Measurement of the $B^0_S$ Meson Lifetime Using Semileptonic Decays


(CDF Collaboration)
Abstract

The lifetime of the $B^0_s$ meson is measured using the semileptonic decay $B^0_s \to D^- \ell^+ \nu X$. The data sample consists of about 110 pb$^{-1}$ of $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV collected by the CDF detector at Fermilab. Four different $D^- \ell^+$ decay modes are reconstructed resulting in approximately 600 $D^- \ell^+$ signal events. The $B^0_s$ meson lifetime is determined to be $\tau(B^0_s) = (1.36 \pm 0.09^{+0.06}_{-0.05})$ ps, where the first and second uncertainties are statistical and systematic, respectively. The $B^0_s$ meson decay length distribution is examined for a lifetime difference $\Delta \Gamma/\Gamma$ between the two mass eigenstates of the $B^0_s$ meson. An upper limit of $\Delta \Gamma/\Gamma < 0.83$ is set at 95% confidence level.

PACS numbers: 13.20 He, 13.25.Hw, 14.40.Nd
I. INTRODUCTION

The lifetime differences between different bottom flavored hadrons probe $B$ decay mechanisms which are beyond the simple quark spectator model. In the case of charm mesons, such differences have been observed to be quite large $(\tau(D^+)/\tau(D^0) \sim 2.5)$ [1]. Among bottom hadrons, the lifetime differences are expected to be smaller due to the heavier bottom quark mass [2,3]. Some QCD inspired models based on the heavy quark expansion [2] predict a difference between the $B^+$ and $B^0$ meson lifetimes of about 5% but expect the $B^0$ and $\bar{B}^0$ lifetimes to differ by no more than 1%. Although some assumptions in Ref. [2] have been questioned in Ref. [3], there is agreement that the models expect a difference between the lifetimes to differ by no more than 1%. These predictions are consistent with previous results of the $B^0$ and $B^+$ meson lifetimes, as well as recent $B^0_\text{s}$ lifetime measurements [1,4,5].

In the Standard Model [6], the $B^0_\text{s}$ meson exists in two $CP$-conjugate states, $|B^0_\text{s}\rangle = |\bar{b}s\rangle$ and $|\bar{B}^0_\text{s}\rangle = |bs\rangle$. The two mass eigenstates of the $B^0_\text{s}$ meson, $B^0_{\text{HS}}$ and $B^0_{\text{LS}}$ ($H$ = ‘heavy’ and $L$ = ‘light’), are not $CP$ eigenstates but are mixtures of the two $CP$-conjugate quark states:

$$|B^0_{\text{HS}}\rangle = p|B^0_\text{s}\rangle - q|\bar{B}^0_\text{s}\rangle \quad \text{and} \quad |B^0_{\text{LS}}\rangle = p|\bar{B}^0_\text{s}\rangle + q|B^0_\text{s}\rangle, \quad \text{with} \quad p^2 + q^2 = 1. \quad (1)$$

The mass and lifetime differences between the $B^0_{\text{HS}}$ and $B^0_{\text{LS}}$ can be defined as

$$\Delta m \equiv m_{H} - m_{L}, \quad \Delta \Gamma \equiv \Gamma_{L} - \Gamma_{H}, \quad \text{and} \quad \Gamma = \frac{\Gamma_{H} + \Gamma_{L}}{2}, \quad (2)$$

where $m_{H,L}$ and $\Gamma_{H,L}$ denote the mass and decay width of $B^0_{\text{HS}}$ and $B^0_{\text{LS}}$. Unlike in the case of the $B^0$ meson, the width difference in the $B^0_\text{s}$ system is expected to be large [7]. Theoretical estimates predict $\Delta \Gamma/\Gamma$ to be on the order of 10% to 20% [8,9]. In the $B^0_\text{s}$ system the ratio $\Delta m/\Delta \Gamma$ is related to the ratio of the Cabibbo-Kobayashi-Maskawa (CKM) [10] matrix elements $|V_{cb}V_{cs}|/|V_{ts}V_{tb}|$, which is quite well known, and depends only on QCD corrections within the Standard Model [9,11]. Currently these QCD corrections are known to next-to-leading order in the $1/m_b$ expansion [9]. A measurement of $\Delta \Gamma$ would therefore imply a determination of $\Delta m$ and thus a way to infer the existence of $B^0_\text{s}$ meson oscillations, which will ultimately determine the ratio of the CKM matrix elements $|V_{td}|/|V_{ts}|$.

It is assumed that $B^0_\text{s}$ mesons are produced as an equal mixture of $B^0_{\text{HS}}$ and $B^0_{\text{LS}}$ [9]. In a search for $\Delta \Gamma$, the $B^0_\text{s}$ meson decay length distribution can be described by a function of the form

$$\mathcal{F}(t) = e^{-\Gamma_{H}t} + e^{-\Gamma_{L}t} \quad \text{with} \quad \Gamma_{L,H} = \Gamma \pm \Delta \Gamma/2, \quad (3)$$

rather than by just one exponential lifetime $e^{-\Gamma t}$ which is the functional form used in the measurement of the $B^0_\text{s}$ lifetime assuming a single lifetime.

In this paper, we present an update of the $B^0_\text{s}$ lifetime measurement at CDF [5] using the semileptonic decay\(^1\) $B^0_\text{s} \rightarrow D^-_\text{s} \ell^+ \nu X$ ($\ell = e, \mu$), where the $D^-_\text{s}$ is identified via the four decay modes $D^-_\text{s} \rightarrow \phi \pi^-, K^+K^-, K^0\bar{K}^0, K^+_sK^-$.\(^1\) and $\phi \mu^-\nu$. We also examine the $B^0_\text{s}$ decay length

\(^1\)Throughout the paper references to a specific charge state imply the charge-conjugate state as well.

5
distribution for a lifetime difference $\Delta \Gamma/\Gamma$ with a fit to two exponential lifetimes. The data sample consists of approximately 110 pb$^{-1}$ of $p\bar{p}$ collisions at $\sqrt{s}=1.8$ TeV collected with the CDF detector during Run I. Of this, approximately 20 pb$^{-1}$ were collected during the 1992-93 running period, while about 90 pb$^{-1}$ were accumulated during the 1994-96 run of the Tevatron Collider. The result presented in this paper supersedes CDF’s previous measurement of the $B_s^0$ lifetime using semileptonic $B_s^0$ decays [5]. That publication was based on 20 pb$^{-1}$ of data and reconstructed the $D_s^-$ meson in the $\phi\pi^-\pi^+$ decay mode only.

The outline of this article is as follows: After a short description of the CDF detector in Section II, the selection of the $D_s^-\ell^+$ candidates is detailed in Section III. The determination of the $B_s^0$ lifetime is the topic of Section IV. We describe the search for a lifetime difference $\Delta \Gamma/\Gamma$ in Sec. V and offer our conclusions in Section VI.

II. THE CDF DETECTOR

The Collider Detector at Fermilab (CDF) is a multi-purpose detector designed to study 1.8 TeV $p\bar{p}$ collisions produced by the Fermilab Tevatron Collider. The detector has a coordinate system with the $z$-axis along the proton beam direction, the $y$-axis pointing vertically upwards, and the $x$-axis pointing horizontally out of the Tevatron ring. Throughout this article $\varphi$ is the azimuthal angle, $\theta$ is the polar angle measured from the proton direction, and $r$ is the radius perpendicular to the beam axis. The CDF detector is described in detail elsewhere [12]. We summarize here only the detector features most relevant to this analysis.

Three devices inside the 1.4 T solenoidal magnetic field are used for the tracking of charged particles: the silicon vertex detector (SVX), a set of vertex time projection chambers (VTX), and the central tracking chamber (CTC). The SVX [13] consists of four layers of silicon microstrip detectors located at radii between 2.9 cm and 7.9 cm from the interaction point. It provides spatial measurements in the $r$-$\varphi$ plane with a track impact parameter resolution of about $(13+40/p_T) \mu$m [13], where $p_T$ is the component of the track momentum $p$ transverse to the $z$-axis ($p_T = p \cdot \sin \theta$) given in GeV/c. The geometric acceptance of the SVX is about 60% as it covers only $\pm 25$ cm from the nominal interaction point whereas the luminous region of the Tevatron beam has an RMS of $\sim 30$ cm along the beam direction.

The VTX, which is located outside the SVX up to a radius of 22 cm, reconstructs track segments in the $r$-$z$ plane and is used to determine the $z$-position of the primary interaction vertex with a resolution of about 0.2 cm on average. Surrounding the SVX and VTX is the CTC, located between radii of 30 cm and 132 cm. The CTC is a 3.2 m long cylindrical drift chamber that contains 84 layers of sense wires grouped into nine alternating superlayers of axial and stereo wires with a stereo angle of $3^\circ$. The outer 54 layers of the CTC are instrumented to record the specific ionization $dE/dx$ of charged particles. The CTC covers the pseudorapidity interval $|\eta|$ less than about 1.1, where $\eta = -\ln[\tan(\theta/2)]$. The $p_T$ resolution of the CTC combined with the SVX is $\sigma(p_T)/p_T = [(0.0066)^2 + (0.0009 p_T)^2]^{1/2}$, with $p_T$ measured in GeV/c.

Outside the solenoid are electromagnetic (CEM) and hadronic (CHA) calorimeters ($|\eta| < 1.1$) that employ a projective tower geometry with a segmentation of $\Delta \eta \times \Delta \varphi \sim 0.1 \times 15^\circ$. The sampling medium is composed of scintillators layered with lead and steel absorbers. A layer of proportional wire chambers (CES) is located near shower maximum in the CEM and provides a measurement of electromagnetic shower profiles in both the $\varphi$- and $z$-directions.
The muon detection system has four of its layers of planar drift chambers (CMU) located beyond the central calorimeters. To reduce the probability of misidentifying penetrating hadrons as muon candidates in the pseudorapidity region $|\eta| \leq 0.6$, four more layers of chambers (CMP) are located outside the magnet return yoke. To reach these two detectors, particles produced at the primary interaction vertex with a polar angle of 90° must traverse material totaling 5.4 and 8.4 pion interaction lengths, respectively. An additional set of muon chambers (CMX) is located in the pseudorapidity interval $0.6 < |\eta| < 1.0$ to extend the polar acceptance of the muon system.

III. THE DATA SELECTION

In this section, we describe the data selection, which begins with the description of the lepton trigger data sets. This is followed by a summary of the selection requirements, which are applied to obtain the $D_S^- \ell^+$ candidate events used for the $B_S^0$ lifetime measurements. At the end of this section, we briefly describe the Monte Carlo simulation of our data.

A. The Lepton Trigger Data

Events containing semileptonic $B_S^0$ decays are collected using inclusive electron and muon trigger data as well as a data set obtained from a dimuon trigger. CDF uses a three-level trigger system. The first two levels are hardware based triggers, while Level 3 is a software trigger based on the offline reconstruction code optimized for computational speed.

At Level 1, inclusive electrons are selected by the presence of a single calorimeter tower above a threshold of 6-8 GeV depending on run conditions, while inclusive muons require the presence of a track in the CMU as well as the CMP. At Level 2, both of these triggers demand a charged track with $p_T > 7.5$ GeV/c reconstructed in the $r$-$\phi$ plane of the CTC by the central fast tracker (CFT) [14], a hardware track processor, which uses fast timing information from the CTC as input. The momentum resolution of the CFT is $\sigma(p_T)/p_T = 3.5\%$ with a high efficiency. In the case of the electron trigger, this track has to be matched to a cluster in the electromagnetic calorimeter with transverse energy $E_T > 8.0$ GeV, where $E_T = E \cdot \sin \theta$, with $E$ being the energy of the calorimeter cluster. In the case of the muon trigger, this track must be matched to a reconstructed track-segment in both the CMU and CMP. At Level 3, a computer farm is used to fully reconstruct the data including three-dimensional track reconstruction in the CTC. However, the fast algorithm used for tracking is only efficient for particles with $p_T > 1.4$ GeV/c. In the third level of the trigger more stringent electron and muon selection criteria, which are similar to those described in the next Section III B, are applied.

The reconstruction of the $D_S^- \rightarrow \phi \mu^- \nu$ decay mode is based on dimuon trigger data where the Level 1 trigger requires two muon candidates be observed in the muon system. The second level trigger requires the detection of at least one CFT track with $p_T > 2$ GeV/c to match in $\phi$ of each muon candidate. The third level trigger requires that two reconstructed CTC tracks are matched with two tracks in the muon chambers and the dimuon invariant mass is less than 2.8 GeV/c$^2$. During Run I about $7.5 \times 10^6$ electron trigger events and
about 2.5 × 10^6 inclusive muon trigger events were recorded by CDF, while about 1.3 × 10^6
dimuon trigger events with a dimuon invariant mass of less than 2.8 GeV/c^2 were recorded.

B. The Lepton Identification

The identification of electron candidates reconstructed after data collection uses information
from both the calorimeters and the tracking chambers. The longitudinal shower profile has to be consistent with an electron shower with a leakage energy from the CEM into the CHA of less than 4% if one track is pointing to the calorimeter tower or less than 10% if more than one track is pointing to the calorimeter tower. The lateral shower profile of the CEM cluster has to be consistent with that from test beam electrons. Additionally, a \( \chi^2 \) comparison of the CES shower profile with that of test beam electrons has to result in \( \chi^2 < 10 \). For the association of a single high \( p_T \) track with the calorimeter shower based on the position matching at the CES plane, it is required that \( r|\Delta \phi| < 1.5 \text{ cm} \) and \( |\Delta z \sin \theta| < 3 \text{ cm} \). In addition, we demand the \( E_T \) of the electron candidate reconstructed offline to be greater than 6 GeV. Electron candidates from photon conversion due to detector material are reduced to less than 10% by looking for oppositely charged tracks which have a small opening angle with the electron candidate.

The reconstruction of muon candidates is described in Ref. [15]. We compute a \( \chi^2 \) characterizing the separation between the track segment in the muon chamber and the extrapolated CTC track, where the uncertainty in this \( \chi^2 \) variable is dominated by the multiple scattering in the detector material. We require \( \chi^2 < 9 \) in the \( r-\phi \) view (CMU, CMP, and CMX) and \( \chi^2 < 12 \) in the \( r-z \) view (CMU and CMX). The transverse muon momentum reconstructed offline is required to be \( p_T > 6 \text{ GeV/c} \). For the dimuon sample this cut is \( p_T > 2 \text{ GeV/c} \) for each muon candidate. Finally, for optimal vertex resolution the electron and muon candidate tracks have to be reconstructed in the SVX detector.

C. The \( D_s^- \) Selection

The \( D_s^- \) candidates are reconstructed in the decay modes

(i) \( D_s^- \rightarrow \phi \pi^-, \phi \rightarrow K^+K^- \),
(ii) \( D_s^- \rightarrow K^{*0}K^-, K^{*0} \rightarrow K^+\pi^- \),
(iii) \( D_s^- \rightarrow K_S^0K^-, K_S^0 \rightarrow \pi^+\pi^- \),
(iv) \( D_s^- \rightarrow \phi \mu^-\nu, \phi \rightarrow K^+K^- \).

The \( D_s^- \) reconstruction is explained in the next section with the example of the \( \phi \pi^- \) decay channel. The other \( D_s^- \) decay channels (ii)-(iv) are reconstructed in a similar way. We then describe only the differences in the selection of these decays (ii)-(iv) as compared to the \( D_s^- \rightarrow \phi \pi^- \) decay mode. The kinematic selection criteria used in this analysis are optimized by maximizing the quantity \( N_S/\sqrt{N_S + N_B} \), where \( N_S \) is the predicted number of signal events based on Monte Carlo calculations (see Sec. III D) and \( N_B \) is the observed number of background events estimated from the \( D_s^- \) sideband regions (see Sec. IV E).
1. The $D^-_s \rightarrow \phi \pi^-$ Mode

The $D^-_s \rightarrow \phi \pi^-$ reconstruction starts with a search for $\phi \rightarrow K^+K^-$ candidates. We first define a search cone around the lepton candidate with a radius $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$ of 0.8. Any two oppositely charged tracks with $p_T > 1.2$ GeV$/c$ within that cone are assigned the kaon mass and combined to form a $\phi$ candidate. Each $\phi$ candidate is required to have a mass within $\pm 10$ MeV$/c^2$ of the world average $\phi$ mass $[1]$. The $\phi$ candidate is then combined with another track of $p_T > 0.8$ GeV$/c$ inside the cone which has the opposite charge of the lepton (we call this the ‘right-sign’ combination). This third track is assigned the pion mass.

To ensure a good decay vertex measurement, track quality cuts requiring a minimum number of hits in the CTC are imposed on the tracks forming the $D^-_s$ candidate. In addition, each of the tracks forming the $D^-_s$ is required to be reconstructed in the SVX with hits in at least three out of the four silicon layers and the $\chi^2$ of the track fit per SVX hit has to be less than 6 to reject badly measured tracks. The same track selection requirements are also applied to the lepton candidate tracks.

The specific ionization information $dE/dx$ from the CTC is used to help identify hadrons in the $D^-_s$ reconstruction. Because of the large Landau tail of the ionization distribution, the 80% truncated mean of the measured charges from the CTC sense wires is taken as the best estimator of the track $dE/dx$. The probabilities, $P(i)$, for a track to be consistent with the $i = e, \mu, \pi, K$, or $p$ hypotheses are then calculated using the measured $dE/dx$ value and the predictions for the assumed particle hypotheses. We define a likelihood ratio, $\ell h_{DE/dx}^K$, for a track being a kaon to be the ratio of $P(K)$ divided by the sum of the probabilities of all particle hypotheses. The quantity $\ell h_{DE/dx}^K$ is defined correspondingly. We require the likelihood ratios $\ell h_{DE/dx}^K > 0.01$ and $\ell h_{DE/dx}^\pi > 0.01$.

Since the $\phi$ has spin 1 and both the $D^-_s$ and $\pi^-$ are spin 0, the helicity angle $\Psi$, which is the angle between the $K^-$ and $D^-_s$ directions in the $\phi$ rest frame, exhibits a distribution $dN/d(\cos \Psi) \sim \cos^2 \Psi$. A cut $|\cos \Psi| > 0.4$ is therefore applied to suppress combinatorial background, which is found to be flat in the $\cos \Psi$ distribution. We also apply an isolation cut $E_T^{iso}/p_T(\phi\pi^-) < 1.0$ on the $D^-_s$ candidate, where $E_T^{iso}$ is the sum of the transverse energy within a cone of radius 0.4 in $\eta$-$\phi$ space around the lepton candidate, excluding the lepton energy. This cut eliminates many of the fake $D^-_s$ combinations from high track multiplicity jets. Finally, the mass of the $D^-_s \ell^+$ system is required to be between 3.0 GeV$/c^2$ and 5.0 GeV$/c^2$ to be consistent with coming from a $B^0_s$ decay.

In Figure 1a) the $\phi\pi^-$ invariant mass distribution for the ‘right-sign’ $D^-_s \ell^+$ combinations is displayed. A $D^-_s$ signal of 220 $\pm 21$ events is observed from fitting a Gaussian with a mean of (1968 $\pm 1$) MeV$/c^2$ and a width of (10.5 $\pm 1.1$) MeV$/c^2$ plus a straight line to describe the combinatorial background. The mass region indicated by a dashed line has not been included in the fit to avoid contributions from $D^-_s$ decays where particles have been missed such as the $\pi^0$ in the decay $D^-_s \rightarrow \phi\pi^-\pi^0$. Evidence of the Cabibbo-suppressed decay $D^- \rightarrow \phi\pi^-$ is also present in Fig. 1a). The shaded distribution shows $\phi\pi^-$ candidates from ‘wrong-sign’ $D^-_s \ell^-$ combinations. Here, no enhancement is seen in the $D^-_s$ mass region. We allow multiple $\phi\pi^-$ combinations per event but the number of multiple entries is found to be less than 1% of the total number of combinations. This is also the case for the $D^-_s \rightarrow K^{*0}\pi^-$ and $K^{*0}_{S}\pi^-K^-$ decay modes.
2. The $D_S^- \rightarrow K^{*0}K^-$ Mode

For the $D_S^- \rightarrow K^{*0}K^-$ decay mode with $K^{*0} \rightarrow K^+\pi^-$, we assign the $K^+$ and $\pi^-$ masses to two oppositely charged tracks found in the cone of $\Delta R < 0.8$ around the lepton. To reflect the different decay kinematics of the $K^{*0} \rightarrow K^+\pi^-$ decay compared to $\phi \rightarrow K^+K^-$, we require $p_T(K^+) > 1.2$ GeV$/c$ and $p_T(\pi^-) > 0.4$ GeV$/c$. Each $K^{*0}$ candidate is required to have a mass within $\pm 40$ MeV$/c^2$ of the world average $K^{*0}$ mass [1]. To further reduce the high combinatorial background in this decay channel, we tighten the helicity cut to $|\cos \Psi| > 0.5$ and introduce a track based isolation requirement $p_T(D_S^e \ell^+)/p_T($cone$) > 0.6$, where $p_T($cone$)$ is the sum of the transverse momenta of all tracks in a cone $\Delta R < 1.0$ in $\eta$-$\phi$ space. The cone has the lepton direction as its axis and the primary event vertex (see Sec. IV A) as its vertex. All other selection requirements discussed in the $\phi\pi^-$ decay mode remain the same with the exception of the $\ell h_{dE/dx}$ cut described below.

The $K^{*0}K^-$ invariant mass distribution for the ‘right-sign’ $D_S^- \ell^+$ combinations is shown as dots with error bars in Figure 1b). A signal at the $D_S^-$ mass is visible. This signal contains also events from a $D^- \rightarrow K^{*0}\pi^-$ reflection, where the $\pi^-$ is incorrectly assigned the kaon mass. This reflection is further discussed in Section IV C. To reduce the effect of this reflection the $dE/dx$ requirement for that track is tightened to $\ell h_{dE/dx} > 0.1$, while we still demand $\ell h_{dE/dx} > 0.01$ for the tracks forming the $K^{*0}$ candidate.

3. The $D_S^- \rightarrow K_S^0K^-$ Mode

The reconstruction of the $D_S^- \rightarrow K_S^0K^-$ decay mode begins with a search for $K_S^0 \rightarrow \pi^+\pi^-$ candidates by assigning the pion mass to any two oppositely charged tracks with $p_T > 0.4$ GeV$/c$ in a cone of $\Delta R < 1.0$ in $\eta$-$\phi$ space around the lepton. These two tracks are constrained to come from a common vertex and their invariant mass has to be within $5\sigma$ of the nominal $K_S^0$ mass [1], where $\sigma$ is the uncertainty on the $\pi^+\pi^-$ mass measurement. Exploiting the long lifetime of the $K_S^0$ meson, we require the $K_S^0$ vertex to be significantly displaced from the primary event vertex, which is further described in Sec. IV A. We determine the transverse decay length $L_{xy}$ (see Sec. IV B) of the $K_S^0$ and require $L_{xy} > 3\sigma$, where $\sigma$ is the measured uncertainty on $L_{xy}$ for each candidate event.

The $K_S^0$ candidate is combined with any kaon candidate with $p_T > 1.2$ GeV$/c$ within $\Delta R < 0.8$ around the lepton to find the $D_S^-$ candidate. The dots with error bars in Figure 1c) show the $K_S^0K^-$ invariant mass distribution for the ‘right-sign’ $D_S^- \ell^+$ combinations. An enhancement at the $D_S^-$ mass is visible. As in the $K^{*0}K^-$ mode, this signal contains events from a $D^- \rightarrow K_S^0\pi^-$ reflection, where the $\pi^-$ is assigned the kaon mass (see Section IV C). To reduce the effect of this reflection, we again require $\ell h_{dE/dx} > 0.1$, while we demand $\ell h_{dE/dx} > 0.01$ for the tracks forming the $K_S^0$.

4. The $D_S^- \rightarrow \phi\mu^-\nu$ Mode

For the $D_S^- \rightarrow \phi\mu^-\nu$ decay mode, we start with two oppositely charged muons with $p_T > 2$ GeV$/c$ utilizing the dimuon data set obtained with a trigger which requires the dimuon invariant mass to be smaller than 2.8 GeV$/c^2$. This requirement is more than
90% efficient for a double semileptonic decay $B_0^s \rightarrow D_s^- \mu^+ \nu X$ followed by $D_s^- \rightarrow \phi \mu^- \nu$. In addition, two oppositely charged tracks with $p_T > 0.8$ GeV/c are assigned the kaon mass and combined to form a $\phi$ candidate. There is an ambiguity in the assignment of one of the two muons to a found $\phi \rightarrow K^+K^-$ candidate. One of the muons comes from the $D_s^-$ semileptonic decay ($\mu_{D_s^-}$) while the other originates from the $B_0^s$ decay ($\mu_{B_0^s}$). In order to resolve this ambiguity, we require $m(KK\mu_{D_s^-})$ to be smaller than the world average $D_s^-$ mass [1], while $m(KK\mu_{B_0^s})$ has to be greater than $m_{D_s^-}$. To reduce combinatorial background in this decay channel, we use the track based isolation quantity $p_T(D_s^- \ell^+)/p_T(\text{cone})$ and require it to be greater than 0.5. As required in the other decay modes, the invariant mass of the $KK\mu\mu$ system has to lie between 3.0 GeV/$c^2$ and 5.0 GeV/$c^2$. The number of multiple $D_s^-\mu^+$ combinations per event is larger compared to the other three decay modes (about 10%). We therefore allow only one $D_s^-\mu^+$ candidate per event by choosing the $D_s^-\mu^+$ combination with the largest probability from the combined vertex fit (see Sec. IV B).

The $K^+K^−$ invariant mass distribution for the $D_s^-\mu^+$ sample is shown in Figure 1d) with the fit result overlaid. To obtain the number of $\phi$ signal events, we fit a second order polynomial together with a Breit-Wigner line shape convoluted with a Gaussian to account for detector resolution. We find $205 \pm 38 \phi$ signal events and measure the $\phi$ mass to be $(1020.1 \pm 0.5)$ MeV/$c^2$ in agreement with the world average $\phi$ mass [1]. The shaded histogram in Fig. 1d) shows the ‘wrong-sign’ $KK$ mass spectrum, where we consider events with same sign $K^±K^±$ or $\mu^±\mu^±$ combinations as ‘wrong-sign’. For display purposes the ‘wrong-sign’ distribution is scaled by a factor of 0.6 to the same area as the combinatorial background of the ‘right-sign’ $K^+K^−$ distribution. No indication of a $\phi$ signal is evident in the ‘wrong-sign’ distribution.

D. The Monte Carlo Simulation

Some quantities in this analysis like efficiencies or the $K$-factor distribution further described in Sec. IV B are determined using a Monte Carlo (MC) calculation of $b$ quark production and $B$ meson decay followed by a simulation of the detector response to the final state particles. Since we extract only kinematic quantities of the $B$ hadron decay from this Monte Carlo study, we do not simulate the underlying event from the $p\bar{p}$ scattering nor include fragmentation products, but generate only $B$ hadrons and their decay products.

The MC simulation begins with a model of $b$ quark production based on a next-to-leading order QCD calculation [16]. This calculation employs the MRSD0 parton distribution function [17] to model the kinematics of the initial state partons. We generate $b$ quarks in the rapidity interval $|y_b| < 1.0$ with a minimum $p_T$ for the $b$ quark of 8 GeV/c and 5 GeV/c to simulate events corresponding to the single lepton and dimuon data samples, respectively. These $p_T$ requirements are chosen in a way to avoid any biases in the $B$ meson kinematic distributions after the application of the kinematic cuts used in the analysis. The $b$ quarks are fragmented into $B$ mesons according to a model using the Peterson fragmentation function [18] with a Peterson parameter of $\epsilon_b = 0.006$. The bottom and charm hadrons are decayed into the various final states using branching ratios and decay kinematics governed by the world average masses and lifetimes of the involved particles [1].
Events with a lepton above a momentum threshold corresponding to the appropriate hardware trigger are kept based on an efficiency parameterization of the CFT trigger that depends on the lepton $p_T$. The accepted events are passed through a simulation of the CDF detector that is based on parameterizations and simple models of the detector response, which are functions of the particle kinematics. After the simulation of the CDF detector, the same selection criteria applied to the data are imposed on the Monte Carlo events.

**IV. THE $B^0_S$ LIFETIME ANALYSIS**

In this Section, we describe the measurement of the $B^0_S$ lifetime starting with the determination of the primary event vertex followed by the reconstruction of the $B^0_S$ decay length. In order to determine the number of $D^-_s$ signal events used as a constraint in the lifetime fit, a reflection from $D^-$ decays has to be considered for the $D^-_s \to K^{*0}K^-$ and $D^-_s \to K^0_SK^-$ decay modes. In Section IV D background from non-$B^0_S$ decays is discussed, while the lifetime fit is detailed in Sec. IV E. The $B^0_S$ lifetime fit results are then presented together with the determination of the systematic uncertainties.

**A. The Primary Event Vertex**

The $B^0_S$ lifetime reported in this paper is based on measuring the distance between the primary $p\bar{p}$ event vertex and the secondary $B^0_S$ decay vertex in the transverse plane. We first identify the $z$-position of the primary interaction vertex using the tracks reconstructed in the VTX detector. These tracks, when projected back to the beam axis, determine the longitudinal location of the primary interaction with an accuracy of about 0.2 cm along the beam direction. The primary vertices are distributed along the beam direction according to a Gaussian function with a width of $\sim 30$ cm. On average during Run I, the number of reconstructed interaction vertices in a given event follows a Poisson distribution with a mean of about 2.5. For the $B^0_S$ lifetime measurement, we determine the $z$-location of the primary event vertex by choosing the $p\bar{p}$ interaction vertex recorded by the VTX which is closest to the intercept of the lepton from the semileptonic $B^0_S$ decay with the beamline. We also require the $z$-coordinates of all tracks from the $D^-_s$ decay to be within 5 cm of the $z$-coordinate of this primary vertex.

The transverse position of the primary event vertex is determined by using the average beam position through the detector together with the knowledge of the longitudinal primary vertex position from the VTX. The average beam position is calculated offline for each data acquisition run. This calculation yields a transverse profile of the Tevatron beam which is circular with an RMS of $\sim 25$ $\mu$m in both the $x$- and $y$-directions. We find that the average beam trajectory is stable over the period that a given $p\bar{p}$ beam is stored in the Tevatron Collider. A detailed description of the determination of the average beamline can be found in Ref. [15]. For the $B^0_S$ lifetime measurement, we consider only events from data runs with a sufficiently large number of collected events to allow a good determination of the run averaged beamline. In this analysis, we choose not to measure the primary vertex event-by-event because the presence of a second $b$ quark decay in the event coupled with
the low multiplicity in semileptonic $B$ decays can lead to a systematic bias in the lifetime determination.

B. The Decay Length Reconstruction

The tracks forming the $D_S^-$ candidate are refit with a common vertex constraint referred to as the tertiary vertex $V_{D_S^0}$. The secondary vertex where the $B_S^0$ decays to a lepton and a $D_S^-$ (referred to as $V_{B_S^0}$) is obtained by simultaneously intersecting the trajectory of the lepton track with the flight path of the $D_S^-$ candidate. Since we fully reconstruct the $D_S^-$ meson in the $\phi\pi^-, K^{*-}K^-$, and $K^0_s K^-$ decay modes, we know the $D_S^-$ flight path. In the $D_S^- \rightarrow \phi\mu^-\nu$ channel, where we do not fully reconstruct the $D_S^-$ meson, we use the $\phi\mu^-$ flight direction as a good estimate of the $D_S^-$ flight path.

The confidence level of the combined vertex fit is required to be greater than 1%. Furthermore, we require that the reconstructed $D_S^-$ decay vertex $V_{D_S^-}$ be positively displaced from the primary vertex as projected along the direction of the $D_S^-$ momentum.

The transverse decay length $L_{xy}(B_S^0)$ is defined as the displacement $\vec{X}$ in the transverse plane of $V_{B_S^0}$ from the primary event vertex projected onto the $D_S^-\ell^+$ momentum:

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

(4)

$L_{xy}$ is a signed variable which can be negative for the configuration where the particle seems to decay before the point where it is produced. The $B_S^0$ meson decay time is given by

$$c t(B_S^0) = L_{xy} \frac{m(B_S^0)}{p_T(B_S^0)},$$

(5)

where $m(B_S^0)$ is the $B_S^0$ mass [1]. Since we do not fully reconstruct the $B_S^0$ meson in our analysis, we define the ‘pseudo-proper decay length’

$$\lambda = L_{xy} \frac{m(B_S^0)}{p_T(D_S^-\ell^+)},$$

(6)

which has a typical uncertainty of $\sim 60 \mu m$ including the contribution from the finite size of the primary event vertex. In addition, we introduce a correction factor

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

(7)

to correct between the reconstructed $p_T(D_S^-\ell^+)$ and the unknown $p_T(B_S^0)$ in the data. The $B_S^0$ meson decay time is then given as

$$c t(B_S^0) = L_{xy} \frac{m(B_S^0)}{p_T(D_S^-\ell^+)} \times K.$$

(8)

The correction between $p_T(D_S^-\ell^+)$ and $p_T(B_S^0)$ is done statistically by smearing an exponential decay distribution with a Monte Carlo distribution of the correction factor $K$ when
extracting $c\tau(B_s^0)$ from the pseudo-proper decay length in the lifetime fit as described in Sec. IV.E. The $K$-distribution is obtained from $D_s^-\ell^+$ combinations which originate from a Monte Carlo simulation (see Sec. III.D) of semileptonic $B_s^0$ decays into $D_s^{(*)-}\ell^+X$ including $D_s^{(*)-}\tau^+X$ with $\tau^+ \to \ell^+X$. As an example, the $K$-distribution is shown for the $D_s^- \to \phi\pi^-$ and $D_s^- \to \phi\mu^-\nu$ decay modes in Figures 2a) and 2b), respectively. The $K$-distributions have mean values of 0.86 and 0.77 with RMS values of 0.10 and 0.12 for the $D_s^- \to \phi\pi^-$ and $D_s^- \to \phi\mu^-\nu$ modes, respectively. The $K$-distribution is approximately constant as a function of $p_T(D_s^-\ell^+)$ for the range of $p_T(D_s^-\ell^+)$ corresponding to our data.

To ensure a precise $B_s^0$ lifetime determination, we consider only $B_s^0$ candidates for which the pseudo-proper decay length is measured with an uncertainty of less than 0.1 cm. We also require that the $D_s^-$ candidates have a proper decay length measured between $V_{D_s}$ and $V_{D_s}$ of less than 0.1 cm and that the uncertainty on this proper decay length is less than 0.1 cm. This requirement removes background events with very long-lived $D_s^-$ candidates, where the long extrapolation back to the $B_s^0$ decay vertex results in a poor vertex measurement. These requirements have already been applied to the $D_s^-$ mass distributions shown in Fig. 1.

### C. The Determination of the Reflection from $D^-$

The reconstructions of the $D_s^-$ decay modes into $K^{*0}K^-$ and $K_s^0K^-$ suffer from reflections of $D^- \to K^{*0}\pi^-$ and $D^- \to K_s^0\pi^-$, respectively, where the $\pi^-$ is incorrectly assigned the kaon mass. We will discuss this reflection from $D^-$ and the determination of the true number of events from the $D_s^-$ decay with the example of the $D_s^- \to K^{*0}K^-$ mode. The effect of this $K^-\pi$ misassignment can be seen in Figure 3; events from a $B \to D^-\ell\nu X$ Monte Carlo simulation with $D^- \to K^{*0}\pi^-$ yield an invariant mass distribution indicated by the shape of the shaded area in Fig. 3c) if they are reconstructed as $B_s^0 \to D_s^-\ell\nu X$ with $D_s^- \to K^{*0}K^-$, misinterpreting the $\pi^-$ as $K^-$. A significant portion of this $D^-$ reflection lies at the $D_s^-$ mass peak.

Although we have already tightened our $dE/dx$ likelihood ratio to better identify the $K^-$ track as a kaon, CDF’s $dE/dx$ capabilities with a $\pi/K$ separation of about 1 sr for tracks with $p_T$ greater than about 1 GeV/c are not sufficient to remove this $D^-$ reflection. Applying a $D^-$ mass veto by rejecting all $K^{*0}K^-$ combinations which are within a $\pm 3\sigma$ window around the nominal $D^-$ mass when reconstructed as $K^{*0}\pi^-$, distorts the $K^{*0}K^-$ mass distribution. It would be very difficult to estimate the remaining $D_s^-$ signal from that distribution and use it as input to the $B_s^0$ lifetime fit. We therefore choose to measure the $D^-$ reflection directly from our data and account for the $D^-$ component in the $B_s^0$ lifetime fit. We use two methods to determine the $D^-$ reflection in our data.

The first method performs a simultaneous fit to the $K^{*0}K^-$ and $K^{*0}\pi^-$ invariant mass distributions, where the $K^{*0}\pi^-$ mass distribution is created by switching the mass assignment on the $K^-$ track to be a pion. Figure 3a) shows the $K^{*0}K^-$ mass distribution, while the corresponding $K^{*0}\pi^-$ mass distribution is displayed in Fig. 3b). Each distribution is described by a Gaussian for the corresponding $D^-$ and $D_s^-$ signal as shown in Figures 3c) and 3d) plus a linear lineshape to parameterize the combinatorial background. The shape of the corresponding $D^-$ or $D_s^-$ reflection as obtained from a Monte Carlo simulation is also included in the fit as displayed in Fig. 3c) and d) as the shaded areas. The two mass
distributions are fit simultaneously with the number of events in the Gaussian $D_S^-$ ($D^-$) signal constrained to the number of events in the corresponding $D^-$ ($D_S^-$) reflection. In addition, the difference between the $D_S^-$ and $D^-$ mass values is fixed based on their nominal mass values [1]. The fit returns $123 \pm 25$ $D_S^-$ signal events and 80 $\pm 10$ events from the $D^-$ reflection within the $D_S^-$ signal region defined in Sec. IV E. The fit result is shown in Figure 3a)–d). We perform studies using MC pseudo-experiments to verify the validity of this method [19]. We find that the simultaneous fitting method returns the number of true $D_S^-$ events in our MC studies with no bias and the error obtained from the fit to the data agrees with the expected uncertainty of this technique for our sample size.

The second method for determining the amount of $D^-$ reflection in our $D_S^-$ signal events exploits the difference between the $D^-$ lifetime ($\tau(D^-) = (1.057 \pm 0.0015)$ ps [1]) and the $D_S^-$ lifetime ($\tau(D_S^-) = (0.467 \pm 0.0017)$ ps [1]). As described in Sec. IV F, we can determine the $D_S^-$ lifetime in our fit for the $B_S^0$ lifetime. We modify the fitting method used to determine the $D_S^-$ lifetime in the following way: We replace the exponential describing the $D_S^-$ signal by the sum of two exponentials, one with the $D_S^-$ lifetime and one with the $D^-$ lifetime (see Sec. IV E about the fitting method and Sec. V for an example of a two-lifetime fit). We fix the $D_S^-$ and $D^-$ lifetimes to their nominal values [1] and allow the relative fractions of $D_S^-$ and $D^-$ to float in the fit. With this method we obtain $129^{+34}_{-31}$ $D_S^-$ events and $84^{+34}_{-31}$ events attributed to the $D^-$ reflection. We again perform studies using MC pseudo-experiments and verify the validity of this method to work without any bias [19].

We determine the weighted average of $D_S^-$ events from both methods and obtain a $D_S^-$ signal of $125 \pm 20$ events for the $D_S^- \rightarrow K^{\ast 0} K^-$ decay. Both methods are also used to calculate the number of $D_S^-$ events and the contribution from the $D^-$ reflection in the $D_S^- \rightarrow K_S^0 K^-$ decay mode. We obtain $33 \pm 8$ $D_S^-$ signal events for the $K_S^0 K$ mode. These numbers are displayed in Table I together with the numbers of $D_S^-$ signal events for the $D_S^- \rightarrow \phi \pi^-$ and $D_S^- \rightarrow \phi \pi^- \nu \bar{\nu}$ decay modes. As further described in Sec. IV E, these event numbers are used as a constraint in the $B_S^0$ lifetime fit.

**D. The Non-$B_S^0$ Background**

There are two possible sources of non-strange $B$ meson decays which can lead to ‘right-sign’ $D_S^- \ell^+$ combinations. The first process originates from the decays $B^0 \rightarrow D_S^{(*)-} D^{(*)+} X$ and $B^- \rightarrow D_S^{(*)-} D^{(*)0} X$, with the $D^0$ or $D^+$ decaying semileptonically. These decays produce softer and less isolated leptons than the leptons from $B_S^0$ semileptonic decays. Therefore we expect the acceptance for this background source relative to the $B_S^0$ signal to be quite small. We use a Monte Carlo simulation of these events and estimate their contribution $f_{D_S^0}$ in the following way:

$$f_{D_S^0} = \epsilon_{rel} \cdot \frac{f_u + f_d}{f_S} \cdot \frac{BR(B \rightarrow D_S^{(*)-} D^{(*)+} X)}{BR(B_S^0 \rightarrow D_S^{(*)-} \ell \nu X)}.$$  \hspace{1cm} (9)

We use the following branching ratios and fragmentation fractions from the Particle Data Group [1]: $BR(B^0 \rightarrow D_S^{(*)-} \ell \nu X) = (7.6 \pm 2.4)\%$, $BR(B \rightarrow D_S^{(*)-} D^{(*)+} X) = (4.9 \pm 1.1)\%$, $f_u = f_d = (37.8 \pm 2.2)\%$, and $f_S = (11.2 \pm 2.2)\%$. $\epsilon_{rel}$ is the ratio of efficiencies and acceptances for both decays obtained from a Monte Carlo simulation:
\[ \epsilon_{\text{rel}} = \frac{\epsilon(B \rightarrow D^{(*)}_S D^{(*)} X)}{\epsilon(B^0 \rightarrow D^{(*)}_S \ell \nu X)}. \] (10)

The values obtained for \( \epsilon_{\text{rel}} \) are in the order of 0.5% to 1%. The calculated fractions \( f_{D^0 S^0} \) for each \( D_S^- \) decay mode are compiled in Table I. The \( f_{D^0 S^0} \) fraction is larger for the \( D_S^- \rightarrow \phi \mu^- \nu \) decay mode because of the on average softer \( B \) meson momentum in the dimuon data sample compared to the single lepton trigger events.

The second process is a four body decay \( B^0 / B^+ \rightarrow D_S^- K \ell^+ \nu X \), where \( K \) denotes any type of strange meson. Because of the low probability of producing \( s \bar{s} \) pairs and the limited phase space, this process is suppressed and has not been observed experimentally \([20]\). Based on the quoted limit \( \text{BR}(B^0 / B^+ \rightarrow D_S^- \ell^+ \nu X) < 0.9 \% \) (90\% CL) \([1,20]\) and our detection efficiency determined from MC simulation, we expect less than 1.0 \% of our \( D_S^- \ell^+ \) combinations to originate from this source.

We also consider events from \( B^0_S \rightarrow D^{(*)}_S + D^{(*)}_S - X \) decays, with one \( D_S \) decaying semileptonically. This contribution to our \( B^0_S \) lifetime sample is determined from Monte Carlo studies in the same way as described above for the \( B \rightarrow D^{(*)}_S D^{(*)} X \) background. The obtained fractions \( f_{D^0 S^0 D^0} \) from these decays are small and compiled in Table I. Finally, backgrounds with a real \( D_S^- \) meson and a fake lepton from decays such as \( B \rightarrow D^{(*)}_S D^{(*)} X \) with a hadron from the \( D^{(*)} \) decay faking a lepton are negligible due to the low probability of a hadron faking a lepton.

In summary, the contribution of all above physics backgrounds is quite small compared to the combinatorial background. We account for contributions from \( B \rightarrow D^{(*)}_S D^{(*)} X \) and \( B^0_S \rightarrow D^{(*)}_S(D^{(*)}_S - X \) decays in our lifetime fit described next, and treat the contribution of \( B^0 / B^+ \rightarrow D_S^- K \ell^+ \nu X \) decays as a source of systematic uncertainty in the \( B^0_S \) lifetime measurement.

**E. The Description of the Lifetime Fit**

As input to the \( B^0_S \) lifetime fit, we define a signal sample using a \( D_S^- \) mass window from 1.944 GeV/c\(^2\) to 1.994 GeV/c\(^2\) for the \( D_S^- \rightarrow \phi \pi^- \), \( \phi_{00}K^- \), and \( K^{00}K^- \) decay modes and a \( \phi \) signal window from 1.0094 GeV/c\(^2\) to 1.0294 GeV/c\(^2\) for the \( D_S^- \rightarrow \phi \mu^- \nu \) decay channel. The numbers of events in the signal samples can be found for the four decay modes in Table I. To model the pseudo-proper decay length distribution of the combinatorial background events contained in the signal sample, we define a background sample which consists of ‘right-sign’ events from the \( D_S^- \) sidebands (1.884 GeV/c\(^2\) – 1.934 GeV/c\(^2\) and 2.004 GeV/c\(^2\) – 2.054 GeV/c\(^2\)) and ‘wrong-sign’ events from the interval 1.884 GeV/c\(^2\) to 2.054 GeV/c\(^2\). For the \( D_S^- \rightarrow \phi \mu^- \nu \) decay mode the \( \phi \) sidebands are defined from 0.9844 GeV/c\(^2\) to 1.0044 GeV/c\(^2\) and from 1.0344 GeV/c\(^2\) to 1.0544 GeV/c\(^2\), while the ‘wrong-sign’ combinations are taken from the region 0.9844 GeV/c\(^2\) to 1.0544 GeV/c\(^2\). We assume the combinatorial background to originate from random track combinations and therefore use the sidebands to model the background in the signal sample. This assumption is supported by the mass distribution of the ‘wrong-sign’ combinations where no enhancement is visible at the \( D_S^- \) mass. By adding the ‘wrong-sign’ combinations to the ‘right-sign’ sideband events, we better constrain the shape of the combinatorial background events.
in the $D_s^-$ signal samples for decay channels with low combinatorial background like the $D_s^- \to \phi \pi^-$ mode.

The pseudo-proper decay length distribution obtained from the signal sample is fit using an unbinned maximum log-likelihood method. Both the $B_s^0$ lifetime, denoted as $c\tau$ below, and the background shape are determined in a simultaneous fit using the signal and background samples. Thus the likelihood function $\mathcal{L}$ is a combination of two parts:

$$
\mathcal{L} = \prod_i N_{S}^{i} \left[ f_{\text{sig}}^{i} F_{\text{sig}}^{i} + (1 - f_{\text{sig}}^{i}) F_{\text{bg}}^{i} \right] \cdot \prod_j N_{B}^{j} F_{\text{bg}}^{j},
$$

(11)

where $N_S$ and $N_B$ are the number of events in the signal and background samples. $f_{\text{sig}}$ is the ratio of $D_s^-$ signal events obtained from the $D_s^-$ mass distributions (see Table I) to the total number of events in the signal sample. To constrain $f_{\text{sig}}$ we factor in an additional $\chi^2$ term to the likelihood function $\mathcal{L}$ above to constrain the number of $D_s^-$ signal events obtained from the invariant mass distributions within their uncertainty.

The signal probability function $F_{\text{sig}}$ consists of a normalized decay exponential function convoluted with a Gaussian resolution function $G$ and is smeared with a normalized $K$-distribution $\mathcal{H}(K)$:

$$
F_{\text{sig}}^{i}(x) = \int dK \mathcal{H}(K) \left[ \frac{K}{c\tau} \exp\left\{-\frac{K x}{c\tau}\right\} \otimes G(\lambda^i | x, \sigma^i) \right].
$$

(12)

Here, $\lambda^i$ is the measured pseudo-proper decay length with uncertainty $\sigma^i$ and $x$ is the true pseudo-proper decay length. Because of systematic uncertainties in the overall scale of the decay length uncertainties, which we estimate on an event-by-event basis, we introduce a scale factor, $s$, which is a free parameter in the $B_s^0$ lifetime fit. We subsequently vary $s$ in the fits to determine the sensitivity of our measurement to this uncertainty. The integration over the momentum ratio $K$ is approximated by a finite sum

$$
\int dK \mathcal{H}(K) \to \sum_i \Delta K \mathcal{H}(K_i),
$$

(13)

where the sum is taken over bin $i$ of a histogrammed distribution $\mathcal{H}(K_i)$ with bin width $\Delta K$ as shown e.g. in Figure 2.

The background probability function $F_{\text{bg}}$ is parameterized by a Gaussian centered at zero, a negative exponential tail, and a positive decay exponential to characterize the contribution of heavy flavor decays in the background sample:

$$
F_{\text{bg}}^{i}(x) = (1 - f_+ - f_-) G(\lambda^i | x, \sigma^i) + \\
\quad + \frac{f_+}{\lambda_+} \exp\left\{-\frac{x}{\lambda_+}\right\} \otimes G(\lambda^i | x, \sigma^i) + \\
\quad + \frac{f_-}{\lambda_-} \exp\left\{-\frac{x}{\lambda_-}\right\} \otimes G(\lambda^i | x, \sigma^i).
$$

(14)

Here, $f_{\pm}$ are the fractions of positive and negative lifetime backgrounds and $\lambda_{\pm}$ are the effective lifetimes of those backgrounds. We verify the parameters $f_{\pm}$ and $\lambda_{\pm}$ agree with the ‘right-sign’ sideband events and the ‘wrong-sign’ combinations separately, allowing us to combine both samples resulting in the background samples described above.
The events originating from the $D^-$ reflection in the $D_s^- \to K^{*0} K^-$ and $D_s^- \to K_S^0 K^-$ decays (see Sec. IV C) are also accounted for in the likelihood function by a term

$$ \int dK \mathcal{H}(K) \left[ f_{D^-} \frac{K}{c\tau(B)} \exp\{-\frac{Kx}{c\tau(B)}\} \otimes G(\lambda^t|x, s^t) \right], $$

(15)

where $f_{D^-}$ refers to the fraction of the $D^-$ reflection in the $D_s^-$ sample and $c\tau(B)$ is taken to be the world average $B^0$ lifetime [1].

**F. The Fit Results**

We first determine the $B_s^0$ lifetime for each of the four $D_s^-$ decay channels individually. The parameters allowed to float in the fit are the $B_s^0$ lifetime, $f_{\text{sig}}$, $\lambda_\pm$, $f_\pm$, and the overall scale factor $s$. The fitted values for $c\tau(B)$ and their statistical uncertainties are shown in Table I, and are in good statistical agreement. The pseudo-proper decay length distribution of the signal sample with the result of the fit superimposed is shown in Figure 4a) for the $D_s^- \to \phi\pi^-$ decay mode. The dashed line represents the $B_s^0$ signal contribution, while the shaded curve shows the sum of the background probability function over the events in the signal sample. The same distribution of the background sample is displayed in Figure 4b) with the result of the fit superimposed. Figures 5, 6, and 7 show the corresponding distributions for the $D_s^- \to K^{*0} K^-$, $D_s^- \to K_S^0 K^-$, and $D_s^- \to \phi\mu^-\nu$ decay modes, respectively. The combined $B_s^0$ lifetime from all four $D_s^-$ decay modes is determined from a simultaneous fit to be $c\tau(B_s^0) = (408 \pm 28 \pm 27) \mu\text{m}$ or $\tau(B_s^0) = (1.36 \pm 0.09) \text{ps}$, where the errors shown are statistical only.

As a consistency check, we use the $D_s^- \to \phi\pi^-$ decay mode to also fit the $D_s^-$ lifetime from the proper decay length measured from the secondary vertex $V_{B_s}$ to the tertiary vertex $V_{D_s}$. Since the $D_s^-$ decay is fully reconstructed, its relativistic boost is known and a convolution with a $p_T$-correction factor distribution in the fit does not apply. The result is $c\tau(D_s^-) = (136 \pm 17) \mu\text{m}$ (statistical error only), which is consistent with the world average $D_s^-$ lifetime [1]. Figures 8a) and 8b) show the $D_s^-$ proper decay length distributions for the signal and background samples, respectively, with the results of the fit superimposed.

**G. The Systematic Uncertainties**

Table II lists all sources of systematic uncertainty considered in this analysis. The major contribution originates from the treatment of the background. In particular, the following evaluations of systematic errors are performed yielding the uncertainties reported in Tab. II:

- **Background treatment**: The combinatorial background in the signal sample is parameterized by the positive and negative lifetimes $\lambda_+$ and $\lambda_-$ as well as their respective fractions $f_+$ and $f_-$ as described in Eq. (14). To evaluate the dependence of the $B_s^0$ lifetime on the background parameterization, we vary $\lambda_\pm$ and $f_\pm$ within $\pm 1\sigma$ of their values returned from the $B_s^0$ lifetime fit. Since there is a correlation between the background parameters $\lambda_\pm$ and $f_\pm$ and the $B_s^0$ lifetime in the simultaneous fit to the signal and background samples, part of this systematic uncertainty is already accounted
for in the statistical error on $\tau(B^0_S)$. However, this correlation is small. We therefore adopt this method as a conservative way to evaluate the systematic uncertainty from background treatment.

- Non-$B^0_S$ backgrounds: These backgrounds (see Sec. IV D) enter our fit as fixed fractions. We vary these fractions within $\pm50\%$ of their reported values (see Tab. I) and repeat the fit. We also evaluate the background from $B^0/B^+ \to D^-_s K^+ \ell^+ \nu X$ decays and consider its fraction to be $2.0\%$, twice the quoted limit [1,20].

- Decay length requirement: There are two requirements that can bias the $B^0_S$ lifetime result. These are the cut on $|c\tau(D^-_s)| < 0.1$ cm and the requirement that the reconstructed $D^-_s$ decay vertex $V_{D^-_s}$ be positively displaced from the primary event vertex. To study the effect of these cuts, we use high statistics Monte Carlo samples. We first fit the lifetime with all the selection requirements, and then remove each cut individually noting the shift in the $B^0_S$ lifetime.

- Momentum estimate: The $B^0_S$ lifetime result is sensitive to the distribution of the correction factor $H(K)$, which can be affected by the lepton $p_T$ cut and the decay kinematics. For the standard fit we require $p_T(\mu) > 2.0$ GeV/c for the $D^-_s \to \phi \mu^- \nu$ decay mode and $p_T(\ell) > 6.0$ GeV/c for the other $D^-_s$ decay channels. To test the effect of the lepton $p_T$ dependence, we generate new $K$-distributions for lower and higher lepton $p_T$ cut values. We also compare the effect on the kinematics of semileptonic $B^0_S$ decays using a pure V-A decay versus semileptonic decays using the ISGW form factor [21]. In addition, an alternative $p_T$ spectral shape of $b$ quark production is considered based on a comparison of the lepton $p_T$ shape in the data and in Monte Carlo events. Finally, the $K$-distribution is somewhat dependent on the electron identification. We study a possible incompleteness in the treatment of the electron selection with our Monte Carlo simulation and assign a systematic error of $\pm3\mu m$.

- Decay length resolution: Our uncertainty on the estimate of the decay length resolution is expressed in the scale factor $s$, which is fitted to $1.29 \pm 0.03$. We fix the scale factor at 1.0, and again at 1.38, the latter corresponding to a $+3\sigma$ upward shift from the fitted value, and repeat the $B^0_S$ lifetime fit.

- $D^-$ reflection: The reflection from $D^-$ in the $D^-_s \to K^{*0}K^-$ and $D^-_s \to K^0_S K^-$ decay modes changes the number of $D^-_s$ signal events in these two channels. We study the influence of the $D^-$ reflection by varying the number of $D^-_s$ signal events within their error as determined in Sec. IV C.

- Detector alignment: We also account for a possible residual misalignment of the SVX and assign an error of $\pm2\mu m$ as further detailed in Ref. [15].

The systematic uncertainties noted above have been combined in quadrature. Quoting the statistical and systematic uncertainties separately, we measure the $B^0_S$ lifetime using semileptonic $B^0_S$ decays to be

$$\tau(B^0_S) = (1.36 \pm 0.09 \pm 0.05) \, \text{ps}, \quad (16)$$
where the first error is statistical and the second systematic. This result is currently the world’s best measurement of the $B_s^0$ lifetime from a single experiment. In comparison, the world average $B_s^0$ lifetime is $(1.57 \pm 0.08)$ ps [1]. This measurement supersedes CDF’s previously published $B_s^0$ lifetime result of $\tau(B_s^0) = (1.42^{+0.27}_{-0.23} \pm 0.11)$ ps using a data sample corresponding to an integrated luminosity of 20 pb$^{-1}$ [5].

Using the CDF average $B^0$ lifetime $\tau(B^0) = (1.513 \pm 0.053)$ ps [22], we determine the $B_s^0/B^0$ lifetime ratio to be $0.899 \pm 0.072$ taking correlated systematic uncertainties into account. However, ignoring the correlated systematic uncertainties increases the error on the $B_s^0/B^0$ lifetime ratio only to $\pm 0.077$, since $\tau(B_s^0)/\tau(B^0)$ is dominated by the statistical error on the $B_s^0$ lifetime measurement.

V. THE DETERMINATION OF $\Delta \Gamma/\Gamma$

We examine the $B_s^0$ meson pseudo-proper decay length distribution from $D_s^- \ell^+$ correlations for a lifetime difference $\Delta \Gamma/\Gamma$ between the two mass eigenstates of the $B_s^0$ meson, $B_{sH}^0$ and $B_{sL}^0$. In the case of a lifetime difference in the $B_s^0$ system, the decay length distribution for events from the semileptonic decay $B_s^0 \to D_s^- \ell^+ \nu X$ is expected to be governed by the sum of two exponentials. We expand the likelihood fit to describe the $B_s^0$ pseudo-proper decay length distribution to a function of the form

$$F(t) = e^{-\Gamma_{H}t} + e^{-\Gamma_{L}t} \quad \text{with} \quad \Gamma_{L,H} = \Gamma \pm \frac{\Delta \Gamma}{2} = \Gamma \cdot (1 \pm \frac{1}{2} \frac{\Delta \Gamma}{\Gamma}),$$

(17)

rather than fitting for just one exponential lifetime $e^{-\Gamma t}$. The parameter $\Delta \Gamma/\Gamma$ is the parameter we fit for. Since $\Delta \Gamma/\Gamma$ is symmetric about zero, it is required to be positive. In the case of a lifetime difference $\Delta \Gamma \neq 0$, the total decay width $\Gamma = 1/2 \cdot (\Gamma_H + \Gamma_L)$ and the mean $B_s^0$ lifetime $\tau_m(B_s^0)$ obtained from a fit assuming a single $B_s^0$ lifetime, are no longer reciprocal to each other but follow the relation

$$\tau_m(B_s^0) = \frac{1}{\Gamma} \cdot \frac{1 + (\Delta \Gamma)^2}{1 - (\Delta \Gamma)^2}.$$

(18)

We incorporate the relation in Eq. (18) into our likelihood fitting function. We follow the suggestion given in Ref. [9] and fix the mean $B_s^0$ lifetime to the world average $B^0$ lifetime since both lifetimes are expected to agree within 1% [2,3]. This theoretical assumption can be verified by the current world average $\tau(B^0) = 1.55 \pm 0.05$ ps and $\tau(B_s^0) = 1.57 \pm 0.08$ ps. The CDF average $B_s^0/B^0$ lifetime ratio derived above also supports this assumption.

The fit returns $\Delta \Gamma/\Gamma = 0.34 \pm 0.31$, where the given error is statistical only. This indicates that with the current statistics of our $D_s^- \ell^+$ sample we are not sensitive to a $B_s^0$ lifetime difference. Based on this fit result, we integrate the normalized likelihood as a function of $\Delta \Gamma/\Gamma$ and find the 95% confidence level (CL) limit at

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

(19)

This is the first experimental result for the lifetime difference in the $B_s^0$ system.
Using a value of $\Delta \Gamma / \Delta m = (5.6 \pm 2.6) \cdot 10^{-3}$ from Ref. [9] and setting $\tau_m(B_0^0)$ to the world average $B^0$ lifetime [1], an upper limit on the $B_0^0$ mixing frequency of $\Delta m_5 < 96 \text{ ps}^{-1}$ (95% CL) can be determined within the Standard Model. Including the dependence on $\Delta \Gamma / \Delta m$ and $\tau_m(B_0^0)$ in our limit, we obtain

$$\Delta m_5 < 96 \text{ ps}^{-1} \times \frac{5.6 \cdot 10^{-3}}{\Delta \Gamma / \Delta m} \times \left( \frac{1.55 \text{ ps}}{\tau_m(B_0^0)} \right) \text{ (95% CL).} \quad (20)$$

VI. CONCLUSION

We have presented a measurement of the $B_0^0$ meson lifetime using semileptonic $B_0^0$ decays, where the $D_S^-$ meson is reconstructed through the four decay modes $D_S^- \rightarrow \phi \pi^-$, $D_S^- \rightarrow K^{*0}K^-$, $D_S^- \rightarrow K_S^0K^-$, and $D_S^- \rightarrow \phi \mu^- \nu$. We obtain

$$\tau(B_0^0) = (1.36 \pm 0.09 ^{+0.06}_{-0.05}) \text{ ps}, \quad (21)$$

where the first error is statistical and the second systematic. This is currently the world’s best measurement of the $B_0^0$ lifetime from a single experiment. This result agrees with an earlier CDF measurement [5], which is superseded by the present measurement. We determine the $B_0^0/B^0$ lifetime ratio to be $0.899 \pm 0.072$ using the CDF average $B^0$ lifetime [22].

In addition, we have examined the $B_0^0$ meson pseudo-proper decay length distribution for a lifetime difference $\Delta \Gamma / \Gamma$ between the two mass eigenstates of the $B_0^0$ meson, $B_0^H$ and $B_0^L$. Using all four $D_S^-$ decay modes, an upper limit of $\Delta \Gamma / \Gamma < 0.83$ is set at 95% CL, corresponding to the Standard Model limit

$$\Delta m_5 < 96 \text{ ps}^{-1} \times \frac{5.6 \cdot 10^{-3}}{\Delta \Gamma / \Delta m} \times \left( \frac{1.55 \text{ ps}}{\tau_m(B_0^0)} \right) \text{ (95% CL).} \quad (22)$$

With considerably increased statistics in the next run of the Tevatron Collider, our sensitivity on the lifetime difference $\Delta \Gamma / \Gamma$ will be significantly improved [23].

Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. It is a pleasure to thank G. Buchalla, I. Dunietz and H.-G. Moser for valuable discussions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Science and Culture of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; and the A. P. Sloan Foundation.
REFERENCES


### TABLE I. Summary of results for the four $D_S^-$ decays:
- the number $N(D_S^-)$ of $D_S^-$ signal events as input to the lifetime fit;
- the expected fraction $f_{DSD}$ of $B\to D_S^*(s) D(s)$ decays;
- the expected fraction $f_{DSD}$ of $B_0^s\to D_S^*(s) D_S^*(s)$ decays;
- the number $N_{ev}$ of events in the signal samples;
- the fitted $B_0^s$ lifetimes $c\tau(B_0^s)$, where the errors shown are statistical only.

<table>
<thead>
<tr>
<th>$D_S^-$ Decay Mode</th>
<th>$N(D_S^-)$</th>
<th>$f_{DSD}$</th>
<th>$f_{DSD}$</th>
<th>$N_{ev}$</th>
<th>$c\tau(B_0^s)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi\pi^-$</td>
<td>220 ± 21</td>
<td>2.6 %</td>
<td>0.8 %</td>
<td>350</td>
<td>$418^{+43}_{-38}$ µm</td>
</tr>
<tr>
<td>$K^{*0}K^-$</td>
<td>125 ± 20</td>
<td>2.5 %</td>
<td>0.8 %</td>
<td>820</td>
<td>$411^{+73}_{-76}$ µm</td>
</tr>
<tr>
<td>$K_0^0K^-$</td>
<td>33 ± 8</td>
<td>1.8 %</td>
<td>0.6 %</td>
<td>146</td>
<td>$397^{+61}_{-52}$ µm</td>
</tr>
<tr>
<td>$\phi\mu^-\nu$</td>
<td>205 ± 38</td>
<td>5.7 %</td>
<td>1.7 %</td>
<td>635</td>
<td>$399^{+50}_{-45}$ µm</td>
</tr>
</tbody>
</table>

### TABLE II. Compilation of systematic uncertainties in the measurement of the $B_0^s$ lifetime combining all four $D_S^-$ decay modes.

<table>
<thead>
<tr>
<th>Error Source</th>
<th>$\Delta c\tau(B_0^s)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background treatment</td>
<td>±11 µm</td>
</tr>
<tr>
<td>Non-$B_0^s$ backgrounds</td>
<td>±6 µm</td>
</tr>
<tr>
<td>Decay length requirement</td>
<td>±1 µm</td>
</tr>
<tr>
<td>Momentum estimate</td>
<td></td>
</tr>
<tr>
<td>Lepton $p_T$ dependence</td>
<td>±6 µm</td>
</tr>
<tr>
<td>$B$ decay model</td>
<td>±3 µm</td>
</tr>
<tr>
<td>$b$ quark $p_T$ spectrum</td>
<td>±1 µm</td>
</tr>
<tr>
<td>Electron selection</td>
<td>±3 µm</td>
</tr>
<tr>
<td>Decay length resolution</td>
<td>±7 µm</td>
</tr>
<tr>
<td>$D^-$ reflection</td>
<td>±2 µm</td>
</tr>
<tr>
<td>Detector alignment</td>
<td>±15 µm</td>
</tr>
</tbody>
</table>

TABLE II. Compilation of systematic uncertainties in the measurement of the $B_0^s$ lifetime combining all four $D_S^-$ decay modes.
FIG. 1. Invariant mass distributions of (a) $D_s^- \to \phi \pi^-$, (b) $D_s^- \to K^*\bar{0}K^-$, (c) $D_s^- \to K^0S K^-$, and (d) $\phi \to K^+K^-$ from $D_s^- \to \phi \mu^-\nu$. The dots with error bars are for 'right-sign' $D_s^- \ell^+$ combinations while the shaded histograms show the corresponding 'wrong-sign' distributions. In (a) evidence of the decay $D^- \to \phi \pi^-$ is present. The results of the fits described in the text are also superimposed. The mass regions indicated by a dashed line have not been included in the fits.
FIG. 2. Normalized $K$-factor distributions $p_T(D^-_s \ell^+)/p_T(B^0_s)$, for $B^0_s \to D^-_s \ell^+\nu X$ Monte Carlo decays with (a) $D^-_s \to \phi\pi^-$ and (b) $D^-_s \to \phi\mu^-.\nu$. 
FIG. 3. (a) Mass distributions for candidates in the $D_s^- \rightarrow K^{*0}K^-$ decay mode. (b) Mass distribution if these candidates are assumed to be $D^− \rightarrow K^{*0}π^-$. (c) Distribution of the $D_s^- \rightarrow K^{*0}K^-$ signal and the reflection from $D^- \rightarrow K^{*0}π^-$ (shaded area) as obtained from Monte Carlo simulations. Normalizations are determined from the simultaneous fit described in the text. (d) Mass distribution of the corresponding $D^- \rightarrow K^{*0}π^-$ signal and the reflection from $D_s^- \rightarrow K^{*0}K^-$ (shaded area).
FIG. 4. (a) $B^0_s$ pseudo-proper decay length distribution for the $D_s^- \rightarrow \phi \pi^-$ signal sample with the result of the fit superimposed. The dashed line is the $B^0_s$ signal contribution, while the shaded curve represents the contribution from the combinatorial background. (b) Pseudo-proper decay length distribution for the background sample with the fit result superimposed.
FIG. 5. (a) $B^0_s$ pseudo-proper decay length distribution for the $D^-$ signal sample with the result of the fit superimposed. The dashed line is the $B^0_s$ signal contribution, while the shaded curve represents the contribution from the combinatorial background. (b) Pseudo-proper decay length distribution for the background sample with the fit result superimposed.
FIG. 6. (a) $B_0^0$ pseudo-proper decay length distribution for the $D_s^- \to K_0^0 K^-$ signal sample with the result of the fit superimposed. The dashed line is the $B_0^0$ signal contribution, while the shaded curve represents the contribution from the combinatorial background. (b) Pseudo-proper decay length distribution for the background sample with the fit result superimposed.
FIG. 7. (a) $B_s^0$ pseudo-proper decay length distribution for the $D_s^- \rightarrow \phi \mu^- \nu$ signal sample with the result of the fit superimposed. The dashed line is the $B_s^0$ signal contribution, while the shaded curve represents the contribution from the combinatorial background. (b) Pseudo-proper decay length distribution for the background sample with the fit result superimposed.
FIG. 8. (a) $D_s^-$ proper decay length distribution $c\tau(D_s^-)$ for the $D_s^- \rightarrow \phi\pi^-$ signal sample with the result of the fit superimposed. The dashed line is the $D_s^-$ signal contribution, while the shaded curve represents the contribution from the combinatorial background. (b) Proper decay length distribution $c\tau(D_s^-)$ for the background sample with the fit result superimposed.