, shown in Figure 93, has almost no spike at zero which indicates that the trigger muon fake rate is small. One can obtain a more quantitative estimate by assuming that: trigger muons are either real muons or punch through hadrons. This means that their likelihood distribution is a linear combination of likelihoods corresponding to real muons and protons. We again assume that the hadrons that decay-in-flight into muons are well described by the real muon distributions. We find

the fraction α of punch through hadrons for which the linear combination:

$$\alpha \times \text{proton template} + (1 - \alpha) \times \text{muon template}$$

best matches the observed distribution of trigger muons. The muon and proton templates are obtained by requiring that the track has at least 4 GeV/c and the track-stub matching cuts are the same as for trigger muons. We find a fake fraction $\alpha \approx 3\%$. By examining the likelihood distribution of CMUP pions from Figure 90, one can estimate that the ratio between decay-in-flight and punch through is roughly 1.5 which means that the DIF rate of the trigger muons is $\approx 5\%$. These results are in reasonable agreement with the findings in Reference [31].

From the likelihood distributions for real muons and fakes presented in Section 3.5, one can extract the efficiencies for rejecting real muons and fakes for a given likelihood cut. Table 43 shows these fractions for different likelihood cuts.

A.6 Trigger Side and Opposite Side Dilutions

A b quark produced in a proton-antiproton collision is always produced together with a $\bar{b}$ quark. The quarks hadronize into B mesons (or baryons). The B meson that triggers the event is called the "trigger B meson" and the second meson in the event is called the "the opposite B meson". If either the b or the $\bar{b}$ quark hadronizes into a neutral B state, this meson may oscillate and its flavor at decay will be different than the flavor at production. We define the probability that the trigger/opposite B meson decays with different/same flavor as the production flavor as $P^{trg/opp}_{mix/unmix}$. We also define the probability to correctly/incorrectly identify the flavor of the trigger/opposite B meson as $P^{trg/opp}_{R/W}$.

As an example, we consider the case in which both trigger and opposite B mesons decay semileptonically. The sign of the charge of each lepton will be the same as the sign of the charge of the parent b or $\bar{b}$ quark. Thus, the lepton charge determines whether the B meson contains a b or a $\bar{b}$ quark. The probability to observe an event with opposite charge leptons is given by:

$$\begin{aligned} P_{OS} &= (P^{trg}_{unmix} P^{opp}_{unmix} + P^{trg}_{mix} P^{opp}_{mix}) (P^{trg}_{R} P^{opp}_{R} + P^{trg}_{W} P^{opp}_{W}) \\ &\quad + (P^{trg}_{unmix} P^{opp}_{mix} + P^{trg}_{mix} P^{opp}_{unmix}) (P^{trg}_{R} P^{opp}_{W} + P^{trg}_{W} P^{opp}_{R}) \end{aligned} \quad (147)$$

and the probability to observe an event with same-sign leptons is given by:

$$\begin{aligned} P_{SS} &= (P^{trg}_{unmix} P^{opp}_{unmix} + P^{trg}_{mix} P^{opp}_{mix}) (P^{trg}_{R} P^{opp}_{W} + P^{trg}_{W} P^{opp}_{R}) \\ &\quad + (P^{trg}_{unmix} P^{opp}_{mix} + P^{trg}_{mix} P^{opp}_{unmix}) (P^{trg}_{R} P^{opp}_{R} + P^{trg}_{W} P^{opp}_{W}) \end{aligned} \quad (148)$$

Using the above equations, one finds that the raw dilution D_{raw} is given by:

$$\begin{aligned} D_{raw} = \frac{P_{OS} - P_{SS}}{P_{OS} + P_{SS}} &= (1 - 2P^{opp}_{W}) (1 - 2P^{opp}_{mix}) (1 - 2P^{trg}_{W}) (1 - 2P^{trg}_{mix}) \\ &= D_{opp} D_{trg}, \end{aligned} \quad (149)$$

where the trigger side dilution is given by:

$$D_{trg} = (1 - 2P_W^{trg}) (1 - 2P_{mix}^{trg}), \tag{150}$$

while the opposite side dilution is given by the similar expression:

$$D_{opp} = (1 - 2P_W^{opp}) (1 - 2P_{mix}^{opp}). \tag{151}$$

We note that both the trigger and opposite side dilution are products of the dilution due to misidentification of the b flavor at decay $1 - 2P_W$ and the dilution due to mixing $1 - 2P_{mix}$.

A.7 Detailed Calculation of $\varepsilon\mathcal{D}^2$

Tables 44 and 45 show the efficiency, raw dilution and $\varepsilon\mathcal{D}^2_{raw}$ in each likelihood range. $\varepsilon\mathcal{D}^2$ is obtained after summing over $p_T^{\rm rel}$ bins.

$\mathcal{L}$ range	type	ε (%)	$\mathcal{D}_{raw}$ [%]	$\varepsilon\mathcal{D}^2_{raw}$ (%)
0.0-0.8	CMU	1.174 ± 0.030	1.9 ± 2.6	0.0036 ± 0.0059
	CMP	0.240 ± 0.013	8.5 ± 5.4	0.0044 ± 0.0034
	CMUP	0.219 ± 0.012	19.7 ± 5.3	0.0136 ± 0.0059
	CMX	0.277 ± 0.014	18.5 ± 4.9	0.0134 ± 0.0061
	BMU	1.305 ± 0.031	5.1 ± 2.4	0.0045 ± 0.0040
0.8-0.95	CMU	0.316 ± 0.014	17.4 ± 4.5	0.0138 ± 0.0059
	CMP	0.149 ± 0.010	31.0 ± 6.3	0.0207 ± 0.0075
	CMUP	0.135 ± 0.009	29.3 ± 6.5	0.0123 ± 0.0054
	CMX	0.176 ± 0.010	26.7 ± 5.6	0.0211 ± 0.0068
	BMU	0.173 ± 0.010	24.2 ± 5.9	0.0167 ± 0.0065
0.95-1.0	CMU	0.606 ± 0.020	24.1 ± 3.2	0.0410 ± 0.0099
	CMP	0.035 ± 0.005	31.8 ± 13.0	0.0057 ± 0.0044
	CMUP	0.455 ± 0.016	34.0 ± 3.3	0.0558 ± 0.0105
	CMX	0.292 ± 0.013	30.2 ± 4.2	0.0305 ± 0.0078
	BMU	0.379 ± 0.015	29.7 ± 3.8	0.0343 ± 0.0087
Sum	CMU	2.096 ± 0.039	–	0.0584 ± 0.0130
	CMP	0.423 ± 0.017	–	0.0308 ± 0.0094
	CMUP	0.809 ± 0.022	–	0.0817 ± 0.0132
	CMX	0.744 ± 0.021	–	0.0650 ± 0.0120
	BMU	1.858 ± 0.036	–	0.0555 ± 0.0116
Total	All	5.929 ± 0.063	–	0.291 ± 0.0266
Corrected				0.708 ± 0.065

Table 44: Efficiency, raw dilutions and $\varepsilon\mathcal{D}^2_{raw}$ in three likelihood bins for the electron-SVT data sample. $\varepsilon\mathcal{D}^2$ is obtained after summing over $p_T^{\rm rel}$ bins.

$\mathcal{L}$ range	type	ε (%)	$\mathcal{D}_{raw}$ (%)	$\varepsilon\mathcal{D}^2_{raw}$ [%]
0.0-0.8	CMU	1.280 ± 0.027	3.8 ± 2.1	0.0040 ± 0.0033
	CMP	0.244 ± 0.011	21.4 ± 4.6	0.0151 ± 0.0054
	CMUP	0.231 ± 0.011	19.6 ± 4.5	0.0113 ± 0.0046
	CMX	0.294 ± 0.012	10.8 ± 4.2	0.0049 ± 0.0031
	BMU	1.356 ± 0.027	6.0 ± 2.0	0.0058 ± 0.0035
0.8-0.95	CMU	0.322 ± 0.013	11.5 ± 3.9	0.0126 ± 0.0045
	CMP	0.167 ± 0.009	24.9 ± 5.1	0.0127 ± 0.0046
	CMUP	0.151 ± 0.008	31.9 ± 5.3	0.0199 ± 0.0056
	CMX	0.173 ± 0.009	24.4 ± 4.9	0.0173 ± 0.0055
	BMU	0.191 ± 0.009	21.4 ± 4.9	0.0130 ± 0.0049
0.95-1.0	CMU	0.612 ± 0.017	20.6 ± 2.7	0.0405 ± 0.0082
	CMP	0.047 ± 0.004	36.7 ± 9.1	0.0062 ± 0.0032
	CMUP	0.450 ± 0.014	29.3 ± 3.0	0.0495 ± 0.0088
	CMX	0.284 ± 0.011	29.1 ± 3.8	0.0261 ± 0.0065
	BMU	0.399 ± 0.014	32.1 ± 3.2	0.0441 ± 0.0085
Sum	CMU	2.214 ± 0.034	–	0.0571 ± 0.0099
	CMP	0.457 ± 0.015	–	0.0340 ± 0.0078
	CMUP	0.832 ± 0.019	–	0.0807 ± 0.0114
	CMX	0.750 ± 0.019	–	0.0483 ± 0.0090
	BMU	1.946 ± 0.031	–	0.0628 ± 0.0104
Total	All	6.200 ± 0.056	–	0.283 ± 0.0219
Corrected				0.688 ± 0.053

Table 45: Efficiency, raw dilutions and $\varepsilon\mathcal{D}^2_{raw}$ in three likelihood bins for the muon-SVT data sample. $\varepsilon\mathcal{D}^2$ is obtained after summing over $p_T^{\rm rel}$ bins.

A.8 Correlations Between Likelihood Variables

Among the five variables used in the likelihood estimator, only ΔX and $\Delta\Phi$ are significantly correlated. Their correlation factor is $\sim 45\%$. Detailed correlation factors for each muon type are determined in Reference [35]. The correlation between ΔX and $\Delta\Phi$ for CMU muons from J/ψ decays is shown in Figure 94. It represents the mean value of $\Delta\Phi$ as a function of ΔX. To first order, the dependence is linear for values of ΔX less than 20 cm and becomes non-linear for ΔX larger than 20 cm. The correlation distribution is fit with a 5th order polynomial. To asses the effect of this correlation on the likelihood estimator, we attempt to "de-correlate" the two variables in the following way. For a given muon with measured ΔX and $\Delta\Phi$, we calculate $\Delta\Phi^{uncorr}$ by shifting $\Delta\Phi$ by an amount that corresponds to the measured ΔX as obtained from the curve in Figure 94. The new correlation factor between ΔX and $\Delta\Phi^{uncorr}$ drops from 45% to 2%. We use the "uncorrelated" pair ΔX, $\Delta\Phi^{uncorr}$ in the likelihood estimator and compare the new distribution with the likelihood distribution in which the correlated variables were used. The two distributions are shown in Figure 94. The distribution with uncorrelated variables shows a higher spike at one and a smaller tail towards zero resulting in a better muon identification

$\mathcal{L}$ cut (%)	5	50	90
Correlated (%)	0.6	8.0	27.9
Uncorrelated (%)	0.4	6.6	24.3
Ideal (%)	0.2	4.7	21.3

Table 46: The fraction of rejected real muons for different likelihood cuts for the three cases: ΔX and $\Delta\Phi$ correlated, uncorrelated and in the ideal case for toy Monte Carlo experiments.

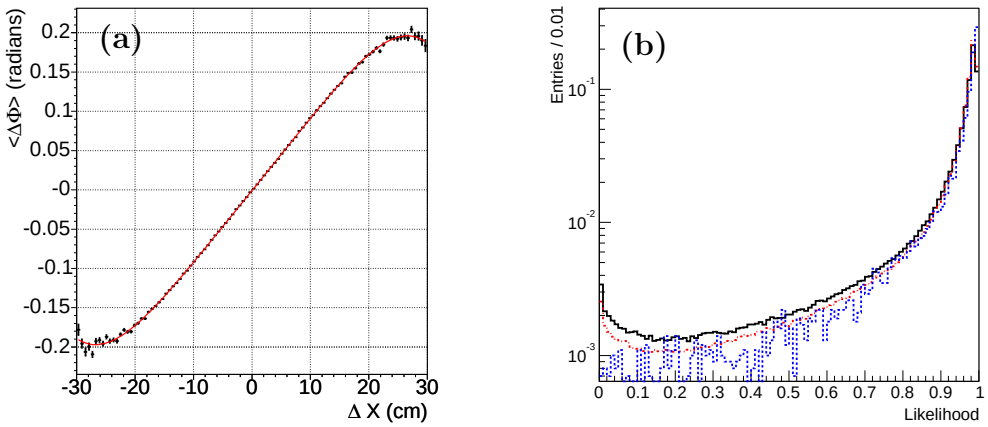

Figure 94: (a) Correlation between $\Delta\Phi$ and ΔX for CMU muons from $J/\psi \to \mu^+\mu^-$ decays shown as mean value of $\Delta\Phi$ along the vertical axis as function of ΔX on the horizontal axis. (b) Likelihood distribution for CMU muons from J/ψ as measured directly from data with correlated ΔX and $\Delta\Phi$ (solid), the same distribution after ΔX and $\Delta\Phi$ are 'uncorrelated' (dot-dashed) and the likelihood distribution obtained from toy Monte Carlo (dashed).

Another method of studying the correlation between the discriminating variables is to generate a toy Monte Carlo sample in which all five variables are randomly generated according to the fitted templates. The likelihood distribution obtained with this method corresponds to the ideal case in which all five variables are totally independent and uncorrelated. Their distributions are perfectly described by the fitted templates. The likelihood distribution obtained with this method is shown with a dashed line in Figure 94. The spike at one is significantly higher than the one corresponding to the distribution obtained from data. In addition, the tail is also strongly suppressed. Table 46 shows the fractions of rejected real muons for different likelihood cuts. The difference between the correlated and the uncorrelated case is a measure of the possible gain from considering the correlation between ΔX and $\Delta\Phi$ in the likelihood. Although the difference is not negligible, it is still small and would have a minor impact on the final tagging power of finding opposite side muons.

A.9 *K*-Factor Distributions

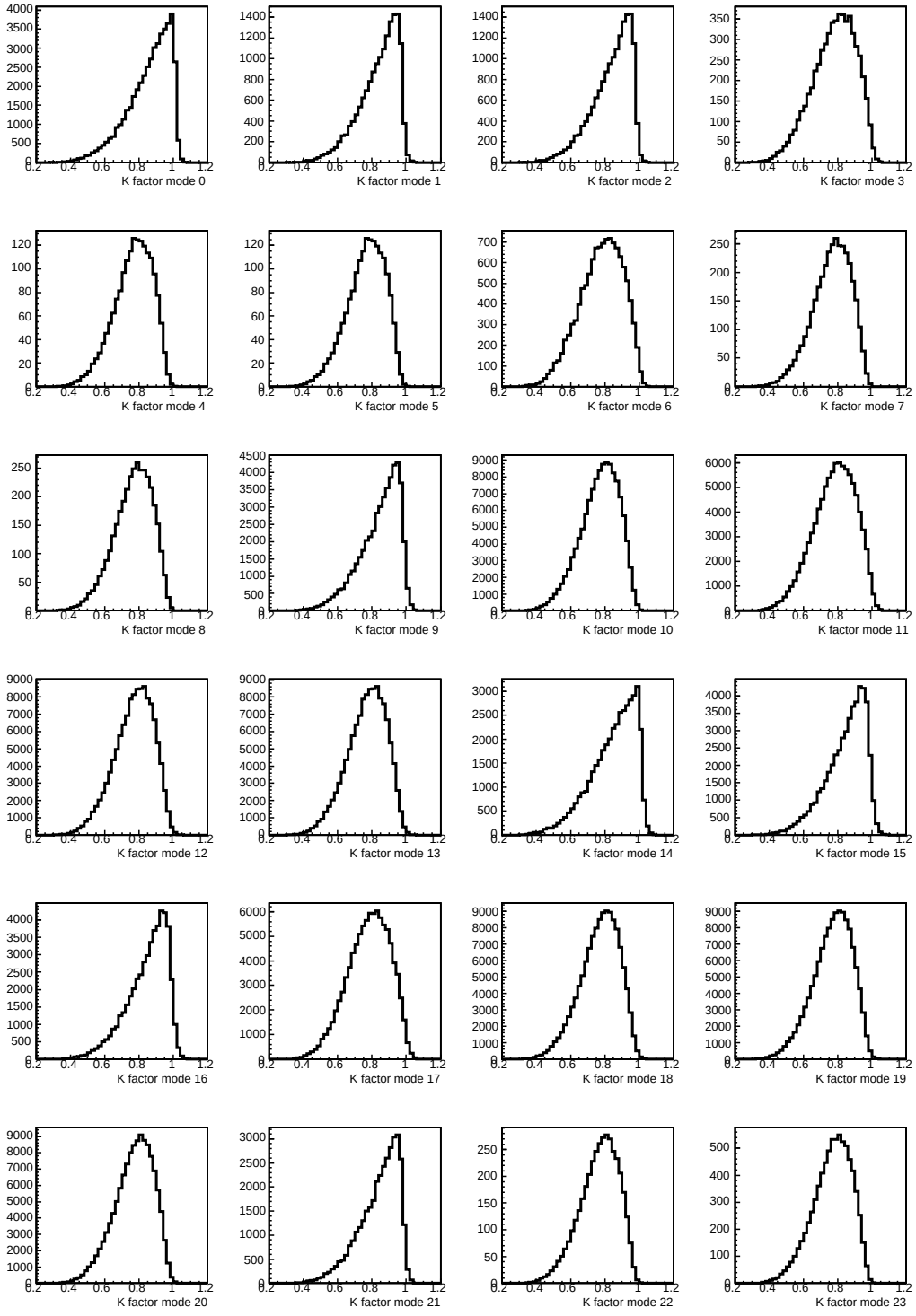

Figure 95: K-factor distributions for the 24 decay chains obtained from the realistic Monte Carlo simulation.

A.10 SVT Trigger Bias Parameterizations

Channel	p_0	p_1	p_2	f_1	λ_1	f_2	λ_2	p_{offset}
0	−0.001116	−1.020901	−0.000061	−17.378164	16.335343	17.609053	78.865740	0.263771
1	−0.005438	−0.853987	−1.930660	−15.286705	16.335343	23.524861	78.865740	0.303549
2	−0.005438	−0.853987	−1.930660	−15.286705	16.335343	23.524861	78.865740	0.303549
3	−0.004733	−0.878096	−0.495597	−15.021563	16.335343	26.639509	78.865740	0.332083
4	−0.001670	−1.036731	0.594412	−17.385505	16.335343	21.389822	78.865740	0.274730
5	−0.001670	−1.036731	0.594412	−17.385505	16.335343	21.389822	78.865740	0.274730
6	0.000254	−0.996314	−0.178622	−16.95246	16.335343	19.492472	78.865740	0.279021
7	−0.002131	−1.069154	1.113166	−17.900133	16.335343	19.933986	78.865740	0.255790
8	−0.002131	−1.069154	1.113166	−17.900133	16.335343	19.933986	78.865740	0.255790
9	−0.009055	−1.031776	−0.173983	−16.875573	16.335343	32.885300	78.865740	0.283834
10	−0.009205	−1.073081	0.637103	−17.536759	16.335343	30.524130	78.865740	0.264111
11	−0.009158	−1.052305	−0.001458	−17.425921	16.335343	29.350328	78.865740	0.255147
12	−0.009279	−1.060635	1.357332	−17.322503	16.335343	33.217036	78.865740	0.289640
13	−0.009279	−1.060635	1.357332	−17.322503	16.335343	33.217036	78.865740	0.289640
14	−0.008895	−1.119704	2.014460	−18.530618	16.335343	31.807476	78.865740	0.250348
15	−0.007713	−1.086215	−0.316557	−18.083313	16.335343	27.732622	78.865740	0.221549
16	−0.007713	−1.086215	−0.316557	−18.083313	16.335343	27.732622	78.865740	0.221549
17	−0.008799	−1.059117	0.665851	−17.561364	16.335343	30.974612	78.865740	0.262554
18	−0.008077	−1.126006	1.022371	−18.595081	16.335343	28.711824	78.865740	0.225515
19	−0.008077	−1.126006	1.022371	−18.595081	16.335343	28.711824	78.865740	0.225515
20	−0.008646	−1.086980	0.925053	−18.156598	16.335343	29.078997	78.865740	0.240811
21	−0.009109	−1.031679	−0.392569	−17.015515	16.335343	30.778366	78.865740	0.270440
22	−0.009330	−1.069113	0.395545	−17.667903	16.335343	28.487848	78.865740	0.248732
23	−0.008433	−1.077233	1.110775	−17.942099	16.335343	29.771884	78.865740	0.251540

Table 47: Fit parameters for the SVT-bias curves for each of the 24 decay channels.

A.11 Convolution Integrals

The calculation of the integrals involved in Equation (83) is not trivial. There are two convolution integrals that have to be evaluated, the convolution between the true decay time distribution $\frac{1}{2\tau}\mathrm{e}^{-t/\tau}\,(1-\cos\,\Delta m_d\,t)$ and the Gaussian detector resolution function $\mathcal{G}(t'-t,\sigma)$ followed by a convolution with the momentum resolution function $\mathcal{F}(K)$. The integral in Equation (83) can be explicitly written as:

$$\int_{K_{min}}^{K_{max}} \mathcal{F}(K) K \mathcal{E}(Kt^*) dK = \int_0^\infty \frac{1}{\tau} e^{-Kt/\tau} \frac{1}{2}(1 \pm \cos \Delta m_d K t) \frac{1}{\sqrt{2\pi}\sigma^*} e^{-(t^*-t)^2/2\sigma^{*2}} dt, \tag{152}$$

where σ^* is the uncertainty on the pseudo proper decay time t^* given in Equation (70). For simplicity we omit the factors $\mathcal{A}\,\mathcal{S}_{\mathcal{D}}\,\mathcal{D}$ which multiply the cosine term and the factor s_σ which accounts for underestimated vertex resolution as they contribute trivially to the evaluation of the integrals.

We first calculate the convolution between the decay time distribution and the Gaussian detector resolution function. We evaluate the first term as:

$$\int_0^\infty e^{-Kt/\tau} e^{-(t^*-t)^2/2\sigma^{*2}} dt = e^{K(K\sigma^{*2}-2\tau t^*)/2\sigma^{*2}} \sqrt{\frac{\pi}{2}}\,\sigma Erfc\frac{K\sigma^{*2}-\tau t^*}{\sqrt{2}\sigma^*\tau}, \tag{153}$$

with $Erfc(x) = 1 - Erf(x)$, where $Erf(x)$ is the error function. The evaluation of the second term is more difficult as in involves complex functions:

$$\int_0^\infty e^{-Kt/\tau} \cos(\Delta m_d K t) e^{-(t^*-t)^2/2\sigma^{*2}} dt = \\ Re\left\{Exp\left(-\frac{K(-i+\Delta m_d\tau)(K\sigma^{*2}(-i+\Delta m_d\tau)+2i\tau t^*)}{2\tau^2}\right)\sqrt{\frac{\pi}{2}}\sigma^* \right. \\ \left. \times Erfc\left(\frac{K\sigma^*}{\sqrt{2}\tau} - \frac{t^*}{\sqrt{2}\sigma^*} + i\frac{\Delta m_d K\sigma^*}{\sqrt{2}}\right)\right\} \qquad (154)$$

We use the complex functions implemented in the `RooFit` package [58] to evaluate Equation (154). In particular, we utilize the function `evalCerfRe`, which in turn calls `FastComplexErrFuncRe` and `evalCerfApprox`.

The second convolution integral with the momentum resolution function $\mathcal{F}(K)$ is performed numerically using the K-factor distributions discussed in Section 4.3.

A.12 Amplitude Scans in Each D_s^- Decay Mode

We perform separate amplitude scans for each of the three D_s decay modes. Figure 96 shows the results of these scans including statistical errors only. We find that the $D_s^- \to \phi\pi^-$ mode is the dominant one as expected since it has the largest number of events and the best signal to background events ratio.

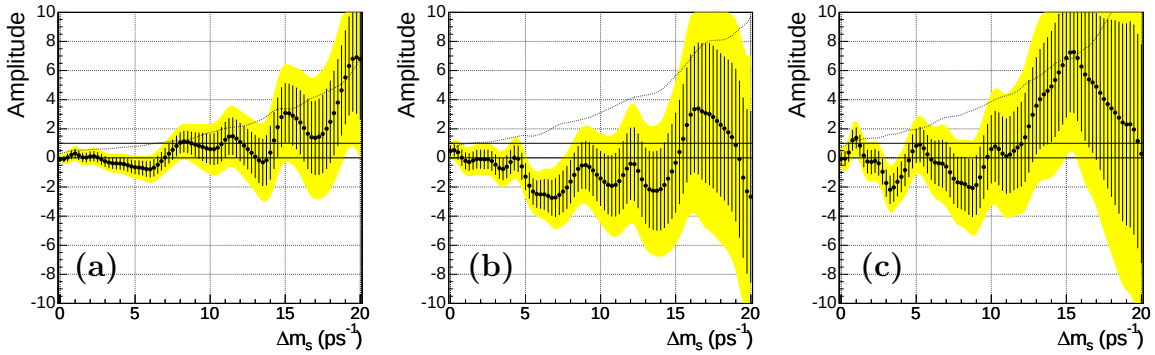

Figure 96: Separate amplitude scans for (a) $D_s^- \to \phi\pi^-$, (b) $D_s^- \to K^{*0}K^-$and (c) $D_s^- \to \pi^+\pi^-\pi^-$.

Decay Mode	Limit (ps^{-1})	Sensitivity (ps^{-1})
$D_s^- \to \phi\pi^-$	6.9	6.3
$D_s^- \to K^{*0}K^-$	0.0	1.3
$D_s^- \to \pi^+\pi^-\pi^-$	0.3	0.5

Table 48: Limit and sensitivity results from each D_s decay mode.

A.13 Amplitude Scan Result

CDF Run II Semileptonic Preliminary (355 pb^-1)

	<--- stat only --->		<--- stat+syst --->		<-- stat+syst -->
Delta_ms	Amplitude	+- error	Amplitude	+- error	Amplitude/error
0.000	0.015	0.159	0.015	0.223	0.066
0.250	0.039	0.182	0.039	0.239	0.163
0.500	0.134	0.223	0.134	0.270	0.497
0.750	0.291	0.248	0.291	0.290	1.004
1.000	0.332	0.260	0.332	0.299	1.107
1.250	0.168	0.267	0.168	0.305	0.549
1.500	0.011	0.272	0.011	0.309	0.034
1.750	-0.029	0.278	-0.029	0.314	-0.092
2.000	0.011	0.287	0.011	0.321	0.036
2.250	0.036	0.294	0.036	0.327	0.111
2.500	-0.039	0.298	-0.039	0.331	-0.117
2.750	-0.228	0.305	-0.228	0.336	-0.678
3.000	-0.446	0.312	-0.446	0.342	-1.304
3.250	-0.585	0.317	-0.585	0.347	-1.683
3.500	-0.616	0.327	-0.616	0.356	-1.728
3.750	-0.563	0.342	-0.563	0.370	-1.521
4.000	-0.467	0.357	-0.467	0.384	-1.218
4.250	-0.390	0.372	-0.390	0.397	-0.982
4.500	-0.388	0.386	-0.388	0.411	-0.942
4.750	-0.464	0.398	-0.464	0.422	-1.098
5.000	-0.579	0.406	-0.579	0.430	-1.348
5.250	-0.709	0.415	-0.709	0.439	-1.615
5.500	-0.849	0.429	-0.849	0.452	-1.880
5.750	-0.979	0.448	-0.979	0.470	-2.081
6.000	-1.040	0.473	-1.040	0.495	-2.102
6.250	-0.997	0.500	-0.997	0.520	-1.917
6.500	-0.867	0.520	-0.867	0.540	-1.607
6.750	-0.700	0.536	-0.700	0.556	-1.259
7.000	-0.517	0.556	-0.517	0.576	-0.898
7.250	-0.334	0.577	-0.334	0.596	-0.561
7.500	-0.165	0.598	-0.165	0.617	-0.267
7.750	0.032	0.619	0.032	0.638	0.050
8.000	0.247	0.637	0.247	0.656	0.376
8.250	0.393	0.651	0.393	0.670	0.586
8.500	0.424	0.667	0.424	0.686	0.618
8.750	0.383	0.689	0.383	0.707	0.542
9.000	0.321	0.718	0.321	0.735	0.437
9.250	0.274	0.753	0.274	0.770	0.356

9.500	0.238	0.789	0.238	0.806	0.295
9.750	0.207	0.823	0.207	0.840	0.246
10.000	0.184	0.856	0.184	0.872	0.211
10.250	0.162	0.887	0.162	0.903	0.179
10.500	0.170	0.920	0.170	0.936	0.182
10.750	0.268	0.959	0.268	0.974	0.275
11.000	0.457	0.997	0.457	1.012	0.452
11.250	0.679	1.027	0.679	1.043	0.651
11.500	0.829	1.050	0.829	1.065	0.778
11.750	0.856	1.067	0.856	1.083	0.790
12.000	0.792	1.088	0.792	1.104	0.718
12.250	0.679	1.116	0.679	1.132	0.600
12.500	0.513	1.150	0.513	1.166	0.440
12.750	0.322	1.184	0.322	1.200	0.269
13.000	0.126	1.222	0.126	1.237	0.102
13.250	-0.043	1.265	-0.043	1.280	-0.034
13.500	-0.112	1.315	-0.112	1.330	-0.084
13.750	0.015	1.379	0.015	1.394	0.011
14.000	0.411	1.457	0.411	1.472	0.279
14.250	1.071	1.538	1.071	1.553	0.690
14.500	1.822	1.603	1.822	1.617	1.126
14.750	2.463	1.642	2.463	1.656	1.487
15.000	2.896	1.654	2.896	1.668	1.736
15.250	3.123	1.655	3.123	1.670	1.870
15.500	3.203	1.666	3.203	1.681	1.905
15.750	3.212	1.695	3.212	1.710	1.879
16.000	3.086	1.748	3.086	1.763	1.751
16.250	2.833	1.820	2.833	1.835	1.544
16.500	2.539	1.895	2.539	1.910	1.329
16.750	2.309	1.963	2.309	1.978	1.168
17.000	2.200	2.027	2.200	2.042	1.078
17.250	2.170	2.094	2.170	2.108	1.029
17.500	2.204	2.164	2.204	2.178	1.012
17.750	2.336	2.238	2.336	2.251	1.038
18.000	2.571	2.318	2.571	2.331	1.103
18.250	2.862	2.407	2.862	2.420	1.182
18.500	3.207	2.505	3.207	2.519	1.273
18.750	3.625	2.616	3.625	2.629	1.379
19.000	4.007	2.743	4.007	2.756	1.454
19.250	4.174	2.891	4.174	2.903	1.438
19.500	4.000	3.051	4.000	3.063	1.306
19.750	3.592	3.209	3.592	3.221	1.115
20.000	3.216	3.364	3.216	3.375	0.953

References

[1] S. Eidelman *et al.* [Particle Data Group], *Review of particle physics*, Phys. Lett. B **592**, 1 (2004).

[2] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967);
A. Salam, p. 367 of *Elementary Particle Theory, ed. N. Svartholm (Almquist and Wiksells, Stokholm, 1969)*;
S. L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. **D2**, 1285 (1970).

[3] N. Cabibbo, *Phys. Rev. Lett.* **10**, 531 (1963);
M. Kobayashi and T. Maskawa, *Prog. Teor. Phys.* **49**, 652 (1973); *Phys. Rev. Lett.*, **51**, 1945 (1983).

[4] L.-L. Chau and W.-Y. Keung, *Phys. Rev. Lett.* **53**, 1802 (1984).

[5] L. Wolfenstein, *Phys. Rev. Lett.* , **51**, 1945 (1983).

[6] C. Jarlskog, Phys. Rev. Lett. **55**, 1039 (1985).

[7] `http://www.slac.stanford.edu/xorg/ckmfitter/help/ckmfitter_help.html`

[8] B. Aubert, *et al.* [BaBar Collab.] Phys. Rev. Lett. **89**, 201802 (2002).

[9] K. Abe *et al.*, Belle-CONF-0353, LP'03 (2003).

[10] A. Bornheim, *et al.* [CLEO Collab.], Phys. Rev. Lett. **88**, 231803 (2002).

[11] B. Aubert, *et al.* [BaBar Collab.], arXiv:hep-ex/0207081.

[12] K. Abe *et al.*, [BELLE Collab.] Belle-CONF-0325, 2003.

[13] M. Battaglia *et al.*, hep-ph/0304132.

[14] A. J. Buras, W. Slominski and H. Steger, Nucl. Phys. **B245**, (1984).

[15] C. Gay, hep-ex/0103016.

[16] T. Inami, C.S. Lim, Prog. Theor. Phys. **65**, 297 (1981).

[17] A. Kronfeld, hep-lat/0310063v1.

[18] M. Gronau, A. Nippe and J. Rosner, Phys. Rev. Lett. **80**, 2057 (1998);
F. Abe *et al.* [CDF Collab.], Phys. Rev. Lett. **80**, 2057 (1998).

[19] H.G. Moser and A. Roussarie, NIM A384 (1997), 491-505.

[20] R. Blair *et al.*, *The CDF-II detector: Technical design report*, FERMILAB-PUB-96/390-E (1996).

[21] D. Acosta *et al.* [CDF Collab.], Phys. Rev. D **71**, 032001 (2005).

[22] A. Sill *et al.*, Nucl. Inst. Meth. A **447**, 1 (2000).

[23] T. Affolder *et al.*, Nucl. Inst. Meth. A **526**, 249 (2004).

[24] C. Grozis *et al.*, Int. J.Mod. Phys. A16S1C:1119-1121, 2001.

[25] H. Wenzel, *Tracking in SVX*, CDF note 1790.

[26] G. Apollinari *et al.*, Nucl. Instrum. Meth. A412:515-526, 1998.

[27] T. Moulik, *Offline Central PreRadiator Reconstruction in RunIIa* , CDF note 6192.

[28] A. Bardi *et al.*, Nucl.Instrum.Meth.A409:658-661,1998.

[29] E. J. Thomson *et al.*, IEEE Trans. Nucl.Sci. **49**, 1063 (2002).

[30] R. G. Wagner, *Understanding and Using Lshr*, CDF note 6249.

[31] M. Jones *et al.*, *Sample Composition of the lepton+SVT Triggers*, CDF note 6480.

[32] M. D. Peters *et al.*, *Soft Muon B-Flavor Tagging in Run 1B*, CDF note 3808.

[33] S. Giagu *et al.*, *The CharmMods/DFinder Reconstruction Package*, CDF note 6158.

[34] J. Marriner, *Secondary Vertex Fit with Mass and Pointing Constraints (CTVMFT)*, CDF note 1996.

[35] L. Cerrito and A. Taffard, *A Soft Muon Tagger for Run II*, CDF note 6305.

[36] J. Bellinger *et al.*, *A Guide to Muon Reconstruction for Run 2*, CDF note 5870.

[37] T. LeCompte, *Efficiency of Momentum Dependent Muon Matching Cuts*, CDF note 6114.

[38] H.-C. Fang *et al.*, *SemiLeptonicB/LeptonSvtSel - An offline filter for lepton-SVT data*, CDF note 6326.

[39] `/CdfCode/source/JetUserObjects/JetUserObjects/ConeClusteringAlg.hh`

[40] P. Sphicas, *Combining Flavor Taggers*, CDF note 3425.

[41] G. Bauer *et al.*, *Improved Jet Charge Tagger for Summer Conferences 2004*, CDF public note 7131,
`http://www-cdf.fnal.gov/physics/new/bottom/040812.blessed-jqt-lsvt`

[42] D. Usynin *et al.*, *B Flavor Identification Using Opposite Side Muons*, CDF note 6483.

[43] G. Giurgiu *et al.*, *Muon B Flavor Tagging - A Likelihood Approach*, CDF public note 7043,
`http://www-cdf.fnal.gov/physics/new/bottom/072204.blessed-like-mu-tag/`

[44] G. Giurgiu *et al.*, *Updated Likelihood Muon Tagger*, CDF note 7644.

[45] V. Tiwari *et al.*, *Likelihood Based Electron Tagging*, CDF public note 7121,
`http://www-cdf.fnal.gov/physics/new/bottom/072204.blessed-like-mu-tag/`

[46] G. Comez-Ceballos *et. al*, *Measurement of B^0 Oscillations Using Same-Side Tagging in Semileptonic B Decays*, CDF note 7011.

[47] The CDF B Group's Physics Analysis Kernel Subgroup Webpage, available at `http://www-cdf.fnal.gov/internal/physics/bottom/bpak/`.

[48] B_s mixing group (M. Casarsa *et al.*), *Performance of the Event-by-Event Primary Vertex Finder on B Physics Data at CDF*, CDF note 7512.

[49] B_s mixing group (A. Belloni *et al.*), *B^0 Mixing and Taggers Performance in Semileptonic B Decays*, CDF note 7461.

[50] S. Gromoll and Ch. Paus, *A Heavy Quark Generator*, CDF note 5985.

[51] P. Nason, S. Dawson and R.K. Ellis, Nucl. Phys. **B303** (1988) 607; Nucl. Phys **B327** (1989) 49;
M. Mangano, P. Nason and G. Ridolfi, Nucl. Phys. **B373** (1992) 295.

[52] C. Peterson *et all.*, Phys. Rev. **D27** (1983) 105.

[53] W. Bell *et al.*, *User Guide For EvtGen @ CDF*, CDF note 5618.

[54] E. Gerchtein and M. Paulini, in *Proceeding of the Conference for Computing in High-Energy and Nuclear Physics (CHEP 03), La Holla, California, 2003*, C0303241, TUMT005 (2003).

[55] K. Anikeev and C. Paus, *An Update of B Lifetime Measurements Using Exclusively reconstructed Decays $B \to J/\psi X$ in Run II Data*, CDF note 6550.

[56] G. Punzi, [e-print physics/0401045], "Comments on Likelihood fits with variable resolution".

[57] F. James, *MINUIT - Function Minimization and Error Analysis Referrence Manual*, CERN Program Library Long Writeup D506.

[58] `http://roofit.sourceforge.net/`

[59] B_s mixing group (A. Belloni *et al.*), *Study of B_s^0 Oscillations in $B_s^0 \to D_s \ell X$*, CDF note 7472.

[60] S. Giagu, private comunication.

[61] B_s mixing group (A. Belloni *et al.*), *Scale Factors for Proper Time Uncertainties*, CDF note 7500.

[62] B_s mixing group (G. Punzi), *Blinding procedure for the measurement of* x_s, CDF note 7544.

[63] K. Anikeev, G. Bauer and C. Paus, *B Meson Lifetime Measurements Using Exclusively Reconstructed Decays* $B \to J/\psi X$, CDF note 6266.

[64] B_s mixing group (A. Belloni *et al.*), *Study of* B_s *Oscillations in* $B_s^0 \to D_s^- \pi^+$, CDF note 7481, CDF Public note 7535, `http://www-cdf.fnal.gov/physics/new/bottom/050310-bsmix-hadronic/`.

[65] B_s mixing group, *Combined Results from* B_s *Mixing Analyses*, CDF note 7555, `http://www-cdf.fnal.gov/physics/new/bottom/050310.bsmix-combined/`.

[66] A. Heister *et al.* [ALEPH Collab.], Eur.Phys.J.C29:143-170, 2003.

[67] B_s mixing group, *Long-Term Sensitivity Projections for* B_s^0 *Mixing*, CDF note 7671.

[68] B_s mixing group (P. Catastini *et. al*), *Same Side Kaon Tagging Studies*, CDF note 7496.