



B_s Mixing studies with $B_s^0 \rightarrow D_s^- \mu^+ X$ ($D_s^- \rightarrow K^{*0} K^-$) Decay using Unbinned Fit

The DØ Collaboration
URL <http://www-d0.fnal.gov>

(Dated: July 21, 2006)

Measurement of the B_s^0 oscillation frequency via B_s^0 mixing provides a powerful constraint on the CKM matrix elements and might reveal new physics phenomena to which B_s^0 oscillations are sensitive. At DØ, we already have set a two sided bound on the mixing parameter using our large semileptonic data sample. This note briefly reviews the motivation behind this analysis and describes the various steps that go into a mixing measurement. Preliminary results obtained using the semileptonic $B_s^0 \rightarrow D_s^- \mu^+ X$ with D_s decaying to $D_s^- \rightarrow K^{*0} K^-$ ($K^{*0} \rightarrow K^+ \pi^-$) final state with $\sim 1.2 \text{ fb}^{-1}$ of data collected with the DØ detector during Run II of the Fermilab Tevatron are presented here.

I. INTRODUCTION

Mixing is the process whereby some neutral mesons change from their particle to their anti-particle state, and vice versa. This kind of oscillation of flavor eigenstates into one another was first observed in the K^0 meson system. It has since been seen for B mesons, first in a mixture of B_d^0 and B_s^0 by UA1[1] and then in B_d^0 mesons by ARGUS[2]. The frequency of the oscillation is proportional to the small difference in mass between the two eigenstates, Δm , and for the $B_d^0 - \bar{B}_d^0$ system can be translated into a measurement of the CKM element $|V_{td}|$. Δm_d has been precisely measured (the world average is $\Delta m_d = 0.502 \pm 0.007 \text{ ps}^{-1}$)[3] but large theoretical uncertainties dominate the extraction of $|V_{td}|$ from Δm_d . This problem can be reduced if the B_s^0 mass difference, Δm_s , is also measured. $|V_{td}|$ can then be extracted with better precision from the ratio:

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m(B_s^0)}{m(B_d^0)} \xi^2 \left| \frac{V_{ts}}{V_{td}} \right|^2 \quad (1)$$

where ξ is estimated from Lattice QCD calculations to be $1.15 \pm 0.05_{-0.00}^{+0.12}$ (Ref. [3]). The recent double sided bound on B_s mixing by the $D\mathcal{O}$ collaboration [4] and it's confirmation later by CDF collaboration [5] has let to a lot of excitement and enthusiasm in the Flavor Physics sector and provided a very good opportunity to these experiments to measure the B_s oscillations as precisely as possible. If the Standard Model is correct, then $\Delta m_s = 18.3_{-1.5}^{+6.5} \text{ ps}^{-1}$ from global fits to the unitarity triangle if the current experimental limits on Δm_s are included in the fit. If information from B_s^0 oscillations limits are not included, global fits give $\Delta m_s = 20.9_{-4.2}^{+4.5} \text{ ps}^{-1}$ [6]. The current measurements indicate a value well within the SM allowed range and small enough to allow us to measure Δm_s with semileptonic decays.

II. DETECTOR DESCRIPTION

The $D\mathcal{O}$ detector is described in detail elsewhere [7, 8]. The following main elements of the $D\mathcal{O}$ detector are essential for this analysis:

- The magnetic central-tracking system, which consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2-T superconducting solenoidal magnet;
- The liquid-argon/uranium calorimeter;
- The muon system located beyond the calorimeter.

The SMT has 800,000 individual strips, with typical pitch of $50 - 80 \mu\text{m}$, and a design optimized for tracking and vertexing capability at $|\eta| < 3$, where $\eta = -\ln(\tan(\theta/2))$ and θ is the polar angle. The CFT has eight thin coaxial barrels, each supporting two doublets of overlapping scintillating fibers of 0.835 mm diameter, one doublet being parallel to the collision axis, and the other alternating by $\pm 3^\circ$ relative to the axis. The resolution of the impact parameter with respect to the collision point is about $20 \mu\text{m}$ for $5 \text{ GeV}/c$ tracks.

The three components of the liquid-argon/uranium calorimeter are housed in separate cryostats. A central section, lying outside the tracking system, covers up to $|\eta| = 1.1$. Two end calorimeters extend the coverage to $|\eta| \approx 4$.

The muon system consists of a layer of tracking detectors and scintillation trigger counters inside a 1.8 T iron toroid, followed by two additional layers outside the toroid. Tracking at $|\eta| < 1$ relies on 10 cm wide drift tubes, while 1-cm mini-drift tubes are used at $1 < |\eta| < 2$.

III. EVENT SELECTION

All tracks in an event were clustered into jets using the DURHAM clustering algorithm [9] with the cutoff parameter of $15 \text{ GeV}/c$. The following requirements were made to identify the $B_s^0 \rightarrow D_s^- \mu^+ X$, $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ \pi^-$ decay chain. The muon was identified using the standard $D\mathcal{O}$ algorithm [10]. It was required to have $p_T > 2 \text{ GeV}/c$ and $p > 3 \text{ GeV}/c$, to have hits in both the CFT and SMT and to have at least 2 measurements in the muon chambers. Particles were assigned the masses of kaons (K_1 and K_2) and pion requiring the charge combination $\mu^+ K_1^+ K_2^- \pi^-$ or its charge conjugate. For the $D_s \rightarrow K^* K$ channel, transverse momenta were required to be: $p_T(K_1) > 0.9 \text{ GeV}/c$, $p_T(K_2) > 1.8 \text{ GeV}/c$ and $p_T(\pi) > 0.5 \text{ GeV}/c$, assuming that K_1 comes from the $K^{*0} \rightarrow K^+ \pi^-$ decay.

For each particle, the transverse[15] ϵ_T and longitudinal[16] ϵ_L projections of track impact parameter with respect to the primary vertex together with the corresponding errors ($\sigma(\epsilon_T)$, $\sigma(\epsilon_L)$) were computed. The combined significance

$(\epsilon_T/\sigma(\epsilon_T))^2 + (\epsilon_L/\sigma(\epsilon_L))^2$ was required to be greater than 4 for K_1 and K_2 , while there was no cut on the significance of the pion.

Three charged particles were required to come from the same D_s^- vertex with the χ^2 of the vertex fit satisfying $\chi^2 < 16$. The D_s^- candidate produced by their combination was required to have a common B vertex with the muon with the χ^2 of the vertex fit such that $\chi^2 < 9$. The mass of the $\mu^+ D_s^-$ system was required to be $2.6 < M(\mu^+ D_s^-) < 5.4$ GeV/ c^2 . The distance d_T^D in the axial plane between the D_s^- vertex and the primary interaction point was required to satisfy $d_T^D/\sigma(d_T^D) > 4$. The angle α_T^D between the momentum direction of the D_s^- candidate and the direction from the primary to the D_s^- vertex in the axial plane was required to fulfill the condition: $\cos(\alpha_T^D) > 0.9$.

If the distance d_T^B between the primary and B_s^0 vertex in the axial plane exceeded $4 \cdot \sigma(d_T^B)$, the angle α_T^B between the B_s^0 momentum and the direction from the primary to B_s^0 vertex in the axial plane was required to satisfy the condition: $\cos(\alpha_T^B) > 0.95$. The distance d_T^B was allowed to be greater than d_T^D , provided that the distance between the B_s^0 and D_s^- vertices d_T^{BD} was less than $2 \cdot \sigma(d_T^{BD})$.

Additionally, the condition $0.82 < M(K_1\pi) < 0.95$ was applied. The final event samples were selected using the likelihood ratio method, described below.

It is assumed that a set of discriminating variables x_1, \dots, x_n can be constructed for a given event. It is also assumed that the probability density functions $f_i^s(x_i)$ for the signal and $f_i^b(x_i)$ for the background can be built for each variable x_i . The combined tagging variable y is defined as:

$$y = \prod_{i=1}^n y_i; \quad y_i = \frac{f_i^b(x_i)}{f_i^s(x_i)}. \quad (2)$$

The selection of the signal is obtained by applying the cut on $y < y_0$. For uncorrelated variables x_1, \dots, x_n , the selection using the combined variable y gives the best possible tagging performance, i.e., maximal signal efficiency for a given background efficiency.

The following discriminating variables were used:

- Helicity angle, defined as the angle between the D_s^- and K_1 momenta in the $(K_1\pi)$ center of mass system;
- Isolation, computed as $Iso = p^{tot}(\mu D_s)/(p^{tot}(\mu D_s) + \sum p_i^{tot})$. The sum $\sum p_i^{tot}$ was taken over all charged particles in the cone $\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$, where $\Delta\eta$ and $\Delta\phi$ are the pseudorapidity and the azimuthal angle with respect to the (μD_s) direction. The μ^+ , K_1 , K_2 and π^- were not included in the sum;
- $p_T(K_2)$;
- Invariant mass, $M(\mu^+ D_s^-)$;
- χ^2 of the D_s^- vertex fit;
- $M(K_1\pi)$.

The probability density functions were constructed using real data events. For each channel, three bands B_1 , B_2 and S were defined as:

$$\begin{aligned} B_1 &: 1.75 < M(D_s^-) < 1.79 \text{ GeV}/c^2 \\ B_2 &: 2.13 < M(D_s^-) < 2.17 \text{ GeV}/c^2 \\ S &: 1.92 < M(D_s^-) < 2.00 \text{ GeV}/c^2 \end{aligned}$$

The background probability density function for each variable was constructed using events from the B_1 and B_2 bands. The signal probability density function was constructed by subtracting the background, obtained as a sum of distributions in the B_1 and B_2 bands, from the distribution of events in band S . The final selection of events for the analysis was done by applying a cut on the combined variable $\log_{10} y < 0.16$. This cut was selected by requiring the maximal value of $S/\sqrt{S+B_1+B_2}$.

Figure 1 shows the $K^{*0}K^-$ invariant mass distribution after all the selection cuts. Distributions for both the ‘‘right-sign’’ $D_s^- \mu^+$ combinations ($Q_\mu * Q_\pi < 0$) and the ‘‘wrong-sign’’ $D_s^- \mu^-$ combinations ($Q_\mu * Q_\pi > 0$) are shown.

IV. MASS FITTING PROCEDURE

In the decays $D^- \rightarrow K^+ \pi^- \pi^-$, when the pion is mis-identified as a kaon it peaks right under the $D_s^- \rightarrow K^{*0}K^-$ signal peak and this peak strongly overlaps. The final state $KK\pi$ mass spectrum could be a

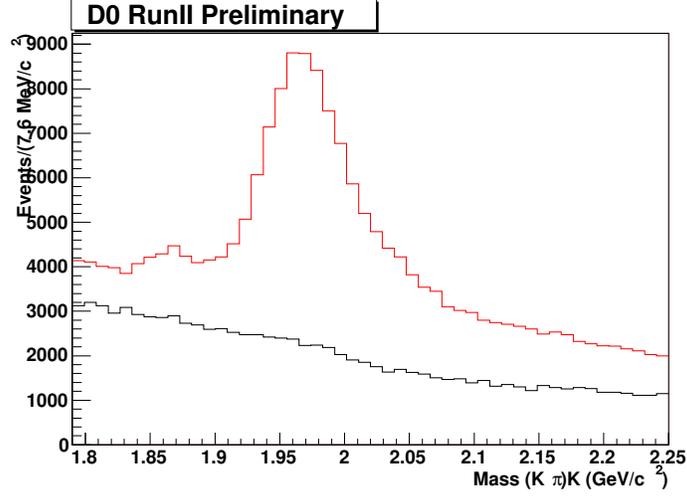


FIG. 1: Distribution of the mass of $D_s^- \rightarrow K^{*0} K^-$ candidates. Both “right-sign” (red) and “wrong-sign” (black) combinations are shown.

mixture of many different decays so the shape of reflection $D^- \rightarrow K^+ \pi^- \pi^-$ strongly depends on kinematical properties of events and changes for different selections. Due to large reflections coming from the mis-identification of pion track as a kaon track, extracting the signal becomes very difficult in $D_s^- \rightarrow K^{*0} K^-$ decays. In order to separate the signal from the large reflection we developed an event-by-event fit based on the kinematic properties of the events. The details of the unbinned mass fitting procedure is as follows:

Consider the decay $D^- \rightarrow K^+ \pi^- \pi^-$. The mass of $K\pi\pi$ system is given as:

$$M_D^2 = (E_{K\pi} + E_\pi)^2 - (\vec{P}_{K\pi} + \vec{P}_\pi)^2 \quad (3)$$

where $E_{K\pi}$ is the energy of $(K^+ \pi^-)$ and E_π , P_π is the energy and momentum of the second pion respectively. $E_\pi = \sqrt{P_{Tr}^2 + M_\pi^2}$, is the energy of track assuming the pion mass hypothesis, When this pion is assigned the mass of Kaon, the shifted mass of $K\pi\pi$ system is

$$M_R^2 = (E_{K\pi} + E_K)^2 - (\vec{P}_{K\pi} + \vec{P}_K)^2 \quad (4)$$

$$\vec{P}_K = \vec{P}_\pi = \vec{P}_{Tr} \quad (5)$$

where, $E_K = \sqrt{P_{Tr}^2 + M_K^2}$ is the energy of the track assuming the kaon mass hypothesis. We can express M_R as:

$$M_R^2 = M_D^2 + (E_{K\pi} + E_K)^2 - (E_{K\pi} + E_\pi)^2 \quad (6)$$

$$M_R^2 = M_D^2 + (M_K^2 - M_\pi^2) + 2E_{K\pi}(E_K - E_\pi) \quad (7)$$

$$M_R^2 = M_D^2 + (1 + 2R)(M_K^2 - M_\pi^2) \quad (8)$$

where,

$$R = \frac{E_{K\pi}(E_K - E_\pi)}{(M_K^2 - M_\pi^2)} \quad (9)$$

is the reflection variable.

A similar equation can be written for reflection due to $\Lambda_c \rightarrow K^+ \pi^- P^+$, where the proton can be mis-identified as kaon.

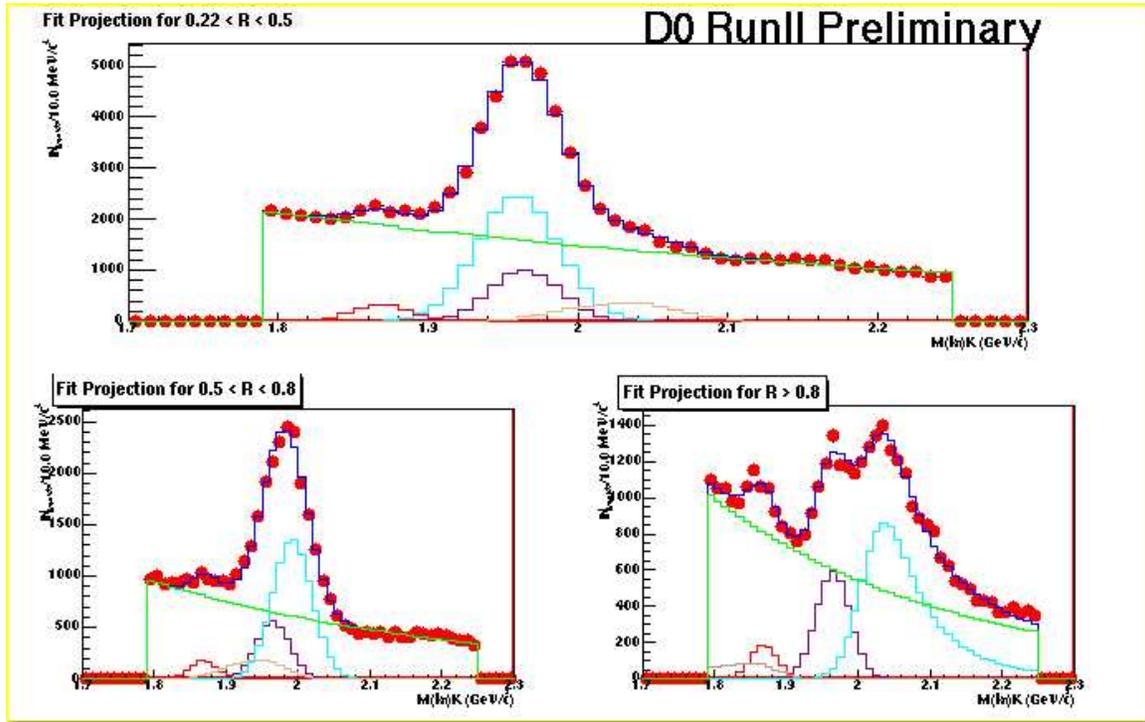


FIG. 2: Distribution of $(K\pi)K$ mass in three different bins of the variable R with the fit results overlaid. The individual histograms at the bottom show the different components separately.

From equation 9, we see that the shifted mass depends on kinematic properties of the events but for a given $E_{K\pi}$ and P_{Tr} , the shift is constant. Also, for a given $E_{K\pi}$ and P_{Tr} the mass distribution of the reflection is determined only by the detector resolution and can be approximated by a Gaussian.

Figure 2 shows the distribution of $(K\pi)K$ mass system in different bins of R variable. From these plots, it can be seen that the shape of the combinatoric background changes significantly with R .

The $(K\pi)K$ mass range of $M_{min} = 1.79 \text{ GeV}$ to $M_{max} = 2.25 \text{ GeV}$ was chosen for fit. Four decay channels which include two signal and two physics background channels and combinatoric background were considered as follows:

- $D_s^- \rightarrow K^{*0}K^-$ (The signal) with fraction f_{sig} .
- $D^+ \rightarrow K\pi\pi$ or $D^+ \rightarrow K^{*0}\pi(K^{*0} \rightarrow K^+\pi^-)$ (Reflection) with fraction f_{Dr} .
- $\Lambda_c^+ \rightarrow K^+\pi^-P^+$ (Reflection) with fraction f_{Lr} .
- $D^+ \rightarrow K^{*0}K^+(K^{*0} \rightarrow K^+\pi^-)$ (Cabbibo suppressed decay) with fraction f_{Dp} .
- Combinatorial background with fraction $f_{bkg}(= 1 - f_{sig} - f_{Dr} - f_{Lr} - f_{Dp})$.

All the fractions were determined from the fit to the data. For each event i with given value of R , the *pdf* of a given channel j is given as:

$$P_i^j(M) = \frac{1}{\sqrt{2\pi}\sigma_j} \cdot \exp\left(-\frac{(M - M_i^j(R_i))^2}{2\sigma_j^2}\right) \quad (10)$$

where M_i^j is the shifted mass defined in equation 8.

The fraction of each channel is given by

$$f_i^j = f_0^j \cdot C(R) \cdot N_i^j \quad (11)$$

Where f_0^j is the free parameter in the fit. Value of this parameter for each channel is listed in table I. We introduce the term $C(R)$ to allow the change of channel fraction f_i^j with R . $C(R)$ is parameterized as

$$C(R) = 1.0 + R \cdot p0 + R^2 \cdot p1 \quad (12)$$

Here $p0$ and $p1$ are parameters returned by the fit. See table II for the values of these parameters. The distribution of $C(R)$ is given in Fig. 3. Normalization coefficient N_i^j takes into account that for a given event i and given channel j , a part of pdf P_i^j can be outside the selection range $1.79 < M(K\pi K) < 2.25$ and hence renormalizes the pdf as,

$$N_i^j = \int_{M_{min}}^{M_{max}} dM P_i^j(M) \quad (13)$$

The background pdf is given as exponential

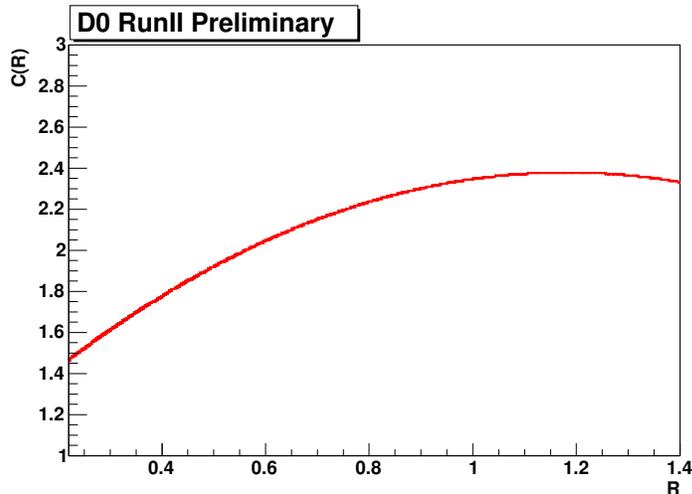


FIG. 3: Distribution of $C(R)$.

$$P_{bkg}(M) = \frac{1}{N_{bkg}} \cdot \exp\left(-\frac{M}{M_0(R)}\right) \quad (14)$$

where we allow the slope $M_0(R)$ to change with R . We parameterize this variation by a polynomial

$$M_0(R) = p2 \cdot (1.0 + R \cdot p3 + R^2 \cdot p4 + R^3 \cdot p5 + R^4 \cdot p6) \quad (15)$$

where, $p2$, $p3$, $p4$, $p5$ and $p6$ are the free parameters in the fit. See table II for the values of these parameters. The distribution of $M_0(R)$ is given in Fig. 4. The total likelihood is given as

$$L_n = \Pi_i \left(\sum_j f_i^j \cdot P_i^j + f_{bkg} \cdot P_{bkg} \right) \quad (16)$$

where,

$$f_{bkg} = (1 - \sum_j f_i^j)$$

The following form is being minimized using the MINUIT program:

$$\mathcal{L} = -2 \cdot \ln L_n \quad (17)$$

For small values of R the kinematical threshold distorts the shape of the background and it can not be described by an exponential any more. That is why in this analysis we use events with $R > 0.22$ only. We checked that for

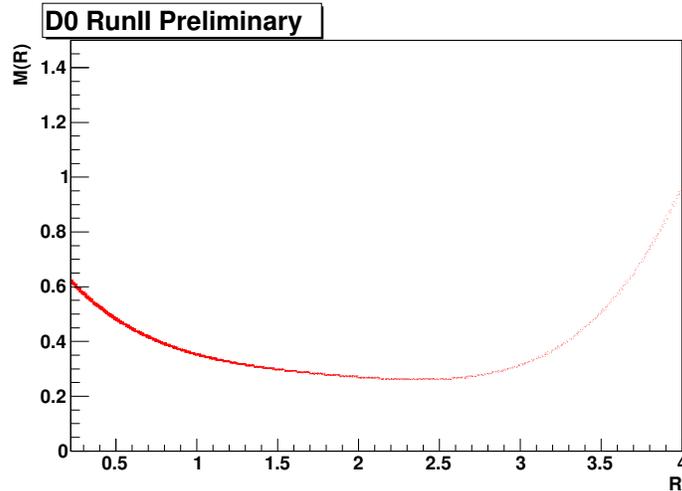


FIG. 4: Distribution of $M_0(R)$.

$R > 0.22$ the distortion of the background by threshold effects is negligible. Unless stated otherwise, all the figures sensitive to the R variable and final results in this note were obtained for $R > 0.22$.

The fitting program was run for all the untagged events and allowing all the parameters to float. The resulting fit can be seen in the Figure 5. We obtain 12647 ± 740 signal candidates, 35937 ± 1856 D^+ reflection candidates, 3232 ± 258 Cabbibo suppressed candidates and 5820 ± 397 Λ_c candidates. Once we obtained the yields for the total untagged sample we then fix all the parameters except the fraction of events in different components and parameter $p2$ in order to estimate the yields for the tagged, unmixed and mixed candidates. We observe that the ratio of events in Cabbibo suppressed decay to D^+ reflection is constant within errors for untagged, tagged, mixed and unmixed samples. This is another cross check of the validity of the fit as this ratio is supposed to be constant irrespective of the sample. We fix this ratio from the total untagged sample in the fit for the tagged sample. Table I shows the masses, widths and fractions obtained from the mass fit for the total untagged sample.

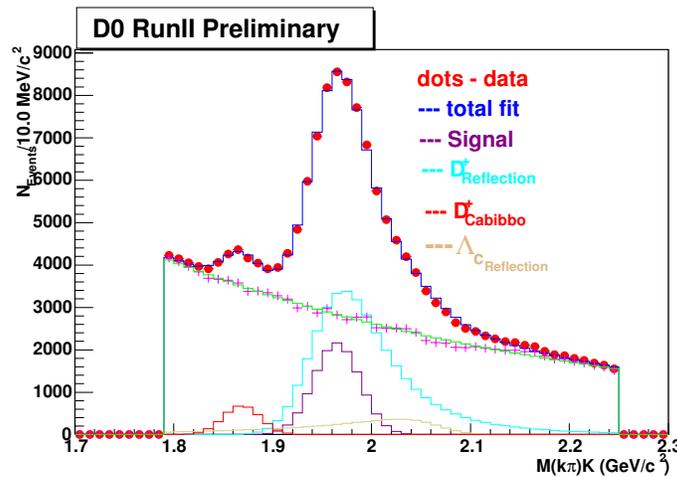


FIG. 5: Fit to the total untagged sample. The red crosses are the signal subtracted background and the green line is the fit to the combinatoric background.

V. INITIAL STATE FLAVOR TAGGING

In order to measure the mixing oscillations, we need to determine whether a B_s^0 has mixed or not. To know the initial flavor of the B_s^0 mesons, an Initial State Flavor Tagging technique was used. The second B meson (or

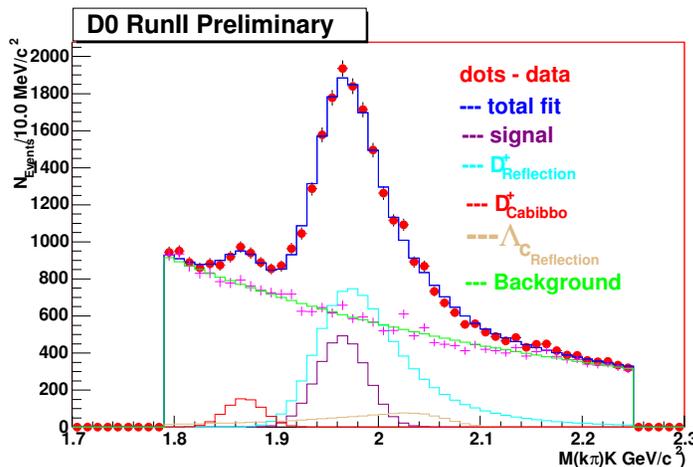


FIG. 6: Fit to the total tagged sample, dots represents the data points and histogram is the fit result.

TABLE I: Fit parameters from the mass fit

Decay Channel	Mass (GeV/c^2)	Width (MeV/c^2)	Fractions
$D_s^- \rightarrow K^{*0} K^-$	1.9647 ± 0.0006	25.76 ± 0.05	0.036 ± 0.0021
$D^+ \rightarrow K^+ \pi^- \pi^+$	1.8603 ± 0.0002	23.34 ± 0.62	0.105 ± 0.0054
$D^+ \rightarrow K^{*0} K^+$	1.8688 ± 0.0013	18.77 ± 1.15	0.009 ± 0.0007
$\Lambda_c^+ \rightarrow K^+ \pi^- P^+$	2.2779 ± 0.0009	22.81 ± 0.92	0.020 ± 0.0013

baryon) in the event was used to tag the initial flavor of the reconstructed B^0 meson. The tagging technique utilized information from identified leptons (muons and electrons) and reconstructed secondary vertices. For reconstructed $B_s^0 \rightarrow D_s^- \mu^+ X$ decays both leptons having the same sign would indicate that one B hadron had oscillated while opposite signs would indicate that neither (or both) had oscillated. The performance of the flavor tagging is characterized by the efficiency, $\epsilon = N_{\text{tag}}/N_{\text{tot}}$, where N_{tag} is the number of tagged B_s^0 mesons, and N_{tot} is the total number; the tag purity η_s , defined as $\eta_s = N_{\text{cor}}/N_{\text{tag}}$, where N_{cor} is the number of B_s^0 mesons with correct flavor identification; and dilution, related to purity as $\mathcal{D} = 2\eta_s - 1$. The details of the initial state flavor tagging is given in reference [11]. A short description is given here.

The flavor of the initial state of the signal B_s candidate was determined using a likelihood ratio method based on the properties of the other b hadron in the event (opposite-side tagging). For events with a reconstructed muon on the opposite side of the B_s candidate, where $\cos \phi(\mathbf{p}_\mu, \mathbf{p}_B) < 0.8$, the muon jet charge was used as a discriminating variable. The muon jet charge was defined as $Q_J^\mu = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}$, where the sum was taken over all charged particles on the opposite side, including the muon. All charged particles were required to be within a cone of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$ around the muon direction. For events without an identified muon but with a reconstructed electron on the opposite side, the electron jet charge $Q_J^e = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}$, defined similarly to the muon jet charge, was used. For events with neither a muon nor an electron, the event charge was used as a discriminating variable, defined as $Q_{EV} = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}$, where the

TABLE II: Parameters for background slope and signal fraction parameterization

Parameter	Value
P0	2.332 ± 0.279
p1	-0.987 ± 0.149
P2	0.781 ± 0.012
P3	-1.087 ± 0.006
P4	0.772 ± 0.003
P5	-0.270 ± 0.001
P6	0.037 ± 0.0005

sum was taken over all charged particles with $0.5 < p_T < 50 \text{ GeV}/c$ and having $\cos\phi(\mathbf{p}, \mathbf{p}_B) < 0.8$. The upper cut on p_T ensures that the event charge will not be determined by a single track with very high p_T , though no optimization of this cut has been made. Finally, in any event with a reconstructed secondary vertex, the secondary vertex charge was also used as a discriminating variable. The secondary vertex charge was defined as $Q_{SV} = \frac{\sum_i (q^i p_L^i)^{0.6}}{\sum_i (p_L^i)^{0.6}}$, where the sum was taken over all particles included in the secondary vertex, and p_L^i is the longitudinal momentum of a given particle with respect to the sum of all the momenta of the particles associated with the secondary vertex.

In the likelihood ratio method, a combined tagging variable y is constructed:

$$y = \prod_{i=1}^n y_i; \quad y_i = \frac{f_i^{\bar{b}}(x_i)}{f_i^b(x_i)}, \quad (18)$$

where the $f_i^b(x_i)$ are the probability density functions (*pdf*) for each discriminating variable described above. The *pdfs* were determined from real $B^+ \rightarrow \mu^+ \nu \bar{D}^0$ events in which the B flavor is given by the sign of the muon. For an oscillation analysis, it is more convenient to define the tagging variable as

$$d = \frac{1 - y}{1 + y}. \quad (19)$$

The parameter d varies between -1 and 1 . An event with $d > 0$ is tagged as b quark and that with $d < 0$ as a \bar{b} quark, with larger $|d|$ values corresponding to higher tagging purities.

An important property of opposite-side tagging is the independence of its performance on the type of the reconstructed B meson, since the hadronization of the two b quarks is not correlated in $p\bar{p}$ interactions. Therefore, the flavor tagging algorithm can be calibrated in data by applying it to the events with the B^0 and B^+ decays. The measured performance can then be used in the study of B_s meson oscillations. This tagging method was tested and calibrated extensively on both Monte Carlo and real $B \rightarrow \mu^+ \nu D^{*-}$ events. Fits to the asymmetry distribution, in various ranges of $|d|$ for these events show clear B_d oscillations with Δm_d values consistent with the world average value [3].

A subsample of all B candidates for which the partial reconstruction on the opposite side was available is called the “total tagged events” sample. B_d^0 mesons oscillate with low frequency while B^+ mesons do not oscillate. Therefore, the B_d^0 and B^+ samples are used to determine the number of “correctly tagged events” and, therefore, to calibrate the OST.

Each tagged B candidate is characterized by a variable d_{pr} , which gives a prediction of the dilution for that candidate using the formulas

$$\begin{aligned} \mathcal{D}(d_{pr}) \Big|_{d_{pr} < 0.6} &= 0.457 \cdot |d_{pr}| + 2.349 \cdot |d_{pr}|^2 - 2.498 \cdot |d_{pr}|^3, \\ \mathcal{D}(d_{pr}) \Big|_{d_{pr} > 0.6} &= 0.6. \end{aligned} \quad (20)$$

Another parameterization, $\mathcal{D}(d_{pr})$, was used to estimate the systematic uncertainty:

$$\mathcal{D}(d_{pr}) = \frac{0.6}{1 + \exp\left(-\frac{d_{pr} - 0.312}{0.108}\right)}. \quad (21)$$

Figure 6 shows the fit for the total tagged sample. We obtain a total of 2997 ± 146 tagged signal candidates, 8208 ± 145 tagged D^+ reflection candidates, 1261 ± 89 tagged Λ_c events and 732 Cabbibo suppressed candidates.

VI. EXPERIMENTAL OBSERVABLES

The proper lifetime of the B_s^0 meson, $c\tau_{B_s^0}$, for semileptonic decays can be written as

$$c\tau_{B_s^0} = x^M \cdot K, \quad \text{where } x^M = \left[\frac{\mathbf{d}_T^B \cdot \mathbf{p}_T^{\mu D_s^-}}{(p_T^{\mu D_s^-})^2} \right] \cdot cM_B. \quad (22)$$

x^M is the *visible proper decay length*, or VPDL, and K is the correction factor, also called the K factor. Semileptonic B decays necessarily have an undetected neutrino present in the decay chain, making a precise determination of the kinematics for the B meson impossible. In addition, other neutral or non-reconstructed charged particles can be present in the decay chain of the B meson. This leads to a bias towards smaller values of the B momentum,

which is calculated using the reconstructed particles. A common practice to correct this bias is to scale the measured momentum of the B candidate by a K factor, which takes into account the effects of the neutrino and other lost or non-reconstructed particles. For this analysis, the K factor was defined as

$$K = p_T(\mu^+ D_s^-) / p_T(B_s^0), \quad (23)$$

where p_T denotes the absolute value of the transverse momentum. The K -factor distributions used to correct the data were obtained from the Monte Carlo (MC) simulation.

VII. FITTING PROCEDURE

All tagged events with $1.79 < M(K^+ K^- \pi^-) < 2.25 \text{ GeV}/c^2$ were used in the unbinned likelihood fitting procedure. The likelihood for an event to arise from a specific source in the sample depends on the x^M , its uncertainty (σ_{x^M}), the mass of the D_s^- meson candidate (m), the predicted dilution (d_{pr}) and the selection variable y described in Section III. All of the quantities used in the unbinned likelihood fitting procedure are known on an event-by-event basis. The pdf for each source can be expressed by the product of the corresponding pdf s:

$$f_i = P_i^{x^M}(x^M, \sigma_{x^M}, d_{pr}) P_i^{\sigma_{x^M}} P_i^m P_i^{d_{pr}} P_i^y. \quad (24)$$

The VPDL pdf $P_i^{x^M}(x^M, \sigma_{x^M}, d_{pr})$ represents a conditional probability, therefore it should be multiplied by $P_i^{\sigma_{x^M}}$ and $P_i^{d_{pr}}$ to have a joint pdf (see ‘‘Probability’’ section in PDG [3]). The pdf s P_i^m and P_i^y are used for separation of signal and background.

The following sources, i , were considered:

- $\mu D_s (\rightarrow K^* K)$ signal with fraction $Fr_{\mu D_s}$.
- $\mu D^\pm (\rightarrow K^* K)$ signal with fraction $Fr_{\mu D^\pm}$.
- $\mu D^\pm (\rightarrow K \pi \pi)$ reflection with fraction $Fr_{\mu D_{refl}^\pm}$.
- $\mu \Lambda_c^\pm (\rightarrow K \pi P)$ reflection with fraction $Fr_{\mu \Lambda_{refl}^\pm}$.
- Combinatorial background with fraction $(1 - Fr_{\mu D_s} - Fr_{\mu D^\pm} - Fr_{\mu D_{refl}^\pm} - Fr_{\mu \Lambda_{refl}^\pm})$.

The fractions $Fr_{\mu D_s}$, $Fr_{\mu D^\pm}$, $Fr_{\mu D_{refl}^\pm}$ and $Fr_{\mu \Lambda_{refl}^\pm}$ were determined from the mass fit (see Fig. 6). The total pdf for a B candidate has the form:

$$F_n = Fr_{\mu D_s} f_{\mu D_s} + Fr_{\mu D^\pm} f_{\mu D^\pm} + Fr_{\mu D_{refl}^\pm} f_{\mu D_{refl}^\pm} + Fr_{\mu \Lambda_{refl}^\pm} f_{\mu \Lambda_{refl}^\pm} + \left(1 - Fr_{\mu D_s} - Fr_{\mu D^\pm} - Fr_{\mu D_{refl}^\pm} - Fr_{\mu \Lambda_{refl}^\pm}\right) f_{bkg} \quad (25)$$

The following form was minimized using the MINUIT [12] program:

$$\mathcal{L} = -2 \sum_n \ln F_n, \quad (26)$$

where n varies from 1 to $N_{total \text{ tagged events}}$.

The pdf s for the VPDL uncertainty ($P_i^{\sigma_{x^M}}$), mass (P_i^m), dilution ($P_i^{d_{pr}}$), and selection variable y (P_i^y) were taken from experimental data. The signal pdf s were also used for the $\mu^+ D^- (\rightarrow K^+ \pi^- \pi^-)$ reflection.

A. pdf for $\mu^+ D_s^-$ Signal

The $\mu^+ D_s^-$ sample is composed mostly of B_s^0 mesons with some contributions from B^+ and B_d^0 mesons. Different species of B mesons behave differently with respect to oscillations. Neutral B_d^0 and B_s^0 mesons do oscillate (with different frequencies) while charged B^+ mesons do not.

The data sample is divided into non-oscillated and oscillated subsamples as determined by the flavor tagging. For a given type of B_q hadron, where q is d , u , or s , or b-baryon, the distribution of the visible proper decay length x for non-oscillated and oscillated cases (p^{nos} and p^{osc}) is given by:

$$p_s^{nos}(x, K, d_{pr}) = \frac{K}{c\tau_{B_s^0}} \exp\left(-\frac{Kx}{c\tau_{B_s^0}}\right) \cdot 0.5 \cdot (1 + \mathcal{D}(d_{pr}) \cos(\Delta m_s \cdot Kx/c)) \quad (27)$$

$$p_s^{osc}(x, K, d_{pr}) = \frac{K}{c\tau_{B_s^0}} \exp\left(-\frac{Kx}{c\tau_{B_s^0}}\right) \cdot 0.5 \cdot (1 - \mathcal{D}(d_{pr}) \cos(\Delta m_s \cdot Kx/c)) \quad (28)$$

$$p_{D_s D_s}^{osc}(x, K) = \frac{K}{c\tau_{B_s^0}} \exp\left(-\frac{Kx}{c\tau_{B_s^0}}\right) \cdot 0.5 \quad (29)$$

$$p_{D_s D_s}^{nos}(x, K) = \frac{K}{c\tau_{B_s^0}} \exp\left(-\frac{Kx}{c\tau_{B_s^0}}\right) \cdot 0.5 \quad (30)$$

$$p_u^{nos}(x, K, d_{pr}) = \frac{K}{c\tau_{B^+}} \exp\left(-\frac{Kx}{c\tau_{B^+}}\right) \cdot 0.5 \cdot (1 - \mathcal{D}(d_{pr})) \quad (31)$$

$$p_u^{osc}(x, K, d_{pr}) = \frac{K}{c\tau_{B^+}} \exp\left(-\frac{Kx}{c\tau_{B^+}}\right) \cdot 0.5 \cdot (1 + \mathcal{D}(d_{pr})) \quad (32)$$

$$p_d^{nos}(x, K, d_{pr}) = \frac{K}{c\tau_{B_d^0}} \exp\left(-\frac{Kx}{c\tau_{B_d^0}}\right) \cdot 0.5 \cdot (1 - \mathcal{D}(d_{pr}) \cos(\Delta m_d \cdot Kx/c)) \quad (33)$$

$$p_d^{osc}(x, K, d_{pr}) = \frac{K}{c\tau_{B_d^0}} \exp\left(-\frac{Kx}{c\tau_{B_d^0}}\right) \cdot 0.5 \cdot (1 + \mathcal{D}(d_{pr}) \cos(\Delta m_d \cdot Kx/c)). \quad (34)$$

$$p_\Lambda^{nos}(x, K, d_{pr}) = \frac{K}{c\tau_\Lambda} \exp\left(-\frac{Kx}{c\tau_\Lambda}\right) \cdot 0.5 \cdot (1 - \mathcal{D}(d_{pr})) \quad (35)$$

$$p_\Lambda^{osc}(x, K, d_{pr}) = \frac{K}{c\tau_\Lambda} \exp\left(-\frac{Kx}{c\tau_\Lambda}\right) \cdot 0.5 \cdot (1 + \mathcal{D}(d_{pr})) \quad (36)$$

Here τ_{B_q} is the lifetime of the B_q hadron, where q is u , d , or s and τ_Λ is the lifetime of b-baryon. Note that there is a sign swap in Eqs. 31–34 with respect to Eqs. 27 and 28 due to anti-correlation of charge for muons from $B \rightarrow DD_s^-$; $D \rightarrow \mu^+ X$ processes.

The translation from real VPDL, x , to the measured VPDL, x^M , is achieved by a convolution of the K factors and resolution functions as specified below.

$$P_j^{osc, nos}(x^M, \sigma_{x^M}, d_{pr}) = \int_{K_{min}}^{K_{max}} dK D_j(K) \cdot \frac{Eff_j(x^M)}{N_j(K, \sigma_{x^M}, d_{pr})} \int_0^\infty dx G(x - x^M, \sigma_{x^M}) \cdot p_j^{osc, nos}(x, K, d_{pr}). \quad (37)$$

Here

$$G(x - x^M, \sigma_{x^M}) = \frac{1}{\sqrt{2\pi}\sigma_{x^M}} \exp\left(-\frac{(x - x^M)^2}{2\sigma_{x^M}^2}\right) \quad (38)$$

is the detector resolution of the VPDL and $Eff_j(x)$ is the reconstruction efficiency for a given decay channel j of this type of B meson as a function of VPDL. The function $D_j(K)$ gives the normalized distribution of the K factor in a given channel j . The normalization factor N_j is calculated by integration over the entire VPDL region:

$$N_j(K, \sigma_{x^M}, d_{pr}) = \int_{-\infty}^\infty dx^M Eff_j(x^M) \cdot \int_0^\infty dx G(x - x^M, \sigma_{x^M}) \cdot (p_j^{osc}(x, K, d_{pr}) + p_j^{nos}(x, K, d_{pr})) \quad (39)$$

The total VPDL *pdf* for the $\mu^+ D_s^-$ signal is a sum of all the contributions that yield the D_s^- mass peak:

$$P_{\mu D_s}^{osc, nos}(x^M, \sigma_{x^M}, d_{pr}) = (1 - \mathcal{F}_{c\bar{c}}) \sum_j Br_j \cdot P_j^{osc, nos}(x^M, \sigma_{x^M}, d_{pr}) + \mathcal{F}_{c\bar{c}} \cdot P_{c\bar{c}}^{osc, nos}(x^M) \quad (40)$$

Here the sum \sum_j is taken over all decay channels that yield a $\mu^+ D_s^-$ final state and Br_j is the branching rate of a given channel j . In addition to the long-lived $\mu^+ D_s^-$ candidates from B meson decays, there is a contribution, with fraction \mathcal{F}_{peak} , of the ‘‘peaking background’’, which consists of combinations of D_s^- mesons and muons originating from different c or b quarks. The direct c production gives the largest contribution to this background and, therefore, the function $P_{peak}^{osc, nos}(x^M)$ was determined from $c\bar{c}$ MC. We assume that this background produces negative and positive flavor tags with equal probability.

The choice of oscillated or non-oscillated VPDL *pdf* for Eq. 24 is made using relative charge of the muon from the B_s^0 meson with respect to the sign of d_{pr} :

$$\begin{aligned} d_{pr} \cdot q_\mu > 0 : P^{x^M}(x^M, \sigma_{x^M}, d_{pr}) &= P_{\mu D_s}^{osc}(x^M, \sigma_{x^M}, d_{pr}), \\ d_{pr} \cdot q_\mu < 0 : P^{x^M}(x^M, \sigma_{x^M}, d_{pr}) &= P_{\mu D_s}^{nos}(x^M, \sigma_{x^M}, d_{pr}). \end{aligned} \quad (41)$$

The branching rates Br_j were taken from the PDG [3]. The functions $D_j(K)$ and $Eff_j(x)$ were taken from the MC simulation, as explained later. The lifetimes of the B^+ and B_d^0 mesons were taken from PDG while the B_s^0 lifetime was measured using the total tagged $\mu^+ D_s^-$ sample.

B. *pdf* for μD^\pm Signal

The μD^\pm Cabbibo suppressed signal which forms a small peak on the left of the signal peak, See section IV, was also considered in the final fit. This peak is mainly due to decays from B_d and has been modeled with the *pdf*

$$p_{B_d}^{osc/nos}(x^M, \sigma_{x^M}, d_{pr}) = \frac{Eff(x^M)}{N} \int_0^\infty dx G(x - x^M, s_{B_d} \sigma_{x^M}) \cdot \exp\left(-\frac{Kx}{c\tau_{B_d}}\right) \cdot (1 \pm \mathcal{D}(d_{pr}) \cos(\Delta m_d \cdot Kx/c)) \quad (42)$$

where, τ_{B_d} is the lifetime of the B_d meson, $G(x - x^M, s_{B_d} \sigma_{x^M})$ is the detector resolution defined in equation 38 and $Eff(x^M)$ is the reconstruction efficiency for $B_d^0 \rightarrow \mu D^\pm X$ decay as a function of VPDL.

C. *pdf* for Combinatorial Background

The following contributions into the combinatorial background were considered:

1. Prompt background with the μD_s vertex coinciding with the PV (described as a Gaussian with width determined by resolution; fraction in the background: \mathcal{F}_0).
2. Background with quasi-vertices distributed around the PV (described as a Gaussian with constant width $\sigma_{c\bar{c}}$; fraction in the background: $\mathcal{F}_{c\bar{c}}$).
3. Long-lived background (exponential with constant decay length $c\tau_{long}$ convoluted with resolution; Fraction in the background: \mathcal{F}_{long}).
4. A negative exponential to take into account the B outliers at negative tail. Fraction in the background: \mathcal{F}_{NegExp} with constant decay length $c\tau_{NegExp}$
5. A positive exponential convoluted with resolution with constant decay length $c\tau_{PosExp}$ to take into account the outliers on positive tail.

The Long-lived background was further divided into three subsamples:

1. insensitive to the tagging (fraction in the long-lived background: $(1 - \mathcal{F}_{tsens})$);
2. sensitive to the tagging and non-oscillating (fraction in the background sensitive to the tagging: $(1 - \mathcal{F}_{osc})$);

3. sensitive to the tagging and oscillating with frequency Δm_d (fraction in the background sensitive to the tagging: \mathcal{F}_{osc}).

The fractions of these contributions and their parameters were determined from the data sample. It was expected that the combinatorial background had a constant mixed/unmixed asymmetry d_{bkg} . The background *pdf* was expressed in the following form:

$$P_{bkg}(x^M, \sigma_{x^M}, d_{pr}) = \mathcal{F}_{c\bar{c}} G(0 - x^M, \sigma_{c\bar{c}}) + \mathcal{F}_{NegExp} \cdot \frac{-1}{c\tau_{NegExp}} \exp\left(-\frac{x}{c\tau_{NegExp}}\right) + (1 - \mathcal{F}_{c\bar{c}} - \mathcal{F}_{NegExp}) \cdot P_{bkg}^{res}(x^M, \sigma_{x^M}), \quad (43)$$

$$P_{bkg}^{res}(x^M, \sigma_{x^M}, d_{pr}) = \frac{Eff(x^M)}{N} \int_0^\infty dx (\mathcal{F}_0 G(x - x^M, s_{bkg} \sigma_{x^M}) \delta(x) + (1 - \mathcal{F}_0) \cdot G(x - x^M, \sigma_{x^M}) \left(\mathcal{F}_{long} \cdot p_{bkg}^{long} + (1 - \mathcal{F}_{long}) \cdot \frac{1}{c\tau_{PosExp}} \exp\left(-\frac{x}{c\tau_{PosExp}}\right) \right)), \quad (44)$$

$$p_{bkg}^{long,osc/nos}(x, d_{pr}) = \frac{1}{c\tau_{long}} \exp\left(-\frac{x}{c\tau_{long}}\right) ((1 - \mathcal{F}_{tsens}) + \mathcal{F}_{tsens} ((1 \pm \mathcal{D})(1 - \mathcal{F}_{osc}) + (1 \pm \mathcal{D} \cos(\Delta m_d \cdot x/c)) \cdot \mathcal{F}_{osc}), \quad (45)$$

where N is the normalization constant and the fit parameters are \mathcal{F}_{peak_bkg} , σ_{peak_bkg} , \mathcal{F}_0 , \mathcal{F}_{tsens} , \mathcal{F}_{long} , \mathcal{F}_{NegExp} , \mathcal{F}_{osc} , $c\tau_{NegExp}$, $c\tau_{PosExp}$ and $c\tau_{long}$. As an efficiency $Eff(x^M)$, the efficiency for the $B_d^0 \rightarrow D^- \mu^+ \nu X$ channel was used.

VIII. FIT INPUTS

We have used the following measured parameters for B mesons from the PDG [3] as inputs for the lifetime fitting procedure: $c\tau_{B^+} = 501 \mu\text{m}$, $c\tau_{B_d^0} = 460 \mu\text{m}$, and $\Delta m_d = 0.502 \text{ ps}^{-1}$. The latest PDG values were also used to determine the branching fractions of decays contributing to the D_s^- sample. We used the event generator EvtGen [13] since this code was developed specifically for the simulation of B decays. For those branching fractions not given in the PDG, we used the values provided by EvtGen, which are motivated by theoretical considerations. Taking into account the corresponding branching rates and reconstruction efficiencies, we calculated the contributions to our signal region from the various processes. The $B_s^0 \rightarrow D_s^- \mu^+ \nu X$ modes (including through D_s^{*-} , D_{s0}^{*-} , and $D_{s1}^{\prime-}$ and μ^+ originating from τ decays) comprise $82.4 \pm 3.3\%$ of our sample, including reconstruction efficiency. Other backgrounds with both a real D_s^- and μ^+ and showing up in the peak, but not expected to oscillate with Δm_s , that are considered are $B \rightarrow D_{(s)}^+ D_s^- X$ decays followed by $D_{(s)}^+ \rightarrow \mu^+ \nu X$. The assigned uncertainty to each channel covers possible trigger efficiency biases. We then determined the efficiency of the lifetime selections for the sample as a function of VPDL, as shown in Fig. 7 for the decay $B_s^0 \rightarrow D_s^- \mu^+ \nu X$.

In determining the K factor distributions, MC generator-level information was used for the computation of p_T . Following the definition used in Eq. 23, the K factor distributions for all considered decays were determined. Figure 8 shows the distributions of the K factors for the semi-muonic decays of the B_s^0 meson. As expected, the K factors for D_s^{*-} , D_{s0}^{*-} and $D_{s1}^{\prime-}$ have lower mean values because more decay products are lost. Note that since the K factors in Eq. 23 were defined as the ratio of transverse momenta, they can exceed unity.

The VPDL uncertainty was estimated by the vertex fitting procedure. A resolution scale factor was introduced to take into account a possible bias. It was determined using a J/ψ sample. Figure 9 shows the pull distribution, $PDL_{J/\psi}/\sigma(PDL_{J/\psi})$, of the J/ψ vertex position with respect to that of the primary vertex, where PDL is the *proper decay length*. The negative tail of the pull distribution of the J/ψ vertex position with respect to that of the primary vertex should be a Gaussian with a sigma of unity if uncertainties assigned to the vertex coordinates are correct. We ignore the positive side of the pull distribution as that tends to be biased towards larger values due to J/ψ mesons from real B meson decays. For this study we exclude muons from J/ψ decays from the primary vertex. The resulting pull distribution was fitted using a double Gaussian: the narrow Gaussian with width $\sigma_{narrow} = 0.998$ comprises 72% of the events, and the wide Gaussian with width $\sigma_{wide} = 1.775$ comprises 28%.

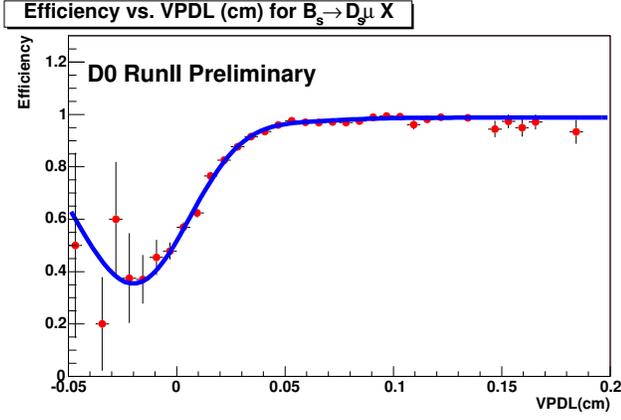


FIG. 7: Efficiency of the lifetime-dependent cuts as a function of VPDL for $B_s^- \rightarrow D_s^- \mu^+ \nu X$.

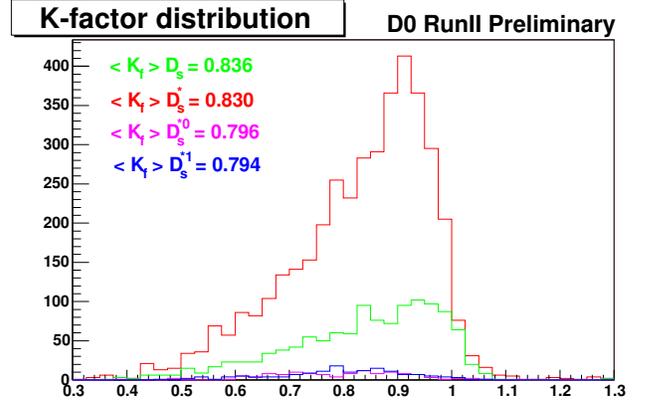


FIG. 8: K factor distributions for $B_s^0 \rightarrow \mu^+ \nu D_s^-$; $B_s^0 \rightarrow \mu^+ \nu D_s^{*-}$; $B_s^0 \rightarrow \mu^+ \nu D_s^-$; $B_s^0 \rightarrow \mu^+ \nu D_{s0}^{*-}$; $B_s^0 \rightarrow \mu^+ \nu D_s^-$; $B_s^0 \rightarrow \mu^+ \nu D_{s1}^-$ processes.

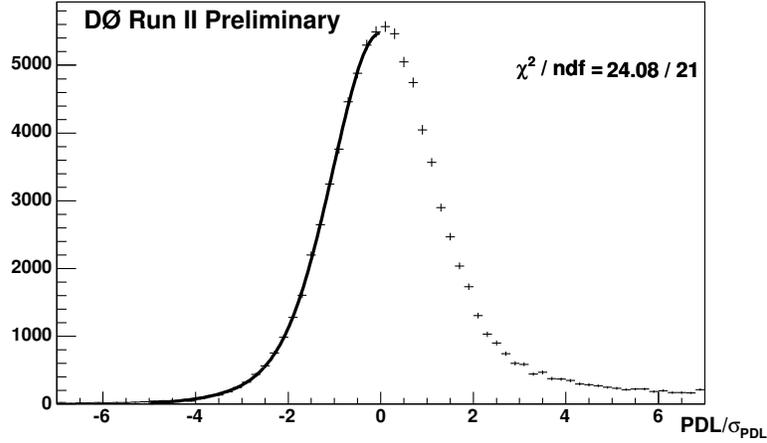


FIG. 9: Pull distribution of the J/ψ vertex position with respect to that of the primary vertex: $PDL_{J/\psi}/\sigma(PDL_{J/\psi})$.

It is known that the scale factor depends on track transverse momenta. We took this dependence into account using the scale factor determined for J/ψ candidates where the leading muon has $p_{t\mu} > 6$ GeV/ c to estimate a contribution to the systematic uncertainty. The corresponding scale factors increased by 2.5%.

The total tagged data sample was used to determine the parameters: $Fr_{quasi} = 0.023 \pm 0.003$, $\sigma_{quasi} = 120 \pm 9$ μm , $Fr_0 = 0.067 \pm 0.002$, $c\tau_{bkg} = 549 \pm 10$ μm , $Fr_{long} = 0.91 \pm 0.013$, $Fr_{Neg} = 0.062 \pm 0.002$, $Fr_{B_d} = 0.51 \pm 0.085$, $Fr_{Mix} = 0.66 \pm 0.096$, $c\tau_{long} = 662 \pm 9$ μm , $c\tau_{Neg} = -53 \pm 2$ μm , $sf_{bkg} = 1.91 \pm 0.034$ and $c\tau_{B_s} = 407 \pm 22$ μm . The discrepancy of this fitted value of $c\tau_{B_s^0}$ from the world average value was included as a systematic uncertainty.

IX. AMPLITUDE FIT METHOD

The amplitude fit method [14] is a technique that can be used to calculate an experimental Δm_s oscillation limit. This technique requires a modification of Eqs. 27 and 28, yielding the form

$$p_s^{nos/osc}(x, K, d_{pr}) = \frac{K}{c\tau_{B_s^0}} \exp\left(-\frac{Kx}{c\tau_{B_s^0}}\right) \cdot 0.5 \cdot (1 \pm \mathcal{D}(d_{pr}) \cos(\Delta m_s \cdot Kx/c) \cdot \mathcal{A}), \quad (46)$$

where \mathcal{A} is now the only fit parameter.

The values of Δm_s were changed from 0.5 ps^{-1} to 25 ps^{-1} with a step size of 0.5 ps^{-1} . By plotting the fitted value of \mathcal{A} as a function of the input value of Δm_s , one searches for a peak of $\mathcal{A} = 1$ to obtain a measurement of Δm_s . For any value of Δm_s not equal to the “true” value of B_s^0 oscillation frequency, the amplitude \mathcal{A} should be zero. If no peak is found, limits can be set on Δm_s using this method. The expected limit (i.e., sensitivity) of a measurement is determined by calculating the probability that at a non-“true” value of Δm_s the amplitude could fluctuate to $\mathcal{A}=1$. This occurs at the lowest value of Δm_s for which $1.645 \sigma_{\mathcal{A}} = 1$ for a 95% CL, where $\sigma_{\mathcal{A}}$ is the uncertainty on the value of \mathcal{A} at the point Δm_s . The limit is determined by calculating the probability that a fitted value of \mathcal{A} could fluctuate to $\mathcal{A} = 1$. This occurs at the lowest value of Δm_s for which $\mathcal{A} + 1.645\sigma_{\mathcal{A}} = 1$.

Figure 10 shows the dependence of the parameter \mathcal{A} from Eq. 46, and its uncertainty, on Δm_s . In the figure, the yellow (light shaded) and green (dark shaded) regions indicate 1.645 times the statistical uncertainty and 1.645 times the statistical plus systematic uncertainties, respectively. A 95% CL limit on the B_s^0 oscillation frequency $\Delta m_s > 9.3 \text{ ps}^{-1}$ and expected limit 11.7 ps^{-1} were obtained with statistical uncertainties only.

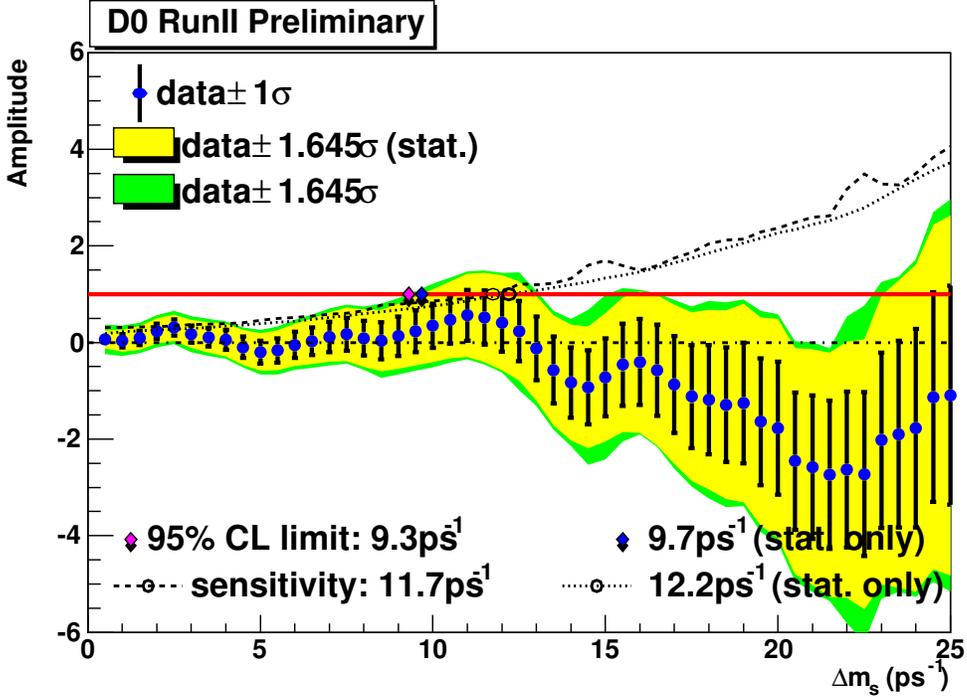


FIG. 10: B_s^0 oscillation amplitude with statistical and systematic errors. The red (solid) line shows the $\mathcal{A} = 1$ axis for reference. The dashed line shows the sensitivity including both statistical and systematic uncertainties.

X. SYSTEMATIC UNCERTAINTIES AND CROSS-CHECKS

All studied contributions to the systematic uncertainty of the amplitude are listed in Table III. For each Δm_s step, the deviations of $\Delta\mathcal{A}$ and $\Delta\sigma_{\mathcal{A}}$ from the default values are given. One can see that the largest deviations come from the uncertainty in the resolution. The resulting systematic uncertainties were obtained using the formula from Ref. [14]

$$\sigma_{\mathcal{A}}^{sys} = \Delta\mathcal{A} + (1 - \mathcal{A}) \frac{\Delta\sigma_{\mathcal{A}}}{\sigma_{\mathcal{A}}} \quad (47)$$

and were summed in quadrature. The effect of the systematic uncertainties is represented by the green (dark shaded) region in Fig. 10. Taking into account the systematic uncertainties, we obtained a 95% confidence level limit on the oscillation frequency $\Delta m_s > 9.3 \text{ ps}^{-1}$ and a expected limit of 11.7 ps^{-1} .

The decays $B_d^0 \rightarrow X\mu^+D^- (\rightarrow K\pi^-\pi^-)$ (the reflection component in Fig. 6) allow for a cross-check of the entire fitting procedure using B_d^0 meson decays present in the same data sample as the signal $B_s^0 \rightarrow D_s^-\mu^+\nu X$, $D_s^- \rightarrow \phi\pi^-$ events. Figure 11 shows the dependence of the parameter \mathcal{A} and its uncertainty on the B_d^0 oscillation frequency, Δm_d ,

TABLE III: Systematic uncertainties on the amplitude. The shifts of both the measured amplitude, $\Delta\mathcal{A}$, and its statistical uncertainty, $\Delta\sigma$, are listed

Osc. frequency		1 ps ⁻¹	2 ps ⁻¹	3 ps ⁻¹	4 ps ⁻¹	5 ps ⁻¹	6 ps ⁻¹	7 ps ⁻¹	8 ps ⁻¹	9 ps ⁻¹	10 ps ⁻¹	11 ps ⁻¹	12 ps ⁻¹	13 ps ⁻¹
\mathcal{A}		0.035	0.235	0.174	0.053	-0.199	-0.049	0.115	0.089	0.129	0.346	0.563	0.409	-0.120
Stat. uncertainty		0.137	0.158	0.176	0.207	0.236	0.268	0.313	0.362	0.411	0.456	0.525	0.595	0.658
$PDG\ c\tau_{B_s}$	$\Delta\mathcal{A}$	-0.020	-0.006	-0.004	+0.002	-0.011	+0.003	+0.011	+0.003	+0.007	+0.012	+0.027	+0.021	+0.002
	$\Delta\sigma$	+0.001	+0.002	+0.002	+0.004	+0.005	+0.006	+0.008	+0.009	+0.011	+0.013	+0.016	+0.019	+0.021
Signal SF variation by 3.5%	$\Delta\mathcal{A}$	+0.002	+0.003	+0.001	+0.002	-0.003	-0.001	+0.005	+0.009	+0.017	+0.031	+0.036	+0.030	+0.015
	$\Delta\sigma$	+0.000	+0.001	+0.001	+0.002	+0.003	+0.005	+0.006	+0.008	+0.010	+0.012	+0.015	+0.018	+0.022
D_{Ref}^+ fraction + 1σ	$\Delta\mathcal{A}$	-0.011	+0.006	+0.007	+0.001	-0.006	-0.006	+0.000	+0.005	+0.004	+0.008	+0.017	+0.016	+0.003
	$\Delta\sigma$	+0.002	+0.002	+0.003	+0.003	+0.004	+0.004	+0.005	+0.005	+0.006	+0.006	+0.008	+0.009	+0.010
D_{Ref}^+ fraction - 1σ	$\Delta\mathcal{A}$	+0.037	-0.020	-0.023	-0.007	+0.015	+0.010	-0.009	-0.019	-0.017	-0.028	-0.052	-0.046	-0.012
	$\Delta\sigma$	-0.007	-0.008	-0.009	-0.010	-0.012	-0.013	-0.015	-0.018	-0.020	-0.022	-0.025	-0.030	-0.032
D_s signal - 1σ	$\Delta\mathcal{A}$	-0.018	+0.011	+0.011	+0.005	-0.003	+0.003	+0.013	+0.014	+0.010	+0.013	+0.020	+0.021	+0.009
	$\Delta\sigma$	+0.004	+0.004	+0.004	+0.005	+0.006	+0.007	+0.008	+0.010	+0.011	+0.012	+0.015	+0.017	+0.019
Signal $c\bar{c}$ fraction changed to 6%	$\Delta\mathcal{A}$	+0.024	+0.039	+0.032	+0.026	+0.008	+0.015	+0.029	+0.028	+0.029	+0.043	+0.064	+0.055	+0.002
	$\Delta\sigma$	+0.004	+0.004	+0.005	+0.006	+0.008	+0.009	+0.012	+0.015	+0.018	+0.021	+0.027	+0.033	+0.037
$Br(D_s D_s) = 4.7\%$	$\Delta\mathcal{A}$	+0.001	-0.003	-0.002	-0.001	+0.002	+0.000	-0.002	-0.001	-0.002	-0.004	-0.006	-0.005	+0.001
	$\Delta\sigma$	-0.001	-0.002	-0.002	-0.002	-0.002	-0.003	-0.003	-0.004	-0.004	-0.005	-0.006	-0.006	-0.007
K-factor decreased by 2%	$\Delta\mathcal{A}$	+0.003	-0.012	+0.017	+0.021	+0.010	-0.027	-0.032	+0.034	-0.045	-0.028	-0.043	+0.037	+0.176
	$\Delta\sigma$	-0.001	-0.000	-0.001	-0.002	-0.003	-0.003	-0.006	-0.007	-0.008	-0.008	-0.017	-0.014	-0.015
$Br(\mu D_s) = 5.5\%$	$\Delta\mathcal{A}$	+0.022	+0.005	+0.004	-0.009	-0.021	-0.015	-0.005	-0.006	+0.001	+0.013	+0.026	+0.025	-0.005
	$\Delta\sigma$	+0.008	+0.009	+0.011	+0.012	+0.014	+0.016	+0.019	+0.022	+0.026	+0.028	+0.032	+0.037	+0.041
Oscillated and Mixed fraction varied in bkg	$\Delta\mathcal{A}$	+0.033	-0.000	-0.009	-0.012	-0.012	-0.011	-0.011	-0.012	-0.010	-0.009	-0.008	-0.006	-0.008
	$\Delta\sigma$	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000
$Br(\mu D_s) = 10.8\%$	$\Delta\mathcal{A}$	-0.011	-0.003	-0.003	+0.004	+0.011	+0.008	+0.002	+0.003	-0.001	-0.008	-0.015	-0.015	+0.003
	$\Delta\sigma$	-0.005	-0.005	-0.006	-0.007	-0.008	-0.009	-0.011	-0.013	-0.014	-0.016	-0.018	-0.021	-0.023
$Br(D_s D_s) = 23\%$	$\Delta\mathcal{A}$	-0.002	+0.007	+0.006	+0.003	-0.004	-0.001	+0.004	+0.003	+0.004	+0.010	+0.016	+0.012	-0.003
	$\Delta\sigma$	+0.004	+0.004	+0.005	+0.005	+0.006	+0.007	+0.008	+0.009	+0.011	+0.012	+0.014	+0.016	+0.017
K-factor increased by 2%	$\Delta\mathcal{A}$	+0.011	+0.011	+0.018	-0.042	+0.012	+0.023	+0.019	-0.052	+0.088	+0.149	-0.006	-0.158	-0.310
	$\Delta\sigma$	+0.002	+0.004	+0.008	+0.012	+0.017	+0.021	+0.028	+0.034	+0.039	+0.048	+0.055	+0.057	+0.067
Bkg SF changed to 2.0	$\Delta\mathcal{A}$	+0.001	+0.001	-0.000	-0.001	-0.001	-0.002	-0.003	-0.003	-0.004	-0.004	-0.004	-0.005	-0.004
	$\Delta\sigma$	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
Bkg $c\bar{c}$ changed to 10.23%	$\Delta\mathcal{A}$	+0.018	+0.009	+0.007	+0.004	-0.001	-0.005	-0.011	-0.014	-0.011	-0.001	+0.009	+0.001	-0.018
	$\Delta\sigma$	-0.000	-0.000	-0.001	-0.001	-0.000	-0.000	+0.000	+0.001	+0.002	+0.002	+0.003	+0.005	+0.005
$P_T(\mu) > 6\text{ GeV}$	$\Delta\mathcal{A}$	-0.014	-0.005	-0.005	+0.005	+0.014	+0.009	+0.001	+0.004	-0.002	-0.012	-0.021	-0.019	+0.007
	$\Delta\sigma$	-0.006	-0.007	-0.008	-0.009	-0.011	-0.012	-0.014	-0.017	-0.019	-0.022	-0.025	-0.028	-0.031
Generator level K factor	$\Delta\mathcal{A}$	+0.004	+0.011	+0.014	-0.019	-0.006	-0.002	+0.001	+0.023	+0.030	-0.018	-0.013	-0.010	+0.014
	$\Delta\sigma$	+0.001	+0.002	+0.003	+0.005	+0.006	+0.009	+0.011	+0.014	+0.017	+0.020	+0.023	+0.026	+0.032
Smoothed K factor	$\Delta\mathcal{A}$	+0.004	+0.013	+0.014	-0.016	-0.004	-0.001	+0.003	+0.028	+0.031	-0.020	-0.005	-0.012	-0.005
	$\Delta\sigma$	+0.001	+0.002	+0.004	+0.006	+0.008	+0.010	+0.012	+0.015	+0.018	+0.022	+0.025	+0.028	+0.034
Total syst.	σ_{tot}^{sys}	0.132	0.128	0.141	0.116	0.152	0.164	0.175	0.164	0.247	0.263	0.173	0.198	0.304
Total	σ_{tot}	0.191	0.203	0.226	0.237	0.280	0.314	0.359	0.398	0.479	0.526	0.552	0.627	0.725

TABLE IV: Systematic uncertainties on the amplitude. The shifts of both the measured amplitude, $\Delta\mathcal{A}$, and its statistical uncertainty, $\Delta\sigma$, are listed (cont'd)

Osc. frequency		14 ps ⁻¹	15 ps ⁻¹	16 ps ⁻¹	17 ps ⁻¹	18 ps ⁻¹	19 ps ⁻¹	20 ps ⁻¹	21 ps ⁻¹	22 ps ⁻¹	23 ps ⁻¹	24 ps ⁻¹	25 ps ⁻¹
\mathcal{A}		-0.828	-0.721	-0.405	-0.865	-1.185	-1.252	-1.777	-2.581	-2.634	-2.024	-1.774	-1.093
Stat. uncertainty		0.729	0.809	0.890	1.004	1.130	1.251	1.377	1.484	1.613	1.817	2.058	2.266
$PDG\ c\tau_{B_s}$	$\Delta\mathcal{A}$	-0.033	-0.025	+0.005	-0.004	-0.022	+0.007	-0.039	-0.094	-0.090	+0.001	-0.065	+0.011
	$\Delta\sigma$	+0.024	+0.028	+0.030	+0.036	+0.043	+0.047	+0.055	+0.059	+0.062	+0.070	+0.081	+0.087
Signal SF variation by 3.5%	$\Delta\mathcal{A}$	-0.019	-0.021	-0.030	-0.057	-0.072	-0.051	-0.063	-0.112	-0.164	-0.219	-0.273	-0.314
	$\Delta\sigma$	+0.026	+0.031	+0.036	+0.043	+0.052	+0.061	+0.071	+0.080	+0.091	+0.106	+0.124	+0.144
D_{Refl}^+ fraction + 1σ	$\Delta\mathcal{A}$	-0.012	-0.008	-0.001	+0.005	+0.018	+0.030	+0.024	+0.010	+0.018	+0.031	+0.046	+0.059
	$\Delta\sigma$	+0.011	+0.011	+0.012	+0.013	+0.016	+0.018	+0.019	+0.020	+0.021	+0.024	+0.027	+0.030
D_{Refl}^+ fraction - 1σ	$\Delta\mathcal{A}$	+0.035	+0.027	+0.002	-0.008	-0.033	-0.055	-0.027	+0.020	-0.007	-0.053	-0.091	-0.129
	$\Delta\sigma$	-0.036	-0.039	-0.041	-0.047	-0.054	-0.061	-0.067	-0.070	-0.074	-0.083	-0.094	-0.105
D_s signal - 1σ	$\Delta\mathcal{A}$	-0.014	-0.013	+0.002	-0.005	-0.013	-0.019	-0.046	-0.083	-0.073	-0.046	-0.035	-0.012
	$\Delta\sigma$	+0.021	+0.023	+0.026	+0.030	+0.035	+0.039	+0.044	+0.047	+0.049	+0.054	+0.062	+0.067
Signal $c\bar{c}$ fraction changed to 6%	$\Delta\mathcal{A}$	-0.081	-0.074	-0.034	-0.098	-0.153	-0.168	-0.267	-0.424	-0.444	-0.355	-0.342	-0.213
	$\Delta\sigma$	+0.044	+0.051	+0.057	+0.071	+0.087	+0.101	+0.118	+0.133	+0.148	+0.176	+0.219	+0.245
$Br(D_s D_s) = 4.7\%$	$\Delta\mathcal{A}$	+0.009	+0.007	+0.004	+0.009	+0.012	+0.013	+0.020	+0.028	+0.029	+0.022	+0.019	+0.012
	$\Delta\sigma$	-0.008	-0.009	-0.010	-0.011	-0.012	-0.014	-0.015	-0.016	-0.017	-0.020	-0.022	-0.025
K-factor decreased by 2%	$\Delta\mathcal{A}$	+0.125	-0.151	-0.011	+0.201	+0.160	-0.017	+0.266	+0.158	+0.004	-0.508	-0.033	-0.547
	$\Delta\sigma$	-0.018	-0.023	-0.025	-0.036	-0.040	-0.043	-0.041	-0.038	-0.053	-0.095	-0.104	-0.081
$Br(\mu D_s) = 5.5\%$	$\Delta\mathcal{A}$	-0.050	-0.055	-0.040	-0.070	-0.087	-0.092	-0.110	-0.164	-0.176	-0.148	-0.135	-0.137
	$\Delta\sigma$	+0.046	+0.050	+0.054	+0.060	+0.067	+0.075	+0.083	+0.088	+0.096	+0.109	+0.122	+0.134
Oscillated and Mixed fraction varied in bkg	$\Delta\mathcal{A}$	-0.011	-0.012	-0.011	-0.010	-0.007	-0.004	+0.000	+0.001	-0.006	-0.012	-0.013	-0.014
	$\Delta\sigma$	+0.000	+0.000	-0.000	-0.001	-0.001	-0.002	-0.002	-0.002	-0.003	-0.002	-0.001	-0.001
$Br(\mu D_s) = 10.8\%$	$\Delta\mathcal{A}$	+0.029	+0.031	+0.022	+0.040	+0.049	+0.051	+0.062	+0.094	+0.101	+0.083	+0.074	+0.075
	$\Delta\sigma$	-0.026	-0.028	-0.031	-0.034	-0.038	-0.043	-0.047	-0.050	-0.055	-0.062	-0.069	-0.076
$Br(D_s D_s) = 23\%$	$\Delta\mathcal{A}$	-0.021	-0.018	-0.008	-0.021	-0.029	-0.032	-0.048	-0.070	-0.070	-0.054	-0.046	-0.028
	$\Delta\sigma$	+0.019	+0.021	+0.023	+0.027	+0.030	+0.033	+0.037	+0.039	+0.043	+0.048	+0.055	+0.060
K-factor increased by 2%	$\Delta\mathcal{A}$	-0.003	+0.412	-0.211	-0.235	-0.685	-0.352	-0.557	-0.126	+0.461	-0.023	+0.121	-0.344
	$\Delta\sigma$	+0.073	+0.085	+0.110	+0.119	+0.107	+0.143	+0.133	+0.143	+0.219	+0.235	+0.212	+0.226
Bkg SF chnged to 2.0	$\Delta\mathcal{A}$	-0.003	-0.004	-0.004	-0.004	-0.004	-0.004	-0.003	-0.001	-0.001	-0.003	-0.004	-0.007
	$\Delta\sigma$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002	-0.002	-0.003	-0.003	-0.003
Bkg $c\bar{c}$ changed to 10.23%	$\Delta\mathcal{A}$	-0.038	-0.023	-0.012	-0.031	-0.041	-0.040	-0.061	-0.099	-0.108	-0.094	-0.104	-0.079
	$\Delta\sigma$	+0.007	+0.008	+0.010	+0.012	+0.015	+0.017	+0.019	+0.020	+0.022	+0.027	+0.033	+0.036
$P_T(\mu) > 6\text{ GeV}$	$\Delta\mathcal{A}$	+0.042	+0.038	+0.027	+0.048	+0.070	+0.074	+0.084	+0.124	+0.135	+0.112	+0.103	+0.088
	$\Delta\sigma$	-0.035	-0.038	-0.042	-0.046	-0.052	-0.058	-0.064	-0.068	-0.074	-0.085	-0.095	-0.105
Generator level	$\Delta\mathcal{A}$	+0.085	-0.028	-0.160	+0.125	-0.132	-0.280	-0.144	+0.112	+0.086	-0.275	-0.095	-0.492
	$\Delta\sigma$	+0.033	+0.037	+0.045	+0.050	+0.050	+0.056	+0.060	+0.074	+0.089	+0.087	+0.092	+0.096
Smoothed K factor	$\Delta\mathcal{A}$	+0.076	-0.018	-0.132	+0.064	-0.088	-0.279	-0.163	+0.099	-0.002	-0.252	-0.182	-0.634
	$\Delta\sigma$	+0.036	+0.040	+0.049	+0.054	+0.055	+0.061	+0.066	+0.079	+0.093	+0.097	+0.102	+0.106
Total syst.	σ_{tot}^{syst}	0.335	0.644	0.181	0.341	0.520	0.359	0.413	0.524	1.054	0.820	0.526	0.981
Total	σ_{tot}	0.802	1.034	0.909	1.060	1.244	1.302	1.438	1.574	1.926	1.994	2.124	2.469

which has a world-average value of $\Delta m_d = 0.502 \pm 0.007 \text{ ps}^{-1}$ [3]. The peak in the amplitude scan at $\Delta m_d \approx 0.5 \text{ ps}^{-1}$ reveals the oscillations in the $B_d^0 - \bar{B}_d^0$ system. The peak amplitude is in good agreement with unity, which confirms that the dilution calibration is correct.

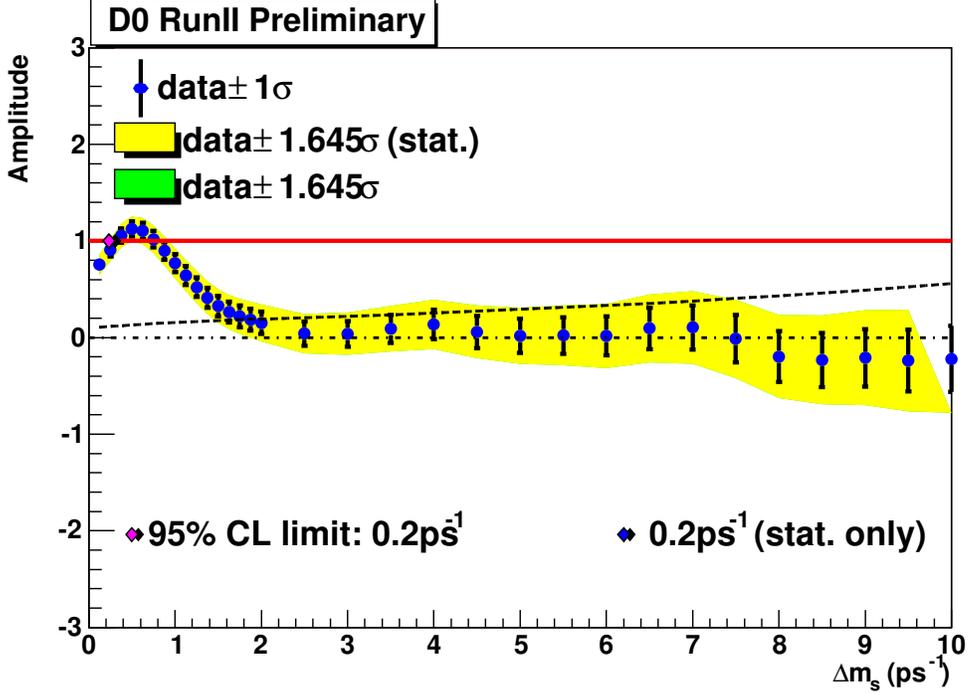


FIG. 11: B_d^0 oscillation amplitude with statistical uncertainty only for events in D^- reflection component. The red (solid) line shows the $\mathcal{A} = 1$ axis for reference. The dashed line shows the expected limit including statistical uncertainties only.

XI. CONCLUSIONS

Using $B_s^0 \rightarrow D_s^- \mu^+ \nu X$ decays, where $D_s^- \rightarrow K^{*0} K^-$, $K^{*0} \rightarrow K^+ K^-$, in combination with an opposite-side flavor tagging algorithm and an unbinned fit, we performed a search for $B_s^0 - \bar{B}_s^0$ oscillations. A 95% confidence level limit on the oscillation frequency $\Delta m_s > 9.3 \text{ ps}^{-1}$ and sensitivity 11.7 ps^{-1} were obtained.

Acknowledgments

Thanks to Prof. D.S. Kulshrestha for his continuous guidance and useful suggestions. We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); PPARC (United Kingdom); MSMT (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); Research Corporation, Alexander von Humboldt Foundation, and the Marie Curie Program.

-
- [1] C.Albajar *et al.* (UA1 Collaboration), “Search for B^0 - \bar{B}^0 Oscillations at the CERN P- \bar{P} Collider,” Phys. Lett. **B186**, 247 (1987).
[2] H.Albrecht *et al.* (ARGUS Collaboration), “Observation of B^0 - \bar{B}^0 Mixing,” Phys. Lett. **B192**, 245 (1987).
[3] S. Eidelman *et al.*, Phys. Lett. **B592**, 1 (2004) .

- [4] V. Abazov *et al.* [DØ Collaboration] First Direct Two-Sided Bound on the B_s^0 Oscillation Frequency, Accepted in Physical Review Letters, arXiv:hep-ex/0603029.
- [5] [CDF Collaboration] Measurement of B_s^0 - Anti- B_s^0 Oscillation Frequency, Submitted in Physical Review Letters, arXiv:hep-ex/0606027.
- [6] CKM Fitter Group, http://www.slac.stanford.edu/xorg/ckmfitter/ckm_welcome.html.
- [7] V. Abazov *et al.* [DØ Collaboration] *The Upgraded DØ Detector*, submitted to Nucl. Instrum. Methods Phys. Res. A., arXiv:hep-physics/0507191, Fermilab-Pub-05/341-E.
- [8] V. M. Abazov *et al.*, Nucl. Instrum. Meth. A **552**, 372 (2005).
- [9] S. Catani, Yu.L. Dokshitzer, M. Olsson, G. Turnock, B.R. Webber, Phys. Lett. **B269** (1991) 432.
- [10] http://www-d0.fnal.gov/computing/algorithms/muon/muon_algo.html.
- [11] G. Borissov *et al.*, *D0 Note 4991*, Bd mixing measurement using Opposite-side Flavor Tagging.
- [12] F. James, *MINUIT - Function Minimization and Error Analysis*, CERN Program Library Long Writeup D506, 1998.
- [13] D.J. Lange, Nucl. Instrum. Meth. A **462**, 152 (2001); for details see <http://www.slac.stanford.edu/~lange/EvtGen>.
- [14] H.G. Moser, A. Roussarie, Nucl.Instrum.Meth.A **384**, 491 (1997).
- [15] in the plane perpendicular to the beam direction.
- [16] parallel to the beam direction.