12-13-2006

Measurement of B_d mixing using opposite-side flavor tagging

V.M. Abazov
Joint Institute for Nuclear Research, Dubna, Russia

Kenneth A. Bloom
University of Nebraska - Lincoln, kbloom2@unl.edu

Gregory Snow
University of Nebraska-Lincoln, gsnow1@unl.edu

D0 Collaboration

Measurement of B_d mixing using opposite-side flavor tagging

V. M. Abazov,[35] B. Abbott,[75] M. Abolins,[65] B. S. Acharya,[28] M. Adams,[51] T. Adams,[49] M. Agelou,[17] E. Aguilo,[5] S. H. Ahn,[30] M. Ahsan,[59] G. D. Alexeev,[35] G. Alkhazov,[39] A. Alton,[64] G. Alverson,[63] G. A. Alves,[2] M. Anastasoaie,[34] T. Andeen,[53] S. Anderson,[45] B. Andrieu,[16] M. S. Anzelc,[53] Y. Arnoud,[13] M. Arov,[52] A. Askew,[49] B. Åsman,[40] A. C. S. Assis Jesus,[3] O. Atramentov,[49] C. Autermann,[20] C. Avila,[7] C. Ay,[23] F. Badaud,[12] A. Baden,[61] L. Bagby,[52] B. Baldin,[50] D. V. Bandurin,[59] P. Banerjee,[28] S. Banerjee,[28] E. Barberis,[63] P. Bargassa,[80] P. Baringer,[58] C. Barnes,[43] J. Barreto,[2] J. F. Bartlett,[50] U. Bassler,[16] D. Bauer,[43] S. Beale,[5] A. Bean,[58] M. Begalli,[3] M. Begel,[71] C. Belanger-Champagne,[5] L. Bellantoni,[50] A. Bellavance,[67] J. A. Benitez,[65] S. B. Beri,[26] G. Bernardi,[16] R. Bernhard,[41] L. Berntzon,[14] I. Bertram,[42] M. Besançon,[17] R. Beuselinck,[43] V. A. Bezzubov,[38] P. C. Bhat,[50] V. Bhatnagar,[26] M. Binder,[24] C. Biscarat,[42] K. M. Black,[62] I. Blackler,[43] G. Blazey,[52] F. Blekman,[43] S. Blessing,[49] D. Bloch,[18] K. Bloom,[67] U. Blumenschein,[22] A. Boehnlein,[50] O. Boeriu,[55] T. A. Bolton,[59] G. Borissov,[42] K. Bos,[33] T. Bose,[77] A. Brandt,[78] R. Brock,[65] G. Brooijmans,[70] A. Bross,[50] D. Brown,[78] N. J. Buchanan,[49] D. Buchholz,[53] M. Buehler,[81] V. Buescher,[22] S. Burdin,[50] S. Burke,[45] T. H. Burnett,[82] E. Busato,[16] C. P. Buszello,[43] J. M. Butler,[62] P. Calfayan,[24] S. Calvet,[14] J. Cammin,[71] S. Caron,[33] W. Carvalho,[3] B. C. K. Casey,[77] N. M. Cason,[55] H. Castilla-Valdez,[32] S. Chakrabarti,[28] D. Chakraborty,[52] K. M. Chan,[71] A. Chandra,[48] F. Charles,[18] E. Cheu,[45] F. Chevallier,[13] D. K. Cho,[62] S. Choi,[31] B. Choudhary,[27] L. Christofek,[77] D. Claes,[67] B. Clément,[18] C. Clément,[40] Y. Coadou,[5] M. Cooke,[80] W. E. Cooper,[50] D. Coppage,[58] M. Corcoran,[80] M.-C. Cousinou,[14] B. Cox,[44] S. Crépé-Renaudin,[13] D. Cutts,[77] M. Ćwiok,[29] H. da Motta,[2] A. Das,[62] M. Das,[60] B. Davies,[42] G. Davies,[43] G. A. Davis,[53] K. De,[78] P. de Jong,[33] S. J. de Jong,[34] E. De La Cruz-Burelo,[64] C. De Oliveira Martins,[3] J. D. Degenhardt,[64] F. Déliot,[17] M. Demarteau,[50] R. Demina,[71] P. Demine,[17] D. Denisov,[50] S. P. Denisov,[38] S. Desai,[72] H. T. Diehl,[50] M. Diesburg,[50] M. Doidge,[42] A. Dominguez,[67] H. Dong,[72] L. V. Dudko,[37] L. Duflot,[15] S. R. Dugad,[28] D. Duggan,[49] A. Duperrin,[14] J. Dyer,[65] A. Dyshkant,[52] M. Eads,[67] D. Edmunds,[65] T. Edwards,[44] J. Ellison,[48] J. Elmsheuser,[24] V. D. Elvira,[50] S. Eno,[61] P. Ermolov,[37] H. Evans,[54] A. Evdokimov,[36] V. N. Evdokimov,[38] S. N. Fatakia,[62] L. Feligioni,[62] A. V. Ferapontov,[59] T. Ferbel,[71] F. Fiedler,[24] F. Filthaut,[34] W. Fisher,[50] H. E. Fisk,[50] I. Fleck,[22] M. Ford,[44] M. Fortner,[52] H. Fox,[22] S. Fu,[50] S. Fuess,[50] T. Gadfort,[82] C. F. Galea,[34] E. Gallas,[50] E. Galyaev,[55] C. Garcia,[71] A. Garcia-Bellido,[82] J. Gardner,[58] V. Gavrilov,[36] A. Gay,[18] P. Gay,[12] D. Gelé,[18] R. Gelhaus,[48] C. E. Gerber,[51] Y. Gershtein,[49] D. Gillberg,[5] G. Ginther,[71] N. Gollub,[40] B. Gómez,[7] A. Goussiou,[55] P. D. Grannis,[72] H. Greenlee,[50] Z. D. Greenwood,[60] E. M. Gregores,[4] G. Grenier,[19] Ph. Gris,[12] J.-F. Grivaz,[15] S. Grünendahl,[50] M. W. Grünewald,[29] F. Guo,[72] J. Guo,[72] G. Gutierrez,[50] P. Gutierrez,[75] A. Haas,[70] N. J. Hadley,[61] P. Haefner,[24] S. Hagopian,[49] J. Haley,[68] I. Hall,[75] R. E. Hall,[47] L. Han,[6] K. Hanagaki,[50] P. Hansson,[40] K. Harder,[59] A. Harel,[71] R. Harrington,[63] J. M. Hauptman,[57] R. Hauser,[65] J. Hays,[53] T. Hebbeker,[20] D. Hedin,[52] J. G. Hegeman,[33] J. M. Heinmiller,[51] A. P. Heinson,[48] U. Heintz,[62] C. Hensel,[58] K. Herner,[72] G. Hesketh,[63] M. D. Hildreth,[55] R. Hirosky,[81] J. D. Hobbs,[72] B. Hoeneisen,[11] H. Hoeth,[25] M. Hohlfeld,[15] S. J. Hong,[30] R. Hooper,[77] P. Houben,[33] Y. Hu,[72] Z. Hubacek,[9] V. Hynek,[8] I. Iashvili,[69] R. Illingworth,[50] A. S. Ito,[50] S. Jabeen,[62] M. Jaffré,[15] S. Jain,[75] K. Jakobs,[22] C. Jarvis,[61] A. Jenkins,[43] R. Jesik,[43] K. Johns,[45] C. Johnson,[70] M. Johnson,[50] A. Jonckheere,[50] P. Jonsson,[43] A. Juste,[50] D. Käfer,[20] S. Kahn,[73] E. Kajfasz,[14] A. M. Kalinin,[35] J. M. Kalk,[60] J. R. Kalk,[65] S. Kappler,[20] D. Karmanov,[37] J. Kasper,[62] P. Kasper,[50] I. Katsanos,[70] D. Kau,[49] R. Kaur,[26] R. Kehoe,[79] S. Kermiche,[14] N. Khalatyan,[62] A. Khanov,[76] A. Kharchilava,[69] Y. M. Kharzheev,[35] D. Khatidze,[70] H. Kim,[78] T. J. Kim,[30] M. H. Kirby,[34] B. Klima,[50] J. M. Kohli,[26] J.-P. Konrath,[22] M. Kopal,[75] V. M. Korablev,[38] J. Kotcher,[73] B. Kothari,[70] A. Koubarovsky,[37] A. V. Kozelov,[38] D. Krop,[54] A. Kryemadhi,[81] T. Kuhl,[23] A. Kumar,[69] S. Kunori,[61] A. Kupco,[10] T. Kurča,[19] J. Kvita,[8] S. Lammers,[70] G. Landsberg,[77] J. Lazoflores,[49] A.-C. Le Bihan,[18] P. Lebrun,[19] W. M. Lee,[52] A. Leflat,[37] F. Lehner,[41] V. Lesne,[12] J. Leveque,[45] P. Lewis,[43] J. Li,[78] Q. Z. Li,[50] J. G. R. Lima,[52] D. Lincoln,[50] J. Linnemann,[65] V. V. Lipaev,[38] R. Lipton,[50] Z. Liu,[5] L. Lobo,[43] A. Lobodenko,[39] M. Lokajicek,[10] A. Lounis,[18] P. Love,[42] H. J. Lubatti,[82] M. Lynker,[55] A. L. Lyon,[50] A. K. A. Maciel,[2] R. J. Madaras,[46] P. Mättig,[25] C. Magass,[20] A. Magerkurth,[64] A.-M. Magnan,[13] N. Makovec,[15] P. K. Mal,[55] H. B. Malbouisson,[3] S. Malik,[67] V. L. Malyshev,[35] H. S. Mao,[50] Y. Maravin,[59] M. Martens,[50] R. McCarthy,[72] D. Meder,[23] A. Melnitchouk,[66] A. Mendes,[14] L. Mendoza,[7] M. Merkin,[37] K. W. Merritt,[50] A. Meyer,[20] J. Meyer,[21] M. Michaut,[17] H. Miettinen,[80] T. Millet,[19] J. Mitrevski,[70] J. Molina,[3] N. K. Mondal,[28] J. Monk,[44] R. W. Moore,[5] T. Moulik,[58] G. S. Muanza,[15] M. Mulders,[50] M. Mulhearn,[70] O. Mundal,[22] L. Mundim,[3] Y. D. Mutaf,[72] E. Nagy,[14] M. Naimuddin,[27] M. Narain,[62] N. A. Naumann,[34] H. A. Neal,[64] J. P. Negret,[7] P. Neustroev,[39] C. Noeding,[22] A. Nomerotski,[50] S. F. Novaes,[4] T. Nunnemann,[24] V. O'Dell,[50] D. C. O'Neil,[5] G. Obrant,[39] V. Oguri,[3] N. Oliveira,[3] D. Onoprienko,[59] N. Oshima,[50] R. Otec,[9] G. J. Otero y Garzón,[51] M. Owen,[44] P. Padley,[80] N. Parashar,[56] S.-J. Park,[71] S. K. Park,[30] J. Parsons,[70] R. Partridge,[77] N. Parua,[72] A. Patwa,[73] G. Pawloski,[80] P. M. Perea,[48] E. Perez,[17] K. Peters,[44]

P. Pétroff,[15] M. Petteni,[43] R. Piegaia,[1] J. Piper,[65] M.-A. Pleier,[21] P. L. M. Podesta-Lerma,[32] V. M. Podstavkov,[50] Y. Pogorelov,[55] M.-E. Pol,[2] A. Pompoš,[75] B. G. Pope,[65] A. V. Popov,[38] C. Potter,[5] W. L. Prado da Silva,[3] H. B. Prosper,[49] S. Protopopescu,[73] J. Qian,[64] A. Quadt,[21] B. Quinn,[66] M. S. Rangel,[2] K. J. Rani,[28] K. Ranjan,[27] P. N. Ratoff,[42] P. Renkel,[79] S. Reucroft,[63] M. Rijssenbeek,[72] I. Ripp-Baudot,[18] F. Rizatdinova,[76] S. Robinson,[43] R. F. Rodrigues,[3] C. Royon,[17] P. Rubinov,[50] R. Ruchti,[55] V. I. Rud,[37] G. Sajot,[13] A. Sánchez-Hernández,[32] M. P. Sanders,[61] A. Santoro,[3] G. Savage,[50] L. Sawyer,[60] T. Scanlon,[43] D. Schaile,[24] R. D. Schamberger,[72] Y. Scheglov,[39] H. Schellman,[53] P. Schieferdecker,[24] C. Schmitt,[25] C. Schwanenberger,[44] A. Schwartzman,[68] R. Schwienhorst,[65] J. Sekaric,[49] S. Sengupta,[49] H. Severini,[75] E. Shabalina,[51] M. Shamim,[59] V. Shary,[17] A. A. Shchukin,[38] W. D. Shephard,[55] R. K. Shivpuri,[27] D. Shpakov,[50] V. Siccardi,[18] R. A. Sidwell,[59] V. Simak,[9] V. Sirotenko,[50] P. Skubic,[75] P. Slattery,[71] R. P. Smith,[50] G. R. Snow,[67] J. Snow,[74] S. Snyder,[73] S. Söldner-Rembold,[44] X. Song,[52] L. Sonnenschein,[16] A. Sopczak,[42] M. Sosebee,[78] K. Soustruznik,[8] M. Souza,[2] B. Spurlock,[78] J. Stark,[13] J. Steele,[60] V. Stolin,[36] A. Stone,[51] D. A. Stoyanova,[38] J. Strandberg,[64] S. Strandberg,[40] M. A. Strang,[69] M. Strauss,[75] R. Ströhmer,[24] D. Strom,[53] M. Strovink,[46] L. Stutte,[50] S. Sumowidagdo,[49] P. Svoisky,[55] A. Sznajder,[3] M. Talby,[14] P. Tamburello,[45] W. Taylor,[5] P. Telford,[44] J. Temple,[45] B. Tiller,[24] M. Titov,[22] V. V. Tokmenin,[35] M. Tomoto,[50] T. Toole,[61] I. Torchiani,[22] S. Towers,[42] T. Trefzger,[23] S. Trincaz-Duvoid,[16] D. Tsybychev,[72] B. Tuchming,[17] C. Tully,[68] A. S. Turcot,[44] P. M. Tuts,[70] R. Unalan,[65] L. Uvarov,[39] S. Uvarov,[39] S. Uzunyan,[52] B. Vachon,[5] P. J. van den Berg,[33] R. Van Kooten,[54] W. M. van Leeuwen,[33] N. Varelas,[51] E. W. Varnes,[45] A. Vartapetian,[78] I. A. Vasilyev,[38] M. Vaupel,[25] P. Verdier,[19] L. S. Vertogradov,[35] M. Verzocchi,[50] F. Villeneuve-Seguier,[43] P. Vint,[43] J.-R. Vlimant,[16] E. Von Toerne,[59] M. Voutilainen,[67] M. Vreeswijk,[33] H. D. Wahl,[49] L. Wang,[61] M. H. L. S Wang,[50] J. Warchol,[55] G. Watts,[82] M. Wayne,[55] G. Weber,[23] M. Weber,[50] H. Weerts,[65] N. Wermes,[21] M. Wetstein,[61] A. White,[78] D. Wicke,[25] G. W. Wilson,[58] S. J. Wimpenny,[48] M. Wobisch,[50] J. Womersley,[50] D. R. Wood,[63] T. R. Wyatt,[44] Y. Xie,[77] N. Xuan,[55] S. Yacoob,[53] R. Yamada,[50] M. Yan,[61] T. Yasuda,[50] Y. A. Yatsunenko,[35] K. Yip,[73] H. D. Yoo,[77] S. W. Youn,[53] C. Yu,[13] J. Yu,[78] A. Yurkewicz,[72] A. Zatserklyaniy,[52] C. Zeitnitz,[25] D. Zhang,[50] T. Zhao,[82] B. Zhou,[64] J. Zhu,[72] M. Zielinski,[71] D. Zieminska,[54] A. Zieminski,[54] V. Zutshi,[52] and E. G. Zverev[37]

(D0 Collaboration)

[1]*Universidad de Buenos Aires, Buenos Aires, Argentina*
[2]*LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*
[3]*Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*
[4]*Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil*
[5]*University of Alberta, Edmonton, Alberta, Canada; Simon Fraser University, Burnaby, British Columbia, Canada; York University, Toronto, Ontario, Canada; and McGill University, Montreal, Quebec, Canada*
[6]*University of Science and Technology of China, Hefei, People's Republic of China*
[7]*Universidad de los Andes, Bogotá, Colombia*
[8]*Center for Particle Physics, Charles University, Prague, Czech Republic*
[9]*Czech Technical University, Prague, Czech Republic*
[10]*Center for Particle Physics, Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic*
[11]*Universidad San Francisco de Quito, Quito, Ecuador*
[12]*Laboratoire de Physique Corpusculaire, IN2P3-CNRS, Université Blaise Pascal, Clermont-Ferrand, France*
[13]*Laboratoire de Physique Subatomique et de Cosmologie, IN2P3-CNRS, Universite de Grenoble 1, Grenoble, France*
[14]*CPPM, IN2P3-CNRS, Université de la Méditerranée, Marseille, France*
[15]*IN2P3-CNRS, Laboratoire de l'Accélérateur Linéaire, Orsay, France*
[16]*LPNHE, IN2P3-CNRS, Universités Paris VI and VII, Paris, France*
[17]*DAPNIA/Service de Physique des Particules, CEA, Saclay, France*
[18]*IPHC, IN2P3-CNRS, Université Louis Pasteur, Strasbourg, France, and Université de Haute Alsace, Mulhouse, France*
[19]*Institut de Physique Nucléaire de Lyon, IN2P3-CNRS, Université Claude Bernard, Villeurbanne, France*
[20]*III. Physikalisches Institut A, RWTH Aachen, Aachen, Germany*
[21]*Physikalisches Institut, Universität Bonn, Bonn, Germany*
[22]*Physikalisches Institut, Universität Freiburg, Freiburg, Germany*
[23]*Institut für Physik, Universität Mainz, Mainz, Germany*
[24]*Ludwig-Maximilians-Universität München, München, Germany*
[25]*Fachbereich Physik, University of Wuppertal, Wuppertal, Germany*
[26]*Panjab University, Chandigarh, India*

[27]*Delhi University, Delhi, India*
[28]*Tata Institute of Fundamental Research, Mumbai, India*
[29]*University College Dublin, Dublin, Ireland*
[30]*Korea Detector Laboratory, Korea University, Seoul, Korea*
[31]*SungKyunKwan University, Suwon, Korea*
[32]*CINVESTAV, Mexico City, Mexico*
[33]*FOM-Institute NIKHEF, Amsterdam, The Netherlands, and University of Amsterdam/NIKHEF, Amsterdam, The Netherlands*
[34]*Radboud University Nijmegen/NIKHEF, Nijmegen, The Netherlands*
[35]*Joint Institute for Nuclear Research, Dubna, Russia*
[36]*Institute for Theoretical and Experimental Physics, Moscow, Russia*
[37]*Moscow State University, Moscow, Russia*
[38]*Institute for High Energy Physics, Protvino, Russia*
[39]*Petersburg Nuclear Physics Institute, St. Petersburg, Russia*
[40]*Lund University, Lund, Sweden; Royal Institute of Technology, Stockholm, Sweden; Stockholm University, Stockholm, Sweden; and Uppsala University, Uppsala, Sweden*
[41]*Physik Institut der Universität Zürich, Zürich, Switzerland*
[42]*Lancaster University, Lancaster, United Kingdom*
[43]*Imperial College, London, United Kingdom*
[44]*University of Manchester, Manchester, United Kingdom*
[45]*University of Arizona, Tucson, Arizona 85721, USA*
[46]*Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA, and University of California, Berkeley, California 94720, USA*
[47]*California State University, Fresno, California 93740, USA*
[48]*University of California, Riverside, California 92521, USA*
[49]*Florida State University, Tallahassee, Florida 32306, USA*
[50]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[51]*University of Illinois at Chicago, Chicago, Illinois 60607, USA*
[52]*Northern Illinois University, DeKalb, Illinois 60115, USA*
[53]*Northwestern University, Evanston, Illinois 60208, USA*
[54]*Indiana University, Bloomington, Indiana 47405, USA*
[55]*University of Notre Dame, Notre Dame, Indiana 46556, USA*
[56]*Purdue University Calumet, Hammond, Indiana 46323, USA*
[57]*Iowa State University, Ames, Iowa 50011, USA*
[58]*University of Kansas, Lawrence, Kansas 66045, USA*
[59]*Kansas State University, Manhattan, Kansas 66506, USA*
[60]*Louisiana Tech University, Ruston, Louisiana 71272, USA*
[61]*University of Maryland, College Park, Maryland 20742, USA*
[62]*Boston University, Boston, Massachusetts 02215, USA*
[63]*Northeastern University, Boston, Massachusetts 02115, USA*
[64]*University of Michigan, Ann Arbor, Michigan 48109, USA*
[65]*Michigan State University, East Lansing, Michigan 48824, USA*
[66]*University of Mississippi, University, Mississippi 38677, USA*
[67]*University of Nebraska, Lincoln, Nebraska 68588, USA*
[68]*Princeton University, Princeton, New Jersey 08544, USA*
[69]*State University of New York, Buffalo, New York 14260, USA*
[70]*Columbia University, New York, New York 10027, USA*
[71]*University of Rochester, Rochester, New York 14627, USA*
[72]*State University of New York, Stony Brook, New York 11794, USA*
[73]*Brookhaven National Laboratory, Upton, New York 11973, USA*
[74]*Langston University, Langston, Oklahoma 73050, USA*
[75]*University of Oklahoma, Norman, Oklahoma 73019, USA*
[76]*Oklahoma State University, Stillwater, Oklahoma 74078, USA*
[77]*Brown University, Providence, Rhode Island 02912, USA*
[78]*University of Texas, Arlington, Texas 76019, USA*
[79]*Southern Methodist University, Dallas, Texas 75275, USA*
[80]*Rice University, Houston, Texas 77005, USA*
[81]*University of Virginia, Charlottesville, Virginia 22901, USA*
[82]*University of Washington, Seattle, Washington 98195, USA*

(Received 21 September 2006; published 15 December 2006)

We report on a measurement of the B_d^0 mixing frequency and the calibration of an opposite-side flavor tagger in the D0 experiment. Various properties associated with the b quark on the opposite side of the reconstructed B meson are combined using a likelihood-ratio method into a single variable with enhanced tagging power. Its performance is tested with data, using a large sample of reconstructed semileptonic $B \to \mu D^0 X$ and $B \to \mu D^* X$ decays, corresponding to an integrated luminosity of approximately 1 fb^{-1}. The events are divided into groups depending on the value of the combined tagging variable, and an independent analysis is performed in each group. Combining the results of these analyses, the overall effective tagging power is found to be $\varepsilon \mathcal{D}^2 = (2.48 \pm 0.21^{+0.08}_{-0.06})\%$. The measured B_d^0 mixing frequency $\Delta m_d = 0.506 \pm 0.020(\text{stat}) \pm 0.016(\text{syst})$ ps^{-1} is in good agreement with the world average value.

DOI: 10.1103/PhysRevD.74.112002 PACS numbers: 12.15.Hh

I. INTRODUCTION

Particle-antiparticle mixing in the B^0 (B_d^0) system has been known for more than a decade now [1] and has been studied at the CERN LEP collider and subsequently at the Fermilab Tevatron collider during run I, between 1992 and 1996. It is currently being measured at the B-factory experiments, Belle and *BABAR*, and the Fermilab Tevatron collider experiments in the run II phase which started in 2002.

Mixing measurements involve identifying the "flavor" of the B^0 meson at production and again when it decays, where flavor indicates whether the meson contained a b or a $\bar{b}$ quark. The decay flavor is identified from the B^0 decay products when the B^0 meson is reconstructed. The determination of the initial flavor is known as flavor tagging.

The B_d^0 meson flavor at its production can be identified with information from the reconstructed side or from the opposite side (see Fig. 1). One can tag the flavor using charge correlation between "fragmentation tracks" associated with the reconstructed B meson. Such correlations were first observed in $e^+e^- \to Z^0 \to b\bar{b}$ events by the OPAL experiment [2]. This is known as "same-side flavor tagging." The flavor can be inferred also from the decay information of the second B meson in the event, assuming that b and $\bar{b}$ are produced in pairs, and thus in the ideal case, the two mesons have opposite flavors. This method is known as "opposite-side flavor tagging." An advantage of the latter method is that its performance should be independent of the type of the reconstructed B meson.

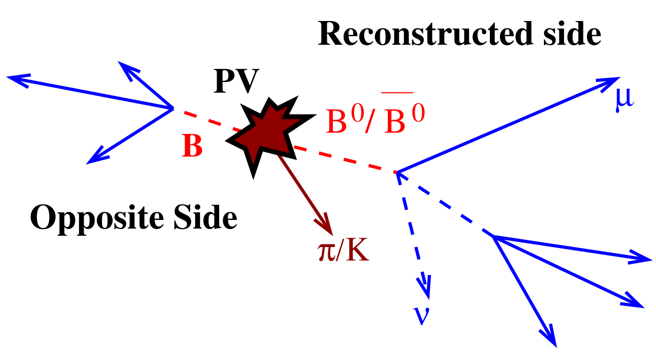

FIG. 1 (color online). Diagram of an event with a reconstructed B^0 candidate. PV indicates the primary vertex for the event.

Measurement of the B_d^0 mixing parameter is an important test of the opposite-side flavor tagging as the same tagger is used for our study of B_s mixing. Studies of tagged B^0 and B^+ samples at hadron colliders could reveal physics beyond the standard model [3]. Flavor tagging performances and optimization in e^+e^- colliders is very different from the performances and optimization techniques at hadron colliders. Hence studies of flavor tagging at Tevatron will be useful also for future hadron collider experiments at the Large Hadron Collider at CERN.

This paper describes the opposite-side flavor tagging algorithm used by the D0 experiment in run II and the measurement of its performance using $B \to \mu^+ \bar{D}^0 X$ and $B \to \mu^+ D^{*-} X$ events. Throughout the paper, a reference to a particular final state also implies its charge conjugated state. B^+ decays represent the main contribution to the $B \to \mu^+ \bar{D}^0 X$ sample, and B^0 decays dominate in the $B \to \mu^+ D^{*-} X$ sample. We measure the flavor tagging purity independently for reconstructed B^+ and B^0 events and then extract the B^0 oscillation frequency. This technique allows us to verify the assumption of independence of the opposite-side flavor tagging on the type of reconstructed B meson. Its performance is described by the two parameters, efficiency and dilution. The efficiency ε is defined as the fraction of reconstructed events (N_{tot}) that are tagged (N_{tag}):

$$\varepsilon = N_{\text{tag}}/N_{\text{tot}}. \tag{1}$$

The dilution $\mathcal{D}$ is a normalized difference of correctly and wrongly tagged events:

$$\mathcal{D} = \frac{N_{\text{cor}} - N_{\text{wr}}}{N_{\text{cor}} + N_{\text{wr}}} = \frac{N_{\text{cor}} - N_{\text{wr}}}{N_{\text{tag}}} = 2P - 1, \tag{2}$$

where $P = N_{\text{cor}}/N_{\text{tag}}$ is called the *purity*. The terms "correctly" and "wrongly" refer to the determination of the reconstructed B meson flavor. The effective tagging power of a tagging algorithm is given by $\varepsilon \mathcal{D}^2$.

II. DETECTOR DESCRIPTION

The D0 detector is described in detail elsewhere [4]. The main features of the detector essential for this analysis are summarized below. Tracks of charged particles are reconstructed from the hits in the central tracking system, which consists of the silicon microstrip tracker (SMT) and the central fiber tracker (CFT), both located in a 2 T superconducting solenoidal magnet. The SMT has $\approx 800\,000$ individual strips, with a typical pitch of 50–80 μm and a design optimized for tracking and vertexing capability for $|\eta| = 3$. The pseudorapidity, $\eta = -\ln[\tan(\theta/2)]$, approximates the true rapidity, $y = \frac{1}{2}\ln[(E + p_z c)/(E - p_z c)]$, for finite angles in the limit of $(mc^2/E) \to 0$, θ being the polar angle. We use the term "forward" to describe the regions at large $|\eta|$. The SMT system consists of six barrels arranged longitudinally (each with a set of four layers of silicon detectors arranged axially around the beam pipe), interspersed with 16 radial disks. The CFT has eight thin coaxial barrels, each supporting two doublets of overlapping scintillating fibers of 0.835 mm in diameter, one doublet being parallel to the beam axis and the other alternating by $\pm 3°$ relative to this axis. Light signals are transferred via clear light fibers to solid-state photon counters (VLPCs) that have $\approx 80\%$ quantum efficiency.

The muon system consists of a layer of tracking detectors and scintillation trigger counters in front of 1.8 T toroids, followed by two additional similar layers after the toroids. Muon tracking for $|\eta| < 1$ relies on 10-cm-wide drift tubes, while 1-cm mini drift tubes are used for $1 < |\eta| < 2$.

Electrons are identified using matching between the tracks identified in the central tracker and energy deposits in a primarily liquid-argon/uranium sampling calorimeter [4]. We also use the energy deposits in the central preshower detector [4], which consists of three concentric cylindrical layers of triangular scintillator strips and is located in a nominal 5 cm gap between the solenoid and the central calorimeter, to provide additional discrimination between electrons and fakes. The calorimeter consists of the inner electromagnetic section followed by the fine and coarse hadronic sections. In this analysis, we only use the central calorimeter ($|\eta| < 1$).

III. DATA SAMPLE AND EVENT SELECTION

This measurement is based on a large semileptonic B decay data sample corresponding to approximately 1 fb^{-1} of integrated luminosity collected with the D0 detector between April 2002 and October 2005.

B mesons are selected using their semileptonic decays $B \to \mu^+ \nu \bar{D}^0 X$ and are divided into two exclusive groups: the D^* sample, containing all events with reconstructed $D^{*-} \to \bar{D}^0 \pi^-$ decays, and the D^0 sample, containing all the remaining events. The D^* sample is dominated by $B_d^0 \to \mu^+ \nu D^{*-} X$ decays, while the D^0 sample is dominated by $B^+ \to \mu^+ \nu \bar{D}^0 X$ decays.

The flavor tagging procedure is developed using events from the D^0 sample. Events from the D^* sample are used to measure the purity of the flavor tagging and the oscillation parameter Δm_d. In addition, the purity is measured in the D^0 sample to test the hypothesis that the flavor tagger is independent of the type of reconstructed B meson.

Muons for this analysis are required to have hits in more than one muon chamber, an associated track in the central tracking system with hits in both SMT and CFT detectors, transverse momentum $p_T^\mu > 2$ GeV$/c$, as measured in the central tracker, pseudorapidity $|\eta^\mu| < 2$, and total momentum $p^\mu > 3$ GeV$/c$.

All charged particles in a given event are clustered into jets using the DURHAM clustering algorithm [5] with the cutoff parameter set to 15 GeV$/c$. Events with more than one identified muon in the same jet or with the reconstructed $J/\psi \to \mu^+\mu^-$ decays are rejected.

D^0 candidates are constructed from two tracks of opposite charge belonging to the same jet as the reconstructed muon. Both tracks are required to have transverse momentum $p_T > 0.7$ GeV$/c$ and pseudorapidity $|\eta| < 2$. They are required to form a common D vertex with a fit $\chi^2 < 9$, number of degrees of freedom being 1. For each track, the projection ϵ_T (onto the axial plane, i.e. perpendicular to the beam direction) and projection ϵ_L (onto the stereo plane, i.e. parallel to the beam direction) of its impact parameter with respect to the primary vertex, together with the corresponding uncertainties $(\sigma(\epsilon_T), \sigma(\epsilon_L))$ are computed. The combined impact parameter significance $S = \sqrt{[\epsilon_T/\sigma(\epsilon_T)]^2 + [\epsilon_L/\sigma(\epsilon_L)]^2}$ is required to be greater than 2. The distance d_T^D between the primary and D vertices in the axial plane is required to exceed 4 standard deviations: $d_T^D/\sigma(d_T^D) > 4$. The accuracy of the d_T^D determination is required to be better than 500 μm. The angle α_T^D between the D^0 momentum vector and the direction from the primary to the D vertex in the axial plane is required to satisfy the condition $\cos\alpha_T^D > 0.9$. The tracks of the muon and D^0 candidate are required to form a common B vertex with a fit $\chi^2 < 9$, with number of degrees of freedom being 1. The mass of the kaon is assigned to the track having the same charge as the muon; the remaining track is assigned the mass of the pion. The mass of the $(\mu^+ \bar{D}^0)$ system is required to fall within the $2.3 < M(\mu^+ \bar{D}^0) < 5.2$ GeV$/c^2$ range.

If the distance d_T^B between the primary and B vertices in the axial plane exceeded $4\sigma(d_T^B)$, the angle α_T^B between the B momentum and the direction from the primary to the B vertex in the axial plane is required to satisfy the condition $\cos\alpha_T^B > 0.95$. The distance d_T^B is allowed to be greater than d_T^D, provided that the distance between the B and D vertices d_T^{BD} is less than $3\sigma(d_T^{BD})$. The uncertainty $\sigma(d_T^B)$ is required to be less than 500 μm. In addition, the cut $p_T(\bar{D}^0) > 5$ GeV$/c^2$ is applied.

To select $\mu^+ D^{*-}$ candidates, we searched for an additional pion track with $p_T > 0.18$ GeV$/c$ and the charge opposite to the charge of the muon. The mass difference $\Delta M = M(K\pi\pi) - M(K\pi)$ for D^* candidates, with $1.75 < M(\bar{D}^0) < 1.95$ GeV$/c^2$, is shown in Fig. 2. The

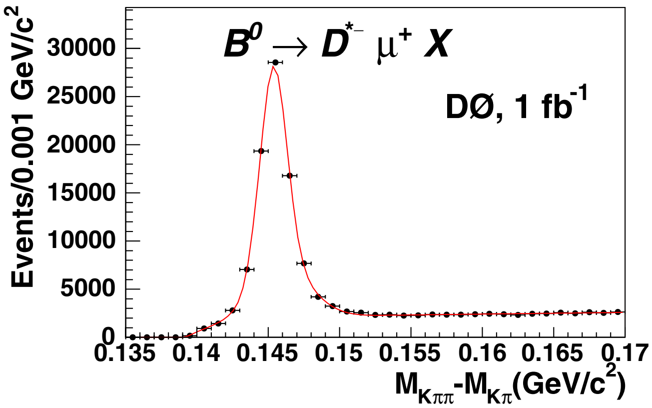

FIG. 2 (color online). The $M(K\pi\pi) - M(K\pi)$ invariant mass distribution for selected μD^* candidates. The curve shows the result of the fit described in Sec. VII.

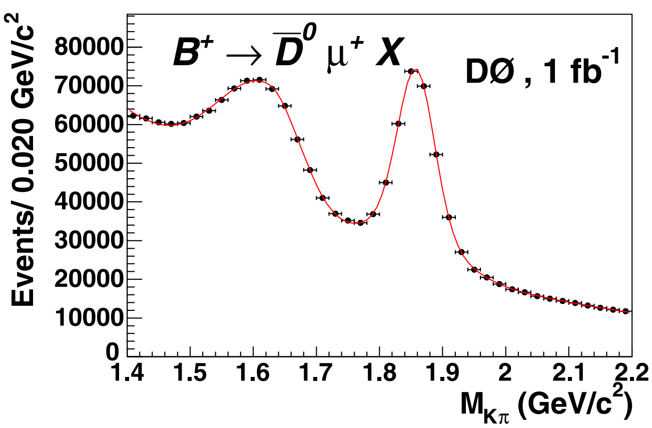

FIG. 3 (color online). The $K\pi$ invariant mass distribution for selected μD^0 candidates. The curve shows the result of the fit described in Sec. VII.

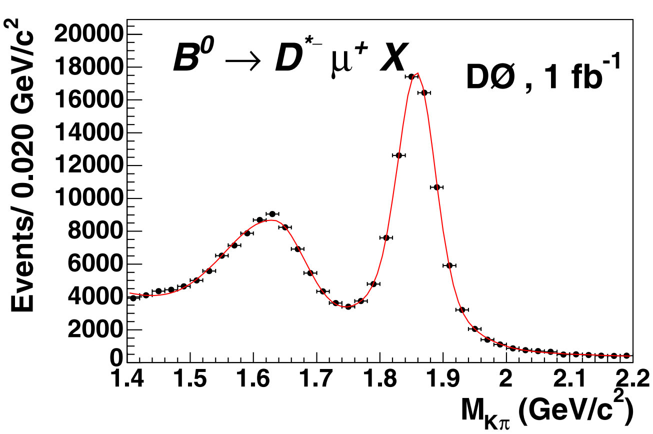

FIG. 4 (color online). The $K\pi$ invariant mass for selected μD^* candidates. The curve shows the result of the fit described in Sec. VII.

peak corresponding to the mass of the soft pion in the $\mu^+ D^{*-}$ sample is clearly seen.

All events with $0.1425 < \Delta M < 0.1490\ \mathrm{GeV}/c^2$ are included in the D^* sample. The remaining events are assigned to the D^0 sample. The $K\pi$ mass distributions for these two samples together with the results of the fits are shown in Figs. 3 and 4. The procedure to fit these mass spectra is described in Sec. VII. In total, $230\,551 \pm 1627$ $B \to \mu^+ \nu \bar{D}^0$ decays and $73\,532 \pm 304$ $B \to \mu^+ \nu \bar{D}^*$ decays are reconstructed.

IV. VISIBLE PROPER DECAY LENGTH

The oscillations of B mesons are usually studied as a function of their proper decay length. Since in semileptonic B decays an undetected neutrino carries away part of the energy, the proper decay length cannot be accurately measured. Instead, a *visible proper decay length* (VPDL) is used in this analysis. It is defined as

$$L = M_B(\boldsymbol{L}_{xy} \cdot \boldsymbol{P}_{xy}^{\mu D^0})/|\boldsymbol{P}_T^{\mu D^0}|^2. \tag{3}$$

Here $\boldsymbol{L}_{xy}$ is a vector in the plane perpendicular to the beam direction from the primary to the B meson decay vertex. The transverse momentum $\boldsymbol{P}_T^{\mu D^0}$ is defined as the vector sum of the transverse momenta of the muon and D^0. M_B is the mass of the B meson.

V. DESCRIPTION AND COMBINATION OF FLAVOR TAGGERS

Many different properties can be used to identify the initial flavor b or $\bar{b}$ of a heavy quark fragmenting into a reconstructed B meson. Some of them have strong separation power, while others are weaker. In all cases, their combination into a single tagging variable gives a significantly better result than that in the case of their separate use. We build such a combination with the likelihood-ratio method described below. We first describe the combination algorithm and then discuss the discriminating variables used.

A. Combination of variables

We construct a set of discriminating variables $x_1, \ldots, x_n$ for a given event. A discriminating variable, by definition, should have different distributions for b and $\bar{b}$ flavors. 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 r is defined as

$$r = \prod_{i=1}^{n} r_i; \qquad r_i = \frac{f_i^{\bar{b}}(x_i)}{f_i^b(x_i)}. \tag{4}$$

A given variable x_i may not be defined for some events. For example, there are events that do not contain an identified muon on the opposite side. In this case, the corresponding variable r_i is set to 1. An initial b flavor is more probable if $r < 1$, and a $\bar{b}$ flavor is more probable if $r > 1$. By construction, an event with $r < 1$ is tagged as a b quark and an event with $r > 1$ is tagged as a $\bar{b}$ quark. For an oscillation analysis, it is more convenient to define the tagging variable as

$$d = \frac{1-r}{1+r}. \tag{5}$$

By construction, the variable d ranges between -1 and 1. An event with $d > 0$ is tagged as a b quark and with $d < 0$ as a $\bar{b}$ quark, with higher $|d|$ values corresponding to higher tagging purities. For uncorrelated variables $x_1, \ldots, x_n$, and perfect modeling of the pdf, d gives the best possible tagging performance, and its absolute value provides a measure of the dilution of the flavor tagging defined in Eq. (2).

Very often, the analyzed events are divided into samples with significantly different discriminating variables and tagging performances. This division would imply making a separate analysis for each sample and combining the results at a later stage. In contrast to this approach, the tagging variable d defined by Eqs. (4) and (5) provides a "calibration" for all events, regardless of their intrinsic differences. Since the absolute value of d gives a measure of the dilution of the flavor tagging, events from different categories but with a similar absolute value of d can be treated in the same way. Thus, another important advantage of this method of flavor tagging is the possibility of building a single variable having the same meaning for different kinds of events. It allows us to classify all events according to their tagging characteristics and use them simultaneously in the analysis.

All of the discriminating variables used in this analysis are constructed using the properties of the b quark opposite to the reconstructed B meson ("opposite-side tagging"). Since an important property of the opposite-side tagging is the independence of its performance of the type of reconstructed B meson, it can be calibrated in data by applying tagging to the events with B^0 and B^+ decays. The measured performance can then be used to study B_s^0 meson oscillations, as an example.

The probability density functions for each discriminating variable discussed below are constructed using events from the D^0 sample with $0 < \mathrm{VPDL} < 500\ \mu\mathrm{m}$. In this sample, the decay $B^+ \to \mu^+ \nu \bar{D}^0$ dominates, see Sec. VII C. The $B_d^0 \to \mu^+ \nu D^{*+}$ events give a 16% contribution to the sample and, due to the cut on VPDL, contains mainly nonoscillated B_d^0 decays, as determined by Monte Carlo (the standard PYTHIA generator v6.2 [6] generation, followed by decay of B mesons with EVTGEN decay package [7], passed through GEANT v3.15 [8] modeling of the detector response, followed by event reconstruction).

The initial flavor of a b quark is therefore determined by the charge of the muon. Estimates based on Monte Carlo simulation indicate that the purity of the initial flavor determination in the selected sample is 0.98 ± 0.01, where the uncertainty is due to the uncertainties in measured branching fractions of B meson decays.

For each discriminating variable, the signal band containing all events with $1.80 < M(K\pi) < 1.92\ \mathrm{GeV}/c^2$ and the background band containing all events with $1.94 < M(K\pi) < 2.2\ \mathrm{GeV}/c^2$ are defined. The pdf's are constructed as the difference in the distributions. The latter distributions are normalized by multiplying them by 0.74 so that the number of events in the background band corresponds to the estimated number of background events in the signal band.

B. Flavor tagger discriminants

We now describe the variables used. An additional muon is searched for in each analyzed event. This muon is required to have at least one hit in the muon chambers and to have $\cos\phi(\boldsymbol{p}_\mu, \boldsymbol{p}_B) < 0.8$, where $\boldsymbol{p}_B$ is the three-momentum of the reconstructed B meson, and ϕ is the angle between the vectors $\boldsymbol{p}_\mu$ and $\boldsymbol{p}_B$. If more than one muon is found, the muon with the highest number of hits in the muon chambers is used. If more than one muon with the same number of hits in the muon chambers is found, the muon with the highest transverse momentum p_T is used. For this muon, a *muon jet charge* Q_J^μ is constructed as

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

where q^i is the charge and p_T^i is the transverse momentum of the ith particle, and the sum is taken over all charged particles, including the muon, satisfying the condition $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$, where $\Delta\phi$ and $\Delta\eta$ are computed with respect to the muon direction. Daughters of the reconstructed B meson are explicitly excluded from this sum. In addition, any charged particle with $\cos\phi(\boldsymbol{p}, \boldsymbol{p}_B) > 0.8$ is excluded. The distribution of the muon jet charge variable is shown in Figs. 5(a) and 5(b). In these plots, q^{rec} gives the charge of the b quark in the reconstructed $B \to \mu^+ \nu \bar{D}^0$ decay, in this case given by the muon charge. We build separate pdf's for muons with hits in all three layers of the muon detector, Fig. 5(a), and for muons with fewer than three hits, Fig. 5(b).

In addition to the muon tag, reconstructed electrons with $\cos\phi(\boldsymbol{p}_e, \boldsymbol{p}_B) < 0.8$ are also used for flavor tagging. The electron is reconstructed by extrapolating a track to the calorimeter and adding up the energy deposited in a narrow tube or "road" around the track. Calorimeter cells are collected around the track extrapolated positions in each layer and the total transverse energy of the cluster is defined by the sum of the energies in each layer. The electrons are required to be in the central region ($|\eta| < 1.1$), with $p_T > 2\ \mathrm{GeV}/c$. They are required to have at

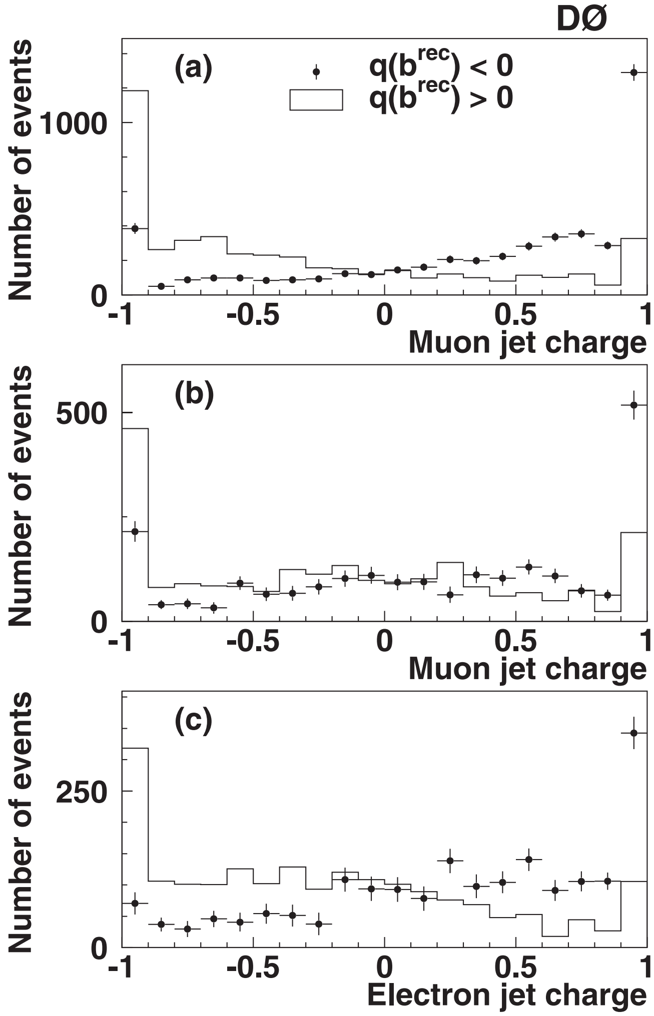

FIG. 5. (a) Distribution of the jet charge Q_J^μ for muons with hits in all three layers of the muon detector. (b) Distribution of the jet charge Q_J^μ, for muons with fewer than three hits. (c) Distribution of the jet charge for electrons Q_J^e. Here $q(b^{\rm rec})$ is the charge of the muon from the reconstruction side.

least one hit each in the CFT and SMT. They are required to have energy deposits in the electromagnetic calorimeter consistent with an electron, $0.55(0.5) < E/p < 1.0(1.1)$, and low energy deposit in the hadron calorimeter, i.e. the electromagnetic fraction $\mathrm{EMF} > 0.8(0.7)$. The cuts are looser for electrons with $p_T > 3.5\ \mathrm{GeV}/c$ and are given in parentheses. EMF and E/p are calculated as below:

$$\mathrm{EMF} = \frac{\sum_{\text{layer number } i=1,2,3} E_T(i)}{\sum_{\text{all layers}} E_T(i)} \tag{6}$$

$$E/p = \frac{\sum_{\text{layer number } i=1,2,3} E_T(i)}{p_T(\text{track})}, \tag{7}$$

where $E_T(i)$ is the transverse energy within the road in the ith layer. We also require a minimum single layer cluster energy of a cluster in the central preshower, $\mathrm{CPS}_E^{\rm SLC} > 4.0(2.0)\ \mathrm{MeV}/c$. The cuts are optimized by studying electrons from conversion decays ($\gamma \to e^+e^-$) and fakes from $K_S^0 \to \pi^+\pi^-$ decays to obtain a 90% purity for electrons. For these electrons, an *electron jet charge* (Q_J^e) is constructed in the same way as the *muon jet charge*, Q_J^μ. The distribution of the electron jet charge variable is shown in Fig. 5(c).

An additional secondary vertex corresponding to the decay of a B hadron is searched for, using all charged tracks in the event excluding those from the reconstructed B hadron. The secondary vertex is also required to contain at least two tracks with an axial impact parameter significance greater than 3. The distance l_{xy} from the primary to the secondary vertex must also satisfy the condition: $l_{xy} > 4\sigma(l_{xy})$. The details of the secondary vertex identification algorithm can be found in Ref. [9].

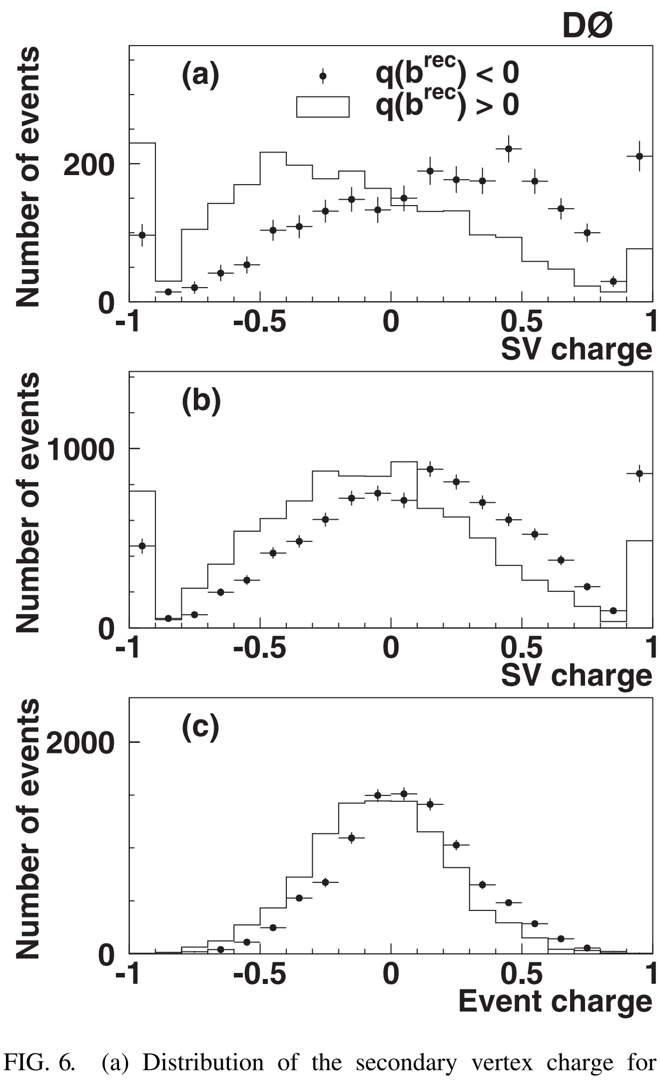

FIG. 6. (a) Distribution of the secondary vertex charge for events with an opposite-side muon. (b) Distribution of the secondary vertex charge for events without an opposite-side muon. (c) Distribution of the event jet charge. $q(b^{\rm rec})$ is the charge of the b quark from the reconstruction side.

The three-momentum of the secondary vertex $\boldsymbol{p}_{\rm SV}$ is defined as the vector sum of the momenta of all tracks included in the secondary vertex. A secondary vertex with $\cos\phi(\boldsymbol{p}_{\rm SV}, \boldsymbol{p}_B) < 0.8$ is used for flavor tagging. A *secondary vertex charge* $Q_{\rm SV}$ is defined as the third discriminating variable

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

where the sum is taken over all tracks included in the secondary vertex. Daughters of the reconstructed B meson are explicitly excluded from this sum. In addition, any charged particle with $\cos\phi(\boldsymbol{p}, \boldsymbol{p}_B) > 0.8$ is excluded. Here p_L^i is the longitudinal momentum of track i with respect to the direction of the secondary vertex momentum $\mathbf{p}_V$. A value of $k = 0.6$ is used, taken from previous studies at LEP [10]. We verified that this value of k results in the optimal performance of the $Q_{\rm SV}$ variable. Figures 6(a) and 6(b) show the distribution of this variable for the events with and without an identified muon flavor tag.

Finally, the *event charge* $Q_{\rm EV}$ is constructed as

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

The sum is taken over all charged tracks with $0.5 < p_T < 50$ GeV$/c$ and having $\cos\phi(\boldsymbol{p}, \boldsymbol{p}_B) < 0.8$. Daughters of the reconstructed B meson are explicitly excluded from this sum. The distribution of this variable is shown in Fig. 6(c).

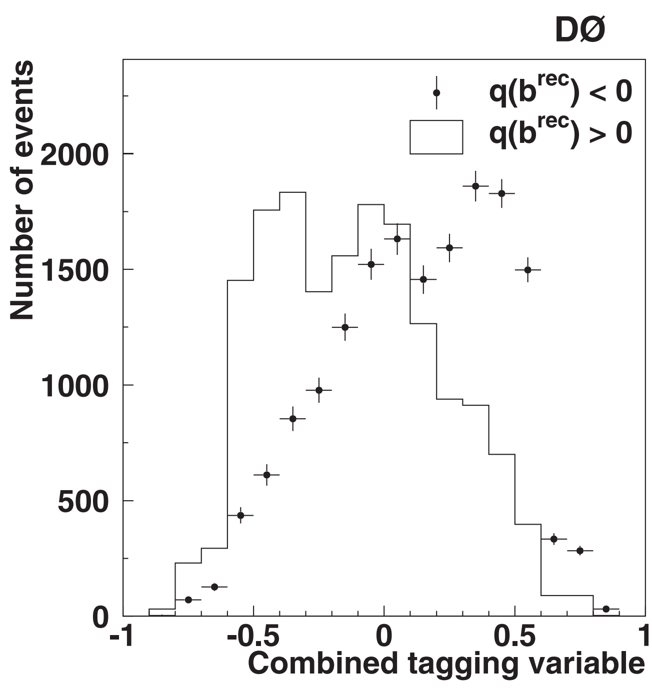

FIG. 7. Normalized distributions of the combined tagging variable. $q(b^{\rm rec})$ is the charge of the b quark from the reconstruction side.

For each event with an identified muon, the muon jet charge Q_J^μ and the secondary vertex charge $Q_{\rm SV}$ are used to construct a *muon tagger*. For each event without a muon but with an identified electron, the electron jet charge Q_J^e and the secondary vertex charge $Q_{\rm SV}$ are 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 $Q_{\rm SV}$ and the event jet charge $Q_{\rm EV}$ are used to construct a 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. 7. The performances of these taggers are discussed in the following sections.

VI. MULTIDIMENSIONAL TAGGER

In addition to the flavor tagger described in Sec. V, an alternative algorithm is developed also and used to measure B^0 mixing. This tagger is *multidimensional*, i.e., the likelihood functions it is based on depends on more than a single variable. In addition, the pdf's are determined from simulated events, while the primary flavor tagger described in Sec. V uses data to construct the pdf's. The multidimensional tagger therefore provides a cross-check of the primary algorithm.

If, as before, we have a set of discriminants $x_1, \ldots, x_n$, the likelihood that the meson has flavor b at the time of creation can be written as $\mathcal{L}(b; x_1, \ldots, x_n)$. A similar expression $\mathcal{L}(\bar{b}; x_1, \ldots, x_n)$ holds for the likelihood for $\bar{b}$. These likelihoods relate to the variable d as

$$d = \frac{\mathcal{L}(b) - \mathcal{L}(\bar{b})}{\mathcal{L}(b) + \mathcal{L}(\bar{b})}. \tag{8}$$

This definition is similar to Eq. (5).

The likelihoods are obtained from the simulated samples of $B^\pm \to J/\psi K^\pm$ with $J/\psi \to \mu^+\mu^-$. This final state does not oscillate and is therefore flavor-pure. The $B^- \to J/\psi K^-$ sample is used to obtain $\mathcal{L}(b)$, while $\mathcal{L}(\bar{b})$ is determined from $B^+ \to J/\psi K^+$ sample. In practice, the likelihoods are stored as multidimensional histograms (with one dimension per discriminating variable) with the bin content 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. (8).

In addition to the discriminating variables introduced in Sec. V, other variables are used for the multidimensional tagger. For each identified opposite-side muon, the transverse momentum p_T relative to the beam axis and transverse momentum $p_T^{\rm rel}$ relative to the nearest jet are computed. (The muon is included in the jet clustering.) Another variable defined for the muon is its impact parameter significance S_μ, where S_μ is the transverse impact parameter significance $\epsilon_T/\sigma(\epsilon_T)$, where ϵ_T is defined in

Sec. III. For each reconstructed opposite-side secondary vertex, the secondary vertex transverse momentum p_T^{SV} is computed by taking the magnitude of the transverse projection of the vector sum of all tracks in that vertex. In principle, all discriminating variables can be combined into a single multidimensional likelihood. However, since a binned likelihood is used, 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 limited simulation statistics, this means that discriminating variables must be chosen wisely when making a combination.

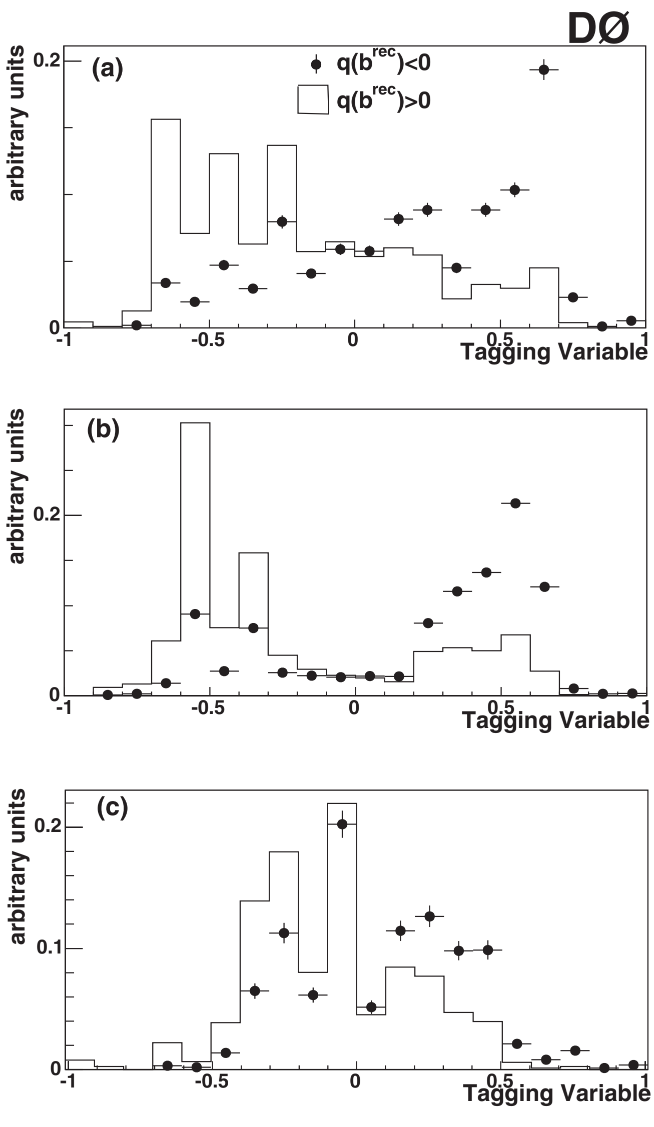

FIG. 8. Normalized distributions of the combined tagging variable for three multidimensional taggers for the simulated samples $B^\pm \to J/\psi K^\pm$. Here $q(b^{\mathrm{rec}})$ is the charge of the b quark from the reconstructed side. (a) Distribution of d for Tag(μ + SV). (b) Distribution of d for Tag(μ − SV). (c) Distribution of d for Tag(SV − μ).

All events are divided into three categories based on their opposite-side content. The following variables for different categories are selected.

(1) Events with muon and secondary vertex: $\mathrm{Tag}(\mu + \mathrm{SV}) = \{Q_J^\mu; p_T^{\mathrm{rel}}; Q_{\mathrm{SV}}\}$.

(2) Events with muon and without secondary vertex: $\mathrm{Tag}(\mu - \mathrm{SV}) = \{Q_J^\mu; p_T^{\mathrm{rel}}; p_T; S_\mu\}$.

(3) Events with secondary vertex without a muon: $\mathrm{Tag}(\mathrm{SV} - \mu) = \{Q_{\mathrm{EV}}; Q_{\mathrm{SV}}; p_T^{\mathrm{SV}}\}$.

Distributions in the tagging variable d for the above three taggers are shown in Fig. 8. They are made by applying the taggers to the simulated $B^\pm \to J/\psi K^\pm$ samples from which they are created.

The final multidimensional tagger used the following logic to decide which of its subtaggers to use. For events containing a muon and a secondary vertex, the Tag(μ + SV) is used. If the opposite-side contained a muon and no secondary vertex, the Tag(μ − SV) is used. If the opposite-side contained an electron, the electron tagger described in Sec. V is used. Note that this tagger is not multidimensional and is not derived from simulation. If the opposite-side contained a secondary vertex, the Tag(SV − μ) is used.

VII. ASYMMETRY FIT PROCEDURE

The performance of the flavor tagging and measurements of the B^0 mixing frequency Δm_d are obtained from a study of the dependence of the flavor asymmetry on the B-meson decay length.

The flavor asymmetry A is defined as

$$A = \frac{N^{\mathrm{nos}} - N^{\mathrm{osc}}}{N^{\mathrm{nos}} + N^{\mathrm{osc}}}. \tag{9}$$

Here N^{nos} is the number of nonoscillated B decays and N^{osc} is the number of oscillated B decays. An event $B \to \mu^+ \nu \bar{D}^0 X$ with $q(\mu) \times d < 0$ is tagged as nonoscillated, and an event with $q(\mu) \times d > 0$ is tagged as oscillated. The flavor tagging variable d is defined in Eq. (5) or (8).

All events in the D^0 and D^* samples are divided into seven groups according to the measured VPDL (L) defined in Eq. (3). The numbers of oscillated N_i^{osc} and nonoscillated N_i^{nos} signal events in each group i are determined from the number of the D^0 signal events given by a fit to the $K\pi$ invariant mass distribution for both samples. The seven VPDL bins (in cm) defined are $-0.025 < L \leq 0.0$, $0.0 < L \leq 0.025$, $0.025 < L \leq 0.050$, $0.050 < L \leq 0.075$, $0.075 < L \leq 0.1$, $0.1 < L \leq 0.125$, and $0.125 < L \leq 0.2$.

A. Mass fit

In this section we describe the mass fitting procedure. The fitting function is chosen to give the best χ^2 of the fit to the $K\pi$ mass spectrum of the entire sample of $B \to \mu^+ \bar{D}^0 X$ events shown in Figs. 3 and 4. The signal peak corresponding to the decay $D^0 \to K^- \pi^+$ can be seen at $1.857\ \mathrm{GeV}/c^2$. The background to the right of the signal region is adequately described by an exponential function:

$$f_1^{\rm bkg}(x) = a_0 \times e^{-(x/b_0)}, \tag{10}$$

where x is the $K\pi$ mass.

The peak in the background to the left of the signal is due to events in which D mesons decay to $K\pi X$ where X is not reconstructed. It is modeled with a bifurcated Gaussian function:

$$\begin{aligned} f_2^{\rm bkg}(x) &= A e^{-[(x-\mu_0)^2/(2\sigma_R^2)]} \quad \text{for } x - \mu_0 \geq 0 \\ &= A e^{-[(x-\mu_0)^2/(2\sigma_L^2)]} \quad \text{for } x - \mu_0 < 0. \end{aligned} \tag{11}$$

Here μ_0 is the mean of the Gaussian, and σ_L and σ_R are the two widths of the bifurcated Gaussian function.

The signal has been modeled by the sum of two Gaussians:

$$f^{\rm sig}(x) = \frac{N^{\rm sig}}{\sqrt{2\pi}}\left(\frac{r_1}{\sigma_1} e^{-[(x-\mu_1)^2/(2\sigma_1^2)]} + \frac{1-r_1}{\sigma_2} e^{-[(x-\mu_2)^2/(2\sigma_2^2)]}\right), \tag{12}$$

where $N^{\rm sig}$ is the number of signal events, μ_1 and μ_2 are the means of the Gaussians, σ_1 and σ_2 are the widths of the Gaussians, and r_1 is the fractional contribution of the first Gaussian.

The complete fitting function, which has 12 free parameters, is

$$f(x) = f^{\rm sig}(x) + f_1^{\rm bkg}(x) + f_2^{\rm bkg}(x). \tag{13}$$

The low statistics in some VPDL bins, which have as few as ten events after flavor tagging, do not permit a free fit to this function. Consequently, some parameters had to be constrained or fixed. In order to do this, it is necessary to show that the constraints on the parameters are valid for all of the VPDL bins. Unconstrained fits are performed to several high statistic samples, and the set of all events is used as a reference fit. Events are divided into VPDL bins and fitted to investigate the VPDL dependence of the fit results. In addition, three samples are made to test whether the presence of a flavor tag changes the mass spectrum: all tagged events over the entire VPDL range, all events in the short VPDL range [0, 0.05] tagged as opposite-sign events, and all events in VPDL range [0, 0.05] tagged as same-sign events.

This study showed that the width, position, and the ratio of the signal Gaussians, as well as the position and widths of the bifurcated Gaussian describing the background can be fixed to the values obtained from the fit to the total D^0 or D^* mass distribution. This left four free parameters: the numbers of events in the signal peak, background peak, and exponential background, and the slope constant of the exponential background. Examples of the fits to the $K\pi$ mass distribution in different VPDL bins are shown in Fig. 9.

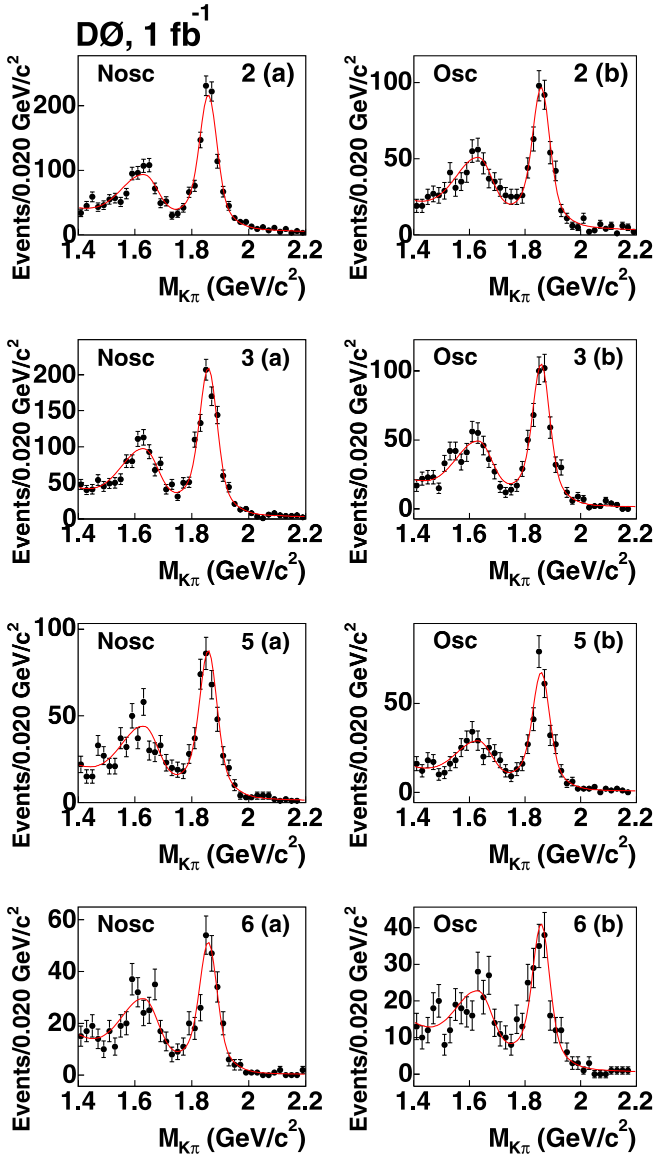

FIG. 9 (color online). The fit to $M(K\pi)$ mass for nonoscillating (left) and oscillating (right) $B \to \mu^+ \nu D^{*-}$ events tagged by the muon tagger with $|d| > 0.3$ in VPDL bins $0.0 < L \leq 0.025$ cm (2a, 2b), $0.025 < L \leq 0.050$ cm (3a, 3b), $0.075 < L \leq 0.100$ cm (5a, 5b), and, $0.100 < L \leq 0.125$ cm (6a, 6b).

The number of D^* candidates is estimated using the distribution of $(M_{K^+\pi^-\pi^-} - M_{K^+\pi^-})$, shown in Fig. 2. In this case, the signal is modeled with two Gaussians as described by Eq. (12) and the background by the product of a linear and exponential function

$$f^{\rm bkg}(x) = a[1 + c(x - x_0)]e^{(x-x_0)/b_0}, \tag{14}$$

where x is the mass difference $(M_{K^+\pi^-\pi^-} - M_{K^+\pi^-})$ in this equation.

B. Expected flavor asymmetry

For a given type of B_q meson ($q = u, d, s$), the distribution of the visible proper decay-length L is given by

$$n_u^{\rm nos}(L, K) = \frac{1}{2} \cdot \frac{K}{c\tau(B^+)} e^{(-(KL)/(c\tau(B^+)))}(1 + \mathcal{D}_u), \quad (15)$$

$$n_u^{\rm osc}(L, K) = \frac{1}{2} \cdot \frac{K}{c\tau(B^+)} e^{(-(KL)/(c\tau(B^+)))}(1 - \mathcal{D}_u), \quad (16)$$

$$n_d^{\rm nos}(L, K) = \frac{1}{2} \cdot \frac{K}{c\tau(B^0)} e^{(-(KL)/(c\tau(B^0)))} \cdot \left[1 + \mathcal{D}_d \cos\left(\Delta m_d \frac{KL}{c}\right)\right], \quad (17)$$

$$n_d^{\rm osc}(L, K) = \frac{1}{2} \cdot \frac{K}{c\tau(B^0)} e^{(-(KL)/(c\tau(B^0)))} \cdot \left[1 - \mathcal{D}_d \cos\left(\Delta m_d \frac{KL}{c}\right)\right], \quad (18)$$

$$n_s^{\rm nos}(L, K) = \frac{1}{2} \cdot \frac{K}{c\tau(B_s^0)} e^{(-(KL)/(c\tau(B_s^0)))}, \quad (19)$$

$$n_s^{\rm osc}(L, K) = \frac{1}{2} \cdot \frac{K}{c\tau(B_s^0)} e^{(-(KL)/(c\tau(B_s^0)))}. \quad (20)$$

Here τ is the lifetime of the B meson, Δm_d is the mixing frequency of B^0 mesons, the factor $K = P_T^{\mu D^0}/P_T^B$ reflects the difference between the measured ($P_T^{\mu D^0}$) and true (P_T^B) momenta of the B meson. The B^+ meson does not oscillate, and it is assumed in these studies that the B_s^0 meson oscillates with infinite frequency. The flavor tagging dilution is given by $\mathcal{D}$. In general, it can be different for B^0 and B^+. In our study we verified the assumption that $\mathcal{D}_d = \mathcal{D}_u$ for our opposite-side flavor tagging.

The transition from the true to the experimentally measured visible proper decay-length L^M is achieved by integration over the K-factor distribution and convolution with the resolution function:

$$N_{q,j}^{\rm nos/osc}(L^M) = \int dL R_j(L - L^M)\varepsilon_j(L)\theta(L) \times \int dK D_j(K) n_{q,j}^{\rm nos/osc}(L, K). \quad (21)$$

Here $R_j(L - L^M)$ is the detector resolution in the VPDL, and $\varepsilon_j(L)$ is the reconstruction efficiency for a given channel j of B_q meson decay. The step function $\theta(L)$ forces L to be positive in the integration. L^M can be negative due to resolution effects. The function $D_j(K)$ is a normalized distribution of the K-factor in a given channel j, obtained from simulated events.

In addition to the main decay channel $B \to \mu^+ \nu \bar{D}^0 X$, the process $c\bar{c} \to \mu^+ \nu \bar{D}^0 X$ contributes to the selected final state. A dedicated analysis is developed to study this process, both in data and in simulation. It shows that the pseudo decay length, constructed from the crossing of the μ and $\bar{D}^0$ trajectories, is distributed around zero with $\sigma \approx 150\ \mu$m. The distribution $N^{c\bar{c}}(L^M)$ of the VPDL for this process is taken from simulation. It is assumed that the production ratio $(c \to D^*)/(c \to D^0)$ is the same as in semileptonic B decays and that the flavor tagging for the $c\bar{c}$ events gives the same rate of oscillated and nonoscillated events. The fraction $f_{c\bar{c}}$ of $c\bar{c}$ events is obtained from the fit.

Taking into account all of the above-mentioned contributions, the expected number of (non)oscillated events in the ith bin of VPDL is

$$N_i^{e,\rm nos/osc} = \int_i dL^M(1 - f_{c\bar{c}})\left(\sum_{q=u,d,s}\sum_j {\rm Br}_j \cdot N_{q,j}^{\rm nos/osc}(L^M)\right) + \int_i dL^M f_{c\bar{c}} N_{c\bar{c}}(L^M). \quad (22)$$

Here the integration $\int_i dL^M$ is taken over a given interval i, the sum $\sum_j$ is taken over all decay channels $B_q \to \mu^+ \nu \bar{D}^0 X$ contributing to the selected sample, and ${\rm Br}_j$ is the branching fraction of channel j.

Finally, the expected value of asymmetry, $A_i^e(\Delta m, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u)$, for the interval i of the measured VPDL is given by

$$A_i^e(\Delta m, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) = \frac{N_i^{e,\rm nos} - N_i^{e,\rm osc}}{N_i^{e,\rm nos} + N_i^{e,\rm osc}}. \quad (23)$$

The expected asymmetry can be computed both for the D^* and the D^0 samples. The only difference between them is due to the different relative contributions of various decay channels of B mesons.

For the computation of A_i^e, the B meson lifetimes and the branching fractions ${\rm Br}_j$ are taken from the Particle Data Group (PDG) [11]. They are discussed in the following section. The functions $D_j(K)$, $R_j(L)$, and $\varepsilon_j(L)$ are obtained from MC simulation. Variations of these inputs within their uncertainties are included in the systematic uncertainties.

C. Sample composition

There is a cross-contamination between the $b \to B^0 \to \mu^+ \nu \bar{D}^0 X$, $b \to B_s^0 \to \mu^+ \nu \bar{D}^0 X$, and $b \to B^+ \to \mu^+ \nu \bar{D}^0 X$ samples. To determine the composition of the selected samples, we studied all possible decay chains for B^0, B_s^0, and B^+ with their corresponding branching fractions, from which we estimated the sample composition in the D^* and D^0 samples.

The following decay channels of B mesons are considered for the D^* sample:

$$B^0 \to \mu^+ \nu D^{*-}, \qquad B^0 \to \mu^+ \nu D^{**-} \to \mu^+ \nu D^{*-} X,$$
$$B^+ \to \mu^+ \nu \bar{D}^{**0} \to \mu^+ \nu D^{*-} X, \qquad B_s^0 \to \mu^+ \nu D^{*-} X;$$

and for the D^0 sample:

$$B^+ \to \mu^+ \nu \bar{D}^0,$$
$$B^+ \to \mu^+ \nu \bar{D}^{*0},$$
$$B^+ \to \mu^+ \nu \bar{D}^{**0} \to \mu^+ \nu \bar{D}^0 X,$$
$$B^+ \to \mu^+ \nu \bar{D}^{**0} \to \mu^+ \nu \bar{D}^{*0} X,$$
$$B^0 \to \mu^+ \nu D^{**-} \to \mu^+ \nu \bar{D}^0 X,$$
$$B^0 \to \mu^+ \nu D^{**-} \to \mu^+ \nu \bar{D}^{*0} X,$$
$$B_s^0 \to \mu^+ \nu \bar{D}^0 X,$$
$$B_s^0 \to \mu^+ \nu \bar{D}^{*0} X.$$

Here, and in the following, the symbol "D^{**}" denotes both narrow and wide D^{**} resonances, as well as nonresonant $D\pi$ and $D^*\pi$ production.

The most recent PDG values [11] are used to determine the branching fractions of decays contributing to the D^0 and D^* samples:

$$\mathrm{Br}(B^+ \to \mu^+ \nu \bar{D}^0) = (2.15 \pm 0.22) \times 10^{-2},$$
$$\mathrm{Br}(B^0 \to \mu^+ \nu D^-) = (2.14 \pm 0.20) \times 10^{-2},$$
$$\mathrm{Br}(B^+ \to \mu^+ \nu \bar{D}^{*0}) = (6.5 \pm 0.5) \times 10^{-2},$$
$$\mathrm{Br}(B^0 \to \mu^+ \nu D^{*-}) = (5.44 \pm 0.23) \times 10^{-2}.$$

$\mathrm{Br}(B^+ \to \mu^+ \nu \bar{D}^{**0})$ is estimated using the following inputs:

$$\mathrm{Br}(B \to \mu^+ \nu X) = (10.73 \pm 0.28) \times 10^{-2},$$
$$\mathrm{Br}(B^0 \to \mu^+ \nu X) = \tau(B^0)/\tau(B^+) \cdot \mathrm{Br}(B^+ \to \mu^+ \nu X),$$
$$\mathrm{Br}(B^+ \to \mu^+ \nu \bar{D}^{**0}) = \mathrm{Br}(B^+ \to \mu^+ \nu X) - \mathrm{Br}(B^+ \to \mu^+ \nu \bar{D}^0) - \mathrm{Br}(B^+ \to \mu^+ \nu \bar{D}^{*0}), \qquad (24)$$

where $\tau(B^0)$ is the B^0 lifetime, and $\tau(B^+)$ is the B^+ lifetime. The following value is obtained:

$$\mathrm{Br}\,(B^+ \to \mu^+ \nu \bar{D}^{**0}) = (2.70 \pm 0.47) \times 10^{-2}. \qquad (25)$$

$\mathrm{Br}(B^0 \to \mu^+ \nu D^{**-})$ is obtained as follows:

$$\mathrm{Br}(B^0 \to \mu^+ \nu D^{**-}) = \mathrm{Br}(B^0 \to \mu^+ \nu X) - \mathrm{Br}(B^0 \to \mu^+ \nu D^-) - \mathrm{Br}(B^+ \to \mu^+ \nu D^{*-}),$$
$$\mathrm{Br}(B^0 \to \mu^+ \nu D^{**-}) = \frac{\tau(B^0)}{\tau(B^+)} \cdot \mathrm{Br}(B^+ \to \mu^+ \nu \bar{D}^{**0})$$

$\mathrm{Br}(B^+ \to \mu^+ \nu \bar{D}^{**0} \to \mu^+ \nu D^{*-} X)$ is estimated from the following inputs [12,13]:

$$\mathrm{Br}(\bar{b} \to l^+ \nu D^{*-} \pi^+ X) = (4.73 \pm 0.8 \pm 0.6) \times 10^{-3},$$
$$\mathrm{Br}(\bar{b} \to l^+ \nu D^{*-} \pi^+ X) = (4.80 \pm 0.9 \pm 0.5) \times 10^{-3},$$
$$\mathrm{Br}(\bar{b} \to l^+ \nu D^{*-} \pi^+ X) = (0.6 \pm 0.7 \pm 0.2) \times 10^{-3}$$

and assuming $\mathrm{Br}(b \to B^+) = (0.397 \pm 0.010)$ [11]. The usual practice in estimating this decay rate is to neglect the contributions of the decays $D^{**} \to D^* \pi\pi$. However, the above data allows us to take these decays into account.

Neglecting the decays $D^{**} \to D^* \pi\pi\pi$, the available measurements can be expressed as

$$\mathrm{Br}(\bar{B} \to l^+ \nu D^{*-} \pi^+ X) = \mathrm{Br}(B^+ \to l^+ \nu D^{*-} \pi^+ X^0), + \mathrm{Br}(B^0 \to l^+ \nu D^{*-} \pi^+ \pi^-),$$
$$\mathrm{Br}(\bar{B} \to l^+ \nu D^{*-} \pi^- X) = \mathrm{Br}(B^0 \to l^+ \nu D^{*-} \pi^+ \pi^-).$$

From these relations and using the above measurements, we obtain

$$\mathrm{Br}(B^+ \to \mu^+ \nu \bar{D}^{**0} \to l^+ \nu D^{*-} X) = (1.06 \pm 0.24) \times 10^{-2}. \qquad (26)$$

All other factors for the $\mathrm{Br}(B \to \mu^+ \nu \bar{D}^{**} \to \mu^+ \nu \bar{D}^* X)$ are obtained assuming the following relations,

$$\frac{\mathrm{Br}(B^0 \to \mu^+ \nu D^{**-} \to \mu^+ \nu D^* \pi)}{\mathrm{Br}(B^+ \to \mu^+ \nu \bar{D}^{**0} \to \mu^+ \nu D^* \pi)} = \tau(B^0)/\tau(B^+),$$
$$\frac{\mathrm{Br}(B \to \mu^+ \nu \bar{D}^{**} \to \mu^+ \nu \bar{D}^* \pi^+)}{\mathrm{Br}(B \to \mu^+ \nu \bar{D}^{**} \to \mu^+ \nu \bar{D}^* \pi^0)} = 2.$$

$\mathrm{Br}(B \to \mu^+ \nu \bar{D}^{**} \to \mu^+ \nu \bar{D} X)$ is estimated from the following inputs:

$$\frac{\mathrm{Br}(B \to \mu^+ \nu \bar{D}^{**} \to \mu^+ \nu \bar{D} \pi^+)}{\mathrm{Br}(B \to \mu^+ \nu \bar{D}^{**} \to \mu^+ \nu \bar{D} \pi^0)} = 2,$$

$$\mathrm{Br}(B \to \mu^+ \nu \bar{D}^{**}) = \mathrm{Br}(B \to \mu^+ \nu \bar{D}^{**} \to \mu^+ \nu \bar{D} X) + \mathrm{Br}(B \to \mu^+ \nu \bar{D}^{**} \to \mu^+ \nu \bar{D}^* X).$$

To estimate branching fractions for B_s^0 decays, $\mathrm{Br}(B_s^0 \to \mu^+ \nu D_s^- X) = (7.9 \pm 2.4) \times 10^{-2}$ is taken from Ref. [11] and the following assumptions are used:

$$\frac{\mathrm{Br}(B_s^0 \to \mu^+ \nu X)}{\mathrm{Br}(B^0 \to \mu^+ \nu X)} = \tau(B_s^0)/\tau(B^0),$$
$$\frac{\mathrm{Br}(B_s^0 \to \mu^+ \nu D_s^{**-} \to \mu^+ \nu D^{*-} X)}{\mathrm{Br}(B_s^0 \to \mu^+ \nu D_s^{**-} \to \mu^+ \nu \bar{D}^{*0} X)} = 1,$$

where $\tau(B_s^0)$ is the B_s^0 meson lifetime. In addition, it is assumed that

$$\frac{\mathrm{Br}(B_s^0 \to \mu^+ \nu D_s^{**-} \to \mu^+ \nu D^* X)}{\mathrm{Br}(B_s^0 \to \mu^+ \nu D_s^{**})} = 0.35. \qquad (27)$$

There is no experimental measurement of this ratio yet and to estimate the corresponding systematic uncertainty, this ratio is varied between 0 and 1.

Since the D^* sample is selected by the cut on the mass difference $\Delta M = M(D^0\pi) - M(D^0)$, there is a small additional contribution of $B \to \mu^+ \nu \bar{D}^0$ events to the D^* sample, when D^0 is randomly combined with a pion from the combinatorial background. The fraction of this contribution is estimated using $\mu^+ \bar{D}^0 \pi^+$ events. These events are selected applying all the criteria for the D^* sample, described in Sec. III, except that the wrong charge correlation of muon and pion is required, i.e., the muon and the pion are required to have the same charge. The number of D^0 events is determined using the same fitting procedure as for the D^* sample, and the additional fraction of $B \to \mu^+ \nu \bar{D}^0$ events in the D^* sample is estimated to be $(4.00 \pm 0.85) \times 10^{-2}$. This fraction is included in the fitting procedure and the uncertainty in this value is taken into account in the overall systematics.

In addition to these branching fractions, various decay chains are affected differently by the B meson selection cuts, and the corresponding reconstruction efficiencies are determined from simulation to correct for this effect. Taking into account these efficiencies, the composition of the D^* sample is estimated to be $(0.86 \pm 0.03)\ B^0$, $(0.13 \pm 0.03)\ B^+$, and $(0.01 \pm 0.01)\ B_s^0$. The D^0 sample contains $0.83 \pm 0.03\ B^+$, $(0.16 \pm 0.04)\ B^0$, and $(0.01 \pm 0.01)\ B_s^0$.

VIII. RESULTS

For each sample of tagged events, the observed and expected asymmetries are determined using Eqs. (9) and (23) in all VPDL bins, and the values of Δm_d, $f_{c\bar{c}}$, $\mathcal{D}_u$, and $\mathcal{D}_d$ are obtained from a simultaneous χ^2 fit:

$$\chi^2(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) = \chi^2_{D^*}(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) + \chi^2_{D^0}(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u); \quad (28)$$

$$\chi^2_{D^*}(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) = \sum_i \frac{[A_{i,D^*} - A^e_{i,D^*}(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u)]^2}{\sigma^2(A_{i,D^*})}; \quad (29)$$

$$\chi^2_{D^0}(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u) = \sum_i \frac{[A_{i,D^0} - A^e_{i,D^0}(\Delta m_d, f_{c\bar{c}}, \mathcal{D}_d, \mathcal{D}_u)]^2}{\sigma^2(A_{i,D^0})}. \quad (30)$$

Here $\sum_i$ is the sum over all VPDL bins. Examples of the χ^2 fit to the flavor asymmetry minimization given in Eq. (28) is shown in Fig. 10.

The performance of the flavor tagging method is studied separately for the muon, electron, and secondary vertex taggers using events with $|d| > 0.3$. Results are given in Tables I, II, and III. All uncertainties in these tables are statistical only and do not include systematic uncertainties. The performances of the combined tagger defined in Sec. V B for events with $|d| > 0.3$ and the alternative multidimensional tagger defined in Sec. VI for events with $|d| > 0.37$ are also shown. The cut on $|d|$ is somewhat different for the multidimensional tagger as the calibration is different and we compare the dilutions for the same tag efficiency for the two taggers. The tagging efficiencies shown in Tables I and II are computed using events with VPDL $= [0.025, 0.250]$. This selection reduces the contribution from $c\bar{c} \to \mu^+ \nu D^0 X$ events, since they have a

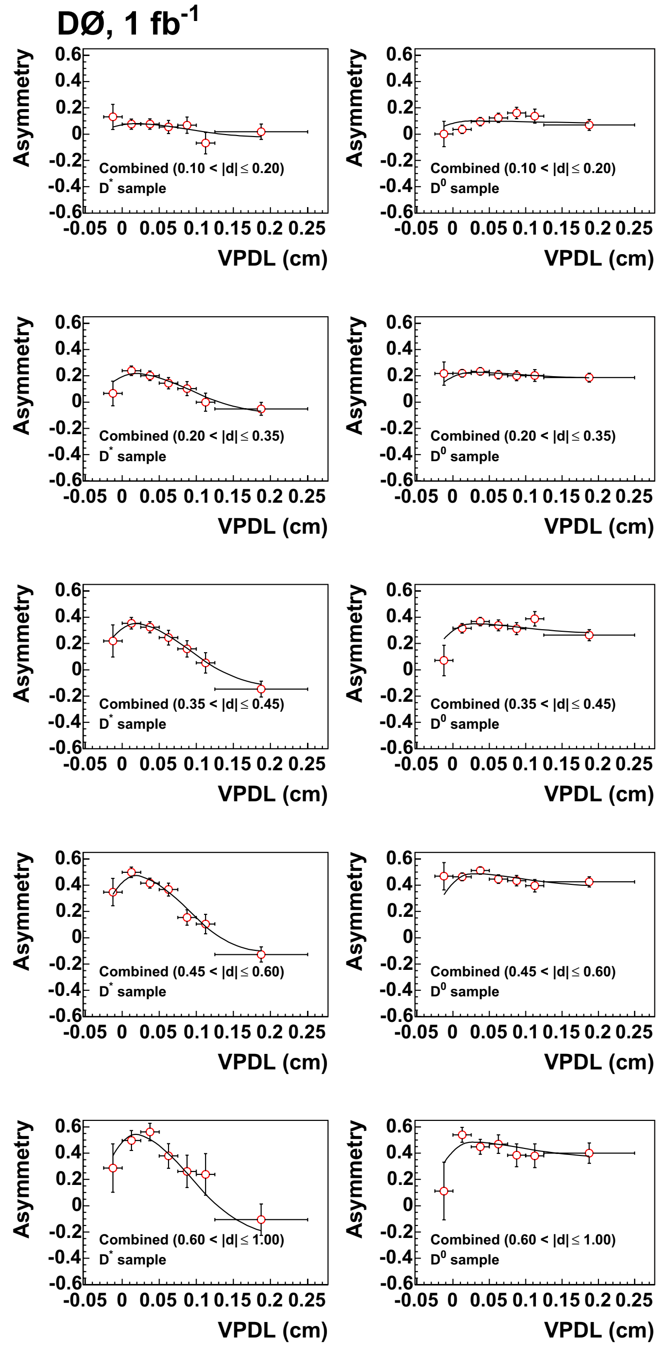

FIG. 10 (color online). Asymmetries obtained in the D^* and D^0 sample with the combined tagger in $|d|$ bins. Circles are data, and the result of the fit is superimposed.

TABLE I. Tagging performance for the D^* sample for different taggers and subsamples. Uncertainties are statistical only.

Tagger	ε (%)	$\mathcal{D}_d$	$\varepsilon\mathcal{D}_d^2$ (%)
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
SV Charge ($\|d\| > 0.3$)	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 ($0.10 < \|d\| \leq 0.20$)	4.63 ± 0.10	0.084 ± 0.031	0.03 ± 0.02
Combined ($0.20 < \|d\| \leq 0.35$)	5.94 ± 0.12	0.236 ± 0.027	0.33 ± 0.08
Combined ($0.35 < \|d\| \leq 0.45$)	3.89 ± 0.09	0.385 ± 0.034	0.58 ± 0.10
Combined ($0.45 < \|d\| \leq 0.60$)	4.36 ± 0.10	0.512 ± 0.032	1.14 ± 0.14
Combined ($0.60 < \|d\| \leq 1.00$)	1.13 ± 0.05	0.597 ± 0.058	0.40 ± 0.08

TABLE II. Tagging performance for the D^0 sample 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. Uncertainties are statistical only.

Tagger	ε (%)	$\mathcal{D}_u$	$\mathcal{D}'_d$	$\varepsilon\mathcal{D}_u^2$ (%)
Muon ($\|d\| > 0.3$)	7.10 ± 0.09	0.444 ± 0.015	0.463 ± 0.028	1.400 ± 0.096
Electron ($\|d\| > 0.3$)	1.88 ± 0.05	0.445 ± 0.032	0.324 ± 0.060	0.372 ± 0.054
SV Charge ($\|d\| > 0.3$)	2.81 ± 0.06	0.338 ± 0.026	0.421 ± 0.049	0.320 ± 0.050
Combined ($\|d\| > 0.3$)	11.74 ± 0.11	0.419 ± 0.012	0.434 ± 0.023	2.058 ± 0.121
Multidim. ($\|d\| > 0.37$)	11.67 ± 0.11	0.363 ± 0.012	0.384 ± 0.023	1.540 ± 0.106
Combined ($0.10 < \|d\| \leq 0.20$)	4.59 ± 0.08	0.104 ± 0.017	0.079 ± 0.029	0.050 ± 0.016
Combined ($0.20 < \|d\| \leq 0.35$)	6.10 ± 0.09	0.234 ± 0.014	0.212 ± 0.024	0.335 ± 0.042
Combined ($0.35 < \|d\| \leq 0.45$)	3.98 ± 0.07	0.361 ± 0.018	0.364 ± 0.032	0.519 ± 0.052
Combined ($0.45 < \|d\| \leq 0.60$)	4.77 ± 0.07	0.504 ± 0.016	0.489 ± 0.030	1.211 ± 0.077
Combined ($0.60 < \|d\| \leq 1.00$)	1.17 ± 0.04	0.498 ± 0.031	0.572 ± 0.056	0.290 ± 0.038

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

Tagger	Δm_d (ps^{-1})	$f_{c\bar{c}}$
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 ($0.10 < \|d\| \leq 0.20$)	0.506 ± 0.209	0.495 ± 0.505
Combined ($0.20 < \|d\| \leq 0.35$)	0.523 ± 0.064	0.021 ± 0.025
Combined ($0.35 < \|d\| \leq 0.45$)	0.531 ± 0.042	0.063 ± 0.038
Combined ($0.45 < \|d\| \leq 0.60$)	0.510 ± 0.032	0.010 ± 0.010
Combined ($0.60 < \|d\| \leq 1.00$)	0.456 ± 0.049	0.032 ± 0.026

VPDL distribution with zero mean and $\sigma \approx 150\ \mu$m as described in Sec. VII.

Individual taggers give compatible values of Δm_d and $f_{c\bar{c}}$, as can be seen in Table III. For the combined tagger with $|d| > 0.3$, the following results are obtained:

$$\begin{aligned} \varepsilon\mathcal{D}_d^2 &= (2.19 \pm 0.22)\%, \\ \Delta m_d &= 0.513 \pm 0.023\ \text{ps}^{-1}, \\ f_{c\bar{c}} &= (3.3 \pm 1.3)\%. \end{aligned} \tag{31}$$

The multidimensional tagger which used simulation for the description of pdf's as described in Sec. VI gives consistent results both for the Δm_d and the fraction $f_{c\bar{c}}$, which is used as a cross-check of the main tagging algorithm.

One of the goals of this analysis is to verify the assumption of independence of the opposite-side flavor tagging on the type of the reconstructed B meson. It can be seen from Tables I and 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 the better selection of $\mu^+\nu D^{*-}$ events due to an additional requirement of the charge correlation between the muon and pion from $D^{*-} \to D^0\pi^-$ 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 suppresses this background and results in a better measurement of the tagging performance.

To verify this hypothesis, a special sample of events satisfying all conditions for the D^* sample, except the requirement of the charge correlation between the muon and the pion, is selected. The dilution $\mathcal{D}'_d$ for this sample is shown in Table II. It can be seen that $\mathcal{D}'_d$ is systematically lower than $\mathcal{D}_d$ for both the individual and the combined taggers. $\mathcal{D}'_d$ is the right quantity to be compared with $\mathcal{D}_u$ and Table II shows that they are statistically compatible. This result, therefore, confirms the expectation of the same performance of the opposite-side flavor tagging for B^+ and B^0 events. It also shows that contribution of background in the D^0 sample reduces the measured dilution for B^+ events. Thus, the dilution measured in the D^* sample can be used for the B_s^0 mixing measurement, where a similar charge correlation between the muon and D_s is required.

By construction, the dilution for each event should strongly depend on the magnitude of the tagging variable d. This property becomes important in the B_s^0 mixing measurement, since in this case the dilution of each event can be estimated using the value of d and can be included in a likelihood function, improving the sensitivity of the measurement. To test the dependence of the dilution on d, all tagged events are divided into subsamples with $0.1 < |d| \leq 0.2$, $0.2 < |d| \leq 0.35$, $0.35 < |d| \leq 0.45$, $0.45 < |d| \leq 0.6$, and $|d| > 0.6$. The overall tagging efficiency for this sample is $(19.95 \pm 0.21)\%$. The dilutions obtained are shown in Table I. Their strong dependence on the value of the tagging variable is clearly seen. This allows us to perform a dilution calibration and obtain the measured dilution $\mathcal{D}_d$ as a function of the predicted value $|d|$. This is used to provide an event-by-event dilution for the B_s mixing analysis and the calibration derived in this analysis is used for the two sided C.L. on B_s mixing, obtained by D0 [14]. The overall tagging power, computed as the sum of the tagging powers in all subsamples, is

$$\varepsilon\mathcal{D}_d^2 = (2.48 \pm 0.21)\%. \tag{32}$$

The measured oscillation parameters Δm_d for all considered taggers and subsamples are given in Table III. They are compatible with the world average value $\Delta m_d = 0.502 \pm 0.007\ \mathrm{ps}^{-1}$ [11] in each instance.

The final mixing parameter Δm_d is obtained from the simultaneous fit of the flavor asymmetry in the various tagging variable subsamples defined above. The fraction $f_{c\bar{c}}$ is constrained to be the same for all subsamples. The result is

TABLE IV. Systematic uncertainties for Δm_d.

	Default	Variation		Δm_d (ps^{-1})	
		(a)	(b)	(a)	(b)
$\mathrm{Br}(B^0 \to D^{*-}\mu^+\nu)$	5.44	-0.23	0.23	0.002	-0.002
$\mathrm{Br}(B \to D^*\pi\mu\nu X)$	1.07	-0.17	0.17	-0.0078	0.0078
R^{**}	0.35	0	1.0	0.0006	-0.0012
B lifetimes	0.05022	-0.00054	0.00054	0.0008	-0.0008
Resolution scale factor	⋯	1.2	0.8	0.0021	-0.0021
Alignment	⋯	$-10\ \mu$m	$+10\ \mu$m	-0.004	$+0.004$
K-factor	⋯	-2%	$+2\%$	0.0098	-0.0094
Efficiency	⋯	-12%	$+12\%$	-0.0054	0.0052
Fraction D^0 in D^*	4%	3.15%	4.85%	-0.0020	$+0.0030$
Fit procedure		See split below		$^{+0.0023}_{-0.0019}$	
Bin width	2 MeV	1.6	2.67	0.0009	0.0014
Parameter μ_0	⋯	-3σ	3σ	-0.0001	0.0001
Parameter $\frac{\sigma_R+\sigma_L}{2}$	⋯	-3σ	3σ	-0.0001	⋯
Parameter $\frac{\sigma_R-\sigma_L}{\sigma_R+\sigma_L}$	⋯	-3σ	3σ	-0.0001	0.0001
Parameter μ_1	⋯	-3σ	3σ	-0.0016	0.0015
Parameter $\frac{\sigma_1+\sigma_2}{2}$	⋯	-3σ	3σ	-0.0006	0.0006
Parameter R	⋯	-3σ	3σ	$-.0005$	0.0004
Parameter $(\mu_2 - \mu_1)$	⋯	-3σ	3σ	0.0006	-0.0007
Parameter $\frac{\sigma_1-\sigma_2}{\sigma_1+\sigma_2}$	⋯	-3σ	3σ	⋯	⋯
Total				± 0.0158	

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

	Default	Variation		$\mathcal{D}(B^0)$ $0.1 < \lvert d\rvert \le 0.2$		$\mathcal{D}(B^0)$ $0.2 < \lvert d\rvert \le 0.3$		$\mathcal{D}(B^0)$ $0.3 < \lvert d\rvert \le 0.45$		$\mathcal{D}(B^0)$ $0.45 < \lvert d\rvert \le 0.6$		$\mathcal{D}(B^0)$ $0.6 < \lvert d\rvert \le 1.0$	
		(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
$\mathrm{Br}(B^0 \to D^{*-}\mu^+\nu)$	5.44	−0.23	0.23	···	···	···	−0.001	0.001	···	0.001	−0.001	0.001	−0.001
$\mathrm{Br}(B \to D^*\pi\mu\nu X)$	1.07	−0.17	0.17	0.0004	−0.0004	−0.0011	0.0011	−0.0019	0.0021	−0.0020	0.0021	−0.0008	0.0028
R^{**}	0.35	0.0	1.0	−0.0009	0.0016	−0.0027	0.0048	−0.0042	0.0079	−0.0057	0.0105	−0.0066	0.0124
B lifetimes	0.05022	−0.00054	0.00054	···	−0.0001	0.0001	−0.0002	0.0003	−0.0001	0.0003	−0.0003	0.0014	−0.0003
Resolution function	···	×1.2	×0.8	0.0005	−0.0006	0.0010	−0.0012	0.0020	−0.0021	0.0024	−0.0028	0.0028	−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	···	···	0.0001	−0.0001	···	···	···
Efficiency	···	−12%	+12%	0.0006	−0.0007	−0.0008	0.0006	−0.0012	0.0011	−0.0013	0.0010	−0.0021	0.0019
Fraction D^0 in D^*	4%	3.15%	4.85%	···	0.0010	−0.0010	···	−0.0010	0.0010	−0.0010	0.0010	−0.0010	0.0010
Fit procedure		See split below		+0.0021 −0.0031		+0.0040 −0.0041		+0.0060 −0.0046		+0.0044 −0.0019		+0.0119 −0.0111	
Bin width	2 MeV	1.6	2.67	−0.0026	0.0002	−0.0024	0.0014	−0.0001	0.0027	0.0037	0.0038	0.0089	0.0087
Parameter μ_0	···	-3σ	3σ	−0.0003	0.0002	0.0001	−0.0001	0.0001	0.0001	−0.0002	0.0001	−0.0007	0.0007
Parameter $\frac{\sigma_R+\sigma_L}{2}$	···	-3σ	3σ	0.0002	−0.0002	0.0001	−0.0001	0.0004	−0.0003	···	−0.0001	−0.0002	0.0001
Parameter $\frac{\sigma_R-\sigma_L}{\sigma_R+\sigma_L}$	···	-3σ	3σ	−0.0005	0.0005	0.0002	−0.0001	0.0002	0.0001	−0.0002	0.0001	−0.0015	0.0011
Parameter μ_1	···	-3σ	3σ	−0.0009	0.0010	−0.0017	0.0018	0.0023	−0.0015	0.0006	−0.0005	−0.0004	−0.0004
Parameter $\frac{\sigma_1+\sigma_2}{2}$	···	-3σ	3σ	0.0008	−0.0005	0.0014	−0.0009	0.0037	−0.0034	−0.0013	0.0017	−0.0099	0.0068
Parameter R	···	-3σ	3σ	0.0015	−0.0011	0.0029	−0.0024	0.0030	−0.0027	0.0013	−0.0011	−0.0046	0.0035
Parameter $\mu_2 - \mu_1)$	···	-3σ	3σ	···	−0.0003	0.0008	−0.0011	−0.0001	0.0006	−0.0003	0.0002	0.0008	−0.0003
Parameter $\frac{\sigma_1-\sigma_2}{\sigma_1+\sigma_2}$	···	-3σ	3σ	−0.0001	···	−0.0004	0.0003	0.0002	−0.0002	−0.0004	0.0004	−0.0006	0.0010
Total				+.0049 −0.0052		+.0077 −0.0066		+.0111 −0.0081		+.0125 −0.0081		+.0182 −0.0140	

$$\Delta m_d = 0.506 \pm 0.020\ \mathrm{ps}^{-1}, \qquad f_{c\bar{c}} = (2.2 \pm 0.9)\%. \tag{33}$$

The statistical precision of Δm_d from the simultaneous fit is about 10% better than that from the fit of events with $|d| > 0.3$. This improvement is directly related to a better overall tagging power [Eq. (32)] for the sum of subsamples as compared to the result [Eq. (31)] for the sample with $|d| > 0.3$.

IX. SYSTEMATIC UNCERTAINTIES

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

These uncertainties are obtained as follows:

(i) The B meson branching fractions and lifetimes used in the fit of the asymmetry are taken from Ref. [11] and are varied by 1 standard deviation.

(ii) The VPDL resolution obtained in simulation is multiplied by factors of 0.8 and 1.2. These factors exceed the uncertainty in the difference of the resolution between data and simulation.

(iii) The variation of K-factors with the change in the B momentum is neglected in this analysis. To check the impact of this assumption on the final result, the computation of K-factors is repeated without the cut on $p_T(D^0)$ or by applying an additional cut on the p_T of muon, $p_T > 4$ GeV$/c$. The change in the average values of the K-factors does not exceed 2%, which is used as the estimate of the systematic uncertainty in their values. This uncertainty is propagated into the variation of Δm_d and tagging purity by repeating the fit with the K-factor distributions shifted by 2%.

(iv) The ratio of the reconstruction efficiencies 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 [15] of semileptonic B decays is used. The uncertainty in the reconstruction efficiency, set at 12%, is estimated by varying the 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 a HQET-motivated model [16] produces a smaller variation. The fit to the asymmetry is repeated with the efficiencies to reconstruct the $B \to \mu^+ \nu D^{**-}$ and $B \to \mu^+ \nu \bar{D}^{**0}$ channels modified by $\pm 12\%$, and the difference is taken as the systematic uncertainty from this source.

(v) The additional fraction of D^0 events contributing to the D^* sample is estimated at $(4.00 \pm 0.85)\%$ (see Sec. VII C). This variation is used to estimate the systematic uncertainty from this source. As a cross-check, the number of D^* events is determined from the fit of the mass difference $M(D^0\pi) - M(D^0)$ and the fit of the flavor asymmetry is repeated. The measured value of $\Delta m_d = 0.507 \pm 0.020\ \mathrm{ps}^{-1}$ is consistent with Eq. (33).

(vi) We also investigated the systematic uncertainty in determining the number of D^* and D^0 candidates in each VPDL bin.

(a) The values of the parameters which had been fixed from the fit to ''all'' events are varied by $\pm 3\sigma$.

The default bin width for the fits in the VPDL bin is 0.020 GeV. We lowered the bin width to 0.016 GeV and increased the bin width to 0.027 GeV and included the resulting variations in the systematic uncertainty.

X. CONCLUSIONS

We have performed a study of a likelihood-based opposite-side tagging algorithm in B^0 and B^+ samples obtained with $\sim 1\ \mathrm{fb}^{-1}$ of run II data. The dilutions $\mathcal{D}(B^+)$ and $\mathcal{D}(B^0)$ are found to be the same within their statistical uncertainties. This result justifies the application of the B_d^0 dilution to the B_s^0 mixing analysis.

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

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

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

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

which is in good agreement with the world average value of $\Delta m_d = 0.502 \pm 0.007\ \mathrm{ps}^{-1}$ [11].

ACKNOWLEDGMENTS

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 and KOSEF (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); The Swedish Research Council (Sweden); Research Corporation; Alexander von Humboldt Foundation; and the Marie Curie Program.

[1] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **192**, 245 (1987); M. Artuso *et al.* (CLEO Collaboration), Phys. Rev. Lett. **62**, 2233 (1989).

[2] R. Akers *et al.* (OPAL Collaboration), Z. Phys. C **66**, 19 (1995).

[3] E. Berger *et al.*, Phys. Rev. Lett. **86**, 4231 (2001).

[4] V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **565**, 463 (2006).

[5] S. Catani *et al.*, Phys. Lett. B **269**, 432 (1991).

[6] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).

[7] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).

[8] R. Brun and F. Carminati, CERN Prog. Library Long Writeup Report No. W5013 (unpublished).

[9] J. Abdallah *et al.* (DELPHI Collaboration), Eur. Phys. J. C **32**, 185 (2004).

[10] J. Abdallah *et al.* (DELPHI Collaboration), Eur. Phys. J. C **35**, 35 (2004).

[11] S. Eidelman *et al.* (Particle Data Group), Phys. Lett. B **592**, 1 (2004).

[12] D. Buskulic *et al.* (ALEPH Collaboration), Z. Phys. C **73**, 601 (1997).

[13] P. Abreu *et al.* (DELPHI Collaboration), Phys. Lett. B **475**, 407 (2000).

[14] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **97**, 021802 (2006).

[15] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995).

[16] M. Neubert, Phys. Rep. **245**, 259 (1994).