Observation of $\pi - B$ meson Charge-flavor Correlations and Measurement of Time Dependent $B^0\bar{B}^0$ Mixing in $p\bar{p}$ Collisions

by

Petar Maksimović

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1998

Author . . . .
Department of Physics
November 17, 1997

Certified by . . . .
Paraskevas A. Sphicas
Professor
Thesis Supervisor

Accepted by . . . .
George Koster
Professor
Chairman, Departmental Committee on Graduate Students

Observation of $\pi - B$ meson Charge-flavor Correlations and Measurement of Time Dependent $B^0\bar{B}^0$ Mixing in $p\bar{p}$ Collisions

by

Petar Maksimović

Submitted to the Department of Physics
on November 17, 1997, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

We present a study of time-dependent $B^0 - \bar{B}^0$ mixing in $p\bar{p}$ collisions at 1.8 TeV using 110 pb^{-1} collected with the CDF detector at the Fermilab Tevatron Collider. B mesons are partially reconstructed using the semileptonic decays $B^0 \to \ell^+ D^{(*)-} X$ and $B^+ \to \ell^+ \overline{D^0} X$ (and their charge conjugates). B meson-charged pion correlations are used in order to determine the flavor of the B meson at $t = 0$. Such correlations are expected to arise from pions produced in the fragmentation chain and also from B^{**} decays. We measure the efficiency and purity of this flavor tagging method for both charged and neutral B mesons.

Thesis Supervisor: Paraskevas A. Sphicas
Title: Professor

Observation of $\pi - B$ meson Charge-flavor Correlations and Measurement of Time Dependent $B^0\bar{B}^0$ Mixing in $p\bar{p}$ Collisions

by

Petar Maksimović

Submitted to the Department of Physics
on November 17, 1997, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

We present a study of time-dependent $B^0 - \bar{B}^0$ mixing in $p\bar{p}$ collisions at 1.8 TeV using 110 pb^{-1} collected with the CDF detector at the Fermilab Tevatron Collider. B mesons are partially reconstructed using the semileptonic decays $B^0 \to \ell^+ D^{(*)-} X$ and $B^+ \to \ell^+ \overline{D^0} X$ (and their charge conjugates). B meson-charged pion correlations are used in order to determine the flavor of the B meson at $t = 0$. Such correlations are expected to arise from pions produced in the fragmentation chain and also from B^{**} decays. We measure the efficiency and purity of this flavor tagging method for both charged and neutral B mesons.

Thesis Supervisor: Paraskevas A. Sphicas
Title: Professor

Acknowledgments

"It ain't no detector, it's a Borg spaceship!" exclaimed a friend of mine when she saw the CDF detector. The CDF detector (who by now has almost achieved a consciousness of its own), as well as all people who built it, kept it alive and well-supplied with the colliding beams, deserve special thanks. Without them, this thesis would not be possible. Within the CDF collaboration, I would like to single out the B group crowd, who provided the intellectual nourishment and transmitted the technical expertise as well as a craving for the CKM measurements to us grad students.

I thank Paris Sphicas for leadership, wisdom, critique and support, for giving me an opportunity to learn from him and for coming up with such a cool thesis. I thank Fritz DeJongh for advice and help, and excellent ideas that are interspersed throughout this analysis. I also thank Ken who pioneered the Same Side Tagging infrastructure, and Gerry and Jeff who helped him. On a textual front, I am especially grateful to Gerry, whose opinions and suggestions had a tremendous impact on the style, the tone and the structure of the thesis, as well as of the two papers accompanying it.

The home front is another matter, and the bizarre and superb life in the MIT house with Wasiq, Troy and Tushar was a 'safe haven' from all things sane. In the absence of the real MIT atmosphere, we somehow managed to occasionally fake it, from gedanken experiments in liberal politics, to "the game of 24," to grandiose parties, to the poetry of Rumi and the language of the Pennsylvania Dutch. I will always also be grateful to Benn, Dejan and Sandra, Paul, Kara, Nancy, Simona, Rocio, Owen, Marc, Andreas, Bill, Mark and Mindy, Randy, Steve, Joe and everybody else for making Fermilab vibrant with life and friendship.

And I will always remember George, who was the kindest and smartest of us all. His tragic death was a real loss to Particle Physics.

Back in Belgrade, I'd like to thank the staff of the Department of Physics, who prepared me well for MIT, and the staff of Matematička Gimnazija (my high school), who prepared me well for Life. I am especially proud to be a student of Nataša

Čaluković, who showed me how to enjoy Physics.

Also, I would like to thank the students who protested last winter in the streets of Belgrade, while I was busy writing this thesis. By this thesis I pay a tribute to their struggle.

Finally, I would like to thank Ana for being equally busy and showing me first-hand how logistically hard it is to be in love with a physicist.

But and above all, I am indebted to my parents, who made me the way I am: to my mother who served me stories of black holes with cauliflower and injected art into my veins, and to my father who engaged me in an eternal dialogue and who was always worried whether I would be able to take care of myself, but did not live to see it happen. To two of you, and to my sister Jelena, whose star is still rising, I dedicate this thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The concept of symmetry has played a fundamental role in modern physics. Particle physics is especially rich in symmetries – some seemingly simple, the others highly non-trivial.

The symmetries that appeared the simplest, however, turned out to be the most intriguing. Consider the following operations:

- charge conjugation, C, representing replacing a particle by its antiparticle,
- parity, P, corresponding to looking in a mirror which reverses all three spatial coordinates, and
- time reversal, T.

One would naively think that physics interactions ought to be invariant under C, P and T, *i.e.*, that C, P and T are *valid symmetries.* In the electromagnetic and strong interactions, each of C, P and T is indeed conserved. This was believed to be true for the weak interactions as well, until in 1956 T. D. Lee and C. N. Yang [1] concluded from the available data that weak decays may violate P. This hypothesis was soon confirmed in an experiment by C. S. Wu [2]. The weak interactions do not conserve C either: applying the charge conjugation to a left-handed neutrino results in a left-handed antineutrino, a particle that does not exist. However, a combination of C and P, the CP, produces a right-handed antineutrino, which does exist.

Hopes that CP is a valid symmetry were dashed in 1964, when J. Cronin and V. Fitch [3] showed that CP is violated in K^0 decays at the level of a fraction of a percent. Since all Quantum Field Theories in existence demand that the combination of all three operations, CPT, is conserved (a conjecture that, so far, has been supported by experiment [5]), the violation of CP must be counterbalanced by a violation of T. In conclusion, the weak interaction violates all three symmetries: C, P and T. Nature distinguishes between particles and antiparticles, between left and right and determines the direction of the time arrow.

At present, the Standard Model of particle interactions describes the world as we know it extremely well. Its electro-weak part incorporates the *large* violation of C and P by construction. However, the violation of CP at a *very small* level, although technically easily accommodated by the CKM matrix (section 1.1), is one of its least understood aspects, probably on par with the spontaneous electro-weak symmetry breaking. Overall, the Standard Model is deemed unsatisfactory, and a dream of every particle physicist is to get a glimpse of the physics beyond it, sometimes called *the New Physics*. Future measurements of the Standard Model parameters, especially those plausibly responsible for CP violation, will check its internal consistency, and hopefully provide some insight on the New Physics. Indeed, any theory more fundamental than the Standard Model will have to predict many of the Standard Model parameters, so measuring them provides a set of constraints on what this new theory can be. Performing these measurements is therefore not only intriguing, but also necessary.

This thesis presents a measurement constraining one of the Standard Model parameters (CKM element V_{td}, using $B^0\bar{B}^0$ mixing). The measurement is done in a novel way, using a technique which can be employed in future experiments likely to shed some light on the problem of CP violation. The result presented herein is fairly competitive in its own right, however the method employed in the measurement probably surpasses it in potential impact. Throughout the pages that follow one should keep in mind that the tools presented are as important as the physics result itself.

In this chapter, the basics of the Standard Model framework for accommodating

CP violation (via the CKM matrix) is described.[1] In addition, it is shown that the the measurements involving B mesons (the mesons with a b-quark bound to a light quark) are the most suitable for measuring many parameters of the CKM matrix.

The second chapter focuses on measuring $B^0\bar{B}^0$ oscillation, especially at a proton-antiproton collider. Main complications caused by the hadronic environment are introduced. The relevant parts of the CDF detector are described in chapter 3. Chapters 4-8 provide the detailed description of the measurement. In chapter 9 we return to the prospects of the CP violation measurements using the methods presented in the previous chapters.

1.1 CP violation in the Standard Model

It is an experimental fact that there is coupling between the quark generations – otherwise the lightest strange particle (the kaon, K) would be absolutely stable, and so would be the lightest beautiful particle, the B-meson, and we know that is not the case.

Cross-generational coupling was first introduced in 1963 by Cabibbo [6], who suggested that the $d \to u + W^-$ vertex carries a multiplicative factor of $\cos\theta_c$, whereas the $s \to u + W^-$ vertex carries a factor of $\sin\theta_c$. The second one is weaker, hence θ_c is small.[2] This model was fairly successful, except that it allowed the K^0 to decay to $\mu^+\mu^-$ via the Feynman diagram shown in fig. 1-1a. According to Cabibbo's prescription, the width would be , $(K^0 \to \mu^+\mu^-) \sim \sin\theta_c \cos\theta_c$, dramatically above the experimentally set limit.

The Cabibbo model was rescued in 1970 by Glashow, Iliopoulos and Maiani [7], who extended it to include another quark called *charm* (and denoted by c). In this model, known as the GIM mechanism, the vertices $d \to c + W^-$ and $s \to c +$

---

[1] The cosmological implication of CP violation is out of the scope of this thesis, and as such is barely mentioned, but the connection between the CP violation and the prevalence of matter in the Universe gives this problem quite a special place in contemporary physics.

[2] Experimentally, $\theta_c = 13.1^o$.

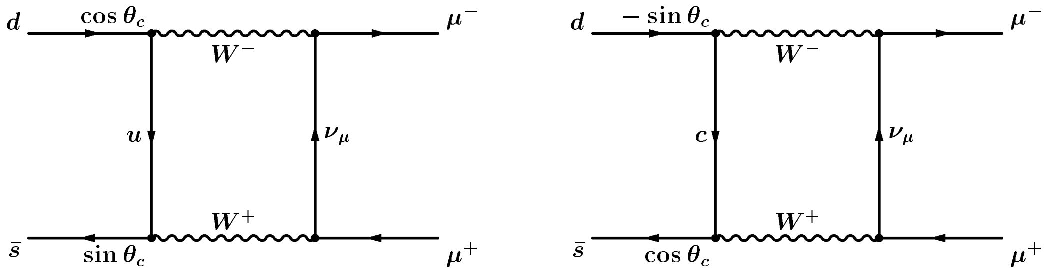


Figure 1-1: Two contributions to the decay $K^0 \to \mu^+\mu^-$ showing the factors present at the quark vertices. If only the diagram (a) were present, the decay rate would be far in excess of the observed rate. The second contribution cancels most of the first; if would cancel all if the c quark had the same mass as the u quark. This cancelation is an illustration of the Glashow-Iliopoulos-Maiani (GIM) mechanism.

W^- carry factors of $-\sin\theta_c$ and $\cos\theta_c$ respectively, so that the superposition of Feynman diagrams with the virtual u and c quarks cancel (fig. 1-1a,b), and the width $\Gamma(K^0 \to \mu^+\mu^-)$ equals almost zero. In early 1970's, the GIM mechanism seemed a little extravagant, until the J/ψ (a bound state of $c\bar{c}$) was discovered in 1974.

1.1.1 The quark mixing

The GIM mechanism suggests that, instead of the physical quarks d and s, the "correct" states to use in the weak interactions are d' and s', given by

$$d' = (\cos\theta_c)d + (\sin\theta_c)s \qquad (1.1)$$

$$s' = (-\sin\theta_c)d + (\cos\theta_c)s \qquad (1.2)$$

This phenomenon is called quark mixing. Equation 1.2 can be rewritten using so-called *mixing matrix*:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \qquad (1.3)$$

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

$$\begin{pmatrix} u \\ d' \end{pmatrix} \text{ and } \begin{pmatrix} c \\ s' \end{pmatrix}$$

in exactly the same way that they couple to the lepton pairs,

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix} \text{ and } \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$$

In Cabibbo-GIM two-generation model, the mixing matrix is just a rotation of the quark basis by the Cabibbo angle, θ_c (where $\sin\theta_c \approx 0.22$). Decays that involve the factor of $\sin\theta_c$ are said to be 'Cabibbo suppressed'.

In 1973, Kobayashi and Maskawa [8] generalized GIM to three generations of quarks (proposing the third generation even before the second one was complete). Their motivation was to explain CP violation, and concluded that since a complex phase can always be redefined in a 2×2 matrix, one needs a 3×3 matrix and thus there ought to be more than two generations of quarks.[3] This hypothesis was proven by later discoveries of the τ lepton in 1975, followed by bottom (b) and top (t) quarks in 1976 and 1995, respectively.

---

[3] In order to provide for CP violation, one needs a complex term in the interaction $J^\dagger_\mu J^\mu$ where $J_\mu = \overline{U}\gamma_\mu V(1-\gamma_5)D$ is the weak current. If there are n families, U represents the column of n charge 2/3 quarks ("upper" members of quark doublets) and D the column of n charge $-1/3$ quarks ("lower" members of quark doublets). The matrix V is unitary and has n^2 complex or $2n^2$ real parameters. Unitarity imposes the conditions $V_{ij}V^*_{kj} = \delta_{ik}$, which give $n(n-1)/2$ complex constraints for $i \neq k$ and n real constraints for $i = k$. Altogether there are n^2 remaining free parameters in V.

It is possible to eliminate some of the complex phases in V by redefining the phases of the $2n$ quark fields. Changing n fields in U or n fields in D by the same phase does not change $J^\dagger_\mu J^\mu$ so $2n-1$ phases from V can be eliminated in this way. Thus the number of real parameters characterizing V is $n^2 - 2n + 1 = (n-1)^2$.

For two families this gives just one parameter, which is the Cabibbo angle θ_c. For three families there are 4 parameters. If V were purely real it would be a 3×3 rotation matrix, which is determined by three real parameters. Thus the fourth parameter of V must necessarily introduce a complex component into V, one that cannot be absorbed into a redefinition of the quark fields.

1.1.2 The Cabibbo-Kobayashi-Maskawa (CKM) matrix

In general, the eigenstates of the weak interactions (quarks that interact with the W and Z bosons) and the strong interaction (those that interact with gluons) need not be the same. By convention, the "lower" part of each quark doublet is assumed to be the superposition of the mass states. We define:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_L \equiv V \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \tag{1.4}$$

Here the subscript L indicates the left-handed quarks. Then, the quark doublets to be used in the weak interactions are

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L.$$

The $W^{\pm}$ couples u_L to d'_L, c_L to s'_L and t_L to b'_L. The matrix V in eq. (1.4) is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix, and by construction is a unitary transformation in the quark space, changing from the "mass" basis (the quark fields d_L, s_L and b_L) to the "weak" basis (the quark fields d'_L, s'_L and b'_L).

The three-generation CKM matrix V is a 3×3 unitary matrix:

$$V \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.5}$$

In Wolfenstein's empirical parameterization [9], V is parameterized by three real numbers, and one complex phase:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.6}$$

λ ($\equiv \sin\theta_c \approx 0.22$), A and $\sqrt{\rho^2+\eta^2}$ are real (A, ρ and λ being of order of unity), while the phase in question is $\arg(\rho, \eta)$. This parameterization allows for CP violation if $\eta \neq 0$.

So far, CP violation has been observed only in the decays of the K^0 meson. Although readily accommodated in the Standard Model, CP violation remains one of the least understood phenomena in physics, since in the case of K^0 the effects that arise from strong interactions make it nearly impossible to tell whether the complex CKM phase is the sole source of CP violation. Furthermore, the CKM matrix is only a phenomenological parameterization: even if it provides an adequate description of Nature, the Standard Model offers no insight as to why the various elements have the values they do.

1.1.3 Matter-antimatter asymmetry

In addition, CP violation is thought to be responsible for the observed excess of matter in the Universe. In the Big-Bang model, the matter and antimatter have been created in the equal amounts, but today there appears to be only matter. In 1966, Sakharov outlined three conditions necessary for this imbalance to occur [11]: the proton must decay; the Universe had to pass through a period of thermal non-equilibrium; and there must be preference for matter over antimatter – that is, CP violation. CP violation is thus one of the most important ingredients of Sakharov's condition.

Recent work [12] suggests that the minimal Standard Model with the CKM matrix from eq. (1.6) cannot provide CP violation sufficient to account for the observed baryon asymmetry, indicating that other sources may be present. Thus, an experimental effort to determine the elements of the CKM matrix and measure the extent of the CP violation is of considerable importance.

1.2 Probing CKM with B-mesons

In principle, the squares of the various elements of the CKM matrix can be deduced from the observations of various weak decays. The comparison of nuclear beta decay and muon decay indicates $|V_{ud}| \approx 0.97$, while the strangeness-changing decays $|V_{us}| \approx 0.22$. These two are just $\cos\theta_c$ and $\sin\theta_c$ in the Cabibbo scheme (eq. (1.2)). The production of charmed particles in neutrino (or antineutrino) nucleus scattering is proportional to $|V_{cd}|^2$. Data from CDHS Collaboration led to a value $|V_{cd}| = 0.21 \pm 0.03$.

The decays of the b-quark involve generation-changing transitions. The decay mechanisms of the B-meson (a bound state of a b and a u or a d quark) are shown in fig. 1-2. The dominant b-decay process is $b \to c$ (fig. 1-2a), the rate being proportional to $|V_{cb}|^2$. However, since $V_{cb} \approx A\lambda^2$ (eq. (1.6)) is relatively small, we can observe processes that are otherwise suppressed like $b \to u$ (fig. 1-2c), $b \to s$ (fig. 1-2d) and $b \to d$ (fig. 1-2g,h), all of them involving various elements of the CKM matrix ($V_{ub} \approx A\lambda^3(\rho - i\eta)$, $V_{ts} \approx -A\lambda^2$, $V_{td} \approx A\lambda^3(1-\rho-i\eta)$ respectively). We can therefore use decays of B mesons to determine the couplings of the third generation quarks (b and t) with the quarks of the other two generations.

The unitarity of the CKM matrix translates into nine linearly independent equations, out of which six have the form of three complex numbers that sum to zero:

$$V_{ub}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \tag{1.7}$$

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.8}$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \tag{1.9}$$

$$V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0 \tag{1.10}$$

$$V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0 \tag{1.11}$$

$$V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0 \tag{1.12}$$

As proposed by Chau, Keung and Bjorken [10], each of these equations can be represented as a triangle in the complex plane. The triangle corresponding to eq. (1.8), called the *Bjorken triangle* is particularly useful from the phenomenological point of view, since it contains the most poorly known entries in the CKM matrix. In addition, all three terms are proportional to λ^3, so the triangle is almost equilateral, thus having relatively large angles. It can be shown that large angles of the Bjorken triangle imply a large CP violation (hopefully large enough to be experimentally observable). This makes the B decays an ideal tool for exploring the physics of the CKM matrix.

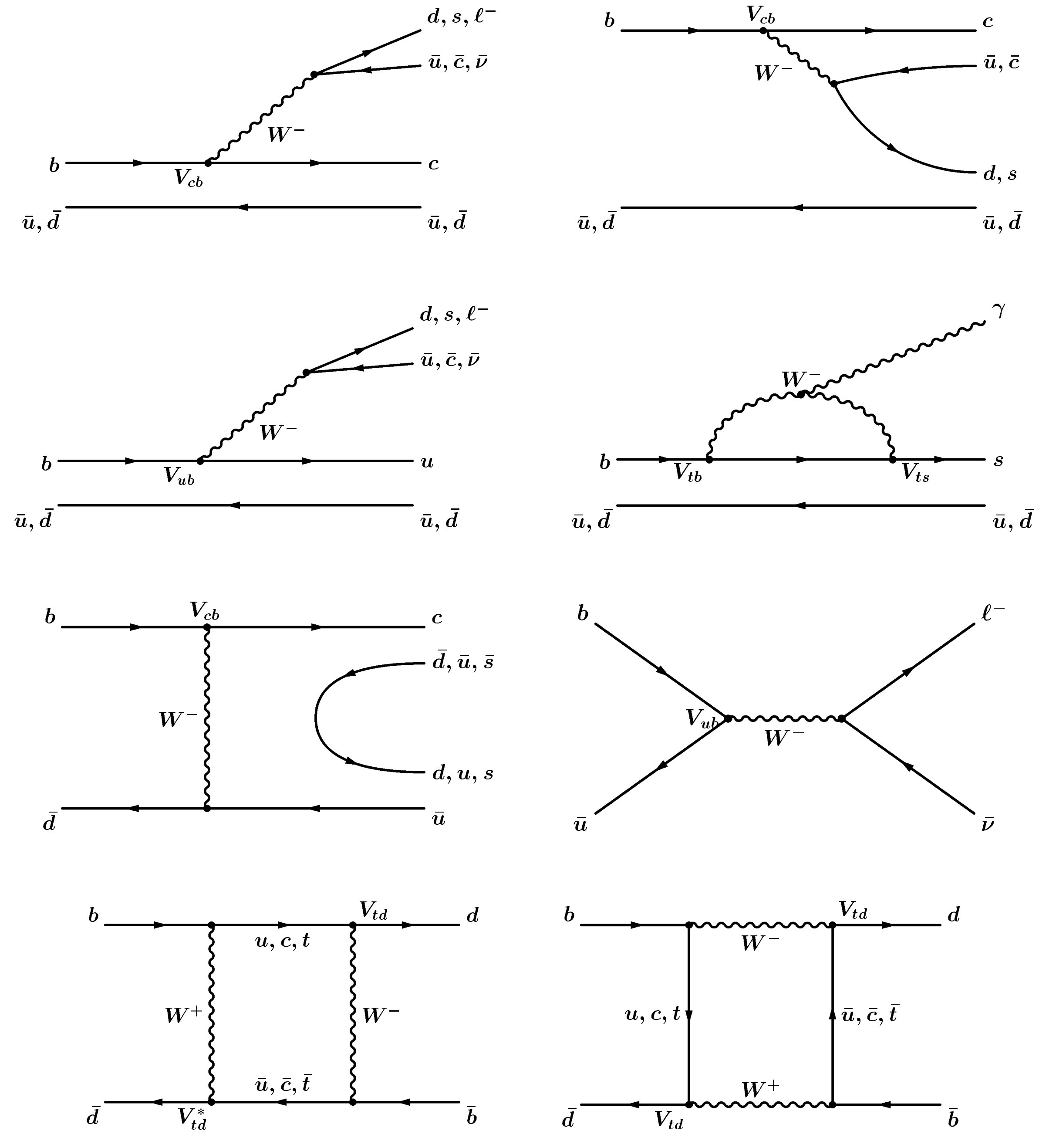
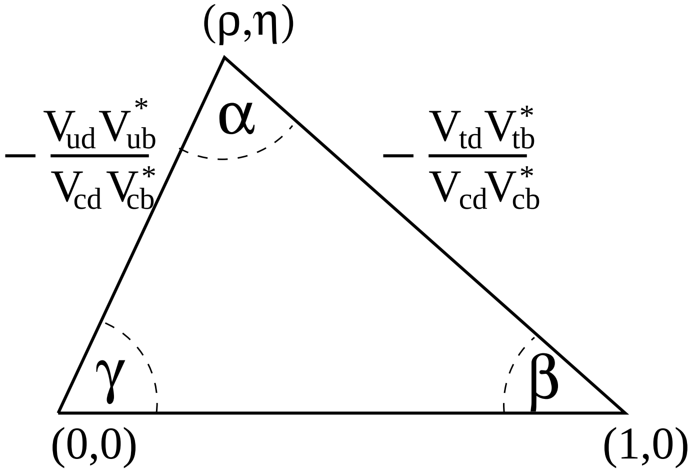


Figure 1-2: B meson decay mechanisms: (a) external and (b) internal spectator diagrams; (c) $b \to u$ spectator diagram; (d) $b \to s\gamma$ electromagnetic penguin; (e) W-exchange and (d) W-annihilation diagrams; and (g), (h) box diagrams for $B^0\bar{B}^0$ mixing.



[The Bjorken triangle.]

Figure 1-3: The Bjorken triangle.

1.3 Constraining the CKM matrix

Checking the unitarity of the CKM matrix is equivalent to verifying that the Bjorken triangle (fig. 1-3), as well as other five triangles derived from the CKM matrix, are *closed*. This is done by independently measuring both the sides and the angles of the Bjorken triangle.

1.3.1 Measuring sides of the Bjorken triangle

The sides of this triangle are proportional to $V_{ub}^*/A\lambda^3$ and $V_{td}/A\lambda^3$ (the third one is by definition of the unit length). The CKM element V_{ub} can be determined either by studying the end-point of the $B \to \nu\ell X$ spectrum, which gives $|V_{ub}| = (3.1{\pm}0.8){\times}10^{-3}$ [4], or by observing $b \to u$ transitions. For instance, CLEO has reconstructed decays like $B \to \nu\ell\rho$ and $B \to \nu\ell\pi$, resulting in $|V_{ub}| = (3.3 \pm 0.8) \times 10^{-3}$ [4]. Nevertheless, although the world averages for both types of measurement are consistent with each other, due to uncertainties on the host of theoretical models employed in extracting $|V_{ub}|$, they are still not combined into one limit. In either case,

$|V_{ub}/V_{cb}| \simeq 0.08 \pm 0.02$.

The same approach is not applicable to V_{td}, which would involve a $t \to d$ transition at the tree-level. Given that the $t \to d$ decay is suppressed ($V_{td} \approx A\lambda^3(1-\rho-i\eta)$), a large sample of top quark decays is needed, and this will be hard to achieve even when the LHC turns on in 2005. Alternatively, one utilizes higher-order processes involving a virtual t-quark interacting with real d-quark. For instance, V_{td} can be inferred from the processes involving radiative penguin diagrams like $b \to s\gamma$ (fig.1-2d) and $b \to d\gamma$ (in which the outgoing s quark is replaced by a d quark). CLEO has a clear signal for[4] $B \to K^*\gamma$, involving V_{ts}, however there is yet no evidence for[5] $B \to \rho\gamma$ and $B \to \omega\gamma$ which depend on V_{td}. The limit on the ratio of branching ratios, $\mathcal{B}(B \to \rho/\omega\gamma)/\mathcal{B}(B \to K^*\gamma) < 0.19$ (at 90% CL) [13] can be converted into a limit on $|V_{td}/V_{ts}|$. Unfortunately, due to small branching ratios for these processes, the number of observed events is still insufficient for a good determination of $|V_{td}/V_{ts}|$.

Instead, the $B_d^0\bar{B}_d^0$ mixing Feynman diagrams (fig. 1-2g,h) are also sensitive to V_{td}, and the high statistics B^0 samples (like the $B \to \nu\ell D^{(*)}$ sample described in chapter 4) are used. Consequently, the measurement of the $B_d^0\bar{B}_d^0$ oscillation frequency Δm_d proved to be the best way of indirectly measuring V_{td}. $B_d^0\bar{B}_d^0$ oscillations are the topic of this thesis, and are discussed in detail in chapter 2.

1.3.2 Measuring angles of the Bjorken triangle

In addition to measuring the sides of the Bjorken triangle, its angles can also be determined independently. The angles α, β and γ (fig. 1-3) are related to the expected asymmetries in the decay rates of the neutral B_d^0 meson and its antiparticle $\bar{B}_d^0$ into a CP eigenstate f (*e.g.* $J/\psi K_S^0$). Since neutral B mesons mix via the Feynman diagrams shown in fig. 1-2g,h, there are two alternative decay paths for $B^0 \to f$ transition:

- $B^0 \to f$

---

[4] B meson contains a b quark and K^* meson an s quark.

[5] Both ρ and ω contain only light quarks u and d.

- $B^0 \to \bar{B}^0 \to f$

In general, there is a relative phase between the decay amplitudes[6] for these two processes. The CP conjugate situation (starting out with $\bar{B}^0$) has the opposite phase. It can be shown that the time-dependent asymmetry, defined as

$$\mathcal{A}_{CP}(t) \equiv \frac{\Gamma(B^0(t) \to f) - \Gamma(\bar{B}^0(t) \to f)}{\Gamma(B^0(t) \to f) + \Gamma(\bar{B}^0(t) \to f)} \tag{1.13}$$

where $\Gamma(X)$ is the the decay rate for the process X, can be expressed as

$$\mathcal{A}_{CP}(t) = \pm \sin(2\phi_M + 2\phi_D) \sin(\Delta m_d t) \tag{1.14}$$

Here Δm is the oscillation frequency for $B^0 \leftrightarrow \bar{B}^0$ mixing (Δm_d for B_d^0, Δm_s for B_s^0), $2\phi_M$ denotes the weak phase of the $B^0 \to \bar{B}^0$ amplitude (fig. 1-2g,h), and $2\phi_D$ is the phase difference between the decay amplitudes for $B^0 \to f$ and $\bar{B}^0 \to f$. Therefore, if one denotes the amplitude of a given process X by $a(X)$, the sum of the relative phases between $B^0 \to f$ and $B^0 \to \bar{B}^0 \to f$ is

$$2(\phi_M + \phi_D) = \arg \left(\frac{a(B^0 \to f)}{a(B^0 \to \bar{B}^0 \to f)} \right) \tag{1.15}$$

From the Bjorken triangle (fig. 1-3) one can deduce (see Appendix B) that

$$2(\phi_M + \phi_D) = -2\beta \quad \text{in the case of} \quad B_d^0 \to J/\psi K_S^0 \tag{1.16}$$

$$2(\phi_M + \phi_D) = -2\alpha \quad \text{in the case of} \quad B_d^0 \to \pi^+\pi^- \tag{1.17}$$

$$2(\phi_M + \phi_D) = -2\gamma \quad \text{in the case of} \quad B_s^0 \to \rho K_S^0 \tag{1.18}$$

These examples demonstrate that the three angles of the Bjorken triangle can, in principle, be measured independently of each other. Moreover, from eq. (1.14) we immediately see that the large angles of the Bjorken triangle imply large CP-asymmetries. As the Bjorken triangle is almost equilateral, none of the angles is small, and the conditions for experimental observation of CP violation in the B system are quite favorable. The feasibility of the measurement of $\sin 2\beta$ (eq. (1.16), which is the easiest of the three) is discussed in Chapter 9.

---

[6] By decay amplitude $a(B^0 \to f)$ we mean the matrix element $\langle f|\mathcal{H}|B^0\rangle$, where $\mathcal{H}$ is the effective Hamiltonian for weak decays.

In all three cases (eqs. (1.16), (1.17) and (1.18)), one compares the rates of , $(B^0_{d,s} \to f)$ and , $(\bar{B}^0_{d,s} \to f)$ using equation (1.13). Experimentally, one counts the decays of the neutral B meson into f, and classifies them based on whether the neutral B meson was a B^0 or a $\bar{B}^0$ at the moment when it was produced in the collision. Thus, one needs to know the *flavor* of the B^0 meson at the time of production – that is, whether the B-meson contains a b or a $\bar{b}$ quark. The process of determination of the b-quark flavor is called *flavor tagging.* Several flavor tagging methods commonly used will be discussed in section 2.2.

The measurement of $B^0\bar{B}^0$ mixing also requires the knowledge of the flavor of the neutral B meson at its production point. This thesis describes a novel flavor tagging algorithm more suitable for the $p\bar{p}$ environment, and its application to the observation of $B^0_d\bar{B}^0_d$ mixing and the measurement of the mixing frequency. In the following chapter we turn to the theoretical and experimental basis of the $B^0_d\bar{B}^0_d$ mixing.

Chapter 2

$B^0\bar{B}^0$ mixing in $p\bar{p}$ collisions

2.1 Mixing in the Standard Model

The mesons $B_d^0(\bar{b}d)$ and $\bar{B}_d^0(b\bar{d})$ are eigenstates of the strong interaction Hamiltonian. They have the same quantum numbers of the parity (P), charge conjugation (C), and angular momentum (J) operators, as well as the same mass (i.e. the energy levels of the bound state of a heavy and a light quark).

However, they differ in the quark flavors (b vs. $\bar{b}$-quark). The weak interactions do not conserve quark flavor, and thus can mix B^0 and $\bar{B}^0$ via second order Feynman diagrams – called *box diagrams* – shown in figure 2-1.[1] This phenomenon of spontaneous turning of a particle into its own antiparticle (and back) is known as *oscillation* or *mixing*. It has been known in the kaon system since 1950's [14, 15]. The first evidence of the mixing in the neutral B^0 system – a combination of $B_d^0 \leftrightarrow \bar{B}_d^0$ and $B_s^0 \leftrightarrow \bar{B}_s^0$ processes – was presented by UA1 [16] in 1987. $B_d^0\bar{B}_d^0$ oscillations (without the B_s^0 contaminations) were observed at ARGUS, using a sample of $B_d^0\bar{B}_d^0$ events[2] [17]. Both measurements were *time-integrated*, meaning that the number of mixed events has been integrated over the proper decay time of the B meson decay. The first *time-dependent* measurement, in which B meson oscillations are studied as a function of the proper time of the B decay, was performed by ALEPH [18]. The

---

[1] Similar diagrams involving the charged Higgs bosons $H^\pm$ would have to be included, if $H^\pm$ exist.

[2] ARGUS operated at $\Upsilon(4S)$ resonance, which subsequently decays into $B_d^0\bar{B}_d^0$

oscillations in the $B_s^0\bar{B}_s^0$ system are very rapid, and have not been observed yet [62].

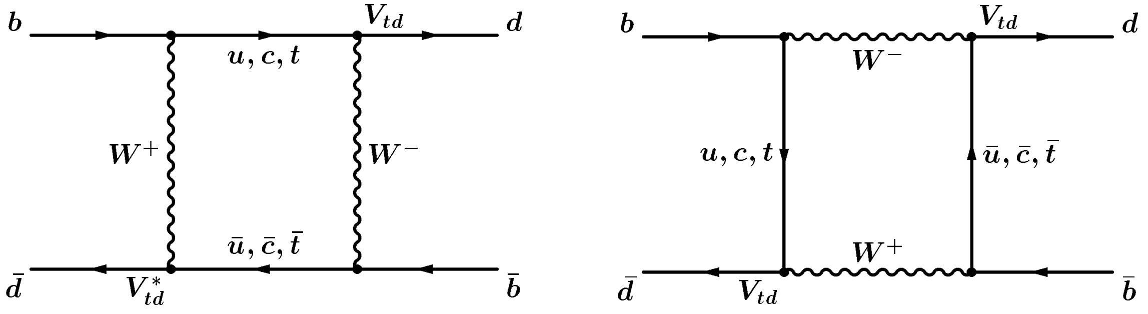


Figure 2-1: The 'box diagrams' for $B^0\bar{B}^0$ mixing.

A sample of initially pure B^0 mesons can be represented by a two-component wave function, $\begin{pmatrix} a(t) \\ \bar{a}(t) \end{pmatrix}$, meaning $a(t)|B^0\rangle + \bar{a}(t)|\bar{B}^0\rangle$, where $a(t)$ and $\bar{a}(t)$ are the amplitudes for finding B^0 and $\bar{B}^0$ respectively, at proper time t. $\bar{B}^0$ is the charge-conjugate of B^0, so that $CP|B^0\rangle = -|\bar{B}^0\rangle$. The Hamiltonian of this system can be phenomenologically expressed as a mass matrix, and the Shrödinger equation has the form

$$i\frac{d}{dt}\begin{pmatrix} a(t) \\ \bar{a}(t) \end{pmatrix} = \begin{pmatrix} m - \frac{1}{2}i, & m_{12} - \frac{1}{2}i,_{12} \\ m_{12}^* - \frac{1}{2}i,_{12}^* & m - \frac{1}{2}i, \end{pmatrix} \begin{pmatrix} a(t) \\ \bar{a}(t) \end{pmatrix} \tag{2.1}$$

The diagonal elements describe the decay of the neutral B mesons with m being the mass of the flavor eigenstates B^0 and $\bar{B}^0$, and , their decay width. The off-diagonal terms describe the mixing from the box diagrams (fig. 2-1). m_{12} and $,_{12}$ can be determined from theory by evaluating the box diagrams. However, in contrast to the $K^0\bar{K}^0$ system,[3] these common decays are Cabbibo suppressed here (meaning that they involve a small CKM matrix element), and represent a tiny fraction of the total B decay rate. The term $,_{12}$ in eq. (2.1) can therefore be neglected.

Diagonalizing the Hamiltonian from eq. (2.1), one obtains the CP eigenstates B_1 and B_2 which are linear combinations of the flavor eigenstates:

$$|B_1\rangle \;=\; \frac{1}{\sqrt{2}}(|B^0\rangle + |\bar{B}^0\rangle) \tag{2.2}$$

---

[3] A K^0 meson, a bound state of a s and a $\bar{d}$ quark, can oscillate into its antiparticle $\bar{K}^0$ for exactly the same reasons as B^0. Furthermore, the particle-antiparticle mixing was discovered in the $K^0 \leftrightarrow \bar{K}^0$ system.

$$|B_2\rangle = \frac{1}{\sqrt{2}}(|B^0\rangle - |\bar{B}^0\rangle) \tag{2.3}$$

with masses $m_{1,2}$ and widths $\Gamma_{1,2}$:

$$m_{1,2} = m \pm \frac{\Delta m}{2} \tag{2.4}$$

$$\Gamma_{1,2} = \Gamma \pm \frac{\Delta\Gamma}{2} \tag{2.5}$$

where

$$\Delta m = 2Re\sqrt{\left(m_{12} - \frac{\Gamma_{12}}{2}\right)\left(m_{12}^* - \frac{\Gamma_{12}^*}{2}\right)} \tag{2.6}$$

$$\Delta\Gamma = 2Im\sqrt{\left(m_{12} - \frac{\Gamma_{12}}{2}\right)\left(m_{12}^* - \frac{\Gamma_{12}^*}{2}\right)} \tag{2.7}$$

The mass difference Δm between B_1 and B_2 is the oscillation frequency of changing from a B^0 to $\bar{B}^0$ and vice versa. This can be verified by explicitly writing the probabilities to find a B^0 or a $\bar{B}^0$ at a time t, $P_{B^0\to B^0}(t)$ and $P_{B^0\to\bar{B}^0}(t)$ respectively. If a particle started out as a B^0 at $t = 0$:

$$P_{B^0\to B^0}(t) = \frac{1}{2}\left[e^{-\Gamma_1 t} + e^{-\Gamma_2 t} + 2e^{-\Gamma t}\cos(\Delta m t)\right] \tag{2.8}$$

$$P_{B^0\to\bar{B}^0}(t) = \frac{1}{2}\left[e^{-\Gamma_1 t} + e^{-\Gamma_2 t} - 2e^{-\Gamma t}\cos(\Delta m t)\right] \tag{2.9}$$

Therefore the measurement of the time dependence of the $B^0 \to \bar{B}^0$ transition allows one to extract the value of Δm_d (we will use the subscript to distinguish it from Δm_s, the mass difference in the $B_s^0\bar{B}_s^0$ system).

The oscillation frequency, Δm_d, is related to the elements of the CKM matrix by explicit evaluation of the box diagrams (fig. 2-1):

$$\Delta m_d = \frac{G_F^2}{6\pi^2} m_B m_t^2 F\left(\frac{m_t^2}{m_W^2}\right)\eta_{QCD} B_{B_d} f_{B_d}^2 |V_{tb}^* V_{td}|^2 \tag{2.10}$$

Here G_F is the weak coupling constant; m_B is the mass of the B_d^0 meson and m_W the mass of the W; m_t is the 'running' top quark mass ($m_t = 167 \pm 6$ GeV) [20]; $F(x)$ is a known function, given to a good approximation by $0.784\ x^{-0.24}$; η_{QCD} is a factor that accounts for the QCD corrections (recently calculated to next-to-leading order and found to be 0.55 ± 0.01); B_{B_d} is the non-perturbative "bag" factor, and f_{B_d} is the decay constant of the B meson, with $f_{B_d}\sqrt{B_{B_d}} = 200 \pm 40$ MeV. All values are taken from Ref. [19].

The experimental goal is to measure Δm_d. Once that is done, knowledge of other quantities in equation (2.10) enables the extraction of V_{td} (as $V_{tb} = 1$ in Wolfenstein's parameterization (1.6)). Since $V_{td} = A\lambda^3(1-\rho-i\eta)$, constraining V_{td} is equivalent to constraining the apex of the Bjorken triangle (the point with the coordinates (ρ,η)) in fig. 1-3, to lie on a circle centered at $(1,0)$:

$$|V_{td}|^2/(A^2\lambda^6) = (1-\rho)^2 + \eta^2$$

Assuming $|V_{tb}| = 1$, the world average[4] value of $\Delta m_d = 0.460 \pm 0.018$ ps^{-1} yields

$$V_{td} = (8.6 \pm 0.2 \pm 0.2 \pm 1.7) \times 10^{-3}$$

where the errors are due to Δm_d, m_t and $f_{B_d}\sqrt{B_{B_d}}$ [19]. Currently, the overall error on V_{td} is dominated by the uncertainty on the product $f_{B_d}\sqrt{B_{B_d}}$.

2.2 Experimental approaches to mixing

To observe mixing in a sample of B decays, we need to determine the fraction of events that have mixed as a function of the proper time t. Therefore, there are three pieces of information necessary for any measurement of the $B^0\bar{B}^0$ oscillations:

1. The proper time of the B-meson decay.

2. The flavor of the B meson at $t = 0$.

3. The flavor of the B meson when it decayed.

---

[4] The world average quoted includes the measurement presented in this thesis.

To date, all measurements of the $B^0\bar{B}^0$ mixing based on time oscillations have been made in experiments at colliders (LEP[5], SLC[6] and Tevatron[7]), in which a pair of b-quarks is produced. First we mention the strategies used at LEP, and then contrast them with the possibilities at CDF.

At LEP running at the Z^0 resonance (91 GeV/c^2), each of the b quarks from the Z decay has the energy of about 45 GeV, and they emerge from the interaction point in the opposite directions. For this reason, if a B meson is found on one side, it is very likely that at least some decay products of a b-flavored hadron are found on the opposite side of the detector. Usually, one first finds a neutral B meson (for simplicity, we call it B^0), and reconstructs its decay point (called *vertex*); the hemisphere of the event containing the B^0 vertex is hence called the *vertexing* side. The proper time is extracted from the reconstructed decay length, *i.e.* the distance between the interaction point where the B meson was produced, and the B^0-meson decay point – the B vertex. The flavor of the B^0 meson at the time of its decay is deduced from the decay products. The flavor of the B^0 meson at $t = 0$ is then inferred from the decay products of the b-flavored hadron containing the *other* b quark, which is usually found in the other hemisphere of the event. This other hemisphere is thus called the *flavor tag* side. The class of tagging methods utilizing the opposite side to determine the flavor of the b-quark at production is therefore called the Opposite Side Tagging (OST). Opposite Side Tagging is schematically shown in fig. 2-2.

---

[5] Large Electron-Positron collider, in CERN, Geneva, Switzerland. e^+ and e^- are collided at the center-of-mass energy of 91 GeV/c^2, corresponding to the Z^0 resonance. Z^0 then decays into a $b\bar{b}$ pair.

[6] Stanford Linear Collider, in SLAC, Palo Alto, CA. e^+ and e^- are collided at the Z^0 resonance, however the beams are polarized, resulting in kinematic correlations between the b and $\bar{b}$ quarks that are different from those at LEP and Tevatron. For this reason we do not compare tagging methods used at SLC with those used at LEP and Tevatron.

[7] At Tevatron, p and $\bar{p}$ beams collide at the center-of-mass energy of $\sqrt{s} = 1.8$ TeV/c^2. The $b\bar{b}$ production mechanisms are discussed in section 2.3.2.

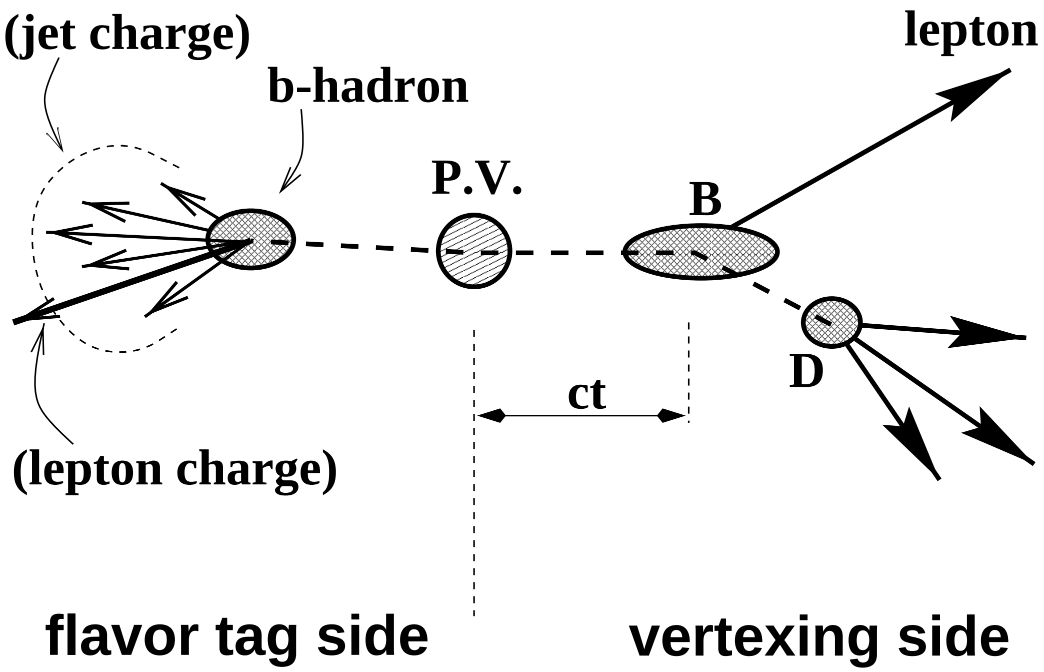


Figure 2-2: A schematic representation of the tagging possibilities. Two b-quarks are produced at the primary vertex (P.V.), resulting in a B^0 meson and a b-hadron recoiling in the opposite directions. The 'vertexing side' is the hemisphere where the B^0 vertex is reconstructed (in this example from ℓ and D meson). On the opposite side (the 'flavor tag side'), decay products of the b-flavored hadron are used to infer the flavor of the B^0 meson at $t = 0$. Usually, two methods are used to determine the flavor of the b-hadron on the flavor tag side: the charge of the lepton from the semileptonic decay of the b-hadron ("lepton tagging"), and the weighted sum of the charges of likely b-hadron decay products ("jet-charge tagging").

A typical LEP mixing measurement includes $B^0 \to \ell + \text{jet}^8$ or a $B^0 \to D^{*-}X$ on the vertexing side, and a flavor tag on the opposite side. The Opposite Side Tagging algorithms that are commonly used [43] include:

- the charge of a lepton, most likely coming from the semileptonic decay of the b-flavored hadron on the opposite side;
- the charge of an identified kaon; the $b \to c \to s$ transition chain results in a strange hadron – usually a kaon (a meson containing the s quark) – that determines the flavor of the b quark;
- the charge of a D^{*+} from $\bar{B} \to D^{*+}X$;
- the sum of the momentum-weighted charges of the opposite jet, also known as the *jet charge*.

The $B^0\bar{B}^0$ oscillation is diluted first by the presence of the charged B mesons on the vertexing side, and also by the mistags on the flavor tags side. In designing a B^0-mixing measurement, the goal is to improve the B^0 purity on the vertexing side and reduce the mistag rate on the flavor tag side, while maximizing statistical precision of the measurement.

In contrast to LEP, the $b\bar{b}$ production cross-section at hadronic colliders is orders of magnitude higher, and therefore the experiments at Tevatron have a tremendous potential advantage in statistics. (For instance, the lepton + charm sample used in this thesis is the largest of its kind in the world.) The precision of the determination of the B decay position at the CDF detector is also comparable to that of the LEP experiments.

However, hadrons have a complicated internal structure (quarks and gluons), and thus hadronic collisions are more complex than leptonic annihilation. As a result, the Opposite Side Tagging methods outlined above usually do not work as well as in the LEP environment. This thesis uses a novel tagging algorithm based on the information available on the vertexing side, called the Same Side Tagging (SST).

---

[8] A jet is stream of particles, coming from a hadronization of a quark into a hadron and subsequent decay of the hadron.

2.3 B^0-meson production in $p\bar{p}$ collisions

We now briefly describe the benefits and the complications of the environment of the proton-antiproton collisions.

2.3.1 Parton model and parton distribution functions

We imagine the proton as composed of virtual states of free partons – three constituent quarks (the 'valence quarks'), virtual gluons, and virtual quark-antiquark pairs (the 'sea quarks') created by splitting of virtual gluons.

At the Tevatron, protons and antiprotons collide at the center-of-mass energy (denoted by $\sqrt{s}$) of 1.8 TeV. At these energies, the collision time (the time spent in interchanging virtual bosons) is much shorter than the lifetime of these virtual states, so we may treat the partons as free during the collision.

Partons do not divide the proton momentum among themselves equally. We define the *parton distribution functions* $f_i^a(x)$ as the probability that a parton i (a quark or a gluon) carries a fraction x of the momentum of the beam particle a (a proton or an antiproton). In the collision, each interacting parton i carries only the fraction x of proton's 900 GeV, for quarks, about $\langle x \rangle = 10\%$ on average. If the effects arising from the Quantum Chromo-Dynamics (QCD) are taken into account, then parton distribution functions also depend on the momentum exchanged in the interaction.

2.3.2 $p\bar{p} \rightarrow b\bar{b}$ processes

In leading order (LO) approximation in QCD, b-quarks are produced via the Feynman diagrams shown in figures 2-3a, 2-3b, 2-3c. In these events, b and $\bar{b}$ quarks are created, and they move away from each other with equal but opposite momenta transverse to the $p\bar{p}$ beam axis.

The leading-order picture is satisfactory when the quark mass, m_q, is similar or greater than the average momentum carried by each parton, $\langle x \rangle \sqrt{s}/2 \approx 90$ GeV.[9]

---

[9] Each parton carries the fraction $\langle x \rangle$ of proton's 900 GeV. Here, $\sqrt{s} = 1.8$ TeV/c^2 is the total collision energy in the center of mass.

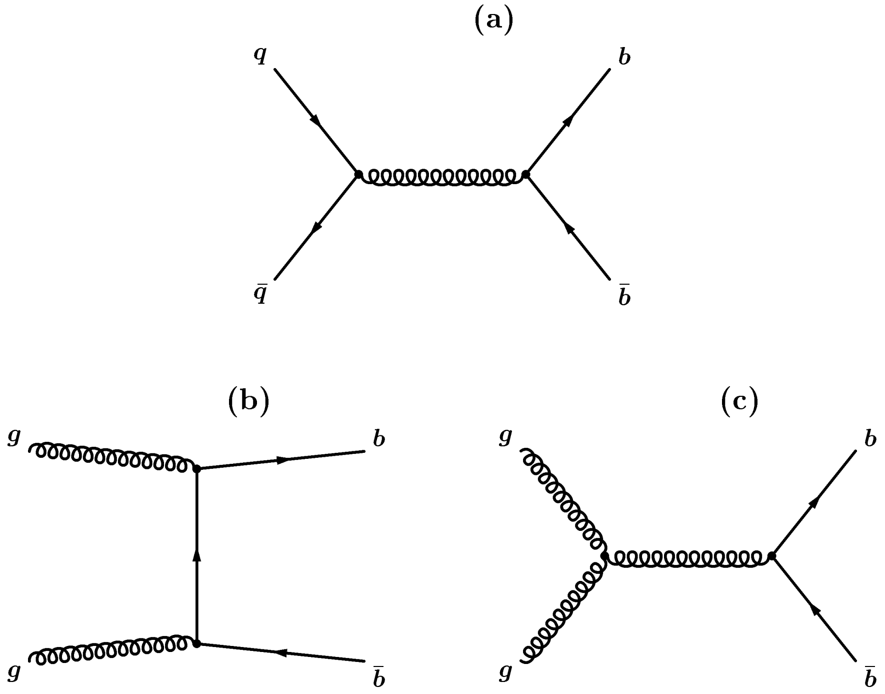


Figure 2-3: Leading order processes contributing to the $b\bar{b}$ production.

While this is true for the top quark ($m_t = 175$ GeV), it certainly is not for the b quark ($m_b \approx 4.75$ GeV), where $2m_b/\sqrt{s} \approx 0.005 \ll 0.1$. In this case, the higher order diagrams (figures 2-4a, 2-4b, 2-4c and 2-4d) may have equally significant contribution as the leading order processes.[10]

Dominant next-to-leading order (NLO) processes involve a two-gluon initial state: gluon splitting (2-4d), and flavor excitation (essentially an initial state gluon splitting) (2-4b,c). From the theoretical standpoint, it is important to understand the contribution of the NLO processes: their contribution is large, which suggests that even higher-order graphs must be included.

Moreover, understanding the NLO contribution is important from a purely practical point of view: the Opposite Side Tagging techniques rely on the b quark produced on the opposite side. However, the kinematic correlations between b and $\bar{b}$ quarks (section 3.2.1) are quite different in the leading order approximation (no gluons, b

---

[10] This is due to the large gluon density at small x, the increased color factor at the 3-gluon vertex, and the cross-section enhancement for diagrams containing a vector exchange in the t-channel.

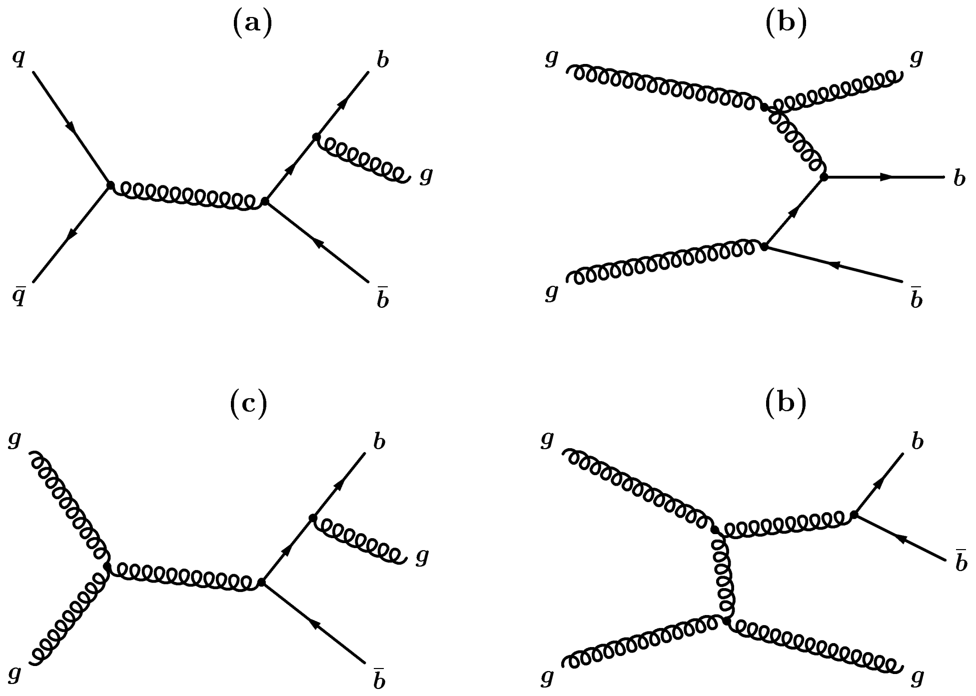


Figure 2-4: Next-to-leading order processes contributing to the $b\bar{b}$ production.

and $\bar{b}$ are back-to-back) compared to the NLO (gluon may take significant amount of the transverse energy), so the ability to use these tagging methods depends upon the understanding of the higher-order diagrams.

The fact that the b and $\bar{b}$ quarks do not emerge in opposite directions has profound implications on the possibility of the Opposite Side Tagging at CDF. Because the momentum of the $b\bar{b}$ pair along the beam direction is not zero, Monte Carlo studies show that if one b-quark is in the central region of the detector, the other b-quark is also central with the probability of only about 40%; this limits the maximum Opposite Side Tagging efficiency. Moreover, because of the high NLO contribution, it is very hard to distinguish the opposite side b-quark jet from gluon jets, so the techniques such as jet-charge tagging do not work nearly as well as at LEP [21]. Finally, the b-quark momentum spectrum is softer (lower average momentum) than LEP's 45 GeV (at the Z^0 mass of 91 GeV/c^2), and consequently the semileptonic decays of the opposite-side b-quark result in leptons that are barely distinguishable from the multitude of the surrounding hadronic tracks. Therefore, even the Opposite Side Tagging techniques based on the leptonic tagging are not as effective as their counterparts at LEP [22, 23, 24].[11]

For all these reasons, it is highly advantageous to use the information available on the vertexing side – the same side where the neutral B meson is reconstructed. For this reason, the tagging algorithms utilizing the information gathered on the vertexing side are called *Same Side Tagging* (SST). The information that can be used to infer the flavor of the b-quark is available in two kinds: the pions from the decays of the higher B resonances, $B^{**+} \rightarrow B^0\pi^+$, as well as the pions that were created in the process of turning the b-quark into a B-meson, called *fragmentation*.

The quarks are never observed free – they can only be found in the structures of mesons (a quark and an antiquark) and baryons (three quarks). The process in which a quark that was just created (in a $p\bar{p}$ interaction or a decay of a particle like Z^0) associates itself with other quarks to form a hadron is called fragmentation or hadronization. During fragmentation, usually several hadrons are created, appearing in a detector as a *jet* of particles. Gluons emitted in a hadronic interaction also fragment (three-jet events were taken as the first experimental evidence for the existence of gluons). More information on fragmentation can be found in Appendix C.

The charge correlations among the fragmentation products, however, are crucial for the Same Side Tagging. When a $\bar{b}$-quark fragments into a B^0 meson, a $d\bar{d}$ pair is created from the $\bar{b}$-quark's kinetic energy. A d quark combines with $\bar{b}$ to form $B^0(\bar{b}d)$, and the remaining $\bar{d}$-quark continues the fragmentation process. If the $\bar{d}$-quark pulls a $u\bar{u}$ pair from vacuum, the $\pi^+(u\bar{d})$ meson is created. Therefore, the first *charged* particle produced after the B^0 meson in the fragmentation chain uniquely determines the flavor of the B meson: π^+ is produced along with the B^0 meson, while π^- is produced together with the $\bar{B}^0$ meson.

The correlation between the charge of first fragmentation pion and the neutral B meson is the same as between the pion and the B meson that are products of the

---

[11] A lepton-based Opposite Side Tagging analysis gains a lot in the case of the dedicated trigger requiring two leptons on the opposite sides.

$B^{**\pm}$ decay. In this thesis we make use of both of these charge-flavor correlations to determine the flavor of a neutral B meson at the production point. The Same Side Tagging is described in detail in chapter 6.

Figure 2-5 shows a schematic representation of a tagging methods used at CDF.

2.3.3 A $b\bar{b}$ event

Now we can picture a typical $b\bar{b}$ event. There are usually two b-jets (jets created in the fragmentation of two b-quarks), and a possible additional gluon jet in the case of a higher order process, as well as many other particles produced in the fragmentation of the remnants of the proton and the antiproton, that are not correlated in any way with the b-quark jets. This remainder of the $p\bar{p}$ collision is called the *underlying event*. In addition, other protons and antiprotons can collide in the same bunch crossing, contributing an independent $p\bar{p}$ event (usually just a soft QCD process) at a different position along the beamline. An additional $p\bar{p}$ interaction of this kind is called a *pile-up event*, and there may be more than one in each bunch crossing. A schematic representation of a process resulting in a b, a $\bar{b}$ and a gluon (g) is shown in fig. 2-6.

Depending on whether the $\bar{b}$-quark initially creates a $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$ or a diquark-antidiquark pair, the B^+, B^0, B_s, B_c mesons and b-flavored baryons (e.g. Λ_b) are produced, respectively. This is characterized with the respective probabilities

$$f_u : f_d : f_s : f_c : f_\Lambda$$

where $f_u = f_d \approx 38\%$, $f_s \approx 11\%$ [53]. Since B_c is produced extremely rarely, roughly 13% of b quarks fragment to baryons. This thesis is concerned only with the behavior of B^+ and B^0 mesons (and their charge-conjugates).

2.4 Basic ingredients of a mixing measurement

In this thesis we apply the Same Side Tagging technique to the sample of $B \to \nu\ell D^{(*)}$ decays. The Same Side Tagging is described in detail in chapter 6. Here it suffices to say that it is based on the correlations between the B meson flavor and the charge of

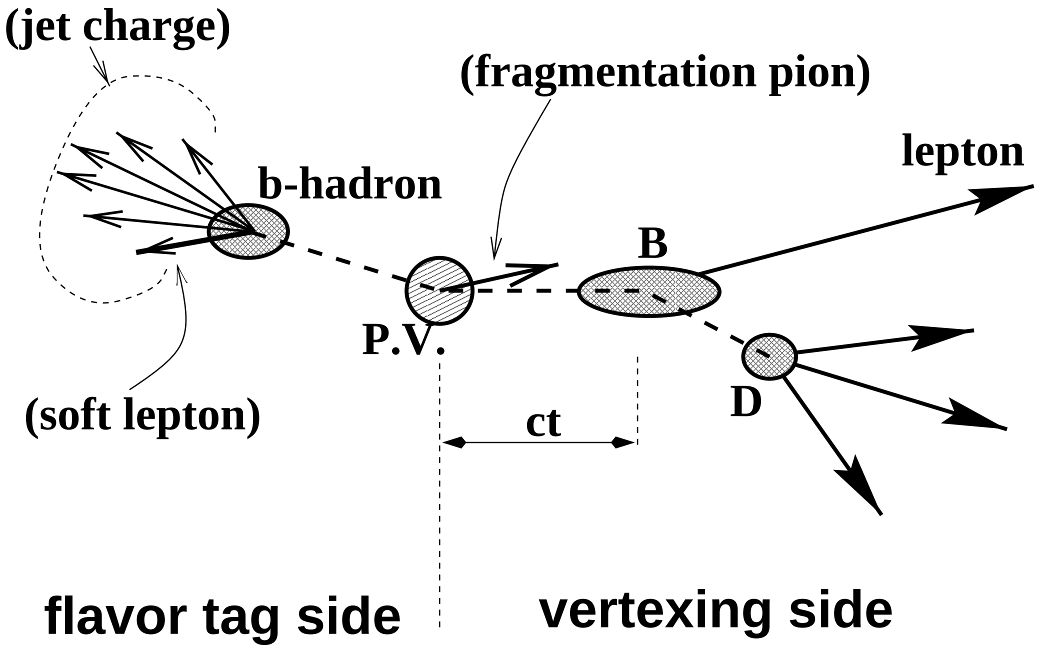


Figure 2-5: A schematic representation of the tagging possibilities at CDF. The $b\bar{b}$ pair is created at the primary vertex (P.V.). However, as opposed to LEP, the b and $\bar{b}$ quarks do not always emerge back-to-back from the primary vertex (cf. fig. 2-2). As a consequence, the usual Opposite Side Tagging methods like lepton tagging (at CDF called 'soft-lepton' tagging, since the lepton used as the tag usually has a low momentum) and jet-charge tagging do not work as well as at LEP. However, the Same Side Tagging methods (using the information on the 'vertexing side') are less affected by the peculiarities of the $b\bar{b}$ production in $p\bar{p}$ environment. In the Same Side Tagging (SST) method used in this thesis, the pions originating from the B^{**} decay and the fragmentation of the b-quark (into the B^0 meson) are used to deduce the flavor of the B^0 meson at $t = 0$.

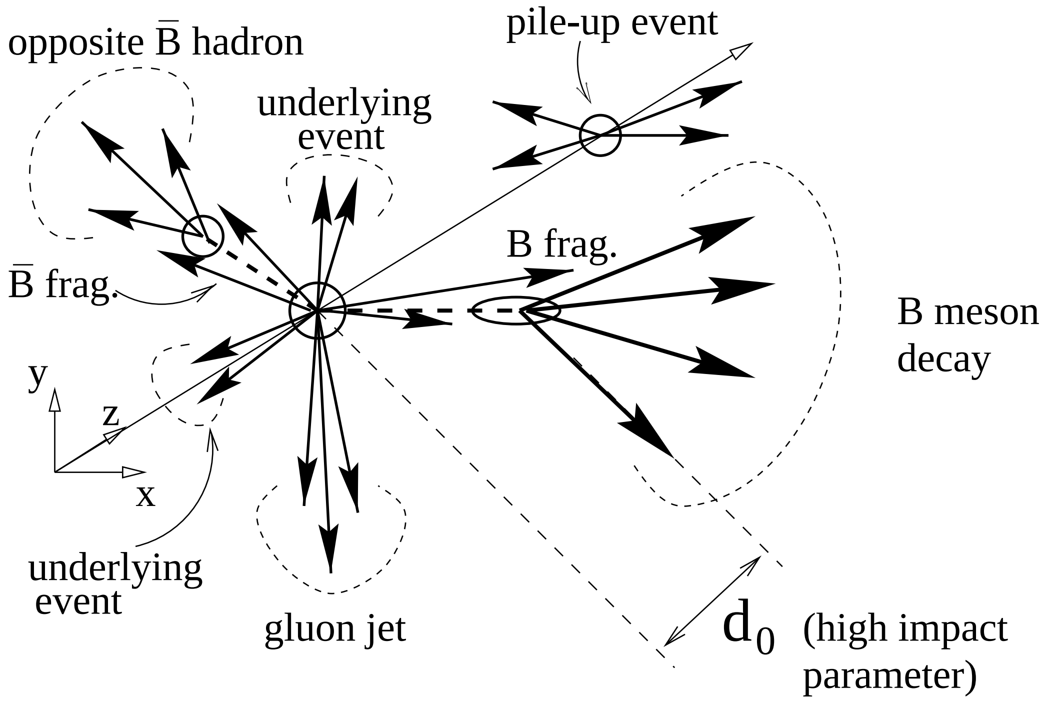


Figure 2-6: A schematic representation of a typical event with a $b\bar{b}$ pair and a gluon (g) in the hadronic environment. A B-meson decay of interest is on the right-hand side. A $\bar{b}$-flavored hadron emerges on the opposite side. A jet that originated from the fragmentation of the soft gluon recoils against the B meson and the $\bar{b}$-flavored hadron. racks that were created in the process of fragmentation of the proton remnants (an underlying event), as well as tracks that come from a secondary $p\bar{p}$ interaction (a pile-up event) are also shown.

the particles produced along with it. We select one such particle and use its charge to tag the flavor of the B meson at its *production* time. Finding a π^+ means that a B^0 was produced, while finding a π^- implies a $\bar{B}^0$. The charge of the lepton tags the flavor of the B meson at *decay*: $B^0 \to \ell^+$ and $\bar{B}^0 \to \ell^-$. Therefore, finding a π^+ along with the B meson decaying into a ℓ^+ means that the B^0 meson did not mix. We define the $\ell^\pm\pi^\pm$ as the right correlation, or *right sign* (RS). On the other hand, finding a π^+ together with a ℓ^- means that the B^0 meson did mix. We define $\ell^\pm\pi^\mp$ as *wrong sign* (WS).

The probabilities that the B^0 will not mix and mix, respectively, are:

$$P(t)_{B^0\to B^0} = \frac{1}{2\tau}e^{-t/\tau}(1+\cos\Delta m_d t) \tag{2.11}$$

$$P(t)_{B^0\to \bar{B}^0} = \frac{1}{2\tau}e^{-t/\tau}(1-\cos\Delta m_d t) \tag{2.12}$$

Wrong sign lepton-pion pairs, however, can also come from *mistagging*, that is, selecting an incorrect pion as the tag. If we define $d \equiv$ the fraction of correctly tagged B's, we can write the numbers of 'right sign' and 'wrong sign' events:

$$N_{RS}(t) = dN(t)_{B^0\to B^0}(t) + (1-d)N(t)_{B^0\to \bar{B}^0}(t) \tag{2.13}$$

$$N_{WS}(t) = (1-d)N(t)_{B^0\to B^0}(t) + dN(t)_{B^0\to \bar{B}^0}(t) \tag{2.14}$$

We then define the asymmetry $\mathcal{A}(t)$ as

$$\mathcal{A}(t) \equiv \frac{N_{RS}(t) - N_{WS}(t)}{N_{RS}(t) + N_{WS}(t)} \tag{2.15}$$

Substituting eq.(2.13) and (2.14) into (2.15), we get the explicit dependence on the proper time:

$$\mathcal{A}(t) = (2d-1)\cos\Delta m_d t = \mathcal{D}_0 \cos\Delta m_d t \tag{2.16}$$

The amplitude of the oscillation is ≤ 1 due to imperfect flavor tagging. The quantity $\mathcal{D}_0$ is called *dilution* and is the measure of the purity of the flavor tag decision.

To perform a mixing measurement, we obtain $\mathcal{A}(t)$ and fit it with a cosine (convoluted with the resolution function for t). The amplitude is $\mathcal{D}_0$, and the frequency is proportional to Δm_d. To get $\mathcal{A}(t)$ we need to determine:

1. The flavor of the B^0 at the time of its decay. For this we will use the charge of the ℓ in $B \to \ell D^{(*)}$ decays. The lepton + charm dataset is described in section 4.1.

2. The proper decay time for the B^0 meson. The necessary corrections to the measured proper time of the B^0 decay are described in section 4.3.

3. The flavor of the B^0 meson at the time of production. For this we will use the Same Side Tagging technique, described in section 6.1.

We start by reviewing the relevant features of the Tevatron and the CDF detector in chapter 3, and then describe the sample of $B \to \ell D^{(*)} X$ decays in chapters 4 and 5. The details of the Same Side Tagging algorithm are given in chapter 6. In chapter 7 we describe the fit for Δm_d and the dilutions. The systematic uncertainties are discussed in chapter 8. We conclude in chapter 9.

Chapter 3

The experimental apparatus

3.1 Tevatron – the source of $p\bar{p}$ collisions

The proton beam begins as gaseous hydrogen, which is ionized to form H^-. H^- ions are extracted from the ion source by the Cockroft-Walton electrostatic accelerator, and accelerated to 750 keV. They continue down the short, straight section, the Linac, where they are accelerated up to 401.5 MeV, first by the Drift Tube Linac, then by the Side-coupled Cavity Linac.

Before entering the third stage, the Booster, the H^- ions pass through a carbon foil which removes the two electrons, leaving only the protons.

The Booster is a rapid cycling synchrotron 500 feet in diameter. The protons travel around the Booster about $20,000$ times and their energy is raised to 8 GeV. The Booster normally cycles twelve times in rapid succession, loading twelve pulses, or bunches of protons, into the Main Ring.

The Main Ring is another proton synchrotron located in the Fermilab tunnel, which is four miles in circumference.[1] Under the current operating mode, the Main Ring accelerates protons to 150 GeV.

In fixed target mode, the proton beam is lowered into the Tevatron, accelerated to 900 GeV, extracted and sent down the Fixed Target beamline to the experimental

---

[1] A tunnel ten feet in diameter, buried 20 feet underground, also houses the Tevatron, whose magnets are positioned below Main Ring's magnets.

areas.

In collider mode, however, the proton beam in the Main Ring is used to produce antiprotons. The protons are accelerated to 120 GeV, and extracted to hit a Tungsten target. Antiprotons are among the secondary particles produced in this collision. They are selected and transported to the Debuncher ring where they are reduced to the same low energy by a process known as stochastic cooling [27], and then stored into the Accumulator ring. Finally, when a sufficient number has been produced, the antiprotons are reinjected into the Main Ring, accelerated to 150 GeV, and moved down into the Tevatron, in six bunches. At the time of the antiproton injection, six bunches of protons are already orbiting in the opposite direction. The number of particles in a proton bunch is $N_p \sim 2 \times 10^{11}$, while in an antiproton bunch it is $N_{\bar{p}} \sim 6 \times 10^{10}$.

After all protons and antiprotons are in the Tevatron, they are accelerated to 900 GeV in the 5.7 T magnetic field of superconducting dipole magnets.[2] The antiproton content of the Tevatron is called a 'store'. Each store lasts for several hours, occasionally even days. Events during each store are grouped into 'runs'. Most parameters of the CDF operation (*e.g.*, the position of the beam) are stored in a database in run-averaged format.

Proton and antiproton bunches cross each other at six straight sections of the Tevatron, with a frequency of 3.5μs. At two of them, B0 and D0, two general purpose detectors, CDF and DØ are situated. The quadrupole magnets squeeze the p and $\bar{p}$ beams into a cross-sectional area of $\sigma_A \sim 5 \times 10^{-5}\text{cm}^2$, so that the p and $\bar{p}$ collide in the geometrical center of each detector.

The instantaneous luminosity is defined as

$$\mathcal{L} = \frac{N_p N_{\bar{p}} N_B f}{4\pi\sigma_A^2} \tag{3.1}$$

where $N_B = 6$ is the number of bunches, and f is the bunch revolution frequency (about 50 kHz). Instantaneous luminosity is highest at the beginning of the store, and declines as the antiprotons are lost to collisions and beam instability.

---

[2] The Tevatron also contains quadrupole and octupole magnets for focusing and beam stabilization.

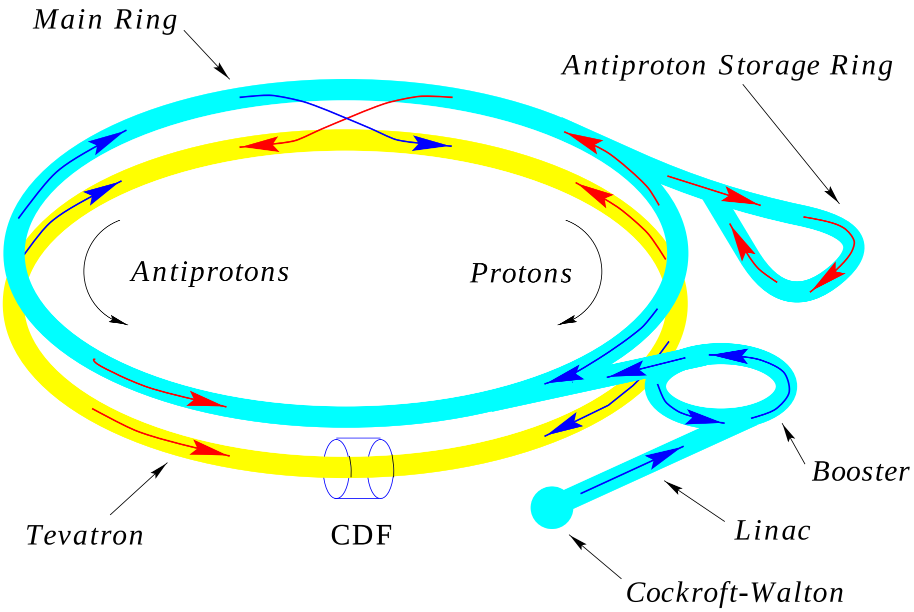


Figure 3-1: A schematic view of the Tevatron at Fermilab. The two experiments CDF and DØ are also shown.

Data used in this thesis have been collected between 1992 and 1995 ('Run 1'). This period has been divided in two parts, different mainly in the Tevatron conditions. The first part, Aug 1992 to May 1993, called Run 1A, is characterized by average instantaneous luminosities of 0.54×10^{31} cm^{-2}s^{-1} and a beam cross-section of about 40μm. In the second part, Jan 1994 to July 1995, called Run 1B, the Tevatron delivered significantly higher luminosities (maximal to 2.8×10^{31} cm^{-2}s^{-1}, average 1.6×10^{31} cm^{-2}s^{-1}) and also a narrower beam ($\sim 25\mu$m). Also, for Run 1B, CDF's Silicon Vertex Detector, was replaced by a radiation-hard version, the "SVX"' [29]). During Run 1, total of 109.4 ± 7.2 pb^{-1} of data have been collected.

3.2 The CDF detector

3.2.1 Kinematic variables in $p\bar{p}$ collisions

In the case of a proton-antiproton collider with *unpolarized* beams, the only preferred axis is the beam direction. Physics processes are hence invariant under rotations around the beam direction, and the polar coordinate system, with the beam direction as the z axis is a natural choice.

When a proton and an antiproton collide, the total kinematic energy available to the collision is $\sqrt{s} = 1.8$ TeV. However, each parton (a constituent or sea quark, or a gluon) carries only a fraction of the proton momentum, determined by the parton distribution function. As a consequence, the physics interactions often have large boosts along the beam direction, and the observed momenta projected to the z axis do not sum to zero. However, the total momentum in the plane transverse to the beam, (p_x, p_y), is zero. In this plane, we choose to represent each particle by its polar coordinates, $p_T = \sqrt{p_x^2 + p_y^2}$ and ϕ (the azimuthal angle, measured from x axis).

The polar angle θ is measured in the $x - z$ plane, starting from the positive z axis. However, a more useful quantity called rapidity is defined as

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

Under a boost in the z direction, β, rapidity transforms as $y' = y + \tanh^{-1}\beta$, so the distribution of rapidity, dN/dy is invariant under Lorentz transformations along the z axis (and thus the same in both the lab frame and in the true center-of-mass frame). For ultrarelativistic particles, $p \gg m$, and rapidity can be approximated by *pseudo*-rapidity

$$\eta \equiv \frac{1}{2} \ln \left(\frac{p + p_z}{p - p_z} \right) = -\ln \tan \frac{\theta}{2}$$

3.2.2 Overview of the CDF detector

Since energetic hadron collisions can be described by the Lorentz-invariant variables, p_T, ϕ and η, the Collider Detector at Fermilab (CDF) [28] was designed to have an approximately cylindrically symmetric layout of detector components with segmentation roughly uniform in η and ϕ.

For the purpose of a precise measurement of the p_T, axial magnetic field of 1.412 T permeates the tracking detectors. In this configuration, the particle tracks are helices, which are represented by five parameters: curvature (inversely proportional to p_T), polar angle θ, and the position of the point of closest approach – its azimuthal angle θ_0, its location in z, z_0, and its *impact parameter* d_0 – the distance of the closest approach. The weakly-decaying B mesons live long enough[3] for their decay position to sometimes be well-separated from the interaction point. Their decay products thus often have impact parameters larger than 150μm (see figure 2-6), and thus can be separated from the particles originating from the interaction point. Thus, a precise determination of d_0 is crucial in selection of B meson decays.

The beam is circular in the $x-y$ plane (with the radius of 40μm in Run 1A, and 25μm in 1B), and gaussian in z (width 30 cm), the center of the beam being very close to the geometrical center of the detector. The exact position of the beam changes from store to store, the beam displacements varying from 400 to 1000μm in vertical (y) and from 200 to 1200μm in the horizontal (x) direction. Moreover, the beam and the z-axis of the CDF detector are not exactly parallel, and have relative slopes of about 3μm/cm and 5μm/cm in x and y directions respectively. Run averages for both beam slopes and beam displacements are kept in the CDF database, and are used in this thesis to calculate the position of the beam.

In a particular event, the $p\bar{p}$ interaction point is called the *primary vertex*. Frequently, especially during the high luminosity conditions of Run 1B, more than one proton and antiproton may interact in the same bunch crossing, and there may be more than one primary vertex in the event. The number of primary interactions per event, averaged over Run 1B, is about 1.5.

The goal of the CDF detector is to provide a large coverage of the solid angle, in order to capture as many tracks as possible. A side view cross-section of the CDF detector is shown in figure 3-2. A particle leaving the interaction point first traverses three tracking chambers (SVX, VTX and CTC), positioned inside a superconducting solenoid. All particles except muons and neutrinos are then stopped inside one of the calorimeter systems, which provides a measurement of their energy. The tracks that pass through the calorimeters and leave tracks in the muon chambers are then identified as the muon candidates. Integrated tracking is available only in the central rapidity region ($|\eta| \leq 1.0$)

---

[3] A B meson that decays after one lifetime will, on average, have traversed the distance of about 2 mm.

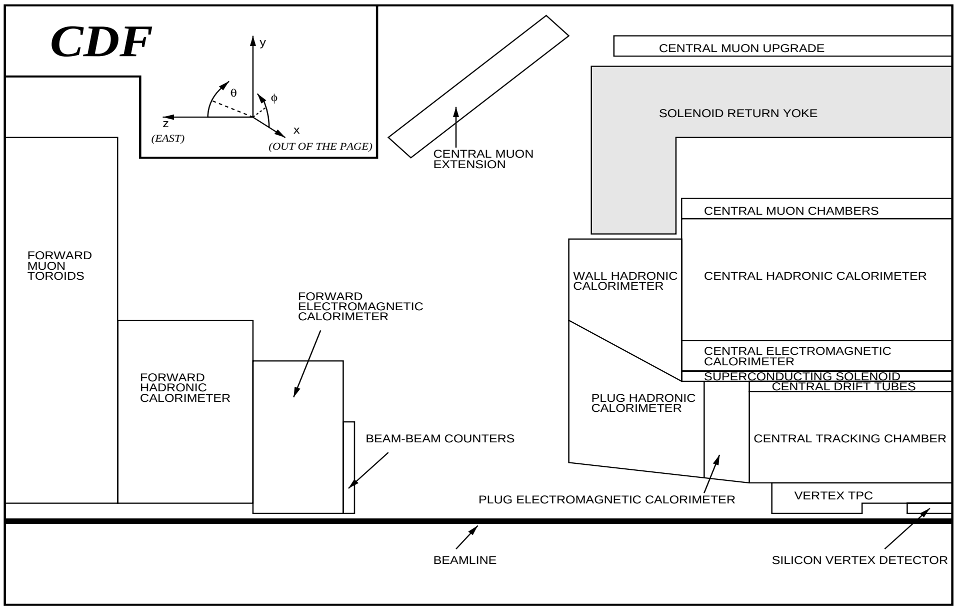


Figure 3-2: A side view cross-section of the CDF detector. The interaction region is in the lower right corner. The detector is forward-backward symmetric about the interaction region. The CDF co-ordinate system is described in the upper left corner.

Neutrinos escape undetected resulting in an apparent non-conservation of energy and momentum in the transverse plane. Their presence in the event can therefore be inferred from the amount of energy missing in the transverse plane.

3.2.3 Tracking systems

There are three tracking systems at CDF: the SVX used for measuring the displacement of tracks from the primary vertex, the VTX used to determine the z position of the primary vertex, and the CTC used to measure tracks' momenta. Combined, they provide a good measurement of all track parameters, especially in the $x - y$ plane.

Measuring d_0 and ϕ with the SVX

The Silicon Vertex Detector (SVX) [29] is intended to provide points very close to the beam spot, thus increasing the resolution on the impact parameter d_0 and ϕ of charged particles. This high precision is needed when trying to discern whether a collection of tracks intersect at a point other than the primary vertex.[4]. The original SVX, installed for Run 1A, because of radiation damage had to be replaced for Run 1B with the SVX'. SVX and SVX' are very similar in construction. Their comparison is show in table 3.1.

The SVX (and the SVX') consists of two barrels aligned along the detector z-axis (the beam direction). An SVX barrel is shown in figure 3-3. At $z = 0$, there is a gap of 2.15 cm between barrels. Total length (along z) of the SVX is 51 cm; since the $p\bar{p}$ interactions are spread around $z = 0$ with $\sigma \sim 30$ cm, the geometric acceptance of the SVX is roughly 60%. η coverage is between ± 1.9.

Each barrel consists of four concentric layers of silicon strip detectors, which are composed of twelve wedges (thus 30^o each). Every layer consists of twelve ladders, each 25.5 cm in length. A ladder is shown in figure 3-4.

Each ladder in turn is divided into three single sided silicon wafers, each 8.5 cm long. Wafers have silicon strips on one side only, and they are aligned along the z axis – thus the SVX is a $r - \phi$ tracking device (2-D tracking in the transverse plane). The width of the strips is 60μm for the three inner layers, 55μm for the fourth one. The individual hit position is obtained by fitting the charge distribution of neighboring strips, and has a resolution of about 10μm. σ_{d_0}, the resolution of the impact parameter, d_0, as a function of the track transverse momentum, p_T, for the SVX' detector is shown in figure 3-6.

SVX strips are read out in so-called 'sparse mode', in which only strips that are significantly over the threshold are read out. The data is read by the SVX chip, which contains 128 channels. There are 2, 3, 4 and 6 chips per ladder on layers 1 to 4 respectively. In total, the entire SVX has 46080 channels. Because of this multitude of channels, the readout time of the SVX detector is about 2 ms (a large value when compared with the readout time of the other CDF detector systems). The SVX chip also allows hardware subtraction of the leakage current, on a strip-by-strip basis.

---

[4] Such a point is called the *secondary vertex*

Feature	SVX	SVX′
Channels	46080	46080
z coverage	51.1 cm	51.1 cm
Gap at z=0	2.15 cm	2.15 cm
Radius of layer 0	3.0049 cm	2.8612 cm
Radius of layer 1	4.2560 cm	4.2560 cm
Radius of layer 2	5.6872 cm	5.6872 cm
Radius of layer 3	7.8658 cm	7.8658 cm
Overlap of layer 0	-1.26deg	0.17deg
Overlap of layer 1	0.32deg	0.32deg
Overlap of layer 2	0.30deg	0.30deg
Overlap of layer 3	0.04deg	0.04deg
Silicon	one-sided	one-sided
Power	DC	AC, FOXFET bias
Passivation	none	polyimide
Atmosphere	Argon/Ethane+H_2O	Dry Nitrogen
Readout chip	SVX IC Rev. D	SVX IC Rev.H3
Sampling	quadruple	double
Noise	2200 electrons	1300 electrons
Gain	15 mV/fc	21 mV/fc
Reset/Integrate	3.5 μs	3.5 μs
Readout time	2.7 μs	2.1μs
Radiation Limit	15-20 KRad	$>$ 1 MRad
Bad channels	2.93%	1.73%
Typical Occupancy	7-10%	5%
Max Occupancy	12-20%	25%

Table 3.1: A comparison of the SVX and SVX′ detectors.

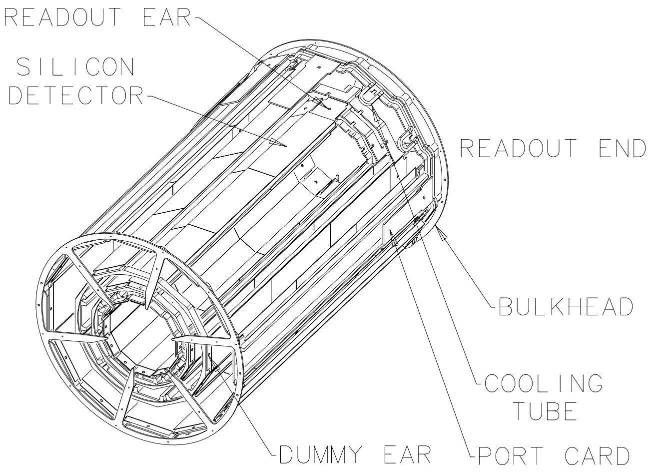


Figure 3-3: An isometric view of a single SVX barrel. Some of the ladders of the outer layer have been removed to allow a view of the inner layers.

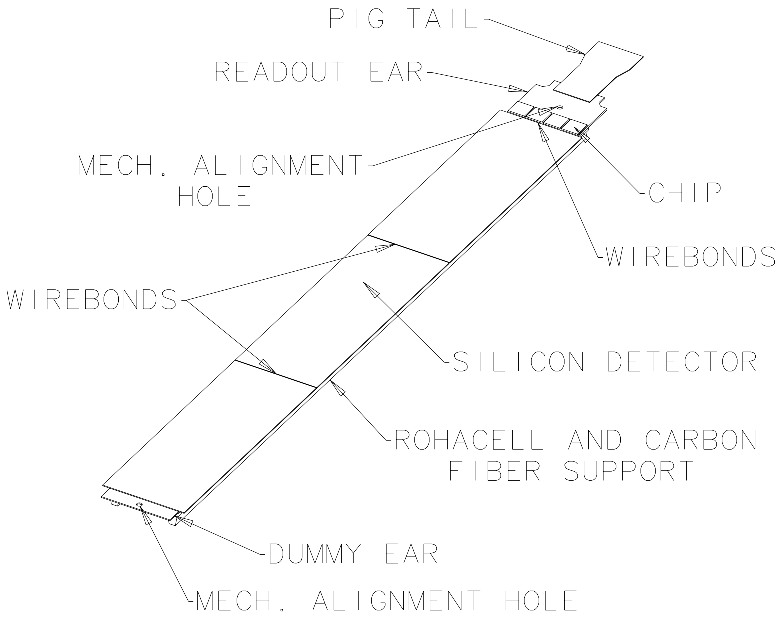


Figure 3-4: An SVX ladder used in the construction of SVX layers.

Measuring z_0 and θ with the VTX

The VTX is a time-proportional drift chamber segmented into 8 modules along the z axis. Each module is octagonal, composed of 8 wedges. A central high voltage grid divides each module into two 15.25 cm long drift regions. This distance was chosen so that the maximum drift time (in 50/50% argon-ethane mixture at atmospheric pressure and $E = 320$ V/cm) is less than the time between two bunch crossings. The electrons drift away from the center grid until they pass through a cathode grid and enter one of the two proportional chamber endcaps. Each endcap is divided into octants, with 24 sense wires and 24 cathode pads in each octant. The arrival times of the electrons at the sense wires give a picture of the event in the $r - z$ plane. Adjacent modules have a relative rotation angle of $\phi_0 = \tan^{-1}(0.2)$ about the z axis. For tracks passing through at least two modules, this eliminates inefficiencies near octant boundaries and provides ϕ information from small angle stereo. The active area of the chamber extents from about $r = 7$ cm to $r = 21$ cm.

Primary vertices in the event are obtained by fitting the tracks in $r - z$ plane. The primary vertex z resolution is $1 - 2$ mm, depending on the number of tracks participating in each fit.

Measuring p_T, θ and ϕ with the CTC

The Central Tracking Chamber (CTC) is a cylindrical drift chamber 3.2 m long, and extends from the VTX, to the superconducting solenoid (at a radius of 130 cm).

84 layers of sense wires are grouped into 9 superlayers, numbered 0 to 8. According to their purpose, the superlayers are divided into axial (track reconstruction in the $r - \phi$ plane) and stereo (reconstruction in the $r - z$ plane). A transverse view of the CTC is shown in figure 3-5. Superlayers 0, 2, 4, 6, and 8 are axial, while 1, 3, 5 and 7 are stereo. The axial superlayers consists of 12 sense-wire layers each, along the z axis, while the stereo superlayers contain 6 sense-wire layers each, rotated by $\pm 3^o$ with respect to the axial superlayers.

Sense-wire layers are tilted by 45^o with respect to the radial direction, in order to compensate for the Lorentz drifting angle.[5] In the CTC, the electrons drift perpendicularly to the radius vector. This design also ensures that high p_T tracks pass close to at least one sense wire. Moreover, large tilt angle resolves the left-right ambiguity arising from the fact that it is impossible to tell which side of the sense wire the electrons drifted from. Wrong assignment results in a 'ghost track', which must also be considered in the pattern recognition. In the CTC, however, this ghost track is rotated by an angle of 70^o with respect to the real track, so the pattern recognition problem is greatly simplified.

The superlayers are further divided into cells, so that the maximum drift time is around 800 ns, which is much shorter than the 3.5μs between two bunch crossings.

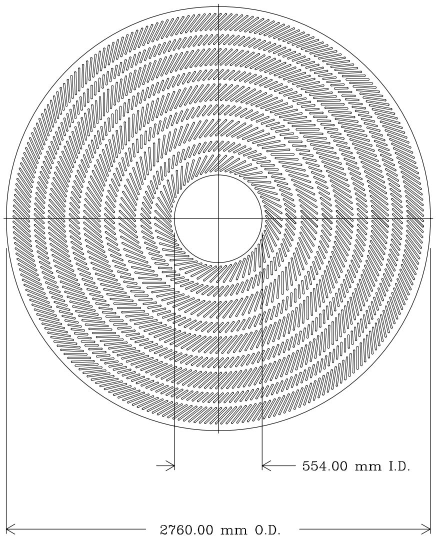


Figure 3-5: A transverse view of the CTC endplate. The nine superlayers are shown.

Integrating the tracking information

The event pattern recognition starts with a set of primary vertices identified by the VTX. The CTC pattern recognition algorithm then loops over superlayers (starting

---

[5] In the presence of crossed electric ($E = 1350$ V/cm) and magnetic ($B = 1.4$ T) fields, electrons drifting in a gas move at an angle – called ***the Lorentz angle*** – with respect to the electric field direction.

from the outer-most), trying to associate hits with an arc of a helix that originated in one of the primary vertices. Tracks with hits in both stereo and axial superlayers are called 3-D tracks.

Once tracks are found in the CTC, they are extrapolated into the SVX. Only hits in a road of a given width around the CTC track are considered. As in the CTC case, hits are added starting from the outer-most layer (layer 4). Apart from providing the 'seed vertices' for the CTC tracks, the VTX is otherwise ignored, even though, in principle, it could improve the resolution on z_0 (currently ~ 4 mm, provided by the CTC).

The combined SVX-CTC tracking gives the momentum resolution of

$$\frac{\delta p_T}{p_T} = \sqrt{0.0066^2 + (0.0009 p_T)^2}\mathrm{GeV}^{-1}$$

and the impact parameter resolution shown on figure 3-6. A typical p_T of the charged particles considered in this analysis is usually below 3 GeV/c. Only the lepton in $B \to \nu \ell D^{(*)}$ decay has a $p_T \sim 10$ GeV/c.

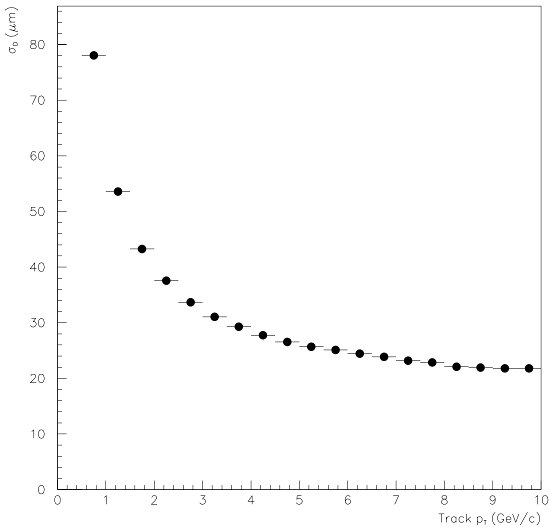


Figure 3-6: Impact parameter resolution of the SVX (σ_D in μm) in the transverse plane as a function of track p_T.

3.2.4 Calorimeters

The CDF calorimeter has a 'projective tower' geometry, i.e. the calorimeter is segmented in η and ϕ 'towers', that point at the interaction region. The coverage is full (2π) in azimuth and $|\eta| < 4.2$ in pseudorapidity. The calorimeter system is divided into three regions, the central, the plug and the forward. The segmentation of the CDF calorimeter is shown in fig 3-7.

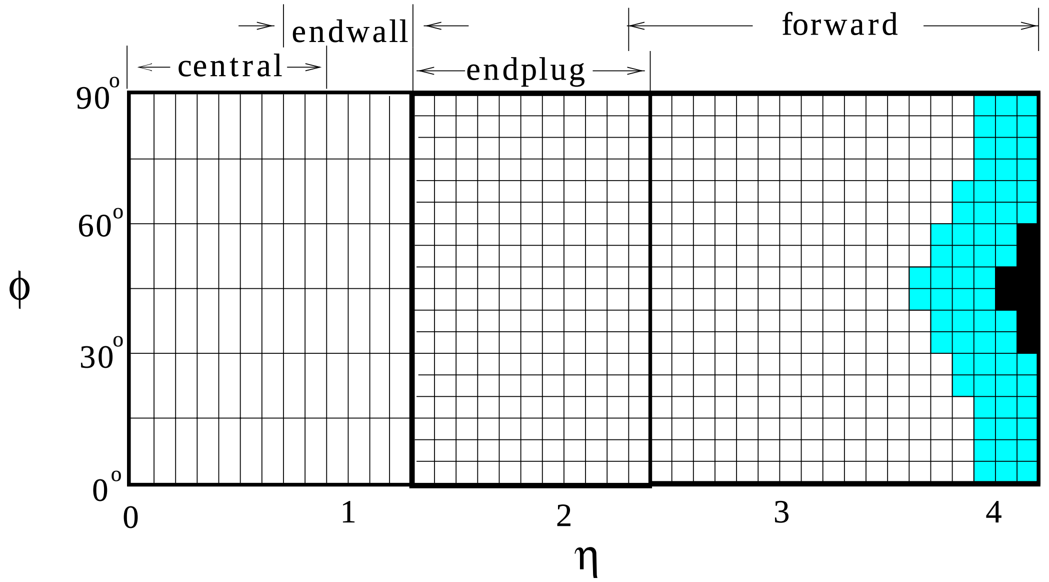


Figure 3-7: A schematic representation of the segmentation of the CDF calorimeter systems.

Each tower has an electromagnetic shower counter in front of a corresponding hadron calorimeter. This allows a detailed comparison of the electromagnetic and hadronic energies deposited in each tower, and thus facilitates the separation of electrons and photons from other hadrons.

There are three types of the electromagnetic shower counters: CEM, PEM, and FEM, corresponding to the central, the plug and the forward regions. The CEM $(0 < |\eta| < 1.1)$ uses lead sheets interspersed with scintillator as the active detector medium. The PEM $(1.1 < |\eta| < 2.4)$ and FEM $(2.2 < |\eta| < 4.2)$ use proportional chambers with cathode pad readout instead. For all three, the spatial resolution is about 2 mm. After 6 radiation lengths in CEM, 3 in PEM and 12 in FEM, there are further proportional chambers giving a detailed measurement of the shower profile.

The hadron calorimeters are CHA and WHA (wall) in the central region, PHA in the plug, and FHA in the forward. They consist of steel plates alternated with active detectors: plastic scintillator in the central ($0 < |\eta| < 1.3$), and gas proportional chambers in the plug ($1.3 < |\eta| < 2.4$) and forward ($2.3 < |\eta| < 4.2$). The pseudorapidity coverage is slightly different from the electrogmagnetic shower counters because of the geometry of the solenoid.

System	η Coverage	Energy Resolution	Thickness
CHA	$\|\eta\| < 0.9$	$50\%/\sqrt{E_T} \oplus 3\%$	$4.5\ \lambda_0$
WHA	$0.7 < \|\eta\| < 1.3$	$75\%/\sqrt{E_T} \oplus 4\%$	$4.5\ \lambda_0$
PHA	$1.3 < \|\eta\| < 2.4$	$90\%/\sqrt{E_T} \oplus 4\%$	$5.7\ \lambda_0$
FHA	$2.4 < \|\eta\| < 4.2$	$130\%/\sqrt{E_T} \oplus 4\%$	$7.7\ \lambda_0$
CEM	$\|\eta\| < 1.1$	$13.7\%/\sqrt{E_T} \oplus 2\%$	$18\ X_0$
PEM	$1.1 < \|\eta\| < 2.4$	$28\%/\sqrt{E_T} \oplus 2\%$	18-21 X_0
FEM	$2.2 < \|\eta\| < 4.2$	$25\%/\sqrt{E_T} \oplus 2\%$	$25\ X_0$

Table 3.2: A summary of the properties of the different CDF calorimeter systems. Energy resolutions for the hadronic calorimeters are for incident pions, and for the electromagnetic calorimeters are for incident electrons and photons. (The symbol $\oplus$ means that the constant term is added in quadrature to the resolution.) λ_0 signifies interaction lengths and X_0 radiation lengths.

3.2.5 Muon systems

Most of the charged particles that penetrate the hadron calorimeters are muons. In the central region, the muons are detected by the central muon system (CMU), located just outside CHA (which acts as a hadron absorber), and the central muon upgrade (CMP), located outside the return magnet yoke (adding 0.6 m of steel). Both CMU and CMP consist of four layers of drift chambers, and cover $|\eta| < 0.6$. To extend the pseudorapidity coverage to $|\eta| < 1.0$, four free-standing conical arches have been added outside the main body of the CDF detector. This detector, called central muon extension (CMX), consists of drift chambers sandwiched between scintillator counters. Central muon systems record only 'muon stubs', i.e. track segments in drift chambers left by muon candidates. In the pattern recognition, these stubs are then linked with the CTC tracks: the hits in the muon chambers are required to match the extrapolated CTC track in both location and slope at the entry into the chamber. The system is almost 100% efficient for central muons with $p_T > 3GeV$.

The forward muon system (FMU) detects muons in the high-η region ($|\eta| > 1.0$). Since there is no tracking in the forward, the FMU has its own toroids and acts as a spectrometer. This thesis makes use only of the muons that leave stubs in both CMU and CMP, since requiring information in both chambers improves the quality of muon candidates.

3.2.6 Triggering

The proton and antiproton bunches cross every 3.5μs. However, not all $p\bar{p}$ interactions are recorded. First, it is not technically possible, since it would require extremely fast electronics and enormous amount of permanent data storage. Moreover, the physically interesting processes (e.g. production of heavy flavor) occur more rarely, and our goal is to record only the events that show signs of interesting physics. This task is achieved by the trigger system.

The CDF has four levels of triggering, numbered from 0 to 3. The first three are implemented in hardware, and introduce no dead time.[6]

The Level 0 trigger requires a bunch crossing (and thus ensures that the event is a $p\bar{p}$ collision and not a cosmic ray). The Level 1 trigger requires a candidate muon in one of the muon chambers, or two neighboring calorimeter towers above the threshold.

---

[6] Dead time is the time interval during which the detector cannot take any new data.

The Level 2 trigger bases its decision on the calorimeter and muon information, as well as the tracking, and the event passes if there is significant jet activity or if there are electron and muon candidates. Calorimeter jets are found by a hardware cluster finder, which first searches for the so-called 'seed trigger towers' (transverse energy $E_T > 3$ GeV) and then considers adding neighboring towers to the cluster. For each cluster, E_T as well as the average η and ϕ are returned.

To find leptons, muon stubs and electromagnetic showers are linked with the CTC tracks found by the Central Fast Tracker (CFT). The CFT is one of the unique features of the CDF – it uses hits from the axial sense-wire layers of the CTC to find tracks at Level 2. The CFT finds tracks with $p_T > 2$ GeV/c that originate from near the origin in the $x-y$ plane and pass through the entire tracking volume of the CTC. As the CTC sense-wire layers are tilted by 45^o with respect to the radial direction (see fig. 3-5), each high-momentum track passes close to at least one wire in each sense-wire layer. In the first pass through the CTC data, the CFT finds hits that occur very near the sense wires, called "prompt" hits. In the second pass, the hits that have drifted most of the way across the cell are found. These hits are called "delayed" hits. High-momentum tracks typically have one prompt hit and two delayed hits per axial superlayer, resulting in a total of 15 hits for the five axial superlayers. Starting from the prompt hit in the superlayer 8, the CFT searches a predefined lookup table of hit patterns and returns the p_T corresponding to the best match. In the case of the inclusive lepton trigger, only higher p_T tracks (> 7 GeV), are matched to the hits in the muon chambers and showers in the electromagnetic calorimeter. For this reason, the Level 2 lepton trigger applies an implicit p_T requirement on the lepton candidates. An approximate Level 2 lepton trigger efficiency will be shown in the following chapter.

If the Level 2 trigger passes the event, the whole detector is read out. The readout time is about 3 ms and results in, on average, a 10% deadtime (which otherwise depends on the trigger rate and instantaneous luminosity). Different parts of the detector are read out in parallel, and passed to the Level 3 trigger. The Level 3 trigger is a software trigger, running on a series of multi-processor Challenge systems.

The event is assembled on one of the CPU's, and then passed to others that execute a highly optimized version of the CDF production executable. This configuration allows for a more sophisticated analysis of each event, and a finer classification of each event into various data 'streams'. The output of the Level 3 is limited by the amount of the temporary data storage (tape in Run 1A, and staging disk in Run 1B).

Chapter 4

Lepton + Charm Sample

As already mentioned in chapter 2, the production rate for $b\bar{b}$ pairs is significantly higher than the current maximum data recording rate achievable with today's technology. The goal of the trigger system is to make a quick decision on whether the event is worth further consideration. Each set of the trigger criteria defines a *dataset*, whose size and characteristic features are set by the specific trigger requirements.

For a mixing measurement, it is advantageous to use a dataset rich in B^0 mesons, however one must trade the number of events versus the purity of the sample. Moreover, one needs both the flavor of the B^0 meson when it decays, as well as the position of the decay.

At a hadron collider, the number of reconstructed tracks in each event is large due to the fragmentation, underlying event and occasional additional interaction (a *pile-up* event), and the B decay products must be distinguished among them. When the B meson decays *hadronically* (all daughters are hadrons; the majority of B decays proceed this way), there is very little to differentiate the B-daughter tracks from the non B-daughter tracks, originating from the primary vertex. Because of the long lifetime of the B meson, one can potentially select events where the B meson lived long enough to decay at the point well-separated from the primary vertex. (Such a point is called a *secondary vertex.*) In order to trigger on the displaced B-daugher tracks, a dedicated track processor is necessary.[1]

---

[1] Such a system did not exist at CDF in Run I, however, the Secondary Vertex Trigger (SVT) is

So while the abundance of B mesons makes B physics attractive at a $p\bar{p}$ collider, not all decays are equally feasible for studying the B meson properties. However, the semileptonic B decays ($B \to \nu\ell X$, where X stands for anything) have proved to be a most useful tool for this purpose. A lepton is easily recognized, and thus it is frequently used for triggering in a hadronic environment.[2] If the lepton is coming from a semileptonic B decay, its charge tells us the flavor of the B meson at the time of the decay. Moreover, the B decay vertex must be along the lepton track, and the goal is to find the remaining B decay products, and use it to determine the exact position of the B meson decay point – the B *vertex.*

Almost all semileptonic B meson decays are $b \to \nu\ell c$ transitions, so the decay products will include a charmed meson. We therefore search for a D meson (a meson composed of a c and a u or d quark) or a D^* meson (an excited state of the D meson) associated with a trigger lepton.

4.1 B candidate selection

For the sample of B^0 mesons we use four decay signatures[3] – three $B^0 \to \nu\ell^+ D^{*-}$ signatures and one $B^0 \to \nu\ell^+ D^-$:

$$B^0 \to \nu\ell^+ D^{*-},\ D^{*-} \to \overline{D^0}\pi_s^-,\ \overline{D^0} \to K^+\pi^- \tag{4.1}$$

$$B^0 \to \nu\ell^+ D^{*-},\ D^{*-} \to \overline{D^0}\pi_s^-,\ \overline{D^0} \to K^+\pi^-\pi_2^-\pi_3^+ \tag{4.2}$$

$$B^0 \to \nu\ell^+ D^{*-},\ D^{*-} \to \overline{D^0}\pi_s^-,\ \overline{D^0} \to K^+\pi^-\pi^0 \tag{4.3}$$

$$B^0 \to \nu\ell^+ D^-,\ D^- \to K^+\pi^-\pi_2^- \tag{4.4}$$

For the charged B mesons, we use only one decay signature,

$$B^+ \to \nu\ell^+\overline{D^0},\ \overline{D^0} \to K^+\pi^- \tag{4.5}$$

---

being built for the CDF's Run II (the data-taking period scheduled to commence in 1999).

[2] CDF's observation of the Top quark was based on the datasets including some kind of lepton trigger.

[3] Throughout this thesis, a reference to the charge conjugate states is implicit, unless explicitly stated otherwise.

In this section we discuss the general selection requirements. A typical $B \to \ell D$ event topology is shown schematically in fig. 4-1

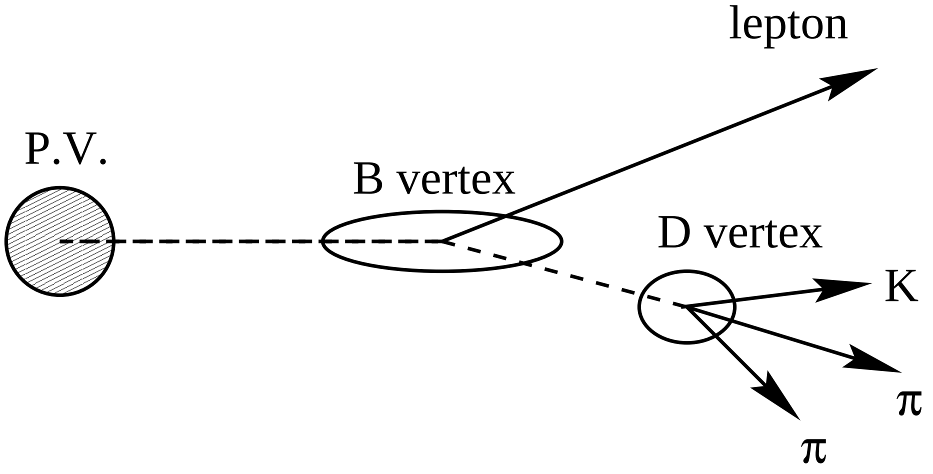


Figure 4-1: A typical $B \to \ell D$ event topology. A B meson is produced at the event primary vertex (P.V.), and decays into a lepton, a D meson, and a neutrino (undetected, and not shown). The D meson decays into a kaon and one or more pions.

In the search for the D mesons accompanying the trigger lepton, we consider sets (pairs, triplets, etc.) of tracks and test the hypothesis that they all are daughters of a D meson. This set of tracks we call a *D candidate.* (If it includes a lepton, it will be a *B candidate.*) There are many tracks in each event, and, in case there was a true D meson in the event, only one set (if at all) will be a true D meson candidate. These candidates comprise the *signal.* All other combinations, that include some or none of the D meson daughters are *not* D mesons – they comprise the *background.* The goal of the search is to find as many true D mesons (maximize the signal), and suppress the false D candidates (minimize the background). Unfortunately, no set of selection requirements will remove all background and leave us with the pure signal, and every set of selection requirements will remove some signal as well.

We start with the preselection of the possible D-meson daughter tracks. All tracks are required to pass standard CTC quality requirements: there must be ≥ 5 hits in at least 2 axial superlayers, and ≥ 2 hits in at least two stereo superlayers. We also extrapolate the track to the z position corresponding to the end of the CTC and require that the radius of that point ("the CTC exit radius") be $R_{exit} > 130$ cm. This requirement guarantees that the track is fully contained within the CTC. Because of the requirement that the D-meson daughters form a vertex we demand that all tracks (except one in the case of $D^0 \rightarrow K3\pi$) are in the SVX (meaning that they have SVX information associated with them; that, however, does not necessarily imply a SVX quality requirement – except in some cases).

For all decay signatures, we first loop over all leptons (selected with loose quality criteria) with[4] $p_T > 6.0$ GeV/c. The candidates for the daughters of the D meson must lie within a cone of $\Delta R < 1.0$ around the lepton. A Monte Carlo study shows that for the B meson energies typical for CDF, the daughters of the D meson associated with a trigger lepton are fully contained within a cone of this size.[5] They must also come from the neighborhood around the primary vertex of the lepton(since the D^0 meson lifetime is $\sim 125\mu$m, and the D^- lifetime is $\sim 200\mu$m): the z of the point of the closest approach to the primary vertex (Z_0) for each track must lie within 5 cm of the lepton's Z_0.

These tracks are then used in the search of the exclusive decay of a D meson. We consider those that also pass a minimal p_T requirement and that are displaced from the primary vertex. The fragmentation and underlying event particles that we are trying to suppress are coming from the primary vertex, and have on average lower momentum than the D meson decay products. The charges of the candidate tracks must also match the hypothesis of the D meson: for example, in $\overline{D^0} \rightarrow K^+\pi^-$, the K and π candidate tracks must have the opposite charge. The invariant mass of the candidate tracks must be in a window around the nominal D-meson mass (1.8645 GeV/c^2): between 1.8 GeV/c^2 and 1.95 GeV/c^2.

Once the tracks for the D-meson candidate have been found, they are constrained

---

[4] The trigger leptons from the inclusive lepton dataset have an implicit p_T cut (see section 3.2.6), so the requirement $p_T > 6.0$ GeV/c passes all trigger leptons while removing non-trigger lepton candidates.

[5] Some other CDF analyses use the cone of $\Delta R < 0.7$.

to intersect at a common point, the D-meson decay vertex.The track parameters are recomputed subject to this constraint, and, for this reason, the invariant mass of the D-daughter candidate tracks may change. After the D-vertex constraint, we verify that the invariant mass of the D meson candidate is still within its limits. The resulting D vertex must satisfy some further quality requirements, including the maximum χ^2 of the fit, and a requirement that the separation from the primary vertex is greater than its uncertainty ($L_{\mathrm{xy}} > \sigma_{L_{\mathrm{xy}}}$). From the list of the primary vertices determined by the VTX, we select the closest one to the D vertex, and use it as the primary vertex for this candidate. The x and y coordinates are taken from the run averaged beam position determined by the SVX.

Once the D meson has been found, we combine it with the lepton to get the B vertex. In the case of the $B \to \ell D^{*-}$ decays, the lepton and the soft pion from the D^{*-} decay, π_s^-, are both coming from the B decay point($D^* \to \overline{D^0}\pi_s^-$ is a strong decay, and the D^* meson decays immediately), so the fit for the B vertex involves intersecting the lepton and the soft pion tracks at the B vertex and requiring that the D meson points back to it. For $B \to \ell D^0$ or $B \to \ell D^-$, there is no additional track emerging from the B decay, so we just point the D meson back towards the lepton track, and the point where they intersect is taken as the B vertex.

In the end, we require that the charges of the lepton and the charm candidates be consistent with the hypothesis that they come from the B meson: we require that the B^0 candidate is neutral, and that the charge of the B^+ candidate is $+1$.Since the charge of the K carries the flavor of the D meson, we require that the lepton and kaon are of the same charge.

The decays $B^0 \to \nu\ell^+ D^{*-}$ followed by $D^{*-} \to \pi_s^- \overline{D^0}$ final state also contribute to the ℓD^0 final state. In order to improve the separation between the B^0 and B^+ samples, we remove all ℓD^0 candidates that are also ℓD^{*-} candidates. We define a D^{*-} candidate as a valid D^0 candidate with another track – a π_s^- candidate – that makes the mass difference $m(D^0, \pi_s^-) - m(D^0)$ consistent with the world average value of $m_{D^*} - m_{D^0}$. Therefore, we remove any D^0 candidate for which there is at least one other track that makes the $m(D^0, \pi) - m(D^0)$ consistent with the world average.

Since the $m(D^0, \pi) - m(D^0)$ distribution for real D^{*-} mesons is very narrow (~ 3 MeV), this removal is very efficient once the π_s is reconstructed.[6]

The narrow $m(D^0, \pi) - m(D^0)$ distribution is also exploited in the last $B^0 \rightarrow \ell D^{*-}$ decay signature, in which the D^0 meson decay $\overline{D^0} \rightarrow K^+\pi^-\pi^0$ is not fully reconstructed (i.e. some particles are missed) because of the unidentified π^0. The identification efficiency for $\pi^0 \rightarrow \gamma\gamma$ decays is low. We proceed as in the case of the fully reconstructed D mesons, however we count the candidates by fitting the mass difference distribution.

This outlines the general selection procedure. The specific selection requirements for each of the decay signatures are presented in Appendix D.

4.1.1 Mass distributions and the number of B meson candidates

The numbers of candidates are extracted from the fit of the invariant mass distributions. In all decay signatures but one ($\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi\pi^0$) the D meson is fully reconstructed. As a result, the distribution of the invariant mass of the real D meson is normally distributed around the true value of its mass. The width of the Gaussian is a measure of the detector resolution. However, there are sets of tracks that also satisfy all D meson selection criteria, and that therefore contribute to the invariant mass distribution of the candidate tracks. The collection of these candidates is called the *combinatorial background.* The invariant mass distribution of the combinatorial background depends on the availability of random tracks that satisfy the selection criteria. In the interval between 1.80 GeV/c^2 and 1.95 GeV/c^2, it gently falls as the invariant mass increases, since it is more likely to randomly pick a set of less energetic particles that forms a candidate at lower invariant mass. Figure 4-2 shows the invariant mass distributions (solid histogram), for the four channels of the exclusively

---

[6] As the energy of these pions in the D^{*-} rest frame is about 40 MeV/c, they are just boosted by the $\beta\gamma$ of the D^* meson; usually, they are very soft (low-p_T) in the laboratory frame. The CDF detector does not reconstruct tracks below 200 MeV/c (they curve inwards before reaching the CTC), so the soft pions, π_s^-, are not always reconstructed.)

reconstructed charm.

The signal component of the D meson candidate invariant mass distribution is modeled by a gaussian. Given that the combinatorial background varies very slowly, we model it with a linear function. The fits of the invariant mass distributions to the sum of the gaussian signal and the linear background are also shown on figure 4-2, as solid curves.

[As an aside, the fact that the signal is confined to a relatively narrow region of the invariant mass distribution allows us to obtain the signal contribution to any distribution. In order to do that we first make a distribution of an arbitrary quantity X for the *signal region*, defined in such a way that most of the signal is within boundaries of such a region. (For instance, for the D^0 meson the signal region is between 1.85 GeV/c^2 and 1.88 GeV/c^2.) Then we define the *sideband region* – two bands in the invariant mass distribution that are well separated from the signal region. (For the D^0 meson, the sidebands are between 1.80 GeV/c^2 and 1.83 GeV/c^2, and between 1.90 GeV/c^2 and 1.93 GeV/c^2.) The sidebands provide a perfect background sample, and the distribution of X for the sideband region can be, after it is scaled properly, subtracted from the distribution of X for the signal region, resulting in the distribution of X for the signal only. This process is known as *background subtraction* or *sideband subtraction*, and is very common in High Energy Physics. A distribution produced in this way is called a *sideband-subtracted* or *background-subtracted* distribution.]

The dashed histograms in figure 4-2 represent the D meson invariant mass distribution for the B candidates where the lepton and the kaon have the wrong charge correlation. The 'wrong charge' events can be combinatorial background, however they could also represent the cases where there was a real D meson and a *fake* lepton. The fact that there is no peak in the 'wrong charge' D mass distribution reassures us that *each* $\ell - D^{(*)}$ pair is actually coming from the decay of the same B meson.

In the case of the decay $\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$, the π^0 is lost, and there is no peak in the $m(K,\pi)$ invariant mass distribution: instead of a gaussian centered at m_{D^0}, there is a very broad and asymmetric distribution peaking around 1.6 GeV/c^2, with

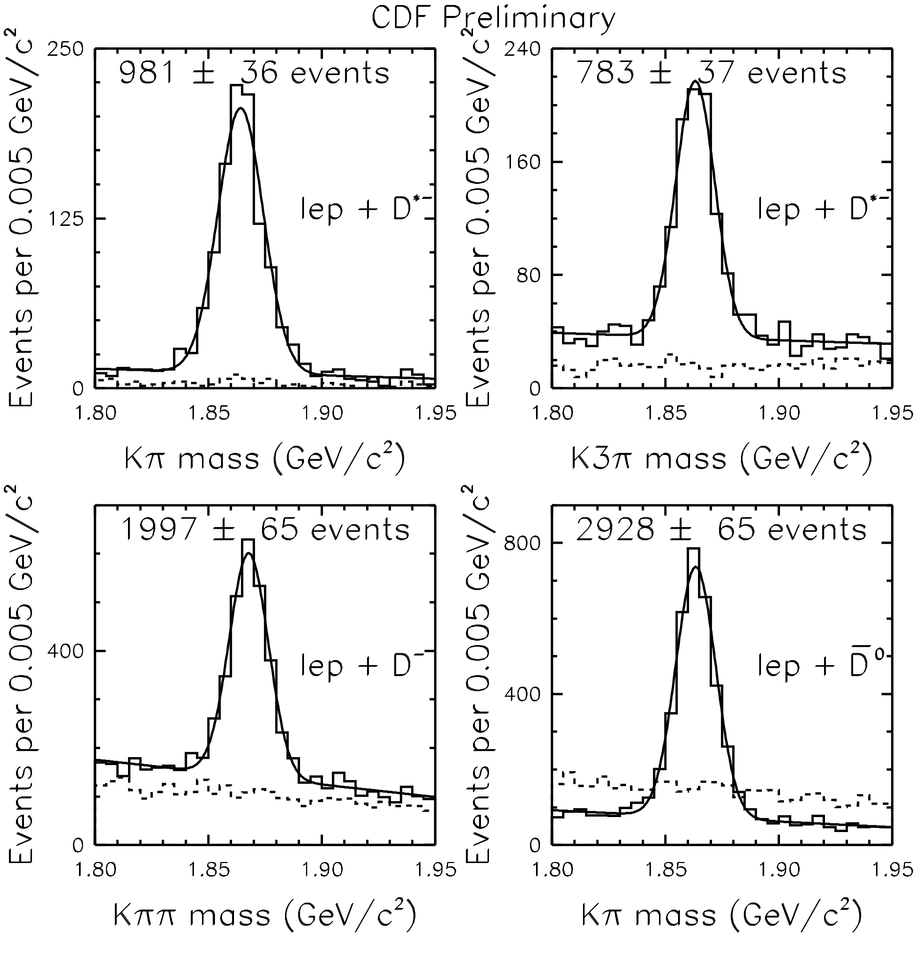


Figure 4-2: Invariant mass distributions of the fully reconstructed D meson candidate (solid histogram), for $\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi$ decay signature (upper left); $\ell^+ D^{*-}, \overline{D^0} \rightarrow K3\pi$ (upper right); $\ell^+ D^-, D^- \rightarrow K\pi\pi$ (lower left) and $\ell^+ \overline{D^0}, \overline{D^0} \rightarrow K\pi$ (lower right). The dashed histograms are the distribution of the 'wrong sign' candidates.

the cut-off at $m_{D^0} - m_{\pi^0}$. This artifact in the $K\pi$ invariant mass distribution is called a *satellite peak*, and, for this reason, the signature $\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$ is also known as the *satellite signature.*

In order to distinguish the signal from the combinatorial background, we resort to the $m(K,\pi,\pi_s) - m(K,\pi)$ distribution. If the D^0 meson were fully reconstructed, this distribution would have had a very narrow peak at $m_{D^{*-}} - m_{D^0}$, a result of the small phase space available in the $D^{*-} \to D^0\pi_s^-$ decay.[7] Even if a part of the D^0 decay was lost, the $m(K,\pi,\pi_s) - m(K,\pi)$ distribution still peaks at the $m_{D^{*-}} - m_{D^0}$ difference (since the same momentum is missing in both terms, the difference is hardly affected), except that the peak is smeared. Figure 4-3 shows the distribution of the mass difference distribution $m(K,\pi,\pi_s) - m(K,\pi)$ for this decay signature.

The complication for this decay signature is that the combinatorial background cannot be treated as a slowly varying function; on the contrary, it rises from zero at the kinematic limit of m_π (since for no combination of tracks (t_1, t_2) can the mass difference $m(t_1,t_2,\pi) - m(t_1,t_2)$ be less than m_π). We overcome this complication by making use of the invariant mass difference distribution for 'wrong sign' sample (ℓD^* pairs where the lepton and the kaon have charges of the opposite sign, since they are also combinatorial background) to give us the *shape* of the combinatorial background in the 'right sign' sample. The normalization, however, must be changed, since there is more combinatorial background in the 'right sign' sample, simply because it's more probable to pick random tracks in the ratio 2 : 2, than 3 : 1 between the 'right sign' and 'wrong sign' combinations.[8]

Therefore, we first fit the 'wrong sign' mass difference distribution. This distribution gives us the shape of the combinatorial background, and it is shown on the

---

[7] The pion mass is $m_\pi = 139$ MeV$/c^2$, while the mass difference is merely $m_{D^{*-}} - m_{D^0} \sim 145.4$ MeV$/c^2$, which is why the π_s is called the *soft pion.*

[8] Assume that, due to the charge conservation, the tracks in the cone around the lepton track are half positive and half negative (including the lepton). Out of $2n$ tracks in total, one can make $n^2(n-1)^2/4$ 2 : 2 combinations, compared with $n^2(n-1)(n-2)/3!$ 3 : 1 (but not 1 : 3) combinations. So, very roughly, the amount of the 'wrong charge' combinatorial background will be 2/3 of the amount of the 'right charge' one.

lower histogram in fig. 4-3, and the result of the fit with a dashed line. Then we fix the background shape (letting the normalization float), add the signal function, and fit for the signal area. The details of both functions used in this fit are given in the Appendix E. The result of the full fit is shown in fig. 4-3.

4.2 Monte Carlo sample

4.2.1 Monte Carlo simulation of a single B meson

In this analysis, two Monte Carlo simulations are used, depending on the context. Information concerning only the B meson is derived from a Monte Carlo simulation that uses the following algorithm:

1. A single b quark is generated according to the p_T spectrum from the inclusive b-quark production calculation by Nason, Dawson and Ellis [30]. The b quark is then transformed into a B meson, but no additional hadronization products are created. The B meson energy, however, is corrected using the Peterson fragmentation model [26], (described in Appendix C). The Peterson parameter $\epsilon = 0.006$ is used. This procedure is done inside the BGENERATOR program [31].

2. The B meson is then decayed using the QQ program [32], developed by the CLEO collaboration.

3. The B-decay products are searched for the presence of a lepton and a charm. All events are required to have a $\ell D^{(*)}$ candidate. Also, the four-body $B \to \nu \ell D^{(*)} \pi$ decays are identified as "D^{**}" decays.

4. A trigger turn-on (described below in sec. 4.2.3) is optionally applied at this stage.

5. A simulation of the CDF detector response is performed, using the QFL' program [33]. To create the CTC tracks, QFL just smears the trajectories of the

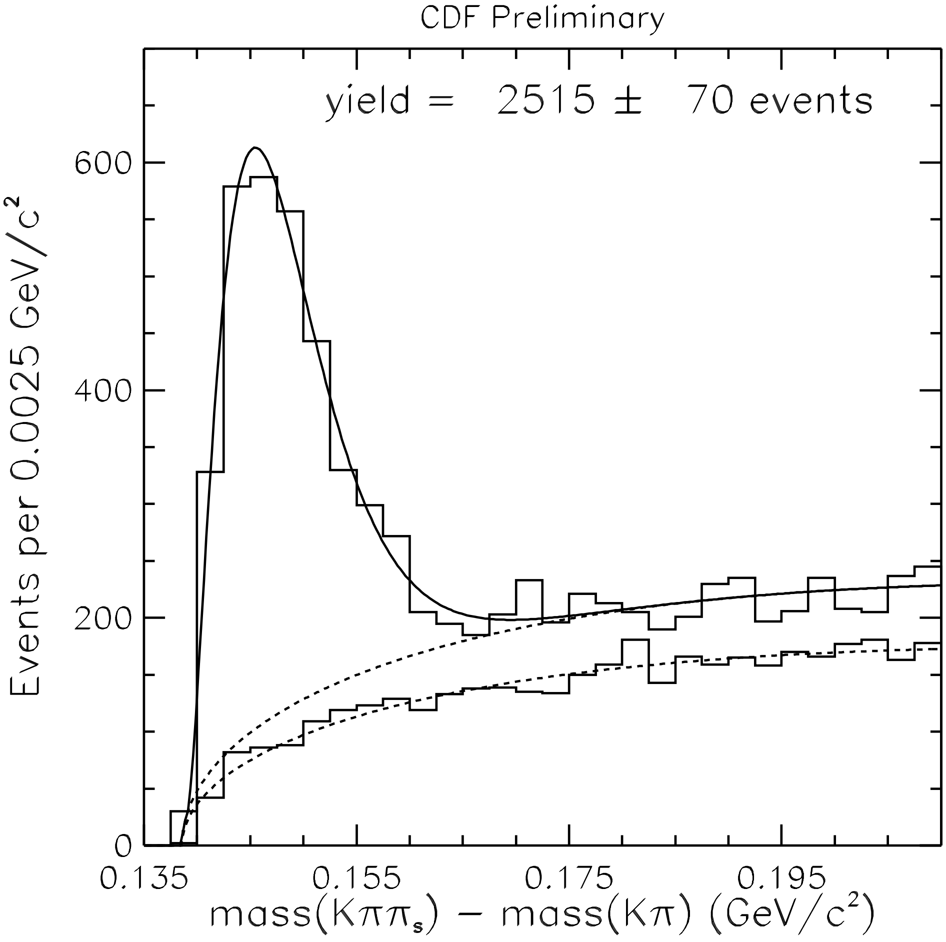


Figure 4-3: Mass distribution of $\Delta m = m(K\pi\pi_s) - m(K\pi)$ for the signature with $D^{*-} \rightarrow \bar{D}^0\pi_s^-$, with $\bar{D}^0 \rightarrow K^+\pi^-\pi^0$ (the π^0 is not reconstructed). The lower histogram is the distribution of the 'wrong sign' candidates, and the lower dashed curve is the fit to it. The higher dashed curve is the result of the fit of the 'right sign' histogram, using the 'wrong sign' distribution as the background shape.

generated particles, using parameterized resolutions. In the SVX, instead of smearing the generated trajectories, the actual SVX hits are created, and then the SVX pattern recognition is used to identify the SVX tracks. The detector simulation also includes conversions and bremsstrahlung.

6. At this point, the generated and simulated Monte Carlo B meson along with the associated decay particles fully resembles the data. To verify that the $\ell D^{(*)}$ candidate can still be found, the $\ell D^{(*)}$ reconstruction program used in the data is run on the event.

Almost all use of the Monte Carlo simulation in this thesis is confined to the simulation described above (*e.g.* in sections 4.3, 5.1.4, and 5.2.2). The uncertainties in the Monte Carlo modeling of the single B meson production and decay, as well as the modeling of the detector response, are all small. They are described in detail in sec. 8.2.

4.2.2 Monte Carlo simulation of the whole event

The other type of the Monte Carlo simulation involves the generation of a complete $p\bar{p}$ interaction: the $b\bar{b}$ pair, the hadronization products, and the underlying event. The $p\bar{p}$ interaction is generated using the PYTHIA program [34], developed by the LUND group. PYTHIA uses an improved string fragmentation model – invented by the LUND group – tuned to available experimental data. PYTHIA also produces the B^{**} states. The remaining steps (2)-(6) are identical to the case of a single b-quark generation.

We note, however, that the analysis described in this thesis has been designed to rely as little as possible on the Monte Carlo simulation of the production of additional particles along with the B meson, because this is what we are trying to study. This version of the Monte Carlo simulation is used only once, in section 7.2.3.

4.2.3 The trigger turn-on

The p_T spectrum of b-quarks produced in $p\bar{p}$ collisions, to a good approximation, decreases exponentially. In the absence of the inclusive lepton trigger, the p_T of the lepton emerging from the semileptonic B meson decay would be similarly distributed. The inclusive lepton trigger, however, applies an implicit p_T requirement (section 3.2.6), drastically modifying the lepton p_T distribution in the region below ~ 9 GeV/c.

The Monte Carlo simulation should reproduce the data distributions of the lepton p_T, the B meson p_T and $m_{\ell D}$ reasonably well. Therefore, the effect of the inclusive lepton trigger must also be simulated. In this analysis, we model the trigger efficiency (the probability that the lepton candidate passes the inclusive lepton requirement) with a single error function:

$$\epsilon_{trig}(p_T(\ell)) = \text{ Errf } (p_T(\ell), p_T^0, \sigma_{p_T}) \tag{4.6}$$

where p_T^0 and σ_{p_T} are the two parameters of the error function. Since we are not concerned with a B meson cross-section, the Monte Carlo simulation only needs to reproduce the *shape* of $\epsilon_{trig}(p_T(\ell)$, and therefore there is no need for an overall normalization factor.

We obtain p_T^0 and σ_{p_T} from the data – namely, we *choose* p_T^0 and σ_{p_T} so that the $p_T(\ell)$ distributions in the Monte Carlo simulation and the data agree. p_T^0 and σ_{p_T} are derived separately for the electron and muon events, because of the differences in the trigger. To measure the dependence of the trigger efficiency on $p_T(\ell)$, we divide the $p_T(\ell)$ distribution from the data by the $p_T(\ell)$ distribution from the Monte Carlo simulation, and then fit the ratio of these two distributions with with the error function from eq. 4.6. An example of such a fit is shown in figure 4-4. Only the region $0 < p_T(\ell) < 20$GeV/c is fitted, since this is where the effect of the trigger turn-on is the most pronounced. The results of the fits in all five decay signatures are summarized in table 4.1. For simplicity, in the final Monte Carlo simulation (step 4 of the algorithm) we use the weighted-average of the trigger turn-on parameters of all five decay signatures (table 4.1, the last row). The comparison of some distributions from the data and Monte Carlo simulation of a typical decay signature ($\ell^+ D^-, D^- \to K\pi\pi$), is given in fig. 4-5.

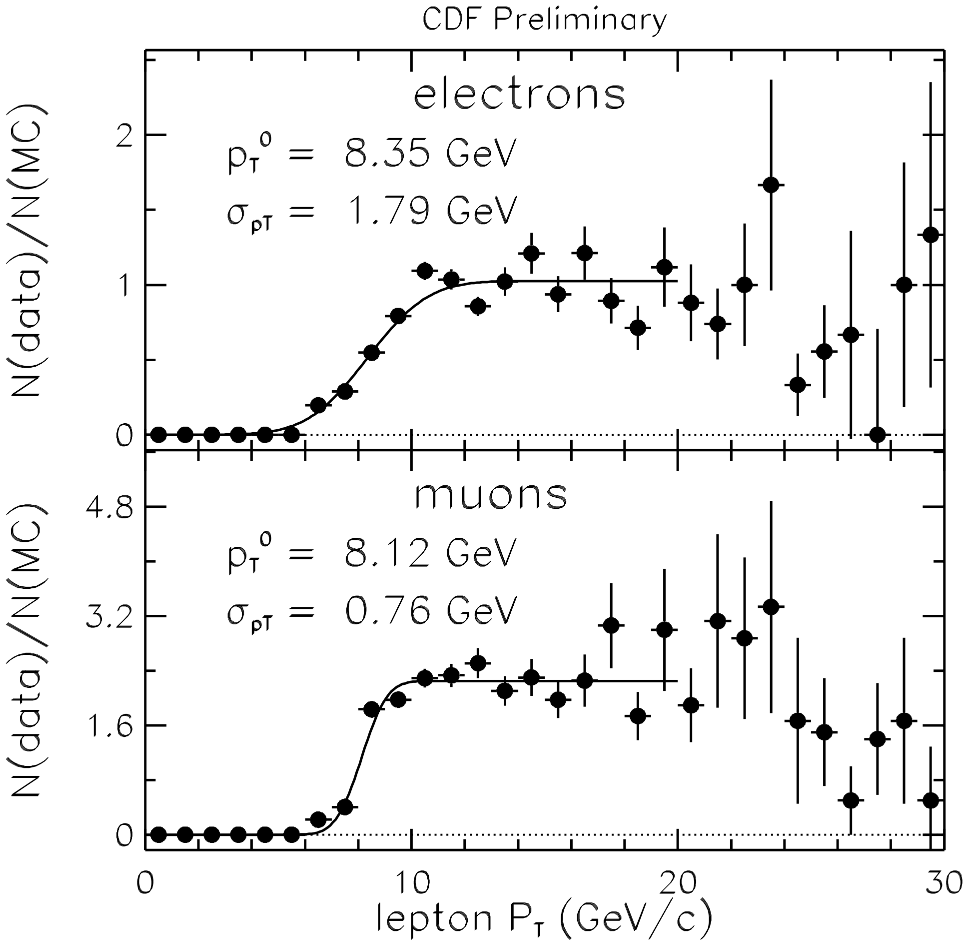


Figure 4-4: A fit to the ratios of the $p_T(\ell)$ distributions from the data and from the Monte Carlo simulation, for electrons (top) and muons (bottom), for the $\ell^+ D^-, D^- \rightarrow K\pi\pi$ signature. The distributions are fit with the error function. As only the parameters describing the function shape (p_T^0 and σ_{p_T}) are used, the overall normalization is unimportant.

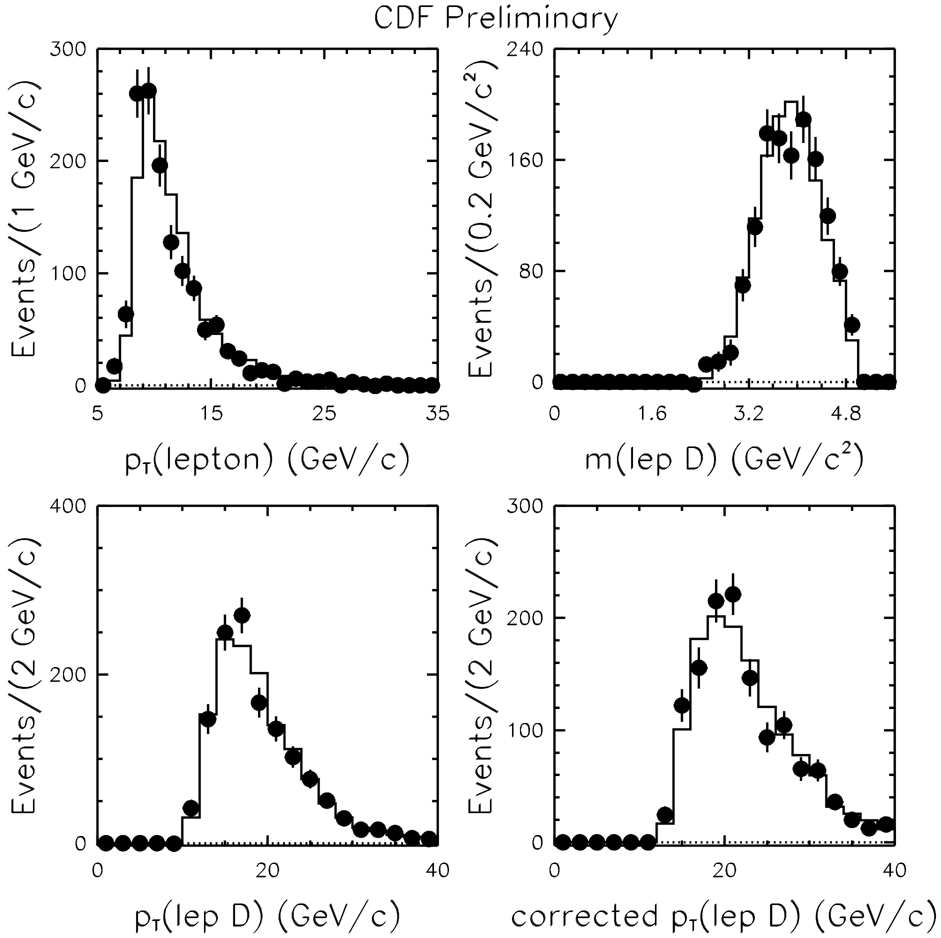


Figure 4-5: The comparison between the data and the single-B-meson Monte Carlo simulation, for the decay signature $\ell^+ D^-, D^- \rightarrow K\pi\pi$. The distributions compared are: p_T of the lepton, e and μ combined (top left), mass and the p_T of the ℓD^- system (top right), the p_T of the ℓD^- system, uncorrected (bottom left), and after correcting for the missing neutrino (bottom right). Ultimately, only the agreement in the corrected $p_T(\ell D^-)$ distribution is important.

Signature	$p_T^0(e)$	$\sigma_{p_T}(e)$	$p_T^0(\mu)$	$\sigma_{p_T}(\mu)$
$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$	8.86 ± 0.07	1.50 ± 0.07	8.10 ± 0.04	0.55 ± 0.03
$\ell^+ D^-, D^- \to K\pi\pi$	8.35 ± 0.10	1.79 ± 0.12	8.12 ± 0.10	0.76 ± 0.14
$\ell^+ D^{*-}, \overline{D^0} \to K\pi$	8.89 ± 0.20	1.66 ± 0.20	8.06 ± 0.09	0.54 ± 0.09
$\ell^+ D^{*-}, \overline{D^0} \to K3\pi$	8.42 ± 0.19	2.30 ± 0.25	8.02 ± 0.09	1.23 ± 0.12
$\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$	8.53 ± 0.21	2.56 ± 0.25	8.23 ± 0.08	0.95 ± 0.12
Weighted average	8.68 ± 0.05	1.66 ± 0.05	8.11 ± 0.03	0.61 ± 0.03

Table 4.1: The fit parameters of the inclusive lepton trigger efficiency turn-on, in all five decay signatures. All units are in GeV/c.

4.3 Measuring the proper decay time

In this section we address the determination of the proper time of the B decay, t (or, alternatively, its proper decay length, ct – in the following text, the two are used interchangeably). As outlined in section 2.4, the proper time is one of the necessary ingredients for a mixing analysis. In the semileptonic B decays $B \to \nu\ell D^{(*)}$, the neutrino escapes undetected, necessitating a correction of the measured ct. In this section, we show how this correction is done, as well as what the effect on the ct-resolution is.

4.3.1 Measuring the proper decay length

To obtain the proper decay length, we measure the two-dimensional decay length L_{xy}^B as the distance in the transverse plane between the primary and the B vertices,[9] and multiply it by the Lorentz boost, $1/(\beta\gamma) \equiv m_B/p_T^B$ (m_B is the mass of the B meson, while p_T^B is its momentum in the transverse plane):

$$ct_{true} \equiv L_{\mathrm{xy}}^B \frac{m_B}{p_T^B} \tag{4.7}$$

---

[9] As the primary vertex, we use the closest primary vertex candidate from the list of the primary vertices determined by the VTX along the beamline . We use the beamline position and slope averaged over the run.

However, since the B meson is not fully reconstructed, we approximate ct by using the transverse momentum of the ℓD system, multiplied by a factor derived from the Monte Carlo simulation:

$$ct_{true} \approx ct \equiv L_{\mathrm{xy}}^{B} \frac{m_B}{p_T^{\ell D}} \left\langle \frac{p_T^{\ell D}}{p_T^{B}} \right\rangle_{MC} \tag{4.8}$$

where p_T^B is the known transverse momentum of the generated B meson. This factor corrects for average missing momentum (e.g. from the neutrino). We call it the *$\mathcal{K}$-factor*:

$$\mathcal{K} \equiv \left\langle \frac{p_T^{\ell D}}{p_T^{B}} \right\rangle_{MC} \tag{4.9}$$

The inclusive lepton trigger (resulting in a minimum p_T requirement on the lepton) favors the decays where the neutrino takes only a small portion of the total B momentum. Since we completely reconstruct the D meson and the only particle that is missing (in principle) is the neutrino, the observed B momentum (i.e., the momentum of the ℓD system) is in fact a fairly good approximation of the initial B momentum. The distribution of the $\mathcal{K}$-factor of a typical decay signature ($\ell^+ D^{*-}, \overline{D^0} \to K\pi$) is shown in figure 4-6. The $\mathcal{K}$-factor distribution has a mean of $\approx 85\%$ and a `RMS` of $\approx 11\%$. The means and RMS for all five decay signatures are given in table 4.2

Signature	mean	RMS
$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$	0.84	0.12
$\ell^+ D^-, D^- \to K\pi\pi$	0.84	0.12
$\ell^+ D^{*-}, \overline{D^0} \to K\pi$	0.88	0.10
$\ell^+ D^{*-}, \overline{D^0} \to K3\pi$	0.88	0.10
$\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$	0.83	0.11

Table 4.2: [The mean and the RMS of the $\mathcal{K}$-factor distribution.] The mean and the RMS of the $\mathcal{K}$-factor distribution for the five decay signatures.

The less energy taken away by the neutrino, the closer the mass of the ℓD cluster, $m_{\ell D}$, is to the B meson mass, m_B. This suggests that we can do better if we allow for the additional dependence of the $\mathcal{K}$-factor on $m_{\ell D}$ and the p_T of the ℓD system, $p_T^{\ell D}$. Both are shown in the bottom of figure 4-6 (again, for the $\ell^+ D^{*-}, \overline{D^0} \to K\pi$ signature). While the $\mathcal{K}$ exhibits very weak dependence on $p_T^{\ell D}$,[10] it clearly rises with $m_{\ell D}$, approaching 1.0 as $m_{\ell D} \to m_B$. We fit the $m_{\ell D}$ distribution with a polynomial. The dependence $\mathcal{K} = \mathcal{K}(m_{\ell D})$ is employed in (4.8) on an event-by-event basis – a different correction $\mathcal{K}$ is used for each $\ell D^{(*)}$ candidate.

4.3.2 Resolution on *ct*

The proper decay distance is smeared by both tracking resolution and by the incomplete knowledge of the kinematics of the B decay. We define Δct as a difference between the reconstructed ("measured") proper decay distance, ct, and the proper decay distance ct_{true}:

$$\Delta ct \equiv ct - ct_{true} \tag{4.10}$$

Therefore,

$$\Delta ct = \Delta(L^B_{\mathrm{xy}} \frac{1}{\beta\gamma}) = \frac{1}{\beta\gamma} \Delta L^B_{\mathrm{xy}} + L^B_{\mathrm{xy}} \frac{1}{\beta\gamma} \frac{\Delta(\frac{1}{\beta\gamma})}{\frac{1}{\beta\gamma}}$$

If we approximate $1/\beta\gamma$ with its mean value, $\langle 1/\beta\gamma \rangle$, and substitute (4.8), we get

$$\Delta ct = \left\langle \frac{1}{\beta\gamma} \right\rangle \Delta L^B_{\mathrm{xy}} + ct \frac{\Delta(\frac{1}{\beta\gamma})}{\frac{1}{\beta\gamma}} \tag{4.11}$$

Equation (4.11) highlights the two resolution effects. The tracking resolution is described by the constant term $\langle 1/\beta\gamma \rangle \Delta L^B_{\mathrm{xy}}$, while the smearing due to using an average $\mathcal{K}$ (from eq. (4.9)) contributes the second term $ct\Delta(1/\beta\gamma)/(1/\beta\gamma)$, proportional to ct.

The width of the ΔL^B_{xy} distribution is a measure of the *residual detector resolution* (the resolution that remains even if we knew the decay kinematics perfectly). For the decay signature $\ell^+ D^{*-}, \overline{D^0} \to K\pi$, it is shown in figure 4-6. We fit it with three Gaussians, one for the peak (the dominant one) and the other two for the tails. The fractions and the widths of each of the Gaussians are given in the table 4.3.

---

[10] Actually, there is virtually none once the full sample composition is included. In the $B \to \nu \ell D^*$ decays $\mathcal{K}$ seems not to depend on $p_T^{\ell D}$, and even the decay signatures $\ell^+ D^-, D^- \to K\pi\pi$ and $\ell^+ \overline{D^0}, \overline{D^0} \to K\pi$ are dominated by $B \to \nu \ell D^*$ decay chains.

Signature	$g_1(\%)$	σ_1 (μm)	$g_2(\%)$	σ_2 (μm)	$g_3(\%)$	σ_3 (μm)
$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$	58	93	32	248	10	888
$\ell^+D^-, D^- \to K\pi\pi$	54	82	36	237	11	925
$\ell^+D^{*-}, \overline{D^0} \to K\pi$	50	98	32	228	18	743
$\ell^+D^{*-}, \overline{D^0} \to K3\pi$	58	106	28	259	14	850
$\ell^+D^{*-}, \overline{D^0} \to K\pi\pi^0$	54	122	34	344	12	1006

Table 4.3: The fraction (g_i) and the width (σ_i) of each of the three gaussians used to fit the ΔL_{xy}^B distribution.

On the other hand, the width of the $\Delta(1/\beta\gamma)/(1/\beta\gamma)$ distribution is a measure of the smearing due to incomplete knowledge of the kinematics and using an average $\mathcal{K}$-factor as a correction. The $\Delta(1/\beta\gamma)/(1/\beta\gamma)$ distribution for the $\ell^+D^{*-}, \overline{D^0} \to K\pi$ signature is shown in figure 4-7, while the values of the RMS are given in the table 4.4.

Signature	RMS of $\Delta(1/\beta\gamma)/(1/\beta\gamma)$	RMS of Δct (cm)
$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$	0.160	0.017
$\ell^+D^-, D^- \to K\pi\pi$	0.168	0.017
$\ell^+D^{*-}, \overline{D^0} \to K\pi$	0.139	0.014
$\ell^+D^{*-}, \overline{D^0} \to K3\pi$	0.141	0.016
$\ell^+D^{*-}, \overline{D^0} \to K\pi\pi^0$	0.144	0.015

Table 4.4: The RMS of the $\Delta(1/\beta\gamma)/(1/\beta\gamma)$ and Δct distributions, for the five decay signatures.

The resolution of the measured proper time ct, $\sigma \equiv \sigma_{ct}$, is defined as the width of the Δct distribution. Because of the second term in eq. (4.11), σ is a function of ct. There are two ways of parameterizing $\sigma = \sigma(ct)$. One is based on decomposing the ΔL_{xy}^B and $\Delta(1/\beta\gamma)/(1/\beta\gamma)$ distributions as a sum of several gaussians.

The method used in this analysis, however, is simpler and consists of making the Δct distributions for each ct bin, and fitting them with a gaussian and a constant (to account for the long tails). The width of each gaussian is taken as a value of σ_{ct} for this ct. Finally, the set of points σ_{ct} vs ct is fit with the functional form for $\sigma(ct)$ and get an analytical dependence of σ on ct. We find that a linear function

$$\sigma = \sigma_0 + b \times ct \tag{4.12}$$

fits well the σ vs ct distributions in all five decay signatures. The fit for the $\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi$ signature is shown in figure 4-7, while the results of the fit for all five signatures are summarized in table 4.5. The intercept σ_0 is the residual detector resolution, and is between $40 - 60$ μm.

Signature	σ_0 (cm)	b
$\ell^+ \overline{D^0}, \overline{D^0} \rightarrow K\pi$	0.0045	0.092
$\ell^+ D^-, D^- \rightarrow K\pi\pi$	0.0039	0.108
$\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi$	0.0052	0.075
$\ell^+ D^{*-}, \overline{D^0} \rightarrow K3\pi$	0.0049	0.073
$\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi\pi^0$	0.0062	0.070

Table 4.5: The parameters of the linear model of the $\sigma(ct)$ dependence (equation 4.12), for the five decay signatures.

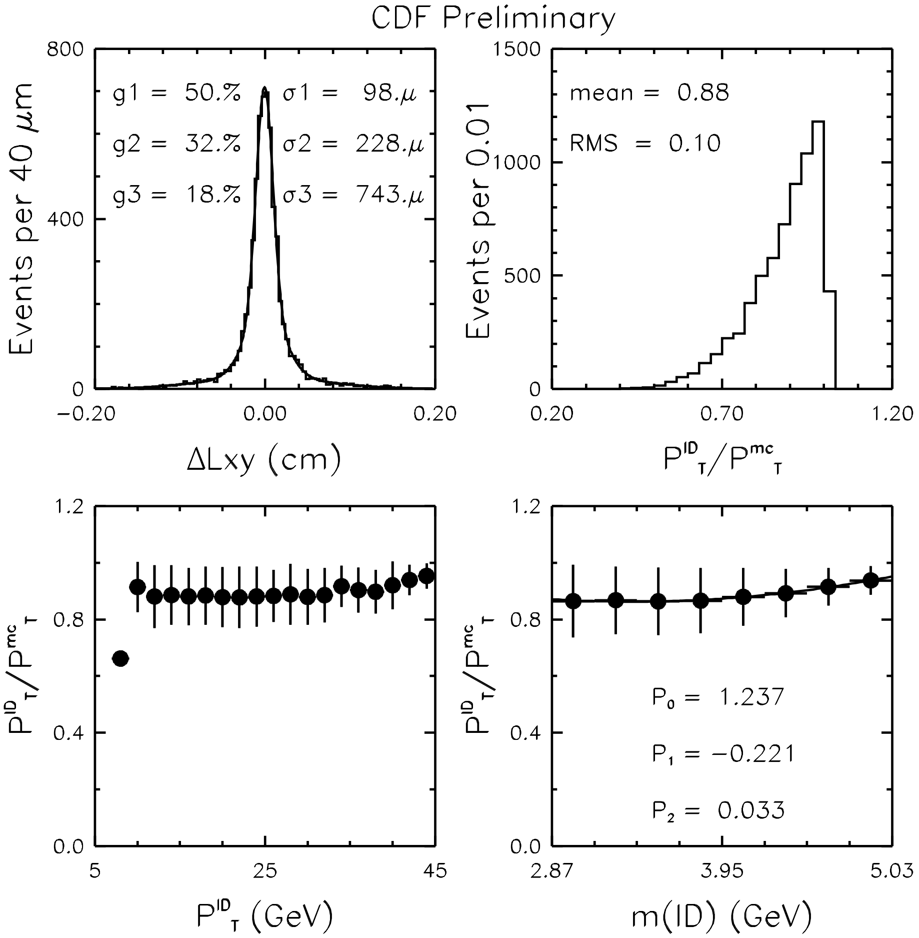


Figure 4-6: Decay signature $\ell^+ D^{*-}$, $D^0 \to K\pi$: ΔL^B_{xy} distribution, fitted with three gaussians. The distribution of $\mathcal{K} \equiv p_T^{\ell D}/p_T^B$. $\mathcal{K}$ vs. $p_T^{\ell D}$ and $\mathcal{K}$ vs $m(\ell D)$, fitted with quadratic function. The error bars in K vs $p_T^{\ell D}$ and $\mathcal{K}$ vs $m(\ell D)$ distributions represent the spread of the $\mathcal{K}$ distribution in each bin, and thus are uncorrelated with the number of events in the bin.

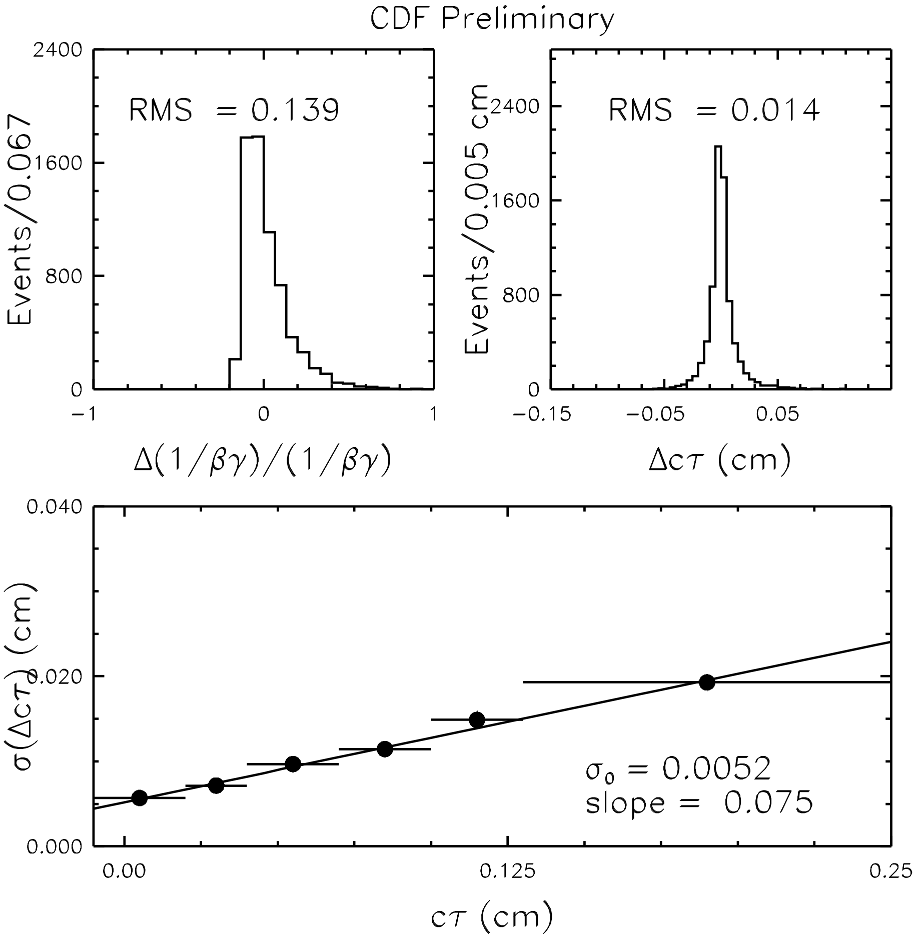


Figure 4-7: Decay signature ℓD^{*+}, $D^0 \to K\pi$: $\Delta(1/\beta\gamma)/(1/\beta\gamma)$ distribution, Δct distribution, and $\sigma(ct)$ distribution.

Chapter 5

The sample composition

5.1 The cross-talk in the $\ell D^{(*)}$ sample

A major complication with $B \to \nu\ell D^{(*)}$ analyses that depend on the distinction between the B^0 and B^+ semileptonic decays, such as the measurement of B^0 mixing, or the measurement of the ratio of the B^+ and B^0 lifetimes, is the imperfect separation between the $B^0 \to \ell D^{(*)}$ and $B^+ \to \ell D^{(*)}$ samples. Since the B meson is not fully reconstructed, one cannot use the B meson invariant mass to guarantee that all B decay products have been accounted for. When a charged daughter particle is missed, the decay of the B^+ meson is sometimes misclassified as a B^0, and vice versa. We say that there is *cross-talk* or *cross-contamination* between the two samples.

The distinction between the B^0 and B^+ samples is often essential. In the case of the measurement of τ_{B^+}/τ_{B^0}, one has to separate two lifetimes that are very close, and this separation is very sensitive to the amount of the cross-contamination. For the mixing measurement using Opposite Side Tagging, the time dependence of the asymmetry (eq.(2.15)) is described by an oscillation plus a constant determined by the sample composition.[1] This thesis documents the analysis of the B^0 mixing using Same Side Tagging, in which the flavor of the B meson is tagged by the charge of the pion preferentially produced with it. As will be discussed in section 6.1, the

---

[1] In Opposite Side Tagging, the B^0 and B^+ mesons are tagged in the same way (they indeed have the same dilution), so the asymmetries add.

correlation between the pion charge and the b-quark flavor depends on whether the b-quark forms a B^0 or a B^+ meson (figure 6-1). So, in the case of Same Side Tagging, the dependence of the asymmetry on the proper time is described by a cosine *minus* a constant. Since the goal of a mixing analysis is the measurement of the oscillation frequency, it is therefore crucial to know the amount of the cross talk, *i.e.* the size of this constant.

This chapter first lists the sources of the cross talk, and then describes the parameters (branching ratios and efficiencies) that govern it. Various cross-talk decay chains are differently affected by the B meson selection (section 4.1), so additional efficiencies are introduced in order to correct for this effect. Finally, one of the efficiencies is derived from the data, and, in the end of this section, we show how this is done. At that point, we will be able to calculate the composition of the $B \to \nu\ell D^{(*)}$ sample if the proper decay time is not needed (*i.e.*, the sample composition is integrated over the proper decay time).

However, both mixing and lifetime measurements require knowledge of the proper time of the B decay. Unfortunately, the amount of the cross-talk (i.e. the sample composition) depends on the proper time, and, conversely, the proper time depends on the sample composition. This complicated inter-dependence is analyzed in section 5.2.

5.1.1 Sources of the cross-talk

There are two causes of the cross-talk:

- **Missing a soft pion from the D^{*-} decay.** For example, the decay

  $$B^+ \to \nu\ell^+\overline{D^0} \tag{5.1}$$

  can be mimicked by the decay chain

  $$\begin{aligned} &B^0 \to \nu\ell^+D^{*-} \\ &D^{*-} \to \overline{D^0}\pi_s^- \end{aligned} \tag{5.2}$$

  if the soft pion π_s^- is not reconstructed. The parameter that quantifies the contamination via this decay chain is the reconstruction efficiency (the probability that the particle is reconstructed in the detector) for the soft pion π_s, $\epsilon(\pi_s)$. A Monte Carlo study suggests $\epsilon(\pi_s) \approx 0.88$. However, it is desirable to derive this number from the data, as the Monte Carlo simulation may not model well the detector response for low-p_T tracks. A method for measuring $\epsilon(\pi_s)$ is described in section 5.1.5.

- ***B* decays to orbitally excited *D*-mesons, D^{**}.** The sequence of decays

$$B^0 \rightarrow \nu\ell^+ D^{**-}$$

$$D^{**-} \rightarrow \overline{D^0}\pi^-_{**} \tag{5.3}$$

will also mimic the $\ell^+\overline{D^0}$ signature of the B^0 meson, since in most cases we cannot fully reconstruct the D^{**} meson (see table 5.1). There are four expected D^{**} resonances, some of which decay into $D\pi$, others to $D^*\pi$, and one to both. The total decay rate to these states is not known well and the proportion of the four possible D^{**} states is almost totally unknown. The predicted D^{**} states are listed in Table 5.1, together with their possible decay modes. So far there is evidence that the $D_1(2420)$ and $D_2^*(2460)$ states are produced at some level [41].

name	J^P	width	decay modes
D_0^*	0^+	wide	$D\pi$
D_1^*	1^+	wide	$D^*\pi$
$D_1(2420)$	1^+	narrow	$D^*\pi$
$D_2^*(2460)$	2^+	narrow	$D\pi$, $D^*\pi$

Table 5.1: The expected D^{**} states with the widths expected from Heavy Quark Effective Theory. (In J^P, J is the total angular momentum, and P is parity.)

There may also be non-resonant $D^{(*)}\pi$ production (from the B meson four-body decays $B \rightarrow \nu\ell D^{(*)}\pi$) that have the same cross-talk effect [41]. It would be extraordinarily difficult to distinguish these decays from the two D^{**} resonances which are predicted to be wide by Heavy Quark Effective Theory [42]. From here onwards, by "D^{**}" we mean both two narrow and two wide resonances, as well as the four-body semileptonic decay of the B meson.

All in all, only the two narrow states can potentially be fully reconstructed and thus explicitly excluded from the sample, however, the removed portion of the D^{**} states would also depend on the poorly known D^{**} composition (D_0^* : D_1^* : $D_1(2420)$: $D_2^*(2460)$: four-body B decays). Since the bulk of the D^{**} would still remain, the removal of the narrow states would not diminish this source of the cross-talk, so we opted not to pursue it.

The full picture of the sample composition is more complicated, since both B^0 and B^+ mesons decay into "D^{**}'s", and D^{**-} and $\overline{D^{**0}}$ decay into both D^{*-} and $\overline{D^{*0}}$, as well as D^- and $\overline{D^0}$. The full complexity of the sample composition is illustrated in the state diagram, shown in figure 5-1. From this diagram it is possible to tabulate all possible decay sequences that feed into a decay signature. Our goal is to estimate the composition of each decay signature, *i.e.*, the fraction due to B^0, and the fraction due to B^+ mesons), and we will do it by summing the contribution of all decay chains to each decay signature that originated from a B^0 or B^+ state. The contribution of each decay chain can be calculated knowing the $\epsilon(\pi_s)$, D^{**} production and branching ratios, as well as a few other branching ratios. These quantities fully describe the sample composition, and we refer to them as the *sample composition parameters.*

5.1.2 Sample composition parameters

In this section we discuss the sample composition parameters – the set of numbers needed to estimate the amount of cross-talk. We defer the calculation of $\epsilon(\pi_s)$ until section 5.1.5, since it requires the full knowledge of the sample composition.[2] Next, we

---

[2] The reconstruction efficiency for the soft pion from D^{*-}, $\epsilon(\pi_s)$, gives the fraction of $\ell^- D^{*-}$ events that 'leaks' into the $\ell^- \overline{D^0}$ signature (that fraction is $1 - \epsilon(\pi_s)$). However, the total amount of 'leakage' is also proportional to the total number of $\ell^- D^{*-}$ candidates, which in turn depends on the D^{**} production and decays (see figure 5-1). Therefore, in order to quantify this 'leakage', we need to know all other details of the sample composition picture.

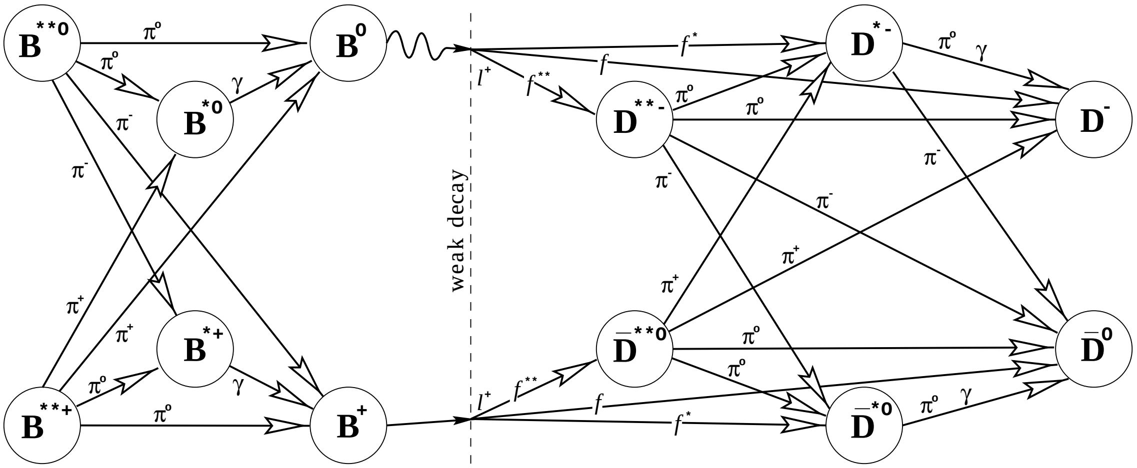


Figure 5-1: The state diagram for all possible $B \to \ell D^{(*)} X$ transitions.

list other necessary branching ratios, and note how the unequal B^0 and B^+ lifetimes can also disturb the sample composition. In section 5.1.4 we correct for the bias of the data selection requirements. At that point we will have treated all sample composition parameters, except $\epsilon(\pi_s)$, so we return to it and show how it can be constrained from the data.

We first concentrate on describing D^{**} production and decays. To estimate the cross-talk due to $B \to \nu\ell D^{**}$ decays, it is necessary to know two numbers: the D^{**} *fraction* parameter – the fraction of B meson semileptonic decays that proceed through a D^{**},

$$f^{**} \equiv \frac{\mathcal{B}(B \to \nu\ell D^{**})}{\mathcal{B}(B \to \nu\ell X)} \tag{5.4}$$

and the D^{**} *composition parameter* – the fraction of D^{**} mesons that decay into $D^*\pi$,

$$P_V \equiv \frac{\mathcal{B}(B \to D^{**} \to D^*\pi_{**})}{\mathcal{B}(B \to D^{**} \to D^*\pi_{**}) + \mathcal{B}(B \to D^{**} \to D\pi_{**})} \tag{5.5}$$

From here onward, by π_{**} we denote the pion coming from the D^{**} decay. As three of D^{**} states decay to either $D^*\pi$ or $D\pi$, P_V is just an effective quantity, depending on the relative composition of $B \to \nu\ell D^{**}$ decays. This number is very poorly known, and most analyses of semileptonic B decays in CDF and elsewhere assume it to be between 0.34 and 0.78 and assign to it huge systematic errors. We let this parameter float freely, since it turns out (section 7.2.3) that P_V can be indirectly constrained from other quantities we also measure from the data. As will be shown in sec. 7.3, our data also suggest a low P_V (≈ 0.35).

The D^{**} fraction parameter, f^{**}, is also not very well known, however there are experimental constraints on its value. The other two B semileptonic fractions,

$$f = \frac{\mathcal{B}(B \to \nu \ell D)}{\mathcal{B}(B \to \nu \ell X)} \tag{5.6}$$

$$f^* = \frac{\mathcal{B}(B \to \nu \ell D^*)}{\mathcal{B}(B \to \nu \ell X)} \tag{5.7}$$

are much better known (measured by CLEO [57]). Given the definition of "D^{**}" we have adopted, *i.e.*, it includes both B semileptonic decays into the orbitally excited D mesons, and the four-body semileptonic B decays, the f^{**} covers all $b \to \nu \ell c$ transitions not included in f and f^*, and all three fractions add to unity: $f + f^* + f^{**} = 1$. Therefore, there are only two independent parameters. Even though f and f^* are directly measured, we elect to use f^{**} and $R_f \equiv f^*/f$, since f^{**} highlights the needed information (and is more commonly used elsewhere). Using the CLEO measurements for $B \to \nu \ell D$ and $B \to \nu \ell D^*$ branching ratios [57], and the inclusive branching ratio to a lowest lying D-meson (a D^- or a $\overline{D^0}$):

$$\mathcal{B}(B \to \nu \ell D) = 1.7 \pm 0.5 \tag{5.8}$$

$$\mathcal{B}(B \to \nu \ell D^*) = 4.4 \pm 0.7 \tag{5.9}$$

$$\mathcal{B}(B \to \nu \ell D X) = 9.6 \pm 1.6 \tag{5.10}$$

we derive[3] $f^{**} = 1 - f - f^* = 0.36 \pm 0.12$ and $R_f = f^*/f = 2.5 \pm 0.6$.

The amount of cross-contamination also depends on the relative semileptonic branching ratios: although the production rates of the B^0 and B^+ mesons are equal, the fraction of the B^0 or B^+ decays that proceed semileptonically may not be – thus affecting the sample composition. According to the spectator model of the B meson

---

[3] Strictly speaking, CLEO reports $f + f^* = 0.64 \pm 0.12$, from which we derive f^{**}. If one starts from eqs. (5.8) and (5.9), one gets a larger error, since the error on $\mathcal{B}(B \to \nu \ell D X)$ is being counted twice.

(in which the b quark decays semileptonically, and the light quark is just a 'spectator'), the semileptonic width, Γ_{sl}, is expected to be the same for the B^0 and B^+ mesons. If the hadronic widths Γ_{had} are different for the B^0 and B^+ mesons, the total widths, defined as $\Gamma_{tot} = \Gamma_{sl} + \Gamma_{had}$, will also be different. The ratio of the semileptonic branching ratios $\mathcal{B}_{sl} = \Gamma_{sl}/\Gamma_{tot}$ for the B^0 and B^+ mesons is then proportional to the ratio of their lifetimes

$$\frac{\mathcal{B}_{sl}(B^+)}{\mathcal{B}_{sl}(B^0)} = \frac{\Gamma_{sl}(B^+)/\Gamma_{tot}(B^+)}{\Gamma_{sl}(B^0)/\Gamma_{tot}(B^0)} = \frac{\Gamma_{tot}(B^0)}{\Gamma_{tot}(B^+)} = \frac{\tau_{B^+}}{\tau_{B^0}} \tag{5.11}$$

since $\Gamma_{tot}\tau = \hbar$. The ratio of B lifetimes, τ_{B^+}/τ_{B^0} is therefore another sample composition parameter. The world-average value is

$$\frac{\tau_{B^+}}{\tau_{B^0}} = (1.02 \pm 0.05)\ \text{ps}^{-1}$$

so the effect of different lifetimes is not large.

The sample composition parameters and their values are summarized in table 5.2. The reconstruction efficiency for the soft pion, $\epsilon(\pi_s)$ is not shown since it is derived from the data and other sample composition parameters (see section 5.1.5). With the use of isospin relations such as

$$\frac{\mathcal{B}(D^{**-} \to D^{*-}\pi^0)}{\mathcal{B}(D^{**-} \to \overline{D^{*0}}\pi^-)} = \frac{1}{2}$$

and the knowledge of the world-average values for the D^{*-} branching ratios [53]

$$\begin{aligned} \mathcal{B}(D^{*-} \to D^-\pi^0) &= 0.317 \pm 0.008 \\ \mathcal{B}(D^{*-} \to \overline{D^0}\pi^-) &= 0.683 \pm 0.013 \end{aligned}$$

specific predictions for the cross-talk can be made. In the following section, we show how that is done.

5.1.3 Calculating the sample composition

To calculate the sample composition means to estimate the amount of cross-talk, i.e. to *estimate the fractions of each decay signature coming from the B^0 and B^+ decays.* In general, knowing the sample composition in an arbitrary sample of the

name	central value	lower limit	upper limit
f^{**}	0.36	0.24	0.46
R_f	2.5	1.9	3.1
τ_{B^+}/τ_{B^0}	1.02	0.97	1.07

Table 5.2: The values of the sample composition parameters and their variations.

semileptonic B-meson decays (e.g. a bin in *ct*), is equivalent to knowing how many decays originated from the B^0 and how many from the B^+ decay.

In order to calculate the sample composition, we tabulate all possible sequences of decays (*decay chains*) that comprise a B^0 or B^+ signature. (Decay sequences (5.2) and (5.3) are examples of decay chains.) Each of the five decay signatures considered in this analysis consists of several decay chains: three for every ℓD^{*-}, nine for the ℓD^- and twelve for the ℓD^0. All decay chains contributing to each of the decay signatures are listed in Appendix F. As an illustration, table 5.3 gives the possible decay chains [4] feeding into the decay signature ℓD^{*-}.

---

[4] We use a compact numerical notation for each of the decay chains. In the following chapters, we could have simply referred to quantities of any decay chain by its equation number, however, this notation is more useful since it has a physical meaning . The *decay chain code* is always a six-digit integer. The chain can involve a maximum of six particles: B, D^{**}, π_{**}, D^*, π_s and D, and for every chain it is sufficient to know whether the particle was present, and whether it was neutral or charged. Thus, for each of them we have only three possibilities: 0 means the particle is not present, 1 that it is neutral, and 2 that it is charged. These numbers are then strung in the above order (B, D^{**}, π_{**}, D^*, π_s and D). Examples:

- 100.002 means that there were a neutral B and a charged D, thus $B^0 \to D^+$.
- xxx.221 means that there were D^{*+}, π_s^+, D^0, thus the chain involves $D^{*+} \to D^0\pi_s^+$ (a frequent combination).
- xxx.202 means that there were D^{*+} and D^+, but *not* π^0. The only remaining possibility is a photon, so the chain involves $D^{*+} \to D^+\gamma$.

Et cetera. A dot is added in the middle for readability.

code	decay chain	$\mathcal{B}/\mathcal{B}(B \to \nu\ell X)$
100.221	$B^0 \to \nu\ell^+ D^{*-}$	f^*
121.221	$B^0 \to \nu\ell^+ D^{**-}$, $D^{**-} \to D^{*-}\pi^0_{**}$ (didn't detect π^0_{**})	$\frac{1}{3}f^{**}P_V$
212.221	$B^+ \to \nu\ell^+ \overline{D}^{**0}$, $\overline{D}^{**0} \to D^{*-}\pi^+_{**}$ (missed π^+_{**})	$\frac{2}{3}f^{**}P_V$

Table 5.3: The list of decay chains contributing to the three ℓD^{*-} decay signatures, and their relative contributions.

In order to use a compact notation, we label the decay signatures with an index k. A decay chain feeding into a given decay signature k, is labeled by ℓ. (Examples of ℓ would be [100.221], [121.221] or [212.221].) Every decay chain from Appendix F also has the branching ratio relative to the total inclusive B semileptonic branching ratio (e.g. the column $\mathcal{B}/\mathcal{B}(B \to \nu\ell X)$ in table 5.3). We denote this quantity by $\phi_{k\ell}$,

$$\phi_{k\ell} \equiv \frac{\mathcal{B}_{k\ell}}{\mathcal{B}(B \to \nu\ell X)} \tag{5.12}$$

The way $\phi_{k\ell}$'s are defined, they do not include the effect of different lifetimes.[5] If, for the moment, we ignore the effect of potentially different B^0 and B^+ semileptonic decay rates, the composition of a $B \to \nu\ell D^{(*)}$ signature k can be expressed using the fractions of the B^0 and B^+ semileptonic decays:

$$\Phi^0_k \equiv \frac{\mathcal{B}(B^0 \to \text{“}k\text{”})}{\mathcal{B}(B^0 \to \nu\ell X)} \tag{5.13}$$

$$\Phi^+_k \equiv \frac{\mathcal{B}(B^+ \to \text{“}k\text{”})}{\mathcal{B}(B^+ \to \nu\ell X)} \tag{5.14}$$

In terms of $\phi_{k\ell}$, Φ^0_k and Φ^+_k are just sums of contributions of each individual decay chain ℓ:

$$\Phi^0_k = \sum_{B^0 \to \ell} \phi_{k\ell} \tag{5.15}$$

$$\Phi^+_k = \sum_{B^+ \to \ell} \phi_{k\ell} \tag{5.16}$$

We denote the summation over the decay chains that originate with B^0 as $\sum_{B^0 \to \ell}$ and over the chains that originate with B^+ as $\sum_{B^+ \to \ell}$. By their nature, Φ^0_k and Φ^+_k

---

[5] The reason for this will be apparent in section 5.2.

are the total contributions of the B^0 and B^+ decays to the decay signature k, so the composition of k can be summarized as:

$$\text{fraction of } k \text{ from } B^0 = \frac{\Phi_k^0}{\Phi_k^0 + \Phi_k^+} \tag{5.17}$$

$$\text{fraction of } k \text{ from } B^+ = \frac{\Phi_k^+}{\Phi_k^0 + \Phi_k^+} \tag{5.18}$$

If we allow that the B^0 and B^+ lifetimes (and thus the semileptonic decay rates) are different, then, according to eq. (5.11), the B^+ component with respect to the B^0 component must be scaled by the lifetime ratio, τ_{B^+}/τ_{B^0}, to account to the larger fraction of the B^+ mesons decaying semileptonically. Explicitly,

$$\text{fraction of } k \text{ from } B^0 = \frac{\tau_{B^0}\sum_{B^0\to\ell}\phi_{k\ell}}{\tau_{B^0}\sum_{B^0\to\ell}\phi_{k\ell} + \tau_{B^+}\sum_{B^+\to\ell}\phi_{k\ell}}$$

$$\text{fraction of } k \text{ from } B^+ = \frac{\tau_{B^+}\sum_{B^+\to\ell}\phi_{k\ell}}{\tau_{B^+}\sum_{B^+\to\ell}\phi_{k\ell} + \tau_{B^0}\sum_{B^0\to\ell}\phi_{k\ell}}$$

In order to preserve the form of equations (5.17) and (5.18), we absorb the B lifetimes into the definition of Φ: the $\phi_{k\ell}$ for a decay chain ℓ that originated from a B^0 is multiplied by τ_{B0}, and, analogously, if ℓ started with a B^+ decay, $\phi_{k\ell}$ is multiplied by τ_{B+}:

$$\Phi_k^0 = \tau_{B^0}\sum_{B^0\to\ell}\phi_{k\ell} \tag{5.19}$$

$$\Phi_k^+ = \tau_{B^+}\sum_{B^+\to\ell}\phi_{k\ell} \tag{5.20}$$

In summary, equations (5.17) and (5.18), using Φ^0 from eq. (5.19) and Φ^+ from eq. (5.20) give us the sample composition of the decay signature k. However, the calculation of the sample composition outlined so far does not take into account a potential bias introduced by the event selection. This bias can be corrected for on a chain-by-chain basis, and the definitions of Φ's must be changed accordingly. These corrections are described in the following section.

5.1.4 Relative charm reconstruction efficiencies

The lepton and the D meson p_T spectra in a semileptonic B decay depend, among other things, on the mass and the spin of the D meson. The D, D^* and the host of

D^{**} mesons have different masses and spins, and one may expect differences in the lepton p_T spectra across the decay channels.

All events in the lepton + charm sample have been recorded with the inclusive lepton trigger, that, on average, selects electrons with $p_T > 8$ GeV/c and muons with $p_T > 9$ GeV/c. The number of B decays that pass the lepton trigger requirement obviously depends on the p_T-distribution of the lepton. Other reconstruction requirements have similar effects.[6]

code	decay chain	$\mathcal{B}/\mathcal{B}(B \to \nu\ell X)$
100.221	$B^0 \to \nu\ell^+ D^{*-}$	f^*
121.221	$B^0 \to \nu\ell^+ D^{**-}$, $D^{**-} \to D^{*-}\pi^0_{**}$ (didn't detect π^0_{**})	$\frac{1}{3}f^{**}P_V$
212.221	$B^+ \to \nu\ell^+\overline{D^{**0}}$, $\overline{D^{**0}} \to D^{*-}\pi^+_{**}$ (missed π^+_{**})	$\frac{2}{3}f^{**}P_V$

Table 5.4: The list of decay chains contributing to the three ℓD^{*-} decay signatures, and their relative contributions.

For example, in the decay signature $\ell^+ D^{*-}$, the probability that an event from a decay chain listed in table 5.4 (the same as table 5.3, here just repeated for convenience) passes the selection requirements (and is thus recognized as a $\ell^- D^{*+}$ candidate) varies across decay chains. This probability we call the *reconstruction efficiency* and denote by it $\epsilon(\ell D^{*-})$. (As some D^{**} decay products are missed, the term 'reconstruction efficiency' applies only to the part of the decay we identify as $\ell^+ D^{*-}$.)

Consider the reconstruction efficiencies for the decay signature $\ell^+ D^{*-}$, $\epsilon(\ell^+ D^{*-}|B \to D^*)$ for $B \to D^*$, and $\epsilon(\ell^+ D^{*-}|B \to D^{**})$ for $B \to D^{**}$. If $\epsilon(\ell^+ D^{*-}|B \to D^{**}) \neq \epsilon(\ell^+ D^{*-}|B \to D^*)$, the sample composition is affected (potentially more or less cross-talk than what the branching fractions and the lifetimes would suggest). Different event selections results in different $\epsilon(D)$'s. (The three ℓD^{*-} decay signatures have somewhat different reconstruction requirements, and, for this reason, we must allow for a different sample composition.)

---

[6] However, a detailed study shows that the dominant effect is caused by the trigger.

The efficiency $\epsilon(k\ell) \equiv \epsilon(D)$ then multiplies $\phi_{k\ell}$ for each decay chain ℓ. The sample composition for decay signature k, now becomes:

$$\text{fraction of } k \text{ from } B^0 = \frac{\tau_{B^0} \sum_{B^0\to\ell} \phi_{k\ell}\epsilon(k\ell)}{\tau_{B^0} \sum_{B^0\to\ell} \phi_{k\ell}\epsilon(k\ell) + \tau_{B^+} \sum_{B^+\to\ell} \phi_{k\ell}\epsilon(k\ell)} \tag{5.21}$$

$$\text{fraction of } k \text{ from } B^+ = \frac{\tau_{B^+} \sum_{B^+\to\ell} \phi_{k\ell}\epsilon(k\ell)}{\tau_{B^+} \sum_{B^+\to\ell} \phi_{k\ell}\epsilon(k\ell) + \tau_{B^0} \sum_{B^0\to\ell} \phi_{k\ell}\epsilon(k\ell)} \tag{5.22}$$

$\epsilon(k\ell)$ is the *absolute* reconstruction efficiency – and includes all possible reasons for not reconstructing a decay signature, some of which (e.g. the tracking mistakes, detector aging) cannot be simulated well in the Monte Carlo program. On the other hand, the *ratio* of two such efficiencies, $\epsilon(k\ell_1)/\epsilon(k\ell_2)$, almost exclusively depends on the simulation of the decay kinematics, and most other effects that are hard to simulate cancel out. We trust the Monte Carlo simulation of the decay kinematics, however we will still derive the systematic uncertainties for these efficiency ratios in section 8.2.

We divide all $\epsilon(k\ell)$ in equations (5.21) and (5.22) by the efficiency for the decay chain in which the B meson decays directly into what we are trying to reconstruct (and thus no particles, except the neutrino, are missed). We call this decay chain the *direct* decay chain. (For example, direct chains are $B^+ \to \ell^+\overline{D^0}$, $B^0 \to \ell^+ D^-$ and $B^0 \to \ell^+ D^{*-}$.) Therefore, the desired quantity is the *relative charm reconstruction efficiency*, the reconstruction efficiency relative to the direct decay chain. Using a general notation, for a decay chain ℓ that contributes to the decay signature k (with the direct chain d), we define the relative charm reconstruction efficiency as

$$\epsilon^D_{k\ell} \equiv \frac{\epsilon(k\ell)}{\epsilon(kd)} \tag{5.23}$$

For example, in $\ell^+ D^{*-}$ (table 5.4), the direct chain is [100.221], and the reconstruction must be defined for the other two decay chains:

$$\epsilon^D_{121.221} = \frac{\epsilon(121.221)}{\epsilon(100.221)}$$

$$\epsilon^D_{212.221} = \frac{\epsilon(212.221)}{\epsilon(100.221)}$$

By definition, for the direct decay chain $\ell = d$, $\epsilon^D_{kd} = 1$.

We need $\epsilon^D_{k\ell}$ for each decay chain ℓ. (Even though some of them are almost identical (e.g. [121.221] and [212.221]), it is simpler to derive them for all chains independently.) We obtain the $\epsilon_{k\ell}^{D}$'s, from the Monte Carlo simulation. Events generated using the sample composition parameters in table 5.5, are first simulated using the CDF simulation package and then reconstructed using the algorithm described in section 4.1. For each decay chain ℓ, the ratio of reconstructed and generated events measures $\epsilon(k\ell)$. Dividing $\epsilon(k\ell)$ by the reconstruction efficiency for the direct signature, $\epsilon(kd)$, we get $\epsilon_{k\ell}^{D}$. A sample calculation of $\epsilon_{k\ell}^{D}$ in one case is provided in Appendix G. The relative charm reconstruction efficiencies are shown in tables 5.6, 5.7, 5.8, 5.9 and 5.10. Some of $\epsilon_{k\ell}^{D}$ significantly deviate from unity. (Breaking down the contribution of the trigger and the D meson selection requirements, we find that most of the effect is due to the inclusive lepton trigger, and very little due to the D meson selection requirements.)

parameter	value
R_f	2.722
f^{**}	0.356
P_V	0.687
τ_{B^+}/τ_{B^0}	1.014

Table 5.5: The values of the sample composition parameters used in the Monte Carlo generation.

We now rewrite the sample composition equations (5.21) and (5.22), using $\epsilon_{k\ell}^{D}$; the fractions of the signature k coming from B^0 and B^+ respectively, are

$$\text{fraction of } k \text{ from } B^0 = \frac{\tau_{B^0}\sum_{B^0\to\ell}\phi_{k\ell}\epsilon_{k\ell}^{D}}{\tau_{B^0}\sum_{B^0\to\ell}\phi_{k\ell}\epsilon_{k\ell}^{D}+\tau_{B^+}\sum_{B^+\to\ell}\phi_{k\ell}\epsilon_{k\ell}^{D}} \qquad (5.24)$$

$$\text{fraction of } k \text{ from } B^+ = \frac{\tau_{B^+}\sum_{B^+\to\ell}\phi_{k\ell}\epsilon_{k\ell}^{D}}{\tau_{B^+}\sum_{B^+\to\ell}\phi_{k\ell}\epsilon_{k\ell}^{D}+\tau_{B^0}\sum_{B^0\to\ell}\phi_{k\ell}\epsilon_{k\ell}^{D}} \qquad (5.25)$$

As in the case of the lifetimes, we absorb the $\epsilon_{k\ell}^{D}$'s into Φ's:

$$\Phi_k^0 = \tau_{B^0}\sum_{B^0\to\ell}\phi_{k\ell}\epsilon_{k\ell}^{D} \qquad (5.26)$$

$$\Phi_k^+ = \tau_{B^+}\sum_{B^+\to\ell}\phi_{k\ell}\epsilon_{k\ell}^{D} \qquad (5.27)$$

decay chain	$\epsilon^D_{k\ell}$
100.221	1.41 ± 0.06
121.221	0.44 ± 0.07
122.001	0.87 ± 0.03
122.101	0.59 ± 0.03
122.111	0.57 ± 0.02
200.001	1.00 ± 0.03
200.101	1.55 ± 0.04
200.111	1.51 ± 0.04
211.001	0.83 ± 0.04
211.101	0.56 ± 0.04
211.111	0.59 ± 0.03
212.221	0.39 ± 0.05

Table 5.6: Relative charm reconstruction efficiencies for the decay signature $\ell^+\overline{D^0}, \overline{D^0} \to K\pi$.

decay chain	$\epsilon^D_{k\ell}$
100.002	1.00 ± 0.04
100.202	1.57 ± 0.19
100.212	1.32 ± 0.05
121.002	0.85 ± 0.06
121.202	0.78 ± 0.32
121.212	0.61 ± 0.06
212.002	0.76 ± 0.04
212.202	0.19 ± 0.11
212.212	0.54 ± 0.04

Table 5.7: Relative charm reconstruction efficiencies for the decay signature $\ell^+D^-, D^- \to K\pi\pi$.

decay chain	$\epsilon^D_{k\ell}$
100.221	1.00 ± 0.02
121.221	0.38 ± 0.02
212.221	0.37 ± 0.01

Table 5.8: Relative charm reconstruction efficiencies for the decay signature $\ell^+ D^{*-}, \overline{D^0} \to K\pi$.

decay chain	$\epsilon^D_{k\ell}$
100.221	1.00 ± 0.03
121.221	0.44 ± 0.03
212.221	0.41 ± 0.02

Table 5.9: Relative charm reconstruction efficiencies for the decay signature $\ell^+ D^{*-}, \overline{D^0} \to K3\pi$.

decay chain	$\epsilon^D_{k\ell}$
100.221	1.00 ± 0.03
121.221	0.37 ± 0.03
212.221	0.44 ± 0.03

Table 5.10: Relative charm reconstruction efficiencies for the decay signature $\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$.

while preserving the sample composition equations (5.17) and (5.18).

This completes the framework for the calculation of the sample composition. However, in section 5.1.2, the derivation of the reconstruction efficiency for the soft pion from the D^{*-} decay, $\epsilon(\pi_s)$, was deferred until all the mechanics of the sample composition calculation was in place. Now we address how $\epsilon(\pi_s)$ can be obtained from the data and the known sample composition.

5.1.5 Deriving $\epsilon(\pi_s)$

In section 4.1, we removed all $\ell\overline{D^0}$ candidates that were also ℓD^{*-} candidates in order to improve the separation between the B^0 and B^+ samples. We defined a D^{*-} candidate as a valid $\overline{D^0}$ candidate with another track, a π_s^- candidate, that made the mass diference $m(\overline{D^0}, \pi_s^-) - m(\overline{D^0})$ consistent with the world average value of $m_{D^{*-}} - m_{\overline{D^0}}$. Therefore, we removed any D^0 candidate for which there was at least one other track that made the $m(\overline{D^0}, \pi) - m(\overline{D^0})$ consistent with the world-average. Since the $m(\overline{D^0}, \pi) - m(\overline{D^0})$ distribution for real D^{*-} mesons is very narrow (~ 3 MeV/c^2), this removal is very efficient once the π_s is reconstructed. As the energy of these pions in the D^{*-} rest frame is about 40 MeV/c, they are esentially boosted by the $\beta\gamma$ of the D^* meson; usually, they are very soft (low-p_T) in the laboratory frame. Unfortunately, the detector response at these energies is poorly understood (not only at CDF, but at all other collider experiments as well), and an apriori derivation of the reconstruction efficiency, $\epsilon(\pi_s)$, is difficult.

A simple-minded approach to measuring $\epsilon(\pi_s)$ would involve a Monte Carlo generation of $B \rightarrow \ell D^{(*)} X$ events followed by the detector simulation and the event reconstruction. In the end, one would count the number of $\ell^+ D^{*-}$ events that were misclassified as $\ell^+ \overline{D^0}$ because of the missing soft pion. This method suffers from many uncertainties, especially in the detector simulation, since it is essential that all effects present in the data are properly included.

We therefore derive $\epsilon(\pi_s)$ from the data and the sample composition. We measure $N(\ell D^0, D^*$'s not removed), the number of events that would end up in the $\ell\overline{D^0}$ sample if there were no removal of D^{*-} candidates. After the removal, a fraction $(1 - \epsilon(\pi_s))$

of the original number of D^{*-} mesons still remains.

Out of $N(\ell D^0, D^*\text{'s not removed})$, we count how many events are found as possible D^{*-} candidates, $N(\ell D^*)$. We define the ratio of these two numbers:

$$R^{*(meas)} \equiv \frac{N(\ell D^*)}{N(\ell D^0, D^*\text{'s not removed})} \tag{5.28}$$

so that $R^{*(meas)}$ is the fraction of the $\ell^+ D^{*-}$ candidates that *are* actually removed. By simultaneously fitting the D^0 mass distributions for ℓD^* and ℓD^0 (without the D^* removal) samples for the respective numbers of candidates, we obtain $R^{*(meas)} = 0.249 \pm 0.008$

R^*, the fraction of $N(\ell D^0, D^*\text{'s not removed})$ that should have been removed, can also be predicted from the sample composition equations. The denominator in (5.28) is the sum of all absolute fractions $\phi_{k\ell}$ for the ℓD^0 decay signature (see Appendix F). The numerator is just a partial sum over the decay chains[7] coming through the D^{*-}, but without the $(1 - \epsilon(\pi_s))$ factor, so that:

$$\begin{aligned} R^* &\equiv \frac{1}{\Phi^0 + \Phi^+} \left[\frac{2}{3} f^{**} P_V \mathcal{B}(D^{*+} \to D^0 \pi^+) \epsilon^D_{100.221} \tau_{B^0} \right. \\ &+ f^* \mathcal{B}(D^{*+} \to D^0 \pi^+) \epsilon^D_{121.221} \tau_{B^0} \\ &+ \left. \frac{1}{3} f^{**} P_V \mathcal{B}(D^{*+} \to D^0 \pi^+) \epsilon^D_{212.221} \tau_{B^+} \right] \end{aligned} \tag{5.29}$$

If the removal were perfect (i.e. $\epsilon(\pi_s) = 1.0$), $R^{*(meas)}$ and R^* would be equal. Since that is not the case, $R^{*(meas)} = \epsilon(\pi_s) R^*$. Knowing the sample composition parameters, one can calculate R^* and in this way extract $\epsilon(\pi_s)$. Using the central values of the sample composition from table 5.2, one obtains $\epsilon(\pi_s) = 0.74 \pm 0.02$. The error of 0.02, however, does not include the much larger systematic uncertainty due to the sample composition. The way of properly handling the correlation between $\epsilon(\pi_s)$ and the sample composition parameters will be addressed in section 7.2.2.

5.1.6 Summary of the sample composition

In order to determine the fraction of the B^0 and B^+ decays in each of the five decay signatures, we get the sample composition parameters R_f, f^{**}, and τ_{B^+}/τ_{B^0} from

---

[7] $\phi_{100.221}$, $\phi_{121.221}$ and $\phi_{212.221}$

other experiments, obtain $\epsilon^D_{k\ell}$ from the Monte Carlo simulation, and derive $\epsilon(\pi_s)$ from $R^{*(meas)}$ (measured in the data) and R^* (calculated from the sample composition parameters and $\epsilon^D_{k\ell}$), as well as P_V from quantities that will be described in sec. 7.2.3. We then calculate the contributions of the B^0 and B^+ decays to the decay signature k,

$$\begin{aligned}\Phi^0_k &= \tau_{B^0} \sum_{B^0 \to \ell} \phi_{k\ell}\epsilon^D_{k\ell} \\ \Phi^+_k &= \tau_{B^+} \sum_{B^+ \to \ell} \phi_{k\ell}\epsilon^D_{k\ell}\end{aligned}$$

and use them to calculate the sample composition, i.e. the fractions of k coming from the decays of the B^0 and B^+ mesons:

$$\begin{aligned}\text{fraction of } k \text{ from } B^0 &= \frac{\Phi^0_k}{\Phi^0_k + \Phi^+_k} \\ \text{fraction of } k \text{ from } B^+ &= \frac{\Phi^+_k}{\Phi^0_k + \Phi^+_k}\end{aligned}$$

For any analysis that does not significantly depend on the proper time of the B decay, this level of knowledge of the composition of the $B \to \nu\ell D^{(*)}X$ sample is sufficient. However, measuring $B^0\bar{B}^0$ mixing and the B^+/B^0 lifetime ratio requires the separation of the B^0 and B^+ meson samples as a function of the proper time of the B decay. This results in a dependence of the sample composition on the measured proper decay time, as will be described in section 5.2.2.

5.2 Sample composition effects in the measurement of the proper time

This section considers effects that become important as one looks into the sample composition in light of the measurement of ct. On one hand, the amount of the cross-talk in each bin of the ct distribution is a function of the value of ct for that bin, while on the other, the corrected proper time for the B^0 and B^+ components of the ct-bin depends on the sample composition.

We first examine the simplest dependence of the sample composition on the proper time: when the B^0 and B^+ lifetimes are different. Next, we analyze the sample composition in a ct bin, and show that the B^0 and B^+ components do not have the same proper time. Hence the sample composition depends on ct even if the lifetimes are identical. We then present an approximate way of feeding this information back into the sample composition, and obtain our final expressions for the sample composition.

5.2.1 Sample composition dependence on ct when $\tau_{B^+} \neq \tau_{B0}$

In the section 5.1.2 we have already described the effect of the different B meson lifetimes on the B meson semileptonic width (equation (5.11)). The effect was global, and equally affected all ct bins. Now we consider how the sample composition depends on the proper time when the lifetimes are different:

- The rates of $B^+ \to \nu\ell D^{(*)}$ and $B^0 \to \nu\ell D^{(*)}$ decays are proportional to the B^+ and B^0 lifetimes, according to equation (5.11)).
- However, even if the rates of $B^+ \to \nu\ell D^{(*)}$ and $B^0 \to \nu\ell D^{(*)}$ decays were the same, if $\tau_{B^+} \neq \tau_{B^0}$, the sample composition would depend on ct, merely because one component would decay sooner than the other. At the proper time ct, the abundance of each component would be scaled down by $e^{-ct/c\tau_B}/\tau_B$: all $\phi_{k\ell}$ corresponding to the decay chains originating in B^0 would be multiplied by $e^{-ct/c\tau_{B^0}}/\tau_{B^0}$, and, analogously, all $\phi_{k\ell}$ corresponding to the decay chains originating in B^0 would be multiplied by $e^{-ct/c\tau_{B^+}}/\tau_{B^+}$.

When both effects are combined, the $\phi_{k\ell}$'s for the B^0 terms must be multiplied by $e^{-ct/c\tau_{B^0}}$, and for the B^+ terms by $e^{-ct/c\tau_{B^+}}$. Consequently, we modify the Φ's (equations (5.26) and (5.27)) to include the variation of ct:

$$\Phi_k^0(ct) = e^{-ct/c\tau_{B^0}} \sum_{B^0 \to \ell} \phi_{k\ell}\epsilon_{k\ell}^D \tag{5.30}$$

$$\Phi_k^+(ct) = e^{-ct/c\tau_{B^+}} \sum_{B^+ \to \ell} \phi_{k\ell}\epsilon_{k\ell}^D \tag{5.31}$$

The sample composition equations (5.17) and (5.18) remain unchanged, except that the fractions of the decay signature k originating from B^0 and B^+ now also depend on the proper time:

$$\text{fraction of } k \text{ from } B^0 = \frac{\Phi_k^0(ct)}{\Phi_k^0(ct) + \Phi_k^+(ct)} \tag{5.32}$$

$$\text{fraction of } k \text{ from } B^+ = \frac{\Phi_k^+(ct)}{\Phi_k^0(ct) + \Phi_k^+(ct)} \tag{5.33}$$

We note that equations (5.26) and (5.27) remain valid wherever the explicit dependence on the proper time is unnecessary (e.g. in derivation of $\epsilon(\pi_s)$).

5.2.2 Correcting the ct scale

In section 4.3.1, we approximated the true decay length, ct_{true}, by (eq. (4.8)):

$$ct_{true} \approx ct \equiv L_{\mathrm{xy}}^B \frac{m_B}{p_T^{\ell D}} \mathcal{K}_{kd}$$

where the average correction factor $\mathcal{K}_{kd}$ was defined as

$$\mathcal{K}_{kd} \equiv \left\langle \frac{p_T^{\ell D}}{p_T^{MC}} \right\rangle$$

The subscripts kd in $\mathcal{K}_{kd}$ stress that it was derived *for the direct decay chain only.* As defined in eq. (4.8), the measured ct does not include any of the sample composition effects. If the $\beta\gamma$ correction is to take into account the sample composition, it is sufficient to add another scale factor multiplying the measured proper time ct,

$$ct_{true} \approx L_{\mathrm{xy}}^B \frac{m_B}{p_T^{\ell D}} \mathcal{K}_{kd} \frac{\bar{\mathcal{K}}_k}{\mathcal{K}_{kd}} = ct \, \frac{\bar{\mathcal{K}}_k}{\mathcal{K}_{kd}} \tag{5.34}$$

Here, $\bar{\mathcal{K}}_k$ is a function of the proper time itself, as well as of the sample composition parameters,

$$\bar{\mathcal{K}}_k = \bar{\mathcal{K}}_k(ct, R_f, f^{**}, P_V, \tau_{B^+}/\tau_{B^0}, \ldots)$$

There is a different $\mathcal{K}$-factor for each B decay chain: the kinematics of various decay chains within the same decay signature are different (analogously to $\epsilon_{k\ell}^D$). We label the $\mathcal{K}$-factor for the decay chain ℓ feeding into the decay signature k by $\mathcal{K}_{k\ell}$. The default $\beta\gamma$ correction (that is, the default $\mathcal{K}$-factor) must be a *weighted average* of corrections for all decay chains, the weight of each decay chain given by its contribution (governed by the sample composition) to the decay signature k:

$$\bar{\mathcal{K}}_k(ct) = \sum_\ell \frac{e^{-ct/c\tau} \phi_{k\ell} \epsilon_{k\ell}^D}{\Phi_k^0(ct) + \Phi_k^+(ct)} \mathcal{K}_{k\ell} \tag{5.35}$$

Equation (5.35) gives us the average $\bar{\mathcal{K}}_k$ factor for the decay signature k, given the specific sample composition. $\Phi_k^0(ct) + \Phi_k^+(ct)$ (eq. (5.30) and (5.31)), as the total contribution of all decay chains, is a normalization factor, so that $e^{-ct/c\tau}\phi_{k\ell}\epsilon_{k\ell}^D/(\Phi_k^0(ct) + \Phi_k^+(ct))$ is a properly normalized weight. By τ we here mean either τ_{B0} or τ_{B+}, depending on whether the decay chain ℓ originated from the B^0 or B^+ meson.

Since no additional decay particles are lost in the direct decay chain ($\ell = d$), one would expect that $\mathcal{K}_{kd}$ is closer to unity than the $\mathcal{K}$-factor of any other decay chain, $\mathcal{K}_{k\ell}$. ($\mathcal{K} = 1.0$ would mean that nothing was lost, and $p_T^{\ell D} = p_T^B$.) Therefore, $\mathcal{K}_{k\ell} < \mathcal{K}_{kd}$ for a decay chain ℓ ($\ell \neq d$), and consequently $\bar{\mathcal{K}}_k < \mathcal{K}_{kd}$. As a result, the ct of the B meson decay is multiplied by a scale factor $\bar{\mathcal{K}}_k/\mathcal{K}_{kd} < 1$.

Our goal is to deduce the amount of the B^0 and B^+ contribution to each ct bin. Since we want to separate the B^0 and B^+ components, there is no need to calculate an overall $\mathcal{K}$-factor. Instead, we treat them separately; the decay chains proceeding via a D^{**} decay dominate the cross-talk,[8] and thus require a different $\mathcal{K}$ correction from the direct decays. We therefore divide the decay chains into those originating from the B^0 decay and those originating from the B^+ decay, and introduce scale factors for ct ($\bar{\mathcal{K}}_k/\mathcal{K}_{kd}$) calculated *independently* for the B^0 and B^+ mesons.[9] We define two ct-scale factors for *each* decay signature k, $\bar{\mathcal{K}}_k^0(ct)/\mathcal{K}_{kd}$ for the B^0 mesons, and $\bar{\mathcal{K}}_k^+(ct)/\mathcal{K}_{kd}$ for the B^+ mesons, and calculate them by summing over decay chains ℓ coming from either the B^0 or B^+ decays:

$$ct_k^0 \equiv ct\frac{\bar{\mathcal{K}}_k^0(ct)}{\mathcal{K}_{kd}} = ct\sum_{B^0\to\ell}\frac{e^{-ct/c\tau_{B^0}}\phi_{k\ell}\epsilon_{k\ell}^D}{\Phi_k^0(ct)}\frac{\mathcal{K}_{k\ell}}{\mathcal{K}_{kd}} \tag{5.36}$$

$$ct_k^+ \equiv ct\frac{\bar{\mathcal{K}}_k^+(ct)}{\mathcal{K}_{kd}} = ct\sum_{B^+\to\ell}\frac{e^{-ct/c\tau_{B^+}}\phi_{k\ell}\epsilon_{k\ell}^D}{\Phi_k^+(ct)}\frac{\mathcal{K}_{k\ell}}{\mathcal{K}_{kd}} \tag{5.37}$$

Here, by ct_k^0 and ct_k^+ we denote the proper times for the B^0 and the B^+ decays. Due to the separation of the B^0 and B^+ factors, the normalization factors are now different: $\Phi_k^0(ct)$ for B^0 and $\Phi_k^+(ct)$ for B^+. Note that the weights $e^{-ct/c\tau_{B^0}}\phi_{k\ell}\epsilon_{k\ell}^D/\Phi_k^0(ct)$ do not depend on the proper time, as the factor $e^{-ct/c\tau_{B^0}}$ is also part of $\Phi_k^0(ct)$ (eqs. (5.30)

---

[8] They account for *all* cross-talk decay chains for the four B^0 decay signatures.

[9] In the case of mixing, only the B^0 mesons mix, so, when the oscillation is calculated, it would be incorrect to use the ct scale factor averaged over all B meson decays.

and (5.31)) and cancels in the ratio. The scale factors for individual decay chains, $\mathcal{K}_{k\ell}/\mathcal{K}_{kd}$, are obtained from the Monte Carlo simulation, and are shown in tables 5.11, 5.12, 5.13, 5.14 and 5.15.

As an example, let us examine a bin from an arbitrary distribution binned in ct, in a decay signature k dominated by B^0 (so the B^+ decays are the cross-talk). Here, we are concerned only with the sample composition of this bin. The fraction of the B^0 decays in this bin is $\Phi_k^0(ct)/(\Phi_k^0(ct)+\Phi_k^+(ct))$. Equations (5.36) and (5.37) give the average proper times of the B^0 and B^+ decays in this bin. In this case, $ct^+ < ct^0$, and both are less than the measured ct (for which the 'direct' $\mathcal{K}$-factor, $\mathcal{K}_{kd}$, was used). Usually, the ct of the dominant component (here ct^0) is close to the measured ct, while the ct of the cross-talk (here ct^+) is $5-10\%$ lower. We note that $ct^0(ct^+)$ is the average *true* proper time for the decay chains originating from the B^0 (B^+) meson – the decays that were reconstructed in the bin with the measured ct, on average had the true proper time of ct^0 (or ct^+). As a consequence, we must use ct^0 and ct^+ in all calculations requiring the true proper time, such as the probability for mixing (ct^0 only) and the sample composition.

In summary, due to the cross-talk, each decay signature is split into two parts: one coming from the B^0 decays, the other from the B^+ decays. Equations (5.36) and (5.37) give us the proper times corresponding to each part. (For each of the five decay signatures, there are two proper-time axes, ct_k^0 and ct_k^+, so there are ten in total.) The scale factors ct_k^0/ct and ct_k^0/ct do not depend on the measured proper time ct, only on the sample composition itself.

5.2.3 Feeding ct^+ and ct^0 back into the sample composition

We have seen in section 5.2.1 that $\tau_{B^+} \neq \tau_{B^0}$ causes the ct-dependence of the sample composition, since one B meson component decays faster than the other one. However, from equations (5.36) and (5.37) we see that $ct_k^0 \neq ct_k^+$ even when the lifetimes are equal.

In order to incorporate the effect of $ct_k^0 \neq ct_k^+$, we calculate $\Phi_k^0(ct_k^0)$ and $\Phi_k^+(ct_k^+)$ (eqs. (5.30) and (5.31)) by substituting ct with ct_k^0 and ct_k^+ respectively. Finally,

decay chain	$\mathcal{K}_{k\ell}/\mathcal{K}_{kd}$	$\sigma_{k\ell}/\sigma_{kd}$
100.221	1.054 ± 0.005	0.88
121.221	0.930 ± 0.020	1.12
122.001	0.954 ± 0.005	1.02
122.101	0.929 ± 0.006	0.92
122.111	0.928 ± 0.005	0.89
200.001	1.000 ± 0.004	1.00
200.101	1.026 ± 0.004	0.94
200.111	1.027 ± 0.003	0.95
211.001	0.942 ± 0.007	0.95
211.101	0.923 ± 0.009	1.08
211.111	0.922 ± 0.007	1.03
212.221	0.954 ± 0.014	1.01

Table 5.11: $\beta\gamma$ correction and `RMS` scale factors for the decay signature $\ell^+\overline{D^0}, \overline{D^0} \to K\pi$.

decay chain	$\mathcal{K}_{k\ell}/\mathcal{K}_{kd}$	$\sigma_{k\ell}/\sigma_{kd}$
100.002	1.000 ± 0.005	1.00
100.202	1.014 ± 0.015	0.82
100.212	1.017 ± 0.005	1.01
121.002	0.955 ± 0.008	1.18
121.202	0.938 ± 0.040	0.43
121.212	0.929 ± 0.011	0.84
212.002	0.949 ± 0.007	1.33
212.202	0.971 ± 0.033	2.44
212.212	0.928 ± 0.008	1.18

Table 5.12: $\beta\gamma$ correction and `RMS` scale factors for the decay signature $\ell^+D^-, D^- \to K\pi\pi$.

decay chain	$\mathcal{K}_{k\ell}/\mathcal{K}_{kd}$	$\sigma_{k\ell}/\sigma_{kd}$
100.221	1.000 ± 0.002	1.00
121.221	0.904 ± 0.006	0.96
212.221	0.910 ± 0.004	0.96

Table 5.13: $\beta\gamma$ correction and `RMS` scale factors for the decay signature $\ell^+ D^{*-}, \overline{D^0} \to K\pi$.

decay chain	$\mathcal{K}_{k\ell}/\mathcal{K}_{kd}$	$\sigma_{k\ell}/\sigma_{kd}$
100.221	1.000 ± 0.003	1.00
121.221	0.910 ± 0.008	1.11
212.221	0.904 ± 0.006	0.93

Table 5.14: $\beta\gamma$ correction and `RMS` scale factors for the decay signature $\ell^+ D^{*-}, \overline{D^0} \to K3\pi$.

decay chain	$\mathcal{K}_{k\ell}/\mathcal{K}_{kd}$	$\sigma_{k\ell}/\sigma_{kd}$
100.221	1.000 ± 0.004	1.00
121.221	0.918 ± 0.011	1.20
212.221	0.889 ± 0.007	1.07

Table 5.15: $\beta\gamma$ correction and `RMS` scale factors for the decay signature $\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$.

we substitute the new values of Φ_k^0 and Φ_k^+ into the sample composition equations (5.32) and (5.33), and obtain the final dependence on the fractions of the B^0 and B^+ components on the measured ct:

$$\text{fraction of } k \text{ from } B^0 = \frac{\Phi_k^0(ct_k^0)}{\Phi_k^0(ct_k^0) + \Phi_k^+(ct_k^+)} \tag{5.38}$$

$$\text{fraction of } k \text{ from } B^+ = \frac{\Phi_k^+(ct_k^+)}{\Phi_k^0(ct_k^0) + \Phi_k^+(ct_k^+)} \tag{5.39}$$

5.2.4 Scaling the ct-resolution

The resolution on ct, σ_{ct}, is also a function of the sample composition.[10] In section 4.3, we have derived the dependence of the ct-resolution for the direct decay chain, σ_{kd}, on ct. In this section, our goal is to determine a method for scaling of the $\sigma_{kd}(ct)$ analogous to the scaling of the measured proper time, ct, by the scale factor $\bar{\mathcal{K}}_k/\mathcal{K}_{kd}$ that depends on the sample composition.

We note that, in section 4.3.1, we approximated the true proper decay length, ct_{true}, by the 'measured' proper decay length ct(eq. (4.8)):

$$ct_{true} \approx ct \equiv L_{\text{xy}}^B \frac{m_B}{p_T^{\ell D}} \mathcal{K}_{kd}$$

and derived $\sigma_{ct}(ct)$ as a width of the distribution

$$\Delta ct \equiv ct - ct_{true}$$

binned in ct_{true}. The subscript kd underscores that both ct and σ have been derived for the 'direct' decay chain.

In the decay signature k, the $\ell D^{(*)}$ candidates that are part of a 'non-direct' decay chain ℓ ($\ell \neq d$) contribute to the Δct distribution in two ways:

- **The Δct is wider ($\sigma_{k\ell} > \sigma_{kd}$).** E.g. the $\beta\gamma$ resolution for $B \to D^{**} \to D^*$ decay chain is poorer than for the direct $B \to D^*$ chain, not only because a decay product (the π_{**}) is lost, but also because the decays $B \to \nu\ell D^*$ and $B \to \nu\ell D^{**}$ are kinematically different (as discussed in sections 5.1.4 and 5.2.2).

---

[10] Sometimes, for brevity, we use σ instead of σ_{ct}.

- **The Δct distribution from the decay chain ℓ is offset from zero due to using an incorrect $\mathcal{K}$-factor.** The correction $\mathcal{K}_{k\ell}$ is separately determined for each decay chain ℓ feeding into the decay signature k. If the $\mathcal{K}$-factor $\mathcal{K}_{k\ell}$ were used for $\ell D^{(*)}$ candidates that come from this decay chain, the distribution of $\Delta ct_{k\ell} \equiv ct - ct_{true}$ would be a gaussian centered at zero of the width $\sigma_{k\ell}$. However, we use the average $\mathcal{K}$-factor for the B^0 (or B^+) component, $\bar{\mathcal{K}}_k^0$ ($\bar{\mathcal{K}}_k^+$). When this average $\mathcal{K}$-factor is applied to the decays from the decay chain ℓ, the measured ct is systematically over- (or under-) corrected. Therefore, the distribution of $\Delta ct_{k\ell}$ is a gaussian of the width $\sigma_{k\ell}$ offset from zero by an offset $\Delta_{k\ell}$.

The dependence $\sigma_{ct}(ct)$ is determined for the 'direct' decay chain, $\sigma_{kd}(ct)$. We also determine the `RMS` of the overall σ_{kd}. The total distribution Δct is a sum of off-center gaussians for each decay chain ℓ. We calculate the `RMS` of the sum of the off-gaussian distributions for the decay signature k, $\overline{RMS}_k$, divide it by the `RMS` of of the overall Δct_{kd} distribution for the direct mode, RMS_{kd}, and use this ratio as a scale factor multiplying $\sigma_{kd}(ct)$:

$$\sigma_{ct}(ct) = \frac{\overline{RMS}_k}{RMS_{kd}} \sigma_{kd}(ct) \tag{5.40}$$

The exact expression for $\sigma_{ct}(ct)$ is somewhat cumbersome, and is not given here since it does not give us any new insight. For completeness, it is provided in Appendix H.

5.2.5 The effect of the sample composition

The table 5.16 summarizes the ratio of $\bar{\mathcal{K}}_k^{(+,0)}$ (the $\mathcal{K}$-factor corresponding to the final sample composition for either B^0 or B^+), and $\mathcal{K}_{kd}$ (the $\mathcal{K}$-factor of the direct channel), calculated for the default sample composition (section 5.1.2, table 5.2).

In the case of the decay signature $\ell^+\overline{D^0}, \overline{D^0} \to K\pi$, the fact that $\bar{\mathcal{K}}_k^0/\mathcal{K}_d > 1.0$ is a consequence of the difference in the semileptonic decays $B \to D$ and $B \to D^*$. The latter emits a less energetic neutrino (which is why it is favored by the trigger), so, somewhat surprisingly, the $\beta\gamma$ correction estimate is actually better for decays going through the D^*, even though a particle (the soft pion) is lost.

Decay signature (k)	$\bar{\mathcal{K}}_k^0/\mathcal{K}_{kd}$	$\bar{\mathcal{K}}_k^+/\mathcal{K}_{kd}$
$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$	1.021	1.012
$\ell^+D^-, D^- \to K\pi\pi$	1.002	0.942
$\ell^+D^{*-}, \overline{D^0} \to K\pi$	0.994	0.910
$\ell^+D^{*-}, \overline{D^0} \to K3\pi$	0.994	0.904
$\ell^+D^{*-}, \overline{D^0} \to K\pi\pi^0$	0.995	0.889

Table 5.16: The table of the ratios of $\bar{\mathcal{K}}_k^{(+,0)}$ (the $\mathcal{K}$-factor corresponding to the final sample composition for either B^0 or B^+), and $\mathcal{K}_{kd}$ (the $\mathcal{K}$-factor of the direct channel).

This concludes the discussion of the sample composition: at this point, we know how to express all sample composition effects in terms of the sample composition parameters (f^{**}, R_f, P_V, τ_{B^+}/τ_{B^0} and $\epsilon(\pi_s)$): we can calculate the fraction of the decay signature k that originated in the B^0 decays (eq. (5.32)) and in the B^+ decays (eq. (5.33)), as well as the average true proper time for the B^0 and B^+ components (ct_k^0 and ct_k^+, equations (5.36) and (5.37)).

In the following section we turn to the description of Same Side Tagging, and its application to the $B \to \nu\ell D^{(*)}$ decays. We measure the dependence of the asymmetry on the proper time (equation (2.15)). However, since Same Side Tagging behaves differently for B^0 and for B^+ mesons, one must know the *sample composition* in order to uncover how the true asymmetries depend on the proper time.

Chapter 6

$\pi - B$ meson Charge-flavor Correlations

6.1 Same-Side Tagging

The Same Side Tagging (SST) technique was first proposed by Gronau, Nippe and Rosner [35]. The basic idea behind the SST is that the flavor of a B meson at production time can be inferred from the charged particles produced along with it. To date, various techniques have already been used to determine the flavor of this second hadron: examples are lepton-tagging or jet charge-tagging [43]. We refer to such methods, which employ the "other" b-flavored hadron in the event, as *opposite side tagging* (OST) methods. Monte Carlo simulations indicate that in a hadron collider detector with central rapidity coverage such as CDF, once one B meson is produced in the central rapidity region, the second b-flavored hadron is also produced in the central region of the detector only $\sim 40\%$ of the time. For lepton tagging, there is an additional loss of efficiency arising from the branching ratio $b \rightarrow \ell$. For jet-charge tagging, the purity of the flavor-tag decision is reduced by the presence of charged tracks from the proton-antiproton remnants. Finally, flavor tagging based on OST suffers from the inevitable dilution arising from mixing of the second b-flavored hadron.

In contrast, the SST ignores the second b-flavored hadron and, instead, considers correlations of charged particles produced along with the B meson of interest. These correlations arise from either the tracks originating from the fragmentation of a b quark into a B meson or from decays of higher B resonances such as B^{**} [49, 48]. Figure 6-1 displays possible fragmentation paths for a $\bar{b}$ quark, assuming a naive view of string fragmentation. If the $\bar{b}$ quark combines with a u quark to form a B^+ meson, then the remaining $\bar{u}$ quark may combine with a d quark to form a π^-. Similarly, if the $\bar{b}$ quark fragments to form a B^0 meson, the correlated pion would be a π^+.[1] Another possible source of correlated pions are B^{**} decays like $B^{**0} \to B^{(*)+}\pi^-$ or $B^{**+} \to B^{(*)0}\pi^+$. The correlations here are the same as for pions produced in fragmentation. In this analysis no attempt is made to differentiate the sources of correlated pions, and only the overall effect of the charge correlation between the B meson and the pion is measured.

In this simple picture of fragmentation, we expect charged B mesons to display a higher degree of correlation with charged particles than neutral B mesons, based on the production of strange quarks in the fragmentation process. The resulting strange particle would be a K^- for a B^+ and a $\bar{K}^0$ for a B^0. While the K^- exhibits the correct charge correlation and can be selected as a tag, the $\bar{K}^0$ either escapes undetected or decays into two oppositely charged pions that are no longer kinematically correlated with the B meson (and therefore are equally likely to be selected as tags). As a result of these considerations, the dilution for B^+ is expected to be higher than the dilution for B^0.

6.1.1 The SST algorithm

In constructing the SST algorithm we start from the candidate track selection. The pions from the b-quark fragmentation or B^{**} decay are expected to be fairly collimated with the direction of the B meson. Therefore, we require that the track is within a $\eta - \phi$ cone of $\Delta R = 0.7$ around the direction of the B meson, approximated by the

---

[1] A complication with fragmentation tagging is that many times the first particle (neutral or charged) will be a ρ, decaying into two pions behaving differently from the pions that originated in the fragmentation.

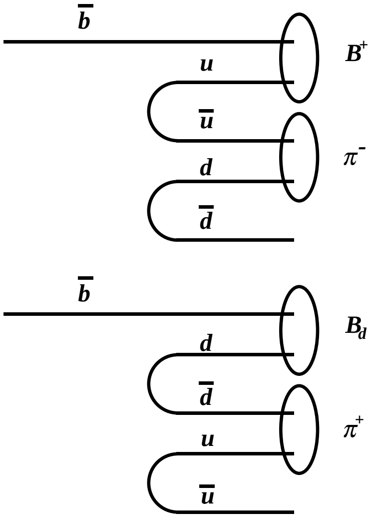


Figure 6-1: Possible fragmentation paths for a $\bar{b}$ quark, assuming a naive view of string fragmentation.

direction of the $\ell D^{(*)}$ system. It has been known in CDF that most of the B-meson fragmentation products for the B's collected by the inclusive semileptonic trigger are contained[2] within $\Delta R \sim 0.6$, so using 0.7 is appropriate.

Moreover, the tagging candidate tracks originate from the primary vertex of the $p\bar{p}$ interaction that produced the $b\bar{b}$ pair.Therefore, we require that the z of the point of the closest approach to the primary vertex of the track is within 5 cm of the primary vertex. In order to constrain the candidate tracks in the transverse $(x - y)$ plane, we require that the impact parameter significance, d_0/σ_0 be less than 3. This requirement is also crucial in rejecting π_{**} from D^{**}, as will be explained in section 7.2.3. Because of the requirement $d_0/\sigma_0 < 3$, the candidate tracks must satisfy the standard CDF SVX quality requirements, since the impact parameter, d_0, is measured precisely

---

[2] This number was derived by comparing the $\Delta\phi$ distribution of tracks around a B meson, with the $\Delta\phi$ distribution of tracks around the lepton coming from a W. The latter distribution is constant in ϕ, since there is no contribution from the b-quark fragmentation, and all tracks are part of the underlying event. In case of the B mesons, however, the ϕ distribution peaks at 0 due to the fragmentation, and then falls down to the constant level at about 0.6.

by the SVX (chapter 3.2.3) . Finally, we also require the basic CDF CTC quality requirements: the track must have a three-dimensional fit, $p_T > 400$ MeV/c and $R_{exit} > 130$ cm. The requirements on p_T and R_{exit} also reduce the charge asymmetry inherent in the CDF tracking system. The charge asymmetry is defined as

$$\mathcal{A}_{CP} \equiv \frac{N_+ - N_-}{N_+ + N_-} \tag{6.1}$$

where N_+ and N_- are numbers of positive and negative tracks in a given sample, and should not be confused with the asymmetry between the right and wrong correlations, eq. (2.15). The sense wires in each of the superlayers are tilted by 45° with respect to the radial direction, in order to compensate for the Lorentz drifting angle (see figure 3-5 in section 3.2.3). That, however, results in an unequal reconstruction efficiency for the low-p_T (i.e. high-curvature) positive and negative tracks: the positive tracks are moving parallel to the sense wires and are more likely to leave hits in the chamber. Previous studies at CDF show that the bulk of the tracking charge asymmetry is bellow 400 MeV/c. In addition to $p_T > 400$ MeV/c, the requirement that the CTC 'exit radius', R_{exit} (defined as the distance from the beam at which the track leaves the volume of the CTC), is equal to the outer CTC radius, guarantees that the track passes through all nine superlayers, and further reduces the tracking charge asymmetry. In the sample of tracks that pass these two requirements, the remaining charge asymmetry is at a level of a few percent.

With the above selection requirements, there are, on average, about 2.2 candidate tracks per B candidate. Obviously, for an event to be tagged there must be at least one track that passes the above cuts, so the ratio of the number of such events and the total number of events defines the tagging efficiency. The tagging efficiencies, as measured in the five decay signatures, are shown in figures 6-2 and 6-3.

String fragmentation models [51] indicate that particles produced in the b-quark hadronization chain have low momenta transverse to the direction of the B meson momentum. We thus select as the tag the track that has the minimum component of momentum, p_t^{rel}, orthogonal to the momentum sum of the track and the $\ell^+ D^{(*)}$ system. (In addition to 'minimum p_t^{rel}', seven other tagging algorithms have been

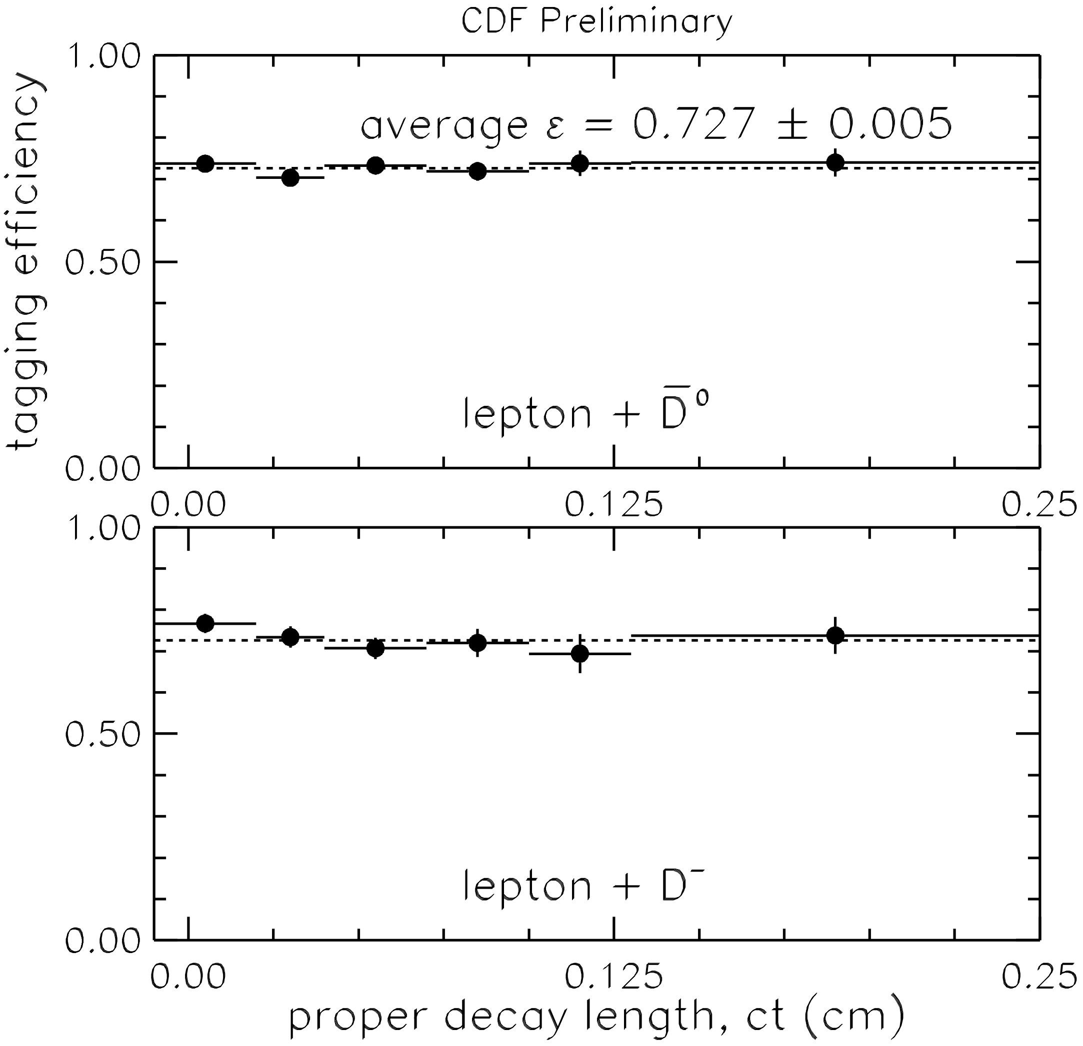


Figure 6-2: The tagging efficiencies versus the measured proper time, ct, for the decay signatures "$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$" and "$\ell^+D^-, D^- \to K\pi\pi$".

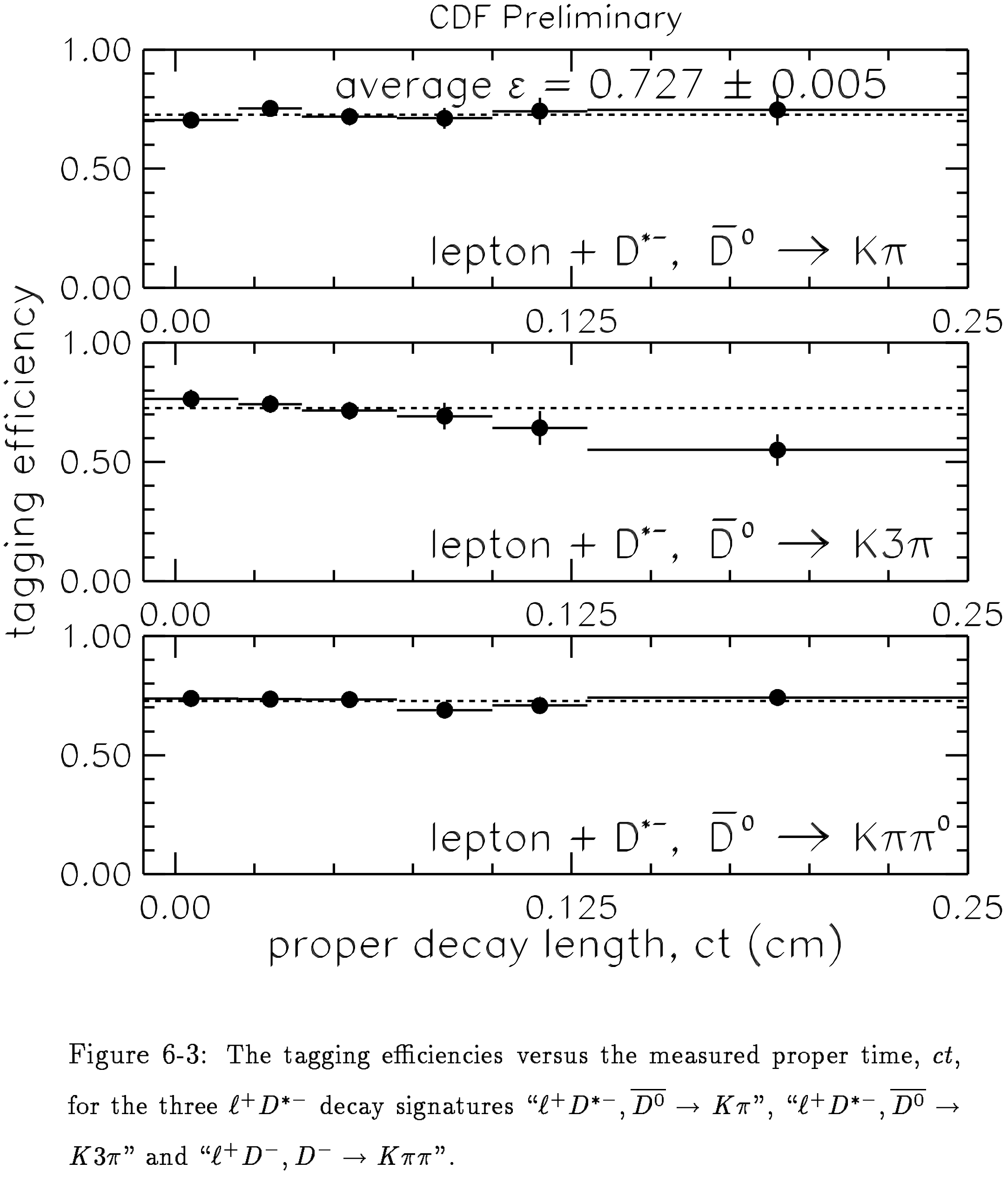


Figure 6-3: The tagging efficiencies versus the measured proper time, *ct*, for the three $\ell^+ D^{*-}$ decay signatures "$\ell^+ D^{*-}, \overline{D^0} \to K\pi$", "$\ell^+ D^{*-}, \overline{D^0} \to K3\pi$" and "$\ell^+ D^-, D^- \to K\pi\pi$".

considered,[3] however the 'minimum p_t^{rel}' has been found to be the best.) The charge of the chosen track determines the flavor of the B, and we compare it to the charge of the lepton from the B decay to determine the right/wrong correlation for the given $\ell^+ D^{(*)}$ candidate. The distributions of the p_T and p_t^{rel} for all SST candidate tracks and selected tags, for all five decay signatures, are shown in figures 6-4 6-6, 6-8, 6-10, and 6-12. The B candidates are divided according to the "right" or "wrong" charge correlation between the lepton and the tagging pion. The distributions of the total number of SST candidate tracks and the number of SST candidate tracks as a function of the proper time are shown in figures 6-5 6-7, 6-9, 6-11, and 6-13. We also show the "right-wrong"asymmetry (in the sense of eq. (2.15)) as a function of the total number of SST tracks and as a function of the proper time.

The distributions of the same quantities, however divided according to the charge of the SST candidate track or the tag ("+" or "−") are given in Appendix I. They indicate that the remaining charge asymmetry (eq. (6.1)) is rather small and that it does not introduce a significant bias into the tagging.

6.1.2 On the *ct* of the asymmetry points

For each decay signature, we divide events into 6 bins in corrected proper decay time, *ct*, and measure the asymmetry in each bin. But before we apply the Same Side Tagging to each bin in proper time, we must determine where on the *ct* axis to put the asymmetry points. Since our goal is to extract the oscillation frequency from the asymmetry distributions, whether each asymmetry point is in the center of the bin or not actually does matter.

We use the *data average ct* of the bin. Let us denote the underlying distribution of

---

[3] They are: 'maximum p_T', 'maximum p_l^{rel}' (longitudinal momentum of the track with respect to the combination of the track and the $\ell^+ D^{(*)}$ system), 'minimum Q' ($Q = m(\ell^+ D^{(*)}, \pi) - m(\ell^+ D^{(*)}) - m_\pi$), 'minimum $\Delta R(\ell^+ D^{(*)}, \pi)$' 'minimum $\cos\theta^*$' and 'maximum $\cos\theta^*$' (where $\cos\theta^*$ is the angle between the track in the center of mass frame and the direction of motion of the $(\ell^+ D^{(*)}, \pi)$ system in the laboratory frame). Since the methods are correlated, it was sufficient to study only three, 'minimum p_t^{rel}', 'maximum p_l^{rel}' and 'minimum ΔR.

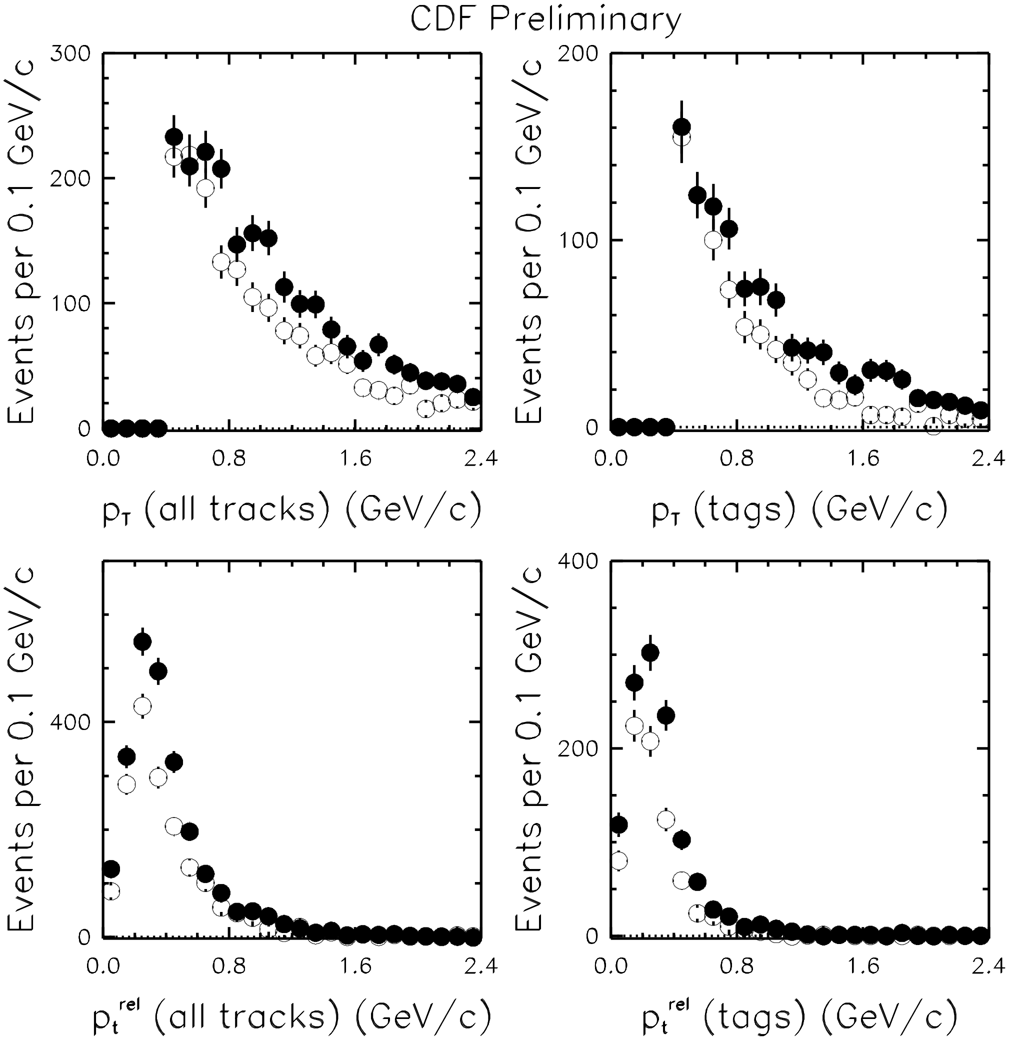


Figure 6-4: Tagging distributions for decay signature “$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$”. Solid points are “right sign” tags, open are “wrong sign” tags Upper left: p_T of the SST candidate tracks. Upper right: p_T of the tag. Lower left: p_t^{rel} (the value of the tagging function) of the SST candidate tracks. Lower right: p_t^{rel} of the tag.

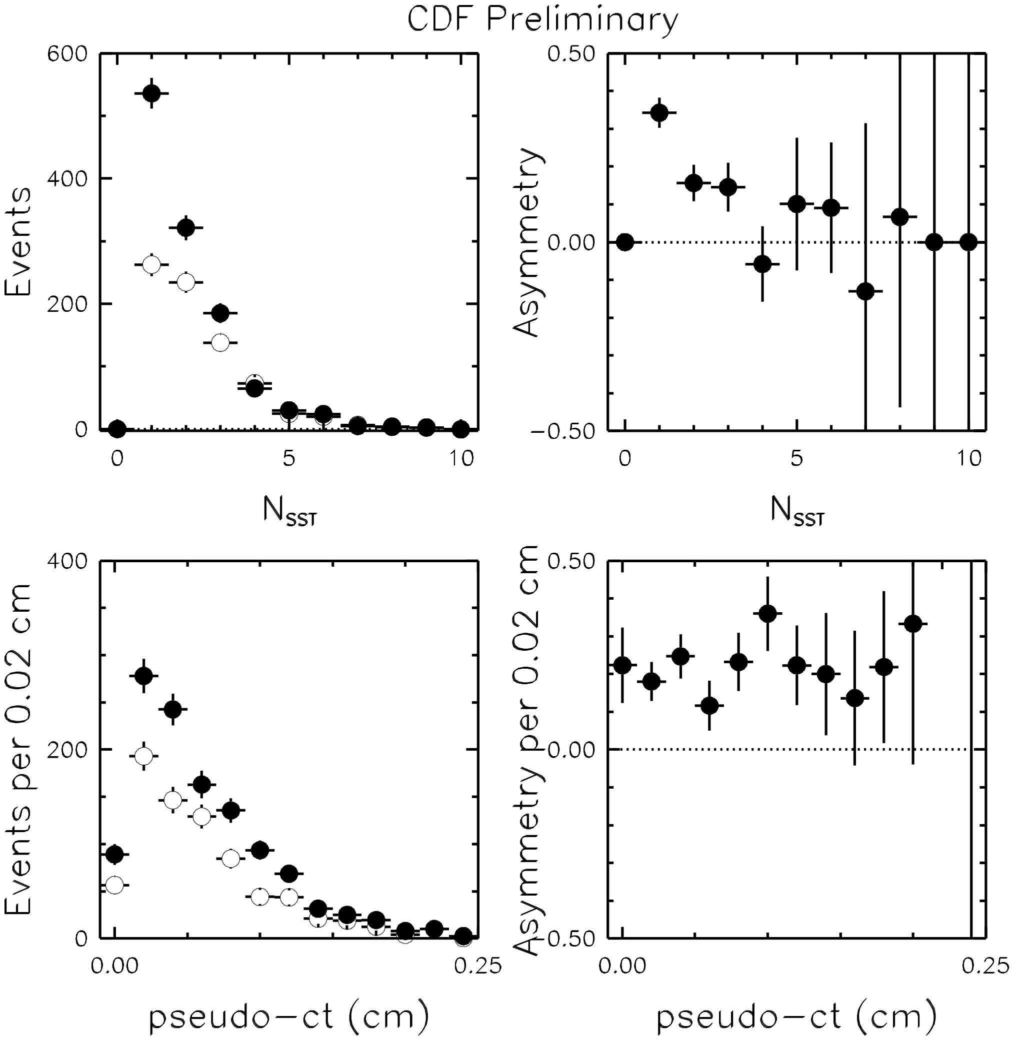


Figure 6-5: Tagging distributions for decay signature "$\ell^+\overline{D^0}, \overline{D^0} \rightarrow K\pi$". Solid points are "right sign" tags, open are "wrong sign" tags. Upper left: the total number of SST candidate tracks. Upper right: the asymmetry vs the total number of the SST candidate tracks. Lower left: the measured proper decay time (ct) corrected using the $\mathcal{K}$-factor of the 'direct' decay chain. Lower right: the asymmetry vs ct.

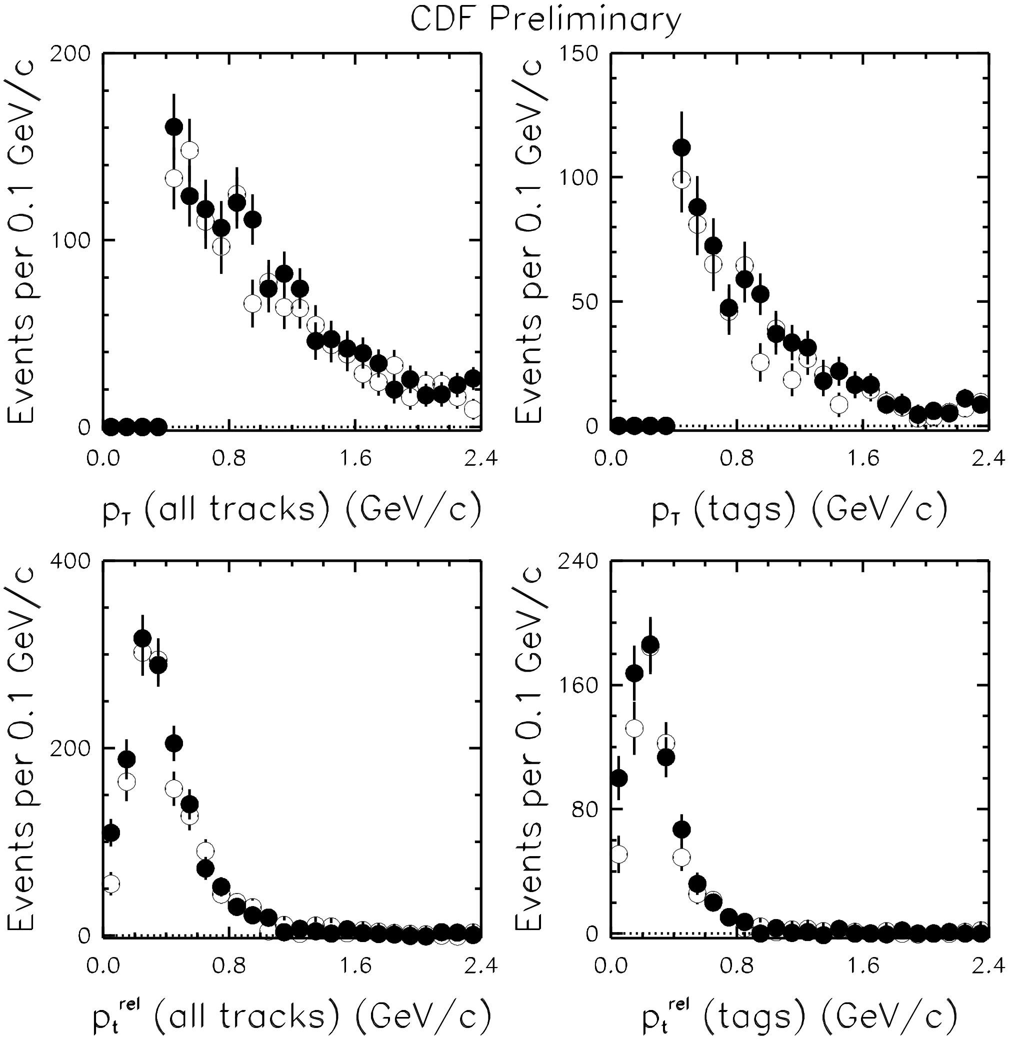


Figure 6-6: Tagging distributions for decay signature "$\ell^+ D^-, D^- \to K\pi\pi$". Solid points are "right sign" tags, open are "wrong sign" tags Upper left: p_T of the SST candidate tracks. Upper right: p_T of the tag. Lower left: p_t^{rel} (the value of the tagging function) of the SST candidate tracks. Lower right: p_t^{rel} of the tag.

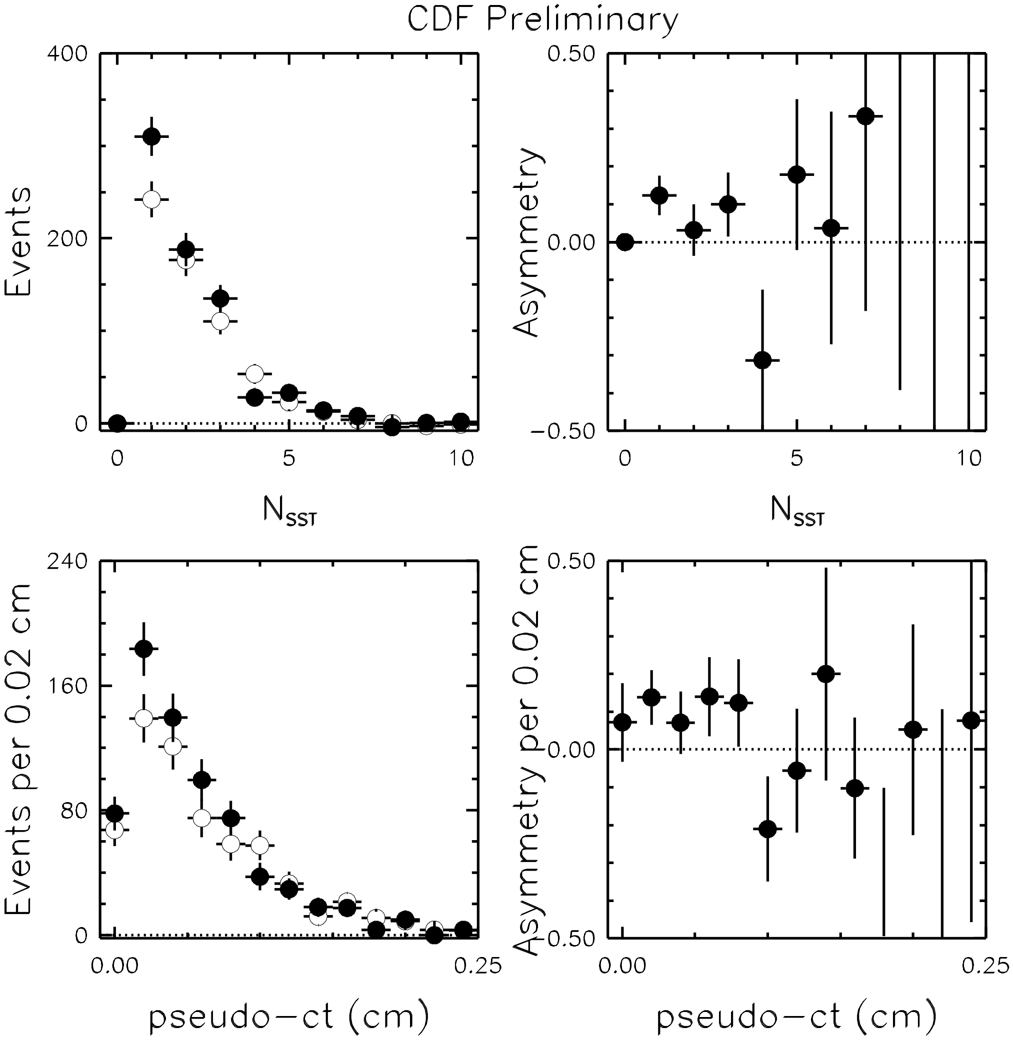


Figure 6-7: Tagging distributions for decay signature "$\ell^+ D^-, D^- \to K\pi\pi$". Solid points are "right sign" tags, open are "wrong sign" tags. Upper left: the total number of SST candidate tracks. Upper right: the asymmetry vs the total number of the SST candidate tracks. Lower left: the measured proper decay time (ct) corrected using the $\mathcal{K}$-factor of the 'direct' decay chain. Lower right: the asymmetry vs ct.

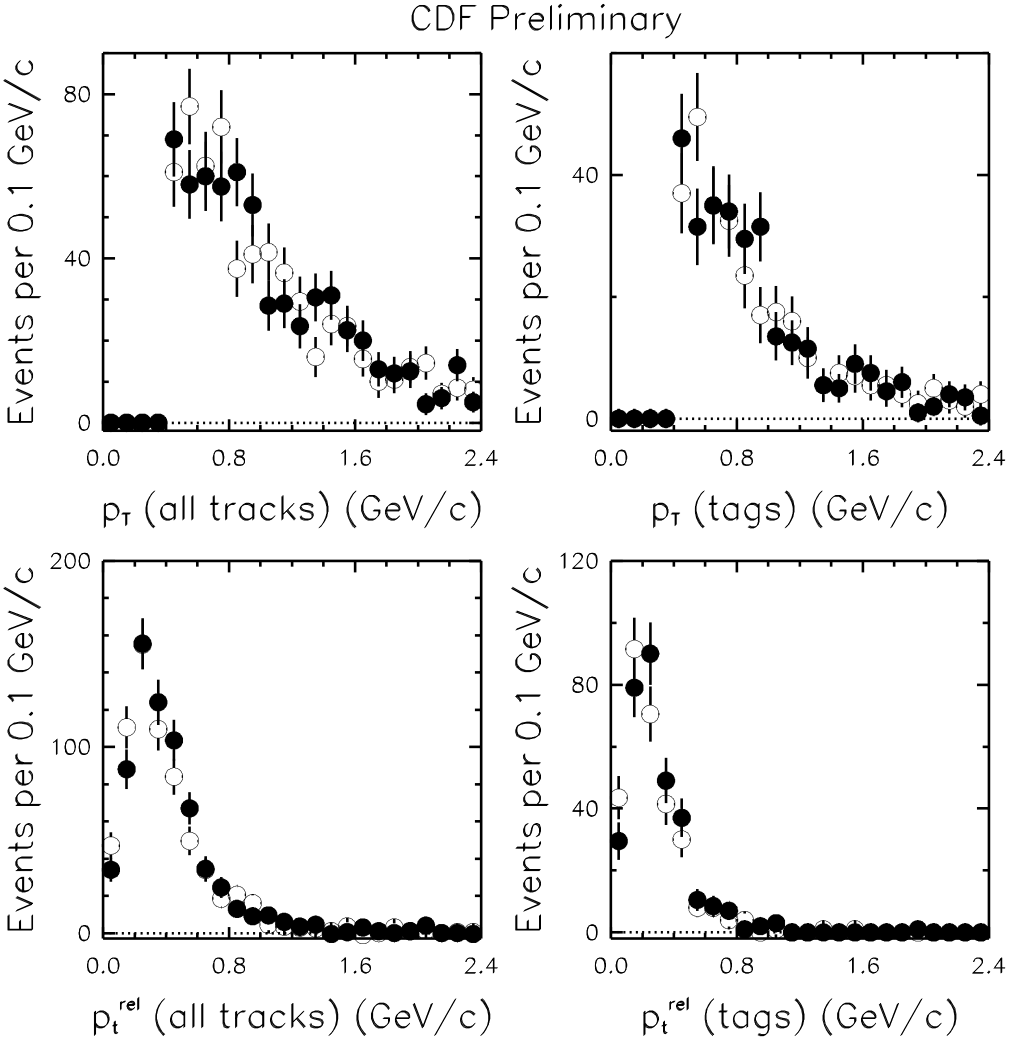


Figure 6-8: Tagging distributions for decay signature "$\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi$". Solid points are "right sign" tags, open are "wrong sign" tags Upper left: p_T of the SST candidate tracks. Upper right: p_T of the tag. Lower left: p_t^{rel} (the value of the tagging function) of the SST candidate tracks. Lower right: p_t^{rel} of the tag.

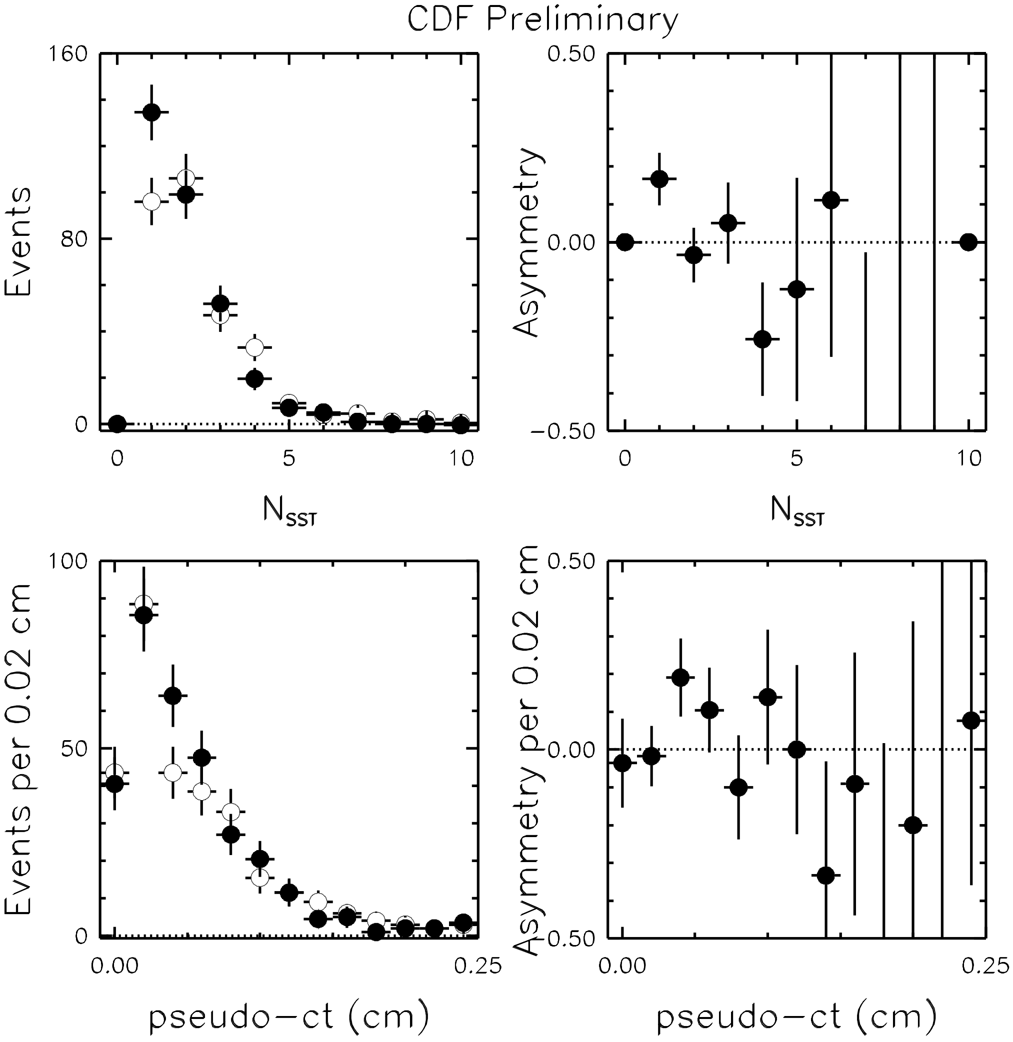


Figure 6-9: Tagging distributions for decay signature "$\ell^+ D^{*-}, \overline{D^0} \to K\pi$". Solid points are "right sign" tags, open are "wrong sign" tags. Upper left: the total number of SST candidate tracks. Upper right: the asymmetry vs the total number of the SST candidate tracks. Lower left: the measured proper decay time (ct) corrected using the $\mathcal{K}$-factor of the 'direct' decay chain. Lower right: the asymmetry vs ct.

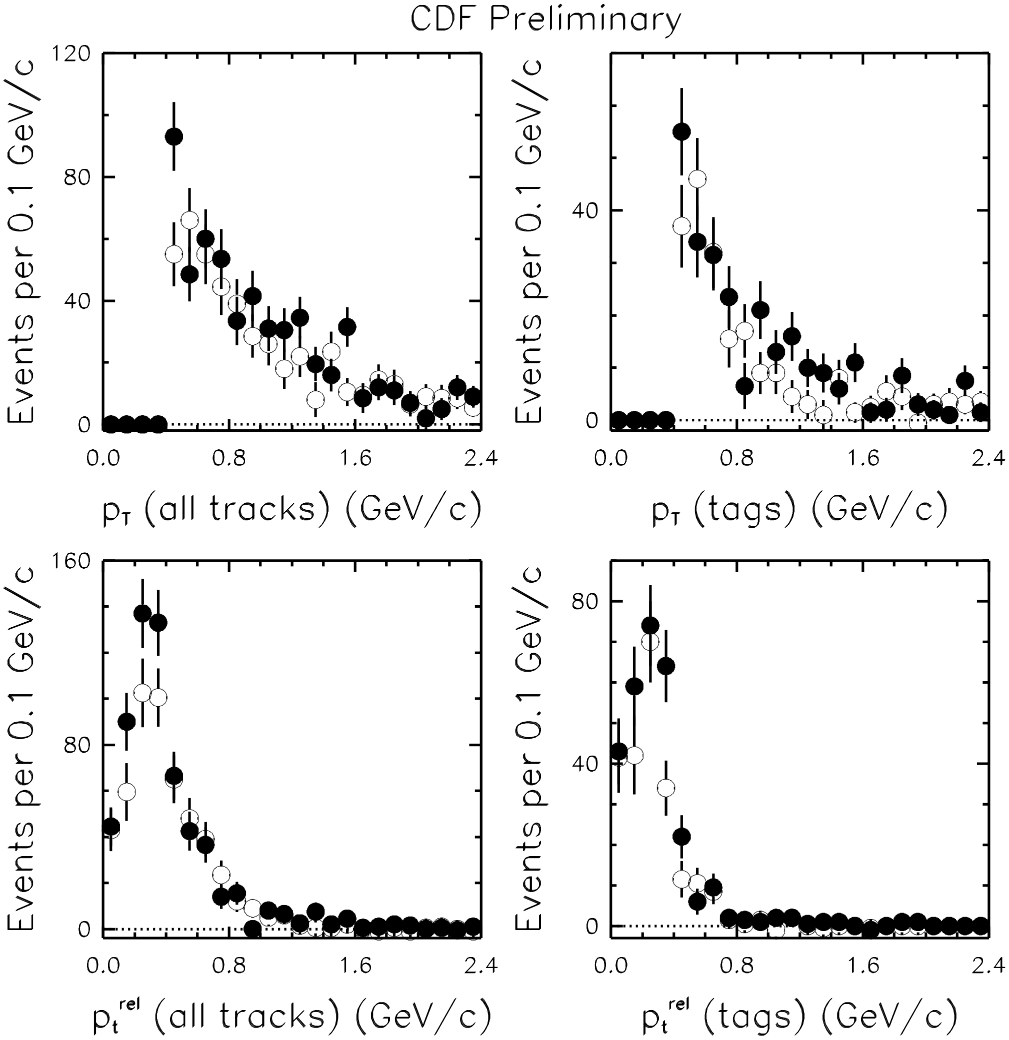


Figure 6-10: Tagging distributions for decay signature "$\ell^+ D^{*-}, \overline{D^0} \to K3\pi$". Solid points are "right sign" tags, open are "wrong sign" tags Upper left: p_T of the SST candidate tracks. Upper right: p_T of the tag. Lower left: p_t^{rel} (the value of the tagging function) of the SST candidate tracks. Lower right: p_t^{rel} of the tag.

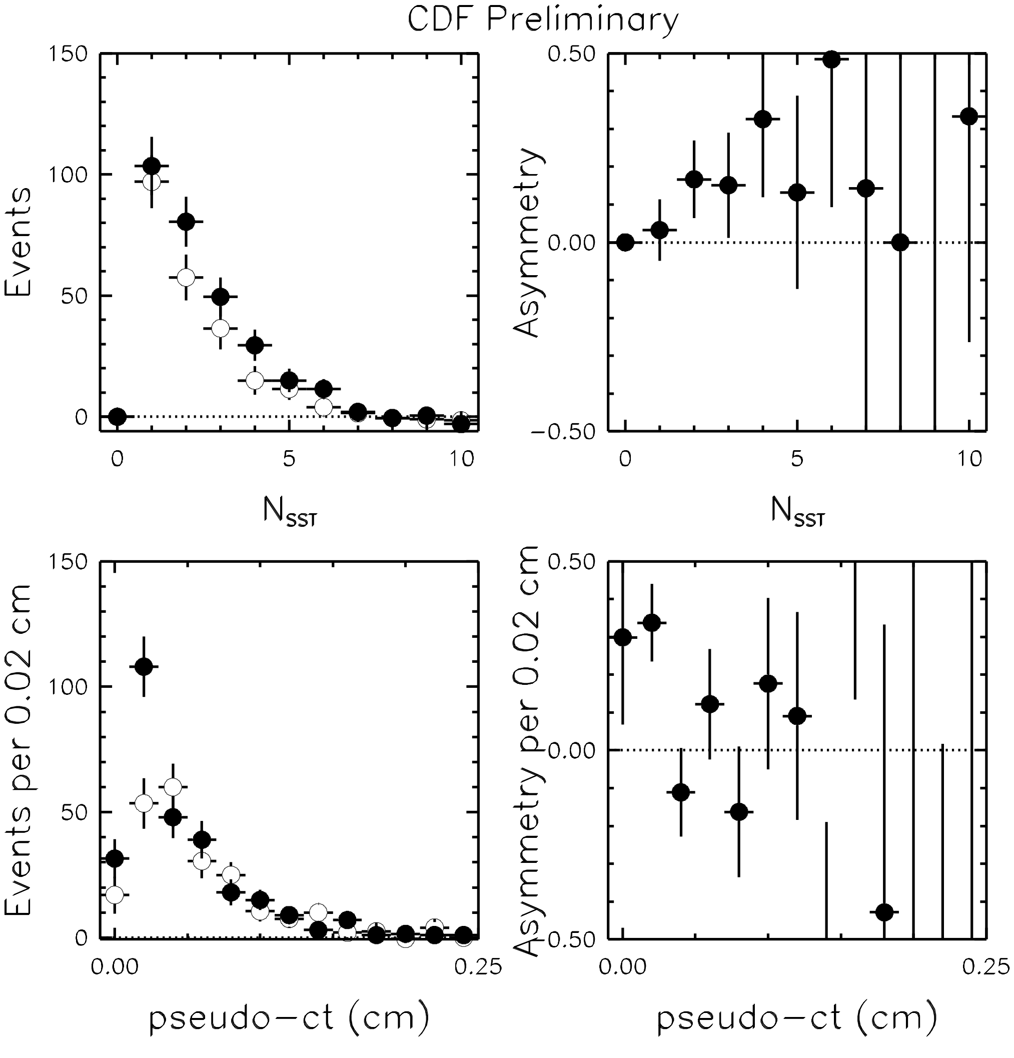


Figure 6-11: Tagging distributions for decay signature "$\ell^+ D^{*-}, \overline{D^0} \to K3\pi$". Solid points are "right sign" tags, open are "wrong sign" tags. Upper left: the total number of SST candidate tracks. Upper right: the asymmetry vs the total number of the SST candidate tracks. Lower left: the measured proper decay time (ct) corrected using the $\mathcal{K}$-factor of the 'direct' decay chain. Lower right: the asymmetry vs ct.

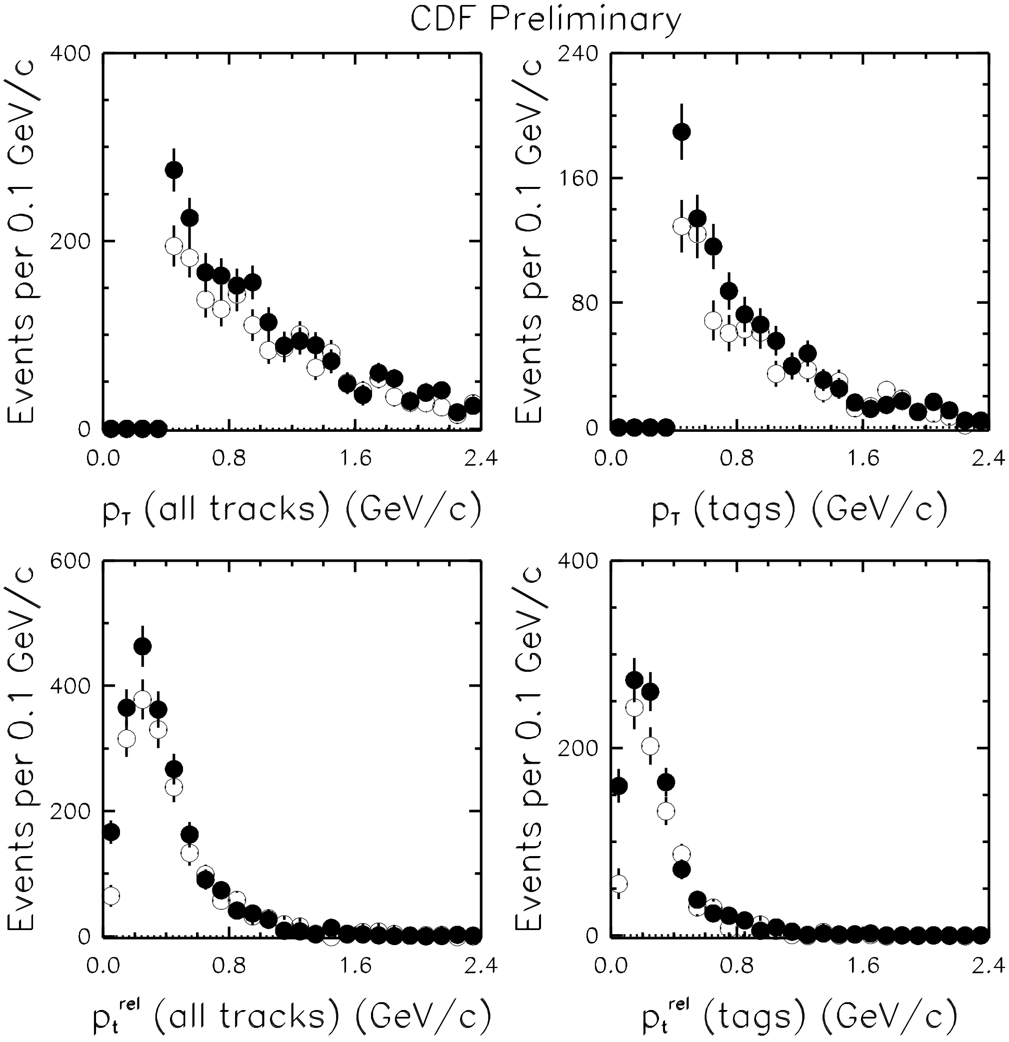


Figure 6-12: Tagging distributions for decay signature "$\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi\pi^0$". Solid points are "right sign" tags, open are "wrong sign" tags Upper left: p_T of the SST candidate tracks. Upper right: p_T of the tag. Lower left: p_t^{rel} (the value of the tagging function) of the SST candidate tracks. Lower right: p_t^{rel} of the tag.

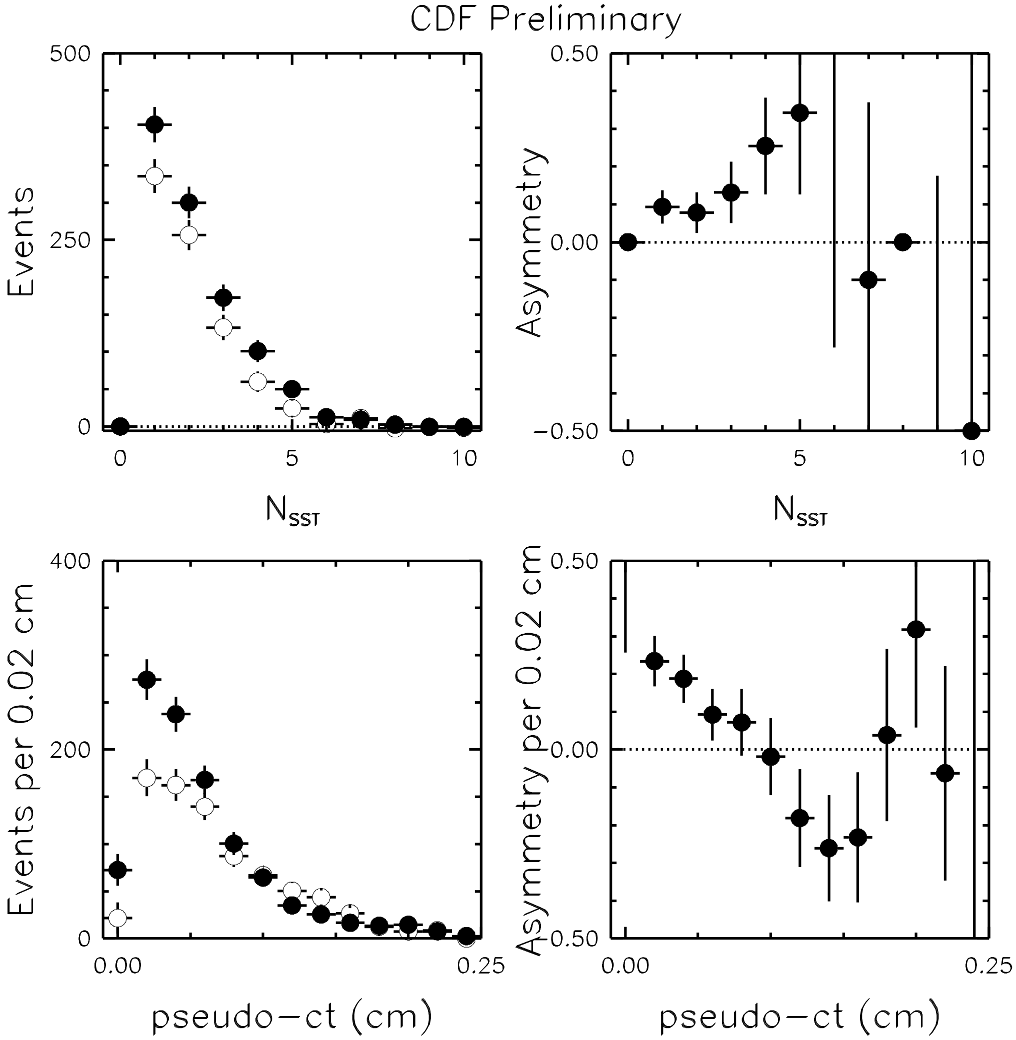


Figure 6-13: Tagging distributions for decay signature "$\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi\pi^0$". Solid points are "right sign" tags, open are "wrong sign" tags. Upper left: the total number of SST candidate tracks. Upper right: the asymmetry vs the total number of the SST candidate tracks. Lower left: the measured proper decay time (ct) corrected using the $\mathcal{K}$-factor of the 'direct' decay chain. Lower right: the asymmetry vs ct.

the data as $g(ct)$. The 'data average' $\overline{ct}$ of the bin is the point where $g(\overline{ct})$ equals the average $g(ct)$ over the bin, $\overline{g(ct)}$. $g(ct)$ is a product of two functions: an exponential $e^{-ct/c\tau}$ convoluted with the resolution function (since ct is not the true proper time), and an efficiency turn-on (since the B meson candidate selection requirements[4] bias the ct distribution, especially at low ct).

Because of the exponential, $g(ct)$ is a rapidly falling function of ct, and because of the finite bin width, the B candidates' ct are unevenly distributed along the bin. Therefore, the true, data-weighted center of the bin is

$$\overline{ct} \equiv \int_{ct_1}^{ct_2} ct\; g(ct) d(ct) \tag{6.2}$$

where ct_1 and ct_2 are the bin boundaries.

In order to derive the data-weighted center of each bin, we use the mean ct of the background-subtracted distribution of the data itself. (The signal is, by definition, distributed in ct as $g(ct)$.) For each bin, we derive the *ct-shift*, the difference of the data-weighted bin center from the nominal bin center. If the $g(ct)$ distribution is falling, the ct-shift is negative (the real bin center is to the left of the nominal bin center). The ct-shifts for the six ct bins, for all five decay signatures, are shown in table 6.1.

6.1.3 Measured asymmetries

Now that we know where to place the asymmetries in each ct bin, we obtain them by fitting the mass distributions for the events with the right and the wrong correlation. The right- and wrong-correlation mass distributions are fitted together, by forcing them to have the same mean and width of the gaussian and the same slope of the background as the combined mass distribution for the whole signature.[5] Therefore,

---

[4] Mainly due to d/σ cuts on the D-meson daughter tracks, and the $L_{\rm xy}^B/\sigma(L_{\rm xy}^B)$ cut.

[5] In the case of the 'satellite' decay signature,"$\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$", the whole idea is generalized into fixing the shape of the signal and the shape of the background in the overall fit first, then fitting the right- and wrong-correlation mass difference distributions of each ct bin simultaneously by letting the number of signal events and the background float.

Decay signature	bin 1	bin 2	bin 3	bin 4	bin 5	bin 6
$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$	0.0059	-0.0003	-0.0009	-0.0008	-0.0003	-0.0208
$\ell^+D^-, D^- \to K\pi\pi$	0.0052	-0.0003	-0.0015	-0.0012	-0.0010	-0.0174
$\ell^+D^{*-}, \overline{D^0} \to K\pi$	0.0047	-0.0009	-0.0016	-0.0016	-0.0020	-0.0197
$\ell^+D^{*-}, \overline{D^0} \to K3\pi$	0.0024	-0.0010	-0.0020	-0.0014	-0.0014	-0.0224
$\ell^+D^{*-}, \overline{D^0} \to K\pi\pi^0$	0.0042	-0.0005	-0.0009	-0.0011	-0.0017	-0.0227

Table 6.1: The ct-shifts: the differences between the data-weighted bin centers and the nominal bin centers, in cm. The ct-shifts are positive in the first bin because the underlying distribution of the data, $g(ct)$, is rising, due to the turn-on of the B candidate reconstruction efficiency. The ct-shifts are only important in the last bin, which is the largest.

the resulting asymmetries have the combinatorial background subtracted – they are the measured asymmetries of the $\ell^+D^{(*)}$ signal only. They are shown in Fig. 6-14 and 6-15. From now on we will denote these measured asymmetries by $\mathcal{A}_k(ct)$, where k is an index running over the decay signatures, and ct is the 'data average' proper time of a given bin.

However, due to the cross-talk between the $B \to \nu\ell^+D^{(*)}$ decay signatures, the observed asymmetries are the combination of the true asymmetries of the B^0 and B^+ component. We address the necessary corrections in section 7.1, and demonstrate that, once the sample composition of each decay signature is known, it is possible to extract $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d from the measured asymmetries, $\mathcal{A}_k(ct)$.

Nevertheless, even without the corrections for the cross-talk, the observed asymmetries are behaving as expected. The predominantly B^+ signature, $\ell^+\overline{D^0}, \overline{D^0} \to K\pi$, (Fig. 6-14, top) is mostly constant, while the asymmetries for the other four decay signatures (dominated by the B^0 decays) start out as positive, decrease as the ct increases, and end up as negative, hinting an $B^0\bar{B}^0$ oscillation.

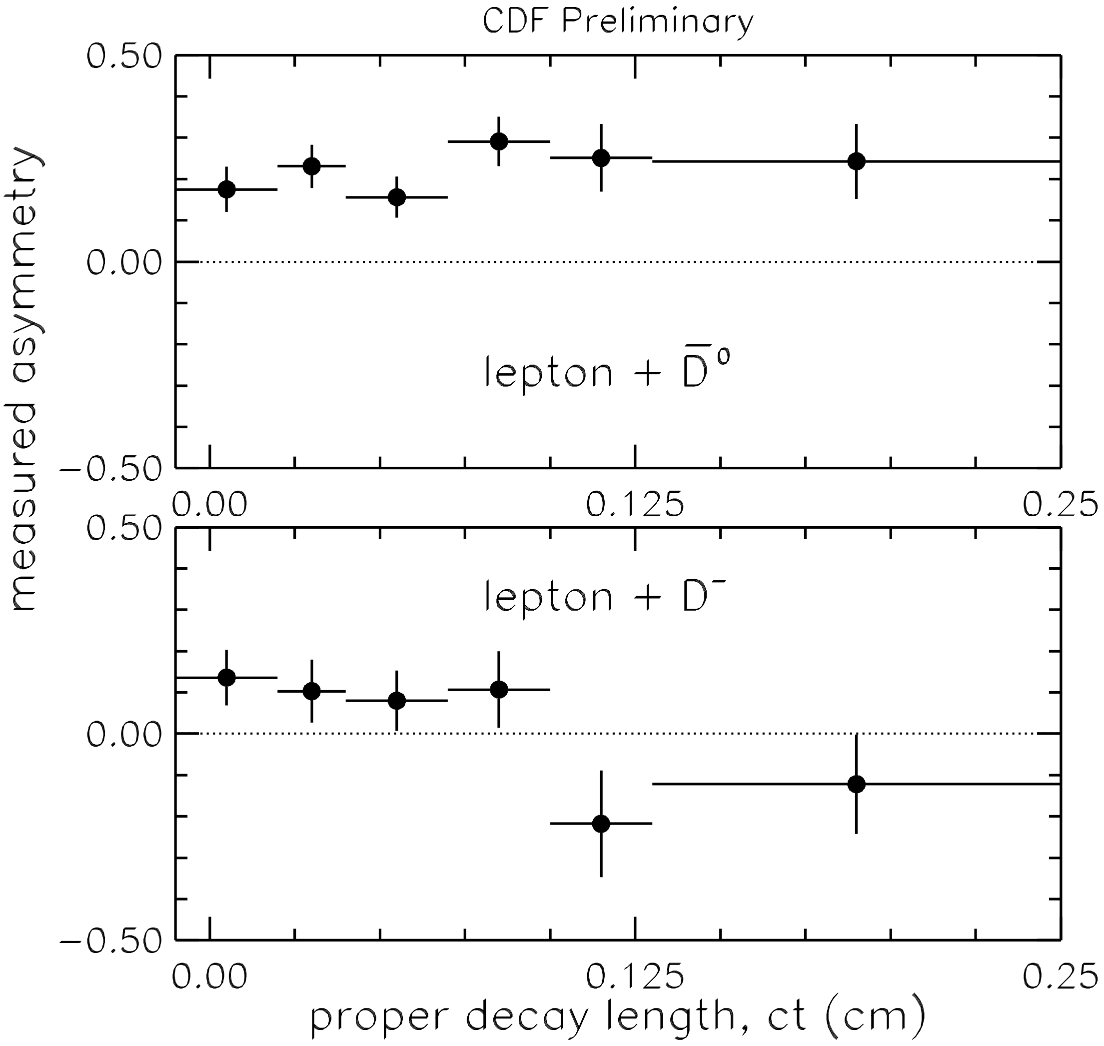


Figure 6-14: The measured asymmetries versus the measured proper time, ct, for the decay signatures "$\ell^+\overline{D^0}, \overline{D^0} \rightarrow K\pi$" and "$\ell^+D^-, D^- \rightarrow K\pi\pi$".

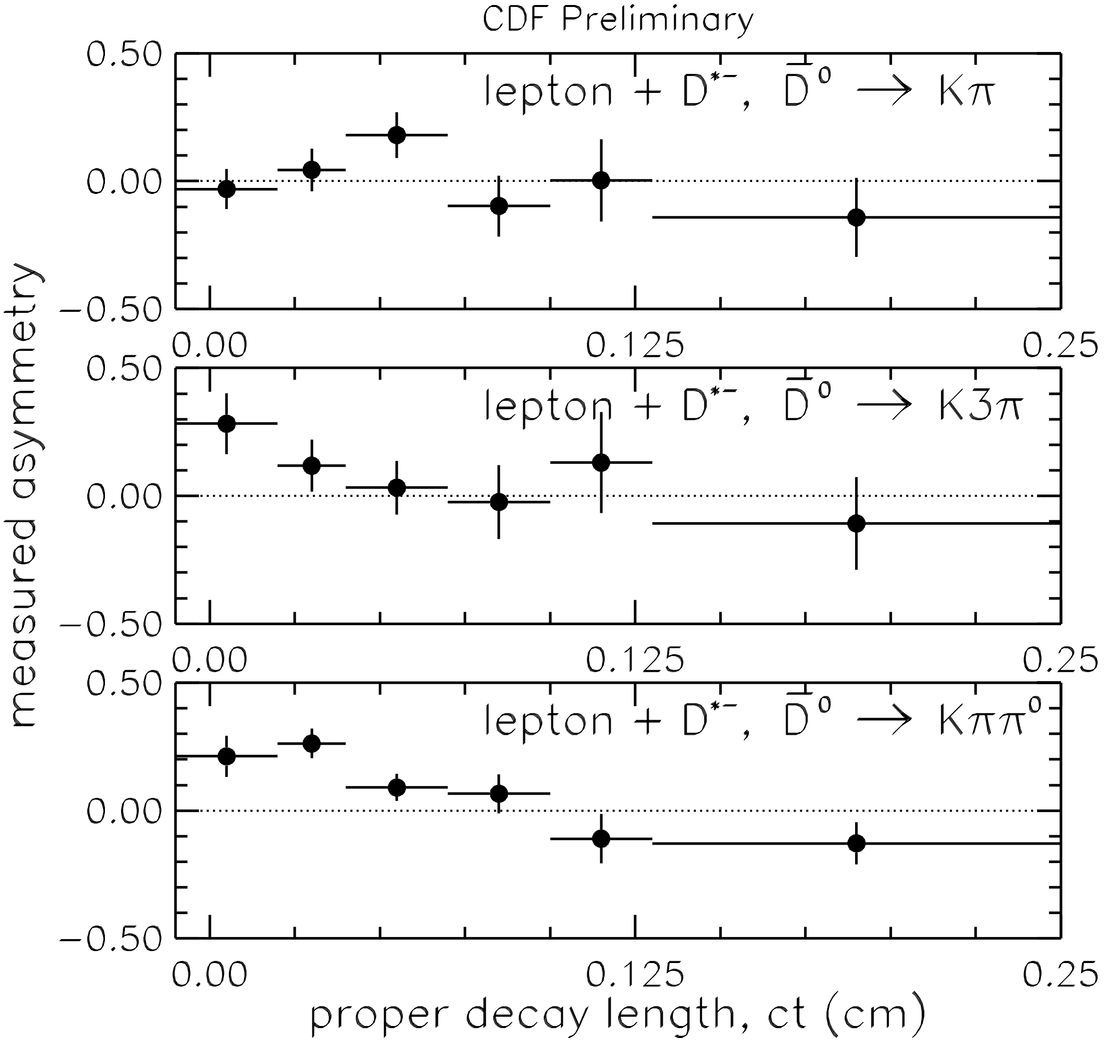


Figure 6-15: The measured asymmetries versus the measured proper time, ct, for the three $\ell^+ D^{*-}$ decay signatures "$\ell^+ D^{*-}, \overline{D^0} \to K\pi$", "$\ell^+ D^{*-}, \overline{D^0} \to K3\pi$" and "$\ell^+ D^-, D^- \to K\pi\pi$".

Chapter 7

Measurement of the SST dilution and the $B^0\bar{B}^0$ oscillation frequency

7.1 Fitting for $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d

As outlined in section 2.4, our goal is to fit the asymmetries as a function of the proper time, the B^+ asymmetry with a constant, and the B^0 asymmetry with a cosine (equation 2.16). In this section, we carry this through. We first (in section 7.1.1) examine the simple case of the pure B^0 and B^+ samples, and sketch the method for extracting $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d from the observed asymmetries. Then, in section 7.1.2, we explain how the observed asymmetries are modified in the presence of the cross-talk. Although there is a new effect that needs to be taken into account (namely, selecting the charged pion from the D^{**} decay, $\pi^{\pm}_{**}$, as a tag), we will show that once the sample composition is known, it is fairly easy to predict the measured asymmetries. In section 7.1.3 we express the observed asymmetries as a linear combination of the true asymmetries, the coefficients being determined by the sample composition. Tagging on the π_{**} is potentially troublesome, since the π_{**} is always correctly correlated with the lepton; correcting for this effect requires one more factor to be derived from the data, and this is done in section 7.2. Finally, in section 7.3 we present the result of the fit – the values for $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d.

7.1.1 Time dependence of the true asymmetries

If the B decay signature $\ell^+\overline{D^0}$ was composed of 100% B^+ mesons, and the other four decay signatures were pure B^0 mesons, then the measured asymmetry for the B^+ signature would be equal to the true B^+ asymmetry, $\mathcal{A}^+$, while the asymmetry for the B^0 signatures would be equal to the true B^0 asymmetry, $\mathcal{A}^0$.

The true asymmetry for B^+ is constant in ct, while for B^0 it displays a cosine dependence on ct:

$$\mathcal{A}^+(ct) = \mathcal{D}_+ = \text{const} \tag{7.1}$$

$$\mathcal{A}^0(ct) = \mathcal{D}_0\,(\cos\otimes G)(ct, \Delta m_d, \sigma_{ct}) \tag{7.2}$$

where $(\cos\otimes G)(ct, \sigma_{ct})$ denotes the cosine convoluted with the gaussian resolution function, $G(ct, \sigma_{ct})$. σ_{ct} is the resolution of the ct (measured as the width of Δct, eq. (4.10)), and is parameterized by eq. (4.12):

$$\sigma_{ct} \equiv \sigma(ct) = \sigma_0 + b \times ct$$

When calculating $(\cos\otimes G)(ct, \sigma)$, we substitute $\sigma_{ct} = \sigma(ct)$, and therefore $\cos\otimes G$ is a function of ct and Δm_d. We note that the measured proper time, ct, is, as always, obtained by correcting by the $\mathcal{K}$-factor of the 'direct' decay chain ($\mathcal{K}_{kd}$). In this simple case the B^0 and B^+ are perfectly separated and all five signatures are composed only of their respective 'direct' decay chains, so there is no need for any further correction, in contrast to the case where the cross-talk is present.

In order to derive $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d from the 30 measured asymmetries $\mathcal{A}_k^{(meas)}(ct)$, and their errors, $\sigma_k^{\mathcal{A}}(ct)$, we build the following χ^2 function, in which we compare the difference between the predicted asymmetry $\mathcal{A}_k(ct) = \mathcal{A}^{0,+}(ct)$ and the observed asymmetry $\mathcal{A}_k^{(meas)}(ct)$ against the measurement error, $\sigma_k^{\mathcal{A}}(ct)$:

$$\chi^2 = \sum_{k,ct}\left(\frac{\mathcal{A}_k^{(meas)}(ct) - \mathcal{A}_k(ct)}{\sigma_k^{\mathcal{A}}(ct)}\right)^2 \tag{7.3}$$

Here, k is the index that runs over the five decay signatures, while ct symbolizes the summation over the proper times for all data points. $\mathcal{A}_k^{0,+}(ct)$ denotes either $\mathcal{A}^0$

or $\mathcal{A}^+$, depending on the decay signature k. Minimizing the χ^2 from equation (7.3) results in the selection of $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d that best describe the data. Of course, $\mathcal{A}_k(ct) = \mathcal{A}^{0,+}(ct)$ is true only when the B^0 and B^+ samples are pure. In the presence of the cross-talk, the prediction $\mathcal{A}_k(ct)$ must also include the effects of the sample composition.

7.1.2 Measured asymmetries and the sample composition

In the presence of cross-talk, the effects described in sections 5.1 and 5.2 will take place. For instance, the proper time used for the calculation of the true asymmetries will not be ct, but rather ct_k^0 and ct_k^+ (equations (5.36) and (5.37)). More important is the change in the observed asymmetries. As it has been described in section 6.1, a B^+ is preferentially produced with a π^-, while a B^0 is produced with a π^+ (figure 6-1): the b quarks of the same flavor are correctly tagged by pions of the *opposite* charge. When a $B \to \nu \ell D^{(*)}$ sample is pure, this does not concern us, since the observed asymmetries are the same as the true B^0 and B^+ asymmetries. However, in section 5.1 we have seen that there is a cross-contamination between the B^0 and B^+ samples, and that each of the five reconstructed B-meson signatures is a combination of B^0 and B^+ decays. When the B decay chain is misidentified, picking the correct track as the tag results in the *wrong* correlation between the tag and the lepton. Adding these decays to the sample of correctly tagged decays causes the observed asymmetry to be *lower* than the true asymmetry.

Furthermore, there is a possibility of selecting the $\pi_{**}^{\pm}$ from the D^{**} decay (see eq. (5.3)) as a tag by mistake. This is a potentially significant effect since the lepton and the π_{**} *always* have the right correlation, hence adding these decays to the sample of correctly tagged decays causes the observed asymmetry to be *higher* than the true asymmetry. This effect is quantified by f^{**} and the probability of selecting the π_{**} as a tag in a tagged event in which the π_{**} was produced, which we call ξ:

$$\xi \equiv \frac{N(\text{tagged on } \pi_{**})}{N(\pi_{**} \text{ was produced and the event is taggable})} \tag{7.4}$$

The parameter ξ is defined in such a way so that the contribution of the tagging algorithm alone can be separated from the D^{**} branching ratios (see Appendix F for more details). It depends on the tagging algorithm, on the kinematics of the π_{**} as well as on the properties of the fragmentation and the underlying event tracks. The dependence of ξ on ct is derived from the Monte Carlo simulation, but the overall normalization factor is derived from the data, since it is hard to model the fragmentation and the underlying event in the Monte Carlo simulation. The details of the normalization procedure are given in section 7.2.3.

Knowing $\xi(ct)$ and the sample composition (expressed in terms of $\epsilon(\pi_s)$, f^{**}, R_f, P_V, τ_{B^+}/τ_{B^0}, $\epsilon^D_{k\ell}$, $\mathcal{K}_{k\ell}$), we can calculate the prediction for the observed asymmetries, for each of the 30 measured asymmetry points, $\bar{\mathcal{A}}_k(ct)$, that includes the effects introduced by the sample composition: wrong correlation for the cross-talk, and tagging on the π_{**}. $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d, together with the sample composition parameters f^{**}, R_f, τ_{B^+}/τ_{B^0} (table 5.2), and the fitted quantities $\epsilon(\pi_s)$ and P_V, are used to predict the asymmetry in the decay signature k, at the proper time ct,

$$\bar{\mathcal{A}}_k(ct) = \bar{\mathcal{A}}_k(ct, \mathcal{D}_+, \mathcal{D}_0, \Delta m_d, f^{**}, P_V, R_f, \tau_{B^+}/\tau_{B^0}, \epsilon(\pi_s), \ldots).$$

The χ^2 function (7.3) in now modified to use $\bar{\mathcal{A}}_k(ct)$ as the predictions for the measured asymmetries, instead of the true asymmetries, $\mathcal{A}_k(ct)$:

$$\chi^2 = \sum_{k,ct} \left(\frac{\mathcal{A}_k^{(meas)}(ct) - \bar{\mathcal{A}}_k(ct)}{\sigma_k^{\mathcal{A}}(ct)} \right)^2 \tag{7.5}$$

We next describe how the asymmetry predictions $\bar{\mathcal{A}}_k(ct)$ are calculated. We start from the true B^0 and B^+ asymmetries, equations (7.1) and (7.2) and then explain how they are combined into the predictions for the measured asymmetries, $\bar{\mathcal{A}}_k(ct)$.

7.1.3 Calculating the predictions for the measured asymmetries in the presence of cross-talk

We now examine how the sample composition affects the asymmetry $\bar{\mathcal{A}}_k(ct)$, in the decay signature k and at the measured proper time ct. As an example, we consider the B^0 decay signature $k =$ "$\ell^+ D^-, D^- \to K\pi\pi$". Here, the B^0 mesons are produced with the correctly correlated pions, while the B^+ mesons comprise the cross-talk, and are produced with the pions with the opposite charge correlation.

The measured asymmetries are binned in the measured ct (figure 6-14). In each ct bin, the B decays are either B^0 or B^+, divided according to equations (5.38) and (5.39). The total contributions of the B^0 and B^+ decays are

$$\begin{aligned} \Phi_k^0(ct_k^0(ct)) &\equiv e^{-ct_k^0(ct)/c\tau_{B^0}} \sum_{B^0\to\ell} \phi_{k\ell}\epsilon_{k\ell}^D && (7.6) \\ \Phi_k^+(ct_k^+(ct)) &\equiv e^{-ct_k^+(ct)/c\tau_{B^+}} \sum_{B^+\to\ell} \phi_{k\ell}\epsilon_{k\ell}^D && (7.7) \end{aligned}$$

The dependence of the true proper times for each component, ct_k^0 and ct_k^0, on the measured proper time, ct, has been explicitly stated. Using (7.6) and (7.7), equations (5.38) and (5.39) can be rewritten as

$$\begin{aligned} \text{fraction of } k \text{ from } B^0 &= \frac{\Phi_k^0(ct_k^0)}{\Phi_k^0(ct_k^0)+\Phi_k^+(ct_k^+)} && (7.8) \\ \text{fraction of } k \text{ from } B^+ &= \frac{\Phi_k^+(ct_k^+)}{\Phi_k^0(ct_k^0)+\Phi_k^+(ct_k^+)} && (7.9) \end{aligned}$$

Here, for simplicity, the dependence of $ct_k^{0,+}$ on ct is implied. Throughout the rest of this section, we will assume that any dependence on ct has been realized via $ct_k^0(ct)$ and $ct_k^+(ct)$ as given by eqs. (5.36) and (5.37).

In the simple case where the π_{**}'s are never selected as tags, the observed asymmetry $\bar{\mathcal{A}}_k(ct)$ would have been just the linear combination of the true asymmetries $\mathcal{A}^0(ct)$ and $\mathcal{A}^+(ct)$[1]:

$$\bar{\mathcal{A}}_m(ct) = \frac{\Phi_m^0(ct_m^0)}{\Phi_m^0(ct_m^0)+\Phi_m^+(ct_m^+)}\mathcal{A}^0(ct_m^0) + \frac{\Phi_m^+(ct_m^+)}{\Phi_m^0(ct_m^0)+\Phi_m^+(ct_m^+)}(-\mathcal{A}^+(ct_m^+)) \quad (7.10)$$

Here the true asymmetries are calculated using (7.1) and (7.2), however the proper times for the B^0 and B^+ subsamples, ct_k^0 and ct_k^+, are used instead of the measured proper time, ct. The true asymmetry for the cross-talk (B^+ sample), $\mathcal{A}^+(ct_k^+)$, comes

---

[1] Suppose that one has N data samples with asymmetry $\mathcal{A}_i$ $(i = 1, ...N)$ contributing to a signal. The total observed asymmetry is the linear sum of the individual asymmetries:

$$\bar{\mathcal{A}} = \sum \alpha_i \mathcal{A}_i^{(true)}$$

where α_i is a fraction of subset i, having asymmetry $\mathcal{A}_i$. ($\sum \alpha_i = 1$.)

with a minus sign, since the correctly selected tagging pions have the wrong correlation. Index m is the same as k, except that it is used for the B^0 decay signatures. Equation (7.10) holds also for the other three B^0 decay signatures. In the case of the only one B^+ decay signature, $\ell^+\overline{D^0}, \overline{D^0} \to K\pi$, the signs 0 and + must be exchanged, but the effect is the same:

$$\bar{\mathcal{A}}_n(ct) = \frac{\Phi_n^+(ct_n^+)}{\Phi_n^0(ct_n^0) + \Phi_n^+(ct_n^+)}\mathcal{A}^+(ct_n^+) + \frac{\Phi_n^0(ct_n^0)}{\Phi_n^0(ct_n^+) + \Phi_n^+(ct_n^+)}(-\mathcal{A}^0(ct_n^0)) \tag{7.11}$$

Index n is the same as k, except that it is used for the B^+ decay signatures.

Now we consider tagging on the π_{**}. In the formalism, we allow that the charged π_{**} (the only one we can tag on) can come from both the B^0 and B^+ (although, in the above example of $k =$"$\ell^+D^-, D^- \to K\pi\pi$", $\pi_{**}^\pm$ can only appear in the decay chains originating from the B^+ meson). In the decays when the π_{**} is coming from B^+, it is always correctly correlated with the lepton, and this subsample carries the asymmetry of +1. In the decays where the π_{**} is coming from B^0, it is always *incorrectly* correlated with the lepton, and this subsample carries the asymmetry of −1. We now need a finer break-down of the B^0 and B^+ decays, into B^0 decays with and without π_{**}, and analogously for the B^+. We define the contribution of the decays with the π_{**} to the B^0 and B^+ subsamples:

$$\Phi_k^{0,**}(ct_k^0(ct)) \equiv e^{-ct_k^0(ct)/c\tau_{B^0}} \sum_{B^0\to\ell,\, \exists\pi_{**}^\pm} \phi_{k\ell}\epsilon_{k\ell}^D \tag{7.12}$$

$$\Phi_k^{+,**}(ct_k^+(ct)) \equiv e^{-ct_k^+(ct)/c\tau_{B^+}} \sum_{B^+\to\ell,\, \exists\pi_{**}^\pm} \phi_{k\ell}\epsilon_{k\ell}^D \tag{7.13}$$

Here the notation "$B^0 \to \ell,\ \exists\pi_{**}^\pm$" implies a summation over all decay chains ℓ originating from the B^0 meson and resulting in a charged pion from D^{**}, the $\pi_{**}^\pm$.

However, not all charged π_{**}'s are selected as tags. Only a fraction $\xi(ct)$ of them are. (ξ is a probability of selecting the π_{**} as a tag in a tagged event in which the π_{**} was produced, eq. (7.4).) Therefore, we split the B^0 and B^+ components of the bin ct into:

$$\begin{aligned} \alpha_k^0(ct) &\equiv \text{fraction of } k \text{ from } B^0 \text{ with no tag on } \pi_{**} \\ &= \frac{\Phi_k^0(ct_k^0) - \xi(ct_k^0)\Phi_k^{0,**}(ct_k^0)}{\Phi_k^0(ct_k^0) + \Phi_k^+(ct_k^+)} \end{aligned} \tag{7.14}$$

$$\begin{aligned}\alpha_k^{0,**}(ct) &\equiv \text{fraction of } k \text{ from } B^0 \text{ with a } \pi_{**} \text{ tag} \\ &= \frac{\xi(ct_k^0)\Phi_k^{0,**}(ct_k^0)}{\Phi_k^0(ct_k^0)+\Phi_k^+(ct_k^+)} \qquad (7.15)\\ \alpha_k^{+}(ct) &\equiv \text{fraction of } k \text{ from } B^+ \text{ with no tag on } \pi_{**} \\ &= \frac{\Phi_k^+(ct_k^+)-\xi(ct_k^+)\Phi_k^{+,**}(ct_k^+)}{\Phi_k^0(ct_k^0)+\Phi_k^+(ct_k^+)} \qquad (7.16)\\ \alpha_k^{+,**}(ct) &\equiv \text{fraction of } k \text{ from } B^+ \text{ with a } \pi_{**} \text{ tag} \\ &= \frac{\xi(ct_k^+)\Phi_k^{+,**}(ct_k^+)}{\Phi_k^0(ct_k^0)+\Phi_k^+(ct_k^+)} \qquad (7.17)\end{aligned}$$

The dependence of ct_k^0 and ct_k^+ on ct is implied. The α parameters are called the asymmetry coefficients, and they contain all information about the sample composition. By definition, they are normalized to unity,

$$\alpha_k^0(ct)+\alpha_k^{0,**}(ct)+\alpha_k^+(ct)+\alpha_k^{+,**}(ct)=1$$

Using equations (7.14), (7.15), (7.16) and (7.17), we can specify the prediction for the measured asymmetry, $\bar{\mathcal{A}}_k(ct)$, in the decay signature k, at the proper time ct:

$$\begin{aligned}\bar{\mathcal{A}}_k(ct) &= \alpha_k^0(ct)\mathcal{A}^0(ct_k^0) \qquad + \alpha_k^{0,**}(ct)(-1) \\ &+ \alpha_k^+(ct)(-\mathcal{A}^+(ct_k^+)) + \alpha_k^{+,**}(ct)(+1) \qquad (7.18)\end{aligned}$$

In summary, the first term in eq. (7.18) corresponds to the decay chains originating from the decay of the B^0 meson, and the third term describes the contribution of the cross-contamination from the B^+ decay chains. Since the correlation between the B^+ meson and the tagging pion is opposite to the correlation in the case of the B^0 meson, the third term is multiplied by a *negative* asymmetry for charged B's, $-\mathcal{A}^+$. The fourth term corresponds to the $B \to \nu\ell D^{**}$ decay chains where the π_{**} was picked as a tag by accident. Because the π_{**} is always correctly correlated with the lepton, these events contribute the asymmetry of exactly $+1$. The second term describes the decay chains where the D^{**} meson is coming from the B^0 meson, and the correlation is always wrong, thus the asymmetry of -1. Of course, this is possible only in the case of the B^+ decay signatures, however this term is, for consistency, also kept in the case of B^0 signatures, even though $\alpha_k^{**,0}(ct) \equiv 0$.

7.2 Constraining the sample composition parameters in the fit

The χ^2 function (7.5) can easily be extended to include fitting for other quantities that can be constrained from our data and the assumed sample composition, in particular $\epsilon(\pi_s)$ and ξ.

We fit for other parameters by letting them float in the fit, and providing additional terms to the χ^2 function (7.5):

$$\chi^2 = \sum_{k,ct} \left(\frac{\mathcal{A}_k^{(meas)}(ct) - \bar{\mathcal{A}}_k(ct)}{\sigma_k^{\mathcal{A}}(ct)} \right)^2 + \sum_j \left(\frac{F_j^{(meas)} - \bar{F}_j(f^{**}, P_V, R_f, \epsilon(\pi_s)\ldots)}{\sigma_j^F} \right)^2 \tag{7.19}$$

In eq. (7.19), $\bar{F}_j(f^{**}, P_V, R_f, \epsilon(\pi_s)\ldots)$ is a prediction of a quantity F_j, and is a function of the sample composition and possibly other parameters (e.g. ξ). $F_j^{(meas)} \pm \sigma_j^F$ is a measurement of F_j. If F_j strongly depends on a parameter, say $\epsilon(\pi_s)$, which is let free in the fit, then minimizing the χ^2 in (7.19) results in the optimal values of $\mathcal{D}_+$, $\mathcal{D}_0$, Δm_d *and* $\epsilon(\pi_s)$.

In section 7.2.1 we describe the case of sample composition parameters, which is the simplest, and then in 7.2.2 and 7.2.3 we explain how $\epsilon(\pi_s)$, ξ and P_V can be derived from the data.

7.2.1 Letting the parameters float in the fit

The most trivial example of F_j is a sample composition parameter itself. Setting, for example, $F_j = f^{**}$, results in a new χ^2 term:

$$\cdots + \left(\frac{f^{**\,(meas)} - f^{**}}{\sigma_{f^{**}}} \right)^2 + \cdots$$

Here, the measured value for f^{**}, $f^{**\,(meas)}$, and its error, $\sigma_{f^{**}}$, are taken from table 5.2. During the fit, f^{**} freely floats around $f^{**\,(meas)}$, within the error $\sigma_{f^{**}}$. In each iteration of the fit, f^{**} assumes a different value, which is used in all calculations dependent on the sample composition (e.g. equations (7.6), (7.7) and (7.18)). The minimal χ^2 will correspond to the value of f^{**} which is 'preferred' by the data.

If several sample composition parameters are floating – for example f^{**} and τ_{B^+}/τ_{B^0} – then the χ^2 will contain the terms

$$\cdots + \left(\frac{f^{**\,(meas)} - f^{**}}{\sigma_{f^{**}}} \right)^2 + \left(\frac{(\tau_{B^+}/\tau_{B^0})^{(meas)} - \tau_{B^+}/\tau_{B^0}}{\sigma_{\tau_{B^+}/\tau_{B^0}}} \right)^2 + \cdots$$

Apart from resulting in the optimal values for f^{**} and τ_{B^+}/τ_{B^0}, this kind of fit will also take into account any correlation between these two parameters. In each iteration of the fit, both f^{**} and τ_{B^+}/τ_{B^0} differ from their initial values; the asymmetry $\bar{\mathcal{A}}_k(ct)$ depends on both (see eq. (7.18)). The effect of a fluctuation in f^{**} on $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d is a function of τ_{B^+}/τ_{B^0}. The final uncertainty on $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d thus includes the correlations between f^{**} and τ_{B^+}/τ_{B^0}.

7.2.2 Measuring $\epsilon(\pi_s)$

A slightly more complicated case is that of the soft pion reconstruction efficiency, $\epsilon(\pi_s)$. In section 5.1.5 we have already explained how this efficiency can be obtained by dividing the measured $R^{*\,(meas)}$ (eq. (5.28)) by its prediction, $\bar{R}^*$ (eq. (5.29)). The measured value is $R^{*\,(meas)} = 0.249 \pm 0.008$.

Since R^* is a function of the sample composition parameters, $\epsilon(\pi_s)$ consequently also depends on them, and therefore should always be recalculated whenever the sample composition parameters change. If they are floating in the fit, $\epsilon(\pi_s)$ must also be recalculated in each iteration, too. We implement this by adding another term to eq. (7.19), with $F_j = R^*(f^{**}, R_f, \ldots)$:

$$\chi^2 \rightarrow \chi^2 + \left(\frac{R^{*\,(meas)} - \bar{R}^*}{\sigma^*} \right)^2 \tag{7.20}$$

and the convergence of the fit yields $\epsilon(\pi_s)$ in addition to $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d.

7.2.3 Measuring ξ and P_V

We extend the same principle to $\xi(ct)$ and P_V. Figure 7-1 shows the distribution of ξ as a function of ct as obtained from the Monte Carlo simulation (section 4.2.2), in the absence (fig. 7-1, top) and the presence (fig. 7-1, bottom) of the impact parameter significance requirement with respect to the primary vertex on the tagging candidate tracks. A topology of a typical event involving a D^{**} decay is schematically shown in fig. 7-2. Pions from the D^{**} decays (π_{**}) emerge from the B vertex. When the primary and the B vertices are well separated, the π_{**} tracks are expected to be inconsistent with coming from the primary vertex. If the tagging candidate tracks are not required to originate from the primary vertex, one does not expect any ξ dependence on ct. We see that in the absence of the d_0/σ_{d_0} cut, ξ is flat and relatively high ($\xi = 0.33$ means that in 1/3 of the B decays in which a π_{**} was produced, it is selected as a tag). However, if the tagging candidate tracks are required to be consistent with the primary vertex, then π_{**}'s will be among them only if the B vertex is sufficiently close to the primary. Consequently, when the requirement $d_0/\sigma_{d_0} < 3.0$ is imposed, ξ falls quickly to almost 0 as ct increases. This validates the requirement $d_0/\sigma_{d_0} < 3.0$ as one of the tagging selection criteria (sec. 6.1).

We use the $\xi(ct)$ shape from the Monte Carlo simulation (we denote it by $\xi_{MC}(ct)$). However, since the number of good tagging candidate tracks may be different in the data and in the Monte Carlo simulation, we use the average ξ from the data and derive a normalization factor, ξ_{norm}:

$$\xi(ct) = \xi_{norm} \cdot \xi_{MC}(ct) \tag{7.21}$$

We proceed analogously to the case of $\epsilon(\pi_s)$: we find a quantity F_j we can measure in the data as well as predict from the sample composition, and then compare the measured and predicted values in another χ^2 term.

The quantity in hand is the ratio of the number of events where the π_{**} was chosen as a tag, $N(\pi_{**}\text{ tags})$, to the total number of tagged events, $N(\text{tagged events})$, in each decay signature:

$$R^{**} \equiv \frac{N(\pi_{**}\text{ tags})}{N(\text{tagged events})} \tag{7.22}$$

For the purpose of counting $N(\pi_{**}\text{tags})$, we do not cut on d/σ: ξ is constant in ct and we measure the integral of ξ over the proper time (see fig. 7-1, top). This also ensures that the number of tagged π_{**}'s is sufficient to determine R^{**} (and thus ξ_{norm}) with

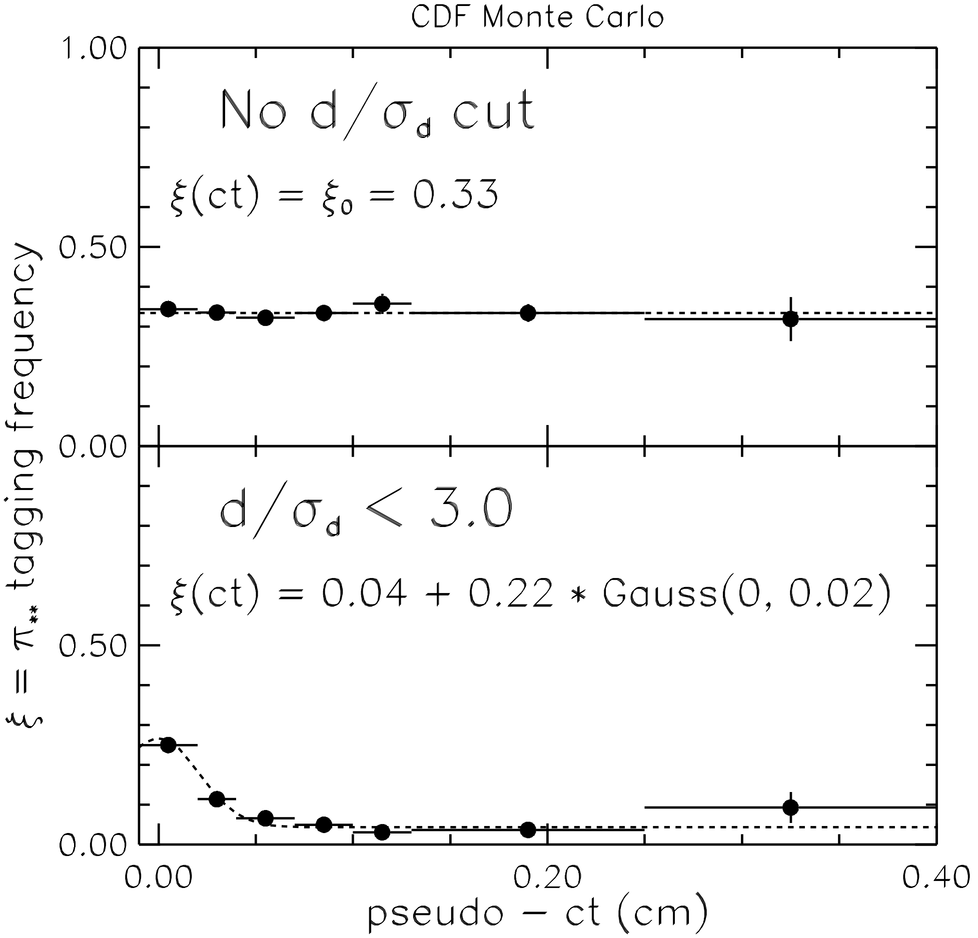


Figure 7-1: ξ as a function of corrected proper time, ct. Upper figure: no d_0/σ_{d_0} cut. Lower figure: $d_0/\sigma_{d_0} < 3.0$.

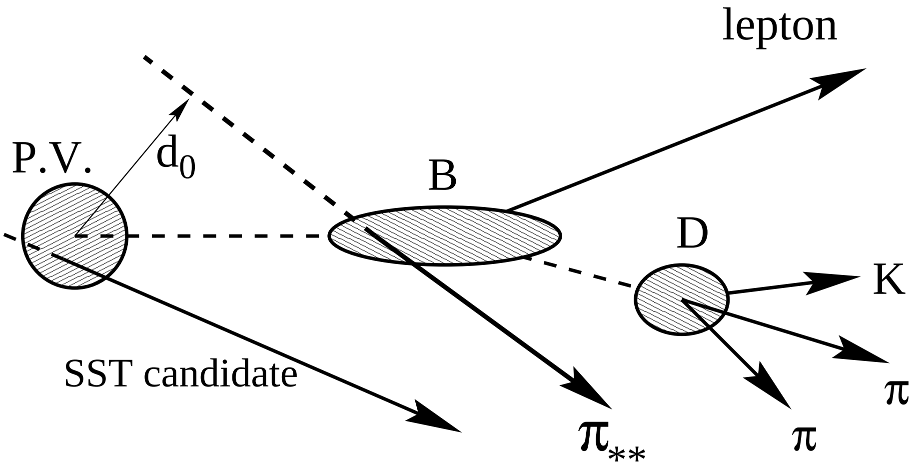


Figure 7-2: A schematic representation of the D^{**} decay. The SST candidate tracks originate from the primary vertex (P.V.), while the π_{**} track originates from the B meson decay vertex. When the B and the primary vertex are well separated, the π_{**} track usually has large impact parameter with respect to the primary vertex, d_0.

a reasonable precision.[2]

To measure $N(\pi_{**}\text{tags})$, we consider the d/σ distribution with respect to the B vertex (denoted by d_B/σ_B) of the tagging tracks. A schematic representation of this measurement is shown in fig. 7-3. π_{**}'s are emerging from the B vertex and always have the right correlation with the lepton. This means the 'right correlation' distribution of d_B/σ_B consists of a Gaussian centered at 0 with unit width[3] and the distribution of d_B/σ_B for the tracks that originated from the primary vertex. One can get the shape of the d_B/σ_B distribution for primary tracks by using the tags with the 'wrong correlation', since among them there are no π_{**}'s and therefore they represent pure combinatorial background. We fit the 'right correlation' distribution with the Gaussian centered at 0 on top of the 'wrong correlation' distribution, and the area

---

[2] A cut on d_0/σ_{d_0}, if applied in this study, would eliminate most of the π_{**} tags, and therefore deteriorate the statistical precision of the $R^{**\,(meas)}$ measurement.

[3] σ_B includes both the error on the track impact parameter and the contribution of the B vertex covariance matrix.

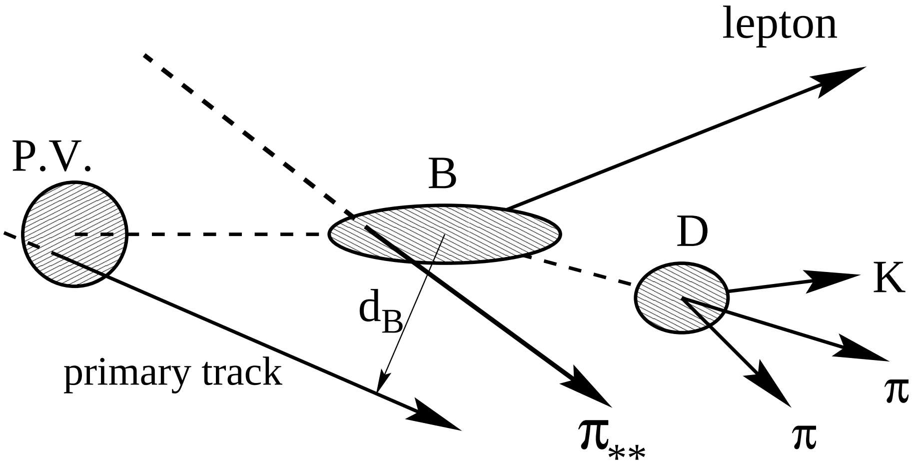


Figure 7-3: A schematic representation of the measurement of $N(\pi_{**})$, the number of π_{**}'s selected as tags. The SST candidate tracks (RS and WS) originate from the primary vertex (P.V.), and thus can have very large values of the impact parameter significance with respect to the B decay vertex, d_B/σ_{d_B}. On the other hand, the π_{**} track originates from the B vertex, and is characterized by $d_B \approx 0$.

under the Gaussian yields $N(\pi_{**}\text{tags})$. Figure 7-4 gives an example of the d_B/σ_{d_B} shapes for the π_{**} (the narrow gaussian) and the tracks originating from the primary vertex (wide gaussians, the higher one for RS and the lower for WS).

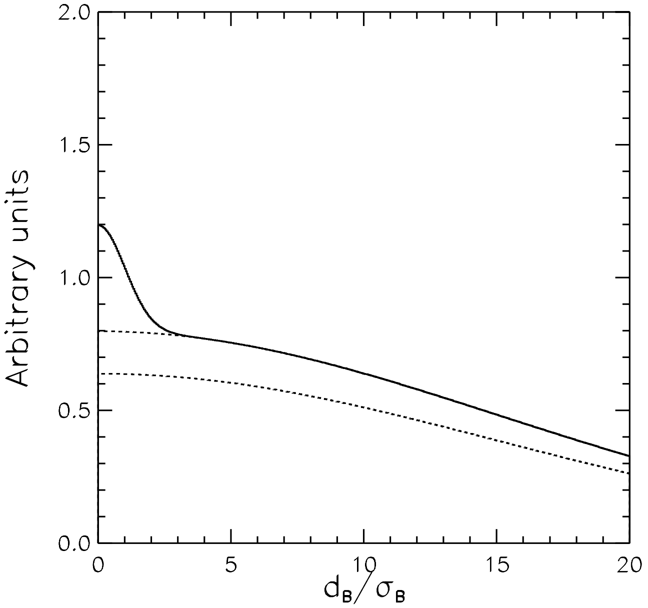


Figure 7-4: A schematic example of the d_B/σ_{d_B} shapes for the π_{**} (the narrow gaussian) and the tracks originating from the primary vertex (wide gaussians, the higher one for RS and the lower one for WS). In the actual measurement of $N(\pi_{**})$, we model the RS background with the scaled shape of the WS distribution.

Figures 7-5, 7-6, 7-7, 7-8 and 7-9 show how we derive the average ξ from a decay signature (the case of $\ell^+\overline{D^0}$ is shown). What happens in the fit can be interpreted as the normalization of the 'wrong correlation' to the 'right correlation' distribution in the range where $d_B/\sigma_B \gg 0$, and one can see a clear excess in the 'right correlation' distribution at $d_B/\sigma_B \to 0$. Table 7.1 shows the values for $R^{**\,(meas)}$ measured in the data.

In order to derive the prediction for R^{**}, we observe that the sum of the normalized second and fourth term in equation (7.18) gives the fraction of B decays where π_{**}'s

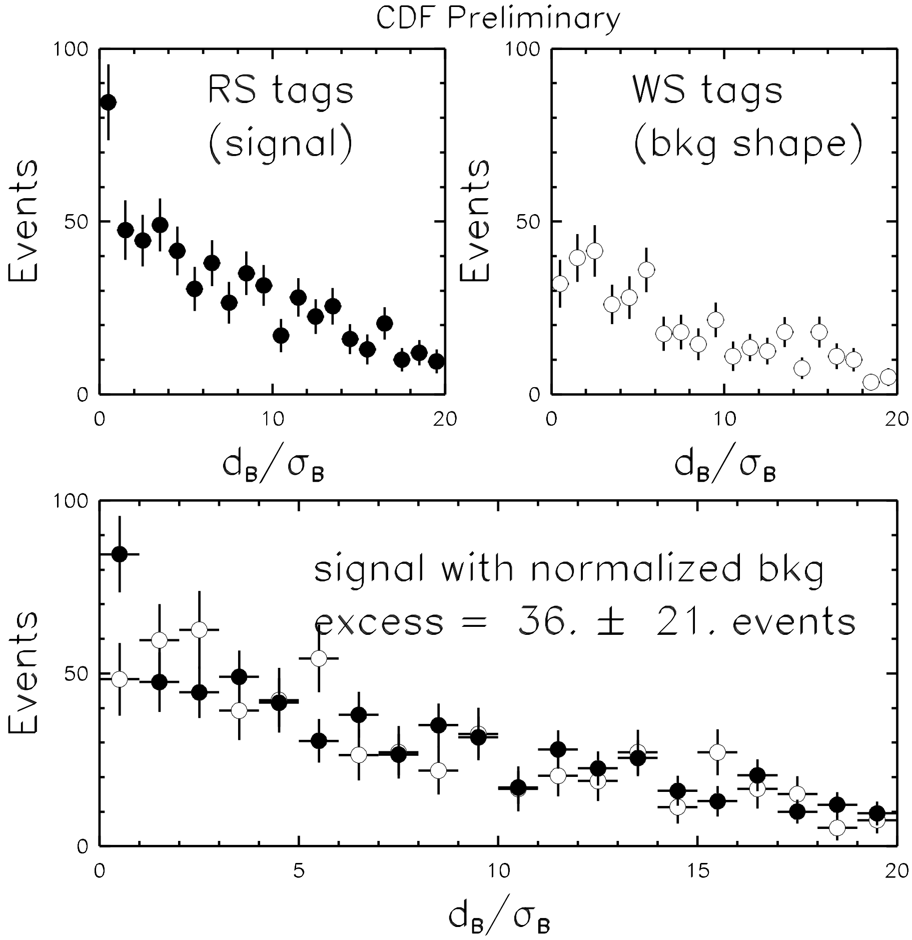


Figure 7-5: Impact parameter distributions w.r.t. the B vertex, $\ell^+\overline{D^0}, \overline{D^0} \rightarrow K\pi$ decay signature, with D-meson sideband subtraction. Top left: right sign tags. The signal + background. Top right: all wrong sign tracks. The background shape. Bottom: we fit the right sign tags with the Gaussian centered at 0 and the histogram for the background shape; from the fit we derive the normalization factor for the background shape. In the bottom figure the signal and the normalized background are overlaid.

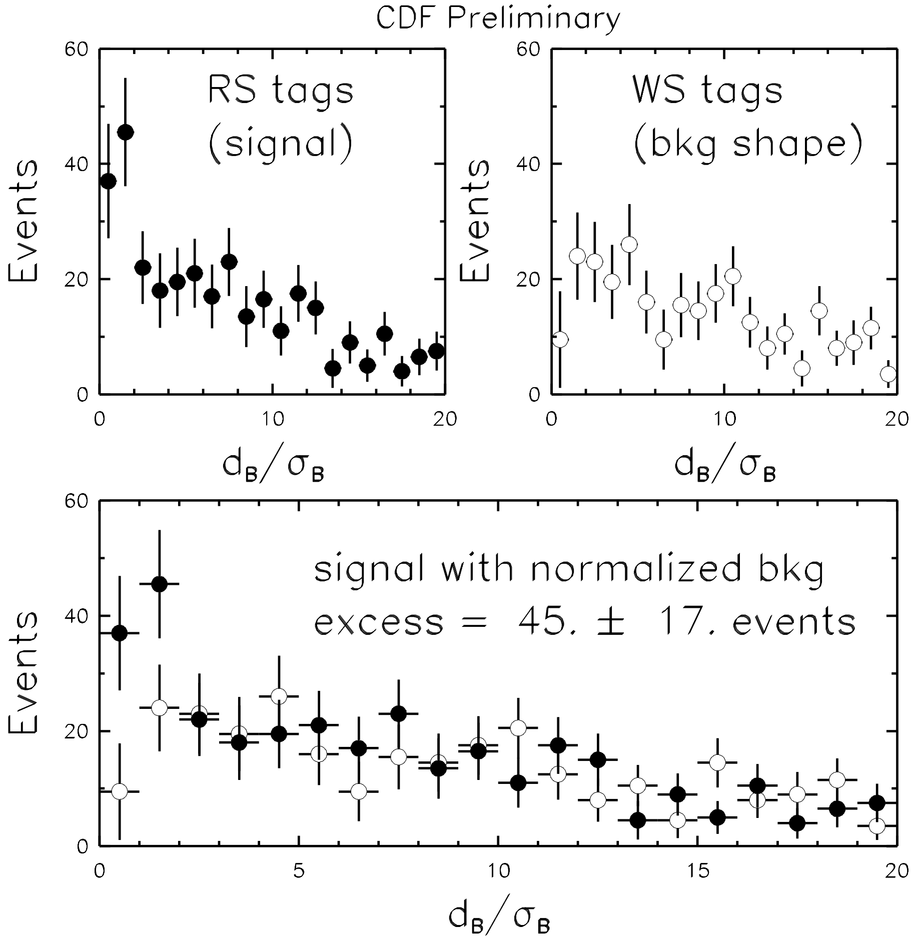


Figure 7-6: Impact parameter distributions w.r.t. the B vertex, $\ell^+ D^-, D^- \rightarrow K\pi\pi$ decay signature, with D-meson sideband subtraction. Top left: right sign tags. The signal + background. Top right: all wrong sign tracks. The background shape. Bottom: we fit the right sign tags with the Gaussian centered at 0 and the histogram for the background shape; from the fit we derive the normalization factor for the background shape. In the bottom figure the signal and the normalized background are overlaid.

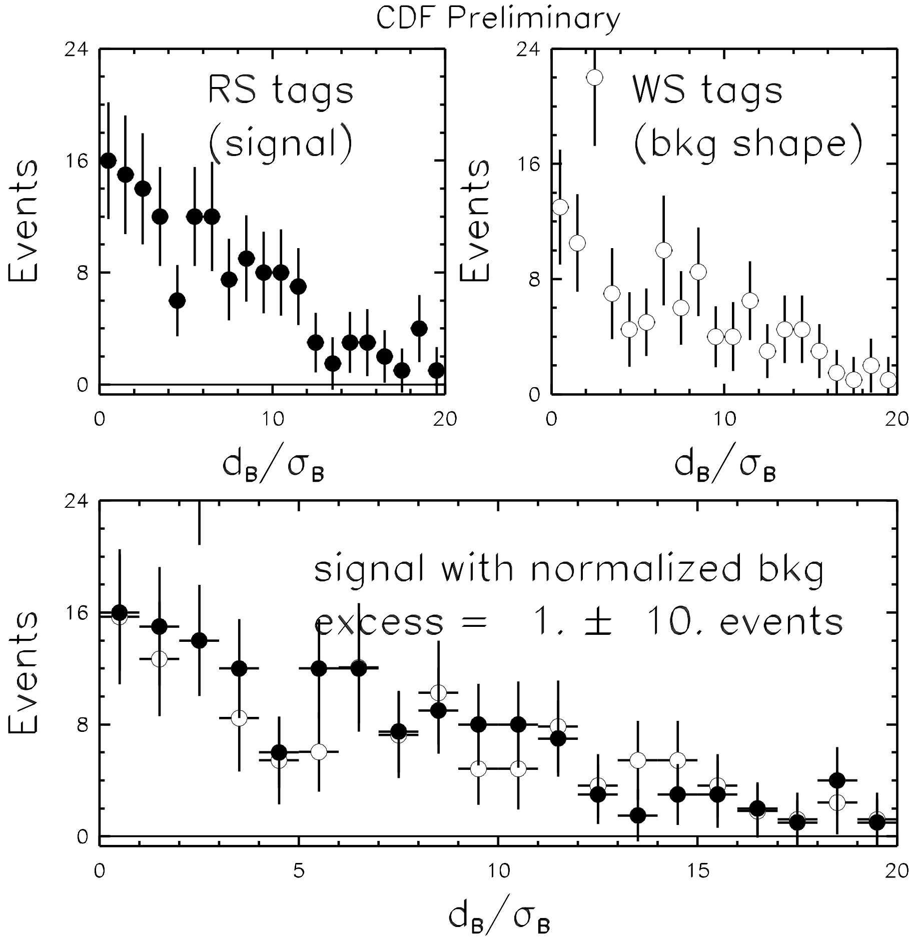


Figure 7-7: Impact parameter distributions w.r.t. the B vertex, $\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi$ decay signature, with D-meson sideband subtraction. Top left: right sign tags. The signal + background. Top right: all wrong sign tracks. The background shape. Bottom: we fit the right sign tags with the Gaussian centered at 0 and the histogram for the background shape; from the fit we derive the normalization factor for the background shape. In the bottom figure the signal and the normalized background are overlaid.

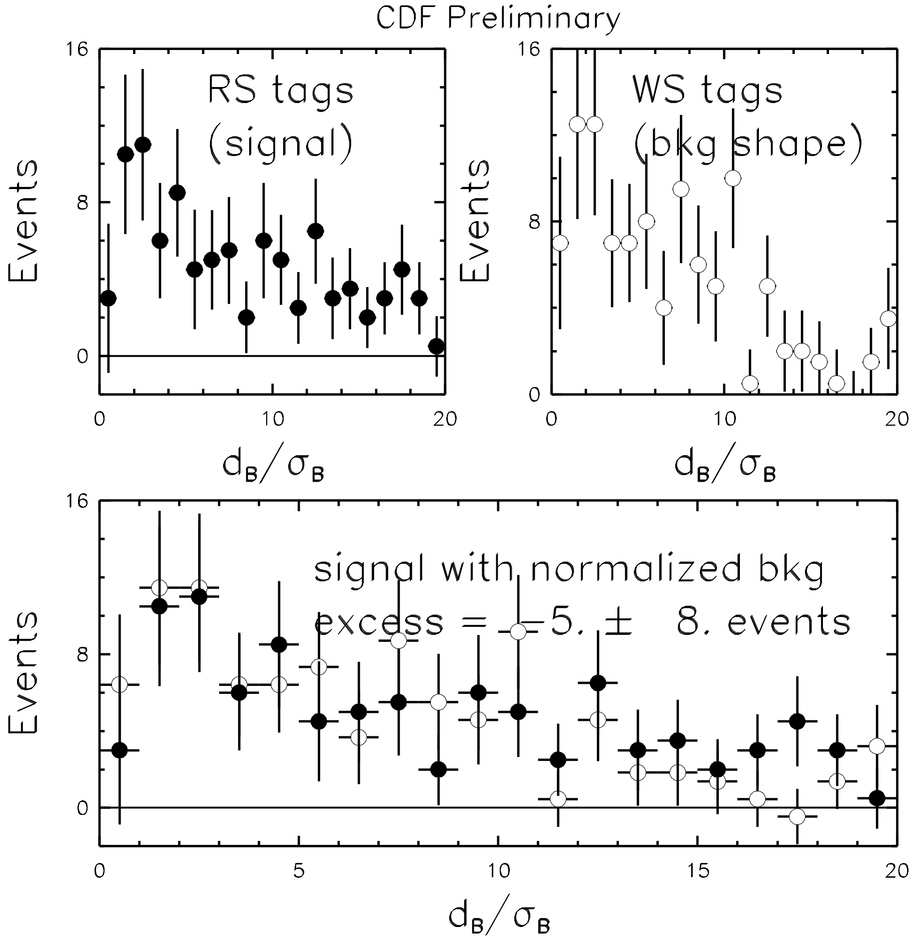


Figure 7-8: Impact parameter distributions w.r.t. the B vertex, $\ell^+ D^{*-}, \overline{D^0} \to K3\pi$ decay signature, with D-meson sideband subtraction. Top left: right sign tags. The signal + background. Top right: all wrong sign tracks. The background shape. Bottom: we fit the right sign tags with the Gaussian centered at 0 and the histogram for the background shape; from the fit we derive the normalization factor for the background shape. In the bottom figure the signal and the normalized background are overlaid.

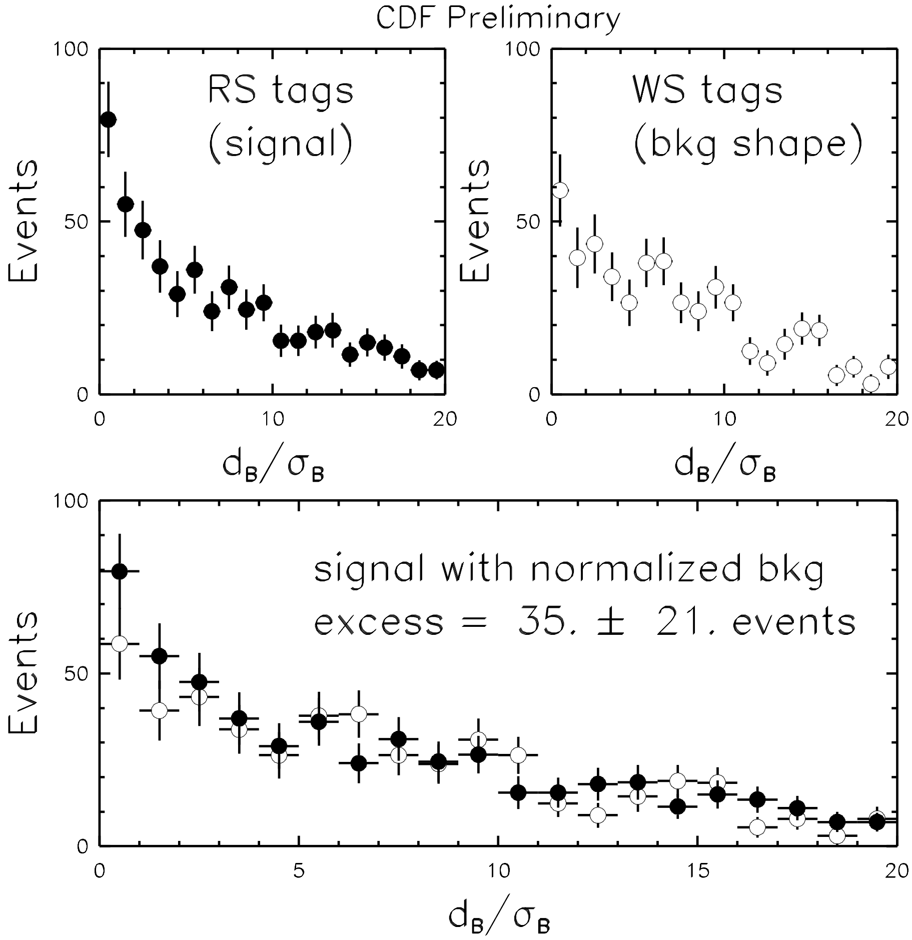


Figure 7-9: Impact parameter distributions w.r.t. the B vertex, $\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$ decay signature, with D-meson sideband subtraction. Top left: right sign tags. The signal + background. Top right: all wrong sign tracks. The background shape. Bottom: we fit the right sign tags with the Gaussian centered at 0 and the histogram for the background shape; from the fit we derive the normalization factor for the background shape. In the bottom figure the signal and the normalized background are overlaid.

Decay signature	$R^{**\,(meas)}$
$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$	0.029 ± 0.018
$\ell^+D^-, D^- \to K\pi\pi$	0.056 ± 0.022
$\ell^+D^{*-}, \overline{D^0} \to K\pi$	0.003 ± 0.029
$\ell^+D^{*-}, \overline{D^0} \to K3\pi$	-0.016 ± 0.026
$\ell^+D^{*-}, \overline{D^0} \to K\pi\pi^0$	0.034 ± 0.021

Table 7.1: The fractions of tags identified as π_{**} candidates, R^{**}, measured in the five decay signatures.

are selected as tags. Thus

$$\bar{R}^{**} = \alpha_k^{**,0} + \alpha_k^{**,+}$$

except that, in this case, the coefficients α_k^{**} are integrated over ct. From equations (7.15) and (7.17), we see that $\bar{R}^{**} \sim \xi_{norm}$

The χ^2 is expanded by five more terms:

$$\chi^2 \to \chi^2 + \sum_k \left(\frac{R_k^{**\,(meas)} - \bar{R}_k^{**}}{\sigma_k^{**}} \right)^2 \tag{7.23}$$

Minimizing the χ^2 now also results in ξ_{norm}, the scale factor for ξ_{MC}.

While ξ_{norm} is essentially determined by the average of all five $R_k^{**\,(meas)}$, P_V is determined by their relative values. P_V tells us how many D^{**} mesons decay into D^* and how many into D. If $P_V = 0$, there would be no $D^{**} \to D^*$ decays, and consequently no $\pi_{**}^{\pm}$'s in the ℓ^+D^{*-} signatures, resulting in $R^{**\,(meas)}$ consistent with zero. Therefore, the relative magnitude of $R^{**\,(meas)}$ between the ℓD and ℓD^* signatures constrains P_V. We note that the errors on $R_k^{**\,(meas)}$ are large, and therefore the error on P_V will also be large. Nevertheless, this method is still better than using just a theoretical estimate for P_V.

7.3 The result of the fit

In this section the result of fit for $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d, is finally presented. In addition, we also fit for $\epsilon(\pi_s)$, ξ_{norm} and P_V, as well as for the optimal values of other sample composition parameters.

7.3.1 The complete χ^2 function

In order to explicitly state the χ^2 function used in the fit for $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d, we start from eq. (7.5) and add the terms for the measurement of $\epsilon(\pi_s)$, eq. (7.20), and the measurements constraining ξ_{norm} and P_V, eq. (7.23), as well as the terms for the sample composition parameters f^*, R_f, τ_{B^+}/τ_{B^0}, and the B^0 lifetime, τ_{B^0}:

$$
\begin{aligned}
\chi^2 &= \sum_{k,ct} \left(\frac{\mathcal{A}_k^{(meas)}(ct) - \bar{\mathcal{A}}_k(ct)}{\sigma_k^{\mathcal{A}}(ct)} \right)^2 \\
&+ \left(\frac{R^{*\,(meas)} - \bar{R}^*}{\sigma^*} \right)^2 + \sum_k \left(\frac{R_k^{**\,(meas)} - \bar{R}_k^{**}}{\sigma_k^{**}} \right)^2 \\
&+ \left(\frac{f^{**\,(meas)} - f^{**}}{\sigma_{f^{**}}} \right)^2 + \left(\frac{R_f^{(meas)} - R_f}{\sigma_{R_f}} \right)^2 \\
&+ \left(\frac{(\tau_{B^+}/\tau_{B^0})^{(meas)} - (\tau_{B^+}/\tau_{B^0})}{\sigma_{(\tau_{B^+}/\tau_{B^0})}} \right)^2 + \left(\frac{(\tau_{B^0})^{(meas)} - (\tau_{B^0})}{\sigma_{(\tau_{B^0})}} \right)^2 \qquad (7.24)
\end{aligned}
$$

The asymmetry predictions $\bar{\mathcal{A}}_k(ct)$ are calculated in eq. (7.18), the prediction for R^* in eq. (7.20) and the predictions for R_k^{**} in eq. (7.23).

7.3.2 The behavior of the fit parameters

The input and the output fit parameters are shown in table 7.2. The fit where all of them, in addition to $\mathcal{D}_+$, $\mathcal{D}_0$, Δm_d, $\epsilon(\pi_s)$, ξ_{norm} and P_V float in the fit we call the 'ten-variable fit', by the number of parameters that are floating (to distinguish it from the 'five-variable fit' in which all the sample composition parameters are fixed; the latter is used in determination of the systematic uncertainties, as described in section 8.1). In table 7.2 we see that the B^0 meson lifetime did not change in the fit, implying that in practice this parameter is completely decoupled from other parameters in the fit.

parameter	input	inp. error	output	out + error	out. − error
f^{**}	0.360	0.120	0.309	0.101	−0.100
R_f	2.50	0.60	2.51	0.60	−0.59
τ_{B^+}/τ_{B^0}	1.020	0.050	1.021	0.049	−0.049
τ_{B^0}	0.0468	0.0018	0.0468	0.0018	−0.0018

Table 7.2: The input and output values of the fit parameters.

The fit results in the following values for dilutions, Δm_d, and auxiliary quantities $\epsilon(\pi_s)$, ξ_{norm} and P_V:

$$\begin{aligned}
\mathcal{D}_0 &= 0.181^{+0.036}_{-0.031} \\
\mathcal{D}_+ &= 0.267^{+0.039}_{-0.034} \\
\Delta m_d &= 0.471^{+0.081}_{-0.063}\ \mathrm{ps}^{-1} \\
\epsilon(\pi_s) &= 0.845^{+0.073}_{-0.058} \\
\xi_{norm} &= 0.747^{+0.470}_{-0.292} \\
P_V &= 0.331^{+0.276}_{-0.298}
\end{aligned}$$

However the errors reported in this fit are the combination of the statistical errors and the systematic uncertainties due to the sample composition. The way the statistical errors are extracted from this combined error will be described in detail in section 8.1. Figure 7-10 shows the result of the fit overlaid on top of the measured asymmetries, where all three $\ell^+ D^{*-}$ signatures are combined. Figure 7-11 gives the three $\ell^+ D^{*-}$ signatures separately. The correlation coefficients of the fit parameters with $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d are shown in table 7.3. Our conjecture that τ_{B^0} is decoupled from other parameters is corroborated, since the correlation coefficients with τ_{B^0} are almost identical to zero. On the other hand, f^{**}, ξ_{norm} and P_V are fairly strongly coupled to $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d, underlying the importance of the π_{**} corrections to this analysis.

The final sample composition is: $\sim$ 82% of the $\ell^+ \bar{D}^0 X$ signature comes from B^+ decays, while $\sim$ 80% of the $\ell^+ D^- X$ and $\sim$ 95% of the $\ell^+ D^{*-} X$ originate from B^0.

Having obtained the values and the statistical uncertainties for $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d, we address the sources of systematic uncertainty in the following chapter.

parameter	$\mathcal{D}_0$	$\mathcal{D}_+$	Δm_d
$\mathcal{D}_0$	1.000	0.372	-0.172
$\mathcal{D}_+$	0.372	1.000	-0.372
Δm_d	-0.172	-0.372	1.000
R_f	0.007	0.126	-0.020
f^{**}	0.504	0.406	-0.385
P_V	-0.310	-0.284	-0.326
$\epsilon(\pi_s)$	0.100	-0.082	-0.031
ξ_{norm}	-0.445	-0.355	0.304
τ_{B0}	-0.001	0.002	-0.005
τ_{B+}/τ_{B0}	0.009	-0.157	-0.051

Table 7.3: Fit parameter correlation coefficients with $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d.

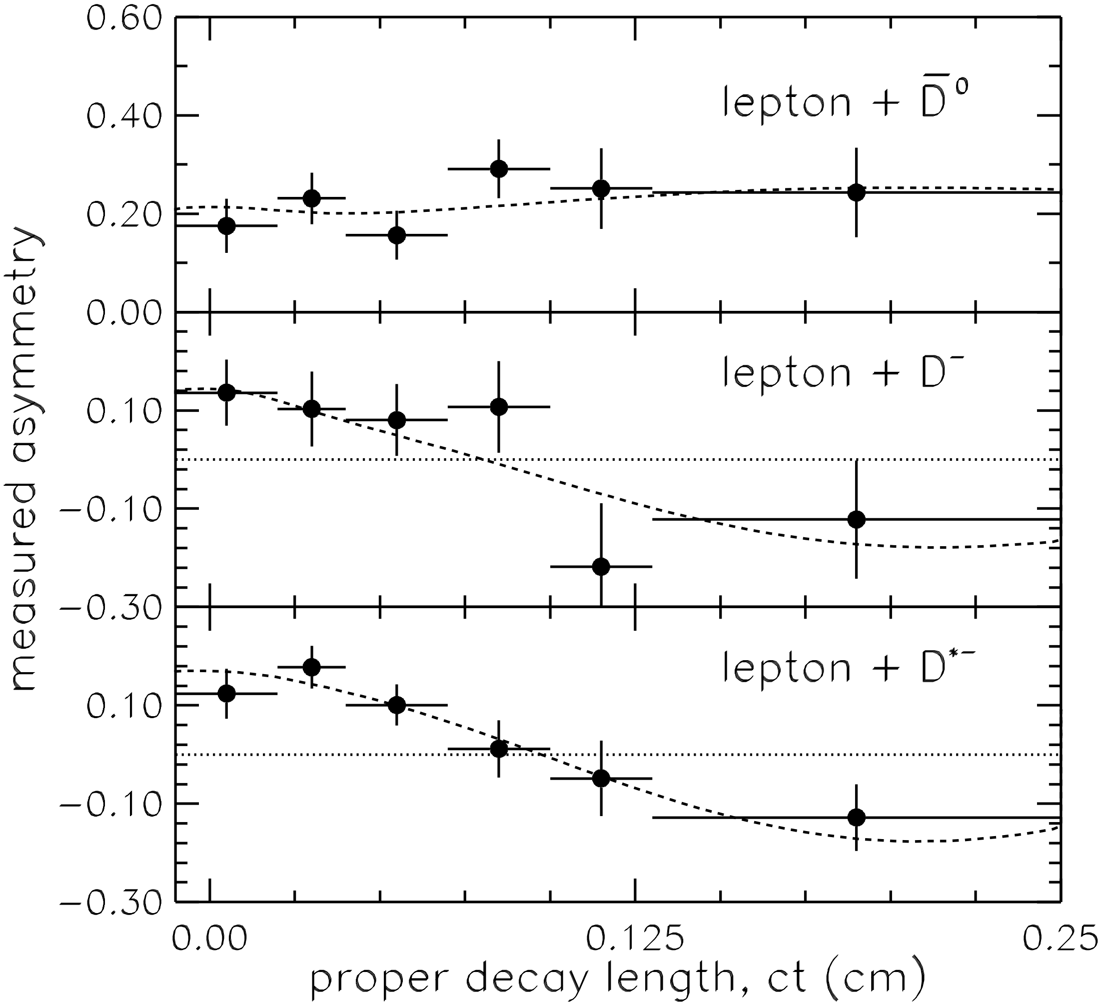


Figure 7-10: Measured asymmetries as a function of the proper decay length, *ct*, for the decay signatures: $\ell^+\bar{D}^0$ (dominated by B^+), ℓ^+D^- and the sum of all three ℓ^+D^{*-} (dominated by B^0). We fit the three ℓ^+D^{*-} signature separately, but combine them for display purposes. The dashed line is the result of the fit.

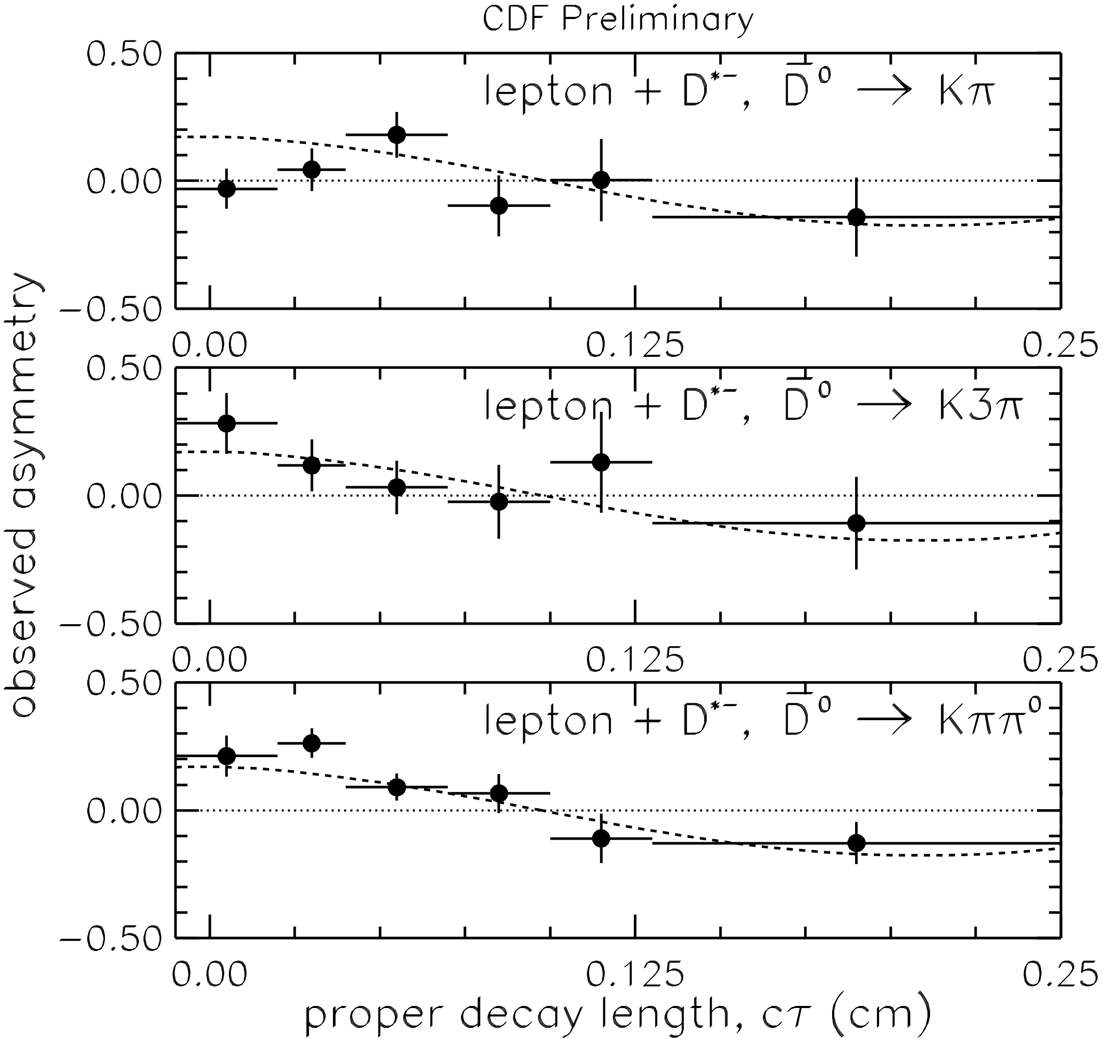


Figure 7-11: The breakdown of $\ell^+ D^{*-}$ of the measured asymmetries into the three $\ell^+ D^{*-}$ decay signatures "$\ell^+ D^{*-}, \overline{D^0} \to K\pi$", "$\ell^+ D^{*-}, \overline{D^0} \to K3\pi$" and "$\ell^+ D^-, D^- \to K\pi\pi$". The result of the fit is overlaid.

Chapter 8

Systematic uncertainties

In the preceding chapter we have built a χ^2 function that incorporates all relevant sample composition effects and allows us to fit for $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d, and applied it to the lepton + charm sample. In this chapter we turn to the study of the systematic uncertainties. The goal behind fitting the observed asymmetries was to let all sample composition parameters float in the fit, in order to simultaneously derive both $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d, and the sample composition preferred by the data. However, this method also enables us to estimate the contribution of the sample composition parameters to the systematic uncertainty on $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d.

In the measurement of $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d, the sources of the systematic uncertainty can be divided into three categories:

- **Correlated:** Parameters of the fit that are coupled to $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d through the sample composition equation (7.18). The parameters that fall into this category are the sample composition parameters: f^{**}, R_f, τ_{B+}/τ_{B0} and τ_{B0}(described in detail in section 5.1.2). These parameters are not correlated among themselves, only the effects of their changes on $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d are.

- **Uncorrelated:** Systematic uncertainties that are caused by imperfect models of the detector, either through the corrections derived from the Monte Carlo simulation (e.g. the $\mathcal{K}$-factor distributions, L_{xy} and ct resolutions), or through errors on the corrections applied to the data (e.g. the detector alignment), as well as by the imperfect Monte Carlo model (e.g. of the B meson decay).

- **Physics backgrounds:** The uncertainties due to other physics processes that contribute to $B \to \nu \ell D^{(*)}$ data sample that have been hitherto neglected. Ideally, these processes would be modeled by additional sample composition equations, and they would be corrected for in the fit. However, as the contribution of these decay chains is very small, we chose not to do that, and just treat them as another kind of uncorrelated systematic uncertainty.

In the following sections we examine these three categories in detail, and determine their contribution to the systematic uncertainty on $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d.

8.1 Correlated systematic uncertainties

Since the contributions of the sample composition parameters to the total systematic uncertainty are correlated through the asymmetry correction equations (7.18), it is not correct just to move the value of each of the parameters by $\pm\sigma$, fix it and refit to derive the shifts in $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d, and then just sum these shifts in quadrature. This method does allow for the situation where the change in the result due to movement of one parameter is offset (or amplified) by simultaneous change of another parameter.

The proper way of taking into account these correlations is to let the sample composition parameters freely float in the fit. Since there are five sample composition parameters in addition to $\mathcal{D}_+$, $\mathcal{D}_0$, Δm_d, $\epsilon(\pi_s)$ and ξ_{norm}, we call this fit the 'ten-variable fit' (ten variables are floating). Equation (7.24) gives the complete χ^2 function. The sample composition parameters (f^{**}, R_f, and τ_{B^+}/τ_{B^0}), as well as τ_{B0}, are constrained to the measured values by their own χ^2 terms.

The errors on $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d that are returned by this fit contain not only the statistical error, but also the contribution of the uncertainties on the sample composition parameters. Let us denote this error as $\sigma_{stat+S.C.}$. The statistical error, σ_{stat}, is derived from the fit where the sample composition parameters are fixed, and only five variables ($\mathcal{D}_+$, $\mathcal{D}_0$, Δm_d, $\epsilon(\pi_s)$ and ξ_{norm}) are floating. For this reason, we call this fit the 'four-variable' fit. The systematic uncertainty due only to the sample composition is then estimated by subtraction in quadrature:

$$\sigma_{S.C.} \equiv \sqrt{\sigma^2_{stat+S.C.} - \sigma^2_{stat}} \tag{8.1}$$

Of course, if this subtraction is to be meaningful, both the ten-variable and the four-variable fits must converge to the same parameter values. Therefore we first perform the ten-variable fit (table 8.1), fix the sample composition parameters to their output values, and then run the four-variable fit (table 8.2) to get the statistical error. The contribution of the sample composition to the systematic uncertainty, estimated by subtracting the columns from tables 8.1 and 8.2 in quadrature, is shown in table 8.3. For comparison, the statistical component of the fit error is also given.

name	value	+error	-error
$\mathcal{D}_0$	0.1810	0.0363	−0.0310
$\mathcal{D}_+$	0.2666	0.0388	−0.0344
Δm_d	0.4714	0.0835	−0.0751
$\epsilon(\pi_s)$	0.8448	0.0812	−0.0632
ξ_{norm}	0.7769	0.4702	−0.2923
P_V	0.3310	0.2761	−0.2984
R_f	2.5128	0.5958	−0.5925
f^{**}	0.3095	0.1007	−0.0998
τ_{B^0}	0.0468	0.0018	−0.0018
τ_{B^+}/τ_{B^0}	1.0208	0.0495	−0.0494

Table 8.1: The result of the 'ten-variable fit'.

name	value	+error	−error
$\mathcal{D}_0$	0.1810	0.0285	−0.0281
$\mathcal{D}_+$	0.2666	0.0322	−0.0318
Δm_d	0.4714	0.0781	−0.0684
$\epsilon(\pi_s)$	0.8448	0.0511	−0.0461
ξ_{norm}	0.7769	0.2458	−0.2428
P_V	0.3310	0.2702	−0.2797

Table 8.2: The result of the 'four-variable fit'.

name	value	$+\sigma_{stat}$	$-\sigma_{stat}$	$+\sigma_{S.C.}$	$-\sigma_{S.C.}$
$\mathcal{D}_0$	0.1810	+0.0285	−0.0281	+0.0216	−0.0131
$\mathcal{D}_+$	0.2666	+0.0322	−0.0318	+0.0225	−0.0131
Δm_d	0.4714	+0.0781	−0.0684	+0.0295	−0.0310

Table 8.3: The contribution of the sample composition to the systematic uncertainty, estimated by subtracting the columns from tables 8.1 and 8.2 in quadrature. For comparison, the statistical component of the fit error is also shown.

8.2 Uncorrelated systematics

In this section, we study the ‘uncorrelated’ systematic uncertainty. We first list possible sources, and then evaluate them one by one. Some types of the ‘uncorrelated’ systematic uncertainty that have been considered in other analyses (e.g. the measurement of the lifetime of the B meson [54]), turn out to be irrelevant here, and we do not take them into account.

The following are the possible sources of the ‘uncorrelated’ systematic uncertainty:

- ***ct*-dependent B meson selection criteria.** The requirement that the D-meson decay is well separated from the primary vertex (L_{xy}^D/σ^D is greater than at least 1.0) coupled with the impact parameter significance requirements (d/σ) on the D-meson daughter tracks (section 4.1) translates into the *ct*-dependent B-meson candidate selection. For this reason, the *ct* distribution is biased by the selection requirements, and a B meson lifetime measurement must correct for it. However, since the asymmetry (defined in equation (2.15)) depends only on the *ratio* of the RS to WS B-meson candidates, the asymmetry itself is not biased, and thus the *ct*-dependent B meson selection criteria do not cause any systematic effect.

- **$\mathcal{K}$-factor distribution.** As explained in section 5.2, the $\mathcal{K}$-factors are an essential ingredient of the measurement of the proper time of each B decay. The $\mathcal{K}$-factors determine the *ct* scale, and therefore directly influence the error on Δm_d. Therefore, in a mixing measurement, the quality of the $\mathcal{K}$-factor distribution deserved a special attention. The $\mathcal{K}$-factor distribution may be incorrect for the following reasons:

  - ***b*-quark spectrum.** If the p_T spectrum of the b-quark used in the Monte Carlo generation is different from the data, the B-meson p_T distribution will be biased, and thus the average observed momentum (that is, the $\mathcal{K}$), will also be biased.

  - **Trigger model.** As explained in the description of the Monte Carlo (section 4.1) we match the Monte Carlo lepton p_T spectrum to the data, so there is no bias.

  - **Electron isolation requirement.** Even if the p_T spectrum of the $\ell D^{(*)}$ system agrees between the Monte Carlo and the data, an inaccurate Monte Carlo simulation of the electron trigger isolation requirement would prefer some B decays where the D daughters are emitted away of the lepton, and thus would select B decays of a particular kinematics, resulting in an incorrect estimate of $\mathcal{K}$.

  - **B decay model.** Similar to the electron isolation requirement, however here the B decay is incorrect merely because of the wrong model of the B decay, rather than because of requiring an isolated electron.

- **Detector alignment.** Significant detector misalignment would result in a bias in ct. However, [54] shows that the effect of misalignment, compared to other sources of the systematic uncertainty, is negligible.

- **Shape of the combinatorial background.** When measuring the asymmetry in each bin in ct, $\mathcal{A}_k(ct)$, we simultaneously fit the RS and WS D-meson mass distributions, and using only the signal component to calculate $\mathcal{A}_k(ct)$. Therefore, the combinatorial background only influences the error on the asymmetry, not the asymmetry itself.

- $L_{\rm xy}$ **resolution.** We fit for Δm_d using the cosine convoluted with the resolution function. If the Monte Carlo model of the residual tracking resolution (defined in sections 4.3) is incorrect, the fit for Δm_d will be biased.

- $\xi_{MC}(ct)$ **shape.** In section 7.2.3, $\xi(ct)$ was factorized into $\xi_{norm} \cdot \xi_{MC}(ct)$. ξ_{norm} is fitted from the data, and thus contributes to the statistical uncertainty through its correlations with $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d. However, $\xi_{MC}(ct)$ is derived from the Monte Carlo simulation, and thus may carry an inherent systematic bias, which must be estimated and included into the systematic uncertainty on $\mathcal{D}_+$, $\mathcal{D}_0$and Δm_d.

In the following subsections, these effects are described in detail, and their contribution to the systematic uncertainty is estimated. It should be noted that the contribution of each of these is rather small, and is almost negligible in comparison to the contribution of the sample composition parameters.

In evaluating the contribution of each of the sources of the 'uncorrelated' systematic uncertainty, we follow this simple method: we establish a 'typical' change in conditions (e.g. by changing a fit parameter, or generating a new Monte Carlo sample using different parameters) for the source of uncertainty in question, and evaluate the shift in $\mathcal{D}_+$, $\mathcal{D}_0$ or $\Delta m_d$from their 'nominal' values from the 'ten-variable' fit (table 8.1). The shifts, denoted by $\Delta\mathcal{D}_+$, $\Delta\mathcal{D}_0$ and $\Delta(\Delta m_d)$, give us estimates of the uncertainties $\sigma(\mathcal{D}_+)$, $\sigma(\mathcal{D}_0)$ and $\sigma(\Delta m_d)$ due to the source of 'uncorrelated' systematic uncertainty that is being considered. (However, since the finite statistics of the Monte Carlo simulation can adversely affect $\epsilon(\pi_s)$ and ξ_{norm}, which would in turn amplify the influence of the systematic effect in question on $\mathcal{D}_+$, $\mathcal{D}_0$ or Δm_d, in the following studies we fix $\epsilon(\pi_s)$ and ξ_{norm} to their values of 0.835 and 0.746, from the 'ten-variable' fit.) The 'uncorrelated' systematics, together with the systematic uncertainties due to sample composition and physics backgrounds, are summarized in table 8.8.

8.2.1 b-quark spectrum

The Monte Carlo simulation used for determination of sample composition fit constants ($\beta\gamma$ corrections for direct decay sequences, $\epsilon^D_{k\ell}$, $K_{k\ell}$, $\sigma_{k\ell}$), described in section 4.2.1, uses the the inclusive b-quark p_T spectrum calculated by Nason, Dawson and Ellis [30] If that spectrum is not correct, the Monte Carlo simulation yields biased K-factor distributions.

In this study we follow the prescription of the CDF measurement of the B-meson lifetime [54]. It has been known in CDF [56] that an additional weight is needed in order to match the lepton p_T spectrum in the data. (Here, we are concerned with the high-p_T end of the lepton spectrum, which is not affected by the trigger turn-on.) All events in the Monte Carlo simulation are weighted by a power function of the b-quark

p_T,

$$\frac{dN}{dp_T^b} \to \frac{dN}{dp_T^b}(p_T^b)^{-A}$$

The constant A has been measured from the data and is $A = 0.8 \pm 0.1$ [56].

After the events in the nominal Monte Carlo simulation have been weighted down using the above method, we rederive the trigger turn-on, apply it to the weighted Monte Carlo simulation, again obtain sample composition fit constants and use it in the tagging and in the fitter. The resulting dilutions and Δm_d are then compared with their nominal values, derived just using the NDE b-quark spectrum. We get $\Delta\mathcal{D}_+ = 0.0052$, $\Delta\mathcal{D}_0 = 0.0017$ and $\Delta(\Delta m_d) = -0.0060 ps^{-1}$.

8.2.2 Electron isolation requirement

The inclusive electron trigger requires a cluster in the electromagnetic calorimeter without a matching cluster in the hadronic calorimeter.Thus the electrons have an implicit isolation requirement – the electrons with hadrons close in ΔR are not recognized as electrons and are not triggered on. Let us assume that the $p_T(\ell D)$ distribution from the Monte Carlo simulation (after applying the trigger turn-on) matches the $p_T(\ell D)$ distribution in the data. If the electron isolation requirement is not simulated well, the resulting distribution of hadrons around the electron – possibly including the D meson decay products – is biased. And since the $p_T(\ell D)$ distribution agrees between the Monte Carlo simulation and the data, our estimate of the energy taken away by the neutrino is incorrect. That is, the K-factor distribution is also biased.

In order to estimate the bias, we derive all sample composition fit constants ($\beta\gamma$ corrections for direct decay sequences, $\epsilon_{k\ell}^D$, $K_{k\ell}$, $\sigma_{k\ell}$) for the electron and muon Monte Carlo simulation separately. We then apply them to the data (in the tagging and in the fit). The muons do not have any isolation requirement, so the difference between the fit results when the muon sample composition fit constants are used versus their electron counterparts gives us an estimate of the electron trigger bias in the detector simulation. The shifts ('electrons – muons') are $\Delta\mathcal{D}_+ = 0.0036$, $\Delta\mathcal{D}_0 = 0.0047$ and $\Delta(\Delta m_d) = -0.0045 ps^{-1}$, and we use them as the systematic uncertainties. They are much smaller than both the statistical and the correlated systematic errors.

8.2.3 B decay model

As we have seen, the relative charm efficiencies $\epsilon^D_{k\ell}$ and the relative K factors (K_ℓ/K_d) take into account the difference in the decay dynamics of the processes $B \to \nu\ell D$, $B \to \nu\ell D^*$ and $B \to \nu\ell D^{**}$. Therefore, the model of the semileptonic B decay used in the Monte Carlo simulation is potentially a significant source of systematic uncertainty.

To estimate this effect, we modify the Monte Carlo simulation to decay B mesons according to phase space. We generate another Monte Carlo sample with this characteristic, and then use it to derive $\epsilon^D_{k\ell}$ K_ℓ/K_d and σ_ℓ/σ_d for all decay chains, as well as the $\beta\gamma$ correction (K-factor) for each decay signature (K_{kd}). As an illustration, for ℓD^+ these numbers are shown in table 8.4. In this situation, there is no preference for ℓD^*, as all decays obey exactly the same dynamics – the only difference being the larger $D^{(*)}$ meson mass. And indeed, in 8.4 we see that all charm reconstruction efficiencies relative to the direct signature are less than one, as compared to the nominal values as given in tables 5.7 and 5.12.

Using the above numbers and the similar values derived for the other four decay signatures, we redo the tagging, run the fitter, keeping all parameters (except the dilutions and Δm_d) fixed to their central values. The fit converges to $\mathcal{D}_+ = 0.2722$, $\mathcal{D}_0 = 0.1921$ and $\Delta m_d = 0.4582 ps^{-1}$. resulting in shifts of $\Delta\mathcal{D}_+ = 0.0005$, $\Delta\mathcal{D}_0 = 0.0045$ and $\Delta(\Delta m_d) = -0.0115$. Only the last one is noticeable, but still smaller than the systematic uncertainties due to the sample composition.

8.2.4 L_{xy} resolution

As explained in section 4.3, for each of the five decay signatures, the resolution on the B-meson proper time ($\sigma(ct)$) is parameterized as a linear function of the residual detector resolution (σ_0) and the B-meson proper time (ct), $\sigma(ct) = \sigma_0 + b \times ct$ (figure 4-7 and table 4.5). The second term describes the contribution of the $\beta\gamma$ correction (the

Decay chain	$\epsilon^{D}_{k\ell}$	K_ℓ/K_d	$\mathrm{RMS}_\ell/\mathrm{RMS}_d$
100.002	1.00 ± 0.03	1.000 ± 0.004	1.00
100.202	0.98 ± 0.13	0.990 ± 0.017	1.32
100.212	0.94 ± 0.03	0.984 ± 0.004	1.07
121.002	0.66 ± 0.04	0.925 ± 0.009	1.24
121.202	0.49 ± 0.22	0.998 ± 0.028	0.12
121.212	0.61 ± 0.05	0.917 ± 0.011	1.09
212.002	0.68 ± 0.03	0.935 ± 0.006	1.13
212.202	0.48 ± 0.15	0.948 ± 0.032	0.93
212.212	0.59 ± 0.03	0.922 ± 0.007	0.82

Table 8.4: The table of the sample composition correction parameters for the ℓD^{+} decay signature, derived from the Monte Carlo simulation in which $B \to \nu \ell D^{(*)}$ decayed according to the phase space.

coefficient b is proportional to the width of the $\Delta\beta\gamma/\beta\gamma$ distribution). All effects due to uncertainty on the $\Delta\beta\gamma/\beta\gamma$ distribution have already been included (scaling K-factors, electron isolation, b-quark spectrum, B-meson decay model). The remaining source of the systematic uncertainty from the measurement of the proper time is the linear term, σ_0, which arises from the L_{xy} resolution. In order to test the sensitivity of $\mathcal{D}_{+}$, $\mathcal{D}_0$ and Δm_d to the changes on σ_0, we scale σ_0 up or down by 20% [61], and as the shifts we take the half of the difference. We get $\Delta\mathcal{D}_{+} = 0.0003$, $\Delta\mathcal{D}_0 = 0.0000$ and $\Delta(\Delta m_d) = 0.0033 ps^{-1}$, and we take these values as the systematic uncertainties due to the L_{xy} resolution. As one would expect, the only noticeable change is on Δm_d, since the ct resolution only affects the convolution with the cosine, and thus Δm_d.

8.2.5 $\xi_{MC}(ct)$ shape

An important ingredient of the correction to the measured asymmetries (section 7.1.2), is ξ, the probability of selecting the π_{**} as a tag in a tagged event in which the π_{**} was produced (eq. 7.4). In section 7.2.3, we broke ξ into two parts: $\xi(ct) = \xi_{norm} \cdot \xi_{MC}(ct)$ (eq. 7.21), where $\xi_{MC}(ct)$ is derived from the Monte Carlo simulation, and ξ_{norm} is fitted for from the data. We have seen in section 7.3.2 that the ξ_{norm} has a significant coupling with $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d. Now we examine the other part of $\xi(ct)$, the $\xi_{MC}(ct)$ shape.

In order to estimate the contribution of uncertainty on the $\xi_{MC}(ct)$ dependence to the systematic uncertainty on $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d, we use a different $\xi_{MC}(ct)$ shape to derive the fit results, and then compare to the 'nominal' fit result. We use a $\ell^+ D^{*-}$ signature instead of $\ell^+ D^-$ (thus having a different D^{**} composition) and a different detector simulation (QFL' vs QFL). The new ξ_{MC} shape used is shown in fig. 8-1. The width of the Gaussian significantly differs the with of the original shape, shown in fig. 7-1.

Nevertheless, we find that the dilutions and Δm_d shift by a small amount when the new $\xi_M C(ct)$ is used. We assign the absolute values of these shifts as another uncorrelated systematic uncertainty. They are: $\Delta\mathcal{D}_+ = -0.0002$, $\Delta\mathcal{D}_0 = -0.0015$, $\Delta(\Delta m_d) = -0.0035$. However, these changes are at least an order of magnitude smaller than the sample composition systematic uncertainty, and thus almost negligible. This fact is reassuring, since it shows that the shape of $\xi_{MC}(ct)$ is not critical to the analysis, due to the additional normalization derived from the data, in the form of ξ_{norm}.

8.3 Physics backgrounds

This section describes the systematic uncertainties due to the presence of the 'physics backgrounds'. These are processes that result in $\nu\ell D^{(*)}X$, with the correct correlation of charges between ℓ and $D^{(*)}$. (We always require that the tracks that comprise $D^{(*)}$ have the correct charges, so here we only ask that the lepton and the kaon have the

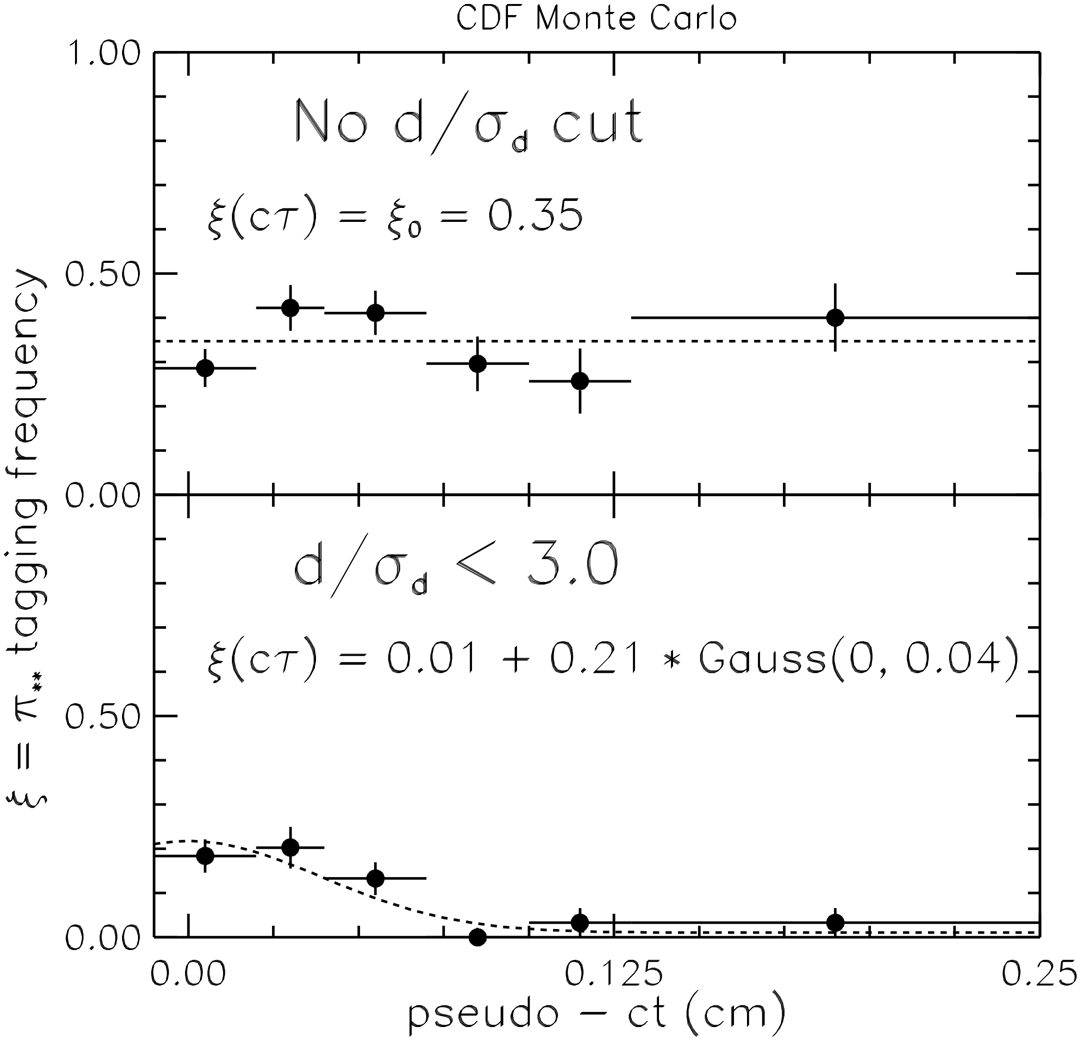


Figure 8-1: ξ_{MC} as a function of the measured ct, for the "$\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi$" decay signature, using the QFL' detector simulation. This $\xi_{MC}(ct)$ is used in the fit to derive the systematic uncertainty on the dilutions and Δm_d due to the uncertainty in the ξ_{MC} shape.

same charge.) The only processes that can actually result in a valid $\ell D^{(*)}$ signature are:

- $B \to D_s^{(*)} D^{(*)} X$, followed by $D_s^{(*)} \to \nu \ell X$
- $B_s \to \nu \ell D_s^{**}$, followed by $D_s^{**} \to D^{(*)} K$
- gluon splitting $g \to c\bar{c}$, followed by $c \to \ell X_1$ and $\bar{c} \to D^{(*)} X_2$

A fraction of these events feeds into each of the five decay signatures of the 'lepton + charm sample'. However, as is shown below, the number of these events is rather small, generally about 1% of the total in any of the five decay signatures. This small fraction warrants the treatment of these backgrounds just as sources of systematic uncertainty, as opposed to the addition of explicit corrections into the fit.

In general, the observed asymmetries ($\mathcal{A}(k, ct)$) are affected due to tagging on:

- **Fragmentation (primary) tracks.** The physics backgrounds involve B or D mesons, in whose fragmentation the flavor-correlated particles are produced. Coming from the primary vertex, these particles are much like the tagging pions the Same Side Tagging is trying to find. (In the case of $B \to D_s^{(*)} D^{(*)} X$, they *are* the SST pions.)
- **Decay (secondary) tracks.** The physics backgrounds also usually result in one or more charged particles emerging from a decay of heavy flavored mesons away from the primary vertex (since only a part of the full decay chain is reconstructed as a $\ell D^{(*)}$ signature). In this respect, these particles are similar to π_{**}'s, and we assume that they exhibit similar properties, most of all that the probability to choose one as a tag rapidly decreases with the proper time of the $\ell - D^{(*)}$ vertex. We assume that the function $\xi(ct)$, derived for π_{**}'s, describes this probability sufficiently well. However, the charge correlation of these particles with the lepton is not obvious, as is discussed below.

To account for these two effects, equation (7.18) is modified by the addition of two extra terms (assuming, for example, that k is a B^0 signature):

$$\bar{\mathcal{A}}_k(ct) \quad = \quad \alpha_k^0(ct)\mathcal{A}^0(ct_k^0) \qquad + \alpha_k^{0,**}(ct)(-1)$$

$$
\begin{aligned}
&+ \quad \alpha_k^{+}(ct)(-\mathcal{A}^{+}(ct_k^{+})) + \alpha_k^{+,**}(ct)(+1) \\
&+ \quad \alpha_k^{frag}(ct)\mathcal{A}^{frag} + \alpha_k^{decay}(ct)\mathcal{A}^{decay} \qquad (8.2)
\end{aligned}
$$

The term $\alpha_k^{frag}(ct)\mathcal{A}^{frag}$ describes the asymmetry due to tagging on the byproducts of the fragmentation of a B or a D meson, while the term $\alpha_k^{decay}(ct)\mathcal{A}^{decay}$ corresponds to tagging on the decay products.

In (8.2), $\alpha_k^{frag}(ct)+\alpha_k^{decay}(ct)$ is the total fraction of the physics background events of a given type, here denoted as $f_k^{ph.bkg}$, in the decay signature k. ($f_k^{ph.bkg} \sim 1\%$, see tables 8.5, 8.6 and 8.7). To divide this fraction into $\alpha_k^{frag}(ct)$ and $\alpha_k^{decay}(ct)$, we use $\xi(ct)$ as the probability to tag on the decay products:

$$
\begin{aligned}
\alpha_k^{frag}(ct) &= f_k^{ph.bkg}(1-\xi(ct)) \qquad (8.3) \\
\alpha_k^{decay}(ct) &= f_k^{ph.bkg}\xi(ct)) \qquad (8.4)
\end{aligned}
$$

The first four coefficients $\alpha_k(ct)$ from eq. (8.2) ought to be normalized to $1 - f_k^{ph.bkg}$, instead of unity.

In order to evaluate the systematic uncertainty due to the presence of a physics background process, we first determine $f_k^{ph.bkg}$ for each decay signature. This is described in subsections below. Then we fit for the dilutions and the Δm_d, using the equation (8.2) to calculate the asymmetry predictions, $\bar{\mathcal{A}}_k(ct)$. The shift of $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d from their nominal values (shown in table 8.1). These shifts, denoted by $\Delta\mathcal{D}_+$, $\Delta\mathcal{D}_0$ and $\Delta(\Delta m_d)$, give us estimates of the uncertainties $\sigma(\mathcal{D}_+)$, $\sigma(\mathcal{D}_0)$ and $\sigma(\Delta m_d)$ due to the process in question.

8.3.1 The process $B \to D_s^{(*)} D^{(*)} X$, followed by $D_s^{(*)} \to \nu\ell X$

When the virtual W decays into a cs quark pair instead of a $\nu\ell$ pair, it forms a $D_s^{(*)}$, which still carries the charge of the W. If the D_s decays semileptonically, resulting in $\nu\ell D^{(*)}X$ (where X is a ϕ or η, and a π^0 or a γ from the $D_s^* \to D_s$ transition), the lepton exhibits the correct charge correlation with the $D^{(*)}$. As a result, the lepton + charm sample may contain events of this type.

They are, however, characterized by a much softer lepton p_T spectrum, so one expects that they will usually have failed the inclusive semileptonic trigger requirement, and their contribution is small. In addition, the position of the B vertex is incorrect, as the D_s semileptonic decay adds an extra kink to the lepton direction.

We modify the Monte Carlo simulation to also include these decays, by including the processes $B \to D_s^{(*)} D^{(*)} X$ to the `QQ` decay table. We simulate and reconstruct all events (a mixture of $B \to D_s^{(*)} D^{(*)} X$ and $B \to \nu \ell D^{(*)}$) using exactly the same algorithm as in the data. In the end, we count how many events of which kind pass the final selection criteria. We find that in all but one of the decay signatures, the total number of $B \to D_s^{(*)} D^{(*)} X$ events does not exceed 1% of the total number of events in the given signature, as shown in the table 8.5.

Signature	$N(B \to D_s^{(*)} D^{(*)} X)$	$N(B \to \nu \ell D^{(*)} X)$	$f_k^{ph.bkg}$
$\ell^+ \overline{D^0}, \overline{D^0} \to K\pi$	201	22328	0.009
$\ell^+ D^-, D^- \to K\pi\pi$	130	7674	0.017
$\ell^+ D^{*-}, \overline{D^0} \to K\pi$	49	10880	0.005
$\ell^+ D^{*-}, \overline{D^0} \to K3\pi$	28	5162	0.005
$\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$	22	3518	0.006

Table 8.5: Contribution of $B \to D_s^{(*)} D^{(*)} X$ to the lepton + charm sample.

A schematic representation of the topology of the $B \to D_s^{(*)} D^{(*)}$ decays is shown in fig. 8-2. Events $B \to D_s^{(*)} D^{(*)} X$ are still B's, therefore they result the same 'fragmentation' asymmetry as the 'direct' decay chains: e.g. $\mathcal{A}^{frag} = \mathcal{A}^0(ct_k^0)$ in equation (8.2).

In $B \to D_s^{(*)} D^{(*)} X$ the X is frequently one or two pions. When the pions are charged, one may be selected as a tag by mistake. However, in decays $B \to D_s^{(*)} D^{(*)} \pi^\pm \pi^\mp$, it is equally likely to choose either of the pions. In the case of the semileptonic D_s decay products, both ϕ^0 or η^0 produce an equal number of π^+ and π^-. Hence we assume that the $\ell D^{(*)}$ candidates with tags on the secondary particles from $D_s \to \nu \ell X$ have the asymmetry $\mathcal{A}^{decay} = 0$. In the case of $D^{(*)}$ (the D-meson that is reconstructed as a part of $\ell D^{(*)}$ candidate), the only possibility for tagging on

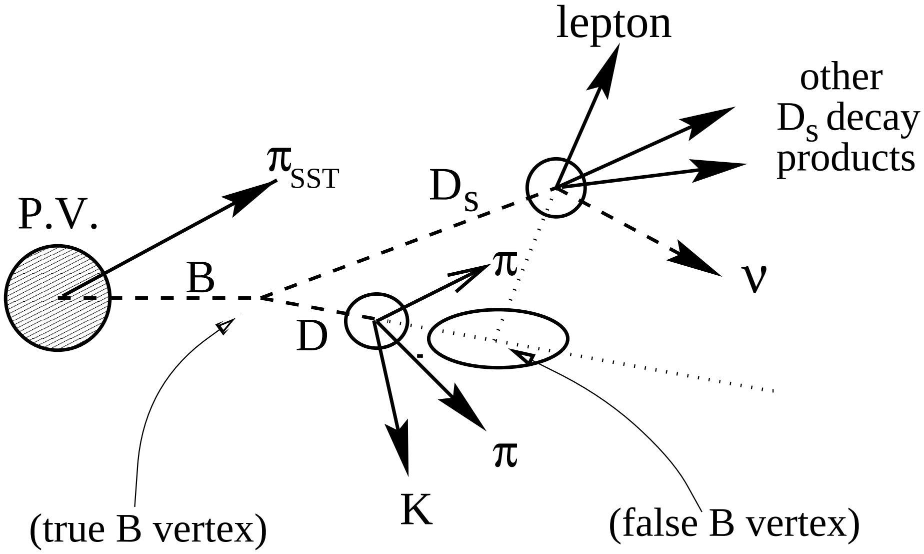


Figure 8-2: A schematic representation of the topology of the $B \to D_s^{(*)} D^{(*)}$ decay. A B meson is produced at the event primary vertex (P.V.), and decays into D_s and D mesons.The lepton originating in the D_s decay does not point back to the position of the B meson decay, resulting in an incorrect position of the B decay vertex.

the secondary particles is when $D^{(*)}$ is a D^{**}. If we assume that $\mathcal{B}(B \to D_s^{(*)} D^{**} X)$ is 36% ($= f^{**}$) of $\mathcal{B}(B \to D_s^{(*)} D^{(*)} X)$, then these decays would contribute $2f^{**}/3$ to the asymmetry,[1] so $\mathcal{A}^{decay} = 0.24$.

The fit results in the shifts $\Delta\mathcal{D}_0 = -0.0004$, $\Delta\mathcal{D}_+ = 0.0006$ and $\Delta(\Delta m_d) = -0.0010 ps^{-1}$, which are, compared to the statistical error and the sample composition systematic uncertainty, negligible. The change is so small since it amounts to adding just another percent to the contribution of the 'same' B's, which is already above 80%.

8.3.2 The process $B_s \to \nu\ell D_s^{**}$, followed by $D_s^{**} \to D^{(*)} K$

In contrast to $B \to D_s^{(*)} D^{(*)} X$, the process $B_s \to \nu\ell D_s^{**}$ is a semileptonic decay of a B-meson, and, apart from the difference between the B_s and $B_{u,d}$ and between the D_s^{**} and $D_{u,d}^{**}$ meson masses, it has exactly the same decay kinematics as the $B \to \nu\ell D^{**}$ decays that have been studied so far.

We estimate the contribution of these decays in the same fashion as for $B \to D_s^{(*)} D^{(*)} X$: we modify the Monte Carlo simulation to allow B to decay semileptonically into $D_s^{(*)}$, then simulate and reconstruct $B_s \to \ell\nu D^{(*)} X$ decays together with all other B decays. In the end, we count the $B_s \to \nu\ell D_s^{**}$ events that pass the selection criteria in each decay signature, and derive the fraction of the total for each, which is shown in table 8.6.

The charge of the kaon from the D_s^{**} decay is always correctly correlated with the lepton charge, and we assume that the probability to pick the $K^{\pm}$ track as a tag is also $\xi(ct)$.[2] As a result, these events contribute $\mathcal{A}^{decay} = +1$ to the asymmetry equation.

---

[1] The factor of 2/3 comes from isospin, since only a charged π_{**} can be selected as a tag.

[2] We neglect any difference due to the higher kaon mass, as the 'minimum p_T^{rel}' tagging method does not explicitly depend on the mass associated with the track. In addition, the kaon tracks have, on average, higher momenta in the laboratory reference frame, and hence are less likely to be picked up by the 'minimum p_T^{rel}' tagging algorithm which favors lower p_T tracks. Therefore, using $\xi(ct)$ to describe the tagging on kaons is conservative.

Signature	$N(B_s \to \nu\ell D_s^{**})$	$N(B \to \nu\ell D^{(*)}X)$	$f_k^{ph.bkg}$
$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$	265	24905	0.011
$\ell^+D^-, D^- \to K\pi\pi$	86	7674	0.011
$\ell^+D^{*-}, \overline{D^0} \to K\pi$	83	10880	0.008
$\ell^+D^{*-}, \overline{D^0} \to K3\pi$	54	5162	0.008
$\ell^+D^{*-}, \overline{D^0} \to K\pi\pi^0$	39	3518	0.009

Table 8.6: Contribution of $B_s \to \nu\ell D_s^{**}$ to the lepton + charm sample.

On the other hand, if the kaon track is not selected as the tag, then the asymmetry is due to the $B_s - K$ flavor-charge correlations. On the proper-time scale appropriate for the B^0, this asymmetry is completely washed away by the B_s mixing. (The B_s oscillation frequency, Δm_s, is expected to be very high: current experimental limit is $\Delta m_s \geq 8.0$ ps^{-1} [62]. Such fast oscillations are not discernible with the ct-binning used in this analysis.) In the fit, we assume that these decays contribute a total of $\mathcal{A}^{frag} = 0$ to the asymmetry correction equation (8.2).

Fitting for $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d yields the shifts $\Delta\mathcal{D}_0 = 0.0008$, $\Delta\mathcal{D}_+ = 0.0019$ and $\Delta(\Delta m_d) = 0.0010 ps^{-1}$. The shift in $\mathcal{D}_+$ is larger than in the other two fit parameters because 1% of the total number of events is effectively added to the coefficient $\alpha^{+,**}$ in the equation (8.2). The dilutions depend the most on the first two bins in proper time, hence a change in ξ will affect them more than it affects Δm_d. Furthermore, the ℓD^0 decay signature, dominated by charged B's, has more cross-talk than the other four dominantly neutral B signatures, so the effect is more pronounced for $\mathcal{D}_+$.

8.3.3 Gluon splitting $g \to c\bar{c}$, followed by $c \to \ell X_1$ and $\bar{c} \to D^{(*)}X_2$

The $g \to c\bar{c}$ events, where a high-p_T gluon creates a $c\bar{c}$ pair, and one c-quark in turn hadronizes into a $D^{(*)}$ meson and the other decays semileptonically, also contribute to the sample of $\nu\ell D^{(*)}$ events with the correctly correlated lepton and charm charges.

This process, however, cannot be easily simulated, as most standard Monte Carlo generators do not perform the gluon splitting. One of our selection requirements is that the combined lepton – D meson invariant mass is less than m_B (otherwise they cannot possibly come from the same B), and this results in a kinematic region where the uncertainty on the QCD calculation is large. Hence the fraction of $g \to c\bar{c}$ must be estimated from the data.

Nevertheless, despite possibly significant $g \to c\bar{c}$ production, the selection requirements (e.g. the cuts on the lepton p_T and on the $L_{xy}/\sigma_{L_{xy}}$ of the D meson) suppress $g \to c\bar{c} \to \ell D^{(*)}$ events, and one does not expect their contribution to the lepton + charm sample to be significant.

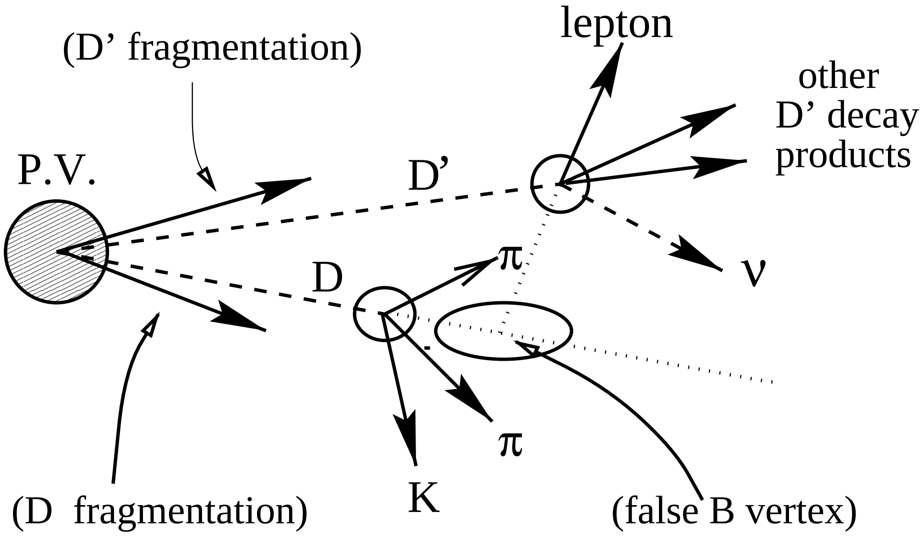


Figure 8-3: A schematic representation of the topology of the gluon splitting $g \to c\bar{c}$, followed by $c \to D^{(*)}$ and $\bar{c} \to D'$ fragmentation (here, by D' we denote a charm hadron that the $\bar{c}$ quark hadronized into). The intersection of the directions of the lepton and the the $D^{(*)}$ meson is mistakenly interpreted as the "B vertex". The position of this false B vertex may be quite inconsistent with the $B \to D$ transition.

To verify this conjecture, one first notes that in $g \to c\bar{c} \to \ell D^{(*)}$ the "B vertex" is not consistent with B meson decay. In this case, what we assume to be the "B vertex" is merely an intersection point of the charm track with the direction of motion of the D meson. A diagram of the $g \rightarrow c\bar{c} \rightarrow \ell D^{(*)}$ topology is shown in fig. 8-3. The vertexing algorithmallows this point to be anywhere along the D direction, not necessarily between the primary and the D vertex, as it would be the case for a properly reconstructed B meson. For this reason, in addition to a good B vertex quality requirements, our event selection demands that the charm proper time with respect to the "B vertex" (denoted by ct_{D-B}) is consistent with the D meson lifetime (ct_{D-B} must be between -500μm and 1000μm for D^0 or 2000μm for D^+)). If this cut is removed, there is a small number of "B vertices" that are in the tails of the ct_{D-B} distribution: possible sources are the CTC tracks wrongly matched to the SVX tracks, as well as $g \rightarrow c\bar{c}$. (For the reader interested in the details, the distributions of the full ct_{D-B} domain (including tails) for all decay signatures, are provided in Appendix J.)

Because of this property of $g \rightarrow c\bar{c}$ events, we expect their ct_{D-B} distribution to be very wide, compared to the real $B \rightarrow D$ events. Our goal is just to estimate the upper limit of the contribution of $g \rightarrow c\bar{c}$ to the lepton + charm sample, so as a prototype of a wide distribution we use a very wide gaussian,[3] and fit the ct_{D-B} distribution with the sum of this gaussian, and a template derived from the $B \rightarrow \nu\ell D^{(*)}$ Monte Carlo simulation.[4] The $B \rightarrow \nu\ell D^{(*)}$ Monte Carlo simulation includes other $B/B_s \rightarrow \nu\ell D^{(*)}X$ processes mentioned above, hence any deviation of the ct_{D-B} distributions from the Monte Carlo model – measured by the area of the wide gaussian – consists of the combination of $g \rightarrow c\bar{c}$ and poorly matched CTC and SVX tracks for the lepton. One can remove most of the latter by imposing a SVX quality requirements. (The χ^2_{total} of an SVX track includes the $\chi^2_{CTC-SVX}$ coming from associating the CTC track to the hits in the SVX.)

---

[3] The choice of the gaussian is fairly arbitrary, since the total number of $g \rightarrow c\bar{c}$ events in the tails is nevertheless small. In the end, to be conservative, we inflate the resulting number of $g \rightarrow c\bar{c}$ events by 2σ.

[4] In order to account for the low statistics in the tails, instead of fitting a sideband-subtracted distribution, we fit the ct_{D-B} distribution of the signal region with three histograms: ct_{D-B} distributions of the sidebands, a ct_{D-B} template from the $B \rightarrow \nu\ell D^{(*)}$ Monte Carlo simulation, and a wide gaussian.

Nevertheless, it is conservative to assume that *all* of the excess comes from the gluon splitting. The details are presented in Appendix J. Table 8.7 summarizes the contribution to each decay signature. As an upper limit on possible gluon splitting events, we assume the fraction which corresponds to the 2σ upward fluctuation, that is, $(N(g \to c\bar{c}) + 2\sigma(g \to c\bar{c}))/N(B \to \nu\ell D^{(*)}X)$.

Signature	$N(g \to c\bar{c} \to \ell D^{(*)})$	$N(B \to \nu\ell D^{(*)}X)$	'$f_k^{ph.bkg} + 2\sigma$'
$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$	4.8 ± 3.0	2192 ± 62	0.005
$\ell^+D^-, D^- \to K\pi\pi$	3.0 ± 2.6	1679 ± 61	0.005
$\ell^+D^{*-}, \overline{D^0} \to K\pi$	0.5 ± 0.5	798 ± 33	0.002
$\ell^+D^{*-}, \overline{D^0} \to K3\pi$	1.3 ± 0.9	667 ± 36	0.005
$\ell^+D^{*-}, \overline{D^0} \to K\pi\pi^0$	18 ± 15	2474 ± 94	0.019

Table 8.7: Contribution of gluon splitting, $g \to c\bar{c}$, to the lepton + charm sample.

The effect of the $c\bar{c}$ gluon splitting events on the mixing measurement differs from the previous two cases. Both c and $\bar{c}$ hadronize at the primary vertex, creating oppositely charged pions. Since the $c\bar{c}$ pair is collimated, these fragmentation pions are close in the phase space, and thus, on average, are equally likely to be selected by the Same Side Tagging. This results in the asymmetry $\mathcal{A}^{frag} \approx 0$.

When the D meson that results in the lepton (let us denote it by D') is charged, the charge is taken away by the lepton, and the total charge of the remaining decay products is zero, resulting in $\mathcal{A}^{decay} \approx 0$. A case when the D' meson is neutral is a more complicated one, since the remaining decay products will have a total charge opposite to the lepton's, making it more likely to pick the oppositely charged particle as the tag. It should be noted that the net asymmetry due to these tags is less than 1, since, on average, there will be more than one charged decay product, not all being correctly correlated with the lepton. The contribution to the asymmetry is decay-signature dependent, as it depends on the definition of which charge combination gives the 'right' correlation. In the ℓD^0 signature we expect the oppositely charged

ℓ-track pair, so $\mathcal{A}^{decay} > 0$. For all other decay signatures, $\mathcal{A}^{decay} < 0$. In estimating the systematic uncertainty due to $g \rightarrow c\bar{c}$, we consider the 'worst-case scenario', in which $\mathcal{A}^{decay}$ is $+1$ for charged B's and -1 for the neutrals.

The requirement that the $D^{(*)}$ proper time is between -500 and $1,500$ microns constrains the position of the $D \rightarrow \nu\ell X$ vertex (where these correlated tracks are coming from), and ensures that it is not too far from the reconstructed "B vertex" (here merely an intersection point of the lepton track and the direction of the $D^{(*)}$ meson). Consequently, one expects that the correlated tracks are coming from the neighborhood of the "B vertex", and the effect is similar to tagging on the π_{**}. Therefore, we again assume that the probability for selecting one of the decay products as the tag follows $\xi(ct)$.

Using the numbers of gluon splitting candidates scaled up by 2σ, we fit for $\mathcal{D}_+$, $\mathcal{D}_0$ and Δm_d and get $\Delta\mathcal{D}_+ = 0.0012$, $\Delta\mathcal{D}_0 = 0.0025$ and $\Delta(\Delta m_d) = -0.0006\,ps^-1$, and therefore assign the systematic uncertainties equal to the absolute values of the shifts. However, compared to the systematic uncertainty due to the sample composition, these shifts are negligible.

8.4 Summary of the systematic uncertainties

Table 8.8 summarizes the effects of various sources of systematic uncertainty. By far the biggest contribution comes from the uncertainties on the input sample composition parameters. The combined systematic uncertainty is still smaller than the statistical uncertainty, especially in the case of Δm_d. As a mixing measurement, the application of the Same Side Tagging on the lepton + charm sample is still limited by statistics. One can summarize the uncertainties of this measurement by

$$\sigma_{uncorr} \approx \sigma_{ph.bkg} \lll \sigma_{S.C.} < \sigma_{stat}$$

Source	$\sigma(\mathcal{D}_+)$	$\sigma(\mathcal{D}_0)$	$\sigma(\Delta m_d)$
Sample Composition	$+0.0216/-0.0131$	$+0.0225/-0.0131$	$+0.0295/-0.0310$
e isolation cuts	± 0.0036	± 0.0047	± 0.0045
b-quark spectrum	± 0.0052	± 0.0017	± 0.0060
$L_{\rm xy}$ resolution	± 0.0003	± 0.0000	± 0.0033
Decay model	± 0.0005	± 0.0045	± 0.0115
$\xi_{MC}(ct)$ shape	± 0.0002	± 0.0015	± 0.0035
$B \to D_s^{(*)} D^{(*)} X$	± 0.0006	± 0.0004	± 0.0010
$B_s \to \nu \ell D_s^{**}$	± 0.0019	± 0.0008	± 0.0010
$g \to c\bar{c} \to \ell D^{(*)}$	± 0.0012	± 0.0025	± 0.0006
Total	$+0.0226/-0.0147$	$+0.0237/-0.0150$	$+0.0329/-0.0343$

Table 8.8: Table of the systematic uncertainties.

Chapter 9

Conclusion and outlook

9.1 Summary of the results

We have applied the Same-Side Tagging technique to the sample of $B \to \ell D^{(*)}$ events. Observed asymmetries are corrected for the sample composition and for the selection of the $\pi^{\pm}_{**}$ as the tag. These corrections are performed simultaneously with the fit of the true B^+ asymmetry, $\mathcal{A}^+$, to a constant, and of the true B^0 asymmetry, $\mathcal{A}^0$ to a cosine convoluted with the ct-resolution function. The final result for the mixing frequency is $\Delta m_d = 0.471^{+0.078}_{-0.068}(\text{stat}) \pm 0.034(\text{syst})\ \text{ps}^{-1}$, and we also obtain the following values for the neutral and charged meson tagging dilutions: $\mathcal{D}_0 = 0.18 \pm 0.03(\text{stat}) \pm 0.02(\text{syst})$ and $\mathcal{D}_+ = 0.27 \pm 0.03(\text{stat}) \pm 0.02(\text{syst})$.

This result unambiguously establishes a significant Same-Side Tagging correlation in $p\bar{p}$ environment. The application of SST to this high statistics semi-exclusive ℓD sample provides a competitive mixing measurement. The uncertainty on Δm_d is, at present, dominated by the statistical error.

9.2 Prospects for observing CP violation using SST

In section B.6 we discussed how the angles of the Bjorken triangle (fig. 1-3), α, β and γ can, in principle, be measured from the asymmetry (equations (1.13) and (1.14)):

$$\mathcal{A}_{CP}(t) \equiv \frac{\Gamma(B^0(t) \to f) - \Gamma(\bar{B}^0(t) \to f)}{\Gamma(B^0(t) \to f) + \Gamma(\bar{B}^0(t) \to f)} = \pm \sin(2(\phi_M + \phi_D)) \sin(\Delta m_d t)$$

where the phase difference $2(\phi_M + \phi_D)$ is given by one of equations (B.6), (B.8) and (B.10), depending on the decay mode used.

If we want to measure an asymmetry $\mathcal{A}_{CP}$ in a sample of N events, then the observed asymmetry $\mathcal{A}_{CP}^{obs}$ simply relates to $\mathcal{A}_{CP}$ by $\mathcal{A}_{CP}^{obs} = \mathcal{D}\mathcal{A}_{CP}$, where $\mathcal{D}$ is the dilution of the tagging method (sec. 2.4). The uncertainty on $\mathcal{A}_{CP}$ is given by

$$\sigma^2_{\mathcal{A}_{CP}} = \frac{\sqrt{1 - \mathcal{D}^2 \mathcal{A}^2_{CP}}}{\mathcal{D}^2(\epsilon N)} \approx \frac{1}{\mathcal{D}^2(\epsilon N)} \tag{9.1}$$

where ϵN is the number of taggable signal events (ϵ is the tagging efficiency, and N is the total number of signal events). Therefore, the figure of merit is the *effective tagging efficiency*, $\epsilon\mathcal{D}^2$. For an optimal asymmetry measurement, this is the quantity one seeks to optimize. The effective tagging efficiencies of the Same Side Tagging algorithm employed in this thesis, for the neutral and charged B mesons respectively, are

$$\epsilon\mathcal{D}_0^2 = 2.4 \pm 0.7(\mathrm{stat})^{+0.6}_{-0.4}(\mathrm{syst})\%$$

and

$$\epsilon\mathcal{D}_+^2 = 5.2 \pm 1.2(\mathrm{stat})^{+0.9}_{-0.6}(\mathrm{syst})\%$$

These values can be compared to the effective tagging efficiencies for the Opposite Side Tagging methods, such as the soft-lepton tagging [22], and the jet charge tagging [21], giving $\epsilon\mathcal{D}^2 \approx 0.6 \pm 0.1$ and $\epsilon\mathcal{D}^2 = 0.3 \approx 0.1$ respectively [65]. From this comparison, it is clear that the Same Side Tagging, at the moment, seems to be the only viable tagging method that can be applied to a small sample of the fully reconstructed B-meson decays.

9.2.1 CKM angle β

Recent fits [63], using as input all available experimental data (from the measurements of CP violation in the K^0 decays, $B_d^0\bar{B}_d^0$ mixing, etc.), have been performed to constrain the angles of the Bjorken triangle, The fits predict that $0.2 \leq \sin(2\beta) \leq 0.9$, supporting the expectation that CP violation outside the kaon system will first be observed through an asymmetry between the rates of $B_d^0, \bar{B}_d^0 \to \psi K_S^0$, which is proportional to $\sin(2\beta)$ (equation (B.6)).

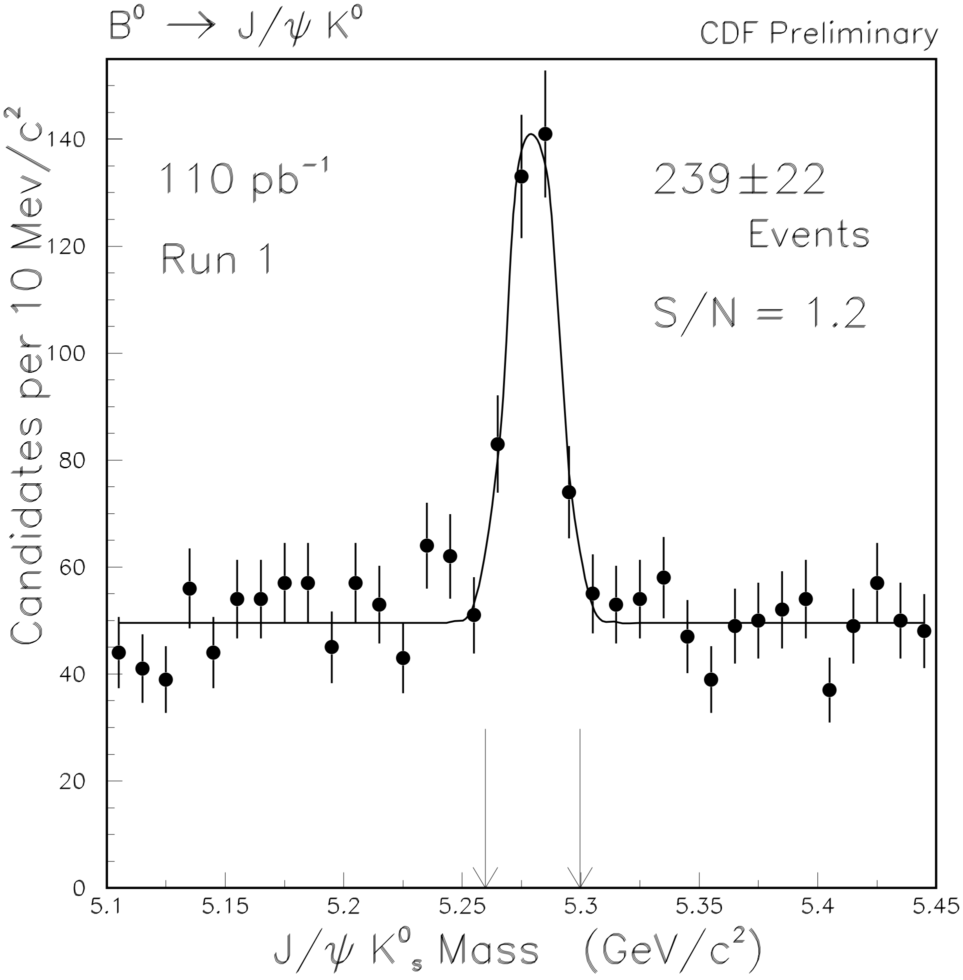


Figure 9-1: CDF's $B^0 \to \psi K_S^0$ signal from Run I. The invariant mass of $\mu\mu\pi\pi$ is shown.

CDF has, at the moment, the world's largest sample of $B^0 \to \psi K_S^0$ decays. This sample can be used for preliminary studies of CP violation in Run I (1992-1995), as well as for extrapolation to Run II (scheduled to begin in late 1999). In 110 pb^{-1} of data collected in Run I, about 240 $B^0 \to \psi K_S^0$ candidates were reconstructed. The B meson invariant mass is shown in fig. 9-1. The statistical error of a time-dependent measurement of $\mathcal{A}_{CP}(ct)$ can be properly evaluated only at the fit level, however an upper limit is given by the error on the time-integrated measurement:

$$\sigma^2_{\mathcal{A}_{CP}} \approx \frac{1+x_d^2}{x_d}\frac{1}{\mathcal{D}^2(\epsilon N)} \tag{9.2}$$

where $x_d \equiv \Delta m_d/\tau_B$, and $x_d/(1+x_d^2) = 0.47$ accounts for an additional dilution effect due to the time evolution of the B^0 signal. Putting all information together, one derives an estimate of the asymmetry error of about 0.9 for Run I. However, in order to use the available statistics in an optimal fashion, a time-dependent unbinned likelihood must be performed [64]. An error on $\sin(2\beta)$ of 0.9 can potentially still rule out a part of physically allowed parameter space if $\sin(2\beta)$ turns out to be greater than 1.0 or less than -1.0 (which may happen, given the low statistics of the $B^0 \to \psi K_S^0$ sample).

Nevertheless, for an observation of CP violation at CDF, a much larger dataset is needed. In Run II, the CDF expects about $10,000$ to $15,000$ $B^0 \to \psi K_S^0$ candidates [65]. Given the estimated improvements in the OST algorithms (section 2.2), a total error on $\sin(2\beta)$ of 0.076 is expected, a precision comparable to that of other experiments that will be trying to measure CP violation in the B-system at the same time (the B-factories BaBar, Belle and CLEO, as well as a dedicated experiment at HERA, HERA-B).

9.2.2 CKM angle α

The rate asymmetry between $B_d^0, \bar{B}_d^0 \to \pi^+\pi^-$ is proportional to $\sin(2\alpha)$ (equation (B.8)). Unfortunately, apart from the fact that these modes have very small branching ratios, there is an uncertainty in the prediction of the CP asymmetry because of penguin diagrams contributing to the decay amplitudes.

Experimentally, $B^0 \rightarrow \pi\pi$ is a formidable challenge. Unlike the clean dimuon signature of $B^0 \rightarrow \psi K^0_S \rightarrow \mu^+\mu^-\pi^+\pi^-$, here there is nothing but the long B-meson lifetime that distinguishes the dipion pair from the overwhelming combinatorial background. For this reason, CDF is building a dedicated *secondary vertex trigger*, SVT [66], which will be able to select events with displaced tracks. In Run II, a yield of about $10,000$ $B^0 \rightarrow \pi^+\pi^-$ candidates is expected. Unfortunately, the physics background processes $B^0 \rightarrow K\pi$, $B^0_s \rightarrow K\pi$ and $B^0_s \rightarrow KK$, where the kaon track is mistaken for a pion, all contribute to the same two-track invariant mass region. Therefore, to extract $B^0 \rightarrow \pi^+\pi^-$ decays, one also needs a system that can distinguish between kaons and pions.

9.2.3 CKM angle γ

The measurement of the CP angle γ will be the hardest to carry out. The modes $B^0_s \rightarrow \rho^0 K^0_S$, whose decay rate asymmetry is proportional to $\sin(2\gamma)$, have very small branching ratios, and, theoretically, they are not clean due to possibly large penguin contributions. A proposed alternative approach [67] is to measure the moduli of the decay amplitudes $B^+ \rightarrow D^0 K^+$, $B^+ \rightarrow \bar{D}^0 K^+$, $B^+ \rightarrow D_{1,2} K^+$ and their charge conjugates. Here D_1 and D_2 are CP-even and CP-odd eigenstates of the D^0 . From the two triangle equations relating the three complex amplitudes for the B^+ and B^0 decays, one can obtain $\sin^2\gamma$. We note that this method does not require tagging of the initial flavor of a B meson, and thus will not use the SST.

9.3 Conclusion

If the Bjorken triangle (fig. 1-3) is closed, $\alpha + \beta + \gamma = \pi$. Any deviation from this requirement would provide evidence for physics beyond the Standard Model, most likely involving additional CP violation interactions. One should note, however, that even if the sum of the angles is π, this does not rule out new CP violation effects in the B system. In addition, further searches can be made in the modes for which the Standard Model predicts very small asymmetries, *e.g.* $B^0_s \rightarrow \psi\phi$. In these channels the Standard Model prediction is that direct CP-violating asymmetries are likely to be at most a few percent, so large effects would suggest the New Physics.

In conclusion, *any* observed deviations from the relationships predicted by the Standard Model will provide a window on the nature of physics beyond it. Still, we first have to *observe* CP violation in the B system. In the world of B-physics, the next few years are going to be quite exciting, and the Same Side Tagging will possibly play an important role in a hadronic environment.

Appendix A

Sakharov's conditions

In section 1.1.3 we mentioned the conditions for the asymmetry between baryons and antibaryons presented by Sakharov in 1966 [11]. Here we outline the experimental evidence for the baryon-antibaryon asymmetry and discuss how it can develop if Sakharov's requirements are met.

Cosmic rays sample material from the entire galaxy. In cosmic rays, protons outnumber antiprotons 10^4 to 1. If there were antimatter galaxies then we should see gamma emissions from matter-antimatter annihilation. Their absence is strong evidence that at least the nearby clusters of galaxies (e.g., Virgo) are matter-dominated.

For the galaxies farther away, there is little experimental proof of this dominance of matter. However, in the Big-Bang model of the early universe, most of the matter and antimatter were causally connected, allowing their interaction and thus mutual annihilation. The annihilation has made the asymmetry much larger today than in the early universe: it has been estimated that at the end of the first microsecond there were 30 million antiquarks for every 30 million *and one* quarks. If antimatter was allowed to interact with matter, then, over time, most of it would have annihilated, leaving a very small initial excess of matter to dominate the Universe.

Sakharov enumerated three necessary conditions for the baryon asymmetry [11]:

1. **Baryon number violation.** If baryon number (the number of baryons *minus* the number of antibaryons) is conserved in all reactions, then the present baryon

asymmetry can only reflect asymmetric initial conditions. This hypothesis does not work with inflation theories, which dilute any initial abundances. Thus, the processes that violate baryon number must exist.

2. **C and CP violation.** Even in the presence of baryon-number-violating reactions, without a preference for matter over antimatter (caused by C and CP violation) baryon number violation will take place at the same rate in *both* directions, leaving no excess.

3. **A period of thermodynamic non-equilibrium.** Because CPT guarantees equal masses for baryons and antibaryons, thermal equilibrium would drive the necessary reactions to correct for any asymmetry otherwise developing.

CP violation is thus one of the most important ingredients of Sakharov's condition.

Appendix B

The angles of the Bjorken triangle

In addition to measuring the sides of the Bjorken triangle (fig. 1-3), the angles α, β and γ can also be determined independently. From the Bjorken triangle (fig. 1-3), we read:

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^{*}}{V_{ud}V_{ub}^{*}}\right), \quad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^{*}}{V_{td}V_{tb}^{*}}\right), \quad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^{*}}{V_{cd}V_{cb}^{*}}\right) \tag{B.1}$$

The angles α, β and γ are related to the expected asymmetries in the decay rates of the neutral B_d^0 meson and its antiparticle $\bar{B}_d^0$ into a CP eigenstate f. Since neutral B mesons mix via the box diagram shown in fig. 2-1, there are two alternative decay paths for $B^0 \to f$ transition:

- $B^0 \to f$
- $B^0 \to \bar{B}^0 \to f$

In general, there is a relative phase between the amplitudes for these two processes. The CP conjugate situation (starting out with $\bar{B}^0$) has the opposite phase. It can be shown that the time-dependent asymmetry, defined as

$$\mathcal{A}_{CP}(t) \equiv \frac{,(B^0(t) \to f) - ,(\bar{B}^0(t) \to f)}{,(B^0(t) \to f) + ,(\bar{B}^0(t) \to f)} \tag{B.2}$$

can be expressed as

$$\mathcal{A}_{CP}(t) = \pm \sin(2(\phi_M + \phi_D)) \sin(\Delta m_d t) \tag{B.3}$$

where $2\phi_M$ denotes the CKM phase of the $B^0 \rightarrow \bar{B}^0$ amplitude (fig. 2-1), and $2\phi_D$ is the phase difference between the decay amplitudes for $B^0 \rightarrow f$ and $\bar{B}^0 \rightarrow f$. Therefore, if one denotes the amplitude of a given process X by $a(X)$, the sum of the relative phases between $B^0 \rightarrow f$ and $B^0 \rightarrow \bar{B}^0 \rightarrow f$ is

$$2(\phi_M + \phi_D) = \arg\left(\frac{a(B^0 \rightarrow f)}{a(B^0 \rightarrow \bar{B}^0 \rightarrow f)}\right) \tag{B.4}$$

Using eq. (B.4), we can now evaluate the decays for which the asymmetries are proportional to the angles of the Bjorken triangle.

- $B_d^0 \rightarrow J/\psi K_S^0$. The B^0 mixing amplitude comes with the phase $\arg((V_{tb}^* V_{td})/(V_{tb} V_{td}^*))$, while the K^0 (from K_S^0) mixing amplitude gives $\arg((V_{cd}^* V_{cs})(V_{cd} V_{cs}^*))$. Putting everything together:

$$\begin{aligned} 2(\phi_M + \phi_D) &= \arg\left(\frac{a(B_d^0 \rightarrow J/\psi K_S^0)}{a(B_d^0 \rightarrow \bar{B}_d^0 \rightarrow J/\psi K_S^0)}\right) && \text{(B.5)} \\ &= \arg\left(\left(\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\right)\left(\frac{V_{cs}^* V_{cb}}{V_{cs} V_{cb}^*}\right)\left(\frac{V_{cd}^* V_{cs}}{V_{cd} V_{cs}^*}\right)\right) \\ &= \arg\left(\left(\frac{V_{td} V_{tb}^*}{V_{cd} V_{cb}^*}\right)\left(\frac{V_{cd} V_{cb}^*}{V_{tb} V_{td}^*}\right)^*\right) \\ &= -2\arg\left(\frac{V_{cd} V_{cb}^*}{V_{tb} V_{td}^*}\right) = -2\beta && \text{(B.6)} \end{aligned}$$

- $B_d^0 \rightarrow \pi^+\pi^-$. The mixing amplitude comes with the phase $\arg((V_{tb}^* V_{td})/(V_{tb} V_{td}^*))$.

$$\begin{aligned} 2(\phi_M + \phi_D) &= \arg\left(\frac{a(B_d^0 \rightarrow \pi^+\pi^-)}{a(B_d^0 \rightarrow \bar{B}_d^0 \rightarrow \pi^+\pi^-)}\right) && \text{(B.7)} \\ &= \arg\left(\left(\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\right)\left(\frac{V_{ud}^* V_{ub}}{V_{ud} V_{ub}^*}\right)\right) = -2\alpha && \text{(B.8)} \end{aligned}$$

- $B_s^0 \rightarrow \rho K_S^0$. In the case of the B_s^0 meson, the mixing amplitude comes with the phase $\arg((V_{tb}^* V_{ts})/(V_{tb} V_{ts}^*))$. Here also the K_S^0 is involved, so the K^0 mixing amplitude $\arg((V_{cd}^* V_{cs})(V_{cd} V_{cs}^*))$ must also be taken into account:

$$\begin{aligned} 2(\phi_M + \phi_D) &= \arg\left(\frac{a(B_s^0 \rightarrow \rho K_S^0)}{a(B_s^0 \rightarrow \bar{B}_s^0 \rightarrow \rho K_S^0)}\right) && \text{(B.9)} \\ &= \arg\left(\left(\frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*}\right)\left(\frac{V_{ud}^* V_{ub}}{V_{ud} V_{ub}^*}\right)\left(\frac{V_{cs}^* V_{cd}}{V_{cs} V_{cd}^*}\right)\right) = -2\gamma && \text{(B.10)} \end{aligned}$$

These examples demonstrate that the three angles of the Bjorken triangle can, in principle, be measured independently of each other.

Appendix C

Fragmentation of b-quarks into B-mesons

The confinement of quarks is thought to be related to "asymptotic freedom", a QCD phenomenon in which the strength of the interaction between two quarks decreases as the $|q^2|$ of the process (momentum transfer squared) increases. Higher $|q^2|$ means a shorter distance between the quarks, and vice versa. As a result, the color force between a quark and an antiquark is weak at very short distances (high $|q^2|$), and therefore the quarks and gluons are treated as free during the interaction. However, as the distance between the interacting partons increases, the color force between them rises, and as they move away from the interaction point, color forces organize them into color-free structures of mesons (a bound state of a quark and an antiquark) and baryons (a bound state of three quarks or three antiquarks). This process is called *hadronization* or *fragmentation*, and usually involves creation of additional quark-antiquark pairs. The newly produced hadron is accompanied by several other hadrons, each carrying a fraction of the energy of the original quark, appearing as if some massive particle crumbled into many fragments. Thus an energetic parton shows in the detector as a *jet* of hadrons.

Consider a fast parton k with energy E_k, producing a hadron h with energy fraction z,

$$z \equiv \frac{E_h}{E_k} \tag{C.1}$$

where $0 \leq z \leq 1$. The probability of finding h in the range z to $z + dz$ is defined to be $D_k^h(z)dz$, where D_k^h is the k-to-h *fragmentation function*. Instead of energy, sometimes the longitudinal momentum p_L (along the quark direction of motion) and the light-cone variable $E + p_L$ are used. However, for relativistic hadrons h traveling close to the direction of the parent parton k, all three definitions coincide.

Fragmentation models

The breakdown of perturbative QCD in the low-q^2 domain characteristic of the fragmentation necessitates the use of phenomenological models. Two are usually distinguished – independent fragmentation and string fragmentation.

The independent fragmentation model of Feynman and Field [25] assumes that the partons leaving the interaction point fragment independently of each other. The model is recursive in nature, and it specifies that the combination of a free quark with an antiquark from a newly created $q'\bar{q}'$ pair is described by a single function $f(z)$. The model, however, is unable to describe the fragmentation of gluons, and cannot account for the baryon production.

String fragmentation, developed by the Lund group [51], pictures the colored partons, say q and $\bar{q}$, as connected by a color field flux tubes. If these flux tubes are uniform,[1] the energy of the color field stretched between two q and $\bar{q}$ is proportional with the distance between them. As the q and $\bar{q}$ move apart, the potential energy increases, and the string breaks with the production of the new $q'\bar{q}'$ pair, so that the system splits into two color singlet systems, $q\bar{q}'$ and $q'\bar{q}$. If the remaining energy is large enough, further breaks may occur. In the Lund string model, the process continues until only hadrons on the mass-shell remain, each hadron corresponding to a small piece of the color string. The baryons are produced by creating a diquark-antidiquark pairs.[2]

---

[1] The tube radius is approximately the typical hadron radius, ~ 1 fm.

[2] A diquark in a color antitriplet state is treated just like an ordinary antiquark.

When several partons are moving away from each other, the details of the string picture become more complicated. In a $q\bar{q}g$ event (as in initial or final gluon radiation, *e.g.* in NLO $b\bar{b}$ production), the color field lines go between q and g, and g and $\bar{q}$ (but not between q and $\bar{q}$, as they carry different colors). Fragmentation along this kinked string proceeds along the same lines as for the simple $q\bar{q}$ string.

Peterson fragmentation

A quark and an antiquark are most likely to combine into a meson when they have about the same velocity. Thus, the heavy b-quark needs to lose only a small fraction of its energy in order to create a number of light $q\bar{q}$ pairs with comparable velocity. When the b-quark combines with a $\bar{q}$, the newly formed B meson carries most of the energy of the b-quark: $z = E_B/E_b \sim 1$ (equation (C.1)). For a heavy quark Q, we thus expect that $D_Q^H(z)$ will peak at high z, approaching delta function at 1 as $m_Q \to \infty$

An explicit model, developed by *Peterson et al.* [26], uses time-ordered perturbation theory in the infinite-momentum frame to derive ΔE, and then assumes that the transition probability is proportional to $1/(\Delta E)^2$, resulting in

$$D_Q^H(z) = (\text{const}) \times \frac{1}{z}\left[1 - \frac{1}{z} - \frac{\epsilon_Q}{1-z}\right]^{-2} \tag{C.2}$$

The parameter ϵ_Q is called *Peterson* ϵ, and is tuned to experimental data. For c quark, the typical value is $\epsilon_c = 0.06$, and for the b quark $\epsilon_b = 0.006$ (smaller ϵ implies the peak closer to 1).

Appendix D

Details of the $B \to \ell D^{(*)} X$ candidate selection

Here we provide the detailed B meson candiate selection requirements.

D.1 Selection requirements for $\ell^+ \overline{D^0}$, $\overline{D^0} \to K^+\pi^-$

The D^0 cuts are the same as in D^{*+} with $D^0 \to K\pi$, except for a higher impact parameter cut (needed to suppress the combinatorial background, which is not as high in D^*'s due to the narrow D^*-D^0 mass difference). In order to improve the separation between the B^+ and B^0 signatures, here we also remove as many D^{*+}-meson candidates (that originated from the B^0 decay) as possible, by rejecting any event that satisfies the $m(K,\pi,\pi_s) - m(K,\pi)$ cut. The other cuts are specifically:

- $p_T(\ell) > 6.0$ GeV
- $1.80 < m(K\pi) < 1.95$
- $m(\ell K\pi) < 5.0$ GeV
- d/σ for $K, \pi > 3.0$
- $p_T(K) > 0.7$ GeV

- $p_T(\pi) > 0.5$ GeV
- $p_T(K\pi) > 2.0$ GeV
- Prob(D vertex fit) $> 1\%$
- $\frac{L_{xy}}{\sigma}(K\pi) > 3.0$
- SVX quality cuts for K, π ($\chi^2/N_{\mathrm{hit}} < 6.0$ and $N_{\mathrm{unshared}} > 0$)
- the D meson proper time: -0.5 mm $< ct_{D-B} < 1$ mm

D.2 Selection requirements for $\ell^+ D^-$, $D^- \to K^+\pi_1^-\pi_2^-$

The algorithm is similar to $\ell^+\overline{D^0}$, except that here there are two pions. Since they have the same charge, they are algorithmically indistinguishable, so for the π_2 loop we try only tracks that haven't been already used as π_1. The remaining requirements are:

- $p_T(\ell) > 6.0$ GeV
- $1.80 < m(K\pi\pi_2) < 1.95$
- $m(\ell K\pi\pi_2) < 5.0$ GeV
- $p_T(K) > 0.6$ GeV
- $p_T(K\pi\pi_2) > 3.0$ GeV
- d/σ for $K, \pi, \pi_2 > 2.0$
- Prob(D vertex fit) $> 1\%$
- $\frac{L_{xy}}{\sigma}(K\pi\pi_2) > 5.0$
- χ^2(B vertex fit) < 25.0
- SVX quality cuts for K, π, π_2 ($\chi^2/N_{\mathrm{hit}} < 6.0$ and $N_{\mathrm{unshared}} > 0$)
- the D-meson proper time: -0.5 mm $< ct_{D-B} < 2$ mm

D.3 Selection Requirements for $\ell^+ D^{*-}$, $D^{*-} \to \overline{D^0}\pi_s^-$, $\overline{D^0} \to K\pi$

Due to a narrow $m_{D^*} - m_{D^0}$ mass difference, this mode is very clean, and other requirements are as loose as possible:

- $p_T(\ell) > 6.0$ GeV
- $1.80 < m(K\pi) < 1.95$
- $|m(K\pi\pi_s) - m(K\pi) - 145.44\ \mathrm{MeV}/c^2| < 3\ \mathrm{MeV}/c^2$
- d/σ for $K, \pi > 1.0$
- Prob(D vertex fit) $> 1\%$
- $\frac{L_{xy}}{\sigma}(K\pi) > 1.0$
- χ^2(B vertex fit) < 60.0
- the D-meson proper time: $-0.5\ \mathrm{mm} < ct_{D-B} < 1\ \mathrm{mm}$

D.4 Selection Requirements for $\ell^+ D^{*-}$, $D^{*-} \to \overline{D^0}\pi_s^-$, $\overline{D^0} \to K^-\pi^+\pi_2^+\pi_3^-$

The requirements here are the same as in D^{*+} with $D^0 \to K\pi$, however the D^0 selection is slightly different, since we allow one of the four decay products of the D^0 not to be in the SVX One of the pions must have the same charge as the K, and the other two the opposite sign. However, since the K and the π_3 tracks have the same charge and we consider both combinations, it might happen that there are two D^0-meson candidates with the K and π_3 tracks swapped. If that happens, we choose the D^0-candidate that results in the mass difference $m_{K\pi\pi_2\pi_3\pi_s} - m_{K\pi\pi_2\pi_3}$ closer to the world average.

- $p_T(\ell) > 6.0$ GeV

- $1.80 < m(K\pi\pi_2\pi_3) < 1.95$
- $|m(K\pi\pi_2\pi_3\pi_s) - m(K\pi\pi_2\pi_3) - 145.44\ \mathrm{MeV}/c^2| < 2\ \mathrm{MeV}/c^2$
- d/σ for $K, \pi, \pi_2, \pi_3 > 0.5$
- Prob(D vertex fit) $> 0.01\%$
- $\frac{L_{xy}}{\sigma}(K\pi\pi_2\pi_3) > 1.0$
- χ^2(B vertex fit) < 60.0
- the D-meson proper time: -0.5 mm $< ct_{D-B} < 1$ mm

D.5 Selection Requirements for $\ell^+ D^{*-}$, $D^{*-} \to \overline{D^0}\pi_s^-$, $\overline{D^0} \to K^+\pi^- X$

We now consider on the satellite D sample defined in the mass region 1.3 GeV $< m(K\pi) < 1.7$ GeV, in which the π^0 is missed and the charm is not fully reconstructed, resulting in a broad $K\pi$ distribution. Since there is no $K\pi$ peak to fit for the number of events, we turn to the mass difference distribution $\Delta m = m(K\pi\pi_s) - m(K\pi)$. The fit to this distribution is described in Appendix E, here it suffices to mention that we will use the events under the peak ($0.139 < \Delta m < 0.155$) as the signal sample, and above the peak as the background sample ($0.16 < \Delta m < 0.19$) (also referred to as the 'sideband'). We impose the following requirements:

- $p_T(\ell) > 6.0$ GeV
- $1.30 < m(K\pi) < 1.70$
- $p_T(K) > 1.0$ GeV
- $p_T(\pi) > 0.8$ GeV
- χ^2(D vertex fit) < 10.0

- d/σ for $K, \pi > 1.0$
- χ^2(B vertex fit) < 10.0
- $\frac{L_{xy}}{\sigma}(K\pi) > 1.0$
- the D-meson proper time -0.5 mm $< ct_{D-B} < 1.5$ mm

Appendix E

Functions to fit the $m_{D^*} - m_{D^0}$ distribution

In order to fit the shape of the combinatorial background, we make use of the mass distibution of the two-body decay phase space, modified in order to account for the kinematic cuts of the D^0 candidate. (The p_T cut on the K and π tracks makes the high end of the mass difference distribution more flat, and the unmodified phase space shape does not fit well.)

We define

$$m_{D^0\pi_s} \equiv m(K\pi\pi_s) - m(K\pi) + m_{D^0} \tag{E.1}$$

where $m(K\pi)$ and $m(K\pi\pi_s)$ are the invariant masses of the D^0 and D^* candidates respectively, and m_{D^0} is the world average value of the D^0 mass, so $m_{D^0\pi_s}$ is (roughly) the energy available to the decay products. The distribution we use for the combinatorial background is

$$\frac{dN}{dm_{D^0\pi_s}} = \frac{C_{norm}}{m^{\omega}_{D^0\pi_s}}\sqrt{(m^2_{D^0\pi_s} - (m_{D^0} - m_\pi)^2)(m^2_{D^0\pi_s} - (m_{D^0} + m_\pi)^2)} \tag{E.2}$$

where the power ω is necessary to take into account the change of the shape of the distribution due to the p_T cuts on the kaon and the pion. The parameters C_{norm} and ω are freely floating in the fit.

In the same way, we take the 4-body phase space as a motivation for the signal shape, but let the parameters float, in order to achieve a good fit. We define

$$x \equiv \frac{m_{D^0\pi_s} - (m_{D^0} + m_\pi)}{E - (m_{D^0} + m_\pi)} \tag{E.3}$$

The distribution we use for the signal is

$$\frac{dN}{dm_{D^0\pi_s}} = N_0 x^{(1/2)(3k-5)}(1-x)^{(3/2)(n-k)-1} \tag{E.4}$$

Appendix F

List of all decay chains in the lepton + charm sample

Below is an exhaustive list of possible decay chains for each decay signature.

A note on the numbering scheme

We have invented a compact numerical notation for each of the decay channels. In the following chapters, we could have simply referred to quantities of any decay channel by its equation number, however, this notation is more useful since it has a physical meaning .

The *channel code* is always a six-digit integer. The channel can involve a maximum of six particles: B, D^{**}, π_{**}, D^{*}, π_s and D, and for every channel it is sufficient to know whether the particle was present, and whether it was neutral or charged. Thus, for each of them we have only three possibilities: '0' means the particle is not present, '1' that it is neutral, and '2' that it is charged. These numbers are then strung in the above order (B, D^{**}, π_{**}, D^{*}, π_s and D).

Few examples:

- 100.002 means that there were a neutral B and a charged D, thus $B^0 \to D^-$.
- xxx.221 means that there were D^{*-}, π_s^-, $\bar{D}^0$, thus the channel involves $D^{*-} \to \bar{D}^0\pi_s^-$ (a frequent combination).

- xxx.202 means that there were D^{*-} and D^-, but *not* π^0. The only remaining possibility is a photon, so the channel involves $D^{*-} \to D^-\gamma$.

Et cetera. (A dot is added in the middle for readability.)

D^{*-} sample:

- **B^0 contribution:**

$B^0 \to \nu\ell^+D^{*-}$ . . . . f^*
[100.221]

and

$B^0 \to \nu\ell^+D^{**-}$
$D^{**-} \to D^{*-}\pi^0_{**}$ (didn't detect π^0_{**}) . . . . $f^{**}P_V(1/3)$
[121.221]

- **B^+ contribution:**

$B^+ \to \nu\ell^+\bar{D}^{**0}$
$\bar{D}^{**0} \to D^{*-}\pi^-_{**}$ (missed π^-_{**})
(a fraction ξ_{D^*} of these has $\mathcal{D} = +1$) . . . . $f^{**}P_V(2/3)$
[212.221]

D^- sample:

- **B^0 contribution:**

$B^0 \to \nu\ell^+D^-$ . . . . f
[100.002]

and

$B^0 \to \nu \ell^+ D^{*-}$

$D^{*-} \to D^- \gamma$ (didn't detect γ) .... $f^* \mathrm{BR}(D^{*-} \to D^- \gamma)$

[100.202]

and

$B^0 \to \nu \ell^+ D^{*-}$

$D^{*-} \to D^- \pi^0_s$ (didn't detect π^0_s) .... $f^* \mathrm{BR}(D^{*-} \to D^- \pi^0)$

[100.212]

and

$B^0 \to \nu \ell^+ D^{**-}$

$D^{**-} \to D^{*-} \pi^0_{**}$(didn't detect π^0_{**})

$D^{*-} \to D^- \gamma$ (didn't detect γ) .... $f^{**} P_V (1/3) \mathrm{BR}(D^{*-} \to D^- \gamma)$

[121.202]

and

$B^0 \to \nu \ell^+ D^{**-}$

$D^{**-} \to D^{*-} \pi^0_{**}$(didn't detect π^0_{**})

$D^{*-} \to D^- \pi^0_s$ (didn't detect π^0_s) .... $f^{**} P_V (1/3) \mathrm{BR}(D^{*-} \to D^- \pi^0)$

[121.212]

and

$B^0 \to \nu \ell^+ D^{**-}$

$D^{**-} \to D^- \pi^0_{**}$ (didn't detect π^0_{**}) .... $f^{**} (1 - P_V)(1/3)$

[121.002]

- **B^+ contribution:**

$B^+ \to \nu\ell^+ \bar{D}^{**0}$
$\bar{D}^{**0} \to D^-\pi^-_{**}$ (missed π^-_{**})
(a fraction ξ_{D^-} of these has $\mathcal{D} = +1$) .... $f^{**}(1 - P_V)(2/3)$
[212.002]

and

$B^+ \to \nu\ell^+ \bar{D}^{**0}$
$\bar{D}^{**0} \to D^{*-}\pi^-_{**}$ (missed π^-_{**})
$D^{*-} \to D^-\gamma$ (didn't detect γ)
(a fraction ξ_{D^-} of these has $\mathcal{D} = +1$) .... $f^{**}P_V(2/3)\mathrm{BR}(D^{*-} \to D^-\gamma)$
[212.202]

and

$B^+ \to \nu\ell^+ \bar{D}^{**0}$
$\bar{D}^{**0} \to D^{*-}\pi^-_{**}$ (missed π^-_{**})
$D^{*-} \to D^-\pi^0_s$ (didn't detect π^0_s)
(a fraction ξ_{D^-} of these has $\mathcal{D} = +1$) .... $f^{**}P_V(2/3)\mathrm{BR}(D^{*-} \to D^-\pi^0)$
[212.212]

$\bar{D}^0$ sample:

- **B^+ contribution:**

$B^+ \to \nu\ell^+ \bar{D}^0$ .... f
[200.001]

and

$B^+ \to \nu\ell^+\bar{D}^{*0}$

$\bar{D}^{*0} \to \bar{D}^0\gamma$ (didn't detect γ) .... $f^*\mathrm{BR}(\bar{D}^{*0} \to \bar{D}^0\gamma)$

[200.101]

and

$B^+ \to \nu\ell^+\bar{D}^{*0}$

$\bar{D}^{*0} \to \bar{D}^0\pi^0_s$ (didn't detect π^0_s) .... $f^*\mathrm{BR}(\bar{D}^{*0} \to \bar{D}^0\pi^0)$

[200.111]

and

$B^+ \to \nu\ell^+\bar{D}^{**0}$

$\bar{D}^{**0} \to \bar{D}^{*0}\pi^0_{**}$(didn't detect π^0_{**})

$\bar{D}^{*0} \to \bar{D}^0\gamma$ (didn't detect γ) .... $f^{**}P_V(1/3)\mathrm{BR}(\bar{D}^{*0} \to \bar{D}^0\gamma)$

[211.101]

and

$B^+ \to \nu\ell^+\bar{D}^{**0}$

$\bar{D}^{**0} \to \bar{D}^{*0}\pi^0_{**}$(didn't detect π^0_{**})

160 $\bar{D}^{*0} \to \bar{D}^0\pi^0_s$ (didn't detect π^0_s) .... $f^{**}P_V(1/3)\mathrm{BR}(\bar{D}^{*0} \to \bar{D}^0\pi^0)$

[211.111]

and

$B^+ \to \nu\ell^+\bar{D}^{**0}$

$\bar{D}^{**0} \to D^o\pi^0_{**}$ (didn't detect π^0_{**}) .... $f^{**}(1 - P_V)(1/3)$

[211.001]

and

$B^+ \rightarrow \nu \ell^+ \bar{D^{**0}}$

$\bar{D^{**0}} \rightarrow D^{*-}\pi^-_{**}$ (missed π^-_{**})

$D^{*-} \rightarrow \bar{D}^0\pi^-_s$ (didn't reconstruct π^-_s)

(a fraction $\zeta_{\bar{D}^0}$ of these has $\mathcal{D} = -1$!)

.... $f^{**}P_V(2/3)\mathrm{BR}(D^{*-} \rightarrow \bar{D}^0\pi^-)(1-\epsilon(\pi_s))$

[212.221]

- **B^0 contribution:**

$B^0 \rightarrow \nu \ell^+ D^{*-}$

$D^{*-} \rightarrow \bar{D}^0\pi^-_s$ (didn't reconstruct π^-_s)

.... $f^{*}\mathrm{BR}(D^{*-} \rightarrow \bar{D}^0\pi^-)(1-\epsilon(\pi_s))$

[100.221]

and

$B^0 \rightarrow \nu \ell^+ D^{**-}$

$D^{**-} \rightarrow D^{*-}\pi^0_{**}$ (didn't detect π^0_{**})

$D^{*-} \rightarrow \bar{D}^0\pi^-_s$ (didn't reconstruct π^-_s)

.... $f^{**}P_V(1/3)\mathrm{BR}(D^{*-} \rightarrow \bar{D}^0\pi^-)(1-\epsilon(\pi_s))$

[121.221]

and

$B^0 \rightarrow \nu \ell^+ D^{**-}$

$D^{**-} \rightarrow \bar{D}^0\pi^-_{**}$ (missed π^-_{**})

(a fraction $\xi_{\bar{D}^0}$ of these has $\mathcal{D} = +1$) .... $f^{**}(1-P_V)(2/3)$

[122.001]

and

$B^0 \to \nu \ell^+ D^{**-}$

$D^{**-} \to \bar{D^{*0}} \pi^-_{**}$ (missed π^-_{**})

$\bar{D^{*0}} \to \bar{D}^0 \gamma$ (didn't detect γ)

(a fraction $\xi_{\bar{D}^0}$ of these has $\mathcal{D} = +1$)

. . . . $f^{**} P_V (2/3) \mathrm{BR}(\bar{D^{*0}} \to \bar{D}^0 \gamma)$

[122.101]

and

$B^0 \to \nu \ell^+ D^{**-}$

$D^{**-} \to \bar{D^{*0}} \pi^-_{**}$ (missed π^-_{**})

$\bar{D^{*0}} \to \bar{D}^0 \pi^0_s$ (didn't detect π^0_s)

(a fraction $\xi_{\bar{D}^0}$ of these has $\mathcal{D} = +1$)

. . . . $f^{**} P_V (2/3) \mathrm{BR}(\bar{D^{*0}} \to \bar{D}^0 \pi^0)$

[122.111]

Appendix G

A sample calculation of $\epsilon^{D}_{k\ell}$

Here we present a sample calculation of a relative charm reconstruction efficiency. We choose to do so for $\ell^+ D^{*-}$ decay signatures. Table G.1 (a copy of table 5.3) shows the list of decay chains feeding into a $\ell^+ D^{*-}$ signature, and their relative branching fractions. Out of three possible decay chains, two proceed through a D^{**} ([121.221] and [212.221]), and thus are reconstructed with a different efficiency compared with the 'direct' decay chain ([100.221]).

code	decay chain	$\mathcal{B}/\mathcal{B}(B \to \nu\ell X)$
100.221	$B^0 \to \nu\ell^+ D^{*-}$	f^*
121.221	$B^0 \to \nu\ell^+ D^{**-}$, $D^{**-} \to D^{*-}\pi^0_{**}$ (didn't detect π^0_{**})	$\frac{1}{3} f^{**} P_V$
212.221	$B^+ \to \nu\ell^+ \bar{D}^{**0}$, $\bar{D}^{**0} \to D^{*-}\pi^+_{**}$ (missed π^+_{**})	$\frac{2}{3} f^{**} P_V$

Table G.1: The list of decay chains contributing to the three ℓD^{*-} decay signatures, and their relative contributions.

From table G.1 we conclude that the decay chains 121.221 and 100.221 are produced in a ratio of $f^{**} P_V / 3 f^*$. We denote this ratio by $N_{proj}(121.221)/N_{proj}(100.221)$. We compare this ratio to the ratio of the measured numbers $N_{meas}(121.221)$ and $N_{meas}(100.221)$, and derive the ratio of two efficiencies:

$$\frac{N_{meas}(121.221)}{N_{meas}(100.221)} = \frac{\epsilon_R(121.221)}{\epsilon_R(100.221)} \frac{N_{proj}(121.221)}{N_{proj}(100.221)} = \epsilon(D)_{121.221} \frac{N_{proj}(121.221)}{N_{proj}(100.221)}$$

Therefore the relative charm reconstruction efficiency for the decay chain 121.221 is:

$$\epsilon(D)_{121.221} = \frac{N_{meas}(121.221)}{N_{meas}(100.221)} \frac{f^*}{f^{**} P_V(1/3)} \tag{G.1}$$

In the calculation of $N_{proj}(121.221)/N_{proj}(100.221)$ we have to use the sample composition that corresponds to the Monte Carlo that was used. The values of the sample composition parameters corresponding to the Monte Carlo simulation are given in table 5.5, which we repeat here in table G.2.

name	value
R_f	2.722
f^{**}	0.356
P_V	0.687
τ_{B^+}/τ_{B^0}	1.014

Table G.2: The values of the sample composition parameters used in the Monte Carlo generation.

Counting the events in the Monte Carlo simulation that pass the reconstruction requirements, we get:

$$\frac{N_{meas}(121.221)}{N_{meas}(100.221)} = 0.069323 \pm 0.002466$$

which we substitute into equation (G.1), and obtain $\epsilon(D)_{121.221} = 0.38 \pm 0.02$.

Appendix H

Dependence of the ct-resolution on the sample composition

Here we derive an explicit expression for the ct-dependent resolution corrected for the sample composition, $\sigma_{ct}(ct)$ (equation (5.40)). From the Monte Carlo simulation we determine the dependence $\sigma_{ct}(ct)$ for the 'direct' decay chain, $\sigma_{kd}(ct)$, as well as the `RMS` of the overall σ_{kd} (integrated over the ct range). As discussed in section 5.2.4, the total distribution Δct is a sum of off-center gaussians for each decay chain ℓ. We calculate the `RMS` of the sum of the off-gaussian distributions for the decay signature k, $\overline{RMS}_k$, divide it by the `RMS` of of the overall Δct_{kd} distributio for the direct mode, RMS_{kd}, and use this ratio as a scale factor multiplying $\sigma_{kd}(ct)$ (equation (5.40)):

$$\sigma_{ct}(ct) = \frac{\overline{RMS}_k}{RMS_{kd}} \sigma_{kd}(ct)$$

We now calculate $\sigma_{ct}(ct)$ in the case of the B^0 component, since the resolution on ct is needed only for the calculation of the cosine convoluted with the resolution function.

The `RMS` of the sum of n distributions of the mean $\Delta_{k\ell}$ and the `RMS` of $\sigma_{k\ell}$ is

$$\overline{\mathrm{RMS}}_k = \sqrt{\sum_\ell \tilde{\phi}_{k\ell} \left[\sigma_{k\ell}^2 + \Delta_{k\ell}^2\right]} \tag{H.1}$$

where $\tilde{\phi}_{k\ell}$ are the normalized weights from section 5.2.2:

$$\tilde{\phi}_{k\ell} \equiv \frac{e^{-ct/c\tau_{B^0}} \phi_{k\ell} \epsilon_{k\ell}^D}{\Phi_k^0(ct)}$$

The only missing ingredient is $\Delta_{k\ell}$, the shift in the Δct distribution for the decay chain ℓ, which we denote by $\Delta ct_{k\ell}$. This shift occurs due to the use of the average correction $\bar{\mathcal{K}}_k$ instead of $\mathcal{K}_{k\ell}$. From the definitions of ct and Δct

$$ct = L_{\mathrm{xy}}^{B} \frac{m_B}{p_T^{\ell D}} \mathcal{K}_{kd}$$

$$\Delta ct \equiv ct - ct_{true}$$

we see that if we are correcting by $\bar{\mathcal{K}}_k$ instead of $\mathcal{K}_{k\ell}$, the ct of the B candidates from the decay chain ℓ is over-(under-)corrected by $\bar{\mathcal{K}}_k/\mathcal{K}_{k\ell}$. While the optimal correction is $\mathcal{K}_{k\ell}/\mathcal{K}_{kd}$ (for which the $\Delta ct_{k\ell}$ would be centered around zero), we consistently use $\bar{\mathcal{K}}_k/\mathcal{K}_{kd}$ instead, thus introducing a systematic shift of

$$\Delta_{k\ell} = ct\frac{\bar{\mathcal{K}}_k^0}{\mathcal{K}_{kd}} - ct\frac{\mathcal{K}_{k\ell}}{\mathcal{K}_{kd}} = ct\frac{\bar{\mathcal{K}}_k^0}{\mathcal{K}_{kd}}\left[\frac{\mathcal{K}_{k\ell}}{\bar{\mathcal{K}}_k^0} - 1\right] \tag{H.2}$$

Therefore, $\Delta ct_{k\ell}$ of the decay chain ℓ is offset from zero by $\Delta_{k\ell}$ given by eq. (H.2). Putting everything together, we get:

$$\sigma_{ct}(ct) = \frac{\sigma_{kd}(ct)}{RMS_{kd}}\sqrt{\sum_{\ell}\frac{e^{-ct/c\tau_{B^0}}\,\phi_{k\ell}\epsilon_{k\ell}^{D}}{\Phi_k^0(ct)}\left[RMS_{k\ell}^2 + \left(ct\frac{\bar{\mathcal{K}}_k^0}{\mathcal{K}_{kd}}\left(\frac{\mathcal{K}_{k\ell}}{\bar{\mathcal{K}}_k^0} - 1\right)\right)^2\right]} \tag{H.3}$$

Appendix I

Charge asymmetries in the lepton + charm sample

Since the CDF detector is charge asymmetric, a severe charge asymmetry, coupled with an unequal number of ℓ^+ and ℓ^- events, may potentially bias the Same Side Tagging result. Here we show that the charge asymmetry is small in the $\ell D^{(*)}$ sample, and include the same types of distributions as shown in section 6.1, but now divided according to the charge of the SST candidate track or the tagging pion ("+" or "−") instead of the charge correlation between the lepton and the tag.

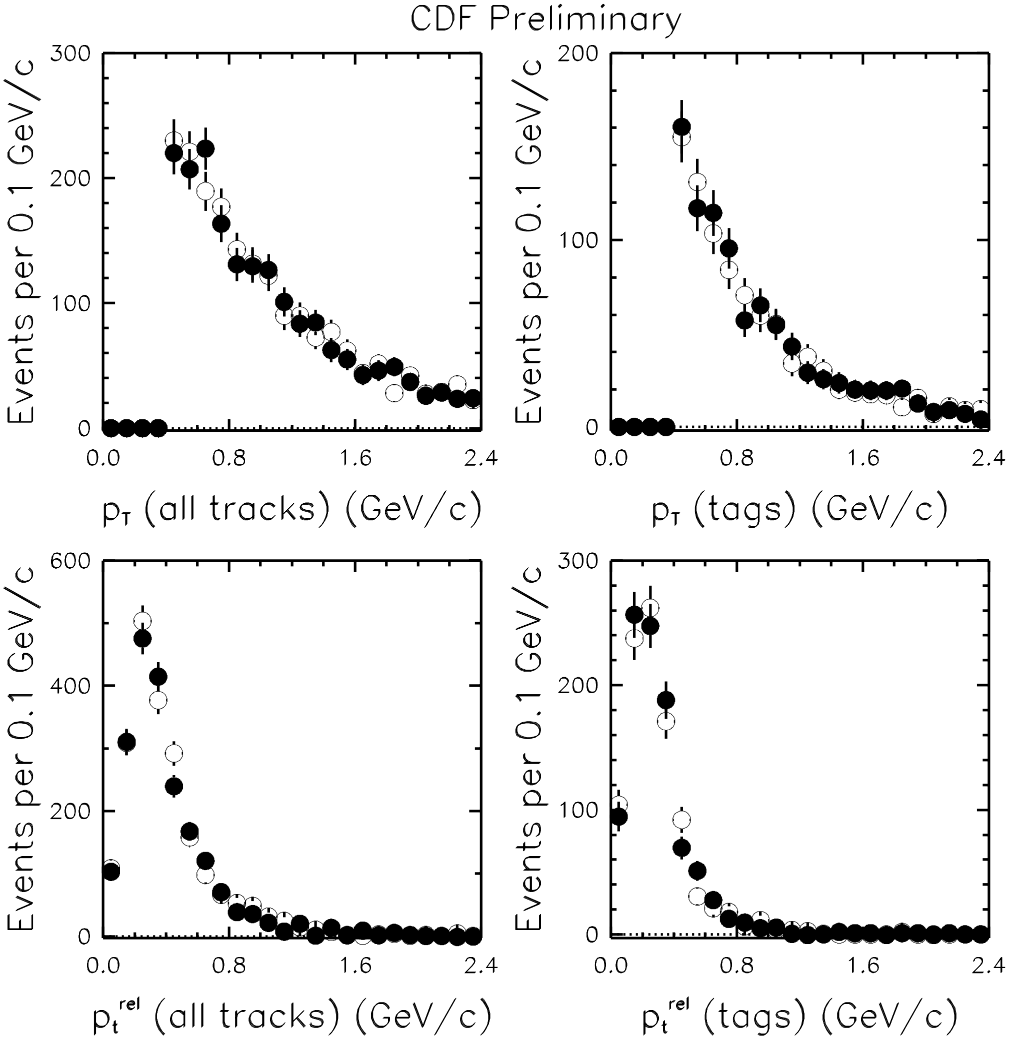


Figure I-1: Tagging distributions for the decay signature "$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$". Solid points are "+ sign" tags, open are "− sign" tags Upper left: p_T of the SST candidate tracks. Upper right: p_T of the tag. Lower left: p_t^{rel} (the value of the tagging function) of the SST candidate tracks. Lower right: p_t^{rel} of the tag.

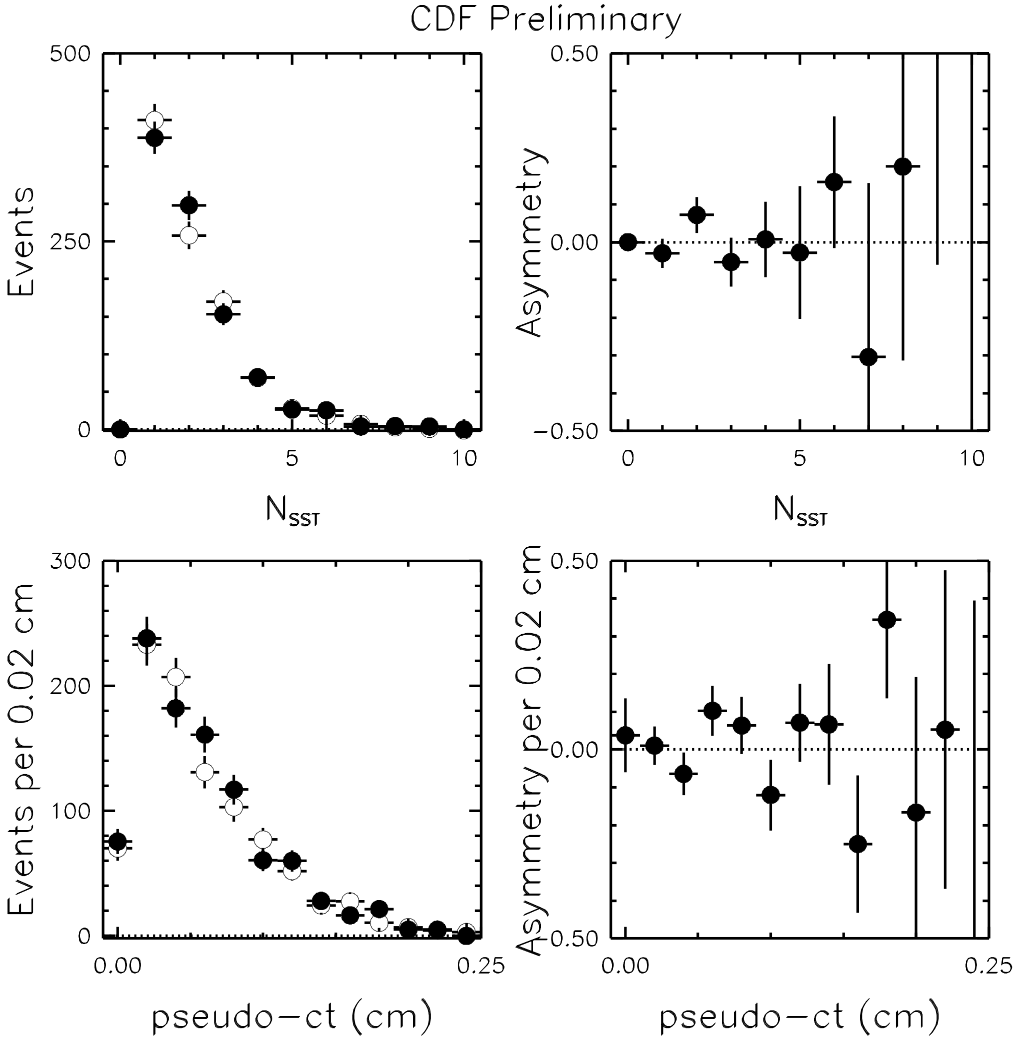


Figure I-2: Tagging distributions for the decay signature "$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$". Solid points are "+ sign" tags, open are "− sign" tags. Upper left: the total number of SST candidate tracks. Upper right: the asymmetry vs the total number of the SST candidate tracks. Lower left: the measured proper decay time (ct) corrected using the $\mathcal{K}$-factor of the 'direct' decay chain. Lower right: the asymmetry vs ct.

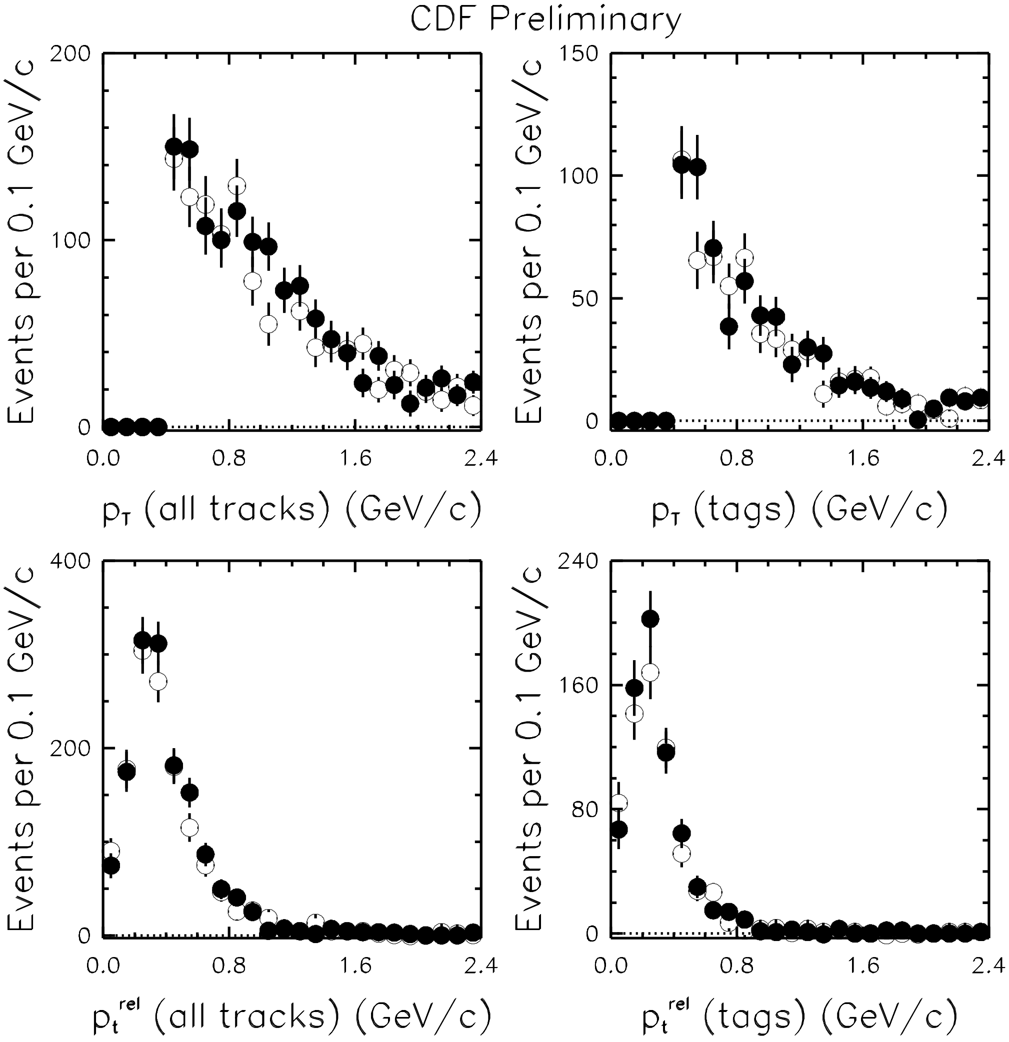


Figure I-3: Tagging distributions for the decay signature “$\ell^+ D^-, D^- \to K\pi\pi$”. Solid points are “+ sign” tags, open are “− sign” tags Upper left: p_T of the SST candidate tracks. Upper right: p_T of the tag. Lower left: p_t^{rel} (the value of the tagging function) of the SST candidate tracks. Lower right: p_t^{rel} of the tag.

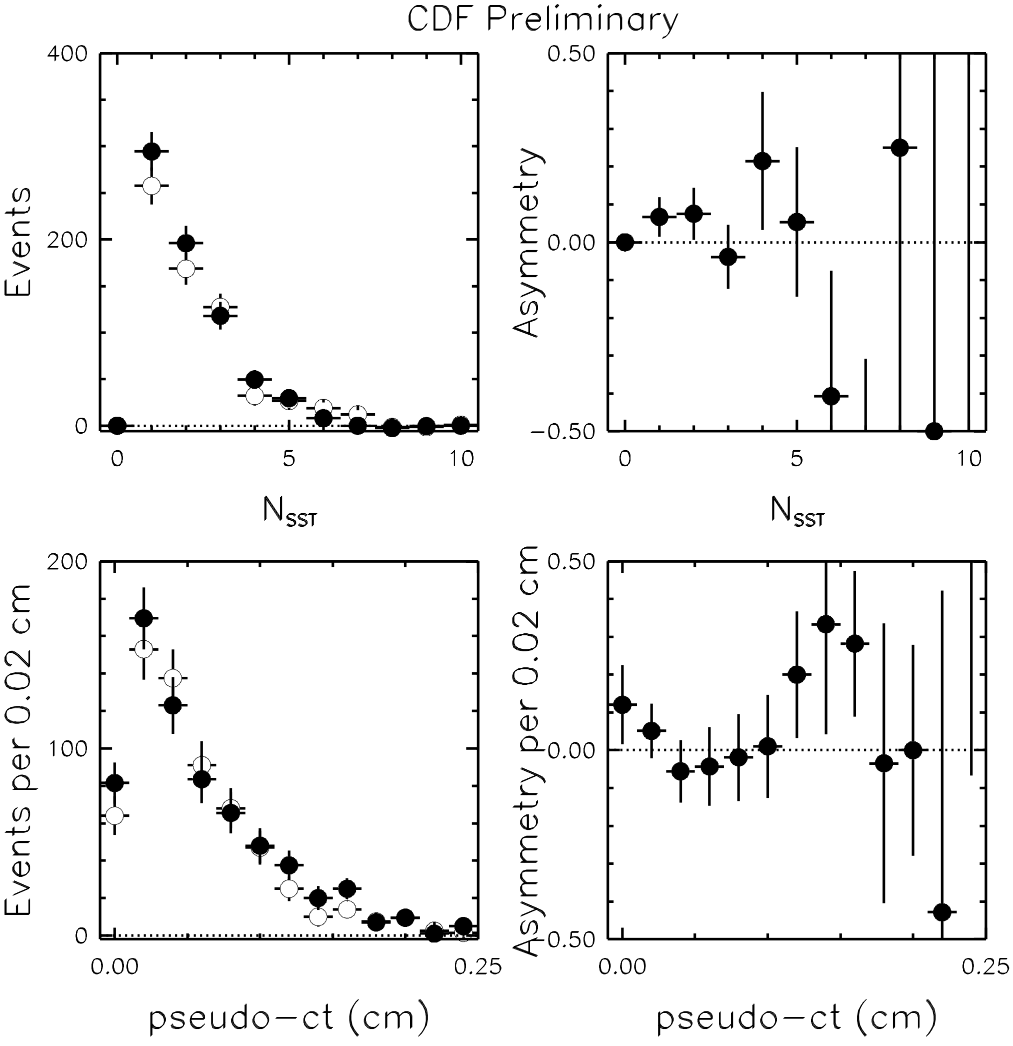


Figure I-4: Tagging distributions for the decay signature "$\ell^+D^-, D^- \rightarrow K\pi\pi$". Solid points are "+ sign" tags, open are "− sign" tags. Upper left: the total number of SST candidate tracks. Upper right: the asymmetry vs the total number of the SST candidate tracks. Lower left: the measured proper decay time (ct) corrected using the $\mathcal{K}$-factor of the 'direct' decay chain. Lower right: the asymmetry vs ct.

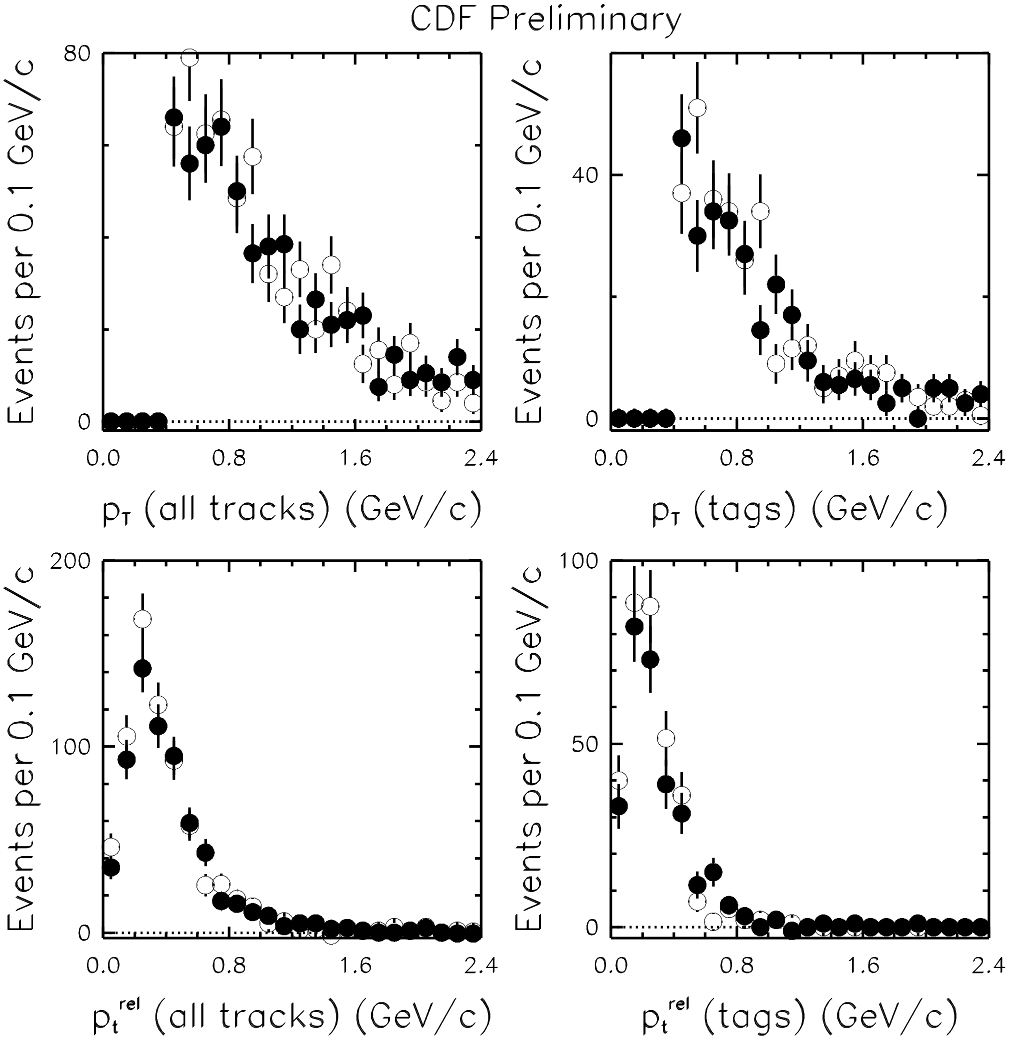


Figure I-5: Tagging distributions for the decay signature "$\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi$". Solid points are "+ sign" tags, open are "− sign" tags Upper left: p_T of the SST candidate tracks. Upper right: p_T of the tag. Lower left: p_t^{rel} (the value of the tagging function) of the SST candidate tracks. Lower right: p_t^{rel} of the tag.

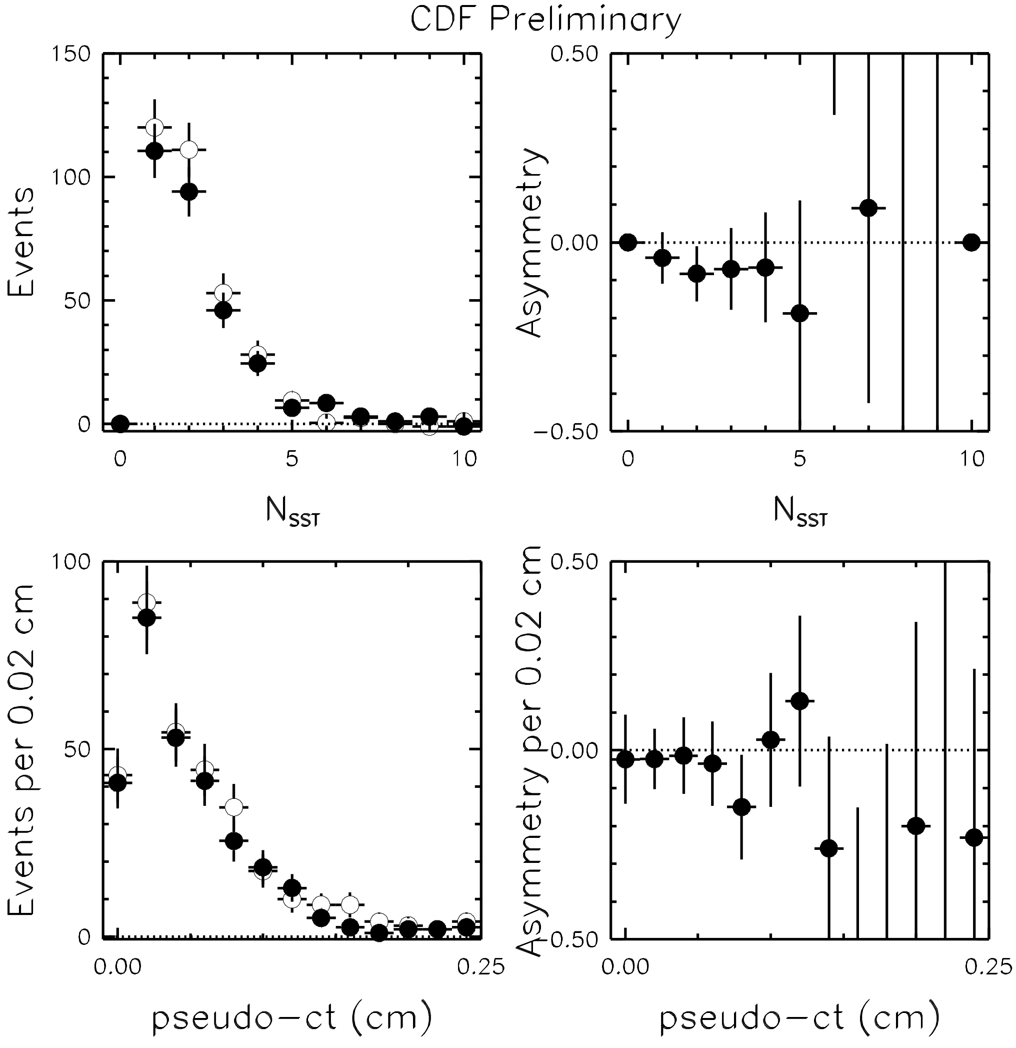


Figure I-6: Tagging distributions for the decay signature “$\ell^+ D^{*-}, \overline{D^0} \to K\pi$”. Solid points are “+ sign” tags, open are “− sign” tags. Upper left: the total number of SST candidate tracks. Upper right: the asymmetry vs the total number of the SST candidate tracks. Lower left: the measured proper decay time (ct) corrected using the $\mathcal{K}$-factor of the ‘direct’ decay chain. Lower right: the asymmetry vs ct.

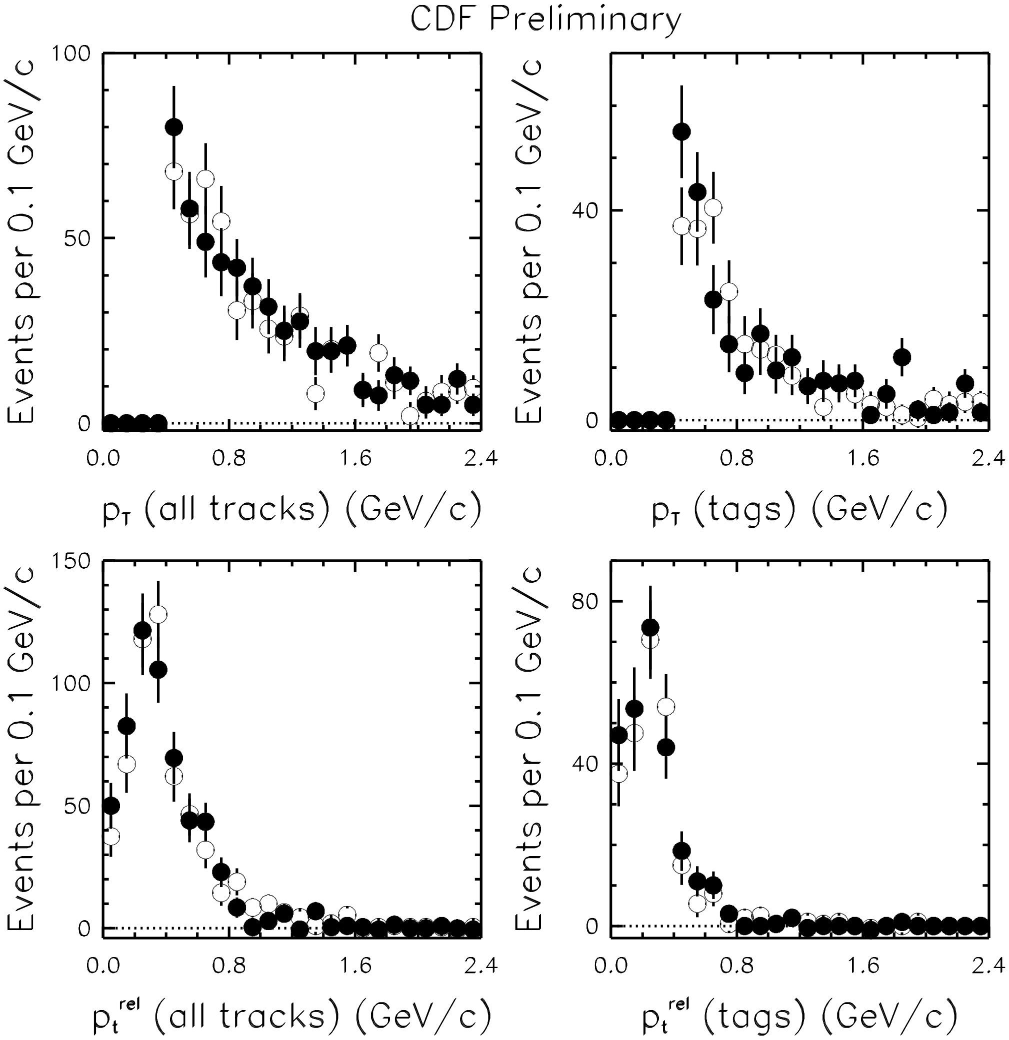


Figure I-7: Tagging distributions for the decay signature "$\ell^+ D^{*-}, \overline{D^0} \to K3\pi$". Solid points are "+ sign" tags, open are "− sign" tags Upper left: p_T of the SST candidate tracks. Upper right: p_T of the tag. Lower left: p_t^{rel} (the value of the tagging function) of the SST candidate tracks. Lower right: p_t^{rel} of the tag.

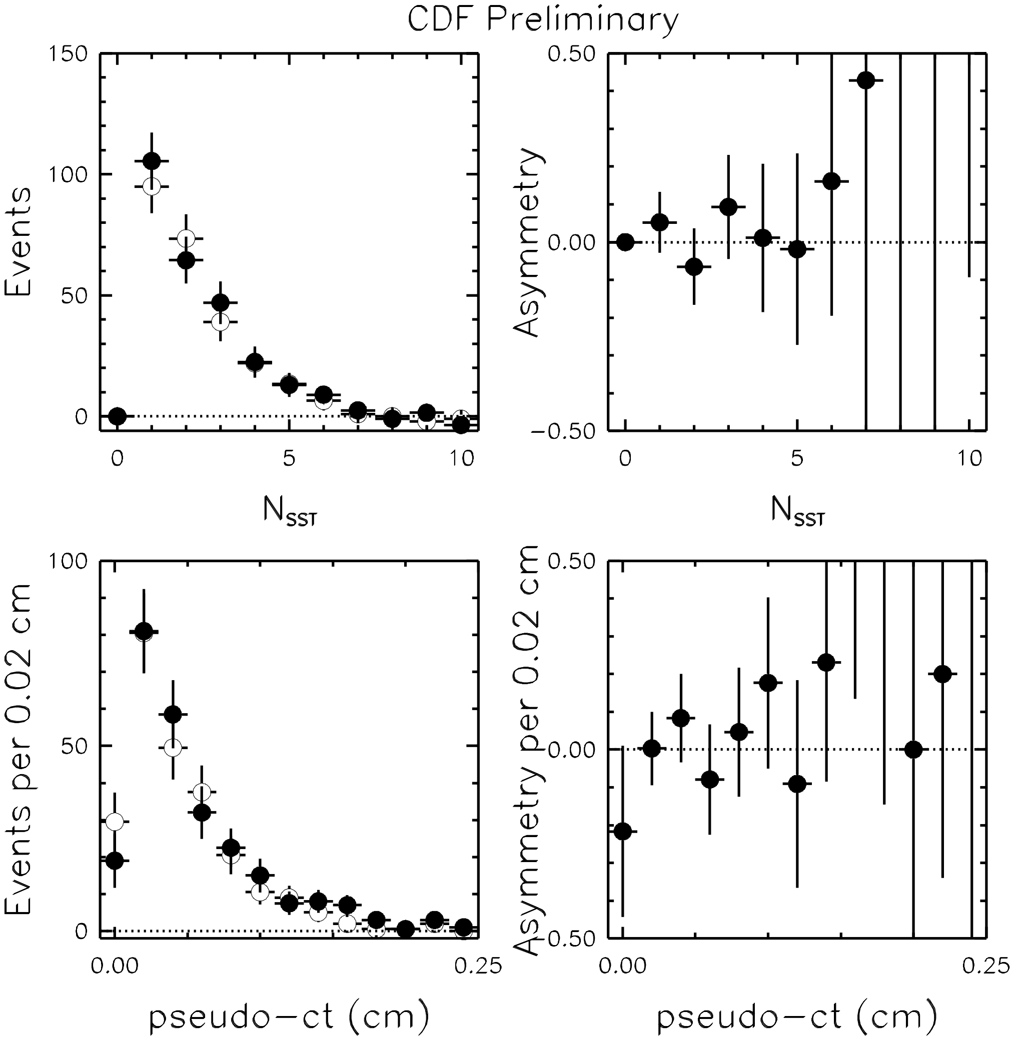


Figure I-8: Tagging distributions for the decay signature "$\ell^+ D^{*-}, \overline{D^0} \rightarrow K3\pi$". Solid points are "+ sign" tags, open are "− sign" tags. Upper left: the total number of SST candidate tracks. Upper right: the asymmetry vs the total number of the SST candidate tracks. Lower left: the measured proper decay time (ct) corrected using the $\mathcal{K}$-factor of the 'direct' decay chain. Lower right: the asymmetry vs ct.

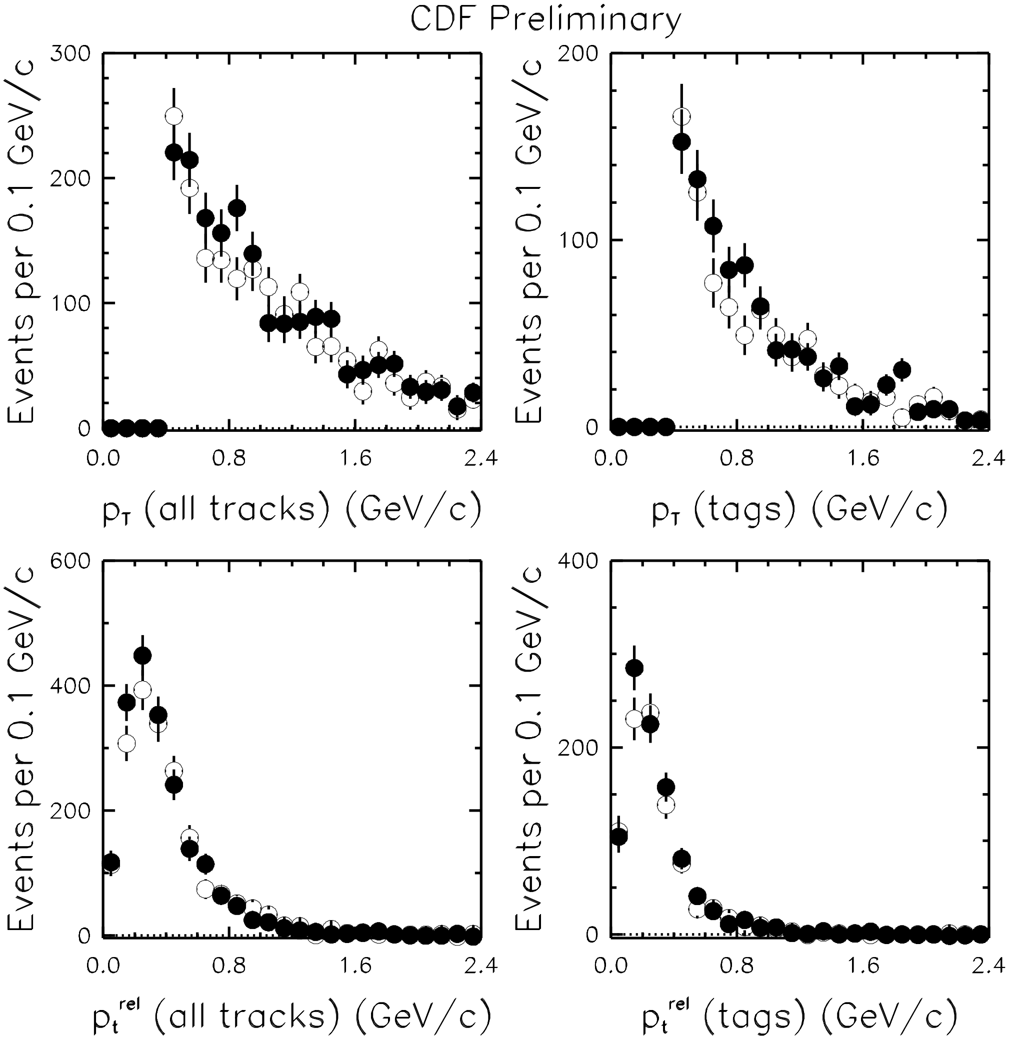


Figure I-9: Tagging distributions for the decay signature "$\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$". Solid points are "+ sign" tags, open are "− sign" tags Upper left: p_T of the SST candidate tracks. Upper right: p_T of the tag. Lower left: p_t^{rel} (the value of the tagging function) of the SST candidate tracks. Lower right: p_t^{rel} of the tag.

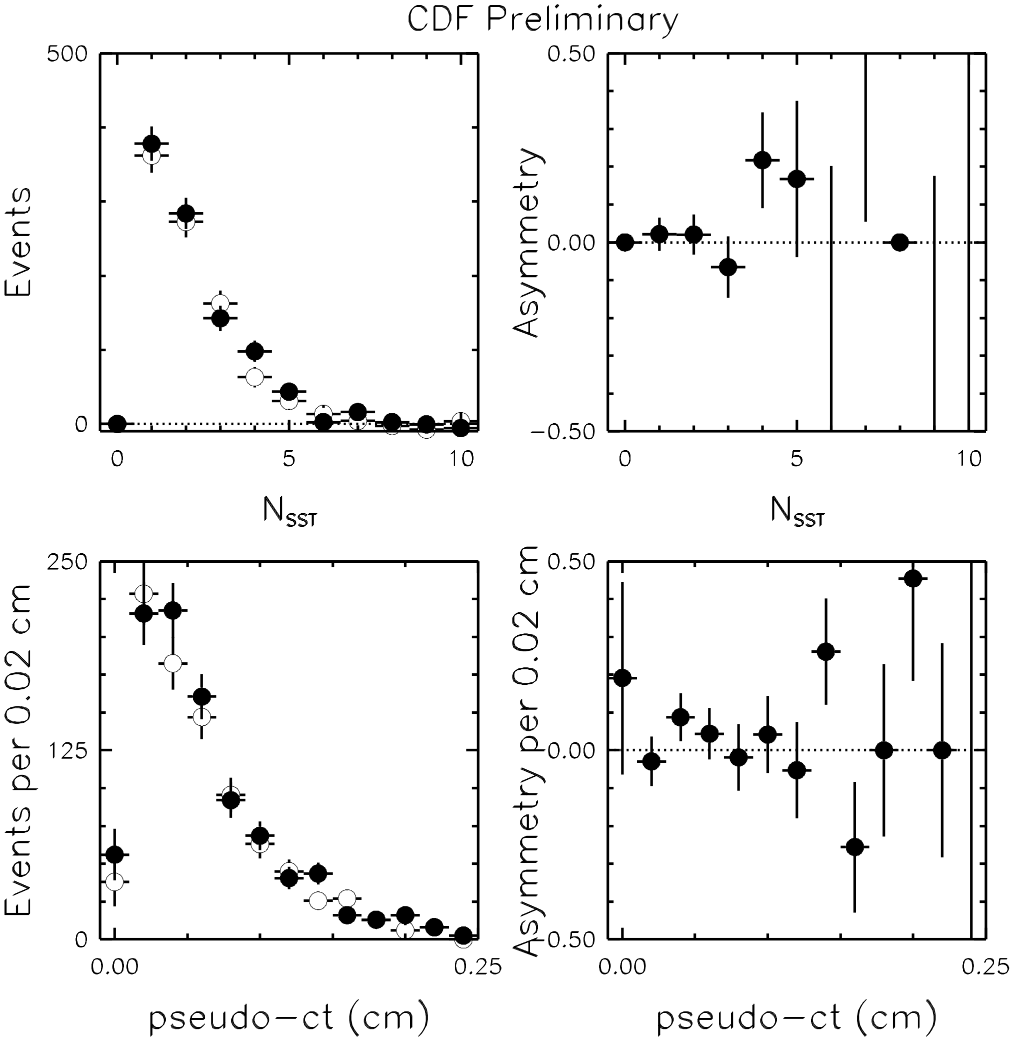


Figure I-10: Tagging distributions for the decay signature "$\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$". Solid points are "+ sign" tags, open are "− sign" tags. Upper left: the total number of SST candidate tracks. Upper right: the asymmetry vs the total number of the SST candidate tracks. Lower left: the measured proper decay time (ct) corrected using the $\mathcal{K}$-factor of the 'direct' decay chain. Lower right: the asymmetry vs ct.

Appendix J

Contribution of $g \to c\bar{c}$ to lepton + charm sample

Here we include the results of the fits to the distribution of the charm lifetime with respect to the "B vertex", ct_{D-B}. We first note that in $g \to c\bar{c} \to \ell D^{(*)}$ the "B vertex" is not consistent with B meson decay. In this case, what we assume to be the "B vertex" is merely an intersection point of the charm track with the direction of motion of the D meson. The vertexing algorithmallows this point to be anywhere along the D direction, not necessarily between the primary and the D vertex, as it would be the case for a properly reconstructed B meson. For this reason, in addition to a good B vertex quality requirements, our event selection demands that the charm proper time with respect to the "B vertex" (denoted by ct_{D-B}) is consistent with the D meson lifetime (ct_{D-B} must be between -500μm and 1000μm for D^0 or 2000μm for D^+)). If this cut is removed, there is a small number of "B vertices" that are in the tails of the ct_{D-B} distribution: possible sources are the CTC tracks wrongly matched to the SVX tracks, as well as $g \to c\bar{c}$. In order to show that in the tails of the ct_{D-B} distribution there is still some lepton + charm signal left, we make the standard D-meson mass distributions, however we require $ct_{D-B} < -1$mm. The D-meson mass distributions are shown in figures J-1, J-3, J-5, J-7 and J-9 (upper left) and they exhibit an excess of the the RS over the WS, showing the existence of *real* $\ell^+ D^{(*)}$ signal. Figures J-1, J-3, J-5, J-7 and J-9 (bottom) show the background-subtracted distribution of ct_{D-PV}, the

D meson proper time with respect to the Primary Vertex. In some decay signatures (e.g. "$\ell^+ D^-, D^- \to K\pi\pi$", "$\ell^+ D^{*-}, \overline{D^0} \to K3\pi$" and "$\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$") a semi-significant excess of D-mesons originating at the primary vertex is visible.

The distributions of the full ct_{D-B} domain (zooming in onto the tails) for all decay signatures, are shown in figures J-2, J-4, J-6, J-8 and J-10. Because of the incorrect B-decay vertex of the $g \to c\bar{c}$ events, we expect their ct_{D-B} distribution to be very wide, compared to the real $B \to D$ events. Our goal is just to estimate the upper limit of the contribution of $g \to c\bar{c}$ to the lepton + charm sample, so as a prototype of a wide distribution we use a very wide gaussian,[1] and fit the ct_{D-B} distribution with the sum of this gaussian, and a template derived from the $B \to \nu\ell D^{(*)}$ Monte Carlo simulation. In order to account for the low statistics in the tails, we use a likelihood fit, and fit the distribution for the signal region with a combination of three shapes:

- the signal template from the Monte Carlo simulation,
- distribution from the sidebands, and
- a wide gaussian, with fixed $\sigma = 400\mu$m.

The width of the gaussian is fixed in order to pick the gluon splitting events from the tail. (Letting the width float in the fit may result in getting a narrow gaussian beneath the center of the ct_{D-B} distribution, resulting in an artificially inflated number of gluon splitting events.) The $B \to \nu\ell D^{(*)}$ Monte Carlo simulation includes other $B/B_s \to \nu\ell D^{(*)}X$ processes mentioned above, hence any deviation of the ct_{D-B} distributions from the Monte Carlo model – measured by the area of the wide gaussian – consists of the combination of $g \to c\bar{c}$ and poorly matched CTC and SVX tracks for the lepton. One can remove most of the latter by imposing a SVX quality requirements. (The χ^2_{total} of an SVX track includes the $\chi^2_{CTC-SVX}$ coming from associating the CTC track to the hits in the SVX.) Nevertheless, it is conservative to assume that *all* of the excess comes from the gluon splitting.

---

[1] The choice of the gaussian is fairly arbitrary, since the total number of $g \to c\bar{c}$ events in the tails is nevertheless small. In the end, to be conservative, we inflate the resulting number of $g \to c\bar{c}$ events by 2σ.

After the fit, the area under the gaussian is integrated in the range of ct_{D-B} that is used in the event selection. Table J.1 summarizes the contribution to each decay signature. As an upper limit on possible gluon splitting events, we assume the fraction which corresponds to the 2σ upward fluctuation, that is, $(N(g \to c\bar{c}) + 2\sigma(g \to c\bar{c}))/N(B \to \nu\ell D^{(*)}X)$.

Signature	$N(g \to c\bar{c} \to \ell D^{(*)})$	$N(B \to \nu\ell D^{(*)}X)$	'2σ fraction'
$\ell^+\overline{D^0}, \overline{D^0} \to K\pi$	4.8 ± 3.0	2192 ± 62	0.005
$\ell^+D^-, D^- \to K\pi\pi$	3.0 ± 2.6	1679 ± 61	0.005
$\ell^+D^{*-}, \overline{D^0} \to K\pi$	0.5 ± 0.5	798 ± 33	0.002
$\ell^+D^{*-}, \overline{D^0} \to K3\pi$	1.3 ± 0.9	667 ± 36	0.005
$\ell^+D^{*-}, \overline{D^0} \to K\pi\pi^0$	18 ± 15	2474 ± 94	0.019

Table J.1: Contribution of gluon splitting, $g \to c\bar{c}$, to the lepton + charm sample.

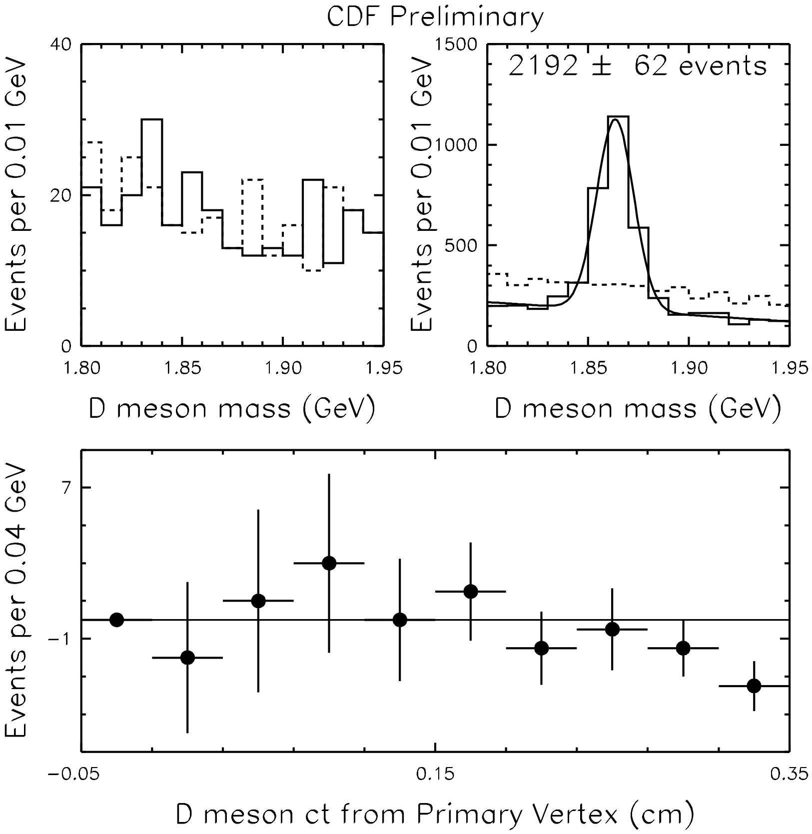
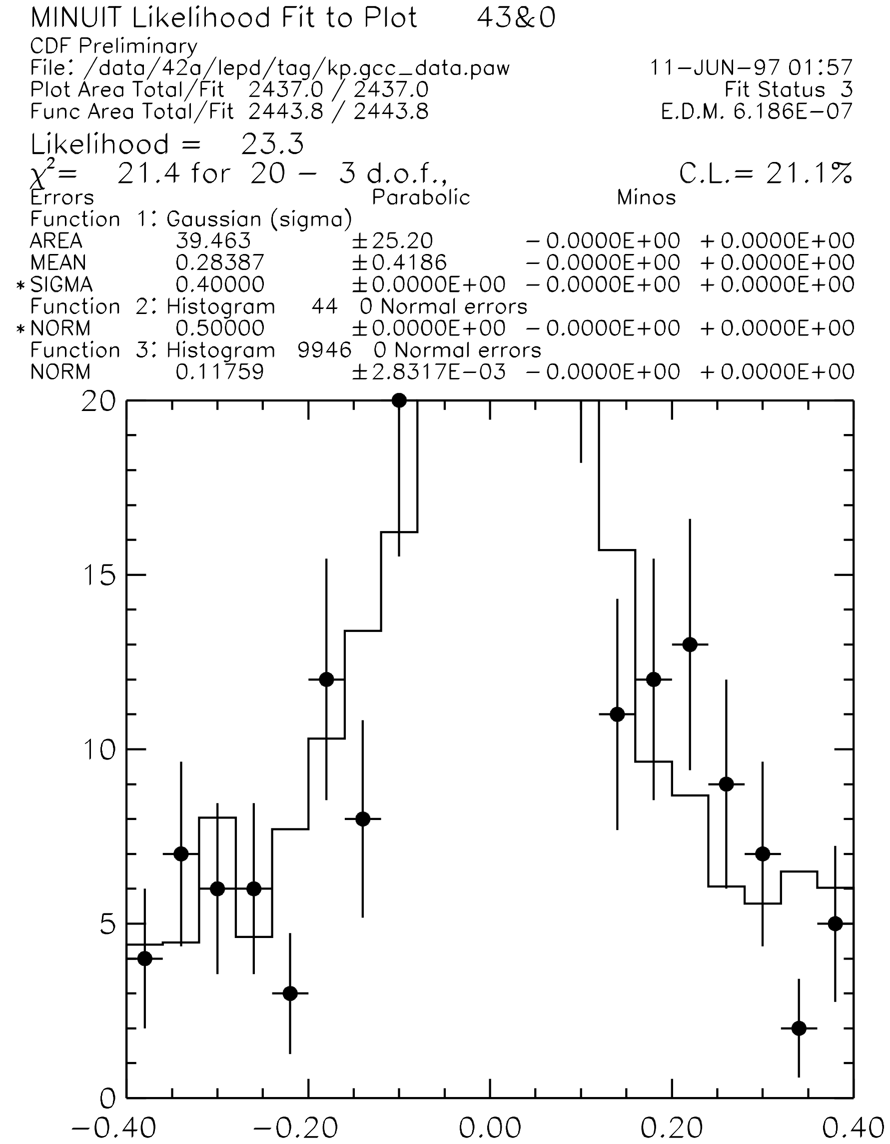


Figure J-1: The $\ell^+\overline{D^0}, \overline{D^0} \rightarrow K\pi$ decay signature: upper left: D mass for $ct_{D-B} < -1$mm, right sign (solid) and wrong sign (dashed). Upper right: D mass for all events that pass the quality cuts. Bottom: ct_{D-PV}, the D meson proper time with respect to the Primary Vertex.



```
MINUIT Likelihood Fit to Plot      43&0
CDF Preliminary
File: /data/42a/lepd/tag/kp.gcc_data.paw                 11-JUN-97 01:57
Plot Area Total/Fit   2437.0 / 2437.0                       Fit Status  3
Func Area Total/Fit  2443.8 / 2443.8                     E.D.M. 6.186E-07
Likelihood =    23.3
χ²=   21.4 for  20 -  3 d.o.f.,                         C.L.= 21.1%
Errors                              Parabolic              Minos
Function  1: Gaussian (sigma)
AREA          39.463          ±25.20          - 0.0000E+00  + 0.0000E+00
MEAN          0.28387         ±0.4186         - 0.0000E+00  + 0.0000E+00
*SIGMA        0.40000         ±0.0000E+00     - 0.0000E+00  + 0.0000E+00
Function  2: Histogram     44   0 Normal errors
*NORM         0.50000         ±0.0000E+00     - 0.0000E+00  + 0.0000E+00
Function  3: Histogram   9946   0 Normal errors
NORM          0.11759         ±2.8317E-03     - 0.0000E+00  + 0.0000E+00
```

Figure J-2: The results of the fits to the distribution of the charm lifetime with respect to the "B vertex", ct_{D-B}, for the $\ell^+\overline{D^0}, \overline{D^0} \to K\pi$ decay signature.

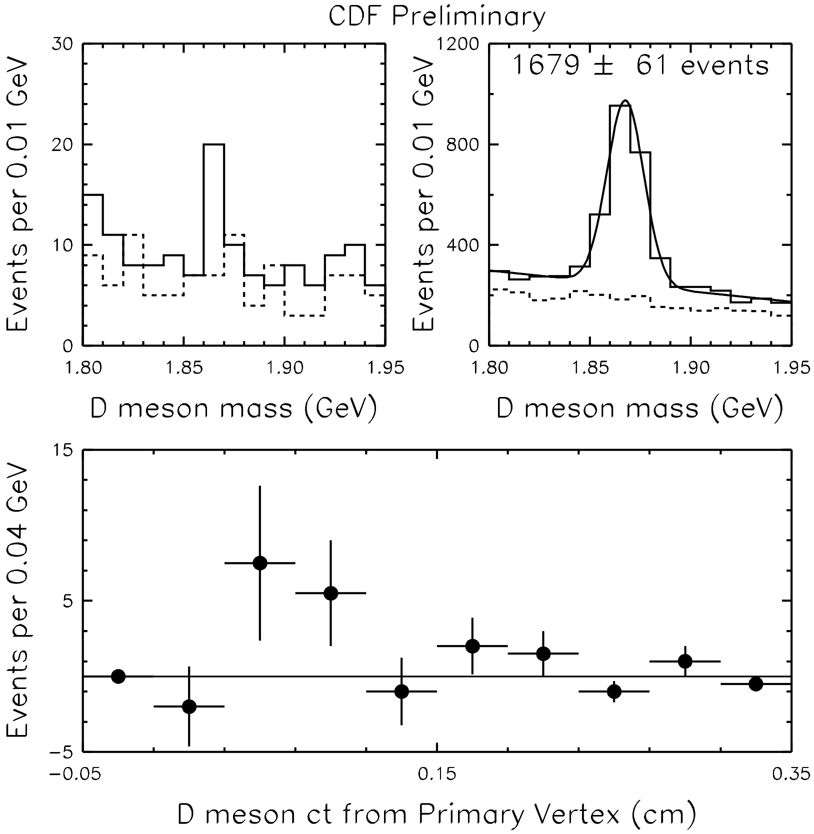
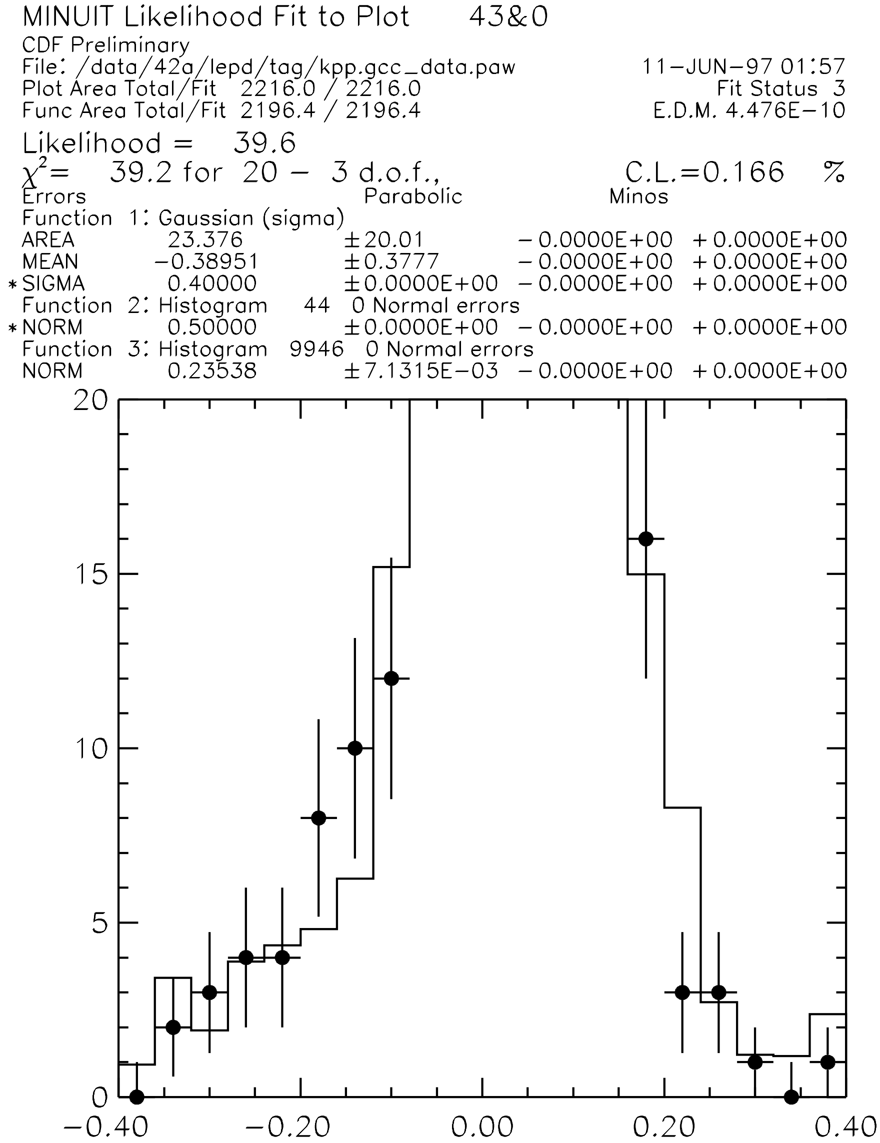


Figure J-3: Distributions of the $g \rightarrow c\bar{c}$ candidates for the $\ell^+ D^-, D^- \rightarrow K\pi\pi$ decay signature: upper left: D mass for $ct_{D-B} < -1$mm, right sign (solid) and wrong sign (dashed). Upper right: D mass for all events that pass the quality cuts. Bottom: ct_{D-PV}, the D meson proper time with respect to the Primary Vertex.



```
MINUIT Likelihood Fit to Plot      43&0
CDF Preliminary
File: /data/42a/lepd/tag/kpp.gcc_data.paw                11-JUN-97 01:57
Plot Area Total/Fit   2216.0 / 2216.0                         Fit Status  3
Func Area Total/Fit  2196.4 / 2196.4                        E.D.M. 4.476E-10
Likelihood =     39.6
χ²=    39.2 for  20 -  3 d.o.f.,                       C.L.=0.166    %
Errors                              Parabolic                   Minos
Function  1: Gaussian (sigma)
AREA          23.376            ±20.01          - 0.0000E+00   + 0.0000E+00
MEAN         -0.38951           ±0.3777         - 0.0000E+00   + 0.0000E+00
*SIGMA        0.40000           ±0.0000E+00     - 0.0000E+00   + 0.0000E+00
Function  2: Histogram     44    0 Normal errors
*NORM         0.50000           ±0.0000E+00     - 0.0000E+00   + 0.0000E+00
Function  3: Histogram   9946    0 Normal errors
NORM          0.23538           ±7.1315E-03     - 0.0000E+00   + 0.0000E+00
```

Figure J-4: The results of the fits to the distribution of the charm lifetime with respect to the "B vertex", ct_{D-B}, for the $\ell^+D^-, D^- \rightarrow K\pi\pi$ decay signature.

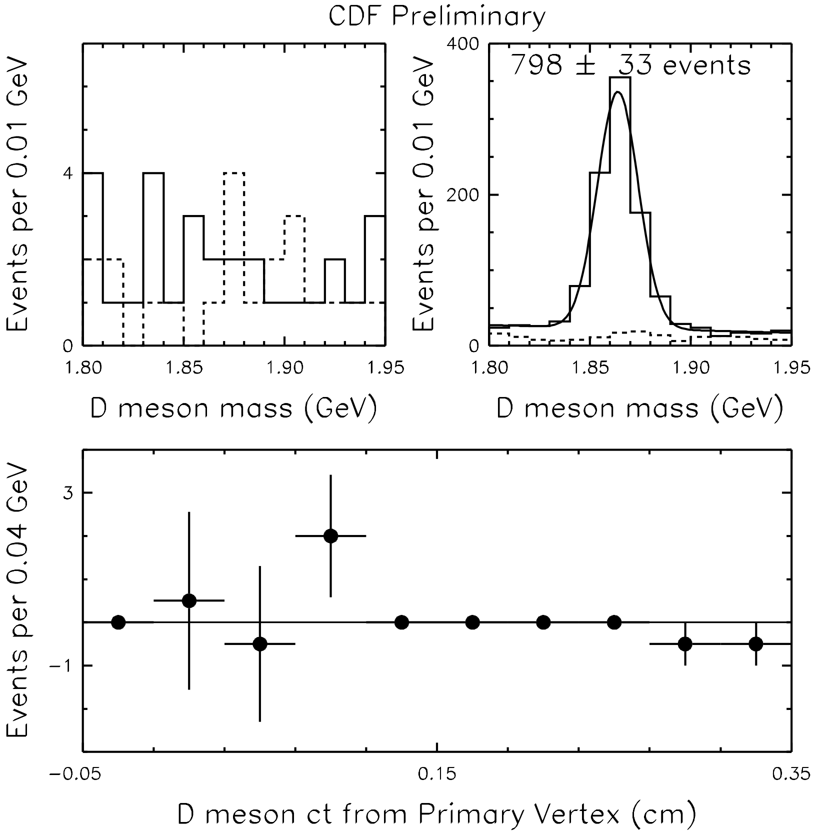
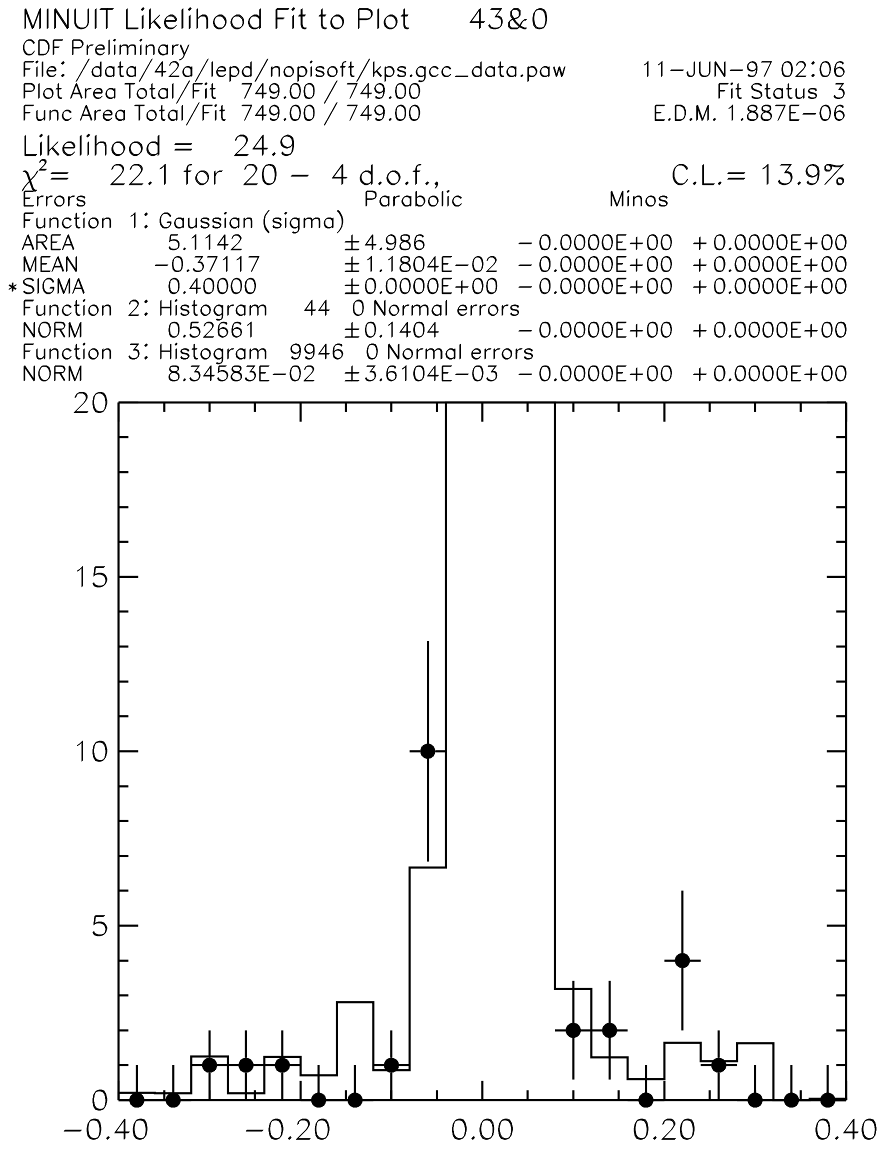


Figure J-5: Distributions of the $g \to c\bar{c}$ candidates for the $\ell^+ D^{*-}$, $\overline{D^0} \to K\pi$ decay signature: upper left: D mass for $ct_{D-B} < -1$mm, right sign (solid) and wrong sign (dashed). Upper right: D mass for all events that pass the quality cuts. Bottom: ct_{D-PV}, the D meson proper time with respect to the Primary Vertex.



```
MINUIT Likelihood Fit to Plot      43&0
CDF Preliminary
File: /data/42a/lepd/nopisoft/kps.gcc_data.paw          11-JUN-97 02:06
Plot Area Total/Fit   749.00 / 749.00                      Fit Status  3
Func Area Total/Fit  749.00 / 749.00                    E.D.M. 1.887E-06
Likelihood =     24.9
χ²=    22.1 for  20 -  4 d.o.f.,                         C.L.= 13.9%
Errors                          Parabolic              Minos
Function  1: Gaussian (sigma)
AREA          5.1142          ±4.986         - 0.0000E+00  + 0.0000E+00
MEAN         -0.37117         ±1.1804E-02    - 0.0000E+00  + 0.0000E+00
*SIGMA        0.40000         ±0.0000E+00    - 0.0000E+00  + 0.0000E+00
Function  2: Histogram    44   0 Normal errors
NORM          0.52661         ±0.1404        - 0.0000E+00  + 0.0000E+00
Function  3: Histogram  9946   0 Normal errors
NORM          8.34583E-02    ±3.6104E-03    - 0.0000E+00  + 0.0000E+00
```

Figure J-6: The results of the fits to the distribution of the charm lifetime with respect to the "B vertex", ct_{D-B}, for the $\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi$ decay signature.

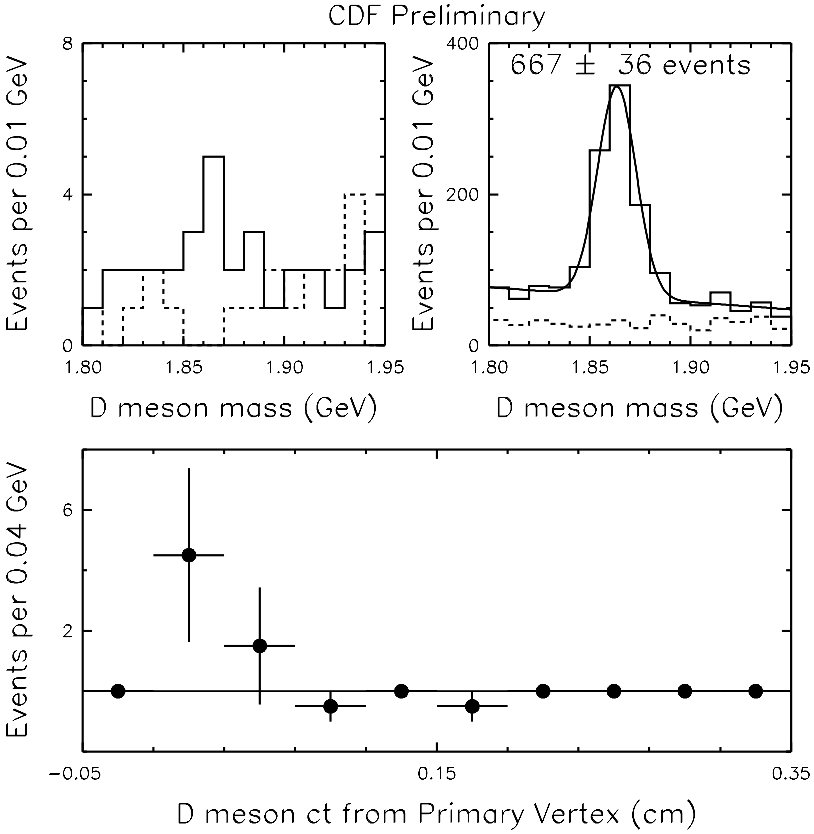
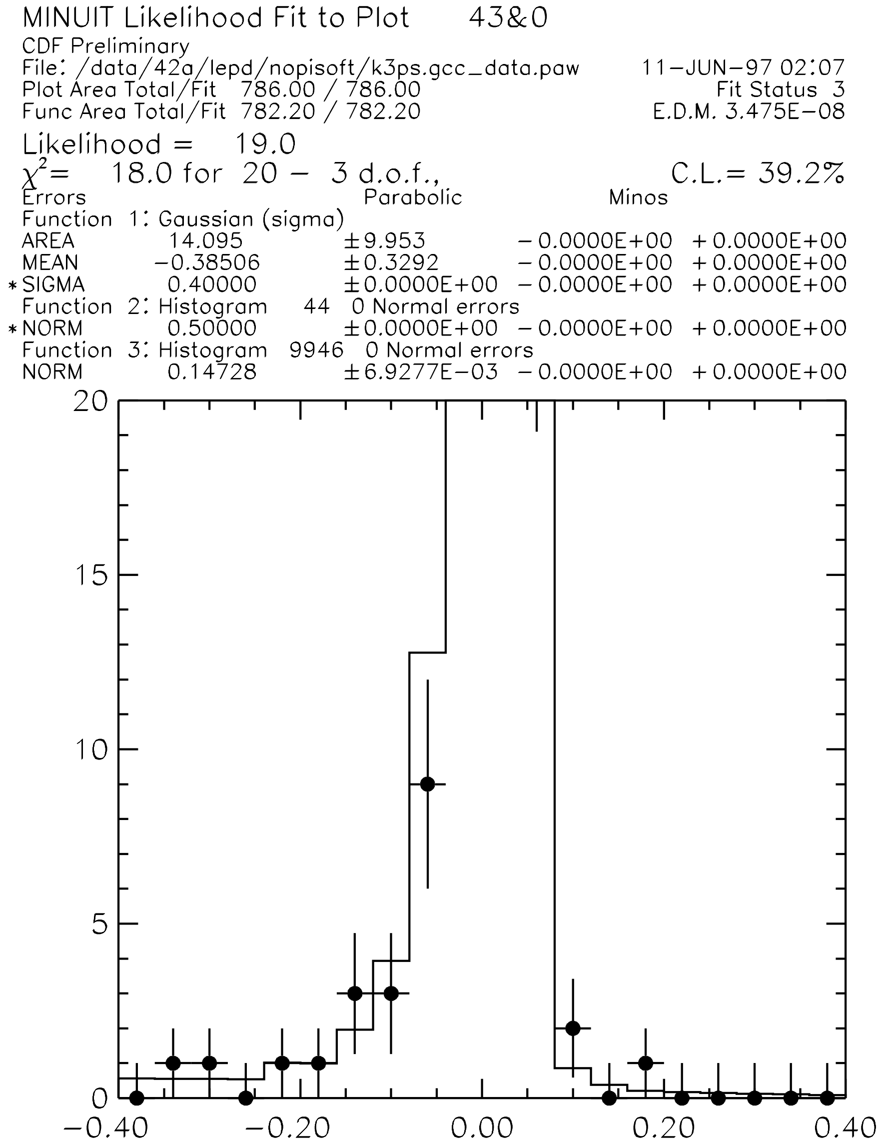


Figure J-7: Distributions of the $g \rightarrow c\bar{c}$ candidates for the $\ell^+ D^{*-}, \overline{D^0} \rightarrow K3\pi$ decay signature: upper left: D mass for $ct_{D-B} < -1\text{mm}$, right sign (solid) and wrong sign (dashed). Upper right: D mass for all events that pass the quality cuts. Bottom: ct_{D-PV}, the D meson proper time with respect to the Primary Vertex.



```
MINUIT Likelihood Fit to Plot      43&0
CDF Preliminary
File: /data/42a/lepd/nopisoft/k3ps.gcc_data.paw        11-JUN-97 02:07
Plot Area Total/Fit   786.00 / 786.00                       Fit Status  3
Func Area Total/Fit  782.20 / 782.20                    E.D.M. 3.475E-08
Likelihood =    19.0
χ²=    18.0 for  20 -  3 d.o.f.,                        C.L.= 39.2%
Errors                          Parabolic                 Minos
Function  1: Gaussian (sigma)
AREA          14.095          ±9.953         - 0.0000E+00  + 0.0000E+00
MEAN         -0.38506         ±0.3292        - 0.0000E+00  + 0.0000E+00
*SIGMA        0.40000         ±0.0000E+00    - 0.0000E+00  + 0.0000E+00
Function  2: Histogram     44   0 Normal errors
*NORM         0.50000         ±0.0000E+00    - 0.0000E+00  + 0.0000E+00
Function  3: Histogram   9946   0 Normal errors
NORM          0.14728         ±6.9277E-03    - 0.0000E+00  + 0.0000E+00
```

Figure J-8: The results of the fits to the distribution of the charm lifetime with respect to the "B vertex", ct_{D-B}, for the $\ell^+ D^{*-}, \overline{D^0} \to K3\pi$ decay signature.

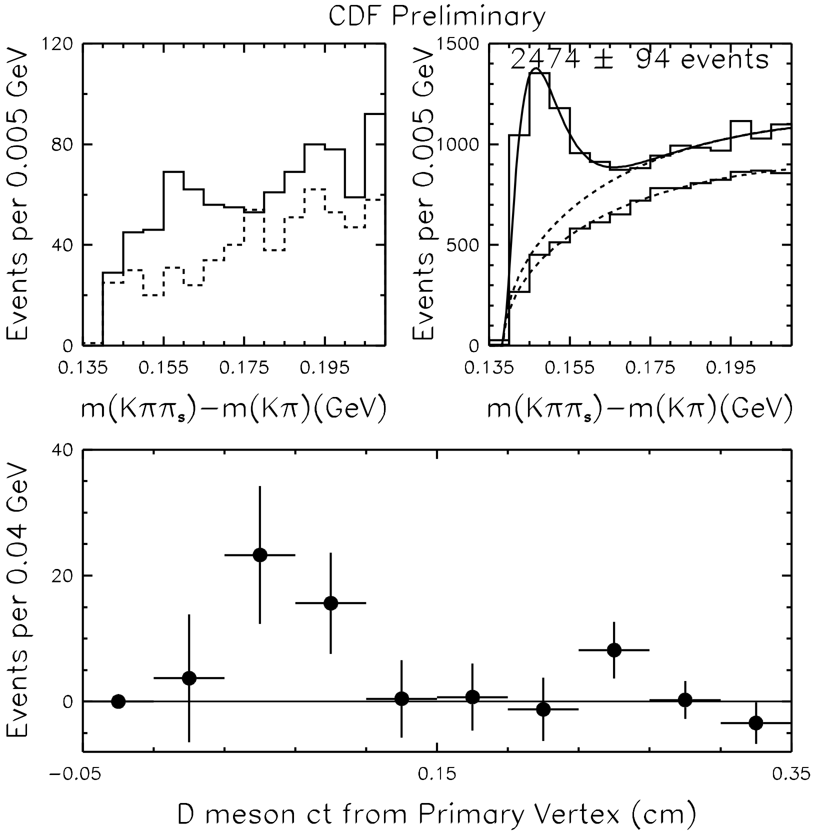
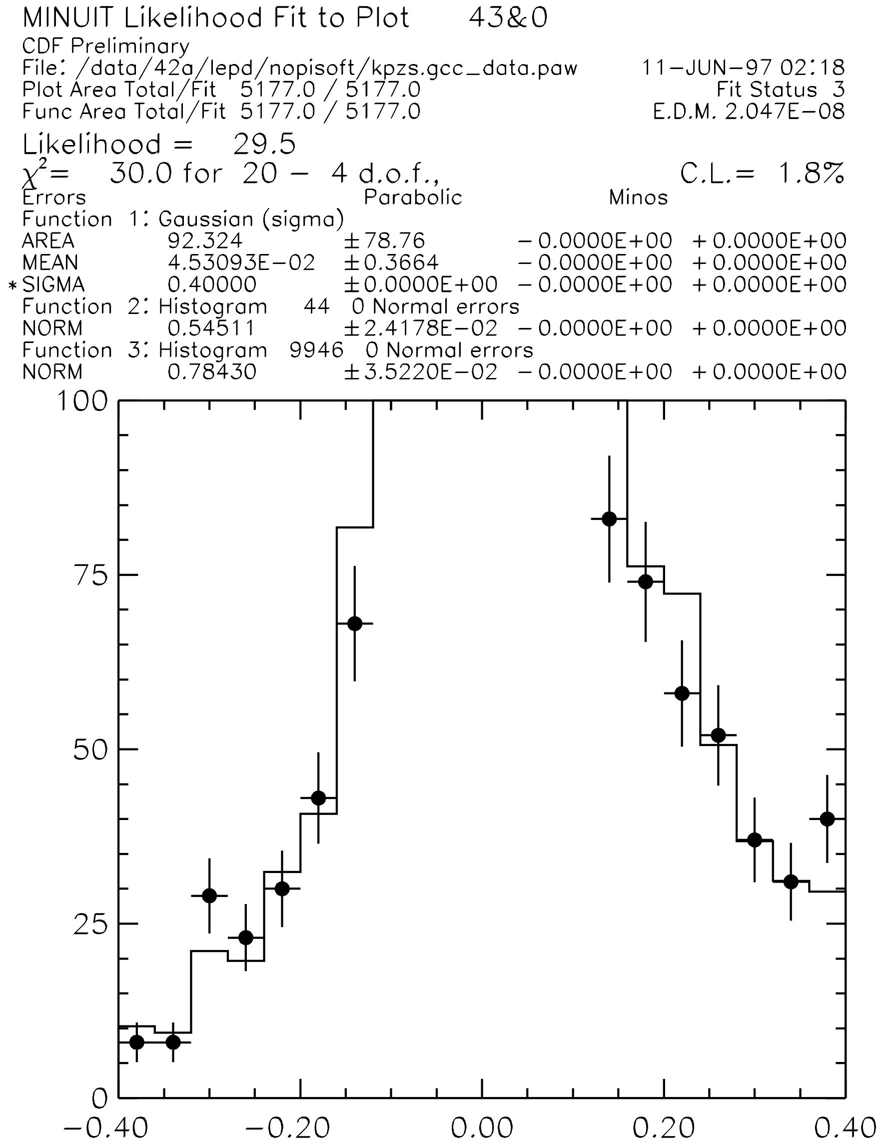


Figure J-9: Distributions of the $g \to c\bar{c}$ candidates for the $\ell^+ D^{*-}, \overline{D^0} \to K\pi\pi^0$ decay signature: upper left: D mass for $ct_{D-B} < -1$mm, right sign (solid) and wrong sign (dashed). Upper right: D mass for all events that pass the quality cuts. Bottom: ct_{D-PV}, the D meson proper time with respect to the Primary Vertex.



```
MINUIT Likelihood Fit to Plot      43&0
CDF Preliminary
File: /data/42a/lepd/nopisoft/kpzs.gcc_data.paw        11-JUN-97 02:18
Plot Area Total/Fit   5177.0 / 5177.0                       Fit Status  3
Func Area Total/Fit  5177.0 / 5177.0                   E.D.M. 2.047E-08
Likelihood =    29.5
χ²=    30.0 for  20 -  4 d.o.f.,                        C.L.=  1.8%
Errors                          Parabolic                Minos
Function  1: Gaussian (sigma)
AREA          92.324        ±78.76         -0.0000E+00  +0.0000E+00
MEAN          4.53093E-02   ±0.3664        -0.0000E+00  +0.0000E+00
*SIGMA        0.40000       ±0.0000E+00    -0.0000E+00  +0.0000E+00
Function  2: Histogram     44   0 Normal errors
NORM          0.54511       ±2.4178E-02    -0.0000E+00  +0.0000E+00
Function  3: Histogram   9946   0 Normal errors
NORM          0.78430       ±3.5220E-02    -0.0000E+00  +0.0000E+00
```

Figure J-10: The results of the fits to the distribution of the charm lifetime with respect to the "B vertex", ct_{D-B}, for the $\ell^+ D^{*-}, \overline{D^0} \rightarrow K\pi\pi^0$ decay signature.

Bibliography

[1] T. D. Lee and C. N. Yang, Phys. Rev. **104**, 254 (1956).

[2] C. S. Wu *et al.*, Phys. Rev. **105**, 1413 (1957).

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

[4] L. Gibbons, *Experimental Status of* $|V_{ub}|$, presented at the 7th International Symposium on Heavy Flavor Physics, Santa Barbara, CA, July 7-11 1997, to appear in the proceedings.

[5] A. Schopper, *Results from CPLEAR on Tests of CP, T, and CPT using* K^0 *and* $\bar{K}^0$, presented at the 2nd International Conference on B Physics and CP Violation, Honolulu, Hawaii, March 24-27, 1997, to appear in the proceedings.

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

[7] S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev.**D2**, 1585 (1970).

[8] M. Kobayashi and T. Maskawa, Progr. Theor. Phys. **49**, 652 (1973).

[9] L. Wolfenstein, Phys. Rev. Lett. **51**, 1915 (1983).

[10] L. L. Chau and W. Y. Keung, Phys. Rev. Lett.**53**, 1802 (1984).

[11] A. Sakharov, JETP Lett. **5**, 24 (1967).

[12] Currently, most theorists favor the hypothesis that the Standard Model cannot generate enough CP violation to explain the baryon asymmetry; for recent estimates, see

P. Huet and E. Sather, Phys. Rev. D **51**, 379 (1995). For the alternative view that the single CKM phase in the Standard Model is sufficient, see G. R. Farrar and M. E. Shaposhnikov, Phys. Rev. Lett. **70**, 2833 (1993).

[13] R. Ammar *et al.*, Phys. Rev. Lett. **71**, 674 (1993).

[14] M. Gell-Mann and A. Pais, Phys. Rev. **97**, 1387 (1955).

[15] K. Lande *et al.*, Phys. Rev. **103**, 1901 (1956).

[16] UA1 Collaboration, C. Albajar *et al.*, Phys. Lett. B **186**, 247 (1987).

[17] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **192**, 245 (1987).

[18] ALEPH Collaboration, D. Buskulic *et al.*, Phys. Lett. B **313**, 498 (1993).

[19] Latest world-average B_s limit can be found in R. W. Forty, *Review of* V_{cb} *and* V_{td}/V_{ts}, presented at the 2nd International Conference on B Physics and CP Violation, Honolulu, Hawaii, March 24-27, 1997, to appear in the proceedings.

[20] A. J. Buras, *Proc. of the XXVIII Int. Conf. on High Energy Physics*, Warsaw, Poland (1996).

[21] O. Long, University of Pennsylvania, Ph.D. dissertation, *A Proper Time Dependent Measurement of* Δm_d *Using Jet Charge and Soft Lepton Flavor Tagging*, in preparation. See also O. Long, J. Kroll, N. Lockyer, M. Paulini, M. Peters, M .Shapiro *An Update of the Vertex Tagged Inclusive Lepton Data Proper Time-Dependent* B^0 *Mixing Analysis*, CDF internal note number 4315 (1997).

[22] M. Peters, University of California at Berkeley, Ph.D. dissertation, *A Measurement of B Meson Oscillations Using Inclusive Leptons in Proton-Antiproton Collisions at 1.8 TeV*, May 1997.

[23] D. Lucchesi, University and Scuola Normale Superiore of Pisa, Italy, Ph.D. dissertation, *Misura dell'andamento temporale del mescolamento* $B_d^0 - \bar{B}_d^0$ *all'esperimento CDF* (1995).

[24] C. Gay, G. Michail, R. Stroehmer, J. de Troconiz, *Measurement of the Time-Dependence of* B^0-$\bar{B}^0$ *Oscillations Using the Low-*p_T *Electron-Muon Data*, CDF internal note 3791 (1996).

[25] R. D. Field and R. P. Feynman, Nucl. Phys. B **136**, 1 (1978).

[26] C. Peterson *et al.*, *Phys. Rev. D* **27**, 105 (1985).

[27] S. van der Meer, *Stochastic Damping of Betatron Oscillations in the ISR*, CERN/ISR-PO/72-31 (1994);
F. Sacherer, *Stochastic Cooling Theory*, CERN-ISR-TH/78-11 (1981).

[28] F. Abe *et al.*, Nucl. Inst. and Meth. Phys. Res., **271A**,387 (1988);
Fermilab-Pub-94/024-E, submitted to Nucl. Instrum. Methods Phys. Res.

[29] D. Amidei *et al.*, Nucl. Inst. and Meth. Phys. Res., **350**,73 (1994)
P. Azzi *et al.*, Nucl. Inst. and Meth. Phys. Res., **360**,137 (1995)

[30] P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. **B 327**, 49 (1988).

[31] P. Sphicas, CDF internal note 2655 (1994).

[32] P. Avery, K. Read, and G. Trahern, Cornell Internal Note CSN-212, March 25, 1985 (unpublished). We use version 9.1.

[33] M. Shapiro *et al.*, CDF internal note 1810 (1992).

[34] H.-U. Bengtsson and T. Sjöstrand, Computer Physics Commun. **46**, 43 (1987).

[35] M. Gronau and J. L. Rosner, Phys. Rev. **D 49**, (1994), 254.

[36] G. Bauer *et al.*, *Same-Side Tagging Using Fully Reconstructed B mesons*, CDF internal note 3066.

[37] P. Maksimović, R. Mattingly, P. Sphicas, F. DeJongh, *Same-Side Tagging using Partially Reconstructed B Mesons*, CDF internal note 3067.

[38] K. Kelley, P. Sphicas, *Same-Side Tagging in Exclusive B decays*, CDF internal note 3638.

[39] F. DeJongh, *Lepton + Charm Data Sets*, CDF internal note 2935.

[40] We generate B mesons according to P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. **B 327**, 49 (1988). We decay them using QQ: P. Avery, K. Read, and G. Trahern, Cornell Internal Note CSN-212, March 25, 1985 (unpublished). We use version 9.1.

[41] For a recent survey of D^{**} production measurements, see C. J. Kreuter, *Measurements of $B \to D^{**}\ell\nu$*, presented at the 2nd International Conference on B Physics and CP Violation, Honolulu, Hawaii, March 24-27, 1997, to appear in the proceedings.

[42] A clear and concise summary of Heavy Quark Effective Theory can be found in M. Neubert, *B physics and CP violation*, lectures in 1995 European School of High Energy Physics, Dubna, Russia.

[43] For a recent review of experimental results on $B^0\bar{B}^0$ oscillations see O. Schneider, *Heavy Quark Spectroscopy, Oscillations, and Lifetime*, presented at the 18th International Symposium on Lepton-Photon Interactions, Hamburg, Germany, to appear in the proceedings.

[44] A. B. Wicklund and K. Byrum, *Self-Tagging Studies Using Semileptonic B Decays*, CDF internal note 2514.

[45] T. LeCompte, private communication.

[46] F. DeJongh, *B-Meson-Pion Charge Correlations*, CDF internal note 2556.

[47] J. Gonzalez, J. Heinrich and N. Lockyer, *Evidence for Self-Tagging using the B_u and B_d Samples*, CDF note 2592.

[48] R. Akers *et al.*, Zeit. für Physik **C66**, 19 (1995); D. Buskulic *et al.*, Zeit. für Physik **C69**, 393 (1996); The Aleph collaboration, *Resonant Structure and*

Flavor-tagging in the $B\pi^{\pm}$ System Using Fully Reconstructed B decays, contribution to the International Europhysics Conference on High Energy Physics, Brussels, Belgium, 27 July 1995.

[49] E. J. Eichten, C. T. Hill, C. Quigg, Phys. Rev. Lett, **71**, 4116 (1993).

[50] I. Dunietz, J. L. Rosner, Phys. Rev. **D 51**, 2471 (1995).

[51] B. Andersson, G. Gustafson, G. Ingelman and T. Sjöstrand, Phys. Rep. **97**, 31 (1983).

[52] F. DeJongh and P. Sphicas, *On the Dilution for Self-Tagging of Neutral B Mesons*, CDF internal note 2850.

[53] R. M. Barnett *et al.*, Phys. Rev. D **54**, 1 (1996).

[54] The CDF Collaboration, *Measurement of the B^- and $\bar{B}^0$ Meson Lifetimes Using Semileptonic Decays*, Phys. Rev. Lett. **76**, 4462 (1996).

[55] F. Ukegawa, Y. Cen, A. B. Wicklund, *Measurment of charged and neutral B meson lifetimes using electron-charm events*, CDF internal note 2598.

[56] F. Ukegawa, A. B. Wicklund, *Electron and b-quark transverse momentum spectra*, CDF internal note 1862.

[57] R. Fulton *et al*, Phys. Rev. D **43**, 651 (1991).

[58] PYTHIA: H.-U. Bengtsson and T. Sjöstrand, Computer Physics Commun. **46**, 43 (1987). We use version 5.7. We decay B mesons using QQ package [40].

[59] F. Bedeschi, M. Caldarelli, D. Lucchesi, *Formulas to describe resolution effects in mixing analysis*, CDF internal note 2859.

[60] The CDF Collaboration, *Measurement of $B^0\bar{B}^0$ Mixing via Time Evolution*, CDF internal note 3255.

[61] M. Shapiro, private communication.

[62] Latest world-average B_s limit can be found in R. W. Forty, *Review of V_{cb} and V_{td}/V_{ts}*, presented at the 2nd International Conference on B Physics and CP Violation, Honolulu, Hawaii, March 24-27, 1997, to appear in the proceedings. For a more comprehensive review, see L. K. Gibbons, *Proc. of the XXVIII Int. Conf. on High Energy Physics*, Warsaw, Poland (1996).

[63] A. Ali and D. London, Nucl. Phys. Proc. Suppl. 54 A, 297 (1997).

[64] K. Kelley, Massachusetts Institute of Technology, Ph.D. dissertation, *Measurement of the CP-violation parameter* $\sin(2\beta)$, in preparation.

[65] The CDF II collaboration, *The CDF Detector: Technical Design Report*, FERMILAB-Pub-96/390-E, October 1996.

[66] *Ibid*, p. 12-17.

[67] A method for extracting the CKM angle γ was originaly developed by M. Gronau, D. London and D. Wyler (and is thus referred to as the GLW method). The GLW method is reviewed in I. Dunietz, in *B decays*, Revised 2nd Edition, edited by S. Stone (World Scientific, Singapore, 1994), p.550. For recent developments, see D. Atwood, I. Dunietz and A. Soni, Phys. Rev. Lett.**78**, 3257 (1997).