Simultaneous measurement of the B^0 meson lifetime and mixing frequency with $B^0 \to D^{*-} \ell^+ \nu_\ell$ decays

B. Aubert,[1] R. Barate,[1] D. Boutigny,[1] J.-M. Gaillard,[1] A. Hicheur,[1] Y. Karyotakis,[1] J. P. Lees,[1] P. Robbe,[1] V. Tisserand,[1] A. Zghiche,[1] A. Palano,[2] A. Pompili,[2] J. C. Chen,[3] N. D. Qi,[3] G. Rong,[3] P. Wang,[3] Y. S. Zhu,[3] G. Eigen,[4] I. Ofte,[4] B. Stugu,[4] G. S. Abrams,[5] A. W. Borgland,[5] A. B. Breon,[5] D. N. Brown,[5] J. Button-Shafer,[5] R. N. Cahn,[5] E. Charles,[5] M. S. Gill,[5] A. V. Gritsan,[5] Y. Groysman,[5] R. G. Jacobsen,[5] R. W. Kadel,[5] J. Kadyk,[5] L. T. Kerth,[5] Yu. G. Kolomensky,[5] J. F. Kral,[5] C. LeClerc,[5] M. E. Levi,[5] G. Lynch,[5] L. M. Mir,[5] P. J. Oddone,[5] T. J. Orimoto,[5] M. Pripstein,[5] N. A. Roe,[5] A. Romosan,[5] M. T. Ronan,[5] V. G. Shelkov,[5] A. V. Telnov,[5] W. A. Wenzel,[5] T. J. Harrison,[6] C. M. Hawkes,[6] D. J. Knowles,[6] S. W. O'Neale,[6] R. C. Penny,[6] A. T. Watson,[6] N. K. Watson,[6] T. Deppermann,[7] K. Goetzen,[7] H. Koch,[7] B. Lewandowski,[7] M. Pelizaeus,[7] K. Peters,[7] H. Schmuecker,[7] M. Steinke,[7] N. R. Barlow,[8] W. Bhimji,[8] J. T. Boyd,[8] N. Chevalier,[8] P. J. Clark,[8] W. N. Cottingham,[8] C. Mackay,[8] F. F. Wilson,[8] C. Hearty,[9] T. S. Mattison,[9] J. A. McKenna,[9] D. Thiessen,[9] S. Jolly,[10] P. Kyberd,[10] A. K. McKemey,[10] V. E. Blinov,[11] A. D. Bukin,[11] A. R. Buzykaev,[11] V. B. Golubev,[11] V. N. Ivanchenko,[11] A. A. Korol,[11] E. A. Kravchenko,[11] A. P. Onuchin,[11] S. I. Serednyakov,[11] Yu. I. Skovpen,[11] A. N. Yushkov,[11] D. Best,[12] M. Chao,[12] D. Kirkby,[12] A. J. Lankford,[12] M. Mandelkern,[12] S. McMahon,[12] R. K. Mommsen,[12] D. P. Stoker,[12] C. Buchanan,[13] H. K. Hadavand,[14] E. J. Hill,[14] D. B. MacFarlane,[14] H. P. Paar,[14] Sh. Rahatlou,[14] G. Raven,[14] U. Schwanke,[14] V. Sharma,[14] J. W. Berryhill,[15] C. Campagnari,[15] B. Dahmes,[15] N. Kuznetsova,[15] S. L. Levy,[15] O. Long,[15] A. Lu,[15] M. A. Mazur,[15] J. D. Richman,[15] W. Verkerke,[15] J. Beringer,[16] A. M. Eisner,[16] M. Grothe,[16] C. A. Heusch,[16] W. S. Lockman,[16] T. Pulliam,[16] T. Schalk,[16] R. E. Schmitz,[16] B. A. Schumm,[16] A. Seiden,[16] M. Turri,[16] W. Walkowiak,[16] D. C. Williams,[16] M. G. Wilson,[16] J. Albert,[17] E. Chen,[17] G. P. Dubois-Felsmann,[17] A. Dvoretskii,[17] D. G. Hitlin,[17] I. Narsky,[17] F. C. Porter,[17] A. Ryd,[17] A. Samuel,[17] S. Yang,[17] S. Jayatilleke,[18] G. Mancinelli,[18] B. T. Meadows,[18] M. D. Sokoloff,[18] T. Barillari,[19] F. Blanc,[19] P. Bloom,[19] W. T. Ford,[19] U. Nauenberg,[19] A. Olivas,[19] P. Rankin,[19] J. Roy,[19] J. G. Smith,[19] W. C. van Hoek,[19] L. Zhang,[19] J. L. Harton,[20] T. Hu,[20] A. Soffer,[20] W. H. Toki,[20] R. J. Wilson,[20] J. Zhang,[20] D. Altenburg,[21] T. Brandt,[21] J. Brose,[21] T. Colberg,[21] M. Dickopp,[21] R. S. Dubitzky,[21] A. Hauke,[21] E. Maly,[21] R. Müller-Pfefferkorn,[21] R. Nogowski,[21] S. Otto,[21] K. R. Schubert,[21] R. Schwierz,[21] B. Spaan,[21] L. Wilden,[21] D. Bernard,[22] G. R. Bonneaud,[22] F. Brochard,[22] J. Cohen-Tanugi,[22] S. T'Jampens,[22] Ch. Thiebaux,[22] G. Vasileiadis,[22] M. Verderi,[22] A. Anjomshoaa,[23] R. Bernet,[23] A. Khan,[23] D. Lavin,[23] F. Muheim,[23] S. Playfer,[23] J. E. Swain,[23] J. Tinslay,[23] M. Falbo,[24] C. Borean,[25] C. Bozzi,[25] L. Piemontese,[25] A. Sarti,[25] E. Treadwell,[26] F. Anulli,[27,*] R. Baldini-Ferroli,[27] A. Calcaterra,[27] R. de Sangro,[27] D. Falciai,[27] G. Finocchiaro,[27] P. Patteri,[27] I. M. Peruzzi,[27,*] M. Piccolo,[27] A. Zallo,[27] S. Bagnasco,[28] A. Buzzo,[28] R. Contri,[28] G. Crosetti,[28] M. Lo Vetere,[28] M. Macri,[28] M. R. Monge,[28] S. Passaggio,[28] F. C. Pastore,[28] C. Patrignani,[28] E. Robutti,[28] A. Santroni,[28] S. Tosi,[28] S. Bailey,[29] M. Morii,[29] G. J. Grenier,[30] U. Mallik,[30] J. Cochran,[31] H. B. Crawley,[31] J. Lamsa,[31] W. T. Meyer,[31] S. Prell,[31] E. I. Rosenberg,[31] J. Yi,[31] M. Davier,[32] G. Grosdidier,[32] A. Höcker,[32] H. M. Lacker,[32] S. Laplace,[32] F. Le Diberder,[32] V. Lepeltier,[32] A. M. Lutz,[32] T. C. Petersen,[32] S. Plaszczynski,[32] M. H. Schune,[32] L. Tantot,[32] G. Wormser,[32] R. M. Bionta,[33] V. Brigljević,[33] D. J. Lange,[33] K. van Bibber,[33] D. M. Wright,[33] A. J. Bevan,[34] J. R. Fry,[34] E. Gabathuler,[34] R. Gamet,[34] M. George,[34] M. Kay,[34] D. J. Payne,[34] R. J. Sloane,[34] C. Touramanis,[34] M. L. Aspinwall,[35] D. A. Bowerman,[35] P. D. Dauncey,[35] U. Egede,[35] I. Eschrich,[35] G. W. Morton,[35] J. A. Nash,[35] P. Sanders,[35] G. P. Taylor,[35] J. J. Back,[36] G. Bellodi,[36] P. Dixon,[36] P. F. Harrison,[36] H. W. Shorthouse,[36] P. Strother,[36] P. B. Vidal,[36] G. Cowan,[37] H. U. Flaecher,[37] S. George,[37] M. G. Green,[37] A. Kurup,[37] C. E. Marker,[37] T. R. McMahon,[37] S. Ricciardi,[37] F. Salvatore,[37] G. Vaitsas,[37] M. A. Winter,[37] D. Brown,[38] C. L. Davis,[38] J. Allison,[39] R. J. Barlow,[39] A. C. Forti,[39] P. A. Hart,[39] F. Jackson,[39] G. D. Lafferty,[39] A. J. Lyon,[39] N. Savvas,[39] J. H. Weatherall,[39] J. C. Williams,[39] A. Farbin,[40] A. Jawahery,[40] V. Lillard,[40] D. A. Roberts,[40] G. Blaylock,[41] C. Dallapiccola,[41] K. T. Flood,[41] S. S. Hertzbach,[41] R. Kofler,[41] V. B. Koptchev,[41] T. B. Moore,[41] H. Staengle,[41] S. Willocq,[41] R. Cowan,[42] G. Sciolla,[42] F. Taylor,[42] R. K. Yamamoto,[42] M. Milek,[43] P. M. Patel,[43] F. Palombo,[44] J. M. Bauer,[45] L. Cremaldi,[45] V. Eschenburg,[45] R. Kroeger,[45] J. Reidy,[45] D. A. Sanders,[45] D. J. Summers,[45] H. Zhao,[45] C. Hast,[46] P. Taras,[46] H. Nicholson,[47] C. Cartaro,[48] N. Cavallo,[48] G. De Nardo,[48] F. Fabozzi,[48,†] C. Gatto,[48] L. Lista,[48] P. Paolucci,[48] D. Piccolo,[48] C. Sciacca,[48] J. M. LoSecco,[49] J. R. G. Alsmiller,[50] T. A. Gabriel,[50] B. Brau,[51] J. Brau,[52] R. Frey,[52] M. Iwasaki,[52] C. T. Potter,[52] N. B. Sinev,[52] D. Strom,[52] E. Torrence,[52] F. Colecchia,[53] A. Dorigo,[53] F. Galeazzi,[53] M. Margoni,[53] M. Morandin,[53] M. Posocco,[53] M. Rotondo,[53] F. Simonetto,[53] R. Stroili,[53] G. Tiozzo,[53] C. Voci,[53] M. Benayoun,[54] H. Briand,[54] J. Chauveau,[54] P. David,[54] Ch. de la Vaissière,[54] L. Del Buono,[54] O. Hamon,[54] Ph. Leruste,[54] J. Ocariz,[54] M. Pivk,[54] L. Roos,[54] J. Stark,[54] P. F. Manfredi,[55] V. Re,[55] V. Speziali,[55] L. Gladney,[56] Q. H. Guo,[56] J. Panetta,[56] C. Angelini,[57] G. Batignani,[57] S. Bettarini,[57] M. Bondioli,[57] F. Bucci,[57] G. Calderini,[57] E. Campagna,[57] M. Carpinelli,[57] F. Forti,[57] M. A. Giorgi,[57] A. Lusiani,[57] G. Marchiori,[57] F. Martinez-Vidal,[57] M. Morganti,[57] N. Neri,[57] E. Paoloni,[57] M. Rama,[57] G. Rizzo,[57] F. Sandrelli,[57] G. Triggiani,[57] J. Walsh,[57] M. Haire,[58] D. Judd,[58] K. Paick,[58] L. Turnbull,[58] D. E. Wagoner,[58] N. Danielson,[59] P. Elmer,[59] C. Lu,[59] V. Miftakov,[59] J. Olsen,[59] A. J. S. Smith,[59] A. Tumanov,[59] E. W. Varnes,[59] F. Bellini,[60] G. Cavoto,[59,60] D. del Re,[60] R. Faccini,[14,60] F. Ferrarotto,[60] F. Ferroni,[60] M. Gaspero,[60] E. Leonardi,[60] M. A. Mazzoni,[60] S. Morganti,[60] G. Piredda,[60] F. Safai Tehrani,[60] M. Serra,[60] C. Voena,[60] S. Christ,[61] G. Wagner,[61] R. Waldi,[61] T. Adye,[62] N. De Groot,[62] B. Franek,[62] N. I. Geddes,[62] G. P. Gopal,[62] E. O. Olaiya,[62] S. M. Xella,[62] R. Aleksan,[63] S. Emery,[63] A. Gaidot,[63] P.-F. Giraud,[63] G. Hamel de Monchenault,[63] W. Kozanecki,[63] M. Langer,[63]

G. W. London,[63] B. Mayer,[63] G. Schott,[63] B. Serfass,[63] G. Vasseur,[63] Ch. Yeche,[63] M. Zito,[63] M. V. Purohit,[64] F. X. Yumiceva,[64] A. W. Weidemann,[64] K. Abe,[65] D. Aston,[65] R. Bartoldus,[65] N. Berger,[65] A. M. Boyarski,[65] O. L. Buchmueller,[65] M. R. Convery,[65] D. P. Coupal,[65] D. Dong,[65] J. Dorfan,[65] W. Dunwoodie,[65] R. C. Field,[65] T. Glanzman,[65] S. J. Gowdy,[65] E. Grauges-Pous,[65] T. Hadig,[65] V. Halyo,[65] T. Himel,[65] T. Hryn'ova,[65] M. E. Huffer,[65] W. R. Innes,[65] C. P. Jessop,[65] M. H. Kelsey,[65] P. Kim,[65] M. L. Kocian,[65] U. Langenegger,[65] D. W. G. S. Leith,[65] S. Luitz,[65] V. Luth,[65] H. L. Lynch,[65] H. Marsiske,[65] S. Menke,[65] R. Messner,[65] D. R. Muller,[65] C. P. O'Grady,[65] V. E. Ozcan,[65] A. Perazzo,[65] M. Perl,[65] S. Petrak,[65] B. N. Ratcliff,[65] S. H. Robertson,[65] A. Roodman,[65] A. A. Salnikov,[65] T. Schietinger,[65] R. H. Schindler,[65] J. Schwiening,[65] G. Simi,[65] A. Snyder,[65] A. Soha,[65] J. Stelzer,[65] D. Su,[65] M. K. Sullivan,[65] H. A. Tanaka,[65] J. Va'vra,[65] S. R. Wagner,[65] M. Weaver,[65] A. J. R. Weinstein,[65] W. J. Wisniewski,[65] D. H. Wright,[65] C. C. Young,[65] P. R. Burchat,[66] C. H. Cheng,[66] T. I. Meyer,[66] C. Roat,[66] W. Bugg,[67] M. Krishnamurthy,[67] S. M. Spanier,[67] J. M. Izen,[68] I. Kitayama,[68] X. C. Lou,[68] F. Bianchi,[69] M. Bona,[69] D. Gamba,[69] L. Bosisio,[70] G. Della Ricca,[70] S. Dittongo,[70] L. Lanceri,[70] P. Poropat,[70] L. Vitale,[70] G. Vuagnin,[70] R. Henderson,[71] R. S. Panvini,[72] Sw. Banerjee,[73] C. M. Brown,[73] D. Fortin,[73] P. D. Jackson,[73] R. Kowalewski,[73] J. M. Roney,[73] H. R. Band,[74] S. Dasu,[74] M. Datta,[74] A. M. Eichenbaum,[74] H. Hu,[74] J. R. Johnson,[74] R. Liu,[74] F. Di Lodovico,[74] A. K. Mohapatra,[74] Y. Pan,[74] R. Prepost,[74] S. J. Sekula,[74] J. H. von Wimmersperg-Toeller,[74] J. Wu,[74] S. L. Wu,[74] Z. Yu,[74] and H. Neal[75]

(BABAR Collaboration)

[1]*Laboratoire de Physique des Particules, F-74941 Annecy-le-Vieux, France*
[2]*Università di Bari, Dipartimento di Fisica and INFN, I-70126 Bari, Italy*
[3]*Institute of High Energy Physics, Beijing 100039, China*
[4]*University of Bergen, Inst. of Physics, N-5007 Bergen, Norway*
[5]*Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720*
[6]*University of Birmingham, Birmingham B15 2TT, United Kingdom*
[7]*Ruhr Universität Bochum, Institut für Experimentalphysik 1, D-44780 Bochum, Germany*
[8]*University of Bristol, Bristol BS8 1TL, United Kingdom*
[9]*University of British Columbia, Vancouver, BC, Canada V6T 1Z1*
[10]*Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom*
[11]*Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia*
[12]*University of California at Irvine, Irvine, California 92697*
[13]*University of California at Los Angeles, Los Angeles, California 90024*
[14]*University of California at San Diego, La Jolla, California 92093*
[15]*University of California at Santa Barbara, Santa Barbara, California 93106*
[16]*University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, California 95064*
[17]*California Institute of Technology, Pasadena, California 91125*
[18]*University of Cincinnati, Cincinnati, Ohio 45221*
[19]*University of Colorado, Boulder, Colorado 80309*
[20]*Colorado State University, Fort Collins, Colorado 80523*
[21]*Technische Universität Dresden, Institut für Kernund Teilchenphysik, D-01062 Dresden, Germany*
[22]*Ecole Polytechnique, LLR, F-91128 Palaiseau, France*
[23]*University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom*
[24]*Elon University, Elon University, North Carolina 27244-2010*
[25]*Università di Ferrara, Dipartimento di Fisica and INFN, I-44100 Ferrara, Italy*
[26]*Florida A&M University, Tallahassee, Florida 32307*
[27]*Laboratori Nazionali di Frascati dell'INFN, I-00044 Frascati, Italy*
[28]*Università di Genova, Dipartimento di Fisica and INFN, I-16146 Genova, Italy*
[29]*Harvard University, Cambridge, Massachusetts 02138*
[30]*University of Iowa, Iowa City, Iowa 52242*
[31]*Iowa State University, Ames, Iowa 50011-3160*
[32]*Laboratoire de l'Accélérateur Linéaire, F-91898 Orsay, France*
[33]*Lawrence Livermore National Laboratory, Livermore, California 94550*
[34]*University of Liverpool, Liverpool L69 3BX, United Kingdom*
[35]*University of London, Imperial College, London SW7 2BW, United Kingdom*
[36]*Queen Mary, University of London E1 4NS, United Kingdom*
[37]*University of London, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, United Kingdom*
[38]*University of Louisville, Louisville, Kentucky 40292*
[39]*University of Manchester, Manchester M13 9PL, United Kingdom*
[40]*University of Maryland, College Park, Maryland 20742*
[41]*University of Massachusetts, Amherst, Massachusetts 01003*
[42]*Massachusetts Institute of Technology, Laboratory for Nuclear Science, Cambridge, Massachusetts 02139*

[43]*McGill University, Montréal, Québec, Canada H3A 2T8*
[44]*Università di Milano, Dipartimento di Fisica and INFN, I-20133 Milano, Italy*
[45]*University of Mississippi, University, Mississippi 38677*
[46]*Université de Montréal, Laboratoire René J. A. Lévesque, Montréal, Québec, Canada H3C 3J7*
[47]*Mount Holyoke College, South Hadley, Massachusetts 01075*
[48]*Università di Napoli Federico II, Dipartimento di Scienze Fisiche and INFN, I-80126 Napoli, Italy*
[49]*University of Notre Dame, Notre Dame, Indiana 46556*
[50]*Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831*
[51]*Ohio State University, 174 W. 18th Ave., Columbus, Ohio 43210*
[52]*University of Oregon, Eugene, Oregon 97403*
[53]*Università di Padova, Dipartimento di Fisica and INFN, I-35131 Padova, Italy*
[54]*Universités Paris VI et VII, Lab de Physique Nucléaire H. E., F-75252 Paris, France*
[55]*Università di Pavia, Dipartimento di Elettronica and INFN, I-27100 Pavia, Italy*
[56]*University of Pennsylvania, Philadelphia, Pennsylvania 19104*
[57]*Università di Pisa, Scuola Normale Superiore and INFN, I-56010 Pisa, Italy*
[58]*Prairie View A&M University, Prairie View, Texas 77446*
[59]*Princeton University, Princeton, New Jersey 08544*
[60]*Università di Roma La Sapienza, Dipartimento di Fisica and INFN, I-00185 Roma, Italy*
[61]*Universität Rostock, D-18051 Rostock, Germany*
[62]*Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, United Kingdom*
[63]*DAPNIA, Commissariat à l'Energie Atomique/Saclay, F-91191 Gif-sur-Yvette, France*
[64]*University of South Carolina, Columbia, South Carolina 29208*
[65]*Stanford Linear Accelerator Center, Stanford, California 94309*
[66]*Stanford University, Stanford, California 94305-4060*
[67]*University of Tennessee, Knoxville, Tennessee 37996*
[68]*University of Texas at Dallas, Richardson, Texas 75083*
[69]*Università di Torino, Dipartimento di Fisica Sperimentale and INFN, I-10125 Torino, Italy*
[70]*Università di Trieste, Dipartimento di Fisica and INFN, I-34127 Trieste, Italy*
[71]*TRIUMF, Vancouver, British Columbia, Canada V6T 2A3*
[72]*Vanderbilt University, Nashville, Tennessee 37235*
[73]*University of Victoria, Victoria, British Columbia, Canada V8W 3P6*
[74]*University of Wisconsin, Madison, Wisconsin 53706*
[75]*Yale University, New Haven, Connecticut 06511*

(Received 5 December 2002; published 24 April 2003)

We measure the B^0 lifetime τ_{B^0} and the B^0-$\overline{B}^0$ oscillation frequency Δm_d with a sample of approximately 14000 exclusively reconstructed $B^0 \to D^{*-} \ell^+ \nu_\ell$ signal events, selected from 23 million $B\overline{B}$ pairs recorded at the $\Upsilon(4S)$ resonance with the BABAR detector at the Stanford Linear Accelerator Center. The decay position of the other B is determined with the remaining tracks in the event, and its b-quark flavor at the time of decay is determined with a tagging algorithm that exploits the correlation between the flavor of the b quark and the charges of its decay products. The lifetime and oscillation frequencies are measured simultaneously with an unbinned maximum-likelihood fit that uses, for each event, the measured difference in decay times of the two B mesons (Δt), the calculated uncertainty on Δt, the signal and background probabilities, and b-quark tagging information for the other B. The results are $\tau_{B^0} = (1.523^{+0.024}_{-0.023} \pm 0.022)$ ps and $\Delta m_d = (0.492 \pm 0.018 \pm 0.013)$ ps^{-1}. The statistical correlation coefficient between τ_{B^0} and Δm_d is -0.22.

DOI: 10.1103/PhysRevD.67.072002 PACS number(s): 13.25.Hw, 11.30.Er, 12.15.Hh, 14.40.Nd

I. INTRODUCTION AND ANALYSIS OVERVIEW

The time evolution of B^0 mesons is governed by the overall decay rate $1/\tau_{B^0}$ and the B^0-$\overline{B}^0$ oscillation frequency Δm_d. The phenomenon of particle-antiparticle oscillations or "mixing" has been observed in neutral mesons containing a down quark and either a strange quark (K mesons) [1] or a bottom quark (B mesons) [2]. In the standard model of particle physics, mixing is the result of second-order charged weak interactions involving box diagrams containing virtual quarks with charge 2/3. In B mixing, the diagrams containing the top quark dominate due to the large mass of the top quark. Therefore, the mixing frequency is sensitive to the Cabibbo-Kobayashi-Maskawa quark-mixing matrix element V_{td} [3]. In the neutral K meson system, mixing also has contributions from real intermediate states accessible to both a K^0 and a $\overline{K}^0$ meson. Real intermediate states lead to a difference in the decay rate for the two mass eigenstates of

*Also with Università di Perugia, Perugia, Italy.
†Also with Università della Basilicata, Potenza, Italy.

the neutral meson system. For the B system, the decay rate difference is expected to be of $O(10^{-2}-10^{-3})$ times smaller [4] than the average decay rate and the mixing frequency, and is ignored in this analysis.

We present a simultaneous measurement of the B^0 lifetime and oscillation frequency based on a sample of approximately 14000 exclusively reconstructed $B^0 \to D^{*-}\ell^+\nu_\ell$ decays [5] selected from a sample of 23 million $B\overline{B}$ events recorded at the $\Upsilon(4S)$ resonance with the BABAR detector [6] at the Stanford Linear Accelerator Center in 1999–2000. In this experiment, 9-GeV electrons and 3.1-GeV positrons, circulating in the SLAC e^+e^- storage ring PEP-II [7], annihilate to produce $B\overline{B}$ pairs moving along the e^- beam direction (z axis) with a Lorentz boost of $\beta\gamma=0.55$, allowing a measurement of the proper time difference between the two B decays, Δt.

The decay-time difference Δt between two neutral B mesons produced in a coherent P-wave state in an $\Upsilon(4S)$ event is governed by the probabilities of observing an unmixed event,

$$P(B^0\overline{B}^0 \to B^0\overline{B}^0) \propto e^{-|\Delta t|/\tau_{B^0}}(1+\cos\Delta m_d \Delta t), \qquad (1)$$

or a mixed event,

$$P(B^0\overline{B}^0 \to B^0B^0 \ \text{ or } \ \overline{B}^0\overline{B}^0) \propto e^{-|\Delta t|/\tau_{B^0}}(1-\cos\Delta m_d \Delta t). \qquad (2)$$

Therefore, if we measure Δt and identify the b-quark flavor of both B mesons at their time of decay, we can extract τ_{B^0} and Δm_d. In this analysis, one B is reconstructed in the mode $B^0 \to D^{*-}\ell^+\nu_\ell$, which has a measured branching fraction of $(4.60\pm0.21)\%$ [8]. Although the neutrino cannot be detected, the requirement of a reconstructed $D^{*-} \to \overline{D}^0\pi^-$ decay and a high-momentum lepton satisfying kinematic constraints consistent with the decay $B^0 \to D^{*-}\ell^+\nu_\ell$ allows the isolation of a signal sample with (65–89)% purity, depending on the D^0 decay mode and whether the lepton candidate is an electron or a muon. The charges of the final-state particles identify the meson as a B^0 or a $\overline{B}^0$. The remaining charged particles in the event, which originate from the other B (referred to as $B_{\rm tag}$), are used to identify, or ''tag,'' its flavor as a B^0 or a $\overline{B}^0$. The time difference $\Delta t = t_{D^*\ell} - t_{\rm tag} \approx \Delta z/\beta\gamma c$ is determined from the separation Δz of the decay vertices for the $D^{*-}\ell^+$ candidate and the tagging B along the boost direction. The average separation is about 250 μm.

The oscillation frequency and the average lifetime of the neutral B meson are determined simultaneously with an unbinned maximum-likelihood fit to the measured Δt distributions of events that are classified as mixed and unmixed. This is in contrast to most published measurements [8,9] in which only τ_{B^0} is measured, or Δm_d is measured with τ_{B^0} fixed to the world average. There are several reasons to measure the lifetime and oscillation frequency simultaneously. The statistical precision of this measurement for both τ_{B^0} and Δm_d is comparable to the uncertainty on the world average; hence, it is appropriate to measure both quantities rather than fixing the lifetime to the world average. Since mixed and unmixed events have different Δt distributions, the separation of mixed and unmixed events gives greater sensitivity to the Δt resolution function; as a result, the statistical uncertainty of τ_{B^0} is improved by approximately 15% [10]. Also, since B^+B^- events do not mix, we can use the Δt distributions for mixed and unmixed events to help discriminate between $B^0\overline{B}^0$ signal events and B^+B^- background events in the lifetime and mixing measurement.

There are three main experimental complications that affect the Δt distributions given in Eqs. (1) and (2). First, the tagging algorithm, which classifies events into categories c depending on the source of the available tagging information, incorrectly identifies the flavor of $B_{\rm tag}$ with a probability ω_c with a consequent reduction of the observed amplitude for the mixing oscillation by a factor $(1-2\omega_c)$. Second, the resolution for Δt is comparable to the lifetime and must be well understood. The probability density functions (PDF's) for the unmixed $(+)$ and mixed $(-)$ signal events can be expressed as the convolution of the underlying $\Delta t_{\rm true}$ distribution for tagging category c,

$$\frac{e^{-|\Delta t_{\rm true}|/\tau_{B^0}}}{4\tau_{B^0}}[1\pm(1-2\omega_c)\cos\Delta m_d \Delta t_{\rm true}], \qquad (3)$$

with a resolution function that depends on a set of parameters determined from the data. A final complication is that the sample of selected $B^0 \to D^{*-}\ell^+\nu_\ell$ candidates includes several types of background for which the Δt distributions must be determined.

To characterize the backgrounds, we select control samples of events enhanced in each type of background and determine the signal and the background probabilities for each event in the signal samples and the background control samples as described in Sec. IV. The measurement of Δz and the determination of Δt and the uncertainty on Δt ($\sigma_{\Delta t}$) for each event is discussed in Sec. V. The b-quark tagging algorithm is described in Sec. VI. In Sec. VII, we describe the unbinned maximum-likelihood fit. The physics model and Δt resolution function used to describe the measured Δt distribution for the signal are given in Sec. VIII. A combination of Monte Carlo simulation and data samples are used to determine the parameterization of the PDF's to describe the Δt distribution for each type of background, as described in Sec. IX. The likelihood is maximized in a simultaneous fit to the signal and background control samples to extract the B^0 lifetime τ_{B^0}, the mixing frequency Δm_d, the mistag probabilities ω_c, the signal Δt resolution parameters $\vec{q}_c$, the background Δt model parameters, and the fraction of $B^+ \to D^{*-}\ell^+\nu_\ell X$ decays in the signal sample. The results of the fit are given in Sec. X. Cross-checks are described in Sec. XI and systematic uncertainties are summarized in Sec. XII.

II. THE BABAR DETECTOR

The BABAR detector is described in detail elsewhere [6]. The momenta of charged particles are measured with a combination of a five-layer silicon vertex tracker (SVT) and a 40-layer drift chamber (DCH) in a 1.5-T solenoidal magnetic

field. A detector of internally reflected Cherenkov radiation (DIRC) is used for charged particle identification. Kaons are identified with a neural network based on the likelihood ratios calculated from dE/dx measurements in the SVT and DCH, and from the observed pattern of Cherenkov light in the DIRC. A finely segmented CsI(Tl) electromagnetic calorimeter (EMC) is used to detect photons and neutral hadrons, and to identify electrons. Electron candidates are required to have a ratio of EMC energy to track momentum, an EMC cluster shape, DCH dE/dx, and DIRC Cherenkov angle all consistent with the electron hypothesis. The instrumented flux return (IFR) contains resistive plate chambers for muon and long-lived neutral hadron identification. Muon candidates are required to have IFR hits located along the extrapolated DCH track, an IFR penetration length, and an energy deposit in the EMC consistent with the muon hypothesis.

III. DATA SAMPLES

The data used in this analysis were recorded with the BABAR detector at the PEP-II storage ring in the period October 1999 to December 2000. The total integrated luminosity of the data set is 20.6 fb^{-1} collected at the $\Upsilon(4S)$ resonance and 2.6 fb^{-1} collected about 40 MeV below the $\Upsilon(4S)$ (off-resonance data). The corresponding number of produced $B\overline{B}$ pairs is 23 million.

Samples of Monte Carlo simulated $B\overline{B}$ and $c\overline{c}$ events, generated with a GEANT3 [11] detector simulation, are analyzed through the same analysis chain as the data to check for biases in the extracted physics parameters and are also used to develop models for describing physics and detector resolution effects. However, the values of the parameters used in these models are determined with data. The equivalent luminosity of this simulated sample is approximately equal to that of the data for $B\overline{B}$ events and about half that of data for $c\overline{c}$ events. In addition, we generate signal Monte Carlo samples in which one neutral B meson in every event decays to $D^{*-}\ell^+\nu_\ell$, with $D^{*-}\to\overline{D}^0\pi^-$, and the other neutral B meson decays to any final state [12]. The D^0 then decays to one of the four final states reconstructed in this analysis (described in the next section). The equivalent luminosity of the simulated signal samples is between 2 and 8 times that of the data, depending on the D^0 decay mode.

IV. EVENT SELECTION AND CHARACTERIZATION

We select events containing a fully reconstructed D^{*-} and an identified oppositely charged electron or muon. This $D^{*-}\ell^+$ pair is then required to pass kinematic cuts that enhance the contribution of semileptonic $B^0\to D^{*-}\ell^+\nu_\ell$ decays. In addition to the signal sample, we select several control samples that are used to characterize the main sources of background.

We define the following classification of the sources of signal and background that we expect to contribute to this sample. The nomenclature shown in italics will be used throughout this paper to define signal and all possible types of background. Events are classified according to the D^{*-} candidate reconstruction status and the source of the lepton candidate.

(1) Events with a correctly reconstructed D^{*-} candidate.
  (a) Events that originate from $B\overline{B}$ events.
    (i) Events with a correctly identified lepton candidate.
      (A) *Signal*—$B^0\to D^{*-}\ell^+\nu_\ell$ (X) decays, where the D^{*-} and lepton originate from a common point. (X) indicates the possibility of one or more pions or photons from the direct decay of the parent B or from the decay of short-lived intermediate resonances (radially and orbitally excited D states).
      (B) *Uncorrelated-lepton background*—events in which the lepton does not come from the primary B decay that produced the D^{*-}: $(B\to D^{*-}X$, other $B\to\ell^+Y)$ or $(B\to D^{*-}X, X\to\ell^+Y)$.
      (C) *Charged B background*—$B^+\to D^{*-}\ell^+\nu_\ell X$.
    (ii) *Fake-lepton background*—events with a misidentified lepton candidate.
  (b) *Continuum background*—$c\overline{c}\to D^{*-}X$.
(2) *Combinatorial-D* background*—events with a misreconstructed D^{*-} candidate.

A. Lepton candidates

Lepton candidates are defined as tracks with momentum greater than 1.2 GeV/c in the $\Upsilon(4S)$ rest frame. For the $D^{*-}e^+$ samples, the electron candidate passes selection criteria with a corresponding electron identification efficiency of about 90% and hadron misidentification less than 0.2%. For the $D^{*-}\mu^+$ samples, the muon candidate passes selection criteria with a corresponding muon identification efficiency of about 70% and hadron misidentification between 2% and 3%. The particle identification criteria in BABAR are described in detail elsewhere [13]. A sample enriched in fake-lepton background is also selected, where $D^{*-}\ell^+$ candidates are accepted if the lepton *fails* both electron and muon selection criteria looser than those required for lepton candidates. This sample is used to determine the fraction and Δt distribution of the fake-lepton background.

B. D^{*-} candidates

D^{*-} candidates are selected in the decay mode $D^{*-}\to\overline{D}^0\pi^-$. The $\overline{D}^0$ candidate is reconstructed in the modes $K^+\pi^-$, $K^+\pi^-\pi^+\pi^-$, $K^+\pi^-\pi^0$ and $K^0_s\pi^+\pi^-$ The daughters of the $\overline{D}^0$ decay are selected according to the following definitions. π^0 candidates are reconstructed from two photons with energy greater than 30 MeV each, and an invariant mass between 119.2 and 150.0 MeV/c^2 and a total energy greater then 200 MeV. The mass of the photon pair is constrained to the π^0 mass and the photon pair is kept as a π^0 candidate if the χ^2 probability of the fit is greater than 1%. K^0_s candidates are reconstructed from a pair of charged particles with invariant mass within 15 MeV/c^2 of the K^0_s mass. The pair of tracks is retained as a K^0_s candidate if the χ^2

probability that the two tracks form a common vertex is greater than 1%. Charged kaon candidates satisfy loose kaon criteria [13] for the $K^+\pi^-$ mode and tighter criteria for the $K^+\pi^-\pi^+\pi^-$ and $K^+\pi^-\pi^0$ modes. For the $K^+\pi^-\pi^0$ and $K^0_s\pi^+\pi^-$ modes, a likelihood is calculated as the square of the decay amplitude in the Dalitz plot for the three-body candidate, based on measured amplitudes and phases [14]. The candidate is retained if the likelihood is greater than 10% of its maximum value across the Dalitz plot. This criterion rejects about 95% (97%) of uniform background and has a signal efficiency of about 62% (48%) for the $K^+\pi^-\pi^0$ ($K^0_s\pi^+\pi^-$) mode if the real signal is described by the results in Ref. [14].

$\overline{D}^0$ candidates in the $K^+\pi^-$, $K^+\pi^-\pi^+\pi^-$, and $K^0_s\pi^+\pi^-$ modes ($K^+\pi^-\pi^0$ mode) are selected if they have an invariant mass within 17 MeV/c^2 (34 MeV/c^2) of the D^0 mass. The invariant mass of the daughters is constrained to the D^0 mass and the tracks are constrained to a common vertex in a simultaneous fit. The $\overline{D}^0$ candidate is retained if the χ^2 probability of the fit is greater than 0.1%.

The low-momentum pion candidates for the $D^{*-}\to\overline{D}^0\pi^-$ decay are selected with total momentum less than 450 MeV/c in the $\Upsilon(4S)$ rest frame and momentum transverse to the beamline greater than 50 MeV/c. The momentum of the D^{*-} candidate in the $\Upsilon(4S)$ rest frame is required to be between 0.5 and 2.5 GeV/c. The requirements on the momenta of the low-momentum pion and D^{*-} candidates retain essentially all signal events and reject higher momentum D^{*-} from continuum events.

D^{*-} candidates are retained if $m(D^*)-m(D^0)$ is less than 165 MeV/c^2, where $m(D^*)$ is the candidate $\overline{D}^0\pi^-$ mass calculated with the candidate $\overline{D}^0$ mass constrained to the true D^0 mass, $m(D^0)$. Note that the $m(D^*)-m(D^0)$ distribution has a kinematic threshold at the mass of the π^-, and a peak at 145.5 MeV/c^2 with a resolution of 1 MeV/c^2 or better. We have retained the sideband of the $m(D^*)-m(D^0)$ distribution for studies of combinatorial-D^* background.

C. $D^{*-}\ell^+$ candidates

$D^{*-}\ell^+$ candidates are rejected if $|\cos\theta^*_{\rm thrust}|\geq 0.85$, where $\theta^*_{\rm thrust}$ is the angle between the thrust axis of the $D^{*-}\ell^+$ candidate and the thrust axis of the remaining charged and neutral particles in the event. The distribution of $|\cos\theta^*_{\rm thrust}|$ is peaked at 1 for jetlike continuum events and is flat for more spherical $B\overline{B}$ events.

$D^{*-}\ell^+$ candidates are retained if the following criteria are met: the χ^2 probability of the fit of the lepton, π^-, and $\overline{D}^0$ candidates to a common vertex is greater than 1%; the decay point of $B_{\rm tag}$ is determined from at least two tracks; the fit that determines the distance Δz between the two B decays along the beamline converges; the time between decays (Δt) calculated from Δz is less than 18 ps; and the calculated error on Δt ($\sigma_{\Delta t}$) is less than 1.8 ps. See Sec. V for details on the determination of the decay point of $B_{\rm tag}$ and the calculation of Δt and $\sigma_{\Delta t}$.

We define two angular quantities for each $D^{*-}\ell^+$ candidate to classify them into a sample enriched in $B^0\to D^{*-}\ell^+\nu_\ell$ signal events, and a sample enriched in *uncorrelated-lepton* background events. The first angle is $\theta_{D^*,\ell}$, the angle between the D^{*-} and lepton candidate in the $\Upsilon(4S)$ rest frame. The second is $\theta_{B,D^*\ell}$, the inferred angle between the direction of the B^0 and the vector sum of the D^{*-} and lepton candidate momenta, calculated in the $\Upsilon(4S)$ rest frame. Since we do not know the direction of the B^0, we calculate the cosine of $\theta_{B,D^*\ell}$ from the following equation, in which we assume that the only B decay particle missed in the reconstruction is a massless neutrino:

$$\cos\theta_{B,D^*\ell}\equiv\frac{-(m^2_{B^0}+m^2_{D^*\ell}-2E_BE_{D^*\ell})}{2|\vec{p}_B||\vec{p}_{D^*\ell}|}. \qquad (4)$$

All quantities in Eq. (4) are defined in the $\Upsilon(4S)$ rest frame. The energy and the magnitude of the momentum of the B are calculated from the e^+e^- center-of-mass energy and the B^0 mass. For true $B^0\to D^{*-}\ell^+\nu_\ell$ events, $\cos\theta_{B,D^*\ell}$ lies in the physical region $[-1,+1]$, except for detector resolution effects. Backgrounds lie inside and outside the range $[-1,+1]$. We also calculate the same angle with the lepton momentum direction reflected through the origin in the $\Upsilon(4S)$ rest frame: $\theta_{B,D^*(-\ell)}$. This angle is used to select samples enriched in uncorrelated-lepton background.

A sample enhanced in $B^0\to D^{*-}\ell^+\nu_\ell$ signal events (called the *opposite-side* sample) is composed of $D^{*-}\ell^+$ candidates with $\cos\theta_{D^*\ell}<0$ and $|\cos\theta_{B,D^*\ell}|<1.1$. Samples are defined for lepton candidates that satisfy the criteria for an electron, a muon and a fake lepton. The first two samples are the signal samples, and the latter is the *fake-lepton* control sample.

An additional background control sample, representative of the uncorrelated-lepton background and called the *same-side* sample, is composed of $D^{*-}\ell^+$ candidates satisfying $\cos\theta_{D^*\ell}\geq 0$ and $|\cos\theta_{B,D^*(-\ell)}|<1.1$. We use $\cos\theta_{B,D^*(-\ell)}$ rather than $\cos\theta_{B,D^*\ell}$ because, in Monte Carlo simulation, the distribution of $\cos\theta_{B,D^*(-\ell)}$ in this control sample is similar to the distribution of $\cos\theta_{B,D^*\ell}$ for uncorrelated-lepton background in the signal sample, whereas the distribution of $\cos\theta_{B,D^*\ell}$ in the background control sample is systematically different.

D. Signal and background subsamples

Approximately 68000 candidates pass the above selection criteria. These candidates are distributed over two signal samples and ten background control samples defined by the following characteristics:

(1) whether the data were recorded on or off the $\Upsilon(4S)$ resonance (two choices);

(2) whether the candidate lepton is *same side* or *opposite side* to the D^{*-} candidate (two choices);

(3) whether the lepton candidate passes the criteria for an electron, a muon, or a fake lepton (three choices).

The signal samples are the electron and muon samples in the opposite-side, on-resonance data.

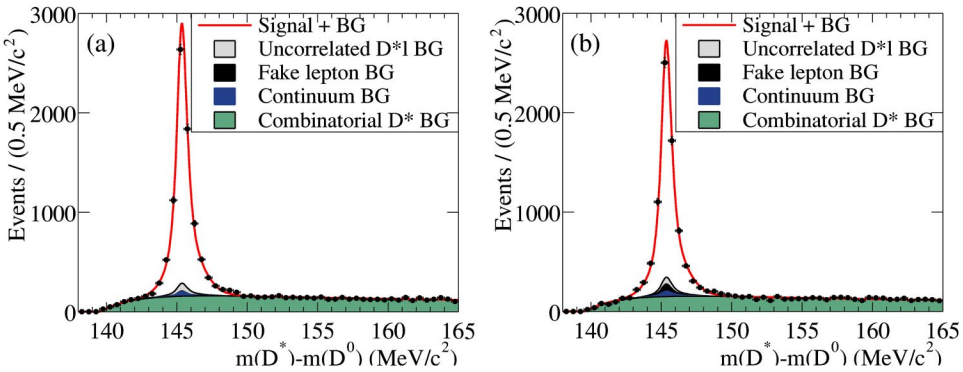

FIG. 1. $m(D^*)-m(D^0)$ distribution for events passing all selection criteria for $B^0 \to D^{*-}\ell^+\nu_\ell$, with (a) an electron or (b) a muon candidate. The points correspond to the data. The curve is the result of a simultaneous unbinned maximum likelihood fit to this sample of events and a number of background control samples. The shaded distributions correspond to the four types of background (BG) described in the text. The charged B background is not shown separately.

The combinatorial-D^* background can be distinguished from events with a real D^{*-} in a plot of the mass difference $m(D^*)-m(D^0)$. The $m(D^*)-m(D^0)$ distributions for the samples of signal events (opposite-side $D^{*-}e^+$ and $D^{*-}\mu^+$ candidates in on-resonance data) are shown as data points in Fig. 1 for (a) electron candidates and (b) muon candidates. The contributions of the three types of background that contain a real D^{*-} (continuum, uncorrelated lepton, and fake lepton, together called the *peaking background*), except for the charged B background, are also shown in the plots. The $m(D^*)-m(D^0)$ distributions for five background control samples used for determining the background levels in the signal sample are shown as data points in Fig. 2: opposite-side (a) $D^{*-}e^+$ and (b) $D^{*-}\mu^+$ candidates in off-resonance data; same-side (c) $D^{*-}e^+$ and (d) $D^{*-}\mu^+$ candidates in on-resonance data; (e) opposite-side D^{*-}-fake-lepton candidates in on-resonance data. The remaining five background control samples are useful for determining the background levels in the first five control samples.

Each of the 12 samples described above is further divided into 30 subsamples according to the following characteristics that affect the $m(D^*)-m(D^0)$ or Δt distributions.

(1) The π^- from the D^{*-} decay reconstructed in the SVT only, or in the SVT and DCH (two choices): The $m(D^*)-m(D^0)$ resolution is worse when the π^- is reconstructed only in the SVT.

(2) The $\overline{D}^0$ candidate reconstructed in the mode $K^+\pi^-$ or $K^+\pi^-\pi^0$ or ($K^+\pi^-\pi^+\pi^-$ or $K^0_S\pi^+\pi^-$) (three choices): The level of contamination from combinatorial-D^* background and the $m(D^*)-m(D^0)$ resolution depend on the $\overline{D}^0$ decay mode.

(3) The b-tagging information used for the other B (five choices; see Sec. VI): The level of contamination from each type of background and the Δt resolution parameters depend on the tagging information.

This allows subdivision into 360 samples. In the unbinned maximum likelihood fits to the $m(D^*)-m(D^0)$ and $(\Delta t, \sigma_{\Delta t})$ distributions, individual fit parameters are shared among different sets of subsamples based on physics motivation and observations from the data.

We fit the $m(D^*)-m(D^0)$ distributions to determine signal and background probabilities for each of the 360 subsamples. The peak due to real D^{*-} candidates is modeled by

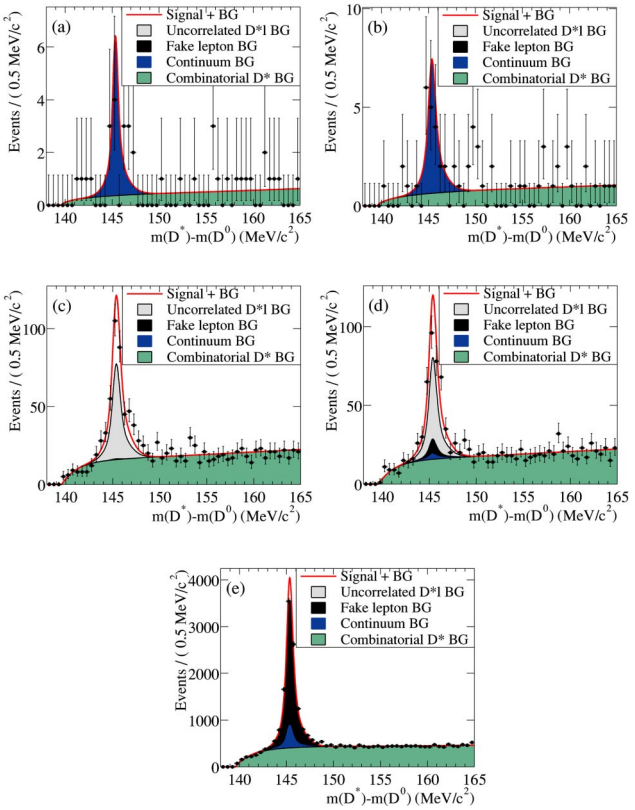

FIG. 2. $m(D^*)-m(D^0)$ distribution for events passing all selection criteria in background control samples: opposite-side (a) $D^{*-}e^+$ and (b) $D^{*-}\mu^+$ candidates in off-resonance data; same-side (c) $D^{*-}e^+$ and (d) $D^{*-}\mu^+$ candidates in on-resonance data; (e) opposite-side D^{*-}-fake-lepton candidates in on-resonance data. The points correspond to the data. The curve is the result of a simultaneous unbinned maximum likelihood fit to this sample of events, the signal sample, and a number of other background control samples. The shaded distributions correspond to the four types of background described in the text. The charged B background is not shown separately.

TABLE I. Peak yields and the fraction of them that are due to continuum, fake-lepton, and uncorrelated-lepton events. Also shown is the combinatorial-D^* fraction of total events in a $m(D^*)-m(D^0)$ signal window for the signal and background control samples in on-resonance data. Peak yields include the peaking backgrounds. The signal window for combinatorial-D^* background fractions is defined as (143–148) MeV/c^2. e, μ, and fake indicate the type of lepton candidate: electron, muon or fake lepton.

Category	Peak yield	$f_{\rm cont}(\%)$	$f_{\rm fake}(\%)$	$f_{\rm uncor}(\%)$	$f_{\rm comb}(\%)$
Opposite side					
e	7008 ± 91	1.5 ± 0.4	0.168 ± 0.004	3.1 ± 0.4	17.9 ± 0.2
μ	6569 ± 88	2.3 ± 0.6	2.67 ± 0.07	2.9 ± 0.5	18.4 ± 0.3
Fake	8770 ± 108	12.8 ± 1.3	72.4 ± 1.8	0.7 ± 1.6	31.4 ± 0.2
Same side					
e	306 ± 21	<5.9[a]	0.53 ± 0.04	56.9 ± 7.0	34.0 ± 1.3
μ	299 ± 20	5.1 ± 3.6	8.9 ± 0.6	48.9 ± 8.0	34.4 ± 1.3
Fake	1350 ± 45	20.4 ± 4.1	74.4 ± 5.4	3.6 ± 7.8	42.6 ± 0.6

[a]90% C.L.

the sum of two Gaussian distributions; the mean and variance of both the Gaussian distributions, as well as the relative normalization of the two Gaussians, are free parameters in the fit. We model the shape of the combinatorial-D^* background with the function

$$\frac{1}{N}\left[1-\exp\left(-\frac{\delta m-m_{\pi^-}}{c_1}\right)\right]\left(\frac{\delta m}{m_{\pi^-}}\right)^{c_2}, \tag{5}$$

where $\delta m \equiv m(D^*)-m(D^0)$, N is a normalization constant, m_{π^-} is the mass of the π^-, and c_1 and c_2 are free parameters in the fit. An initial unbinned maximum likelihood fit is performed to determine the shape parameters describing the peak and combinatorial-D^* background. Separate values of the five parameters describing the shape of the peak are used for the six subsamples defined by (1) whether the π^- candidate is tracked in the SVT only or in both the SVT and DCH (two choices), and (2) the three types of $\overline{D}^0$ decay modes. Each of these six groups that use separate peak parameters is further subdivided into 12 subgroups that each uses a different set of the two combinatorial-D^* shape parameters but the same set of peak parameters. Ten of these 12 subgroups are defined by the five tagging categories for the large signal samples and for the fake-lepton control samples, in on-resonance data. The other two subgroups are defined as same-side, on-resonance samples and all off-resonance samples.

Once the peak and combinatorial-D^* shape parameters have been determined, we fix the shape parameters and determine the peak and combinatorial-D^* yields in each of the 360 subsamples with an unbinned extended maximum-likelihood fit.

The total peak yields in the signal sample and each background control sample are then used to determine the amount of true signal and each type of peaking background in the $m(D^*)-m(D^0)$ peak of each sample as follows.

(1) *Continuum background*—For each subsample in on-resonance data, the peak yield of the corresponding subsample in off-resonance data is scaled by the relative integrated luminosity for on- and off-resonance data, to determine the continuum-background yields in on-resonance data.

(2) *Fake-lepton background*—Particle identification efficiencies and misidentification probabilities for the electron, muon, and fake-lepton selection criteria are measured in separate data samples as a function of laboratory momentum, polar angle, and azimuthal angle, for true electrons, muons, pions, kaons, and protons. $B^0\overline{B}^0$ and B^+B^- Monte Carlo simulations are used to determine the measured laboratory momentum, polar angle, and azimuthal angle distributions for true electrons, muons, pions, kaons and protons that pass all selection criteria for $D^{*-}\ell^+$ candidates, except the lepton (or fake-lepton) identification criteria. These distributions are combined with the measured particle identification efficiencies and misidentification probabilities to determine the momentum- and angle-weighted probabilities for a true lepton or true hadron to pass the criteria for a lepton or a fake lepton in each of the $D^{*-}\ell^+$ signal and background control samples. We then use these efficiencies and misidentification probabilities, and the observed number of electron, muon and fake-lepton candidates in each subsample in data, after removing the continuum background contribution, to determine the number of true leptons and fake leptons (hadrons) in each control sample.

(3) *Uncorrelated-lepton background*—The relative efficiencies for signal and uncorrelated-lepton events to pass the criteria for same-side and opposite-side samples are calculated from Monte Carlo simulation. These efficiencies and the $m(D^*)-m(D^0)$ peak yields, after removing the continuum and fake-lepton background contributions, are used to determine the number of uncorrelated-lepton events in each subsample.

The peak yields and continuum, fake-lepton, and uncorrelated-lepton fractions of the peak yield, as well as the combinatorial-D^* fraction of all events in a $m(D^*)-m(D^0)$ signal window, are shown in Table I for the signal and background control samples in on-resonance data. The peak yields include the peaking backgrounds. The signal window is defined as (143–148) MeV/c^2 for the calculation of the combinatorial-D^* background fractions. Table II

TABLE II. Peak yields and the combinatorial-D^* background fraction of total events in a $m(D^*)-m(D^0)$ signal window for different divisions of the signal sample (opposite-side lepton candidates in on-resonance data). In the first block, the signal sample is divided according to the reconstruction status of the π^- from the D^{*-} decay; the second block by the $\overline{D}^0$ decay mode; and the third block by the *b*-tagging information (see Sec. VI). The signal window for combinatorial-D^* background fractions is defined as (143–148) MeV/c^2.

Category	Peak yield	$f_{\rm comb}$(%)
e		
SVT only	5427±81	19.5±0.3
DCH and SVT	1581±41	11.8±0.4
μ		
SVT only	5053±78	20.3±0.3
DCH and SVT	1517±41	11.1±0.4
e		
$K\pi$	2623±53	7.0±0.3
$K\pi\pi\pi$ and $K^0_s\pi\pi$	2219±54	28.6±0.5
$K\pi\pi^0$	2166±51	16.9±0.5
μm		
$K\pi$	2491±52	7.4±0.3
$K\pi\pi\pi$ and $K^0_s\pi\pi$	1939±51	30.9±0.5
$K\pi\pi^0$	2139±50	16.1±0.4
e		
lepton	783±29	8.2±0.6
kaon	2565±55	17.9±0.4
NT1	630±27	14.3±0.8
NT2	921±33	20.9±0.7
NT3	2108±51	20.7±0.5
μ		
lepton	746±28	8.3±0.6
kaon	2393±53	18.6±0.4
NT1	545±25	15.1±0.8
NT2	958±34	19.4±0.7
NT3	1928±49	21.8±0.5

shows the peak yields and the combinatorial-D^* background fractions for different divisions of the signal sample (opposite-side lepton candidates in on-resonance data). This table demonstrates that the background levels vary significantly among subgroups of the signal sample.

Finally, we use the calculated fractions and fitted shapes of the background sources in each control sample to estimate the probability of each candidate to be signal or each type of background (combinatorial D^*, continuum, fake lepton, or uncorrelated lepton) when we fit the $(\Delta t, \sigma_{\Delta t})$ distribution to determine the lifetime and mixing parameters. We take advantage of the fact that charged and neutral B decays have different decay-time distributions (because the charged B does not mix) to determine the fraction of charged B background events in the fit to $(\Delta t, \sigma_{\Delta t})$.

V. DECAY-TIME MEASUREMENT

The decay-time difference Δt between B decays is determined from the measured separation $\Delta z \equiv z_{D^*\ell} - z_{\rm tag}$ along the z axis between the $D^{*-}\ell^+$ vertex position ($z_{D^*\ell}$) and the flavor-tagging decay $B_{\rm tag}$ vertex position ($z_{\rm tag}$). This measured Δz is converted into Δt with the use of the $\Upsilon(4S)$ boost, determined from the beam energies [15] for each run. Since we cannot reconstruct the direction of the B meson for each event, we use the approximation $\Delta t \approx \Delta z/(\beta\gamma c)$. Without detector resolution effects, this approximation has a bias that depends on the sum of the proper decay times (t_1+t_2) of the two B mesons and their direction in the $\Upsilon(4S)$ rest frame [16]. Neither of these quantities can be measured because the $\Upsilon(4S)$ production point is not known and the momentum of the B is not fully reconstructed due to a missing neutrino. After integrating over t_1+t_2 and the B meson direction, the mean and rms of the bias are 0 and 0.2 ps, respectively.

The momentum and position vectors of the $\overline{D}^0$, π^-, and lepton candidates, and the average position of the e^+e^- interaction point (called the beam spot) in the plane transverse to the beam are used in a constrained fit to determine the position of the $D^{*-}\ell^+$ vertex. The beam-spot constraint is about 100 μm in the horizontal direction and 30 μm in the vertical direction, corresponding to the rms size of the beam in the horizontal direction and the approximate transverse flight path of the B in the vertical direction. The beam-spot constraint improves the resolution on $z_{D^*\ell}$ by about 20% in Monte Carlo simulation; the rms spread on the difference between the measured and true position of the $D^{*-}\ell^+$ vertex is about 70 μm (0.4 ps).

We determine the position of the $B_{\rm tag}$ vertex from all tracks in the event except the daughters of the $D^{*-}\ell^+$ candidate, using $K^0_S \to \pi^+\pi^-$ and $\Lambda \to p\pi^-$ candidates in place of their daughter tracks, and excluding tracks that are consistent with photon conversions. The same beam-spot constraint applied to the $B_{D^*\ell}$ vertex is also applied to the $B_{\rm tag}$ vertex. To reduce the influence of charm decay products, which bias the determination of the vertex position, tracks with a large contribution to the χ^2 of the vertex fit are iteratively removed until no track has a χ^2 contribution greater than 6 or only one track remains. The RMS spread on the difference between the measured and true position of the $B_{\rm tag}$ vertex in Monte Carlo simulation is about 160 μm (1.0 ps). Therefore, the Δt resolution is dominated by the z resolution of the tag vertex position.

For each event, we calculate the uncertainty on Δz ($\sigma_{\Delta z}$) due to uncertainties on the track parameters from the SVT and DCH hit resolution and multiple scattering, our knowledge of the beam-spot size, and the average B flight length in the vertical direction. The calculated uncertainty does not account for errors in pattern recognition in tracking, errors in associating tracks with the B vertices, the effects of misalignment within and between the tracking devices, or the error on the approximation we use to calculate Δt from Δz. The calculated uncertainties will also be incorrect if our assumptions for the amount of material in the tracking detectors or the beam-spot size or position are inaccurate. We use parameters in the Δt resolution model, measured with data, to account for uncertainties and biases introduced by these effects.

VI. FLAVOR TAGGING

All tracks in the event, except the daughter tracks of the $D^{*-}\ell^+$ candidate, are used to determine whether the $B_{\rm tag}$ decayed as a B^0 or a $\overline{B}^0$. This is called flavor tagging. We use five different types of flavor tag, or tagging categories, in this analysis. The first two tagging categories rely on the presence of a prompt lepton, or one or more charged kaons, in the event. The other three categories exploit a variety of inputs with a neural-network algorithm. The tagging algorithms are described briefly in this section; see Ref. [17] for more details.

Events are assigned a `lepton` tag if they contain an identified lepton with momentum in the $\Upsilon(4S)$ rest frame greater than 1.0 or 1.1 GeV/c for electrons and muons, respectively, thereby selecting mostly primary leptons from the decay of the b quark. If the sum of charges of all identified kaons is nonzero, the event is assigned a `kaon` tag. The final three tagging categories are based on the output of a neural network that uses as inputs the momentum and charge of the track with the maximum center-of-mass momentum, the number of tracks with significant impact parameters with respect to the interaction point, and the outputs of three other neural networks, trained to identify primary leptons, kaons, and low momentum pions from D^* decays. Depending on the output of the main neural network, events are assigned to an `NT1` (most certain), `NT2`, or `NT3` (least certain) tagging category. About 30% of events are in the `NT3` category, which has a mistag rate close to 50%. Therefore, these events are not sensitive to the mixing frequency, but they increase the sensitivity to the B^0 lifetime.

Tagging categories are mutually exclusive due to the hierarchical use of the tags. Events with a `lepton` tag and no conflicting `kaon` tag are assigned to the `lepton` category. If no `lepton` tag exists, but the event has a `kaon` tag, it is assigned to the `kaon` category. Otherwise events are assigned to corresponding neural network categories. The mistag rates are free parameters in the final fit. The final results are shown in Table III in Sec. X.

VII. FIT METHOD

We perform an unbinned fit simultaneously to events in each of the 12 signal and control samples (on or off resonance, opposite-or same-side lepton, electron or muon or fake lepton—indexed by s) that are further subdivided into 30 subsamples (tagging category, D^0 decay mode, with or without DCH hits for the pion from the D^* decay—indexed by c). We maximize the likelihood

$$\mathcal{L}=\prod_{s=1}^{12}\prod_{c=1}^{30}\prod_{k=1}^{N(s,c)} P_{s,c}(\delta m_k,\vec{x}_k;\vec{\eta}), \tag{6}$$

where k indexes the $N(s,c)$ events $\vec{x}_k$ in each of the 360 subsamples. The probability $P_{s,c}(\delta m_k,\vec{x}_k;\vec{\eta})$ of observing an event $(\delta m_k,\vec{x}_k)$, where $\vec{x}_k=(\Delta t,\sigma_{\Delta t},g)$, is calculated as a function of the parameters

$$\vec{\eta}=(f^{\rm comb}_{s,c},\vec{p}^{\,\rm comb}_{s,c},\vec{p}^{\,\rm peak}_{c},\vec{q}^{\,\rm comb}_{s,c},f^{\rm pkg}_{s,c,1},f^{\rm pkg}_{s,c,2},f^{\rm pkg}_{s,c,3},f_{B^+},\vec{q}^{\,\rm pkg}_{s,c,1},\overline{q}^{\,\rm pkg}_{s,c,2},\vec{q}^{\,\rm pkg}_{s,c,3},\vec{q}^{\,\rm sig}_{c},\vec{q}^{\,\rm ch}_{c}) \tag{7}$$

as

$$P_{s,c}(\delta m_k,\vec{x}_k;\vec{\eta})=f^{\rm comb}_{s,c}\mathcal{F}^{\rm comb}(\delta m;\vec{p}^{\,\rm comb}_{s,c})\mathcal{G}^{\rm comb}(\vec{x}_k;\vec{q}^{\,\rm comb}_{s,c})+(1-f^{\rm comb}_{s,c})\mathcal{F}^{\rm peak}(\delta m;\vec{p}^{\,\rm peak}_{c})$$
$$\times\left\{\sum_{j=1}^{3} f^{\rm pkg}_{s,c,j}\mathcal{G}^{\rm pkg}_{j}(\vec{x}_k;\vec{q}^{\,\rm pkg}_{s,c,j})+\left(1-\sum_{j=1}^{3} f^{\rm pkg}_{s,c,j}\right)[(1-f_{B^+})\mathcal{G}^{\rm sig}(\vec{x}_k;\vec{q}^{\,\rm sig}_{c})+f_{B^+}\mathcal{G}^{\rm ch}(\vec{x}_k;\vec{q}^{\,\rm ch}_{c})]\right\}, \tag{8}$$

where δm is the mass difference $m(D^*)-m(D^0)$ defined earlier. The symbol ''comb'' in the first term signifies combinatorial-D^* background. In the second term, the symbol ''pkg'' denotes peaking background and j indexes the three sources of peaking background (continuum, fake lepton, and uncorrelated lepton). In the last term, the parameter f_{B^+} describes the charged-B fraction in the sample after all other types of background are subtracted, and ''sig'' and ''ch'' label functions and parameters for the signal and charged-B background, respectively. The charged-B fraction is assumed to be identical for all categories. The index g is $+1$ (-1) for unmixed (mixed) events. By allowing different effective mistag rates for apparently mixed or unmixed events in the background functions $\mathcal{G}^{\rm comb}$ and $\mathcal{G}^{\rm pkg}$, we accommodate the different levels of background observed in mixed and unmixed samples. Functions labeled with $\mathcal{F}$ describe the probability of observing a particular value of δm while functions labeled with $\mathcal{G}$ give probabilities for values of Δt and $\sigma_{\Delta t}$ in category g. Parameters labeled with f describe the relative contributions of different types of events. Parameters labeled with $\vec{p}$ describe the shape of a δm distribution, and those labeled with $\vec{q}$ describe a $(\Delta t,\sigma_{\Delta t})$ shape. The parameters labeled with $\vec{p}$ and f have been determined by a set of fits to $m(D^*)-m(D^0)$ distributions described in Sec. IV, and are kept fixed in the fit to $(\Delta t,\sigma_{\Delta t})$.

Note that we make explicit assumptions that the δm peak shape, parametrized by $\vec{p}^{\,\rm peak}_{c}$, and the signal and charged-B background $(\Delta t,\sigma_{\Delta t})$ shapes, parametrized by $\vec{q}^{\,\rm sig}_{c}$ and $\vec{q}^{\,\rm ch}_{c}$, depend only on the subsample index c and not on the control sample index s. The first of these assumptions is supported by data, and simplifies the analysis of peaking background contributions. The second assumption reflects our expecta-

tion that the Δt distribution of signal and charged-B background events does not depend on whether they are selected in the signal sample or appear as a background in a control sample.

The ultimate aim of the fit is to obtain the B^0 lifetime and mixing frequency, which by construction are common to all sets of signal parameters $\vec{q}_c^{\,\mathrm{sig}}$. Most of the statistical power for determining these parameters comes from the signal sample, although the fake and uncorrelated background control samples also contribute due to their signal content (see Table I).

We bootstrap the full fit with a sequence of initial fits using reduced likelihood functions restricted to a partial set of samples, to determine the appropriate parameterization of the signal resolution function and the background Δt models, and to determine starting values for each parameter in the full fit.

(1) We first find a model that describes the Δt distribution for each type of event: signal, combinatorial-D^* background, and the three types of backgrounds that peak in the $m(D^*)-m(D^0)$ distribution. To establish a model, we use Monte Carlo samples that have been selected to correspond to only one type of signal or background event based on the true Monte Carlo information. These samples are used to determine the Δt model and the categories of events (e.g., tagging category, fake or real lepton) that can share each of the parameters in the model. Any subset of parameters can be shared among any subset of the 360 subsamples. We choose parametrizations and sharing of parameters that minimizes the number of different parameters while still providing an adequate description of the Δt distributions.

(2) We then find the starting values for the background parameters by fitting to each of the background-enhanced control samples in data, using the model (and sharing of parameters) determined in the previous step. Since these background control samples are not pure, we start with the purest control sample [combinatorial-D^* background events from the $m(D^*)-m(D^0)$ sideband] and move on to less pure control samples, always using the models established from earlier steps to describe the Δt distribution of the contamination from other backgrounds.

The result of the above two steps is a Δt model for each type of event and a set of starting values for all parameters in the fit. When we do the final fit, we fit all signal and control samples simultaneously (approximately 68000 events), leaving all parameters in the $\mathcal{G}$ functions free in the fit, except for a few parameters that either are highly correlated with other parameters or reach their physical limits. The total number of parameters that are free in the fit is 72. The physics parameters τ_{B^0} and Δm_d were kept hidden until all analysis details and the systematic errors were finalized, to minimize experimenter's bias. However, statistical errors on the parameters and changes in the physics parameters due to changes in the analysis were not hidden.

VIII. SIGNAL Δt MODEL

For signal events in a given tagging category c, the probability density function for Δt consists of a model of the intrinsic time dependence convolved with a Δt resolution function:

$$\begin{aligned}\mathcal{G}^{\mathrm{sig}}(\Delta t,\sigma_{\Delta t},g;\vec{q}_c^{\,\mathrm{sig}}) &= \left\{\frac{1}{4\tau_{B^0}}e^{-|\Delta t_{\mathrm{true}}|/\tau_{B^0}}[1+g(1-2\omega_c)\cos(\Delta m_d\Delta t_{\mathrm{true}})]\right\}\\ &\otimes\mathcal{R}(\delta\Delta t,\sigma_{\Delta t};\vec{q}_c^{\,\mathrm{sig}}),\end{aligned} \tag{9}$$

where $\mathcal{R}$ is a resolution function, which can be different for different event categories, g is $+1$ (-1) for unmixed (mixed) events, $\delta\Delta t$ is the residual $\Delta t-\Delta t_{\mathrm{true}}$, and ω_c is the mistag fraction for category c. To account for an observed correlation between the mistag rate and $\sigma_{\Delta t}$ in the kaon category (described in Sec. VIII A), we allow the mistag rate in the kaon category to vary as a linear function of $\sigma_{\Delta t}$:

$$\omega_{\mathrm{kaon}}=\alpha_{\mathrm{kaon}}\sigma_{\Delta t}+\omega_{\mathrm{kaon}}^{\mathrm{offset}}, \tag{10}$$

and allow both the slope α_{kaon} and the offset $\omega_{\mathrm{kaon}}^{\mathrm{offset}}$ to be free parameters. In addition, we allow the mistag fractions for B^0 tags and $\overline{B}^0$ tags to be different. We define $\Delta\omega=\omega_{B^0}-\omega_{\overline{B}^0}$ and $\omega=(\omega_{B^0}+\omega_{\overline{B}^0})/2$, so that

$$\omega_{B^0/\overline{B}^0}=\omega\pm\frac{1}{2}\Delta\omega. \tag{11}$$

The model for the intrinsic time dependence has 13 parameters: ω_c and $\Delta\omega_c$ for each of the five tagging categories, α_{kaon}, Δm_d and τ_{B^0}.

For the Δt resolution model, we use the sum of a single Gaussian distribution and the same Gaussian convolved with a one-sided exponential to describe the core part of the resolution function, plus a single Gaussian distribution to describe the contribution of ''outliers''—events in which the reconstruction error $\delta\Delta t$ is not described by the calculated uncertainty $\sigma_{\Delta t}$:

$$\begin{aligned}\mathcal{R}_{\mathrm{GExp+G}}(\delta\Delta t,\sigma_{\Delta t};s,\kappa,f,b^{\mathrm{out}},s^{\mathrm{out}},f^{\mathrm{out}}) &= fG(\delta\Delta t;0,s\sigma_{\Delta t})+(1-f-f^{\mathrm{out}})G(u-\delta\Delta t;0,s\sigma_{\Delta t})\\ &\otimes E(u;\kappa\sigma_{\Delta t})+f^{\mathrm{out}}G(\delta\Delta t;b^{\mathrm{out}},s^{\mathrm{out}}),\end{aligned} \tag{12}$$

where u is an integration variable in the convolution $G\otimes E$. The functions G and E are defined by

$$G(x;x_0,\sigma)\equiv\frac{1}{\sqrt{2\pi}\sigma}\exp[-(x-x_0)^2/(2\sigma)^2] \tag{13}$$

and

$$E(x;a)\equiv\begin{cases}\dfrac{1}{a}\exp(x/a) & \text{if } x\leqslant 0,\\ 0 & \text{if } x>0.\end{cases} \tag{14}$$

The exponential component is used to accommodate a bias due to tracks from charm decays on the B_{tag} side.

Since the outlier contribution is not expected to be described by the calculated error on each event, the last Gaussian term in Eq. (12) does not depend on $\sigma_{\Delta t}$. However, in the terms that describe the core of the resolution function [the first two terms on the right-hand side of Eq. (12)], the Gaussian width s and the constant κ in the exponential are scaled by $\sigma_{\Delta t}$. The scale factor s is introduced to accommodate an overall underestimate ($s>1$) or overestimate ($s<1$) of the errors for all events. The constant κ is introduced to account for residual charm decay products included in the $B_{\rm tag}$ vertex; κ is scaled by $\sigma_{\Delta t}$ to account for a correlation observed in Monte Carlo simulation between the mean of the $\delta\Delta t$ distribution and the measurement error $\sigma_{\Delta t}$.

The correlation between $\delta\Delta t$ and $\sigma_{\Delta t}$ is due to the fact that, in B decays, the vertex error ellipse for the D decay products is oriented with its major axis along the D flight direction, leading to a correlation between the D flight direction and the calculated uncertainty on the vertex position in z for the $B_{\rm tag}$ candidate. In addition, the flight length of the D in the z direction is correlated with its flight direction. Therefore, the bias in the measured $B_{\rm tag}$ position due to including D decay products is correlated with the D flight direction. Taking into account these two correlations, we conclude that D mesons that have a flight direction perpendicular to the z axis in the laboratory frame will have the best z resolution and will introduce the least bias in a measurement of the z position of the $B_{\rm tag}$ vertex, while D mesons that travel forward in the laboratory will have poorer z resolution and will introduce a larger bias in the measurement of the $B_{\rm tag}$ vertex.

The mean and rms spread of Δt residual distributions in Monte Carlo simulation vary significantly among tagging categories. We find that we can account for these differences by allowing the fraction of core Gaussian, f, to be different for each tagging category. In addition, we find that the correlations among the three parameters describing the outlier Gaussian ($b^{\rm out}, s^{\rm out}, f^{\rm out}$) are large and that the outlier parameters are highly correlated with other resolution parameters. Therefore, we fix the outlier bias $b^{\rm out}$ and width $s^{\rm out}$, and vary them over a wide range to evaluate the systematic uncertainty on the physics parameters due to fixing these parameters (see Sec. XII). The signal resolution model then has eight free parameters: s, κ, $f^{\rm out}$, and five fractions f_c (one for each tagging category c).

As a cross-check, we use a resolution function that is the sum of a narrow and a wide Gaussian distribution, and a third Gaussian to describe outliers:

$$\begin{aligned}\mathcal{R}_{\rm G+G+G}(\delta\Delta t,\sigma_{\Delta t};b,s,f,b^w,s^w,b^{\rm out},s^{\rm out},f^{\rm out})&\\ =fG(\delta\Delta t;b\sigma_{\Delta t},s\sigma_{\Delta t})+(1-f-f^{\rm out})&\\ \times G(\delta\Delta t;b^w\sigma_{\Delta t},s^w\sigma_{\Delta t})+f^{\rm out}G(\delta\Delta t;b^{\rm out},s^{\rm out}).&\end{aligned} \tag{15}$$

This resolution function has two more parameters than $\mathcal{R}_{\rm GExp+G}$. It accommodates a bias due to tracks from charm decays on the $B_{\rm tag}$ side by allowing the means of the Gaussian distributions to be nonzero.

A. Vertex-tagging correlations

A correlation $d\omega_c/d\sigma_{\Delta t}\approx 0.12\ {\rm ps}^{-1}$ is observed between the mistag rate and the Δt resolution for kaon tags. This effect is modeled in the resolution function for signal as a linear dependence of the mistag rate on $\sigma_{\Delta t}$, as shown in Eq. (10). In this section, we describe the source of this correlation.

We find that both the mistag rate for kaon tags and the calculated error on Δt depend inversely on $\sqrt{\Sigma p_t^2}$, where p_t is the transverse momentum with respect to the z axis of tracks from the $B_{\rm tag}$ decay. Correcting for this dependence of the mistag rate removes most of the correlation between the mistag rate and $\sigma_{\Delta t}$. The mistag rate dependence originates from the kinematics of the physics sources for wrong-charge kaons. The three major sources of mistagged events in the kaon category are wrong-sign D^0 mesons from B decays to double charm ($b\to c\bar{c}s$), wrong-sign kaons from D^+ decays, and kaons produced directly in B decays. All these sources produce a spectrum of tracks that have smaller $\sqrt{\Sigma p_t^2}$ than B decays that produce a correct tag. The $\sigma_{\Delta t}$ dependence originates from the $1/p_t^2$ dependence of σ_z for the individual contributing tracks.

IX. Δt MODELS FOR BACKGROUNDS

Although the true Δt and resolution on Δt are not well defined for background events, we still describe the total Δt model as a ''physics model'' convolved with a ''resolution function'' since an exponential or oscillatory behavior is preserved in some backgrounds.

The background Δt physics models we use in this analysis are all a linear combination of one or more of the following terms, corresponding to prompt, exponential decay, and oscillatory distributions:

$$\mathcal{G}^{\rm prmt}_{\rm phys}(\Delta t_{\rm true},g)=\frac{1}{2}\,\delta(\Delta t_{\rm true})[1+g(1-\omega^{\rm prmt})], \tag{16}$$

$$\mathcal{G}^{\rm life}_{\rm phys}(\Delta t_{\rm true},g)=\frac{1}{4\tau^{\rm bg}}\exp(-|\Delta t_{\rm true}|/\tau^{\rm bg})[1+g(1-\omega^{\rm life})], \tag{17}$$

$$\begin{aligned}\mathcal{G}^{\rm osc}_{\rm phys}(\Delta t_{\rm true},g)=&\frac{1}{4\tau^{\rm bg}}\exp(-|\Delta t_{\rm true}|/\tau^{\rm bg})\\ &\times[1+g(1-\omega^{\rm osc})\cos(\Delta m^{\rm bg}\Delta t_{\rm true})],\end{aligned} \tag{18}$$

where $\delta(\Delta t)$ is a δ function, $g=+1$ for unmixed and -1 for mixed events, and $\tau^{\rm bg}$ and $\Delta m^{\rm bg}$ are the effective lifetime and mixing frequency for the particular background.

For backgrounds, we use a resolution function that is the sum of a narrow and a wide Gaussian distribution:

$$\begin{aligned}\mathcal{R}_{\rm G+G}(\delta\Delta t,\sigma_{\Delta t};b,s,f,b^w,s^w)&\\ =fG(\delta\Delta t;b\sigma_{\Delta t},s\sigma_{\Delta t})+(1-f)G(\delta\Delta t;b^w\sigma_{\Delta t},s^w\sigma_{\Delta t}).&\end{aligned} \tag{19}$$

A. Combinatorial-D^* background

Events in which the D^{*-} candidate corresponds to a random combination of tracks (called the combinatorial-D^* background) constitute the largest background in the signal sample. We use two sets of events to determine the appropriate parameterization of the Δt model for the combinatorial-D^* background: events in data that are in the upper $m(D^*)-m(D^0)$ sideband (above the peak due to real D^{*-} decays); and events in Monte Carlo simulation that are identified as combinatorial-D^* background, based on the true information for the event, in both the $m(D^*)-m(D^0)$ sideband and peak regions. The data and Monte Carlo Δt distributions are described well by a prompt plus oscillatory term convolved with a double-Gaussian resolution function:

$$\mathcal{G}^{\rm comb}=[f^{\rm osc}\mathcal{G}^{\rm osc}_{\rm phys}(\Delta t_{\rm true},g;\tau^{\rm comb},\Delta m^{\rm comb},\omega^{\rm osc}) + (1-f^{\rm osc})\mathcal{G}^{\rm prmt}_{\rm phys}(\Delta t_{\rm true},g;\omega^{\rm prmt})]\otimes\mathcal{R}_{\rm G+G}. \tag{20}$$

Approximately 60% of combinatorial-D^* background events are from $B^0\overline{B}^0$ events according to Monte Carlo simulation. Although the D^{*-} is not correctly reconstructed, the identified lepton is very likely to be a primary lepton. The tagging algorithm can still identify the flavor of $B_{\rm tag}$ with a reasonable mistag probability, especially for the `lepton` category, and for the kaon category if the tracks swapped between the $D^{*-}\ell^+$ candidate and $B_{\rm tag}$ are pions. Therefore, the combinatorial-D^* background also exhibits oscillatory behavior.

The parameters $\omega^{\rm prmt}$, $\Delta m^{\rm comb}$, $\tau^{\rm comb}$, f, b^w, and s^w are shared among all subsamples. The parameters $\omega^{\rm osc}$, $f^{\rm osc}$, b, and s are allowed to be different depending on criteria such as tagging category, whether the data were recorded on or off resonance, whether the candidate lepton passes real- or fake-lepton criteria, and whether the event passes the criteria for same-side or opposite-side D^{*-} and ℓ. The total number of free parameters in the combinatorial-D^* background Δt model is 24.

The relative fraction of $B^0\overline{B}^0$ and B^+B^- events in the combinatorial-D^* background depends slightly on $m(D^*)-m(D^0)$. However, no significant dependence of the parameters of the Δt model on $m(D^*)-m(D^0)$ is observed in data or Monte Carlo simulation. The sample of events in the $m(D^*)-m(D^0)$ sideband is used to determine the starting values for the parameters in the final full fit to all data samples.

To reduce the total number of free parameters in the fit, parameters that describe the shape of the wide Gaussian (bias and width) are shared between combinatorial-D^* background and the three types of peaking background: continuum, fake lepton, and uncorrelated lepton. The wide Gaussian fraction is allowed to be different for each type of background.

B. Continuum peaking background

All $c\overline{c}$ events that have a correctly reconstructed D^{*-} are defined as continuum peaking background, independent of whether the associated lepton candidate is a real lepton or a fake lepton. The $c\overline{c}$ Monte Carlo sample and off-resonance data are used to identify the appropriate Δt model and sharing of parameters among subsamples. The combinatorial-D^* background Δt model and parameters described in the previous section are used to model the combinatorial-D^* background contribution in the off-resonance Δt distribution in data.

The decay vertex of a real D^{*-} from continuum $c\overline{c}$ production always coincides with the primary vertex. If the lepton candidate also originates from the primary vertex, we can use a prompt physics model convolved with a resolution function that can accommodate a bias due to tracks from charm decays other than the D^{*-} candidate. If the lepton candidate is from a charm decay, the measured vertices of the $D^{*-}\ell^+$ candidate and the remaining tracks are both likely to be between the primary vertex and the charm vertex; hence the measured Δz is likely to be very small. Both types of events can be modeled with a prompt model convolved with a double-Gaussian resolution function:

$$\mathcal{G}^{\rm cont}=\mathcal{G}^{\rm prmt}_{\rm phys}(\Delta t_{\rm true},g;\omega^{\rm prmt})\otimes\mathcal{R}_{\rm G+G}. \tag{21}$$

Dependence on the flavor tagging information is included to accommodate any differences in the amount of background events classified as mixed and unmixed.

By fitting to the data and Monte Carlo control samples with different sharing of parameters across subsets of the data, we find that the apparent ''mistag fraction'' for events in the `kaon` category is significantly different from the mistag fraction for other tagging categories. We also find that the core Gaussian bias is significantly different for opposite-side and same-side events. We introduce separate parameters to accommodate these effects.

The total number of parameters used to describe the Δt distribution of continuum peaking background is six. The off-resonance control samples in data are used to determine starting values for the final full fit to all data samples.

C. Fake-lepton peaking background

To determine the Δt model and sharing of parameters for the fake-lepton peaking backgrounds, we use $B^0\overline{B}^0$ and B^+B^- Monte Carlo events in which the D^{*-} is correctly reconstructed but the lepton candidate is misidentified. In addition, we use the fake-lepton control sample in data. The combinatorial-D^* and continuum peaking background Δt models and parameters described in the previous two sections are used to model their contribution to the fake-lepton Δt distribution in data. For this study, the contribution of signal is described by the signal parameters found for signal events in the Monte Carlo simulation.

Since the fake-lepton peaking background is due to B decays in which the fake lepton and the D^{*-} candidate can originate from the same B or different B mesons, and the charge of the fake lepton can carry correct flavor information of the reconstructed B candidate, we include both prompt and oscillatory terms in the Δt model:

$$\mathcal{G}^{\rm fake}=[f^{\rm osc}\mathcal{G}^{\rm osc}_{\rm phys}(\Delta t_{\rm true},g;\tau^{\rm fake},\Delta m^{\rm fake},\omega^{\rm osc}) + (1-f^{\rm osc})\mathcal{G}^{\rm prmt}_{\rm phys}(\Delta t_{\rm true},g;\omega^{\rm prmt})]\otimes\mathcal{R}_{\rm G+G}. \quad (22)$$

We find that the apparent mistag rates for both the prompt and mixing terms, and the bias of the core Gaussian of the resolution function, are different between some tagging categories. The total number of parameters used to describe the fake-lepton background is 14. The fake-lepton control samples in data are used to determine starting values for the final full fit to all data samples.

D. Uncorrelated-lepton peaking background

To determine the Δt model and sharing of parameters for the uncorrelated-lepton peaking backgrounds, we use $B^0\overline{B}^0$ and B^+B^- Monte Carlo events in which the D^{*-} is correctly reconstructed but the lepton candidate is from the other B in the event or from a secondary decay of the same B. In addition, we use the same-side control sample in data, which is only about 30% uncorrelated-lepton background in the $m(D^*)-m(D^0)$ peak region due to significant contributions from combinatorial-D^* background and signal. The combinatorial-D^* and peaking background Δt models and parameters described in the previous two sections are used to model their contribution to the same-side Δt distribution in data. For this initial fit, the contribution of signal is described by the signal parameters found for signal events in the Monte Carlo simulation.

Physics and vertex reconstruction considerations suggest several features of the Δt distribution for the uncorrelated-lepton sample. First, we expect the reconstructed Δt to be systematically smaller than the true Δt value since using a lepton and a D^{*-} from different B decays will generally reduce the separation between the reconstructed $B_{D^*\ell}$ and $B_{\rm tag}$ vertices. We also expect that events with small true Δt will have a higher probability of being misreconstructed as an uncorrelated lepton candidate because it is more likely that the fit of the D^{*-} and ℓ to a common vertex will converge for these events. Finally, we expect truly mixed events to have a higher fraction of uncorrelated-lepton events because in mixed events the charge of the D^* is opposite that of primary leptons on the tagging side. These expectations are confirmed in the Monte Carlo simulation.

We do not expect the uncorrelated-lepton background to exhibit any mixing behavior and none is observed in the data or Monte Carlo control samples. We describe the Δt distribution with the sum of a lifetime term and a prompt term, convolved with a double-Gaussian resolution function:

$$\mathcal{G}^{\rm uncor}=[f^{\rm life}\mathcal{G}^{\rm life}_{\rm phys}(\Delta t_{\rm true},g;\tau^{\rm uncor},\omega^{\rm life}) + (1-f^{\rm life})\mathcal{G}^{\rm prmt}_{\rm phys}(\Delta t_{\rm true},g;\omega^{\rm prmt})]\otimes\mathcal{R}_{\rm G+G}. \quad (23)$$

The effective mistag rates $\omega^{\rm prmt}$ and $\omega^{\rm life}$ accommodate different fractions of uncorrelated-lepton backgrounds in events classified as mixed and unmixed. We find that the apparent mistag rate for the lifetime term is different between some tagging categories. All other parameters are consistent among the different subsamples. The total number of parameters used to describe the uncorrelated-lepton background is six. The uncorrelated-lepton control samples in data are used to determine starting values for the final full fit to all data samples.

E. Charged B peaking background

The charged-B peaking background is due to decays of the type $B^\pm\to D^*\ell\,\nu_\ell X$. Since charged B's do not exhibit mixing behavior, we use the Δt and tagging information to discriminate charged-B peaking background events from neutral-B signal events, in the simultaneous fit to all samples. We use the same resolution model and parameters as for the neutral-B signal since the Δt resolution is dominated by the $z_{\rm tag}$ resolution and the B decay dynamics are very similar. The simulation does not show any significant difference between the signal and the charged-B background Δt residual distributions. The charged B background contribution is described by

$$\mathcal{G}^{\rm ch}=\frac{1}{4\tau_{B^+}}e^{-|\Delta t_{\rm true}|/\tau_{B^+}}[1+g(1-2\omega^c_{B^+})] \otimes\mathcal{R}(\delta\Delta t,\sigma_{\Delta t};\vec{q}_c), \quad (24)$$

where $\omega^c_{B^+}$ is the mistag fraction for charged B mesons for tagging category c.

Given that the ratio of the charged B to neutral B lifetime is close to 1 and the fraction of charged B mesons in the peaking sample is small, we do not have sufficient sensitivity to distinguish the lifetimes in the fit. We parameterize the physics model for the B^+ in terms of the lifetime ratio τ_{B^+}/τ_{B^0}, and fix this ratio to the *Review of Particle Properties 2002* world average of 1.083 [8]. We vary the ratio by the error on the world average (±0.017) to estimate the corresponding systematic uncertainties on τ_{B^0} and Δm_d (see Sec. XII).

In each tagging category, the fit is sensitive to only two parameters among ω_{B^+}, the neutral B mistag fraction (ω_{B^0}) and the charged B fraction (f_{B^+}). Therefore we fix the ratio of mistag rates, $\omega_{B^+}/\omega_{B^0}$, to the value of the ratio measured with fully reconstructed charged and neutral hadronic B decays in data, for each tagging category.

X. FIT RESULTS

The total number of free parameters in the final fit is 72: 21 in the signal model, one for the charged B fraction, 24 in the combinatorial-D^* background model, and 26 in peaking background models. The fitted signal Δt model parameters are shown in Table III.

The statistical correlation coefficient between τ_{B^0} and Δm_d is $\rho(\Delta m_d,\tau_{B^0})=-0.22$. The global correlation coefficients (the largest correlation between a variable and every possible linear combination of other variables) for τ_{B^0} and Δm_d, and some of the correlation coefficients between τ_{B^0} or Δm_d and other parameters, are shown in Table IV.

Figure 3 shows the Δt distributions for unmixed and

TABLE III. Results of full fit to data—signal model and resolution function parameters. A correction, described in Sec. XI A, has been applied to τ_{B^0} and Δm_d. The uncertainties are statistical only.

Parameter	Value	Parameter	Value
Δm_d (ps^{-1})	0.492 ± 0.018	$\Delta\omega_{\rm NT2}$	-0.112 ± 0.028
τ_{B^0} (ps)	$1.523^{+0.024}_{-0.023}$	$\Delta\omega_{\rm NT3}$	-0.023 ± 0.019
f_{B^+}	0.082 ± 0.029	s	1.201 ± 0.063
$\omega_{\rm lepton}$	0.071 ± 0.015	κ	0.86 ± 0.17
$\omega^{\rm offset}_{\rm kaon}$	0.002 ± 0.024	$f_{\rm lepton}$	0.72 ± 0.10
$\alpha_{\rm kaon}$ (ps^{-1})	0.229 ± 0.036	$f_{\rm kaon}$	0.609 ± 0.088
$\omega_{\rm NT1}$	0.212 ± 0.020	$f_{\rm NT1}$	0.69 ± 0.13
$\omega_{\rm NT2}$	0.384 ± 0.018	$f_{\rm NT2}$	0.70 ± 0.10
$\omega_{\rm NT3}$	0.456 ± 0.012	$f_{\rm NT3}$	0.723 ± 0.078
$\Delta\omega_{\rm lepton}$	-0.001 ± 0.022	$f^{\rm out}$	0.0027 ± 0.0017
$\Delta\omega_{\rm kaon}$	-0.024 ± 0.015	$b^{\rm out}$ (ps)	-5.000 (fixed)
$\Delta\omega_{\rm NT1}$	-0.098 ± 0.032	$s^{\rm out}$ (ps)	6.000 (fixed)

mixed events in a sample in which the probability of each event being a signal is higher than a threshold chosen so that the sample is 80% pure in signal events. The points correspond to data. The curves correspond to the sum of the projections of the appropriate relative amounts of signal and background Δt models for this 80%-pure signal sample. Figure 4 shows the time-dependent asymmetry

$$A(\Delta t) = \frac{N_{\rm unmixed}(\Delta t) - N_{\rm mixed}(\Delta t)}{N_{\rm unmixed}(\Delta t) + N_{\rm mixed}(\Delta t)}. \qquad (25)$$

The unit amplitude for the cosine dependence of A is diluted by the mistag probabilities, the experimental Δt resolution, and backgrounds.

TABLE IV. Global correlation coefficients for Δm_d and τ_{B^0} from the full fit to data and other correlation coefficients for pairs of key parameters in the fit.

Δm_d global correlation	0.74
τ_{B^0} global correlation	0.69
$\rho(\Delta m_d, \tau_{B^0})$	-0.22
$\rho(\Delta m_d, f_{B^+})$	0.58
$\rho(\tau_{B^0}, s_{\rm sig})$	-0.49
$\rho(\tau_{B^0}, f^{\rm out}_{\rm sig})$	-0.26

Figure 5 shows the Δt distributions for unmixed and mixed events, and the asymmetry $A(\Delta t)$ for data samples in which events are selected based on the background probabilities such that the sample contains 99.5%-pure combinatorial background events (left plots), or 60%-pure fake-lepton background events (right plots). The observed oscillatory behaviors are expected as explained in Sec. IX.

Since many parameters in the model are free, it is interesting to see how the errors on τ_{B^0} and Δm_d, and their correlation, change when different parameters are free in the fit, or fixed to their best value from the full fit. We perform a series of fits, fixing all parameters at the values obtained from the default fit, except (a) Δm_d and τ_{B^0}, (b) Δm_d, τ_{B^0}, and all mistag fractions in the signal model, (c) Δm_d, τ_{B^0}, and f_{B^+}, (d) Δm_d, τ_{B^0}, f_{B^+}, and all mistag fractions in the signal model, (e) all parameters in the signal Δt model. The one-sigma error ellipses for these fits and for the default fit are shown in Fig. 6.

We can see that the error on τ_{B^0} changes very little until we float the signal resolution function. Floating the background parameters adds a very small contribution to the er-

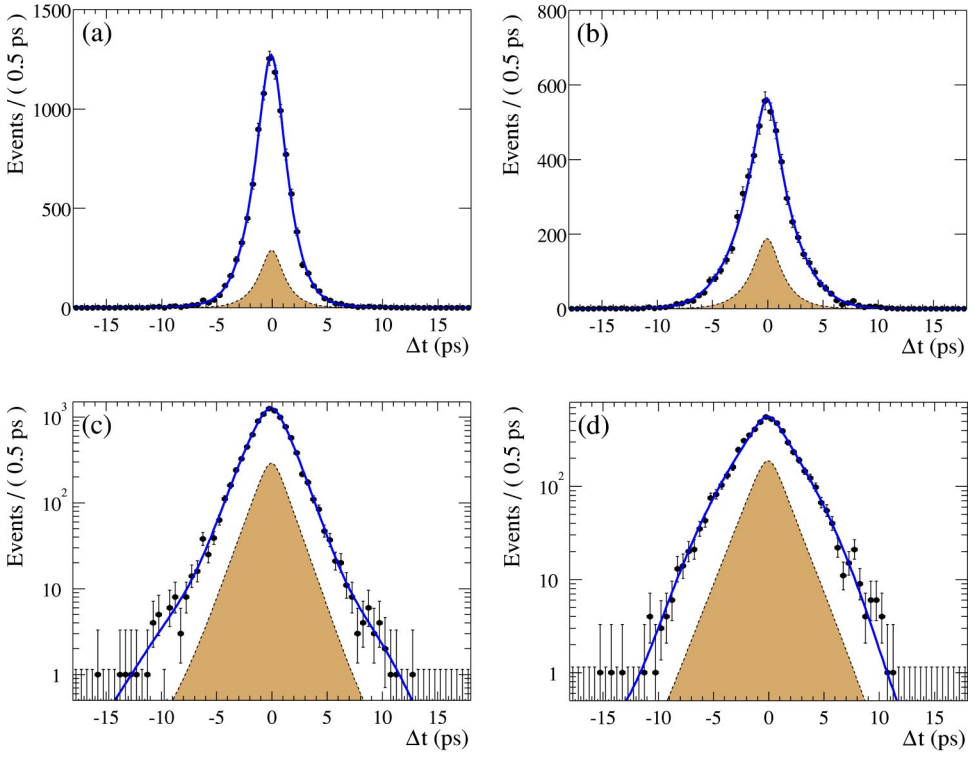

FIG. 3. The Δt distribution on linear (a), (b) and logarithmic (c), (d) scale for (a), (c) unmixed and (b), (d) mixed events in an 80%-pure signal sample, and the projection of the fit results. Each event in this sample has a probability of being a signal higher than a threshold chosen so that the sample is 80% pure in signal events. The shaded area shows the background contribution to the distributions.

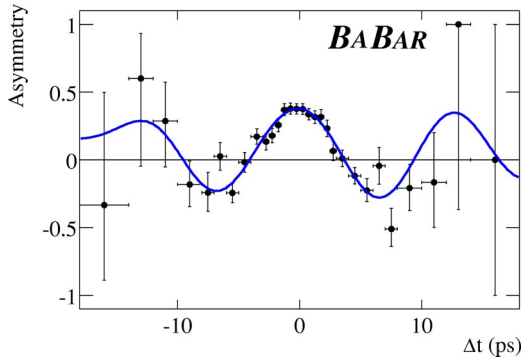

FIG. 4. The asymmetry plot for mixed and unmixed events in an 80%-pure signal sample and the projection of the fit results. Errors on the data points are computed by considering the binomial probabilities for observing different numbers of mixed and unmixed events while preserving the total number.

ror. The contribution from the charged B fraction and mistag fractions to the τ_{B^0} error is negligible. On the other hand, the charged B fraction changes the error on Δm_d the most. The contributions from floating the mistag fractions, resolution functions, and background Δt models are relatively small.

We also check the statistical errors on data by measuring the increase in negative log likelihood in data in the two-dimensional $(\tau_{B^0}, \Delta m_d)$ space in the vicinity of the minimum of the negative log likelihood. We find that the positive error on τ_{B^0} is about 6% larger than that determined by the fitting program, whereas the other errors are the same as those determined by the fit. To take this into account, we increase the positive statistical error on τ_{B^0} by 6%.

XI. VALIDATION AND CROSS-CHECKS

In Sec. XI A, we describe several tests of the fitting procedure that were performed with both fast parameterized Monte Carlo simulations and full detector simulations. In Sec. XI B, we give the results of performing cross-checks on data, including fitting to different subsamples of the data and fitting with variations to the standard fit.

A. Tests of fitting procedure with Monte Carlo simulations

A test of the fitting procedure is performed with fast parameterized Monte Carlo simulations, where 87 experiments are generated with signal and background control sample sizes and compositions corresponding to that obtained from the full likelihood fit to data. The mistag rates and Δt distributions are generated according to the model used in the likelihood fit. The full fit is then performed on each of these experiments. We find no statistically significant bias in the average values of τ_{B^0} and Δm_d for the 87 fits. The rms spread in the distribution of results is consistent with the mean statistical error from the fits and the statistical error on the results in data, for both τ_{B^0} and Δm_d. We find that 20% of the fits result in a value of the negative log likelihood that is smaller (better) than that found in data.

We also fit two types of Monte Carlo samples that include full detector simulation: pure signal and signal plus background. To check whether the selection criteria introduce any bias in the lifetime or mixing frequency, we fit the signal physics model to the true lifetime distribution, using true tagging information, for a large sample of signal Monte Carlo events that pass all selection criteria. We also fit the

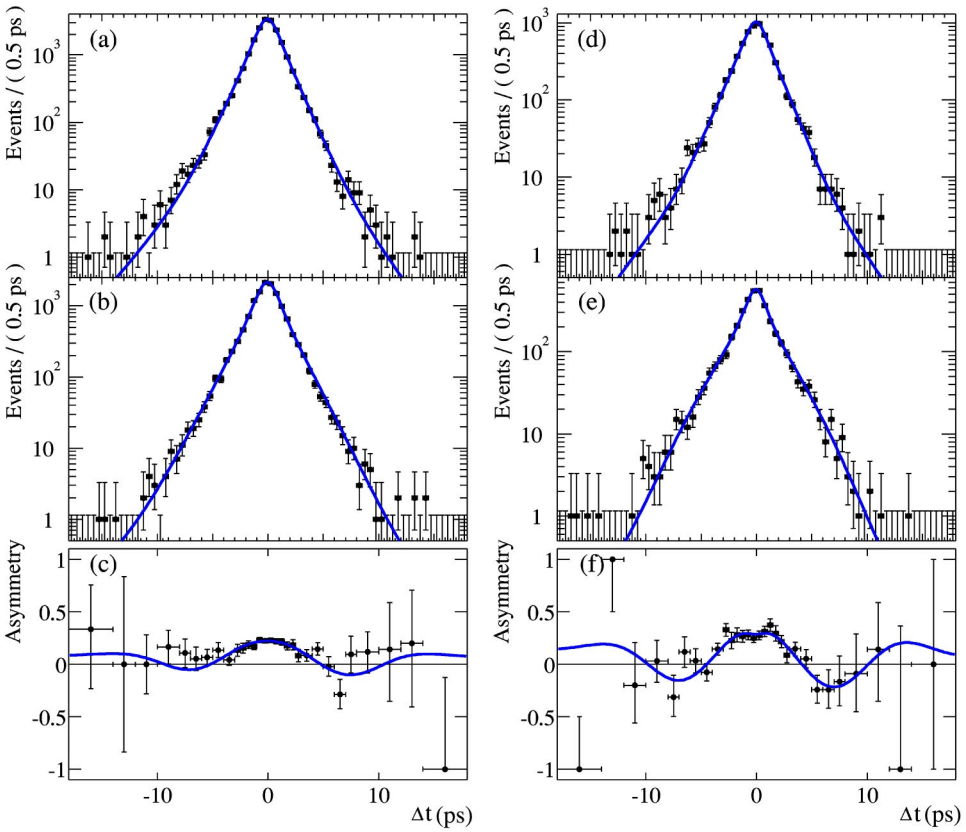

FIG. 5. The Δt distributions for (a), (d) unmixed and (b), (e) mixed events, and (c), (f) the asymmetry plot in a 99.5%-pure combinatorial-D^* sample (a), (b) and (c) and in a 60%-pure D^{*-}-fake-lepton event sample (d), (e), and (f). Events are selected based on the background probabilities, such that the sample contains 99.5%-pure combinatorial-D^* events, or 60%-pure D^{*-}-fake-lepton background events. The projection of the fit results is overlaid on top of the data points. Errors on the data points in the asymmetry plots are computed by considering the binomial probabilities for observing different numbers of mixed and unmixed events while preserving the total number.

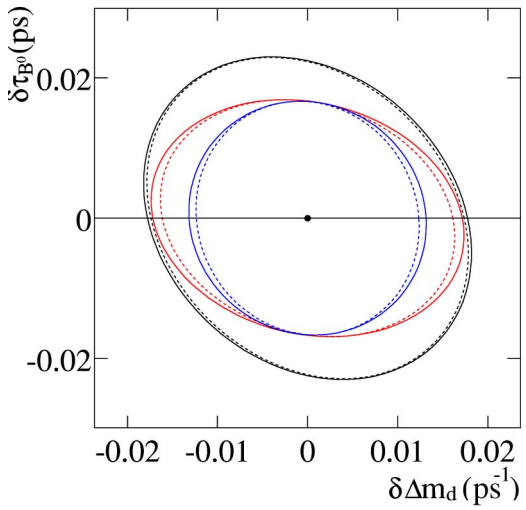

FIG. 6. Comparison of one-sigma error ellipses in the $\Delta m_d - \tau_{B^0}$ plane for fits in which different sets of parameters are free. From the innermost to the outermost ellipse, the floating parameters are ($\Delta m_d, \tau_{B^0}$), (Δm_d, τ_{B^0}, mistag fractions), ($\Delta m_d, \tau_{B^0}, f_{B^+}$), ($\Delta m_d$, τ_{B^0}, f_{B^+}, mistag fractions), all signal Δt parameters, and the default fit (72 floating parameters).

measured Δt distribution, using measured tagging information, with the complete signal Δt model described in Sec. VIII. We find no statistically significant bias in the values of τ_{B^0} or Δm_d extracted in these fits.

The $B^0\overline{B}^0$, B^+B^-, and $c\overline{c}$ Monte Carlo samples that provide simulated background events along with signal events are much smaller than the pure signal Monte Carlo samples. In addition, they are not much larger than the data samples. In order to increase the statistical sensitivity to any bias introduced when the background samples are added to the fit, we compare the values of τ_{B^0} and Δm_d from the fit to signal plus background events, and pure signal events from the same sample. We find that when background is added, the value of τ_{B^0} increases by (0.022 ± 0.009) ps and the value of Δm_d increases by (0.020 ± 0.005) ps^{-1}, where the error is the difference in quadrature between the statistical errors from the fit with and without background. We correct our final results in data for these biases, which are each roughly the same size as the statistical error on the results in data. We conservatively apply a systematic uncertainty on this bias equal to the full statistical error on the measured result in Monte Carlo simulation with background: ± 0.018 ps for τ_{B^0} and ± 0.012 ps^{-1} for Δm_d.

B. Cross-checks in data

We perform the full maximum-likelihood fit on different subsets of the data and find no statistically significant difference in the results for different subsets. The fit is performed on datasets divided according to tagging category, b-quark flavor of the $B^0 \to D^{*-}\ell^+\nu_\ell$ candidate, b-quark flavor of the tagging B, and D^0 decay mode. We also vary the range of Δt over which we perform the fit (maximum value of $|\Delta t|$ equal to 10, 14, and 18 ps), and decrease the maximum allowed value of $\sigma_{\Delta t}$ from 1.8 ps to 1.4 ps. Again, we do not find statistically significant changes in the values of τ_{B^0} or Δm_d.

TABLE V. Summary of systematic uncertainties on the two physics parameters, τ_{B^0} and Δm_d.

Source	$\delta(\Delta m_d)$ (ps^{-1})	$\delta(\tau_{B^0})$ (ps)
Selection and fit bias	0.0123	0.0178
z scale	0.0020	0.0060
PEP-II boost	0.0005	0.0015
Beam spot position	0.0010	0.0050
SVT alignment	0.0030	0.0056
Background/signal probability	0.0029	0.0032
Background Δt models	0.0012	0.0063
Fixed B^+/B^0 lifetime ratio	0.0003	0.0019
Fixed B^+/B^0 mistag ratio	0.0001	0.0003
Fixed signal outlier shape	0.0010	0.0054
Signal resolution model	0.0009	0.0034
Total systematic error	0.013	0.022

XII. SYSTEMATIC UNCERTAINTIES

We estimate systematic uncertainties on the parameters τ_{B^0} and Δm_d with studies performed on both data and Monte Carlo samples, and obtain the results summarized in Table V.

The largest source of systematic uncertainty on both parameters is the limited statistical precision for determining the bias due to the fit procedure (in particular, the background modeling) with Monte Carlo events. We assign the statistical errors of a full fit to Monte Carlo samples including background to estimate this systematic uncertainty. See. Sec. XI A for more details.

The calculation of the decay-time difference Δt for each event assumes a nominal detector z scale, PEP-II boost, vertical beam-spot position, and SVT internal alignment. The PEP-II boost has an uncertainty of 0.1% [6] based on our knowledge of the beam energies. The z-scale uncertainty is determined by reconstructing protons scattered from the beam pipe and comparing the measured beam pipe dimensions with the optical survey data. The z-scale uncertainty is less than 0.4%. We shift the vertical beam-spot position by up to 80 μm, or vary the position randomly with a Gaussian distribution with a width of up to 80 μm, and assign the variation in the fitted parameters as a systematic uncertainty. The systematic uncertainty due to residual errors in SVT internal alignment is estimated by reprocessing the simulated sample with different internal alignment errors. We assign the shift of the fitted parameters as a systematic uncertainty.

The modeling of the $m(D^*)-m(D^0)$ distribution determines the probability we assign for each event to be due to signal. We estimate the uncertainty due to the signal probability calculations by repeating the full fit using an ensemble of different signal and background parameters for the $m(D^*)-m(D^0)$ distributions, varied randomly according to the measured statistical uncertainties and correlations between the parameters. We assign the spread in each of the resulting fitted physics parameters as the systematic uncertainty.

The modeling of the background Δt distribution affects the expected background contributions to the sample. The systematic uncertainty due to the assumed background Δt distributions is estimated as the shift in the fitted parameters when the model for the largest background (due to combinatorial-D^* events) is replaced by the sum of a prompt term and a lifetime term.

The model of the charged B background assumes fixed B^+/B^0 ratios for the mistag rates and lifetimes. We vary the mistag ratio by the uncertainty determined from separate fits to hadronic B decays. We vary the lifetime ratio by the statistical uncertainty on the world average [8]. The resulting change in the fitted physics parameters is assigned as a systematic uncertainty.

The final category of systematic uncertainties is due to assumptions about the resolution model for signal events. We largely avoid assumptions by floating many parameters to describe the resolution simultaneously with the parameters of interest. However, two sources of systematic uncertainty remain: the shape of the outlier contribution, which cannot be determined from data alone, and the assumed parameterization of the resolution for nonoutlier events. We study the sensitivity to the outlier shape by repeating the full fit with outlier Gaussian functions of different means and widths. The mean is varied between -1 ps and -10 ps, and the width is varied from 4 ps to 12 ps. We assign the spread of the resulting fitted values as a systematic uncertainty. We estimate the uncertainty due to the assumed resolution parameterization by repeating the full fit with a triple-Gaussian resolution model and assigning the shift in the fitted values as the uncertainty.

The total systematic uncertainty on τ_{B^0} is 0.022 ps and on Δm_d is 0.013 ps^{-1}.

XIII. SUMMARY

We use a sample of approximately 14000 exclusively reconstructed $B^0 \to D^{*-} \ell^+ \nu_\ell$ signal events to simultaneously measure the B^0 lifetime τ_{B^0} and oscillation frequency Δm_d. We also use samples of events enhanced in the major types of backgrounds. The lifetime and oscillation frequency are determined with an unbinned maximum-likelihood fit that uses, for each event, the measured difference in decay times of the two B mesons and its uncertainty, the signal and background probabilities, and b-quark tagging information for the other B. In addition to the lifetime and oscillation frequency, we extract the parameters describing the signal Δt resolution function, the background Δt models, the mistag fractions, and the B^+ background fraction, in the simultaneous fit to signal and background samples. The results for the physics parameters are

$$\tau_{B^0} = (1.523^{+0.024}_{-0.023} \pm 0.022) \text{ ps}$$

and

$$\Delta m_d = (0.492 \pm 0.018 \pm 0.013) \text{ ps}^{-1}.$$

The statistical correlation coefficient between τ_{B^0} and Δm_d is -0.22.

Both the lifetime and mixing frequency have combined statistical and systematic uncertainties that are comparable to those of the most precise previously published experimental measurements [8]. The results are consistent with the world average measurements of $\tau_{B^0} = (1.542 \pm 0.016)$ ps and $\Delta m_d = (0.489 \pm 0.008)$ ps^{-1} [8].

This analysis demonstrates the feasibility of measuring the B^0 lifetime and mixing frequency simultaneously at B factory experiments, realizing the advantages of better determinations of the Δt resolution function and the amount of B^+ background. All background fractions, Δt resolution parameters for signal and background, and mistag fractions are determined from the data. The lifetime is most correlated with the Δt resolution parameters for signal, while the mixing frequency is most correlated with the B^+ background fraction. The largest systematic uncertainty on both parameters is the limited statistical precision for determining any bias due to the fit procedure (in particular, the background modeling) with Monte Carlo simulation.

Both the statistical and systematic uncertainties on these parameters can be reduced with the larger data and Monte Carlo simulation samples already available at the B factories. Other physics parameters, such as the difference in decay rates of the neutral B mass eigenstates, can also be included in a simultaneous fit in future data samples.

ACKNOWLEDGMENTS

We are grateful for the extraordinary contributions of our PEP-II colleagues in achieving the excellent luminosity and machine conditions that have made this work possible. The success of this project also relies critically on the expertise and dedication of the computing organizations that support BABAR. The collaborating institutions wish to thank SLAC for its support and the kind hospitality extended to them. This work is supported by the U.S. Department of Energy and National Science Foundation, the Natural Sciences and Engineering Research Council (Canada), Institute of High Energy Physics (China), the Commissariat *à* l'Energie Atomique and Institut National de Physique Nucléaire et de Physique des Particules (France), the Bundesministerium für Bildung und Forschung and Deutsche Forschungsgemeinschaft (Germany), the Istituto Nazionale di Fisica Nucleare (Italy), the Research Council of Norway, the Ministry of Science and Technology of the Russian Federation, and the Particle Physics and Astronomy Research Council (United Kingdom). Individuals have received support from the A. P. Sloan Foundation, the Research Corporation, and the Alexander von Humboldt Foundation.

[1] K. Lande, L. M. Lederman, and W. Chinowsky, Phys. Rev. **105**, 1925 (1957).
[2] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **192**, 245 (1987).
[3] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963); M. Kobayashi and K. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[4] See, for example, I. Bigi and A. Sanda, *CP Violation* (Cambridge University Press, Cambridge, England, 2000), Chap. 10.
[5] Throughout this paper, charge conjugate modes are always implied.
[6] BABAR Collaboration, A. Palano *et al.*, Nucl. Instrum. Methods Phys. Res. A **479**, 1 (2002).
[7] ''PEP-II: An Asymmetric *B* Factory,'' Conceptual Design Report, SLAC-418, LBL-5379, 1993.
[8] Particle Data Group, K. Hagiwara *et al.*, Phys. Rev. D **66**, 010001 (2002).
[9] The OPAL Collaboration used partially reconstructed $B^0 \rightarrow D^{*-}\ell^+\nu_\ell$ decays at LEP to simultaneously measure τ_{B^0} and Δm_d. G. Abbiendi *et al.*, Phys. Lett. B **493**, 266 (2000).
[10] This improvement is estimated with a large Monte Carlo simulation of signal events.
[11] ''GEANT, Detector Description and Simulation Tool,'' CERN program library long writeup W5013, 1994.
[12] D. J. Lange, Nucl. Instrum. Methods Phys. Res. A **462**, 152 (2001).
[13] BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. D **66**, 032003 (2002).
[14] E687 Collaboration, P. L. Frabetti *et al.*, Phys. Lett. B **331**, 217 (1994).
[15] See Ref. [6], Sec. 3.3.
[16] See Ref. [13], Sec. V.
[17] See Ref. [13], Sec. IV.