Measurement of the Lifetime of the B_S^0 Meson and the Width Difference in the B_S^0-$\overline{B}_S^0$ System in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV

by

Kevin Alan Burkett

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in The University of Michigan
1997

Doctoral Committee:

Associate Professor Myron K. Campbell, Chair
Associate Professor Dante E. Amidei
Associate Professor August Evrard
Professor Peter G. Hinman
Professor Gordon L. Kane

“ *Midway on our life's journey, I found myself*
In dark woods, the right road lost. ”

Dante Alighieri, Inferno *I. 1-9* *[1]*

— ⋄ —

“ *Ah, but I was so much older then,*
I'm younger than that now. ”

Bob Dylan, My Back Pages *[2]*

— ⋄ —

“ *Anyone who keeps the ability to see beauty never grows old* ”

Franz Kafka

Acknowledgments

The accumulation of the data used in this thesis was a large collaborative effort. This dissertation can not adequately describe the hard and often thankless work of many members of CDF and the Fermilab Accelerator Division. Without them it would not have been possible, and I am very grateful to those involved.

I am grateful to my advisor, Myron Campbell, for giving me the opportunity to join CDF and work in High Energy Physics, and for supporting me in my work. I must thank Myron and the rest of the Michigan group for allowing me to pursue an analysis topic that lay somewhat outside the interests of the rest of the group.

For teaching me almost everything I knew about the trigger at CDF, I thank Bill Badgett. He was very patient in training me, and sympathetic about, as he said, "being caught in the maelstrom." I must thank Soo-Bong Kim for all that he taught me and for all the work on the trigger that we did together. Thanks also go to Gene Guillian for working so dilligently to take over for me, so that I could pursue my analysis.

I am deeply indebted to Manfred Paulini for helping me with this work. He taught me a great deal, and without his patient guidance this analysis would never have been possible. I would also like to thank Joe Kroll for his advice, and Fumi Ukegawa for giving this analysis a very careful reading along the way.

I would like to thank my family for all their support and encouragement over the years. Finally, I am especially grateful to Carrie. Without her patience and support I might never have finished.

Table of Contents

List of Tables

Table

List of Figures

Figure

List of Appendices

Appendix

Chapter 1

Introduction

Over the past century the search for an understanding of the fundamental constituents of matter and the forces that hold them together has seen many breakthroughs. In recent years a clear theory of the constituents and their interactions has emerged. This theory is known as the Standard Model.

1.1 The Standard Model

In the Standard Model interactions in matter can be described by four fundamental forces: the strong force, the weak force, electromagnetism, and gravity. Since gravity is much weaker it can be neglected while discussing the other forces. Each of these interactions can be described by a gauge theory. These are theories where the observables of the system are invariant under a phase translation. The strong force is described by an SU(3) gauge theory, known as Quantum Chromodynamics (QCD). The weak force is described by an SU(2) gauge theory, while electromagnetism is a U(1) gauge theory. The weak and electromagnetic forces were successfully unified by the work of Glashow [3], Weingberg [4], and Salam [5], to form the current electroweak theory. Combining all of these, the Standard Model is often referred to as an SU(3)×SU(2)×U(1) gauge theory.

Each of the forces is mediated by a gauge boson. The gauge bosons are particles of varying masses with integer spin that "carry" the forces. Eight massless gluons carry the strong force, while the weak force is carried by the massive W^+, W^-, and Z^0 bosons, and the electromagnetic force is mediated by the massless photon (γ).

The constituents of matter in the Standard Model are of two types: quarks and leptons. Both types are spin-$\frac{1}{2}$ fermions. There are six quarks which have been observed experimentally: up, down, charm, strange, top, and bottom. They are identified by the first letter of their names. There are also six leptons: the electron (e), muon (μ), and tau (τ), and their corresponding neutrinos, ν_e, ν_μ, and ν_τ. The quarks and leptons can be organized into three doublets each. The doublets of quarks and leptons are shown below.

$$\text{Quarks}: \begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \qquad \text{Leptons}: \begin{pmatrix} e^- \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}$$

The magnitude of the charge of the electron ($|e|$) is used as the basis for measuring the electric charges of other particles. The e, μ, and τ all have electric charge -1$|e|$, while their corresponding neutrinos are neutral. The quarks have fractional electric charges. The u, c, and t all have electric charge $+\frac{2}{3}|e|$, while the d, s, and b have electric charge $-\frac{1}{3}|e|$.

Quarks and leptons both experience the electroweak interaction, but only quarks experience the strong interaction. Each quark carries a color charge of the strong force, similar to the electric charge of electromagnetism. Quarks are bound together by the strong force to form hadrons. Hadrons made up of a quark anti-quark pair are called mesons, while hadrons made up of a triplet of quarks or anti-quarks are called baryons. As the leptons have no color charge, the strong force does not act on them.

1.2 Heavy Quark Physics

The study of heavy quark physics began with the discovery of the J/ψ simultaneously at Brookhaven [6] and SLAC [7] in 1974. The J/ψ is a $c\bar{c}$ bound state with a mass of 3.1 GeV/c^2. A year later, the discovery of the τ lepton [8][9] demonstrated the existence of a third doublet of leptons, suggesting the existence of a third quark doublet.

The existence of a third quark doublet was confirmed in 1977 by the discovery of the Υ resonances at Fermilab[10]. These are a series of narrow resonances with masses around

10 GeV/c^2, consisting of a $b\bar{b}$ bound state. The first three upsilon resonances, known as Υ, Υ', and Υ'', or $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$, were all found first at Fermilab [11] and later confirmed elsewhere. The $\Upsilon(4S)$ was discoverered at Cornell in 1980 [12] [13]. The $\Upsilon(4S)$ has a mass of 10.58 GeV/c^2 and is a broader resonance than the lower Υ resonances. The importance of the $\Upsilon(4S)$ is that it lies above the threshold for decays into pairs of B mesons.

1.3 B Mesons

The B mesons are $\bar{b}q$ pairs, where the b anti-quark is paired with a lighter quark q. The lightest B mesons are the B^+, a $\bar{b}u$ pair, and the B^0, a $\bar{b}d$ pair. The B^+ and the B^0 both have masses of 5.28 GeV/c^2. The B_S^0 was discovered in 1993 nearly simultaneously by the ALEPH experiment at CERN [14] and by the CDF collaboration at Fermilab [15]. It is a $\bar{b}s$ pair, with a mass of 5.37 GeV/c^2. The existence of the B_C^+ meson, a $\bar{b}c$ pair, has been predicted, but it has yet to be confirmed experimentally.

1.4 B Production at Hadron Colliders

The study of B mesons has traditionally been the domain of e^+e^- colliders which exploit the $\Upsilon(4S)$ resonance, such as ARGUS and CLEO. The B production cross section at such colliders is $\sigma \sim 1\ nb$. In such an environment the B cross section is a significant fraction of the total cross section, so there is little background to obscure the events of interest.

More recently, hadron colliders have proven successful in many B physics measurements. One of the advantages of a hadron collider compared to the e^+e^- colliders operating at the $\Upsilon(4S)$ is that all B species are produced in the hadron colliders. The energy at the $\Upsilon(4S)$ is only sufficient to produce either $B^0\overline{B^0}$ or B^+B^- pairs, while B_S^0 mesons and Λ_b baryons can be produced at the higher energies of a hadron collider. In addition, the B's are produced at rest at the $\Upsilon(4S)$ while they are produced with momentum at a hadron collider. This fact is important for performing measurements which are time-dependent.

Another advantage of a hadron collider is the enormous production cross section. At the Tevatron, with a center-of-mass energy of 1.8 TeV, the b quark production cross section is $\sigma_b \sim 50$ μb. However, the disadvantage is that the total inelastic cross section is roughly three orders of magnitude larger, meaning that the background is much higher. In addition, the b production cross section is a steeply falling cross section. Figure 1.1 shows the b cross section measured at CDF as a function of the transverse momentum (p_T) of the b quark. This cross section drops nearly two orders of magnitude from the cross section for $p_T(b) >$ 7.5 $GeV/$c to the cross section for $p_T(b) > 20$ $GeV/$c. Nevertheless the cross section is still large enough to produce a large sample of B mesons for study.

1.5 B Decays

The decays of B mesons involve the b quark decaying to a lighter quark via the weak interaction. This decay proceeds by the b emitting a virtual W boson and decaying to a lighter quark q. The W boson then decays to a pair of either leptons or quarks.

In semi-leptonic B decays, the virtual W decays to a lepton-neutrino pair. Such decays are well-described by the Spectator Model. In the Spectator Model, the decay of the B meson is governed entirely by the decay of the b quark. The light quark q of the $\overline{b}q$ pair acts as a spectator to the decay of the b quark. Feynman diagrams of the spectator decays $B^0 \to D^- \ell^+ \nu$ and $B^0_S \to D^-_S \ell^+ \nu$ are shown in Figure 1.2. As the diagrams differ only in the spectator quark, comparisons of the B^0 and B^0_S lifetimes provide a test of the validity of the Spectator Model.

The Lagrangian for weak decays of quarks has the form

$$\mathcal{L} \sim (\overline{u_L},\ \overline{c_L},\ \overline{t_L})\ \gamma^\mu\ V_{CKM} \begin{pmatrix} \overline{d_L} \\ \overline{s_L} \\ \overline{b_L} \end{pmatrix} W^\mu$$

where W^μ is the propagator of the W boson and V_{CKM} is the Cabibbo-Kobayashi-Maskawa

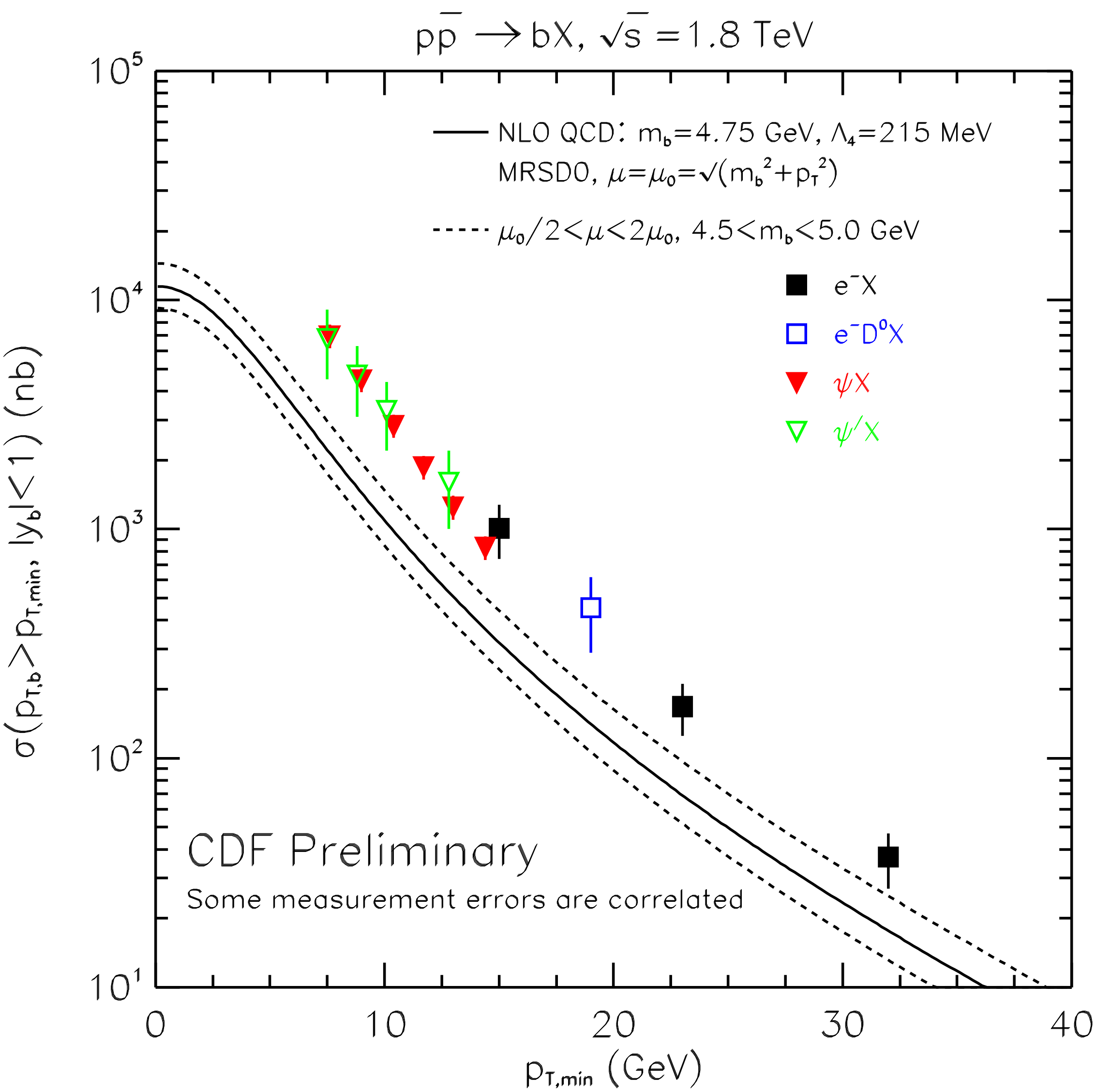


Figure 1.1: b quark production cross section as a function of the b quark p_T. The points show the cross section measured at CDF in several decay modes. The curve shows a next-to-leading order QCD calculation.

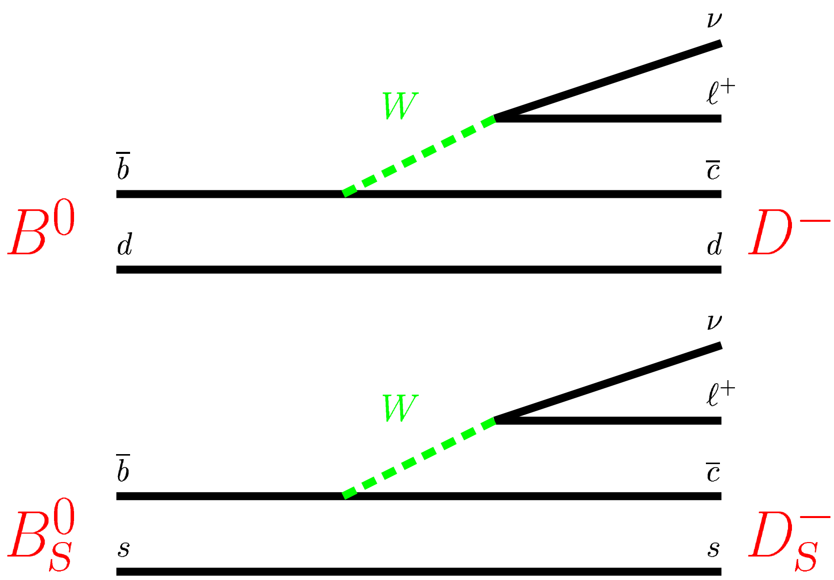


Figure 1.2: Feynman diagrams of the semi-leptonic decays of B^0 (top) and B_S^0 (bottom).

(CKM) [16][17] matrix, which can be written as

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}.$$

The CKM matrix is a unitary matrix, and the coupling strength between two different flavor quarks is given by the corresponding CKM matrix element.

The phases of each of the quark wave functions are arbitrary. Using this fact and applying the constraints of unitarity reduces the CKM matrix to 4 independent parameters. Thus the CKM matrix can be re-written in the Wolfenstein parameterization [18] as

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

The parameters of the CKM matrix must be determined experimentally, providing an important motivation for the study of B decays.

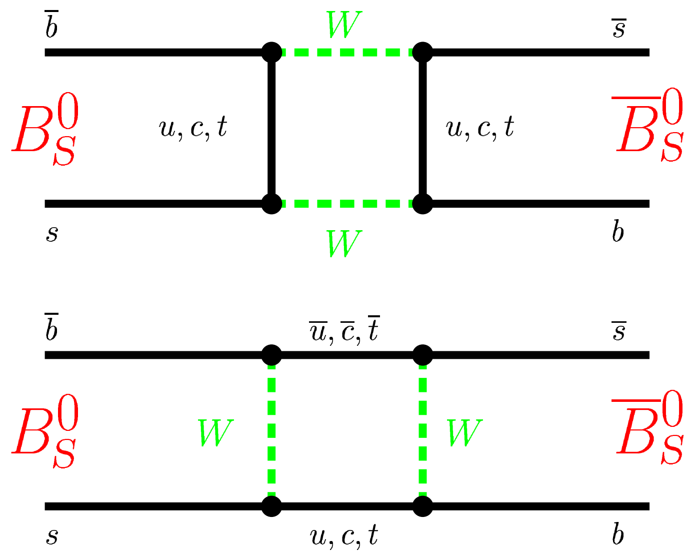


Figure 1.3: Feynman diagrams showing $B_S^0\overline{B}_S^0$ mixing.

1.6 B Mixing

One of the most interesting phenomena involving B mesons is $B\overline{B}$ mixing. This refers to both $B^0\overline{B^0}$ and $B_S^0\overline{B_S^0}$ mixing. These processes can occur since quark flavor is not conserved in the weak interaction. The Feynman diagrams which describe $B_S^0\overline{B_S^0}$ mixing are the second-order box diagrams shown in Figure 1.3. The diagrams governing $B^0\overline{B^0}$ mixing are the same, except for the substitution of a $d(\overline{d})$ quark in place of the $s(\overline{s})$ quark, though the mixing amplitude is reduced by the CKM factors at the vertices.

Simple quantum mechanics can be used to understand the mixing. The strong(flavor) eigenstates of the system are $|B_S^0>$ and $|\overline{B}_S^0>$. The eigenstates of the weak interaction are a superposition of these two states. The weak eigenstates can be written as

$$|B_S^L> = \frac{1}{\sqrt{2}}(|B_S^0> + |\overline{B}_S^0>)$$
$$|B_S^H> = \frac{1}{\sqrt{2}}(|B_S^0> - |\overline{B}_S^0>)$$

where the L and H refer to "light" and "heavy", since these are also the eigenstates of the mass matrix. This is analogous to the neutral K system where K^0_L and K^0_S refer to the "K Long" and "K Short", emphasizing the large difference in decay rates rather than the mass difference. The B^L_S and B^H_S have masses and decay widths which are denoted by $m_{L,H}$ and $\Gamma_{L,H}$ respectively, leading to the definitions

$$\Delta m = m_H - m_L$$

$$\Delta\Gamma = \Gamma_H - \Gamma_L$$

$$\Gamma = \frac{1}{2}(\Gamma_H + \Gamma_L).$$

For a state of pure $|B^0_S>$ at time $t = 0$, the state can be expressed as

$$|B^0_S >= \frac{1}{\sqrt{2}}(|B^H_S > + |B^L_S >).$$

The states $|B^L_S >$ and $|B^H_S >$ evolve over time according to the equation

$$|B^H_S(t) >= |B^H_S(t=0) > e^{-iHt},$$

where

$$H = m - \frac{i}{2}\Gamma.$$

Thus the evolution of $|B^0_S >$ with time can be expressed as

$$\begin{aligned} |B^0_S(t)> &= \frac{1}{\sqrt{2}}(e^{-im_H t - \frac{1}{2}\Gamma_H t}|B^H_S > + e^{im_L t - \frac{1}{2}\Gamma_L t}|B^L_S >) \\ &= \frac{1}{2}e^{-i(m+\frac{1}{2}\Delta m)t}e^{-\frac{1}{2}(\Gamma - \frac{1}{2}\Delta\Gamma)t} \times \\ &\quad \{(e^{-\frac{1}{2}\Delta\Gamma t} + e^{i\Delta m t})|B^0_S > + (e^{-\frac{1}{2}\Delta\Gamma t} - e^{i\Delta m t})|\overline{B}^0_S >\}. \end{aligned}$$

Starting with an initially pure B^0_S state, the probability of finding a B^0_S at a later time t is

$$P(t) = \frac{1}{2}\Gamma e^{-\Gamma t}(\cosh(\frac{\Delta\Gamma}{2}t) + \cos(\Delta m t)).$$

From this it can be seen that the frequency of oscillations will be Δm. It is convenient to normalize to the total width, so the mixing parameters x_d and x_S are defined as

$$x_d = \frac{\Delta m_d}{\Gamma_d} \qquad \text{and} \qquad x_S = \frac{\Delta m_S}{\Gamma_S}.$$

The ratio of the mixing parameters is given by the equation

$$\frac{x_d}{x_S} = \frac{\tau_{B^0}}{\tau_{B_S^0}} \frac{\eta_{B^0}}{\eta_{B_S^0}} \frac{f_{B^0}^2 B_{B^0}}{f_{B_S^0}^2 B_{B_S^0}} \frac{|V_{td}|^2}{|V_{ts}|^2} \text{ [19][20]},$$

where τ_{B^0} and $\tau_{B_S^0}$ are the B^0 and B_S^0 lifetimes, f_B and B_B are related to the B meson wave function, and η_B contains the QCD corrections which are of order 1. It can be seen from this expression that a measurement of $B_S^0\overline{B}_S^0$ mixing implies a measurement of the ratio of CKM matrix elements $|V_{td}|/|V_{ts}|$.

While $B_S^0\overline{B}_S^0$ mixing has never been observed, it is expected that Δm_S is much larger than Δm_d. The current world average value for Δm_d is $\Delta m_d = 0.472 \pm 0.018\ ps^{-1}$ [21], while the best limit on Δm_S is $\Delta m_S > 10.4\ ps^{-1}$ [22].

In the CKM model, the ratio $\Delta m/\Delta\Gamma$ is a constant. It contains no ratios of CKM matrix elements and depends only on QCD corrections [23]. This ratio can be expressed as

$$\frac{\Delta m}{\Delta\Gamma} = \frac{-2}{3\pi} \frac{m_t^2 h(m_t^2/M_W^2)}{m_b^2} \left(1 - \frac{8}{3}\frac{m_c^2}{m_b^2}\right)^{-1}$$

with

$$h(y) = 1 - \frac{3y(1+y)}{4(1-y)^2}\left\{1 + \frac{2y}{1-y^2}\ln y\right\}.$$

If x_S is large, implying a large value of Δm and a rapid frequency of oscillations, then $\Delta\Gamma$ will also be large. For large values of x_S where oscillations may be too rapid to resolve, a measurement of $\Delta\Gamma$ provides a way to find evidence of $B_S^0\overline{B}_S^0$ mixing, and an indirect measurement of x_S. It would also provide a measurement of the ratio $|V_{td}|/|V_{ts}|$, as explained earlier. Current theory predicts $\Delta\Gamma/\Gamma < 30\%$ [23].

The decay $B_S^0 \to D_S^- \ell^+ \nu$ is expected to be an equal mixture of B_S^L and B_S^H. Thus one method for measuring $\Delta\Gamma(B_S^0)$ is to describe the proper lifetime distributions from these decays by a function of the form

$$(\Gamma + \frac{\Delta\Gamma}{2})e^{-(\Gamma+\frac{\Delta\Gamma}{2})t} + (\Gamma - \frac{\Delta\Gamma}{2})e^{-(\Gamma-\frac{\Delta\Gamma}{2})t},$$

where Γ is the value obtained when the distribution is described by a single exponential [24][25]. Since the decay mode $B_S^0 \to J/\psi\phi$ is expected to be dominated by one CP eigenstate of the B_S^0, $\Delta\Gamma(B_S^0)$ could also be measured by taking the difference between the average lifetime in $B_S^0 \to D_S^- \ell^+ \nu$ events and the lifetime from $B_S^0 \to J/\psi\phi$ events [26][27].

1.7 Overview of Analysis

This analysis measures the lifetime of the B_S^0 meson. Decays involving a lepton and a D_S^- are used to reconstruct the B_S^0. (Throughout this thesis a reference to a particular charge state also implies its charge conjugate.) A sample of events with high p_T leptons is selected, and in those events D_S^- candidates are reconstructed through their decays to 3 other hadrons. Precision measurements of the paths of the hadrons near the collision point are provided by a silicon microstrip detector. The mean proper decay length of the b hadron is 465 μm [52] and the mean proper decay length of the D_S^- is 140 μm [52], providing sufficient separation to reconstruct the B_S^0 decay vertex and the D_S^- decay vertex.

The D_S^- vertex, or tertiary vertex, is found by intersecting the tracks of the 3 hadrons. Once the tertiary vertex has been found, the path of the D_S^- is traced back to find its intersection with the lepton in the event. This is the B_S^0 vertex, or secondary vertex. Using the distance travelled by the B_S^0 from the primary vertex to its decay vertex, and knowing the approximate momentum of the B_S^0, the proper decay length of the B_S^0 can be found.

With only events coming from decays of a B_S^0, and with perfect resolution, the proper decay length distribution would be described by an exponential. However, this exponential is smeared by the gaussian resolution of the detector. In addition, other products of the B_S^0 decay may be missed, further altering the exponential distribution. As well, there are background events which mimic the signature of the signal. The contributions of several sources of background must be calculated and included. Separate samples of signal and representative background events are selected and a simultaneous fit to the proper decay length distributions is performed to extract the B_S^0 lifetime. Sources of systematic uncertainty in the measurement of the B_S^0 lifetime must be considered, and the resulting uncertainty evaluated.

Once the B_S^0 lifetime has been extracted, the complementary search for a lifetime difference between B_S^L and B_S^H can be performed by changing the exponential form of the function used to describe the signal to a sum of two exponentials. The potential for measuring two different lifetimes is evaluated in Monte Carlo based on sample size and the lifetime difference between the two components of the signal. The proper decay length distributions of the data are then fit again to extract the lifetime difference.

Chapter 2

Experimental Apparatus

There are two distinct components used in the collection of the data for this analysis. The first is the Tevatron, a synchrotron accelerator at the Fermi National Accelerator Larab-oratory (Fermilab) that collides beams of protons (p) and anti-protons ($\overline{p}$) at a center-of-mass energy of 1.8 TeV. The second is the Collider Detector at Fermilab (CDF), a general-purpose detector designed to study the results of the $p\overline{p}$ collisions.

In the discussion that follows, first the Tevatron and the associated components used in generating and accelerating the colliding beams are described. The CDF detector is then described, with an emphasis on those parts of the detector that are important to this analysis.

The data used in this analysis was taken at Fermilab during two periods. The first period lasted from June 1992 until May 1993, with approximately 20 pb^{-1} of data accumulated. This period is referred to as Run Ia. After a brief shutdown, data-taking resumed in September 1993 and continued until February 1996. This period is referred to as Run Ib and produced 90 pb^{-1} of data.

2.1 The Tevatron

Before achieving their final center of mass energy of 1.8 TeV, the proton and anti-proton beams are accelerated in several stages. A schematic diagram of the Tevatron [28] [29] and associated components is shown in Figure 2.1.

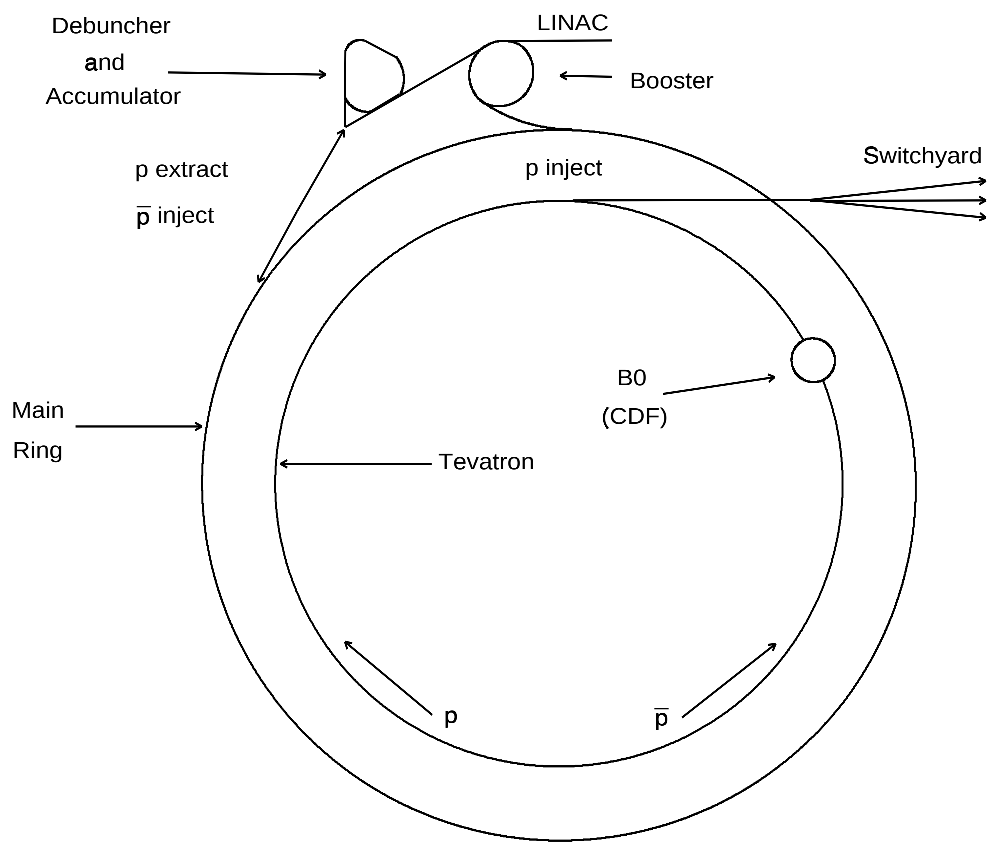


Figure 2.1: Schematic diagram of the Tevatron accelerator at Fermilab and its associated components.

The protons begin as negatively charged hydrogen ions, composed of two electrons and one proton. They are accelerated to 750 keV by a Cockroft-Walton accelerator (not shown in Figure 2.1), and then enter a 150 m linear accelerator (LINAC) where they are accelerated to an energy of 200 MeV. The electrons are stripped from the H^- ions by passing them through a carbon foil and the beam is passed to the Booster, a 500 m circumference synchrotron, which accelerates the protons to 8 GeV. The protons then enter a 6.3 km synchrotron called the Main Ring. Here the protons are focused into bunches and accelerated to an energy of 150 GeV, before being injected into the Tevatron for acceleration to a final energy of 900 GeV. Though they are shown separately in Figure 2.1, the Main Ring and the Tevatron lie in the same 1 km radius tunnel.

At the same time that the proton beam is accelerated to 150 GeV by the Main Ring, a 120 GeV beam of protons is extracted and strikes a copper target, creating antiprotons and other secondary particles. The antiprotons which emerge from the proton-nucleus collisions have a large angular divergence, and a focusing magnetic field is used to collect them. The antiprotons are then passed to the Debuncher. The antiprotons have a longitudinal momentum spread and have the same bunched spatial structure as the protons did which struck the copper target. The Debuncher performs a phase space rotation which increases the spatial spread while reducing the momentum spread. Out of the Debuncher the antiproton beam is passed to the Accumulator, where pulses of antiprotons are received and stored until the intensity is high enough for collisions.

To bring the beams into collision, the antiprotons are transferred to the Main Ring and boosted to an energy of 150 GeV and then passed to the Tevatron. The protons and antiprotons circulate in opposite directions due to their opposite charges. Though both beams circulate in the same beam pipe, they are kept in separate helical orbits by electrostatic separators. In the final step the bunches are squeezed by quadrupole magnets to increase their densities and thus the probability for collisions.

The $p\overline{p}$ collisions occur in two interaction regions at the Tevatron. The first is at B0, where the CDF detector lies. The other is at D0, approximately one-third of the way around the ring from CDF, though not shown in Figure 2.1.

2.2 The CDF Detector

The CDF detector is a general-purpose detector designed to study $p\overline{p}$ collisions [30]. It is cylindrically symmetric due to the inherently cylindrical nature of the colliding beam environment. Figure 2.2 shows a wide view of the entire detector.

There is a central barrel region, with plug regions closing the barrel at both ends, and two forward regions. Figure 2.3 shows one quadrant of the detector. The coordinate

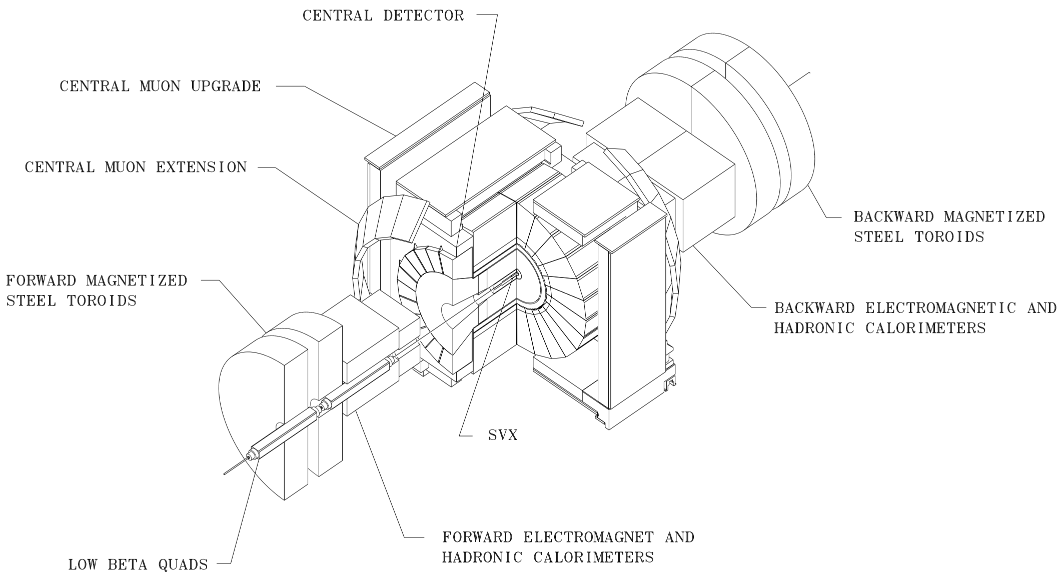


Figure 2.2: Isometric view of the CDF detector.

system used at CDF is shown here. The origin is defined at the center of the detector, at the midpoint of the interaction region, in the lower right-hand corner of Figure 2.3. The positive z-axis is defined to point along the beamline in the direction of the proton beam. The azimuthal angle (ϕ) of the cylindrical coordinate system is defined with respect to the horizontal, while the polar angle (θ) is defined relative to the proton direction. Further, the pseudorapidity (η) is defined by the relation:

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

2.3 Tracking Detectors

The system for tracking charged particles at CDF consists of three components which all lie inside a 1.4 Tesla solenoidal magnet of length 4.8 m and radius 1.5 m. Closest to the beampipe is the silicon vertex detector (SVX and its replacement, SVX$'$) (see Figure 2.3), designed to track particles very near the interaction point and to separate sequential decay vertices of long-lived particles. As shown in Figure 2.3, outside the SVX is a vertex drift chamber (VTX) which reconstructs tracks in the $r-z$ plane. The VTX is designed to measure

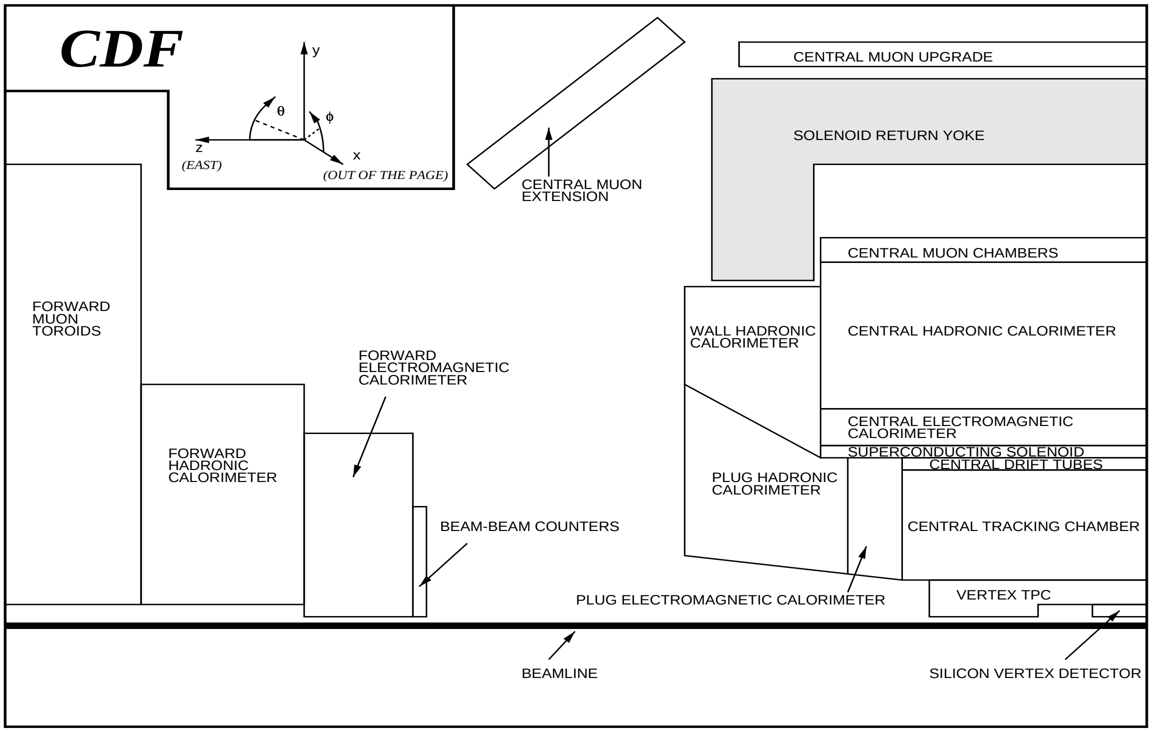


Figure 2.3: Side view of a cross section of the CDF detector. The interaction region is in the lower right corner of the figure.

the z position of the interaction point with a resolution of 1 mm. The SVX and VTX both lie inside the Central Tracking Chamber (CTC), a cylindrical drift chamber extending to a radius of 132 cm. The physical properites of all of the tracking subsystems are listed in Table 2.1.

2.3.1 SVX

The SVX [32] and SVX′ [33] are silicon microstrip detectors designed to provide precision tracking information close to the interaction point. The SVX detector was first used during Run Ia. Due to radiation damage incurred during running, it was replaced in Run Ib by a radiation-hardened version, SVX′.

Both detectors consist of two cylindrical barrels positioned end-to-end with a 2.15

Tracking System	Polar Angle Coverage	Radial Coverage(cm)	Length (cm)	Layers	Spatial Resolution(μm)
SVX	$\lvert\eta\rvert < 1.2$	$3.0 < r < 7.9$	30	4	15
SVX$'$	$\lvert\eta\rvert < 1.2$	$2.9 < r < 7.9$	30	4	13
VTX	$\lvert\eta\rvert < 3.25$	$7 < r < 21$	143.5	24	200-500
CTC	$\lvert\eta\rvert < 1.5$	$30.9 < r < 132$	160.7	60,24	200

Table 2.1: Physical properties of the components of the tracking system.

cm gap between the barrels at z=0. A diagram of one barrel is shown in Figure 2.4. Each barrel consists of four concentric layers, labelled 0 to 3, which extend to 7.9 cm from the beamline, with each layer divided into twelve 30° wedges. In the SVX the innermost layer, layer 0, was positioned at a radius of 3.0 cm, but for SVX$'$ it was moved in to a radius of 2.86 cm. The pitch of the strips is 60 μm on the three inner layers and 55 μm on the outer layer. As each barrel extends to $\pm$ 25 cm from the interaction point, while the luminous region extends approximately 30 cm along the z-axis, the geometric acceptance of the SVX is about 60%.

The basic element of the SVX is the ladder. Each 30° wedge contains four 25.5 cm ladders, one for each layer. A ladder is composed of three 8.5 cm single-sided silicon strip detectors, bonded together. At the outer end of each ladder is a small circuit board, referred to as the ear, which contains the readout electronics (see Figure 2.4). Each ladder is rotated by 3° about its longitudinal axis to allow an overlap between neighboring ladders. In the SVX, a 1.26° gap remained at layer 0, but this was solved in the SVX$'$ by the smaller radius of layer 0 and an additional 1° rotation of the ladders in layer 0.

The raw data collected in the SVX begins as the charge deposited on a strip. First, a pedestal subtraction is applied on a strip-by-strip basis, and then neighboring strips with signals above a threshold are grouped together to form clusters. Different signal-to-noise thresholds are applied to the cluster based on the number of strips used, where the thresholds have been optimized for good efficiency for real hits and rejection of noise. The position of the cluster is determined by a charge-weighting of the positions of the strip centers using

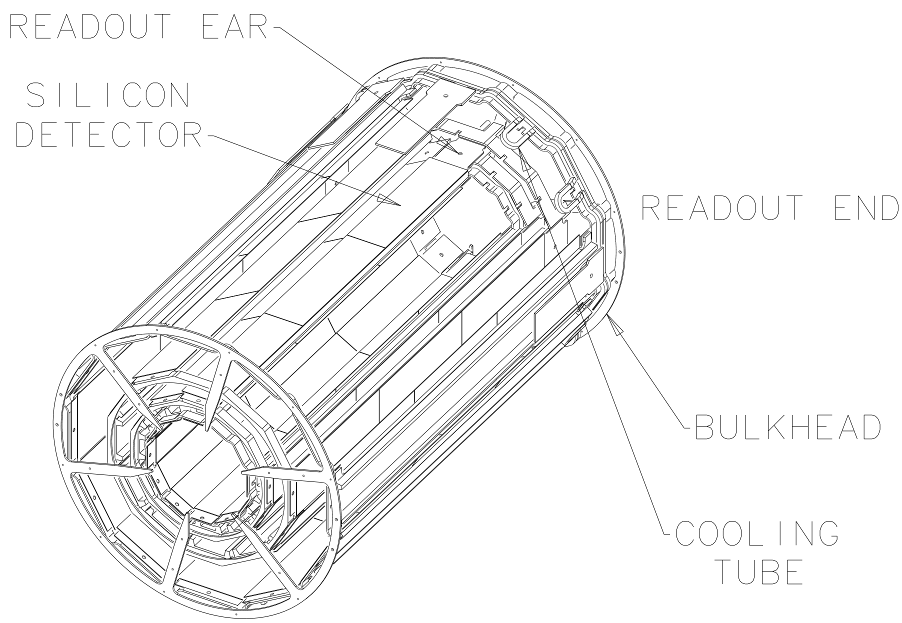


Figure 2.4: Drawing of a single barrel of the SVX detector. The detector is made up of two barrels joined end-to-end, which are readout from the outer ends.

the individual strip charges. The resolution of the cluster position is then based on the total charge and the total number of strips.

2.3.2 CTC

The CTC [31] is a large drift chamber containing 6156 sense wires. These wires are arranged into cells containing 6 or 12 wires each, and these cells are combined to form nine superlayers. The cells are tilted at an angle with respect to the radial direction such that wires from neighboring cells will overlap. Five axial superlayers, comprised of cells of 12 wires each, run parallel to the beamline. These are alternated with four stereo superlayers, comprised of cells of 6 wires. The stereo layers are rotated at an angle of 3° with respect to the beamline in order to provide z information for the tracks. Figure 2.5 is a transverse view of the CTC endplate, showing the cell geometry of the CTC. The nine layers of the CTC

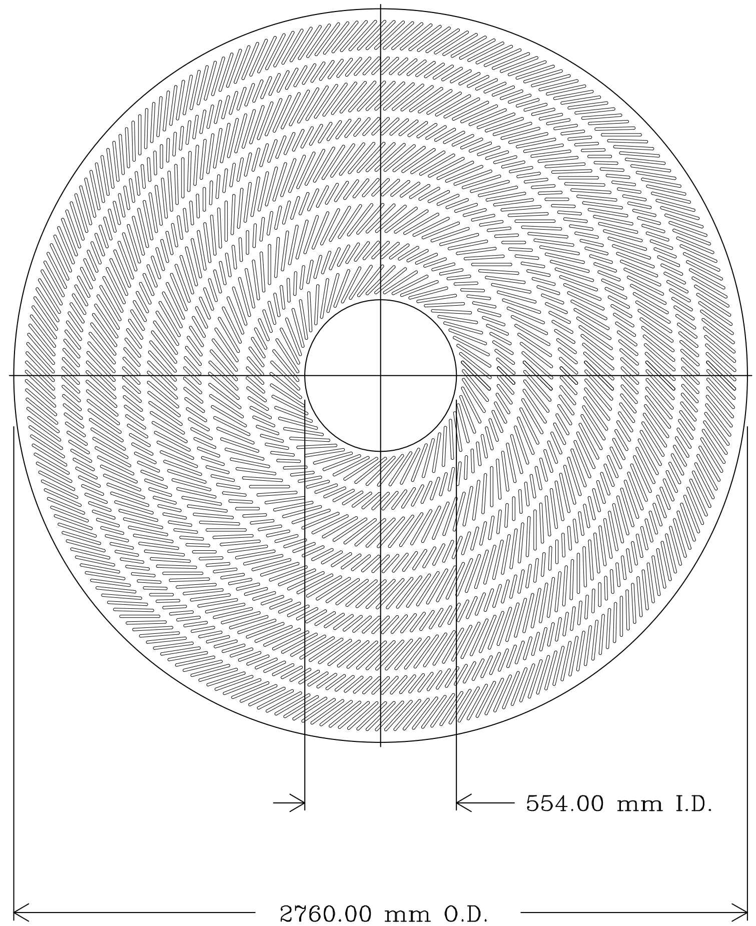


Figure 2.5: Transverse view of the CTC endplate, showing the superlayer geometry.

can be seen, with the smaller cells being the 6 wire cells of the stereo superlayers, while the larger cells are the 12 wire cells of the axial superlayers.

2.3.3 Track Reconstruction

In CDF the paths of particles in the tracking chamber are treated as helices and can therefore be described by the following five parameters:

- $\cot(\theta)$: The cotangent of the polar angle at the point of minimum approach to the origin of the helix

- C: The half curvature of the path, also 1/(radius of curvature)

- z_0: The z coordinate at the point of minimum approach

- D: The signed impact parameter distance between the track and the primary event vertex

- ϕ_0: The azimuthal angle of the track at the point of minimum approach.

Other track parameters, such as p_T, can be derived from these five parameters.

Offline, tracks are reconstructed first in the CTC. A two-dimensional fit in the $r - \phi$ plane is done first, using hits from the five axial superlayers of the CTC. A three-dimensional fit is then performed, using the additional information from the VTX and the four stereo layers of the CTC. This fit returns the five track parameters listed above.

Once a track has been found in the CTC, clusters in the SVX are assigned to the track based on an extrapolation inwards toward the interaction point. The reconstruction begins with the track at the outer wall of the CTC. The track is then extrapolated back to the outer layer of the SVX and corrected for multiple scattering and energy loss in the material in the detector. Based on the track fit from the CTC, a road in the $r - \phi$ plane is calcluated in which to look for clusters in the outer layer of the SVX. If a cluster is found, then that cluster is added to the track, the fit is updated, and the search moves invward to the next layer of the SVX.

The algorithm for reconstructing tracks prefers tracks with more associated clusters. A first pass is made that requires clusters in all four layers of the SVX, followed by a second pass in which a cluster may be missing from one of the four layers. The χ^2 of the fit, based on the errors from the fit, is used to choose the best candidate. If there are no tracks with three or four associated clusters, a search is performed for tracks in which clusters are missing from two layers. A minimum of two clusters must be found for a track to be considered an SVX track. The efficiency for reconstructing tracks in the SVX can be measured from CTC tracks that are well-measured and extrapolate into the fiducial volume of the SVX. The track finding efficiency has been measured at 98.0% for the SVX [32] and 98.7% for the SVX$'$ [34]. This is the efficiency for finding a track with two, three, or four associated

# Clusters	Fraction Reconstructed(%)	
	SVX	SVX′
4	70.5	81.3
3	23.0	15.3
2	4.5	2.1
Total	98.0	98.7

Table 2.2: Fraction of tracks reconstructed with 2, 3, or 4 associated SVX clusters. The fractions are listed separately for SVX and SVX′.

clusters. The separate efficiencies for finding two, three, and four associated clusters are shown in Table 2.2. The single hit(cluster) efficiency is 93% for the SVX and 96% for SVX′, with resolutions of 13 and 12 μm.

Combining tracking information from the CTC and SVX, the transverse momentum resolution for reconstructed tracks is $\frac{\delta p_T}{p_T} = \sqrt{(0.0066)^2 + (0.0009 p_T)^2)}$, where p_T is expressed in $GeV/$c. The impact parameter of a track is defined as the distance of closest approach of the track with respect to the interaction point. The impact parameter resolution using the SVX is (13 + 40/p_T) μm. At low p_T this resolution is dominated by multiple scattering, while at high p_T it is dominated by the intrinsic resolution of the detector.

It is the precise reconstruction of tracks near the collision point by the SVX that has proven to be crucial in performing precision measurements of b hadron lifetimes at CDF. The average b hadron lifetime from decays to $J/\psi X$ has been measured with a 2.5% uncertainty [35]. Using the decays to $D^{*-}\ell^+\nu$ and $\bar{D}^0\ell^+\nu$, the B^0 and B^+ lifetimes have been measured with an uncertainty of 5% [36], while using fully reconstructed decays the lifetimes have been measured with uncertainties of 6% [35] and 4% [35], respectively.

2.4 Calorimetry

The solenoidal magnet is surrounded by calorimeters which are used to measure the electromagnetic and hadronic energies of particles and jets. Extending from -4.2 to 4.2 in η and covering a full 2π in azimuth, they are constructed in a projective tower geometry which

Calorimeter System	Coverage ($\|\eta\|$)	Segmentation ($\Delta\eta \times \Delta\phi$)	Energy Resolution(GeV)	Thickness
CEM	$0.0 < \|\eta\| < 1.1$	$0.1 \times 15°$	$13.7\%/\sqrt{E_T} \oplus 2\%$	$18\ X_0$
CHA	$0.0 < \|\eta\| < 0.9$	$0.1 \times 15°$	$50\%/\sqrt{E_T} \oplus 3\%$	$4.5\ \lambda_0$
WHA	$0.7 < \|\eta\| < 1.3$	$0.1 \times 15°$	$75\%/\sqrt{E} \oplus 4\%$	$4.5\ \lambda_0$
PEM	$1.1 < \|\eta\| < 2.4$	$0.1 \times 5°$	$22\%/\sqrt{E} \oplus 2\%$	18-21 X_0
PHA	$1.3 < \|\eta\| < 2.4$	$0.1 \times 5°$	$106\%/\sqrt{E} \oplus 6\%$	$5.7\ \lambda_0$
FEM	$2.2 < \|\eta\| < 4.2$	$0.1 \times 5°$	$26\%/\sqrt{E} \oplus 2\%$	$25\ X_0$
FHA	$2.4 < \|\eta\| < 4.2$	$0.1 \times 5°$	$137\%/\sqrt{E} \oplus 3\%$	$7.7\ \lambda_0$

Table 2.3: Summary of calorimeter properties. The two terms in the resolution are added in quadrature. Energy is given in GeV. The thicknesses are given in terms of radiation lengths(X_0) or interaction lengths (λ_0).

points back to the nominal interaction point. Each tower covers 0.1 units in η and no more than 15° in ϕ.

The calorimeter is divided into subsystems covering different η regions: central, covering $|\eta| < 1.1$; plug, covering $1.1 < |\eta| < 2.4$; and forward, covering $2.4 < |\eta| < 4.2$. Each of the regions has separate electromagnetic (CEM, PEM, FEM) and hadronic (CHA/WHA, PHA, and FHA) calorimeters, with the hadronic calorimeters lying behind the electromagnetic calorimeters in the radial direction. (See Figure 2.3.) Physical properties of the calorimeter subsystems are summarized in Table 2.3. In this table, the energy resolutions have two components, one energy-dependent and one constant, with the symbol $\oplus$ denoting that the two components are added in quadrature. The thicknesses are given in terms of radiation lengths (X_0) for the electromagnetic calorimeters and interaction lengths (λ_0) for the hadronic calorimeters.

A drawing of one CEM wedge [37] is shown in Figure 2.6. Each wedge is divided into ten towers, each covering $\Delta\eta \times \Delta\phi = 0.1 \times 15°$. The CEM consists of 31 layers of polystyrene scintillator interleaved with 30 layers of lead absorber. The layers lie parallel to the beamline, meaning that a particle passes through more material per layer as a function of the polar angle θ. To account for this, acrylic has been substituted for lead in some layers of certain towers. Light from the scintillators is converted in wavelength shifters and read out through two photomultiplier tubes (shown in Figure 2.6) for each tower.

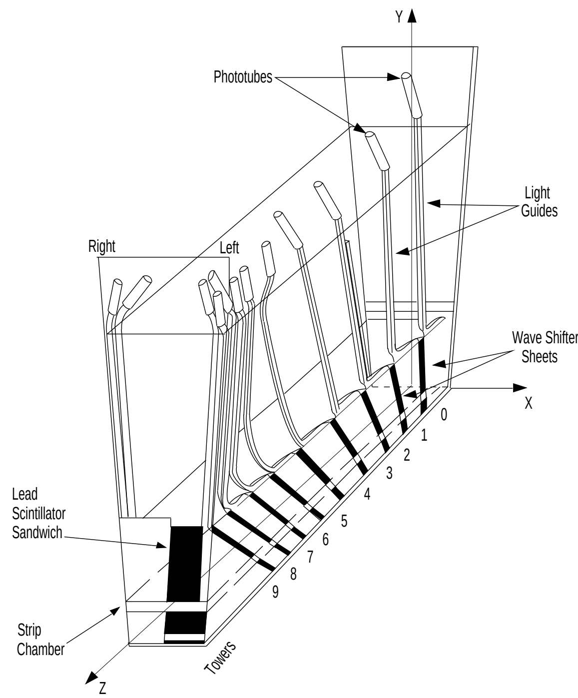


Figure 2.6: Schematic drawing of one CEM wedge, covering $\Delta\eta \times \Delta\phi = 1.0 \times 15°$. There are a total of 48 wedges in the CEM.

A wire proportional chamber (CES) lies inside the CEM at a depth which corresponds to the maximum average transverse shower development. (Shown near the bottom of Figure 2.6.) The chamber provides position measurements in both the z and $r - \phi$ views, allowing precise measurement of the shower profile.

The CHA [38] actually refers to both the central (CHA) and endwall (WHA) hadron calorimeters. The η coverage of the two components can be seen in Table 2.3. Figure 2.3 shows the positioning of the two components outside the CEM and shows their overlapping coverage. The CHA has the same segmentation as the CEM and is composed of layers of acrylic scintillators alternated with layers of iron absorber. Like the CEM it is read out by pairs of phototubes.

The PEM [39] and FEM [40] are both gas-based calorimeters, alternating layers of lead absorber panels with conductive plastic proportional tube arrays. Each layer is read out by cathode pads and arranged to form towers, leading to a position resolution of approximately 2 mm in the plug region with similar resolution in the forward. Similarly, the PHA [41] and FHA [42] alternate steel absorber plates with conductive plastic proportional tubes with cathode readout.

2.5 Muon Detectors

The muon detectors at CDF consist of several subsystems which cover different ranges of η. The two sets of muon chambers in the central region are the Central Muon Chambers (CMU) and Central Muon Upgrade (CMP). Extending further out in η is the Central Muon Extension (CMX), and in the far forward region is the Forward Muon detector (FMU).

The CMU [43] lies directly outside the central calorimeters, forming a cylinder. It is divided into 24 wedges, each covering 15° in ϕ, with each wedge divided into three 5° towers. Each wedge is also divided into east and west sections at $\theta = 90^o$. One 5° section of a wedge is shown in Figure 2.7, with the solid line representing a muon track passing through the wedge. Each wedge contains four layers of drift chambers, giving measurements at four points along the trajectory of a particle. The deviation from the radial direction (indicated by the dashed line in Figure 2.7) is due to the curvature of the track in the magnetic field preceding the muon chambers. The individual drift cells shown in Figure 2.7 contain an argon/ethane gas mixture with a high voltage sense wire at the center of the cell. The difference in drift times measured at different layers of the chamber can be converted to an angle and therefore a measurement of the p_T of the track, useful at the trigger level. The z position of the track can be determined by measuring the relative charge at the two ends of a sense wire, giving an accuracy of 1.2 mm on each point. Figure 2.8 shows the coverage of one 15° CMU wedge in ϕ (left) and z (right). This figure shows the gaps at the edges of each wedge. Due to these small gaps and the gap at the boundary at $\theta = 90^o$ the CMU chambers only cover approximately 84% of the solid angle for $|\eta| < 0.6$.

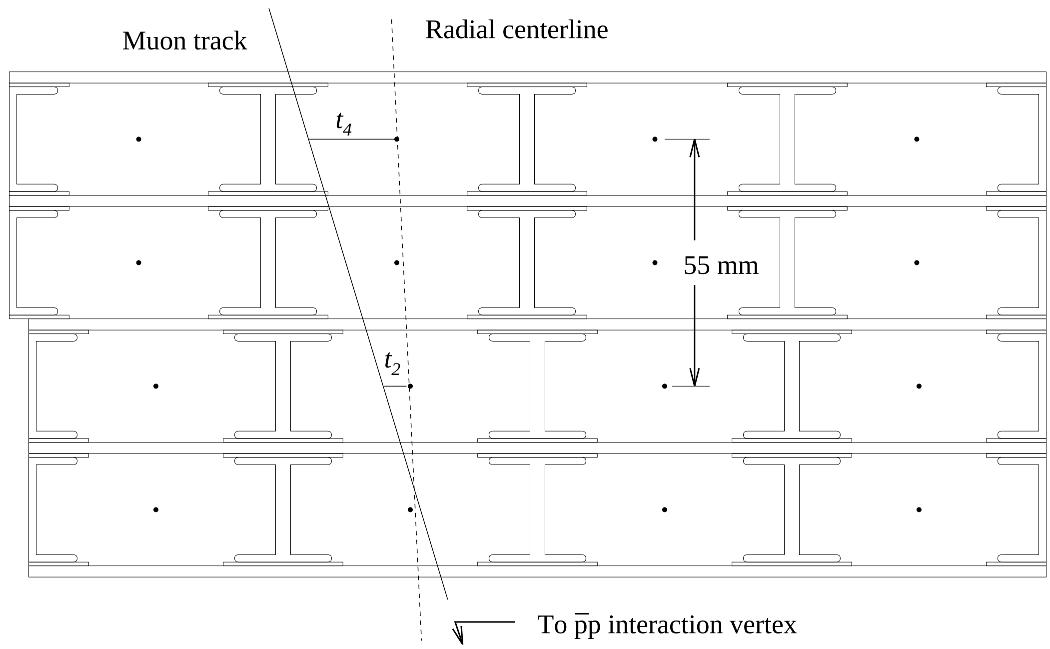


Figure 2.7: Drawing of a 5° section of the Central Muon Drift Chambers

Though the CHA acts as an absorber for the CMU, earlier data-taking showed that pion punch-through remained a significant background. For the higher luminosity achieved during Run I, a layer of steel 0.6 m thick was added behind the CMU for additional hadron absorption, and the CMP chambers were added behind this steel.

Like the CMU, the CMP chambers [44] are made up of four layers of drift chambers aligned along the z axis, though no z information is read out for the CMP. In contrast to the cylindrical shape of the CMU, the CMP chambers are mounted on four flat planes that surround the central detector. Due to gaps in the coverage of the CMP, it covers only 63% of the solid angle for $|\eta| < 0.6$, and together the CMU and CMP cover only 53%.

Additional muon coverage extends to $\eta \sim 1.0$ with the CMX chambers. These chambers are arranged in a conical shape at the ends of the central detector, as shown in Figures 2.2 and 2.3. Muon coverage in the far forward region is achieved with the FMU. These chambers are disk-shaped, lying beyond the forward calorimeters, shown at the far

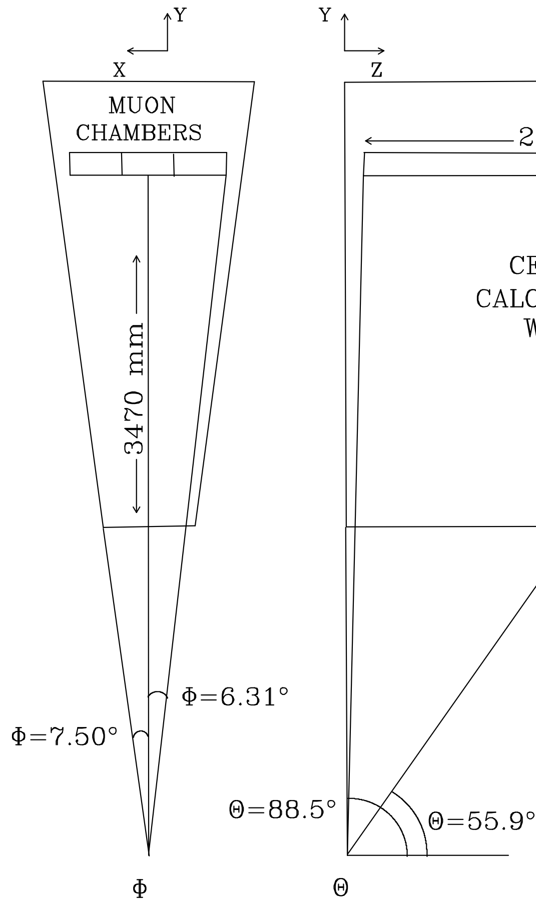


Figure 2.8: One 15° Central Muon Chamber, showing the coverage in the transverse plane (left) and in the z direction (right).

left of Figure 2.3. Because of the high background rates in the muon samples for the CMX and FMU, muons identified in those regions are not used in this analysis.

2.6 Trigger

One of the greatest challenges of hadron collider physics lies in the trigger. The total cross section is $\sigma \sim$100 mb, while the b quark cross section is $\sigma_b \sim$50 μb, three orders of magnitude smaller. With a time between beam crossings of only 3.5 μsec, it is impossible to record every interaction. Thus it is necessary to quickly identify events which might be of interest, while rejecting others. At CDF this is done through a three-level trigger system, with each level making a more careful decision and requiring more time to make it. Beginning with an input rate of 286 kHz, the event rate is reduced to $\sim$10 Hz which can be written onto tape.

The Level 1 trigger decision must be made within the 3.5 μsec between crossings in order to avoid incurring any deadtime, where other crossings are missed. The Level 2 trigger then takes $\sim$30 μsec to make its decision. During this time deadtime accumulates, but since the rate into Level 2 is $\sim$100 times lower than the rate into Level 1, this deadtime is small compared to the livetime.

Once an event has been accepted by Level 2, signals from the entire detector are read out [45] and that data is sent to Level 3. Once this is done the detector becomes live again to examine subsequent crossings.

Whereas the Level 1 and Level 2 triggers are hardware based [46], the Level 3 trigger is software based. The Level 3 trigger [47] consists of a farm of 48 commercial processors, each containing two buffers. A given event is sent to only one buffer, allowing the processing of multiple events in parallel. Level 3 performs a nearly complete reconstruction of the event, using a version of the offline software package that has been simplified for increased speed, such that the processing of a single event takes only 1-2 seconds. The final rate out of Level 3 is $\sim$10 Hz, which is written to 8 mm tape.

The events used in this analysis are triggered by the lepton from the B_S^0 decay. The following discussion of the trigger requrirements at each level focuses on the single lepton triggers that the events must pass.

2.6.1 Level 1

For electrons, the primary trigger at Level 1 is the calorimeter trigger, which also serves as the primary trigger for photons and hadronic jets. Analog signals from the calorimeters are combined in towers of size $(\Delta\phi = 15°) \times (\Delta\eta = 0.2)$. The signals are first corrected for variations in gain and offset and are then multiplied by sinθ, where θ is the polar angle. This allows triggering to be based on transeverse energy, E_T, as opposed to E.

The trigger decision is based on a single tower over threshold. Thresholds are set separately for each region of the calorimeter, and as well for the EM and HAD calorimeters. In Run I the lowest threshold was used in the CEM, where the threshold was set at 6-8 GeV. The highest thresholds are used in the forward EM calorimeters and in the plug and forward HAD calorimeters. During Run I the thresholds in these regions were often set to 51 GeV, corresponding to full scale for a single tower.

There is a second calorimeter trigger whose thresholds are lowered in all regions except for the PHA and FHA, but this trigger operates in only one crossing out of every 40. Thus the contribution of events coming from this trigger is much lower in comparison to the main calorimeter trigger.

For muons there are two primary types of triggers at Level 1: single muon and dimuon triggers. These triggers require either one or two stubs in the muon chambers above a momentum threshold, with a lower threshold used in the dimuon triggers. In the regions of the CMU that are covered by the CMP, a corresponding coincidence of hits in the CMP is required as confirmation of the hits in the CMU.

The momentum measurement in the muon chambers is based on the timing difference between hits in 2 layers of a single tower, as described in Section 2.5 and illustrated in Figure 2.7. The difference in times, $\Delta t = t_4 - t_2$ in Figure 2.7, is approximately inversely proportional to the transverse momentum of the track. For the CMU the momentum thresholds during Run I were 6 $GeV/$c for the single muon trigger and 3.3 $GeV/$c for dimuons.

The cross section at Level 1 for the calorimeter trigger was σ ~20 μb and for the single muon CMU-CMP trigger was σ ~40 μb, while the total Level 1 cross section was σ ~100 μb. This corresponds to a Level 1 accept rate of ~1-2 kHz, a reduction of over 2 orders of magnitude from the input rate.

2.6.2 Level 2

Once an event has passed the Level 1 trigger, the detector signals are held while a Level 2 decision is made. As mentioned above, the time required to make a Level 2 trigger decision is $\sim$ 30 μsec, allowing for a more intelligent decision.

The most important additional information used at Level 2 comes from tracking. The Central Fast Tracker (CFT) [48] uses the $r-\phi$ information from the five axial superlayers of the CTC to find tracks and to make a fast measurement of the p_T and ϕ of the tracks. The CFT makes two passes looking for hits in a superlayer. The first looks for hits very near a wire, called prompt hits, while the second pass looks for hits that have drifted a greater distance, called delayed hits. A track typically has one prompt hit and two delayed hits on each of the axial layers. This yields a total of 15 hits for a given track.

The pattern of hits is compared with lookup tables for pattern recognition. The prompt hit on the outermost superlayer is used to start the lookup, and also defines the exit ϕ of the track. Tracks are binned in eight p_T ranges and by charge, with a resolution of $\delta p_T/p_T^2 \sim 3.5\%$. The nominal p_T threshold of a given bin is defined by the 90% efficiency point for that bin. For Run Ia the p_T bins were at 3.0, 3.7, 4.8, 6.0, 9.2, 13.0, 16.7, and 25.0 $GeV/$c. For Run Ib, the values were 2.2, 2.7, 3.4, 4.7, 7.5, 12.0, 18.0, and 27.0 $GeV/$c, allowing for triggering on tracks at lower p_T.

The single muon trigger at Level 2 requires that a CFT track above a p_T threshold matches to a muon stub. A lookup table is used to search for matches. For each wire in the outermost superlayer and for each p_T bin, a straight-line extrapolation of the track out to the muon chambers has been performed. Some tolerance is allowed to account for multiple scattering, resulting in a window of $\Delta\phi = 5^\circ - 10^\circ$. Thus for each CFT track above the p_T threshold, the list of muon towers with stubs is compared to the list of expected hits. If a match is found then the track is considered to have matched to the muon stub. For Run Ia the p_T threshold for the lowest p_T CMU trigger was 6.0 $GeV/$c. For Run Ib the threshold was increased to 7.5 $GeV/$c to keep the trigger rate manageable in the high luminosity running conditions of Run Ib.

At Level 2 rather than using the information from single calorimeter towers to make a decision, the trigger builds calorimeter clusters. Two sets of thresholds are applied to all of the towers. The higher threshold is referred to as the seed threshold and the lower threshold is referred to as the shoulder threshold. A cluster begins with a tower above the seed threshold. Then the four nearest neighbors in η-ϕ space are checked to see if they are above the shoulder threshold. If a tower above the shoulder threshold is found, it is added to the cluster and its neighbors are examined. This process continues until all neighbors have been examined.

Once a cluster has been found, it is added to a list of all clusters, along with several associated quantities such as E_T, $E_T \sin\phi$, $E_T \cos\phi$, η, and ϕ. There are several passes of cluster finding at Level 2, including jet cluster finding and two passes of electron/photon cluster finding. In electron/photon cluster finding only the information from the EM calorimeters in used to build the cluster. Two passes are made, with the second pass using lower seed/shoulder thresholds to allow detection of lower E_T clusters.

Using the ϕ information for the cluster, electrons can be found by matching CFT tracks to the clusters. This is almost identical to the matching done for muons, except that the 15° calorimeter wedge is used in place of the 5° CMU tower. The E_T threshold used for the single electron trigger was 6.0 GeV in Run Ia and 8.0 GeV in Run Ib, while the corresponding track p_T thresholds were identical to those in the single muon triggers.

In Run Ib an additional requirment was placed on the 8 GeV single electron trigger. The shower position information from the CES was used by matching CFT tracks to a group of four adjacent CES wires (corresponding to $\sim 2°$ in ϕ) whose summed energy was above a threshold. The electron triggers then had the added requirement that tracks matching to an electron cluster also matched to a CES cluster.

The single electron trigger and the single CMU-CMP muon trigger each accounted for 10-15% of the total Level 2 cross section of $\sim$2 μb. Due to high background rates at higher luminosity, these triggers were both assigned variable prescales, meaning a lower fraction of events passing the trigger was written out at high luminosity. For the electron trigger, the trigger was prescaled by as much as a factor of 4, while the muon trigger was prescaled by as much as a factor of 8. Higher p_T electron and muon triggers were unprescaled, but the higher p_T thresholds on these triggers give them low acceptance for the events of interest in this analysis.

2.6.3 Level 3

The Level 3 trigger is used to not only select events, but also to split the data into several output streams based on the reconstructed objects in the events. These streams allow for easy filtering of the data for later offline processing. As there are many Level 3 triggers, the discussion is restricted to the triggers relevant to this analysis.

The inclusive muon trigger at Level 3 in Run I required $p_T > 8.0$ $GeV/$c and applied a cut on the quality of the match between the CMU stub and the CTC track extrapolated to the CMU chambers. The primary inclusive electron trigger at Level 3 required $E_T >$ 7.5 GeV and $p_T > 6.0$ $GeV/$c, and required tracks to match within 3 cm in x and 10 cm in z. Further selection cuts can be made at Level 3 using offline quantities, such as the ratio of hadronic to electromagnetic energy ($E_{\mathrm{had}}/E_{\mathrm{em}}$) and the quality of the match of the CTC track to the cluster position information from the strip and wire chambers. These cuts are described in further detail in the following chapter. The cross section for each of these triggers at Level 3 was $\sim$10 nb, corresponding to a rate of less than 1 Hz.

Chapter 3

Event Selection

In the decay of interest, the B_S^0 meson decays to $D_S^- \ell^+ X$, where the D_S^- is reconstructed through its decay to other hadrons. The high momentum lepton in the event is the key signature of the B_S^0 decay and is the starting point for the reconstruction of the event. In this chapter the selection of the electron and muon samples is described first. Then the reconstruction of the event is discussed and the selection of D_S^- candidates from the decay $D_S^- \to \phi\pi^-$, with $\phi \to K^+K^-$, is described.

3.1 Lepton Selection

The lepton datasets are made up of the events that have passed the low p_T inclusive electron and muon triggers at Level 3. The events in these datasets are first processed with the standard offline reconstruction software. Track-finding is repeated, starting from the hits in the SVX and CTC, using the best alignment constants for the SVX [49]. Tighter selection requirements are then applied to select good lepton candidates.

3.1.1 Electron Selection

The list of the selection requirements for electrons is shown in Table 3.1. Quantities used in the selection are described here.

- E_T: The transverse energy of the cluster, defined as $E \times \sin\theta$. Electrons produced in the decays of the B_S^0 should have higher E_T than electrons coming from the sequential decays of charmed hadrons or other background processes.

- E_T/p_T: The ratio of transverse energy in the cluster to the transverse momentum of the associated track. For electrons, E_T/p_T is expected to be near 1.0.

- $E_{\mathrm{had}}/E_{\mathrm{em}}$: The ratio of hadronic energy to electromagnetic energy in the cluster. Electron showers should be well contained in the electromagnetic calorimeters. Hadronic showers will deposit energy in both the EM and HAD calorimeters. Different values for this cut may be appropriate if the cluster is isolated and only one track points to the cluster, or if several tracks point to the cluster.

- LSHR: The lateral shower profile. For electrons this has been measured with test beam electrons. The lateral energy sharing of the candidate can be compared with the measured shape for electrons. Low LSHR values indicate a good match between the observed shape and the expected shape for electrons. Wider clusters coming from multiple particles will give higher LSHR values.

- Δx, Δz of track vs. CES wire: The position of the candidate track, extrapolated from the CTC, is required to match well to the position of the shower in the CES in both x $(r-\phi)$ and z. Overlapping backgrounds created by tracks from charged hadrons and energy deposited by neutrals should not match as well.

- χ^2_{strip}, χ^2_{wire}: The pulse height shape of electrons has been determined from test beam electrons. Candidate electron showers must be consistent with the expected shape. A χ^2 requirement is made in both the wire $(r-\phi)$ and strip (z) views.

In addition, electrons are removed if they are consistent with being produced by the conversion of photons in the detector material prior to entering the CTC. Photon conversions will produce an electron-positron pair, so the track belonging to the conversion partner must be located. As these tracks should have few VTX hits, the ratio of the number of found VTX hits to the expected number of VTX hits must be less than 0.5. Conversion candidates

Variable		Cut Value
E_T	$>$	6.0 GeV
$E_{\text{had}}/E_{\text{em}}$	$<$	0.04 (1 track pointing to cluster)
$E_{\text{had}}/E_{\text{em}}$	$<$	0.10 (>1 track pointing to cluster)
LSHR	$<$	0.2
Δx (track-CES)	$<$	1.5 cm
Δz (track-CES)	$<$	3.0 cm
strip profile χ^2	$<$	10
wire profile χ^2	$<$	15

Table 3.1: Selection criteria for the electron sample.

and their partners are then selected based on three variables relating to the two tracks: (1) S, the separation in cm between the two tracks in the $r-\phi$ plane at the point where they are tangent; (2) $\Delta\cot\theta$ between the two tracks; and (3) R_c, the radius in cm of the tangent point, assumed to be the most likely conversion point. An electron is tagged as a conversion if

- $|S| < 0.02$,
- $|\Delta\cot\theta| < 0.06$, and
- $18.0 < R_c < 50.0$.

The electron dataset initially contains approximately 7.5 million events. The selection cuts reduce the sample to approximately 2.8 million events.

3.1.2 Muon Selection

The selection of candidate muons is based primarily on the matching of the position of the track extrapolated from the CTC to the muon chambers and the position of the stub in the muon chambers. As stated earlier, muons candidates in the CMX and FMU are not included. The muon candidate must have a stub in both the CMU and the CMP. A χ^2 requirement is made on the match between the extrapolated CTC track and the muon stub in the $r-\phi$ plane for both the CMU and CMP, and in z for the CMU. The uncertainty

Variable		Cut Value
p_T	$>$	6.0 $GeV/$c
χ^2_{x}(CMU)	$<$	9
χ^2_{z}(CMU)	$<$	12
χ^2_{x}(CMP)	$<$	9

Table 3.2: Selection criteria for the muon sample.

on the position of the track is dominated by multiple scattering, which to first order varies inversely with the track momentum. A p_T requirement is also made, analagous to the E_T requirement in the electron selection. The list of the selection requirements for muons is shown in Table 3.2. The initial muon dataset contains approximately 2.4 million events. Approximately 1.7 million events remain after applying the muon selection requirements.

3.2 Track Selection

Once a candidate lepton has been found, a search is performed for candidate D_S^- tracks near the lepton. In order to use only tracks which are well-measured, a sample of tracks is selected which pass the following requirements:

- $\geq$ 4 hits on 2 CTC axial layers
- $\geq$ 2 hits on 2 CTC stereo layers
- SVX-CTC link required
- $\geq$ 3 SVX hits
- χ^2_{SVX} / # SVX hits < 6.

The track associated with the lepton is required to pass the same cuts.

3.3 Primary Vertex Reconstruction

Events in which a B_S^0 meson is produced typically have few tracks other than those associated with the B_S^0. There are few tracks coming from the primary vertex, and since the tracks coming from the B_S^0 decay are displaced with respect to the primary vertex, it is difficult to determine the coordinates of the primary vertex without being biased toward the B decay vertex. Rather than reconstructing the primary vertex on an event-by-event basis, the average beam position as determined in the SVX on a run-by-run basis is used as the primary vertex [35].

The beam position is found by first fitting the primary vertex in all events in a data acquisition run. Using only those events with at least three tracks with $p_T >$ 1.0 $GeV/$c and four associated SVX hits, a straight line is fit through the vertices to determine the beamline for that run. A minimum of a few hundred events in a run is required, and the x and y positions of the beam at $z = 0$ as well as the slope of the beam as a function of z are stored in a database. The error on the beam position is $\sim$ 30 μm, and it is roughly constant as a function of z. The z distribution of primary vertices is approximately gaussian, with a sigma of roughly 30 cm [50].

To determine the primary vertex in an event in one of the lepton samples, the z coordinate of the primary vertex is determined first by using the z coordinate of the primary vertex found by the VTX closest to the z position of the lepton. Using this z coordinate, the x and y coordinates of the primary vertex can be determined using the information in the database.

3.4 B_S^0 Reconstruction

A schematic view of the decay $B_S^0 \to D_S^- \ell^+ \nu$ with $D_S^- \to \phi\pi^-$ is shown in Figure 3.1. The reconstruction starts with the candidate lepton. Looking near the lepton, oppositely charged pairs of tracks are used, each assigned the K mass, in order to reconstruct the ϕ

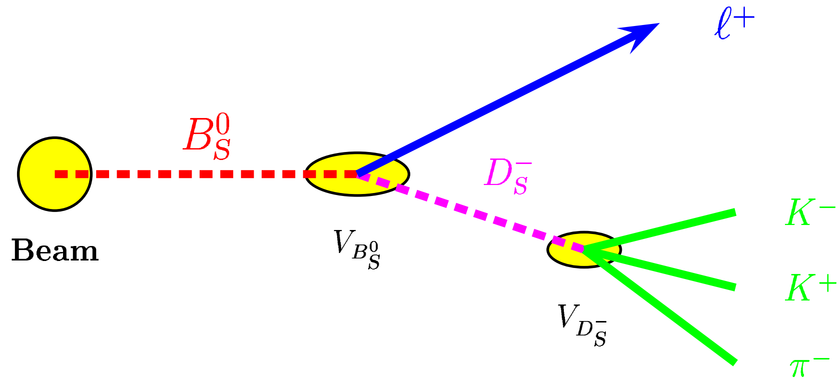


Figure 3.1: Sketch of the decay mode $B^0_S \to D^-_S \ell^+ \nu$.

candidate. If the invariant mass of the pair of tracks corresponds to the ϕ mass, a third track is added, assigned the π mass, to form the D^-_S candidate.

To find the D^-_S decay vertex, all three tracks are required to intersect at a common point, and to have an invariant mass that corresponds to the D^-_S mass. The vertex is reconstructed using a χ^2 minimization that varies the input track parameters within their errors. Using the momenta of the candidate tracks, the D^-_S candidate is extrapolated back along its flight path to intersect it with the lepton flight direction to find the B^0_S decay vertex. A similar χ^2 minimization is used again. Based on the returned χ^2 for the combined two-vertex fit, a probability can be computed. Combinations with low probabilities, where the vertex fit is poor, can then be removed.

3.5 $D^-_S \to \phi\pi^-$ Selection

Having selected a candidate lepton, all tracks in a cone around the lepton with radius $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$=0.8 are examined as candidate K and π tracks. The selection cuts for D^-_S candidates were chosen to optimize the signal-to-noise ratio, with the signal modelled by Monte Carlo events. D^-_S candidates are selected with the following requirements:

- $p_T(K) > 1.2\ GeV/\mathrm{c}$
- $p_T(\pi) > 0.8\ GeV/\mathrm{c}$
- $\Delta R < 0.8$
- $|z_{vertex} - z_{track}| < 5$ cm
- $|\cos\psi| > 0.4$, where ψ is the helicity angle between the K and D_S^- in the ϕ rest frame. The ϕ has spin 1, while the D_S^- and π both have spin 0. The angle ψ will then have a distribution $dN/d(\cos\psi) \sim \cos^2\psi$. The background from random combinations of tracks will be flat vs. $\cos\psi$.
- Prob(χ^2_{vertex}) $> 1\%$ for the two-vertex system.
- $I_{Cal} < 1.0$, where I_{Cal} is a calorimetry-based measure of the isolation, defined as the ratio of the E_T in a cone of radius $\Delta R < 0.4$ around the lepton candidate, excluding the lepton energy, to the p_T of the D_S^-. This isolation requirement removes many fake D_S^- candidates from jets with high track multiplicities.
- SVX beam position information must be available.
- $L_{\mathrm{xy}}(D_S^{PV}) > 0.0$, where $L_{\mathrm{xy}}(D_S^{PV})$ refers to the distance between the D_S^- vertex and the primary vertex in the transverse plane. This cut removes fake D_S^- combinations generated by tracks from the primary vertex.
- $|\mathrm{m}(K^+K^-)$ - $1.0194| < 0.010\ GeV/\mathrm{c}^2$, where 1.0194 is the nominal PDG [52] ϕ mass
- $3.0 < \mathrm{m}(D_S^-\ell^+) < 5.0\ GeV/\mathrm{c}^2$
- Level$(K/\pi) > 0.01$ (see Section 3.5.1)

The D_S^- decay length, denoted by $L_{\mathrm{xy}}(D_S^-)$ is the distance between the B_S^0 and D_S^- vertices. The proper decay length, denoted by $c\tau$, is defined as $c\tau = L_{\mathrm{xy}} \cdot m/p_T$. (See Section 4.2 for further explanation.) It is the proper decay length that is used to reconstruct the lifetime. A cut is applied on the errors on the proper decay lengths to remove events with large errors. In addition, a cut is made on the proper decay length of the D_S^- candidate to remove events where the extrapolation from the D_S^- vertex to the B_S^0 vertex is very long and results in a large uncertainty on the B_S^0 vertex position. The cuts used are:

- $\sigma_{c\tau}(D_S^-) < 0.1$ cm
- $\sigma_{c\tau}(B_S^0) < 0.1$ cm
- $|c\tau(D_S^-)| < 0.1$ cm.

3.5.1 *Use of dE/dx for K, π Selection*

It is possible to distinguish between different particles based on the energy loss per unit distance (dE/dx) in matter. The most probable value of the ionization energy loss in a gas is a function of the $\beta\gamma$ ($= p/mc$) of the particle, as opposed to simply its momentum. Thus a simultaneous measurement of a particle's momentum and dE/dx should determine the mass of the particle. dE/dx is expressed in nanoseconds at CDF, as the information is recorded as the digital pulse width between the leading and trailing edge times for the CTC hits in superlayers 3 through 8. The dE/dx data must be corrected for variations due to path length, wire gain saturation, the number of measurements used for the track, and the instantaneous luminosity [51].

For a given track the predicted dE/dx value, as well as the error on the prediction, can be computed for the hypothesis that the particle is either a π, K, p, e, or μ. The difference between the measured and predicted dE/dx values, divided by the error, is used to compute a probability for each hypothesis. In order to make use of all available information, the probabilities $P(i)$ for the five particle hypotheses are used to compute the following confidence level:

$$Level(i) = \frac{P(i)}{P(\pi) + P(K) + P(p) + P(e) + P(\mu)}.$$

In Figure 3.2(a) the Level(K) distribution is plotted for K candidates from $\phi \rightarrow K^+K^-$ taken from the $D_S^- \rightarrow \phi\pi^-$ signal region, which contains little background (see

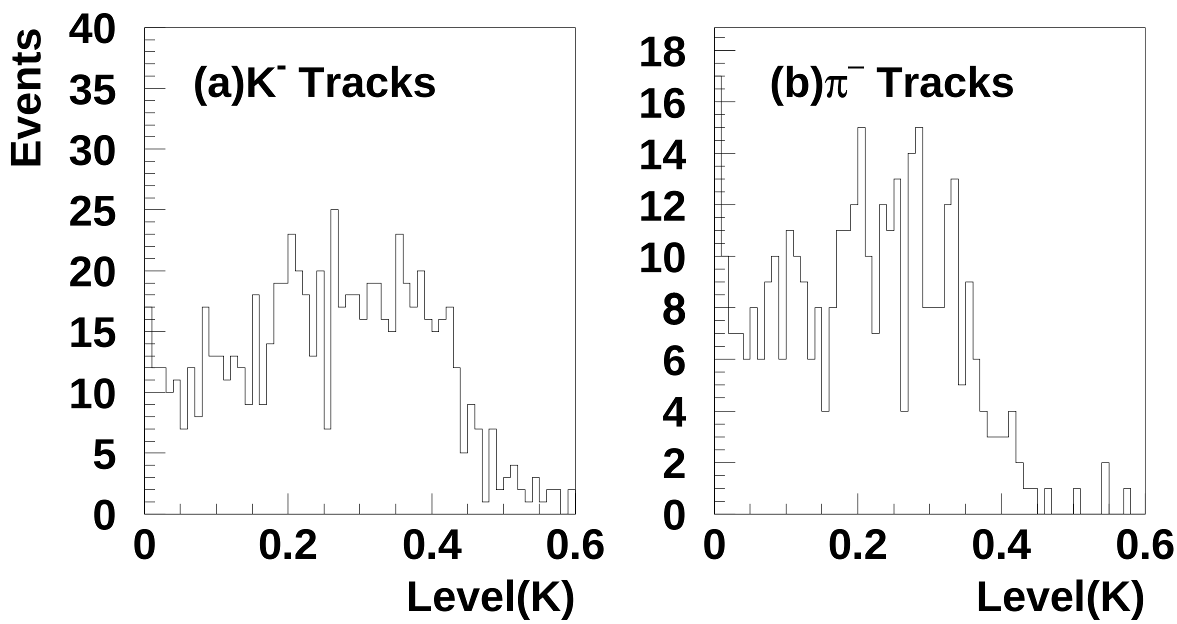


Figure 3.2: Level(K) distribution for the K hypothesis for (a) kaon tracks and (b) pion tracks from $D_S^- \to \phi\pi^-$, $\phi \to K^+K^-$.

Figure 3.3). In Figure 3.2(b) the same Level(K) distribution is shown for π^- tracks from the $D_S^- \to \phi\pi^-$ decay assuming the kaon hypothesis.

It can be seen that the dE/dx information doesn't provide much discriminating power. However, some of the π tracks with very small Level(K) values can be eliminated. In the $D_S^- \to \phi\pi^-$ selection, Level(K/π) > 0.01 is required for each of the K^-, K^+, and π^- tracks. Though these cuts are not crucial in the $D_S^- \to \phi\pi^-$ decay mode, they will be important in the addition of the other D_S^- decay modes.

3.6 D_S^- Mass Results

For real B_S^0 decays, the lepton and the D_S^- should have opposite charges. Such events are referred to as "right-sign" events, while events where the lepton and D_S^- have the same charge are referred to as "wrong-sign". A spurious enhancement of the background in the

D_S^- mass spectrum due to the selection cuts should be uncorrelated with the relative charge of the lepton and should therefore appear in both the right-sign and wrong-sign D_S^- mass distributions.

The D_S^- mass distribution is fit with a gaussian to describe the signal, and a first-order polynomial to describe the combinatoric background. A second gaussian is included for the Cabibbo-suppressed $D^- \to \phi\pi^-$ decay. The right-sign $K^-K^+\pi^-$ mass distribution for the $D_S^- \to \phi\pi^-$ sample is shown in Figure 3.3 with the fit result overlaid. The shaded histogram shows the wrong-sign $K^-K^+\pi^-$ mass spectrum. As expected, there is no evidence of any enhancement in the wrong-sign mass spectrum. The difference of the means between the D_S^- and D^- peaks is fixed to 99.2 MeV/c^2 [52], and the ratio of the widths is fixed to the ratio of their masses. In addition, the region 1750 - 1820 MeV/c^2 has been excluded in the fit to avoid events with a partially reconstructed D_S^- where a π^0 has been missed. The fit yields a signal of 220 $\pm$ 21 $\phi\pi^-$ events in the peak. The D_S^- mass is 1968 $\pm$ 1 MeV/c^2, in agreement with the PDG value of 1969 MeV/c^2 [52], and the width of the gaussian is 10.5 $\pm$ 1.1 MeV/c^2. For the D^- the fit returns 48 ± 15 events in the peak.

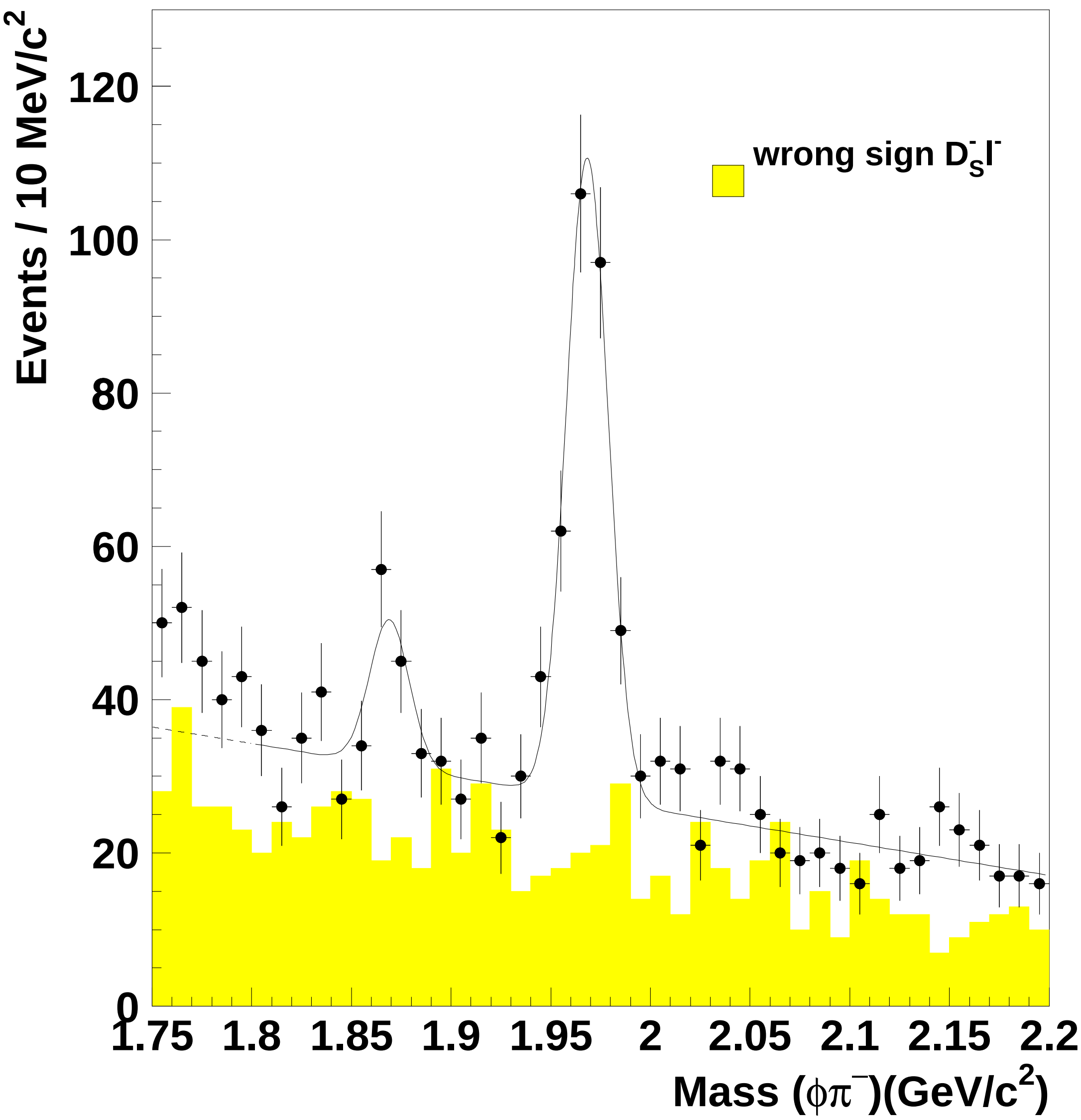


Figure 3.3: Right-sign $\phi\pi^-$ mass distribution with the result of the fit superimposed. The shaded histogram shows the wrong-sign distribution.

Chapter 4

Lifetime Measurement in $D_S^- \to \phi\pi^-$ Sample

This chapter describes the measurement of the B_S^0 lifetime using the events in the $D_S^- \to \phi\pi^-$ sample. First the variables used for the lifetime measurement are defined, including those introduced to account for the partial reconstruction of the B_S^0. Then the fitting procedure is discussed and the lifetime measured in this sample is presented.

4.1 Decay Length Reconstruction

As described in Sections 3.3 and 3.4, the primary vertex in the event is obtained from the beam position, while the B_S^0 and D_S^- vertices (denoted by $V_{B_S^0}$ and $V_{D_S^-}$ in Figure 3.1) are found by a fit to the candidate tracks. The decay length of the B_S^0 in the transverse plane, referred to as $L_{\rm xy}(B_S^0)$, is then defined as the displacement of $V_{B_S^0}$ from the primary vertex, projected onto the transverse momentum of the $D_S^-\ell^+$ system. If the vector $\vec{X}$ points from the primary vertex to the secondary vertex in the transverse plane, then

$$L_{\rm xy}(B_S^0) = \frac{\vec{X} \cdot \vec{p_T}(D_S^-\ell^+)}{|\vec{p_T}(D_S^-\ell^+)|}.$$

Similarly, the decay length of the D_S^- in the transverse plane, denoted by $L_{\rm xy}(D_S^-)$, is the displacement of $V_{D_S^-}$ from $V_{B_S^0}$ in the transverse plane, projected onto the transverse momentum of the D_S^-.

The uncertainty on the transverse decay length is dominated by the uncertainties on the positions of the primary and secondary vertices. Neglecting the contributions from uncertainties in the transverse momenta, the uncertainty on $L_{\mathrm{xy}}(B_S^0)$ is given by

$$\sigma^2_{L_{\mathrm{xy}}} = \frac{1}{p_T^2} \times [(\sigma_x^{PV} p_x)^2 \ + \ (\sigma_y^{PV} p_y)^2 \ + \ (\sigma_x^{SV} p_x)^2 \ + \ (\sigma_y^{SV} p_y)^2 \ + \ 2((\sigma_{xy}^{PV})^2 + (\sigma_{xy}^{SV})^2) p_x p_y],$$

where σ_x^{PV}, σ_y^{PV}, and σ_{xy}^{PV} are the errors on the primary vertex coordinates; σ_x^{SV}, σ_y^{SV}, and σ_{xy}^{SV} are the errors on the secondary vertex coordinates; and p_x and p_y are the x and y components of the transverse momenta of the $D_S^- \ell^+$ system. The uncertainty on $L_{\mathrm{xy}}(D_S^-)$ can be obtained from the same formula by substituting the errors on the secondary and tertiary vertices for the errors on the primary and secondary vertices and using the momemtum components of the D_S^- in place of those of the $D_S^- \ell^+$ system.

4.2 Definition of Pseudo-Proper Decay Length

The decay length, L, is related to the lifetime, τ, by the equation

$$L = c\tau\beta\gamma = c\tau\frac{p}{m},$$

where $c\tau$ is called the proper decay length. In the transverse plane this becomes

$$L_{\mathrm{xy}} = c\tau\frac{p_T}{m}.$$

When the B_S^0 decays semi-leptonically, it is not fully reconstructed and thus $p_T(B_S^0)$ is not accurately measured. Instead the p_T of the $D_S^- \ell^+$ system must be used as the best approximation. A correction factor, K, must then be introduced, defined by

$$K = \frac{p_T(D_S^- \ell^+)}{p_T(B_S^0)}.$$

The quantity used to extract the B_S^0 lifetime is then referred to as the pseudo-proper decay length, denoted by $c\tau^*$, and defined as

$$c\tau^* = L_{\mathrm{xy}} \frac{m(B_S^0)}{p_T(D_S^- \ell^+)} = c\tau\frac{1}{K}.$$

4.3 Determination of $\beta\gamma$ Correction

The distribution of the correction factor K must be determined in Monte Carlo. A Monte Carlo generator is used which generates single b quarks according to the $p_T(b)$ calculation of Nason, Dawson, and Ellis [53]. The b quarks form mesons and are then decayed by the CLEO Monte Carlo package(QQ) [54]. The QQ package uses the models of Isgur and others [55] for semi-leptonic decays to determine the lepton and charm momentum spectra.

B_S^0 mesons decay semi-leptonically through D_S^-, D_S^{*-}, and D_S^{**-}. The D_S^{*-} has a mass of 2112 MeV/c^2 [52], slightly greater than the mass of the D_S^-, and decays almost exclusively to a D_S^- and a photon. The D_S^{**-} refers not to a single meson, but rather to the higher mass charmed, strange mesons which complete the semi-leptonic decays of the B_S^0, including non-resonant $\bar{c}s$ states.

As the B_S^0 branching ratios have not been well-measured, in QQ the assumption is made that the semi-leptonic branching ratios of the B_S^0 are similar to the semi-leptonic branching ratios of the B^0. The semi-leptonic branching ratio of the B_S^0, $BR(B_S^0 \to \ell^- X)$, is assumed to be 10.4%, in agreement with the PDG measurement of $BR(B^0 \to \ell^- X)$. It is further assumed that the fraction of D_S^{**-} mesons in B_S^0 semi-leptonic decays is the same as the fraction of D^{**} mesons in B^0 decays. This has been measured by CLEO to be 0.36 ± 0.12 [56]. The fraction of these D_S^{**-} mesons which decay to a D_S^- is 9% in QQ. In summary, the following branching fractions are used in the Monte Carlo:

$$BR(B_S^0 \to D_S^- \ell^- X) = 1.8\%$$

$$BR(B_S^0 \to D_S^{*-} \ell^- X) = 4.9\%$$

$$BR(B_S^0 \to D_S^{**-} \ell^- X) = 3.7\%$$

$$BR(D_S^{**-} \to D_S^- X) = 9\%.$$

For the sample used to determine the K distribution, b quarks are generated with a minimum $p_T(b)$ of 10 GeV/c and then form B_S^0 mesons. The mesons are decayed semi-leptonically through D_S^-, D_S^{*-}, and D_S^{**-}, and those events which include a D_S^- in the decay

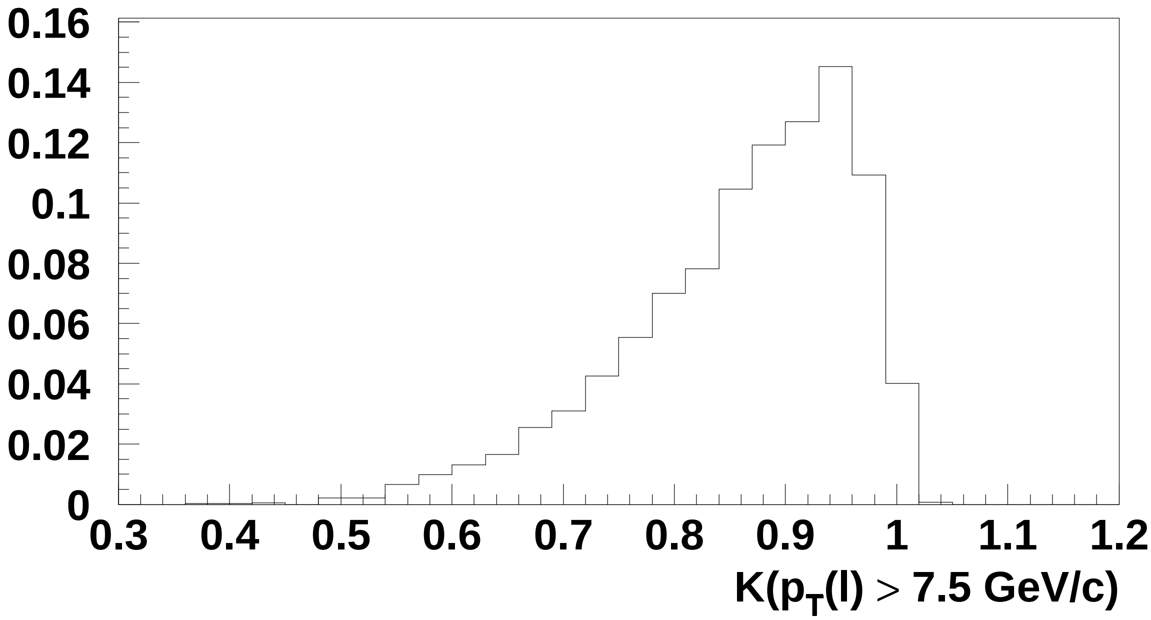


Figure 4.1: Monte Carlo K distribution for the $D_S^- \rightarrow \phi\pi^-$ sample.

are retained. To account for the effect of the trigger, a simple model of the turn-on of the trigger based on the generator level p_T of the lepton is used. The Monte Carlo sample is processed with the standard reconstruction and the same selection requirements are used as in the data. A cut of $p_T(\ell) > 7.5$ $GeV/$c, identical to the requirements of the trigger in Run Ib, is used in generating the final K distribution. The K distribution is shown in Figure 4.1, and has a mean of 0.860 and an RMS of 0.103.

4.4 Fitting Procedure

4.4.1 Signal and Background Samples

The lifetime is obtained by an unbinned maximum-likelihood fit to the observed pseudo-proper decay length distribution. A sample of events with good D_S^- candidates, referred to as the signal sample, is chosen by requiring the reconstructed D_S^- mass to be within the peak region shown in Figure 3.3. The signal sample is defined as the right-sign combinations in the invariant mass region

$$1.944 < m(D_S^-) < 1.994 \; GeV/c^2,$$

representing roughly a 2.5σ window around the nominal D_S^- mass.

In the $D_S^- \to \phi\pi^-$ sample there are 350 events in this region. Since virtually all the D_S^- signal events should be contained in this region, the fraction of signal events is estimated to be 220/350 or 63%. The ratio of signal events to background events in this region is estimated to be 220/130 or about 5/3.

To model the contribution of the combinatorial background events under the mass peak to the shape of the pseudo-proper decay length distribution, a sample of representative background events is taken from events whose reconstructed D_S^- mass is separated from the D_S^- peak. Right-sign combinations in the "sideband" regions

$$1.884 < m(D_S^-) < 1.934 \; GeV/c^2 \qquad \text{and} \qquad 2.004 < m(D_S^-) < 2.054 \; GeV/c^2$$

are used. These regions are chosen so that they are sufficiently separated from the D_S^- peak to avoid any contamination from the signal, and also contamination in the low-mass sideband from $D^- \to \phi\pi^-$ decays. To increase the size of the background sample, wrong-sign combinations in the region

$$1.884 < m(D_S^-) < 2.054 \; GeV/c^2$$

are used as well, since all wrong-sign events are expected to be due to combinatoric background. Figure 4.2 shows the right-sign and wrong-sign D_S^- mass distributions, with the shaded areas showing the regions that make up the background sample.

4.4.2 Likelihood Function

The B_S^0 lifetime and the background shape are determined from a simultaneous fit to both the signal and background samples. The true distribution of the signal is an exponential, however this is smeared by the resolution of the detector and by the mismeasurement of the

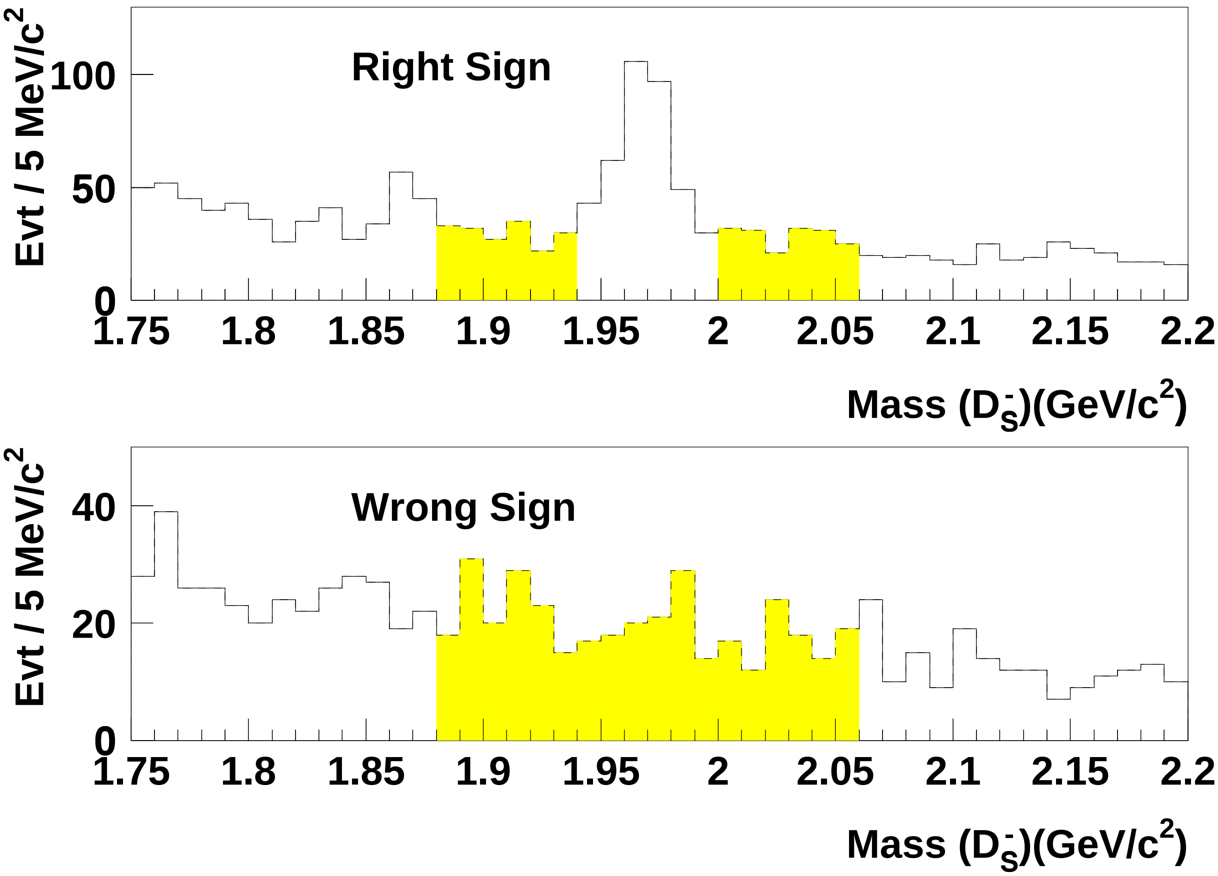


Figure 4.2: D_S^- mass distribution for right-sign(top) and wrong-sign(bottom) combinations, with the shaded areas representing the sideband samples.

B_S^0 momentum. The background distribution has several components. First there is a zero-lifetime component that is smeared by the detector resolution. Second, there is a component with a positive lifetime, due to events containing real b and c decays. Lastly, there is a component with a negative lifetime due to background events which extrapolate behind the primary vertex. In fitting the lifetime, the fractions and lifetimes of each component of the background must also be determined, as well as the resolution of the detector.

The likelihood function then has two parts:

$$L = \prod_{i}^{N_{sig}} [(1 - f_{bkg})\mathcal{F}_{sig}^{i}) + f_{bkg}\mathcal{F}_{bkg}^{i}] \cdot \prod_{i}^{N_{bkg}} \mathcal{F}_{bkg}^{i},$$

where N_{sig} and N_{bkg} are the numbers of events in the signal and background samples, f_{bkg} is the estimated fraction of background events in the signal sample, and $\mathcal{F}_{sig}$ and $\mathcal{F}_{bkg}$ are the probability distributions of the signal and background. In practice the fit minimizes the quantity $\mathcal{L} = -2\ln(L)$.

$\mathcal{F}_{sig}$ consists of an exponential function, convoluted with the correction factor distribution, $H(K)$, convoluted with a gaussian resolution function. It can be written as

$$\mathcal{F}_{sig} = \frac{K}{c\tau} \cdot exp(\frac{-Kx}{c\tau}) \ \otimes \ G(x, s\sigma) \ \otimes \ H(K)$$

where x is the event-by-event pseudo-proper decay length, σ is the error on x, and s is a scale factor to account for an overall under- or over-estimate of the errors. The symbol "$\otimes$" denotes the convolution. The convolution with the gaussian can be performed analytically, but the convolution with the K distribution is performed as a finite sum. (See Appendix A).

$\mathcal{F}_{bkg}$ has three components. There is a gaussian centered at zero, and two exponentials for positive and negative lifetime backgrounds, each convoluted with gaussians. The functional form of $\mathcal{F}_{bkg}$ can be written as

$$\begin{aligned} \mathcal{F}_{bkg} &= (1 - f_+ - f_-)G(x, s\sigma) \\ &+ \frac{f_+}{\lambda_+} exp(\frac{-x}{\lambda_+}) \ \otimes \ G(x, s\sigma) \\ &+ \frac{f_-}{\lambda_-} exp(\frac{x}{\lambda_-}) \ \otimes \ G(x, s\sigma) \end{aligned}$$

where f_+ and f_- are the fractions of positive and negative lifetime backgrounds, and λ_+ and λ_- are the lifetimes of those backgrounds. The values of $c\tau$, s, f_+, f_-, λ_+, and λ_- are all determined in the lifetime fit.

In order to avoid a flat background component which has not been accounted for, the fit is restricted to the region

$$-0.2 \ < \ c\tau^* \ < \ 0.5 \ cm$$

and the probabilities are normalized over the same range.

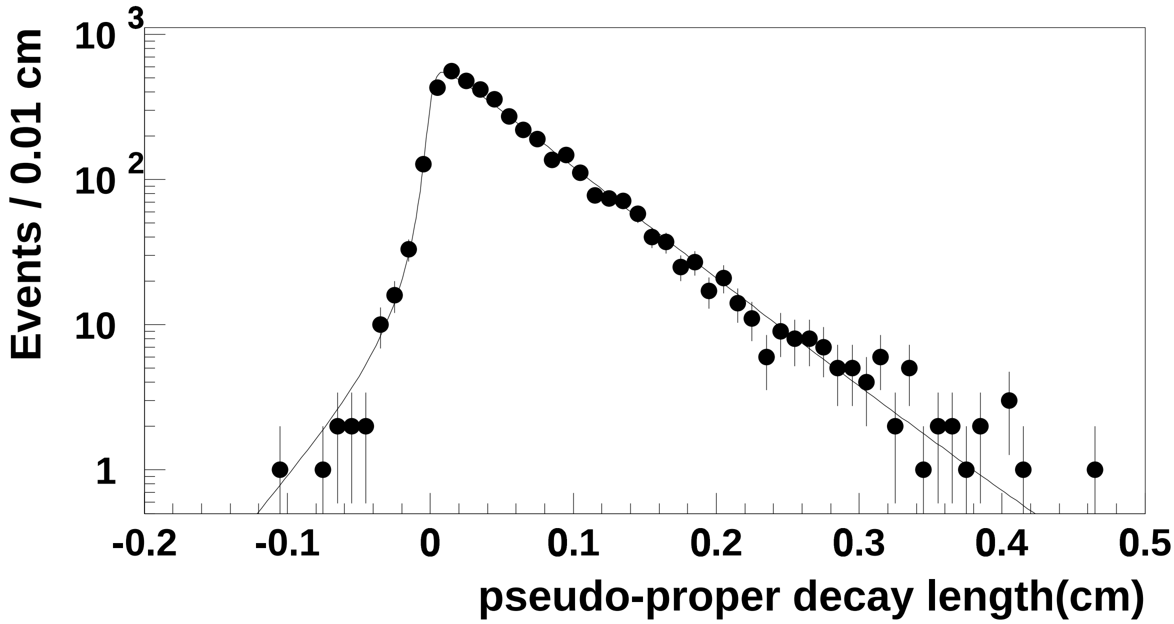


Figure 4.3: Results of the fit to the Monte Carlo B_S^0 pseudo-proper decay length distribution.

In addition, the fraction of background events in the signal sample is constrained to the value determined from the fit to the D_S^- mass distribution. The χ^2 term

$$\frac{(f_{bkg} \; - \; f_{bkg}^{init})^2}{\sigma_{f_{bkg}}^2}$$

is added to $\mathcal{L}$, where f_{bkg} is returned from the fit, f_{bkg}^{init} is the background fraction from the fit to the D_S^- mass distribution, and $\sigma_{f_{bkg}}$ is the error.

4.4.3 Test of Fitting Procedure

The fitting procedure is tested using Monte Carlo events. First the sample described in Section 4.3 is processed using the standard kinematic cuts and the lifetime is determined. The sample has an input lifetime of $c\tau = 465$ μm, and the fit returns $c\tau = 470 \pm 8$ μm, in agreement with the input value. Figure 4.3 shows the pseudo-proper decay length distribution of the Monte Carlo, with the result of the fit superimposed.

As a further test, 1000 samples of 250 events each are generated using the shape of the fitted pseudo-proper decay length distribution from the Monte Carlo sample, corresponding

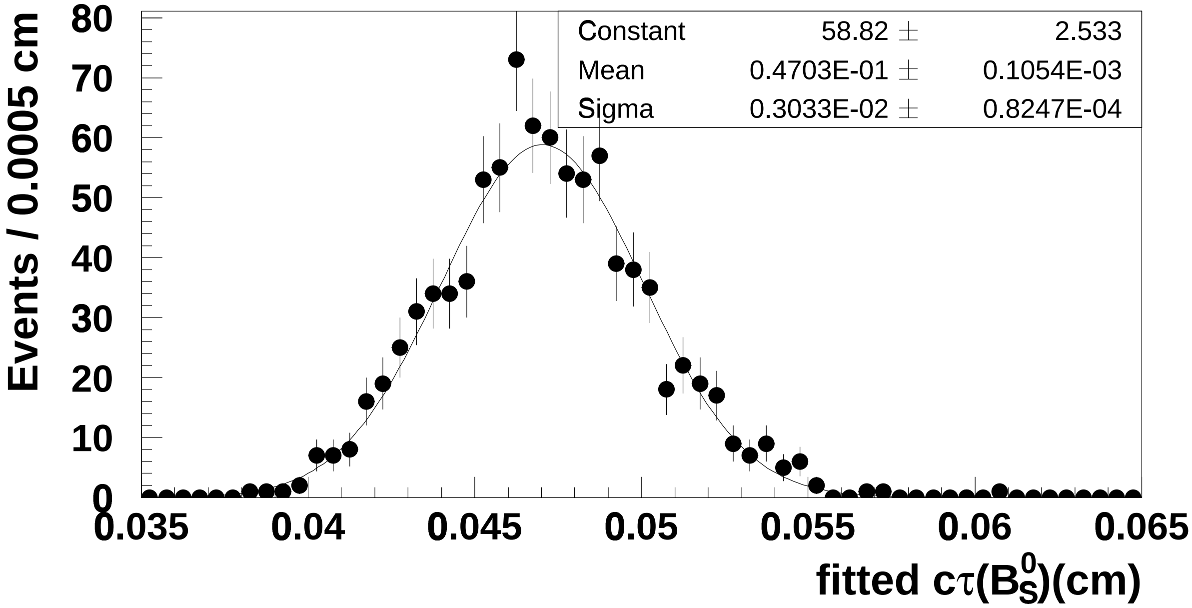


Figure 4.4: Distribution of the fitted $c\tau$ values from the toy Monte Carlo samples with input lifetimes of 470 μm. The results of the gaussian fit are superimposed.

to a lifetime of 470 μm. Each sample is fit and the measured lifetime is recorded. The distribution of fitted lifetimes is fit to a gaussian, and a mean of 470 $\pm$ 1 μm is obtained, in excellent agreement with the input value. Figure 4.4 shows the distribution of the fitted lifetimes with the result of the gaussian fit superimposed.

These two tests demonstrate that the fitting procedure works well and does not introduce a systematic bias in the measurement of the lifetime.

4.4.4 Inclusion of Physics Backgrounds

Up to this point all backgrounds have been assumed to be made up of a real lepton and a fake D_S^- which comes from the combinatoric background. However, there are physics processes which will produce a real lepton and a real D_S^-, where neither comes from the semi-leptonic decay of a B_S^0 meson. Such events will end up in the signal sample, but will not have the correct pseudo-proper decay length distribution.

There are three sources of these backgrounds that must be considered. The first is the contribution from the decays $B^0 \to D_S^{(*)+} D^{(*)-} X$ and $B^+ \to D_S^{(*)+} D^{(*)0} X$, with the $D^{(*)-}$ or $D^{(*)0}$ decaying semi-leptonically. These events can be reconstructed as signal events, though the momentum spectrum of the lepton coming from the decay of the $D^{(*)}$ should be much softer since it comes from the secondary decay of a charm hadron. In order to estimate the contribution of these processes to the signal, Monte Carlo events from these decays are generated and reconstructed with the standard analysis method.

Over 100 million $B \to D_S^{(*)} D^{(*)}$ decays are generated, using the same Monte Carlo as in the determination of the K distribution, again with a minimum $p_T(b)$ of 10 $GeV/$c. All decays which include both a D_S^- and a D are used, with their relative branching fractions as given in QQ.

First, the efficiency for reconstructing these decays relative to the efficiency for reconstructing the signal events must be determined. Since the lepton from the decay of the D is much softer, the efficiency for finding the lepton in these decays is much lower. The ratio of efficiencies for the reconstructing the lepton is found to be

$$\frac{\epsilon_\ell(B \to D_S^{(*)} D^{(*)} X)}{\epsilon_\ell(B_S^0 \to D_S^- \ell^+ X)} = 0.0050.$$

Once a good lepton has been found, the ratio of efficiencies for reconstructing the D_S^- is fairly close. This ratio is measured to be

$$\frac{\epsilon_{D_S^-}(B \to D_S^{(*)} D^{(*)} X)}{\epsilon_{D_S^-}(B_S^0 \to D_S^- \ell^+ X)} = 1.20.$$

The combined ratio of efficiencies is then

$$\epsilon_{rel} = \frac{\epsilon(B \to D_S^{(*)} D^{(*)} X)}{\epsilon(B_S^0 \to D_S^- \ell^+ X)} = 0.0060.$$

The fraction of $B \to D_S^{(*)} D^{(*)}$ decays that make it into the signal sample is given by the following expression:

$$f_{D_S D} = \epsilon_{rel} \cdot \frac{f_u + f_d}{f_S} \cdot \frac{BR(B \to D_S^- X) \cdot BR(D \to \ell^+ X)}{BR(B_S^0 \to D_S^- \ell^+ X)}$$

where f_u, f_d, and f_S are the fractions of B_u, B_d, and B_S out of all B hadrons (including Λ_b). Note that $BR(B_S^0 \to D_S^- \ell^+ X)$ refers to all semi-leptonic B_S^0 decays to a D_S^-, including both direct decays and those that go through D_S^{*-} and D_S^{**-}. That is,

$$\begin{aligned} BR(B_S^0 \to D_S^- \ell^+ X) &= BR(B_S^0 \to D_S^- \ell^+ \nu) \\ &+ BR(B_S^0 \to D_S^{*-} \ell^+ \nu) \\ &+ BR(B_S^0 \to D_S^{**-} \ell^+ \nu) \cdot BR(D_S^{**-} \to D_S^- X) \end{aligned}$$

As explained in Section 4.3,

$$\frac{BR(B_S^0 \to D_S^{(*)-} \ell^+ X)}{BR(B_S^0 \to \ell^+ X)} = 0.64,$$

with the other 36% decaying to $D_S^{**-}\ell^+ X$, and 9% of D_S^{**} decaying to a D_S, resulting in

$$\frac{BR(B_S^0 \to D_S^{**-} \ell^+ X) \cdot BR(D_S^{**-} \to D_S^- X)}{BR(B_S^0 \to \ell^+ X)} = 0.03.$$

Summing these two contributions gives

$$\frac{BR(B_S^0 \to D_S^- \ell^+ X)}{BR(B_S^0 \to \ell^+ X)} = 0.67.$$

Following the assumption in QQ that $BR(B_S^0 \to \ell^+ X) = BR(B^0 \to \ell^+ X)$ (see Section 4.3), where the PDG value of $BR(B^0 \to \ell^+ X)$=10.4% is used, gives

$$BR(B_S^0 \to D_S^- \ell^+ X) = 0.07.$$

For $BR(D \to \ell^+ X)$, the average of $BR(D^0 \to \ell^+ X)$ and $BR(D^+ \to \ell^+ X)$ is used. The PDG values are 7.65% and 17.2%, respectively, which gives

$$BR(D \to \ell^+ X) = 0.125.$$

Using the following PDG values:

$$f_u = f_d = 0.378,$$

$$f_S = 0.112,$$

$$BR(B \to D_S^- X) = 8.9\%,$$

and the results from above, the estimated fraction of events in the signal sample coming from $B \to D_S^{(*)} D^{(*)} X$ decays is

$$f_{D_S D} = 0.0063.$$

The second physics background is the decay $B_S^0 \to D_S^{(*)-} D_S^{(*)+}$, with one $D_S^{(*)}$ decaying semi-leptonically. This background is treated in the same manner as for $B \to D_S^{(*)} D^{(*)}$ decays. The fraction of these events that make it into the signal sample is given by the expression:

$$f_{D_S D_S} = \epsilon_{rel} \cdot \frac{BR(B_S^0 \to D_S^{(*)-} D_S^{(*)+} X) \cdot BR(D_S^+ \to \ell^+ X) \cdot 2}{BR(B_S^0 \to D_S^- \ell^+ X)}$$

The ratio of efficiencies for reconstructing the lepton is measured to be

$$\frac{\epsilon_\ell(B_S^0 \to D_S^{(*)-} D_S^{(*)+} X)}{\epsilon_\ell(B_S^0 \to D_S^- \ell^+ X)} = 0.0095$$

and the ratio of efficiencies for reconstructing the D_S^- is measured to be

$$\frac{\epsilon_{D_S^-}(B_S^0 \to D_S^{(*)-} D_S^{(*)+} X)}{\epsilon_{D_S^-}(B_S^0 \to D_S^- \ell^+ X)} = 1.34.$$

The combined ratio of efficiencies is then

$$\epsilon_{rel} = \frac{\epsilon(B_S^0 \to D_S^{(*)-} D_S^{(*)+} X)}{\epsilon(B_S^0 \to D_S^- \ell^+ X)} = 0.0127.$$

To calculate $f_{D_S D_S}$, the QQ values for $BR(B_S^0 \to D_S^{(*)-} D_S^{(*)+} X) = 0.045$ and $BR(D_S^- \to \ell^- X) = 0.079$ are used. The denominator remains the same as in the calculation of $f_{D_S D}$, so the fraction of events in the signal sample coming from $B_S^0 \to D_S^{(*)-} D_S^{(*)+} X$ decays is estimated to be

$$f_{D_S D_S} = 0.0013.$$

To include the contribution of these events in the lifetime fit, an effective lifetime of these events must be found. A term similar to the signal term is then added to the likelihood,

$$f_{D_S D} \cdot \frac{K}{c\tau(D_S D)} \cdot exp(\frac{-Kx}{c\tau(D_S D)}) \otimes G(x, s\sigma).$$

Fitting for the effective lifetime of both of these samples, Figure 4.5(a) shows the pseudo-proper decay length distribution from the $B \to D_S^{(*)} D^{(*)} X$ Monte carlo, with the results of the fit overlaid. Figure 4.5(b) shows the same distribution for the $B_S^0 \to D_S^{(*)-} D_S^{(*)+} X$ Monte Carlo, with the results of the fit overlaid. For these samples we find

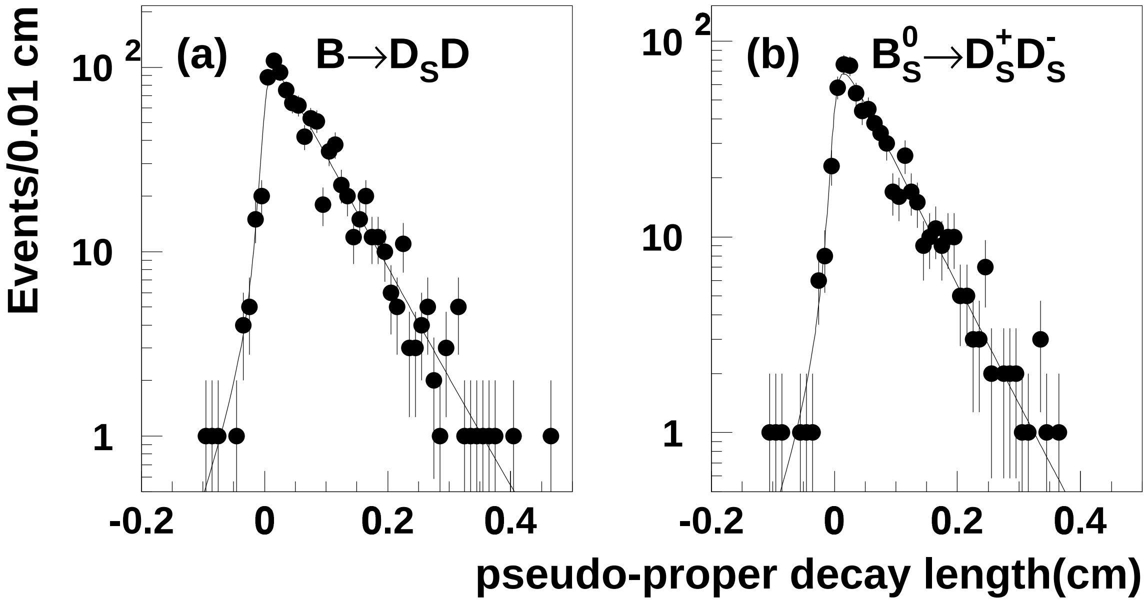


Figure 4.5: Results of the fit to the Monte Carlo pseudo-proper decay length distributions for the effective lifetimes in the (a) $B \to D_S^{(*)} D^{(*)} X$ and (b) $B_S^0 \to D_S^{(*)-} D_S^{(*)+} X$ samples.

$$c\tau(D_S D) = 590 \ \pm 21 \ \mu m.$$

and

$$c\tau(D_S D_S) = 577 \ \pm 24 \ \mu m.$$

The final physics background is the process $B \to D_S^{(*)+} K \ell^- X$. This process has never been observed, and an upper limit of 1.2 % on the corresponding branching ratio has been measured by ARGUS [57]. Because this process has never been observed, the fraction of these events in the sample is assumed to be 0. Using the upper limit on the branching ratio results in an upper limit on the fraction of these events in the signal of 0.03 [58]. This upper limit is included in the systematic uncertainties.

4.5 Lifetime Results from $D_S^- \to \phi\pi^-$ Sample

Using the values listed above for the physics background contributions, a simultaneous fit to the signal and background samples is performed. The result

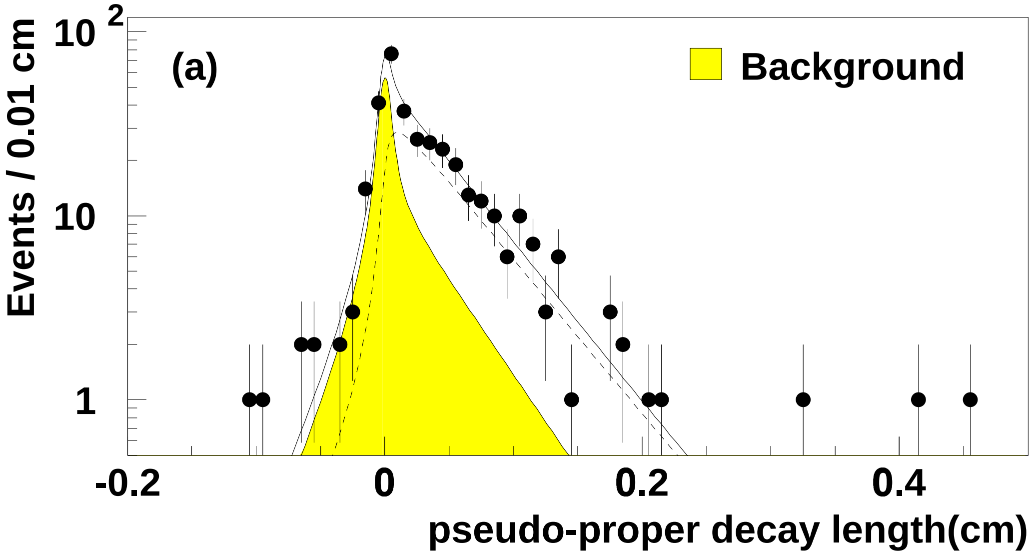
$$c\tau(B_S^0) \;=\; 427\,^{+41}_{-37}\;\mu m$$

is obtained, where the errors shown are statistical only. Figure 4.6(a) shows the pseudo-proper decay length distribution for the signal sample with the fit results superimposed. The shaded curve represents the sum of the background probability function over the events in the signal sample. The signal contribution is represented by the dashed curve. Figure 4.6(b) shows the pseudo-proper decay length distribution for the background sample with the fit result superimposed.

As a cross check of the fitting method, a fit for the lifetime of the D_S^- is performed. Since the D_S^- decay is fully reconstructed, its $\beta\gamma$ is known and the convolution with the K distribution in the fit does not apply. The result of the fit is

$$c\tau(D_S^-) \;=\; 136\,^{+17}_{-15}\;\mu m,$$

which agrees well with the PDG value of 140 μm. Figure 4.7(a) shows the proper decay length distribution for the signal region, with the results of the fit overlaid. The dashed curve represents the signal contribution, while the shaded curve shows the background contribution to the signal region. Figure 4.7(b) shows the proper decay length distribution for the sideband sample with the fit result superimposed.



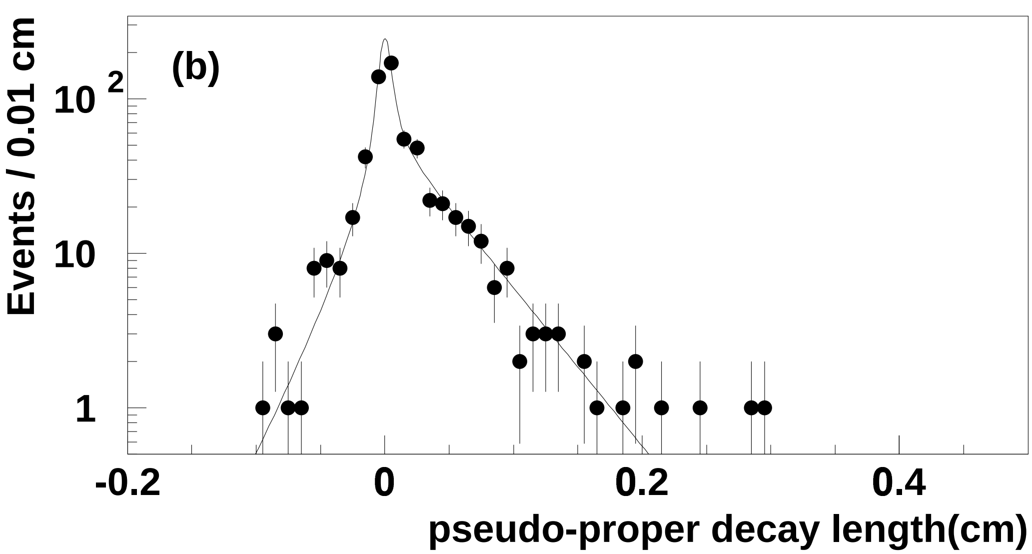


Figure 4.6: Pseudo-proper decay length distributions for (a) signal and (b) background in the $\phi\pi^-$ sample. The shaded histogram in (a) represents the background contribution to the signal sample. The signal contribution is represented by the dashed curve.

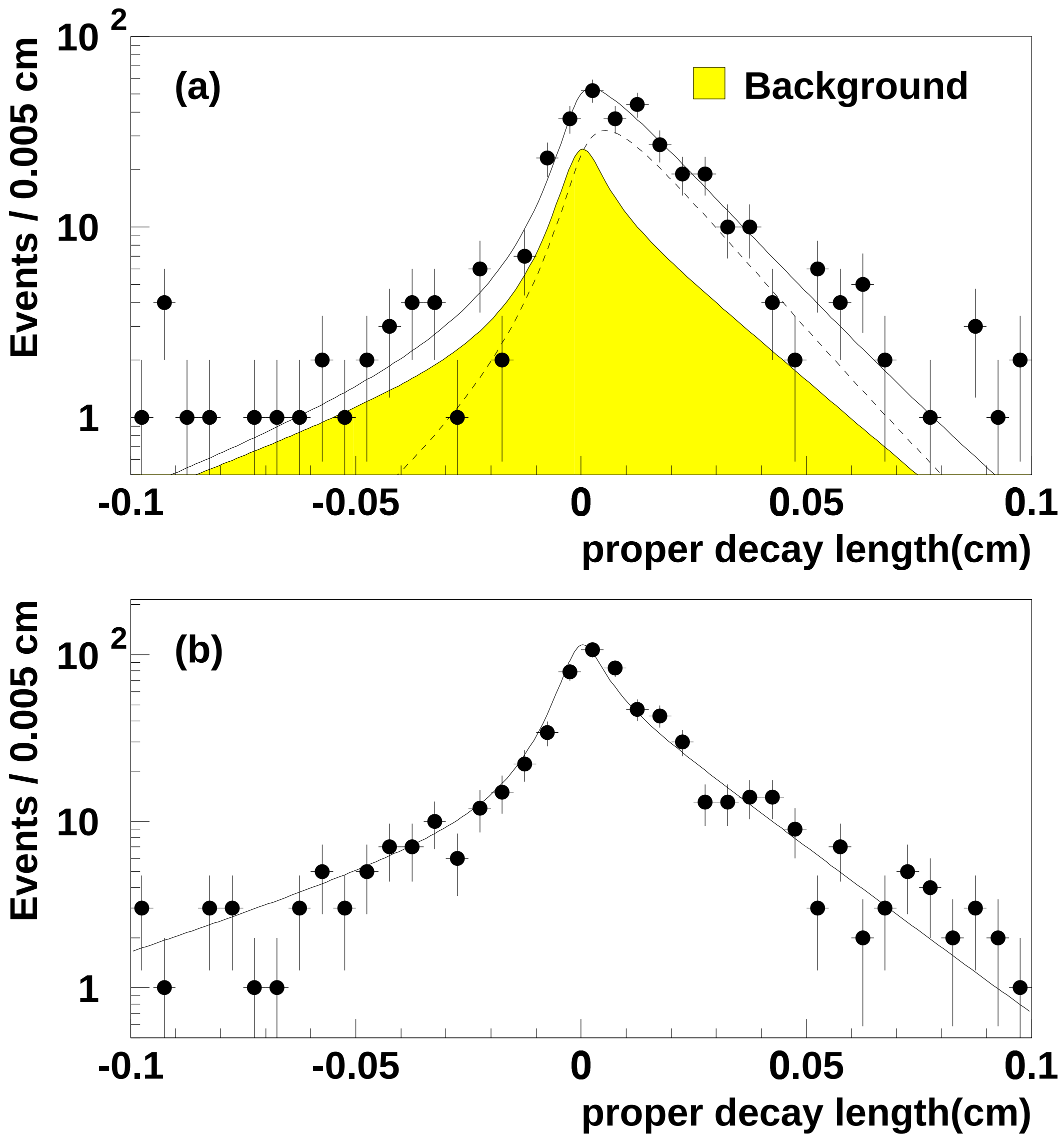


Figure 4.7: Results of the fit to the D_S^- proper decay length distribution in the $\phi\pi^-$ sample for (a) signal region and (b) sideband region. The shaded histogram in (a) represents the background contribution to the signal sample. The signal contribution is represented by the dashed curve.

Chapter 5

Addition of $D_S^- \to K^{*0}K^-$ and $D_S^- \to K_s^0K^-$ Decays

To increase the size of the data sample and increase the precision of the measurement of $\tau(B_S^0)$, decays of the D_S^- meson to $K^{*0}K^-$ and $K_s^0K^-$ are reconstructed. This chapter first explains the expected yield of D_S^- events from these decays and then describes the selection of D_S^- candidates in these channels. There is a significant background in these channels from $D^- \to K^{*0}\pi^-$ and $D^- \to K_s^0\pi^-$. This background is discussed, as well as ways to reduce it. Finally the fraction of real D_S^- decays in the sample is measured.

5.1 Expected Data Sample

The first decay of the D_S^- meson that is considered is $D_S^- \to K^{*0}K^-$ where $K^{*0} \to K^+\pi^-$. Since

$$BR(D_S^- \to \phi\pi^-) \approx BR(D_S^- \to K^{*0}K^-)$$

and

$$\frac{BR(\phi \to K^+K^-)}{BR(K^{*0} \to K^+\pi^-)} \approx \frac{0.49}{0.67},$$

there should be approximately as many $D_S^- \to K^{*0}K^-$ events as there are $D_S^- \to \phi\pi^-$. However, the combinatoric background in this channel will be higher due to the broader natural width of the K^{*0}. In addition there will be some contamination from $D^- \to K^{*0}\pi^-$ decays, where the π^- has been misidentified as a K^-.

The second decay of the D_S^- meson that is considered is the decay $D_S^- \to K^0 K^-$ where $K^0 \to K_s^0$ and $K_s^0 \to \pi^+\pi^-$. Since

$$BR(D_S^- \to \phi\pi^-) \approx BR(D_S^- \to K^0 K^-)$$

and

$$BR(\phi \to K^+K^-) \approx 1.5 \cdot BR(K^0 \to K_s^0 \to \pi^+\pi^-)$$

there should be approximately 1/2 - 1/3 as many $D_S^- \to K_s^0 K^-$ events as $D_S^- \to \phi\pi^-$, depending on the K_s^0 reconstruction efficiency. In this D_S^- decay mode, the combinatoric background will not be as large as in the $D_S^- \to K^{*0}K^-$ sample, but there will again be the problem of contamination from $D^- \to K_s^0\pi^-$ decays.

5.2 Event Selection for $D_S^- \to K^{*0}K^-$

The reconstruction of $D_S^- \to K^{*0}K^-$ decays is very similar to the reconstruction of $D_S^- \to \phi\pi^-$ decays described in Section 3.4. Again the reconstruction begins with the candidate lepton. Nearby pairs of oppositely charged tracks are chosen to reconstruct the K^{*0} candidate. The first track is assigned the K mass and the second is assigned the π mass, with the K track required to have the same charge as the lepton. If the invariant mass of the pair of tracks corresponds to the K^{*0} mass, a third track is added, assigned the K mass. All three tracks are required to intersect at a common vertex, as in the $D_S^- \to \phi\pi^-$ reconstruction, and no charge correlation is required between this track and the other candidate tracks. The B_S^0 decay vertex is found in the same way as in the $D_S^- \to \phi\pi^-$ reconstruction.

The lepton selection is identical for this sample. Selection of D_S^- candidates remains largely the same here as in Section 3.5. The following requirements have changed:

- $p_T(\pi) > 0.4\ GeV/c$
- $|\cos\psi| > 0.5$, where ψ is the helicity angle between the K^+ and D_S^- in the K^{*0} rest frame

- $|\mathrm{m}(K^+\pi^-)$ - $0.896| < 0.040$ GeV/c^2, where 0.896 is the nominal PDG [52] K^{*0} mass.

The differences are that the $p_T(\pi)$ cut has been loosened, while the $|\cos\psi|$ cut has been tightened. To further reduce combinatorial background, a track-based isolation variable is used. The quantity I_{Trk} is defined as

$$I_{Trk} = \frac{\sum\limits_{D_S^- \ell^+} p_T}{\sum\limits_{Cone} p_T}$$

for a cone of $\Delta R < 1.0$ in η-ϕ space. Events are required to have $I_{Trk} > 0.6$. Lastly, Level$(K/\pi) > 0.01$ is required for the K^+ and π^- from the K^{*0}, while Level$(K) > 0.10$ is required for the K^- from the D_S^-.

In this sample events where the lepton and the K^- from the D_S^- have opposite charges are referred to as "right-sign" events, while combinations where they have the same charge are called "wrong-sign" events. The right-sign D_S^- mass distribution for the $K^{*0}K^-$ sample is shown in Figure 5.1 with the fit result overlaid. The shaded histogram shows the wrong-sign D_S^- mass spectrum. The right-sign mass distribution is again fit with a gaussian to describe the signal and a first-order polynomial to describe the combinatoric background. The second gaussian for $D^- \to K^{*0}K^-$ decays is not included as it can not be observed above the background. The fit yields 179 $\pm$ 35 events and returns a D_S^- mass of 1969 $\pm$ 2 MeV/c^2, which is in good agreement with the PDG D_S^- mass of 1969 MeV/c^2. The width is 10.5 $\pm$ 2.6 MeV/c^2, which agrees with the width found in the $D_S^- \to \phi\pi^-$ sample.

5.3 Event Selection for $D_S^- \to K_s^0 K^-$

The reconstruction of this decay is slightly different from the method decribed for the previous decay modes. A schematic diagram of the decay is shown in Figure 5.2. Because of the lifetime of the K_s^0, the system now has three decay vertices. The reconstruction begins with oppositely charged pairs of tracks near the lepton. Each track is assigned the π mass and the pair of tracks alone is vertexed to form the K_s^0. The mass of the reconstructed K_s^0

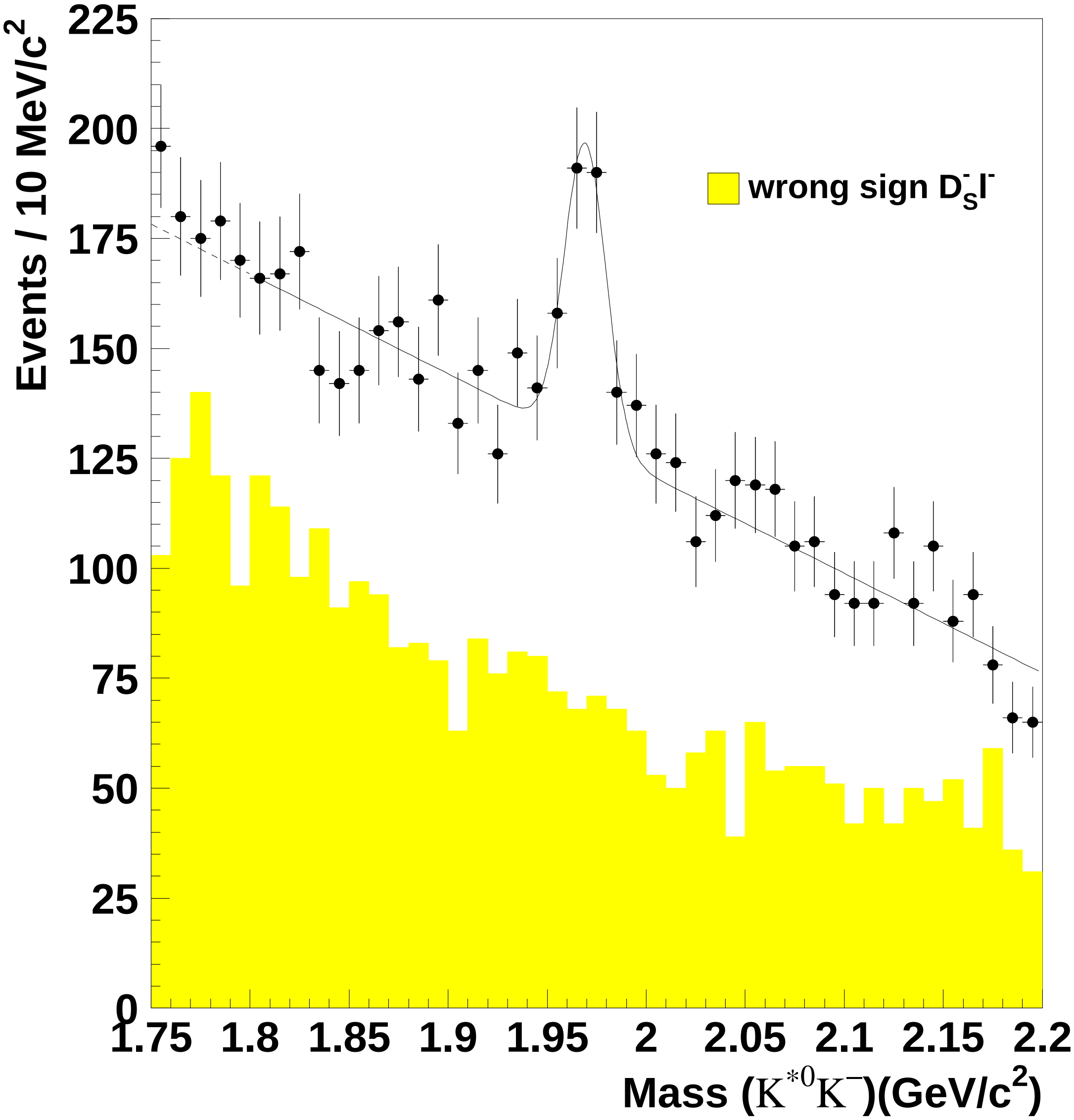


Figure 5.1: Right-sign $K^{*0}K^-$ mass distribution with the results of the fit superimposed. The shaded histogram shows the wrong-sign distribution.

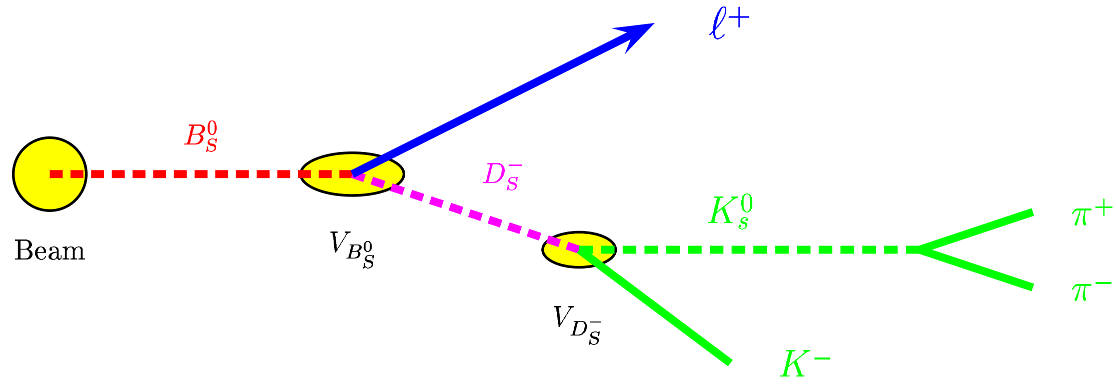


Figure 5.2: Sketch of the decay mode $B_S^0 \to D_S^- \ell^+ \nu$, $D_S^- \to K_s^0 K^-$.

is constrained to its PDG value and the K_s^0 is extrapolated back to intersect with the K^- to find the D_S^- vertex. The reconstructed D_S^- is then extrapolated back to intersect with the lepton to find the B_S^0 vertex.

Lepton selection is unchanged, and the D_S^- selection requirements are the same as described in Section 5.2 with the following exceptions:

- $\Delta R(K) < 0.8$, $\Delta R(\pi) < 1.0$
- $|m(\pi^+\pi^-) - 0.496|/\sigma < 5.0$, where 0.496 is the nominal PDG K_s^0 mass
- $L_{\mathrm{xy}}(K_s^0)/\sigma > 3.0$.

Due to the long lifetime of the K_s^0 ($c\tau = 2.6762$ cm [52]), in many events the K_s^0 will decay outside the SVX. In the $D_S^- \to \phi\pi^-$ and $D_S^- \to K^{*0}K^-$ samples all three tracks from the D_S^- are required to be SVX tracks. This would be inefficient for $D_S^- \to K_s^0 K^-$ decays. However, events that have two CTC tracks will have a different resolution than those with all SVX tracks. For these reasons events where the K_s^0 is reconstructed in the SVX and events where the K_s^0 is reconstructed in the CTC must be treated differently.

For the "SVX" sample, those events where the K_s^0 decay vertex is within the SVX, the requirement

- $L_{\rm xy}(D_S^{PV}) + L_{\rm xy}(K_s^0) < 8.0$ cm

is applied first to effectively require that the position of the K_s^0 vertex is inside the outer layer of the SVX. The requirements are then the same as in Section 5.2 :

- All SVX tracks with at least 3 SVX hits
- χ^2_{SVX}/hit < 6.0
- $L_{\rm xy}(D_S^{PV}) > 0.0$.

For the "CTC" sample, those events where the K_s^0 is reconstructed outside the SVX, the following requirements are made:

- $L_{\rm xy}(D_S^{PV}) + L_{\rm xy}(K_s^0) > 8.0$ cm
- K^- must be SVX track, ≥ 3 SVX hits, χ^2_{SVX}/hit < 6.0
- π's must be either 2-hit SVX track or CTC track.
- CTC track fit used for SVX tracks with only 2 SVX hits.

Figures 5.3(a) and (b) show the distributions of the error on the B_S^0 pseudo-proper decay length for (a) the "SVX" events and (b) the "CTC" events. The mean error on the pseudo-proper decay length for the SVX sample is 103 μm, while the mean error for the CTC sample is 462 μm. Figures 5.3(c) and (d) show the distributions of the error on the D_S^- proper decay length for (c) the "SVX" events and (d) the "CTC" events. In the case of the D_S^- proper decay length, the mean errors are 132 μm and 584 μm. The errors are smaller on the B_S^0 pseudo-proper decay length because in that case the lepton is an additional SVX track which improves the resolution. In both cases the errors are more than 4 times larger for events with the K_s^0 reconstructed outside the SVX. To include these events in the fit would require treating them as a separate sample, with a second scale factor to account for the larger errors on the CTC events. Because the errors on the CTC K_s^0 events are so much larger, and because of the complications in the fit which would be necessary to properly

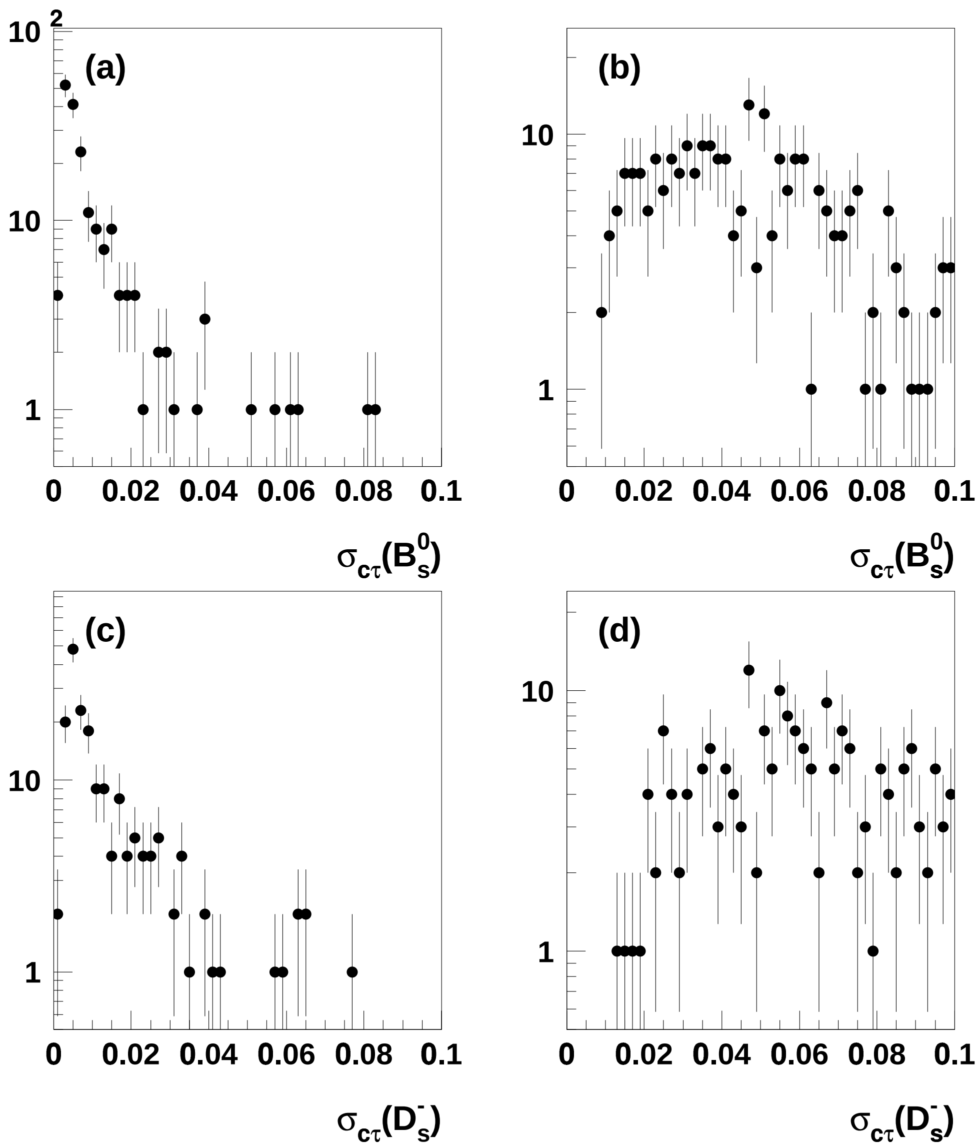


Figure 5.3: Distribution of errors on the reconstructed (pseudo-)proper decay length for: the B_S^0 for events with the K_s^0 reconstructed in the (a)SVX and (b)CTC, and for the D_S^- when the K_s^0 is reconstructed in the (c)SVX and (d)CTC.

include these events, only events with all tracks reconstructed in the SVX are used for the B_S^0 lifetime fit.

Figure 5.4 shows the right-sign D_S^- mass distribution for the final $K_s^0K^-$ sample, where right-sign and wrong-sign events are defined as before. The shaded histogram shows the wrong sign D_S^- mass distribution. The fit to the right-sign mass distribution yields 41 $\pm$ 14 events. The measured D_S^- mass is 1969 $\pm$ 5 MeV/c^2, which agrees with both the previous D_S^- mass distributions and the PDG value. The width is 12.5 $\pm$ 3.8 MeV/c^2, also in agreement with the previous fits to the D_S^- mass distributions.

5.4 Background from $D^- \to K^{*0}\pi^-$ and $D^- \to K_s^0\pi^-$

A significant background to the D_S^- signal will be events where a B^0 meson decays $B^0 \to D^-\ell^+X$ with $D^- \to K^{*0}\pi^-$ or $D^- \to K_s^0\pi^-$. Since the particle identification at CDF does not distinguish well between K's and π's in this momentum range, the π^- can easily be misidentified as a K^-. Such a misassignment of the particle would result in the same signature as a $B_S^0 \to D_S^-\ell^+X$, $D_S^- \to K^{*0}K^-$ or $D_S^- \to K_s^0K^-$ decay. This is referred to as the "D^- reflection".

Figure 5.5(a) shows the effect of misidentifying the π^- as a K^- in a Monte Carlo sample of $B^0 \to D^-\ell^+X$ decays with $D^- \to K^{*0}\pi^-$. The shaded histogram shows the D^- mass distribution when the π^- is correctly identified. The unshaded histogram shows the D_S^- mass distribution when the π^- has been incorrectly assigned the K^- mass.

Figure 5.5(b) shows the similar effect in a Monte Carlo sample of $B_S^0 \to D_S^-\ell^+X$ decays with $D_S^- \to K^{*0}K^-$. Here the shaded histogram shows the D_S^- mass distribution when the K^- is correctly identified while the unshaded histogram shows the D^- mass distribution when the K^- is incorrectly assigned the π^- mass.

From Figure 5.5(a) it can be seen that the reflected D^- mass distribution is peaked in the region of the D_S^- signal, but is smeared out and extends to well above the signal region. About 35% of the D^- reflection is expected to overlap with a $\pm 3\sigma$ mass window around the D_S^- peak. Similarly from Figure 5.5(b) it can be seen that the reflected D_S^- distribution will

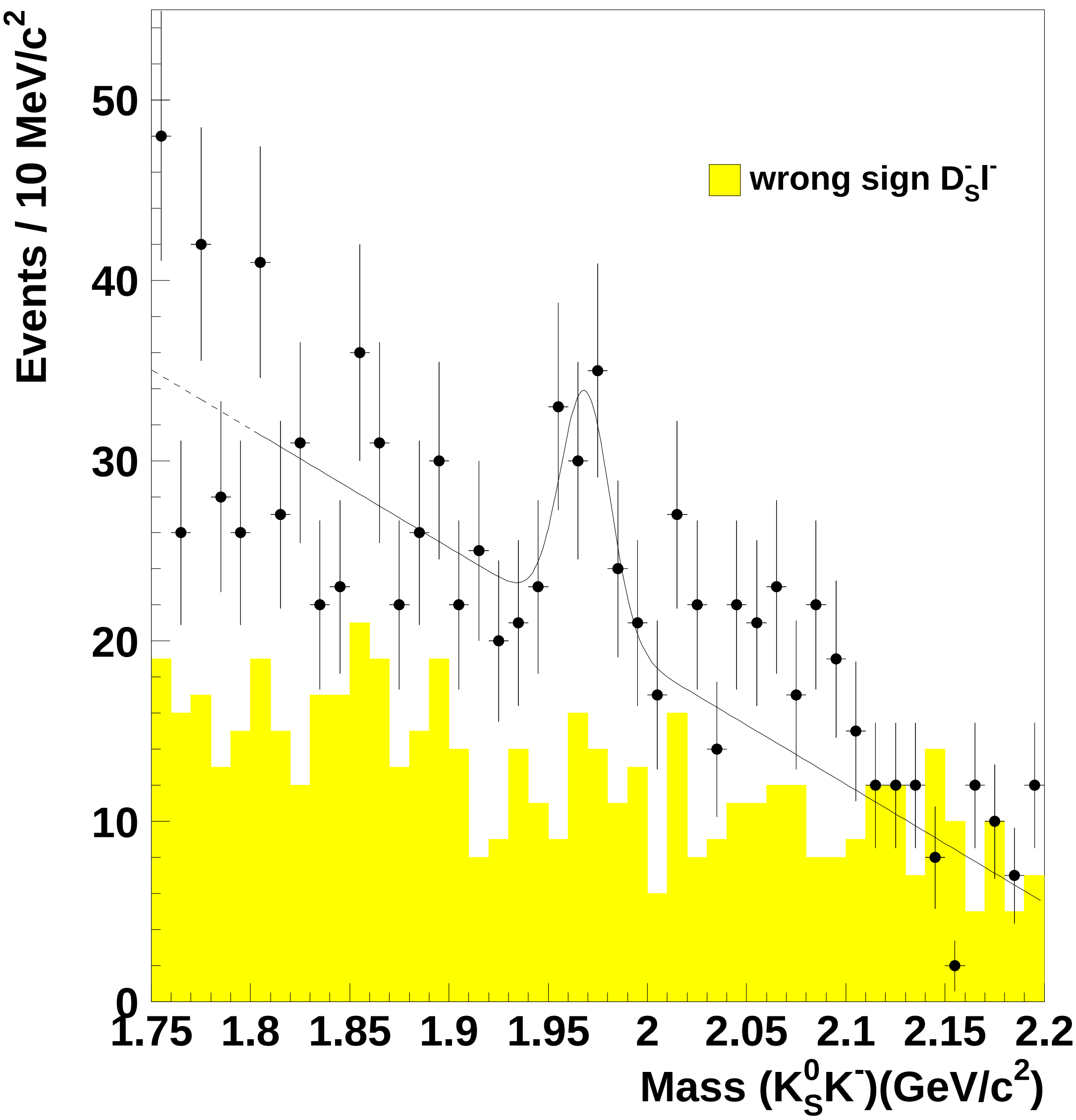


Figure 5.4: Right-sign $K_s^0 K^-$ mass distribution with the results of the fit superimposed. The shaded histogram shows the wrong-sign distribution.

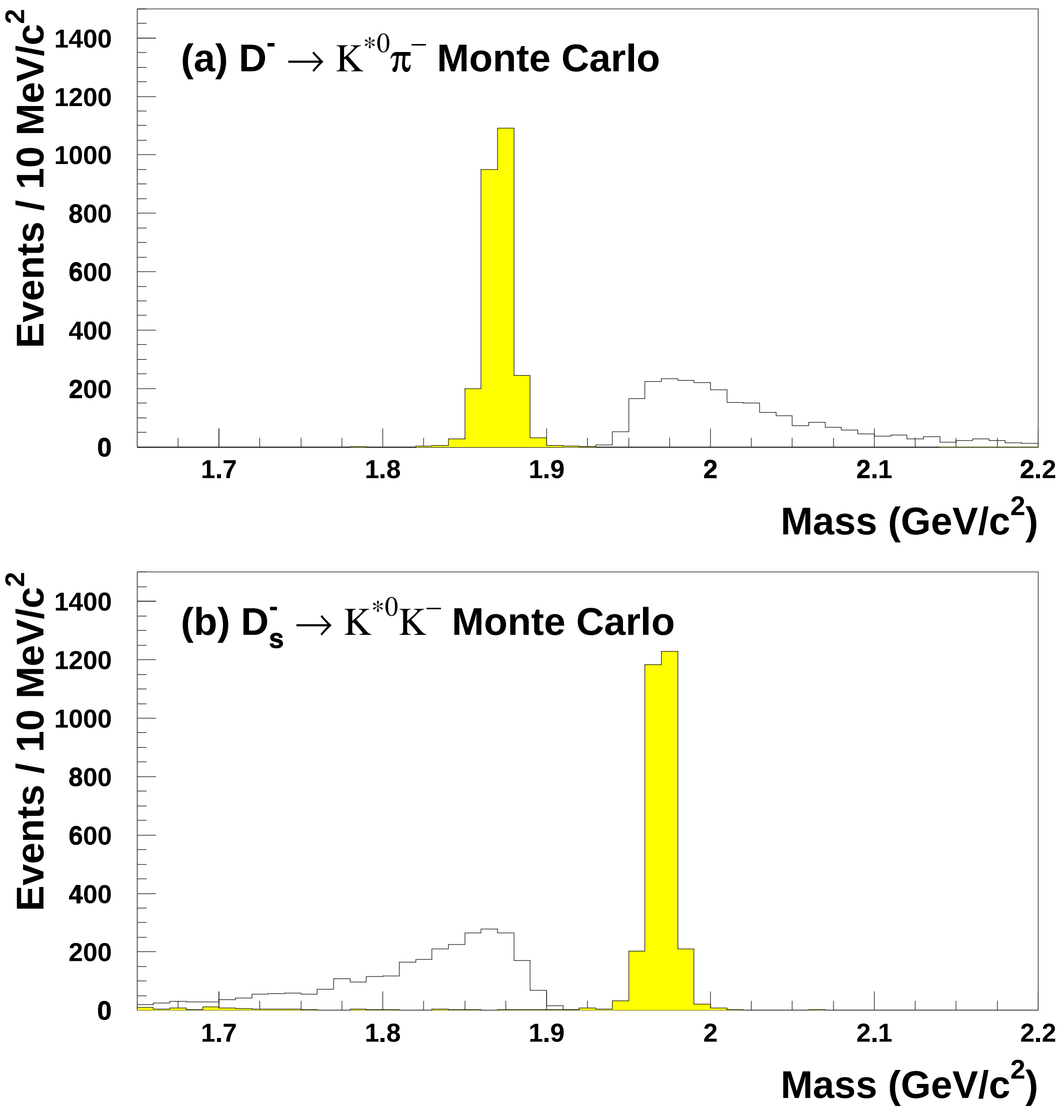


Figure 5.5: Illustration of the D^- reflection: (a) for $D^- \to K^{*0}\pi^-$ Monte Carlo events and (b) for $D_S^- \to K^{*0}K^-$ Monte Carlo events.

peak in the D^- signal region. It is clear that information from the reconstructed D_S^- and D^- mass distributions alone will not be sufficient to discriminate between the D_S^- signal and the D^- reflection.

5.4.1 Attempts to Reduce the D^- Reflection

Though the particle identification is not sufficient to remove all of the background from the D^- reflection, a tighter dE/dx cut is applied on the K^- in the reconstruction of $D_S^- \rightarrow K^{*0}K^-$ and $D_S^- \rightarrow K_s^0 K^-$. The requirement is Level(K) > 0.10 for the K^- from the D_S^- compared to a cut of 0.01 on the tracks from the K^{*0} or K_s^0. However this requirement is not sufficient to reduce the D^- reflection to a negligible amount.

To reduce the remaining D^- reflection, first a mass veto is considered. The simplest solution would be to reconstruct all events as both D_S^- and D^- and veto those events whose reconstructed D^- mass lies within $\pm 3\sigma$ of the PDG D^- mass. From Monte Carlo studies it is estimated that about 35% of real D_S^- events would be removed by this selection.

Figure 5.6 shows the $D_S^- \rightarrow K^{*0}K^-$ mass distribution before and after a mass veto. The mass distribution after the veto has been drastically sculpted, and it would be very difficult to estimate the remaining D_S^- signal. The number of signal events from the mass fit is used to determine the background fraction in the signal sample, which is a constraint in the B_S^0 lifetime fit. Vetoing on the D^- mass would render that constraint unusable and therefore a D^- mass veto is not used.

Another way to reduce the remaining D^- reflection would be to cut on the proper decay length of the reconstructed D. Since the D^- lifetime ($c\tau = 317\mu$m [52]) is more than twice as long as the D_S^- lifetime ($c\tau = 140\mu$m [52]), it might be possible to remove D^- events with a proper decay length cut. Table 5.1 shows the efficiency of a cut on the reconstructed D proper decay length for D_S^- and D^- Monte Carlo events. For cut values which would remove most of the D^- background, the cut is also inefficient for real D_S^- decays. In addition, such a cut could bias the B_S^0 lifetime measurement. For this reason no cut on the D_S^- proper decay length is made.

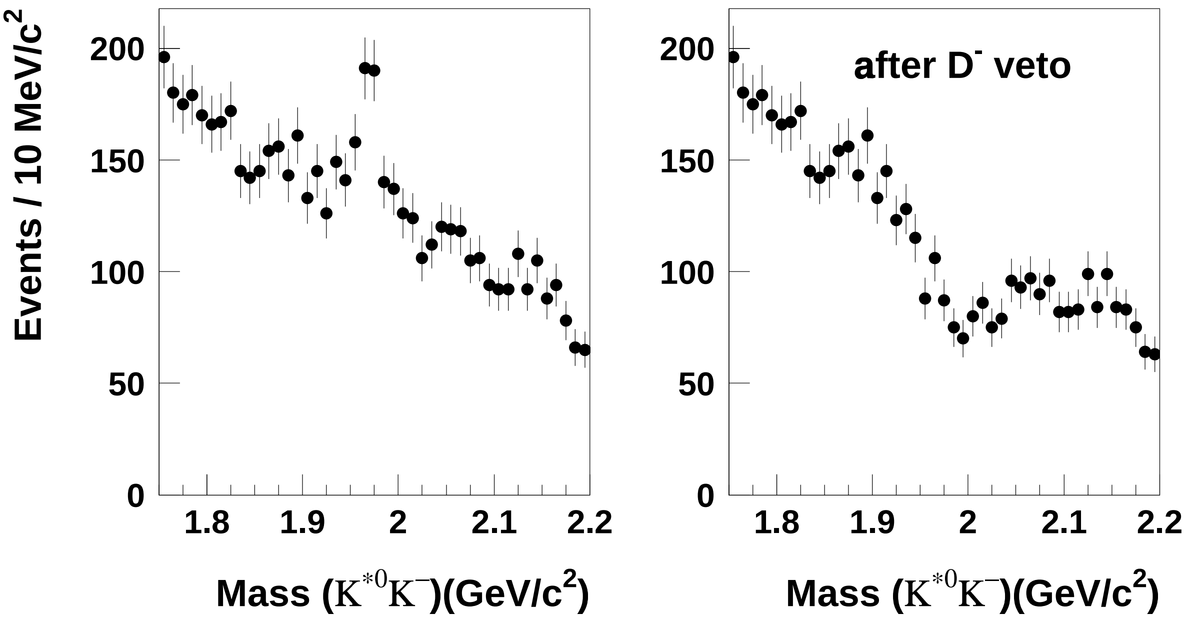


Figure 5.6: Illustration of the effect of a D^- veto on the $K^{*0}K^-$ mass distribution.

proper decay length cut(mm)	D_S^- Eff.	D^- Eff.
1.0	0.98	0.94
0.5	0.93	0.69
0.3	0.81	0.55
0.2	0.66	0.41

Table 5.1: Efficiencies for a cut on the proper decay length for D_S^- and D^-.

Since the remaining D^- component in the $D_S^- \to K^{*0}K^-$ and $D_S^- \to K_s^0K^-$ samples can not be reduced to a negligible level without also drastically reducing the size of the real D_S^- signal, it is better to measure the size of the remaining D^- component in the data and account for its contribution in fitting the B_S^0 lifetime. The fraction of the signal coming from real D_S^- decays, $f_{D_S^-}$, can be measured from two different distributions. First, because of the large difference between the D_S^- and D^- lifetimes, both well-measured, the D_S^- proper decay length distribution can be used to fit for the fractions of each in the signal. Second, the shapes of the reflection in the mass distributions of D_S^- and D^- events from Monte Carlo can be used to extract the amount of each in a simultaneous fit to the D_S^- and D^- mass distributions from the data. Both methods are described here.

5.5 Two Lifetime Fit for the D_S^- Fraction

The D_S^- fraction in the data is first measured using the proper decay length distribution. The technique is to start with the same likelihood that was used to fit the D_S^- lifetime in the $\phi\pi^-$ sample. In place of the exponential that describes the signal distribution, a sum of two exponentials is used, one with the D_S^- lifetime and one with the D^- lifetime. The relative fraction of D_S^-:D^- is allowed to float in order to determine the amount of D^- remaining in the sample. In addition, the right-sign sideband region above the D_S^- peak is excluded from the background sample to avoid any contamination from the tail of the D^- reflection in the sidebands.

5.5.1 Test of Two Lifetime Fit for D_S^- Fraction in $D_S^- \to K^{*0}K^-$ Monte Carlo

The two lifetime fit for the D_S^- fraction is first tested with Monte Carlo events. Beginning with Monte Carlo samples of B_S^0 and B^0 semi-leptonic decays including a D_S^- and D^-, respectively, the individual D lifetimes are fit after applying all the selection criteria. Figure 5.7(a) shows the D_S^- proper decay length distribution with the fit result superimposed. The fitted value of $c\tau(D_S^-)$ is 139 $\pm$ 2 μm, in agreement with the input value of 140 μm. Figure 5.7(b) shows the D^- proper decay length distribution with the fit result superimposed. The fitted value for $c\tau(D^-)$ is 311 $\pm$ 14 μm, which agrees well with the input value of 315 μm. It should be noted that the D^- here has been reconstructed as a D_S^-.

To fit for the D_S^- fraction, these two samples are combined. There are 2867 D_S^- events and 1000 D^- events, for a total of 3867 events. The true D_S^- fraction in this sample is then 74.1%. The D_S^- and D^- lifetimes are fixed at the Monte Carlo input values and the fit returns a D_S^- fraction of $f_{D_S^-} = (72.9 \pm 2.3)$ %, which agrees quite well with the true

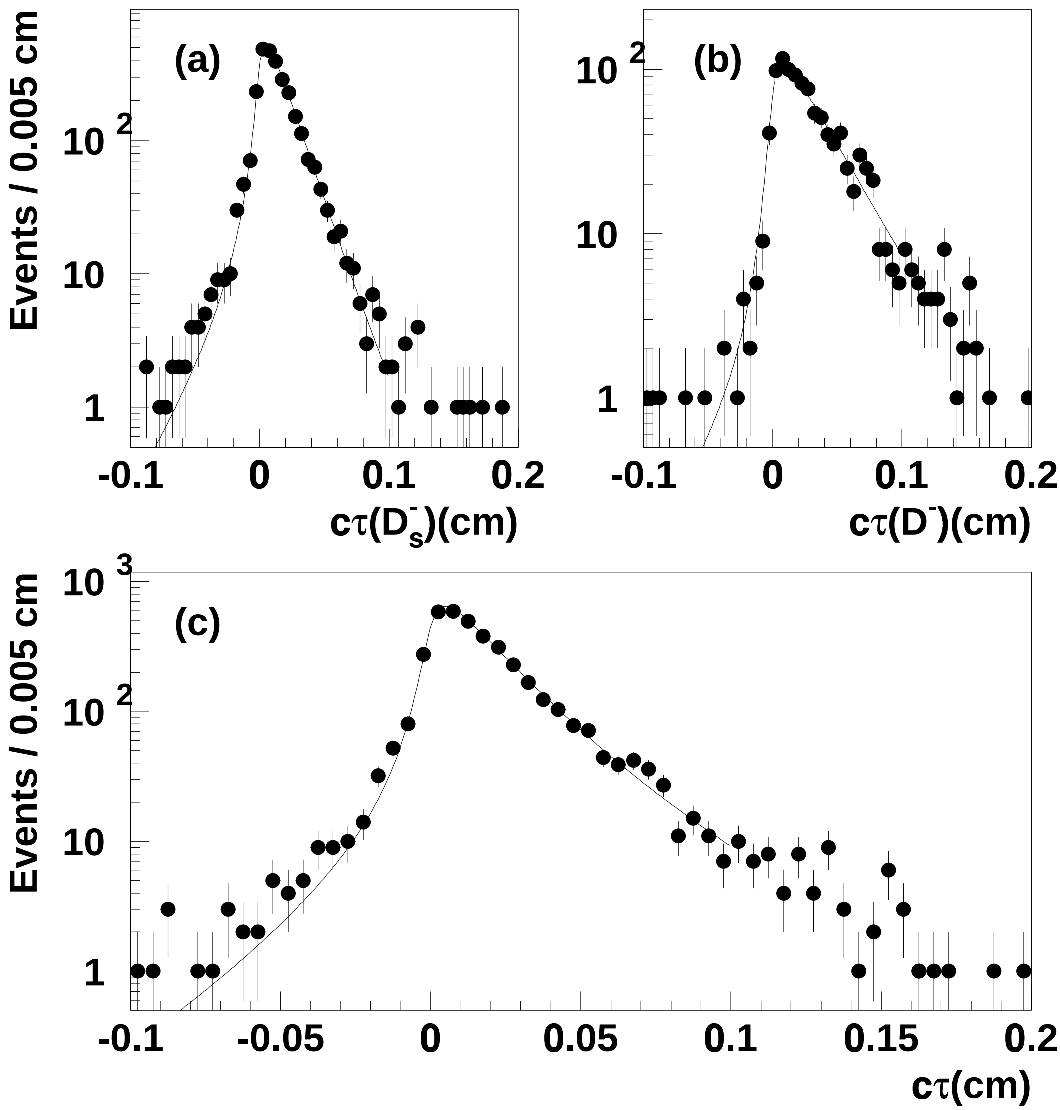


Figure 5.7: Proper decay length distributions for (a) $D_S^- \to K^{*0}K^-$ Monte Carlo events, (b) $D^- \to K^{*0}\pi^-$ Monte Carlo events, and (c) the combined Monte Carlo sample. The results of the fits for (a) $c\tau(D_S^-)$, (b) $c\tau(D^-)$, and (c) the D_S^- fraction have been superimposed. The plots extend to 0.2 cm, but the fit extends only over the range $\pm$ 0.1 cm.

D_S^- fraction in the sample. Figure 5.7(c) shows the proper decay length distribution for the combined Monte Carlo sample, with the result of the D_S^- fraction fit superimposed.

5.5.2 Two Lifetime Fit for D_S^- Fraction in $D_S^- \rightarrow K^{*0}K^-$ Data

The two lifetime fit for the D_S^- fraction is now applied to the $D_S^- \rightarrow K^{*0}K^-$ data sample. The D_S^- and D^- lifetimes are fixed to their PDG values of 140 μm and 317 μm, respectively, and a D_S^- fraction of

$$f_{D_S^-} = (60.6 \, ^{+16.5}_{-19.5}) \ \%$$

is obtained. Figure 5.8(a) shows the proper decay length distribution for the signal region with the result of the fit overlaid. The D_S^- contribution is represented by the dashed curve, while the D^- contribution is represented by the dotted curve. The shaded curve shows the background contribution to the signal sample. Figure 5.8(b) shows the proper decay length distribution for the sideband sample with the fit result overlaid.

5.5.3 Toy Monte Carlo Test of Two Lifetime Fit for D_S^- Fraction

In the first Monte Carlo test of the two lifetime fit for the D_S^- fraction, samples were used that were much larger than the true size of the $D_S^- \rightarrow K^{*0}K^-$ data sample. To test the stability of the technique on sample sizes similar to the data, a toy Monte Carlo is used.

D_S^- and D^- samples are generated with mean numbers of events of 140 and 60, respectively. The numbers of events are fluctuated about the means, and the shapes of the individual proper decay length distributions are taken from the results of the fit to the $D_S^- \rightarrow K^{*0}K^-$ data. Each event is assigned an error where the distribution of errors is also taken from the $D_S^- \rightarrow K^{*0}K^-$ data sample. 5000 samples are generated and the D_S^- fraction is fitted in each sample. The fitted fraction is compared to the true fraction based on the numbers of generated events. Figure 5.9 shows the distribution of the difference between the fitted fraction and the true fraction. The distribution has a mean of -0.4%, showing that the method works properly.

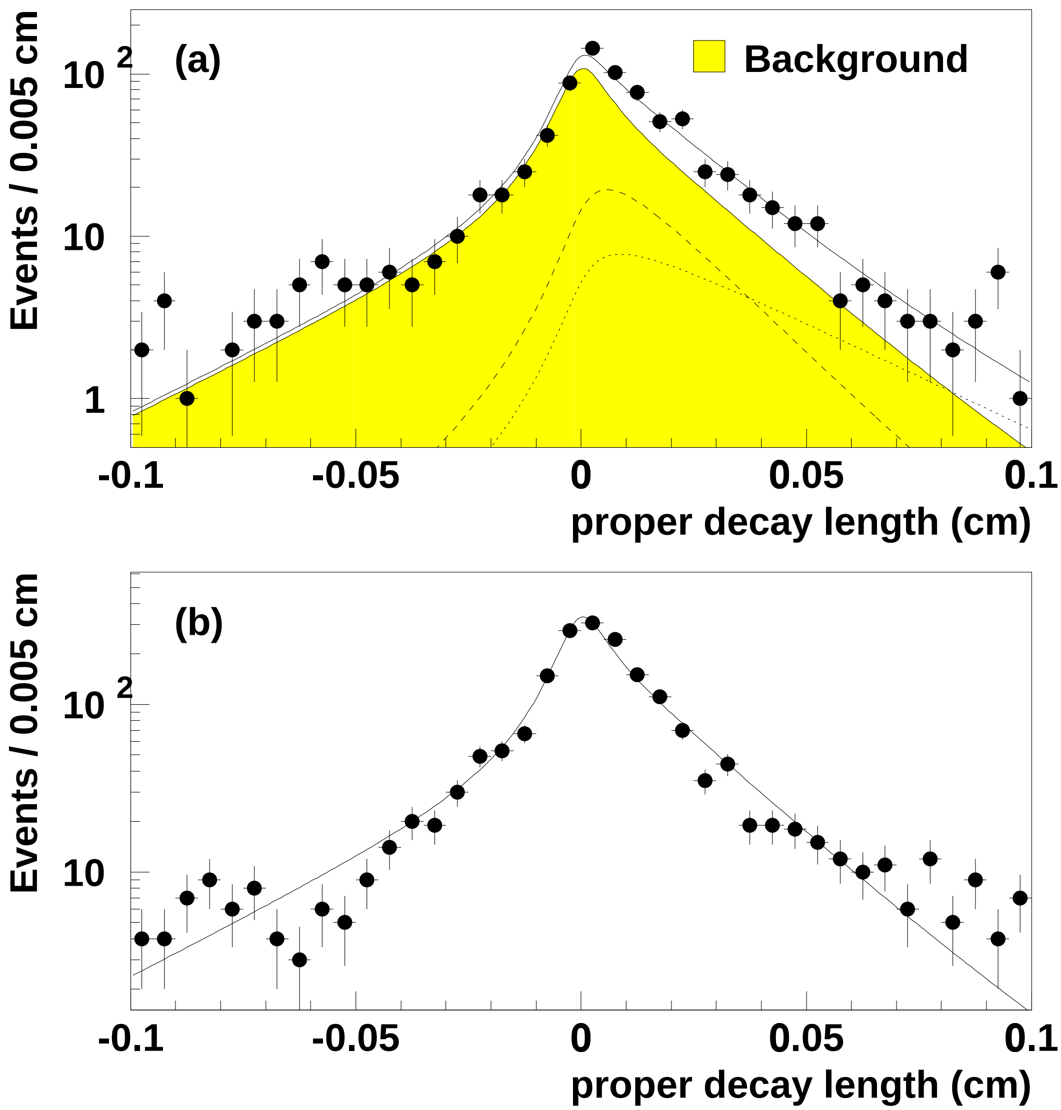


Figure 5.8: D_S^- proper decay length distribution in the $D_S^- \to K^{*0}K^-$ sample, for (a) signal region and (b) sideband region. Results of the two-component fit for the D_S^- fraction are superimposed. The dashed curve in (a) represents the D_S^- contribution, while the dotted curve represents the D^- contribution.

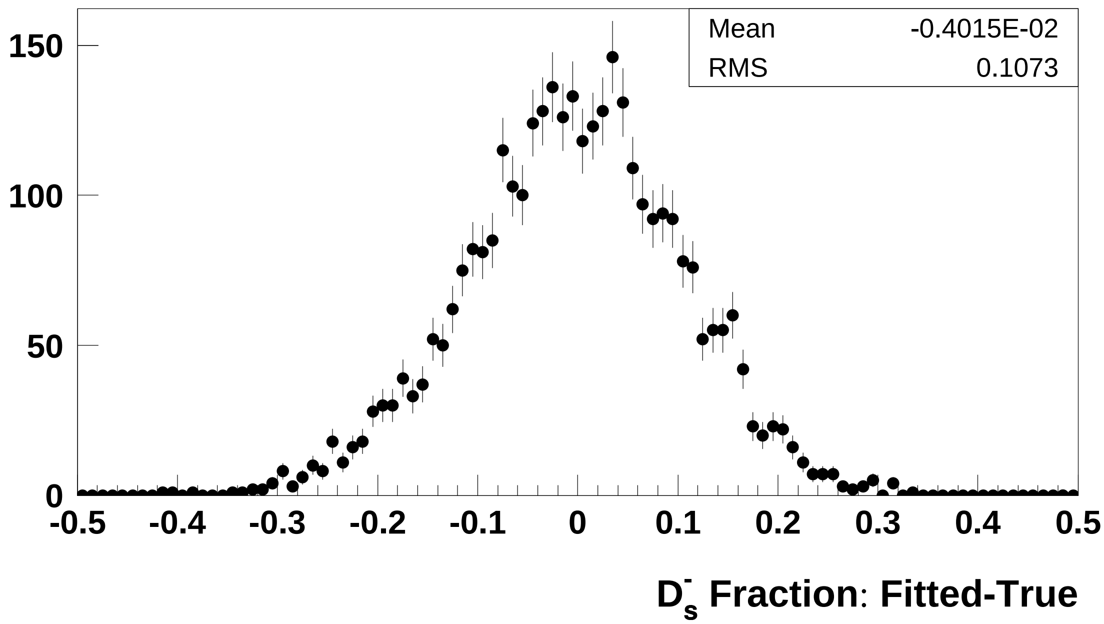


Figure 5.9: Difference between the fitted and true D_S^- fractions in the toy Monte Carlo, where no background is present.

As a further test, background events are included. The signal samples are generated using the same method as before while background events are generated based on the fitted shape of the background distribution in the $D_S^- \to K^{*0}K^-$ sample. A background sample with a mean of 2000 events is generated. A second sample of background events with a mean of 600 events is also generated and included in the signal sample to model the background contribution to the signal sample. The D_S^- fraction is fit and compared with the true fraction. Figure 5.10(a) shows the distribution of the difference between the fitted fraction and the true fraction. The distribution has a mean of -0.9% and an RMS of 17.5%. Figure 5.10(b) shows the distribution of the fitted error on the D_S^- fraction. This distribution shows a mean error of 19.7%, in good agreement with the error returned from the D_S^- fraction fit to the data.

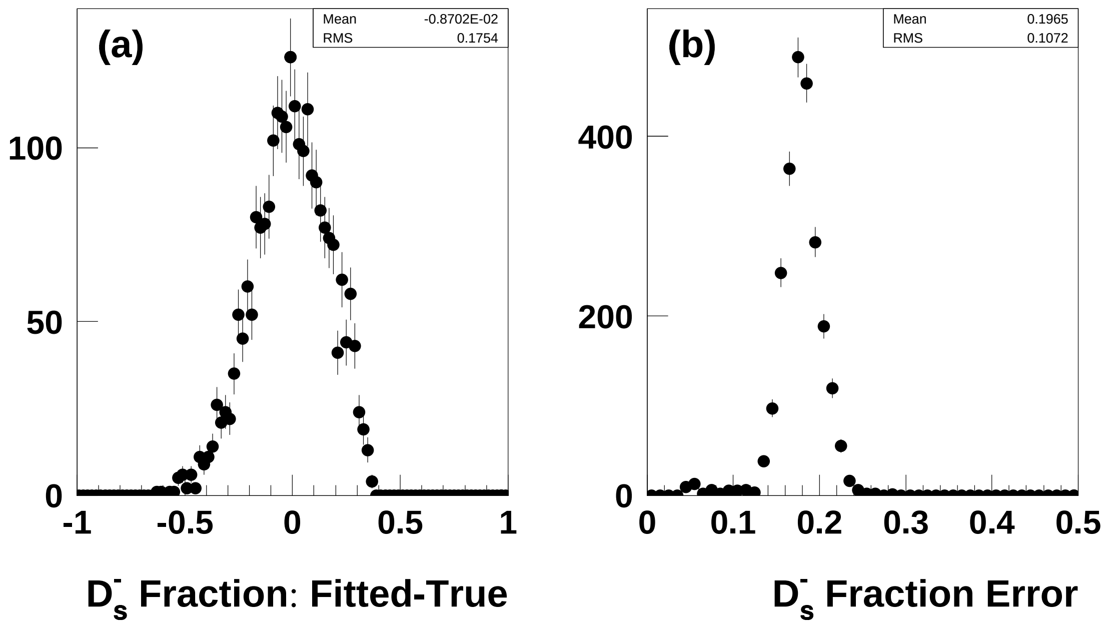


Figure 5.10: (a) Difference between the fitted and true D_S^- fractions in the toy Monte Carlo after background is added. (b) Distribution of errors from the fits for the D_S^- fraction in the toy Monte Carlo.

5.5.4 Test of Two Lifetime Fit for D_S^- Fraction in $D_S^- \to K_s^0 K^-$ Monte Carlo

Though the method of obtaining the D_S^- fraction from the proper decay length distribution has already been tested, it must be checked that the D_S^- and D^- lifetimes are correctly reproduced in the $D_S^- \to K_s^0 K^-$ sample. Monte Carlo samples are generated that are similar to those in Section 5.5.1, except that the D_S^- and D^- decays now involve a K_s^0 in place of a K^{*0}.

Figure 5.11(a) shows the proper decay length distribution in the Monte Carlo $D_S^- \to K_s^0 K^-$ sample, with the fit result superimposed. The fitted value of $c\tau(D_S^-)$ is 139 $\pm$ 4 μm, in agreement with the input value of 140 μm. Figure 5.11(b) shows the proper decay length distribution in the Monte Carlo $D^- \to K_s^0 \pi^-$ sample, with the fit result superimposed. The fitted value for $c\tau(D^-)$ is 312 $\pm$ 21 μm, which agrees with the input value of 315 μm. Again the D^- here has been reconstructed as a D_S^-.

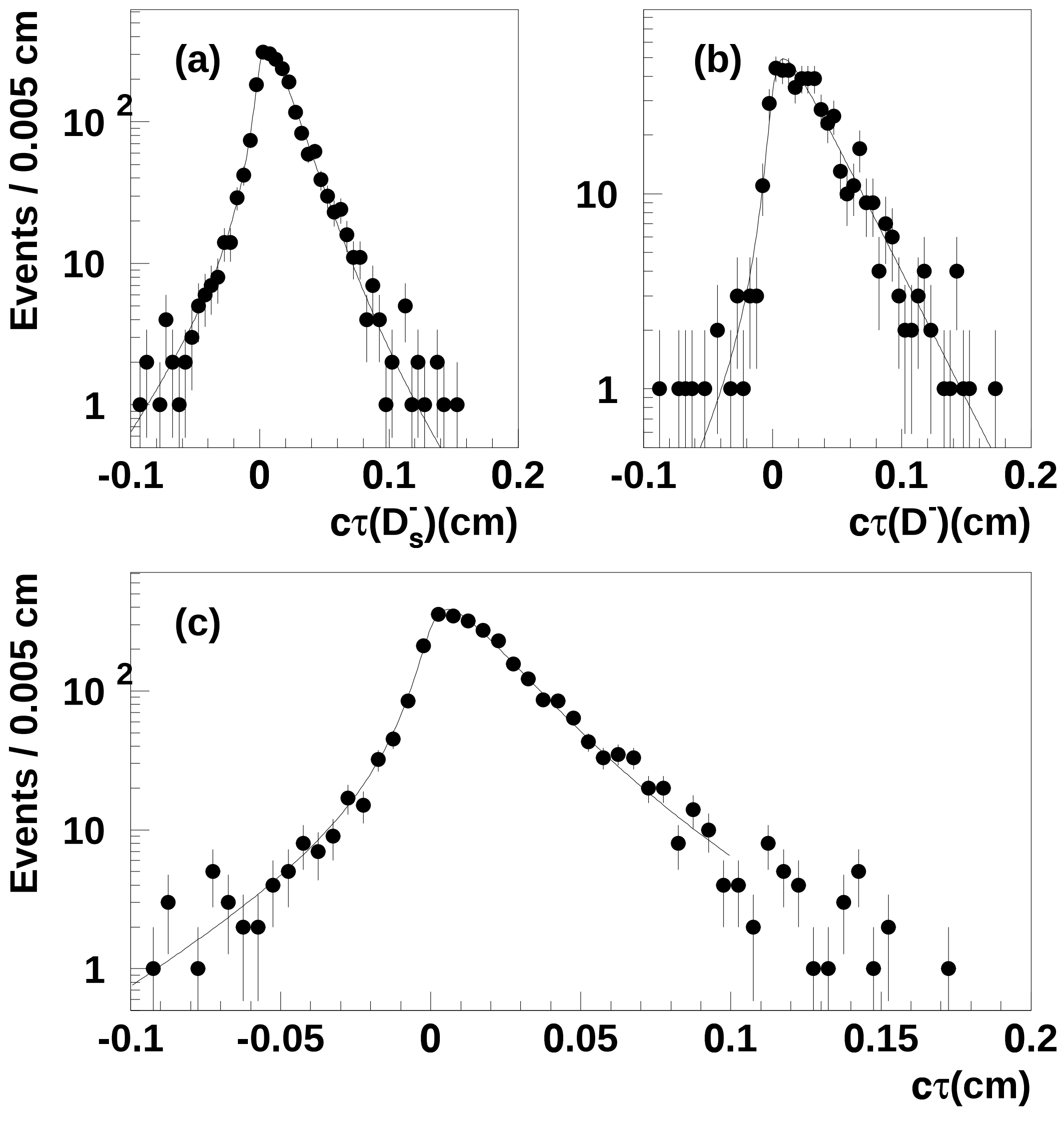


Figure 5.11: Proper decay length distributions for (a) $D_S^- \to K_s^0 K^-$ Monte Carlo events, (b) $D^- \to K_s^0 \pi^-$ Monte Carlo events, and (c) the combined Monte Carlo sample. The results of the fits for (a) $c\tau(D_S^-)$, (b) $c\tau(D^-)$, and (c) the D_S^- fraction have been superimposed. The plots extend to 0.2 cm, but the fit extends only over the range $\pm$ 0.1 cm.

Combining the two samples, there are 2207 D_S^- events and 504 D^- events, for a total of 2711 events in the $K_s^0K^-$ and $K_s^0\pi^-$ Monte Carlo samples. This gives a D_S^- fraction in this sample of 81.4%. This time the fit returns a D_S^- fraction of $f_{D_S^-} = (82.0 \pm 2.8)$ %, in agreement with the true D_S^- fraction in the sample. Figure 5.11(c) shows the proper decay length distribution of the combined Monte Carlo sample, with the result of the two lifetime fit superimposed.

5.5.5 Two Lifetime Fit for D_S^- Fraction in $D_S^- \to K_s^0K^-$ Data

The two lifetime fit for the D_S^- fraction is now applied to the $D_S^- \to K_s^0K^-$ data. Using the same D_S^- and D^- lifetimes as in Section 5.5.2, a D_S^- fraction of

$$f_{D_S^-} = (92.7\ ^{+7.3}_{-44.7})\ \%$$

is obtained. This value is higher than what was measured for the $D_S^- \to K^{*0}K^-$ sample. However, the errors are quite large and the results are still in agreement with the previous measurement of the D_S^- fraction.

Figure 5.12(a) shows the proper decay length distribution for the signal region with the result of the fit overlaid. The D_S^- contribution is represented by the dashed curve, while the D^- contribution is represented by the dotted curve. The shaded curve shows the background contribution to the signal sample. Figure 5.12(b) shows the proper decay length distribution for the sideband sample with the fit result overlaid.

5.6 D_S^- Fraction Fit from Mass Distributions

As a second independent measurement of the D_S^- fractions, a simultaneous fit to the D_S^- and D^- mass distributions is performed, where the D^- mass distribution has been created by switching the mass assignment on the third track of the D_S^- candidate from a K to a π. Each of the two mass distributions is fit with a gaussian for the corresponding signal, plus a linear background, plus the shape of the events coming from the reflection. This shape is taken from the mass distributions of the reflected Monte Carlo events in Figure 5.5.

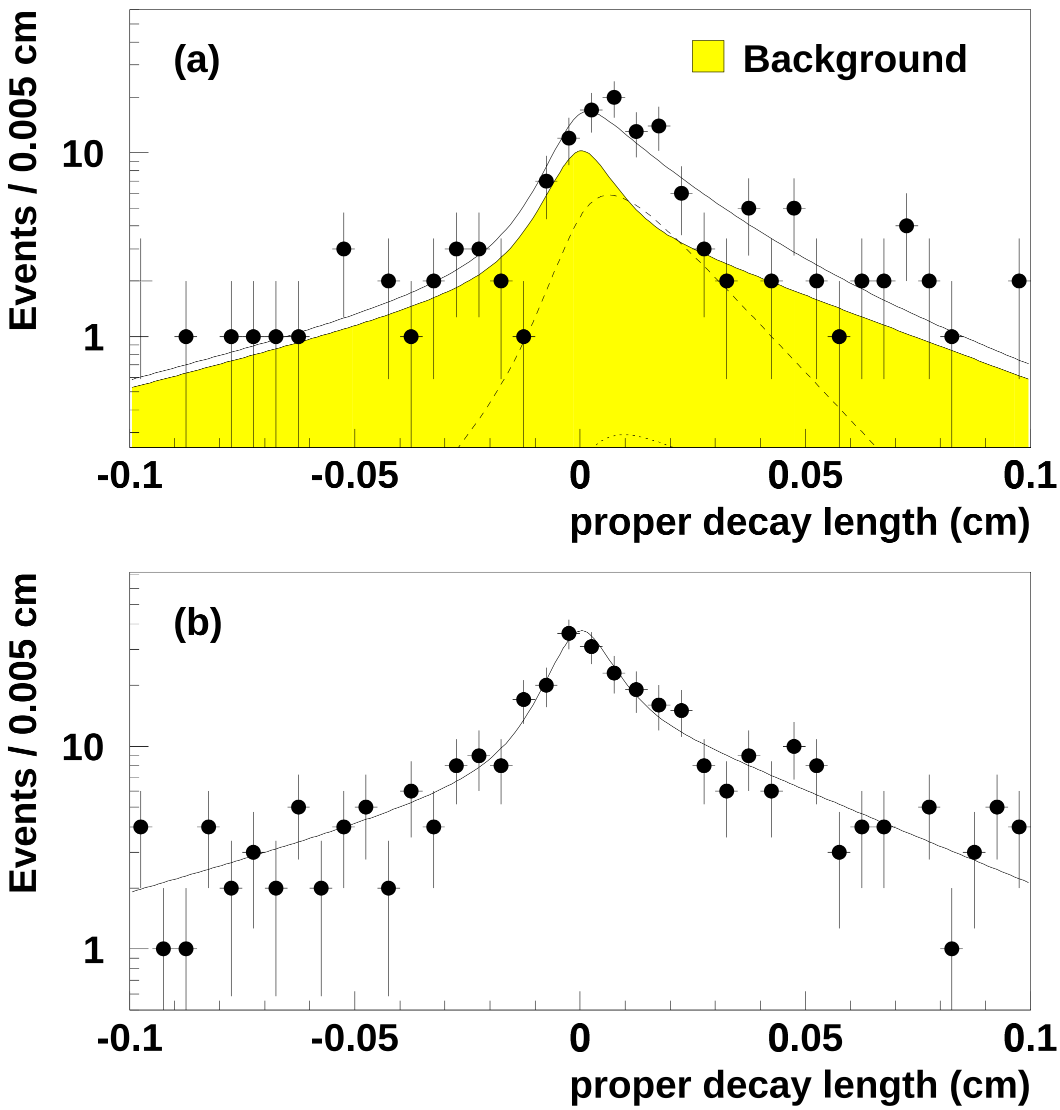


Figure 5.12: D_S^- proper decay length distribution in the $D_S^- \to K_s^0 K^-$ sample, for (a) signal region and (b) sideband region. Results of the two-component fit for the D_S^- fraction are superimposed. The dashed curve in (a) represents the D_S^- contribution, while the dotted curve represents the D^- contribution.

The D_S^- and D^- mass distributions are fit simultaneously, with the number of events in the gaussian signal of the $D_S^-(D^-)$ mass distribution constrained to the number of events in the reflection in the $D^-(D_S^-)$ mass distribution. In addition, the difference between the D_S^- and D^- mass values is fixed based on their PDG values, and the ratio of the widths is fixed to the ratio of the masses.

5.6.1 D_S^- Fraction Fit from Mass Distributions in $D_S^- \to K^{*0}K^-$ Data

Figure 5.13(a) shows the D_S^- mass distribution of the $D_S^- \to K^{*0}K^-$ data sample with the result of the fit superimposed. The unshaded gaussian distribution represents the contribution of the D_S^- signal, while the shaded histogram represents the contribution of events from the D^- reflection. Similarly, Figure 5.13(b) shows the corresponding D^- mass distribution with the results of the fit superimposed. The unshaded gaussian distribution represents the contribution of the D^- signal events, while the shaded histogram represents the contribution of events from the D_S^- reflection. The area of the gaussian in Figure 5.13(a(b)) is equal to the area of the shaded histogram in Figure 5.13(b(a)).

The fit finds 123 $\pm$ 25 D_S^- events and 204 $\pm$ 25 D^- events in the sample. Using the value of 123 $\pm$ 25 D_S^- events here, and the 179 $\pm$ 35 events from the gaussian plus linear background fit shown in Figure 5.1, the D_S^- fraction is

$$f_{D_S^-} \;=\; (68.7 \;\pm\; 19.4)\;\%.$$

This agrees well with the value obtained with the two lifetime fit.

5.6.2 Test of D_S^- Fraction Fit from Mass Distributions with Toy Monte Carlo

To test the method of obtaining the D_S^- fraction from the mass distributions, a toy Monte Carlo is used again. Separate mass distributions are generated for D_S^- and D^-. Each distribution is the sum of three components: (1) a gaussian for the signal, (2) the component from the reflection, and (3) a linear background. The shape of the reflection again comes from the Monte Carlo events shown in Figure 5.5. For each event that is generated for

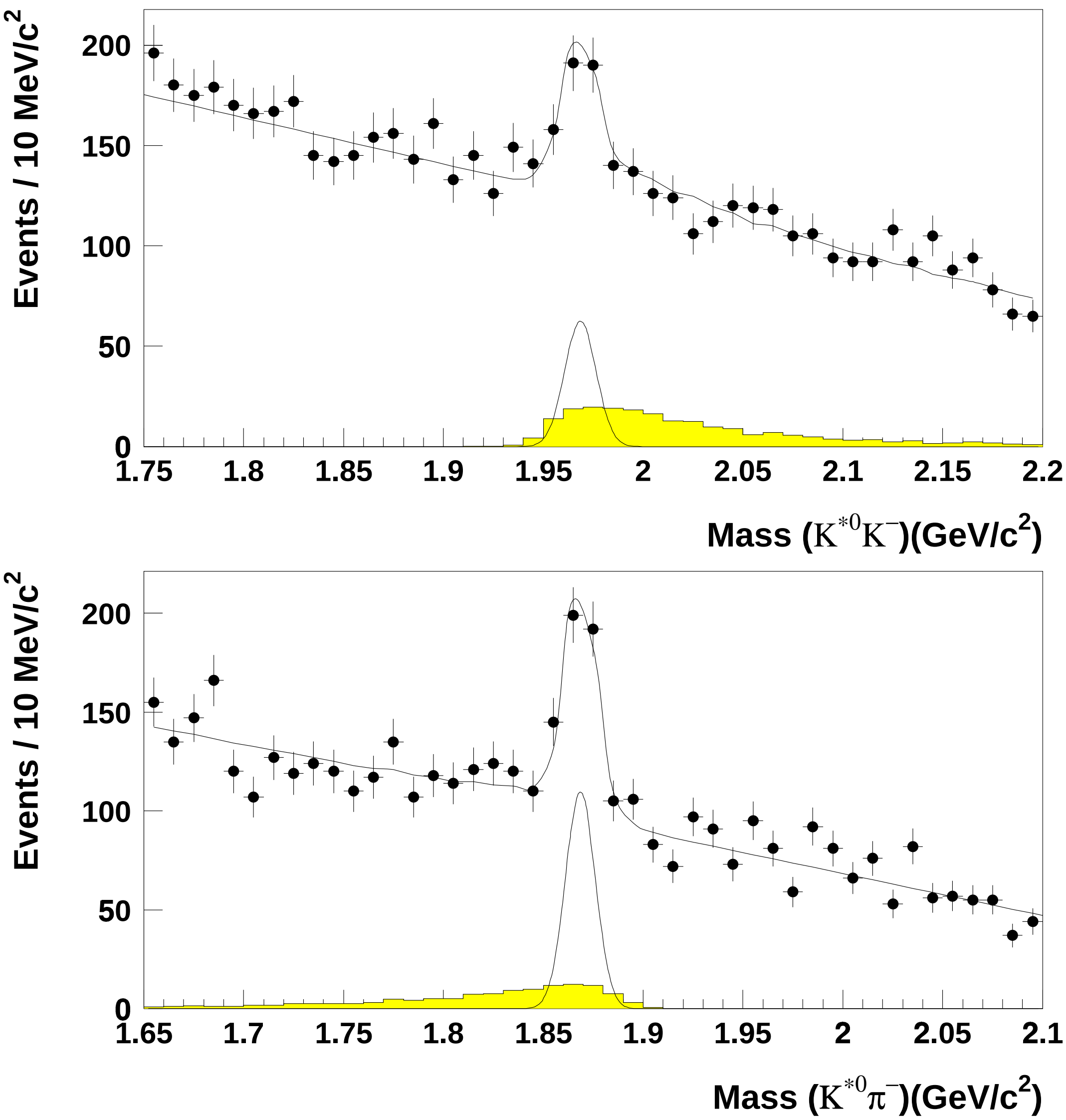


Figure 5.13: Mass distributions for events in the $D_S^- \to K^{*0}K^-$ sample when the events are assumed to be (a)D_S^-, or (b)D^-. The gaussian curve shows the fit result for the real D_S^- or D^- contribution, while the shaded histogram shows the fit result for the contribution from reflected events.

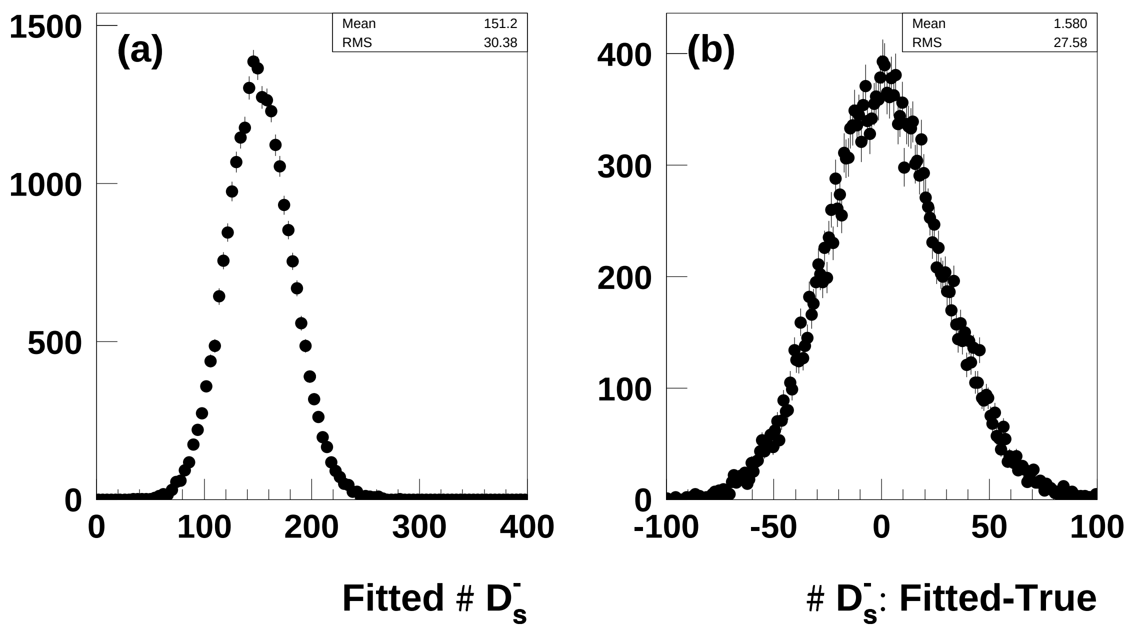


Figure 5.14: (a)Fitted number of D_S^- events in the toy Monte Carlo samples. (b)Difference between the fitted number of D_S^- events and the true number of D_S^- generated.

the gaussian component of one sample, an event is generated for the reflection in the other sample. Samples are generated with means of 150 D_S^- events, 200 D^- events, and 5700 background events, with the actual numbers of events in each sample allowed to fluctuate.

To extract the D_S^- fraction both distributions are first simultaneously fit for the numbers of D_S^- and D^- events. Figure 5.14(a) shows the distribution of the fitted number of D_S^- events, while Figure 5.14(b) shows the difference between the fitted number of D_S^- events and the true number of events generated. The mean difference is only 1.6 events.

Next the D_S^- mass distribution is fit with just a gaussian and a linear background as in the data. Figure 5.15(a) shows the distribution of the fitted total number of events above the background, while Figure 5.15(b) shows the difference between the fitted number of events and the true number of events. Here the true number of events is determined by adding the number of D_S^- events generated and the number of D^- events whose reflected mass lies in the D_S^- signal region. In this case the mean difference is 10.2 events.

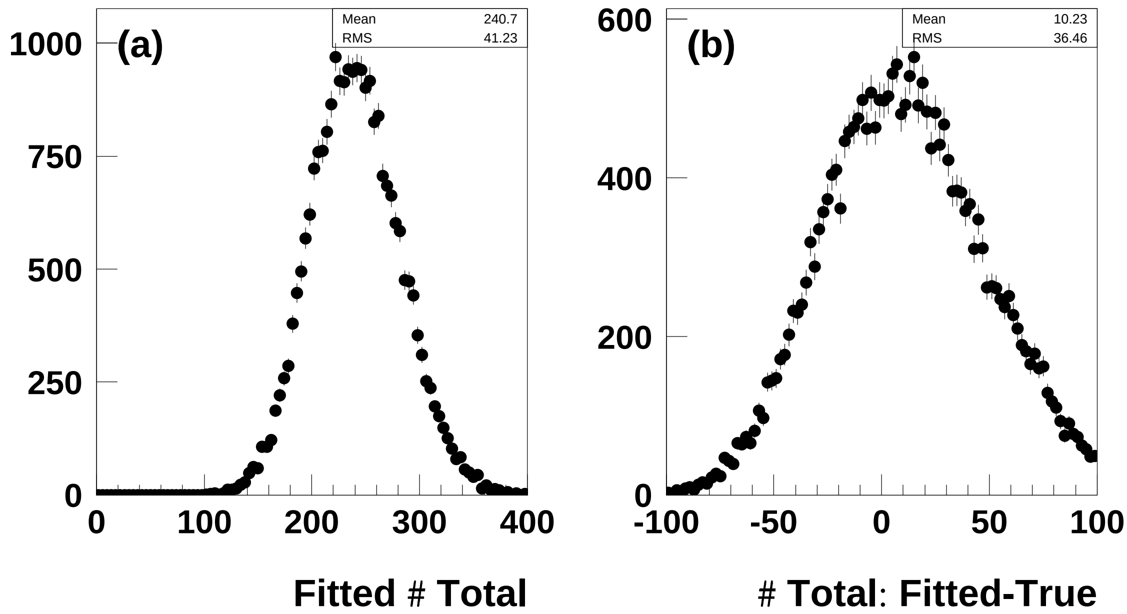


Figure 5.15: (a)Fitted total number of events above background in the toy Monte Carlo samples. (b)Difference between the fitted total number of events and the true number.

To determine the shift from the true D_S^- fraction, first note that the generated samples have means of 150 D_S^- events and 230 total events above background, which translates to a D_S^- fraction of 65.0%. Using the means of the distributions of fitted numbers of events yields a mean D_S^- fraction of 62.8%. Thus a relatively small overall shift of -2.2% is observed, leading to the conclusion that this method also works correctly.

5.6.3 D_S^- Fraction Fit from Mass Distributions in $D_S^- \to K_s^0 K^-$ Data

As with the $D_S^- \to K^{*0} K^-$ sample, a simultaneous fit to the D_S^- and D^- mass distributions from the $D_S^- \to K_s^0 K^-$ sample is performed, where the D^- mass distribution has been created by switching the mass assignment on the third track of the D_S^- candidate from a K to a π. The distributions are fit with the same shapes as in Section 5.6.1, and the same constraints are used.

Figure 5.16(a) shows the D_S^- mass distribution from the $D_S^- \to K_s^0 K^-$ sample with the results of the fit superimposed. The D_S^- signal is represented by the unshaded gaussian distribution, while the component from the D^- reflection is represented by the shaded histogram. Figure 5.16(b) shows the corresponding D^- mass distribution with the results of the fit superimposed. Here the gaussian represents the D^- contribution and the shaded histogram represents the D_S^- reflection. The area of the gaussian in Figure 5.16(a(b)) is equal to the area of the shaded histogram in Figure 5.16(b(a)).

The fit returns 19 $\pm$ 13 D_S^- events and 52 $\pm$ 16 D^- events. Using the number of D_S^- events from this fit, and the 41 $\pm$ 14 events from the gaussian plus linear background fit shown in Figure 5.4, the D_S^- fraction is measured to be

$$f_{D_S^-} = (46.3 \pm 35.4)\ \%.$$

This value is different from the D_S^- fraction in this sample obtained with the two lifetime fit, but both measurements are consistent within their large errors.

5.7 Final D_S^- Fraction Result

Based on the results of the tests using Monte Carlo events, both methods for measuring the fraction of real D_S^- events in the signal are shown to work correctly. To use the most information in determining the D_S^- fraction, the average of the two methods is taken for the value of $f_{D_S^-}$ that will be used in fitting the B_S^0 lifetime. A simple average is used since the statistical errors from the two measurements are nearly equal. The D_S^- fractions measured with the two lifetime fit to the proper decay length distribution are

$$\begin{aligned} f_{D_S^-} &= (60.6\ ^{+16.5}_{-19.5})\ \% \quad (D_S^- \to K^{*0}K^-) \\ f_{D_S^-} &= (92.7\ ^{+7.3}_{-44.7})\ \% \quad (D_S^- \to K_s^0 K^-). \end{aligned}$$

The D_S^- fractions measured with a simultaneous fit to the mass distributions are

$$\begin{aligned} f_{D_S^-} &= (68.7 \pm 19.4)\ \% \quad (D_S^- \to K^{*0}K^-) \\ f_{D_S^-} &= (46.3 \pm 35.4)\ \% \quad (D_S^- \to K_s^0 K^-). \end{aligned}$$

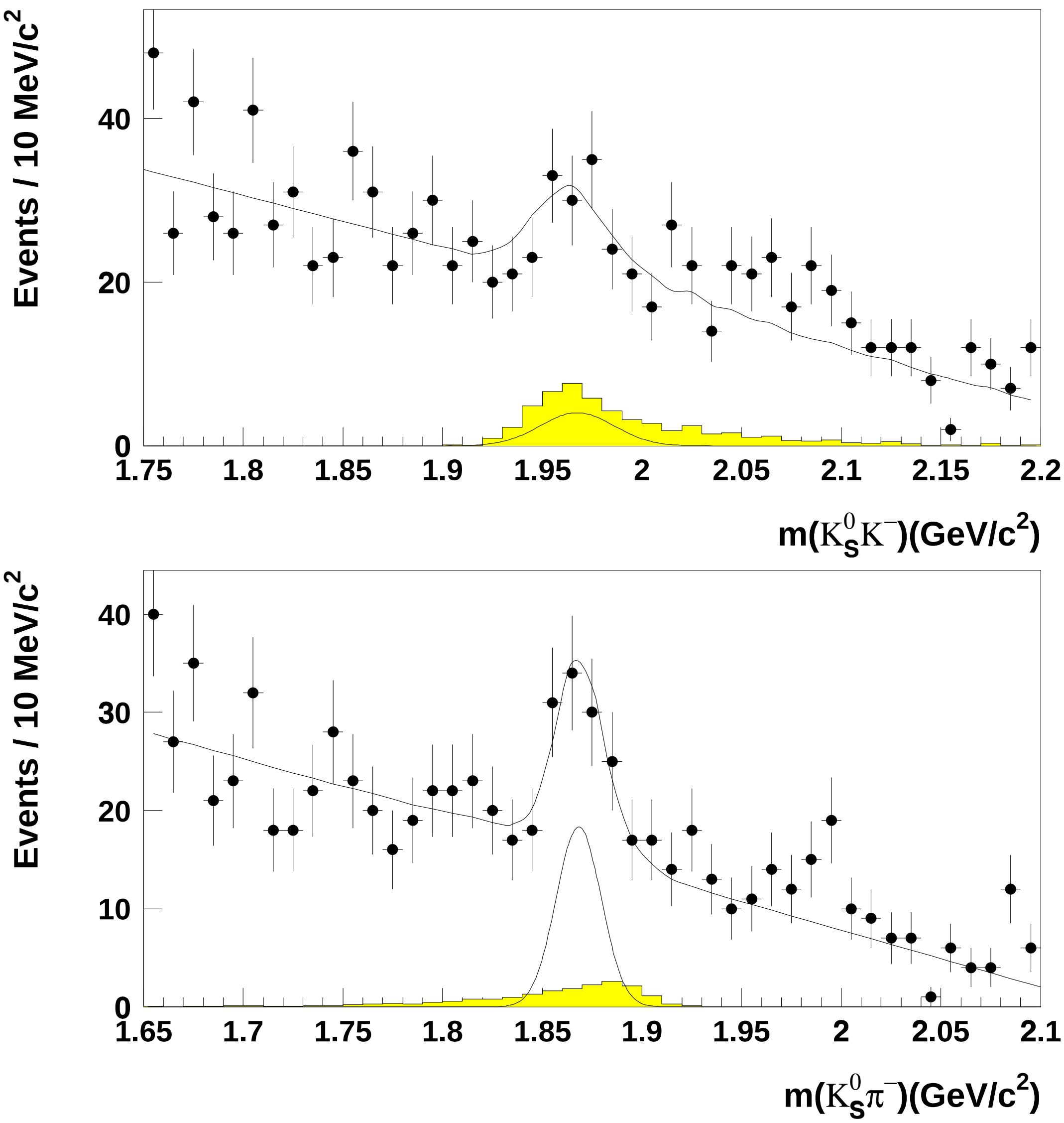


Figure 5.16: Mass distributions for events in the $D_S^- \to K_s^0 K^-$ sample when the events are assumed to be (a)D_S^-, or (b)D^-. The gaussian curve shows the fit result for the real D_S^- or D^- contribution, while the shaded histogram shows the fit result for the contribution from reflected events.

Averaging the two measurements for each sample gives the following D_S^- fractions, which will be used in the B_S^0 lifetime fit (see Chapter 6):

$$\begin{aligned} f_{D_S^-} &= 64.7\ \% \quad (D_S^- \to K^{*0}K^-) \\ f_{D_S^-} &= 69.5\ \% \quad (D_S^- \to K_s^0 K^-). \end{aligned}$$

The effect on the measurement of the B_S^0 lifetime due to a mismeasurement of $f_{D_S^-}$ is considered among the systematic uncertainties (see Section 6.4).

Chapter 6

Combined B_S^0 Lifetime Measurement

The techniques for measuring the B_S^0 lifetime have been described in detail in Chapter 4. In this chapter those techniques are briefly reviewed first. For comparison, the components of the lifetime fit are first presented separately for each D_S^- decay mode. The B_S^0 lifetime is measured in each sample individually, and then a combined fit for the lifetime is performed using all the samples simultaneously. Finally, sources of systematic uncertainty in the measurement of the B_S^0 lifetime are considered and the magnitude of those uncertainties is estimated.

6.1 Review of Lifetime Measurement Technique

The quantity used to extract the B_S^0 lifetime is called the pseudo-proper decay length and is defined as

$$c\tau^* = L_{\mathrm{xy}} \frac{m(B_S^0)}{p_T(D_S^- \ell^+)}.$$

A correction factor K is used to relate this quantity back to the true $c\tau$. K is defined by the expression

$$K = \frac{p_T(D_S^- \ell^+)}{p_T(B_S^0)}.$$

Since the distribution of the correction factor is potentially shaped by the D_S^- selection requirements, separate K distributions are generated for each of the three D_S^- decay modes. Figure 6.1 shows the three K distributions for (a) $D_S^- \to \phi\pi^-$, (b) $D_S^- \to K^{*0}K^-$, and (c)

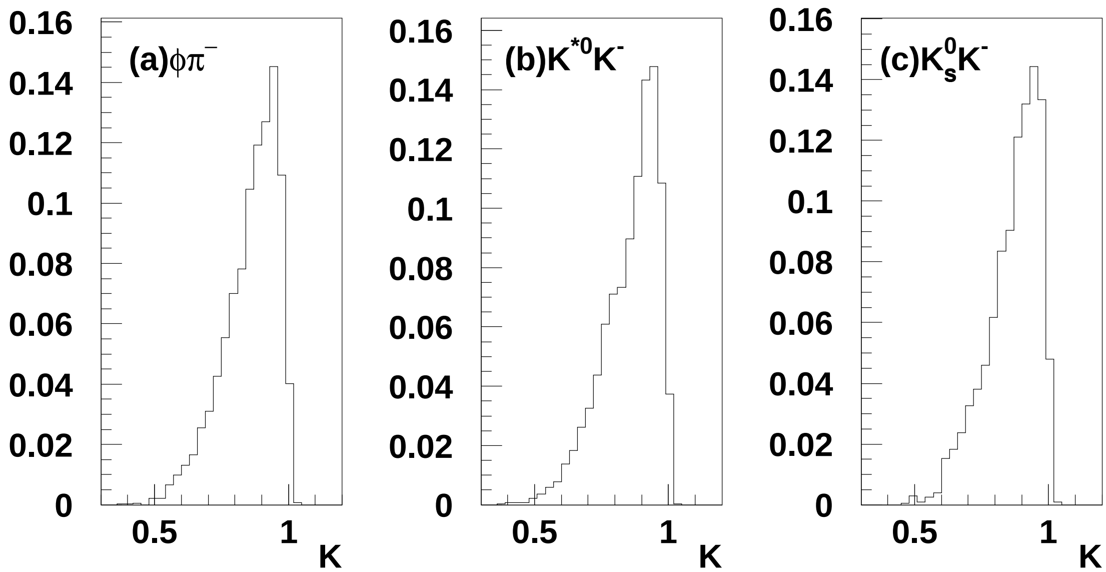


Figure 6.1: K Distributions for the three D_S^- decay modes: (a) $\phi\pi^-$, (b) $K^{*0}K^-$, and (c) $K_s^0K^-$. The mean and RMS of each distribution are listed in Table 6.1.

Decay Mode	Mean	RMS
$\phi\pi^-$	0.860	0.103
$K^{*0}K^-$	0.858	0.106
$K_s^0K^-$	0.868	0.101

Table 6.1: Mean and RMS of the K distributions for each D_S^- decay mode.

$D_S^- \to K_s^0K^-$, and Table 6.1 lists the mean and RMS of each distribution. It can be seen that the three K distributions are all quite similar.

6.2 Physics Backgrounds

Background events which produce a real D_S^- and a real lepton, where neither comes from the semi-leptonic decay of a B_S^0 meson, were described in Chapter 4. The two primary sources of these events are the following decays: (1) $B^0 \to D_S^{(*)+}D^{(*)-}X$ and $B^+ \to$

Decay Mode	$f_{D_S D}$	$c\tau(D_S D)$	$f_{D_S D_S}$	$c\tau(D_S D_S)$
$\phi\pi^-$	0.0063	590 ± 21	0.0013	577 ± 24
$K^{*0}K^-$	0.0061	583 ± 25	0.0012	647 ± 33
$K^0_s K^-$	0.0043	578 ± 27	0.00089	537 ± 30

Table 6.2: Expected fractions and effective lifetimes of $B \rightarrow D_S^{(*)} D^{(*)}$ and $B_S^0 \rightarrow D_S^{(*)} D_S^{(*)}$ events for each D_S^- decay mode.

$D_S^{(*)+} D^{(*)0} X$, with the $D^{(*)-}$ or $D^{(*)0}$ decaying semi-leptonically, and (2) $B_S^0 \rightarrow D_S^{(*)-} D_S^{(*)+}$, with one $D_S^{(*)}$ decaying semi-leptonically.

To account for the presence of these events, Monte Carlo samples are used to first estimate the fractions of events from these decays in the D_S^- signal sample. Using the same Monte Carlo samples, an effective lifetime for these events is determined. Additional terms must be added to the likelihood to account for the contribution of these events to the pseudo-proper decay length distribution. For the $B \rightarrow D_S D$ decays the term

$$f_{D_S D} \cdot \frac{K}{c\tau(D_S D)} \cdot exp(\frac{-Kx}{c\tau(D_S D)}) \otimes G(x, s\sigma) \otimes H(K)$$

is added, where $f_{D_S D}$ is the estimated fraction of the events in the signal sample and $c\tau(D_S D)$ is the effective lifetime of those events. A similar term is added for $B_S^0 \rightarrow D_S^{(*)-} D_S^{(*)+}$ decays.

The explicit calculation of $f_{D_S D}$ and $f_{D_S D_S}$ in the $D_S^- \rightarrow \phi\pi^-$ sample was performed in Section 4.4.4. The same procedure is used to calculate the expected fraction of $D_S D$ and $D_S D_S$ events in each of the three D_S^- decay modes. The expected fractions for each decay mode are listed in Table 6.2 along with the reconstructed effective lifetimes. Figure 6.2 shows the pseudo-proper decay length distributions for the $D_S D$ and $D_S D_S$ samples with the results of the effective lifetime fits superimposed.

In the B_S^0 lifetime fit both f and $c\tau$ are fixed to the values measured in the Monte Carlo, with the same values for $c\tau(D_S D)$ and $c\tau(D_S D_S)$ used for all D_S^- decay modes. The effective lifetime values are determined by fits to the combined $D_S^- \rightarrow \phi\pi^-$ and $D_S^- \rightarrow K^{*0}K^-$ Monte Carlo samples. The results of these fits are

$$c\tau(D_S D) = 586\ \mu m \qquad \text{and} \qquad c\tau(D_S D_S) = 604\ \mu m.$$

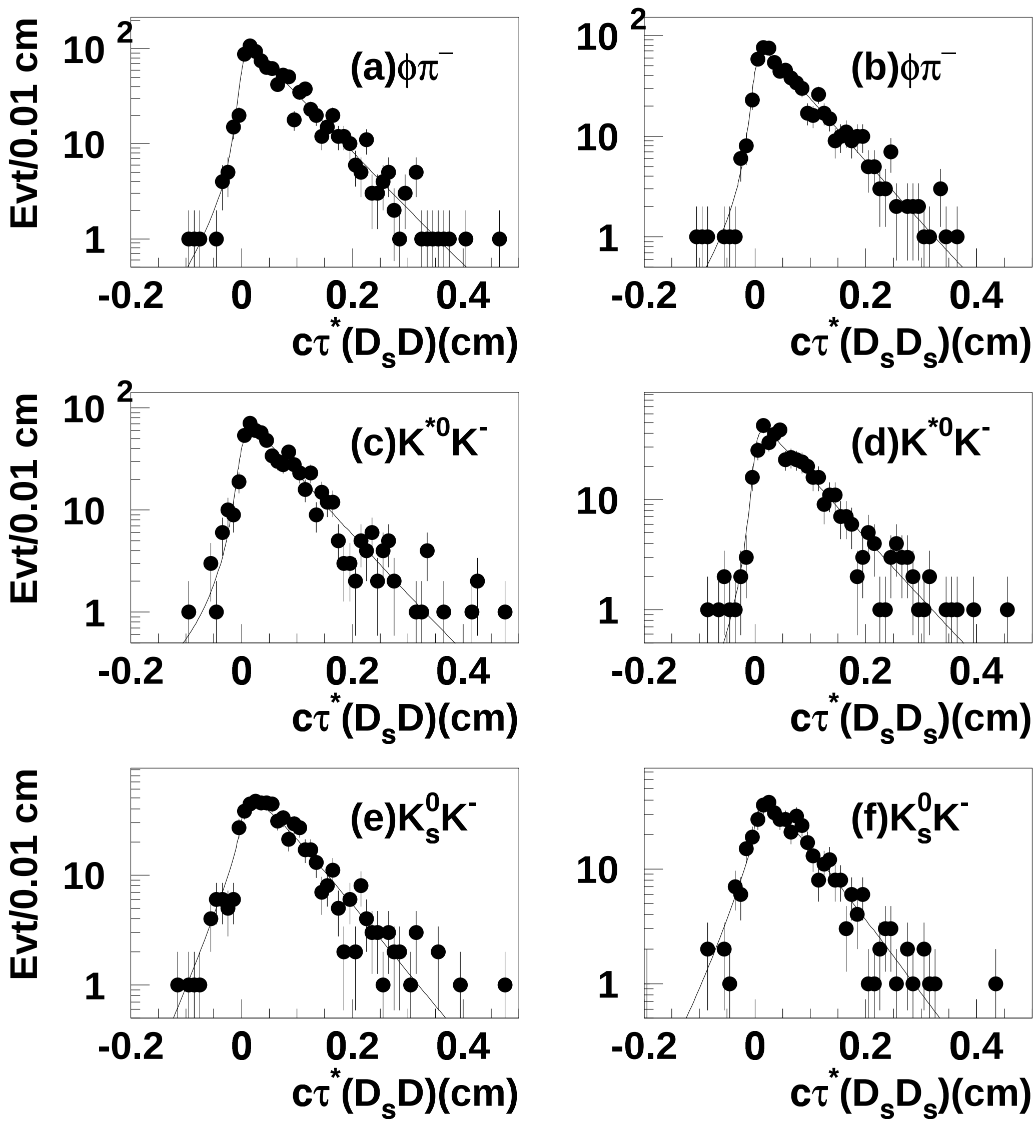


Figure 6.2: Pseudo-proper decay length distributions for Monte Carlo $B \to D_S^{(*)}D^{(*)}$ and $B_S^0 \to D_S^{(*)}D_S^{(*)}$ decays for the three D_S^- decay modes. The fit results have been superimposed and the fitted $c\tau$ values are listed in Table 6.2.

6.3 Inclusion of D^- Reflection

In the previous chapter, the presence in the $D_S^- \to K^{*0}K^-$ and $D_S^- \to K_s^0K^-$ samples of events which come from misidentified $D^- \to K^{*0}\pi^-$ and $D^- \to K_s^0\pi^-$ decays was described. The contribution of these events to the pseudo-proper decay length distribution must be included in the likelihood function.

The inclusion of these events is handled in a similar manner to that described above for the physics backgrounds. The fractions of D^- events in the D_S^- signal samples were measured in Chapter 5, and the D^- fractions are

$$\begin{aligned} f_{D^-} &= 35.3\ \% \quad (D_S^- \to K^{*0}K^-), \\ f_{D^-} &= 30.5\ \% \quad (D_S^- \to K_s^0K^-). \end{aligned}$$

To include the contribution of these events in the likelihood, the term

$$f_{D^-} \cdot \frac{K}{c\tau(B)} \cdot exp(\frac{-Kx}{c\tau(B)}) \otimes G(x, s\sigma) \otimes H(K)$$

is added, where f_{D^-} is fixed at the measured value for each sample and the average B lifetime of 468 μm [52] is used for $c\tau(B)$.

6.4 Final Lifetime Results

To obtain the measurement of $c\tau(B_S^0)$, a simultaneous fit to the signal and background samples in each of the three D_S^- samples is performed. Each D_S^- decay mode is first fit separately, using the values for the physics background contributions and D^- fractions listed above. The results of the fits for the B_S^0 lifetime in the individual samples are

$$\begin{aligned} c\tau(B_S^0) &= 427\ ^{+41}_{-37}\ \mu m \qquad (\phi\pi^-), \\ c\tau(B_S^0) &= 433\ ^{+79}_{-68}\ \mu m \qquad (K^{*0}K^-), \\ c\tau(B_S^0) &= 413\ ^{+164}_{-149}\ \mu m \qquad (K_s^0K^-), \end{aligned}$$

where the errors listed are statistical only. Figures 6.3(a), 6.4(a), and 6.5(a) show the pseudo-proper decay length distributions for the signal samples with the fit results superimposed.

Decay Mode	scale	$\lambda_+(\mu\text{m})$	f_+	$\lambda_-(\mu\text{m})$	f_-
$\phi\pi^-$	1.22	419	0.39	193	0.08
$K^{*0}K^-$	1.22	389	0.40	201	0.07
$K^0_s K^-$	1.40	566	0.56	211	0.17

Table 6.3: Fit results for the free parameters (excluding $c\tau$) in the B^0_S lifetime fit

The shaded curves represent the sum of the background probability function over the events in the signal sample. The dashed curves represent the sum of the signal probablity function over the same events. Figures 6.3(b), 6.4(b), and 6.5(b) show the pseudo-proper decay length distributions for the background samples with the fit results superimposed.

Table 6.3 shows the fit results for all free parameters in the B^0_S lifetime fit, excluding $c\tau$, listed separately for each decay mode. The parameters which describe the background pseudo-proper decay length distributions (f_+, λ_+, f_-, and λ_-) agree well in the $\phi\pi^-$ and $K^{*0}K^-$ samples, so those background distributions are quite similar. As expected, the background distribution for the $K^0_s K^-$ sample is somewhat different.

A combined fit to all three D_S^- decay modes is performed to extract the final value of $c\tau(B^0_S)$. All three samples are fit simultaneously, with the background distributions allowed to float separately for each decay mode. The signal lifetime, $c\tau$, and the scale factor are common to all three decay modes. The measured B^0_S lifetime is then

$$c\tau(B^0_S) = 427\ ^{+35}_{-32}\ \mu m$$

or

$$\tau(B^0_S) = 1.42\ ^{+0.12}_{-0.11}\ ps,$$

where the error shown is statistical only.

6.5 Systematic Uncertainties

The systematic uncertainty on the measurement of the B^0_S lifetime is estimated by evaluating the influence of many effects on the fitted $c\tau$. Contributions are evaluated individually and then summed in quadrature to obtain the final systematic uncertainty.

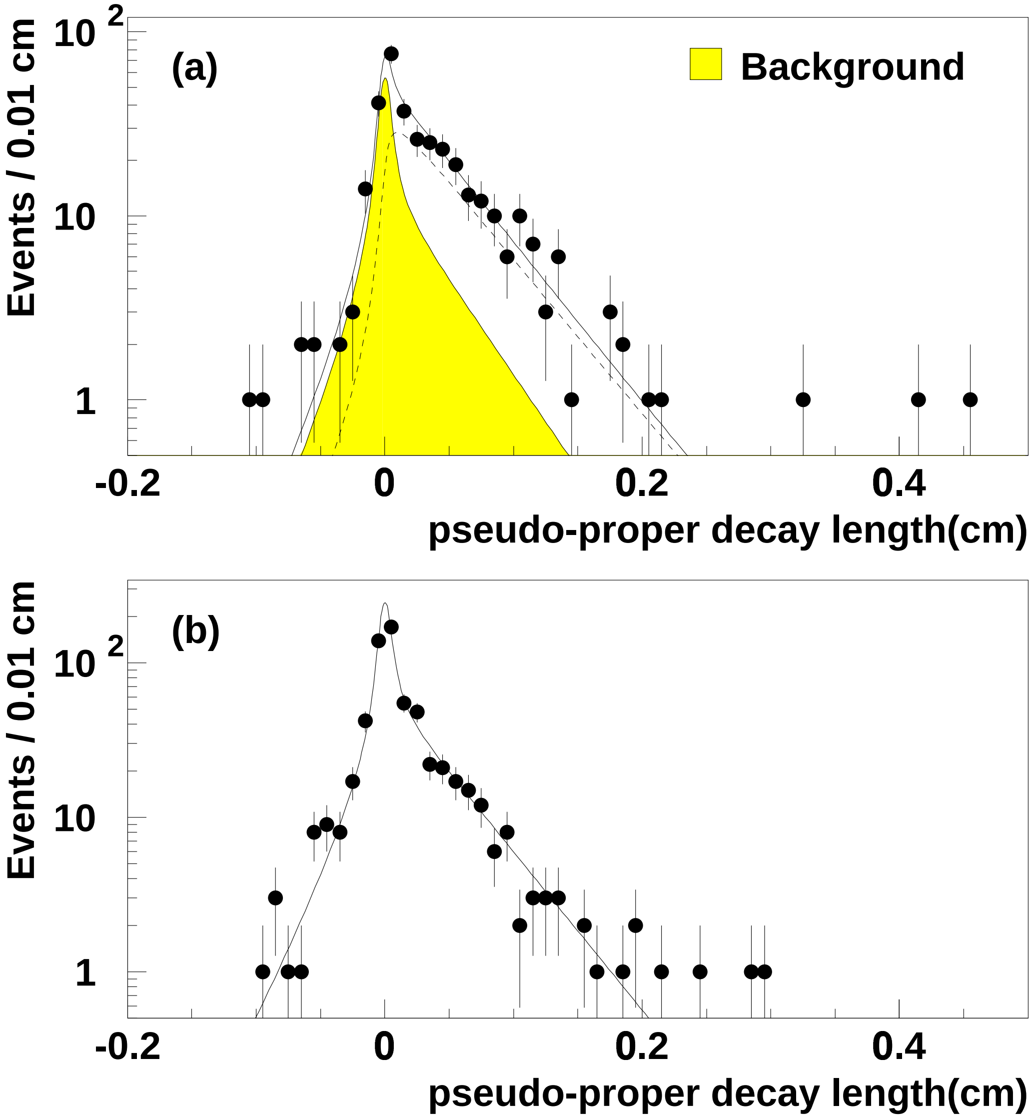


Figure 6.3: Pseudo-proper decay length distributions for (a) signal and (b) background in the $D_S^- \to \phi\pi^-$ sample. The shaded histogram in (a) represents the background contribution to the signal sample, and the dashed curve represents the signal contribution.

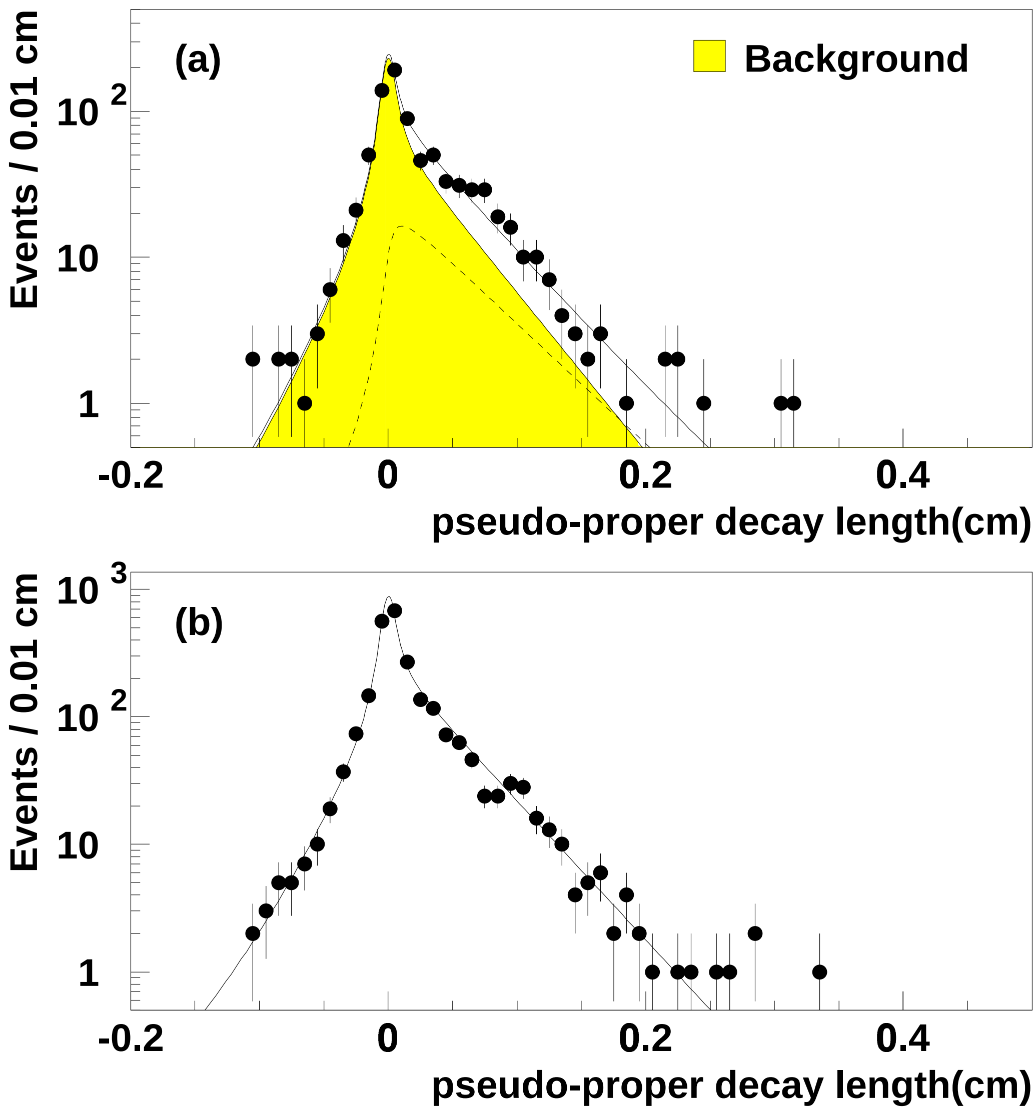


Figure 6.4: Pseudo-proper decay length distributions for (a) signal and (b) background in the $D_S^- \to K^{*0}K^-$ sample. The shaded histogram in (a) represents the background contribution to the signal sample, and the dashed curve represents the signal contribution.

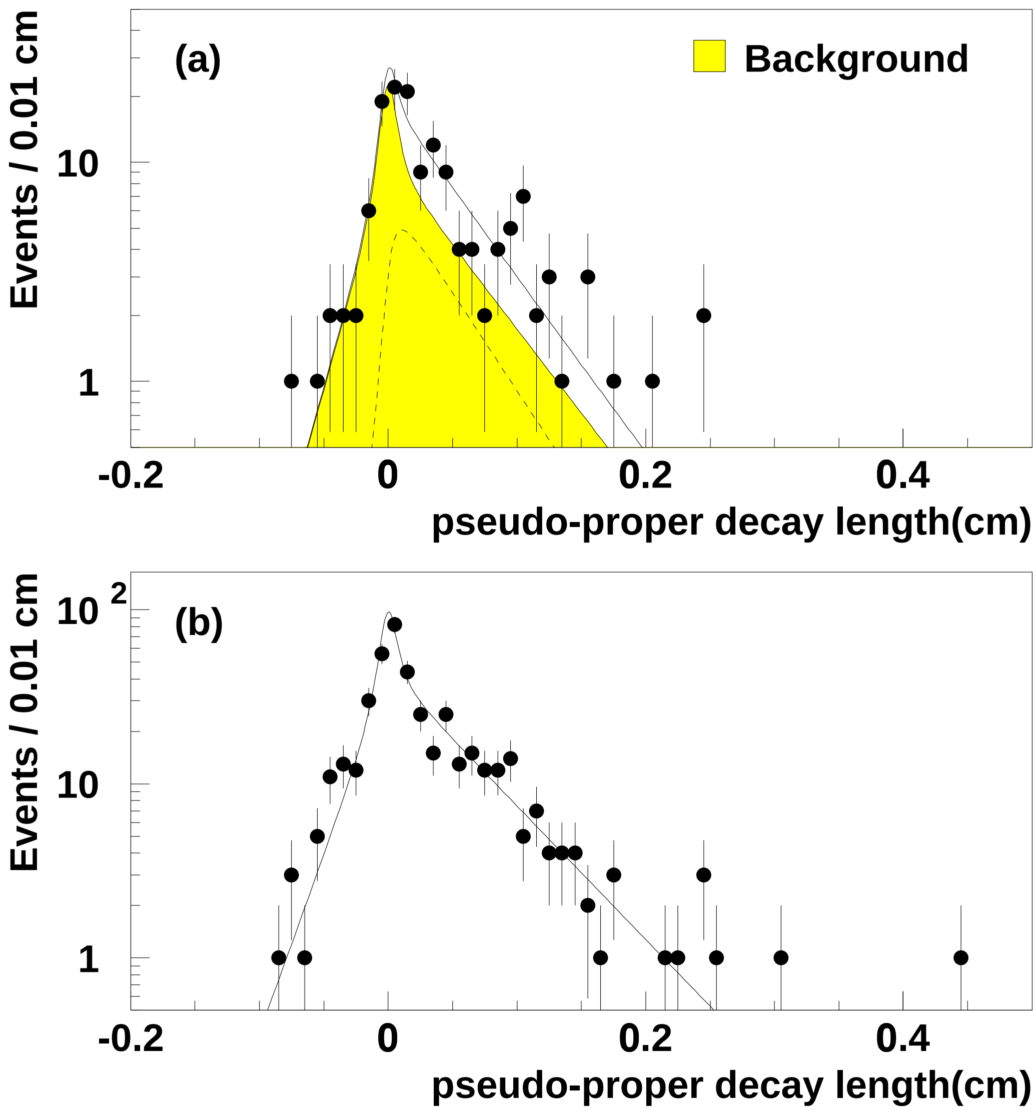


Figure 6.5: Pseudo-proper decay length distributions for (a) signal and (b) background in the $D_S^- \to K_s^0 K^-$ sample. The shaded histogram in (a) represents the background contribution to the signal sample, and the dashed curve represents the signal contribution.

6.5.1 *Background Shape*

The pseudo-proper decay length distribution of the combinatoric background is modelled by wrong-sign events and sidebands away from the D_S^- peak in right-sign events. The two samples are combined in the lifetime fit to increase the size of the background sample and decrease the possibility of statistical fluctuations distorting the background pseudo-proper decay length distribution.

To test the sensitivity of the fit to the shape of the background distribution, the lifetime fit is repeated using the right-sign and wrong-sign background samples separately. Using only the wrong-sign background sample gives $c\tau = 425$ μm, a change of $\Delta c\tau =$ -2 μm. Using only the right-sign background sample the fit result is $c\tau = 429$ μm, a shift of $\Delta c\tau =$ +2 μm. The results are shifted 2 μm in opposite directions from the central value, leading to the assignment of a systematic uncertainty of $\pm$ 2 μm for the background shape.

6.5.2 *D_S^- Fraction*

Because the D_S^- fraction changes the contributions of the $K^{*0}K^-$ and $K_s^0K^-$ samples, different input values for the D_S^- fraction can change the final lifetime result. To account for the effect of a mismeasurement of the D_S^- fractions, the input values of the D_S^- fractions in the B_S^0 lifetime fit are varied by the approximate errors on their fitted values and the lifetime fit is repeated.

In Chapter 5 the D_S^- fraction in the $K^{*0}K^-$ sample was measured to be $(60.6^{+16.5}_{-19.5})\%$ using the two-lifetime fit and $(68.7 \pm 19.4)\%$ using the fit to the mass distributions. These values were averaged to obtain 64.7% for the input to the lifetime fit. Based on the errors on the individual measurements, the D_S^- fraction for the $K^{*0}K^-$ sample is varied by $\pm$20% from a central value of 65%.

In the $K_s^0K^-$ sample the D_S^- fraction was measured to be $(92.7^{+7.3}_{-44.7})\%$ with the two-lifetime fit and $(46.3 \pm 35.4)\%$ with the fit to the mass distributions, which led to an average of 69.5% for the input to the lifetime fit. Using the errors on the two separate measurements would lead to a variation of the D_S^- fraction by $\pm 40\%$ from a central value of 70%. Since the D_S^- fraction must be less than 100%, it is varied from 30% to 100%.

Using D_S^- fractions of 85% and 100% for the $K^{*0}K^-$ and $K_s^0K^-$ samples, the fit result is $c\tau = 432 \pm 31$ μm, a shift of $\Delta c\tau = +5$ μm. Using D_S^- fractions of 45% and 30% the result is $c\tau = 428 \pm 35$ μm, a shift of $\Delta c\tau = +1$ μm. These results show that as the D_S^- fraction decreases, the final result is dominated by the measurement from the $D_S^- \to \phi\pi^-$ sample. The increase in the statistical error is due to the effective decrease of the sample size when more events are assumed to be coming from the D^- reflection.

Based on the measured shifts a systematic uncertainty of $^{+5}_{-0}$ μm is assigned for the D_S^- fraction.

6.5.3 Selection Bias

There are two cuts that could potentially bias the measured value of $c\tau(B_S^0)$. Requiring $L_{\text{xy}}(D_S^{PV}) > 0.0$ and $|c\tau(D_S^-)| < 0.1$ cm could bias the final result by altering the shape of the pseudo-proper decay length distribution. To study the effect of these cuts, Monte Carlo samples from all three D_S^- decay modes are used. The lifetimes are first fit in each sample after applying all of the selection cuts. Each of the two cuts is then removed individually, the fits are repeated, and the shifts of the fit results are recorded.

Removing only the $L_{\text{xy}}(D_S^{PV})$ requirement causes shifts of $\Delta c\tau = -5$ μm, -4 μm, and -5 μm, in the $\phi\pi^-$, $K^{*0}K^-$, and $K_s^0K^-$ samples, respectively. Removing only the $|c\tau(D_S^-)| < 0.1$ cm requirement results in shifts of $\Delta c\tau = +1$ μm, 0 μm, and $+1$ μm.

Using the largest of the individual variations, a systematic uncertainty of $^{+1}_{-5}$ μm is assigned for the possible selection bias.

6.5.4 $\beta\gamma$ Correction

The lifetime result is sensitive to the distributions of the correction factor K. These distributions are altered by variations in the momentum spectra of the B_S^0 decay products.

The first effect considered is the effect of the cut on lepton p_T and and the turn-on of the trigger, also based on the lepton p_T. The K distributions used in the standard fit were obtained from Monte Carlo where $p_T(e) > 7.5\ GeV/c$ was required and the shape of the turn-on of the trigger was simulated. To test the effect of the lepton p_T cuts on the final result, new K distributions are generated for $p_T(e) > 6.0$ where there is no simulation of the trigger turn-on and for $p_T(e) > 9.0$ where the shape of the trigger turn-on is included. Using these new K distributions lifetimes of $c\tau = 424\ \mu$m and $c\tau = 432\ \mu$m are obtained.

Since Monte Carlo electron events are used to generate the standard K distributions, the effects of the electron specific selection requirements must be considered, in particular the $E_{\text{had}}/E_{\text{em}}$ requirement. To examine these effects K distributions similar to those used in the standard fit are generated, except that muon samples are used rather than electron samples. Using K distributions from these samples, the fit result is unchanged.

Lastly, the possible effect of the D_S^{**} contribution must be considered. D_S^{**} can decay to D_S^- but the $p_T(D_S^-)$ spectrum will be different, so the fraction of semi-leptonic B_S^0 decays including a D_S^{**} must be considered. The D_S^{**} fraction of semi-leptonic B_S^0 decays is varied by $\pm 12\%$ from the default value of 36%, corresponding to the errors on the CLEO measurement of f^{**} that was assumed in the Monte Carlo. Since D_S^{**} decays to D_S^- only $\sim$10% of the time, this is expected to be a small effect. After generating new Monte Carlo samples with varied fractions of D_S^{**} decays, new K distributions are obtained. With 48% D_S^{**} the lifetime is shifted by $\Delta c\tau = +3\ \mu$m, and with 24% D_S^{**} the lifetime is shifted by $\Delta c\tau = +5\ \mu$m.

Based on all observed variations in the $c\tau$ results obtained from fits using different K distributions, a systematic uncertainty of $^{+5}_{-3}\ \mu$m is assigned for the $\beta\gamma$ correction.

6.5.5 Error Scale

The error scale factor was introduced in the fit to account for an overall under- or over-estimate of the event-by-event error on the pseudo-proper decay length. The fitted value of the error scale is 1.22 ± 0.03, indicating that the errors were under-estimated by 22%. To estimate the uncertainty in the measured lifetime due to the uncertainty in the scale factor, the value of the scale factor is fixed at $\pm\ 3\sigma$ from the fitted value and the lifetime fit is repeated. The fit results shift by $\Delta c\tau = +4\ \mu$m and $\Delta c\tau = -3\ \mu$m, leading to the assignment of a systematic uncertainty of $\pm$ 4 μm.

6.5.6 *Physics Backgrounds*

These backgrounds enter the lifetime fit as fixed fractions. Since the pseudo-proper decay length distribution of these events is different from the distribution from semi-leptonic B_S^0 decays, a mismeasurement of these background fractions will change the final lifetime result.

To test the effect of the background fractions on the measured lifetime, the values of the background fractions are varied between zero and twice the default values and the fits are repeated. First, $f_{D_S D}$ is doubled and the measured lifetime is shifted by $\Delta c\tau = -2\ \mu$m. Next $f_{D_S D_S}$ is doubled and the fit result remains unchanged. Doubling both $f_{D_S D}$ and $f_{D_S D_S}$ results in a shift of $\Delta c\tau = -2\ \mu$m.

Since the process $B \rightarrow D_S^{+(*)} K \ell^- X$ has never been observed, the fraction of events coming from this process in the signal sample is assumed to be zero. (See Chapter 4.) Based on the published limit from ARGUS[57], an upper limit of 0.03 is assumed for the background fraction from this process. Setting $f_{D_S D}$ and $f_{D_S D_S}$ to their nominal values and setting $f_{D_S K \ell}$ to 0.03 shifts the final result by $\Delta c\tau = -2\ \mu$m.

Lastly all of the background fractions are set to zero, resulting in a shift of $\Delta c\tau = +2\ \mu$m. Based on the observed variations in the fitted $c\tau$ when varying the physics background fractcions, a systematic uncertainty of $\pm$ 2 μm is assigned.

6.5.7 *Silicon Vertex Detector Alignment*

The procedure for aligning the SVX and the checks of the alignment have been described in detail in references [32], [33], and [49]. However, the effect of a residual mis-alignment of the SVX must be considered. In measuring the inclusive b lifetime at CDF [35], this effect was estimated by comparing the lifetime obtained when using alignment constants

Error Source	$\Delta c\tau(B_S^0)$ (μm)
Background Shape	± 2
D_S^- Fraction	$^{+5}_{-0}$
Selection Bias	$^{+1}_{-5}$
$\beta\gamma$ Correction	$^{+5}_{-3}$
Error Scale	± 4
Physics Backgrounds	± 2
Detector Alignment	± 2
Total	$^{+9}_{-8}$

Table 6.4: Systematic errors in the measurement of the B_S^0 lifetime.

derived with two different methods. The systematic uncertainty assigned for this effect was $\pm$ 2 μm, and the same uncertainty is included here.

6.5.8 Total Systematic Uncertainty

The individual contributions to the systematic uncertainty are added in quadrature, resulting in a total systematic uncertainty of $^{+9}_{-8}$ μm. The individual contributions are summarized in Table 6.4.

Chapter 7

Search for $\Delta\Gamma(B_S^0)$

In this chapter the difference in the decay widths of the CP eigenstates of the B_S^0 is considered. The theory behind the width difference in the B_S^0 system is reviewed first, and then the method of extracting $\Delta\Gamma$ is described. Using this technique the sensitivity for measuring $\Delta\Gamma$ in the current sample is considered, and then $\Delta\Gamma$ is measured in the data. Based on the measured value of $\Delta\Gamma$, limits on $\frac{\Delta\Gamma}{\Gamma}$, Δm_s, and the mixing parameter x_S can be set.

7.1 Theoretical Review of $\Delta\Gamma$

As explained briefly in Section 1.6, the two mass eigenstates of the B_S^0 meson, B_S^H and B_S^L (L = 'light', H = 'heavy'), are related to the flavor eigenstates by

$$
\begin{aligned}
|B_S^L\rangle &= \frac{1}{\sqrt{2}}(|B_S^0\rangle + |\overline{B}_S^0\rangle) \text{ and} \\
|B_S^H\rangle &= \frac{1}{\sqrt{2}}(|B_S^0\rangle - |\overline{B}_S^0\rangle).
\end{aligned}
$$

$m_{L,H}$ and $\Gamma_{L,H}$ denote the mass and decay width of B_S^L and B_S^H, leading to the definitions

$$
\begin{aligned}
\Delta m &= m_H - m_L, \\
\Delta\Gamma &= \Gamma_H - \Gamma_L, \text{ and} \\
\Gamma &= \frac{\Gamma_H + \Gamma_L}{2},
\end{aligned}
$$

where the B_S^0 is assumed to be an equal mixture of B_S^L and B_S^H. In the Standard Model of the CKM mixing matrix, the ratio $\Delta m/\Delta\Gamma$ contains no ratio of CKM matrix elements. It depends only on QCD corrections. If the error on this QCD calculation is understood and not too large, a measurement of $\Delta\Gamma$ implies a determination of Δm and thus a way to infer the mixing parameter x_S of the B_S^0 meson, given by $x_S = \Delta m/\Gamma$.

One way to measure $\Delta\Gamma$ is to describe the pseudo-proper decay length distribution from $B_S^0 \to D_S^- \ell^+ X$ decays by a sum of two exponentials, such as

$$(\Gamma + \frac{1}{2}\Delta\Gamma)e^{-(\Gamma+\frac{1}{2}\Delta\Gamma)t} \; + \; (\Gamma - \frac{1}{2}\Delta\Gamma)e^{-(\Gamma-\frac{1}{2}\Delta\Gamma)t}.$$

Before attempting to measure the difference in decay widths, the details of the system should first be considered.

7.2 Functional Form in Fit

Since the B_S^0 and $\overline{B}_S^0$ are assumed to be made up of the two states B_S^L and B_S^H with total widths Γ_L and Γ_H, respectively, the pseudo-proper decay length distributions should then be described by the exponential distributions

$$\Gamma_L e^{-\Gamma_L t} \qquad \text{and} \qquad \Gamma_H e^{-\Gamma_H t}.$$

Since the $B_S^0 \to D_S^- \ell^+ X$ decays are assumed to be an equal mix of B_S^L and B_S^H, the pseudo-proper decay length distribution should be described by the function

$$\frac{1}{2}[\Gamma_L e^{-\Gamma_L t} \; + \; \Gamma_H e^{-\Gamma_H t}].$$

Using $\Gamma_{L,H} = \Gamma \pm \frac{\Delta\Gamma}{2}$ this can be rewritten as

$$\frac{1}{2}[(\Gamma + \frac{\Delta\Gamma}{2})e^{-(\Gamma+\frac{\Delta\Gamma}{2})t} + (\Gamma - \frac{\Delta\Gamma}{2})e^{-(\Gamma-\frac{\Delta\Gamma}{2})t}].$$

Since the variable $\frac{\Delta\Gamma}{\Gamma}$ is more appropriate for comparison with theory than $\Delta\Gamma$, this expression is rewritten as

$$\frac{1}{2}[\Gamma(1 + \frac{1}{2}\frac{\Delta\Gamma}{\Gamma})e^{-\Gamma(1+\frac{1}{2}\frac{\Delta\Gamma}{\Gamma})t} + \Gamma(1 - \frac{1}{2}\frac{\Delta\Gamma}{\Gamma})e^{-\Gamma(1-\frac{1}{2}\frac{\Delta\Gamma}{\Gamma})t}],$$

so that Γ and $\frac{\Delta\Gamma}{\Gamma}$ can be obtained directly from the fit. Because this expression is symmetric about zero in $\Delta\Gamma$, there is no reason to allow $\Delta\Gamma < 0.0$. For this reason $\frac{\Delta\Gamma}{\Gamma}$ is restricted to be positive in the fit.

The width Γ in this expression is the mean Γ. However, the mean τ and the mean Γ are not reciprocals of each other, as it is normally assumed. Since the lifetime that is measured when fitting for a single lifetime is the mean lifetime, τ_{mean}, this is not the same as Γ^{-1}. This can be seen in the following, where τ_{mean} is defined by

$$\tau_{mean} = \frac{1}{2}(\tau_L + \tau_H).$$

Substituting $\tau_L = 1/\Gamma_L$ and $\tau_H = 1/\Gamma_H$,

$$\begin{aligned}
\tau_{mean} &= \frac{1}{2}(\frac{1}{\Gamma_L} + \frac{1}{\Gamma_H}) \\
&= \frac{1}{2}(\frac{1}{\Gamma + \frac{1}{2}\Delta\Gamma} + \frac{1}{\Gamma - \frac{1}{2}\Delta\Gamma}) \\
&= (\frac{\Gamma}{\Gamma^2 - \frac{1}{4}(\Delta\Gamma)^2}) \\
&= (\frac{1}{\Gamma(1 - \frac{1}{4}(\frac{\Delta\Gamma}{\Gamma})^2})
\end{aligned}$$

Thus τ_{mean} and Γ are related by the expression

$$\tau_{mean} \cdot \Gamma = (\frac{1}{1 - \frac{1}{4}(\frac{\Delta\Gamma}{\Gamma})^2}).$$

From this expression it can be seen that for small values of $\Delta\Gamma/\Gamma$, τ_{mean} and Γ^{-1} will be nearly the same. However, for large values of $\Delta\Gamma/\Gamma$ they will be quite different. Figure 7.1 illustrates the dependence of $\tau_{mean} \cdot \Gamma$ as a function of the ratio $\Delta\Gamma/\Gamma$. The points are from a toy Monte Carlo study. The plotted curve shows the function described above, which fits the Monte Carlo points very well, as expected.

Since the lifetime obtained in this fit for $\frac{\Delta\Gamma}{\Gamma}$ should be compared with the lifetime obtained when fitting for a single lifetime, the above expression is used to relate the measured τ_{mean} back to Γ. In the expression used to describe the signal as the sum of two exponentials,

$$\Gamma(1 \pm \frac{1}{2}\frac{\Delta\Gamma}{\Gamma}) \qquad \text{is replaced by} \qquad \frac{1}{\tau}(\frac{1}{1 - \frac{1}{4}(\frac{\Delta\Gamma}{\Gamma})^2})(1 \pm \frac{1}{2}\frac{\Delta\Gamma}{\Gamma}).$$

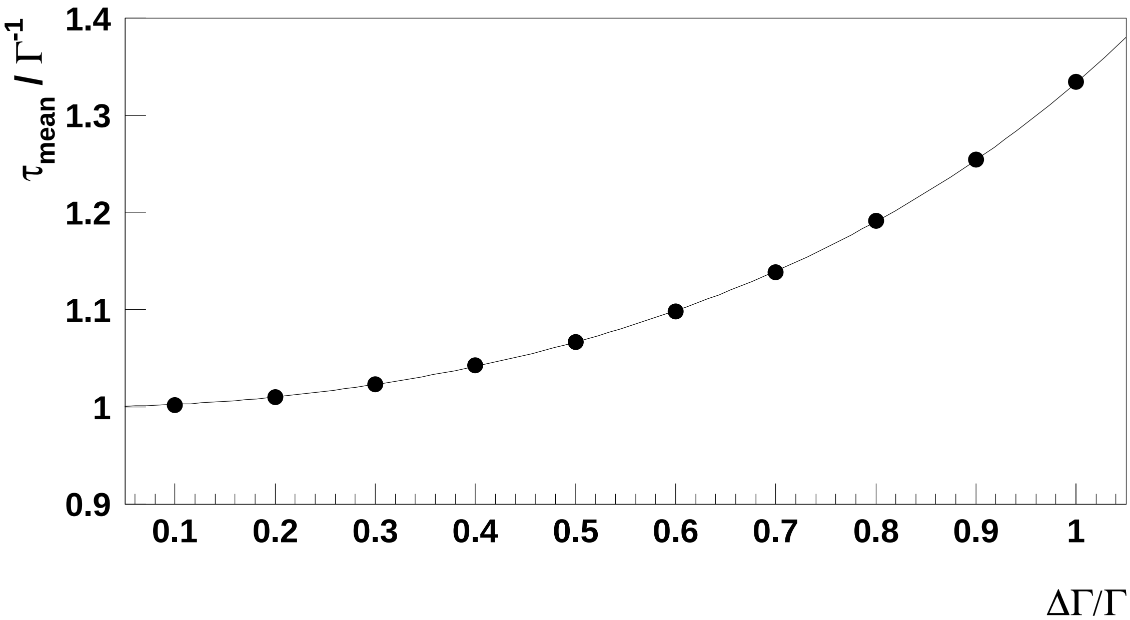


Figure 7.1: Illustration of the dependence of the ratio τ_{mean}/Γ^{-1} on $\frac{\Delta\Gamma}{\Gamma}$. The points are from a toy Monte Carlo study, and the plotted curve shows the expected dependence.

The final form of the expression used to describe the signal is then

$$\frac{1}{2}\left[\frac{1}{\tau}\left(\frac{1}{1-\frac{1}{4}(\frac{\Delta\Gamma}{\Gamma})^2}\right)\left(1+\frac{1}{2}\frac{\Delta\Gamma}{\Gamma}\right)e^{-\frac{1}{\tau}\left(\frac{1}{1-\frac{1}{4}(\frac{\Delta\Gamma}{\Gamma})^2}\right)\left(1+\frac{1}{2}\frac{\Delta\Gamma}{\Gamma}\right)t} + \right.$$
$$\left.\frac{1}{\tau}\left(\frac{1}{1-\frac{1}{4}(\frac{\Delta\Gamma}{\Gamma})^2}\right)\left(1-\frac{1}{2}\frac{\Delta\Gamma}{\Gamma}\right)e^{-\frac{1}{\tau}\left(\frac{1}{1-\frac{1}{4}(\frac{\Delta\Gamma}{\Gamma})^2}\right)\left(1-\frac{1}{2}\frac{\Delta\Gamma}{\Gamma}\right)t}\right],$$

and τ_{mean} and $\frac{\Delta\Gamma}{\Gamma}$ can be obtained simultaneously from the fit.

7.3 Toy Monte Carlo Test of Sensitivity for $\frac{\Delta\Gamma}{\Gamma}$

A toy Monte Carlo is used to test the functional form of the signal in the fit and to evaluate the sensitivity to measuring $\Delta\Gamma$. Samples of 400 events each (roughly equal to the combined number of $D_S^-\ell^+$ events in all three D_S^- samples) are generated for $\frac{\Delta\Gamma}{\Gamma}$ in the range 0.1 to 1.0. Approximately 50% of each sample is generated with the longer lifetime and 50% is generated with the shorter lifetime. As a first step, purely exponential distributions are

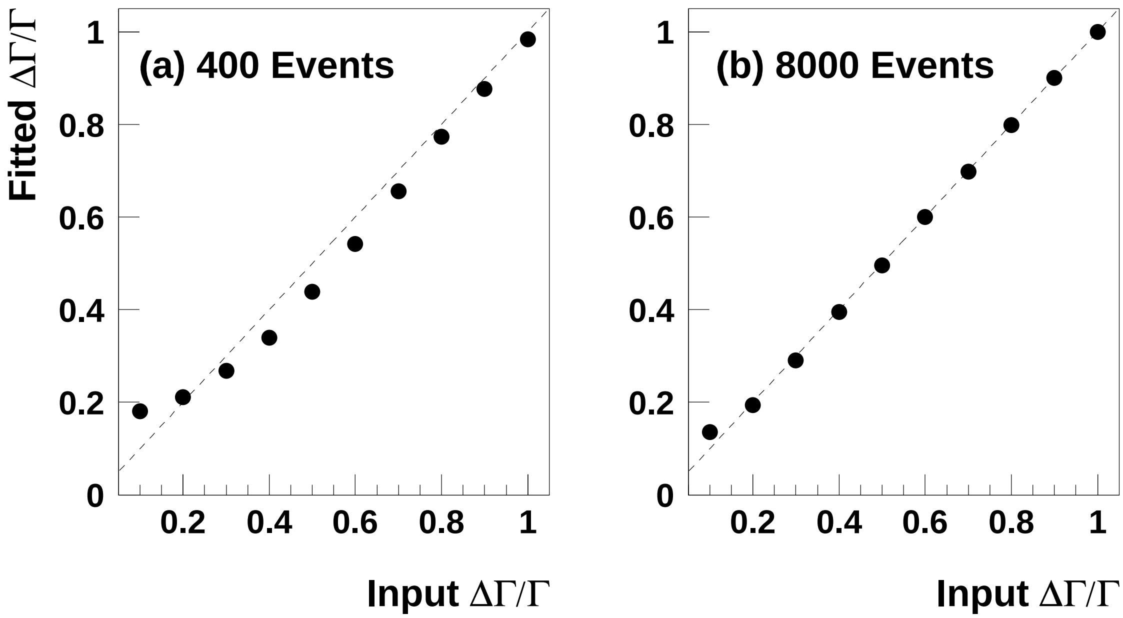


Figure 7.2: Mean fitted $\frac{\Delta\Gamma}{\Gamma}$ as a function of the input $\frac{\Delta\Gamma}{\Gamma}$ in the toy Monte Carlo. Samples of (a) 400 events and (b) 8000 events are used. Note that the x-axis starts at $\Delta\Gamma/\Gamma = 0.05$.

generated with no smearing. In each sample both $\frac{\Delta\Gamma}{\Gamma}$ and τ_{mean} are obtained from the fit. The fitted values as well as their errors are recorded for each sample.

Figure 7.2(a) shows the mean fitted $\frac{\Delta\Gamma}{\Gamma}$ from the toy Monte Carlo samples as a function of the input $\frac{\Delta\Gamma}{\Gamma}$. The points should lie along the diagonal line where the fitted($\frac{\Delta\Gamma}{\Gamma}$) = input($\frac{\Delta\Gamma}{\Gamma}$), represented by the dashed line in Figure 7.2(a). The points largely do lie on this line, except for a divergence at low values of $\frac{\Delta\Gamma}{\Gamma}$. The reason for this divergence is that the statistical size of the 400 event sample is insufficient to correctly resolve the lifetime difference at low values of $\frac{\Delta\Gamma}{\Gamma}$. This can be seen from Figure 7.2(b) which again shows the mean fitted $\frac{\Delta\Gamma}{\Gamma}$ as a function of the input $\frac{\Delta\Gamma}{\Gamma}$, now with each sample containing 8000 events. With the increased statistics of these samples, there is almost no divergence from the diagonal line at low $\frac{\Delta\Gamma}{\Gamma}$.

As a next step the exponential distributions in the toy Monte Carlo are smeared by a Gaussian for the resolution and by the K distribution. Samples of 400 events each are used again and $\frac{\Delta\Gamma}{\Gamma}$ and τ_{mean} are fit in each sample. Figure 7.3(a) shows the mean fitted $\frac{\Delta\Gamma}{\Gamma}$ as a function of the input $\frac{\Delta\Gamma}{\Gamma}$ for these samples. Note that the smearing of the exponentials has caused greater divergence from the expected diagonal line at low $\frac{\Delta\Gamma}{\Gamma}$. Figure 7.3(b) shows the mean average error on the fitted $\frac{\Delta\Gamma}{\Gamma}$. Here, the diagonal (represented by the dashed line) indicates where the error on the fitted($\frac{\Delta\Gamma}{\Gamma}$) is equal to the input($\frac{\Delta\Gamma}{\Gamma}$).

The place where the diagonal crosses the points in Figure 7.3(b) shows the point where $\frac{\Delta\Gamma}{\Gamma}$ can be resolved to 1σ from zero. Based on this plot it is concluded that with the current data sample, values of $\frac{\Delta\Gamma}{\Gamma} \gtrsim 50\%$ can be resolved at a 1σ level.

To improve the resolution on $\frac{\Delta\Gamma}{\Gamma}$, the value of τ_{mean} can be fixed in the fit. This implies precise knowledge of τ_{mean}, in which case the world average value of $\tau(B_S^0)$ should be used. Figure 7.3(c) shows the mean fitted $\frac{\Delta\Gamma}{\Gamma}$ as a function of the input $\frac{\Delta\Gamma}{\Gamma}$ when τ_{mean} has been fixed, and Figure 7.3(d) shows the mean average error on the fitted $\frac{\Delta\Gamma}{\Gamma}$. The divergence of the measured $\frac{\Delta\Gamma}{\Gamma}$ from the true value at low $\frac{\Delta\Gamma}{\Gamma}$ is decreased when τ_{mean} is fixed. The mean error on $\frac{\Delta\Gamma}{\Gamma}$ is also reduced. Based on Figure 7.3(d), it is estimated that values of $\frac{\Delta\Gamma}{\Gamma} \gtrsim 45\%$ can be resolved to 1σ from zero when τ_{mean} is fixed.

7.4 Fit for $\frac{\Delta\Gamma}{\Gamma}$ in Data

The fit is now applied to the data. As a first step, both $c\tau_{mean}$ and $\frac{\Delta\Gamma}{\Gamma}$ are allowed to float. A combined fit to all three D_S^- data samples is performed, with all input parameters fixed to the same values used in fitting the lifetime (see Chapter 6). The result of the fit to the data is

$$\begin{aligned} c\tau_{mean} &= 424\ ^{+36}_{-33}\ \mu m \\ \frac{\Delta\Gamma}{\Gamma} &= 0.45\ ^{+0.26}_{-0.45}. \end{aligned}$$

This result for $c\tau(B_S^0)$ is very close to the value of $427\ ^{+35}_{-32}$ μm obtained when fitting only the lifetime. If $c\tau_{mean}$ is fixed to the earlier measurement of 427 μm (see Chapter 6), the result is the same, $\frac{\Delta\Gamma}{\Gamma} = 0.45\ ^{+0.26}_{-0.45}$.

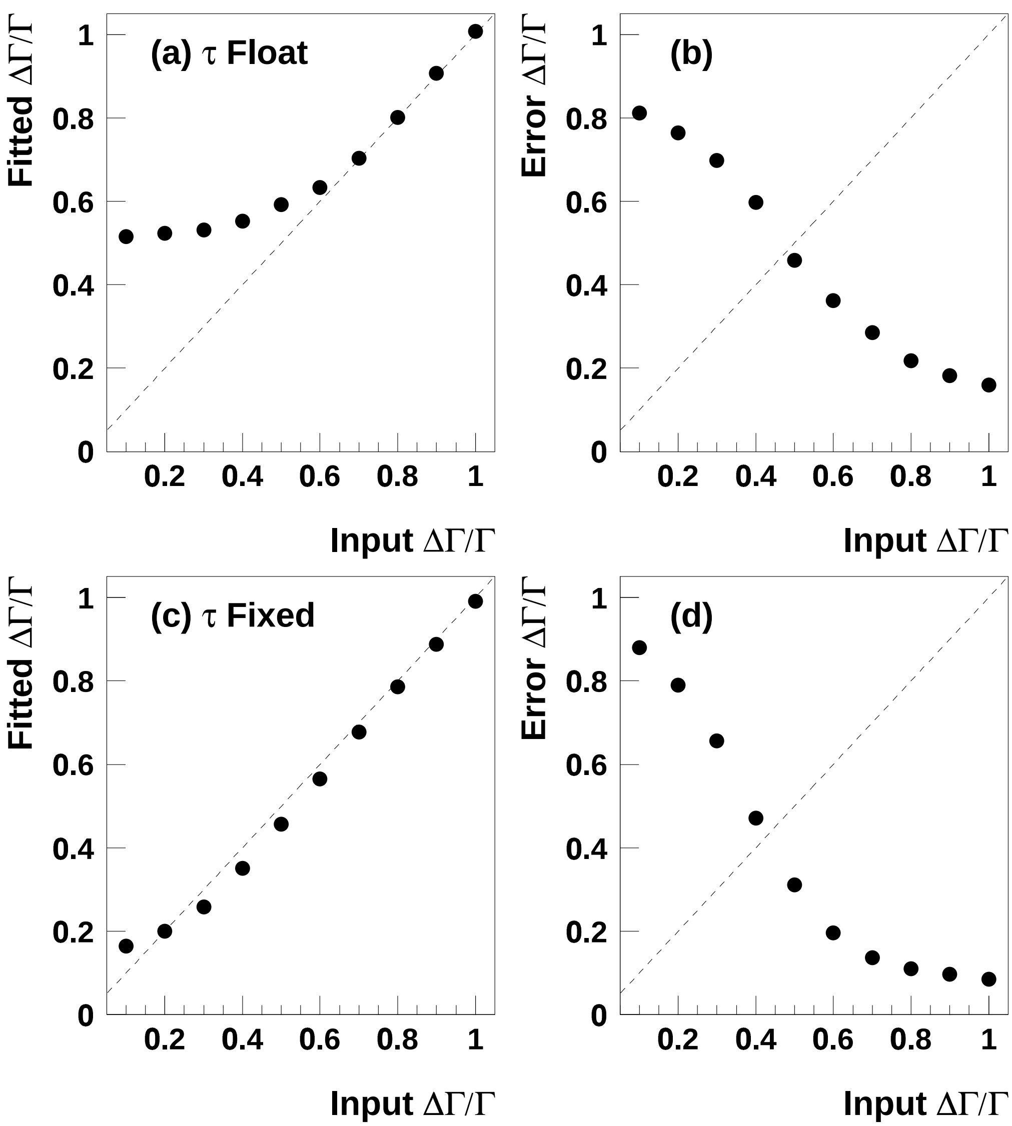


Figure 7.3: (a),(c) Mean fitted $\frac{\Delta\Gamma}{\Gamma}$ and (b),(d) the average fitted error on $\frac{\Delta\Gamma}{\Gamma}$ as a function of the input $\frac{\Delta\Gamma}{\Gamma}$ in the toy Monte Carlo. In (a) and (b) both $\frac{\Delta\Gamma}{\Gamma}$ and τ_{mean} are obtained from the fit, while in (c) and (d) τ is fixed.

Next a fit for $\frac{\Delta\Gamma}{\Gamma}$ is performed with $c\tau_{mean}$ fixed to the current PDG value, $c\tau(B_S^0) = 471 \pm 24$ μm [52]. Again a combined fit to all three D_S^- samples is performed, using the same input parameters as before. This time the fit result for $\frac{\Delta\Gamma}{\Gamma}$ is

$$\frac{\Delta\Gamma}{\Gamma} = 0.45 \ {}^{+0.27}_{-0.45}.$$

It should be noted that in this case the fit determines the number of $D_S^- \ell^+$ signal events to be smaller by a few %. For this reason, the uncertainty on $\frac{\Delta\Gamma}{\Gamma}$ does not improve by fixing τ_{mean} as expected from the toy Monte Carlo studies. Based on the above results it is concluded that the fit for $\frac{\Delta\Gamma}{\Gamma}$ is statistically limited by the current sample and thus it is not sensitive to $\Delta\Gamma$.

7.5 Limits on $\frac{\Delta\Gamma}{\Gamma}$ and x_S

Based on the fit results above, a limit can be set on $\frac{\Delta\Gamma}{\Gamma}$. Figure 7.4(a) shows $-2\ln(L)$ vs. $\frac{\Delta\Gamma}{\Gamma}$ obtained from a scan of the likelihood from the fit with $c\tau_{mean}$ fixed to the world average value. Though the plot extends only to $\frac{\Delta\Gamma}{\Gamma} = 1.0$, the fit extends to $\frac{\Delta\Gamma}{\Gamma} = 2.0$. It appears that the minimum is quite shallow. Only this region has been plotted to better examine the shape of the function near the minimum.

Figure 7.4(b) shows the normalized likelihood as a function of $\frac{\Delta\Gamma}{\Gamma}$. Again, the plot extends to only $\frac{\Delta\Gamma}{\Gamma} = 1.0$, though the fit covers the full allowed range of $\frac{\Delta\Gamma}{\Gamma}$. To find the 95% CL limit, this normalized likelihood is integrated to obtain

$$\frac{\Delta\Gamma}{\Gamma} < 0.83 \quad (95\% \text{ CL}),$$

as indicated by the arrow in Figure 7.4(b). Figure 7.4(c) shows this integrated likelihood as a function of $\frac{\Delta\Gamma}{\Gamma}$.

To set a limit on $x_S = \Delta m_s/\Gamma(B_S^0)$, the calculated value

$$\frac{\Delta\Gamma}{\Delta m} = (5.6 \ \pm \ 2.6) \ \times \ 10^{-3}$$

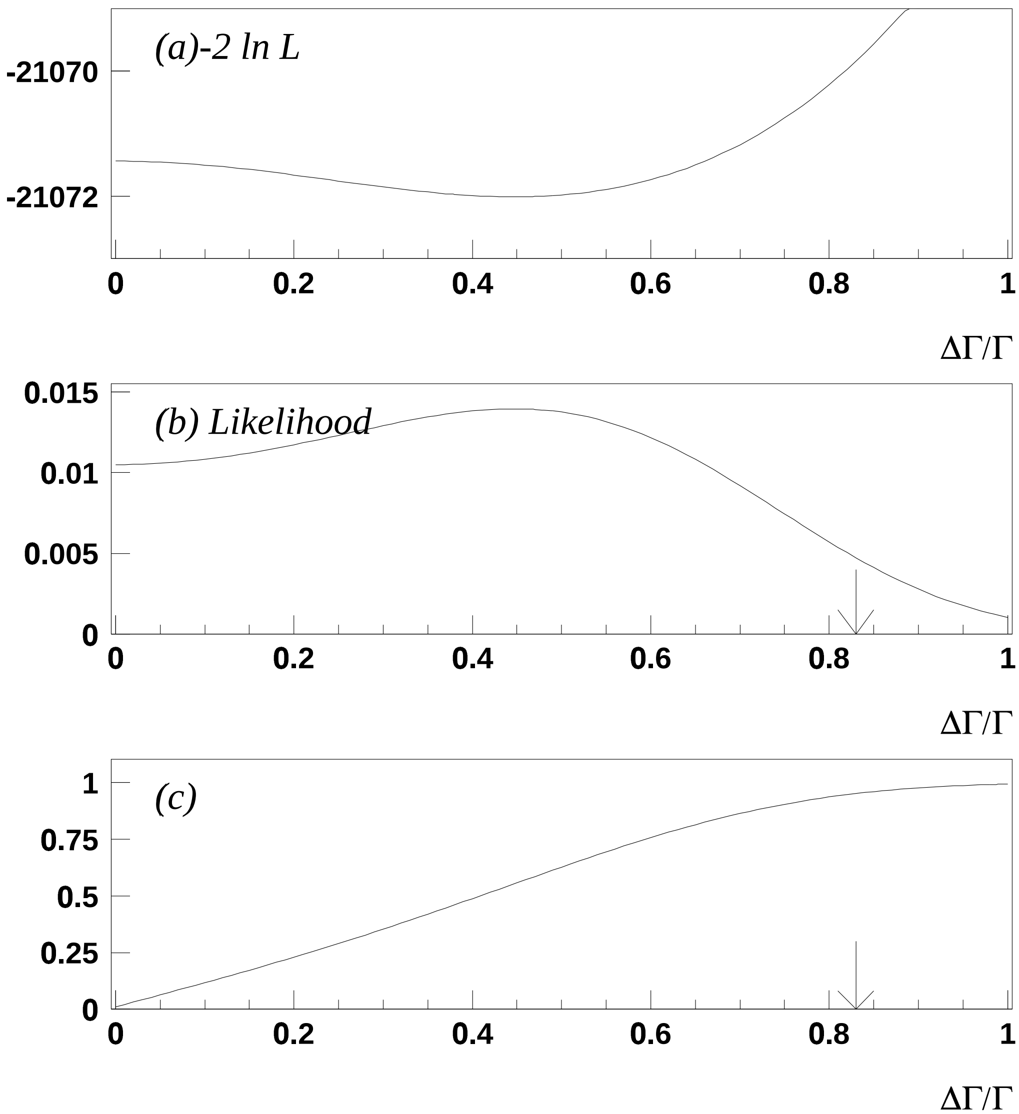


Figure 7.4: (a) $-2\ln(L)$ vs. $\frac{\Delta\Gamma}{\Gamma}$, (b) Normalized L vs. $\frac{\Delta\Gamma}{\Gamma}$, and (c) $\int L$ vs. $\frac{\Delta\Gamma}{\Gamma}$, when τ_{mean} is fixed. The plots extend to only 1.0 though the fit covers the full range, $0 < \frac{\Delta\Gamma}{\Gamma} < 2$. The arrows in (b) and (c) show the 95% CL limit based on the normalized L.

from Beneke, Buchalla, and Dunietz [59] is used. Using the central value gives the limit

$$x_S \; < \; 148 \; \cdot \; \left(\frac{5.6 \times 10^{-3}}{\Delta\Gamma/\Delta m} \right) \quad (95\% \text{ CL}),$$

where only the error on $\frac{\Delta\Gamma}{\Gamma}$ has been considered.

Finally, the limit on x_S is converted to a limit on Δm_s using the PDG value of $\tau(B_S^0)$ that was used in the fit for $\frac{\Delta\Gamma}{\Gamma}$. The limit on Δm_s is then

$$\Delta m_s \; < \; 94 \text{ ps}^{-1} \cdot \left(\frac{5.6 \times 10^{-3}}{\Delta\Gamma/\Delta m} \right) \cdot \left(\frac{1.57 \; ps}{\tau(B_S^0)} \right) \quad (95\% \text{ CL}).$$

Chapter 8

Results and Conclusions

This thesis presents a measurement of the lifetime of the B_S^0 meson in semi-leptonic decays at CDF. The B_S^0 is partially reconstructed in the decay $B_S^0 \to D_S^- \ell^+ X$ where the D_S^- is reconstructed in three decay modes: $D_S^- \to \phi\pi^-$, with $\phi \to K^+K^-$; $D_S^- \to K^{*0}K^-$, with $K^{*0} \to K^+\pi^-$; and $D_S^- \to K_s^0 K^-$, with $K_s^0 \to \pi^+\pi^-$. Using a combined fit to all three D_S^- data samples, the lifetime of the B_S^0 meson is measured to be

$$c\tau(B_S^0) = 427 \; ^{+35}_{-32} \; ^{+9}_{-8} \; \mu m$$

or

$$\tau(B_S^0) = 1.42 \; ^{+0.12}_{-0.11} \; \pm 0.03 \; ps,$$

where the first errors are statistical and the second errors are systematic.

Figure 8.1 summarizes the current B_S^0 lifetime measurements from experiments around the world as of August 1997 [61]. The CDF measurement of the lifetime in the fully reconstructed decay $B_S^0 \to J/\psi\phi$ is also presented. The measurement of the lifetime presented here is included in this figure and it can be seen that it is now the most precise measurement of the lifetime from a single experiment, surpassing the ALEPH result where 277 $D_S^- \ell^+$ events were reconstructed in 7 D_S^- decay modes.

An attempt has been made to observe the width difference, $\Delta\Gamma$, between the CP eigenstates of the B_S^0 system, B_S^L and B_S^H. This could be observed by comparing the lifetime measured in semi-leptonic B_S^0 decays with the lifetime measured in the fully reconstructed

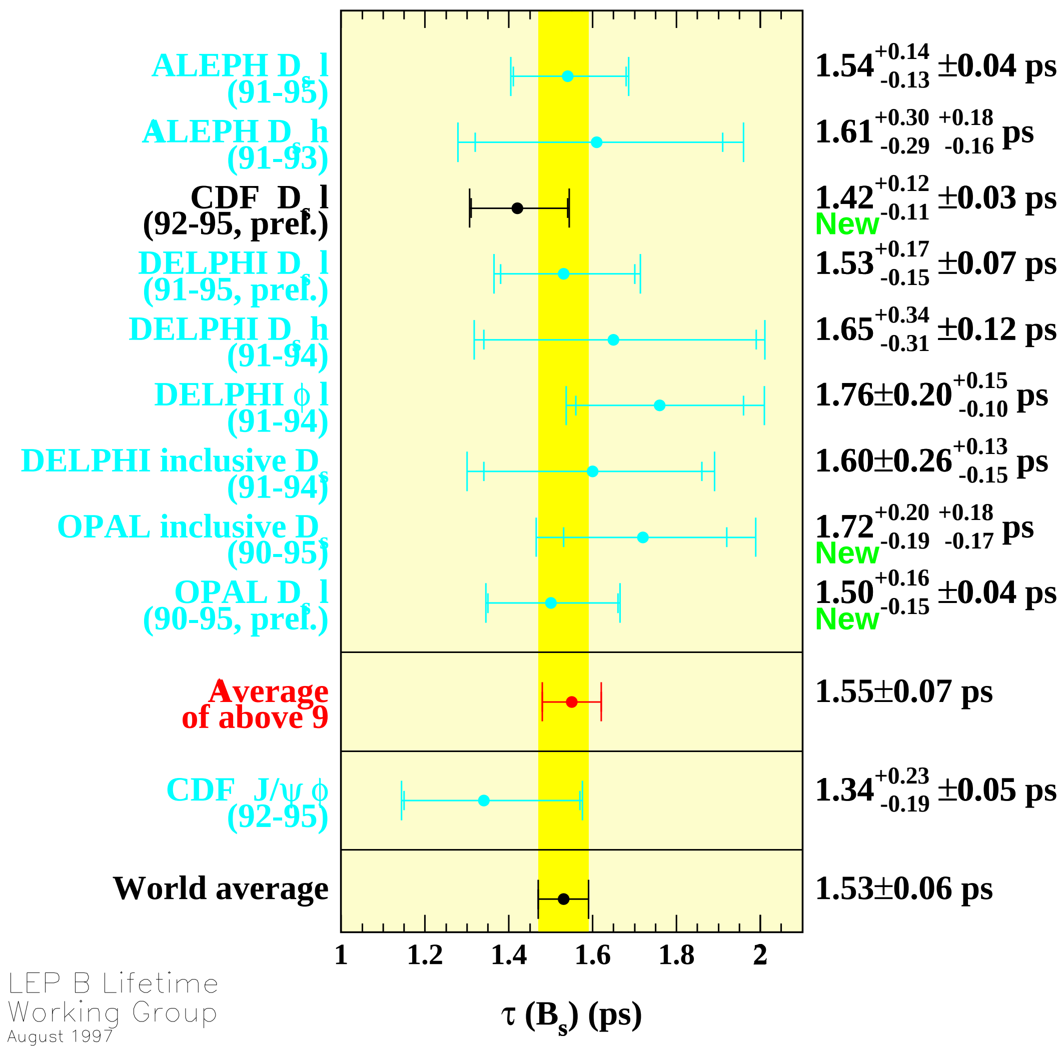


Figure 8.1: Survey of world B_S^0 lifetime measurements.

decay $B_S^0 \to J/\psi\phi$. Since the measurement presented in this paper agrees within errors with the published lifetime measured in the fully reconstructed sample, no attempt has been made to extract a limit on $\Delta\Gamma$ from this technique.

By attempting to fit the sum of two exponentials to the pseudo-proper decay length distribution from semi-leptonic B_S^0 decays, a limit on $\Delta\Gamma/\Gamma$ has been obtained. At 95% CL the limit

$$\frac{\Delta\Gamma}{\Gamma} < 0.83$$

is obtained. Some recent theoretical estimates for $\Delta\Gamma/\Gamma$ are

$$\begin{aligned} \frac{\Delta\Gamma}{\Gamma} &< 0.30 \text{ [23]}, \\ \frac{\Delta\Gamma}{\Gamma} &\geq 0.3 \pm 0.4 \text{ [59], and} \\ \frac{\Delta\Gamma}{\Gamma} &= 0.16\,^{+0.11}_{-0.09} \text{ [60]}. \end{aligned}$$

The result presented here shows no discrepancy with any of these estimates.

Combining the limit on $\Delta\Gamma/\Gamma$ with the calculated value $\frac{\Delta\Gamma}{\Delta m} = (5.6 \pm 2.6) \times 10^{-3}$, leads to the limit

$$x_S < 148 \cdot \left(\frac{5.6 \times 10^{-3}}{\Delta\Gamma/\Delta m}\right) \quad (95\% \text{ CL}).$$

Finally, using the the PDG value for $\tau(B_S^0)$ that was used in the fit for $\frac{\Delta\Gamma}{\Gamma}$ yields a limit on Δm_s,

$$\Delta m_s < 94 \text{ ps}^{-1} \cdot \left(\frac{5.6 \times 10^{-3}}{\Delta\Gamma/\Delta m}\right) \cdot \left(\frac{1.57\, ps}{\tau(B_S^0)}\right) \quad (95\% \text{ CL}).$$

Though the measurement of the B_S^0 lifetime can be improved slightly with increased statistics, the measurement of $\Delta\Gamma/\Gamma$ would greatly benefit from a larger sample. A factor of 20 increase in statistics is expected at CDF in Run II, which could allow probing $\Delta\Gamma/\Gamma \sim$ 20% or better. Further improvements could be gained by using fully reconstructed B_S^0 decays such as $B_S^0 \to D_S^- + \mathrm{n}\pi$, where the uncertainty in the $\beta\gamma$ correction is removed.

Appendices

Appendix A

Convolution of Functions in Likelihood

As explained in Chapter 4, the B_S^0 lifetime is determined from a simultaneous fit to the signal and background samples. The likelihood has two parts:

$$L = \prod_{i}^{N_{sig}} [(1 - f_{bkg})\mathcal{F}^i_{sig}) + f_{bkg}\mathcal{F}^i_{bkg}] \cdot \prod_{i}^{N_{bkg}} \mathcal{F}^i_{bkg},$$

where N_{sig} and N_{bkg} are the numbers of events in the signal and background samples, f_{bkg} is the estimated fraction of background events in the signal sample, and $\mathcal{F}_{sig}$ and $\mathcal{F}_{bkg}$ are the probability distributions of the signal and background.

$\mathcal{F}_{sig}$ consists of an exponential function, convoluted with the correction factor distribution, $H(K)$, convoluted with a gaussian resolution function. The functional form of $\mathcal{F}_{sig}$ is written

$$\mathcal{F}_{sig} = \frac{K}{c\tau} \cdot exp(\frac{-Kx}{c\tau}) \ \otimes \ G(x, s\sigma) \ \otimes \ H(K),$$

where x is the event-by-event pseudo-proper decay length, σ is the error on x, and s is a scale factor to account for an overall under- or over-estimate of the errors. $\mathcal{F}_{bkg}$ has three components. There is a gaussian centered at zero, and two exponentials for positive and negative lifetime backgrounds, each convoluted with gaussians. The functional form of $\mathcal{F}_{bkg}$ is written

$$\begin{aligned} \mathcal{F}_{bkg} &= (1 - f_+ - f_-)G(x, s\sigma) \\ &+ \frac{f_+}{\lambda_+} exp(\frac{-x}{\lambda_+}) \ \otimes \ G(x, s\sigma) \end{aligned}$$

$$+ \quad \frac{f_-}{\lambda_-} exp(\frac{x}{\lambda_-}) \ \otimes \ G(x, s\sigma),$$

where f_+ and f_- are the fractions of positive and negative lifetime backgrounds, and λ_+ and λ_- are the lifetimes of those backgrounds.

The symbol "$\otimes$" used in the equations above denotes a convolution. The convolution with the gaussian is performed analytically, but the convolution with the K distribution must be performed as a finite sum.

The exponential and gaussian functions, E and G, when properly normalized, have the explicit forms:

$$\begin{aligned} E(x) &= \frac{1}{c\tau} \cdot e^{-x/c\tau} \\ G(x,\sigma) &= \frac{1}{\sqrt{2\pi\sigma}} \cdot e^{-x^2/2\sigma^2}. \end{aligned}$$

Then the convolution in $\mathcal{F}_{bkg}$ and $\mathcal{F}_{sig}$ of the exponential with the gaussian is defined as:

$$E(x) \otimes G(x,\sigma) = \frac{1}{\sqrt{2\pi}\sigma c\tau} \int_{-\infty}^{x} e^{[-(x-x')/c\tau]} \ e^{[-x'^2/2\sigma^2]} \ dx'.$$

The exponent in this expression can then be rewritten.

$$\begin{aligned} \frac{-(x-x')}{c\tau} + \frac{-x'^2}{2\sigma^2} &= -\frac{1}{2\sigma^2 c\tau}[2\sigma^2(x-x') + c\tau x'^2] \\ &= -\frac{1}{2\sigma^2 c\tau}[2\sigma^2 x - 2\sigma^2 x' + c\tau x'^2 + \frac{\sigma^4}{c\tau} - \frac{\sigma^4}{c\tau}] \\ &= -\frac{x}{c\tau} + \frac{\sigma^2}{2(c\tau)^2} - \frac{1}{2\sigma^2} \cdot \left(x' - \frac{\sigma^2}{c\tau}\right)^2 \end{aligned}$$

Using this exponent gives

$$E(x) \otimes G(x,\sigma) = \frac{1}{\sqrt{2\pi}\sigma c\tau} \ e^{\frac{\sigma^2}{2(c\tau)^2} - \frac{x}{c\tau}} \int_{-\infty}^{x} e^{-\frac{1}{2}(\frac{x'}{\sigma} - \frac{\sigma}{c\tau})^2} dx'.$$

Now the substitutions $v = \frac{x'}{\sigma} - \frac{\sigma}{c\tau}$ and $dx' = \sigma dv$ are made, which yields

$$E(x) \otimes G(x,\sigma) = \frac{1}{\sqrt{2\pi}c\tau} \ e^{\frac{\sigma^2}{2(c\tau)^2} - \frac{x}{c\tau}} \int_{-\infty}^{\frac{x}{\sigma} - \frac{\sigma}{c\tau}} e^{-\frac{v^2}{2}} dv.$$

The normal frequency function is defined as:

$$\mathrm{freq}(x) \quad = \quad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Substituting this function for the integral in the previous expression gives

$$E(x) \otimes G(x,\sigma) \quad = \quad \frac{1}{c\tau}\, e^{\frac{\sigma^2}{2(c\tau)^2} - \frac{x}{c\tau}} \left[1 - \mathrm{freq}\left(\frac{\sigma}{c\tau} - \frac{x}{\sigma}\right)\right],$$

which can be substituted in the expressions for $\mathcal{F}_{sig}$ and $\mathcal{F}_{bkg}$. The normal frequency function is related to the error function, erf(x), by

$$\mathrm{freq}(x) \quad = \quad \begin{cases} \frac{1}{2} + \frac{1}{2}\mathrm{erf}(x/\sqrt{2}) & (x \geq 0) \\ \frac{1}{2}\mathrm{erfc}(|x|/\sqrt{2}) & (x < 0) \end{cases},$$

and thus it can be computed by rational Chebyshev approximation to the error function [62].

The other convolution in the likelihood is in $\mathcal{F}_{sig}$, where the exponential and gaussian distributions are convoluted with the K distribution $H(K)$. This convolution can not be handled in the same manner as before, but instead is replaced by a finite sum.

$$\int_0^{\infty} P(K)dK \longrightarrow \sum_i P_i(K_i),$$

where the K distribution P_i is given in the form of the normalized histograms shown in Figures 4.1 and 6.1. The P_i are then the probabilities for each value of K.

Using the convolution of the exponential and gaussian functions explained earlier, $\mathcal{F}_{sig}$ can be written

$$\mathcal{F}_{sig} = \frac{K}{c\tau} \cdot e^{\frac{(K\sigma)^2}{2(c\tau)^2} - \frac{Kx}{c\tau}} \left[1 - \mathrm{freq}\left(\frac{K\sigma}{c\tau} - \frac{x}{\sigma}\right)\right] \otimes\ H(K).$$

Substituting the sum over the K values for the convolution gives the final form of $\mathcal{F}_{sig}$ in the likelihood,

$$\mathcal{F}_{sig} = \sum_i P_i(K_i) \cdot \frac{K_i}{c\tau} \cdot e^{\frac{(K_i\sigma)^2}{2(c\tau)^2} - \frac{K_i x}{c\tau}} \left[1 - \mathrm{freq}\left(\frac{K_i\sigma}{c\tau} - \frac{x}{\sigma}\right)\right].$$

Appendix B

The CDF Collaboration

F. Abe,[17] H. Akimoto,[39] A. Akopian,[31] M. G. Albrow,[7] A. Amadon,[5] S. R. Amendolia,[27] D. Amidei,[20] J. Antos,[33] S. Aota,[37] G. Apollinari,[31] T. Arisawa,[39] T. Asakawa,[37] W. Ashmanskas,[18] M. Atac,[7] P. Azzi-Bacchetta,[25] N. Bacchetta,[25] S. Bagdasarov,[31] M. W. Bailey,[22] P. de Barbaro,[30] A. Barbaro-Galtieri,[18] V. E. Barnes,[29] B. A. Barnett,[15] M. Barone,[9] G. Bauer,[19] T. Baumann,[11] F. Bedeschi,[27] S. Behrends,[3] S. Belforte,[27] G. Bellettini,[27] J. Bellinger,[40] D. Benjamin,[35] J. Bensinger,[3] A. Beretvas,[7] J. P. Berge,[7] J. Berryhill,[5] S. Bertolucci,[9] S. Bettelli,[27] B. Bevensee,[26] A. Bhatti,[31] K. Biery,[7] C. Bigongiari,[27] M. Binkley,[7] D. Bisello,[25] R. E. Blair,[1] C. Blocker,[3] S. Blusk,[30] A. Bodek,[30] W. Bokhari,[26] G. Bolla,[29] Y. Bonushkin,[4] D. Bortoletto,[29] J. Boudreau,[28] L. Breccia,[2] C. Bromberg,[21] N. Bruner,[22] R. Brunetti,[2] E. Buckley-Geer,[7] H. S. Budd,[30] K. Burkett,[20] G. Busetto,[25] A. Byon-Wagner,[7] K. L. Byrum,[1] M. Campbell,[20] A. Caner,[27] W. Carithers,[18] D. Carlsmith,[40] J. Cassada,[30] A. Castro,[25] D. Cauz,[36] A. Cerri,[27] P. S. Chang,[33] P. T. Chang,[33] H. Y. Chao,[33] J. Chapman,[20] M. -T. Cheng,[33] M. Chertok,[34] G. Chiarelli,[27] C. N. Chiou,[33] L. Christofek,[13] M. L. Chu,[33] S. Cihangir,[7] A. G. Clark,[10] M. Cobal,[27] E. Cocca,[27] M. Contreras,[5] J. Conway,[32] J. Cooper,[7] M. Cordelli,[9] D. Costanzo,[27] C. Couyoumtzelis,[10] D. Cronin-Hennessy,[6] R. Culbertson,[5] D. Dagenhart,[38] T. Daniels,[19] F. DeJongh,[7] S. Dell'Agnello,[9] M. Dell'Orso,[27] R. Demina,[7] L. Demortier,[31] M. Deninno,[2] P. F. Derwent,[7] T. Devlin,[32] J. R. Dittmann,[6] S. Donati,[27] J. Done,[34] T. Dorigo,[25] N. Eddy,[20] K. Einsweiler,[18] J. E. Elias,[7] R. Ely,[18] E. Engels, Jr.,[28] W. Erdmann,[7] D. Errede,[13] S. Errede,[13] Q. Fan,[30] R. G. Feild,[41] Z. Feng,[15] C. Ferretti,[27] I. Fiori,[2] B. Flaugher,[7] G. W. Foster,[7] M. Franklin,[11] J. Freeman,[7] J. Friedman,[19] H. Frisch,[5] Y. Fukui,[17] S. Gadomski,[14] S. Galeotti,[27] M. Gallinaro,[26] O. Ganel,[35] M. Garcia-Sciveres,[18] A. F. Garfinkel,[29] C. Gay,[41] S. Geer,[7] D. W. Gerdes,[15] P. Giannetti,[27] N. Giokaris,[31] P. Giromini,[9] G. Giusti,[27] M. Gold,[22] A. Gordon,[11] A. T. Goshaw,[6] Y. Gotra,[25] K. Goulianos,[31] H. Grassmann,[36] L. Groer,[32] C. Grosso-Pilcher,[5] G. Guillian,[20] J. Guimaraes da Costa,[15] R. S. Guo,[33] C. Haber,[18] E. Hafen,[19] S. R. Hahn,[7] R. Hamilton,[11] T. Handa,[12] R. Handler,[40] F. Happacher,[9] K. Hara,[37] A. D. Hardman,[29] R. M. Harris,[7] F. Hartmann,[16] J. Hauser,[4] E. Hayashi,[37] J. Heinrich,[26] W. Hao,[35] B. Hinrichsen,[14] K. D. Hoffman,[29]

M. Hohlmann,[5] C. Holck,[26] R. Hollebeek,[26] L. Holloway,[13] Z. Huang,[20] B. T. Huffman,[28] R. Hughes,[23] J. Huston,[21] J. Huth,[11] H. Ikeda,[37] M. Incagli,[27] J. Incandela,[7] G. Introzzi,[27] J. Iwai,[39] Y. Iwata,[12] E. James,[20] H. Jensen,[7] U. Joshi,[7] E. Kajfasz,[25] H. Kambara,[10] T. Kamon,[34] T. Kaneko,[37] K. Karr,[38] H. Kasha,[41] Y. Kato,[24] T. A. Keaffaber,[29] K. Kelley,[19] R. D. Kennedy,[7] R. Kephart,[7] D. Kestenbaum,[11] D. Khazins,[6] T. Kikuchi,[37] B. J. Kim,[27] H. S. Kim,[14] S. H. Kim,[37] Y. K. Kim,[18] L. Kirsch,[3] S. Klimenko,[8] D. Knoblauch,[16] P. Koehn,[23] A. Köngeter,[16] K. Kondo,[37] J. Konigsberg,[8] K. Kordas,[14] A. Korytov,[8] E. Kovacs,[1] W. Kowald,[6] J. Kroll,[26] M. Kruse,[30] S. E. Kuhlmann,[1] E. Kuns,[32] K. Kurino,[12] T. Kuwabara,[37] A. T. Laasanen,[29] I. Nakano,[12] S. Lami,[27] S. Lammel,[7] J. I. Lamoureux,[3] M. Lancaster,[18] M. Lanzoni,[27] G. Latino,[27] T. LeCompte,[1] S. Leone,[27] J. D. Lewis,[7] P. Limon,[7] M. Lindgren,[4] T. M. Liss,[13] J. B. Liu,[30] Y. C. Liu,[33] N. Lockyer,[26] O. Long,[26] C. Loomis,[32] M. Loreti,[25] D. Lucchesi,[27] P. Lukens,[7] S. Lusin,[40] J. Lys,[18] K. Maeshima,[7] P. Maksimovic,[19] M. Mangano,[27] M. Mariotti,[25] J. P. Marriner,[7] A. Martin,[41] J. A. J. Matthews,[22] P. Mazzanti,[2] P. McIntyre,[34] P. Melese,[31] M. Menguzzato,[25] A. Menzione,[27] E. Meschi,[27] S. Metzler,[26] C. Miao,[20] T. Miao,[7] G. Michail,[11] R. Miller,[21] H. Minato,[37] S. Miscetti,[9] M. Mishina,[17] S. Miyashita,[37] N. Moggi,[27] E. Moore,[22] Y. Morita,[17] A. Mukherjee,[7] T. Muller,[16] P. Murat,[27] S. Murgia,[21] H. Nakada,[37] I. Nakano,[12] C. Nelson,[7] D. Neuberger,[16] C. Newman-Holmes,[7] C.-Y. P. Ngan,[19] L. Nodulman,[1] S. H. Oh,[6] T. Ohmoto,[12] T. Ohsugi,[12] R. Oishi,[37] M. Okabe,[37] T. Okusawa,[24] J. Olsen,[40] C. Pagliarone,[27] R. Paoletti,[27] V. Papadimitriou,[35] S. P. Pappas,[41] N. Parashar,[27] A. Parri,[9] J. Patrick,[7] G. Pauletta,[36] M. Paulini,[18] A. Perazzo,[27] L. Pescara,[25] M. D. Peters,[18] T. J. Phillips,[6] G. Piacentino,[27] M. Pillai,[30] K. T. Pitts,[7] R. Plunkett,[7] L. Pondrom,[40] J. Proudfoot,[1] F. Ptohos,[11] G. Punzi,[27] K. Ragan,[14] D. Reher,[18] M. Reischl,[16] A. Ribon,[25] F. Rimondi,[2] L. Ristori,[27] W. J. Robertson,[6] T. Rodrigo,[27] S. Rolli,[38] L. Rosenson,[19] R. Roser,[13] T. Saab,[14] W. K. Sakumoto,[30] D. Saltzberg,[4] A. Sansoni,[9] L. Santi,[36] H. Sato,[37] P. Schlabach,[7] E. E. Schmidt,[7] M. P. Schmidt,[41] A. Scott,[4] A. Scribano,[27] S. Segler,[7] S. Seidel,[22] Y. Seiya,[37] F. Semeria,[2] T. Shah,[19] M. D. Shapiro,[18] N. M. Shaw,[29] P. F. Shepard,[28] T. Shibayama,[37] M. Shimojima,[37] M. Shochet,[5] J. Siegrist,[18] A. Sill,[35] P. Sinervo,[14] P. Singh,[13] K. Sliwa,[38] C. Smith,[15] F. D. Snider,[15] J. Spalding,[7] T. Speer,[10] P. Sphicas,[19] F. Spinella,[27] M. Spiropulu,[11] L. Spiegel,[7] L. Stanco,[25] J. Steele,[40] A. Stefanini,[27] R. Ströhmer,[7a] J. Strologas,[13] F. Strumia, [10] D. Stuart,[7] K. Sumorok,[19] J. Suzuki,[37] T. Suzuki,[37] T. Takahashi,[24] T. Takano,[24] R. Takashima,[12] K. Takikawa,[37] M. Tanaka,[37] B. Tannenbaum,[22] F. Tartarelli,[27] W. Taylor,[14] M. Tecchio,[20] P. K. Teng,[33] Y. Teramoto,[24] K. Terashi,[37] S. Tether,[19] D. Theriot,[7] T. L. Thomas,[22] R. Thurman-Keup,[1] M. Timko,[38] P. Tipton,[30] A. Titov,[31] S. Tkaczyk,[7] D. Toback,[5] K. Tollefson,[19] A. Tollestrup,[7] H. Toyoda,[24] W. Trischuk,[14] J. F. de Troconiz,[11] S. Truitt,[20] J. Tseng,[19] N. Turini,[27] T. Uchida,[37] F. Ukegawa,[26] J. Valls,[32] S. C. van den Brink,[28] S. Vejcik, III,[20] G. Velev,[27] R. Vidal,[7] R. Vilar,[7a] D. Vucinic,[19] R. G. Wagner,[1] R. L. Wagner,[7] J. Wahl,[5] N. B. Wallace,[27] A. M. Walsh,[32] C. Wang,[6] C. H. Wang,[33] M. J. Wang,[33] A. Warburton,[14] T. Watanabe,[37] T. Watts,[32] R. Webb,[34] C. Wei,[6] H. Wenzel,[16] W. C. Wester, III,[7] A. B. Wicklund,[1] E. Wicklund,[7] R. Wilkinson,[26] H. H. Williams,[26] P. Wilson,[5] B. L. Winer,[23] D. Winn,[20] D. Wolinski,[20] J. Wolinski,[21] S. Worm,[22] X. Wu,[10] J. Wyss,[27] A. Yagil,[7] W. Yao,[18]

K. Yasuoka,[37] G. P. Yeh,[7] P. Yeh,[33] J. Yoh,[7] C. Yosef,[21] T. Yoshida,[24] I. Yu,[7] A. Zanetti,[36] F. Zetti,[27] and S. Zucchelli[2]

(CDF Collaboration)

[1] *Argonne National Laboratory, Argonne, Illinois 60439*

[2] *Istituto Nazionale di Fisica Nucleare, University of Bologna, I-40127 Bologna, Italy*

[3] *Brandeis University, Waltham, Massachusetts 02254*

[4] *University of California at Los Angeles, Los Angeles, California 90024*

[5] *University of Chicago, Chicago, Illinois 60637*

[6] *Duke University, Durham, North Carolina 27708*

[7] *Fermi National Accelerator Laboratory, Batavia, Illinois 60510*

[8] *University of Florida, Gainesville, FL 32611*

[9] *Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*

[10] *University of Geneva, CH-1211 Geneva 4, Switzerland*

[11] *Harvard University, Cambridge, Massachusetts 02138*

[12] *Hiroshima University, Higashi-Hiroshima 724, Japan*

[13] *University of Illinois, Urbana, Illinois 61801*

[14] *Institute of Particle Physics, McGill University, Montreal H3A 2T8, and University of Toronto, Toronto M5S 1A7, Canada*

[15] *The Johns Hopkins University, Baltimore, Maryland 21218*

[16] *Institut für Experimentelle Kernphysik, Universität Karlsruhe, 76128 Karlsruhe, Germany*

[17] *National Laboratory for High Energy Physics (KEK), Tsukuba, Ibaraki 305, Japan*

[18] *Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720*

[19] *Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

[20] *University of Michigan, Ann Arbor, Michigan 48109*

[21] *Michigan State University, East Lansing, Michigan 48824*

[22] *University of New Mexico, Albuquerque, New Mexico 87131*

[23] *The Ohio State University, Columbus, OH 43210*

[24] *Osaka City University, Osaka 588, Japan*

[25] *Universita di Padova, Istituto Nazionale di Fisica Nucleare, Sezione di Padova, I-35131 Padova, Italy*

[26] *University of Pennsylvania, Philadelphia, Pennsylvania 19104*

[27] *Istituto Nazionale di Fisica Nucleare, University and Scuola Normale Superiore of Pisa, I-56100 Pisa, Italy*

[28] *University of Pittsburgh, Pittsburgh, Pennsylvania 15260*

[29] *Purdue University, West Lafayette, Indiana 47907*

[30] *University of Rochester, Rochester, New York 14627*

[31] *Rockefeller University, New York, New York 10021*

[32] *Rutgers University, Piscataway, New Jersey 08855*

[33] *Academia Sinica, Taipei, Taiwan 11530, Republic of China*

[34] *Texas A&M University, College Station, Texas 77843*

[35] *Texas Tech University, Lubbock, Texas 79409*

[36] *Istituto Nazionale di Fisica Nucleare, University of Trieste/ Udine, Italy*

[37] *University of Tsukuba, Tsukuba, Ibaraki 315, Japan*
[38] *Tufts University, Medford, Massachusetts 02155*
[39] *Waseda University, Tokyo 169, Japan*
[40] *University of Wisconsin, Madison, Wisconsin 53706*
[41] *Yale University, New Haven, Connecticut 06520*

Bibliography

[1] R. Pinsky, *The Inferno of Dante: A New Verse Translation*, Farrar, Strauss, and Giroux (1994).

[2] B. Dylan, *My Back Pages*, Warner Bros. Inc. (1964).

[3] S.L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22**, 579-588 (1961).

[4] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 1264-1266 (1967).

[5] A. Salam, *Elementary Particle Physics*, ed. by N. Svarthold (Almqvist and Wiksells, Stockholm, 1968), p. 368.

[6] J.J. Aubert, et al, *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33**, 1404-1406 (1974).

[7] J.E. Augustin, et al, *Discovery of a Narrow Resonance in* $e^+ - e^-$ *Annihilation*, Phys. Rev. Lett. **33**, 1406-1408 (1974).

[8] M. Perl, et al, *Evidence for Anomalous Lepton Production in* $e^+ - e^-$ *Annihilation*, Phys. Rev. Lett. **35**, 1489-1492 (1975).

[9] M. Perl, et al, *Properties of Anomalous* $e - \mu$ *Events Produced in* $e^+ - e^-$ *Annihilation*, Phys. Lett. **63B**, 466 (1976).

[10] S.W. Herb, et al, *Observation of a Dimuon Resonance at 9.5 GeV/*c^2 *in 400 GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39**, 252-255 (1977).

[11] W.R. Innes, et al, *Observation of Structure in the Υ Region*, Phys. Rev. Lett. **39**, 1240 (1977).

[12] D. Andrews, et al, *Observation of a Fourth Υ State in $e^+ - e^-$ Annihilations*, Phys. Rev. Lett. **45**, 219 (1980).

[13] D. Andrews, et al, *Observation of the Υ''' at CESR*, Phys. Rev. Lett. **45**, 222 (1980).

[14] D. Buskulic, et al, (The ALEPH Collaboration) *First Measurement of the B_S^0 Meson Mass*, Phys. Lett. **B311**, 425-430 (1993).

[15] F. Abe, et al, (The CDF Collaboration) *Observation of the Decay $B_S^0 \rightarrow J/\psi\phi$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **71**, 1685-1689 (1993).

[16] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10**, 531-532 (1963).

[17] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49**, 652-657 (1973).

[18] L. Wolfenstein, *Parameterization of the Kobayashi- Maskawa Matrix*, Phys. Rev. Lett. **51**, 1945 (1983).

[19] A. Ali, D. London, *CKM Parameter Fits, the $B_S^0 - \overline{B}_S^0$ Mixing Ratio x_S and CP Violating Phases in B Decays*, DESY-93-022, Feb 1993.

[20] T.E. Browder, S. Pakvasa, *A Comment on the Experimental Determination of $|V_{ts}/V_{td}|^2$*, Phys. Rev. **D52**, 3123-3124 (1995).

[21] LEP B Oscillations Working Group. Results updated for the 1997 HEP-EPS conference, Jerusalem, 19-26 August 1997. (See URL: http://www.cern.ch/LEPBOSC/).

[22] D. Buskulic, et al, (The ALEPH Collaboration) *Search for B_S^0 oscillations using inclusive lepton events*, Submitted to the 1997 EPS-HEP conference, Jerusalem.

[23] I. Dunietz, B_S^0-$\overline{B}_S^0$ *Mixing, CP Violation and Extraction of CKM Phases from Untagged* B_S^0 *Data Samples*, Phys. Rev. Lett. **78**, 3257-3260 (1997).

[24] I. Dunietz and J. Rosner, *Time Dependent CP Violation Effects in* $B^0 - \bar{B}^0$ *Systems*, Phys. Rev. **D34**, 1404 (1986).

[25] A. Datta, E. A. Paschos, U. Turke, *Are There* B_S^0 *Mesons with Two Different Lifetimes?*, Phys. Lett. **196B**, 382 (1987).

[26] I. I. Bigi, CERN Report, CERN-TH.7282/94, May 1994; CERN Report CERN-TH.7050/93, October 1993.

[27] I. Bigi, B. Blok, M. Shifman, N. Uraltsev, A. Vainshtein, CERN Report, CERN-TH.7132/94, January 1994, in the second edition of the book *B Decays*, p. 132, ed. S. Stone, World Scientific, 1994.

[28] R. R. Wilson, *The Tevatron*, Phys. Today **10**, 23-30 (1977).

[29] Melvyn J. Shochet, *The Physics of* $p\bar{p}$ *Collisions*, FERMILAB-CONF-91-341-E, Dec 1991. Lectures given at Les Houches Summer School, Particles in the 90's, Les Houches, France, Jun 30 - Jul 26, 1991.

[30] F. Abe, et al, (The CDF Collaboration), *The CDF Detector: An Overview*, Nucl. Instr. and Meth. **A271**, 387-403 (1988).

[31] F. Bedeschi, et al, *Design and Construction of the CDF Central Tracking Chamber*, Nucl. Instr. and Meth. **A268**, 50-74 (1988).

[32] D. Amidei, et al, *The Silicon Vertex Detector of the Collider Detector at Fermilab*, Nucl. Instr. and Meth. **A350** 73 (1994).

[33] S. Cihangir, et al, *SVX′: The New CDF Silicon Vertex Detector*, Nucl. Instr. and Meth. **A360** 137 (1995).

[34] C. Miao, et al, *SVX′ Tracking Efficiency Study*, CDF Internal Note 2878, November 1994 (unpublished).

[35] F. Abe et al. (CDF Collaboration), *Measurement of B Hadron Lifetimes Using J/ψ Final States at CDF*, FERMILAB-PUB-97-352-E (Oct. 1997), Submitted to Phys. Rev. D.

[36] CDF Collaboration,
Measurement of the B^- and $\overline{B^0}$ Meson Lifetimes Using Semileptonic Decays.
(See URL http://www-cdf.fnal.gov/physics/new/bottom/cdf3786/cdf3786.html/).

[37] L. Balka, et al, *The CDF Central Electromagnetic Calorimeter*, Nucl. Instr. and Meth. **A267** 272-279 (1988).

[38] S. Bertolucci, et al, *The CDF Central and Endwall Hadron Calorimeter*, Nucl. Instr. and Meth. **A267** 301-314 (1988).

[39] Y. Fukui, et al, *CDF End Plug Electromagnetic Calorimeter Using Conductive Plastic Proportional Tubes*, Nucl.. Instr. and Meth. **A267**, 280-300 (1988).

[40] G. Brandenburg, et al, *An Electromagnetic Calorimeter for the Small Angle Regions of the Collider Detector at Fermilab*, Nucl. Instr. and Meth. **A267**, 257-271 (1988).

[41] *The CDF End Plug Hadronic Calorimeter*, internal note in progress, to be submitted to Nucl. Instr. and Meth.

[42] S. Cihangir, et al, *The CDF Forward/Backward Hadron Calorimeter* , Nucl. Instr. and Meth. **A267**, 249-256 (1988)

[43] G. Ascoli, et al, *CDF Central Muon Detector*, Nucl. Instr. and Meth. **A268** 33-40 (1988).

[44] A. Gauthier, et al, *Design and Performance of Drift Chambers for the Central Muon Upgrade of the CDF Detector*, to be submitted to Nucl. Instr. and Meth.

[45] E. Barsoti, et al, *Fastbus Data Acquisition for CDF*, Nucl. Instr. and Meth. **A269**, 82-92 (1988).

[46] D. Amidei, et al, *A Two Level Fastbus Based Trigger System for CDF*, Nucl. Instr. and Meth. **A269**, 51-62 (1988).

[47] J.T. Carrol, et al, *The CDF Level 3 Trigger*, Nucl. Instr. and Meth. **A300**, 552 (1991).

[48] G. W. Foster, et al, *A Fast Hardware Track Finder for the CDF Central Tracking Chambers*, Nucl. Instr. and Meth. **A269**, 93-100 (1988).

[49] C.H. Wang, et al, *Alignment of SVX′ Using Run Ib Data*, CDF Internal Note 3002, February 1995 (unpublished).

[50] W. Badgett, *Measurement of $\sigma \cdot B(W \to \mu\nu)$, $\sigma \cdot B(Z \to \mu^+\mu^-)$, and $R_\mu = \sigma \cdot B(W \to \mu\nu)/\sigma \cdot B(Z \to \mu^+\mu^-)$ and $\Gamma(W)$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV*, Ph.D. thesis, University of Michigan (1994).

[51] M. Peters, *Run Ib dE/dx Corrections*, CDF Internal Note 3807, August 1996 (unpublished).

[52] R.M. Barnett et al., Physical Review **D54**, (1996) 1, and 1997 off-year partial update for the 1998 edition available on the PDG WWW pages (URL: http://pdg.lbl.gov/).

[53] P. Nason, S. Dawson, and R. K. Ellis, *The Total Cross-Section for the Production of Heavy Quarks in Hadronic Collisions*, Nucl. Phys. **B303** (1988) 607; *The One Particle Inclusive Differential Cross-Section for Heavy Quark Production in Hadronic Collisions*, Nucl. Phys. **B327** (1988) 49.

[54] P. Avery, K. Read, G. Trahern, Cornell Internal Note CSN-212, March 25, 1985, (unpublished).

[55] N. Isgur, D. Scora, B. Grinstein, and M. Wise, *Semileptonic B and D Decays in the Quark Model*, Phys. Rev. **D39**, (1989) 799.

[56] R. Fulton et al. (CLEO Collaboration), *Exclusive and Inclusive Semileptonic Decays of B Mesons to D Mesons*, Phys. Rev. **D43**, (1991) 651.

[57] H. Albrecht, et al (ARGUS Collaboration) *Search for Rare B Meson Decays into D_S^- Mesons*, Z. Phys. **C60**, 11 (1993).

[58] F. Abe, et al (CDF Collaboration) *Measurement of the B_S^0 Lifetime*, Phys. Rev. Lett. **74** 4988 (1995).

[59] M.Beneke, G.Buchalla, and I. Dunietz, *The Width Difference in the B_S^0-$\overline{B}_S^0$ System*, Phys.Rev. **D54** 6782 (1996).

[60] M.Beneke, *Lifetime Difference of B_S^0 Mesons and its Implications*, SLAC-PUB-7258 (Aug. 1996), Presented at 28th International Conference on High-Energy Physics (ICHEP96), Warsaw, Poland, 25-31 Jul 1996.

[61] LEP B Lifetime Working Group.
Results updated for the 1997 HEP-EPS conference, Jerusalem, 19-26 August 1997.
(See URL: http://wwwcn.cern.ch/~claires/lepblife.html/).

[62] W.J. Cody, *Rational Chebyshev Approximations for the Error Function*, Math. Comp. **22** 631-637 (1969).