



Bd mixing measurement using Opposite-side Flavor Tagging

The DØ Collaboration

(Dated: March 3, 2006)

This note describes the measurement of the B_d mixing frequency and construction of an opposite-side flavor tagger at the DØ experiment. Various properties associated with the b quark opposite to the reconstructed B hadron were combined together into a single variable with an enhanced tagging power. Its performance was tested in data using a large sample of reconstructed semileptonic $B \rightarrow \mu \bar{D}^0 X$ events and $B \rightarrow \mu \bar{D}^* X$ events corresponding to an integrated luminosity $\sim 0.9 \text{ fb}^{-1}$. By dividing events into groups depending on the value of the combined tagging variable, performing an independent analysis in each group and combining results, the tagging power was found to be $\varepsilon \mathcal{D}^2 = (2.48 \pm 0.21_{-0.06}^{+0.08})(\%)$. The measured B_d mixing frequency $\Delta m_d = 0.506 \pm 0.020$ (stat) ± 0.016 (syst) ps^{-1} is in a good agreement with the world average value.

Preliminary Results for Winter 2006 Conferences

I. INTRODUCTION

A B_d mixing measurement is an important tool to calibrate our flavor tagger, and its measurement at a hadron collider could also reveal “new physics” in some SUSY scenarios (see for example [1]). Flavor tagging is an important ingredient of oscillation and CP-violation analyzes involving B mesons. Its performance is described by the combination of two quantities, efficiency and dilution. The efficiency ε is defined as the fraction of reconstructed events that are tagged:

$$\varepsilon = N_{tag}/N_{tot}. \quad (1)$$

Here N_{tag} is the number of tagged B mesons and N_{tot} is their total number. The tag purity η is defined as:

$$\eta = N_{cor}/N_{tag}, \quad (2)$$

where N_{cor} is the number of tagged B mesons with correct original flavor identification. The dilution \mathcal{D} is related with the purity η as

$$\mathcal{D} = 2\eta - 1, \quad (3)$$

The tagging power is given by $\varepsilon\mathcal{D}^2$.

This note describes the construction of the *combined flavor tagging* and the measurement of the B_d mixing frequency using $B \rightarrow \mu^+\nu\bar{D}^0$ and $B \rightarrow \mu^+\nu D^{*-}$ events collected by the DØ experiment in RunII. B^+ decays give the main contribution into the first sample, and B^0 decays dominate in the second sample. The flavor tagging purity of B^0 depends on its decay length due to the $B^0 - \bar{B}^0$ mixing, while the tagging purity of B^+ events remains constant. The B^0 oscillation frequency, given by the parameter Δm_d , is measured with high precision elsewhere [2]. Using this value, the flavor tagging purity can be extracted directly from data both for B^+ and B^0 decays. Alternatively, the value of Δm_d can be measured and compared with the world average [2] to test the flavor tagging for a possible lifetime-dependent bias.

II. DETECTOR DESCRIPTION AND EVENT SELECTION

The DØ detector, which was used to collect and analyze the semileptonic B meson decays, is described in [3]. This measurement exploits the large semileptonic data sample corresponding to approximately 0.9 fb^{-1} of integrated luminosity, accumulated by the DØ detector during the period from April 2002 to October 2005.

The decays $B \rightarrow \mu^+\nu\bar{D}^0 X$ with $\bar{D}^0 \rightarrow K^+\pi^-$ were selected using criteria described in [4]. The same criteria [4] were used to obtain two non-overlapping samples: the D^0 sample with the main contribution from $B^+ \rightarrow \mu^+\bar{D}^0 X$ decays and the D^* sample containing mainly $B^0 \rightarrow \mu^+ D^{*-} X$ decays.

All events with $0.1425 < \Delta M < 0.1490 \text{ GeV}/c^2$ were included in the D^* sample. The remaining events were included in the D^0 sample. The $K\pi$ mass distribution for these two samples together with the results of the fit are shown in Figs. 1 and 2. In total, 230551 ± 1627 $B \rightarrow \mu^+\nu\bar{D}^0$ decays and 73532 ± 304 $B \rightarrow \mu^+\nu\bar{D}^*$ decays were reconstructed.

The mass difference $\Delta M = M(\bar{D}^0\pi) - M(\bar{D}^0)$ for all events in D^* sample is shown in Fig. 3. A peak corresponding to the production of $\mu^+ D^{*-}$ events is clearly seen. The fit to this distribution using the sum of two Gaussians, describing the signal, and a background function is also shown. The total number of D^* candidates in the peak is 73532 ± 304 .

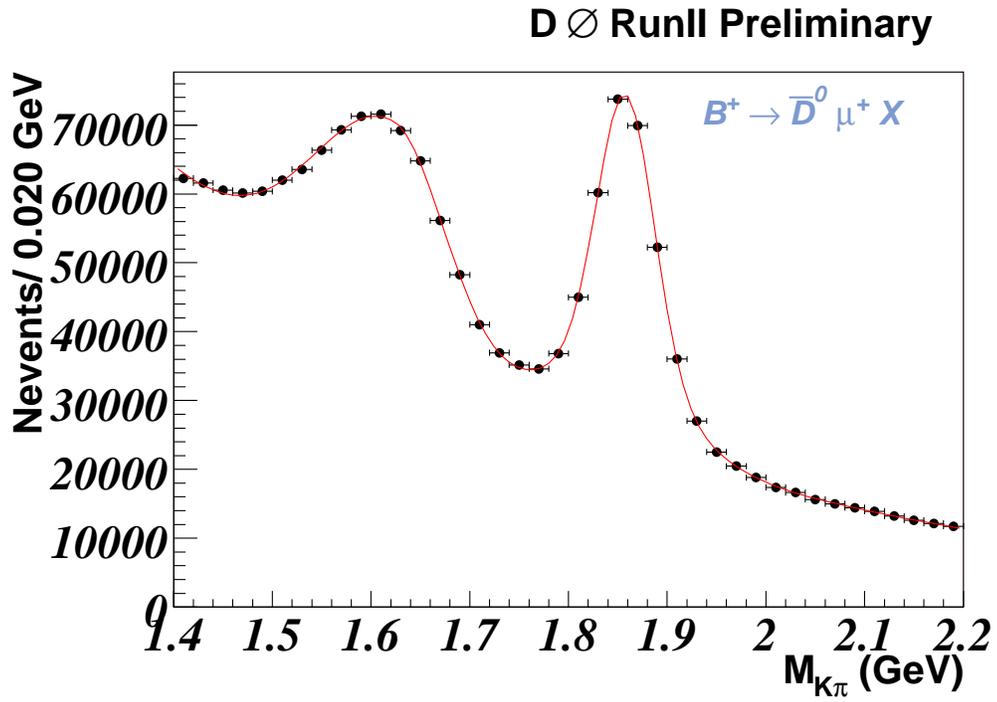


FIG. 1: The invariant mass of the $K\pi$ system for selected $\mu^+K^+\pi^-$ candidates.

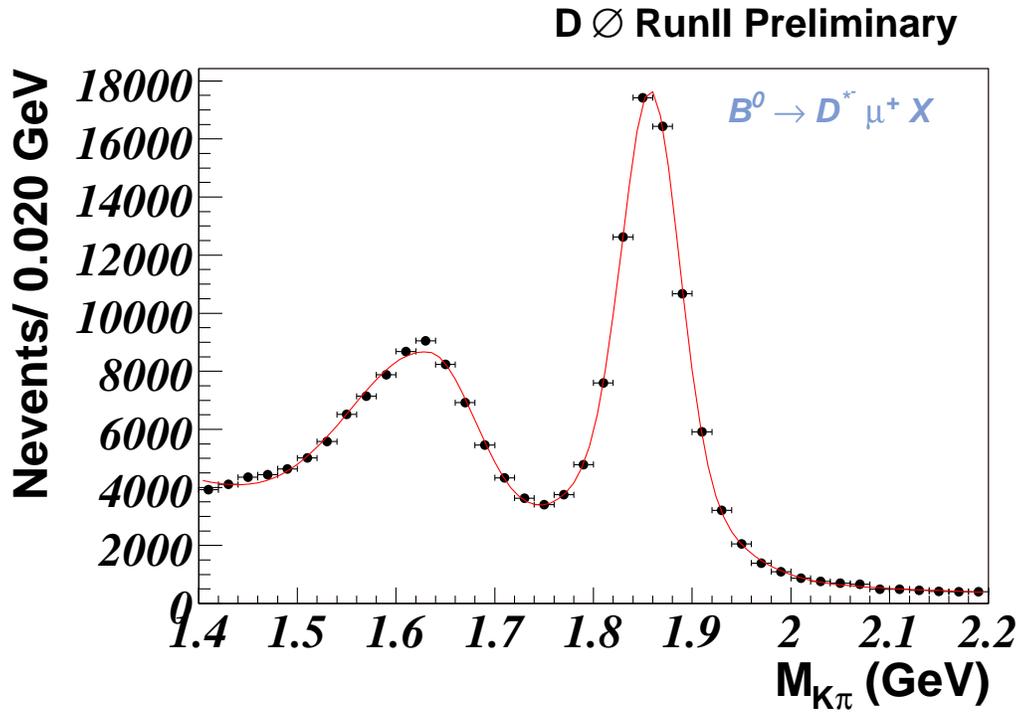


FIG. 2: The $K\pi$ invariant mass for selected μD^* candidates.

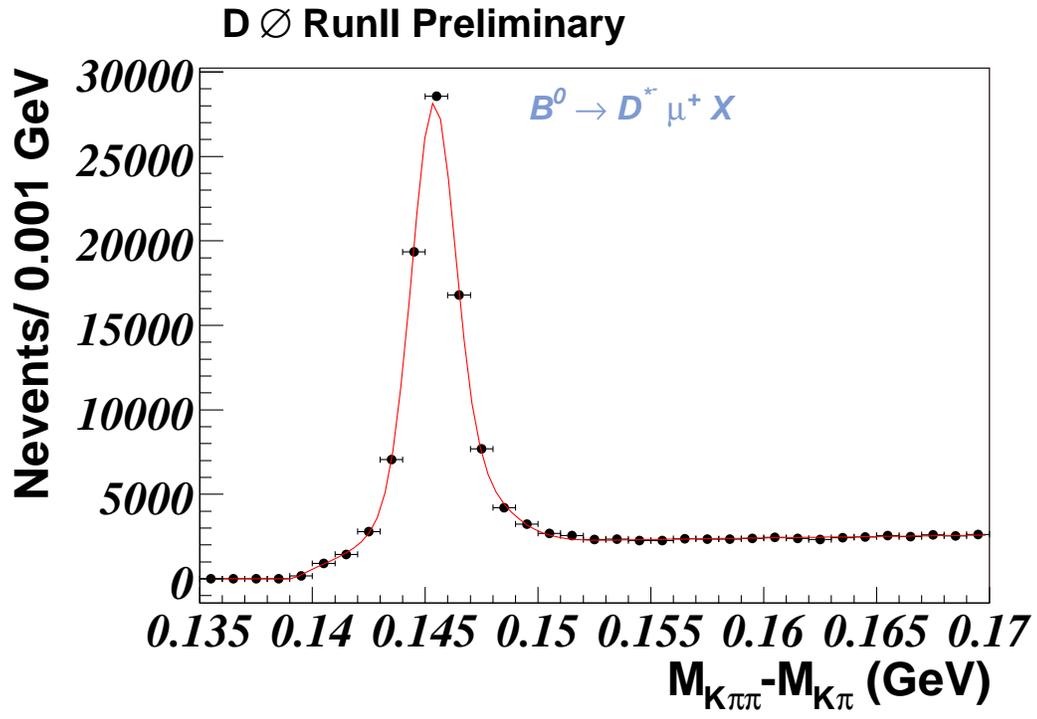


FIG. 3: The mass difference $M(D^0\pi) - M(D^0)$ for events with $1.75 < M(D^0) < 1.95$ GeV/ c^2 . Total number of D^* candidates is found to be 73532 ± 304 .

III. FLAVOR TAGGING METHOD

Many different properties can be used to identify the original flavor – b or \bar{b} – of a heavy quark producing a reconstructed B meson. Some of them perform well by themselves; other properties give a weak separation between flavors. In all cases, their combination into a single tagging variable gives a significantly better result. We obtain such a combination with 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. The discriminating variable, by definition, should have a different distribution for b and \bar{b} flavors. It can be either continuous, like the jet charge defined below, or discrete, like the charge of the electron from the side opposite to the reconstructed B meson. For the initial b quark, the probability density function (PDF) for a given variable x_i is denoted as $f_i^b(x_i)$, while for the initial \bar{b} quark it is denoted as $f_i^{\bar{b}}(x_i)$. The combined tagging variable y is defined as:

$$y = \prod_{i=1}^n y_i; \quad y_i = \frac{f_i^{\bar{b}}(x_i)}{f_i^b(x_i)} \quad (4)$$

A given variable x_i can be undefined for some events. For example, there are events which don't contain an identified muon from the opposite side. In this case, the corresponding variable y_i is set to 1. The initial b flavor is more probable if $y < 1$, and \bar{b} flavor is more probable if $y > 1$. Correspondingly, an event with $y < 1$ is tagged as b quark and the event with $y > 1$ is tagged as \bar{b} quark. For an oscillation analysis, it is more convenient to define the tagging variable as $d = (1 - y)/(1 + y)$. The variable d ranges between -1 and 1. An event with $d > 0$ is tagged as b quark and with $d < 0$ as \bar{b} quark. Higher $|d|$ value corresponds to a higher tagging purity. For uncorrelated variables x_1, \dots, x_n , and perfect modeling in the PDF, d gives the best possible tagging performance and its absolute value gives a dilution of a given event.

Currently, all of our discriminating variables are constructed using properties of the b quark opposite to the reconstructed B hadron (“opposite side tagging”). It is assumed that every event with b quark also contains a \bar{b} quark. Therefore, the b flavor at the opposite side determines the b flavor at the reconstruction side. An important property of opposite side tagging is the independence of its performance on the type of the reconstructed B hadron, since the hadronization of two b quark 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. After that, the measured performance can be used for other purposes, such as studying B_s meson oscillations.

Another set of variables, which exploit properties of hadronization $b \rightarrow B$ at the reconstruction side, can also be constructed (“same side tagging”). The tagging with these variables depends on the type of B meson. Its performance can only be obtained from the simulation and is therefore model dependent. Currently, it is not used for the B_s mixing measurement and is not described here.

The probability density function for each discriminating variable discussed below was constructed using real data $B \rightarrow \mu^+\nu D^0$ events with the visible proper decay length less than 500 μm . The definition of the visible proper decay length is given in section V. In this sample, the non-oscillating decays $B^0 \rightarrow \mu^+\nu D^{*-}$ dominate and the initial state of a b -quark is determined by the charge of the muon. According to MC estimates, the purity of such identification of the initial flavor in the selected sample is 0.98 ± 0.01 , where the error reflects the uncertainty in branching ratios of B decay. The background under the D^* peak was subtracted for each distribution using the $\mu^+\bar{D}^0\pi^+$ events with the wrong sign of π .

In each analyzed event, an additional identified muon was searched for. This muon was used for flavor tagging if $\cos\phi(\mathbf{p}_\mu, \mathbf{p}_B) < 0.8$, where \mathbf{p}_B is a three-momentum of the reconstructed B hadron. If more than one muon was found, the muon with the highest number of hits in the muon chambers was used. If more than one muon with the same number of hits in the muon chambers was found, the muon with the highest p_T was used.

For each such muon, a *muon jet charge*

$$Q_J^\mu = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}$$

was constructed. The sum was taken over all charged particles, including the muon, satisfying the condition $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$. $\Delta\phi$ and $\Delta\eta$ were computed with respect to the muon direction. Daughters of the reconstructed B hadron were explicitly excluded from the sum. The distribution of the muon jet charge variable is shown in Fig. 4a and 4b.

An additional identified electron [5] was used for the flavor tagging if $\cos\phi(\mathbf{p}_e, \mathbf{p}_B) < 0.8$. For this electron, an *electron jet charge* Q_J^e was constructed as:

$$Q_J^e = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}$$

The sum was taken over all charged particles, including the electron, satisfying the condition $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$. $\Delta\phi$ and $\Delta\eta$ were computed with respect to the electron direction. Daughters of the reconstructed B meson were explicitly excluded from the sum. In addition, any charged particle with $\cos\phi(\mathbf{p}, \mathbf{p}_B) > 0.8$ was excluded. The distribution of the electron jet charge variable is shown in Fig. 4c.

A secondary vertex corresponding to the decay of B hadrons was searched for using all charged particles in the event. The secondary vertex should contain at least 2 particles with the transverse impact parameter significance greater than 3. The distance l_{xy} from the primary to the secondary vertex should satisfy the condition: $l_{xy} > 4\sigma(l_{xy})$. The details of the secondary vertex search can be found in [6].

The momentum of the secondary vertex \mathbf{p}_{SV} was defined as the sum of all momenta of particles included in the secondary vertex. The secondary vertex with $\cos\phi(\mathbf{p}_{SV}, \mathbf{p}_B) < 0.8$ was used in the flavor tagging. The secondary vertex containing any particle from the decay of the reconstructed B hadron was excluded from the tagging. A *secondary vertex charge* Q_{SV} was defined as the third discriminating variable:

$$Q_{SV} = \frac{\sum_i (q^i p_L^i)^{0.6}}{\sum_i (p_L^i)^{0.6}}$$

where the sum was taken over all particles included in the secondary vertex. The p_L^i is the longitudinal momentum of a given particle with respect to the direction of the secondary vertex momentum. Figs. 5a and 5b show the distribution of this variable for events with and without an identified muon.

Finally, the *event charge*

$$Q_{EV} = \frac{\sum_i q^i p_T^i}{\sum_i p_T^i}$$

was constructed. The sum was taken over all charged particles with $p_T > 0.5$ GeV/ c and having $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} > 1.5$. $\Delta\phi$ and $\Delta\eta$ were computed with respect to the reconstructed B -hadron direction. Due to a strong correlation with the muon jet charge, this variable was not used for events with an identified muon. The distribution of this variable is shown in Fig. 5d.

For each event with an identified muon, the muon jet charge Q_J^μ and the secondary vertex charge Q_{SV} were used to construct a *muon tagger*. For each event without a muon but with an identified electron, the electron charge Q_J^e and the secondary vertex charge Q_{SV} were used to construct an *electron tagger*. Finally, for events without a muon or an electron but with a reconstructed secondary vertex, the secondary vertex charge Q_{SV} and the event jet charge Q_{EV} were used to construct a secondary vertex tagger. The resulting distribution of the tagging variable d for the combination of all three taggers, called in the following as a combined tagger, is shown in Fig. 6.

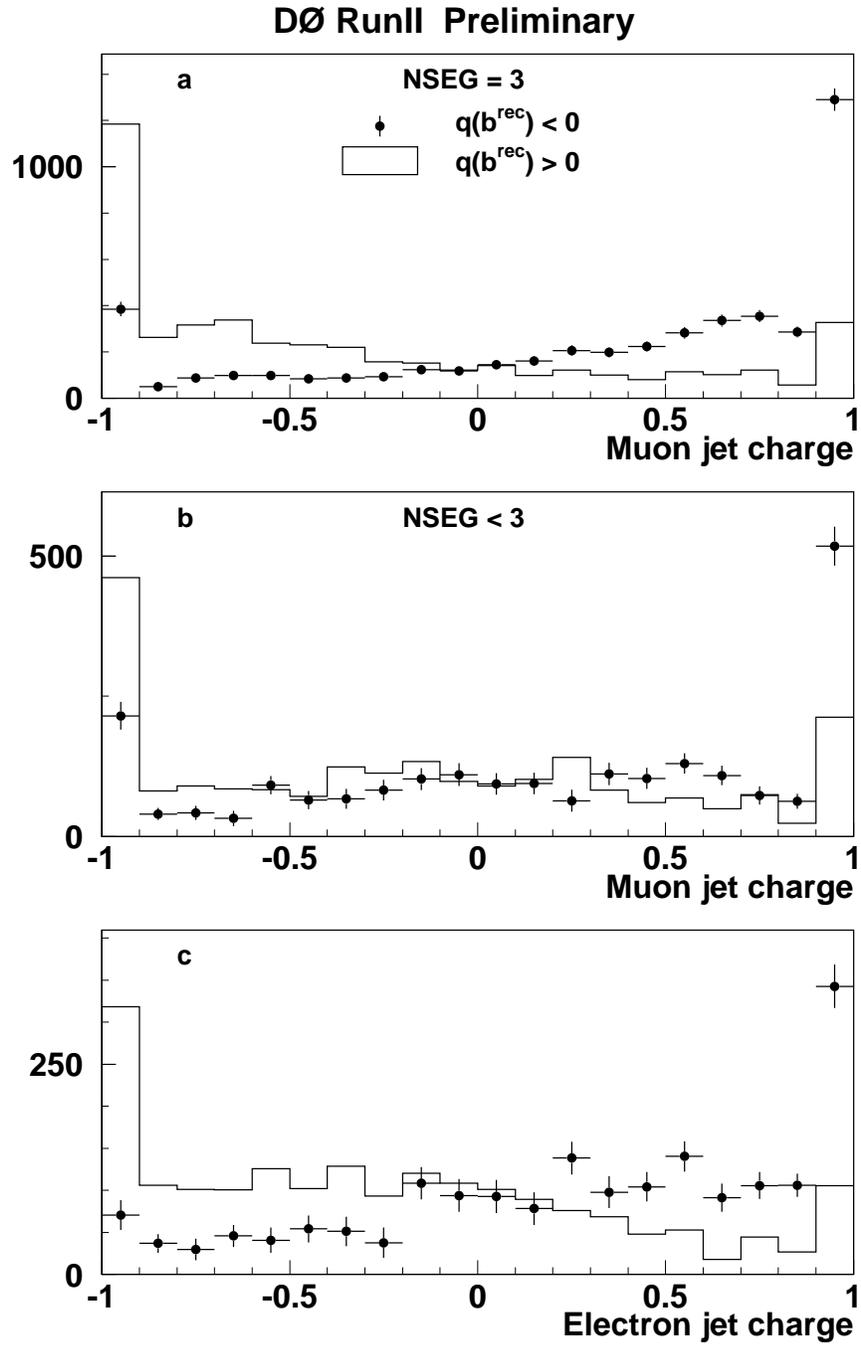


FIG. 4: a) Distribution of muon jet charge for muons with $NSEG = 3$. b) Distribution of muon jet charge for muons with $NSEG < 3$. c) Distribution of electron jet charge. The $q(b^{\text{rec}})$ is the charge of the b quark from the reconstruction side.

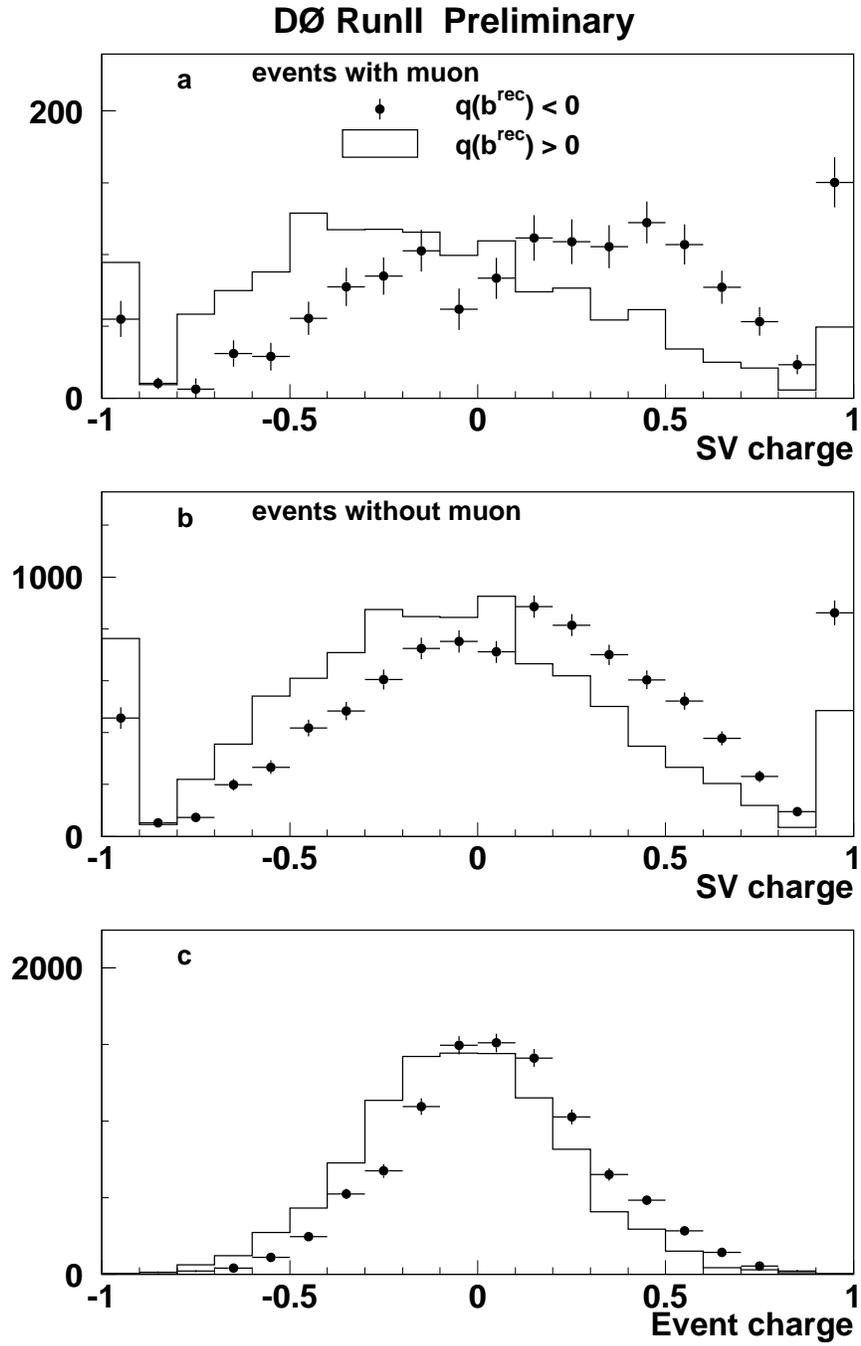


FIG. 5: a) Distribution of secondary vertex charge for events with muon. b) Distribution of secondary vertex charge for events without muon. c) Distribution of event jet charge. The $q(b^{rec})$ is the charge of the b quark from the reconstruction side.

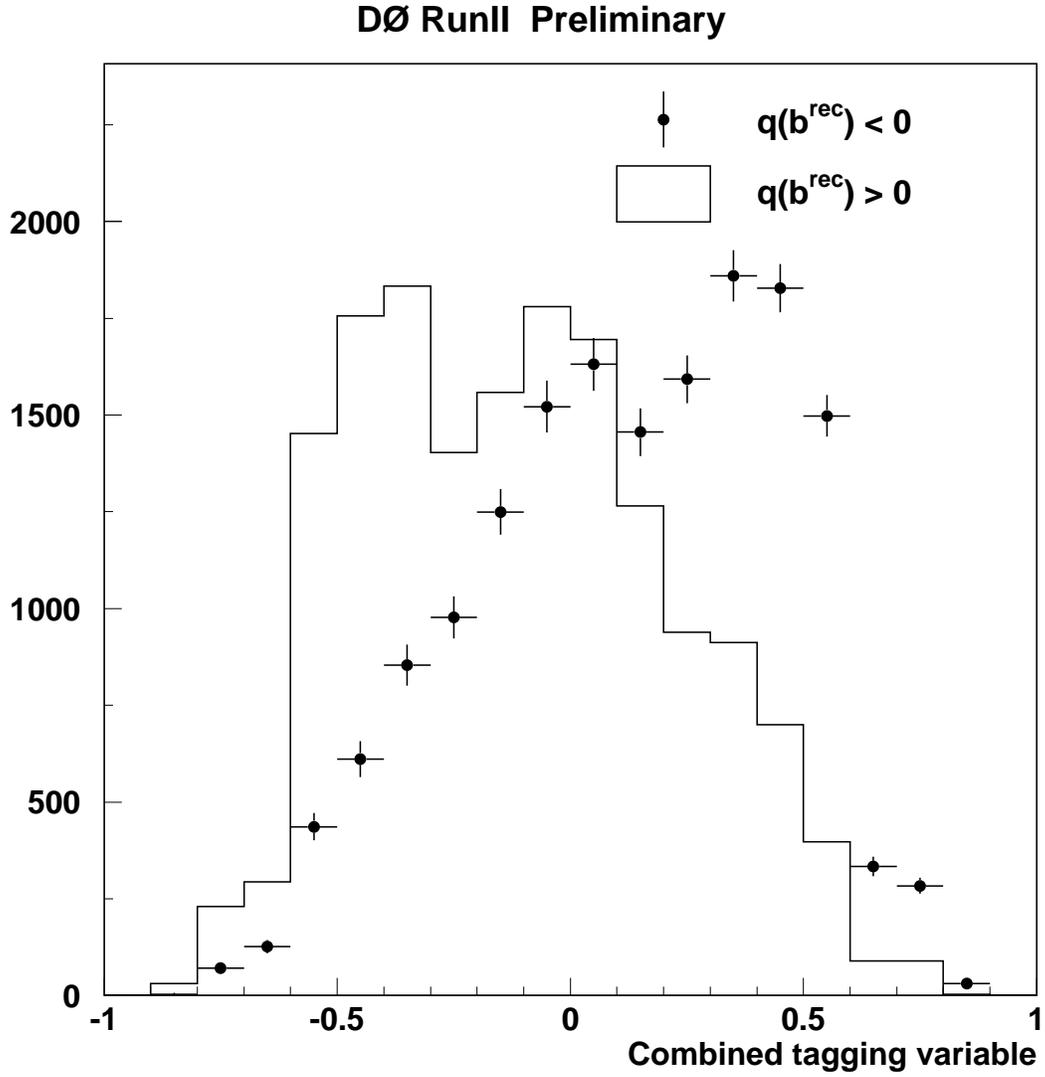


FIG. 6: Normalized distributions of the combined tagging variable. The $q(b^{\text{rec}})$ is the charge of the b quark from the reconstruction side.

IV. MUTIDIMENSIONAL TAGGER

In addition to the primary flavor tagger described in section III, an alternative flavor tagging algorithm has also been developed and used to measure B_d^0 mixing. Because this alternative tagger is trained on Monte Carlo events and the primary flavor tagger is trained on data, it can be used to cross-check the primary tagging algorithm.

The alternative tagger is *multidimensional*; the likelihood functions of which it is composed can be functions of more than a single variable. Thus if, as before, we have a set of discriminants x_1, \dots, x_n , the likelihood that the meson has the flavor B^0 at the time of creation can be written $\mathcal{L}(B^0, \vec{x})$. A similar equation holds for the likelihood for $\overline{B^0}$, namely $\mathcal{L}(\overline{B^0}, \vec{x})$. These likelihoods relate to the variable d as,

$$d = \frac{\mathcal{L}(\overline{B^0}) - \mathcal{L}(B^0)}{\mathcal{L}(\overline{B^0}) + \mathcal{L}(B^0)} \quad (5)$$

where the likelihoods are on the reconstructed side of the event. Note that this definition is identical to the one in section III.

A. Samples and Tagging Logic

We obtain our likelihoods from Monte Carlo samples of $B^\pm \rightarrow J/\psi K^\pm$ with the J/ψ decaying to $\mu^+ \mu^-$. This final state does not oscillate and is therefore flavor pure. The Monte Carlo sample $B^+ \rightarrow J/\psi K^+$ is used to create $\mathcal{L}(B^0, \vec{x})$ and $B^- \rightarrow J/\psi K^-$ is used to create $\mathcal{L}(\overline{B^0}, \vec{x})$. The selections for reconstruction are the same as in [7]. In practice, the likelihoods are histograms that have one dimension per discriminant whose bin contents have been normalized to the total number of events in the sample. For a given event, the tagger output d is obtained by substituting the appropriate normalized bin contents into Eq. (5).

In principle, one could combine all of the discriminants mentioned in section III into a single multidimensional likelihood and use that as a flavor tagger. However, one must remember that these are binned likelihoods and that in order to achieve a reasonable resolution in any given discriminant, the binning must be fine enough to resolve its useful features. In practice, because of finite Monte Carlo statistics, this means that one must choose discriminants wisely when attempting to make a combination.

We therefore divide events into three categories based on their opposite side contents:

1. μ and a secondary vertex.
2. μ without a secondary vertex.
3. Secondary vertex without a μ .

and choose the following sets of discriminants for these taggers:

1. $\text{Tag}(\mu + \text{SV}) = \{Q_J^\mu; p_T^{rel}; Q_{SV}\}$
2. $\text{Tag}(\mu \text{ without SV}) = \{Q_J^\mu; p_T^{rel}; p_T; \text{impact parameter significance}\}$
3. $\text{Tag}(\text{SV without } \mu) = \{Q_{EV}; Q_{SV}; p_T^{SV}\}$

The primary reason for the above grouping is that when Monte Carlo statistics are limited we want to use the best discriminants available. Thus, for the case where both a μ and secondary vertex exist on the opposite side, we want to be sure to use both the μ jet charge as well as the secondary vertex charge because we know that these two variables are the strongest discriminants. For the case where there is a μ but no secondary vertex, we have more freedom to add in discriminants. Finally, for the last case, we simply do not have many discriminants available.

Distributions of the tagging variable d for the above three taggers are shown in Fig. 7. The distributions shown in this figure are made by applying the taggers to the Monte Carlo $B^\pm \rightarrow J/\psi K^\pm$ samples from which they are created.

The final multidimensional tagger employed the following logic to decide which of its sub-taggers to use:

1. If the opposite side contains a $\mu +$ secondary vertex, use $\text{Tag}(\mu + \text{SV})$.
2. If the opposite side contains a μ and no secondary vertex, use $\text{Tag}(\mu \text{ without SV})$.
3. If the opposite side contains an electron, use the electron tagger described in section III. Note that this tagger is not multidimensional and is not derived from Monte Carlo.
4. If the opposite side contains a secondary vertex, use $\text{Tag}(\text{SV without } \mu)$.

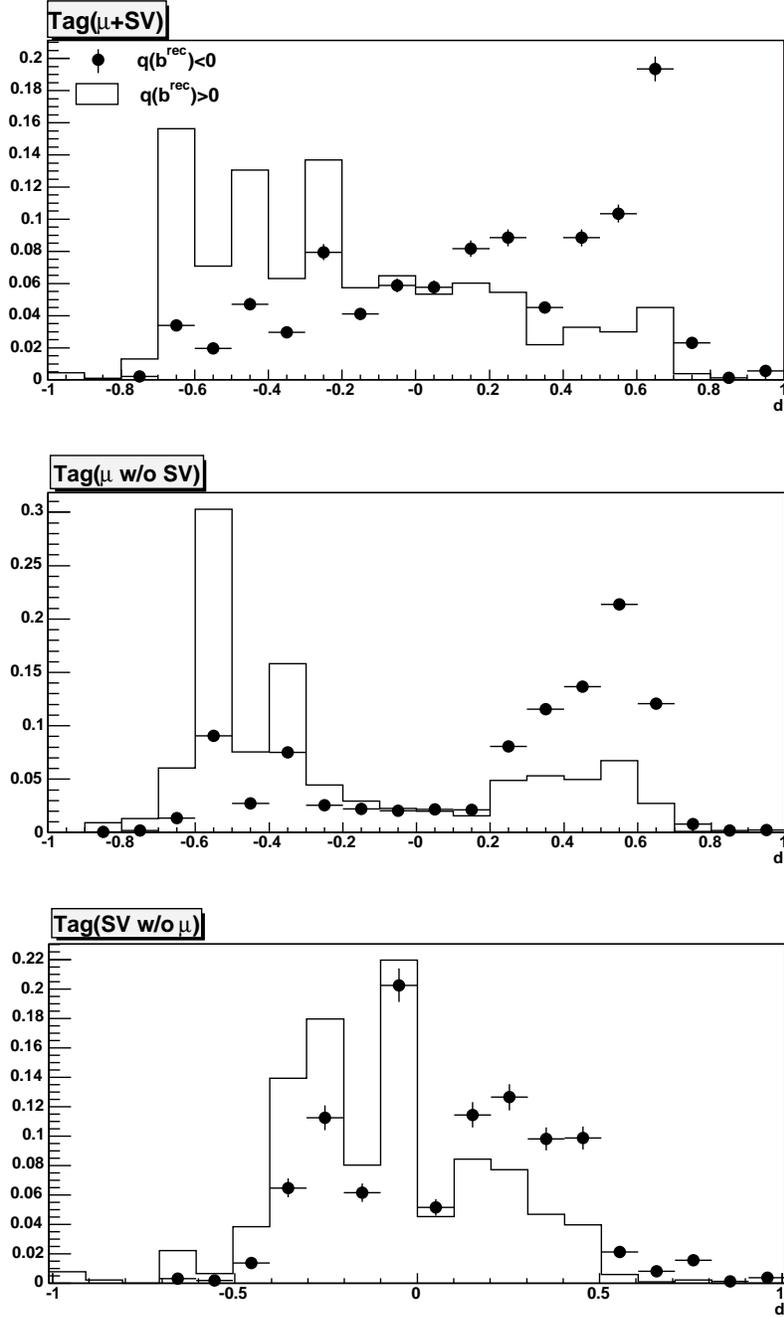


FIG. 7: Normalized distributions of the combined tagging variable for the 3 multidimensional taggers over the Monte Carlo samples $B^\pm \rightarrow J/\psi K^\pm$ from which they are created. $q(b^{rec})$ is the charge of the b quark from the reconstructed side. a) Distribution of d for Tag(μ +SV). b) Distribution of d for Tag(μ without SV). c) Distribution of d for Tag(SV without μ).

V. B_d^0 MIXING AND TAGGER CALIBRATION

The performance of the flavor tagging was studied separately for the muon, electron and secondary vertex taggers. Events with $|d| > 0.3$ were used. Also, the performance of the combined tagger defined in section III for events with $|d| > 0.3$ and of the alternative multidimensional tagger defined in section IV for events with $|d| > 0.37$ are also shown. The cuts on the $|d|$ values are chosen so as to keep same efficiency for the combined and multidimensional tagger.

For optimal performance, the combined tagger tagged events were divided into subsamples with $0.1 < |d| < 0.2$, $0.2 < |d| < 0.35$, $0.35 < |d| < 0.45$, $0.45 < |d| < 0.6$, and $|d| > 0.6$ and were fitted simultaneously.

To do this, two experimental observables, the flavor asymmetry and the visible proper decay length (VPDL), were defined. The visible proper decay length x^M was defined as:

$$x^M = (\mathbf{L}_{xy} \cdot \mathbf{P}_{xy}^{\mu D^0}) / (P_T^{\mu D^0})^2 \cdot M_B. \quad (6)$$

Here \mathbf{L}_{xy} is defined as the vector in transverse plane from the primary to the B -meson decay vertex. The transverse momentum of a B -meson $\mathbf{P}_T^{\mu D^0}$ was defined as the vector sum of transverse momenta of muon and D^* . The algorithm to find a secondary vertex is described in [6].

All events in the D^* and \bar{D}^0 sample were divided into 7 groups according to their measured VPDL. Tagged events, where the reconstructed muon and flavor tag variable $|d|$ have opposite sign were tagged as non-oscillating (N^{nos}) and events where they have the same sign were tagged oscillating (N^{osc}).

The number of oscillating N_i^{osc} and non-oscillating N_i^{nos} μD^* events in each interval i was determined from a fit of the D^* signal in the mass $M(D^0\pi) - M(D^0)$ distribution for the D^* sample. The number of $\mu^+\bar{D}^0$ events was determined from a fit of the \bar{D}^0 signal in the $K\pi$ invariant mass distribution for the \bar{D}^0 sample.

The flavor asymmetry A_i in each VPDL bin was defined as:

$$A_i = \frac{N_i^{nos} - N_i^{osc}}{N_i^{nos} + N_i^{osc}}, \quad (7)$$

The fitting procedure as described in [8] was used to fit the measured asymmetries A_i . The number of \bar{D}^0 events was determined using the same fitting procedure as for the D^* sample and the additional fraction of $B \rightarrow \mu^+\nu\bar{D}^0$ events in the D^* sample was estimated to be $4 \pm 0.85\%$. This fraction was included in the fitting procedure and the uncertainty in this value was taken into account in the systematic error of obtained results.

For any sample of tagged events, the values of Δm_d , $f_{c\bar{c}}$, $|d|$ and \mathcal{D}_d were obtained from the simultaneous χ^2 fit:

$$\begin{aligned} \chi^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, |d|) &= \chi_{D^*}^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, |d|) + \chi_{\bar{D}^0}^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, |d|) \quad (8) \\ \chi_{D^*}^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, |d|) &= \sum_i \frac{(A_{i,D^*} - A_{i,D^*}^e(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, |d|))^2}{\sigma^2(A_{i,D^*})} \\ \chi_{\bar{D}^0}^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, |d|) &= \sum_i \frac{(A_{i,\bar{D}^0} - A_{i,\bar{D}^0}^e(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, |d|))^2}{\sigma^2(A_{i,\bar{D}^0})}. \end{aligned}$$

Here \sum_i is the sum over all VPDL bins.

Figures 8 - 12 show the measured asymmetry for different taggers. A clear oscillation pattern for D^* events, and the reduced oscillation for \bar{D}^0 events are clearly seen. These distributions are described reasonably well by the oscillation functions, which are superimposed in these figures.

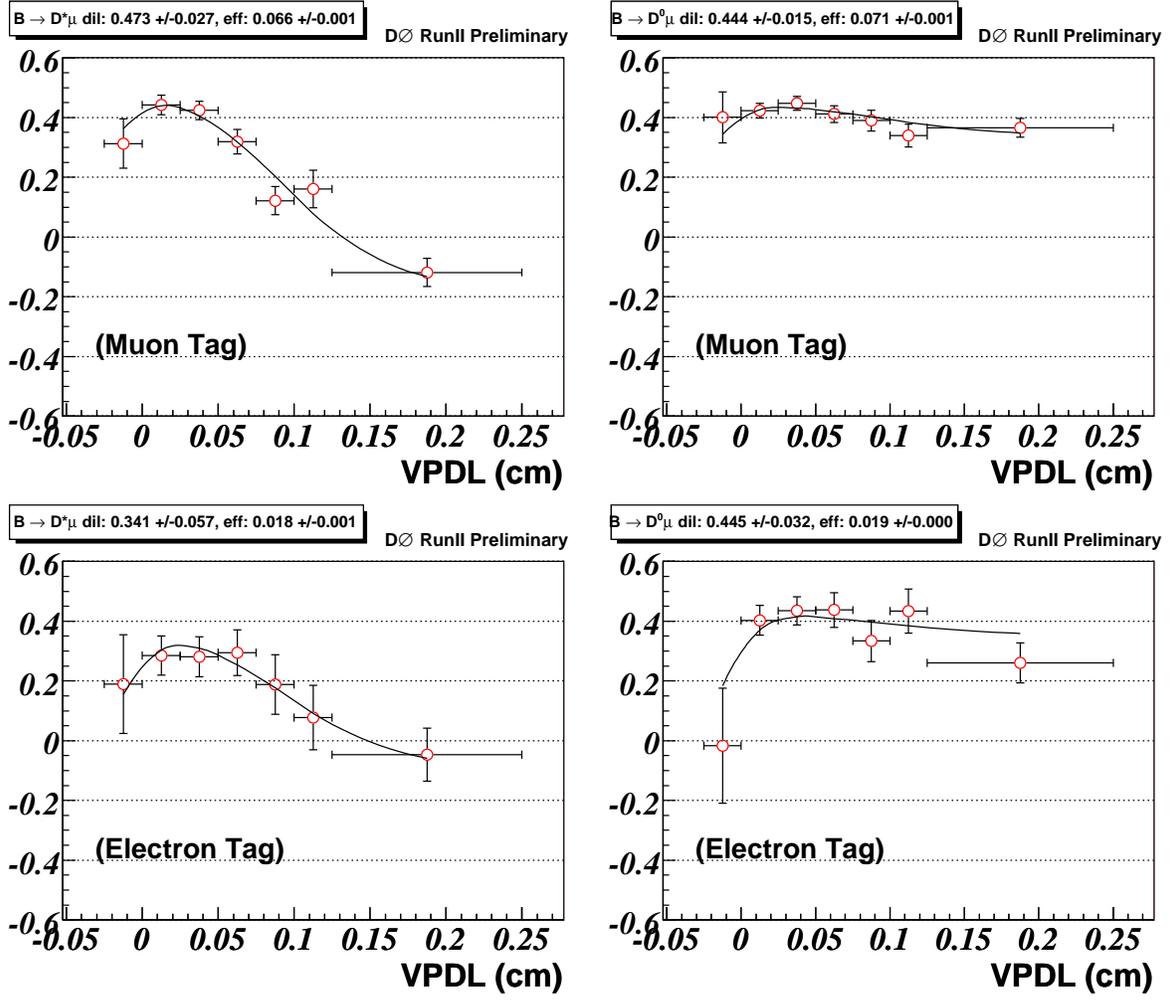


FIG. 8: The asymmetries obtained in the D^* and D^0 sample with the result of the fit superimposed for the Muon and electron tagger. For the individual taggers, $|d| > 0.3$ was required.

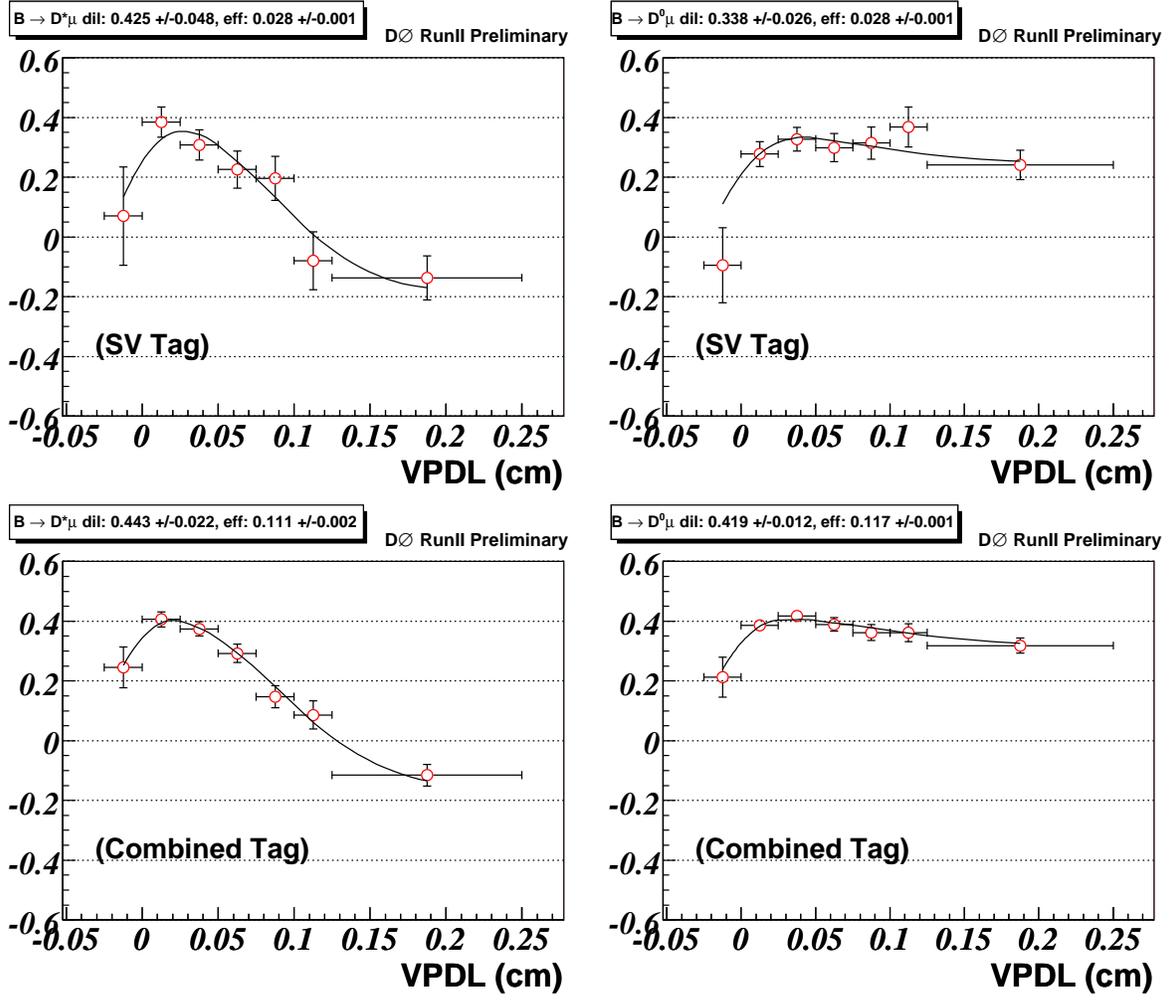


FIG. 9: The asymmetries obtained in the D^* and D^0 sample with the SV and the combined tagger and the result of the fit superimposed. The samples required $|d| > 0.3$.

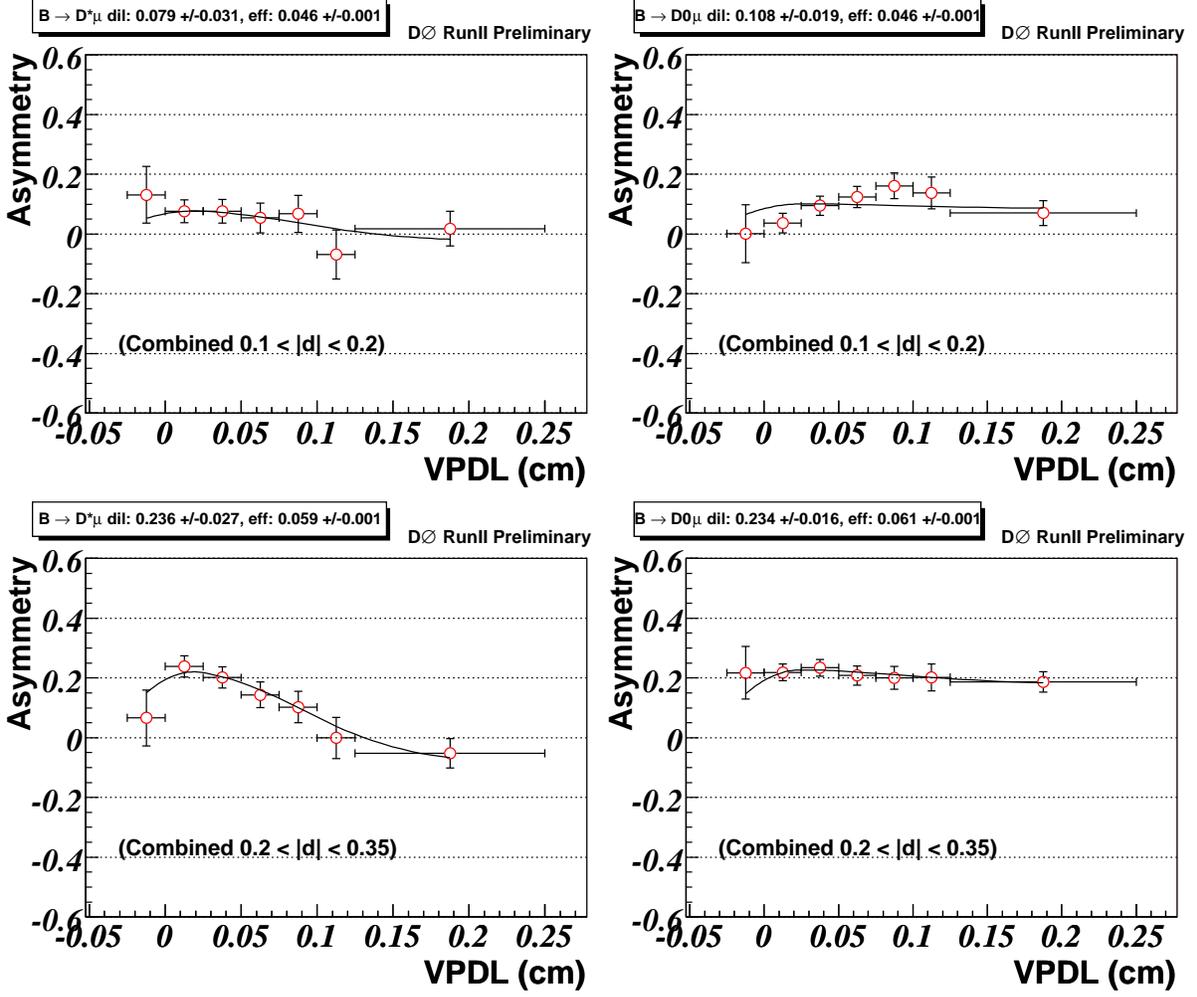


FIG. 10: The asymmetries obtained in the D^* and D^0 sample with the combined tagger in $|d|$ bins, 0.1-0.2 and 0.2-0.35. The result of the fit superimposed

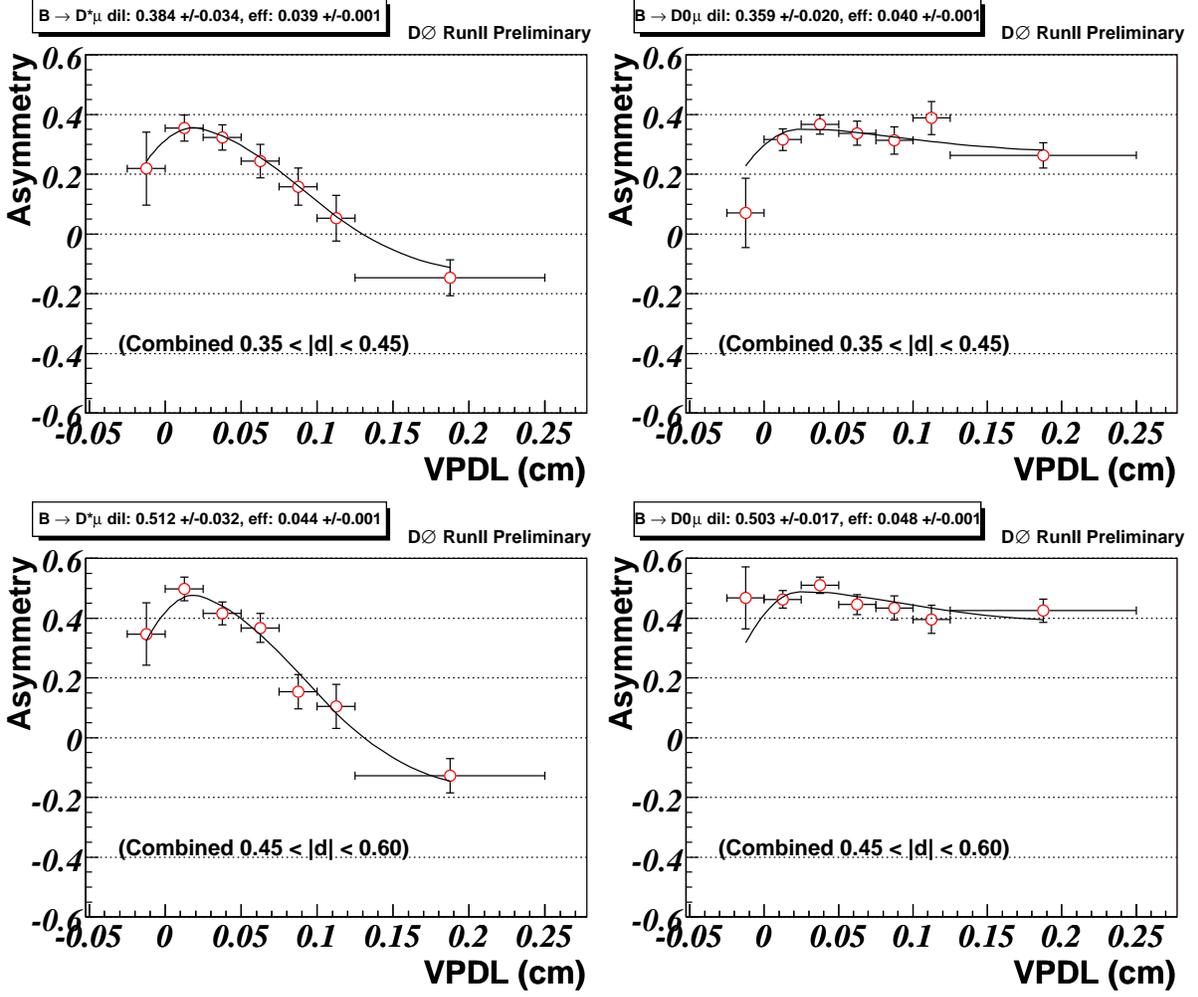


FIG. 11: The asymmetries obtained in the D^* and D^0 sample with the combined tagger in $|d|$ bins, 0.35-0.45 and 0.45-0.6. The result of the fit superimposed

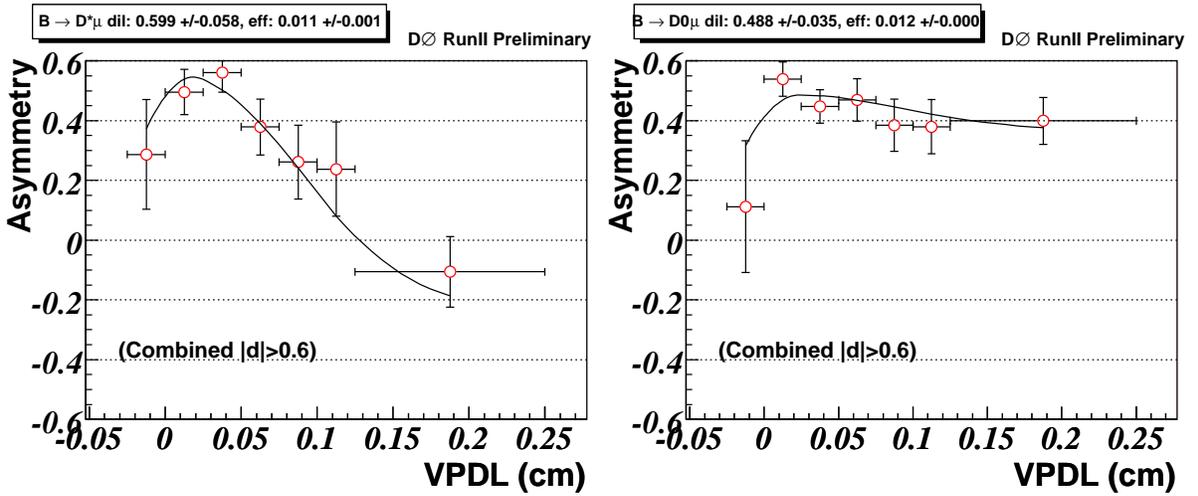


FIG. 12: The asymmetries obtained in the D^* and D^0 sample with the combined tagger for bin $|d| > 0.6$. The result of the fit superimposed

VI. RESULTS

Tables I-III give results obtained for each considered tagger. All errors are statistical and don't include systematic uncertainties. The performance of the combined tagger defined in section III for events with $|d| > 0.3$ and of the alternative multidimensional tagger defined in section IV for events with $|d| > 0.37$ are also shown. The tagging efficiencies shown in Tables I and II were computed using events with the VPDL=[0.025,0.250]. This selection reduces the contribution from $c\bar{c} \rightarrow \mu^+\nu D^0 X$ events since they have a VPDL distribution with zero mean and $\sigma \sim 150\mu\text{m}$ according to our study.

For the combined tagger with $|d| > 0.3$ the following results were obtained:

$$\begin{aligned}\varepsilon\mathcal{D}_d^2 &= (2.19 \pm 0.22)(\%) \\ \Delta\text{m}_d &= 0.513 \pm 0.023 \\ f_{c\bar{c}} &= (3.3 \pm 1.3)(\%)\end{aligned}\tag{9}$$

For the multidimensional tagger with $|d| > 0.37$, which has the same efficiency as the combined tagger, the following results were obtained:

$$\begin{aligned}\varepsilon\mathcal{D}_d^2 &= (1.71 \pm 0.19)(\%) \\ \Delta\text{m}_d &= 0.502 \pm 0.026 \\ f_{c\bar{c}} &= (3.1 \pm 1.4)(\%)\end{aligned}\tag{10}$$

One of the purposes of this measurement is to extract the dilutions for reconstructed B_d and B_u mesons and validate the assumption that these dilutions do not depend on the type of B meson. It can be seen from Tables I,II that the measured flavor tagging performance for B^0 events is slightly better than for B^+ events, both for individual and combined taggers. This difference can be explained by a better selection of $\mu^+\nu D^{*-}$ events due to an additional requirement of the charge correlation between the muon and the pion from $D^{*-} \rightarrow D^0\pi^-$ decay. The D^0 sample can contain events with a wrongly selected muon. Since the charge of the muon determines the flavor asymmetry, such a background can reduce the measured B^+ dilution. The charge correlation between the muon and the pion can suppress this background and result in a better measurement of the tagging performance. To test this hypothesis, a special sample of events satisfying all conditions for D^* sample, except the requirement of the charge correlation between the muon and the pion, was selected. The dilution \mathcal{D}'_d for such sample is shown in Table II. It can be seen that \mathcal{D}'_d is statistically compatible with $|d|$ for all samples and all taggers. The χ^2 for the difference in dilutions is found to be 1.06 to be compared to 1.27 in the case where only RS events are considered for the D^* sample.

By construction of the combined tagging, the dilution for any event should strongly depend on the magnitude of the variable d . This property becomes important in the B_s mixing measurement, since in this case the dilution of each event can be estimated using the value of d and can be included in the likelihood function, improving the sensitivity of the measurement.

To test the dependence of dilution on d , and to obtain the final result for the mixing parameter, all tagged events were divided into subsamples with $0.1 < |d| < 0.2$, $0.2 < |d| < 0.35$, $0.35 < |d| < 0.45$, $0.45 < |d| < 0.6$, and $|d| > 0.6$. The overall efficiency of this sample is $(19.95 \pm 0.21)(\%)$. The obtained dilutions are shown in Table I. Their strong dependence on the value of the tagging variable is clearly seen. The overall tagging power was computed as the sum of tagging powers. The fraction $f_{c\bar{c}}$ was constrained to be the same for all subsamples. We obtain the following final result:

$$\varepsilon\mathcal{D}_d^2 = (2.48 \pm 0.21)(\%)\tag{11}$$

$$\Delta\text{m}_d = 0.506 \pm 0.020\tag{12}$$

$$f_{c\bar{c}} = (2.2 \pm 0.9)(\%)$$

VII. SYSTEMATIC UNCERTAINTIES

The systematic uncertainties are summarized in Tables IV and V. Table IV shows the contributions to the systematic uncertainty of Δm_d . Table V shows the corresponding contributions in $\mathcal{D}(B^0)$.

The estimate of different systematic effects is described below:

The B meson branching rates and lifetimes used in the fit of the asymmetry were taken from [9] and were varied by one standard deviation.

Tagger	$\varepsilon(\%)$	\mathcal{D}_d	$\varepsilon\mathcal{D}_d^2(\%)$
Muon ($ d > 0.3$)	6.61 ± 0.12	0.473 ± 0.027	1.48 ± 0.17
Electron ($ d > 0.3$)	1.83 ± 0.07	0.341 ± 0.058	0.21 ± 0.07
SVCharge ($ d > 0.3$)	2.77 ± 0.08	0.424 ± 0.048	0.50 ± 0.11
Combined ($ d > 0.3$)	11.14 ± 0.15	0.443 ± 0.022	2.19 ± 0.22
Multidim ($ d > 0.37$)	10.98 ± 0.15	0.395 ± 0.022	1.71 ± 0.19
Combined($0.10 < d < 0.20$)	4.63 ± 0.10	0.084 ± 0.031	0.03 ± 0.02
Combined($0.20 < d < 0.30$)	5.94 ± 0.12	0.236 ± 0.027	0.33 ± 0.08
Combined($0.30 < d < 0.45$)	3.89 ± 0.09	0.385 ± 0.034	0.58 ± 0.10
Combined($0.45 < d < 0.60$)	4.36 ± 0.10	0.512 ± 0.032	1.14 ± 0.14
Combined($0.60 < d < 1.00$)	1.13 ± 0.05	0.597 ± 0.058	0.40 ± 0.08

TABLE I: Tagging performance for events with reconstructed B^0 for different taggers and subsamples.

Tagger	$\varepsilon(\%)$	$ d $	$\varepsilon d ^2 (\%)$	\mathcal{D}'_d
Muon ($ d > 0.3$)	7.10 ± 0.09	0.444 ± 0.015	1.400 ± 0.096	0.463 ± 0.028
Electron ($ d > 0.3$)	1.88 ± 0.05	0.445 ± 0.032	0.372 ± 0.054	0.324 ± 0.060
SVCharge ($ d > 0.3$)	2.81 ± 0.06	0.338 ± 0.026	0.320 ± 0.050	0.421 ± 0.049
Combined ($ d > 0.3$)	11.74 ± 0.11	0.419 ± 0.012	2.058 ± 0.121	0.434 ± 0.023
Multidim ($ d > 0.37$)	11.67 ± 0.11	0.363 ± 0.012	1.540 ± 0.106	0.384 ± 0.023
Combined($0.10 < d < 0.20$)	4.59 ± 0.08	0.104 ± 0.017	0.050 ± 0.016	0.079 ± 0.029
Combined($0.20 < d < 0.30$)	6.10 ± 0.09	0.234 ± 0.014	0.335 ± 0.042	0.212 ± 0.024
Combined($0.30 < d < 0.45$)	3.98 ± 0.07	0.361 ± 0.018	0.519 ± 0.052	0.364 ± 0.032
Combined($0.45 < d < 0.60$)	4.77 ± 0.07	0.504 ± 0.016	1.211 ± 0.077	0.489 ± 0.030
Combined($0.60 < d < 1.00$)	1.17 ± 0.04	0.498 ± 0.031	0.290 ± 0.038	0.572 ± 0.056

TABLE II: Tagging performance for events with reconstructed B^+ for different taggers and subsamples. For comparison, the dilution \mathcal{D}'_d measured in the D^* sample with addition of wrong sign $\mu^+ \nu \bar{D}^0 \pi^+$ events is also shown.

The VPDL resolution, obtained in simulation, was multiplied by a large factor, from 0.8 to 1.2, which significantly exceeds the estimated difference in the resolution between data and simulation.

The variation of K -factors with the change of B momentum was neglected in this analysis. The K -factor is an input in the asymmetry fit and is defined as $K = P_T^{\mu D^0} / P_T^B$ and reflects the difference between the measured ($P_T^{\mu D^0}$) and true (P_T^B) momenta of B meson. Since we only partially reconstruct the B because of the presence of the neutrino, we have to correct the visible proper decay length with this factor, and the VPDL we use, is related to the true B decay length, $x_B = Kx$. See [8] for more details. To check the impact of this assumption on the final result, their computation was repeated without the cut on $p_T(D^0)$ or by applying an additional cut on p_T of muon, $p_T > 4$ GeV/ c . The change of average value of K -factors did not exceed 2%, which was used as the estimate of the systematic uncertainty in their values. This uncertainty was afterwards propagated into the variation of Δm_d and tagging purity by repeating the fit with the K -factor distributions shifted by 2%.

The reconstruction efficiency in different B -meson decay channels depends only on the kinematic properties of corresponding decays and can therefore be reliably estimated in the simulation. The $ISGW2$ model [10] of the

Tagger	Δm_d	$f_{c\bar{c}}$
Muon	0.502 ± 0.028	0.013 ± 0.010
Electron	0.481 ± 0.067	0.058 ± 0.045
SV Charge	0.553 ± 0.053	0.096 ± 0.050
Multidim	0.502 ± 0.026	0.031 ± 0.014
Combined($ d > 0.3$)	0.513 ± 0.023	0.033 ± 0.013
Combined($0.10 < d < 0.20$)	0.506 ± 0.209	0.495 ± 0.505
Combined($0.20 < d < 0.35$)	0.523 ± 0.064	0.021 ± 0.025
Combined($0.35 < d < 0.45$)	0.531 ± 0.042	0.063 ± 0.038
Combined($0.45 < d < 0.60$)	0.510 ± 0.032	0.010 ± 0.010
Combined($0.60 < d < 1.00$)	0.456 ± 0.049	0.032 ± 0.026

TABLE III: Measured value of Δm_d and $f_{c\bar{c}}$ for different taggers and subsamples.

semileptonic B decays was used. The uncertainty of the reconstruction efficiency, set at 12%, was estimated by varying kinematic cuts on the p_t of the muon and D^0 in a wide range. Changing the model describing semileptonic B decay from $ISGW2$ to $HQET$ [11] produces a smaller variation. The fit of asymmetry was repeated with the efficiencies to reconstruct $B \rightarrow \mu^+ \nu D^{*-}$ and $B \rightarrow \mu^+ \nu \bar{D}^{*0}$ channels modified by 12%, and the difference was taken as the systematic uncertainty from this source.

The additional fraction of D^0 events contributing to the D^* signal was estimated at $4 \pm 0.85\%$, see section V. This variation was used to estimate the uncertainty from this source. As a crosscheck, the number of D^* events was determined from the fit of the mass difference $M(D^0\pi) - M(D^0)$ and the fit of the flavor asymmetry was repeated. The measured value of $\Delta m_d = 0.507 \pm 0.020$ is consistent with (13).

We also investigated the systematic uncertainty of measuring the number of D^* and D^0 candidates in each VPDL bin. The values of the parameters which had been fixed from the fit to “all” events, were varied by $\pm 3\sigma$. This corresponds to the studies of the mass fits in the 7 VPDL bins, where it was found that deviations from the “all” event fit values were not significant above the 3σ level.

	default	variation		$\Delta m_d/\text{ps}^{-1}$	
		(a)	(b)	(a)	(b)
$Br(B^0 \rightarrow D^{*-} \mu^+ \nu)$	5.44	-0.23	0.23	.002	-.002
$Br(B \rightarrow D^* \pi \mu \nu X)$	1.07	-0.17	0.17	-.0078	.0078
R^{**}	0.35	.0	1.0	.0006	-.0012
B lifetimes	.05022	-.00054	.00054	.0008	-.0008
Resolution function	—	$\div 1.2$	$\div 0.8$.0021	-.0021
Alignment	—	$-10\mu m$	$+10\mu m$	± 0.004	-
K -Factor	—	-2%	+2%	.0098	-.0094
Efficiency	—	-12%	+12%	-.0054	.0052
Fraction D^0 in D^*	4%	3.15%	4.85%	-.0020	+.0030
Fit Procedure	See split below				
Bin width	2 MeV	1.6	2.67	.0009	.0014
Parameter μ_0	—	-3σ	3σ	-.0001	.0001
Parameter $\frac{\sigma_R + \sigma_L}{2}$	—	-3σ	3σ	-.0001	—
Parameter $\frac{\sigma_R - \sigma_L}{\sigma_R + \sigma_L}$	—	-3σ	3σ	-.0001	.0001
Parameter μ_1	—	-3σ	3σ	-.0016	.0015
Parameter $\frac{\sigma_1 + \sigma_2}{2}$	—	-3σ	3σ	-.0006	.0006
Parameter \bar{R}	—	-3σ	3σ	-.0005	.0004
Parameter $(\mu_2 - \mu_1)$	—	-3σ	3σ	.0006	-.0007
Parameter $\frac{\sigma_1 - \sigma_2}{\sigma_1 + \sigma_2}$	—	-3σ	3σ	—	—
Fit Procedure		Overall		+.0023	-.0019
Total				+.0158	-.0158

TABLE IV: Systematic uncertainties Δm_d .

VIII. CONCLUSIONS

We have performed a study of the likelihood-based opposite-side tagging algorithm in B^0 and B^+ samples. The dilutions $\mathcal{D}(B^+)$ and $\mathcal{D}(B^0)$, are consistent within their statistical error.

Splitting the sample into bins according to the tagging variable $|d|$ and measuring the tagging power as the sum of individual tagging power in all bins we obtained

$$\varepsilon \mathcal{D}^2 = (2.48 \pm 0.21 \text{ (stat.)}_{-0.06}^{+0.08} \text{ (syst)}) (\%)$$

From a simultaneous fit to events in all $|d|$ bins we measured the mixing parameter Δm_d parameter:

$$\Delta m_d = 0.506 \pm 0.020 \text{ (stat)} \pm 0.016 \text{ (syst)} \text{ ps}^{-1}$$

. which is in good agreement with the world average value of $\Delta m_d = 0.509 \pm 0.004 \text{ ps}^{-1}$ [2].

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), A.P. Sloan Foundation, Research Corporation, Texas Advanced Research Program, Alexander von Humboldt Foundation, and the Marie Curie Fellowships.

-
- [1] E.Berger *et al.*, Phys. Rev. Lett. **86**, 4231 (2001).
 - [2] Heavy Flavor Averaging Group, arXiv:hep-ex/0505100.
 - [3] V. Abazov *et al.*, DØ collaboration, “The upgraded DØ Detector”, in preparation for submission to Nucl. Instrum. Methods Phys. Res. A;
 - [4] DØ Collab. Phys. Rev. Lett. **94**, 182001 (2005).
 - [5] The Road Method (an algorithm for the identification of electrons in jets), F. Beaudette and J.-F. Grivaz, DØ Note 3976, Nov 4 2002.
 - [6] DELPHI Collab., *b-tagging in DELPHI at LEP*, Eur. Phys. J. **C32**, 185 (2004).
 - [7] G. Borissov et al., Reconstructions of B hadron signals at DØ DØ Note 4481.
 - [8] G. Borissov, S. Burdin, A. Nomerotski, *B_d Flavor oscillations with opposite side muon tagging*, DØ Note 4370.
 - [9] S. Eidelman *et al.* (Particle Data Group), Phys. Lett. B **592**, 1 (2004).
 - [10] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995).
 - [11] M. Neubert, Phys. Rep. **245**, 259 (1994).

	default	variation		$\mathcal{D}(B^0)$		$\mathcal{D}(B^0)$		$\mathcal{D}(B^0)$		$\mathcal{D}(B^0)$		$\mathcal{D}(B^0)$	
				$0.1 < d < 0.2$		$0.2 < d < 0.3$		$0.3 < d < 0.45$		$0.45 < d < 0.6$		$0.6 < d < 1.0$	
		(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
$Br(B^0 \rightarrow D^{*-}\mu^+\nu)$	5.44	-0.23	0.23	—	—	—	-0.001	.001	—	.001	-0.001	.001	-0.001
$Br(B \rightarrow D^*\pi\mu\nu X)$	1.07	-0.17	0.17	.0004	-0.0004	-0.0011	.0011	-0.0019	.0021	-0.0020	.0021	-0.0008	.0028
R^{**}	0.35	.0	1.0	-0.0009	.0016	-0.0027	.0048	-0.0042	.0079	-0.0057	.0105	-0.0066	.0124
B lifetimes	.05022	-0.00054	.00054	—	-0.0001	.0001	-0.0002	.0003	-0.0001	.0003	-0.0003	.0014	-0.0003
Resolution function	—	$\div 1.2$	$\div 0.8$.0005	-0.0006	.0010	-0.0012	.0020	-0.0021	.0024	-0.0028	.0028	-0.0032
Alignment	—	-10 μm	10 μm	-0.004	0.004	-0.004	0.004	-0.004	0.004	-0.004	0.004	-0.004	0.004
K -Factor	—	-2%	+2%	—	—	-0.0001	—	—	.0001	-0.0001	—	—	—
Efficiency	—	-12%	+12%	.0006	-0.0007	-0.0008	.0006	-0.0012	.0011	-0.0013	.0010	-0.0021	.0019
Fraction D^0 in D^*	4%	3.15%	4.85%	—	.0010	-0.0010	—	-0.0010	.0010	-0.0010	.0010	-0.0010	.0010
Fit Procedure	See split below												
Bin width	2 MeV	1.6	2.67	-0.0026	.0002	-0.0024	.0014	-0.0001	.0027	.0037	.0038	.0089	.0087
Parameter μ_0	—	-3 σ	3 σ	-0.0003	.0002	.0001	-0.0001	.0001	.0001	-0.0002	.0001	-0.0007	.0007
Parameter $\frac{\sigma_R+\sigma_L}{2}$	—	-3 σ	3 σ	.0002	-0.0002	.0001	-0.0001	.0004	-0.0003	—	-0.0001	-0.0002	.0001
Parameter $\frac{\sigma_R-\sigma_L}{\sigma_R+\sigma_L}$	—	-3 σ	3 σ	-0.0005	.0005	.0002	-0.0001	.0002	.0001	-0.0002	.0001	-0.0015	.0011
Parameter μ_1	—	-3 σ	3 σ	-0.0009	.0010	-0.0017	.0018	.0023	-0.0015	.0006	-0.0005	-0.0004	-0.0004
Parameter $\frac{\sigma_1+\sigma_2}{2}$	—	-3 σ	3 σ	.0008	-0.0005	.0014	-0.0009	.0037	-0.0034	-0.0013	.0017	-0.0099	.0068
Parameter \bar{R}	—	-3 σ	3 σ	.0015	-0.0011	.0029	-0.0024	.0030	-0.0027	.0013	-0.0011	-0.0046	.0035
Parameter $(\mu_2 - \mu_1)$	—	-3 σ	3 σ	—	-0.0003	.0008	-0.0011	-0.0001	.0006	-0.0003	.0002	.0008	-0.0003
Parameter $\frac{\sigma_1-\sigma_2}{\sigma_1+\sigma_2}$	—	-3 σ	3 σ	-0.0001	—	-0.0004	.0003	.0002	-0.0002	-0.0004	.0004	-0.0006	.0010
Fit Procedure		Overall		+0.0021 -0.0031		+0.0040 -0.0041		+0.0060 -0.0046		+0.0044 -0.0019		+0.0119 -0.0111	
Total				+0.0049 -0.0052		+0.0077 -0.0066		+0.0111 -0.0081		+0.0125 -0.0081		+0.0182 -0.0140	

TABLE V: Systematic uncertainties $\mathcal{D}(B^0)$.