Measurement of B^0 and B^+ Lifetimes using Semileptonic Decays at CDF

David K. Clark,[1] *Brandeis University, MA, USA*

Abstract

Lifetime measurements for B^0 and B^+ are presented using semileptonic decays of B mesons from 360 pb^{-1} of data collected by CDF's lepton plus displaced track trigger. The decays $B \to \ell\nu DX$, where D is either D^+, D^0, or D^{*+}, are partially reconstructed from a muon or electron, a displaced track, and a fully reconstructed charm meson. The B^0 and B^+ lifetimes are obtained from an unbinned maximum-likelihood fit to the proper decay length distributions. The measured lifetimes are $\tau_{B^0} = 1.527 \pm 0.012(stat.) \pm 0.023(syst.)ps$ and $\tau_{B^+} = 1.629 \pm 0.013(stat.) \pm 0.025(syst.)ps$.

[1] dkclark@fnal.gov

Acknowledgements

There are many difficult aspects of completing a thesis in experimental high energy physics including the complexity, the massive infrastructure required, and the frustrations when hitting a roadblock. I have learned, though, that the most difficult task is to stay on task and motivated. Without the help and encouragement of many, I would not have completed this thesis. I would like to thank my advisor, Craig, for help in so many ways; for help in building the framework, for his redirections, and especially for his patience.

There were many who helped get the analysis started. Masa supplied the basic beginnings, Sinead and Guillelmo gave impetus and helped sustain the work, Vivek helped me understand how to identify a lepton, and Burkhart helped in removing a tricky background.

Along the way, discussions and work with others kept the analysis moving forward. I thank Matt for his knowledge of hits and tracks and Karen for her signal optimization study. Reid and Satyajit were pivotal with their clever trigger correction method. Gavril stepped in to help just when I needed it.

I was fortunate to be surrounded by so many experienced and professional physicists who asked the right questions, made good suggestions and gave welcome encouragement. Rob, Giovanni, Manfred, Michal, Diego, Michael, and Farrukh, all had their part in helping me push on.

I have gained many friendships from colleagues at CDF. Sharing numerous lunches with Marcel, Ankush, Ben and Antonio kept me in the loop with the CDF community. Discussing anything and everything over a coffee with John, Michelle, or Sebastian raised my spirits. I am grateful for my office mate, Natasa, who challenged me to think in new ways.

The support of friends and family, without I could not have achieved completion, was unwavering. Jarret's encouragement and especially his light and fun distractions were like fresh air. I am grateful for my mother-in-law's confidant expectations for me and especially her cooking.

Most of all, I thank my wife, Becky, whose love and encouragement was expressed through her involvement, her patience, and her understanding. She is a true partner in every way.

Contents

1 Introduction

In 1973 Makoto Kobayashi and Toshihide Maskawa theorized [1] that the existence of a third generation of quarks (top and bottom) could explain CP violation. Only two generations of leptons and quarks had been confirmed at the time but the existence of a third was under investigation. According to Kobayashi and Maskawa, a third generation was necessary. The bottom quark was discovered four years later by an experiment at the Fermi National Laboratory [2].

Physicists at Fermilab have a long history of studying B mesons at the Tevatron collider. There are many other experiments throughout the world that have also contributed in pinning down the properties of these mesons, including the lifetimes of the various b quark mesons. Lifetimes are fundamental parameters with intrinsic value. This thesis describes the measurement of the lifetimes of two B mesons, B^0 and B^+, using semileptonic decays from data collected by the Collider Detector at Fermilab (CDF).

CDF is a general purpose experiment for the study of proton-antiproton collisions at the Fermilab Tevatron Collider. B mesons are abundantly produced in these collisions so precise measurements of lifetimes are feasible. The experiment uses complex trigger algorithms and electronics to pre-select events out of trillions that occur each day. This is done in real time, while the detector is collecting data so that interesting events are kept and uninteresting events are ignored. There are several different components of the experimental apparatus that rely on diverse technologies. Some of the important parts for this measurement are the solenoid magnet, silicon strip sensors, drift chamber and calorimeters.

The data used in this analysis totals 360 pb^{-1} of integrated luminosity accumulated from 2002 to 2005. Given the large b production cross section at the Tevatron [3], this represents around 500 million B hadrons.

Semileptonic decays of B^0 and B^+ provide a substantially large sample (thanks to high branching ratios and a dedicated trigger) which are primarily used as a control sample or benchmark for the B_s lifetime and mixing measurements at CDF. Accurate and precise measurements of the lifetimes are necessary as input to the B_s mixing amplitude measurement. This requires a thorough understanding of the trigger's effect on the measured lifetimes, the missing energy, the decay length resolution, and background contributions.

B meson decay modes are partially reconstructed and separated into three samples. Fully reconstructed $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$ are used to produce separate B^0 and B^+ dominant samples. Then a third sample of D^{*+}'s is retrieved from D^0 candidates by locating a soft pion in the event and examining the difference between the $K^-\pi^+\pi^+$ and $K^-\pi^+$ masses.

The analysis begins with the selection of events containing either a muon or electron and a track of another particle which does not point back to the primary vertex where the collision occured. B meson decays are partially reconstructed due to the missing neutrino and other decay products along the chain, so a pseudo decay length, ct^*, and

its measurement uncertainty are calculated from the decay profile.

Background templates are made from several sources: the sidebands of the invariant D mass distribution determines the combinitorial background, the wrong sign combinations (the lepton and kaon charges are not consistent with a B decay) are used to determine the fake rate, and realistic Monte Carlo simulations are generated for all relevant physics backgrounds. Monte Carlo data are also used to determine the bias caused by the trigger and the missing transverse energy. These are folded into a probability density function for the signal. Finally, the signal probability functions, functions for the D mass and background templates are evaluated in an unbinned log-likelihood function that is maximized to find the B^0 and B^+ lifetimes simultaneously.

This thesis gives a description of the measurement of B lifetimes and the lifetime ratio using semileptonic decays collected from 360 pb^{-1} of integrated luminosity at CDF. The first section explains the theoretical framework which predicts the hierarchy and magnitude of bottom meson lifetimes. The experimental apparatus is described with special emphasis on those critical components of the detector relied on in this analysis. Then the data selection method is detailed along with the criteria for a candidate signal event. A Monte Carlo data sample is described with a comparison to data. This is used to generate some unknown quantities as well as validate the likelihood fitter. The following chapter details the mathematical construction of the maximum likelihood function and reports the fit results. Finally, the systematic uncertainties of the measurements are described and quantified.

2 Theory

Lifetimes are fundamental parameters therefore their measurements have intrinsic value. Precise measurements are also necessary to test the heavy quark expansion theory [5].

Quarks are bound together by the strong force. The B^0 meson is a bound state of a bottom quark, b, and an anti-down quark, $\bar{d}$. Or, conversely, a $\bar{b}$ and d. The quarks that makeup the B^+ (B^-) meson are $\bar{b}u$ ($b\bar{u}$). B mesons are unstable and will decay via the weak interaction, usually like $b \to Wq$ where q is either a u or c quark. Lifetimes are calculated by the sum of interaction processes that lead to the decay of the b to different final states. The final states of interest in this analysis are those that involve decays to a c quark and a virtual W boson that decays to a lepton and a neutrino.

2.1 The Spectator Model

A simple and naive model is sometimes a good starting point to understand the interactions involved in the decay and arrive at predictions for the B meson lifetimes. The b quark is much heavier than u and d so to first order, the light quark can be ignored or treated as a "spectator". The motions of the spectator quark are considered independant of the motion of the heavy quark much like in the atom, the electron's motion with respect to the proton's.

The calculation of the particle width from the Feynman diagram of a muon decay to an electron can help determine the lifetimes of the b within the spectator model since their diagrams are similar. However, the main difference is associated with the phase space of the decay. As the lightest unstable particle, the muon may only decay into an electron and neutrinos. On the other hand, the b quark may decay into $e\nu_e$, $\mu\nu_\mu$, $\tau\nu_\tau$ and the quark generations, ud and cs. After also considering the color changes of the quark products, the number of fundamental decay modes of the b quark is nine times larger than the number of muon decay modes.

The decay width of a muon is approximated (ignoring higher level diagrams) by []

$$\Gamma_\mu \simeq \frac{m_\mu^5 G_F}{192\pi^3},$$

where m_μ is the mass of the muon. The Fermi coupling constant, G_F, is,

$$G_F = \frac{\sqrt{2}}{8}\frac{g^2}{m_W^2},$$

where g is the weak interaction coupling constant and m_W is the mass of the W boson. The lifetime of the b quark may be estimated using the expression for the width of the muon by adding a couple of terms. A factor of $|V_{cb}|^2$ is necessary to account for the different couplings between a b and c quark versus a muon and neutrino. Factoring in the phase space considerations mentioned previously gives the expression for the b

decay width,

$$\Gamma_b \simeq \frac{m_b^5 G_F |V_{cb}|^2}{192\pi^3} \times 9.$$

Since precise measurements of the muon lifetime have been made [4], the expression can be rewritten for the lifetime of the b in terms of the muon lifetime,

$$\tau_b \simeq \tau_\mu \times \frac{m_\mu^5}{m_b^5} \times \frac{1}{9|V_{cb}|^2}.$$

Using published values for τ_μ, m_μ, m_b, and $|V_{cb}|^2$ [4], the calculated lifetime ranges from 1.3 to 1.5 ps.

Each B meson lifetime would then be identical since it would only depend on the mass of the b quark. However, this is not the case so there must be other non-negligible effects that cause the lifetimes to vary depending on which quark the bottom quark is paired with.

2.2 Non-spectator effects

The possibility of more than one path for the B meson to decay to same final state causes interference between them and thus shortens or lengthens the meson's lifetime depending on whether the interference is constructive or destructive. For example, the B^- has two ways to decay to the final state $D^0\pi$.The calculated interactions destructively interfere with one another which reduces the lifetime. The b in B^+ may decay in the same way, however, the final states are different so the lifetimes remain unchanged when calculated. This effect is referred to as Pauli interference.

Another effect that can explain the differences in lifetimes, although not as dominant as Pauli interference, is weak annihilation also referred to as weak exchange.Weak exchange may occur at the Cabibbo allowed level for the B^0 but only at the Cabibbo suppressed level for the B^+. This effect also lengthens the B^+ lifetime compared to the B^0. Baryons suffer from additional interferences due to weak exchange between the two lighter quarks. Since the B mesons only have one light quark this effect is suppressed.

Considering these non-spectator effects the lifetime hierarchy may be predicted as,

$$\tau(B^+) > \tau(B^0) > \tau(B_s) > \tau(\Lambda_b).$$

The next sections discusses a quantitative approach to further understand the difference in lifetimes.

2.3 The Heavy Quark Expansion Technique

The heavy quark effective theory [5] describes and quantifies the way spectator quarks give rise to the difference in B meson lifetimes. This theory is an effective field theory

which expands in powers of the quark mass. The leading term of the expansion at order $1/m_b$ describes the spectator model. The first correction term at order $1/m_b^2$ differentiates baryon lifetimes from mesons due to the helicity suppression. At this point it is not sensitive to flavors so the meson lifetimes are still the same. The next correction term at order $1/m_b^3$ describes both the Pauli interference and weak annihilation effects and it is at this order in which the meson lifetimes split.

This allows for quantitative predictions of the lifetimes. These are usually expressed as ratios and are listed in Table 1.

ratio	HQE prediction
τ_{B^+}/τ_{B^0}	1.06 ± 0.02
τ_{B^s}/τ_{B^0}	1.00 ± 0.01
$\tau_{\Lambda_b}/\tau_{B^0}$	0.88 ± 0.05

Table 1: Predicted lifetime ratios.[6][7][8]

2.4 Semileptonic B Decays

B mesons decay semileptonically to charmed mesons in complicated fashion as seen in Fig. 1. The dominant decay modes of B^0 are $B \to \ell D^{*+}$ and $B \to \ell D^+$ while the B^+ dominant mode is $B \to \ell D^0$ (Table 2). Since none of these decay modes are pure B^0 or B^+, all three modes are simultaneously included in the lifetimes fit in order to disentangle and measure the two lifetimes.

decay mode	$\%B^0/\%B^+$ produced	$\%B^0/\%B^+$ accepted
ℓD^+	80/20	85/15
ℓD^{*+}	88/12	84/16
ℓD^0	30/70	20/80

Table 2: Approximate B meson composition for each decay signature.

The CDF detector does not directly detect neutrinos, and the analysis does not reconstruct photons or neutral pions from the decays of the resonant states. Soft pions from resonant decay are kept, but only those that satisfy the selection criteria. The energy from these objects or their tracks is not included in the calculated B momentum. This missing energy is accounted for in the K factor described below. It represents all of the unreconstructed decay products from the subsequent decays of the B meson.

The K factor is defined by the following,

$$K = \frac{p_T(\ell D)}{p_T(B)} \frac{L_{xy}(B)}{L_{xy}(\ell D)}, \tag{1}$$

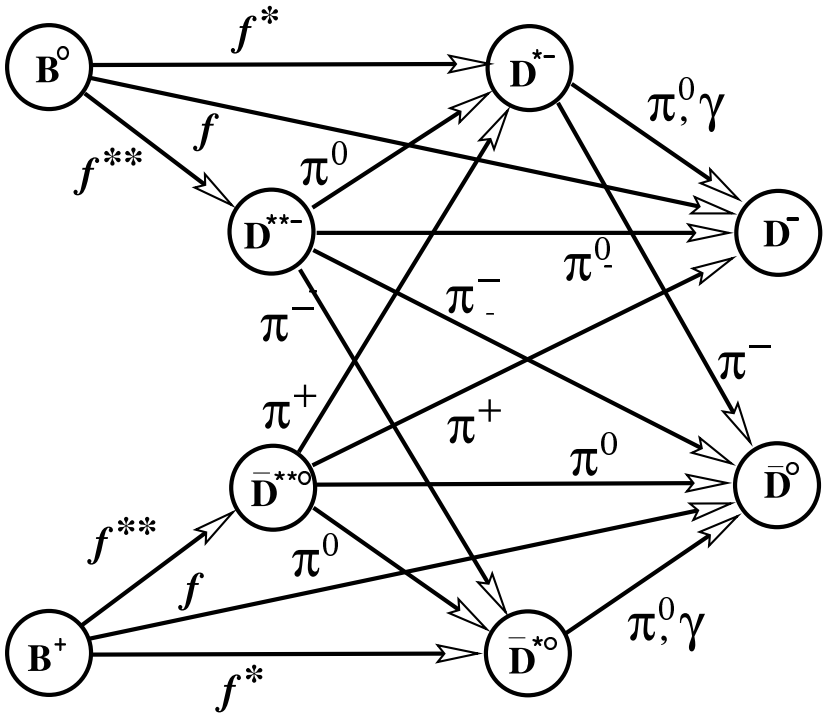


Figure 1: Decay diagram.

where L_{xy} is the distance between the decay vertex and the primary vertex in the xy plane. It is essentially a scale factor which scales the decay length calculated from the vertex of the lepton and D meson in order to recover the actual decay length of the B meson.

3 The Experiment

3.1 Fermilab Accelerator

The Fermilab accelerator includes several components to enable it to accelerate a beam of protons and a beam of antiprotons to 980 GeV/c^2 each before focusing them onto one another to force collisions. Figure 2 shows a layout of the accelerator complex. The first component, the Cockcroft-Walton pre-accelerator, produces negatively charged hydrogen ions, H-, from a bottle of hydrogen gas, then accelerates these to an energy of 750 keV/c^2. A linear accelerator, the Linac, then accelerates the ions up to 400 MeV/c^2 using resonant radio frequency cavities. Oscillating electric fields within the cavities cause the hydrogen ions to accumulate into bunches.

The next stage of acceleration, a circular synchotron ($r = 75$ m) called the Booster, strips off the two electrons and adjusts the bunches into phase with its 18 resonant frequency cavities. A bunch structure of 37.8 MHz is formed and the new beam is accelerated to 8 GeV/c^2 which is then transferred to the Main Injector. Like the Booster, the Main Injector is a circular sychrotron with 18 accelerating cavities but it is much larger with a circumference of over three kilometers. The proton beam is accelerated to 150 GeV/c^2 and then transferred to the largest synchrotron at Fermilab, the Tevatron, with a circumference of 6.28 km.

The antiprotons are similarly accelerated by the Main Injector and the Tevatron, but the process of generating and storing these are quite different. Bunches of 120 GeV/c^2 protons are focused onto a nickel target. Antiprotons are among the products from the collisions with the target. These are separated from other particles with magnets, then transferred to a debuncher, a triangular-shaped synchrotron. The particle beam then undergoes stochastic cooling to reduce the transverse momentum spread.

After reduction of the beam's phase space, the antiprotons are accelerated to 8 GeV/c^2 and transferred to a storage ring called the Accumulator, also a triangular-shaped synchrotron. Here, they await transfer into the Main Injector. Within the Main Injector, the antiprotons and protons are accelerated to 150 GeV/c^2 while they move in opposite directions. Now the beam is ready for transfer to the Tevatron for its final stage of acceleration.

Once inside the Tevatron, the beams, or more precisely, the bunches of protons and antiprotons are accelerated to 980 GeV/c^2. The Tevatron collides the bunches of protons and antiprotons at a center-of-mass energy of 1.96 TeV/c^2 at a rate of one bunch per 396 ns. There are two collision points on the Tevatron ring and the CDF experiment is located at one of these.

3.2 CDF

CDF Run II - operating since 2001 - is a general purpose detector [9] designed to study $p\bar{p}$ collisions at the Tevatron. The detector is shown in a cutaway view in Fig. 3. Three dimensional charged particle tracking is achieved through the use of an open cell

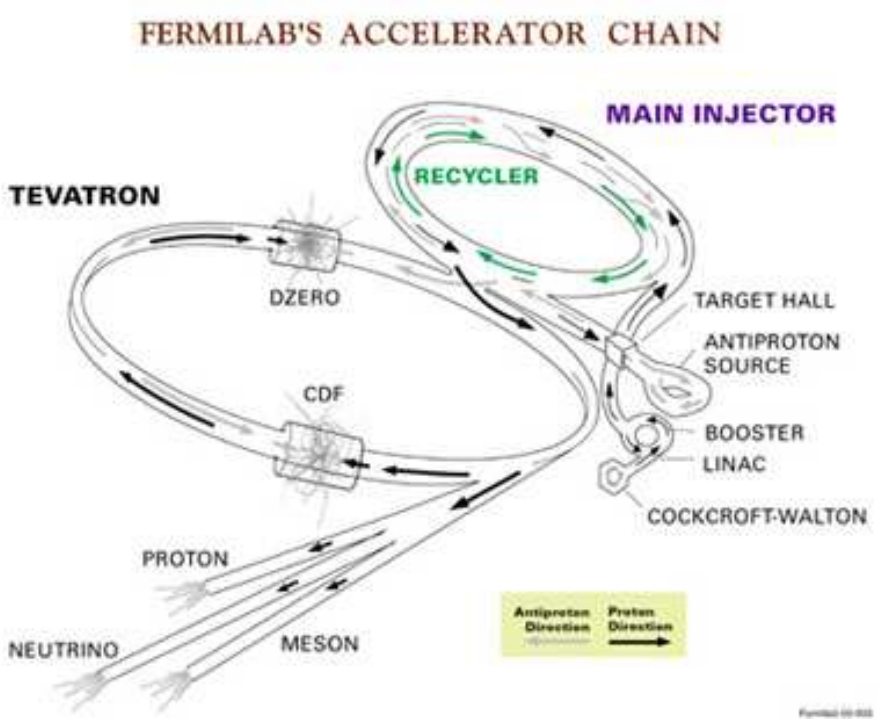


Figure 2: The Fermilab accelerator complex.

drift chamber and a silicon microstrip detector which are enveloped by a solenoid coil magnet (Fig. 4). A plastic scintillator based time of flight detector, electromagnetic and hadronic sampling calorimeters, and muon detectors (planar drift chambers) lie outside of the magnet. A gas Cherenkov counter is employed to measure the average number of inelastic $p\bar{p}$ collisions, also called the beam luminosity. An elaborate trigger and data aquisition system operates in three stages to select from and reduce the high rate of collisions (events) and large data volume.

The following sections describe in more detail the components used for this measurement.

CDF uses a right-handed cylindrical coordinate system. The $+y$ direction is vertically up. The $+x$ direction is the horizontal pointing away from the center of the Tevatron ring while the azimuthal angle ϕ is defined from this direction. The polar angle θ is defined from the $+z$ direction which is the direction of the beam of protons through the detector. This angle is usually expressed in terms of the pseudorapidity, η,

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

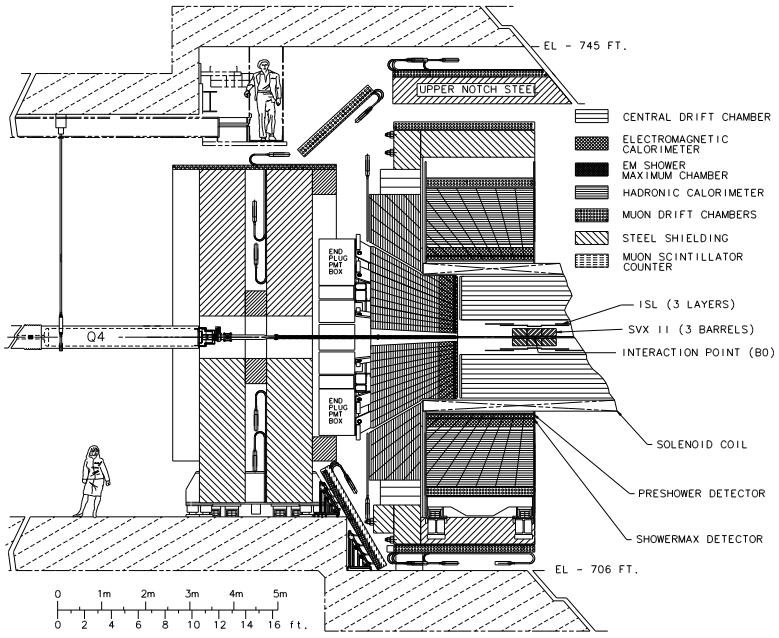


Figure 3: Elevation view of one half of the CDF detector.

3.2.1 Silicon Vertex Detector

The Silicon Vertex Detecter consists of nine concentric layers of AC coupled silicon micro-strip sensors with over 722,000 readout channels (Table 3). The innermost layer [10] is mounted directly on the beam pipe at 1.3 cm from the beam and the outer most layer lies at a radius of 29 cm.

The readout pitch, the distance between two strips which are read, varies from 60 to 150 μm depending on which layer and side. All but the first layer are double sided giving both $r\phi$ and rz measurements. On one side, the orientation of the strips allows for $r\phi$ views. The orientation of the strips on the opposite side is rotated by 90° or 1.2° (small angle stereo view) to measure the rz coordinates. Together, these measurements enable three dimensional reconstruction of secondary vertices.

The length of the silicon detector for the first six layers is nearly one meter allowing for coverage up to z_0^{max}=45 cm and $\eta^{max} = 2$. The combined length of the intermediate layers at 1.74 m also allows for coverage up to $\eta = 2$. The first layer is primarily used for improvement of impact parameter resolution. The next five layers are essential in this experiment since the displaced track trigger uses hits from these layers to pre-select events in real time. The last three layers are referred to as the intermediate silicon layers as they improve the ability to link tracks between the silicon detector and the drift chamber. Each of these intermediate layers overlap one another but do not by themselves span the entire length (see Fig. 4).

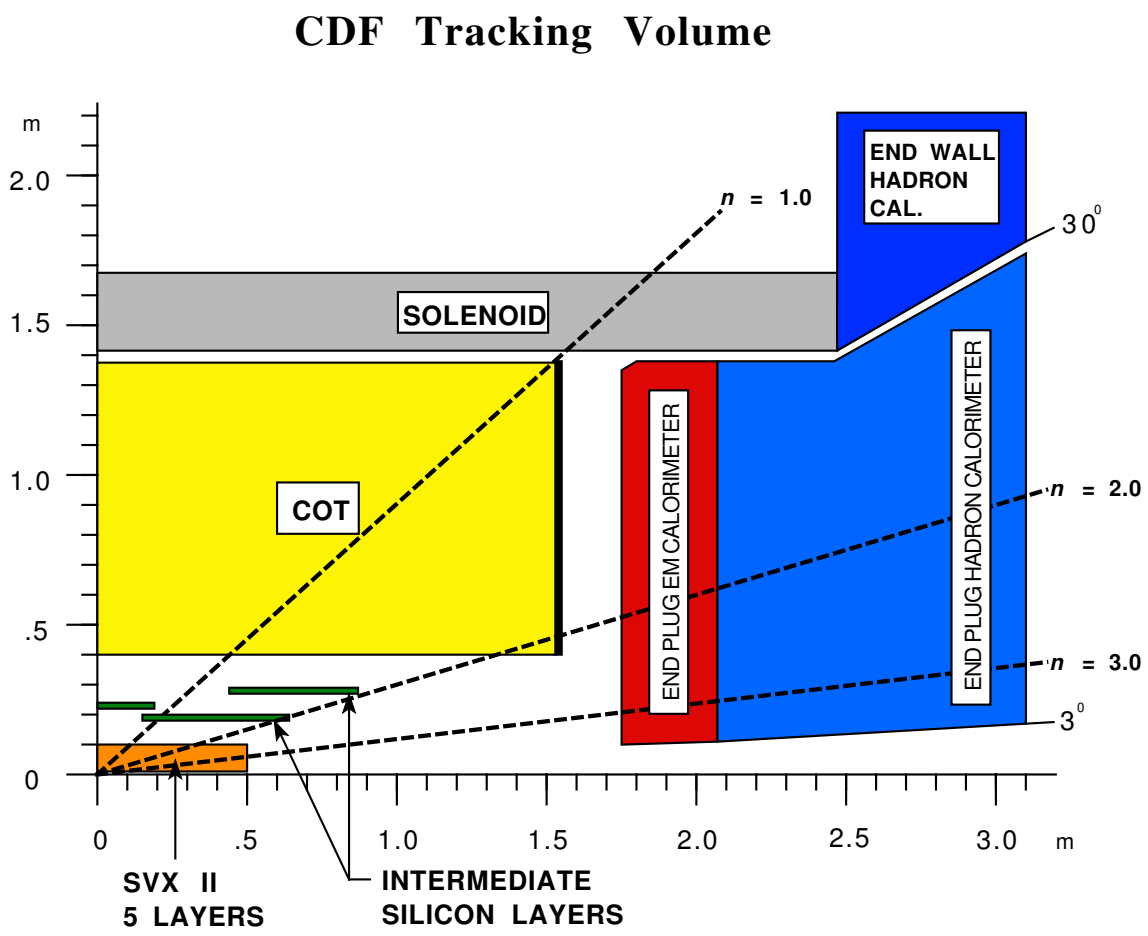


Figure 4: Tracking volume diagram.

3.2.2 Central Outer Tracker

The Central Outer Tracker is an open cell wire drift chamber with 30,240 sense wires (readout channels) and 96 layers. It surrounds the Silicon Detector, extending radially from 44 cm to 132 cm and has a length along the z direction of 3.1 m. This allows good coverage up to $|\eta| = 1.0$.

The concentric layers are grouped into eight superlayers that each contain twelve sense wires and thirteen potential wires within each cell. The number of cells depends on the layer's radial position. The sense wires within the cells (Fig. 5) of even number superlayers - superlayer 1 is the closest to the beam line - are oriented parallel to the beam line so the readout coordinates give $r\phi$-plane measurements whereas odd number superlayers are tilted $\pm 2°$ in order to provide information of the z direction or stereo measurements.

Potential wires are strung on both sides of the sense wires. The chamber is filled with a 50/50 mixture of argon and ethane gas. A charged particle traveling through

Layer 1	
radial Coverage	1.35-1.62 cm
readout coord.	$r\phi$
stereo side	N/A
readout channels	13,824
readout pitch	25 μm
rapidity coverage	$\|\eta\| \leq 2.0$
total length	95.0 cm
resolution	11 μm $r\phi$
Layers 2-6	
radial Coverage	2.4-10.7 cm
readout coord.	$r\phi$
stereo side	rz (2,3,5), ruv (1.2°) (4,6)
readout channels	405,504
readout pitch	60-65 μm $r\phi$; 60-150 μm stereo
rapidity coverage	$\|\eta\| \leq 2.0$
total length	96.0 cm
resolution	12 μm $r\phi$
Intermediate layers	
radial Coverage	20-29 cm
readout coord.	$r\phi$
stereo side	ruv (1.2°)
readout channels	303,104
readout pitch	110 μm $r\phi$; 146 μm stereo
rapidity coverage	$\|\eta\| \leq 1.9$
total length	174.0 cm
resolution	16 μm $r\phi$

Table 3: Parameters of the silicon vertex detector.

the chamber ionizes the gas along its path and the electric gradient produced by the potential between the wires causes the freed electrons to drift to the sense wires. The time it takes for the ionized particles to drift to the sense wire, the drift time, is less than 100 ns. This keeps particle tracks from more than one beam crossing being recorded at the same time.

Transverse momentum resolution is measured using cosmic ray events to be [17] $\sigma_{p_T}/p_T^2 = 0.1\%$ per GeV/c^2 and the spatial hit resolution is approximately 150 μm.

3.2.3 Magnetic Field

The tracking volume - the drift chamber and silicon detector - is encompassed by a superconducting solenoid magnet to provide a 1.41 Tesla field in the direction of the

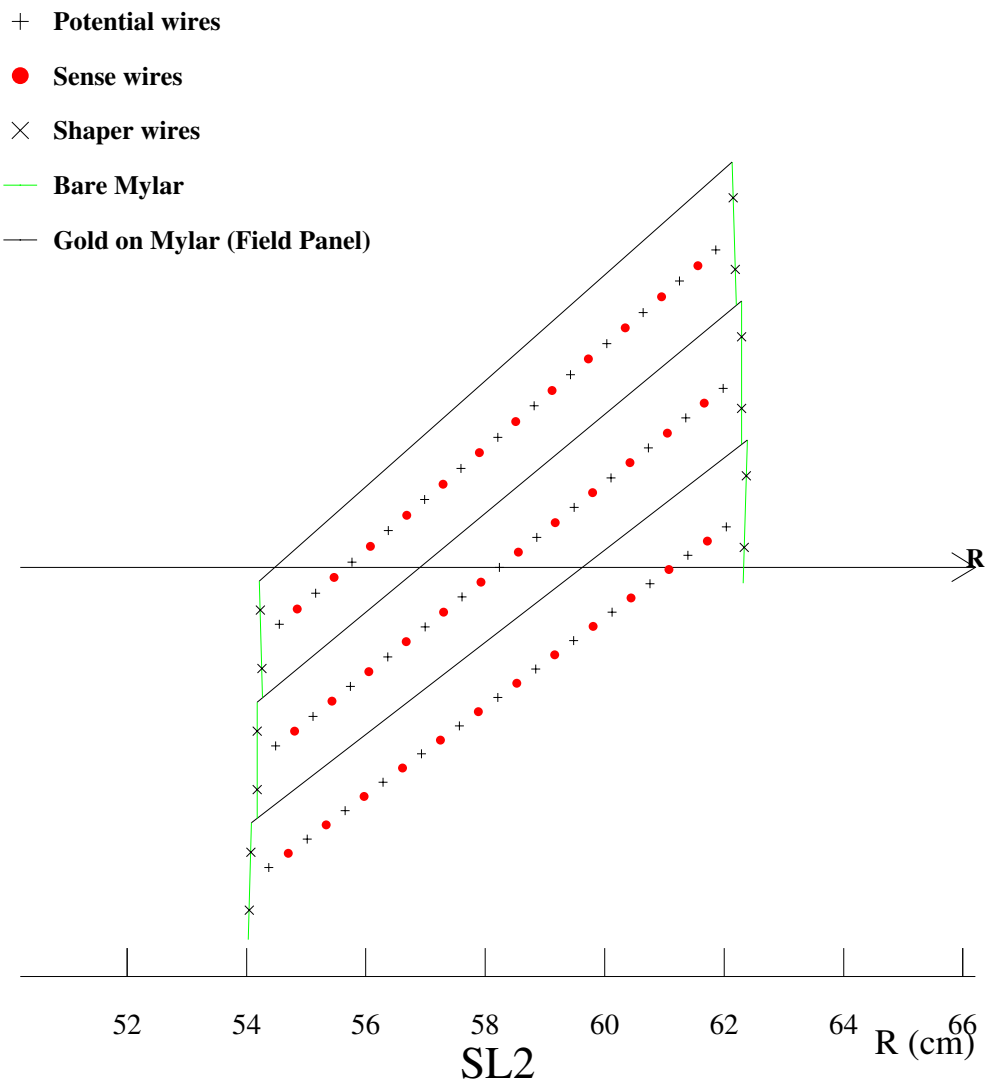


Figure 5: Super layer 2 cells diagram.

beamline ($-z$). This magnetic field enables CDF to measure the charge and momentum of a particle traveling through the tracking volume.

3.2.4 Calorimeters

Beyond the central outer tracker and solenoid magnet stand the electromagnetic and hadronic calorimeters (Fig. 3). These calorimeters are made up of alternating layers of lead or steel and scintillator materials. The energy deposited within them by showering

particles is used to determine what type of particle entered the calorimeter. They are designed to absorb all the energy of a particle as opposed to the small perturbation of particles by the tracking detectors. When an electron enters the calorimeter it radiates via the Bremsstrahlung process whereas a photon produces an electron-positron pair which in turn produces Bremsstrahlung radiation. These processes lead to the development of an electromagnetic shower where nearly all the initial energy is absorbed. Hadrons undergo different processes within the calorimeters. Both elastic and inelastic scattering of the hadron off of a nucleon generates further scattering and hence a shower of hadrons. An electromagnetic shower also accompanies a hadronic shower due to the decay of secondary π^0's.

There are several varieties of calorimeters within CDF. The central electromagnetic calorimeter is made up of 23 alternating layers of lead and polystyrene scintillator and its size allows for coverage of all azimuthal angles and pseudorapidity up to 1.1. The other main calorimeter is the central hadronic calorimeter. This is also made of 23 alternating layers but thick steel is used instead of lead and the scintillator is acrylic rather than polystyrene. It has the same coverage of ϕ but slightly lower coverage of η. Other central calorimeters include the pre-radiator, located between the solenoid magnet and the electromagnetic calorimeter, and the electromagnetic shower detector located within the electromagnetic calorimeter.

The energy deposited in the electromagnetic calorimeter relative to the hadronic calorimeter can help distinguish electrons from hadrons. Electrons and photons will deposit nearly all of there energy within the electromagnetic calorimeter whereas hadrons will lose very little there and instead be absorbed within the hadronic calorimeter. Making appropriate cuts on the measured energy within each calorimeter will help separate one type of particle from the other.

3.2.5 Muon Chambers

Muons used in this measurement are of considerable energy ($p_T > 4$ GeV/c) and they are 200 times heavier than electrons. They have no strong interactions with nuclei and have less bremsstrahlung radiation due to there larger mass. Therefore, they are not absorbed within the hadronic or electromagnetic calorimeters. Coulomb scattering can affect a muon's direction, although this can be accounted for since the scattering is approximately Gaussian with respect to the true direction.

Beyond the calorimeters lie the muon chambers. There are four scintillator and proportional chamber systems available for muon identification. Only two are used in this measurement - the central muon detector and the central muon upgrade. These are made up of four layers of rectangular cells. These cells are essentially drift tubes filled with a 50/50 mixture of argon and ethane and operate under a similar principle to the central outer drift chamber's. Each drift cell has one stainless steel sense wire strung from end to end but has no potential wire. Instead, the four walls of the cell are used to construct a uniform electric field.

Hits on the sense wires within multiple layers are used to generate short tracks

called stubs which are then matched to tracks from the tracking detectors. All muons used in this measurement are required to have stubs in both central muon chambers. The central muon detector provides pseudorapidity coverage up to 0.6 (Fig. 6).

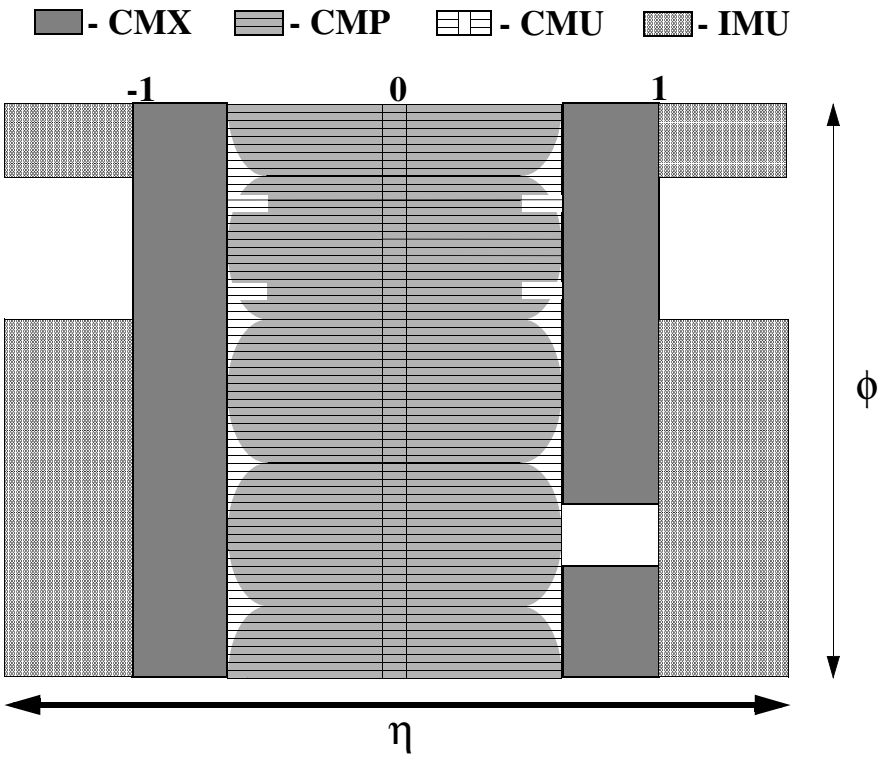


Figure 6: Coverage of the muon system.

Additionally, there is heavy material which shields the muon detectors from energetic hadrons which were not stopped in the calorimeters, although, some are detected in the muon chambers giving a fake muon signal (these are called "punch-throughs").

3.2.6 The Trigger

The rate of $p\bar{p}$ collisions is too high to allow every event to be recorded so CDF applies triggers in real time to accept interesting events for further analysis and ignore uninteresting events. The triggering system has three levels of decision making, each level more selective than the previous. The first level accepts events at the rate of 20 kHz. At this level, triggers simultaneously process information from the central outer tracker, calorimetry, and the muon system. An extremely fast tracking algorithm reconstructs tracks from the central outer tracker within 2.7 microseconds. These tracks are extrapolated to the calorimetry and muon system and linked to triggered objects like electrons, photons, and muons. An event with four or more tracks is automatically

accepted whereas other events may be accepted depending on the number and type of objects.

At the second level, more information is added including hits from the silicon detector. The secondary vertex trigger is applied at level two. Improved electron and photon identification is also achieved at this level. The acceptance rate is reduced to 300 Hz.

The entire detector is readout upon a level two accept decision. This information is then passed to level three where software algorithms act on the complete event information for further rate reduction (75 Hz) before storing an event.

3.2.7 The Track Parameters

The tracking system utilizes two detector components, the silicon detector and central outer tracker, plus the solenoid magnet to enable measurement of a particles momentum.

A charged particle moves through the CDF tracking volume on a helical path due to the magnetic field. The circular motion in the $r-\phi$ plane is described by three parameters; the impact parameter, d_0, azimuthal angle ϕ_0, and the curvature, C. Two more parameters, λ and z_0, describe the straightline motion in the $r-z$ plane. Combined, these components of motion describe a helix [18].

The impact parameter is the distance between the closest approach of the track and the primary interaction point, or primary vertex. The parameters, ϕ_0 and z_0, are the ϕ and z positions of the track at minimum approach to the z axis. The curvature is defined as $C = q/(2R)$ where q is the charge of the track (in units of the proton's charge) and R is the radius of the circle formed by the projection of the helix on the $r-\phi$ plane. Finally, $\lambda = cot\theta$, where θ is the polar angle between the momentum of the track and the z axis. All other track parameters are calculated from these five parameters.

For example, the transverse momentum is measured by the relationship

$$p_T = 0.3BR$$

where the magnetic field, B, and the radius, R, is given in Tesla and meters respectively so that p_T is in GeV/c units. Another important quantity previously introduced is the pseudo-rapidity or the "forwardness" of a track, $\eta = -ln(tan(\theta/2))$.

4 Data Sample

The data come from integrated luminosity of 360 pb^{-1} which was recorded from February 4, 2002 to August 22, 2004. Data were collected in separate operating segments of time, each called a run. The range of run numbers in the data sample is 138425-186598. Each run is compared with a set of minimum criteria to determine whether the data were taken under proper conditions. This removes data runs in which parts of the detecter were off or not operating nominally at the time the data were being collected.

4.1 The Lepton plus Displaced Track Trigger

The trigger system is essential to eliminating most of the unwanted events from uninteresting collisions. It is a three level system where each succeeding level's trigger is more selective. In order to select events with a muon and a displaced track, the first trigger pulls events containing a muon in the central region of the detector with $p_T > 4\ GeV/c$. Next, the second level trigger flags the event if it contains a track with an impact parameter between 120 μm and 1 mm and $p_T > 2\ GeV/c$. At the third level, the event is flagged if it passes the first two triggers, the displaced track is not associated with the muon, and the invariant mass of the muon and track combination is less than 5 GeV/c^2.

Triggering on events containing electrons and displaced tracks is conducted in the same way except the condition that the level one trigger finds an electron in the central calorimeters.

There are two types of quantities at CDF, online and offline. Online quantities are measurements made during detector operation. These quantities, like track impact parameter and p_T, are used to make triggering and event selection in real time. Offline quantities are those calculated by the event reconstruction software after the event has been selected and stored. These quantities are used to select the data sample. There may be slight differences between the two sets of quantities so they must be compared to provide confidence that the events in the data sample are the same events that were triggered.

A displaced track is confirmed by matching a candidate track to a trigger track and re-applying the impact parameter criteria to the offline quantity [21]. A track with an offline impact parameter between 120 μm and 1 mm, $p_T > 2$ GeV/C, and has at least 4 r-ϕ hits in the layers two through six of a single r-ϕ wedge in the silicon detector is confirmed to be a triggered track.

The triggered muon in the event is confirmed by matching its track to the best matching offline track. The offline matched muon must have stubs in both of the central muon detectors and have track $p_T > 4$ GeV/C. In addition, the difference between the muon stub and the track, Δx, must be less than 15 cm and 20 cm for the central muon system and the upgrade respectively.

A triggered electron must have a cluster in the central electromagnetic calorimeter and have an associated track with $p_T > 4$ GeV/c.

4.2 B and D reconstruction

All tracks are refit using the Kalman fitting method. The charm mesons, D^0 and D^-, are reconstructed by determining kaons and pion(s) first. Either a kaon or a pion has to be the confirmed trigger track. The lepton is explicitly prohibited from being the triggered track in order to optimize signal. Without a displaced track from a kaon or pion, more background and unrelated tracks would be accepted.

D^{*-} candidates are chosen from a subset of the D^0 sample (see section 4.3). Wrong sign candidates (the relation of the charge of the kaon and lepton imply they do not derive from the same B meson decay) for all samples are separately examined to determine two sources of background contributions (see section 6.2).

The B and D decay vertices are reconstructed via a vertex/mass constrained iterative fit method [11].

4.3 Selecting D^* candidates

Any additional tracks are temporarily kept to determine if they belong to a soft pion, π_s, from the D^{*-} decay. Any track with p_T greater than 400MeV/c is considered as a soft pion candidate. A D^{*-} candidate is selected by comparing the three-body invariant mass to the D^0 candidate's mass. The mass difference between the two combinations,

$$\delta M = M(K\pi\pi_s) - M(K\pi)$$

is calculated for each additional track. The track associated with the lowest value of δM is kept and further considered. The cut is applied to the value of δM to separate the D^{*-} candidates from the D^0's. Figure 7 shows the δM distribution with a loose cut on the D^0 mass.

Note that there is no vertex requirement for the soft pion and D^0 tracks of the D^* candidate. A three-body vertex requirement would result in better decay length resolution and mass resolution, however these do not compensate for the loss of signal D^* events (events which end up in the D^0 sample). Although the absolute value of signal events would not change, the B^+ dominant D^0 sample would be diluted with B^0's from the D^* sample thus decreasing the ability to separate the two lifetimes.

4.4 Signal Candidate Selection

The data is filtered using criteria (Table 4) derived from a signal optimization study [22]. Some variations from the study were necessary. For example, The lower limit allowed for the pseudo decay length of the B is moved to 0 from 200 μm for this analysis in order to broaden the range available to the B_s mixing measurement.

The mass ranges of the D mesons include background on either side of the signal peak. Some of this background will be used to model the background under the peak. Since the B meson is partially reconstructed, a wide mass range is set. This range lies above the mass of the D^{*+} (2.01 GeV/c^2) and below the B^+ mass (5.279 GeV/c^2) to exclude fully reconstructed hadronic decays of either meson.

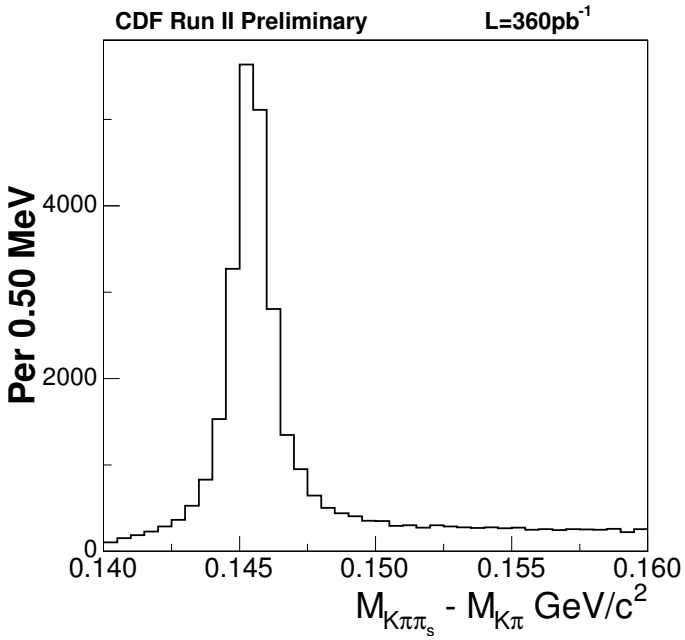


Figure 7: Mass difference, δM.

In order to improve the quality of the signal D mesons, limits are imposed on the transverse momentum of its decay products, the decay length, and the χ^2 of the decay vertex fit. These limits are applied to the data as well as the Monte Carlo sample to ensure equal treatment.

parameter	ℓD^0	ℓD^{*-}	ℓD^+
$M_{\ell D}$	$2.3 \rightarrow 5.2$ GeV/c^2	$2.3 \rightarrow 5.2$ GeV/c^2	$2.3 \rightarrow 5.2$ GeV/c^2
M_D	$1.76 \rightarrow 1.96$ GeV/c^2	$1.76 \rightarrow 1.96$ GeV/c^2	$1.76 \rightarrow 1.96$ GeV/c^2
$ct^*(B)$	$0 \rightarrow 0.4$ cm	$0 \rightarrow 0.4$ cm	$0 \rightarrow 0.4$ cm
σ_{ct^*}	$< 400\,\mu m$	$< 400\,\mu m$	$< 400\,\mu m$
$ct(D)$	$-0.015 \rightarrow 0.1$ cm	$-0.015 \rightarrow 0.1$ cm	$-0.01 \rightarrow 0.2$ cm
$L_{xy}/\sigma_{L_{xy}}$	> 6	> 6	> 11
$P_V(B)$	$> 10^{-6}$	$> 10^{-6}$	$> 10^{-5}$
$\chi^2(V_D)$	< 10	< 10	< 10
$p_T(\pi)$	> 0.4 GeV/c	> 0.4 GeV/c	> 0.4 GeV/c
$p_T(K)$	> 0.5 GeV/c	> 0.5 GeV/c	> 0.6 GeV/c
$p_T(\pi_s)$	N/A	> 0.4 GeV/c	N/A
δM	$< 0.1435, > 0.1475$	$0.1435 \rightarrow 0.1475$ GeV/c^2	N/A

Table 4: Selection criteria.

Table 4 lists the allowed quantities for a candidate B meson but includes a wide

range around the D meson mass peak. In order to select a sample for the signal, any D mesons whose mass is more than three standard deviations from the mean of the mass peak are excluded. For the samples containing D^0 candidates, the signal region is $1.84 < M_D < 1.888$ GeV/c^2 and for the D^+ sample is $1.844 < M_D < 1.892$ GeV/c^2.

4.5 Signal Yields and D Mass distributions

Since the B meson is not fully reconstructed, signal yields are determined from the invariant D mass distributions (Fig. 8). The D meson mass distributions are fit with a double Gaussian and a first order polynomial to represent the background. Yields and significant signal versus background quantities are presented in Table 5. These values represent the number of candidates under the D meson mass peak minus the background derived from the shape of the combinatorial background under the peak. The numbers of "signal" include all real D mesons whether they come from semileptonic B decay or not. Background events containg a D meson are dealt with in section 6. The amounts for these backgrounds are about 3-6%, depending on the sample.

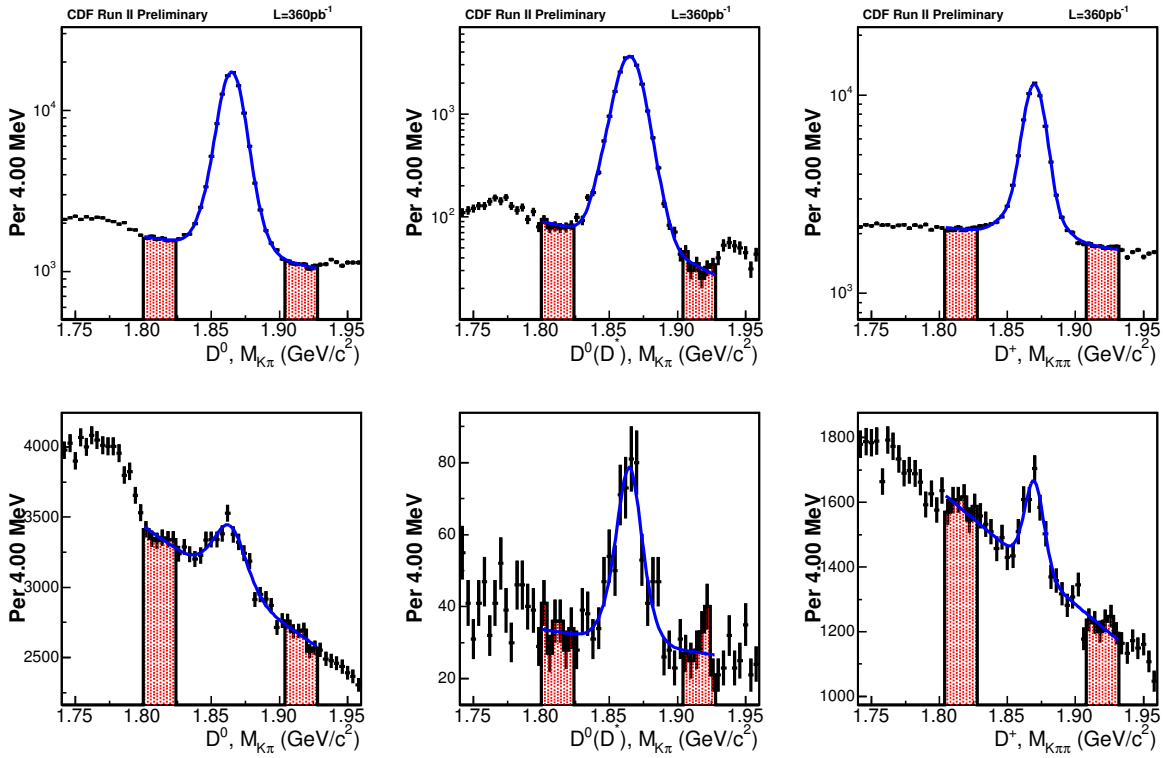


Figure 8: Mass distributions of D candidates (top) and wrong sign D candidates (bottom) with defined sideband regions.

Data sample	N_S	N_S/N_B	$N_S/\sqrt{N_S+N_B}$
ℓD^0	$87.5K \pm 408$	5.80	274
ℓD^*	$19.6K \pm 154$	30.3	138
ℓD^+	$48.4K \pm 514$	2.29	184

Table 5: Signal events.

4.6 Decay Length Reconstruction

The lifetimes of B^0 and B^+ are measured using the distance traveled by the B's before they decay. The decay length of the vertex is calculated using CDF's high resolution transverse quantities. $\vec{V}$ is the reconstructed B decay vertex as determined by the vertex fitter, $\vec{p_T}$ the transverse momentum of the B.

$$L_{xy} = \vec{V} \bullet \frac{\vec{p_T}}{|\vec{p_T}|}$$

Decay length,

$$ct = \frac{L_{xy}}{\beta\gamma} = \frac{L_{xy}(B)M_B}{p_T(B)}$$

Since the neutrino is not reconstructed, only the $l + D$ quantities are known so a new quantity is defined,

$$ct^* = \frac{L_{xy}(\ell D)M_B}{p_T(\ell D)}.$$

This is called the pseudo-decaylength. Similarly for the measurement error,

$$\sigma_{ct^*}^{STD} = \frac{\sigma_{L_{xy}}(\ell D)M_B}{p_T(\ell D)}$$

where $\sigma_{L_{xy}}(\ell D)$ is the error returned from the vertex fitter. The PDG values of M_B are used since the B meson masses are not fully reconstructed. Furthermore, since it is not known whether the vertex candidate is a B^0 or B^+, a weighted average of the two masses is used for M_B.

For each event then, there is a factor, K, which will convert the decay length of the $l + D$ to the decay length of the B;

$$K = \frac{p_T(\ell D)}{p_T(B)} \frac{L_{xy}(B)}{L_{xy}(\ell D)}. \tag{2}$$

However, K is unknown for each event so a realistic Monte Carlo sample is required to model it.

In the early period of this analysis, the p_T error contribution was ignored while calculating σ_{ct^*} since σ_{p_T}/p_T is less than 0.01 and therefore contributed very little to the

overall uncertainty . This uncertainty is denoted with the STD superscript. However, for large ct (corresponding to large L_{xy}), there can be significant contribution from the momentum resolution. Including the error on p_T (assuming weakly correlated $\sigma_{L_{xy}}$ and σ_{p_T} since L_{xy} is measured by the silicon detector and p_T comes from the central outer tracker), the ct^* uncertainty can be written in terms of the standard error calculation,

$$\sigma_{ct^*} = \sigma_{ct^*}^{STD} \sqrt{1 + (\frac{L_{xy}}{\sigma_{L_{xy}}} \frac{\sigma_{p_T}}{p_T})^2}$$

For example, a typical event can have $L_{xy} = 0.5$ cm, $\sigma_{L_{xy}} = 0.005$ cm and $\sigma_{p_T}/p_T = 0.01$. Then the corrected σ_{ct^*} becomes

$$\sigma_{ct^*} = \sigma_{ct^*}^{STD} \sqrt{2}$$

This is not a small effect and so must be included. Before any event is chosen using the selection criteria, σ_{ct*} is recalculated.

4.7 ct^* Error Scale Factor

The error on the L_{xy} measurement is underestimated in the data. A globally defined scale factor is applied to each σ_{ct*} to account for this. This is also done prior to signal candidate selection. The value of the scale factor is fixed to 1.35 which is based on the $b \to J/\psi K^X$ lifetime studies at CDF [12] [13]. This scale factor is separate from the scale factors used in determining the non-Gaussian shapes of the ct^* resolution. See section 7.3 for a discussion of those scale factors.

5 Monte Carlo Data Generation

The K factors and trigger effects are quantified directly from fully realistic Monte Carlo events. In order to understand the trigger effect on pertinent variables, two separate samples were generated; a triggered sample to model the data and an unbiased sample to generate the denominator of the trigger efficiency distributions. Both were generated with the same decays and run range. However, the biased sample was processed with the CDF trigger simulation, whereas the unbiased sample was not. In addition, all event selection criteria suspect of causing lifetime bias were relaxed for the unbiased sample.

The common features of the two samples include the same quark and B meson generator software [15] and table of decay widths[19]. The good run criteria used for the data was also used for both sets of Monte Carlo samples. Finally, the same software package was used to reconstruct the events.

The biased sample comes from Monte Carlo datasets on tape; *oboti0* ($B \to \mu DX$), and *oboto0* ($B \to eDX$). All analysis selection criteria for this sample mirror those of the data.

The untriggered sample is shown to be unbiased in lifetime (Fig.'s 9 and 10), that is, the distribution of generated ct matches a true exponential with the correct input average decay constants.

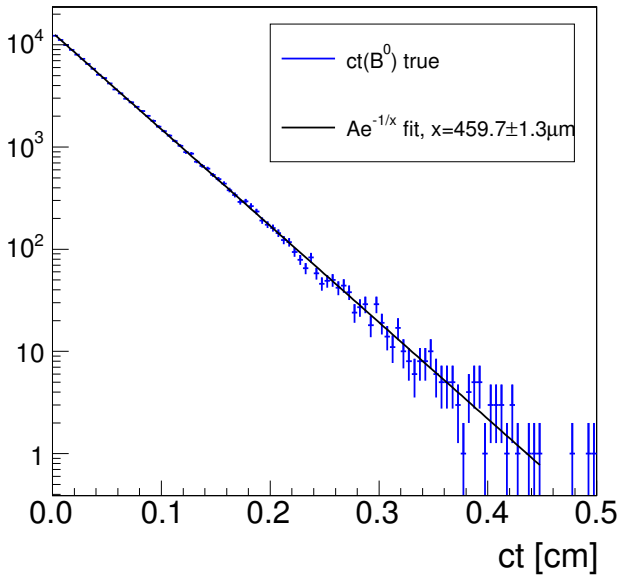


Figure 9: Generated decay length distributions from the unbiased Monte Carlo B^0 sample.

The samples are modified to account for the disagreement with the L_{xy}, σ_{Lxy} and $p_T(\ell D)$ variables in the data samples. Since these quantities are directly applied to

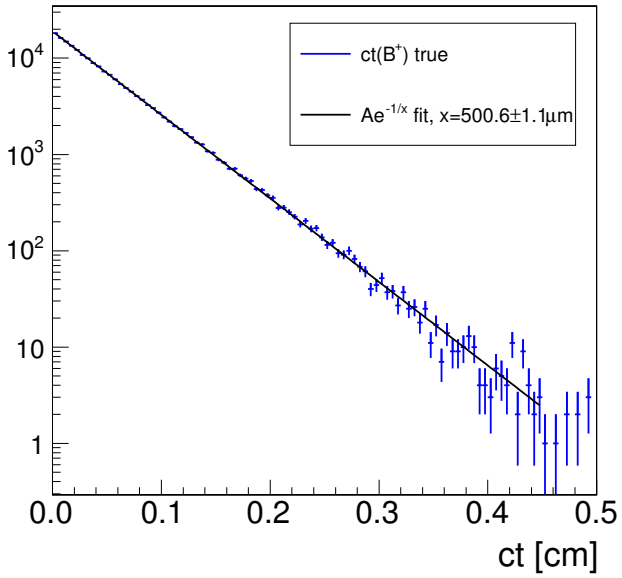


Figure 10: Generated decay length distributions from the unbiased Monte Carlo B^+ sample.

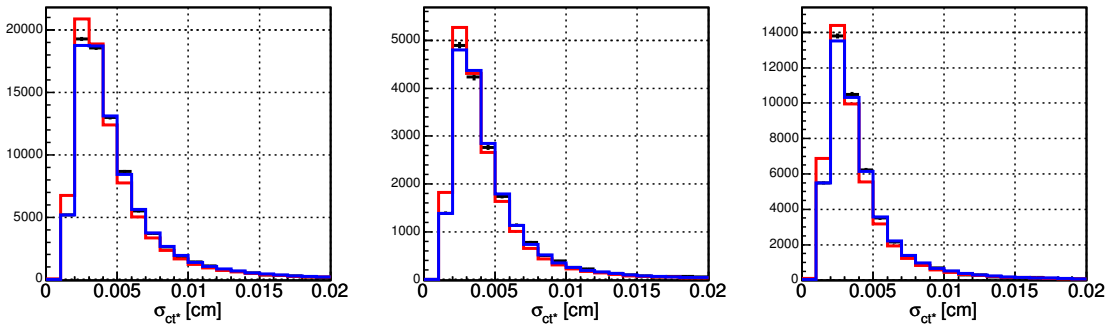
generate the K factor and trigger bias distributions, it is critical that they agree with the data and that their behavior is well understood.

5.1 Rescaling σ_{Lxy}

The first modification of the simulated quantities involves rescaling the secondary vertex error matrix [14] to account for additional smearing from the underlying event. However, unlike the procedure outlined in [14], modification of the primary vertex is not necessary since event-by-event primary vertexing is not applied. The primary vertex is determined from the beam line. A new σ_{Lxy} is calculated from the rescaled error matrix which improves the σ_{ct^*} distribution so that it agrees with data (Fig. 11).

5.2 Soft Pion Reconstruction

The soft pion reconstruction efficiency is higher in the Monte Carlo sample compared to data so this requires adjustment. Comparing the soft pion p_T in the Monte Carlo and data (Fig. 12) shows low p_T tracks from pions are more likely to be reconstructed in the Monte Carlo. Agreement with the data is achieved by "sculpting" (selectively removing soft pions) the Monte Carlo sample to match the data. The soft pion p_T spectrum and the fraction of reconstructed D^* versus D^0 are corrected in this way (these must match data since B^0/B^+ fractions used to fit the lifetimes in the data are derived from Monte Carlo). The ratio of Monte Carlo to data (lower plot in Fig. 12)



(a) (b) (c)

Figure 11: The error on ct^*, σ_{ct^*}, from Monte Carlo before (red) and after (blue) rescaling vs data (black) for modes (a) ℓD^0, (b) ℓD^*,and (c) ℓD^+.

is fit with a parametric function, an 8th order polynomial, which is used to select and then remove D^* candidates, but retain them as D^0 candidates. This method forces agreement for the Monte Carlo $p_T(\pi_s)$ spectrum, and leads to an improved $p_T(\ell D)$ distribution.

5.3 Reweighting $p_T(\ell D)$

Although the $p_T(\ell)$ and $p_T(D)$ distributions in the Monte Carlo agree quite well with the data, the $p_T(\ell D)$ distribution does not. This disagreement is most pronounced in the low p_T region (Fig. 13). Reweighting of the Monte Carlo p_T distribution is therefore necessary before generating the K factor distributions. A weight function - similar to the form used in the soft pion p_T comparison - is applied to $p_T(\ell D)$ prior to generating the K factor and efficiency distributions in order to gain agreement with data (Fig. 14).

5.4 Generating K-factors

K factor distributions are extracted from the unbiased sample. The K factor is defined by the following,

$$ct^* = \frac{L_{xy}(\ell D) M_B}{p_T(\ell D)} = \frac{ct}{K} \tag{3}$$

where

$$K = \frac{p_T(\ell D)}{p_T(B)} \frac{L_{xy}(B)}{L_{xy}(\ell D)} \tag{4}$$

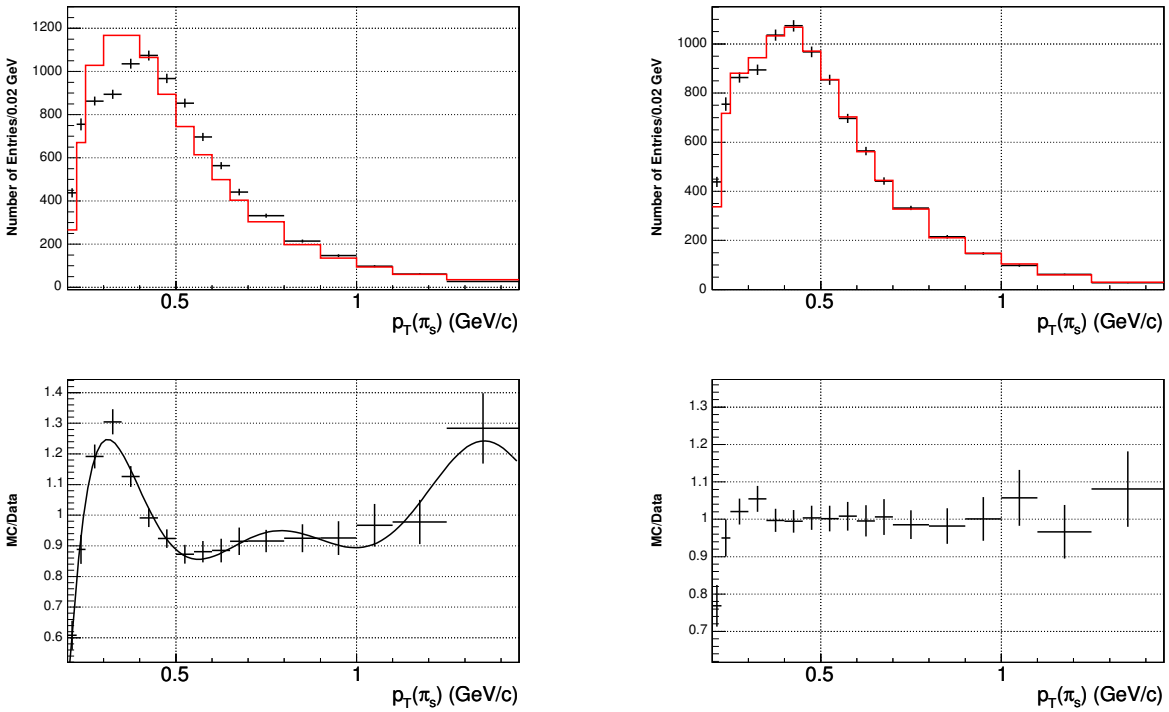


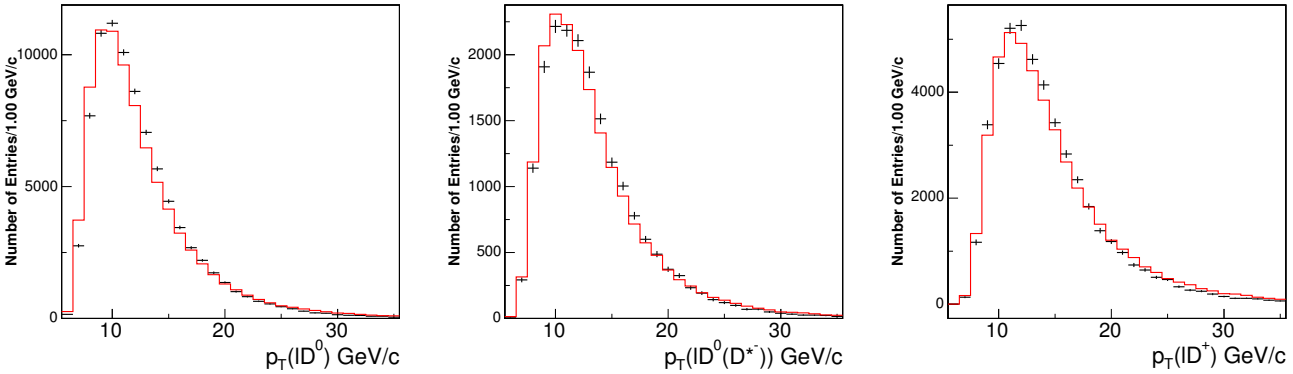
Figure 12: Top plots: Soft pion p_T spectrum in Monte Carlo (red) and data (black) before (left) and after (right) sculpting. The Monte Carlo distribution is normalized to the data. Bottom plots: Ratio of Monte Carlo and data.

K is extracted from each event in the unbiased sample and reweighted according to section 5.3. All values in equation 4 are the true generated quantity. The ratio of L_{xy} has a neglegible effect but is included for completeness. Figure 15 shows all six K factor distributions for each major decay mode in the unbiased sample. Figure 16 displays the distributions in the biased sample.

K is strongly correlated with ct (see Fig. 42) by the trigger and the selection criteria. In order to lessen this correlation and correctly determine the K factor distribution, which is applied to the likelihood function and the trigger bias function, all ct^* biasing cuts are removed. This unbiased K factor distribution is applied to the denominator of the efficiency function (see next section) and to the signal probability function in the likelihood function (see 7.1).

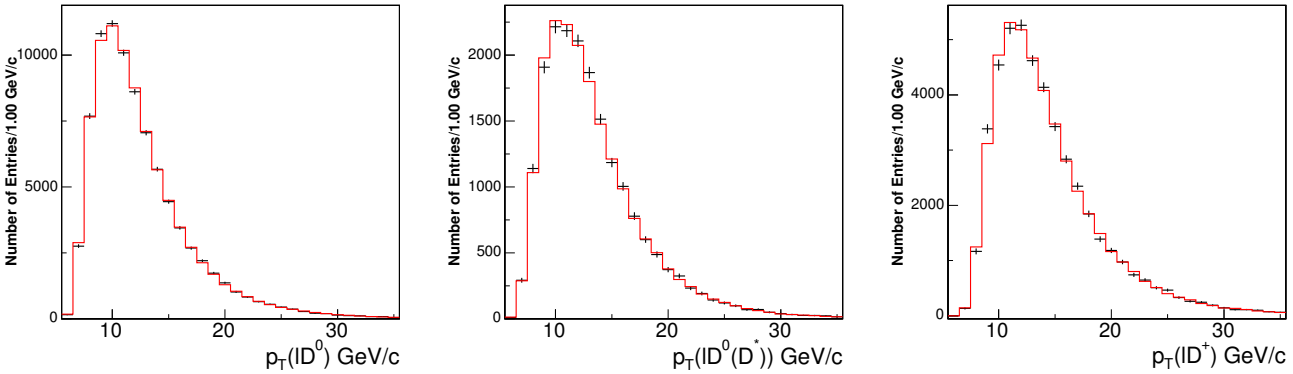
5.5 The Trigger Bias on Lifetime

The trigger bias effect is quantified by a ratio: the number of events that pass the trigger and cuts versus the total number of events generated. The numerator is a histogram filled with events from the Monte Carlo with realistic trigger simulation and all



(a) (b) (c)

Figure 13: Transverse momentum of $\ell + D$ for data (black) and Monte Carlo (red) in (a) ℓD^0, (b) ℓD^*, and (c) ℓD^+ samples.



(a) (b) (c)

Figure 14: Reweighted transverse momentum of $\ell + D$ for data (black) and Monte Carlo (red) in (a) ℓD^0, (b) ℓD^*, and (c) ℓD^+ samples.

applied selection criteria. The denominator histogram is generated by randomly filling it according to a convolution of the decay exponential with the resolution function, assumed to be a Gaussian. Additionally, the events in the denominator are smeared by the K factor from an unbiased event taken from the realistic Monte Carlo samples.

For each $B_x \to lD^x$ mode,

$$\epsilon_i(Kct^*) = \frac{N_{triggered}}{N_{produced}} = \frac{Histo^{SVT+cuts}(Kct^*)}{\sum_i exp(ct_i, c\tau_{MC}) \otimes Gauss(\sigma^i_{ct})} \tag{5}$$

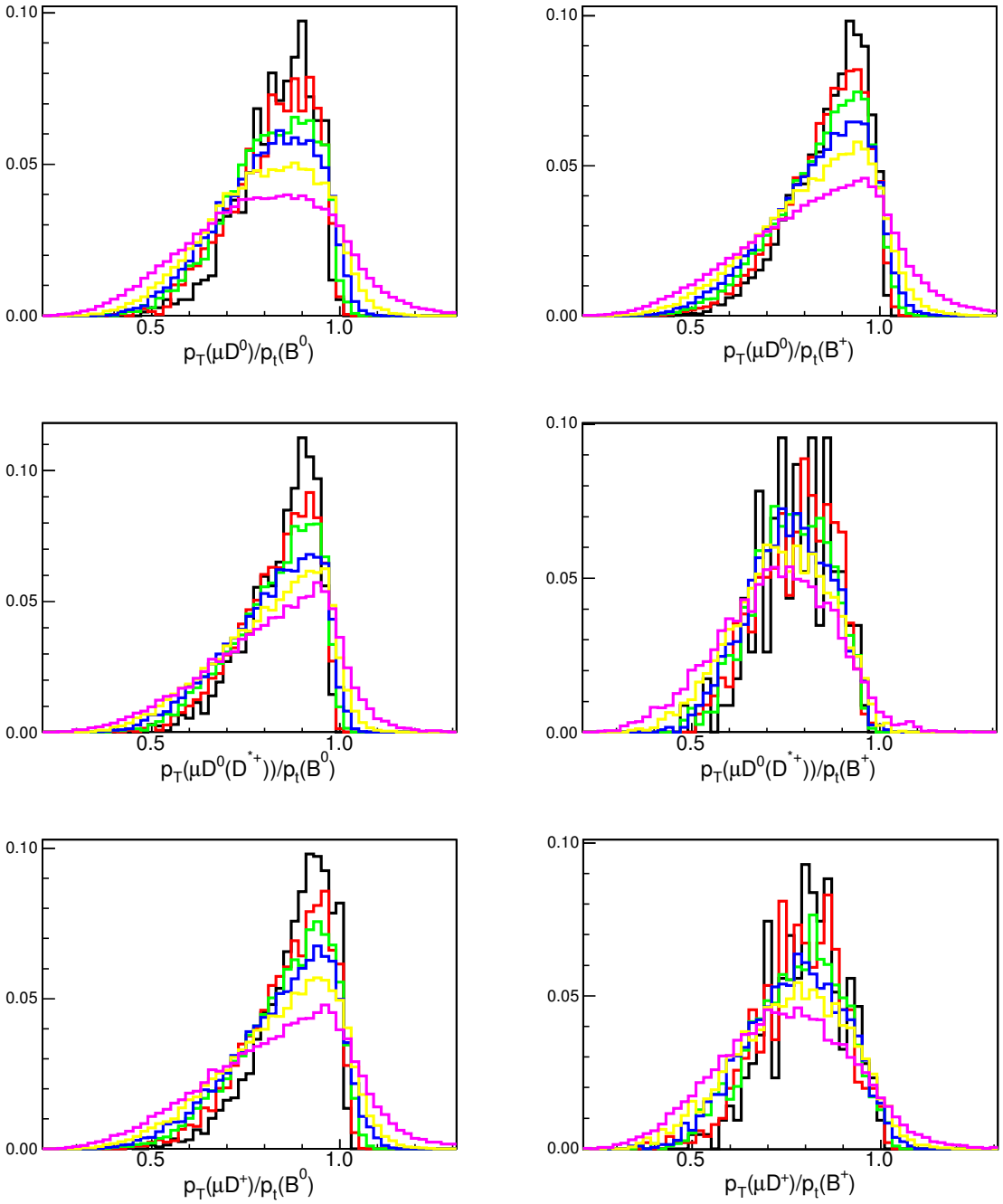


Figure 15: K factor distributions shown for all σ_{ct*} bins from the unbiased Monte Carlo sample.

The resulting distribution (Fig. 17) is then parameterized with the following function,

$$\epsilon(x) = a + (p_0 + p_1 x + p_2 x^2)(f_1 e^{-\lambda_1 x} + f_2 e^{-\lambda_2 x}) \quad (6)$$

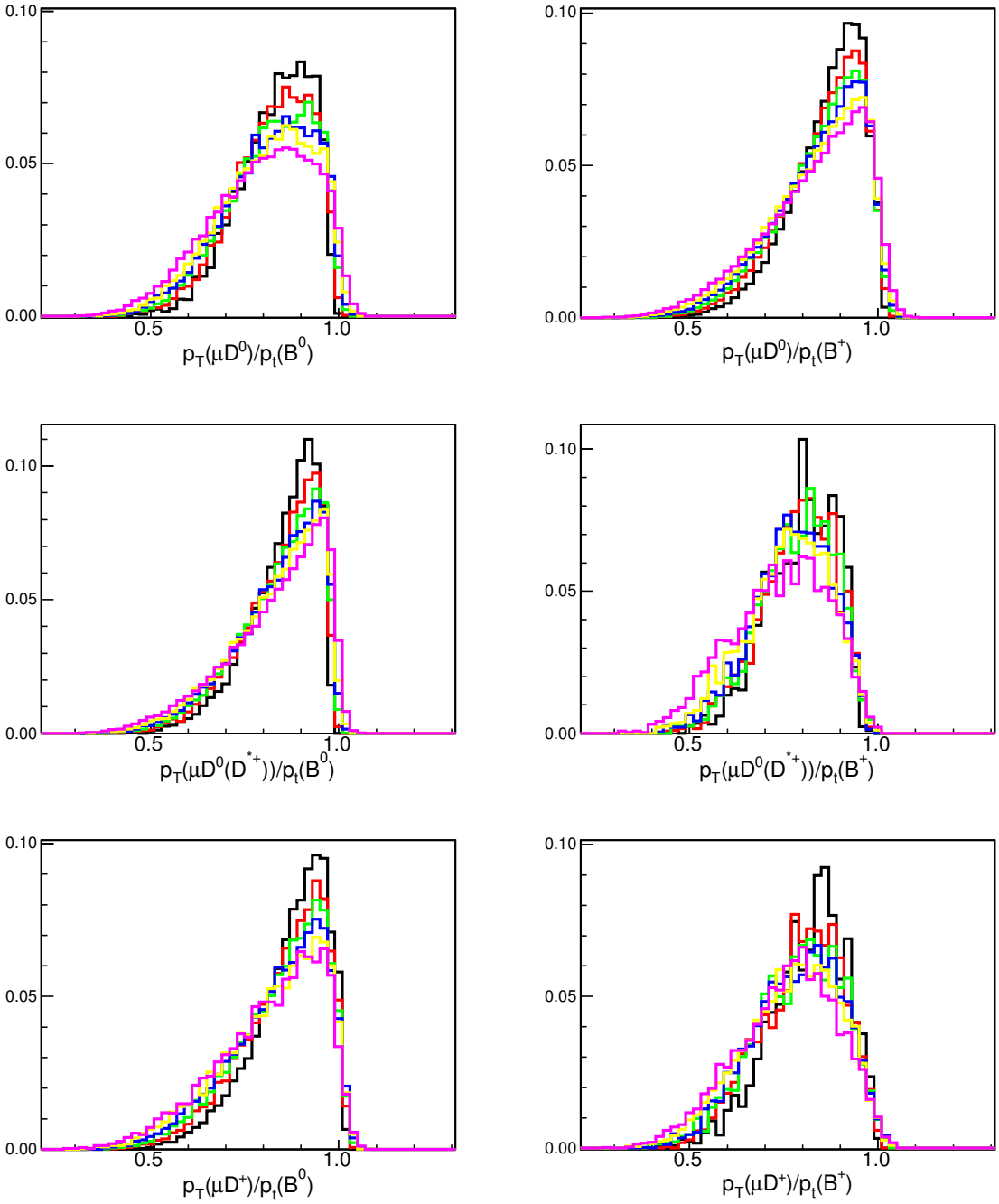


Figure 16: K factor distributions shown for all σ_{ct*} bins from the biased Monte Carlo sample.

where

$$x = Kct^*$$

Correlations caused by the trigger are explained and examined in the appendix 11.1. The correlations of K, σ_{ct^*}, and the trigger are treated by binning the efficiencies and K factor distributions in reasonable bins of σ_{ct^*} (Table 6). These correlations must also be considered when constructing the signal function (section 7.1).

Bin no.	σ_{ct^*} Low	σ_{ct^*} High
1	$0.0\mu m$	$24.0\mu m$
2	$24.0\mu m$	$30.0\mu m$
3	$30.0\mu m$	$37.2\mu m$
4	$37.2\mu m$	$46.8\mu m$
5	$46.8\mu m$	$66.6\mu m$
6	$66.6\mu m$	$400.0\mu m$

Table 6: The six bin ranges for σ_{ct^*}.

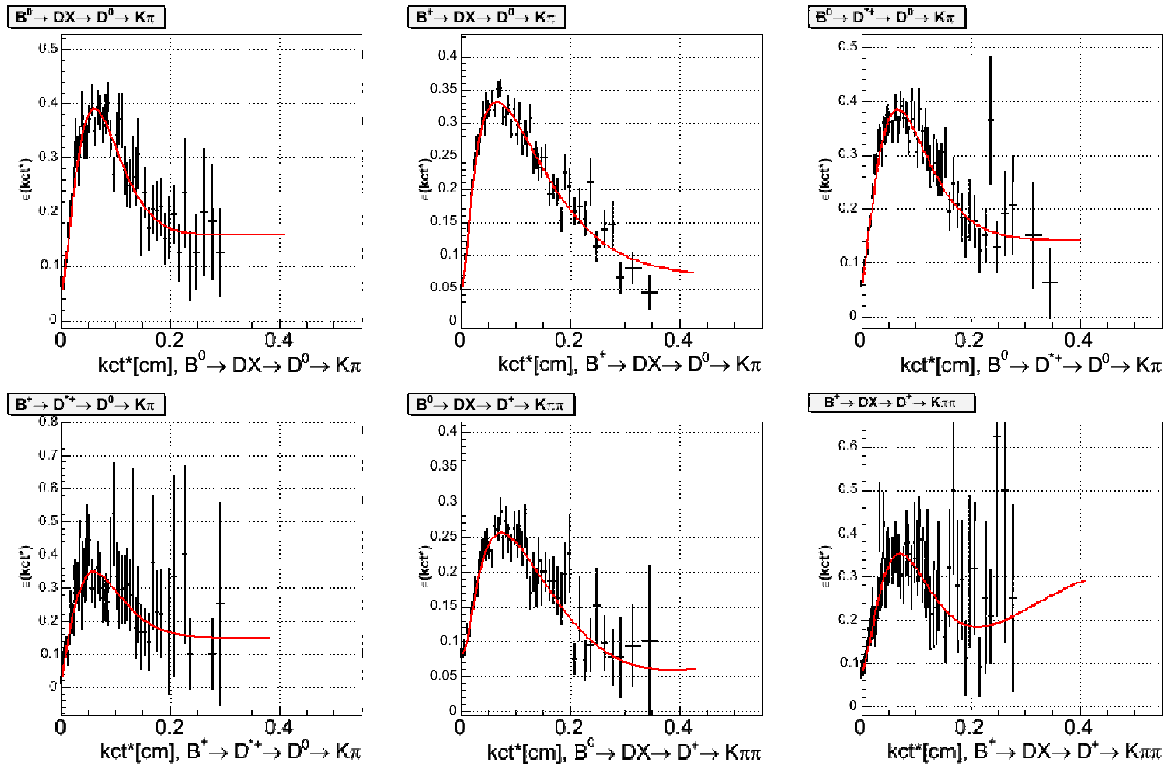


Figure 17: Trigger efficiency curves as functions of reconstructed decay length for one bin of σ_{ct^*}.

5.6 Lepton+Secondary Vertex Trigger efficiency correction

It has been shown in other lifetime analyses [16] utilizing a lifetime biased trigger at CDF that the Monte Carlo does not correctly reproduce the bias towards L_{xy}. There are two likely explanations for the disagreement between the data and Monte Carlo. The first is topological differences between the samples. The second is the difference in track acceptance. The silicon hit deposition model in the simulation does not accurately represent the real detector.

The disagreement can be removed by a method which compares the efficiency in data of the $B \to J/\psi X$ sample, where $J/\psi \to \mu^+\mu^-$, to the efficiency derived from a Monte Carlo sample. The J/ψ sample is well suited since the quantity of interest is the ratio of efficiencies and any kinematic differences in the decays will cancel out. The ratio, R, is defined as,

$$R = \frac{\epsilon(L_{xy})_{Data}}{\epsilon(L_{xy})_{MC}} = \frac{(\frac{L_{xy}(Data;\ell+SVT)}{L_{xy}(Data;pretrigger)})}{(\frac{L_{xy}(MC;\ell+SVT)}{L_{xy}(MC;pretrigger)})} \tag{7}$$

where the data distributions have been sideband subtracted.

The procedure outlined in section 8.2 of [16] is applied in this analysis, however, the event selection criteria and applied triggers are changed to match the lepton and displaced track scenario. A ratio of the efficiencies is calculated and a function fit to the distribution (Fig. 18). This correction function is then applied to the triggered ℓD Monte Carlo samples during generation of K factors and ct efficiency functions in the same fashion as the $p_T(\ell D)$ reweighting method.

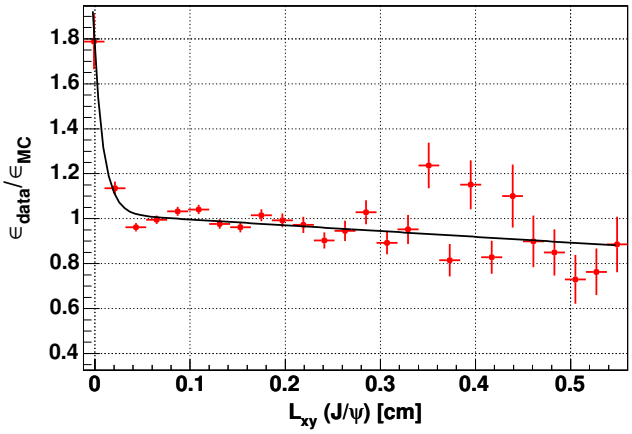


Figure 18: Trigger efficiency ratio and a polynomial fit function for the sideband subtracted signal region.

6 Background Samples

There are several sources of background events requiring separate methods to determine. Combinatorial backgrounds which can be fully reconstructed in the data must be removed. Additional combinatorial background is accounted for with a sideband subtracting method, prompt and $b\bar{b}$ are derived from understanding wrong sign events, and other B meson backgrounds are estimated from realistic Monte Carlo events. However, these backgrounds are treated similarly in the likelihood fitter: templates for the probability distributions of the relevant variables, ct^* and σ_{ct^*}, are constructed and fixed in the fit for the lifetimes along with calculated fractions of background to signal.

The shapes of the backgrounds in the ct^* variable generally share characteristics. Long lifetime events are seen in all source samples. This is expected since they represent decay products of a B or D hadrons. The wrong sign events can be separated into charm decays with a fake lepton attached (prompt decays), and exponential decays with long lifetimes from $b\bar{b}$ decays. Templates are contructed by combining two shapes to describe the prompt and long lived components; a Gaussian shape and exponential functions convoluted with Gaussians. The parametric function templates for each background are included in the relevant subsection below.

Background % of signal	ℓD^0	ℓD^*	ℓD^+
Fake leptons	0.98	0.69	1.71
$b\bar{b}$	0.66	0.21	0.22
$B \to \tau DX$ $D \to K\pi, K\pi\pi,\ \tau \to \mu\nu\nu$	1.73	1.31	1.83
$B \to D1D2X$ $D1 \to K\pi, K\pi\pi,\ D2 \to \mu\nu X$	0.68	0.29	1.10
$B \to D1D2K,$ $D1 \to K\pi, K\pi\pi,\ D2 \to \mu\nu X$	0.68	0.50	1.46
Total	4.74	3.01	6.32

Table 7: Table of background real D decays.

Two exponentials convoluted with a Gaussian function plus one other Gaussian.

$$\begin{aligned} P_{Bg} &= \frac{f_1 f_2}{N\lambda_1} e^{-\frac{ct^*}{\lambda_1}} \otimes \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(ct^*-\mu_1)^2}{2\sigma_1^2}} + \frac{f_1(1-f_2)}{N\lambda_2} e^{-\frac{ct^*}{\lambda_2}} \otimes \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(ct^*-\mu_2)^2}{2\sigma_2^2}} \\ &+ \frac{(1-f_1)}{N\sqrt{2\pi}\sigma_3} e^{-\frac{(ct^*-\mu_3)^2}{2\sigma_3^2}} \end{aligned}$$

One exponential convoluted with a Gaussian function plus one other Gaussian.

$$P_{Bg}(ct^*) = \frac{f_1}{N\lambda_1} e^{-\frac{ct^*}{\lambda_1}} \otimes \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(ct^*-\mu_1)^2}{2\sigma_1^2}} + \frac{(1-f_1)}{N\sqrt{2\pi}\sigma_2} e^{-\frac{(ct^*-\mu_2)^2}{2\sigma_2^2}}$$

One Gaussian.

$$P_{Bg} = \frac{1}{N\sqrt{2\pi}\sigma} e^{-\frac{(ct^*-\mu)^2}{2\sigma^2}}$$

6.1 Combinatorial: Sideband subtraction method

The sidebands of the D meson mass distribution (Fig. 8) are used to make probability functions for combinatorial background ct^*, and σ_{ct^*}. The sideband regions are defined as the D mass candidates within $5\sigma_{M(D^0)} \rightarrow 8\sigma_{M(D^0)}$ from the mean value of the D mass peak. The value of $\sigma_{M(D^0)}$ is determined by CDF's mass resolution and is derived from the narrower Gaussian of the D mass fit; it is about 8 MeV/c^2. Since the background shape is sloped and the left and right sidebands slightly differ in their shapes of ct^*, each sideband distribution is separately scaled such that their contribution to the signal region is correctly proportioned.

Two exponentials convoluted with Gaussians plus another Gaussian shape are used to model the background ct^* shape,

$$\begin{aligned} P_{Bg}(ct^*) &= \frac{f_1 f_2}{N\lambda_1} e^{-\frac{ct^*}{\lambda_1}} \otimes \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(ct^*-\mu_1)^2}{2\sigma_1^2}} + \frac{f_1(1-f_2)}{N\lambda_2} e^{-\frac{ct^*}{\lambda_2}} \otimes \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(ct^*-\mu_2)^2}{2\sigma_2^2}} \\ &+ \frac{(1-f_1)}{N\sqrt{2\pi}\sigma_3} e^{-\frac{(ct^*-\mu_3)^2}{2\sigma_3^2}} \end{aligned}$$

Figure 19 shows the sideband ct^* distributions for all bins of σ_{ct^*} together and the corresponding fits. Once a good fit is achieved for each sample, the fit parameters are set as seeds to fit background distributions for each bin of σ_{ct^*} separately (see Fig. 20).

A Landau function plus an exponential convoluted with a Gaussian are fit to the σ_{ct^*} distributions (Fig. 21).

$$P_{Bg}(\sigma_{ct^*}) = \frac{(1-f_1)}{N} \phi_L \left(\frac{\sigma_{ct^*} - mpv}{s_L} \right) + \frac{f_1}{N\lambda_1} e^{-\frac{\sigma_{ct^*}}{\lambda_1}} \otimes \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(\sigma_{ct^*}-\mu_1)^2}{2\sigma_1^2}}$$

where ϕ_L is a function that computes the probability density $\phi_L(x)$ at x for the Landau distribution using an approximation to the non-integratible Landau function, mpv is the most probable value, and s_L is a scale parameter.

All template parameters are fixed in the lifetime likelihood fit for the lifetimes. The parameters are varied by their uncertainties for calculating systematics (see section 9.5).

6.2 Fake leptons and $b\bar{b}$ background

A wrong sign combination candidate - the kaon and lepton have opposite signs - suggests that the lepton did not come from the same B meson as the D or possibly is a

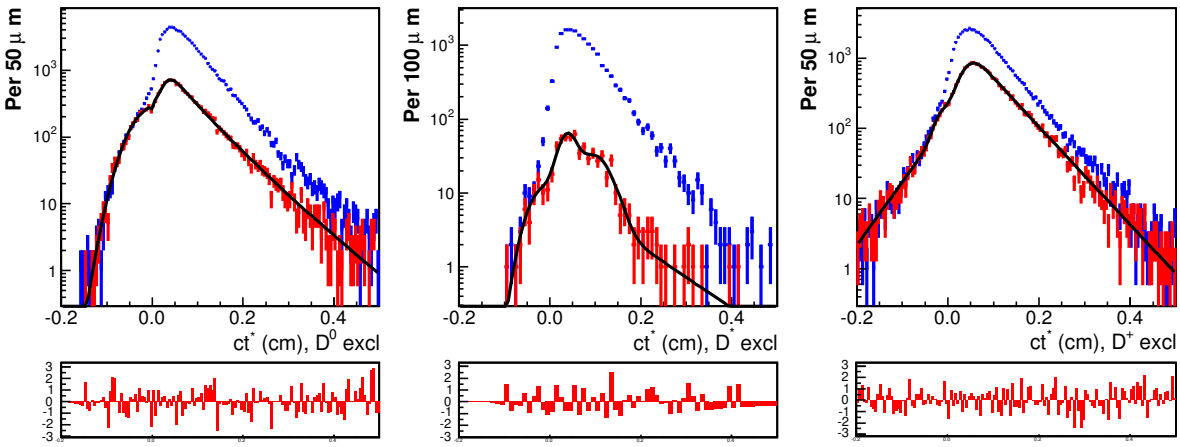


Figure 19: Pseudo decay length distributions of signal (blue), sideband regions (red), and fits to sidebands (black). Residuals of the fits are shown in the lower plot.

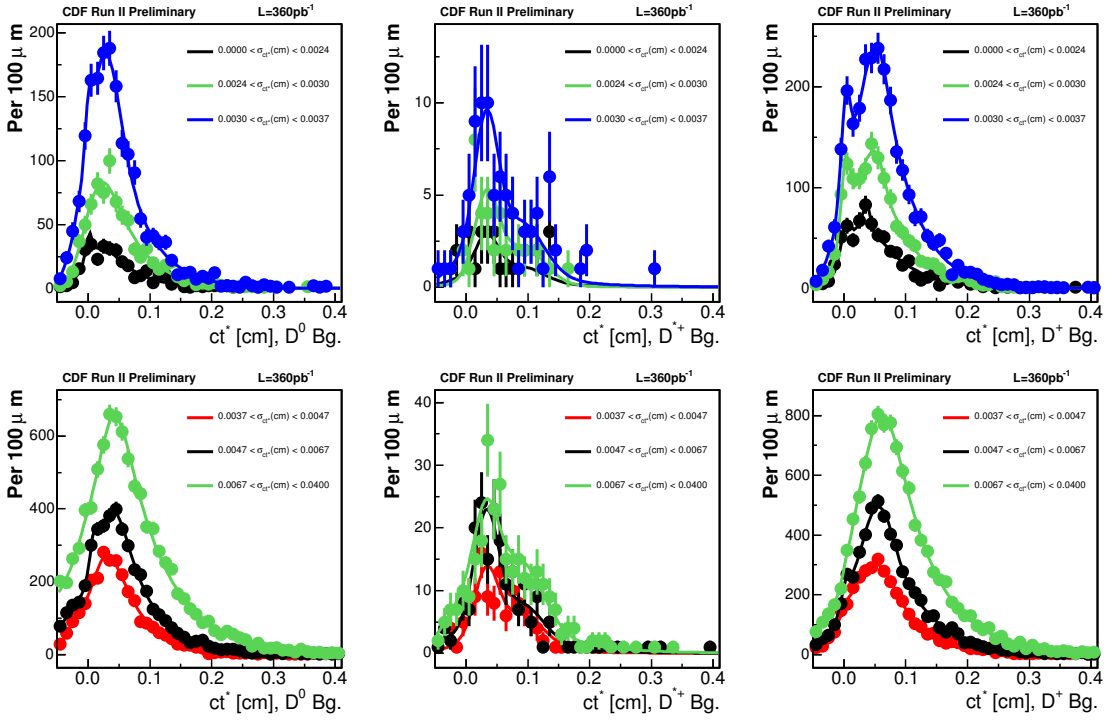


Figure 20: Mass sideband ct^* distributions and their corresponding fit templates for D^0 (left), D^* (middle) and D^0 (right). For clearer viewing the 6 binned distributions for each sample have been split onto 2 plots.

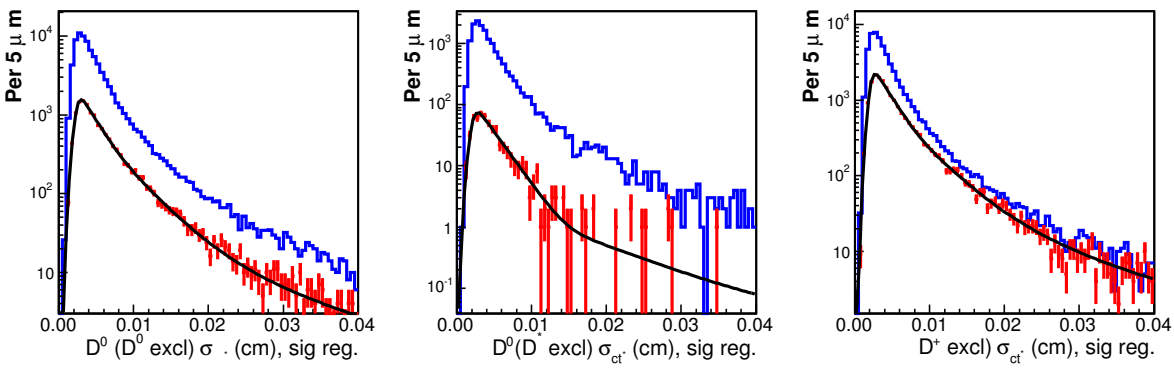


Figure 21: Pseudo decay length error distributions of signal (blue), sideband regions (red), and fits to sidebands (black).

fake lepton. If it is a product of a B, like $b\bar{b}$ production, then it will be long lived. A prompt decay event can arise from a charm decay to a real D meson at the primary vertex combined with a fake lepton.

It is expected that any D mass peak seen in the wrong sign distribution will be a scaled or direct reflection of what is happening in the right sign but which is hidden under the signal peak. Due to branching fractions and suppression, it is expected that about 75% fewer $b\bar{b}$ events in right sign $\ell + D^0$ signal events. The prompt component is due to fake leptons attached to real D's. The expectation is that the same quantity will be present in right sign candidates due to charged track asymmetry.

Figure 22 shows the ct^* distributions for wrong sign distributions with their corresponding fits. The fit functions include a Gaussian component for the fake leptons and an exponential convoluted with a Gaussian for the long lived shape,

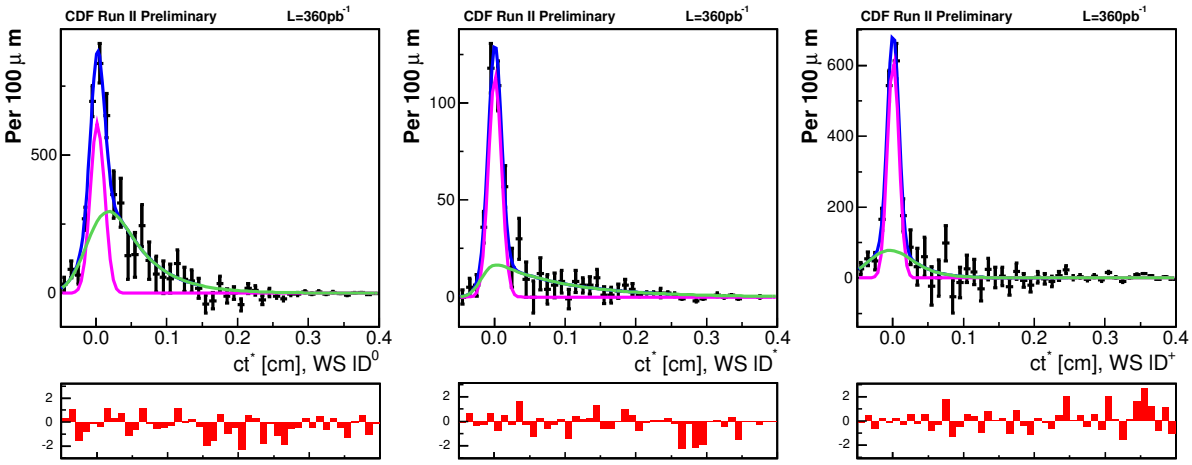
$$P_{Bg}(ct^*) = \frac{f_1}{N\lambda_1} e^{-\frac{ct^*}{\lambda_1}} \otimes \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(ct^*-\mu_1)^2}{2\sigma_1^2}} + \frac{(1-f_1)}{N\sqrt{2\pi}\sigma_2} e^{-\frac{(ct^*-\mu_2)^2}{2\sigma_2^2}}.$$

The long lived component of the $ct*$ fit template is scaled by 1/4th to produce the background template for $b\bar{b}$ but the prompt component fit template is transferred to the likelihood fitter unchanged. The parameters for these functions are held fixed during the likelihood minimization for the two lifetimes.

6.3 Physics Backgrounds

Physics backgrounds are defined as any events hidden under the right sign D mass peak that come from B meson decay, but are not the decays of interest. Realistic Monte Carlo data are generated to model the shapes and expected yields of these backgrounds.

The backgrounds considered are all B decay channels but do not contain the same decay signature as the signal channels. Each background event contains a real lepton



(a) (b) (c)

Figure 22: Pseudo decay length distribution of sideband subtracted signal region for wrong sign candidates in (a) ℓD^0, (b) ℓD^*, and (c) ℓD^+ samples including the unscaled fit templates. The residuals between the data points and the combined fit (blue) are shown in the lower plots.

and D meson and show up as right sign. In each case however, the lepton comes not from the B, but from one of its daughters, τ or second D meson. Hence, any secondary vertex would not be physical. Table 8 gives the yields retrieved from the Monte Carlo background samples when properly scaled by the triggered Monte Carlo signal sample.

Background decay channel yields (% signal)	lD^0	lD^*	lD^+
$B \to \tau DX,\ D \to K\pi, K\pi\pi,\ \tau \to \mu\nu\nu$	1.73	1.31	1.83
$B \to D1D2X,\ D1 \to K\pi, K\pi\pi,\ D2 \to \mu\nu X$	0.68	0.29	1.10
$B \to D1D2K,\ D1 \to K\pi, K\pi\pi,\ D2 \to \mu\nu X$	0.68	0.50	1.46
$Total(\%)$	3.09	2.10	4.39

Table 8: Background contribution estimates from Monte Carlo.

The shape of the ct^* distributions are fit to parametric functions of the same form as the combinatorial background. The $B \to \tau DX$ background is binned by σ_{ct^*}, however, the other two samples have small Monte Carlo statistics which makes it difficult to gain reasonable parametric fits to the distributions. These two samples are treated like the fake leptons and $b\bar{b}$ backgrounds in that they are not binned. Figures 23, 24 and 25 include the ct^* distributions from the D mass signal region and the associated fit functions.

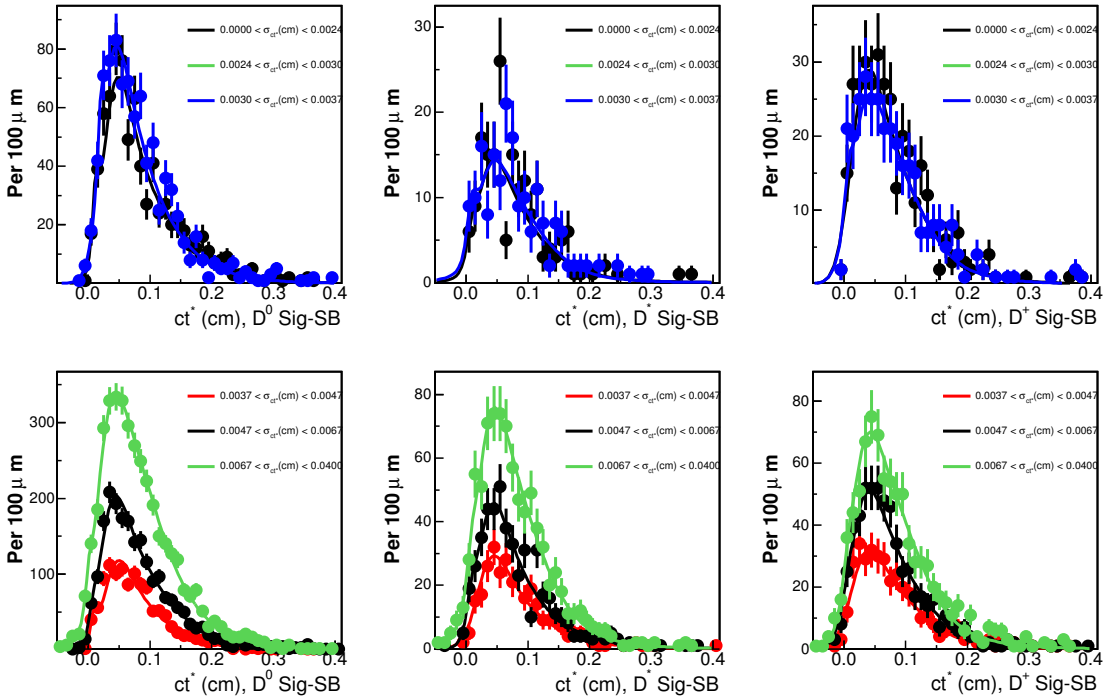


Figure 23: $B \to \tau DX$ background ct^* distributions and their corresponding fit templates for D^0 (left), D^* (middle) and D^0 (right). For clearer viewing the 6 binned distributions for each sample have been split onto 2 plots.

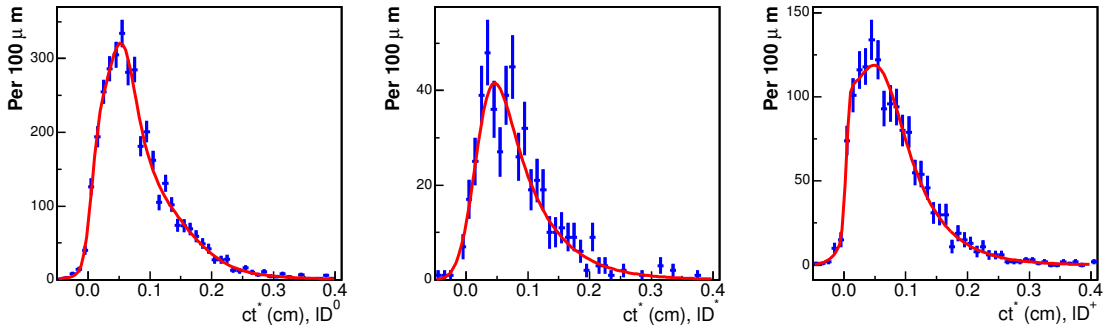


Figure 24: $B \to D1D2X$ background ct^* distributions and their corresponding fit templates for D^0 (left), D^* (middle) and D^0 (right).

6.4 Removed Backgrounds

The assumption that the $M(D^x)$ sideband regions correctly describe the combinatorial background under the signal peak requires the shape of the background ct^* to be

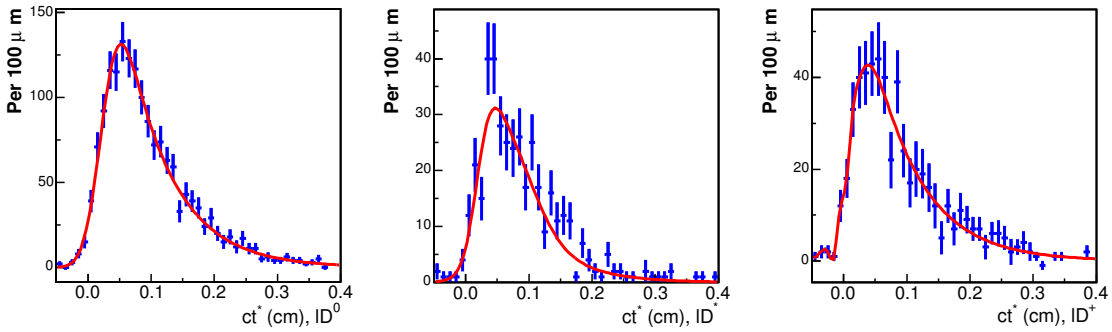


Figure 25: $B \to D1D2K$ background ct^* distributions and their corresponding fit templates for D^0 (left), D^* (middle) and D^0 (right).

symmetric between the two sidebands or that the contents of the sidebands must also lie within the signal region in the same proportion. Any asymmetric background should be removed if possible. There are a few cases where background events can be reconstructed using the particle kinematics and then effectively removed. Figure 26 shows the components - mostly asymmetric - that are removed in the ℓD^+ sample.

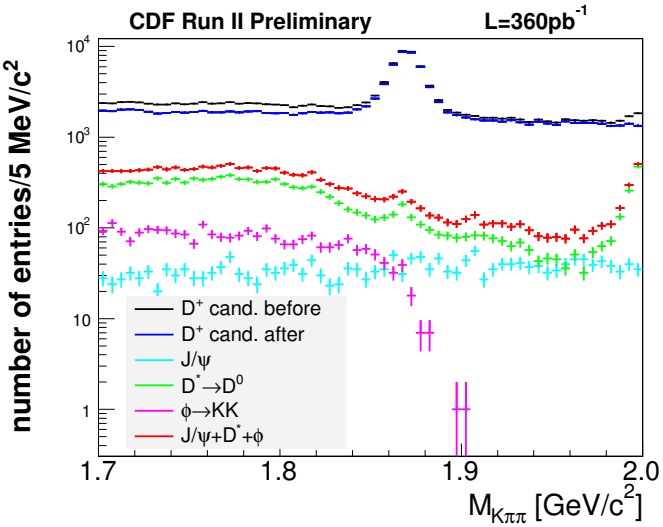


Figure 26: Removal of background in the ℓD^+ sample.

6.4.1 The case of a muon mistaken for a pion

There are many J/ψ's in each of the data samples. These are possibly from $\psi' \to J/\psi X$. In the $\mu + D^0$ mass there is a peak feature around 3.8 GeV/c^2 (near the ψ' mass) which can be almost entirely eliminated with little effort. The same feature is not seen in

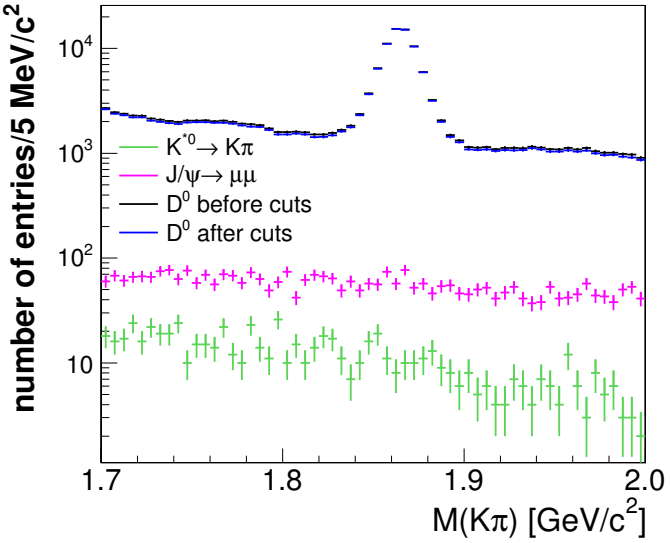


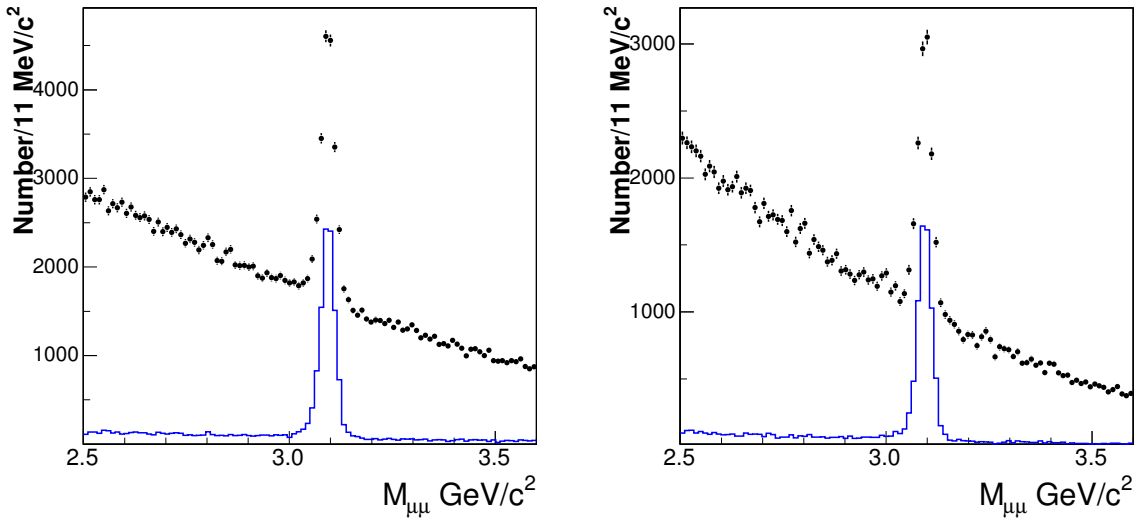
Figure 27: Removal of J/ψ's and K^{*0} decays in the ℓD^0 inclusive sample.

$e + D^0$ mass distribution due to the radiation of photons by the electron. Without the photons, the $\psi' \rightarrow$ mass peak is broadened.

The J/ψ is reconstructed by guessing that some "pion" tracks come from leptons and then calculating the invariant mass of the "dimuon" using the confirmed lepton. A size-able peak is seen at the J/ψ mass of 3.1 GeV/c^2 in Fig. 28 for the ℓD^0 sample. Although removing all candidates within 50 MeV/c^2 of the J/ψ mass is reasonable, many signal candidates would also be removed from the sample. Reduction of signal below the J/ψ peak can be obtained by checking if the "pion" track has lepton information associated with it, in other words, it is identified as an electron candidate or a muon. Any J/ψ candidate whose mass is within 50 MeV/c^2 of the J/ψ mass is then removed.

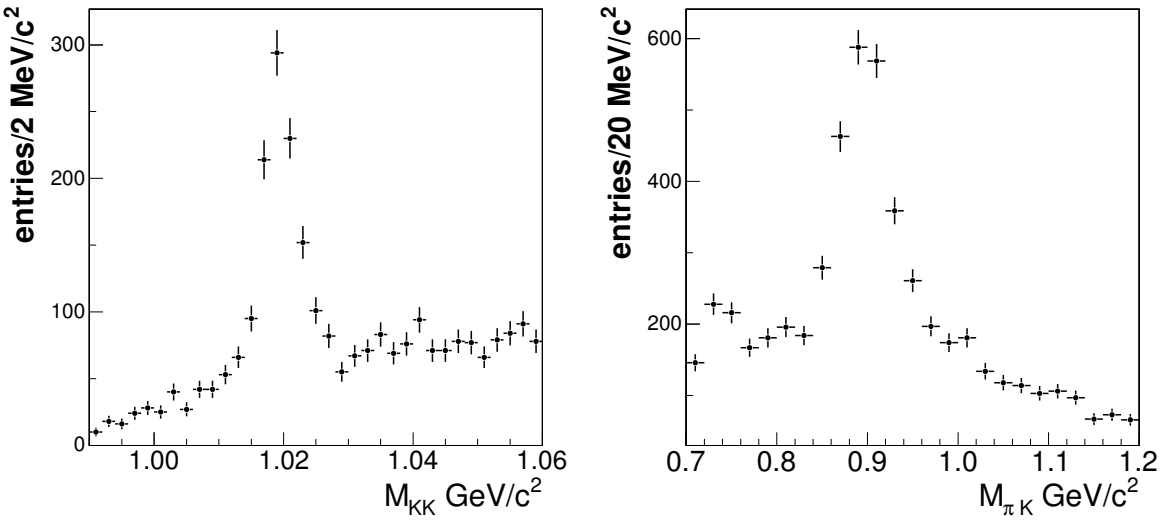
6.4.2 Kaon or Pion?

The assumption is also made that some pions are actually kaons. Calculating the invariant mass of two kaons results in a peak at the ϕ mass (Fig. 29) for the D^+ sample. Remove all events which fall within 10 MeV/c^2 of the ϕ mass with a cut also on the $M(KK\pi)$ to reduce the removal of signal events. Although the inclusive D^0 sample does not have a substantial ϕ peak, it does contain a small contribution from K^{*0} decays (Fig. 29).



(a) (b)

Figure 28: Dimuon mass in (a) ℓD^0 including ℓD^* and (b) ℓD^+ samples. Requiring the "pion" track to have lepton information results in a cleaner sample (blue) for event removal.



(a) (b)

Figure 29: (a) Invariant mass of K^+K^- from the μD^+ sample. (b) Invariant mass of πK from the μD^0 inclusive sample.

7 The Analysis

In this section the mathematical framework for extracting the lifetimes using an unbinned likelihood fit to the data is presented.

7.1 Signal Probability Function

The probability density function for the measured pseudo-decay length of exponential decays must properly include the effects due to the trigger, the error on the measurement, and the missing energy.

Let t be the true lifetime of a B meson and t_m the measured lifetime. The average lifetime, τ, is the quantity of interest. The pseudo-decay times are related to the true by, $t = Kt^*$, and $t_m = Kt_m^*$, where K is defined in equation 4. Let $F(K)$ be the normalized probability distribution (expressed as a histogram) for unbiased K. The measurement offset, $\delta = ct^* - ct_m^*$, is distributed like a Gaussian with mean zero and width σ. Finally, $G(\sigma)$ is the normalized probability distribution (expressed as a histogram) for unbiased σ.

The formulation starts with a joint probability function that has variables as independent as possible. The probability function is derived for the variables that are measured and then integrate over ones that are not. A lifetime measurement resolution σ that varies event-by-event is included. Since the K factor and lifetime resolution σ depend on common kinematical variables, they are correlated, their joint distribution is written as $P(K, \sigma)$. It would be possible to use a realistic Monte Carlo to determine $P(K, \sigma)$ as a two-dimensional histogram. However, this would require prohibitively large Monte Carlo statistics. Instead, a theorem of joint probability $[P(a, b) = P(a|b)P(b)]$ is used to write $P(K, \sigma) = F(K|\sigma)G(\sigma)$. Then, the starting variables are the lifetime (exponentially distributed), the measurement error on a given event (Gaussian distributed with mean zero and width σ), the lifetime resolution σ (probability density function $G(\sigma)$), and the K factor (conditional probability density $F(K|\sigma)$).

Then, the joint probability function is

$$\begin{aligned} P(ct, K, \delta, \sigma) &= P(ct)P(\delta|\sigma)P(K, \sigma) && (8) \\ &= \left[\frac{1}{c\tau}e^{-\frac{ct}{c\tau}}\right]\left[\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\delta^2}{2\sigma^2}}\right][F(K|\sigma)G(\sigma)], && (9) \end{aligned}$$

The measured variable is not ct, but ct_m^*, which is given by $ct = K(ct_m^* + \delta)$. Thus, it is necessary for a probability function that depends on ct_m^*, K, and δ instead of ct, K, and δ. When there is a change of variables in a probability distribution, the Jacobian needs to be included. That is, if a function $P(x, y, z)$ along with $u(x, y, z)$, $v(x, y, z)$, and $w(x, y, z)$ are defined and $P(u, v, w)$ is wanted, then

$$P(u, v, w) = \frac{P(x, y, z)}{J},$$

where

$$J = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix}.$$

In the case at hand, let $x = ct$, $y = K$, $z = \delta$, $u = ct^*_m$, $v = K$, and $w = \delta$. The Jacobian is

$$J = \begin{vmatrix} \frac{1}{K} & -\frac{ct}{K^2} & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \frac{1}{K}.$$

Thus, the probability distribution becomes

$$P(ct^*_m, K, \delta, \sigma) = \frac{K}{c\tau} e^{-\frac{(Kct^*_m + K\delta)}{c\tau}} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{\delta^2}{2\sigma^2}} \theta(Kct^*_m + K\delta) F(K|\sigma) G(\sigma).$$

Next, the efficiency as a function of Kct^*_m must be incorporated. The probability is multiplied by it, but then the resulting distribution must be normalized. Let N be the normalization constant. In addition, δ or K are not measured, so following steps must integrate over them. Since the distribution for K is determined as a histogram, the function is not continuous. Thus, the integration over K is approximated by a weighted sum over i. The weight is $F(K_i|\sigma)$ and includes the bin width.

$$P'(ct^*_m, \sigma) = \sum_i \frac{N_i K_i F(K_i|\sigma) G(\sigma) \epsilon(K_i ct^*_m, \sigma)}{c\tau\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} d\delta e^{-\frac{K_i ct^*_m + K_i\delta}{c\tau}} e^{-\frac{\delta^2}{2\sigma^2}} \theta(K_i ct^*_m + K_i\delta).$$

A change of variable in the integral from δ to $ct = K(ct^*_m + \delta)$ $[d(ct) = Kd\delta]$ gives

$$\begin{aligned} P'(ct^*_m, \sigma) &= \sum_i \frac{N_i F(K_i|\sigma) G(\sigma) \epsilon(K_i ct^*_m, \sigma)}{c\tau\sqrt{2\pi}\,\sigma} \int_0^{\infty} d(ct)\, e^{-\frac{ct}{c\tau}} e^{-\frac{(K_i ct^*_m - ct)^2}{2K_i^2\sigma^2}} \\ &= \sum_i \frac{N_i K_i F(K_i|\sigma) G(\sigma) \epsilon(K_i ct^*_m, \sigma)}{2c\tau} e^{-\frac{K_i ct^*_m}{c\tau}} e^{\frac{K_i^2\sigma^2}{2(c\tau)^2}} erfc\left[\frac{K_i^2\sigma^2 - K_i ct^*_m\, c\tau}{\sqrt{2}\,K_i\sigma\, c\tau}\right]. \end{aligned}$$

$$\frac{1}{N_i} = \sum_j \int_0^{K_i ct^*_{mmax}} d(K_i ct^*_m) \frac{F(K_i|\sigma_j) G(\sigma_j) \epsilon(K_i ct^*_m, \sigma_j)}{2c\tau} e^{-\frac{K_i ct^*_m}{c\tau}} e^{\frac{K_i^2\sigma_j^2}{2(c\tau)^2}} erfc\left[\frac{K_i^2\sigma_j^2 - K_i ct^*_m\, c\tau}{\sqrt{2}\,K_i\sigma_j\, c\tau}\right] \quad (10)$$

Integration over σ is done for the normalization. Since there is a binned histogram for the σ distribution, the integral is again approximated by a weighted sum over j, where the weight $G(\sigma_j)$ includes the bin width.

Now there is a normalized probability density function, P', that includes correlation of variables caused by kinematics and triggers, and also includes effects that smear the exponential decay curve, namely, the missing energy, the error on the measurement, and of course, the acceptance efficiency.

7.2 Modified Efficiency Function

The efficiency is defined (section 5.5) for each B candidate depending on type and decay mode. Furthermore, the efficiency is defined for Kct^*_m and σ. It helps to think of the efficiency in three dimensions; (1) the efficiency of keeping a particular decay mode, the efficiency of (2) collecting an event with a particular decay length with (3) a given resolution.

$$\epsilon(B^x \to \ell D^y, Kct^*_m, \sigma) = \frac{N_{Trig+Cuts}(B^x \to \ell D^y, Kct^*_m, \sigma)}{N_{Produced}(B^x \to \ell D^y, Kct^*_m, \sigma)} \tag{11}$$

The efficiency is parameterized in the Kct^*_m dimension, but not in the other two dimensions. For these dimensions, the efficiency parameterization of Kct^*_m is done separately for each decay mode and for six subranges of σ which gives, 6 modes $\times$ 6 σ subranges = 36 efficiency curves. It is important that the correct efficiency function is not only used in the probability function but also while normalizing. In the later case, all six efficiency functions for a particular decay mode must be included in the normalization since the sum over σ includes the total range. The same care must be given when summing $F(K_i|\sigma)$ since it is approximated in the same way, in that $F(K)$ is defined separately for each mode and each subrange of σ. For a given decay mode, equation 10 effectively then becomes,

$$\frac{1}{N_i} = \sum_j F_j(K_i)G(\sigma_j) \int_0^{K_i ct^*_{mmax}} d(K_i ct^*_m)\, \epsilon_j(K_i ct^*_m) e^{-\frac{K_i ct^*_m}{c\tau}} \frac{e^{\frac{K_i^2\sigma_j^2}{2(c\tau)^2}}}{2c\tau} erfc\left[\frac{K_i^2\sigma_j^2 - K_i ct^*_m\, c\tau}{\sqrt{2}\, K_i\sigma_j\, c\tau}\right] \tag{12}$$

where the subscript on ϵ determines which parameters are used and the subscript on $F(K)$ determines which histogram is used.

7.3 Treating Non-Gaussian resolution effects

The joint probability is constructed assuming a Gaussian distribution for the measurement uncertainty on ct^*. In CDF data the vertex resolutions do not obey a simple Gaussian shape so a method is required to correct for this.

The realistic Monte Carlo sample produces non-Gaussian resolutions which can be used to produce a resolution model. The pull, p, is defined for each event from the generated and measured values,

$$p = \frac{(ct_m - ct_{gen})}{\sigma_{ct}}$$

The distribution of pulls are fit with a triple Gaussian since it returns a good fit (Fig. 30). Define the standard deviation of the pulls to be one and then calculate a scale factor, s_y, for each of three σ_y, and the weight, f_y, of each Gaussian from the fit.

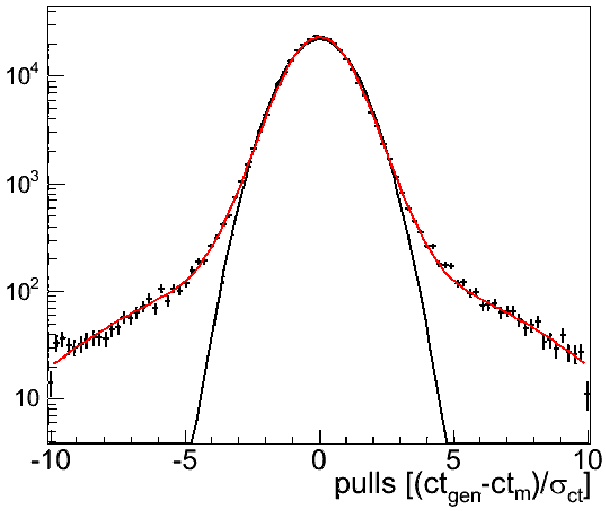


Figure 30: Pulls with a triple Gaus fit function in the ℓD^0 inclusive Monte Carlo sample.

The probability function is already calculated for one Gaussian for σ_{ct^*} (equation 9). Adding two more is straightforward by summing over the signal PDF for each gaussian shape using the correct s_y and f_y. So,

$$P(\delta|\sigma) \rightarrow \sum_y f_y P(\delta|s_y\sigma).$$

From the signal probability function, K factor distributions, efficiencies, background templates and yields, an unbinned likelihood function is formed. This function will be maximized as a function of the two lifetimes.

7.4 Signal Likelihood Construction

The likelihood signal function is constructed for each of the three semileptonic decay signatures, $B \rightarrow \ell D^0$, ℓD^*, ℓD^+ in the following way,

$$\begin{aligned} L_{B\rightarrow\ell D^y}(ct^*_m, \sigma, c\tau_0, c\tau_+) &= f_{B^0\rightarrow\ell D^y} P'_{B^0\rightarrow\ell D^y}(c\tau_0, ct^*_m, \sigma) P(m) \\ &+ (1 - f_{B^0\rightarrow\ell D^y}) P'_{B^+\rightarrow\ell D^y}(c\tau_+, ct^*_m, \sigma) P(m) \end{aligned}$$

where $f_{B^0\rightarrow\ell D^y}$ is the fraction of events that are B^0 in a lifetime biased mixture of $B^0/B^+ \rightarrow \ell D^y$. The shape of the D meson mass distribution, $P(m)$, is included here since it helps distinguish signal from background. The lifetime dependence has been explicitely included since these are the values of interest. The value to be maximized is then (assuming no background for now),

$$\ln L(c\tau_0, c\tau_+) = \sum_{n=1}^{N_{\ell D^0}} \ln L_{B\rightarrow\ell D^0}(ct^*_n, \sigma_n, c\tau_0, c\tau_+) + \sum_{n=1}^{N_{\ell D^*}} \ln L_{B\rightarrow\ell D^*}(ct^*_n, \sigma_n, c\tau_0, c\tau_+) \quad (13)$$

$$+ \sum_{n=1}^{N_{\ell D^+}} \ln L_{B\to\ell D^+}(ct_n^*, \sigma_n, c\tau_0, c\tau_+) \quad (14)$$

where $N_{\ell D^y}$ is the number of events for each decay signature. The subscript m on t^* has been dropped to keep things tidy.

7.5 Sample Composition, Defining f_{B^0}

In the ℓD^y sample, the expected number of B^0's is

$$N_{B^0} = \sigma_{B^0}\mathcal{L}BR(B^0 \to \ell D^y)\sum_i\sum_j \int \epsilon(K_i ct^*, \sigma_j)P^*(c\tau_0, ct^*, K_i, \sigma_j)d(ct^*)$$

where σ_{B^0} is the B^0 production cross section, $\mathcal{L}$ is the integrated luminosity, BR is the branching ratio for $B^0 \to \ell D^y$, ϵ is the efficiency, and P^* is the probability P' before the sum over i. This defines the number of B^0's in the sample as the number produced times the fraction that decay to the given mode times the efficiency weighted by the probability of getting a given ct^*, K and σ. A similar expression gives N_{B^+}. The fraction of B^0 in the sample is then given by

$$f_{B^0} = \frac{N_{B^0}}{N_{B^0} + N_{B^+}}$$

This makes f_{B^0} a function of τ_0 and τ_+ since, for example, if you reduce the B^0 lifetime, fewer of these events will pass the trigger cuts. The equation for f_{B^0} can be rewritten as,

$$f_{B^0}(\tau_0, \tau_+) = \frac{S_{B^0}(\tau_0)}{S_{B^0}(\tau_0) + RS_{B^+}(\tau_+)}$$

where

$$S_{B^x}(\tau_x) = \sum_i\sum_j \int \epsilon(K_i ct^*, \sigma_j)P^*(c\tau_x, ct^*, K_i, \sigma_j)d(ct^*)$$

and

$$\begin{aligned} R &= \frac{\sigma_{B^+}BR(B^+ \to \ell D^y)}{\sigma_{B^0}BR(B^0 \to \ell D^y)} \\ &= \frac{BR(B^+ \to \ell D^y)}{BR(B^0 \to \ell D^y)} \end{aligned}$$

Luminosity has canceled out and the assumption made that the production cross sections for B^0 and B^+ are the same at the Tevatron. The branching ratios are sums of individual decay channels and so this description of f_{B^0} assumes a pure ℓD^y sample. However, there is contamination of ℓD^* in the ℓD^0 sample ($D^{*-} \to D^0\pi_s^-$ without soft

pion reconstruction) so effective branching ratios from the fractions are derived from the composition in the unbiased Monte Carlo samples,

$$R = \frac{1 - f_{B^0}^{UN}}{f_{B^0}^{UN}}.$$

This is consistent with the efficiency definition - treating the unbiased samples as the generated events. This keeps the likelihood from the added complexity of dealing with the contamination explicitly, but does not allow simple means to determine a lD^* contamination systematic later.

The lifetime dependence of f_{B^0} turns out to be quite small. Varying the lifetimes by $50\mu m$ gives rise to a shift in the fractions of less than ± 0.005. This expression for f_{B^0}, evaluated with the actual generated lifetime values, also predicts quite well the fractions for each Monte Carlo sample after trigger and selection criteria are applied (Table 9). However, since there is more concern with the change in fractions over the lifetimes phase space than their absolute values, f_{B^0} is corrected to achieve precise agreement with the biased Monte Carlo sample at the actual lifetimes. This involves simply adding a constant, Δ, to the fraction expression

decay mode	unbiased MC	biased MC	f_{B^0} actual τ_0, τ_+	Δ
ℓD^0	0.292	0.222	0.221	$+0.001$
ℓD^*	0.881	0.908	0.912	-0.004
ℓD^+	0.798	0.833	0.837	-0.004

Table 9: B^0 fractions. The biased and unbiased fractions are determined by simple counting in the realistic Monte Carlo samples.

The lifetimes are changed during each minimization step, and so f_{B^0} must be recalculated. A fast way to calculate f_{B^0} is to parameterize for the two lifetime variables. Scanning the value of f_{B^0} over the range which will come into play during the minimization shows how the fraction changes (Fig. 31). The resulting shape of f_{B^0} is approximated with a two-dimensional second order polynomial. This polynomial is then used to calculate f_{B^0} during the likelihood calculation instead of the more complex and time consuming integrations of the true function.

7.6 Background Likelihood Functions

There is a model - a set of parametric functions - for every background mode (see sections 6.1-6.3). Each model is made up of parametric fits, P_{bg} to real or simulated data corresponding to different bins of σ_{ct*}. These functions serve as templates within the likelihood function and their corresponding parameters are fixed for the lifetimes fit. The following describes the likelihood formulation used for each background mode.

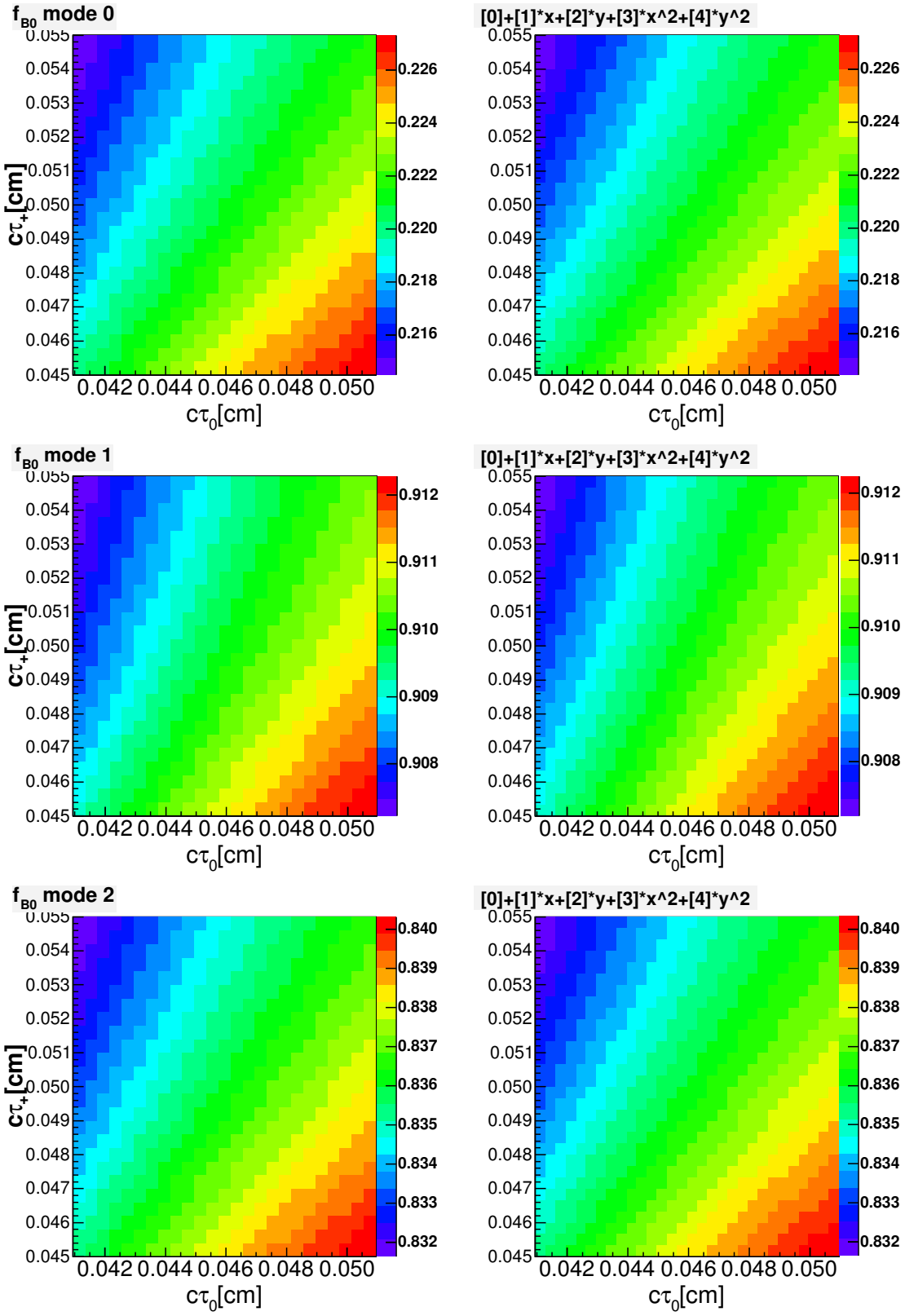


Figure 31: The value of $f_{B^0}(c\tau_0, c\tau_+)$ over a wide range (left) and the corresponding polynomial fit (right) of the shape for each mode.

As in the signal, the ct^*_m of a background event is correlated with its uncertainty but the mass of the D is not. The likelihood function is

$$L_{bg}(ct^*_m, \sigma, m) = P_{bg}(ct^*_m, \sigma) P_{bg}(m), \tag{15}$$

where $P_{bg}(ct^*_m, \sigma)$ and $P_{bg}(m)$, are normalized probability distributions of ct^*_m, σ, and the D mass, m. The correlation of ct^*_m and σ is included in the same way as the signal in that the sample is split into six bins according to σ and separate $P_{bg}(ct^*_m, \sigma)$ are calculated for each bin. The correlation within each bin is negligible so using the conditional probability again, the likelihood may be rewritten as,

$$L_{bg}(ct^*_m, \sigma, m) = P_{bg}(ct^*_m|\sigma) P_{bg}(\sigma) P_{bg}(m) \tag{16}$$

where $P_{bg}(ct^*_m|\sigma)$ represents six normalized functions for each background and $P_{bg}(\sigma)$ is a single normalized probability function for σ.

7.7 Combined Likelihood Function

The general form of the likelihood includes all the signal (equation 14) and background (equation 16) contributions for the three event types. Taking the natural logarithm, the log likelihood is given as

$$\begin{aligned}
\ln L(c\tau_0, c\tau_+) &= \sum_{n=1}^{N_{\ell D^0}} \log[f^0_s L_{B\to\ell D^0}(ct^*_n, \sigma_n, c\tau_0, c\tau_+) + \sum_i f^0_i L_{bg_i\to\ell D^0}(ct^*_n, \sigma_n)] \\
&+ \sum_{n=1}^{N_{\ell D^*}} \log[f^*_s L_{B\to\ell D^*}(ct^*_n, \sigma_n, c\tau_0, c\tau_+) + \sum_i f^*_i L_{bg_i\to\ell D^*}(ct^*_n, \sigma_n)] \\
&+ \sum_{n=1}^{N_{\ell D^+}} \log[f^+_s L_{B\to\ell D^+}(ct^*_n, \sigma_n, c\tau_0, c\tau_+) + \sum_i f^+_i L_{bg_i\to\ell D^+}(ct^*_n, \sigma_n)]
\end{aligned}$$

The fraction of signal and background for mode x are f^x_s and f^x_i respectively and $f^x_s + \sum_i f^x_i = 1$ by definition.

The MINUIT minimization package [23] then minimizes $-2\ln L$ to determine the two average decay lengths, $c\tau_0$ and $c\tau_+$, and the statistical error (using the MIGRAD subroutine) on each. Since the signal normalization constant, N, depends on the lifetimes, it must be recalculated for each decay mode for all values used in the minimization. Similarily, the fractions of B^0 for each sample must be recalculated for each minimization iteration since they too depend on the lifetime values.

7.8 Fit to data

The resulting average decay lengths from the likelihood scan are,

$$\begin{aligned}
c\tau_{B^0} &= 457.9 \pm 3.5(stat.)\mu m \\
c\tau_{B^+} &= 488.5 \pm 3.8(stat.)\mu m
\end{aligned}$$

The correlation coefficient for the two lifetimes is -0.424. Figures 32, 33 and 34 show the likelihood fit projection on pseudo-decaylength along with the data and background templates for each sample. The residuals are defined as (data-fit)/fit. A selection of parameters which are fixed in the likelihood function are given in Table 10.

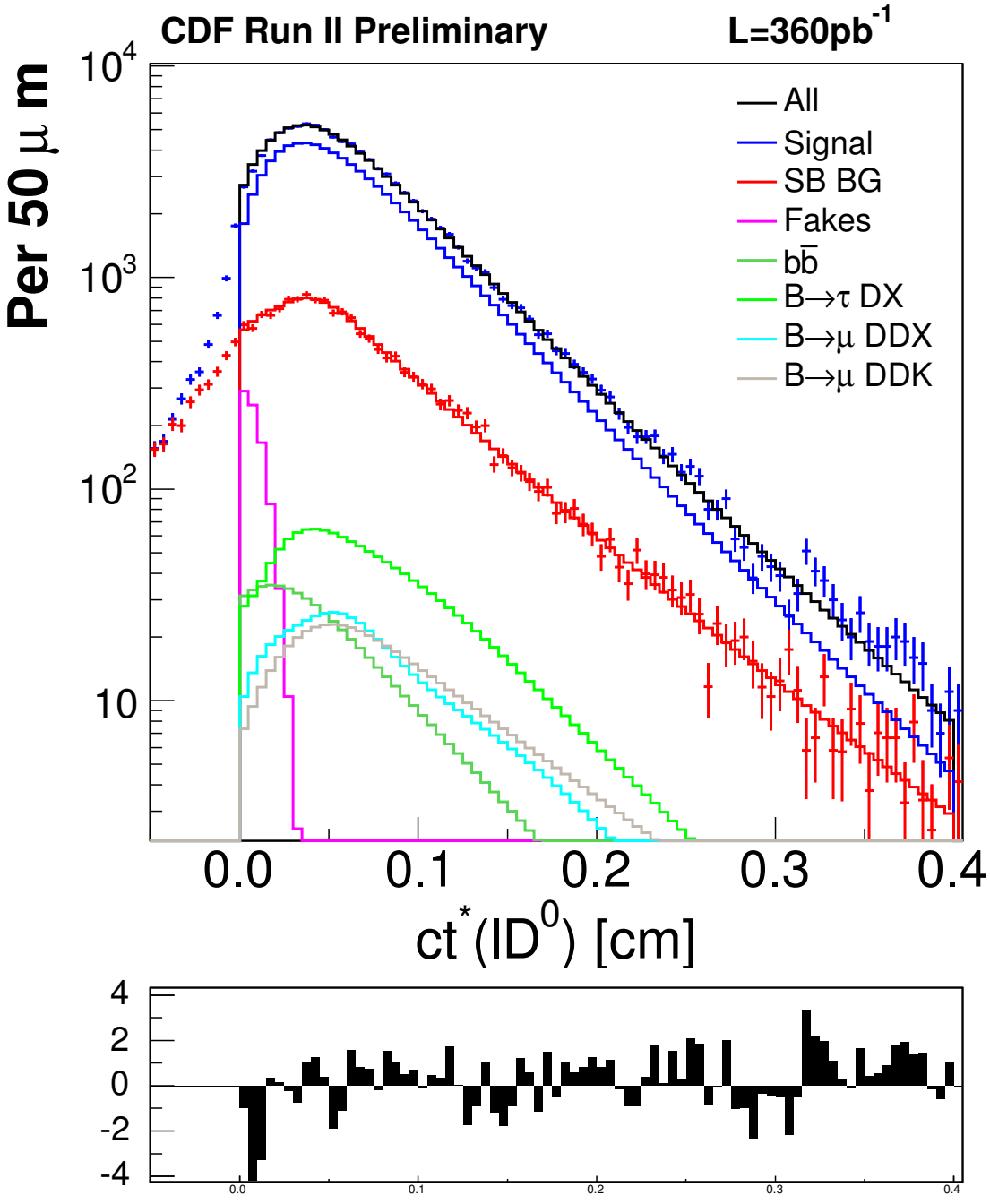


Figure 32: Fit projection of decay lengths for the ℓD^0 sample. Residuals are shown in lower plots.

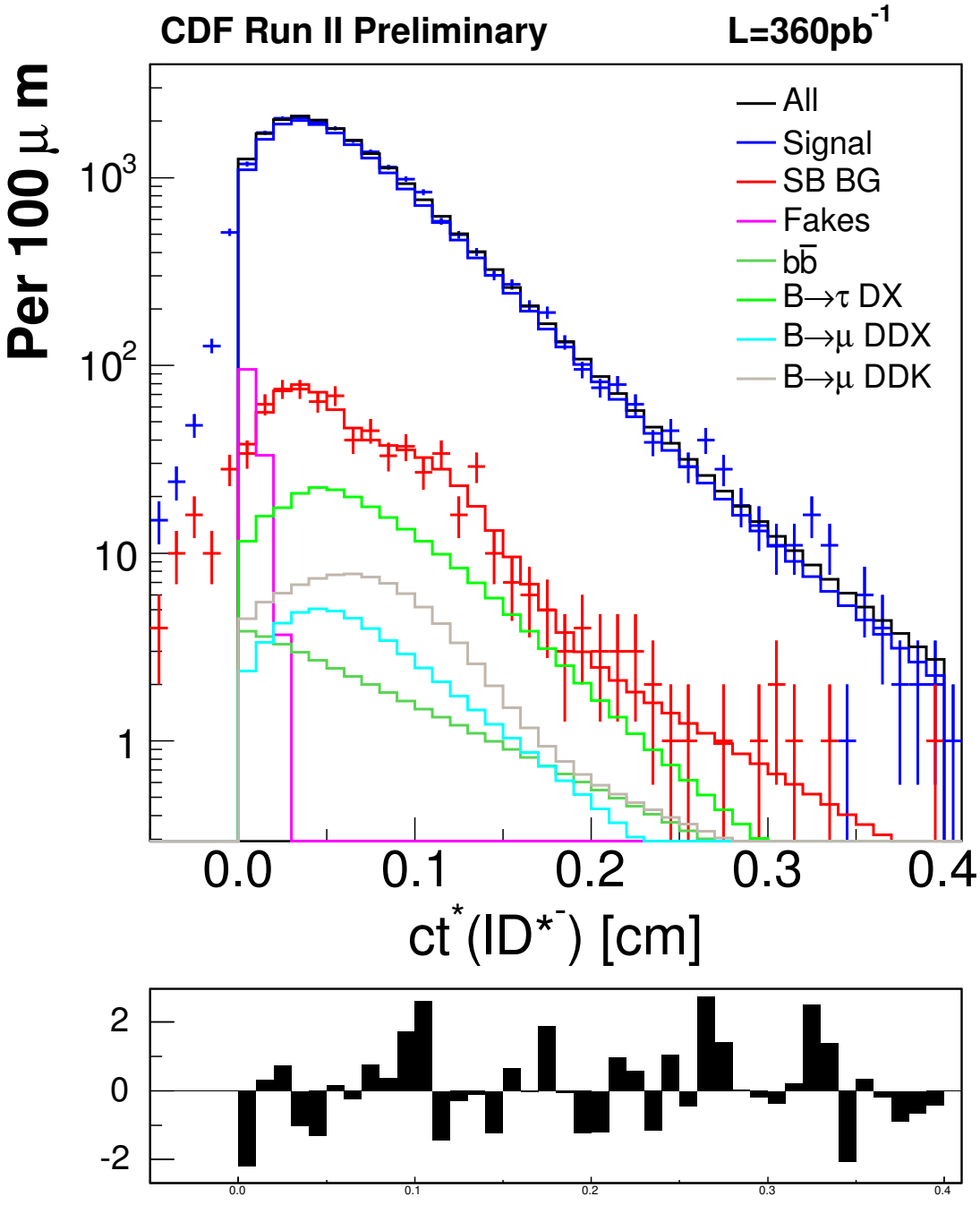


Figure 33: Fit projection of decay lengths for the ℓD^{*-} sample. Residuals are shown in lower plots.

8 Monte Carlo Validation

8.1 Toy Monte Carlo Validation

A toy Monte Carlo model was developed to validate the fitting procedure. The toy model uses templates and yields from the unbiased realistic Monte Carlo (K,f_{B^0}), data

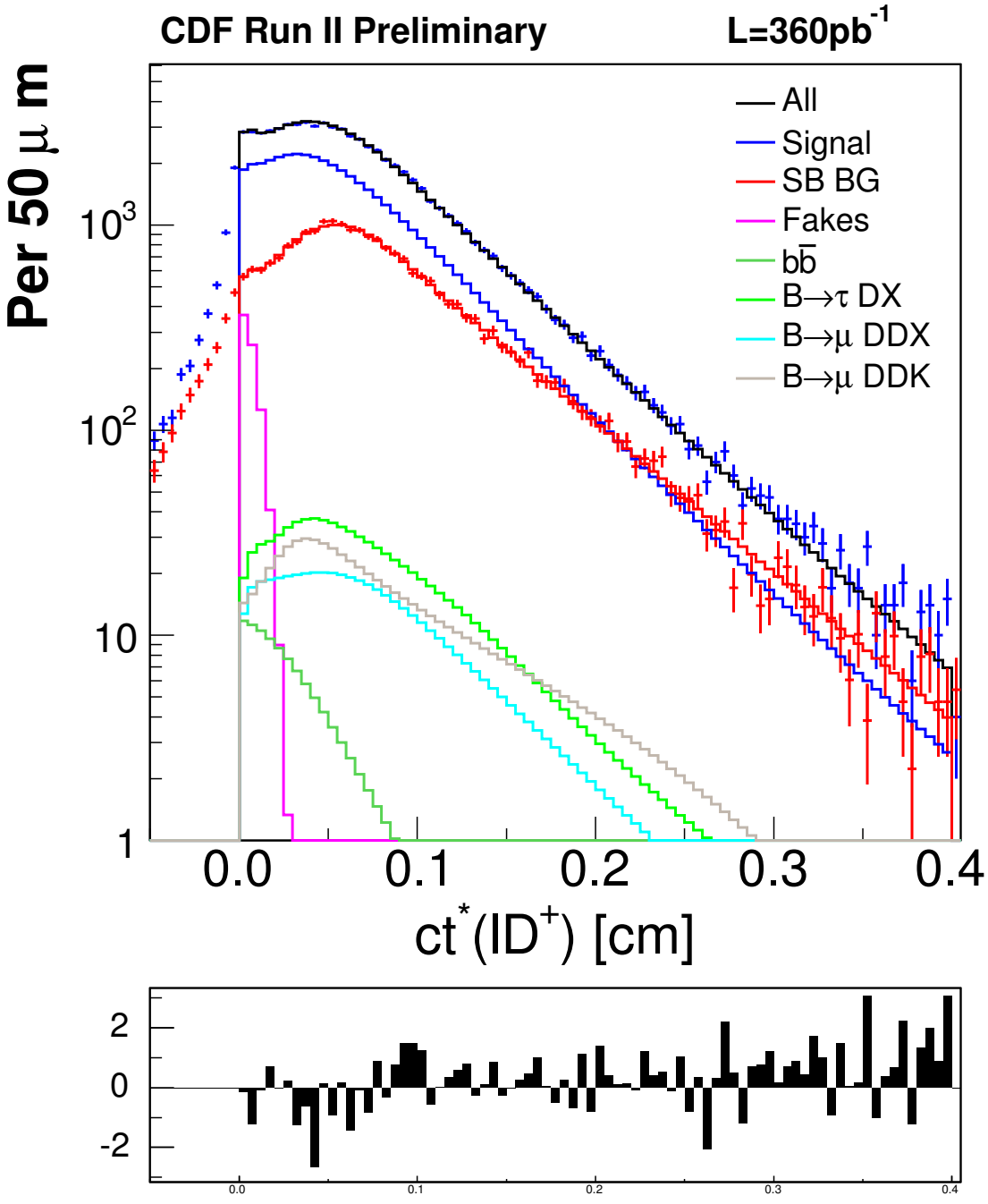


Figure 34: Fit projection of decay lengths for the ℓD^+ sample. Residuals are shown in lower plots.

(M_D, background quantities), and physics background Monte Carlo.

Each toy sample is generated in the following way. First, all parameters and def-

Parameter	ℓD^0	ℓD^{*-}	ℓD^+
f_{sig}	0.8023	0.9366	0.6312
f_{B^0}	0.2217	0.9079	0.8331
Background fractions			
$f_{\tau\to\ell DX}$	0.0139	0.0122	0.0116
$f_{\mu\to DDX}$	0.0055	0.0028	0.0069
$f_{\mu\to DDX}(D\to KX)$	0.0055	0.0047	0.0092
f_{prompt}	0.0083	0.0066	0.0115
$f_{b\bar{b}}$	0.0055	0.0021	0.0015
D mass templates			
f_m	0.6006	0.5656	0.7546
m_1	1.8648	1.8649	1.8699
σ_1	0.0072	0.0072	0.0071
m_2	1.8642	1.8638	1.8677
σ_2	0.0127	0.0119	0.0155
p_0	1342.5	57.505	1900.3
p_1	-4741.9	-490.22	-3837.8
WS D mass templates			
f_1	0.0717	0.2807	0.3389
m_1	1.8648	1.8649	1.8699
σ_1	0.0072	0.0072	0.0071
m_2	1.8634	1.8639	1.7128
σ_2	0.0127	0.0119	0.0155
σ_{ct^*} Bg. templates			
f_G	0.04098	1.00000	0.04413
μ	0.00295	0.00269	0.00271
λ	0.00429	0.00420	0.00428
σ_G	0.00070	0.00067	0.00067
MPV	0.01858	0.00500	0.01962
s_L	0.00438	0.00200	0.00398

Table 10: Table of parameters.

initions are initialized to match those of the data or RMonte Carlo. These includes background templates, signal fractions, f_{B^0} for each sample, and lifetimes. The input average decay lengths are 460.5 μm and 501.0 μm for B^0 and B^+ respectively. Next, the K factor distributions and efficiency functions are defined. Then, a candidate event is generated with the following steps;

1. Choose the D meson sample (D^0, D^*, or D^+) based on relative statistics of data samples.

2. Decide whether it is a background or signal event

3. If background, randomly choose what type of background based on relative fractions expected in sample and fill the quantities - $M(D), ct^*, \sigma_{ct^*}, \delta(M_D)(D^*only)$ - based on distributions from data.

4. if signal, select B^0 or B^+ based on expected fractions within chosen sample.

5. randomly choose true decay length, ct, based on exponential with appropriate B lifetime.

6. randomly choose σ_{ct^*} from unbiased σ_{ct^*} distribution derived from RMonte Carlo.

7. randomly choose the K factor using the bin of σ_{ct^*} to select the correct distribution.

8. calculate $ct^* = ct/K + R_G S_f \sigma_{ct^*}$ where R_G is a random value from a Gaussian shape of unit width and S_f is the same scale factor applied in data.

9. decide to keep or throw away the event depending on "trigger" and cuts by applying the appropriate efficiency function as a filter; randomly accepting Kct^* based on shape of the efficiency curve.

10. if do not accept, return to step 4, otherwise, choose $M(D)$, $\delta(M_D)$ (for D^* events), based on signal distributions.

Repeat steps 1-10 until desired number of signal events is obtained.

Two hunded toy samples were generated, each with the same number of signal events as in the data samples. Each sample was then fit with the same fit procedure as applied to the data. No bias was found in the lifetime fit results (Fig.'s 35, 36). The mean values of all the lifetimes from the likelihood scans are within a tenth of a micron of the generated lifetimes.

The pulls,

$$p = \frac{c\tau_{fit} - c\tau_{input}}{\sigma_{c\tau_{fit}}}$$

are fit with a Gaussian shape (Fig.'s 37, 38). The mean values from Gaussian fits to the pull distributions are,

$$\begin{aligned} \mu_{p_0} &= 0.024 \pm 0.080 \\ \mu_{p_+} &= -0.023 \pm 0.074 \end{aligned}$$

and widths,

$$\begin{aligned} \sigma_{p_0} &= 1.036 \pm 0.066 \\ \sigma_{p_+} &= 0.977 \pm 0.064. \end{aligned}$$

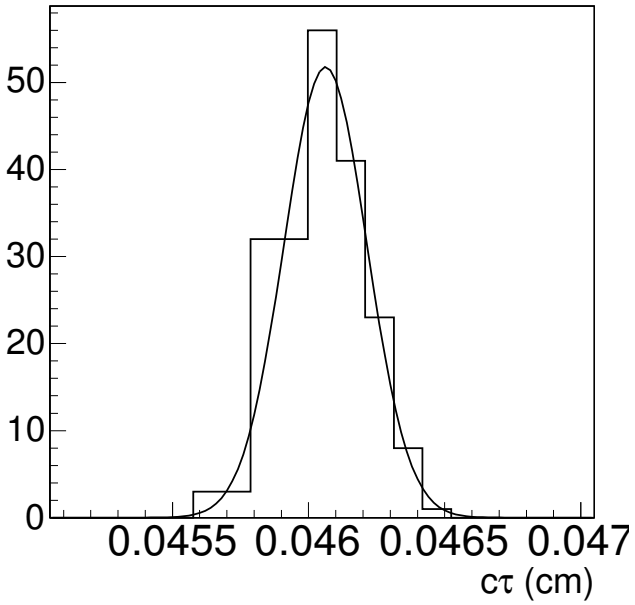


Figure 35: Likelihood fit values for B^0 decay length in 200 toy Monte Carlo samples.

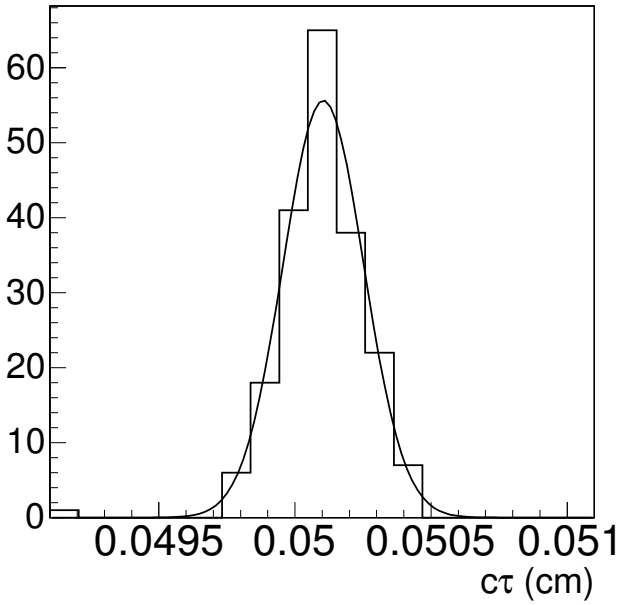


Figure 36: Likelihood fit values for B^+ decay length in 200 toy Monte Carlo samples.

These show good agreement with unit width and means which are consistent with zero. Non-unitary widths would suggest the statistical uncertainties returned by the fitter

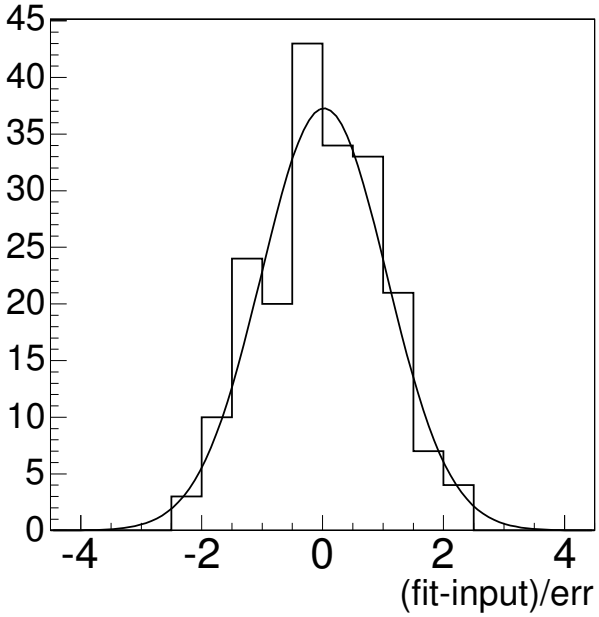


Figure 37: Pulls of B^0 decay length fits in toy Monte Carlo.

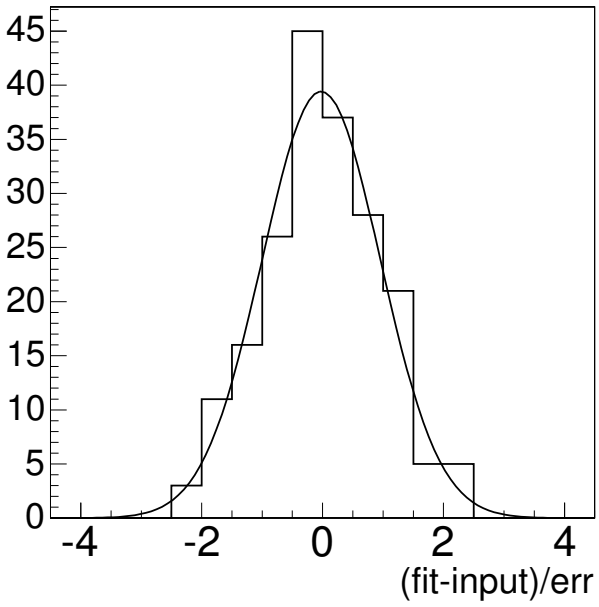


Figure 38: Pulls of B^+ decay length fits in toy Monte Carlo.

are incorrect. A mean value other than zero would imply a lifetime bias.

The toy models were also used to validate the fitter under different true lifetime

scenarios. Samples were generated with substantially lower and higher lifetimes than PDG values. The following lifetimes were used for B^0 and B^+; $360\mu m$ and $550\mu m$. No significant shift was seen in the resulting likelihood fits for either case. In another test, the lifetimes for B^0 and B^+ were switched and toy samples generated. Again, the average lifetime fit values did not deviate from the generated values.

8.2 Realistic Monte Carlo Validation

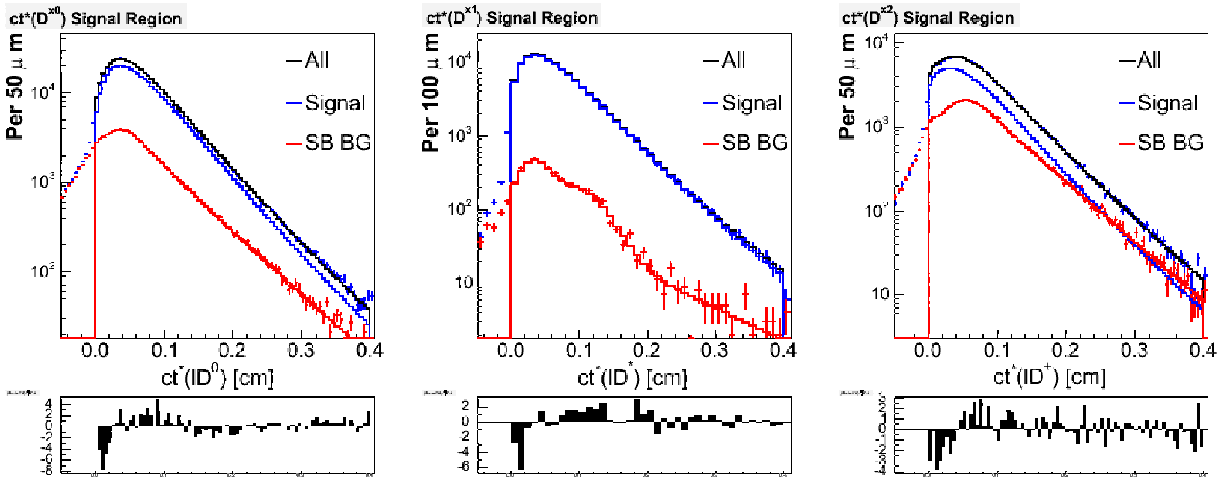


Figure 39: Likelihood fit projections of ct^* for the full Monte Carlo sample with data modeled background. The residuals are in the lower plots.

The likelihood projection fit to the Monte Carlo is shown in Fig. 39. The resulting fit values for the decay lengths are,

$$\begin{aligned} c\tau_{B^0} &= 465.5 \pm 1.4(stat.)\mu m \\ c\tau_{B^+} &= 500.3 \pm 1.4(stat.)\mu m \end{aligned}$$

Comparing these to the generated values, 460.5 μm and 501.0 μm for B^0 and B^+ respectively, shows disagreement with the B^0 lifetime. This disagreement is a 3.5 σ effect. However, it should be asked how the effect might change if background was added to the sample to make it more realistic? The idea being, perhaps the bias is caused by a sensitivity that is enhanced by the absence of background. After including background events which were randomly sampled from the combinatorial background templates derived from data, the resulting lifetimes are,

$$\begin{aligned} c\tau_{B^0} &= 461.5 \pm 1.6(stat.)\mu m \\ c\tau_{B^+} &= 502.6 \pm 1.6(stat.)\mu m \end{aligned}$$

This improves our confidence that the bias is not as large in data as is seen in the signal only Monte Carlo.

9 Systematic Uncertainties

A lifetime measurement in a triggered semileptonic mixture of B mesons contains many sources of uncertainties. The dominant systematic uncertainties on the lifetimes are from the correction of the efficiency function, the σ_{ct*} scale factor, and backgrounds which contain a real lepton and D meson. Uncertainties on the branching ratios are reflected in the shape of the K factor distribution as well as the fraction of B^0s in each sample. Even the detector alignment is relatively large which is not often the case for B lifetime measurements in a hadron collider. However, the size of the uncertainty is less important to the lifetimes ratio. For the ratio, it is the correlation coefficient which determines the weight of the systematic uncertainty. Table 11 lists the systematic uncertainties of the B^0 and B^+ lifetimes for each source along with the correlation coefficient in the fourth column. The correlation between the uncertainties of the two lifetimes allows a determination of the uncertainty on the ratio of lifetimes. These are given in column five. The systematic uncertainties are summed in quadrature for each lifetime and the ratio.

Error source	$\delta_{c\tau_{B0}}(\mu m)$	$\delta_{c\tau_{B+}}(\mu m)$	Correlation	δ_R
Alignment	2.0	2.0	1	0.0003
L_{xy} resolution function	0.3	0.3	-1	0.0014
Signal fractions	0.8	0.3	-0.39	0.0021
$\epsilon(ct\|\sigma)$ shape	0.5	0.3	0	0.0013
σ_{ct*} Scale factor	2.6	2.2	1	0.0013
$b\bar{b}$ BG fraction	0.1	0.3	0.5	0.0005
ℓD^+ WS$\rightarrow$ RS Bg. model	1.6	0.4	-0.66	0.0032
Fake lepton asymmetry	1.8	1.8	1	0.0003
τ decay BG fraction	1.5	2.2	1	0.0013
$B \rightarrow D1D2X$ BG frac.	1.7	1.8	1	0.0000
f_{B^0} per sample	0.9	0.3	-0.5	0.0025
$\epsilon(L_{xy})$ correction	4.5	4.6	1	0.0004
K factor distributions	1.0	3.3	1	0.0049
$M(D)$ Sideband selection	1.2	1.2	1	0.0002
$M(D)$ Sideband templates	0.3	0.2	1	0.0003
Totals	6.8	7.4		0.0072

Table 11: Table of systematic uncertainties on lifetimes and the ratio with lifetime correlation coefficients for each systematic source.

9.1 Signal Fraction

The signal fractions for each sample are derived from the D mass templates and take into account the fractions of backgrounds under the signal peak. These fractions remain

fixed in the likelihood fitter thus, any statistical uncertainty on the fraction is not folded into the error calculation of lifetimes.

A systematic uncertainty is determined for each sample by varying the signal fraction by $\pm 1\%$. This value is derived from the largest statistical error of the three samples (see Table 5). Given the adjusted value, the likelihood function is refit for the lifetimes in the data. Six sets of toy Monte Carlo samples are generated in which one of the three samples is moved +0.01 or -0.01. The average shifts in lifetimes from the nominal fit values are summed in quadrature for each lifetime to give the following uncertainties: ± 0.75 μm and ± 0.3 μm for B^0 and B^+ respectively.

9.2 σ_{ct*} Scale Factor

The σ_{ct*} scale factors have been calculated from fits to $B \to J\psi X$ data [12] and the largest uncertainties are about 5%. However, the resulting scale factors varied from 1.17 to 1.54, depending on the sample, so a conservative uncertainty of $\pm 15\%$ is used here to cover the possible range.

Toy Monte Carlo samples were generated with $S_f \times 1.15$, but the lifetimes were fit using the standard scale factor. This was repeated for 100 independent samples and the same was done for $S_f \times 0.85$. In the first case, the average shifts in the lifetimes are +2.4 μm and +2.2 μm for B^0 and B^+ respectively. In the second case, the scale factor is overestimated and the lifetime shifts for B^0 and B^+ are -2.6 μm and -1.9 μm respectively. The larger of the shifts is taken as the systematic.

9.3 L_{xy} Resolution Model

The triple Gaussian L_{xy} resolution model is derived from a best fit to the realistic Monte Carlo data. Two alternative models were tried: A double Gaussian model could not fit the "tails" of the non-Gaussian resolution, and two exponential convoluted with a Gaussian functions failed to fit the central region of the pulls. Each bin in σ_{ct*} has its own set of weights and parameters for each Gauss function, each with its own set of parameter uncertainties from the fits to the pulls. The fits are performed with high statistics in the samples, much higher than in the bins of the data, so when the uncertainties are applied to the parameters within the lifetimes fitter, their is no shift (< 0.1 μm) in lifetimes. In fact, a fit to the lifetimes using only a single Gaussian resolution model with only the σ_{ct*} scale factor gives lifetime shifts of -0.3 μm and +0.3 μm for B^0 and B^+ respectively. Clearly the resolution model has little effect. The shifts seen with the single Gauss function are used as the systematic.

9.4 Efficiency functions

The decay length acceptance efficiencies are parametric equations that best fit the MC samples. As such, the parameters each have their own statistical uncertainties.

Since these parameters are not allowed to float during the lifetimes fit, a systematic uncertainty must be derived instead.

There are thirty-six functions each having eight parameters. Varying the parameters by their uncertainty one by one and fitting with a new efficiency against data or toy samples is reasonable. However, for some changed parameter values the signal likelihood normalization behaves in a bad way when machine precision limits are met.

There is a method which avoids this issue. In short, the change in efficiency due to a change in parameter can be calculated for all ct which gives a new efficiency shape but not a new parameter (see the appendix 11.3 for a more detailed description). This new efficiency shape is applied in toy Monte Carlo and the fit uses the original parameters for the efficiency functions.

Then these new toy samples are refit using the efficiency functions with the original parameters as used in the data. Lifetime fits are performed on 500 toy samples of which 250 were made from efficiencies where the parameter uncertainties are positive and 250 toy samples in which the efficiencies were modified by the negative uncertainties. The mean lifetime shifts within the two sets of samples are used as the systematics.

9.5 Background Systematics

All background models have two main sources of uncertainty. Each background is modeled by a template that results from a best fit to data or Monte Carlo samples so there are statistical uncertainties to the template parameters. Treating these uncertainies is straight forward. Simply modify the parameters by their uncertainties and compare the resulting refit lifetimes with the unmodified fit. The other source - the assumption that our samples correctly predict the background fraction or shape under the signal peak is more dependant on particular issues and is considered separately for each background model.

9.5.1 Mass Sideband Selection

The D^0 and D^+ sidebands are chosen sufficiently away from the signal region, but not too far as to include significant reflections from signal. Of course, there is some arbitrariness to the width of the sidebands so it is necessary to calculate what effect this may have on the lifetime values. There is a quite conservative method to calculate the systematic which entails using the combinatorial background models from only one of the sidebands, fit the lifetimes and then do the same using the other sideband. The shifts in the lifetimes provide the systematics. However, this method assumes an unrealistic background though; why should the background events under the signal peak behave the same as under only one sideband as opposed to an average of the two?

Since the data samples are quite large for the D^0 and D^+, we can simply split the sidebands in half and refit the distibutions to get new templates. The lower half of the sidebands are defined as the first set and the upper half the second. Refit the lifetimes with the updated templates in the likelihood, once for the first set and then again for

the second set. As before, the shifts in the lifetime fit results provide the systematic error - $\pm 1.2\ \mu m$ for both lifetimes.

This was not done for the D^* sample since their is insufficient background in the sideband regions to establish confident fits of the resulting distributions after splitting the sideband regions.

9.5.2 Other B Background Systematics

The expected fractions of physics backgrounds that are approximated by realistic Monte Carlo templates are listed in Table 8. The statistical uncertainties on these fractions are calculated from the proper scaling of the number of events in each sample and comparing numbers generated in both the background mode and our signal realistic Monte Carlo. Since these samples represent more events than our data, a simple transfer of the uncertainties to the data is not enough. A simplified process is used to generate more conservative systematics: double the uncertainties from the RMC and apply modified fractions accordingly in the lifetimes fit. The shift in lifetimes supplies the systematic error.

9.5.3 Using the Wrong Sign candidates

Early studies of the systematic uncertainties in using the wrong sign candidates produced relatively large uncertainties on the lifetimes and these were the largest contributors to the lifetimes ratio systematic uncertainty. More sophisticated studies were required to better understand these uncertainties and, more importantly for the ratio, their correlations.

The fake lepton component within the wrong sign candidates is expected to be the same in the right sign signal region due to track charge symmetry. No preference for positive or negative tracks is seen in CDF data, nevertheless, a systematic uncertainty is considered. Varying the fractions of fakes by $\pm 10\%$ is a reasonable amount to calculate the uncertainty of the charge symmetry assumption. The data were fit with modified fractions which resulted in lifetime shifts of $\pm 1.8\ \mu m$ for each lifetime. Since the prompt decays add a short lived component to long lived B decays and there are approximately the same number within the B^0 dominant samples as there are in the sample dominant in B^+, any change in the fraction of fakes moves the two lifetimes in the same direction. This is seen in the lifetime shifts which implies correlation.

Expectations of the $b\bar{b}$ component on the other hand are derived from branching ratios with uncertainties so more careful consideration is necessary. Uncertainties on the branching ratios which govern the $b\bar{b}$ decays, including sequential decays, are summed in quadrature for each sample to give 12-17% uncertainties. The $b\bar{b}$ background fractions in each ℓD sample are conservatively modified by $\pm 20\%$ and sets of 200 toy samples are generated for each change in fraction. The likelihood fit function then uses the nominal value for each fraction. Fits to the lifetimes produce an average shift in the B^+ lifetime of 0.25 μm and less than 0.05 μm shift in the B^0 lifetime. The shifts

in both lifetimes move in the same direction for both cases. The calculated correlation coefficient turns is 0.5.

The ct^* distribution in the wrong sign D^+ sample appears to have little to no $b\bar{b}$ shape (see Fig. 22) and so additional tests are done. The $b\bar{b}$ fit function for the wrong sign sideband subtracted ct^* was constrained to match the ℓD^0 and the ℓD^*+ samples separately. The relative fractions of fakes to $b\bar{b}$ were allowed to float to gain a good fit. Using the new functions and fractions, the data was fit for the lifetimes. For the ℓD^0 function constraint, the lifetimes shifted -0.7 μm and +0.2 μm for B^0 and B^+ respectively. The constraint of the ℓD^*+ function the fit values for the lifetimes in the data moved by -1.6 μm and +0.4 μm for B^0 and B^+ respectively. These higher values are quoted as the uncertainty on the shape and content of the right sign D^+ background.

The most extreme case is also considered for the D^+ sample. What if there is little or no $b\bar{b}$ background in the wrong sign, instead, it is all prompt? If this is the case, then removing the 0.25 scaling factor when modeling these events in the wrong sign, will treat these events as fake decays. This is done for ℓD^+ in 200 Toy samples. Fits to the lifetimes in these samples produced lifetime shifts of -1.8 μm and +0.5 μm for B^0 and B^+ respectively which is entirely consistent with the previous results. There is some overlap of this test with the test of the $b\bar{b}$ fraction so it is not directly quoted as a systematic. However, the value in this case is that it is an additional test to include in the calculation of a correlation coefficient. Combining the results of this extreme case with those of the two constrained models results in a correlation coefficient of -0.66.

9.6 L_{xy} Efficiency Correction

The correction of the L_{xy} efficiency using the $J/\psi \to \mu\mu$ sample contains uncertainties on the parametric fit parameters. The shape of the correction is adjusted in the same way as the ct efficiency. The modified correction shape is used to generate toy Monte Carlo samples which are then fit with the nominal correction shape. See section 9.4 and 11.3 for the details.

9.7 K factor Systematics

The generated K factor distribution is mostly sculpted by the 4 GeV/c p_T trigger requirement for the lepton as determined by the unbiased Monte Carlo sample The triggered Monte Carlo sample is reweighted by sculpting the lepton p_T distribution to match the data, thus the K factor distribution in the Monte Carlo is properly weighted.

The uncertainty in the K factor distribution rises from the uncertainty of the branching ratios of the different decay modes. The average missing p_T of a decay like $B^- \to D^{**0}\ell\nu X \to D^{*+}X \to D^0$ may be quite different than the direct decay $B^- \to \ell D^0 \nu$. We can again reweigh the Monte Carlo by changing the relative fractions of each mode using the uncertainties given in the PDG. Simply reweighing event by event according to the decay mode will produce modified K factor distributions.

Then produce new toy MC samples with these modified K factors and then refit the lifetimes with the original distributions. The B^0 fractions in each sample are fixed to the generated values to avoid the systematic effect of a change in the fractions. The mean lifetime displacement from the generated value results in the systematics; ± 1.0 $\mu m(c\tau_0)$ and ± 3.3 $\mu m(c\tau_+)$.

9.8 Sample composition

The same branching ratio uncertainties that were applied in the K factor systematic calculation also affect the fraction of B^0's in a sub-sample. Any change in the fractions are determined while generating modified K factor distributions as explained in the previous section. These modified fractions are then supplied to and fixed in the likelihood fitter and any shifts of the lifetime fit values within the realistic Monte Carlo samples are defined as the systematic uncertainties.

9.9 Ratio Systematics

The measured ratio of the lifetimes suffers from the anticorrelated lifetimes (correlation coefficient of lifetimes is approximately -0.42) but is aided by many of the correlated systematics. Many of the systematic uncertainties approximately or completely cancel each other.

Any misalignment of the detector geometry has no bias towards one B meson over another and so any effect will be seen equally, thus the systematic uncertainties are correlated. Likewise, the L_{xy} efficiency correction is applied to all samples in the same way so any change in the correction function produces effects on the lifetimes in the same direction.

The toy studies of the signal fraction systematic resulted in a measured correlation coefficient of 0. While studying the systematic effect of the L_{xy} resolution model the lifetimes moved in equal but opposite directions from the nominal values implying anticorrelation. However, lifetime fit values for modified σ_{ct*} scale factors moved in the same direction with nearly the same amplitude. This implies strong correlation of the lifetimes when subjected to scale factor systematics.

The systematics summary table includes approximated or calculated lifetime correlation coefficients for each source where correlated lifetime effects are given the value 1, anticorrelated given value -1, and the value 0 is for uncorrelated effects.

10 Summary

An unbinned likelihood fit for B^0 and B^+ lifetimes in semileptonic decays has been performed and the results agree with other measurements [4]. The correlations due to the lepton plus secondary vertex trigger have been included in this measurement and biases to L_{xy} are effectively derived from the data and realistic Monte Carlo simulations. The fit results are,

$$\begin{aligned} c\tau_{B^0} &= 457.9 \pm 3.5(stat.) \pm 6.8(syst.)\mu m \\ \tau_{B^0} &= 1.527 \pm 0.012(stat.) \pm 0.023(syst.)ps \end{aligned}$$

$$\begin{aligned} c\tau_{B^+} &= 488.5 \pm 3.8(stat.) \pm 7.4(syst.)\mu m \\ \tau_{B^+} &= 1.629 \pm 0.013(stat.) \pm 0.025(syst.)ps \end{aligned}$$

The ratio of the lifetimes is,

$$\frac{\tau_{B^+}}{\tau_{B^0}} = 1.067 \pm 0.013(stat.) \pm 0.007(syst.).$$

Although the lifetime values' uncertainties are dominated by systematics, the ratio's uncertainty is dominated by statistics. This is due to the anticorrelated nature of the lifetimes in mixed samples and the correlated nature of many of the lifetime systematics.

The measured lifetimes ratio is in agreement with the theoretical prediction in Table 1. However, the uncertainty of the measurement is smaller (compare the theoretical uncertainty with this measurement's statistical and systematic uncertainties, summed in quadrature).

This measurement is also in agreement with the world average, 1.071±0.009 [4]. The best measurement among those used to calculate the world average was performed by the Bell collaboration in 2005, which uses a electron-positron collider. They reported a lifetimes ratio of 1.066±0.008(stat.)±0.008(syst.). The other Tevatron collider experiment, D0, reported the ratio 1.080±0.016(stat.)±0.014(syst.) also in 2005. These results are consistent with this thesis.

Any improvements to this analysis would include the addition of more data. Currently at CDF, there is over 4 fb^{-1} of integrated luminosity available. The statistical uncertainty on the lifetimes and the ratio will drastically improve with such a large sample.

This thesis shows that it is possible to perform precise measurements of lifetimes in CDF data which are biased in their lifetimes due to the triggers. Accurate measurements in large samples with high signal to background ratio are necessary for other time dependant B analysis at CDF. The B mixing measurements rely on accurate lifetimes as parameters within the fit. The first B_s mixing observation [24] could not have been made without the large semileptonic B samples (the decays provided over 90% of the events).

11 Appendix

11.1 Correlations introduced by SVT

The most difficult aspect of a lifetime measurement using data that is biased in the track's impact parameter, is due to the correlation introduced between K, ct^* and σ_{ct^*}. Considerable time and effort has been spent in understanding and accounting for the dependencies.

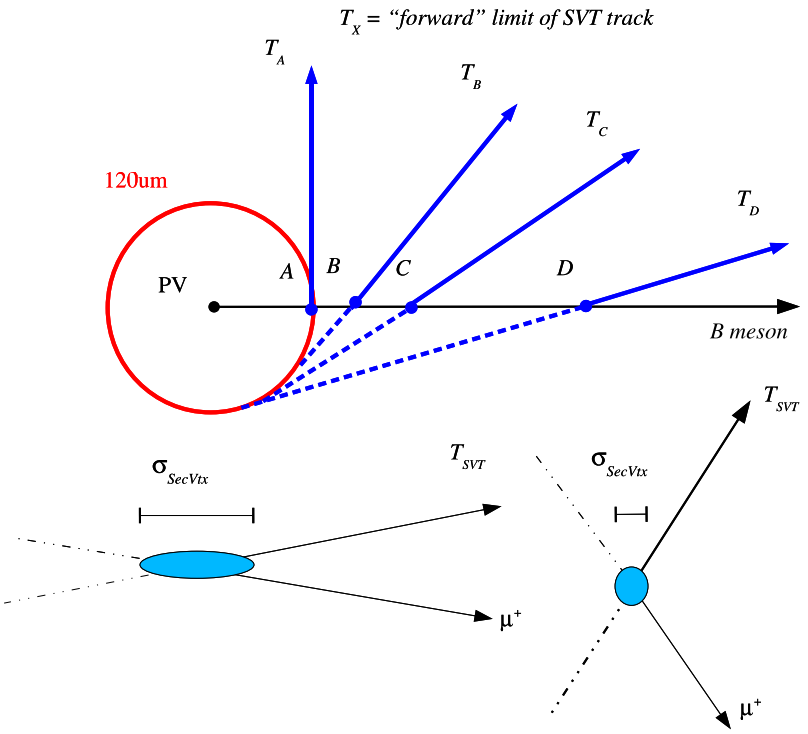


Figure 40: Diagram of ct vs σ_{ct}.

A simple model may help illucidate the principle at work here. Consider a decay in two dimensions in which one of the tracks used in the vertex reconstruction is selected by the trigger because its impact parameter is greater than $120\mu m$. The illustration (Fig. 40) shows four possible decay vertices (A-D) of the B meson. A displaced track, T, has a minimum allowable forward angle (to the direction of the B) in order to have an impact parameter greater than $120\mu m$. Therefore, for shorter decay lengths (vertex A and B), in other words, smaller L_{xy} (ignoring boost for now), there are relatively more events where the decay products produce near back-to-back tracks. For large opening angles between tracks there is better L_{xy}/vertex resolution in the xy plane.

This all translates to; the shorter the lifetime, the more precise the measurement. So, σ_{ct^*} depends on ct^*.

The decay length is also dependent on the missing energy. This arises from the transfer of a large proportion of the forward momentum carried by the B meson to the neutrino. In the case of an event with a shorter decay length, the trigger prefers those events with large opening angles which implies that much of the energy was carried away by the neutrino. Therefore, the spread in missing energy is wider for smaller decay lengths than it is for longer lived B's.

Of course, this model overemphasizes the correlation. In this analysis the displaced trigger track is a decay product of the charm meson and does not necessarily point back to the B decay vertex and so the impact parameter of the B candidate may be smaller than the cutoff for the triggered track, smearing the effect. Also, consider the boost factor which smears this effect. Figures 41 and 42 show the expected correlations in the realistic μ+SVT Monte Carlo sample.

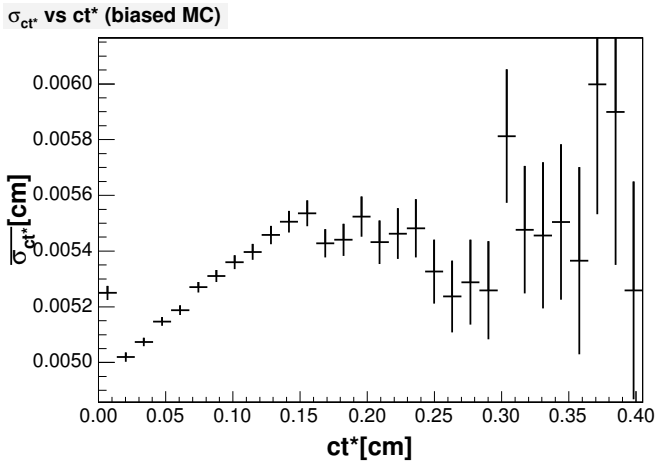


Figure 41: Profile of decay length with the average error in the l+SVT Monte Carlo ℓD^0 sample.

11.2 Differences in e and μ K factors

Since the K factor distributions from the muon Monte Carlo sample are used for both leptons, any differences of K between the leptons must be quantified and if necessary, accounted for in the samples. No difference is expected at meson decay, but electrons and muons behave differently in the detector. A simple comparison of K factors within separate Monte Carlo samples was performed to show any difference. The muon sample is about fourteen times larger than the electron sample but statistics are high enough to determine if there are significant differences. The comparisons (Fig.'s 43 and 44) shows very little difference and thus any difference is neglected in this analysis.

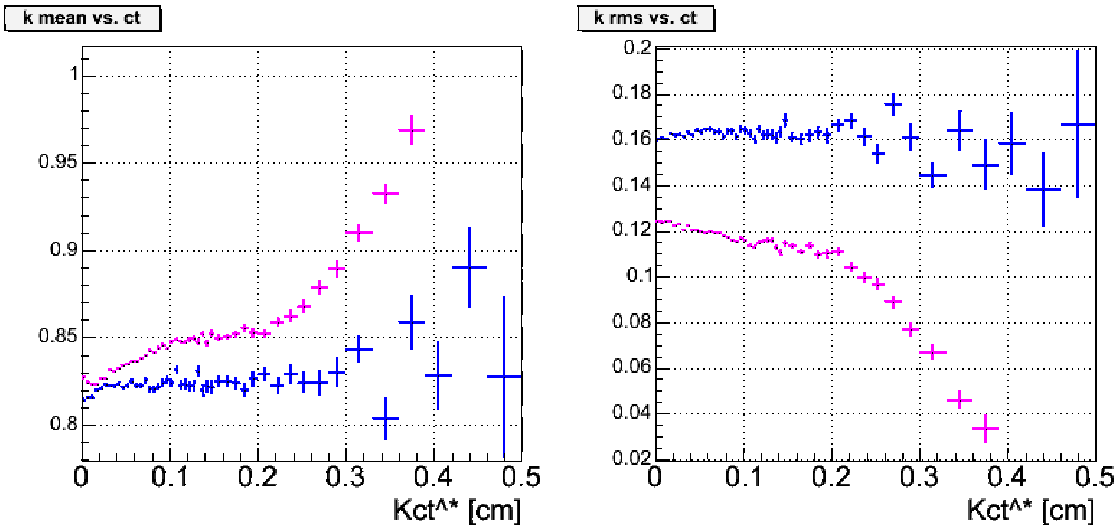


Figure 42: Profile of ct with the mean and RMS of K factors in the Monte Carlo $\ell D^0/D^*$ samples. The unbiased Monte Carlo sample (blue) and the l+SVT Monte Carlo (magenta) are both shown.

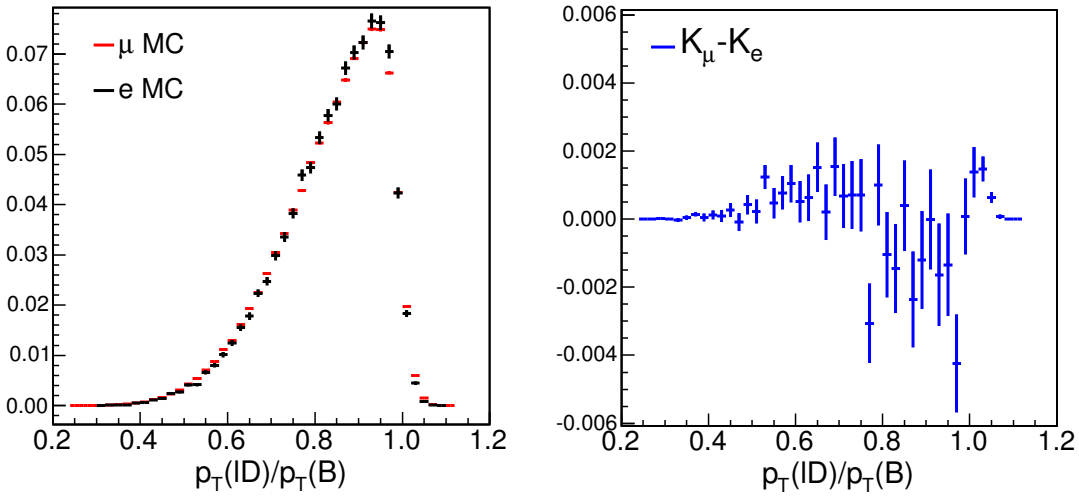


Figure 43: K factor from eD^0 (black) and μD^0 (red) Monte Carlo samples.

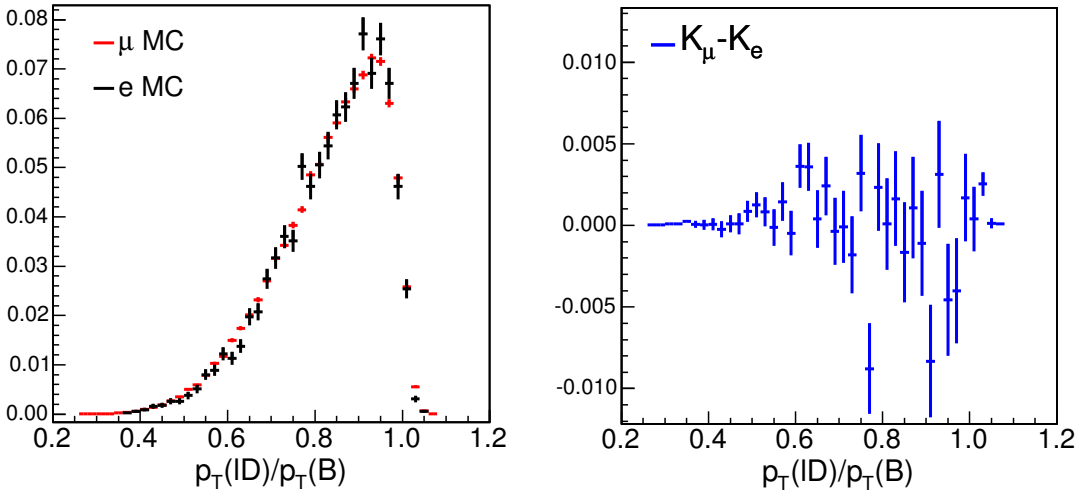


Figure 44: K factor from eD^+ (black) and μD^+ (red) Monte Carlo samples.

11.3 Calculating lifetime systematics from change in efficiency parameters

Let $f(\alpha_i)$ be the efficiency function, a function of N parameters α_i, and let V_{ij} be the error matrix such that the uncertainty on f is given by

$$\sigma_f^2 = \frac{\partial f}{\partial \alpha_i} V_{ij} \frac{\partial f}{\partial \alpha_j}$$

where there are implied sums over i and j (this convention will be used throughout this section unless otherwise stated).

Define σ_i to be $\sqrt{V_{ii}}$ (no sum). The partial derivatives are approximated by making a change $\Delta\alpha_i = \sigma_i$, giving

$$\begin{aligned} \frac{\partial f}{\partial \alpha_i} &\approx \frac{f(\alpha_i + \Delta\alpha_i) - f(\alpha_i)}{\Delta\alpha_i} \\ &= \frac{f(\alpha_i + \sigma_i) - f(\alpha_i)}{\sigma_i} \equiv \frac{\Delta f}{\sigma_i} \end{aligned}$$

Now, the uncertainty on f is approximately given by

$$\begin{aligned} \sigma_f^2 &\approx \frac{\Delta f_i}{\sigma_i} V_{ij} \frac{\Delta f_j}{\sigma_j} = \Delta f_i \frac{V_{ij}}{\sigma_i \sigma_j} \Delta f_j \\ &= \Delta f_i \rho_{ij} \Delta f_j \end{aligned} \tag{17}$$

where ρ_{ij} is the correlation matrix (as returned by Minuit, for example).

Thus, to get the uncertainty on f, calculate $\Delta f_i = f(\alpha_i + \sigma_i) - f(\alpha_i)$ for each i (keeping track of sign) and plug into equation 17.

This is applied to the actual efficiency function. First we must include the proper decay time, $x = Kct^*$, so $f(\alpha_i)$ becomes $f(x;\alpha_i)$. The uncertainty on the efficiency must be calculated for all x within the acceptance range as defined in the analysis cuts. Using a binned approach, the uncertainty, equation 17, is added to each value of the efficiency, $f_b(x_c, \alpha_i)$, where x_c is the center of each bin, b, and thus resulting in a new efficiency distribution. This is illustrated in Fig. 45. The resulting distributions are then used to produce toy Monte Carlo samples which are fit for the lifetimes using the standard lifetime likelihood function with the original efficiency functions.

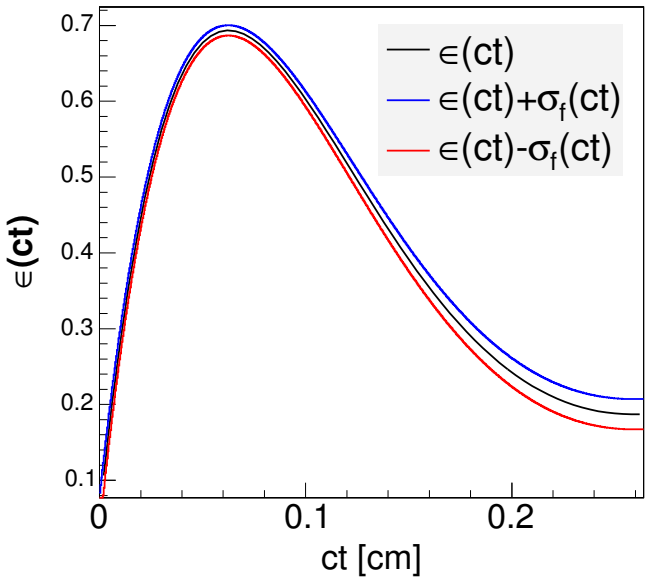


Figure 45: The modified efficieny distributions are shown with the original efficiency function in this example plot. The bin width is 0.0004 cm.

References

[1] M. Kobayashi, T. Maskawa, Prog. Theor. Phys. **49-2** 652, (1973). 5

[2] S. Herb, D. Hom, L. Lederman *et. al.*, Phys. Rev. Lett. **39-5** 252, (1977). 5

[3] T. Aaltonen *et. al.*, CDF-II Collaboration, Phys. Rev. D, **79, 092003** (2009). 5

[4] W.-M. Yao *et. al.*, Particle Data Group, Journal of Physics **G33**, (2006). 8, 68

[5] I. Bigi, M. Schifman, N. Uraltsev, Annu. Rev. Nucl. Part. Sci. **47** 591, (1997). 7, 8

[6] C. Tarantino, Eur. Phys. J. **C33**, S895, (2004). 9

[7] F. Gabbiani *et. al.*, Phys. Rev. **D68**, 114006, (2003). 9

[8] F. Gabbiani *et. al.*, Phys. Rev. **D70**, 094031, (2004). 9

[9] R. Blair *et. al.*, CDF-II Collaboration, FERMILAB-Pub-96/390-E, (1996). 11

[10] C. Hill, L00: Operational Experience and Performance of the CDFII Silicon Detector, Nucl. Instrum. Meth. **A530** 1-6 (2004). 13

[11] J. Marriner, Secondary Vertex Fit with Mass and Pointing Constraints (CTVMFT), CDF internal note 1996, (1993). 21

[12] K. Anikeev, G. Bauer, Ch. Paus, B Meson Lifetime Measurements Using Exclusively Reconstructed Decays B –¿ JPsi X, CDF internal note 6266, (2003). 25, 63

[13] K. Anikeev, Ch. Paus, New measurement of lifetimes of B mesons in the exclusively reconstructed decays, CDF internal note 7409, (2004). 25

[14] K.Anikeev, A.Belloni, G.Gomez-Ceballos, *et. al.*, Monte Carlo Samples for Delta ms Measurement, CDF note 8303 (2006) 27

[15] S. Grommel, Ch. Paus, A Heavy Quark Generator, CDF internal note 5985, (2002). 26

[16] R.Mumford, S.Behari, *et. al.*, Λ_b Lifetime Using fb^{-1} Data Taken With Two Displaced Track Triggers, CDF note 8578 (2009) 34

[17] D. Acosta *et. al.*, FERMILAB-Pub-040440-E, (2004). 15

[18] H. Wenzel, Tracking in the SVX, CDF internal note 1790, (1998). 19

[19] W.Bell *et. al.*, User Guide For EvtGen at CDF, CDF internal note 5618, (2002). 26

[20] J. P. Done, GEANT3 Users Guide for the CDF Tracking Upgrade, CDF internal note 2029, (1993).

[21] H.-C. Fang, *et. al.*, An offline filter module for lepton-SVT data, CDF note 6326 (2003) 20

[22] K. Gibson, *et. al.*, Measurement of B Hadron Relative Fragmentation Fractions, CDF note 7911 (2006) 21

[23] F. James, *Minuit - function minimization and error analysis - reference manual*, http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html (2000), URL 51

[24] A. Abulencia, *et. al.*, CDF-II Collaboration, Phys. Rev. Lett. **97** 062003 (2006) 68