



Combined Opposite-side Flavor Tagging

The DØ Collaboration
URL: <http://www-d0.fnal.gov>
(Dated: July 19, 2005)

This note describes a combined opposite-side flavor tagging algorithm used by the DØ experiment. Various properties of the b quark on the opposite side of the reconstructed signal B hadron are combined together into a single variable which gives enhanced tagging power with respect to the usual simple single property tags. The combined tagging performance is tested in data using a large sample of reconstructed semileptonic $B \rightarrow \mu \bar{D}^0 X$ events. The tagging power is determined from this sample to be $\varepsilon \mathcal{D}^2 = 2.17 \pm 0.13 \pm 0.08\%$, and the measured B_d mixing parameter $\Delta m_d = 0.498 \pm 0.026$ (stat) ± 0.016 (syst) ps^{-1} is in a good agreement with the world average value. We also find that the dilutions of the combined tagger are the same, within statistical errors, for reconstructed B^+ and B^0 mesons.

I. INTRODUCTION

Flavor tagging is an essential ingredient for B meson oscillation and CP-violation measurements. Its performance is described by the combination of efficiency and dilution.

The efficiency ε is defined as the fraction of reconstructed events that are tagged:

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

N_{tag} is the number of tagged B mesons and N_{tot} is the total number of B mesons.

The tagging purity η is defined as:

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

where N_{cor} is the number of tagged B mesons with the correct initial flavor identification.

The dilution \mathcal{D} is related to the purity η by:

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

The figure of merit for a flavor tagged measurement is the tagging power, which is given by $\varepsilon\mathcal{D}^2$.

This note describes a combined opposite side flavor tagging algorithm and the measurement of its performance using reconstructed $B \rightarrow \mu^+\nu\bar{D}^0$ and $B \rightarrow \mu^+\nu D^{*-}$ events [1] collected by the DØ experiment in Run II. B^+ decays give the main contribution into the first sample, and B^0 decays dominate the second sample. The flavor tagging purity measured in B^0 decays depends on the B^0 decay length due to $B^0 - \bar{B}^0$ mixing, while the tagging purity in B^+ events remains constant. The B^0 oscillation frequency, Δm_d , has been measured with high precision elsewhere [2]. Using this value, the flavor tagging purity can be extracted directly from both data samples. Alternatively, the value of Δm_d can be measured and compared with the world average [2] to test the flavor tagging algorithm for a possible lifetime-dependent bias.

II. DETECTOR DESCRIPTION AND EVENT SELECTION

The RunII DØ detector is described in [3]. This analysis exploits the large semileptonic data sample corresponding to approximately 460 pb^{-1} of integrated luminosity, accumulated by the DØ experiment during the period from April 2002 to August 2004.

The $B \rightarrow \mu^+\nu\bar{D}^0 X$ with $\bar{D}^0 \rightarrow K^+\pi^-$ event sample is selected using criteria described in [4]. The same criteria are 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.

The mass spectrum of the $(K\pi)$ system in the D^0 sample is shown in Fig. 1. A fit of this distribution to the sum of a Gaussian, describing the signal, and a background function is also shown. The number of \bar{D}^0 candidates in the narrow peak is 81912 ± 511 . The wide peak at the lower mass corresponds to the partially reconstructed D meson decays.

The mass difference $\Delta M = M(\bar{D}^0\pi) - M(\bar{D}^0)$ for the events in the D^* sample is shown in Fig. 2. The peak, corresponding to the production of $\mu^+ D^{*-}$ events, is clearly seen. A fit to this distribution to the sum of two Gaussians describing the signal and a background function is also shown. The number of D^* candidates in the peak is equal to 39735 ± 341 .

DØ RunII Preliminary

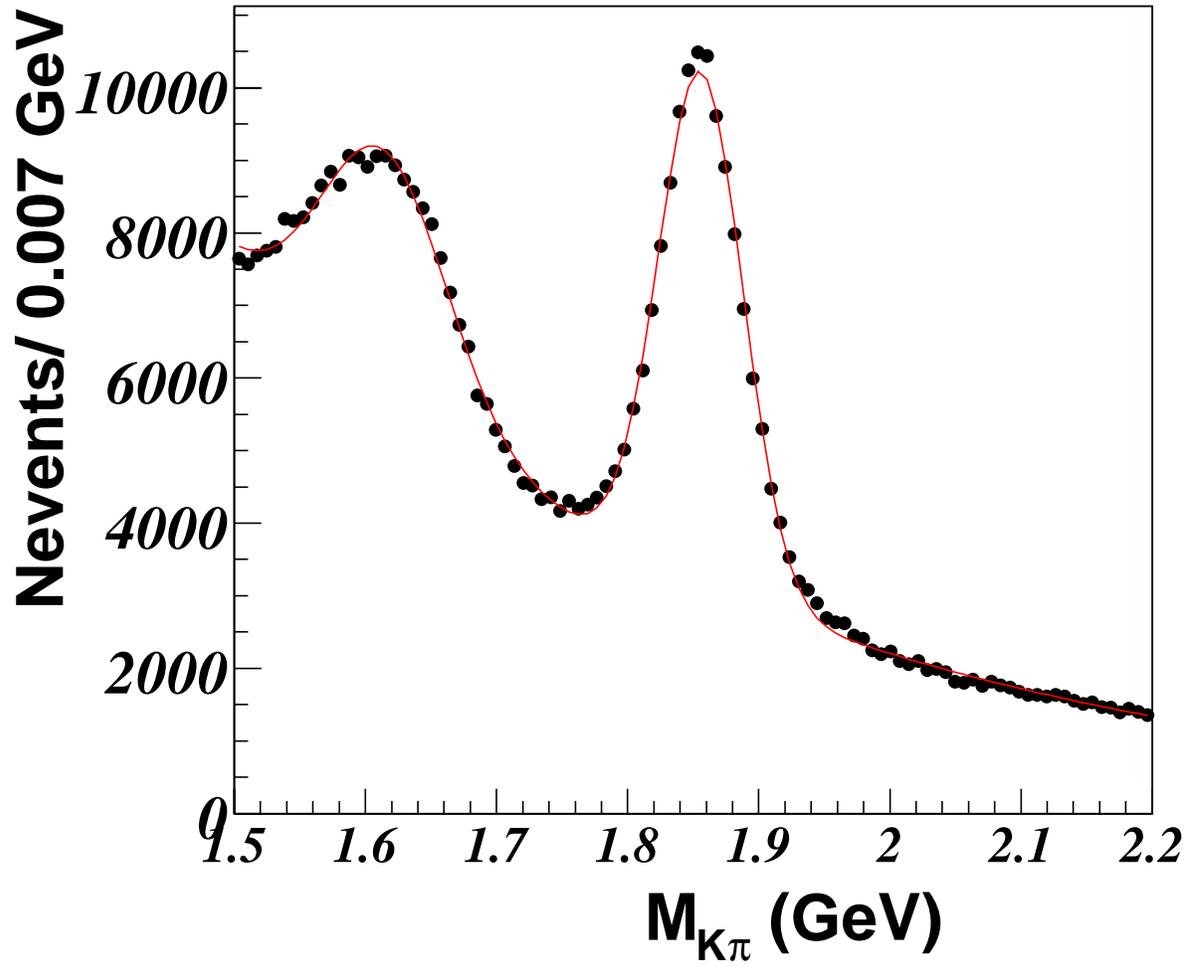


FIG. 1: The invariant mass of the $K\pi$ system for selected $\mu^+K^+\pi^-$ candidates. The curve shows the result of a fit of the distribution to a sum of a Gaussian, describing the signal, and a background function. The number of \bar{D}^0 candidates in the peak is 81912 ± 511 .

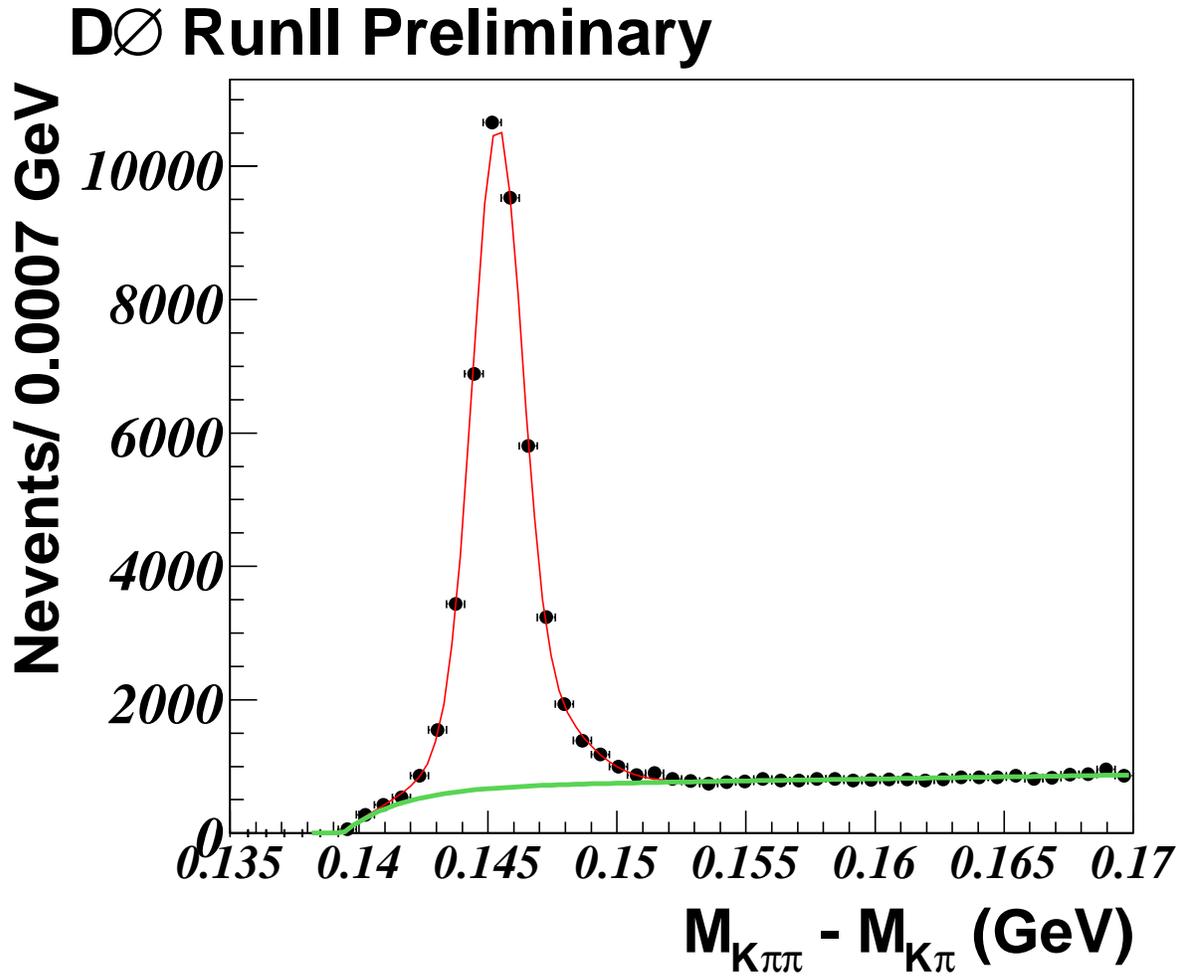


FIG. 2: The mass difference $M(D^0\pi) - M(D^0)$ for events with $1.75 < M(D^0) < 1.95 \text{ GeV}/c^2$. The curve shows the result of a fit of the distribution to a sum of 2 Gaussians, describing the signal, and a background function. The number of D^* candidates is 39735 ± 341 .

III. FLAVOR TAGGING METHOD

Many different event 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, while 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 using a 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 variables, by definition, should have different distributions for b and \bar{b} flavors. They can be either continuous, like the average charge of a jet, or discrete, like the charge of an electron or muon. For an initial b quark, the probability density function (PDF) for a given variable x_i is denoted as $f_i^b(x_i)$, while for an initial \bar{b} quark it is denoted as $f_i^{\bar{b}}(x_i)$. The combined tagging variable y is defined as:

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

Any given variable x_i can be undefined for some events. For example, there are events which do not contain an identified muon at 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 a b quark, and the event with $y > 1$ is tagged as a \bar{b} quark. For an oscillation analysis, it is more convenient to define the tagging variable as:

$$d = (1 - y)/(1 + y). \quad (5)$$

The variable d changes between -1 and 1. An event with $d > 0$ is tagged as a b quark and with $d < 0$ as a \bar{b} quark, and larger $|d|$ values correspond to a higher tagging purity. For uncorrelated variables x_1, \dots, x_n and perfect modeling in the PDF, d is the best possible discriminant and its absolute value gives the 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”). In high energy particle collisions heavy quarks are produced in particle/antiparticle pairs. The flavor of the opposite side, therefore, determines the initial flavor of the reconstructed B meson. 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 quarks is not correlated in $p\bar{p}$ interactions. The flavor tagging algorithm can, therefore, be calibrated with data by applying it to the B^0 and B^+ event samples, and then used to study such quantities as the B_s meson oscillations.

Another set of variables, which exploit the properties of the hadronization of the reconstructed B meson, can also be used for flavor tagging (“same side tagging”). Flavor tagging with these variables depends on the type of reconstructed B meson, and its performance can, therefore, only be obtained from simulation which introduces significant uncertainties. Because of this, same side tagging is not currently used for our B_s mixing measurement, nor is it described further in this note.

The probability density functions for each discriminating variable used in this tagging algorithm are constructed using a real data sample of $B \rightarrow \mu^+ \nu D^{*-}$ events which have a visible proper decay length less than $500 \mu\text{m}$. The definition of the visible proper decay length is given in section IV. Events with $0.143 < M(D^0\pi) - M(D^0) < 0.148 \text{ GeV}/c^2$ and the correct charge correlation between μ and π were selected for signal distributions. The background under the D^* peak is subtracted in each distribution using the $\mu^+ \bar{D}^0 \pi^+$ events with the wrong combination of charges of muon and pion. In the selected sample, the non-oscillating decays $B^0 \rightarrow \mu^+ \nu D^{*-}$ dominate and the initial state of the b -quark is determined by the charge of the signal muon. According to MC estimates, the purity of such identification in the selected sample is 0.956 ± 0.007 , where the error reflects the uncertainty in branching ratios of the decays involved.

A. Opposite Side Muon Tagging

Opposite side muons are used for flavor tagging if they have $\cos\phi(\mathbf{p}_\mu, \mathbf{p}_B) < 0.8$, where \mathbf{p}_B is the three-momentum of the reconstructed B hadron. If more than one muon is found with this condition, the muon with the highest number of hits in the muon chambers is used. If both muons have the same number of hits, the muon with the highest p_T is used.

For each such muon, a *muon jet charge* is constructed:

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

The sum is 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 $\Delta\eta$ are computed with respect to the muon direction. Daughters of the reconstructed B hadron are explicitly excluded from the sum. The distribution of the muon jet charge variable is shown in Fig. 3a,b.

Another discriminating variable associated with the muon, is its *transverse momentum relative to the nearest jet*, p_T^{rel} . Jets are defined from reconstructed tracks using the DURHAM clustering algorithm [5] with a cut-off parameter of 15 GeV/c [6]. The muon is included in the jet clustering. Similarly to the muon jet charge variable, only muons with $\cos\phi(\mathbf{p}_\mu, \mathbf{p}_B) < 0.8$ are used to construct p_T^{rel} . Muons from the cascade $b \rightarrow c \rightarrow \mu$ decay have the same charge as the b quark from the reconstructed side (b^{rec}), i.e $q(\mu) \times q(b^{rec}) > 0$, and give a wrong tagging of the initial flavor, decreasing the tagging performance. Such muons, in average, have a smaller p_T^{rel} value, compared to the muons from direct $b \rightarrow \mu$ decays. This property is actually included in the flavor tagging algorithm by the addition of p_T^{rel} as a discriminating variable. Normalized distributions of the p_T^{rel} for events with $q(\mu) \times q(b^{rec}) > 0$ and $q(\mu) \times q(b^{rec}) < 0$ are shown in Fig. 4c,d. The softer muon p_T^{rel} spectrum for $q(\mu) \times q(b^{rec}) > 0$ is clearly seen.

B. Opposite Side Electron Tagging

Opposite side electrons with $\cos\phi(\mathbf{p}_e, \mathbf{p}_B) < 0.5$ are used for flavor tagging, with the charge of the electron used as a discriminating variable [7].

C. Opposite Secondary Vertex and Event Charge Tagging

Opposite side secondary vertices are used for flavor tagging. The details of the vertex finding algorithm can be found in [8]. The secondary vertex is required to have at least two particles associated with it that have a transverse impact parameter significance greater than 3. The distance l_{xy} from the primary to the secondary vertex must satisfy the condition $l_{xy} > 4\sigma(l_{xy})$. The momentum of the secondary vertex \mathbf{p}_{SV} is defined as the sum of all momenta of particles associated to the secondary vertex, and only secondary vertices with $\cos\phi(\mathbf{p}_{SV}, \mathbf{p}_B) < 0.8$ are used in the flavor tagging. In addition, secondary vertices containing any particle from the decay of the reconstructed B hadron are excluded from the tagging. The *secondary vertex charge* Q_{SV} is used as a discriminating variable:

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

where the sum is 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. Fig. 4a,b show the distribution of this variable for events with and without an identified muon.

The *transverse momentum of the secondary vertex* p_T^{SV} is also used as the discriminating variable. Events with fake vertices are not sensitive to the charge of the B meson from the reconstructed side. Their contribution decreases the tagging purity. Usually, they are constructed from the low momentum tracks and their p_T^{SV} is softer. The distribution of $\log_{10}(p_T^{SV})$ is shown in Fig. 4c. The events with $q(Q_{SV}) \times q(b^{rec}) > 0$ have a softer p_T^{SV} distribution due to a larger fraction of fake vertices in this sample. For the flavor tagging, we use the ratio of PDF for events with $q(Q_{SV}) \times q(b^{rec}) > 0$ and $q(Q_{SV}) \times q(b^{rec}) < 0$. Including such a ratio into combined tagging improves its purity, although the discriminating power of this variable is quite small as can be seen in Fig. 4c.

Finally, the *event charge*

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

is also used for flavor tagging. The sum is 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 $\Delta\eta$ are computed with respect to the reconstructed B -hadron direction. Due to a strong correlation with the muon jet charge, this variable is not used for events with the identified muon. The distribution of this variable is shown in Fig. 4d.

D. The Combined Tagger

If a muon is found, the muon jet charge, muon p_T^{rel} and secondary vertex jet charge are used to construct a *muon tagger*. For events without a muon but with an identified electron, the electron charge is used to construct an *electron tagger*.

tagger. Finally, for events without a muon or an electron, the secondary vertex jet charge, the p_T^{SV} and the event jet charge are used to construct the secondary vertex tagger. The resulting distribution of the tagging variable d for the combination of all three taggers, called the *combined tagger*, is shown in Fig. 5.

DØ RunII Preliminary

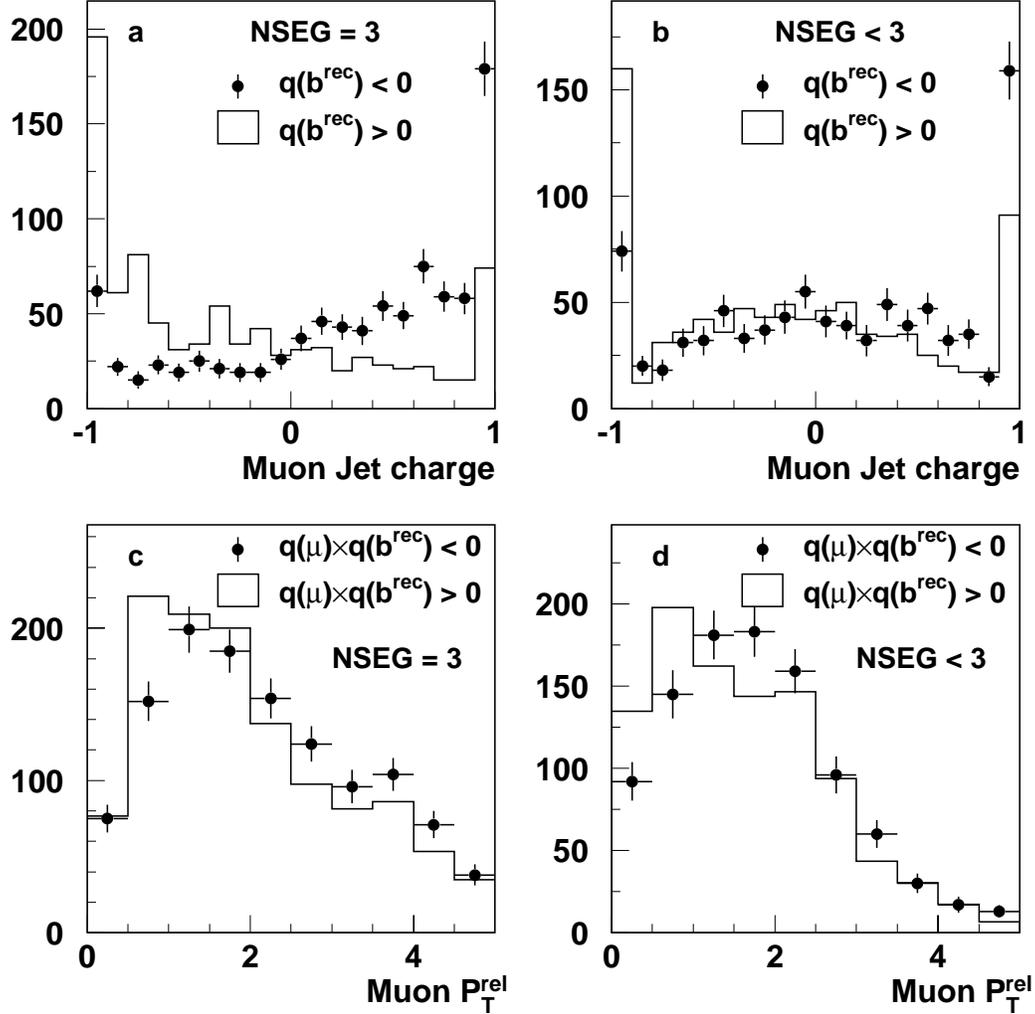


FIG. 3: Normalized distributions of muon jet charge and the p_T^{rel} , $q(b^{\text{rec}})$ is the charge of the b quark from the reconstruction side. Figures a) and c) correspond to the muon with hits in 3 layers of the muon detector and Figures b) and d) correspond to the muon with hits in less than 3 layers of the muon detector.

DØ RunII Preliminary

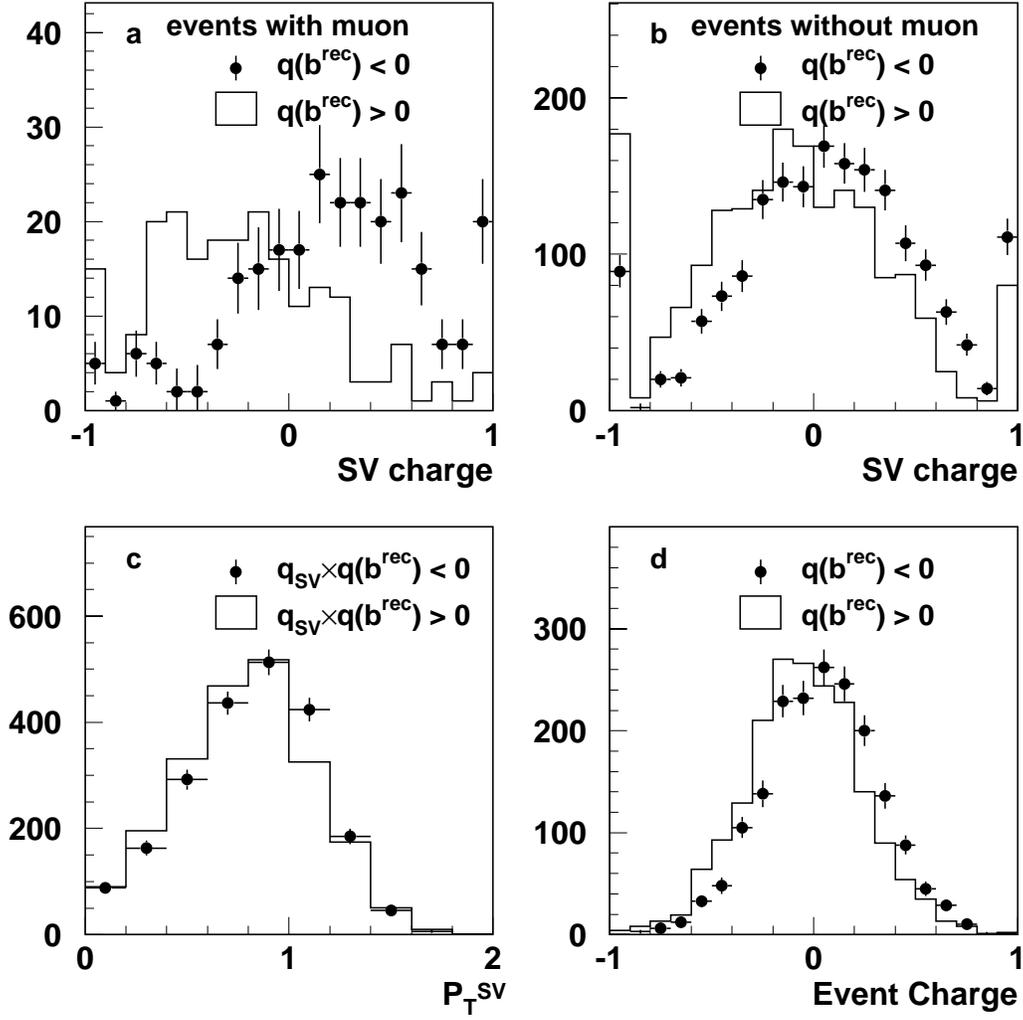


FIG. 4: Normalized distribution of Q_{SV} , p_T^{SV} and Q_{EV} , $q(b^{\text{rec}})$ is the charge of the b quark from the reconstruction side.

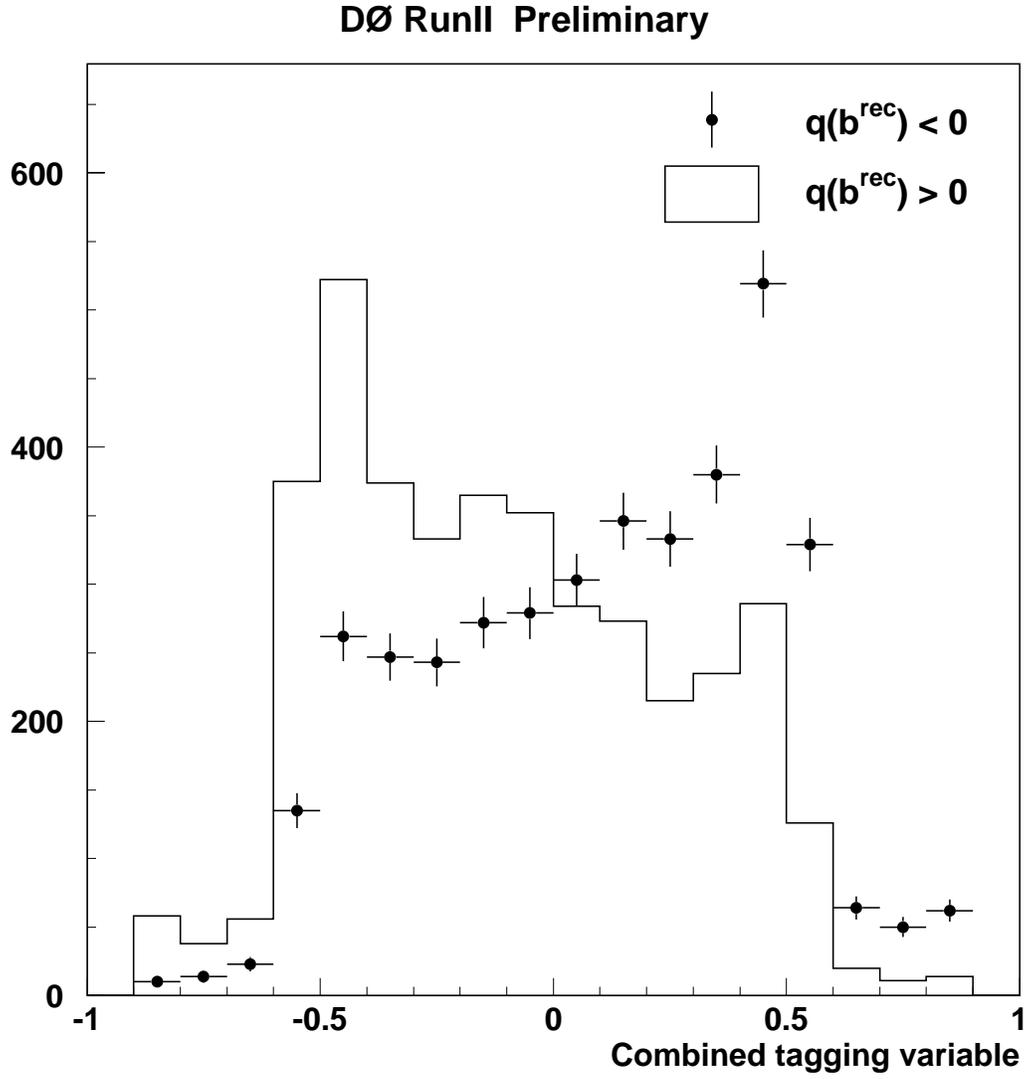


FIG. 5: Normalized distributions of the combined tagging variable, $q(b^{\text{rec}})$ is the charge of the b quark from the reconstruction side.

IV. B_d^0 MIXING MEASUREMENT AND TAGGER CALIBRATION

The visible proper decay length (VPDL) is defined as:

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

The \mathbf{L}_{xy} is the vector in the transverse plane from the primary to the B -meson decay vertex, and $\mathbf{P}_T^{\mu\bar{D}^0}$ is the vector sum of transverse momenta of muon and D^0 .

$B \rightarrow \mu\bar{D}^0 X$ events in which the signal muon has the opposite sign of the flavor tagging variable d are considered non-oscillating, and events with the d having the same sign of the signal muon charge are tagged as oscillating.

All events are divided into 7 groups according to the measured VPDL. The number of oscillating N_i^{osc} and non-oscillating N_i^{nos} μD^* events in each interval i is 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 oscillating N_i^{osc} and non-oscillating N_i^{nos} $\mu^+ \bar{D}^0$ events is determined from a fit of the \bar{D}^0 signal in the $K\pi$ invariant mass distribution for the D^0 sample.

The performance of the flavor taggers is determined by fitting the flavor asymmetry

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

of the D^0 and D^* samples in each VPDL bin.

The same procedure as described in [9] is used to fit the measured asymmetries A_i and determine the values of Δm_d , tagging purity η_{B^0} of events with reconstructed B^0 and the tagging purity η_{B^+} of events with reconstructed B^+ . The composition of D^0 and D^* samples, efficiencies to reconstruct D^0 and soft pion as well as K -factors for various B -meson decays are obtained as in [4].

The asymmetries are determined and fit for events using the combined tagger, and for the muon, electron and secondary vertex taggers separately. For each tagger, the variable d is computed using (4,5) and events with $|d| > 0.3$ are used. Four samples of events having $|d| > 0.3$, $0.22 < |d| < 0.3$, $0.3 < |d| < 0.45$, and $|d| > 0.45$ are used for the combined tagger study.

Figures 6 and 7 show the measured asymmetry for different taggers. A significant oscillation pattern for D^* events, and a reduced oscillation for D^0 events are clearly seen. These distributions are described reasonably well by the expected oscillation functions, which are superimposed in these figures.

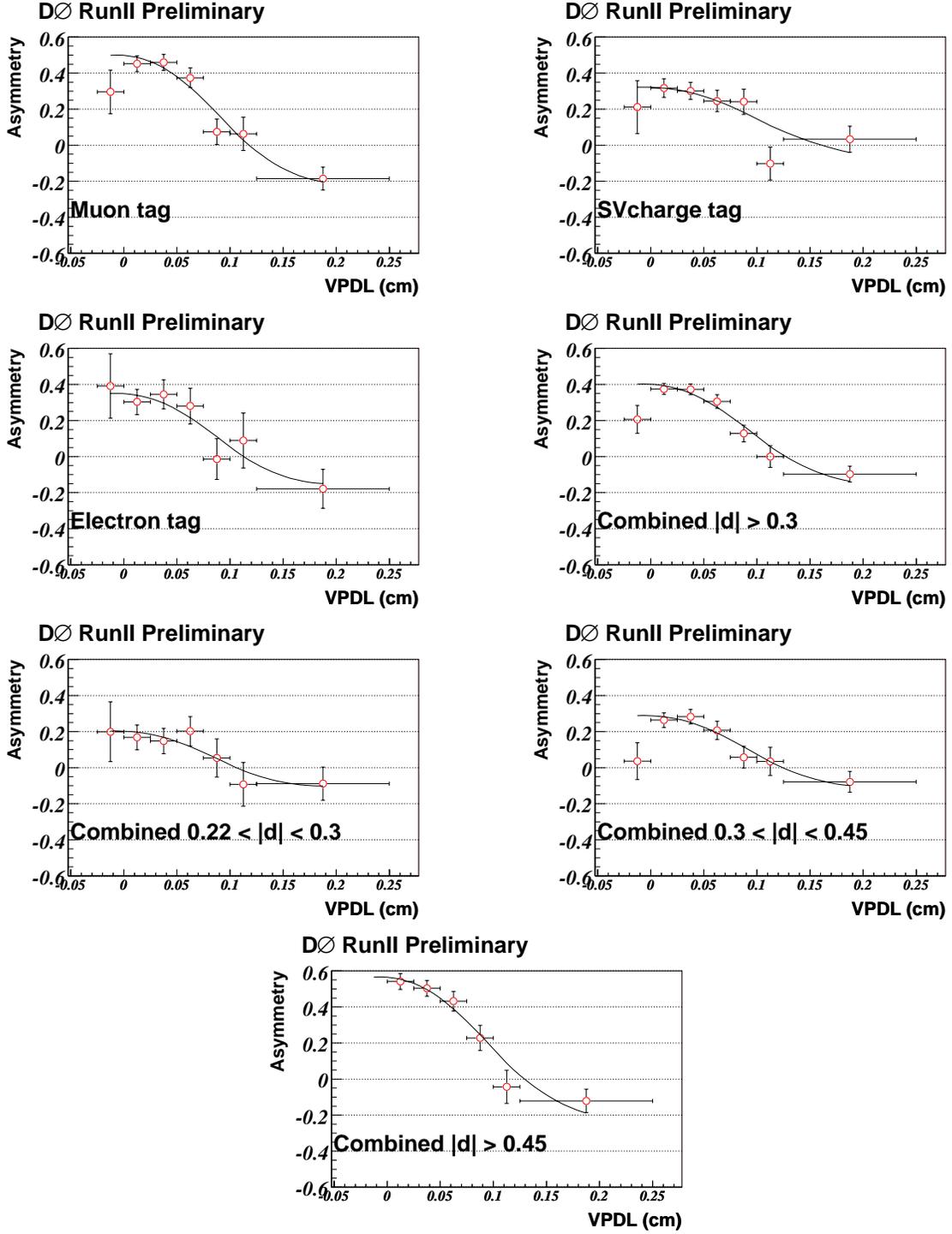


FIG. 6: The asymmetries obtained in the D^* sample with the result of the fit superimposed. For the individual taggers, $|d| > 0.3$ is required.

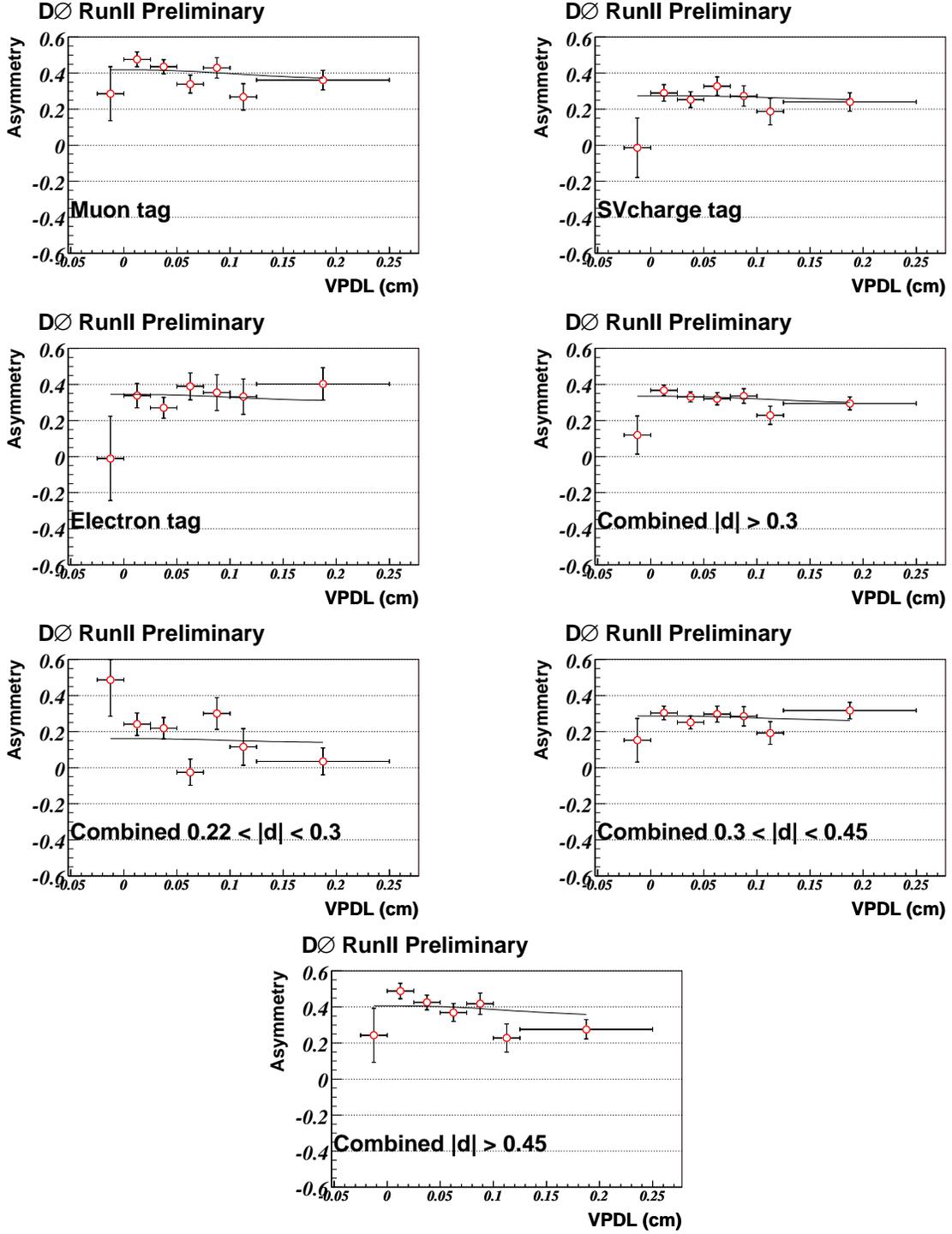


FIG. 7: The asymmetries obtained in the D^0 sample with the result of the fit superimposed. For the individual taggers, $|d| > 0.3$ is required.

V. RESULTS

B_d oscillations are clearly seen using the combined tagger (and in each tagger separately) showing that the algorithm is working well. The fitted values of Δm_d are given in table I for the cases where the B^+ and B^0 tagging purities are allowed to be different, and when they are constrained to be equal. Combining results from the three non-overlapping samples with $0.22 < |d| < 0.3$, $0.3 < |d| < 0.45$, and $|d| > 0.45$, a value of $\Delta m_d = 0.498 \pm 0.026$ (stat) is obtained, which is in a good agreement with the world average value of $\Delta m_d = 0.509 \pm 0.004$ [2].

One of the purposes of this measurement is to extract the flavor tagging efficiencies (shown in Table II) and dilutions for these taggers on reconstructed B_d and B_u mesons and to validate the assumption that the dilutions do not depend on the type of signal B meson.

We consider four different cases:

- Δm_d , η_{B^0} and η_{B^+} are free parameters.
- Δm_d , η_{B^0} and η_{B^+} are free parameters, with $\eta_{B^0} = \eta_{B^+}$.
- η_{B^0} and η_{B^+} are free parameters, Δm_d is fixed to the world average $0.502 \pm 0.005 \text{ps}^{-1}$.
- Δm_d is fixed to the world average and η_{B^0} and η_{B^+} are free parameters, with $\eta_{B^0} = \eta_{B^+}$.

Table III shows the measured dilutions for all these cases and for each considered tagger. The dilutions for B^+ and B^0 are consistent within the statistical error. Using the combined tagging variable $|d| > 0.3$, we measure a B_d mixing parameter of $\Delta m_d = 0.501 \pm 0.030$ and a dilution of $\mathcal{D} = 0.384 \pm 0.013$.

The tagging power $\varepsilon \mathcal{D}^2$ for different taggers is shown in Table IV. The background processes, like the $c\bar{c}$ production or the misidentification of the muon, tend to decrease the observed tagging efficiency. These processes contribute mainly to events with small VPDL. Therefore for computation of the tagging power, the tagging efficiency for events with the VPDL=[0.025,0.250] is used. The value of dilution is taken from the fit where the η_{B^+} and η_{B^0} are constrained to be equal and the Δm_d is fixed to the world average. For events with $|d| > 0.3$, the $\varepsilon \mathcal{D}^2 = 1.94 \pm 0.14\%$ is obtained. Selecting separately events with $0.22 < |d| < 0.3$ and $|d| > 0.3$ and adding their tagging power, a total $\varepsilon \mathcal{D}^2 = 2.03 \pm 0.13\%$ is obtained. Selecting separately events with $0.22 < |d| < 0.3$, $0.3 < |d| < 0.45$ and $|d| > 0.45$ and adding their tagging power, a total $\varepsilon \mathcal{D}^2 = 2.17 \pm 0.13\%$ is obtained.

Tagger	Δm_d	Δm_d (Purities constrained)
Muon	0.526 ± 0.037	0.520 ± 0.039
SV Charge	0.426 ± 0.055	0.421 ± 0.057
Electron	0.561 ± 0.088	0.563 ± 0.086
Combined($ d > 0.3$)	0.501 ± 0.030	0.498 ± 0.031
Combined($0.22 < d < 0.3$)	0.582 ± 0.131	0.578 ± 0.137
Combined($0.3 < d < 0.45$)	0.519 ± 0.057	0.522 ± 0.055
Combined($ d > 0.45$)	0.489 ± 0.030	0.479 ± 0.034

TABLE I: Fit results for Δm_d with B^0 and B^+ purities floating independently and for B^0 and B^+ purities constrained to be equal.

Tag Type	$N_{tag}(D^*)$	$\varepsilon(\%)(D^*)$	$N_{tag}(D^0)$	$\varepsilon(\%)(D^0)$
Muon	2085 ± 55	5.25 ± 0.14	4270 ± 112	5.21 ± 0.14
Secondary Vertex Charge	1916 ± 52	4.82 ± 0.13	3737 ± 119	4.56 ± 0.15
Electron	816 ± 37	2.05 ± 0.09	1650 ± 89	2.01 ± 0.11
Combined ($ d > 0.3$)	4810 ± 83	12.10 ± 0.21	9633 ± 183	11.76 ± 0.22
Combined ($0.22 < d < 0.3$)	1063 ± 39	2.67 ± 0.10	2041 ± 108	2.49 ± 0.13
Combined ($0.3 < d < 0.45$)	2837 ± 63	7.14 ± 0.16	5470 ± 138	6.68 ± 0.17
Combined ($ d > 0.45$)	1966 ± 50	4.95 ± 0.12	4183 ± 123	5.11 ± 0.13

TABLE II: Tag efficiencies obtained from a fit to the total tagged sample in all VPDL bins for the different taggers.

Tag	$\mathcal{D}(B^0), \mathcal{D}(B^+)$ Δm_d floating	$\mathcal{D}(B^0) = \mathcal{D}(B^+)$ Δm_d floating	$\mathcal{D}(B^0), \mathcal{D}(B^+)$ Δm_d fixed	$\mathcal{D}(B^0) = \mathcal{D}(B^+)$ Δm_d fixed
Muon	0.513 ± 0.035	-	0.507 ± 0.034	-
	0.459 ± 0.024	0.477 ± 0.019	0.456 ± 0.024	0.474 ± 0.018
SV Charge	0.329 ± 0.037	-	0.367 ± 0.037	-
	0.302 ± 0.025	0.311 ± 0.020	0.307 ± 0.025	0.317 ± 0.019
Electron	0.354 ± 0.060	-	0.346 ± 0.058	-
	0.384 ± 0.037	0.375 ± 0.030	0.380 ± 0.037	0.370 ± 0.029
Combined $ d > 0.3$	0.413 ± 0.024	-	0.414 ± 0.023	-
	0.367 ± 0.017	0.383 ± 0.013	0.368 ± 0.016	0.384 ± 0.013
Combined $0.22 < d < 0.3$	0.210 ± 0.054	-	0.203 ± 0.053	-
	0.176 ± 0.035	0.189 ± 0.028	0.173 ± 0.035	0.183 ± 0.027
Combined $0.3 < d < 0.45$	0.292 ± 0.032	-	0.290 ± 0.031	-
	0.320 ± 0.021	0.310 ± 0.017	0.319 ± 0.021	0.309 ± 0.016
Combined $ d > 0.45$	0.589 ± 0.035	-	0.592 ± 0.034	-
	0.441 ± 0.025	0.492 ± 0.018	0.441 ± 0.024	0.496 ± 0.019

TABLE III: Dilutions for Δm_d floating and Δm_d fixed to the world average. For columns 3 and 5, the purities for the B^+ and B^0 are constrained to be equal.

Tagger	$\varepsilon(\%)$	\mathcal{D}	$\varepsilon\mathcal{D}^2(\%)$
Muon	5.64 ± 0.16	0.474 ± 0.018	1.27 ± 0.13
Electron	2.03 ± 0.10	0.370 ± 0.029	0.28 ± 0.06
SVCharge	5.42 ± 0.16	0.317 ± 0.019	0.54 ± 0.08
Combined ($ d > 0.3$)	13.13 ± 0.24	0.384 ± 0.013	1.94 ± 0.14
Combined($0.22 < d < 0.3$)	2.82 ± 0.11	0.183 ± 0.027	0.094 ± 0.028
Combined($0.3 < d < 0.45$)	7.65 ± 0.19	0.309 ± 0.016	0.730 ± 0.077
Combined($ d > 0.45$)	5.45 ± 0.15	0.496 ± 0.018	1.341 ± 0.100

TABLE IV: Tagging power of individual taggers and the combined tagger in 3 bins.

VI. SYSTEMATIC UNCERTAINTIES

The systematic uncertainties of these measurements are summarized in Table V, which shows the variation of Δm_d and dilution for the cases where the B^0 and B^+ dilutions are independent, and where they are constrained to be equal. The estimate of different systematic effects is described below.

The B meson branching rates and lifetimes used in the fit of the asymmetry are taken from [10] and are varied by one standard deviation.

The VPDL resolution, obtained in simulation, is multiplied by a large factor, from 0.2 to 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 is neglected in this analysis. To check the impact of this assumption on the final result, their computation is 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 is used as the estimate of the systematic uncertainty in their values. This uncertainty is 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 [11] of the semileptonic B decays is used. The uncertainty of the reconstruction efficiency, set at 12%, is estimated by varying kinematic cuts on p_t of the muon and D^0 in a wide range. Changing the model describing semileptonic B decay from *ISGW2* to *HQET* [12] produces a smaller variation. The fit of asymmetry is 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 is taken as the systematic uncertainty from this source.

Possible peaking background, i.e. events with (μD^0) not coming from $B \rightarrow \mu^+ \bar{D}^0 X$ decays, can bias the measurement of the mixing parameter. Such events come from e.g. the $c\bar{c}$ contribution or the misidentification of the muon. The contribution of this background is varied from 3.5% to 10% and the difference in the result is taken as the systematic uncertainty from this source.

We also investigated the systematic uncertainty of measuring the number of D^* and D^0 candidates in each VPDL bin. We changed the background parametrization for the D^0 mass fit and varied the parameters of background

parametrization within one standard deviation. Since the background contribution in the D^* sample is small, its influence on the estimate of the number of D^* candidates is reduced. We performed cross-checks using other functions, but the chosen description gives the best description. We varied the parameters of the background parametrization by one standard deviation. We also varied the bin width of histograms used to measure the number of D^* and D^0 events. The results of these studies are presented in Table V.

	variation	$\delta(\Delta m_d)$	$\delta(\mathcal{D}(B^0))$	$\delta(\mathcal{D}(B^+))$	$\delta(\mathcal{D})$
$Br(B^0 \rightarrow D^{*-} \mu^+ \nu)$	$5.53 \pm 0.023\%$	0.002 ps^{-1}	0.001	0.001	0.001
$Br(B \rightarrow D^* \pi \mu \nu X)$	$1.07 \pm 0.17\%$	0.008 ps^{-1}	0.002	0.001	0.001
B lifetimes	$\pm 1\sigma$	0.001 ps^{-1}	0.000	0.000	0.000
Resolution function	$\times [0.2 \div 2]$	0.005 ps^{-1}	0.002	0.000	0.000
Alignment	$\pm 10 \mu\text{m}$	0.007 ps^{-1}	0.004	0.000	0.004
K -factor	$\pm 2\%$	0.009 ps^{-1}	0.000	0.000	0.000
Peaking background	$[0.035 \div 0.1]$	0.002 ps^{-1}	0.002	0.000	0.002
Efficiency	$\pm 12\%$	0.006 ps^{-1}	0.001	0.001	0.001
Fit procedure	see text	${}_{-0.004}^{+0.002} \text{ ps}^{-1}$	0.016	0.007	0.006
Total		0.016 ps^{-1}	0.017	0.008	0.008

TABLE V: Systematic uncertainties.

VII. CONCLUSIONS

We have developed a likelihood-based opposite-side flavor tagging algorithm which has been calibrated using our large semileptonic B^0 and B^+ samples. The dilutions of the combined tagger $\mathcal{D}(B^+)$ and $\mathcal{D}(B^0)$ are consistent within statistical errors. By dividing the sample into small bins in the combined variable d , the following tagging power and Δm_d values are obtained:

$$\varepsilon \mathcal{D}^2 = 2.17 \pm 0.13 \text{ (stat)} \pm 0.09 \text{ (syst)} \% \quad (8)$$

$$\Delta m_d = 0.498 \pm 0.026 \text{ (stat)} \pm 0.016 \text{ (syst)} \text{ ps}^{-1}. \quad (9)$$

For the cut $|d| > 0.3$, used in the B_s mixing studies, the following values are obtained:

$$\mathcal{D}(B^0) = 0.414 \pm 0.023 \text{ (stat)} \pm 0.017 \text{ (syst)} \quad (10)$$

$$\mathcal{D}(B^+) = 0.368 \pm 0.016 \text{ (stat)} \pm 0.008 \text{ (syst)} \quad (11)$$

$$\mathcal{D} = 0.384 \pm 0.013 \text{ (stat)} \pm 0.008 \text{ (syst)} \quad (\mathcal{D}(B^0) = \mathcal{D}(B^+)) \quad (12)$$

$$\varepsilon \mathcal{D}^2 = 1.94 \pm 0.14 \text{ (stat)} \pm 0.09 \text{ (syst)} \% \quad (13)$$

$$\Delta m_d = 0.501 \pm 0.030 \text{ (stat)} \pm 0.016 \text{ (syst)} \text{ ps}^{-1} \quad (14)$$

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

The measured dilution $\mathcal{D} = 0.384 \pm 0.013 \text{ (stat)} \pm 0.008 \text{ (syst)}$ will be used as the central value in the B_s mixing study, and the variation of dilution between B^+ and B^0 events will be used for the systematic error estimate in the B_s mixing measurement.

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] Charge conjugate states are always implied in this note.

- [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] S. Catani *et al.*, Phys. Lett. B **269**, 432 (1991).
- [6] T. Sjöstrand *et al.*, Comp. Phys. Commun. **135**, 238 (2001).
- [7] DØ Collaboration, *B flavor tagging with soft electrons*, Conference Note 4848.
- [8] DELPHI Collab., *b-tagging in DELPHI at LEP*, Eur. Phys. J. **C32** (2004), 185-208
- [9] DØ Collaboration, *Flavor oscillations in Bd mesons with opposite side muon tagging*, Conference Note 4370, 2004.
- [10] S. Eidelman *et al.* (Particle Data Group), Phys. Lett. B **592**, 1 (2004).
- [11] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995). N. Isgur, D. Scora, B. Grinstein and M.B. Wise, Phys.Rev. D **39**, 799 (1989).
- [12] M. Neubert, Phys. Rep. **245**, 259 (1994).