PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/84049

Please be advised that this information was generated on 2019-04-01 and may be subject to change.
Search for CP violation in $B_s^0 \to \mu^+ \mu^-$ decays in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

We have performed a search for CP violation in a sample of $B_0^s \rightarrow \mu^+ \mu^- X$ decays corresponding to 5 fb$^{-1}$ of proton-antiproton collisions collected by the D0 detector in Run II at the Fermilab Tevatron Collider. New physics in $B_0^s$ mixing could contribute a significant CP violating weak phase, which would be observed as a difference in the decay-time distribution for $B^0_s \rightarrow B^0_s$ oscillated states versus that for $B^0_s \rightarrow B^0_s$. A fit to the decay-time distributions of the $B^0_s = B^0_s$ candidates yields the flavor-specific asymmetry $a_{fs} = [(-1.7 \pm 9.1)$ (stat)$ ^{+1.2}_{-1.3}$ (syst) $] \times 10^{-3}$, which excludes CP violation due to new physics within the experimental sensitivity.

DOI: 10.1103/PhysRevD.82.012003 PACS numbers: 11.30.Er, 13.20.He, 14.40.Nd
I. INTRODUCTION

The search for large CP violating (CPV) effects in the $B_s^0 \rightarrow \bar{B}_s^0$ system is of special interest since their observation would be a direct indication of new physics. A nonzero CPV weak phase $\phi_s$ arises from the phase difference between the absorptive and dispersive parts of the $B_s^0 - \bar{B}_s^0$ mixing amplitude: $\phi_s = \arg(-M_{12}/\Gamma_{12})$ [1], where $M_{12}$ and $\Gamma_{12}$ are the off-diagonal elements of the $B_s^0 - \bar{B}_s^0$ mass and decay matrices, respectively. A global fit to various measurements interpreted in the context of the standard model (SM) yields the prediction $\phi_s^{SM} = (4.2 \pm 1.4) \times 10^{-3}$ [1]. However, new physics, such as the existence of a fourth generation [2], could contribute an additive phase $\phi_s^{NP}$ such that $\phi_s = \phi_s^{SM} + \phi_s^{NP}$. Recent measurements of $\phi_s^{NP}$ in $B_s^0 \rightarrow J/\psi \phi$ decays by the CDF [3] and D0 [4] collaborations differ from zero by approximately 2 standard deviations, which motivates further CP violation studies in $B_s^0$ decays.

The CPV weak phase $\phi_s$ can be obtained from the flavor-specific asymmetry

$$a_{fs} = \frac{\Gamma_{B_s^0 \rightarrow f} - \Gamma_{B_s^0 \rightarrow \bar{f}}}{\Gamma_{B_s^0 \rightarrow f} + \Gamma_{B_s^0 \rightarrow \bar{f}}},$$

according to $a_{fs}^s = \frac{\Delta m_s}{\Delta m_s} \tan \phi_s$, where $\Delta m_s$ and $\Delta m_s$ are the width and mass differences, respectively, between the heavy and light eigenstates of the mixed $B_s^0$ system. World average values of these quantities [5] yield $a_{fs}^s = (-8.4^{+5.7}_{-6.2}) \times 10^{-3}$ [6]. Improved precision is needed to establish evidence of physics beyond the SM, which predicts $a_{fs}^s = (0.0206 \pm 0.0057) \times 10^{-3}$ [1].

We present a measurement of $a_{fs}^s$ using $B_s^0 \rightarrow \mu^+ \mu^- D_s^0$ decays (charge conjugate states are assumed throughout) reconstructed in proton-antiproton collisions collected by the D0 detector between April 2002 and August 2008, corresponding to about 5 fb$^{-1}$ of integrated luminosity [7]. This CPV study is complementary to those using inclusive dimuon events and $B_s^0 \rightarrow J/\psi \phi$ decays. The time-integrated inclusive dimuon analysis [8] does not distinguish between the various $B$ hadrons and therefore depends heavily on $B_s^0$ asymmetry results from the $B$ factories and the determination of the $B^+/B^0_s/B^0_s/b$-baryon production fractions. In contrast, the present measurement allows a straightforward determination of the sample composition, due to the partial reconstruction of the $B_s^0$ meson. The $B_s^0 \rightarrow J/\psi \phi$ CPV measurements [3,4] involve an analysis of the decay product transversity angles to separate the $CP$-even and $CP$-odd components. The present measurement does not require any angular analysis. Furthermore, it uses all the $B_s^0$ production and decay information available in an event, the former via initial-state flavor tagging, when possible, and the latter via an unbinned fit to the decay-time distribution.

II. EVENT RECONSTRUCTION AND SELECTION

The D0 detector is described in detail elsewhere [9]. Charged particles are reconstructed using the central tracking system, which consists of a silicon microstrip tracker and a central fiber tracker, both located within a 2 T superconducting solenoidal magnet. An additional single-layer silicon microstrip detector called Layer 0 [10] installed immediately outside the beam pipe provides improved impact parameter resolution and vertexing efficiency. Electrons are identified by the preshower detector and liquid-argon/uranium calorimeter. Muons are identified by the muon system, which consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8 T iron toroids, followed by two similar layers after the toroids [11]. The solenoid and toroid polarities are reversed regularly, the latter allowing a determination of the muon charge asymmetries induced by the detector. The 5 fb$^{-1}$ data sample used in this analysis is divided into two subsamples; the first 1.3 fb$^{-1}$ is referred to as Run Ia and the remaining 3.7 fb$^{-1}$ collected after the installation of the Layer 0 detector is referred to as Run Ib.

Most of the sample was collected with single muon triggers. The reconstruction of the $B_s^0 \rightarrow \mu^+ \mu^- D_s^0 X$ candidates is as follows. The tracks were required to have signals in both the central fiber tracker and silicon microstrip tracker. Muons were required to have measurements in at least two layers of the muon system. The muon segment was required to be matched to a track in the central tracking system and to have momentum $p(\mu^+) > 3.0 \text{ GeV}/c$ and $p_T(\mu^+) > 2.0 \text{ GeV}/c$, where $p_T$ is the momentum component transverse to the proton beam direction.

All the tracks in each event were clustered into jets using the algorithm described in Ref. [12]. The $D_s^0$ candidates were then formed from tracks found in the same jet as the muon candidate. Two $\mu^+ D_s^0$ final-state samples were reconstructed: $\mu^+ \phi \pi^-$ where $\phi \rightarrow K^+ K^-$ and $\mu^+ K^{0 \pi^0} K^-$ where $K^{0 \pi^0} \rightarrow K^+ \pi^-$. The $\mu^+ \phi \pi^-$ reconstruction follows the technique described in Ref. [13]. The $\phi$ candidate was formed from two oppositely charged particles assigned the kaon mass (the D0 detector is unable to distinguish between kaons, pions, and protons). The kaon candidates were required to have $p_T(K^0) > 0.7 \text{ GeV}/c$. The $K^+ K^-$ invariant mass distribution for the $\phi$ candidates in the final selected sample is shown in Fig. 1. The $\phi$ candidate was required to have an invariant mass in the range $1.004 < M(K^+ K^-) < 1.034 \text{ GeV}/c^2$, consistent with that of a $\phi$ meson. A pion candidate with charge opposite to that of the muon and $p_T(\pi^-) > 0.5 \text{ GeV}/c$ was then added to form the $D_s^0$ meson candidate. In the $\mu^+ K^{0 \pi^0} K^-$ decay mode, the $D_s^0$ candidate was formed from three charged particles, one with the same charge as the muon and two with a charge opposite to that of the muon. The particle with the same charge as the muon was assigned the kaon mass and required to have...
The primary proton-antiproton interaction vertex was assigned the pion mass and required to have \( p_T^{(K^-)} > 0.9 \text{ GeV}/c \). A more stringent requirement of \( p_T^{(K^-)} > 1.8 \text{ GeV}/c \) was imposed on the second kaon candidate to reduce combinatorial background. The third particle was assigned the pion mass and required to have \( p_T^{(\pi^-)} > 0.5 \text{ GeV}/c \). The \( K^+ \pi^- \) invariant mass distribution for the \( K^{*0} \) candidates in the final selected sample is shown in Fig. 2. \( K^{*0} \) candidates were required to have an invariant mass in the range \( 0.82 < M(K^+ \pi^-) < 0.95 \text{ GeV}/c^2 \), consistent with the \( K^{*0} \) mass. Details about the \( \mu^+ K^{*0} \) analysis are available in Ref. [14]. The sharp edge below 0.94 GeV/c^2 in Fig. 2 is an artifact of the selection criteria that are dependent on the mass resolution. Since the events in this mass range are primarily background, the measured asymmetry is not affected.

The primary proton-antiproton interaction vertex was determined for each event. The average position of the collision point in the plane transverse to the beam was measured for each run and was included as a constraint. The precision of the primary vertex reconstruction for each event was on average about 20 \( \mu \text{m} \) in the plane perpendicular to the beam direction and about 40 \( \mu \text{m} \) along the beam direction.

In both decay modes, a common \( D_s^- \) decay vertex was formed from the three \( D_s^- \) daughter tracks using the algorithm described in Ref. [15]. To reduce combinatorial background, the \( D_s^- \) vertex was required to have a displacement from the primary vertex in the transverse plane of at least 4 standard deviations. The cosine of the angle between the \( D_s^- \) momentum and the direction from the primary vertex to the \( D_s^- \) decay vertex was required to be greater than 0.9. The trajectories of the muon and \( D_s^- \) candidates were required to be consistent with originating from a common vertex (used as the \( B_s^0 \) decay vertex) and the \( \mu^+ D_s^- \) system was required to have an invariant mass between 2.6 and 5.4 GeV/c^2, consistent with coming from a \( B_s^0 \) semileptonic decay. We define an angle between the combined \( \mu^+ D_s^- \) momentum (an approximation of the \( B_s^0 \) momentum) and the direction from the primary vertex to the \( B_s^0 \) decay vertex. The cosine of this angle was required to be greater than 0.95 for \( B_s^0 \) candidates displaced from the primary vertex in the transverse plane by at least 4 standard deviations. These angular criteria ensure that the \( D_s^- \) momentum is sufficiently aligned with that of its \( B_s^0 \) parent. The displacement and angular criteria give rise to a decay-time dependent reconstruction efficiency, which is discussed later.

The \( B_s^0 \) selection was further improved using a likelihood ratio method [13,16] that combines a number of discriminating variables: the helicity angle between the \( D_s^- \) and \( K^\pm \) momenta in the \( \phi \) or \( K^{*0} \) center-of-mass frame; the isolation of the \( \mu^+ D_s^- \) system, defined as \( I = p(\mu^+ D_s^-)/[p(\mu^+ D_s^-) + \Sigma p_i] \), where the sum is over all tracks in the cone \( \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2} < 0.5 \) around the \( \mu^+ D_s^- \) direction (\( \phi \) is the azimuthal angle of the track, \( \eta = -\ln[\tan(\theta/2)] \) is the pseudorapidity and \( \theta \) is the polar angle between the track momentum and the beam axis); the \( \chi^2 \) of the \( D_s^- \) vertex; the invariant masses \( M(\mu^+ D_s^-) \), \( M(K^+ K^-) \) (\( \mu^+ \phi \pi^- \) sample) or \( M(K^+ \pi^-) \) (\( \mu^+ K^{*0} \pi^- \) sample); and \( p_T(K^+ K^-) \) (\( \mu^+ \phi \pi^- \) sample) or \( p_T(K^+ \pi^-) \) (\( \mu^+ K^{*0} \pi^- \) sample). The final requirement on the like-
likelihood ratio variable, $y_{\text{sel}}$, was chosen to maximize the predicted ratio $S/\sqrt{S+B}$ in a data subsample corresponding to 20% of the full data sample, where $S$ is the number of signal events and $B$ is the number of background events determined from signal and sideband regions of the $M(K^+K^-\pi^-)$ distributions. Since $y_{\text{sel}}$ is independent of the muon charge information, this optimization does not influence the measured asymmetry and, therefore, the subsample was also used in the complete analysis. The numbers of $D_s^-$ signal events, determined from a fit to the $K^+K^-\pi^-$ invariant mass distributions (see Figs. 3 and 4), are $N(\mu^+\phi\pi^-) = 81, 394 \pm 865$ and $N(\mu^+K^{0}\bar{K}^-) = 33, 557 \pm 1, 200$, where the uncertainties are statistical. In approximately 1% of events, both a $\mu^+K^{0}\bar{K}^-$ candidate and a $\mu^+\phi\pi^-$ candidate were found. To avoid double counting, these events were removed from the $\mu^+K^{0}\bar{K}^-$ sample. The events with two $\mu^+\phi\pi^-$ or $\mu^+K^{0}\bar{K}^-$ candidates were removed from the analysis.

### III. FLAVOR TAGGING

In order to measure the flavor-specific asymmetry $a_{fs}$, it is necessary to distinguish between $B_s^0 \to B_s^0$ and $B_s^0 \to B_s^0$ oscillated states, which requires knowledge of the initial-state (production) and final-state (decay) flavors of the reconstructed $B_s^0$ meson. The final-state $b$ quark flavor is correlated with the charge of the muon in $B_s^0 \to \mu^+D^-\bar{K}^0$ semileptonic decays. The initial-state flavor, which provides additional information in the likelihood fit that is used to extract $a_{fs}$, was determined using an opposite-side tagging (OST) [13,17] algorithm. This algorithm relies in most cases on the reconstruction of a second lepton (muon or electron) from the decay of the other $b$ quark produced in the proton-antiproton interaction. This lepton appears on the side of the detector opposite to the reconstructed $B_s^0$ meson (hence the term “opposite-side” tag) and its charge provides the flavor tag. When a second lepton cannot be identified, the OST algorithm attempts to reconstruct an opposite-side secondary vertex, in which case the charge is determined from the tracks comprising the vertex. Only 21% of the events are tagged; the remaining 79% of events have neither a lepton nor a secondary vertex on the opposite side. Properties of the tagging lepton and secondary vertex tracks are incorporated in the tagging variable, $d_{\text{tag}}$, which is assigned to each $B_s^0$ candidate. By definition, the variable $d_{\text{tag}}$ is defined in the interval $[-1, 1]$. An event with $d_{\text{tag}} > 0$ is tagged as an initial $b$ quark and an event with $d_{\text{tag}} < 0$ is tagged as an initial $\bar{b}$ quark. A higher magnitude $|d_{\text{tag}}|$ corresponds to a higher tagging confidence. Samples of reconstructed $B_s^0 \to \mu^+D^0\bar{K}$ and $B_s^0 \to \mu^+D^0X$ decays were used to empirically determine the calibration function, $D(d_{\text{tag}})$, called dilution (see Tables I, II and V in Ref. [17]). This function is used to calculate the probability $p_{\text{cor}} = (D(d_{\text{tag}}) + 1)/2$ that a given $B_s^0$ candidate has been tagged correctly. In events where no tagging information is available, the dilution is set to zero.

### IV. PROPER DECAY TIME

The proper decay time of each $B_s^0$ candidate is derived from the measured displacement $\bar{L}_T$ of the $B_s^0$ decay vertex from the primary vertex in the transverse plane. A Lorentz transformation of $\bar{L}_T$ into the $B_s^0$ rest frame would yield the desired decay time. However, the undetected neutrino and other nonreconstructed particles in the semileptonic $B_s^0$ decay prevent the precise determination of $p_T$ needed to calculate the Lorentz boost factor. Instead, the combined transverse momentum of the $\mu^+D^0_s$ pair, $p_T(\mu^+D^-_s)$, is used to calculate the visible proper decay length (VPDL)

$$I = M(B_s^0) \cdot [\bar{L}_T \cdot \bar{p}_T(\mu^+D^-_s)]/[p_T(\mu^+D^-_s)]^2, \tag{2}$$

where $M(B_s^0) = 5.3663$ GeV/c$^2$ [18]. The proper decay length of each $B_s^0$ meson is then $c\tau(B_s^0) = I/K$, where $K = p_T(\mu^+D^-_s)/p_T(B_s^0)$ is a correction factor that accounts for the missing momentum. Since $K$ is not known on an event-by-event basis, it was estimated from a Monte Carlo simulation, which included the PYTHIA generator [19] interfaced with the EVTGEN decay package [20], followed by full GEANT [21] modeling of the detector response and event reconstruction. As large samples were required to obtain sufficient statistical precision, only generator-level information was used to determine the $K$ distributions. However, a sample of fully simulated events was used to verify that the difference between generator-level $K$ distributions and those obtained using fully simulated/reconstructed events is negligible. A model of the muon trigger efficiency dependence on $p_T(\mu^+)$ was included in the construction of the $K$ distributions, which were obtained for each decay channel contributing to the signal sample. $B_s^0$ semileptonic decays yielding an invariant $\mu^+D^-_s$ mass that is close to the actual $B_s^0$ mass have
less missing momentum than those with lower \( M(\mu^+ D_s^-) \). Distributions of \( \mathcal{K} \) for a given decay channel in ten bins of \( M(\mu^+ D_s^-) \) were used to exploit this fact, thereby reducing the uncertainty of the proper decay time associated with \( \mathcal{K} \).

The probability density function (PDF) for the \( B_s^0 \) decay time is convoluted with the PDF describing the VPDL detector resolution and the PDF for the \( \mathcal{K} \) factor. The decay-time PDF is then scaled by the detector resolution and the PDF for the decay time. This efficiency is shown as a function of VPDL, \( \tau_B = 1.470 \text{ ps} [18] \), and releasing the signal efficiency parameters in the decay-time fit to the data. In the \( \mu^+ K^0 \bar{K}^- \) sample, the reconstruction efficiency for the decay channels contributing to the signal was then tuned to data by fixing the \( B_s^0 \) lifetime, \( \tau_{B_s^0} \), and releasing the signal efficiency.

For both decay modes, the reconstruction efficiency for combinatorial background was determined from the data, as an adequate Monte Carlo model was not available.

![Graph 1](image1.png)

**FIG. 5.** The reconstruction efficiency as a function of VPDL for \( B_s^0 \to \mu^+ D_s^- X \) decays in the \( \mu^+ \phi \pi^- \) data samples. Plot (a) is for Run IIa and plot (b) is for Run IIb. The width of the grey band indicates the range of the systematic uncertainty.

![Graph 2](image2.png)

**FIG. 6.** The reconstruction efficiency as a function of VPDL for \( B_s^0 \to \mu^+ D_s^- X \) decays in the \( \mu^+ K^0 \bar{K}^- \) Monte Carlo samples. Plot (a) is for Run IIa and plot (b) is for Run IIb. The width of the grey band indicates the range of the systematic uncertainty.

The decay-time dependent component of the signal efficiency was obtained by considering the events passing all the selection criteria except those dependent on the decay time. This efficiency is shown as a function of VPDL for the \( \mu^+ \phi \pi^- \) and the \( \mu^+ K^{*0} K^- \) modes in Figs. 5 and 6, respectively. The VPDL range extends below zero to account for the VPDL resolution. The efficiency improvement visible in the low VPDL region in Run IIb is due to the addition of the Layer 0 detector.

The functional form that best describes the efficiency in the positive VPDL region was chosen for the fit. An analytic function is required to properly normalize the \( B_s^0 \) decay-time PDF. In the \( \mu^+ K^{*0} K^- \) mode, where the signal efficiency was obtained from the simulation, the VPDL value for the minimum was allowed to float in the fit. In the \( \mu^+ \phi \pi^- \) mode, where the signal efficiency was tuned to the data, it was necessary to fix the minimum efficiency point to be at VPDL = 0. A systematic uncertainty accounting for this was obtained by forcing a large negative variation in the efficiency in the negative VPDL region, as indicated by the width of the bands in Fig. 5. In the \( \mu^+ K^{*0} K^- \) mode, the systematic uncertainty was obtained...
by simply varying the fit parameters within their uncertainties, as indicated by the width of the bands in Fig. 6.

The VPDL uncertainty \(\sigma_i\) for a given event depends on the uncertainties of the production and decay vertex positions, which in turn depend on the track parameter uncertainties. The procedure described in Ref. [22] was followed to tune the tracking uncertainties using dijet and dimuon data samples. Realistic uncertainties were extracted from the track impact parameter pull distributions, which were fitted by a single Gaussian function. The range for the fit was varied from between \(\pm 1\sigma\) and \(\pm 3\sigma\). Parametrizations of the dependence of these uncertainties on the track momentum and polar angle were thus obtained and applied to tune the track uncertainties in the analysis. The dijum and dijet data samples yielded slightly different results. Using the tuned track parameter uncertainties, the vertex position uncertainty was also tuned, taking into account the tails, using the pull distributions of the vertex positions of \(J/\psi \rightarrow \mu^+ \mu^-\) decays. Only the negative side of the pull distributions was fitted, as the positive side includes a long-lived component due to \(J/\psi\) candidates from \(B\) decays. Separate vertex uncertainty tunings were obtained for data before and after the Layer 0 installation. The tracking uncertainty differences observed in the dimuon and dijet samples, combined with those associated with the Gaussian fit range variation, yield a 5% variation in the resulting VPDL uncertainty. A 5% systematic uncertainty on the VPDL resolution was therefore included in the analysis.

V. SAMPLE COMPOSITION

There are two contributions to the \(D_s^-\) peaks in Figs. 3 and 4: the \(B_{u,d,s} \rightarrow \mu^+ D_s^- X\) decays (including muons originating from \(D_{s(0)}\) and \(\tau^-\) decays) and the background occurring when the \(D_s^-\) meson originates from one \(b\) or \(c\) quark and the muon arises from the decay of another quark. The fraction of the peak that is background was determined from the data, as described later. The contributions from each \(B\) decay channel, accounting for branching fractions, were determined from the simulation. Measured values for the branching fractions were used [18]. The exclusive branching fractions for semileptonic \(B^0\) decays to \(D_s^-, D_s^{*0}, D_s^{*+}\), and \(D_s^{*-}\) have not been measured. They were therefore calculated from the measured branching fractions of the corresponding decays for \(B^0\) mesons, assuming the spectator model. The uncertainty associated with this assumption is expected to be negligible [23] compared to the experimental branching fraction uncertainties.

The contributions to the \(\mu^+ D_s^-\) signal prior to the application of the decay-time dependent criteria in each mode are shown in Table I. The relative contribution of each source varies with the invariant mass \(M(\mu^+ D_s^-)\) of the reconstructed \(\mu^+ D_s^-\) system. For example, an event with a high \(M(\mu^+ D_s^-)\) is more likely to originate from a direct \(B_s^0 \rightarrow \mu^+ \nu \mu D_s^-\) decay, rather than from a \(B_s^0\)

<table>
<thead>
<tr>
<th>(\mu^+ ) Source (\mu^+ D_s^-)</th>
<th>(\mu^+ K^{0})</th>
<th>Fraction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(B_s^0 \rightarrow \mu^+ D_s^- \nu\mu)</td>
<td>(B_s^0 \rightarrow \mu^+ D_s^{*-} \nu\mu)</td>
<td>(B_s^0 \rightarrow \mu^+ D_s^{*0} \nu\mu)</td>
</tr>
<tr>
<td>20.69</td>
<td>63.26</td>
<td>1.56</td>
</tr>
<tr>
<td>23.76</td>
<td>60.22</td>
<td>1.65</td>
</tr>
</tbody>
</table>

\(\mu^+ \nu \mu D_s^{*-}\) decay, where the intermediate \(D_s^{*-}\) decays to \(D_s^- X\). The relative contributions for each source were therefore binned by invariant mass \(M(\mu^+ D_s^-)\) for an improved model of the sample composition.

The mass PDF models the expected mass and width of the \(K^+ K^-\) signal for each source. Four sources were considered in the \(\mu^+ \phi\) sample (see Fig. 3): the signal \(\mu^+ D_s^- (\phi \pi^-);\) the accompanying mass peak due to \(\mu^+ D^- (\phi \pi^-);\) a small reflection (less than 1% and therefore not visible in the figure) due to \(\mu^+ D^- (\rightarrow K^+ \pi^- \pi^-)\), where the kaon mass is mis-assigned to one of the pions; and combinatorial background. Five sources were considered in the \(\mu^+ K^{0}\) sample (see Fig. 4): the signal \(\mu^+ D_s^- (\rightarrow K^{0}\) K\(^- };\) the mass peak due to \(\mu^+ D^- (\rightarrow K^{0}\) K\(^- };\) a reflection due to \(\mu^+ D^- (\rightarrow K^{0}\) \(\pi^-);\) where the pion is mistaken for a kaon; a reflection due to \(\mu^+ \Lambda_s^- (\rightarrow K^{0}\) \(\bar{p}\); where the antiproton is mistaken for a kaon; and combinatorial background. The fractional contributions of these sources were determined from the mass fits to the data. The mass distributions shown in Figs. 3 and 4 are averages of the mass PDFs for each event.

The analysis of the \(\mu^+ K^{0}\) mode is more challenging due to the large \(D^-\) reflection in the \(D_s^-\) signal region (see Fig. 4). The \(\mu^+ D^-\) candidates arise from decays of all three \(B\) mesons, although the \(B_s^0\) contribution is negligible. The fraction of \(\mu^+ D^-\) candidates originating from \(B_s^0\) decays was assumed to be 80% with 20% arising from \(B^+\) decays [20]. The \(B_s^0\) fraction can be measured experimentally by exploiting the oscillating and nonoscillating characteristics of the \(B_s^0\) and \(B^+\) mesons, respectively. A \(B_s^0\) fraction of 0.93 \(\pm 0.04\) was thus determined from the opposite-side tagged \(\mu^+ K^{0}\) data subsample. The difference between these two \(B_s^0\) fractions was included in the \(a_t\) systematic uncertainty as a possible bias.

VI. LIKELIHOOD FUNCTION

All events with \(1.72 < M(\phi \pi^-) < 2.22\) GeV/c\(^2\) \((1.79 < M(K^{0}\) K\(^-) < 2.25\) GeV/c\(^2\)) were used in an unbinned fitting procedure. While the signal events in both

012003-8
modes are confined to a narrower mass range, the events with masses outside the signal region are needed to accurately describe the VPDL distribution for the combinatorial background under the signal peaks. The total likelihood $\mathcal{L}$ for $N$ selected events is the product of likelihoods $L_j$ determined for each event $j$:

$$\mathcal{L} = \prod_{j=1}^{N} L_j,$$

(3)

where

$$L_j = \sum_{i} \left[ f_i \mathcal{P}^i \mathcal{P}^{\text{pdf}} i \mathcal{P}^{M(k_1, k_2, \ldots, k_n)} \mathcal{P}^{d_{\text{tag}}} \right].$$

(4)

The sum is taken over products of the probability density functions for different sources of $K^+ K^- \pi^-$ candidates with fractions $f_i$. The distribution of the VPDL uncertainty is described by $\mathcal{P}^i$. The distribution of the likelihood ratio selection variable is given by $\mathcal{P}^{\text{pdf}}$. The mass PDF is given by $\mathcal{P}^{M(k_1, k_2, \ldots, k_n)}$. Finally, $\mathcal{P}^{d_{\text{tag}}}$ is the distribution of the tagging variable $d_{\text{tag}}$. These PDFs were determined from data.

The PDF $\mathcal{P}^{i}$ for the measured VPDL $l$ was constructed as follows. The formulae for the decay rates of neutral $B^0$ mesons were taken from Ref. [24] assuming no direct CP violation (i.e., $|\mathcal{A}_f| = |\mathcal{A}_d|$), where $\mathcal{A}_f$ and $\mathcal{A}_d$ are the decay amplitudes). The asymmetry $\alpha_{fs}^d$ modifies the decay rates of mixed $B^0$ mesons as follows:

$$\Gamma_{B^0(t) \rightarrow f} = N_f |\mathcal{A}_f|^2 e^{-\Gamma_i/2} \left[ \cosh(\Delta \Gamma_i t/2) + \cos(\Delta m_i t) \right]/2,$$

(5)

$$\Gamma_{B^0(t) \rightarrow f} = N_f |\mathcal{A}_d|^2 e^{-\Gamma_i/2} \left[ \cosh(\Delta \Gamma_i t/2) - \cos(\Delta m_i t) \right]/2,$$

(6)

$$\Gamma_{B^0(t) \rightarrow f} = N_f |\mathcal{A}_f|^2 e^{-\Gamma_i/2} \left[ \cosh(\Delta \Gamma_i t/2) + \cos(\Delta m_i t) \right]/2,$$

(7)

where $N_f$ is the normalization to the total number of $B^0$ mesons. Using the relationship $ct(B^0) = 2\mathcal{K}$ discussed previously, these decay-rate PDFs can be written in terms of VPDL. All reconstructed events were divided into four samples corresponding to the final- and initial-state tags. The dilution of the initial-state tagging leads to a mixture of $B^0$ and $\bar{B}^0$ initial states in these samples, e.g., the sample tagged as $B^0_{\text{in}}$ in the initial state and $\bar{B}^0_{\text{in}}$ in the final state has the PDF

$$\mathcal{P}^i_{B^0_{\text{in}}/\bar{B}^0_{\text{in}}} = \Gamma_{B^0(t) \rightarrow f} P_{\text{cor}} + \Gamma_{\bar{B}^0(t) \rightarrow f} (1 - P_{\text{cor}}),$$

(9)

where $\Gamma_{B^0(t) \rightarrow f}$ is the PDF in Eq. (6) and $\Gamma_{\bar{B}^0(t) \rightarrow f}$ is the PDF in Eq. (7). As an example, the average VPDL PDFs for unmixed events (i.e., $\mathcal{P}^{i}_{B^0_{\text{in}}/\bar{B}^0_{\text{in}}}$ and $\mathcal{P}^{i}_{B^0_{\text{in}}/\bar{B}^0_{\text{in}}}$) and mixed events (i.e., $\mathcal{P}^{i}_{B^0_{\text{in}}/\bar{B}^0_{\text{in}}}$ and $\mathcal{P}^{i}_{B^0_{\text{in}}/\bar{B}^0_{\text{in}}}$) for the Run IIb $\mu^+ \phi \pi^-$ sample are shown in Figs. 7 and 8, respectively. For the figures, the combinatorial background was subtracted using sidebands of the $K^+ K^- \pi^-$ invariant mass distributions. A contribution describing the background due to fake vertices around the primary vertex was included into the PDF. The slight differences in the PDFs for the mixed events in Fig. 8 give rise to a nonzero asymmetry $\alpha_{fs}^d$.

The decay-rate PDF for the semileptonic decays $B^0 \rightarrow \mu^+ D^+ X$ (the events comprising the peaks at 1.87 GeV/$c^2$ in Figs. 3 and 4) is the same as for the $B^0$ decays with the corresponding parameters changed accordingly. In particular, the $B^0_d$ semileptonic asymmetry $\alpha_{fs}^d$ was introduced and determined from the fit.
V. M. ABAZOV et al.

VII. DETECTOR ASYMMETRIES

The PDFs are modified to account for the detector charge asymmetries [8,25]:

\[
A_\mu = (1 + \gamma A_{\text{det}})(1 + q \beta A_{\text{ro}})(1 + q \gamma A_{\text{mb}}) \\
\cdot (1 + \beta \gamma A_{\beta \gamma})(1 + q \beta A_{q \beta}),
\]

(10)

where \( \beta \) is the toroid polarity, \( \gamma \) is the sign of the muon pseudorapidity, and \( q \) is the muon charge. \( A_{\text{det}} \) accounts for any asymmetry due to differences in the north \((z < 0)\) and south \((z > 0)\) ends of the detector (the proton beam travels from north to south). The range-out asymmetry \( A_{\text{ro}} \) reflects the difference in acceptance of muons that bend towards as opposed to away from the beam line (for details, see Ref. [8]). The forward-backward asymmetry \( A_{\text{fs}} \) reflects the fact that positively charged muons tend to go in the direction of the proton beam whereas negatively charged muons typically go in the antiproton beam direction. \( A_{\beta \gamma} \) is a second-order correction to \( A_\mu \) that is nonzero only if \( a_{fs}^1 \) and \( A_{\text{ro}} \) are both nonzero. \( A_{q \beta} \) is a detector asymmetry between tracks bending towards \( \eta < 0 \) and tracks bending towards \( \eta > 0 \). The likelihood for each event was multiplied by the detector charge asymmetry corrections for the muon from the \( B^0 \) semileptonic decay and for the muon used for the OST when it is present. Both kaons in the \( \mu^+ \phi \pi^- \) signal sample originate from the \( \phi \) decay and therefore have the same transverse momentum threshold. Since this is not the case in the \( \mu^+ K^{*-0} K^- \) sample, the momentum-dependent kaon reconstruction asymmetry [26] was taken into account.

VIII. BACKGROUND DESCRIPTION

The background where the \( D_s^- \) meson and the muon do not come from the same parent particle gives rise to fake vertices and a VPDL distribution that peaks around zero. Its shape was modeled by two Gaussian functions, and its contribution was estimated from decay-time fits to be approximately 8% for the total sample and 3% for the opposite-side tagged subsample for the \( \mu^+ \phi \pi^- \) mode. The corresponding contributions are 4% and 1% for the \( \mu^+ K^{*-0} K^- \) mode. The fake vertex background fraction is lower in the \( \mu^+ K^{*-0} K^- \) mode due to the higher kinematic selection criteria. Several contributions to the combinatorial background that have different VPDL distributions were considered. True prompt background was modeled with a Gaussian function. The long-lived combinatorial background is dominated by misreconstructed heavy flavor decays. This background was modeled with an exponential function convoluted with the VPDL resolution, including a component \((\approx 60\% \) of all long-lived background contributions) oscillating with a frequency of \( \Delta m_d \). An asymmetry parameter \( a_{bg} \) was introduced into the PDFs for this component by analogy with Eqs. (6) and (7). This parameter absorbs possible asymmetries in the combinatorial background that were unaccounted for. The unbinned likelihood fit of the total sample was used to determine the various fractions of signal and backgrounds and the background VPDL parametrizations.

IX. RESULTS

The \( B^0 \) oscillation frequency \( \Delta m_s = 17.77 \pm 0.12 \text{ ps}^{-1} \) [27] was fixed in the fit. In the \( \mu^+ \phi \pi^- \) mode, the width difference was fixed to \( \Delta \Gamma_s = 0.09 \pm 0.05 \text{ ps}^{-1} \) [18]. In the \( \mu^+ K^{*-0} K^- \) mode, the ratio \( \Delta \Gamma_s / \Gamma_s = 0.069^{+0.058}_{-0.062} \) was fixed in the fit. First, the \( \mu^+ \phi \pi^- \) and \( \mu^+ K^{*-0} K^- \) samples were fitted separately (see Table II). The results for the charge asymmetries are consistent between these samples. The results of a fit to the combined \( \mu^+ \phi \pi^- \) and \( \mu^+ K^{*-0} K^- \) samples are shown in the last column of Table II. The measured value for the flavor-specific asymmetry is \( a_{fs}^1 = [-1.7 \pm 9.1] \times 10^{-3} \), where the uncertainty is statistical. The detector charge asymmetry \( A_{\text{ro}} \) is observed to be nonzero, as expected [8]. The other detector asymmetries are consistent with zero, as is the physics asymmetry \( a_{bg} \). There is no correlation between the detector charge asymmetries, either amongst themselves or with the physics asymmetries. There is a 5.2% (13.6%) correlation between \( a_{fs}^1 \) and \( a_{fs}^3 \) (\( a_{bg} \)). The nonzero value of \( a_{fs}^1 \) is discussed in the next section.

X. SYSTEMATIC UNCERTAINTIES

The sources contributing to the total \( a_{fs}^1 \) systematic uncertainty of \( +1.4 \times 10^{-3} \) are listed in Table III. The largest contribution is due to the momentum-dependent kaon asymmetry. The fitted value of \( a_{fs}^1 \) is sensitive to this asymmetry due to the large \( D^- \) reflection under the \( D_s^- \) signal peak in the \( \mu^+ K^{*-0} K^- \) mode. To account for this, and for the correlation between \( a_{fs}^1 \) and \( a_{fs}^3 \), the kaon asymmetry was varied between zero and twice its nominal value and the likelihood fit was repeated in each case. Setting the kaon asymmetry to zero yielded a fitted value of \( a_{fs}^1 = [-0.2 \pm 16.5] \times 10^{-3} \), whereas scaling it by two yielded \( a_{fs}^1 = [81.2 \pm 16.6] \times 10^{-3} \). The resulting change in the fitted \( a_{fs}^1 \) value is taken as the associated systematic uncertainty, given in Table III.

The same procedure was followed for the remaining sources of uncertainty. The signal fraction, that is, the

<table>
<thead>
<tr>
<th>TABLE II. Asymmetries with statistical uncertainties.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu^+ \phi \pi^- )</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>( a_{fs}^1 \times 10^3 )</td>
</tr>
<tr>
<td>( a_{fs}^2 \times 10^3 )</td>
</tr>
<tr>
<td>( a_{bg} \times 10^3 )</td>
</tr>
<tr>
<td>( a_{\beta \gamma} \times 10^3 )</td>
</tr>
<tr>
<td>( A_{\text{det}} \times 10^3 )</td>
</tr>
<tr>
<td>( A_{\text{fs}} \times 10^3 )</td>
</tr>
<tr>
<td>( A_{q \beta} \times 10^3 )</td>
</tr>
</tbody>
</table>
TABLE III. Systematic uncertainties.

<table>
<thead>
<tr>
<th>Source of Uncertainty</th>
<th>σ(Δαs) × 10^3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaon asymmetry set to 0</td>
<td>1.24</td>
</tr>
<tr>
<td>Kaon asymmetry scaled by 2</td>
<td>1.30</td>
</tr>
<tr>
<td>Signal fraction − 1σ</td>
<td>−0.76</td>
</tr>
<tr>
<td>Signal fraction + 1σ</td>
<td>0.47</td>
</tr>
<tr>
<td>Dilution scaled by 0.9</td>
<td>−0.19</td>
</tr>
<tr>
<td>Dilution scaled by 1.1</td>
<td>0.21</td>
</tr>
<tr>
<td>μ trigger efficiency low</td>
<td>−0.03</td>
</tr>
<tr>
<td>μ trigger efficiency high</td>
<td>0.00</td>
</tr>
<tr>
<td>Decay-time dependent efficiency low</td>
<td>0.15</td>
</tr>
<tr>
<td>Decay-time dependent efficiency high</td>
<td>−0.01</td>
</tr>
<tr>
<td>VPDL resolution scaled by 0.95</td>
<td>0.03</td>
</tr>
<tr>
<td>VPDL resolution scaled by 1.05</td>
<td>−0.03</td>
</tr>
<tr>
<td>BF B_d^0 → D_s^− D_s^+</td>
<td>0.00</td>
</tr>
<tr>
<td>BF B_d^0 → μ^+D_s^- X</td>
<td>−0.10</td>
</tr>
<tr>
<td>Relative BF B_d^0 → μ^+νμD_s^- low</td>
<td>0.01</td>
</tr>
<tr>
<td>Relative BF B_d^0 → μ^+νμD_s^- high</td>
<td>−0.05</td>
</tr>
<tr>
<td>B_d^0 fraction in μ^+D^- candidates set to 93%</td>
<td>−0.24</td>
</tr>
<tr>
<td>Fake vertex background low</td>
<td>−0.13</td>
</tr>
<tr>
<td>Fake vertex background high</td>
<td>−0.04</td>
</tr>
<tr>
<td>Prompt combinatorial background low</td>
<td>0.01</td>
</tr>
<tr>
<td>Prompt combinatorial background high</td>
<td>−0.01</td>
</tr>
<tr>
<td>ΔΓ_s − 1σ</td>
<td>0.00</td>
</tr>
<tr>
<td>ΔΓ_s + 1σ</td>
<td>−0.01</td>
</tr>
<tr>
<td>Δm_s − 1σ</td>
<td>−0.01</td>
</tr>
<tr>
<td>Δm_s + 1σ</td>
<td>0.02</td>
</tr>
<tr>
<td>Total</td>
<td>−1.50</td>
</tr>
</tbody>
</table>

number D_s^- and D^- signal events divided by the total number of events, obtained from the mass fits was varied up and down by 1σ. The dilution of the opposite-side tagging algorithm was scaled by 0.9 and 1.1. The muon trigger efficiency was varied within its experimental uncertainties. The decay-time dependent reconstruction efficiency as a function of VPDL was varied, as previously described. The VPDL resolution was scaled by 0.95 and 1.05 to account for the tracking uncertainty differences previously discussed. The fake vertex background, where the D^- meson and the muon do not come from the same parent particle, was varied within its experimental range as was the prompt component of the combinatorial background. The B_d^0 − B_d^0 width difference ΔΓ_s and the oscillation frequency Δm_s were varied within their experimental uncertainties. The relative branching fractions (BF) for the exclusive semileptonic B_d^0 decays were varied within their predicted uncertainties in such a way as to keep the total inclusive fraction constant. To model a reduced signal fraction, the B_d^0 → D_s^− D_s^+ branching fraction was increased by 1σ and the B_d^0 → μ^+D_s^- X inclusive branching fraction was decreased by 1σ. To account for a possible bias associated with the fraction of μ^+D^- candidates in the μ^-K^-K^- sample originating from B_d^0 decays (see Fig. 4), this fraction was changed from 0.80 to 0.93 in the likelihood fit. The total systematic uncertainty was determined by adding all the signed contributions in quadrature.

XI. CONCLUSION

In summary, using B_d^0 → μ^+D^- X decays with D_s^- → ϕπ^−, ϕ → K^+K^- and D_s^- → K^-0K^−, K^-0 → K^+π^−, in combination with a decay-time analysis including initial-state flavor tagging, we measure the asymmetry in mixed semileptonic B_d^0 decays to be \( a_{fs} = [-1.7 \pm 9.1(\text{stat})^{+0.8}_{-1.5}(\text{syst})] \times 10^{-3} \). This measurement supersedes the D_0 time-integrated analysis of semileptonic B_d^0 decays [25], which yielded \( a_{fs} = [24.5 \pm 19.3(\text{stat}) \pm 3.5(\text{syst})] \times 10^{-3} \). Our result is also consistent with the value \( a_{fs} = (-6.4 \pm 10.1) \times 10^{-3} \) extracted [28] from the D_0 time-integrated analysis of inclusive same-sign dimuon events [8]. While the present result is the most precise measurement of the semileptonic B_d^0 asymmetry, improved precision is needed to establish evidence of CP violation due to new physics in B_d^0 mixing.

ACKNOWLEDGMENTS

We thank the staff at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); STFC and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); FOM (The Netherlands); STFC and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); CAS and CNSF (China); and the Alexander von Humboldt Foundation (Germany).


[6] This $a_{N}^{a}$ value is calculated by the authors and neglects correlations between systematic uncertainties.


