Study of time-dependent *CP*-violating asymmetries and flavor oscillations in neutral *B* decays at the $\Upsilon(4S)$

B. Aubert, D. Boutigny, J.-M. Gaillard, A. Hicheur, Y. Karyotakis, J. P. Lees, P. Robbe, and V. Tisserand
Laboratoire de Physique des Particules, F-74941 Annecy-le-Vieux, France

A. Palano and A. Pompili
Università di Bari, Dipartimento di Fisica and INFN, I-70126 Bari, Italy

G. P. Chen, J. C. Chen, N. D. Qi, G. Rong, P. Wang, and Y. S. Zhu
Institute of High Energy Physics, Beijing 100039, China

G. Eigen and B. Stugu
University of Bergen, Institute of Physics, N-5007 Bergen, Norway

G. S. Abrams, A. W. Borgland, A. B. Breon, D. N. Brown, J. Button-Shafer, R. N. Cahn, A. R. Clark, M. S. Gill, A. V. Gritsan, Y. Groysman, R. G. Jacobsen, R. W. Kadel, J. Kadyk, L. T. Kerth, Yu. G. Kolomensky, J. F. Kral, C. LeClerc, M. E. Levi, G. Lynch, P. J. Oddone, M. Pripstein, N. A. Roe, A. Romosan, M. T. Ronan, V. G. Shelkov, A. V. Telnov, and W. A. Wenzel
Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720

T. J. Harrison, C. M. Hawkes, D. J. Knowles, S. W. O'Neale, R. C. Penny, A. T. Watson, and N. K. Watson
University of Birmingham, Birmingham, B15 2TT, United Kingdom

T. Deppermann, K. Goetzen, H. Koch, M. Kunze, B. Lewandowski, K. Peters, H. Schmuecker, and M. Steinke
Ruhr Universität Bochum, Institut für Experimentalphysik 1, D-44780 Bochum, Germany

N. R. Barlow, W. Bhimji, N. Chevalier, P. J. Clark, W. N. Cottingham, B. Foster, C. Mackay, and F. F. Wilson
University of Bristol, Bristol BS8 1TL, United Kingdom

K. Abe, C. Hearty, T. S. Mattison, J. A. McKenna, and D. Thiessen
University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1

S. Jolly and A. K. McKemey
Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom

V. E. Blinov, A. D. Bukin, D. A. Bukin, A. R. Buzykaev, V. B. Golubev, V. N. Ivanchenko, A. A. Korol, E. A. Kravchenko, A. P. Onuchin, S. I. Serednyakov, Yu. I. Skovpen, V. I. Telnov, and A. N. Yushkov
Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia

D. Best, M. Chao, D. Kirkby, A. J. Lankford, M. Mandelkern, S. McMahon, and D. P. Stoker
University of California at Irvine, Irvine, California 92697

K. Arisaka, C. Buchanan, and S. Chun
University of California at Los Angeles, Los Angeles, California 90024

D. B. MacFarlane, S. Prell, Sh. Rahatlou, G. Raven, and V. Sharma
University of California at San Diego, La Jolla, California 92093

C. Campagnari, B. Dahmes, P. A. Hart, N. Kuznetsova, S. L. Levy, O. Long, A. Lu, M. A. Mazur, J. D. Richman, and W. Verkerke
University of California at Santa Barbara, Santa Barbara, California 93106

J. Beringer, A. M. Eisner, M. Grothe, C. A. Heusch, W. S. Lockman, T. Pulliam, T. Schalk, R. E. Schmitz, B. A. Schumm, A. Seiden, M. Turri, W. Walkowiak, D. C. Williams, and M. G. Wilson
University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, California 95064

 ©2002 The American Physical Society

E. Chen, G. P. Dubois-Felsmann, A. Dvoretskii, D. G. Hitlin, S. Metzler, J. Oyang, F. C. Porter, A. Ryd, A. Samuel, M. Weaver, S. Yang, and R. Y. Zhu
California Institute of Technology, Pasadena, California 91125

S. Devmal, T. L. Geld, S. Jayatilleke, G. Mancinelli, B. T. Meadows, and M. D. Sokoloff
University of Cincinnati, Cincinnati, Ohio 45221

T. Barillari, P. Bloom, M. O. Dima, W. T. Ford, U. Nauenberg, A. Olivas, P. Rankin, J. Roy, J. G. Smith, and W. C. van Hoek
University of Colorado, Boulder, Colorado 80309

J. Blouw, J. L. Harton, M. Krishnamurthy, A. Soffer, W. H. Toki, R. J. Wilson, and J. Zhang
Colorado State University, Fort Collins, Colorado 80523

T. Brandt, J. Brose, T. Colberg, M. Dickopp, R. S. Dubitzky, A. Hauke, E. Maly, R. Müller-Pfefferkorn, S. Otto, K. R. Schubert, R. Schwierz, B. Spaan, and L. Wilden
Technische Universität Dresden, Institut für Kern-und Teilchenphysik, D-01062 Dresden, Germany

D. Bernard, G. R. Bonneaud, F. Brochard, J. Cohen-Tanugi, S. Ferrag, S. T'Jampens, Ch. Thiebaux, G. Vasileiadis, and M. Verderi
Ecole Polytechnique, F-91128 Palaiseau, France

A. Anjomshoaa, R. Bernet, A. Khan, D. Lavin, F. Muheim, S. Playfer, J. E. Swain, and J. Tinslay
University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom

M. Falbo
Elon University, Elon, North Carolina 27244-2010

C. Borean, C. Bozzi, and L. Piemontese
Università di Ferrara, Dipartimento di Fisica and INFN, I-44100 Ferrara, Italy

E. Treadwell
Florida A&M University, Tallahassee, Florida 32307

F. Anulli,* R. Baldini-Ferroli, A. Calcaterra, R. de Sangro, D. Falciai, G. Finocchiaro, P. Patteri, I. M. Peruzzi,* M. Piccolo, Y. Xie, and A. Zallo
Laboratori Nazionali di Frascati dell'INFN, I-00044 Frascati, Italy

S. Bagnasco, A. Buzzo, R. Contri, G. Crosetti, M. Lo Vetere, M. Macri, M. R. Monge, S. Passaggio, F. C. Pastore, C. Patrignani, M. G. Pia, E. Robutti, A. Santroni, and S. Tosi
Università di Genova, Dipartimento di Fisica and INFN, I-16146 Genova, Italy

M. Morii
Harvard University, Cambridge, Massachusetts 02138

R. Bartoldus, R. Hamilton, and U. Mallik
University of Iowa, Iowa City, Iowa 52242

J. Cochran, H. B. Crawley, P.-A. Fischer, J. Lamsa, W. T. Meyer, and E. I. Rosenberg
Iowa State University, Ames, Iowa 50011-3160

G. Grosdidier, C. Hast, A. Höcker, H. M. Lacker, S. Laplace, V. Lepeltier, A. M. Lutz, S. Plaszczynski, M. H. Schune, S. Trincaz-Duvoid, and G. Wormser
Laboratoire de l'Accélérateur Linéaire, F-91898 Orsay, France

R. M. Bionta, V. Brigljević, D. J. Lange, M. Mugge, K. van Bibber, and D. M. Wright
Lawrence Livermore National Laboratory, Livermore, California 94550

A. J. Bevan, J. R. Fry, E. Gabathuler, R. Gamet, M. George, M. Kay, D. J. Payne, R. J. Sloane, and C. Touramanis
University of Liverpool, Liverpool L69 3BX, United Kingdom

M. L. Aspinwall, D. A. Bowerman, P. D. Dauncey, U. Egede, I. Eschrich, N. J. W. Gunawardane, J. A. Nash, P. Sanders, and D. Smith
University of London, Imperial College, London, SW7 2BW, United Kingdom

D. E. Azzopardi, J. J. Back, G. Bellodi, P. Dixon, P. F. Harrison, R. J. L. Potter, H. W. Shorthouse, P. Strother, and P. B. Vidal
Queen Mary, University of London, E1 4NS, United Kingdom

G. Cowan, S. George, M. G. Green, A. Kurup, C. E. Marker, P. McGrath, T. R. McMahon, S. Ricciardi, F. Salvatore, and G. Vaitsas
University of London, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, United Kingdom

D. Brown and C. L. Davis
University of Louisville, Louisville, Kentucky 40292

J. Allison, R. J. Barlow, J. T. Boyd, A. C. Forti, J. Fullwood, F. Jackson, G. D. Lafferty, N. Savvas, J. H. Weatherall, and J. C. Williams
University of Manchester, Manchester M13 9PL, United Kingdom

A. Farbin, A. Jawahery, V. Lillard, J. Olsen, D. A. Roberts, and J. R. Schieck
University of Maryland, College Park, Maryland 20742

G. Blaylock, C. Dallapiccola, K. T. Flood, S. S. Hertzbach, R. Kofler, V. B. Koptchev, T. B. Moore, H. Staengle, and S. Willocq
University of Massachusetts, Amherst, Massachusetts 01003

B. Brau, R. Cowan, G. Sciolla, F. Taylor, and R. K. Yamamoto
Massachusetts Institute of Technology, Laboratory for Nuclear Science, Cambridge, Massachusetts 02139

M. Milek and P. M. Patel
McGill University, Montréal, Quebec, Canada H3A 2T8

F. Palombo
Università di Milano, Dipartimento di Fisica and INFN, I-20133 Milano, Italy

J. M. Bauer, L. Cremaldi, V. Eschenburg, R. Kroeger, J. Reidy, D. A. Sanders, and D. J. Summers
University of Mississippi, University, Mississippi 38677

J. Y. Nief and P. Taras
Université de Montréal, Laboratoire René J. A. Lévesque, Montréal, Quebec, Canada H3C 3J7

H. Nicholson
Mount Holyoke College, South Hadley, Massachusetts 01075

C. Cartaro, N. Cavallo,[†] G. De Nardo, F. Fabozzi, C. Gatto, L. Lista, P. Paolucci, D. Piccolo, and C. Sciacca
Università di Napoli Federico II, Dipartimento di Scienze Fisiche and INFN, I-80126, Napoli, Italy

J. M. LoSecco
University of Notre Dame, Notre Dame, Indiana 46556

J. R. G. Alsmiller and T. A. Gabriel
Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

J. Brau, R. Frey, E. Grauges, M. Iwasaki, N. B. Sinev, and D. Strom
University of Oregon, Eugene, Oregon 97403

F. Colecchia, F. Dal Corso, A. Dorigo, F. Galeazzi, M. Margoni, G. Michelon, M. Morandin, M. Posocco, M. Rotondo, F. Simonetto, R. Stroili, E. Torassa, and C. Voci
Università di Padova, Dipartimento di Fisica and INFN, I-35131 Padova, Italy

M. Benayoun, H. Briand, J. Chauveau, P. David, Ch. de la Vaissière, L. Del Buono, O. Hamon, F. Le Diberder, Ph. Leruste, J. Ocariz, L. Roos, and J. Stark
Universités Paris VI et VII, Lab de Physique Nucléaire H. E., F-75252 Paris, France

P. F. Manfredi, V. Re, and V. Speziali
Università di Pavia, Dipartimento di Elettronica and INFN, I-27100 Pavia, Italy

E. D. Frank, L. Gladney, Q. H. Guo, and J. Panetta
University of Pennsylvania, Philadelphia, Pennsylvania 19104

C. Angelini, G. Batignani, S. Bettarini, M. Bondioli, F. Bucci, E. Campagna, M. Carpinelli, F. Forti, M. A. Giorgi, A. Lusiani, G. Marchiori, F. Martinez-Vidal, M. Morganti, N. Neri, E. Paoloni, M. Rama, G. Rizzo, F. Sandrelli, G. Simi, G. Triggiani, and J. Walsh
Università di Pisa, Scuola Normale Superiore and INFN, I-56010 Pisa, Italy

M. Haire, D. Judd, K. Paick, L. Turnbull, and D. E. Wagoner
Prairie View A&M University, Prairie View, Texas 77446

J. Albert, P. Elmer, C. Lu, V. Miftakov, S. F. Schaffner, A. J. S. Smith, A. Tumanov, and E. W. Varnes
Princeton University, Princeton, New Jersey 08544

G. Cavoto, D. del Re, F. Ferrarotto, F. Ferroni, E. Lamanna, M. A. Mazzoni, S. Morganti, G. Piredda, F. Safai Tehrani, M. Serra, and C. Voena
Università di Roma La Sapienza, Dipartimento di Fisica and INFN, I-00185 Roma, Italy

R. Faccini
University of California at San Diego, La Jolla, California 92093
and Università di Roma La Sapienza, Dipartimento di Fisica and INFN, I-00185 Roma, Italy

S. Christ and R. Waldi
Universität Rostock, D-18051 Rostock, Germany

T. Adye, N. De Groot, B. Franek, N. I. Geddes, G. P. Gopal, and S. M. Xella
Rutherford Appleton Laboratory, Chilton, Didcot, Oxon, OX11 0QX, United Kingdom

R. Aleksan, S. Emery, A. Gaidot, S. F. Ganzhur, P.-F. Giraud, G. Hamel de Monchenault, W. Kozanecki, M. Langer, G. W. London, B. Mayer, B. Serfass, G. Vasseur, Ch. Yèche, and M. Zito
DAPNIA, Commissariat à l'Energie Atomique/Saclay, F-91191 Gif-sur-Yvette, France

M. V. Purohit, H. Singh, A. W. Weidemann, and F. X. Yumiceva
University of South Carolina, Columbia, South Carolina 29208

I. Adam, D. Aston, N. Berger, A. M. Boyarski, G. Calderini, M. R. Convery, D. P. Coupal, D. Dong, J. Dorfan, W. Dunwoodie, R. C. Field, T. Glanzman, S. J. Gowdy, T. Haas, V. Halyo, T. Himel, T. Hryn'ova, M. E. Huffer, W. R. Innes, C. P. Jessop, M. H. Kelsey, P. Kim, M. L. Kocian, U. Langenegger, D. W. G. S. Leith, S. Luitz, V. Luth, H. L. Lynch, H. Marsiske, S. Menke, R. Messner, D. R. Muller, C. P. O'Grady, V. E. Ozcan, A. Perazzo, M. Perl, S. Petrak, H. Quinn, B. N. Ratcliff, S. H. Robertson, A. Roodman, A. A. Salnikov, T. Schietinger, R. H. Schindler, J. Schwiening, A. Snyder, A. Soha, S. M. Spanier, J. Stelzer, D. Su, M. K. Sullivan, H. A. Tanaka, J. Va'vra, S. R. Wagner, A. J. R. Weinstein, W. J. Wisniewski, D. H. Wright, and C. C. Young
Stanford Linear Accelerator Center, Stanford, California 94309

P. R. Burchat, C. H. Cheng, T. I. Meyer, and C. Roat
Stanford University, Stanford, California 94305-4060

R. Henderson
TRIUMF, Vancouver, British Columbia, Canada V6T 2A3

W. Bugg and H. Cohn
University of Tennessee, Knoxville, Tennessee 37996

J. M. Izen, I. Kitayama, and X. C. Lou
University of Texas at Dallas, Richardson, Texas 75083

F. Bianchi, M. Bona, and D. Gamba
Università di Torino, Dipartimento di Fisica Sperimentale and INFN, I-10125 Torino, Italy

L. Bosisio, G. Della Ricca, S. Dittongo, L. Lanceri, P. Poropat, and G. Vuagnin
Università di Trieste, Dipartimento di Fisica and INFN, I-34127 Trieste, Italy

R. S. Panvini
Vanderbilt University, Nashville, Tennessee 37235

C. M. Brown, P. D. Jackson, R. Kowalewski, and J. M. Roney
University of Victoria, Victoria, British Columbia, Canada V8W 3P6

H. R. Band, E. Charles, S. Dasu, A. M. Eichenbaum, H. Hu, J. R. Johnson, R. Liu, F. Di Lodovico, Y. Pan, R. Prepost, I. J. Scott, S. J. Sekula, J. H. von Wimmersperg-Toeller, S. L. Wu, and Z. Yu
University of Wisconsin, Madison, Wisconsin 53706

T. M. B. Kordich and H. Neal
Yale University, New Haven, Connecticut 06511

(BABAR Collaboration)
(Received 10 January 2002; published 15 August 2002)

We present a measurement of time-dependent CP-violating asymmetries in neutral B meson decays collected with the *BABAR* detector at the PEP-II asymmetric-energy e^+e^- collider at the Stanford Linear Accelerator Center. The data sample consists of 29.7 fb^{-1} recorded at the $\Upsilon(4S)$ resonance and 3.9 fb^{-1} off resonance. One of the neutral B mesons, which are produced in pairs at the $\Upsilon(4S)$, is fully reconstructed in the CP decay modes $J/\psi K_S^0$, $\psi(2S)K_S^0$, $\chi_{c1}K_S^0$, $J/\psi K^{*0}$ ($K^{*0}\to K_S^0\pi^0$) and $J/\psi K_L^0$, or in flavor-eigenstate modes involving $D^{(*)}\pi/\rho/a_1$ and $J/\psi K^{*0}$ ($K^{*0}\to K^+\pi^-$). The flavor of the other neutral B meson is tagged at the time of its decay, mainly with the charge of identified leptons and kaons. A neural network tagging algorithm is used to recover events without a clear lepton or kaon tag. The proper time elapsed between the decays is determined by measuring the distance between the decay vertices. Wrong-tag probabilities, the time-difference resolution function, and the B^0-$\overline{B}^0$ oscillation frequency Δm_d are measured with a sample of about 6350 fully-reconstructed B^0 decays in hadronic flavor-eigenstate modes. A maximum-likelihood fit to this flavor eigenstate sample finds $\Delta m_d = 0.516 \pm 0.016(\text{stat}) \pm 0.010(\text{syst})$ ps^{-1}. The value of the asymmetry amplitude $\sin 2\beta$ is determined from a simultaneous maximum-likelihood fit to the time-difference distribution of the flavor-eigenstate sample and about 642 tagged B^0 decays in the CP-eigenstate modes. We find $\sin 2\beta = 0.59 \pm 0.14(\text{stat}) \pm 0.05(\text{syst})$, demonstrating that CP violation exists in the neutral B meson system. We also determine the value of the CP violation parameter $|\lambda| = 0.93 \pm 0.09(\text{stat}) \pm 0.03(\text{syst})$, which is consistent with the expectation of $|\lambda| = 1$ for no direct CP violation.

DOI: 10.1103/PhysRevD.66.032003 PACS number(s): 13.25.Hw, 11.30.Er, 12.15.Hh

I. INTRODUCTION

CP violation has been a central concern of particle physics since its discovery in 1964 [1]. Interest was heightened by Sakharov's observation [2] in 1967 that without CP violation, a universe that began as matter-antimatter symmetric could not have evolved into the asymmetric one we now see. An elegant explanation of the CP-violating effects in K_L^0 decays is provided by the CP-violating phase of the three-generation Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix [3]. However, existing studies of CP violation in neutral kaon decays and the resulting experimental con-

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

straints on the parameters of the CKM matrix [4] do not provide a stringent test of whether the CKM phase describes *CP* violation [5]. Moreover, the standard model does not, through the CKM phase, incorporate enough *CP* violation to explain the current matter-antimatter asymmetry [6]. Understanding *CP* violation thus remains a pressing challenge.

An excellent testing ground for *CP* violation is provided by *B* mesons through particle-antiparticle mixing. A particle that is purely B^0 at time $t=0$ will oscillate between that state and $\overline{B}^0$ with a frequency Δm_d, the difference between the masses of the two neutral *B* mass eigenstates. If decays to a *CP* eigenstate *f* are observed, any difference between the rates when starting with a B^0 or with a $\overline{B}_0$ is a manifestation of *CP* violation. In some circumstances, including those in the experiment described here, the fundamental parameters of *CP* violation in the CKM model can be measured from such time-dependent rate asymmetries, unobscured by strong interactions. For example, a state initially produced as a B^0 ($\overline{B}^0$) can decay to $J/\psi K_S^0$ directly or can oscillate into a $\overline{B}^0$ (B^0) and then decay to $J/\psi K_S^0$. With little theoretical uncertainty in the standard model, the phase difference between these amplitudes is equal to twice the angle $\beta=\arg[-V_{cd}V_{cb}^*/V_{td}V_{tb}^*]$ of the unitarity triangle. The *CP*-violating asymmetry can thus provide a crucial test of the standard model.

The unitarity of the three-generation CKM matrix can be expressed in geometric form by six triangles of equal area in the complex plane. A nonzero area [7] directly implies the existence of a *CP*-violating CKM phase. The most experimentally accessible of the unitarity relations, involving the two smallest elements of the CKM matrix, V_{ub} and V_{td}, has come to be known as the unitarity triangle. Because the lengths of the sides of the unitarity triangle are comparable, the angles can be large, leading to potentially large *CP*-violating asymmetries from relative phases between CKM matrix elements.

In e^+e^- storage rings operating at the $\Upsilon(4S)$ resonance, a $B^0\overline{B}^0$ pair produced in an $\Upsilon(4S)$ decay evolves in a coherent *P*-wave state. If one of the *B* mesons, referred to as $B_{\rm tag}$, can be ascertained to decay to a state of known flavor, i.e., B^0 or $\overline{B}^0$, at a certain time $t_{\rm tag}$, the other *B*, referred to as $B_{\rm rec}$, *at that time* must be of the opposite flavor as a consequence of Bose symmetry. Consequently, the oscillatory probabilities for observing $B^0\overline{B}^0$, B^0B^0 and $\overline{B}^0\overline{B}^0$ pairs produced in $\Upsilon(4S)$ decays are a function of $\Delta t=t_{\rm rec}-t_{\rm tag}$, allowing mixing frequencies and *CP* asymmetries to be determined if Δt is known. The charges of identified leptons and kaons are the primary indicators of the flavor of the tagging *B*, but other particles also carry flavor information that can be exploited with a neural network algorithm. The reconstructed neutral *B* is found either in a flavor eigenstate ($B_{\rm rec}=B_{\rm flav}$) or a *CP* mode ($B_{\rm rec}=B_{CP}$) by full reconstruction of its observed long-lived daughters.

At the SLAC PEP-II asymmetric e^+e^- collider [8], resonant production of the $\Upsilon(4S)$ provides a copious source of $B^0\overline{B}^0$ pairs moving along the beam axis (*z* direction) with an average Lorentz boost of $\langle\beta\gamma\rangle=0.55$. Therefore, the proper decay-time difference Δt is, to an excellent approximation, proportional to the distance Δz between the two B^0-decay vertices along the axis of the boost, $\Delta t\approx\Delta z/c\langle\beta\gamma\rangle$. The average separation between the two *B* decay vertices is $\Delta z=\langle\beta\gamma\rangle c\tau_B=260\ \mu$m, while the rms Δz resolution of the detector is about 180 μm.

A. Measurement of B^0 flavor oscillations

The phenomenon of particle-antiparticle mixing in the neutral *B* meson system was first observed almost fifteen years ago [9,10]. The oscillation frequency in B^0-$\overline{B}^0$ mixing has been extensively studied with both time-integrated and time-dependent techniques [11]. By interchanging $b\overline{d}$ with $\overline{b}d$, B^0-$\overline{B}^0$ mixing changes the additive bottom quantum number by two units, i.e., $|\Delta B|=2$. In the standard model, such a process is the result of ordinary $|\Delta B|=1$ weak interactions in second order involving the exchange of virtual charge-2/3 quarks, with the top quark contributing the dominant amplitude. A measurement of Δm_d is therefore sensitive to the value of the CKM matrix element V_{td}. At present the sensitivity to V_{td} is not limited by experimental precision on Δm_d, but by theoretical uncertainties in the calculation, in particular the quantity $f_B^2B_B$, where f_B is the B^0 decay constant, and B_B is the so-called bag factor, representing the $\Delta B=2$ strong-interaction matrix element. There may also be contributions from interactions outside the standard model.

Beyond these questions of fundamental interest, since the measurement of Δm_d incorporates all elements of the analysis for time-dependent *CP* asymmetries, including *B* reconstruction, tagging, and Δt determination and resolution, it is an essential test of our understanding of these aspects of the $\sin 2\beta$ measurement.

For the measurement of Δm_d, one neutral *B* ($B_{\rm flav}$) is fully reconstructed in a flavor eigenstate [12] as $D^{(*)-}\pi^+/\rho^+/a_1^+$ or $J/\psi K^{*0}$ ($K^{*0}\to K^+\pi^-$), while the second is tagged by its decay products. For the neutral *B* system produced on the $\Upsilon(4S)$, the probability for obtaining a *mixed*, B^0B^0 or $\overline{B}^0\overline{B}^0$, or *unmixed*, $B^0\overline{B}^0$, final state is a function of Δm_d and the proper time difference Δt between the two *B* decays:

$$\mathrm{Prob}(B^0\overline{B}^0\to B^0B^0\ \text{ or }\ \overline{B}^0\overline{B}^0)=\frac{\Gamma}{4}e^{-\Gamma|\Delta t|}(1-\cos\Delta m_d\Delta t) \tag{1}$$

$$\mathrm{Prob}(B^0\overline{B}^0\to B^0\overline{B}^0)=\frac{\Gamma}{4}e^{-\Gamma|\Delta t|}(1+\cos\Delta m_d\Delta t), \tag{2}$$

where $\tau_{B^0}=1/\Gamma$ is the B^0 lifetime. The final state can be classified as mixed or unmixed depending on whether the reconstructed flavor-eigenstate $B_{\rm rec}=B_{\rm flav}$ has the same or the opposite flavor as the tagging $B=B_{\rm tag}$. If the Δt resolution and flavor tagging were perfect, the asymmetry as a function of Δt,

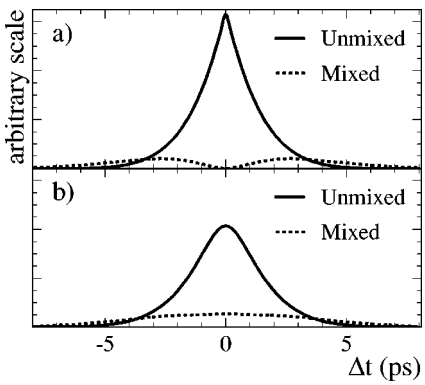

FIG. 1. Expected Δt distribution for mixed and unmixed events (a) with perfect tagging and Δt resolution, and (b) with typical mistag rates and Δt resolution.

$$A_{\text{mixing}}(\Delta t)=\frac{N_{\text{unmix}}(\Delta t)-N_{\text{mix}}(\Delta t)}{N_{\text{unmix}}(\Delta t)+N_{\text{mix}}(\Delta t)}, \tag{3}$$

would describe a cosine function with unit amplitude. The asymmetry goes through zero near 2.1 B^0 proper lifetimes and the sensitivity to Δm_d, which is proportional to $\Delta t^2 e^{-\Gamma|\Delta t|}\sin^2\Delta m_d\Delta t$, reaches a maximum in this region. If the tagging algorithm incorrectly identifies the tag with a probability w, the amplitude of the oscillation is reduced by a dilution factor $\mathcal{D}=(1-2w)$. When more than one type of flavor tag is used, each has its own mistag rate w_i.

Neglecting any background contributions, the probability density functions (PDFs) for the mixed $(-)$ and unmixed $(+)$ events, $\mathcal{H}_\pm$, can be expressed as the convolution of the underlying oscillatory physics distribution

$$h_\pm(\Delta t;\Gamma,\Delta m_d,w)=\frac{\Gamma}{4}e^{-\Gamma|\Delta t|}[1\pm\mathcal{D}\cos\Delta m_d\Delta t] \tag{4}$$

with a time-difference resolution function $\mathcal{R}(\delta_t=\Delta t-\Delta t_{\text{true}};\hat{a})$ to give

$$\begin{aligned}\mathcal{H}_\pm(\Delta t;\Gamma,\Delta m_d,w,\hat{a})&\\ =h_\pm(\Delta t_{\text{true}};\Gamma,\Delta m_d,w)\otimes\mathcal{R}(\delta_t;\hat{a}),&\end{aligned} \tag{5}$$

where Δt and Δt_{true} are the measured and the true time differences, and $\hat{a}$ are parameters of the resolution function. Figure 1 illustrates the impact of typical mistag and Δt resolution effects on the Δt distributions for mixed and unmixed events.

A full likelihood function is then constructed by summing $\mathcal{H}_\pm$ over all mixed and unmixed events in a given uniquely-assigned tagging category i and over all tagging categories

$$\begin{aligned}\ln\mathcal{L}_{\text{mix}}=\sum_i^{\text{tagging}}\Bigg[&\sum_{\text{unmixed}}\ln\mathcal{H}_+(\Delta t;\Gamma,\Delta m_d,w_i,\hat{a}_i)\\ &+\sum_{\text{mixed}}\ln\mathcal{H}_-(\Delta t;\Gamma,\Delta m_d,w_i,\hat{a}_i)\Bigg].\end{aligned} \tag{6}$$

This can be maximized to extract the mistag fractions w_i and resolution parameters $\hat{a}_i$ and, simultaneously, the mixing rate Δm_d. The correlation between w_i and Δm_d is small because the rate of mixed events near $\Delta t=0$, where the B^0-$\overline{B}^0$ mixing probability is small, is principally governed by the mistag rate. Conversely, the sensitivity to Δm_d increases at larger values of Δt; when Δt is approximately twice the B lifetime, half of the neutral B mesons will have oscillated.

B. Measurement of *CP* asymmetries

For the measurement of CP asymmetries, one B (B_{CP}) is fully reconstructed in a CP eigenstate with eigenvalue $\eta_{CP}=-1$ $[J/\psi K_S^0,\ \psi(2S)K_S^0,$ or $\chi_{c1}K_S^0]$ or $+1$ $(J/\psi K_L^0)$, while the second is tagged with its decay products just as for the mixing measurement. The B_{CP} sample is further enlarged by including the mode $J/\psi K^{*0}$ $(K^{*0}\to K_S^0\pi^0)$. However, due to the presence of even $(L=0,2)$ and odd $(L=1)$ orbital angular momenta in the $J/\psi K^{*0}$ system, there are $\eta_{CP}=+1$ and -1 contributions to its decay rate, respectively. When the angular information in the decay is ignored, the measured CP asymmetry in $J/\psi K^{*0}$ is reduced by a dilution factor $D_\perp=1-2R_\perp$, where $R_\perp$ is the fraction of the $L=1$ component. We have measured $R_\perp=0.160\pm0.032\pm0.014$ [13] which, after acceptance corrections, leads to an effective $\eta_{CP}=+0.65\pm0.07$ for the $J/\psi K^{*0}$ mode.

The expected time evolution for the tagged B_{CP} sample depends both on B^0-$\overline{B}^0$ mixing and on the decay amplitudes of B^0 and $\overline{B}^0$ to the final state f through a single complex parameter λ. Mixing generates a lifetime difference as well as a mass difference between the two neutral B meson mass eigenstates, but the lifetime difference is expected to be small since it is a consequence of common final states in B^0 and $\overline{B}^0$ decays. Such common states, which include the CP eigenstates studied here, make up a very small fraction of the decay width. Dropping these, and thus ignoring any lifetime difference, results in a simple expression for λ in terms of the $|\Delta B=1|$ and $|\Delta B=2|$ interactions,

$$\lambda=-\frac{|\langle B^0|\mathcal{H}_{\Delta B=2}|\overline{B}^0\rangle|}{\langle B^0|\mathcal{H}_{\Delta B=2}|\overline{B}^0\rangle}\frac{\langle f|\mathcal{H}_{\Delta B=1}|\overline{B}^0\rangle}{\langle f|\mathcal{H}_{\Delta B=1}|B^0\rangle}. \tag{7}$$

Redefining the states for B^0 and $\overline{B}^0$ by multiplying them by two different phases has no effect on λ, which is thus phase-convention independent, as every physical observable must be. The decay distributions are

$$\begin{aligned}f_\pm(\Delta t)=\frac{\Gamma}{4}e^{-\Gamma|\Delta t|}\Bigg\{1\pm\mathcal{D}\Bigg[&\frac{2\,\text{Im}\,\lambda}{1+|\lambda|^2}\sin\Delta m_d\Delta t\\ &-\frac{1-|\lambda|^2}{1+|\lambda|^2}\cos\Delta m_d\Delta t\Bigg]\Bigg\},\end{aligned} \tag{8}$$

where the $+$ or $-$ sign indicates whether the B_{tag} is tagged as a B^0 or a $\overline{B}^0$, respectively. The dilution factor $\mathcal{D}=1-2w$ accounts for the probability w that the flavor of the tagging B is identified incorrectly.

The distributions are much simpler when $|\lambda|=1$, which is the expectation of the standard model for decays like B^0

$\to J/\psi K_S^0$. If all the mechanisms that contribute to the decay have the same weak phase then the ratio of the weak decay amplitudes in Eq. (7) is just $\eta_{CP} e^{2i\phi_{\rm dec}}$, where $\phi_{\rm dec}$ is the weak phase for $\overline{B}^0 \to f$; $\phi_{\rm dec}$ is convention dependent and unobservable. The remaining factor introduces a phase due to B^0-$\overline{B}^0$ mixing. The combination of these phases is convention independent and observable.

For decays such as $B^0 \to J/\psi K_S^0$, or more generally $(c\overline{c})K_S^0$ and $(c\overline{c})K_L^0$, an explicit representation for λ can be found from the ratio of the amplitude for $\overline{B}^0 \to (c\overline{c})\overline{K}^0$ to the interfering process $\overline{B}^0 \to B^0 \to (c\overline{c})K^0 \to (c\overline{c})\overline{K}^0$. The decay $B^0 \to (c\overline{c})K^0$ involves a $\overline{b} \to \overline{c}c\overline{s}$ transition with an amplitude proportional to $[V_{cb}^* V_{cs}]$, while $\overline{B}^0 \to (c\overline{c})\overline{K}^0$ provides analogously a factor $\eta_{CP}[V_{cb}V_{cs}^*]$. Because $\overline{B}^0 \to B^0$ mixing is dominated by the loop diagram with a t quark, it introduces a factor $[V_{td}^* V_{tb}/V_{td}V_{tb}^*]$, while $K^0 \to \overline{K}^0$ mixing, being dominated by the c-quark loop, adds a factor $[V_{cd}V_{cs}^*/V_{cd}^* V_{cs}]$. Altogether, for transitions of the type $b \to c\overline{c}s$,

$$\begin{aligned}\lambda &= \eta_{CP}\left(\frac{V_{td}V_{tb}^*}{V_{td}^*V_{tb}}\right)\left(\frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}}\right)\left(\frac{V_{cd}^*V_{cs}}{V_{cd}V_{cs}^*}\right)\\ &= \eta_{CP}\left(\frac{V_{cb}V_{cd}^*}{V_{tb}V_{td}^*}\right)\left(\frac{V_{tb}^*V_{td}}{V_{cb}^*V_{cd}}\right)\\ &= \eta_{CP}e^{-2i\beta}. \end{aligned} \tag{9}$$

The time-dependent rate for decay of the B_{CP} final state is then given by

$$\begin{aligned} & f_\pm(\Delta t;\Gamma,\Delta m_d,w,\sin 2\beta)\\ &\quad = \frac{\Gamma}{4}e^{-\Gamma|\Delta t|}[1 \mp \eta_{CP}\mathcal{D}\sin 2\beta \sin \Delta m_d \Delta t]. \end{aligned} \tag{10}$$

In the limit of perfect determination of the flavor of the fully reconstructed B in the $B_{\rm flav}$ sample, which we assume throughout, the dilution here and in the mixed and unmixed amplitudes of Eq. (4) arises solely from the $B_{\rm tag}$ side, allowing the values of the mistag fractions w_i to be determined by studying the time-dependent rate of B^0-$\overline{B}^0$ oscillations.

To account for the finite resolution of the detector, the time-dependent distributions $f_\pm$ for B^0 and $\overline{B}^0$ tagged events [Eq. (10)] must be convolved with a time resolution function $\mathcal{R}(\delta_{\rm t} = \Delta t - \Delta t_{\rm true};\hat{a})$ as described above for mixing,

$$\begin{aligned} & \mathcal{F}_\pm(\Delta t;\Gamma,\Delta m_d,w,\sin 2\beta,\hat{a})\\ &\quad = f_\pm(\Delta t_{\rm true};\Gamma,\Delta m_d,w,\sin 2\beta)\otimes \mathcal{R}(\delta_{\rm t};\hat{a}), \end{aligned} \tag{11}$$

where $\hat{a}$ represents the set of parameters that describe the resolution function. In practice, events are separated into the same tagging categories as in mixing, each of which has a different mistag fraction w_i, determined individually for each category. Figure 2 illustrates the impact of typical mistag and Δt resolution effects on the Δt distributions for B^0- and $\overline{B}^0$-tagged CP events.

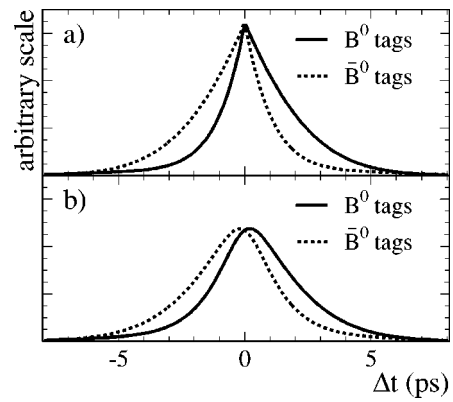

FIG. 2. Expected Δt distribution for B^0- and $\overline{B}^0$-tagged CP events (a) with perfect tagging and Δt resolution, and (b) with typical mistag rates and Δt resolution.

It is possible to construct a CP-violating observable

$$\mathcal{A}_{CP}(\Delta t) = \frac{\mathcal{F}_+(\Delta t) - \mathcal{F}_-(\Delta t)}{\mathcal{F}_+(\Delta t) + \mathcal{F}_-(\Delta t)}, \tag{12}$$

which, neglecting resolution effects, is proportional to $\sin 2\beta$:

$$\mathcal{A}_{CP}(\Delta t) \propto -\eta_{CP}\mathcal{D}\sin 2\beta \sin \Delta m_d \Delta t. \tag{13}$$

Since no time-integrated CP asymmetry effect is expected, an analysis of the time-dependent asymmetry is necessary. The interference between the two amplitudes, and hence the CP asymmetry, is maximal after approximately 2.1 B^0 proper lifetimes, when the mixing asymmetry goes through zero. However, the maximum sensitivity to $\sin 2\beta$, which is proportional to $e^{-\Gamma|\Delta t|}\sin^2 \Delta m_d \Delta t$, occurs in the region of 1.4 lifetimes.

The value of the free parameter $\sin 2\beta$ can be extracted with the tagged B_{CP} sample by maximizing the likelihood function

$$\begin{aligned} \ln \mathcal{L}_{CP} = \sum_i^{\rm tagging} \Bigg[& \sum_{B^0\ {\rm tag}} \ln \mathcal{F}_+(\Delta t;\Gamma,\Delta m_d,\hat{a},w_i,\sin 2\beta)\\ & + \sum_{\overline{B}^0\ {\rm tag}} \ln \mathcal{F}_-(\Delta t;\Gamma,\Delta m_d,\hat{a},w_i,\sin 2\beta)\Bigg], \end{aligned} \tag{14}$$

where the outer summation is over tagging categories i and the inner summations are over the B^0 and $\overline{B}^0$ tags within a given uniquely-assigned tagging category. In practice, the fit for $\sin 2\beta$ is performed on the combined flavor-eigenstate and CP samples with a likelihood constructed from the sum of Eqs. (6) and (14), in order to determine $\sin 2\beta$, the mistag fraction w_i for each tagging category, and the vertex resolution parameters $\hat{a}_i$. Additional terms are included in the likelihood to account for backgrounds and their time dependence.

The mistag rates can also be extracted with a time-integrated analysis as a cross-check. Neglecting possible background contributions and assuming the flavor of $B_{\rm flav}$ is correctly identified, the observed time-integrated fraction of mixed events χ_{obs} can be expressed as a function of the B^0-$\overline{B}^0$ mixing probability χ_d:

$$\chi_{obs}=\chi_d+(1-2\chi_d)w, \tag{15}$$

where $\chi_d=\frac{1}{2}x_d^2/(1+x_d^2)=0.174\pm0.009$ [11] and $x_d=\Delta m_d/\Gamma$. Taking advantage of the available decay time information, the statistical precision on w can be improved by selecting only events that fall into an optimized time interval $|\Delta t|<t_0$, where t_0 is chosen so that the integrated number of mixed and unmixed events are equal outside this range. With the use of such an optimized Δt interval the time-integrated method achieves nearly the same statistical precision for the mistag rates as a full time-dependent likelihood fit.

C. Overview of the analysis

This article provides a detailed description of our published measurement of flavor oscillations [15] and CP-violating asymmetry [16] in the neutral B meson system. These measurements have six main components:

(1) Selection of the B_{CP} sample of signal events for neutral B decays to CP modes $J/\psi K_S^0$, $\psi(2S)K_S^0$, $\chi_{c1}K_S^0$, $J/\psi K^{*0}$ ($K^{*0}\to K_S^0\pi^0$), and $J/\psi K_L^0$; selection of the $B_{\rm flav}$ sample of signal events for neutral flavor-eigenstate decays to $D^{(*)-}\pi^+/\rho^+/a_1^+$ and $J/\psi K^{*0}$ ($K^{*0}\to K^+\pi^-$); selection of the B^+ control sample in the modes $\overline{D}^{(*)0}\pi^+$, $J/\psi K^{(*)+}$, $\psi(2S)K^+$, $\chi_{c1}K^+$; and selection of a semileptonic neutral B sample in the mode $D^{*+}l^-\overline{\nu}$, as described in Sec. III.

(2) Determination of the flavor of the $B_{\rm tag}$, as described in Sec. IV.

(3) Measurement of the distance Δz between the two B^0 decay vertices along the $\Upsilon(4S)$ boost axis, and its conversion to Δt, as described in Sec. V.

(4) Construction of a log-likelihood function to describe the time evolution of signal and background events in the presence of mixing and CP asymmetries, as described in Sec. VI.

(5) Measurement of the mixing rate Δm_d, mistag fractions w_i, and vertex resolution parameters $\hat{a}_i$ for the different tagging categories i, with an unbinned maximum-likelihood fit to the $B_{\rm flav}$ sample, as described in Sec. VII.

(6) Extraction of a value of $\sin 2\beta$, or more generally $\mathrm{Im}\,\lambda/|\lambda|$ and $|\lambda|$, from the amplitude of the CP asymmetry, the mistag fractions w_i, and the vertex resolution parameters $\hat{a}_i$ for the different tagging categories i, with an unbinned maximum-likelihood fit to the combined $B_{\rm flav}$ and B_{CP} samples, as described in Sec. VIII.

Whenever possible, we determine time and mass resolutions, efficiencies and mistag fractions from the data. The measurement of Δm_d is performed with a slightly reduced subset of the full $B_{\rm flav}$ sample, which is optimized for such a precision measurement. The B_{CP} sample is not included, since this would add additional assumptions about the resolution function without significantly improving the precision of Δm_d. The measurement of $\sin 2\beta$ is performed with the full $B_{\rm flav}$ and B_{CP} samples, with a fixed value for Δm_d and the B^0 lifetime. This strategy allows us to account correctly for the small correlations among the mistag rates, Δt resolutions parameters, and $\sin 2\beta$. The same $B_{\rm flav}$ sample and vertex separation techniques have been used to determine precision values for the charged and neutral B lifetimes [17].

II. THE *BABAR* DETECTOR AND DATA SETS

The data used in this analysis were recorded with the *BABAR* detector [18] at the PEP-II collider [8] in the period October 1999–June 2001. The total integrated luminosity of the data set is equivalent to 29.7 fb^{-1} collected near the $\Upsilon(4S)$ resonance and 3.9 fb^{-1} collected 40 MeV below the $\Upsilon(4S)$ resonance (off-resonance data). The corresponding number of produced $B\overline{B}$ pairs is estimated to be about 32 million. The $\Upsilon(4S)$ sample is sometimes divided into two subsamples for comparison purposes: data recorded in 1999–2000, about 20.7 fb^{-1} and referred to as "run 1," and data recorded in 2001, about 9.0 fb^{-1} and referred to as "run 2." These subsamples differ primarily in the quality of the tracking system alignment and on the track-finding efficiency. The former requires a separate treatment of the Δt resolution for the two periods, as discussed in Sec. V F, while the latter results in substantially improved yields in run 2 for reconstructed B mesons.

A. The *BABAR* detector

The *BABAR* detector is a charged and neutral spectrometer with large solid-angle coverage. For this analysis, the most important detector capabilities include charged-particle tracking, vertex reconstruction, and particle identification. Charged particles are detected and their momenta measured by a combination of a 40-layer, small-celled drift chamber (DCH) filled with a 80:20 helium:isobutane gas mixture, and a five-layer silicon vertex tracker (SVT), consisting of 340 ac-coupled double-sided silicon microstrip sensors. The cells of the DCH are organized into 10 superlayers within which the sense wires all have the same orientation, thereby allowing segment-based tracking. Both the DCH and the SVT lie inside a 1.5-T solenoidal magnetic field. Beyond the outer radius of the DCH is a detector of internally reflected Cherenkov radiation (DIRC), which is used primarily for charged-hadron identification. The device consists of 144 fused silica quartz bars in which relativistic charged particles above the Cherenkov threshold radiate photons while traversing the material. The light is transported by total internal reflection down the length of the bars to an array of 10752 photomultiplier tubes mounted on the rear of the detector, where the opening angle of the Cherenkov ring is measured. A finely segmented electromagnetic calorimeter (EMC), consisting of 6580 CsI(T1) crystals, is used to detect photons and neutral hadrons, and also to identify electrons. The EMC is surrounded by a thin cylindrical superconducting coil and a segmented iron flux return, organized into a hexagonal bar-

rel and two endcaps. The instrumented flux return (IFR) consists of multiple layers of resistive plate chambers (RPCs) interleaved with the flux-return iron and is used in the identification of muons and neutral hadrons.

B. Charged particle reconstruction

Charged track finding starts with pattern recognition in the DCH, based on three different algorithms. The first uses the same fast algorithm employed by the level-3 trigger for finding and linking superlayer-based track segments from moderate-to-high p_T tracks originating from the interaction point. Two subsequent track finders then work on superlayer segments not already attached to a reconstructed track. They are designed to find tracks with lower p_T, passing through fewer than the full ten superlayers of the chamber, or originating away from the interaction point. At the end of this process, all tracks are refit with a Kalman-filter fitter [19] that takes into account the detailed distribution of material in the detector and the nonuniformities in the detector magnetic field. These tracks are then projected into the SVT, and silicon-strip hits are added if they are consistent within the extrapolation errors through the intervening material and field. A search is performed for tracks that are reconstructed with the remaining unused silicon clusters, again with two different algorithms. At the end of the SVT-only track finding, an attempt is made to match SVT- and DCH-only track segments, which may result when a hard scatter occurs in the support tube material between the two devices.

Charged-particle transverse momenta p_T are determined with a resolution parametrized by $\sigma(p_T)/p_T = 0.0013(p_T/\mathrm{GeV}/c) + 0.0045$. The SVT, with typical single-hit resolution of 10 μm, provides vertex information in both the transverse plane and in z, as well as the decay angles at the interaction point. Decay vertices for B meson candidates are typically reconstructed with a resolution of 50 μm in z for fully reconstructed modes and about 100 to 150 μm for the vertex of the (unreconstructed) tagging B meson in the event. The efficiency for finding tracks in hadronic events that traverse the full DCH radius ($p_T > 200$ MeV/c) is about 90% for run 1 and 95% for run 2.

C. Neutral reconstruction

EMC clusters are formed around initial seed crystals containing at least 10 MeV of deposited energy. Neighboring crystals are added to the cluster if their energy exceeds 1 MeV. If the newly added crystal has energy greater than 3 MeV, its contiguous neighbors (including corners) are also considered for inclusion in the cluster. In order to identify cases where several showers are in close proximity, such as unresolved photons from high-energy π^0 decays, local maxima within a cluster are identified. These local maxima are defined as candidate crystals that have an energy exceeding each of its neighbors by a fraction that depends on the number of crystals in the local neighborhood. Clusters are then divided into as many "bumps" as there are local maxima. The division is based on a two-dimensional weighting scheme that assumes electromagnetic shower shapes to divide up the cluster energy. The position of each bump is calculated with a logarithmic weighting of crystal energies.

We determine whether a bump is associated with a charged or neutral particle by projecting all tracks in the event to the inner face of the calorimeter. A bump is determined to be neutral, and therefore a photon candidate, if no track intersects any of its crystals. A track intersection is determined by computing the two-dimensional distance on the face of the calorimeter from the projected track impact point to the bump centroid. A requirement is made on the difference between the measured intersection distance and the Monte Carlo expectation for different particle species based on the measured track parameters.

The energy resolution in the EMC is measured directly with a radioactive source at low energy under ideal low-background conditions and with electrons from Bhabha scattering at high energy, from which we determine $\sigma(E)/E = (5.0 \pm 0.8)\%$ at 6.13 MeV and $(1.9 \pm 0.07)\%$ at 7.5 GeV. The energy resolution can also be extracted from the observed mass resolutions for π^0 and η decays to two photons, which are measured to be around 7 MeV and 16 MeV, respectively. A fit to the observed resolutions obtained from the π^0, η, and Bhabha samples gives a photon energy resolution parametrized by $\sigma(E)/E = 0.023(E/\mathrm{GeV})^{-1/4} \oplus 0.019$.

D. Particle identification

Identification of electrons, muons and kaons is an essential ingredient in both B reconstruction and flavor tagging. Particle species can be distinguished by measurements of the specific energy loss (dE/dx) in the SVT layers and in the DCH gas along the particle trajectory, the number of Cherenkov photons and the Cherenkov angle in the DIRC, the electromagnetic shower energy in the EMC, and the particle penetration length in the IFR. Selection criteria are based on these quantities, on likelihood ratios derived from them, or on neural network algorithms combining different detector likelihoods. Typically, looser selection criteria are applied for B reconstruction than for B-flavor tagging. Efficiencies and particle misidentification probabilities are determined from data control samples with similar characteristics.

1. Electron identification

Electron candidates are identified primarily by the ratio of the bump energy in the electromagnetic calorimeter to the track momentum, E/p. They also must have a measured mean dE/dx in the DCH that is consistent with the electron hypothesis. In addition, for some applications, the lateral and azimuthal shape of the EMC shower [20,21] and the consistency of the observed and expected Cherenkov angle in the DIRC are used for identification. Four different categories of electron candidates (`VeryLoose`, `Loose`, `Tight`, and `VeryTight`) are defined with the criteria listed in Table I. Candidates that are not matched to an EMC bump are retained as `noCal` electron candidates if their measured dE/dx satisfies the same requirements as the `VeryTight` selection. Electron identification efficiencies in the momentum range $0.5 < p < 3.0$ GeV/c vary between 88% and 98%

TABLE I. Criteria used for selecting the available categories of electron candidates. The difference between the measured mean dE/dx and the expectation for an electron is required to lie within the interval specified in terms of the expected dE/dx resolution σ.

Category	dE/dx	E/p	Cumulative additional requirements
`VeryLoose`	$[-3\sigma, 7\sigma]$	>0.50	
`Loose`	$[-3\sigma, 7\sigma]$	>0.65	
`Tight`	$[-3\sigma, 7\sigma]$	[0.75, 1.3]	Lateral shower shape
`VeryTight`	$[-2.2\sigma, 4\sigma]$	[0.89, 1.2]	Azimuthal shower shape; consistency of DIRC Cherenkov angle (3σ)

for the criteria in Table I, while the pion misidentification rates are below 0.3% for the `VeryTight` selection.

2. Muon identification

Muon candidates are primarily identified by the measured number of hadronic interaction lengths n_λ traversed from the outside radius of the DCH through the IFR iron, and the difference Δn_λ between n_λ and the predicted penetration depth for a muon of the same momentum and angle. Contamination from hadronic showers is rejected by a combination of the average number $\overline{n}_{hits}$ and the variance $\sigma_{n_{hits}}$ of hits per RPC layer, the χ^2 for the geometric match between the track extrapolation into the IFR and the RPC hits, χ^2_{trk}, and the χ^2 of a polynomial fit to the RPC hits, χ^2_{fit}. In addition, for those muons within the acceptance of the EMC, we require the calorimeter bump energy E to be consistent with a minimum ionizing particle. Four different categories of muon candidates (`VeryLoose`, `Loose`, `Tight`, and `VeryTight`) are selected with the criteria listed in Table II. In the forward region, which suffers from some machine background, additional requirements are made on the fraction of RPC layers with hits. Muon identification efficiencies in the momentum range $1.1<p<3.0$ GeV/c vary between 60% to 92% for the criteria in Table II, while pion misidentification rates are about 3% for the `Tight` selection.

3. Kaon identification

Kaons are distinguished from pions and protons on the basis of specific energy-loss measurements dE/dx in SVT and DCH and the number of Cherenkov photons and the Cherenkov angle in the DIRC. The difference between the measured truncated-mean dE/dx in the DCH and the expected mean for the pion, kaon and proton hypothesis, with typical resolution of 7.5%, is used to compute likelihoods $\mathcal{L}_\pi$, $\mathcal{L}_K$ and $\mathcal{L}_p$ assuming Gaussian distributions. Similarly, the difference between the measured 60% truncated-mean dE/dx in the SVT and the expected dE/dx is described by an asymmetric Gaussian distribution. For minimum-ionizing particles the resolution on the SVT truncated mean is about 14%. In the DIRC, a likelihood is obtained for each particle hypothesis from the product of two components: the expected number of Cherenkov photons, with a Poisson distribution, and the difference between the measured average Cherenkov angle to the expected angle for a given mass hypothesis, assuming a Gaussian distribution.

For B-flavor tagging the likelihood variables from SVT, DCH and DIRC are combined as inputs to a neural network whose output is a single discriminating variable for kaon selection. The network is trained with Monte Carlo simulation of generic B decays. The average efficiency of the selection is about 85% for a pion-misidentification probability of about 2.5%. Further details are described in Sec. IV B.

The exclusive reconstruction of many B meson final states does not generally require explicit kaon identification. For some channels a `VeryLoose` kaon selection based on likelihood ratios is imposed to reduce backgrounds to acceptable levels. The combined likelihood uses the individual likelihoods from SVT and DCH for momenta below 0.5 GeV/c, from DCH only for momenta between 0.5 and 0.6 GeV/c, and from DIRC only for momenta above 0.6 GeV/c. Kaon candidates are rejected if the likelihood ratios satisfy $\mathcal{L}_K/\mathcal{L}_\pi<r$ and $\mathcal{L}_K/\mathcal{L}_p<r$, where $r=0.1$ for $p<0.5$ GeV/c and $r=1$ for $p\geqslant 0.5$ GeV/c. Tracks with no particle information are assumed to be kaons. This `VeryLoose` kaon requirement has a nearly constant kaon-identification efficiency of about 96% and a pion-misidentification probability of at most 15% for tracks in the transverse momentum range 1 to 2.5 GeV/c. Tighter kaon selections require $\mathcal{L}_K/\mathcal{L}_\pi>r$, with r typically greater than one. For a loose pion selection, candidates are rejected if they satisfy tighter kaon or lepton criteria.

TABLE II. Criteria used for selecting available categories of muon candidates.

Category	n_λ	Δn_λ	$\overline{n}_{hits}$	$\sigma_{n_{hits}}$	χ^2_{trk}/n_{layers}	χ^2_{fit}/n_{layers}	E_{EMC} [GeV]
`VeryLoose`	>2.0	<2.5	<10	<6			<0.5
`Loose`	>2.0	<2.0	<10	<6	<7	<4	<0.5
`Tight`	>2.2	<1.0	<8	<4	<5	<3	[0.05, 0.4]
`VeryTight`	>2.2	<0.8	<8	<4	<5	<3	[0.05, 0.4]

TABLE III. Signal yields for the different samples used in this analysis, before any tagging or tagging vertex requirements. The signal size and purity for B decays to hadronic final states are obtained from a fit to the $m_{\rm ES}$ distribution described in Sec. III C, after selection on ΔE. Purities are quoted for $m_{\rm ES} > 5.27$ MeV/c^2. The results for $J/\psi K_L^0$ are obtained from a fit to the ΔE distribution described in Sec. III D. The purity for $J/\psi K_L^0$ is quoted for events with $\Delta E < 10$ MeV. The results for $D^{*-} l^+ \nu$ are obtained from a fit to the $\cos\theta_{B-D^*l}$ distribution described in Sec. III E. Purity is quoted for $-1.1 < \cos\theta_{B-D^*l} < 1.1$.

Sample	Final state	Signal	Purity (%)
B_{CP}	$J/\psi K_S^0$ ($K_S^0 \to \pi^+\pi^-$)	461 ± 22	99
	$J/\psi K_S^0$ ($K_S^0 \to \pi^0\pi^0$)	113 ± 12	93
	$\psi(2S) K_S^0$	86 ± 17	96
	$\chi_{c1} K_S^0$	44 ± 8	98
	$J/\psi K^{*0}$ ($K^{*0} \to K_S^0 \pi^0$)	64 ± 10	74
	$J/\psi K_L^0$	257 ± 24	60
	Total	1025 ± 41	83
$B_{\rm flav}$	$D^{*-}\pi^+$	2380 ± 57	92
	$D^{*-}\rho^+$	1438 ± 52	84
	$D^{*-}a_1^+$	1146 ± 45	80
	$D^-\pi^+$	2685 ± 65	83
	$D^-\rho^+$	1421 ± 57	74
	$D^-a_1^+$	845 ± 44	67
	$J/\psi K^{*0}$ ($K^{*0} \to K^+\pi^-$)	1013 ± 36	95
	Total	10941 ± 133	83
B^+	$\overline{D}^0\pi^+$	6850 ± 102	83
	$\overline{D}^{*0}\pi^+$	1708 ± 51	91
	$J/\omega K^+$	1921 ± 46	97
	$\psi(2S) K^+$	292 ± 18	98
	$\chi_{c1} K^+$	195 ± 29	95
	$J/\psi K^{*+}$ ($K^{*+} \to K^+\pi^0$)	384 ± 25	87
	Total	11343 ± 129	86
Semi-leptonic B^0	$D^{*-} l^+ \nu$	29042 ± 1500	78

III. RECONSTRUCTION OF *B* MESONS

Neutral B mesons in flavor eigenstates are reconstructed in the hadronic final states $B^0 \to D^{(*)-}\pi^+$, $D^{(*)-}\rho^+$, $D^{(*)-}a_1^+$, and $J/\psi K^{*0}(K^+\pi^-)$, and the semileptonic decay mode $B^0 \to D^{*-} l^+ \nu$. The CP sample is reconstructed in the channels $B^0 \to J/\psi K_S^0$, $\psi(2S)K_S^0$, $\chi_{c1}K_S^0$, $J/\psi K^{*0}$ ($K^{*0} \to K_S^0\pi^0$) and $J/\psi K_L^0$. In some cases, control samples of charged B decays are studied, where the hadronic final states $B^+ \to \overline{D}^{(*)0}\pi^+$, $J/\psi K^{(*)+}$, $\psi(2S)K^+$ and $\chi_{c1}K^+$ are used. All final-state particles, with the exception of the neutrino in the semileptonic decay, are reconstructed. A number of $\overline{D}^0$ and D^- decay modes are used to achieve reasonable reconstruction efficiency despite the typically small branching fractions for any given B or D decay channel. A summary of the various reconstructed B samples and purities is provided in Table III. In the following, kinematic quantities and selection criteria are given in the laboratory frame, unless otherwise specified.

A. Event selection

Multihadron events are selected by demanding a minimum of three reconstructed charged tracks in the polar angle range $0.41 < \theta_{lab} < 2.54$ rad. Charged tracks must be reconstructed in the DCH and are required to originate within 1.5 cm in xy and 10 cm in z of the nominal beam spot. A primary vertex is formed on an event-by-event basis from a vertex fit to all charged tracks in the fiducial volume. Tracks with a large χ^2 contribution to the vertex fit are removed until an overall χ^2 probability greater than 1% is obtained or only two tracks remain. The resolution achieved by this method is about 70 μm in x and y for hadronic events. Events are required to have a primary vertex within 0.5 cm of the average position of the interaction point in the plane transverse to the beamline, and 6 cm longitudinally. Electromagnetic bumps in the calorimeter in the polar angle range $0.410 < \theta_{lab} < 2.409$ rad that are not associated with charged tracks, have an energy greater than 30 MeV, and a shower shape consistent with a photon interaction are taken as neutrals. A total energy greater than 4.5 GeV in the fiducial regions for charged tracks and neutrals is required. To reduce continuum background, we require the normalized second Fox-Wolfram moment [22] R_2 of the event, calculated with both charged tracks and neutrals, to be less than 0.5 (0.45) in hadronic (semileptonic) decay modes. The lth Fox-Wolfram

moment is the momentum-weighted sum of Legendre polynomial of the lth order computed from the cosine of the angle between all pairs of tracks. The ratio R_2 provides good separation between jet-like continuum events and more spherical $B\overline{B}$ events.

B. Reconstruction of decay daughters

The reconstruction of B mesons typically involves the summation of a set of related decay modes, with multiple decay chains for the charm daughters or other short-lived decay products. To simplify analysis of such complex decay chains, virtual *composite* particles and their error matrices are constructed from the original daughter particles. The composite particle then replaces the daughters in subsequent fits and analysis. The three-momentum of the virtual particle is fit directly, rather than computed from the updated daughters, improving speed and numerical accuracy.

Vertex and kinematic fitting is used to improve four-momenta and position measurements, as well as to measure the time difference between decaying B hadrons in the $\Upsilon(4S)\rightarrow B\overline{B}$ decay. For example, in the case of $B^0 \rightarrow J/\psi K_S^0$, the position measurement of the B^0 can be improved with the constraint that the line-of-flight of the K_S^0 intersects the J/ψ vertex. The energy resolution of the B^0 can also be improved by applying a mass constraint to the J/ψ and K_S^0 daughters. Generalized procedures have been developed and tested with constraints implemented by the Lagrange-multiplier technique. Possible constraints include a common decay vertex, mass, energy, momentum, beam energy (with and without smearing), beam-spot position and line-of-flight.

Nonlinearities in the fits require the use of an iterative procedure, where convergence is defined by demanding that the change in χ^2 between two successive iterations is less than 0.01, within a maximum of six iterations. Simple fits involving only vertex constraints (except long-lived particles) are, however, accurate enough with a single iteration.

1. π^0 selection

Neutral pion candidates are formed from pairs of EMC bumps with energy greater than 30 MeV, assumed to be photons originating from the interaction point. The invariant mass of the photon pair is required to be within $\pm 20\ \mathrm{MeV}/c^2$ (2.5σ) of the nominal π^0 mass, with a minimum summed energy of 200 MeV. Selected candidates are subjected to a kinematic fit with a π^0 mass constraint. Within the acceptance of the EMC, efficiencies for this selection vary from about 55 to 65% for π^0 energies from 0.3 to 2.5 GeV, typical of B decays.

2. K_S^0 selection

Candidates in the $K_S^0 \rightarrow \pi^+\pi^-$ mode are selected by requiring an invariant $\pi^+\pi^-$ mass, computed at the vertex of the two tracks, between 462 and 534 MeV/c^2. The χ^2 of the vertex fit must have a probability greater than 0.1%. The angle between the flight direction and the momentum vector for the K_S^0 candidate is required to be smaller than 200 mrad. Finally, the transverse flight distance from the primary vertex in the event, r_{xy}, must be greater than 2 mm.

Optimization for the reconstruction of the CP sample has produced slightly different K_S^0 selection criteria. The $\pi^+\pi^-$ invariant mass, determined at the vertex of the two tracks, is required to lie between 489 and 507 MeV/c^2 and the three-dimensional flight length with respect to the vertex of the charmonium candidate is required to be greater than 1 mm.

Pairs of π^0 candidates, each in the mass range $100-155\ \mathrm{MeV}/c^2$ (-5σ, $+3\sigma$) and formed from nonoverlapping EMC bumps, are combined to construct $K_S^0 \rightarrow \pi^0\pi^0$ candidates. For each K_S^0 candidate with an energy greater than 800 MeV and a mass between 300 and 700 MeV/c^2 at the interaction point, we determine the most probable K_S^0 decay point along the path defined by the initial K_S^0 momentum vector and the J/ψ vertex by maximizing the product of probabilities for the daughter π^0 mass-constrained fits. Allowing for vertex resolution, we require the distance from the decay point to the J/ψ vertex to be between -10 and $+40$ cm and the K_S^0 mass, using the measured decay point, to be between 470 and 536 MeV/c^2.

3. K_L^0 selection

Candidates for K_L^0 mesons are identified in the EMC and IFR detectors as reconstructed clusters that cannot be associated with any charged track in the event. EMC candidates must have a cluster energy between 200 MeV and 2 GeV and a polar angle θ that satisfies $\cos\theta < 0.935$. To suppress backgrounds from π^0 decay, K_L^0 candidates consistent with a photon are paired with other neutrals with $E_\gamma > 30$ MeV. Any candidate with $100 < m(\gamma\gamma) < 150\ \mathrm{MeV}/c^2$ is rejected. Likewise, clusters with more than 1 GeV energy that contain two bumps are rejected if the bump energies and shower shapes are consistent with two photons from a π^0 decay. Monte Carlo simulation shows that clusters due to true K_L^0 mesons are easily distinguished from π^0 candidates by these criteria. The remaining background consists primarily of photons and overlapping showers. Isolated clusters produced by charged hadrons are removed by the basic clustering algorithm, which requires a minimum separation of about 20 cm between clusters.

IFR candidates are defined as clusters with hits in two or more RPC layers that are not matched to any reconstructed charged track. To reduce beam-related backgrounds and to avoid regions where the charged tracking efficiency is low, we require that the polar angle θ of the IFR cluster satisfy $-0.75 < \cos\theta < 0.93$, and eliminate clusters that begin in the outer 25% of the forward IFR endcap. Due to the irregular structure of hadronic showers, some hits from charged tracks are missed by the tracking association. We suppress these clusters by rejecting K_L^0 candidates that lie within ± 350 mrad in polar angle and in the range $-750(-300)$ to $+300(+750)$ mrad in azimuth of the EMC intersection of any positively (negatively) charged track in the event. The remaining background is predominantly from charged particles and detector noise.

Some K_L^0 candidates satisfy both the EMC and IFR selection requirements. In the reconstruction of $B^0 \rightarrow J/\psi K_L^0$, ad-

TABLE IV. Particle identification and invariant mass requirements for J/ψ and $\psi(2S)\to l^+l^-$ candidates. The minimal particle identification criteria are applied to both daughters, while only one daughter must pass the restrictive requirement. Electron and muon selection requirements are defined in Sec. II D. Mass ranges are quoted in GeV/c^2 and MIP refers to a minimum-ionizing particle.

	e^+e^- candidates			$\mu^+\mu^-$ candidates		
B channel	Minimal	Restrictive	$m(e^+e^-)$	Minimal	Restrictive	$m(\mu^+\mu^-)$
$J/\psi K_S^0$	None	`Tight` or `noCal`	2.95–3.14	`MIP`	`Loose`	3.06–3.14
$\psi(2S)K_S^0(l^+l^-)$	`VeryLoose`	`Tight`	3.436–3.736	`VeryLoose`	`Loose`	3.636–3.736
$\psi(2S)K_S^O(J/\psi\pi^+\pi^-)$	`VeryLoose`	`Tight`	2.95–3.14	`VeryLoose`	`Loose`	3.06–3.14
$\chi_{c1}K_S^0(J/\psi\gamma)$	`Loose`	`Tight`	2.95–3.14	`VeryLoose`	`Loose`	3.06–3.14
$J/\psi K^*$	`Tight`	`Tight`	2.95–3.14	`Loose`	`Loose`	3.06–3.14
$J/\psi K_L^0$	`Loose`	`VeryTight`	3.00–3.13	`Loose`	`Tight`	3.06–3.13

ditional criteria described in Sec. III D are applied to resolve the classification of the corresponding B candidates. Extensive studies of K_L^0 detection efficiencies have been conducted with a control sample of radiatively produced ϕ mesons, decaying to $K_L^0K_S^0$.

4. Selection of light resonances

For ρ^- candidates, the $\pi^-\pi^0$ mass is required to lie within ± 150 MeV/c^2 of the nominal ρ^- mass. The π^0 from the ρ^- decay is required to have an energy greater than 300 MeV. We reconstruct K^{*0} candidates in the $K^+\pi^-$ and $K_S^0\pi^0$ modes, while K^{*+} candidates are reconstructed in the $K^+\pi^0$ and $K_S^0\pi^+$ modes. The invariant mass of the two daughters is required to be within ± 100 MeV/c^2 of the nominal K^* mass. Candidates in the mode $a_1^+\to\pi^+\pi^-\pi^+$ are reconstructed by combining three charged pions, with invariant mass in the range of 1.0 to 1.6 GeV/c^2. In addition, the χ^2 probability of a vertex fit of the a_1^+ candidate is required to be greater than 0.1%.

5. Charmed meson and charmonium selection

The decay channels $K^+\pi^-$, $K^+\pi^-\pi^0$, $K^+\pi^+\pi^-\pi^-$ and $K_S^0\pi^+\pi^-$ are used to reconstruct $\overline{D}^0$ candidates, while D^- candidates are selected in the $K^+\pi^-\pi^-$ and $K_S^0\pi^-$ modes. Charged and neutral kaons are required to have a momentum greater than 200 MeV/c. The same criterion is applied to the pion in $B^0\to D^{(*)-}\pi^+$, $B^0\to D^{(*)-}\rho^+$ decay. For the decay modes $B^0\to D^{(*)-}a_1^+$, the pions are required to have momentum larger than 150 MeV/c. We require $\overline{D}^0$ and D^- candidates to lie within $\pm 3\sigma$ of the nominal masses, where the error σ is calculated event by event. The distributions of the difference between measured and nominal $\overline{D}^0$ and D^- meson masses, normalized by the measured error on the candidate masses, are found to have an rms in the range 1.1–1.2 when fit with a Gaussian distribution. For $\overline{D}^0\to K^+\pi^-\pi^0$, we only reconstruct the dominant resonant mode $\overline{D}^0\to K^+\rho^-$, followed by $\rho^-\to\pi^-\pi^0$. The angle $\theta^*_{D^0\pi}$ between the π^- and $\overline{D}^0$ in the ρ^- rest frame must satisfy $|\cos\theta^*_{D^0\pi}|>0.4$. Finally, all $\overline{D}^0$ and D^- candidates are required to have a momentum greater than 1.3 GeV/c in the $\Upsilon(4S)$ frame and a χ^2 probability for the topological vertex fit greater than 0.1%. A mass-constrained fit is applied to candidates satisfying these requirements.

We form D^{*-} candidates in the decay $D^{*-}\to\overline{D}^0\pi^-$ by combining a $\overline{D}^0$ with a pion that has momentum greater than 70 MeV/c. The soft pion is constrained to originate from the beam spot when the D^{*-} vertex is computed. To account for the small energy release in the decay $\Upsilon(4S)\to B\overline{B}$ (resulting in a small transverse flight of the B candidates), the effective vertical size of the beam spot is increased to 40 μm. Monte Carlo simulation was used to verify that this does not introduce any significant bias in the selection or in the Δt measurement. After applying a mass-constrained fit to the $\overline{D}^0$ daughter, D^{*-} candidates are required to have $m(\overline{D}^0\pi^-)$ within ± 1.1 MeV/c^2 of the nominal D^{*-} mass for the $\overline{D}^0\to K^+\pi^-\pi^0$ mode and ± 0.8 MeV/c^2 for all other modes. This corresponds to about ± 2.5 times the rms width of the signal distribution, which is estimated by taking a weighted average of the core and broad Gaussian components of the observed $m(\overline{D}^0\pi^-)$ distributions.

We form $\overline{D}^{*0}$ candidates by combining a $\overline{D}^0$ with a π^0 with momentum less than 450 MeV/c in the $\Upsilon(4S)$ frame. $\overline{D}^{*0}$ candidates are required to have $m(\overline{D}^0\pi^0)$ within ± 4 MeV/c^2 of the nominal value, after applying a mass-constrained fit to the $\overline{D}^0$ daughter.

Candidates for J/ψ and $\psi(2S)$ mesons are reconstructed in their e^+e^- and $\mu^+\mu^-$ decay modes, while $\psi(2S)$ mesons are also reconstructed in the $J/\psi\pi^+\pi^-$ channel. Table IV shows the particle identification and invariant mass requirements for the e^+e^- and $\mu^+\mu^-$ daughters. These vary with reconstructed B decay channel due to the differing levels of background encountered. For $J/\psi\to e^+e^-$ and $\psi(2S)\to e^+e^-$ decays, where the electron may have radiated bremsstrahlung photons, the missing energy is recovered by identifying clusters with more than 30 MeV lying within 35 mrad in polar angle and 50 mrad in azimuth of the electron direction projected onto the EMC.

For the $\psi(2S)\to J/\psi\pi^+\pi^-$ mode, J/ψ candidates are constrained to the nominal mass and then combined with pairs of oppositely-charged tracks considered as pions, with invariant mass between 400 and 600 MeV/c^2. Candidates

with $0.574 < m(J/\psi\pi^+\pi^-) - m(J/\psi) < 0.604\ \mathrm{GeV}/c^2$ are retained.

Photon candidates used for the reconstruction of $\chi_{c1} \to J/\psi\gamma$ are required to lie within the calorimeter fiducial volume ($0.41 < \theta_\gamma < 2.41$ rad) and have an energy greater than 150 MeV. In addition, the candidate should not form, in combination with any other photon in the event having at least 70 MeV of energy, a π^0 candidate with mass between 120 and 150 MeV/c^2. The invariant mass of the χ_{c1} candidates is required to be greater than 3.476 and smaller than 3.546 GeV/c^2.

C. *B* meson selection in fully reconstructed modes

We reconstruct B candidates in all modes except $B^0 \to J/\psi K_L^0$ and $B^0 \to D^{*-} l^+ \overline{\nu}$ using a pair of nearly uncorrelated kinematic variables, the difference ΔE between the energy of the B candidate and the beam energy in the $\Upsilon(4S)$ center-of-mass frame, and the beam-energy substituted mass, m_{ES}, defined as

$$m_{\mathrm{ES}} = \sqrt{\left(\frac{\frac{1}{2}s + \mathbf{p}\cdot\mathbf{p}_i}{E_i}\right)^2 - p^2}, \tag{16}$$

where s is the square of the center-of-mass energy, E_i and $\mathbf{p}_i$ are the total energy and the three momentum of the initial state in the laboratory frame, and $\mathbf{p}$ is the three momentum of the B candidate in the same frame. For the purpose of determining event counts and purities, a signal region is defined in the $(m_{\mathrm{ES}}, \Delta E)$ plane as $5.27 < m_{\mathrm{ES}} < 5.29\ \mathrm{GeV}/c^2$ and $|\Delta E| < 3\sigma(\Delta E)$, where $\sigma(\Delta E)$ is the resolution on ΔE. Likewise, a sideband region is defined as $5.20 < m_{\mathrm{ES}} < 5.26\ \mathrm{GeV}/c^2$ and $|\Delta E| < 3\sigma(\Delta E)$. The value of $\sigma(\Delta E)$ is mode-dependent and varies between 7 and 40 MeV as measured in the data. When multiple B candidates (with $m_{\mathrm{ES}} > 5.20\ \mathrm{MeV}/c^2$) are found in the same event, the candidate with the smallest value of $|\Delta E|$ is selected.

Two types of background in the sample of selected B^0 candidates are distinguished. The first background, called combinatorial, arises from random combinations of charged tracks and neutral showers from both B mesons in $B\overline{B}$ events or from continuum events. This background is smoothly distributed in m_{ES} and does not peak near the B mass. The second, so-called "peaking" background, consists of events in which, for example, a slow pion from the reconstructed B meson is replaced by a slow pion from the tagging B, causing an enhancement near the nominal B mass. The loss or addition of a soft pion track does not significantly alter the vertex for the B_{rec} candidate, since such tracks undergo significant multiple scattering and have large impact parameter errors. The peaking background from charged B decays is considered as a specific background source in the construction of the full likelihood function for B^0-$\overline{B}^0$ mixing, since these events have a particular time structure and set of effective dilutions. In this case, the peaking background from other neutral B decays has time-dependent properties and dilutions that are essentially identical to the signal and is treated as such. For the likelihood describing the CP sample, the peaking background is simply assumed to have zero effective CP.

Suppression of continuum background, in addition to a general requirement on R_2, is typically provided by restricting the thrust angle θ_{th}, defined as the angle between the thrust axis of the particles that form the reconstructed B_{rec} candidate and the thrust axis of the remaining tracks and unmatched clusters in the event, computed in the $\Upsilon(4S)$ frame. The two thrust axes are almost uncorrelated in $B\overline{B}$ events, because the B^0 mesons are nearly at rest in the $\Upsilon(4S)$ rest frame. In continuum events, which are more jet-like, the two thrust axes tend to have small opening angles. Thus, a requirement on the maximum value of $|\cos\theta_{\mathrm{th}}|$ is effective in continuum rejection.

Signal yields and sample purities are extracted from fits to the m_{ES} distributions of B candidates with a Gaussian distribution for the signal and an ARGUS background shape [23] for the combinatorial background with a functional form given by

$$\mathcal{A}(m_{\mathrm{ES}}; m_0, \xi) = A_B m_{\mathrm{ES}} \sqrt{1 - x_{\mathrm{ES}}^2}\, e^{\xi(1 - x_{\mathrm{ES}}^2)}, \tag{17}$$

for $x_{\mathrm{ES}} = m_{\mathrm{ES}}/m_0 < 1$, where m_0 represents the kinematic upper limit and is held fixed at the center-of-mass beam energy $E_b^* = 5.291$ GeV, and ξ and A_B are free parameters.

We assign background and signal probabilities to each event included in the likelihood fit based on the measured value for m_{ES}. However, it is the m_{ES} sideband region, where the background probabilities are essentially 100%, that dominates the determination of the combinatorial background fraction and Δt structure for background events under the B^0 signal peak. Monte Carlo simulation shows a modest m_{ES} dependence on the composition of the combinatorial background over the sideband range $m_{\mathrm{ES}} > 5.2\ \mathrm{GeV}/c^2$ through the B signal region, due to variation of the fraction of continuum versus $B\overline{B}$ contributions. Since these two sources have different Δt behaviors, the changing composition leads to a small correction and systematic error on the precision mixing measurement, but is negligible for the $\sin 2\beta$ extraction. The fraction of peaking backgrounds from charged B decays are estimated with Monte Carlo simulation as described in the following sections.

1. B^0 decays to flavor eigenstates

Candidates in the B_{flav} sample of neutral flavor-eigenstate B mesons are formed by combining a D^{*-} or D^- with a π^+, ρ^+ ($\rho^+ \to \pi^+\pi^0$), a_1^+ ($a_1^+ \to \pi^+\pi^-\pi^+$), or by combining a J/ψ candidate with a K^{*0} ($K^{*0} \to K^+\pi^-$). As described in Sec. II D, kaon identification is used to reject background. For most B^0 modes, it is possible to achieve signal purities of at least 90% with the `VeryLoose` selection, or no particle identification at all. However, for the mode $B^0 \to D^- a_1^+$, the tighter kaon identification is required to reduce large combinatorial backgrounds.

For final states with a D^* and 2 (3) pions we require $|\cos\theta_{\mathrm{th}}| < 0.9$ (0.8) for the $\overline{D}^0 \to K^+\pi^-$ and $K^+\pi^-\pi^0$ modes

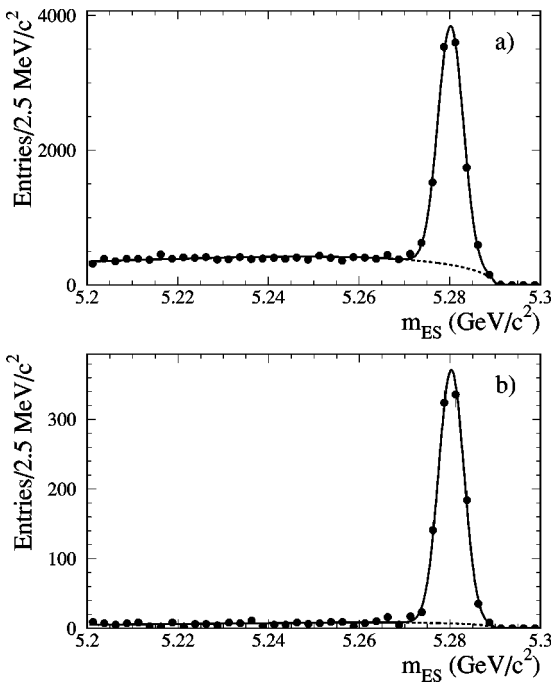

FIG. 3. Distribution of $m_{\rm ES}$ for all selected flavor-eigenstate B^0 candidates in hadronic decays to (a) open charm and (b) charmonium final states. Overlaid in both cases is the result of a fit with a Gaussian distribution for the signal and an ARGUS function for the background.

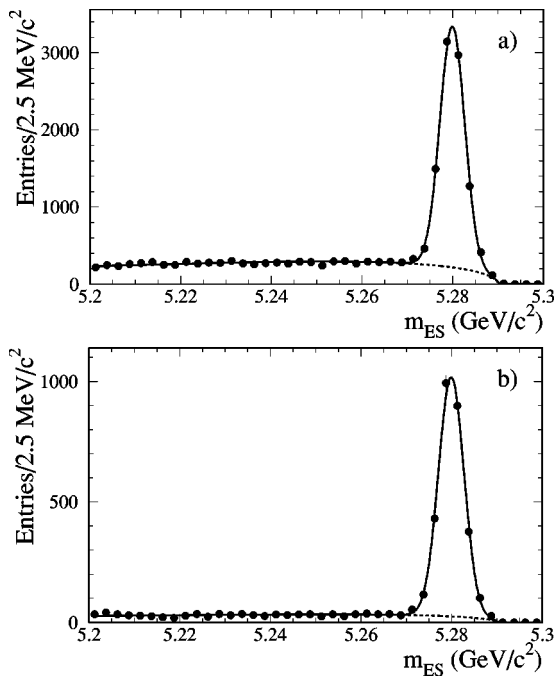

FIG. 4. Distribution of $m_{\rm ES}$ for all selected flavor-eigenstate B^+ candidates in hadronic decays to (a) open charm and (b) charmonium final states. Overlaid in both cases is the result of a fit with a Gaussian distribution for the signal and an ARGUS function for the background.

and 0.8 (0.7) for $\overline{D}^0 \to K^+\pi^+\pi^-\pi^-$ and $K_S^0\pi^+\pi^-$, while no requirement is made for the $B^0 \to D^{*-}\pi^+$ mode. In modes which contain a D^- and a π^+, ρ^+, or a_1^+ in the final state, we require $|\cos\theta_{\rm th}| < 0.9$, 0.8, or 0.7, respectively.

The B^0 signal yield and sample purity extracted from fits to the $m_{\rm ES}$ distribution are summarized in Table III. The net B^0 signal sample, before applying any decay vertex requirements, consists of 9922 ± 129 signal candidates in open charm decays with a purity of about 82%, and 1013 ± 36 in the decay $B^0 \to J/\psi K^{*0}$ ($K^{*0} \to K^+\pi^-$), with a purity of about 95%. Figure 3 shows the combined $m_{\rm ES}$ distribution for all the hadronic B^0 modes. Superimposed is the result of a fit with a Gaussian distribution for the signal and an ARGUS background form [23].

The signal obtained by this method includes a small fraction of peaking background from other charged and neutral B decay modes. However, only the charged B component needs to be determined, since it alone has a time structure that differs from the signal events. Therefore, the fraction of peaking background is estimated with a sample of $\Upsilon(4S) \to B^+B^-$ Monte Carlo events. The B^+ mesons are forced to decay in the decay modes $\overline{D}^{*0}$ or $\overline{D}^0$ with a π^+, ρ^+, or a_1^+, since the main source is decay channels that have one more or one fewer pion in the final state than the signal modes of interest. We then attempt to reconstruct neutral B mesons in the channels used for the $B_{\rm flav}$ sample in data. A small peak at the B^0 mass, obtained with the charged B Monte Carlo sample, leads to an estimate of $(1.3 \pm 0.3^{+0.2}_{-0.5})\%$ as the peaking component in the $B_{\rm flav}$ signal. This result is obtained from a fit with a Gaussian distribution, whose mean and width are fixed by the B^0 signal parameters. The Δt structure of the peaking background in Monte Carlo is found to be consistent with the lifetime of the B^+, as expected.

2. B^+ control samples

The B^+ control sample of charged B candidates is formed by combining a $\overline{D}^{*0}$, $\overline{D}^0$, J/ψ, or $\psi(2S)$ candidate with a π^+ or K^+. For the $\overline{D}^0\pi^+$ final state, we require $|\cos\theta_{\rm th}| < 0.9$ for the $\overline{D}^0 \to K^+\pi^-$ mode and 0.8 for all other $\overline{D}^0$ channels. In modes that contain a $\overline{D}^{*0} \to \overline{D}^0\pi^0$, the requirement is $|\cos\theta_{\rm th}| < 0.9$ for $\overline{D}^0 \to K^+\pi^-$, 0.8 for the $K^+\pi^-\pi^0$ and $K^+\pi^-\pi^-\pi^+$, and 0.7 for $K_S^0\pi^+\pi^-$.

The B^+ signal yield and sample purity extracted from fits to the $m_{\rm ES}$ distribution are summarized in Table III. The net B^+ signal sample in open charm modes, before applying any decay-vertex requirements, consists of 2797 ± 62 signal candidates in charmonium modes, with a purity of about 94%, and 8547 ± 115 signal candidates in open charm modes, with a purity of about 84%. Figure 4 shows the combined $m_{\rm ES}$ distribution for all the hadronic B^+ modes. Superimposed on the data is the result of a fit with a Gaussian distribution for the signal and an ARGUS background form [23].

3. B^0 decays to CP modes involving K_S^0

We form the B_{CP} sample of neutral B candidates in charmonium modes with a K_S^0 by combining mass-constrained J/ψ, $\psi(2S)$ or χ_{c1} candidates with mass-constrained K_S^0 candidates, following the techniques of our recent branching-

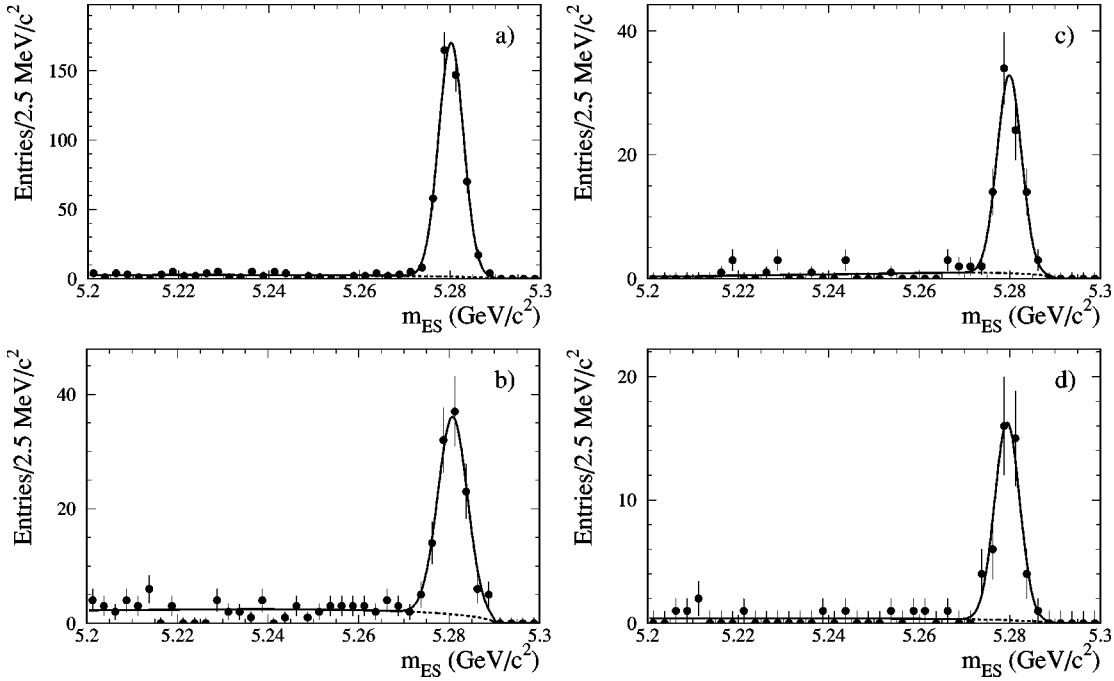

FIG. 5. Candidates for $B^0 \to J/\psi K_S^0$ where K_S^0 decays to (a) $\pi^+\pi^-$ or (b) $\pi^0\pi^0$; candidates for (c) $B^0 \to \psi(2S)K_S^0$ and (d) $B^0 \to \chi_{c1}K_S^0$ ($K_S^0 \to \pi^+\pi^-$). Overlaid in each case is the result of a fit with a Gaussian distribution for the signal and an ARGUS function for the background.

fraction study [24]. The helicity angle θ_h of the J/ψ daughters with respect to the J/ψ flight direction in the B candidate rest frame should have a $\sin^2\theta_h$ distribution. Therefore, we require that $|\cos\theta_h| < 0.8$ for the e^+e^- mode and 0.9 for the $\mu^+\mu^-$ mode, as an efficient way of rejecting backgrounds. For the $\psi(2S)K_S^0$ candidates, $|\cos\theta_h|$ of the $\psi(2S)$ must be smaller than 0.9 for both leptonic modes.

Distributions of m_{ES} are shown in Fig. 5 for the CP samples. Signal event yields and purities, determined from a fit to the m_{ES} distributions after selection on ΔE, are summarized in Table III.

The fraction of peaking background has been estimated with a sample of $B \to J/\psi X$ Monte Carlo events. The main source is decay channels that have one more or one less pion in the final state than the signal mode. The fractions are obtained by fitting the misreconstructed $B \to J/\psi X$ sample with a Gaussian distribution, whose mean and width are fixed by the B^0 signal parameters. The estimated contributions are $(0.41 \pm 0.09)\%$, $(1.2 \pm 0.2)\%$, $(2.9 \pm 1.7)\%$, and $(1.1 \pm 1.1)\%$ for the $J/\psi K_S^0$ ($K_S^0 \to \pi^+\pi^-$), $J/\psi K_S^0$ ($K_S^0 \to \pi^0\pi^0$), $\psi(2S)K_S^0$ and $\chi_{c1}K_S^0$ channels respectively.

In the case of the $\chi_{c1}K_S^0$ mode we have also explored the possibility of contamination from $\chi_{c2}K_S^0$ events. These would have a very similar final-state signature, but opposite CP. However, this decay mode has never been observed and the rate is expected to be highly suppressed due to angular momentum considerations. Figure 6 shows the invariant mass difference, $m(ll\gamma) - m(ll)$, for the χ_{c1} daughters of the $B^+ \to \chi_{c1}K^+$ and $B^0 \to \chi_{c1}K_S^0$ candidates. The distribution is background subtracted with the m_{ES} sideband and a fit with two Crystal Ball distributions [25] is superimposed, where the means have been fixed to the known χ_{c1} and χ_{c2} masses and the widths are forced to be equal. The fraction of $\chi_{c2}K$ events in the selected sample is found to be consistent with zero and, from the fit, an upper limit of 3.5% at 95% C.L. is set on the fraction of $B \to \chi_{c2}K$ candidates in the selected sample.

4. B^0 decays to the CP mode $J/\psi K^{*0}$

The B_{CP} sample is further enlarged by the addition of B^0 candidates in the mode $J/\psi K^{*0}$ ($K^{*0} \to K_S^0\pi^0$). For this purpose, mass-constrained J/ψ candidates are combined with $K^{*0} \to K_S^0\pi^0$ candidates to form a B^0 candidate. To reduce the combinatorial background, the angle between the flight direction of the K_S^0 and the vector connecting the reconstructed vertices of the J/ψ and the K_S^0 candidates is required to be less than 200 mrad. Cross-feed background from other $B \to J/\psi X$ modes involving a π^0 (which includes cross-feed

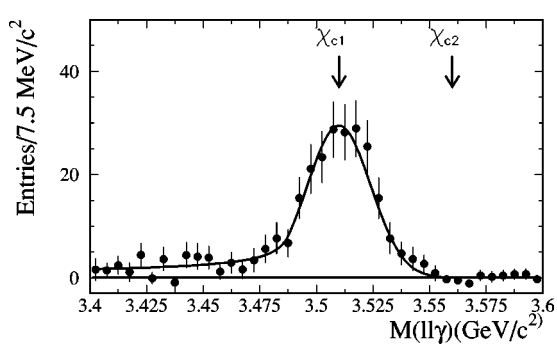

FIG. 6. Distribution of $m(ll_\gamma)$ for the χ_{c1} daughters of $B^+ \to \chi_{c1}K^+$ and $B^0 \to \chi_{c1}K_S^0$ candidates. The expected location of a χ_{c2} signal is indicated by the arrow.

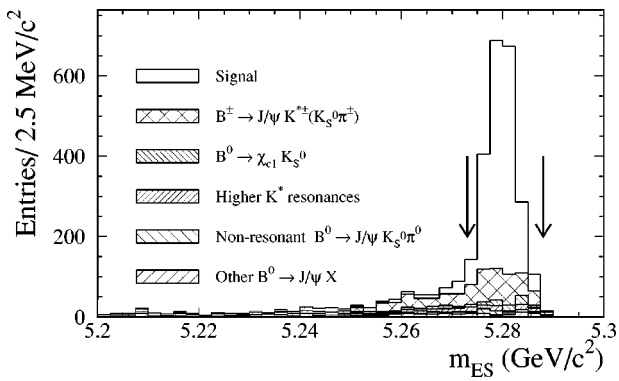

FIG. 7. Distribution of $m_{\rm ES}$ for selected $J/\psi K^{*0}$ combinations, where $K^{*0}\to K_S^0\pi^0$. The arrows indicate the region between 5.273 and 5.288 GeV/c^2 that is used to define the sample of B^0 candidates. Monte Carlo estimates of the various background contributions are also indicated.

from the *CP* mode itself) is suppressed by requiring the cosine of the helicity angle of the K^{*0} in the B^0 meson rest frame to be smaller than 0.95. Further details of the selection and analysis of this sample can be found in Ref. [13].

The $m_{\rm ES}$ distribution for $J/\psi K^{*0}$ ($K^{*0}\to K_S^0\pi^0$) combinations in data is shown in Fig. 7. Given the relatively tight criteria applied in the lepton identification of the daughters of the J/ψ candidates (see Table IV), the background is dominated by true J/ψ mesons from B decays. Its composition can therefore only be estimated with Monte Carlo simulation and the $m_{\rm ES}$ distribution is not expected to follow the phase-space form of Eq. (17). Monte Carlo simulation of events with true J/ψ candidates has been adjusted to match recent results for charmonium branching fractions in B decays and takes into account the indication of S-wave $B^0\to J/\psi K_S^0\pi^0$ decays and contributions due to higher K^* resonances reported in Ref. [13].

As a result, backgrounds are not estimated with a fit to the observed $m_{\rm ES}$ distribution (Fig. 7), but rather by extrapolation of Monte Carlo background distributions, normalized to the number of produced B mesons in the data. All $J/\psi K_S^0\pi^0$ combinations in the range $5.273<m_{\rm ES}<5.288$ GeV/c^2 are considered as candidates for this purpose. Estimates of the signal and background contributions in the candidate sample, and the corresponding effective *CP*, after acceptance correction for the signal selection, is provided in Table V, while the signal yields and purities in data are listed in Table III. The dominant source of cross-feed background, with zero effective *CP*, is $B^+\to J/\psi K^{*+}$ ($K^{*+}\to K_S^0\pi^+$), where the daughter π^+ is exchanged for a background π^0.

TABLE V. Signal and background estimates for the selected $B^0\to J/\psi K^{*0}$ ($K^{*0}\to K_S^0\pi^0$) sample. All the events within the range $5.273<m_{\rm ES}<5.288$ GeV/c^2 are considered as B^0 candidates and the background contributions are estimated with Monte Carlo simulation. The quoted errors are derived from conservative bound on the branching fractions and represent the size of the variation used to estimate the systematic error on $\sin 2\beta$ due to backgrounds.

Event type	Fraction (%)	Effective *CP*
Signal	73.6 ± 7.4	$+0.65\pm0.07$
$B^+\to J/\psi K^{*+}(K_S^0\pi^+)$	17.4 ± 1.7	0
$B^0\to\chi_{c1}K_S^0$	2.4 ± 0.7	-1
Higher K^* resonances	2.6 ± 1.3	0 ± 1
Non-resonant $B^0\to J/\psi K_S^0\pi^0$	1.8 ± 0.9	0 ± 1
Other $B^0\to J/\psi X$	2.4 ± 1.2	0 ± 1
Non $B\to J/\psi X$	0	0

D. B^0 decays to the *CP* mode $J/\psi K_L^0$

Candidates for the B_{CP} sample in the mode $B^0\to J/\psi K_L^0$ are obtained by combining mass-constrained J/ψ and K_L^0 candidates, following the methods in Ref. [24]. The J/ψ candidates are required to have a momentum in the $\Upsilon(4S)$ frame between 1.4 and 2.0 GeV/c. As the K_L^0 energy is not well measured by the EMC or IFR detectors, the laboratory momentum of the K_L^0 is determined by its flight direction as measured from the EMC or IFR cluster and the constraint that the invariant mass of the $J/\psi K_L^0$ system has the known B^0 mass. The production angle θ_B of a B meson with respect to the z axis in the $\Upsilon(4S)$ frame follows a $\sin^2\theta_B$ distribution. We require that $|\cos\theta_B|<0.9$. The J/ψ helicity angle is required to satisfy $|\cos\theta_h|<0.9$ and the sum of $|\cos\theta_B|$ and $|\cos\theta_h|$ must be less than 1.3. Events with a reconstructed charged or neutral B decay to $J/\psi K_S^0$ ($K_S^0\to\pi^+\pi^-$ or $\pi^0\pi^0$), $J/\psi K^{*0}$ ($K^{*0}\to K^+\pi^-$ or $K_S^0\pi^0$), $J/\psi K^+$, or $J/\psi K^{*+}$ ($K^{*+}\to K_S^0\pi^+$ or $K^+\pi^0$) are explicitly removed. The total missing transverse momentum projected along the K_L^0 direction, where the total momentum is calculated with all charged tracks and neutral clusters (without the K_L^0), must be no more than 0.25 (0.40) GeV/c lower than the calculated K_L^0 transverse momentum for EMC (IFR) K_L^0 candidates.

Events where multiple $J/\psi K_L^0$ combinations with $\Delta E<80$ MeV satisfy these requirements are treated as a special case. A hierarchy is imposed where the highest energy EMC cluster for multiple EMC combinations, or the IFR cluster with the largest number of layers for multiple IFR combinations is selected. In cases where there are both an EMC and IFR combination, the EMC combination is selected because it is expected to have better angular resolution. Although the EMC information is used, such events are counted as IFR events, since they have the same relatively high signal purity.

The difference ΔE between the energy of the $J/\psi K_L^0$ system and the beam energy in the $\Upsilon(4S)$ frame is used to discriminate between signal and backgrounds. The ΔE distribution of selected $B^0\to J/\psi K_L^0$ combinations for the $\Upsilon(4S)$ data is shown in Fig. 8. Signal events are peaked within ±10 MeV of $\Delta E=0$ while background events extend towards positive values of ΔE. The small signal width and the asymmetric distribution of the background in comparison with the K_S^0 modes are both consequences of the mass constraint used to determine the K_L^0 momentum.

The purity of the $B^0\to J/\psi K_L^0$ sample is the lowest of the *CP* modes (60%). Irreducible backgrounds are dominantly from $B\to J/\psi K_L^0 X$ modes, which cannot be distinguished from signal due to imposition of the m_B mass constraint in determining the momentum of the K_L^0 candidate. The largest

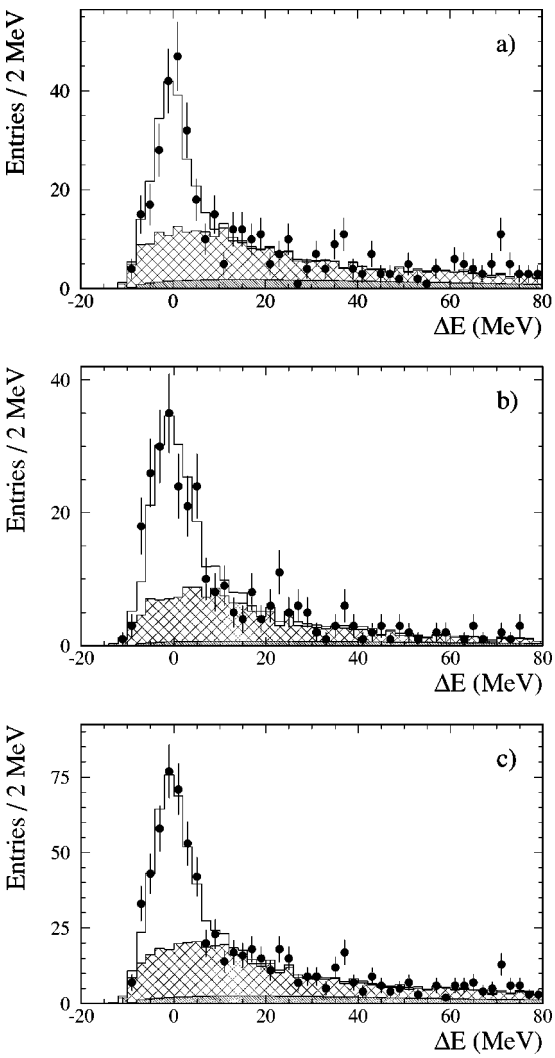

FIG. 8. Distribution of ΔE for selected $B^0 \to J/\psi K_L^0$ combinations where the K_L^0 is identified (a) in the EMC, (b) in the IFR, or (c) either subsample combined. The points with error bars are the data. The open histogram is the result of a three-component binned likelihood fit, where the three components are signal (open), inclusive J/ψ background (cross hatched), and non-J/ψ combinatorial background (dark shading). The shapes of the signal and inclusive J/ψ background are taken from Monte Carlo simulation. The shape of the non-J/ψ combinatorial background is taken from the $m(ll)$ sideband in data. Candidates for the B_{CP} sample are selected in the region $|\Delta E| < 10$ MeV.

single background contribution is from $B \to J/\psi K^*$, where the K^* decays to $K_L^0 \pi$. This mode and backgrounds from other $B \to J/\psi X$ decays are studied with Monte Carlo simulation. The composition of the events that are included in the $J/\psi K_L^0$ sample is given in Table VI. The effective CP, after acceptance correction for the signal selection, is also provided. The additional background from events with a misreconstructed $J/\psi \to ll$ candidate is studied with the $m(ll)$ sidebands.

A binned likelihood fit to the ΔE distribution is performed separately for the EMC and IFR categories to determine the composition of the $B^0 \to J/\psi K_L^0$ sample. There are three fit components: the fraction of the data that is signal, the number of $B \to J/\psi X$ background events, and the number of non-J/ψ background events. The ΔE shapes for the signal and the $J/\psi X$ background are determined from Monte Carlo simulation. The ΔE shape of the non-J/ψ component is taken from the $m(ll)$ sideband in the data. A Poisson term, with mean given by the expected number of non-J/ψ events in the $m(ll)$ signal region, is included in the likelihood to constrain the normalization of the non-J/ψ component. The result of the fit is shown in Fig. 8, and the corresponding signal and background fractions are reported in Table VII for the $B^0 \to J/\psi K_L^0$ combinations that are selected as B_{CP} candidates in the interval $|\Delta E| < 10$ MeV. As expected from Monte Carlo studies, the purity of the IFR sample is significantly better than the EMC sample, mainly because the $B^0 \to J/\psi K_S^0$ ($K_S^0 \to \pi^0 \pi^0$) background is significantly larger in the EMC sample. Since the purities of the two subsamples are quite different, we obtain better statistical sensitivity in the $\sin 2\beta$ fit by treating the EMC and IFR categories separately.

TABLE VI. Monte Carlo prediction for the composition of background channels containing a true J/ψ that pass the $B^0 \to J/\psi K_L^0$ selection criteria. Events are required to have $|\Delta E| < 10$ MeV. The quoted errors are derived from conservative bound on the branching fractions and represent the size of the variation used to estimate the systematic error on $\sin 2\beta$ due to backgrounds.

Event type	EMC (%)	IFR (%)	Effective CP
$B^0 \to J/\psi K^{*0}(K_L^0 \pi^0)$	23 ± 3	26 ± 3	-0.68 ± 0.07
$B^+ \to J/\psi K^{*\pm}(K_L^0 \pi^\pm)$	28 ± 4	45 ± 6	0
$B^0 \to J/\psi K_S^0$	13 ± 2	2 ± 1	-1
$B^0 \to \chi_{c1} K_L^0$	3 ± 1	5 ± 1	$+1$
$B \to J/\psi K_L^0 \pi$	1^{+2}_{-1}	1^{+2}_{-1}	0
Other $B^0 \to J/\psi X$	32 ± 16	21 ± 10	0 ± 0.25

E. Semileptonic B^0 decays

The semileptonic decay $B^0 \to D^{*-} l^+ \nu$, with a measured branching fraction of $(4.60 \pm 0.27)\%$ [11], is potentially a copious source of reconstructed B^0 mesons. However, since the neutrino cannot be detected, the background levels in

TABLE VII. Results of the binned likelihood fit to the full ΔE distribution of the $B^0 \to J/\psi K_L^0$ combinations. All signal yields and background estimates are reported for the B_{CP} candidate range $|\Delta E| < 10$ MeV.

	K_L^0 reconstruction type		
	EMC & IFR	EMC	IFR
Data events	427	228	199
Signal	257 ± 24	128 ± 17	129 ± 17
$J/\psi X$ bkgd	154 ± 15	89 ± 11	65 ± 10
Non-J/ψ bkgd	19 ± 2	14 ± 2	5 ± 1
Signal fraction	0.60 ± 0.04	0.56 ± 0.05	0.65 ± 0.05

selected samples are generally larger and more difficult to characterize. Likewise, the Δz determination cannot take advantage of the beam spot and reconstructed B^0 direction. As a consequence, we use a large sample of $D^{*-}l^+\overline{\nu}$ candidates only as a cross check on our determination of the mistag rates from the time structure of the $B_{\rm flav}$ and B_{CP} events. The selection criteria for this control sample and the characterization of backgrounds is described here, while the analysis of the mistag rates is reported in Sec. VIII C 4. The semileptonic B^0 sample is obtained by reconstructing the D^{*-} through its decay to $\overline{D}^0\pi^-$, where the three decay modes $K^+\pi^-$, $K^+\pi^+\pi^-\pi^-$ and $K^+\pi^-\pi^0$ are used to reconstruct the $\overline{D}^0$.

1. Event selection

All reconstructed $\overline{D}^0$ candidates are required to have an invariant mass within $\pm 2.5\sigma$ of the nominal D^0 mass, based on the observed rms width of the signal. A vertex fit to the $\overline{D}^0$ candidate is required to have a χ^2 probability greater than 0.001. There are no additional requirements for $\overline{D}^0 \to K^+\pi^-$. For $\overline{D}^0 \to K^+\pi^+\pi^-\pi^-$ and $\overline{D}^0 \to K^+\pi^-\pi^0$ we require a `VeryLoose` kaon and a Loose pion particle identification as described in Sec. II D, and a minimum π^0 momentum of 200 MeV/c in the laboratory frame. In addition, the K and π candidates are required to have momenta greater than 200 and 150 MeV/c, respectively, for the mode $\overline{D}^0 \to K^+\pi^+\pi^-\pi^-$. The decay $\overline{D}^0 \to K^+\pi^-\pi^0$ occurs mostly through quasi-two-body channels. The ρ and K^* resonances dominate and we use weights calculated from the Dalitz-plot position for each candidate [26] to construct a probability per D^0 and select events using this quantity as a way of suppressing combinatorial background.

$\overline{D}^0$ candidates satisfying these requirements are combined with all charged tracks with a minimum transverse momentum of 50 MeV/c and charge opposite to that of the kaon from the $\overline{D}^0$ to form D^{*-} candidates. The mass difference $m(\overline{D}^0\pi^-) - m(\overline{D}^0)$ is required to lie within $\pm 2.5\sigma$ of the nominal value, based on the observed rms width of the signal.

Finally, D^{*-} candidates are combined with electron or muon candidates satisfying the `Tight` lepton-identification requirements described in Sec. II D. The lepton is required to have charge opposite to that of the D^* and momentum greater than 1.2 GeV/c in the $\Upsilon(4S)$ frame. A vertex fit to the $D^{*-}l^+$ candidate is required to converge and have a χ^2 probability greater than 0.01. The D^{*-} and lepton from a true B^0 decay tend to be back-to-back in the B^0 rest frame, so we require $\cos\theta_{D^{*-}l} < 0$ where $\theta_{D^{*-}l}$ is the angle between the D^{*-} and the lepton in the $\Upsilon(4S)$ frame. The cosine of the angle between the thrust axes of the $(D^{*-}l^+)$ pair and the rest of the event is required to satisfy $|\cos\theta_{\rm th}| \leq 0.85$, in order to reduce background from $e^+e^- \to c\overline{c}$ events.

The neutrino cannot be reconstructed in the detector, but we can determine whether the missing four-momentum of the candidate is consistent with a particle of zero mass:

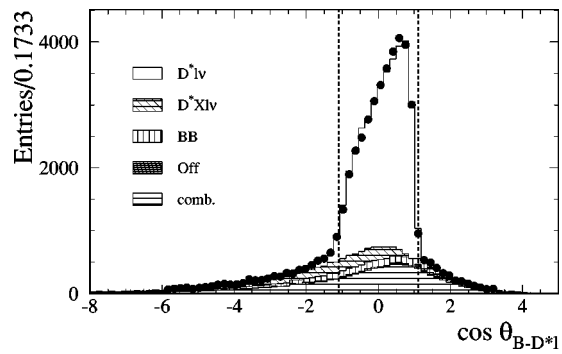

FIG. 9. Extraction of $B\overline{B}$ and D^{**} backgrounds from the data. The points show the $\cos\theta_{B-D^*l}$ distribution for the reconstructed signal. The histogram shows the result of the fit described in the text. The different background contributions are indicated by the various hatchings.

$$(p_B - p_{D^*} - p_l)^2 = p_\nu^2 = 0. \tag{18}$$

Applying this relation in the $\Upsilon(4S)$ frame, we obtain a constraint on the angle between the B^0 and the $D^{*-}l^+$ system:

$$\cos\theta_{B-D^*l} = \frac{2E_B E_{D^*l} - m_b^2 - m_{D^*l}^2}{2|\mathbf{p}_B||\mathbf{p}_{D^*l}|}. \tag{19}$$

The energy E_B and magnitude of the momentum $|\mathbf{p}_B|$ of the initial-state B^0 are known in the $\Upsilon(4S)$ frame from the boosted beam energies. The energy E_{D^*l}, the magnitude of the momentum $|\mathbf{p}_{D^*l}|$, and the invariant mass m_{D^*l} of the $D^{*-}l^+$ system are obtained from the four-momenta of the D^{*-} and the lepton. The cosine of the angle $\cos\theta_{B-D^*l}$ should lie in the physical region $[-1, +1]$ for true $D^{*-}l^+\overline{\nu}$ events. Allowing for detector resolution effects in the reconstructed momenta and angles, we require $|\cos\theta_{B-D^*l}| < 1.1$.

After applying these selection criteria, we obtain a sample of about 37500 $B^0 \to D^{*-}l^+\overline{\nu}$ candidates. The $\cos\theta_{B-D^*l}$ distribution of these candidates is shown in Fig. 9.

2. Sample composition

The final sample contains a fraction $f_{\rm sig}$ of $B^0 \to D^{*-}l^+\overline{\nu}$ signal events, as well as fractions $f_{\rm comb}$ of D^{*-} combinatorial background, $f_{\rm off}$ of continuum background (a true D^{*-} and an identified lepton), $f_{B^+\to D^*Xl\overline{\nu}}$ of $B \to D^*\pi l\overline{\nu}$ events, where the $D^*\pi$ can come from a radially or orbitally-excited $L=1$ state or nonresonant decay, and finally $f_{B\overline{B}}$ of background from $B\overline{B}$ events with a true D^{*-} and an identified lepton ($B\overline{B} \to D^*Yl$). Examples in this last category are cases where the D^{*-} and the lepton come from two different B mesons or where the lepton and the D^{*-} are from the same B but the lepton is produced in a charm decay.

The combinatorial and continuum backgrounds can be extracted directly from the data. The former is determined with the $m(\overline{D}^0\pi^-) - m(\overline{D}^0)$ distribution itself. The latter is estimated with the off-resonance data sample, weighted by the ratio of the relative integrated luminosities for on- and off-resonance data. The remaining three contributions can be distinguished by their different distributions in $\cos\theta_{B-D^*l}$.

The $B^0 \to D^{*-} l^+ \overline{\nu}$ signal events are expected to lie in the region $-1 < \cos\theta_{B-D^*l} < 1$, while contributions from $B \to D^* \pi l \overline{\nu}$ semileptonic decay, due to missing particles, must extend below the kinematic threshold $\cos\theta_{B-D^*l} < -1$. Finally, $B\overline{B}$ background events populate the full $\cos\theta_{B-D^*l}$ distribution. Thus, the region $\cos\theta_{B-D^*l} > 1$ contains mainly $B\overline{B}$ background, while the region $\cos\theta_{B-D^*l} < -1$ is mostly populated by $B \to D^* \pi l \overline{\nu}$. The shape of the $\cos\theta_{B-D^*l}$ distributions for these three components is obtained from Monte Carlo simulation and a fit to the full $\cos\theta_{B-D^*l}$ range is used to determine the two background fractions in the signal region.

The orbitally-excited resonances that can decay to $D^*\pi$ are the two narrow states D_1, D_2^* (observed with masses around 2420 and 2460 MeV/c^2) and the broad state D_1^* (not yet seen, but with mass expected to be about 2420 MeV/c^2 and $\Gamma \geqslant 250$ MeV). Contributions from $B \to D^* X l \overline{\nu}$ decays with more than one pion are expected to be small and are more easily separated from the signal with $\cos\theta_{B-D^*l}$. Isospin symmetry requires that the charged B contribution be 2/3 of the total $D^{*-}\pi$ pairs from $B \to D^* \pi l \overline{\nu}$ decays, either from orbitally-excited states or nonresonant decays. The $\cos\theta_{B-D^*l}$ distribution obtained from Monte Carlo simulation for the different channels is modeled with a general function that is sufficiently flexible to describe both individual channels as well as a superposition of excited charm modes.

After subtraction of continuum and combinatorial backgrounds, a fit is performed to the resulting $\cos\theta_{B-D^*l}$ distribution over the full observed range $[-8, +5]$. However, it is only the relative fraction of the various backgrounds in the signal window $\cos\theta_{B-D^*l} \in [-1.1, 1.1]$ that we require. Furthermore, in the case of D^{**}, only the charged B contribution is a background for the measurement of the mistag fraction and is assumed to be 2/3 of the total D^{**} contribution. The fitted fractions are defined by

$$\begin{aligned} g_{**} &= N(B^+ \to D^* X l \overline{\nu}) / [N(B^0 \to D^{*-} l^+ \overline{\nu}) \\ &\quad + N(B^0 \to D^* X l^+ \overline{\nu}) + N(B^+ \to D^* X l \overline{\nu})] \\ g_{BB} &= N(B\overline{B} \to D^* Y l) / [N(B^0 \to D^{*-} l^+ \overline{\nu}) \\ &\quad + N(B^0 \to D^* X l^+ \overline{\nu}) + N(B\overline{B} \to D^* Y l)] \end{aligned} \tag{20}$$

where the N is the number of events from a given process that survives the selection requirements.

The result of the fit to the full untagged sample is shown in Fig. 9, along with the Monte Carlo model for the D^{**} component. The χ^2 of the fit in the full $\cos\theta_{B-D^*l}$ range is 82 for 69 degrees of freedom. The fitted contributions are $g_{**} = (4.5 \pm 0.3 \pm 2.2)\%$ and $g_{BB} = (4.8 \pm 0.4 \pm 2.2)\%$, where the first error is statistical and the second systematic. To estimate the systematic error on these fractions, three extreme assumptions have been made concerning the $B \to D^* X l \overline{\nu}$ background: all narrow D^{**} states, all broad D^{**} states, or all nonresonant decays. The largest deviation comes from the nonresonant model. Another source of systematic uncertainty is the assumed form for the $\cos\theta_{B-D^*l}$ distributions in the Monte Carlo simulation. The contribution from this effect has been estimated by incorporating a 30% fraction with a uniform distribution.

TABLE VIII. Sample composition in data as determined from fits to the $\cos\theta_{B-D^*l}$ distributions. The fractions have been computed without requiring tagging information. The dominant errors are systematic except for $f_{c\overline{c}}$, which is limited by the statistics.

$f_{\rm comb}$	$f_{\rm off}$	$f_{B\overline{B}}$	$f_{B^+ \to D^* X l \overline{\nu}}$
0.139 ± 0.028	0.008 ± 0.002	0.039 ± 0.018	0.037 ± 0.018

The absolute background fractions in the untagged sample are given in Table VIII, where the uncertainties include both statistical and systematic contributions.

The sum of the fractions of signal and background contributions is constrained to unity. On this basis the signal component is found to be $f_{\rm sig} = (78 \pm 4)\%$ leading to an estimated yield of 29042 ± 1500 $B^0 \to D^{*-} l^+ \overline{\nu}$ signal events.

IV. FLAVOR TAGGING

After the daughter tracks of the reconstructed B are removed, the remaining tracks are analyzed to determine the flavor of the $B_{\rm tag}$, and this ensemble is assigned a tag flavor, either B^0 or $\overline{B}^0$.

We use four different types of flavor tag, or tagging categories, in this analysis. The first two tagging categories rely upon the presence of a prompt lepton, or one or more charged kaons in the event. The next two categories exploit a variety of inputs with a neural-network algorithm. These tagging categories are hierarchical and mutually exclusive.

To quantify the discriminating power of each tagging category, we use as a figure of merit the effective tagging efficiency $Q_i = \epsilon_i (1 - 2w_i)^2$, where ϵ_i is the fraction of events associated to the tagging category i and w_i is the mistag fraction, the probability of incorrectly assigning the tag to an event in this category. The statistical errors in the measurements of $\sin 2\beta$ and Δm_d are inversely proportional to $\sqrt{\Sigma_i Q_i}$.

The mistag fractions are measured with the $B_{\rm flav}$ data sample. The results are shown in Sec. VII. The performance of the tagging algorithm, $\sqrt{\Sigma_i Q_i}$, was optimized and the neural networks were trained with Monte Carlo simulations only. Differences between the tagging inputs in data and in simulation may make the actual tagging algorithm somewhat nonoptimal, but would not lead to a bias because the wrong-tag fractions w_i are measured directly with data, both for the mixing and CP-violation measurements.

A. Lepton and kaon tags

The `Lepton` and `Kaon` tagging categories use the correlation between the flavor of the decaying b quark and the charge of a primary lepton from a semileptonic decay or the charge of a kaon from the chain $b \to c \to s$.

For the `Lepton` category we use both electrons and muons, which are required to pass the `VeryTight` and `Tight` selection, respectively (see Tables I and II). A mini-

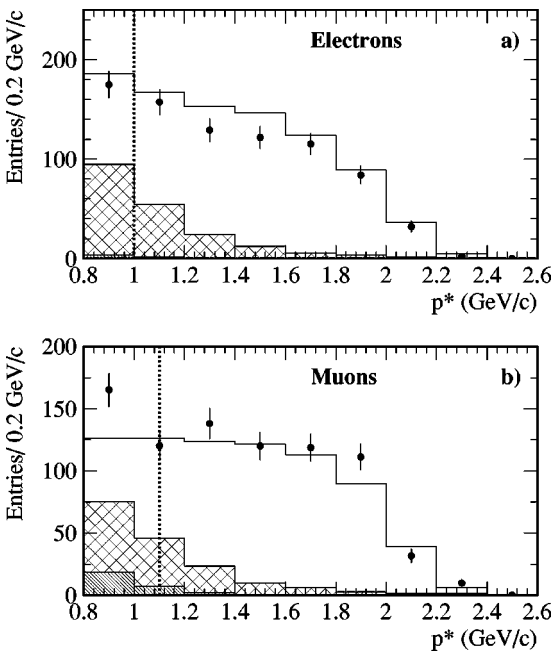

FIG. 10. Center-of-mass momentum distribution for (a) electrons and (b) muons. Data from the $B_{\rm flav}$ sample, after background subtraction based on the $m_{\rm ES}$ sideband, are shown as points. The open histogram shows primary leptons, the cross-hatched histogram cascade leptons, and the diagonally-hatched histogram fake leptons, all from simulation. The simulation is normalized, with a residual overall systematic error of 5%, to the total number of B^0 decays in data after background subtraction, not to the number of observed leptons. The vertical lines at 1.0 GeV/c for electrons and 1.1 GeV/c for muons indicate the requirement on center-of-mass momentum for the lepton category.

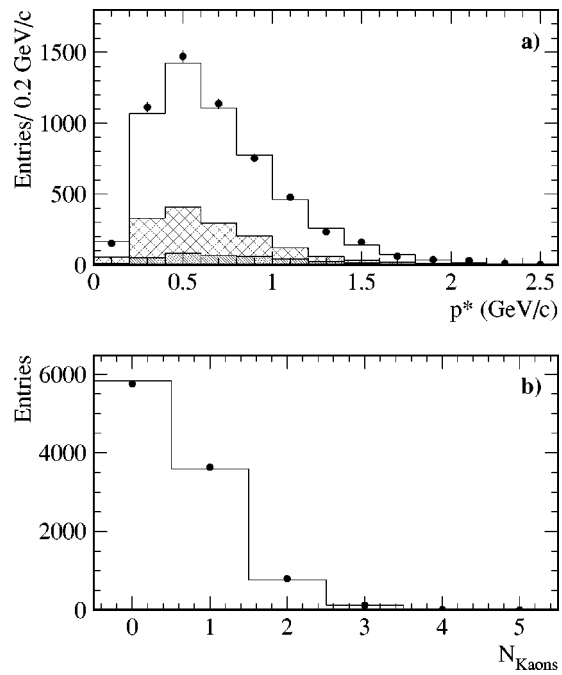

FIG. 11. (a) Center-of-mass momentum distribution for kaons and (b) kaon multiplicity per event. Data from the $B_{\rm flav}$ sample, after background subtraction based on the $m_{\rm ES}$ sideband, are shown as points. The histograms are from simulation. In (a), the diagonally hatched histogram is from fake kaons, the cross-hatched histogram is from kaons that have the wrong-sign charge, and the open histogram is from kaons with the right-sign charge, all from simulation. The simulation is normalized to the total number of B^0 flavor candidates after background subtraction, not to the number of observed kaons.

mum requirement of 1.0 (1.1) GeV/c on the electron (muon) center-of-mass momentum is applied to reduce the contamination from softer, opposite-sign leptons coming from cascade semileptonic decays of charm mesons. The center-of-mass momentum spectra for electrons and muons are compared to simulation in Fig. 10 for the $B_{\rm flav}$ sample, after background subtraction based on the $m_{\rm ES}$ sideband events. In each event, the electron or muon with the greatest center-of-mass momentum is used for flavor tagging; for the tiny fraction of events with both an electron and muon, the electron is used due to its smaller misidentification rate.

The kaon content of the event is evaluated by taking the sum of the charges of all kaons identified with a neural network algorithm (K subnet described below in Sec. IV B). The kaon identification algorithm has been set to maximize the effective tagging efficiency Q. There are 0.8 charged kaons per B decay, and roughly 15% of these have the wrong sign (e.g. K^- from B^0, rather than the expected K^+). Wrong-sign kaons occur primarily in B decays to a charmed-anticharmed pair of mesons. The momentum distributions are quite similar for right- and wrong-sign kaons; we find no kinematic quantity that usefully distinguishes between them. The center-of-mass momentum spectrum for charged kaons and the distribution of charged kaon multiplicity are shown in Fig. 11 for the $B_{\rm flav}$ sample.

An event with an identified high-momentum lepton is assigned to the Lepton category unless the sum of the charges of any kaons present has the opposite sign. Next, events are assigned to the Kaon category if the sum of the kaon charges is nonzero. The charge of the lepton or sum of kaon charges is used to assign the flavor of the $B_{\rm tag}$. All remaining events, approximately 55% of the total including those with inconsistent lepton and kaon charge (0.5% of all events in simulation) and those with two oppositely-charged kaons (4.6% of all events in simulation), are passed to the neural-network-based categories.

B. Neural-network tags

Besides identified high-momentum leptons and charged kaons, there are other features that can be used to determine the flavor of the $B_{\rm tag}$, although not as easily or cleanly distinguishing. These include soft pions from D^* decays, high-momentum primary leptons that are not selected by the electron or muon identification algorithms, lower-momentum primary leptons, and charged kaons that are not selected by the kaon identification algorithm. These sources are combined with a multivariate method; we use a sequence of neural networks to flavor-tag those events not assigned to the `Lepton or Kaon` categories.

Three different track-based neural networks, called ''subnets,'' are trained, each with a specific goal. The `L`, `K`, and

SoftPi subnets are sensitive to the presence of primary leptons, charged kaons and soft pions from D^* decays, respectively. Each subnet is applied to all tracks from the $B_{\rm tag}$.

The L subnet uses the binary output of the electron and muon identification algorithms on the input track, the center-of-mass momentum of the input track, and a pair of kinematic variables, E_W^{90} and $\cos\theta_{l\nu}$, that separate primary leptons from cascade leptons and other tracks.

The isolation variable, E_W^{90}, is given by the sum of the energies of all tracks within 90° of the W direction. The W momentum is inferred as the sum of the input track momentum and the neutrino momentum, which we take to be the missing momentum in the center-of-mass frame using all charged tracks in the $B_{\rm tag}$. This variable is effective because in a semileptonic decay the hadrons recoiling against the virtual W would generally go off in the opposite direction.

The other kinematic variable used, $\cos\theta_{l\nu}$, is the cosine of the angle between the input track and the neutrino direction. The distributions in the $B_{\rm flav}$ sample and simulation of E_W^{90} and $\cos\theta_{l\nu}$ are shown in Figs. 12a and 12b, for all events not in the Lepton or Kaon category. The corresponding distribution of the L subnet output is shown in Fig. 13a.

The K subnet uses the input track momentum in the laboratory frame, together with the three relative likelihoods $\mathcal{L}_K/(\mathcal{L}_\pi+\mathcal{L}_K)$ for the SVT, the DCH and the DIRC. The SVT and DCH likelihoods are derived from dE/dx measurements, and the DIRC likelihood is calculated from a global fit to the number of photons detected, their positions and arrival times relative to the corresponding track. The distribution of the K subnet output, again for events not in the Lepton or Kaon category, is shown in Fig. 13b.

The SoftPi subnet uses the center-of-mass momentum of the input track, the cosine of the angle of the input track with the thrust axis $\cos\theta_{\rm th}$, and the center-of-mass momentum of the minimum momentum track. The thrust axis is determined from all charged tracks and neutral clusters in the $B_{\rm tag}$. The direction of any D^* in the decay of the $B_{\rm tag}$ is approximated by the direction of the thrust axis. Thus soft pions from D^* decays, which are aligned with the D^* direction in the center-of-mass frame, tend to be correlated with the thrust axis. The distribution of $\cos\theta_{\rm th}$ is shown for low center-of-mass momentum tracks in Fig. 12c, comparing the $B_{\rm flav}$ sample with simulation for all events not in the Lepton or Kaon category. The corresponding distribution of the SoftPi subnet output is shown in Fig. 13c.

The outputs of the three subnets are among the inputs to a final neural network, which is trained to distinguish between B^0 and $\overline{B}^0$. The variables used as inputs to the final network include the maximal values of the L and SoftPi subnet outputs, each multiplied by the charge of the corresponding input track, and the highest and the second-highest values of the K subnet output again multiplied by the charge of the corresponding input tracks. The two other inputs to the final neural network are the center-of-mass momentum of the maximum momentum track multiplied by its charge, and the number of tracks with significant impact parameters. The latter is an indicator of the presence of K_S^0 mesons. The distribution of the center-of-mass momentum for all tracks is shown in Fig. 12d.

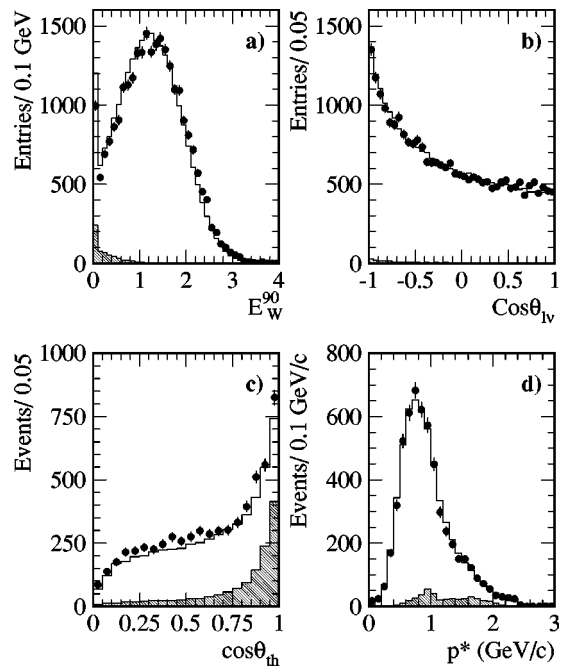

FIG. 12. Inputs to the subnets: (a) E_W^{90}, (b) $\cos\theta_{l\nu}$, (c) $\cos\theta_{\rm th}$ for low center-of-mass momentum tracks ($p^*<0.18$ GeV/c), and (d) the center-of-mass momentum for all tracks. The points are data from the $B_{\rm flav}$ sample after background subtraction based on the $m_{\rm ES}$ sideband, and the histogram is simulation. For $\cos\theta_{\rm th}$ the diagonally hatched histogram shows the contribution from soft π coming from D^* decays, and for the other distributions shows the component from primary leptons. The simulation is normalized to the total number of B^0 flavor candidates after the background subtraction.

The output from the final neural network, x_{NT}, is mapped onto the interval $[-1, 1]$. The assigned flavor tag is B^0 if x_{NT} is negative, and $\overline{B}^0$ otherwise. Events with $|x_{NT}|>0.5$ are assigned to the NT1 tagging category and events with $0.2<|x_{NT}|<0.5$ to the NT2 tagging category. Events with $|x_{NT}|<0.2$, approximately 30% of the total, have very little tagging power and are rejected. The distribution of x_{NT} for all events not assigned to the Lepton or Kaon category is shown in Fig. 14a.

Most of the separation between B^0 and $\overline{B}^0$ in the NT1 and NT2 tagging categories derives from primary leptons that are not identified as electrons or muons and from soft pions from D^* decays. Simulation studies indicate that roughly 37% of the effective tagging efficiency Q is due to events with unidentified primary leptons, 28% is due to events with a soft pion, a further 11% from events with a lower momentum primary lepton, and the remainder from a mixture of the various inputs. This classification is shown for the NT1 and NT2 categories in Fig. 14b for a simulation of B^0 decays. The modest disagreements between Monte Carlo simulation and data that are evident in the distributions of the subnet output variables shown in Fig. 13 lead to a difference in the predicted value of $Q=(3.0\pm0.1)\%$ and $(1.4\pm0.1)\%$ for NT1 and NT2 categories in simulation versus $(2.5\pm0.4)\%$ and $(1.2\pm0.3)\%$ as measured in data.

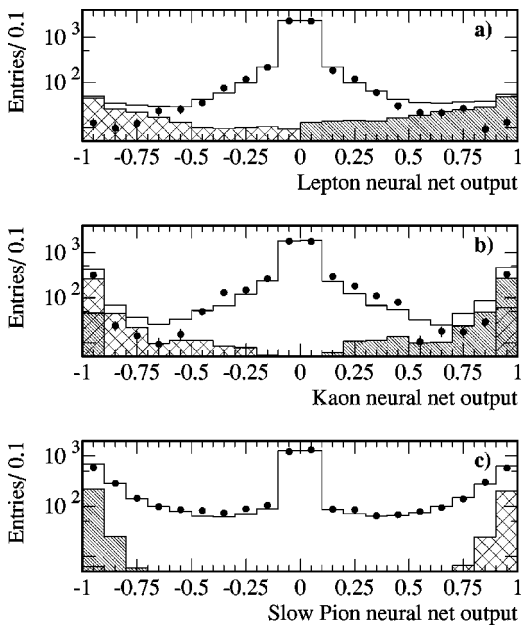

FIG. 13. Output of the subnets for events not assigned to the `Lepton` or `Kaon` categories: (a) `L` subnet, (b) `K` subnet, and (c) `SoftPi` subnet. The points are data from the $B_{\rm flav}$ sample after background subtraction based on the $m_{\rm ES}$ sideband, and the histogram is simulation. For each distribution, the filled portion of the histogram shows the component with a B^0 (singly hatched) or $\overline{B}^0$ (cross-hatched) tag from the full neural network algorithm that arises from true primary leptons, kaons, or soft pions respectively. Note that the latter has the opposite charge correlation with the B^0 tag. The simulation is normalized to the total number of B^0 flavor candidates after background subtraction.

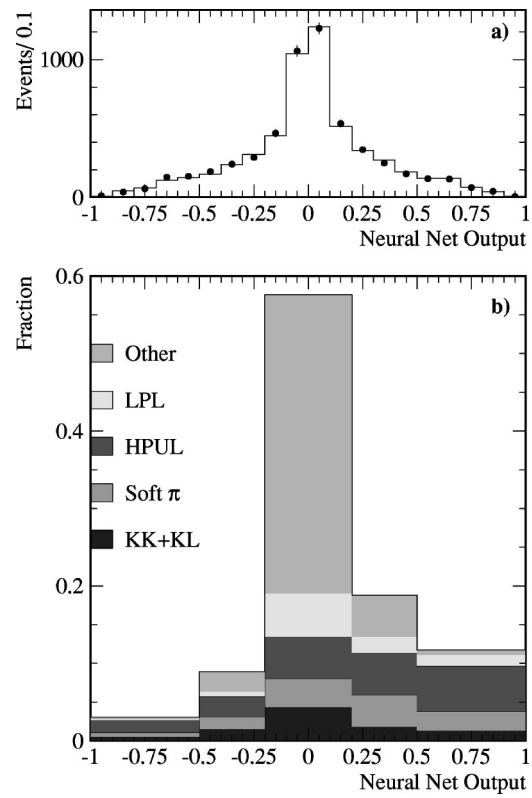

FIG. 14. (a) Output of the final neural network for $B_{\rm flav}$ events that are not assigned to the `Lepton` or `Kaon` category, where the points are from the $B_{\rm flav}$ data after a background subtraction and the histogram is simulation; (b) neural network output from simulation of single B^0 decays with no time evolution, again for events not in the Lepton or Kaon category. The breakdown from bottom to top is events with two kaons or a kaon and lepton (KK+KL), events with a soft pion (soft π), events with a high momentum unidentified lepton (HPUL), events with a lower momentum lepton (LPL), and all remaining events. The outermost bins correspond to the category `NT1` and the next to `NT2`. Entries for $x_{NT}>0.0$ represent correct tags, while those for $x_{NT}<0$ are mistags in each of the categories. The center bin contains events for which no tag is assigned.

V. TIME DIFFERENCE MEASUREMENT

The difference between B decay times, $\Delta t = t_{\rm rec} - t_{\rm tag}$, is determined from the measured separation Δz between the vertex of the reconstructed B meson ($B_{\rm rec}$) and the vertex of the flavor-tagging B meson ($B_{\rm tag}$) along the z axis. The Δz resolution is dominated by the z position resolution for the $B_{\rm tag}$ vertex.

A. Δz reconstruction

In the reconstruction of the $B_{\rm rec}$ vertex, we use all charged daughter tracks. Daughter tracks from K_S^0 and D candidates are first fit to a separate vertex and the resulting parent momentum and position are used in the fit to the $B_{\rm rec}$ vertex. Mass constraints, which include neutral daughters, are used for D candidates but not for D^{*-}, J/ψ and $\psi(2S)$ candidates. The rms resolution in z for the $B_{\rm rec}$ vertex in Monte Carlo simulation is about 65 μm for more than 99% of the B candidates, and 40 μm for about 80% of the candidates. As described in Sec. V E, the resolution is about 5% worse in data than in Monte Carlo simulation.

The vertex for the $B_{\rm tag}$ decay is constructed from all tracks in the event except the daughters of $B_{\rm rec}$. For fully reconstructed modes, an additional constraint is provided by the calculated $B_{\rm tag}$ production point and three-momentum, with its associated error matrix. This is determined from the knowledge of the three momentum of the fully reconstructed $B_{\rm rec}$ candidate, its decay vertex and error matrix, and from the knowledge of the average position of the interaction point and the $\Upsilon(4S)$ average boost (see Fig. 15). The average beam spot position and the angle of the beam in the detector are updated once per run, while the beam energies are recorded more frequently for any change above 0.05 MeV. These $B_{\rm tag}$ parameters are used as input to a geometrical fit to a single vertex, including all other tracks in the event except those used to reconstruct $B_{\rm rec}$. In order to reduce bias and tails due to long-lived particles, K_S^0 and Λ^0 candidates are used as input to the fit in place of their daughters. In addition, tracks consistent with photon conversions ($\gamma \to e^+e^-$) are excluded. To reduce contributions from charm decay products, which bias the determination of the vertex position, the track with the largest vertex χ^2 contribution greater than 6 is removed and the fit is redone until no track fails the χ^2 requirement. In Monte Carlo simulation, the rms of the core and tail Gaussian components of the

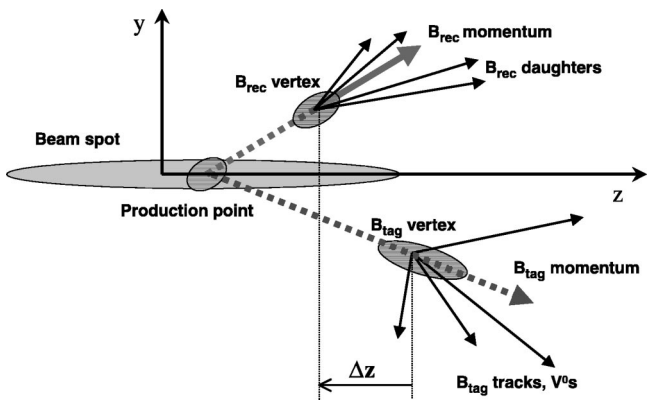

FIG. 15. Schematic view of the geometry in the *yz* plane for a $\Upsilon(4S)\rightarrow B\overline{B}$ decay. For fully reconstructed decay modes, the line of flight of the B_{tag} can be estimated from the (reverse) momentum vector and the vertex position of B_{rec}, and from the beam spot position in the *xy* plane and the $\Upsilon(4S)$ average boost. Note that the scale in the *y* direction is substantially magnified compared to that in the *z* direction.

residual Δz distribution (measured Δz minus true Δz) is 190 μm. We fit this residual distribution to the sum of three Gaussian distributions and find that the rms of the narrowest Gaussian, which contains 70% of the area, is about 100 μm. Only 1% of the area is in the widest Gaussian.

The absolute scale of the measurement of Δz depends on the assumed positions of the silicon wafers in the SVT. These positions are determined from a combination of survey measurements made before the SVT was installed and measurements of the positions of individual SVT modules, each containing several silicon wafers, made with high-momentum charged particles that traverse the SVT. We check the absolute scale comparing the known positions of distinct mechanical features at each end of the beampipe (about 18 cm apart) with the apparent position measured with charged tracks in the SVT. The locations of these mechanical features are measured from the positions of track vertices at least 2 cm from the interaction point that contain a well-identified proton, which are mainly due to $e^{\pm}$-nucleon interactions in material. The measured distance in *z* between these mechanical features is in agreement with the known distance to a precision of 0.2%. We conservatively enlarge this to 0.4% to account for any additional uncertainty in extrapolating to the interaction point.

B. Δt measurement

By far the dominant limitation on the accuracy with which Δt is determined from the measured decay length difference Δz is the experimental resolution on the Δz measurement. The next most significant limitation is the *B* meson momentum of about 340 MeV/c in the $\Upsilon(4S)$ rest frame. We partially correct for this effect, as described below. The impact on the Δt measurement of the spread in the two beam energies, which results in a distribution of $\Upsilon(4S)$ momenta with a Gaussian width of about 6 MeV/c, is negligible. Finally, we correct for the 20 mrad angle between the $\Upsilon(4S)$ boost direction (the *z* axis in the following discussion) and the axis of symmetry of the detector, along which we measure the separation between vertices.

Neglecting the *B* momentum in the $\Upsilon(4S)$ frame, we can write

$$\Delta z = \beta\gamma c \Delta t, \tag{21}$$

where $\beta\gamma$ is the $\Upsilon(4S)$ boost factor. The average value for the boost factor is $\beta\gamma = 0.55$. The boost factor is calculated directly from the beam energies, which are monitored every 5 seconds, and has an uncertainty of 0.1%.

In the case of a fully reconstructed B_{rec}, we measure with good precision the momentum direction of the reconstructed candidate, which can be used to correct for the *B* momentum in the $\Upsilon(4S)$ frame. However, the correction depends on the sum of the decay times, $t_{rec}+t_{tag}$, which can only be determined with very poor resolution. We use the estimate $t_{rec}+t_{tag}=\tau_B+|\Delta t|$ to correct for the measured B_{rec} momentum direction and extract Δt from the following expression:

$$\Delta z = \beta\gamma\gamma^*_{rec} c\Delta t + \gamma\beta^*_{rec}\gamma^*_{rec}\cos\theta^*_{rec} c(\tau_B+|\Delta t|), \tag{22}$$

where θ^*_{rec}, β^*_{rec}, and γ^*_{rec} are the polar angle with respect to the beam direction, the velocity, and the boost factor of the B_{rec} in the $\Upsilon(4S)$ frame. The difference between Δt calculated with Eq. (21) and Eq. (22) is very small because $\gamma^*_{rec}=1.002$ and $\beta^*_{rec}=0.064$. The event-by-event difference in Δt calculated with the two methods has an rms of 0.20 ps. Equation (22) improves the Δt resolution by about 5%. In addition, it removes a correlation between the resolution on Δt and the true value of Δt. This correlation is due to the fact that the rms of the second term in Eq. (22) depends on the expectation value of $(t_{rec}+t_{tag})^2$, which in turn depends on $|\Delta t|$. Equation (22) is used for all *B* decays to hadronic final states, while Eq. (21) is used for semileptonic modes since the *B* direction cannot be measured for these decays.

C. Vertex quality requirements

A number of requirements are made in order to ensure a well-determined vertex separation. The fit for both the B_{rec} and B_{tag} vertex must converge. Also, the error on Δt determined from the vertex fit must be less than 2.4 ps and $|\Delta t|$ must be less than 20 ps. The efficiency for passing these requirements in data and Monte Carlo simulation is about 97% for all B_{rec} modes. From the Monte Carlo simulation, we find that the reconstruction efficiency does not depend on the true value of Δt.

The B_{rec} sample is used both to extract the B^0-$\overline{B}^0$ mixing frequency and to measure the mistag probabilities for the analysis of time-dependent *CP*-violating asymmetries. While the *CP* measurement is statistically limited, the mixing measurement has a statistical precision of a few percent. Therefore, in order to reduce systematic uncertainties in the mixing measurement, more restrictive vertex criteria are imposed for the B_{rec} sample used for the mixing measurement than for the *CP* and B_{rec} samples used for the *CP* measurement. However, as described in Sec. VIII C 5, the more restrictive criteria are also used as a cross-check in the *CP* measurement. In

order to reduce further the contributions from charm decay products in the mixing analysis, a track is not used in the reconstruction of the $B_{\rm tag}$ vertex if it is identified as a kaon according to the kaon identification algorithm used for tagging (see Sec. IV A). The maximum allowed error on Δt determined from the vertex fit is decreased from 2.4 ps for the samples used for the *CP* measurement to 1.4 ps for the sample used for the mixing measurement. The efficiency to pass these two additional criteria is about 87% in data. All figures in this section are obtained with the vertex selection criteria applied in the *CP* analysis.

D. Δt resolution function

The Δt resolution function is represented in terms of $\delta_{\rm t} = \Delta t - \Delta t_{\rm true}$ by a sum of three Gaussian distributions (called the core, tail and outlier components) with different means and widths:

$$\mathcal{R}(\delta_{\rm t};\hat{a}) = \sum_{k=1}^{2} \frac{f_k}{S_k \sigma_{\Delta t}\sqrt{2\pi}} \exp\left(-\frac{(\delta_{\rm t} - b_k \sigma_{\Delta t})^2}{2(S_k \sigma_{\Delta t})^2}\right) + \frac{f_3}{\sigma_3\sqrt{2\pi}} \exp\left(-\frac{\delta_{\rm t}^2}{2\sigma_3^2}\right). \tag{23}$$

For the core and tail Gaussians, we use the measurement error $\sigma_{\Delta t}$ derived from the vertex fit for each event but allow two separate scale factors S_1 and S_2 to accommodate an overall underestimate ($S_k > 1$) or overestimate ($S_k < 1$) of the errors for all events. Figure 16a illustrates the correlation between the rms of δ_t and $\sigma_{\Delta t}$ in Monte Carlo simulation. The core and tail Gaussian distributions are allowed to have a nonzero mean to account for residual charm decay products included in the $B_{\rm tag}$ vertex. In the resolution function, these mean offsets are scaled by the event-by-event measurement error $\sigma_{\Delta t}$ to account for an observed correlation shown in Fig. 16b between the mean of the $\delta_{\rm t}$ distribution and the measurement error $\sigma_{\Delta t}$ in Monte Carlo simulation. This correlation 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 inclusion of *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.

Monte Carlo simulations confirm the expectation that the resolution function is less biased for events with a primary lepton tag than those with a kaon tag. Therefore, the mean of the core Gaussian is allowed to be different for each tagging category. One common mean is used for the tail component.

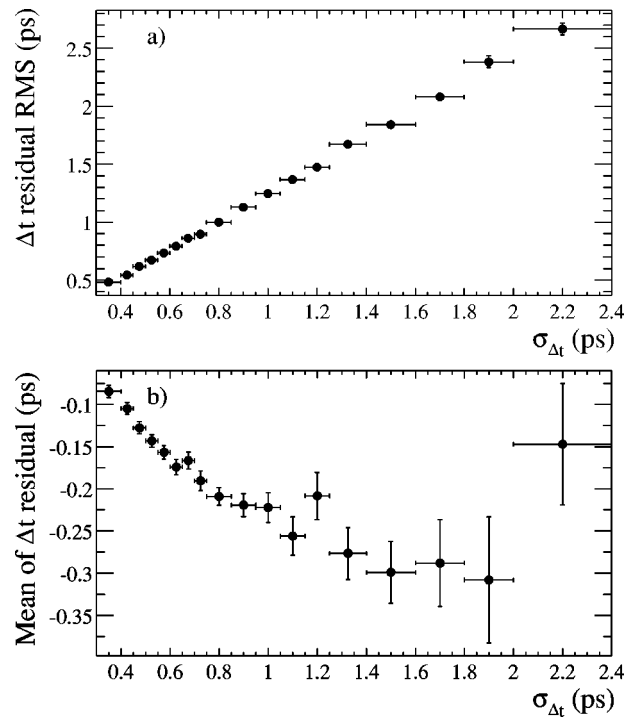

FIG. 16. Correlation between the event-by-event error on Δt ($\sigma_{\Delta t}$) and (a) the observed rms and (b) offset of the mean for $\delta_{\rm t} = \Delta t - \Delta t_{\rm true}$ from Monte Carlo simulation.

The third Gaussian has a fixed width of 8 ps and no offset; it accounts for the fewer than 1% of events with incorrectly reconstructed vertices. The resolution parameters extracted from the full likelihood fits to the Δt distributions are shown in Table XIII for the mixing analysis and in Table XVI for the *CP* analysis.

Since the $B_{\rm tag}$ vertex precision dominates the Δt resolution, no significant differences between the Δt resolution function for the flavor-eigenstate sample and the *CP*-eigenstate sample are expected. Hence, identical resolution functions are used for all modes. This assumption is supported by Monte Carlo simulation and addressed in the evaluation of systematic errors. Figure 17 shows the distribution of the uncertainties on Δt ($\sigma_{\Delta t}$) from the fit to Δz for the flavor-eigenstate sample, and the combined $\eta_{CP} = -1[J/\psi K_S^0, \psi(2S)K_S^0, \chi_{c1}K_S^0]$ and $J/\psi K^{*0}$ samples. Also shown is a comparison between data and Monte Carlo prediction for the two samples, since Monte Carlo simulation is used to evaluate the systematic error contribution. The flavor-eigenstate and *CP*-eigenstate distributions need not be the same under our assumption of a common resolution function, since the topologies and multiplicities of the $B_{\rm rec}$ decays are different. The agreement between the data and Monte Carlo simulation is satisfactory, particularly given the statistical uncertainties for the *CP* sample.

E. Checks and control samples

Two of the fundamental assumptions in this analysis are that the Δt resolution function for the sample of flavor-eigenstate modes is the same as that for *CP* events, and that

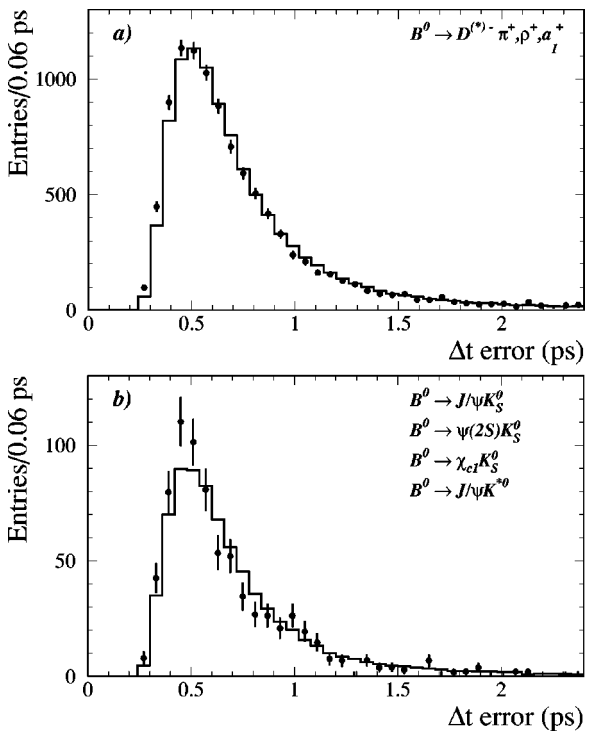

FIG. 17. Distribution of event-by-event uncertainties on Δt ($\sigma_{\Delta t}$) for (a) the sample of neutral B decays to flavor eigenstates other than $J/\psi K^{*0}(K^+\pi^-)$ and (b) the combined $\eta_{CP}=-1$ $[J/\psi K_S^0, \psi(2S)K_S^0, \chi_{c1}K_S^0]$ and $J/\psi K^{*0}$ samples. The histogram corresponds to Monte Carlo simulation and the points with error bars to data. All distributions have been background-subtracted with events from the $m_{\rm ES}$ sideband. The Monte Carlo distribution has been normalized to the same area as the data distribution.

the event-by-event vertex errors provide a good measure of the relative uncertainty on the Δz measurement from event to event. In this section, we describe several studies that have been done to validate these assumptions. We compare various distributions for the CP and flavor-eigenstate samples, in both data and Monte Carlo simulation. We take advantage of the small vertical size of the beam to measure the resolution for the $B_{\rm rec}$ and the $B_{\rm tag}$ vertices in the vertical direction. We also study the vertex resolution for a sample of D^{*+} candidates from $c\bar{c}$ events and for continuum events in data and Monte Carlo simulation.

1. Comparison of flavor-eigenstate and CP samples

In Fig. 18, we compare various properties of the flavor-eigenstate sample with the combined $\eta_{CP}=-1$ and $J/\psi K^{*0}$ samples. These include the χ^2 probability for the vertex fits, the number of tracks used in the $B_{\rm tag}$ vertex, and the momentum in the $\Upsilon(4S)$ rest frame and polar angle in the laboratory frame of tracks used in the $B_{\rm tag}$ vertex. Good agreement in all variables is observed between the two data samples.

A similar comparison of the momentum and polar-angle distribution of tracks in data and Monte Carlo simulation also shows good agreement. However, there are modest discrepancies for the χ^2 probability for the vertex fits and the number of tracks used in the $B_{\rm tag}$ vertex. The agreement improves when we include residual misalignments between the SVT silicon modules in the Monte Carlo simulation. Systematic uncertainties due to residual SVT misalignments in data are discussed in Secs. VII B 1 and VIII B.

As expected, there are no significant differences observed in comparisons between the CP modes used in the $\sin 2\beta$ analysis. However, comparisons between the CP and flavor-eigenstate samples, in data as well as in the Monte Carlo simulation, show that the CP events have a slightly better Δz resolution. For example, in Monte Carlo simulation the most probable value for $\sigma_{\Delta t}$ is about 0.017 ps (3%) worse for the $B_{\rm flav}$ sample, as can be seen by comparing the distributions in Figs. 17a and 17b. This is due to the fact that the B_{CP} vertex is better determined because tracks in the lower-multiplicity CP final states generally have higher momentum. We account for this effect in the likelihood fit by using the calculated event-by-event errors, as described in Sec. V D. Indeed, for Monte Carlo simulation, the pull distributions for Δt (defined as the difference between the fitted and generated value divided by the calculated error) are nearly Gaussian with unit width for both the B_{CP} and $B_{\rm flav}$ samples. Any residual effect due to differences in the observed scale factors in data is included as a systematic uncertainty (see Sec. VIII B) and found to be negligible.

2. Vertex resolution in vertical direction

Since the size of the PEP-II beam is only about 10 μm in the vertical (y) direction, the measured distance Δy between the $B_{\rm rec}$ or $B_{\rm tag}$ vertex and the nominal beam spot position in the y direction can be used to compare the resolution for the CP and flavor-eigenstate samples, and to evaluate the accuracy of the event-by-event errors $\sigma(\Delta y)$. The average beam-spot y position is determined with a precision of better than a few microns with two-track events for each data run (approximately one hour of recorded data). There is a non-negligible contribution to Δy of $\approx$25 μm (rms) due to the B lifetime and the transverse momentum of the B.

The distance in y between the $B_{\rm tag}$ vertex and the beam spot is used to measure the $B_{\rm tag}$ vertex resolution and bias in y. In Fig. 19, we show the distribution of $\Delta y/\sigma(\Delta y)$ for the $B_{\rm tag}$ vertex for the flavor-eigenstate and CP samples, in data and Monte Carlo simulation. The rms of the $\Delta y/\sigma(\Delta y)$ distribution is 1.3 and 1.4 for Monte Carlo simulation and data, respectively. No statistically significant biases are observed.

Similar results are obtained for the $B_{\rm rec}$ vertex resolution. In addition, good agreement in the resolution on the y position is observed between the flavor-eigenstate sample and the CP sample. The resolution is about 5% worse in data than in Monte Carlo simulation.

3. Vertex resolution in continuum events

Two samples have been used to cross-check the reliability of the resolution function extracted from the likelihood fit as well as the discrepancies between data and Monte Carlo simulation: a sample of 109000 D^{*+} candidates from $c\bar{c}$ events and a sample of off-resonance data.

For the first sample, we reconstruct high-momentum D^{*+} candidates in the mode $D^{*+}\to D^0\pi^+$, followed by $D^0 \to K^-\pi^+$, $K^-\pi^+\pi^0$, or $K^-\pi^+\pi^-\pi^+$, and then use the re-

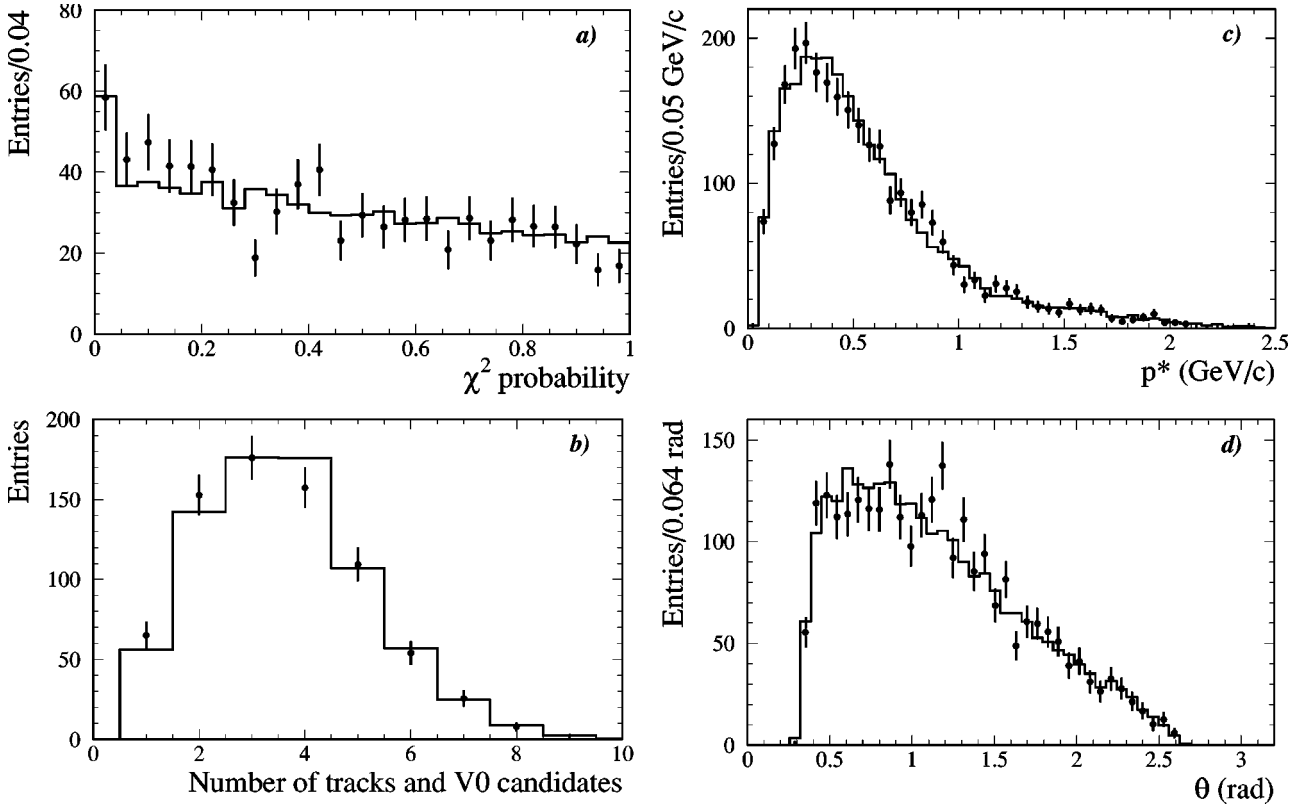

FIG. 18. Distributions of (a) χ^2 probability of the $B_{\rm tag}$ vertex fit, (b) number of charged tracks and V^0 candidates used in the $B_{\rm tag}$ vertex, (c) momentum in the center-of-mass frame, and (d) polar angle in the laboratory frame for tracks in the $B_{\rm tag}$ vertex, for the flavor-eigenstate (histograms) and the combined $\eta_{CP}=-1$ and $J/\psi K^{*0}$ (points with errors bars) data samples. All distributions have been background-subtracted with events from the $m_{\rm ES}$ sideband. The flavor-eigenstate distributions have been normalized to the same area as the distributions from the combined $\eta_{CP}=-1$ and $J/\psi K^{*0}$ samples.

mainder of the charged tracks in the event (fragmentation particles and recoil charm decay products) to determine a vertex position with the standard $B_{\rm tag}$ vertex algorithm. Since position information for the D^{*+} vertex is poor, due to scattering of the slow pion, and the D^{*+} decay point coincides with the e^+e^- interaction point, a beam-spot constraint is used for the D^{*+}. In Fig. 20, we show the distribution of the distance along the z axis between the D^{*+} vertex and the vertex formed from the rest of the tracks in the event Δz, as well as Δz divided by the event-by-event error on Δz, for both data and Monte Carlo simulation.

In Monte Carlo simulation, the resolution on z for the D^{*+} candidate is $\approx$90 μm, very similar to that for $B_{\rm rec}$ vertices. However, the momentum spectrum of fragmentation tracks in $c\bar{c}$ events is softer than that for tracks from B decays, while D mesons are more energetic in the D^{*+} control sample than in B decays. Therefore, a slightly more asymmetric resolution function is expected for the D^{*+} control sample compared to that for B events, as shown in Fig. 20a. Comparison of distributions of several sensitive variables (such as the number of tracks used in the vertex, and the momentum and polar angle of the tracks) shows small differences between D^{*+} and B events.

The rms of the distance between the D^{*+} vertex and the vertex formed from the rest of the tracks in the events is about 220 μm in the Monte Carlo simulation. Fitting this distribution to the sum of three Gaussians, we find a resolution of about 140 μm for 97% of the events, compared to 150 μm for 99% of the $B^0\overline{B}^0$ events. Only small differences are observed in the distribution of $\Delta z/\sigma(\Delta z)$, as illustrated in Fig. 20b. Therefore, the sample can be used to confirm the resolution and scale factors extracted from the likelihood fit, as well as to compare how well the Monte Carlo simulation reproduces the data.

The distributions are fit to the sum of three Gaussian distributions with different widths and means. The width of the third Gaussian is fixed to 2.0 mm. From the fit results, we come to the following conclusions:

(1) The event-by-event errors on Δz are underestimated by about 10% in data (Fig. 20b).
(2) The bias in the resolution function due to charm decay products that is observed in data is well reproduced by the Monte Carlo simulation as shown in Figs. 20a and 20b.
(3) The resolution measured in the data is about 5% worse than that predicted by the Monte Carlo simulation (Fig. 20a).

These results will be compared with those obtained from the likelihood fit to the B events, described in Sec. VIII A.

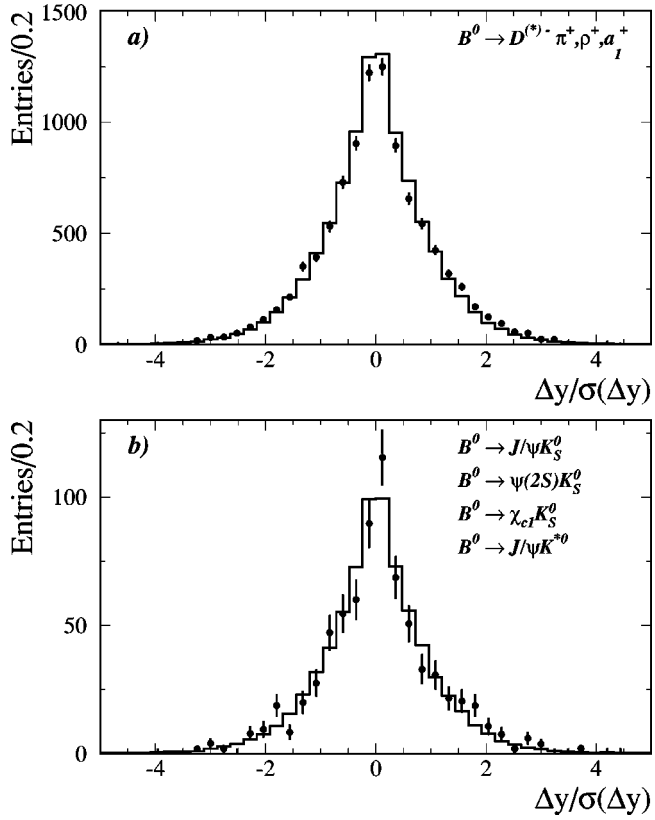

FIG. 19. Distributions of the measured distance in the vertical direction Δy between the $B_{\rm tag}$ vertex and the beam spot position, divided by the event-by-event error on the measured distance $\sigma(\Delta y)$ for each event: (a) Monte Carlo simulation (histogram) and data (points with error bars) for the flavor-eigenstate sample; (b) flavor-eigenstate (histogram), and $\eta_{CP}=-1$ and $J/\psi K^{*0}$ (points with error bars) samples in data. All distributions have been background-subtracted with events from the $m_{\rm ES}$ sideband. In (a), the data distribution has been normalized to the same area as the Monte Carlo simulation distribution; in (b) the combined $\eta_{CP}=-1$ and $J/\psi K^{*0}$ data distribution has been normalized to the same area as the flavor-eigenstate distribution.

The second control sample is obtained from off-resonance data alone. Charged tracks from these continuum events are randomly split into two sets, and the vertex of each set is found with the same algorithm used to determine the $B_{\rm tag}$ vertex. In this case the $B_{\rm tag}$ vertex reconstruction strategy is applied to both vertices in the event, so that this sample provides an unbiased estimation of the resolution, suitable for comparisons between data and Monte Carlo simulation. Results from this study are compatible with those reported above.

F. Comparison of 1999–2000 and 2001 performance

The internal alignment of the SVT has improved significantly for the reconstruction of the 2001 data set (run 2) compared to 1999–2000 (run 1). Therefore, we expect better resolution and event-by-event errors on Δt for run 2, which requires the use of separate resolution functions for the two data sets.

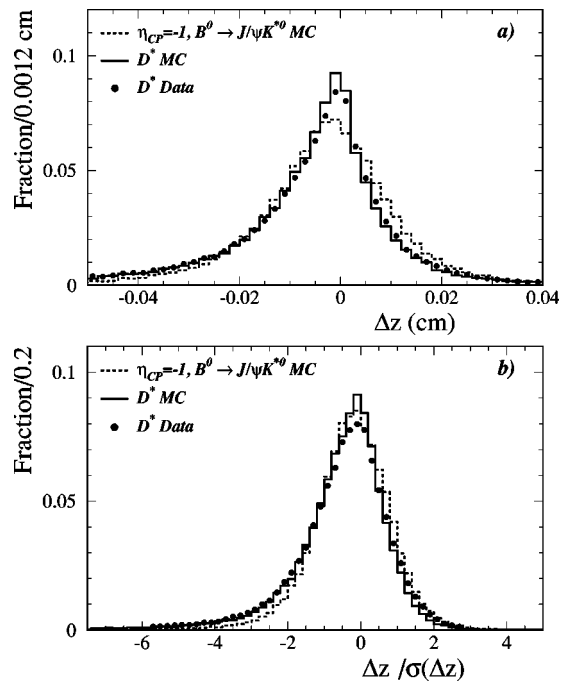

FIG. 20. (a) Δz and (b) $\Delta z/\sigma(\Delta z)$ distributions for the D^{*+} control sample in data (points with error bars) and Monte Carlo simulation (solid histogram). For comparison, the difference between the measured Δz and true Δz for $\eta_{CP}=-1$ and $J/\psi K^{*0}$ events in Monte Carlo simulation is also shown (dashed histogram). All distributions are normalized to unit area.

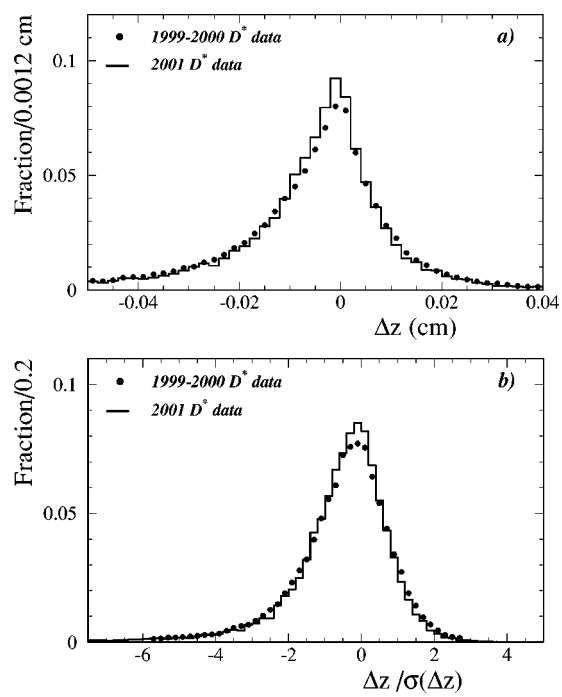

FIG. 21. Comparison of the distributions of (a) Δz and (b) $\Delta z/\sigma(\Delta z)$ for the D^{*+} control sample described in Sec. V E 3, for run 1 (points) and run 2 (histogram) data. All distributions are normalized to unit area.

The differences in resolution and event-by-event errors for run 1 and run 2 are illustrated in Fig. 21, where a comparison of the distributions for Δz and $\Delta z/\sigma(\Delta z)$ in the D^{*+} control sample described in Sec. V E 3 is shown. From the separate analysis of the two data sets, we conclude the following:

(1) The event-by-event errors on Δz are underestimated by 15% (5%) for the run 1 (run 2) data set (Fig. 21b).
(2) There is no statistically significant difference in the bias between the two data sets.
(3) The resolution for the run 1 data set is about 15% worse than that for run 2 (Fig. 21a).

The improved quality of the event-by-event errors in run 2 is also illustrated in Fig. 22, where we compare the distributions of χ^2 probability for the $B_{\rm tag}$ vertex fit with the flavor-eigenstate data sample selected from the two different data periods.

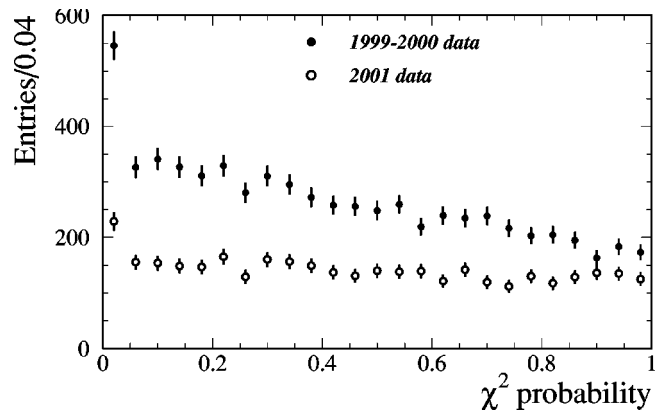

FIG. 22. Comparison of the χ^2 probability distributions of the $B_{\rm tag}$ vertex fit for the flavor-eigenstate data samples in run 1 and run 2. The distributions have been background-subtracted with events from the $m_{\rm ES}$ sideband. The area of each distribution equals the total number of events in the corresponding sample.

VI. LIKELIHOOD FIT METHOD

The value of $\sin 2\beta$ is extracted from the tagged B_{CP} sample with an unbinned maximum-likelihood technique based on $\ln \mathcal{L}_{CP}$ and the probability density functions $\mathcal{F}_\pm$ of Eq. (11). However, the dilutions $\mathcal{D}_i$ and Δz resolution parameters $\hat{a}_i$ are also needed for the measurement. Assuming that mistag rates and vertex resolutions do not depend on the particular channel used to reconstruct the B meson, these parameters are best determined with the much larger mixing sample, since they also appear in $\mathcal{L}_{\rm mix}$. In order to properly incorporate the correlations between these parameters and $\sin 2\beta$, the fit is performed by simultaneously maximizing the sum

$$\ln \mathcal{L}_{CP} + \ln \mathcal{L}_{\rm mix} \tag{24}$$

for the combined tagged $B_{\rm flav}$ and B_{CP} samples. The values of B^0 lifetime and Δm_d are kept fixed in extracting $\sin 2\beta$.

The value of Δm_d is obtained with an unbinned maximum-likelihood fit to the tagged $B_{\rm flav}$ sample alone, where the log-likelihood $\ln \mathcal{L}_{\rm mix}$ is maximized while keeping the B^0 lifetime fixed.

A. Mistag asymmetries

The probabilities of mistagging a B^0 or $\overline{B}^0$ meson are expected to be very nearly, but not exactly, equal. For example, the response of the detector to positive pions and kaons differs from its response to negative pions and kaons due to differences in total and charge-exchange cross sections. To account for any possible mistag differences, we introduce separate mistag probabilities w for B^0 and $\overline{w}$ for $\overline{B}^0$ with the conventions

$$\langle w \rangle = \tfrac{1}{2}(w+\overline{w}); \quad \Delta w = (w-\overline{w})$$

$$\mathcal{D} = 1-2w; \qquad \overline{\mathcal{D}} = 1-2\overline{w}$$

$$\langle \mathcal{D} \rangle = \tfrac{1}{2}(\mathcal{D}+\overline{\mathcal{D}}); \quad \Delta\mathcal{D} = (\mathcal{D}-\overline{\mathcal{D}}).$$

The time distributions for the mixing and CP samples will thus depend on whether the tag was identified as a B^0 or a $\overline{B}^0$, resulting in modifications to the expressions for mixing time development [Eq. (4)]

$$\begin{aligned} h_{\pm,tag=B^0} &\propto [(1+\tfrac{1}{2}\Delta\mathcal{D}) \pm \langle\mathcal{D}\rangle \cos\Delta m_d \Delta t] \\ h_{\pm,tag=\overline{B}^0} &\propto [(1-\tfrac{1}{2}\Delta\mathcal{D}) \pm \langle\mathcal{D}\rangle \cos\Delta m_d \Delta t], \end{aligned} \tag{25}$$

where the $\pm$ in the index refers to mixed $(-)$ and unmixed $(+)$ events as before, and for CP violation time development [Eq. (8)]

$$f_\pm \propto [\left(1 \pm \tfrac{1}{2}\Delta\mathcal{D}\right) \mp \langle\mathcal{D}\rangle \eta_{CP} \sin 2\beta \sin\Delta m_d \Delta t], \tag{26}$$

where the $\pm$ in the index refers to events where $B_{\rm tag}$ is a B^0 $(+)$ and $\overline{B}^0$ $(-)$ and we have taken $|\lambda|=1$.

B. Background modeling

In the presence of backgrounds, the probability distribution functions $\mathcal{H}_\pm$ of Eq. (5) and $\mathcal{F}_\pm$ of Eq. (11) must be extended to include a term for each significant background source. The backgrounds for the flavor eigenstates and $\eta_{CP}=-1$ modes are quite small and are mostly combinatoric in nature. However, for the $B^0 \to J/\psi K_L^0$ and $B^0 \to J/\psi K^{*0}$ channels the backgrounds are substantial and originate mainly from other $B \to J/\psi X$ modes that have, to a very good approximation, the same flavor tagging and Δt resolution properties as the signal. The background properties of the flavor eigenstates, $\eta_{CP}=-1$ modes, and the non-J/ψ background in the $B^0 \to J/\psi K_L^0$ channel are determined empirically from sideband events in the data.

1. Background formulation for flavor eigenstates and $\eta_{CP}=-1$ modes

The background parametrizations are allowed to differ for each tagging category. Each event belongs to a particular tagging category i. In addition, the event is classified as ei-

ther mixed $(-)$ or unmixed $(+)$ for a flavor eigenstate or by whether $B_{\rm tag}$ was a B^0 $(+)$ or a $\overline{B}^0$ $(-)$ for a CP eigenstate. Thus background distributions j must be specified for each possibility $(+/-,i)$, so that the full likelihood function becomes

$$\begin{aligned}\mathcal{H}_{\pm,i}=&f^{\rm flav}_{i,\rm sig}\mathcal{H}_{\pm}(\Delta t;\Gamma,\Delta m_d,w_i,\hat{a}_i)\\&+f^{\rm flav}_{i,\rm peak}\mathcal{B}^{\rm flav}_{\pm,i,\rm peak}(\Delta t;\hat{a}_i)\\&+\sum_{j=\rm bkgd}f^{\rm flav}_{i,j}\mathcal{B}^{\rm flav}_{\pm,i,j}(\Delta t;\hat{b}_i)\end{aligned}\tag{27}$$

for flavor eigenstates, and

$$\begin{aligned}\mathcal{F}_{\pm,i}=&f^{CP}_{i,\rm sig}\mathcal{F}_{\pm}(\Delta t;\Gamma,\Delta m_d,w_i,\sin 2\beta,\hat{a}_i)\\&+f^{CP}_{i,\rm peak}\mathcal{B}^{CP}_{\pm,i,\rm peak}(\Delta t;\hat{a}_i)\\&+\sum_{j=\rm bkgd}f^{CP}_{i,j}\mathcal{B}^{CP}_{\pm,i,j}(\Delta t;\hat{b}_i)\end{aligned}\tag{28}$$

for CP eigenstates.

The fraction of background events for each source and tagging category is a function of $m_{\rm ES}$ and is given by $f_{i,j}$. The peaking and combinatorial background PDFs, $\mathcal{B}_{\pm,i,\rm peak}$ and $\mathcal{B}_{\pm,i,j}$, provide an empirical description of the Δt distribution of the background events in the sample, including a resolution function parametrized by $\hat{a}_i$ and $\hat{b}_i$, respectively. These distributions are normalized such that, for each i and j,

$$\int_{-\infty}^{\infty}(\mathcal{B}_{+,i,j}+\mathcal{B}_{-,i,j})d\Delta t=1.\tag{29}$$

The probability that a B^0 candidate is a signal or a background event is determined from a separate fit to the observed $m_{\rm ES}$ distributions of $B_{\rm flav}$ or B_{CP} candidates with $\eta_{CP}=-1$. We describe the $m_{\rm ES}$ shape with a single Gaussian distribution $\mathcal{S}(m_{\rm ES})$ for the signal and an ARGUS parametrization $\mathcal{A}(m_{\rm ES})$ for the background [Eq. (17)]. Based on this fit, the event-by-event signal and background probabilities that appear as the relative weights for the various signal and background terms in Eqs. (27) and (28) are given by

$$\begin{aligned}f_{i,\rm sig}(m_{\rm ES})&=\frac{(1-\delta_{\rm peak})\mathcal{S}(m_{\rm ES})}{\mathcal{S}(m_{\rm ES})+\mathcal{A}(m_{\rm ES})}\\f_{i,\rm peak}(m_{\rm ES})&=\frac{\delta_{\rm peak}\mathcal{S}(m_{\rm ES})}{\mathcal{S}(m_{\rm ES})+\mathcal{A}(m_{\rm ES})}\\\sum_{j=\rm bkgd}f_{i,j}(m_{\rm ES})&=\frac{\mathcal{A}(m_{\rm ES})}{\mathcal{S}(m_{\rm ES})+\mathcal{A}(m_{\rm ES})}.\end{aligned}\tag{30}$$

The fraction $\delta_{\rm peak}$ of the signal Gaussian distribution that is due to peaking backgrounds is determined from Monte Carlo simulation.

Backgrounds arise from many different sources. Rather than attempting to determine the various physics contributions we use an empirical description in the likelihood fit, allowing for background components with various time dependencies. For the $B_{\rm flav}$ sample, the background time distributions considered, each with its own effective dilution factor $\mathcal{D}_i$ and either a common resolution function $\mathcal{R}(\delta_t;\hat{b}_i)$ or the signal resolution function $\mathcal{R}(\delta_t=\Delta t-\Delta t_{\rm true};\hat{a}_i)$, are

$$\begin{aligned}\mathcal{B}^{\rm flav}_{\pm,i,1}&=\tfrac{1}{2}(1\pm\mathcal{D}^{\rm flav}_{i,1})\delta(\Delta t_{\rm true})\otimes\mathcal{R}(\delta_t;\hat{b}_i),\\\mathcal{B}^{\rm flav}_{\pm,i,2}&=\tfrac{1}{4}\Gamma^{\rm flav}_{i,2}(1\pm\mathcal{D}^{\rm flav}_{i,2})\\&\quad\times e^{-\Gamma^{\rm flav}_{i,2}|\Delta t_{\rm true}|}\otimes\mathcal{R}(\delta_t;\hat{b}_i),\\\mathcal{B}^{\rm flav}_{\pm,i,3}&=\tfrac{1}{4}\Gamma^{\rm flav}_{i,3}(1\pm\mathcal{D}^{\rm flav}_{i,3}\cos\Delta m_{i,3}\Delta t_{\rm true})\\&\quad\times e^{-\Gamma^{\rm flav}_{i,3}|\Delta t_{\rm true}|}\otimes\mathcal{R}(\delta_t;\hat{b}_i),\\\mathcal{B}^{\rm flav}_{\pm,i,\rm peak}&=\tfrac{1}{4}\Gamma_{i,\rm peak}(1\pm\mathcal{D}^{\rm flav}_{i,\rm peak}\cos\Delta m_{i,\rm peak}\Delta t_{\rm true})\\&\quad\times e^{-\Gamma^{\rm flav}_{i,\rm peak}|\Delta t_{\rm true}|}\otimes\mathcal{R}(\delta_t;\hat{a}_i),\end{aligned}\tag{31}$$

corresponding to prompt, nonprompt, and mixing background components, as well as a peaking contribution. For the $\eta_{CP}=-1$ sample, the possible background contributions are

$$\begin{aligned}\mathcal{B}^{CP}_{\pm,i,1}&=\tfrac{1}{4}\delta(\Delta t_{\rm true})\otimes\mathcal{R}(\delta_t;\hat{b}_i),\\\mathcal{B}^{CP}_{\pm,i,2}&=\tfrac{1}{4}\Gamma_{i,2}(1\pm\mathcal{D}^{CP}_{i,2}\sin\Delta m_d\Delta t_{\rm true})\\&\quad\times e^{-\Gamma^{CP}_{i,2}|\Delta t_{\rm true}|}\otimes\mathcal{R}(\delta_t;\hat{b}_i),\\\mathcal{B}^{CP}_{\pm,i,\rm peak}&=\tfrac{1}{4}\Gamma_{i,\rm peak}(1\pm\mathcal{D}^{CP}_{i,\rm peak}\sin\Delta m_d\Delta t_{\rm true})\\&\quad\times e^{-\Gamma^{CP}_{\rm peak}|\Delta t_{\rm true}|}\otimes\mathcal{R}(\delta_t;\hat{a}_i),\end{aligned}\tag{32}$$

corresponding to prompt and CP background components, as well as a peaking contribution.

The likelihood fit includes as free parameters the relative fractions of prompt versus nonprompt background, as well as apparent lifetimes, mixing frequencies and dilutions, and common effective resolution parameters $\hat{b}_i$ that best describe the events with high weights for being background in the $B_{\rm flav}$ and B_{CP} samples. Roughly, *uds* continuum, short-lived charm continuum, and short-lived $B\overline{B}$ backgrounds fall into the prompt category, while some fraction of long-lived charm and $B\overline{B}$ are treated as nonprompt. The standard vertex algorithms are of course applied to all candidates in the $B_{\rm flav}$ and B_{CP} samples. Thus, the $B_{\rm rec}$ vertex fit must converge and the usual procedure is applied to obtain a self-consistent vertex for the remaining tracks on the tag side of the event, irrespective of whether the candidate event is eventually classified as signal or background. To maintain a parallel treatment with the signal PDF, an effective resolution function $\mathcal{R}(\delta_t;\hat{b}_i)$ is used to scale the result of the calculated uncertainty on Δz for background events. The actual choice of background parameters is described in Sec. VI D below, along with additional assumptions.

2. Background formulation for $B^0 \to J/\psi K_L^0$

The higher background level in the $B^0 \to J/\psi K_L^0$ channel requires a more extensive treatment of its properties. As discussed in Sec. III D, the data are used to determine the relative fraction of signal, background from $B \to J/\psi X$ events, and events with a misreconstructed $J/\psi \to ll$ candidate. Along with a Monte Carlo simulation of the channels that contribute to the $B \to J/\psi X$ background, this information is used to formulate the PDF model. In addition, some of the $J/\psi X$ background modes, such as $B^0 \to J/\psi K^{*0}$ and $B^0 \to J/\psi K_S^0$, have a nonzero *CP* asymmetry (η_{CP}), as given in Table VI. The value of the asymmetry in $B^0 \to J/\psi K^{*0}(K_L^0 \pi^0)$ is taken from the measurement of $R_\perp = 0.160 \pm 0.032 \pm 0.014$ in Ref. [13]. The probability density functions $\mathcal{F}_\pm$ of Eq. (11) are modified to include contributions for each of the $B \to J/\psi X$ channels α specified in Table VI and the non-J/ψ background component. The complete PDF is given by

$$\begin{aligned}\mathcal{F}_{\pm,i} = & f_{i,k,\mathrm{sig}}(\Delta E)\mathcal{F}_\pm(\Delta t;\Gamma,\Delta m_d,w_i,\sin 2\beta,\hat{a}_i) \\ & + \sum_{\alpha = J/\psi X} f_{i,k,\alpha}(\Delta E) \\ & \times \mathcal{F}_\pm(\Delta t;\Gamma,\Delta m_d,\eta_{CP,\alpha},w_i,\sin 2\beta,\hat{a}_i) \\ & + f_{i,k,\mathrm{non}-J/\psi}(\Delta E)\mathcal{B}_\pm^{KL}(\Delta t;\hat{b}). \end{aligned} \tag{33}$$

Each event is classified according to its flavor tagging category (i), flavor tag value ($\pm$), and the K_L^0 reconstruction category (k), which is either EMC or IFR. The signal fraction $f_{i,k,\mathrm{sig}}$ and background fractions $f_{i,k,\alpha}$ and $f_{i,k,\mathrm{non}\text{-}J/\psi}$ are determined as a function of ΔE and are the same for all tagging categories. The shape of the signal and background ΔE functions are determined either from data (non-J/ψ contribution) or from Monte Carlo samples (signal and $J/\psi X$ background). The normalizations $\int_{-10\,\mathrm{MeV}}^{10\,\mathrm{MeV}} f d(\Delta E)$ are determined from Tables VI and VII so that

$$\int_{-10\,\mathrm{MeV}}^{10\,\mathrm{MeV}} \left[f_{i,k,\mathrm{sig}}(\Delta E) + f_{i,k,\mathrm{non}\text{-}J/\psi}(\Delta E) + \sum_{J/\psi X} f_{i,k,\alpha}(\Delta E) \right] d(\Delta E) = 1. \tag{34}$$

The non-J/ψ background PDF is given by

$$\mathcal{B}_\pm^{KL} = F_{\tau=0}\mathcal{B}_{\pm,i,1}^{CP} + (1 - F_{\tau=0})\mathcal{B}_{\pm,i,2}^{CP} \tag{35}$$

where the dilutions $\mathcal{D}_{i,2}^{CP} = 0$ and the prompt fraction $F_{\tau=0}$, effective decay width $\Gamma_{i,2}$, and Δt resolutions parameters $\hat{b}$ are fixed to values obtained from an external fit to the $m(ll)$ sideband events as given in Table IX. The resolution function $\mathcal{R}(\Delta t;\hat{b})$ is defined in Eq. (23) with $f_3 = 0.005$ and with core bias δ_1 equal for all tagging categories.

The $J/\psi K_L^0$ sample has significant background, primarily from other J/ψ modes. The Monte Carlo simulation is used to check the flavor tagging efficiency of the inclusive J/ψ background relative to the signal for the K_L^0 mode. The inclusive J/ψ background fraction in the simulation is consistent across the flavor tagging categories to within a few percent. The flavor tagging efficiency for the fake-J/ψ background, determined from the J/ψ sideband, is also roughly consistent with signal. The composition of the $J/\psi K_L^0$ sample is determined from a fit of the ΔE spectrum before flavor tagging. We assume the inclusive J/ψ and fake-J/ψ background fractions are independent of flavor tag in the nominal fit and adjust the fractions as a function of tagging category, based on the Monte Carlo simulation and J/ψ sideband, in order to determine systematic errors.

TABLE IX. Parameters of the probability distribution function for the non-J/ψ background contribution in the $B^0 \to J/\psi K_L^0$ channel.

Parameter	Fit result
$F_{\tau=0}$	0.16 ± 0.49
$\Gamma_{i,2}$ [ps^{-1}]	1.25 ± 0.45
S_1	1.12 ± 0.26
b_1	-0.11 ± 0.20
S_2	3.9 ± 0.8
b_2	-1.2 ± 1.0
f_2	0.23 ± 0.14
f_3	0.005 (fixed)

Some of the decay modes in the inclusive J/ψ background, such as $J/\psi K^{*0}$ and $J/\psi K_S^0$, have an expected *CP* asymmetry. The mistag fractions for all *CP* modes in the inclusive J/ψ background are determined with the Monte Carlo simulation and found to be consistent with the values for the signal. We assume that the signal mistag fractions apply to the *CP* modes in the inclusive J/ψ background.

The Δt resolution for the $B \to J/\psi X$ background should be very similar to the signal resolution. However, extra tracks associated with $B^+ \to J/\psi X^+$ decay, such as the charged π from the K^{*+} decay in $B^+ \to J/\psi K^{*+}$, could bias the measurement of Δt since they are not associated with the B_{CP} vertex and therefore can be used in the B_{tag} vertex. In the Monte Carlo simulation, we find that extra tracks in the $B \to J/\psi X$ decay modes have a negligible effect on the Δtresolution. Therefore, we assume that all $B \to J/\psi X$ background has the same resolution as the signal.

The Δt resolution of the non-J/ψ background is measured with the J/ψ sideband sample. The non-J/ψ Δt resolution parameters are varied by their statistical uncertainties to estimate the systematic uncertainty.

3. Background formulation for $B^0 \to J/\psi K^{*0}(K_S^0 \pi^0)$

Monte Carlo simulation is used to construct the probability density function for the $B^0 \to J/\psi K^{*0}(K_S^0 \pi^0)$ channel. As shown in Table V, the background for this channel is due to true $B \to J/\psi X$ decays. Thus, we assume that the background has the same resolution function and tagging performance as the signal. The probability density functions $\mathcal{F}_\pm$ of Eq. (11) are modified to include contributions for each of the $B \to J/\psi X$ channels α specified in Table V. The complete PDF is given by

$$\mathcal{F}_{\pm,i}=f_{\rm sig}\mathcal{F}_{\pm}(\Delta t;\Gamma,\Delta m_d,\eta_{CP,\rm signal},w_i,\sin 2\beta,\hat{a}_i)+\sum_{\alpha=\rm bkgd} f_\alpha\mathcal{F}_{\pm}(\Delta t;\Gamma,\Delta m_d,\eta_{f,\alpha},w_i,\sin 2\beta,\hat{a}_i), \tag{36}$$

where each event is classified according to its flavor tagging category (i) and flavor tag value ($\pm$). The signal and background fractions as well as η_{CP} are taken from Table V.

C. Extensions for direct *CP* search

While the main likelihood fits are performed with the standard model expectation that $|\lambda|=1$, a search for the effects of direct CP violation is also made. Such a measurement is also particularly sensitive to possible differences in the fraction of B^0 or $\overline{B}^0$ meson that are tagged. Defining $\epsilon_{\rm tag}$ and $\overline{\epsilon}_{\rm tag}$ as the tagging efficiencies for B^0 and $\overline{B}^0$, and ϵ_r and $\overline{\epsilon}_r$ as the reconstruction efficiencies for B^0 and $\overline{B}^0$ in the $B_{\rm flav}$ sample, it is useful to construct

$$\mu_i=\frac{\epsilon_{{\rm tag},i}-\overline{\epsilon}_{{\rm tag},i}}{\epsilon_{{\rm tag},i}+\overline{\epsilon}_{{\rm tag},i}},\quad \langle\epsilon_{\rm tag}\rangle_i=\frac{\epsilon_{{\rm tag},i}+\overline{\epsilon}_{{\rm tag},i}}{2} \tag{37}$$

$$\nu_i=\frac{\epsilon_{r,i}-\overline{\epsilon}_{r,i}}{\epsilon_{r,i}+\overline{\epsilon}_{r,i}},\qquad \langle\epsilon_r\rangle_i=\frac{\epsilon_{r,i}+\overline{\epsilon}_{r,i}}{2}. \tag{38}$$

For the B_{CP} sample, the time-dependent decay rate [Eq. (10)] becomes

$$f_{\pm,i}(\Delta t)=\frac{\Gamma}{4}e^{-\Gamma|\Delta t|}\left[\frac{(1+|\lambda|^2)(1\pm X_i)}{1+|\lambda|^2+\xi_i}+(\mu_i\pm X_i')\left(\frac{2\,{\rm Im}\,\lambda}{1+|\lambda|^2+\xi_i}\sin\Delta m_d\Delta t-\frac{1-|\lambda|^2}{1+|\lambda|^2+\xi_i}\cos\Delta m_d\Delta t\right)\right] \tag{39}$$

where $\xi_i=\mu_i(|\lambda|^2-1)/(1+x_d^2)$, $X_i=\mu_i\langle\mathcal{D}\rangle_i+\Delta\mathcal{D}_i/2$, and $X_i'=\langle\mathcal{D}\rangle_i+\mu_i\Delta\mathcal{D}_i/2$. Likewise, for the $B_{\rm flav}$ sample the time-dependent decay rate [Eq. (4)] becomes

$$h(\Delta t)=\frac{\Gamma}{4}e^{-\Gamma|\Delta t|}\frac{1+s_1\nu_i}{\left[1-\dfrac{\mu_i\nu_i}{1+x_d^2}\right]}[1+s_2X_i-s_i(\mu_i+s_2X_i')\cos\Delta m_d\Delta t] \tag{40}$$

where

$$s_1=1(-1)\ \text{ if the reconstructed } B \text{ is a } B^0(\overline{B}^0)$$

$$s_2=1(-1)\ \text{ for a } B^0(\overline{B}^0)\ \text{ tag.}$$

The parameters ν_i, $\langle\epsilon_{\rm tag}\rangle_i$, and μ_i can be extracted from time-integrated numbers of events in the $B_{\rm flav}$ sample. Defining integrated samples of events by

$$\begin{aligned}
N_i^{\rm tag}&=N(B^0/\overline{B}^0\ \text{ tag in } i\text{th category},B_{\rm flav}=B^0)\\
\overline{N}_i^{\rm tag}&=N(B^0/\overline{B}^0\ \text{ tag in } i\text{th category},B_{\rm flav}=\overline{B}^0)\\
N_i^{\rm no\ tag}&=N(\text{no tag in } i\text{th category},B_{\rm flav}=B^0)\\
\overline{N}_i^{\rm no\ tag}&=N(\text{no tag in } i\text{th category},B_{\rm flav}=\overline{B}^0),
\end{aligned} \tag{41}$$

it can be shown that

$$\begin{aligned}
\nu_i&=\frac{N_i^{\rm tag}-\overline{N}_i^{\rm tag}+N_i^{\rm no\ tag}-\overline{N}_i^{\rm no\ tag}}{N_i^{\rm tag}+\overline{N}_i^{\rm tag}+N_i^{\rm no\ tag}+\overline{N}_i^{\rm no\ tag}}\\
\langle\epsilon_{\rm tag}\rangle_i&=\frac{2N_i^{\rm tag}\overline{N}_i^{\rm tag}+\overline{N}_i^{\rm tag}N_i^{\rm no\ tag}+N_i^{\rm tag}\overline{N}_i^{\rm no\ tag}}{2(N_i^{\rm tag}+N_i^{\rm no\ tag})(\overline{N}_i^{\rm tag}+\overline{N}_i^{\rm no\ tag})}\\
\mu_i&=\frac{(1+x_d^2)(\overline{N}_i^{\rm tag}N_i^{\rm no\ tag}-N_i^{\rm tag}\overline{N}_i^{\rm no\ tag})}{2N_i^{\rm tag}\overline{N}_i^{\rm tag}+\overline{N}_i^{\rm tag}N_i^{\rm no\ tag}+N_i^{\rm tag}\overline{N}_i^{\rm no\ tag}}
\end{aligned} \tag{42}$$

under the assumption that nearly all B mesons decay to final states that can be reached from either B^0 or $\overline{B}^0$, but not both. The results for $\langle\epsilon_{\rm tag}\rangle_i$ and μ_i are shown in Table X. The value of ν_i, averaged over all four tagging categories, is 0.004 ± 0.012. While there is no statistically significant difference in the tagging efficiencies or the reconstruction efficiencies given by μ_i and ν_i, we use the central values obtained from the $B_{\rm flav}$ sample in performing the fit for $|\lambda|$.

D. Free parameters for the $\sin 2\beta$ and Δm_d fits

The unbinned likelihood fit for $\sin 2\beta$ has a total of 45 free parameters:

(1) *Value of* $\sin 2\beta$
(2) *Signal resolution function*: Sixteen parameters $\hat{a}_i$ to describe the resolution function for the signal. Due to improvements in the reconstruction algorithms, the run 1 and run 2 resolution functions are found to be different, as described in Sec. V F. Thus, we allow for separate

TABLE X. Values of $\langle \epsilon_{\rm tag} \rangle_i$ and μ_i for the four tagging categories, as determined by counting numbers of tagged and untagged events in the $B_{\rm flav}$ sample.

Tagging category	$\langle \epsilon_{\rm tag} \rangle_i$	μ_i
Lepton	0.095 ± 0.002	0.069 ± 0.032
Kaon	0.358 ± 0.003	-0.005 ± 0.014
NT1	0.080 ± 0.002	0.061 ± 0.035
NT2	0.139 ± 0.002	0.017 ± 0.026

resolution function parameters for these data samples, each with eight free parameters, being a scale factor S_1 for the event-by-event Δz resolution errors of the core Gaussian components, individual core bias scale factors $b_{1,i}$ for the four tagging categories and a common tail bias b_2, and the tail f_2 and outlier f_3 fractions; the scale factor of the tail component is fixed to 3.0 and the width of the outlier component is fixed to 8 ps with zero bias.

(3) *Signal dilutions*: Eight parameters to describe the measured average dilutions $\langle \mathcal{D} \rangle_i$ and dilution differences $\Delta \mathcal{D}_i$ in each tagging category.

(4) *Background resolution function*: Six parameters are used to describe a common resolution function for all nonpeaking backgrounds. As with the signal resolution function, we include separate resolution function parameters for the run 1 and run 2 data samples. The resolution function is taken as a single Gaussian distribution with a scale factor S_1 for the event-by-event Δz errors and a common bias scale factor b_1, and an outlier fraction f_3; the width of the outlier component is taken to be a fixed 8 ps with zero bias.

(5) B_{flav} *background composition parameters*: A total of 13 parameters describe the $B_{\rm flav}$ background composition. We make several assumptions to simplify the parametrization shown in Eq. (31), such as removing the mixing background contribution by setting $f_{i,3}^{\rm flav}=0$, and assign a corresponding systematic uncertainty. The size of the peaking background is determined from Monte Carlo simulation to be $\delta_{\rm peak}^{\rm flav}=(1.5\pm 0.5)\%$ of the signal contribution in each tagging category. This contribution is predominately from B^+ events, so $\Delta m_{i,\rm peak}=0$, $\Gamma_{i,\rm peak}^{\rm flav}=\Gamma_{B^+}$ and $D_{i,\rm peak}^{\rm flav}$ are taken from the B^+ data sample (Table XI). The effective dilutions for the prompt ($D_{i,1}^{\rm flav}$, 4 parameters) and non-prompt ($D_{i,2}^{\rm flav}$, 4 parameters) contributions are allowed to vary. The relative amount of these two contributions is allowed to vary independently in each tagging category (4 parameters). For the nonprompt contribution, $\Gamma_{i,2}^{\rm flav}$ is assumed to be the same for all tagging categories, giving one free parameter.

(6) *CP background composition parameters*: One parameter, the fraction of prompt relative to nonprompt background, assumed to be the same for each tagging category, is allowed to float to describe the *CP* background properties. The effective dilutions of the nonprompt and peaking contribution are set to zero ($D_{i,2}^{CP}=D_{i,\rm peak}^{CP}=0$), corresponding to no *CP* asymmetry in the background. The size and parameters of the peaking background are determined from Monte Carlo simulation. The fraction of peaking background is $\delta_{\rm peak}^{CP}=(1\pm 1)\%$ of the signal contribution, independent of tagging category. This contribution is assumed to have lifetime parameters in common with the signal. Finally, the lifetime of the nonprompt background is assumed to be τ_{B^0} in all tagging categories.

The unbinned likelihood fit for Δm_d has 44 free parameters, removing $\sin 2\beta$ and the parameter for fraction of prompt background in the *CP* sample and leaving Δm_d to float.

E. Blind analysis

A blind analysis technique was adopted for the extraction of $\sin 2\beta$ and Δm_d in order to eliminate possible experimenter's bias. We used a method that hides not only the central value for these parameters from the unbinned maximum-likelihood fit, but also the visual *CP* asymmetry in the Δt distribution. The error on both the asymmetry and Δm_d is not hidden.

The amplitude of the asymmetry $\mathcal{A}_{CP}(\Delta t)$ from the fit was hidden by a one-time choice of sign flip and arbitrary offset based on a user-specified key word. The sign flip hides whether a change in the analysis increases or decreases the resulting asymmetry. However, the magnitude of the change is not hidden. The visual *CP* asymmetry in the Δt distribution is hidden by multiplying Δt by the sign of the tag and adding an arbitrary offset.

With these techniques, systematic studies can be performed while keeping the numerical value of $\sin 2\beta$ or Δm_d hidden. In particular, we can check that the hidden Δt distributions are consistent for B^0 and $\overline{B}^0$ tagged events. The same is true for all the other checks concerning tagging, vertex resolution and the correlations between them. For instance, fit results in the different tagging categories can be compared to each other, since each fit is hidden in the same way. The analysis procedure for extracting $\sin 2\beta$ and Δm_d were frozen prior to unblinding.

VII. B^0 FLAVOR OSCILLATIONS AND MISTAG RATES

A. Likelihood fit results for Δm_d

We extract Δm_d, the dilution factors $\mathcal{D}_i$, the Δt resolution parameters $\hat{a}_i$, and the background Δt parametrization

TABLE XI. Average mistag fractions $\langle w_i \rangle$ and mistag differences Δw_i for each tagging category i from a maximum-likelihood fit to the distribution for the B^+ control sample.

Tagging category	$\langle w \rangle_i$ [%]	Δw_i [%]
Lepton	4.6 ± 0.6	1.1 ± 1.2
Kaon	11.8 ± 0.5	-0.3 ± 1.0
NT1	21.3 ± 1.6	-5.9 ± 3.2
NT2	37.2 ± 1.3	-0.7 ± 2.7

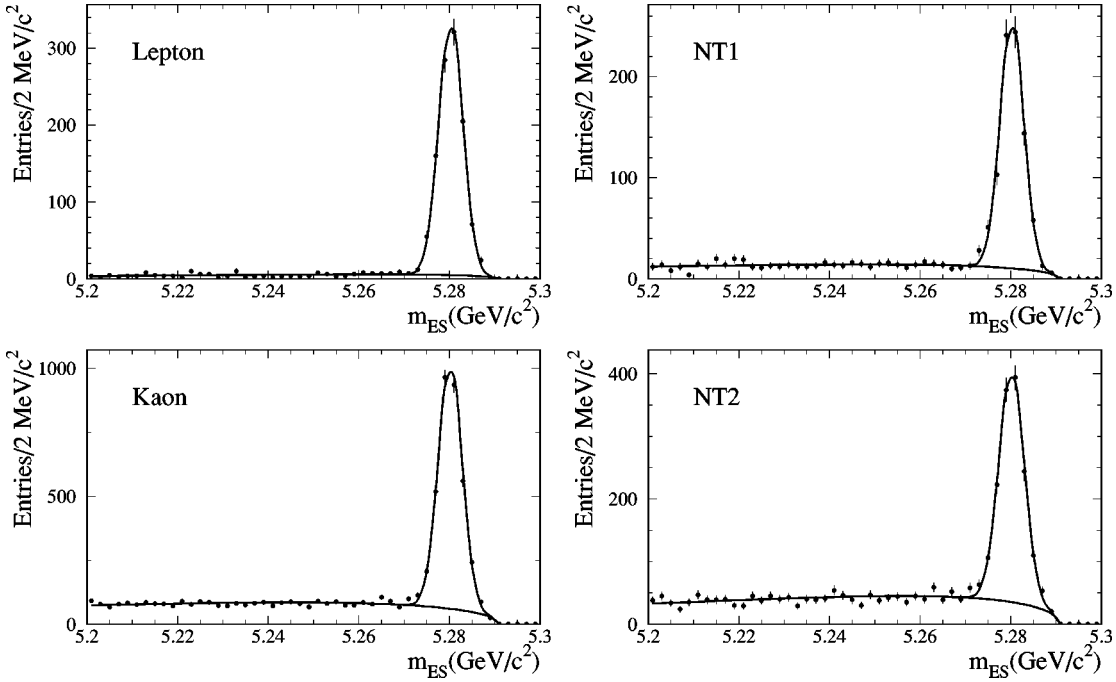

FIG. 23. Distribution of $m_{\rm ES}$ for mixing $B_{\rm flav}$ candidates in separate tagging categories (`Lepton`, `Kaon`, `NT1` and `NT2`), overlaid with the result of a fit with a Gaussian distribution for the signal and an ARGUS function for the background.

by fitting the Δt distributions of the flavor-eigenstate B^0 sample with $m_{\rm ES}>5.2\ {\rm GeV}/c^2$ with the likelihood function described in Sec. VI. The selection of the $B_{\rm flav}$ data sample is described in Sec. III C 1. We also demand a valid tag and Δt determination for the event, based on the algorithms described in Secs. IV and V. The more restrictive requirements $|\Delta t|<20$ ps and $\sigma_{\Delta t}<1.4$ ps are applied to the proper time-difference measurement. In addition, identified kaons in the $B_{\rm tag}$ decay are rejected in the reconstruction of the tagging vertex. These requirements are intended to reduce systematic errors on the precision Δm_d measurement. The final sample consists of 12310 fully reconstructed and tagged B^0 candidates with $m_{\rm ES}>5.2\ {\rm GeV}/c^2$, of which 7399 are in the signal region $m_{\rm ES}>5.27\ {\rm GeV}/c^2$.

The breakdown of this mixing $B_{\rm flav}$ sample into individual tagging categories is shown in Fig. 23 as a function of $m_{\rm ES}$. Superimposed on the observed mass spectra are the results of the fits with a Gaussian distribution for the signal and the ARGUS background function for the background. The tagging efficiency and signal purity for the individual tagging categories in data are extracted from fits to the $m_{\rm ES}$ distributions and are listed in Table XII. The efficiency for each tagging category is defined as the ratio of the number of signal events for each tag over the total number of signal events after imposition of vertex cuts.

The results from the likelihood fit to the mixing sample are summarized in Table XIII. The probability to obtain a likelihood smaller than the observed value, evaluated with fast parametrized Monte Carlo simulation of a large number of similar experiments, is $(44\pm1)\%$. The Δt distributions of the signal ($m_{\rm ES}>5.27\ {\rm GeV}/c^2$) and background ($m_{\rm ES}<5.27\ {\rm GeV}/c^2$) candidates, overlaid with the projection of the likelihood fit, are shown in Fig. 24. In Fig. 25 the mixing asymmetry of Eq. (3) is plotted; the time dependence of the mixing probability is clearly visible.

The tagging separation $Q=\epsilon_{tag}(1-2w)^2$ is calculated from the efficiencies and the mistag rates quoted in Tables XII and XIII respectively. Summing over all tagging categories, we measure a combined effective tagging efficiency $Q\approx27\%$.

Two small corrections, which are described in more detail in Secs. VII B 2 and VII B 4 together with their assigned systematic errors, are applied to the output of the fit. The value of Δm_d obtained after applying these corrections is

$$\Delta m_d=0.516\pm0.016\pm0.010\ {\rm ps}^{-1},$$

where the first error is statistical and the second systematic.

We have also examined the fitted value for Δm_d with various subsamples of the full data set, including individual B decay channels, separate tagging categories, the state of the reconstructed $B_{\rm rec}$ or tagging $B_{\rm tag}$, and different time

TABLE XII. Tagging efficiencies for hadronic B^0 decays and signal purities in data, shown separately for the four tagging categories. Signal purities are estimated for $m_{\rm ES}>5.27\ {\rm GeV}/c^2$.

Tagging category	Efficiency [%]	Signal	Purity [%]
`Lepton`	11.8 ± 0.3	1097 ± 34	96.0 ± 0.7
`Kaon`	33.9 ± 0.5	3156 ± 63	84.6 ± 0.7
`NT1`	8.6 ± 0.3	798 ± 31	88.9 ± 1.2
`NT2`	13.9 ± 0.4	1293 ± 43	79.4 ± 1.3
Full sample	68.1 ± 0.4	6347 ± 89	85.8 ± 0.5

TABLE XIII. Results from the likelihood fit to the Δt distributions of the hadronic B^0 decays. The value for Δm_d includes small corrections as described in the text. The first major column contains the fit results, while the second major column contains the correlation coefficients with respect to Δm_d for each fit parameter.

	Fit result		Correlation	
Parameter	Run 1	Run 2	Run 1	Run 2
Δm_d [ps^{-1}]	0.516 ± 0.016			
Signal resolution function				
S_1 (core)	1.37 ± 0.09	1.18 ± 0.11	0.25	0.16
$b_1(\Delta t)$ `lepton` (core)	0.06 ± 0.13	-0.04 ± 0.16	0.08	0.00
$b_1(\Delta t)$ `kaon` (core)	-0.22 ± 0.08	-0.25 ± 0.09	0.03	0.00
$b_1(\Delta t)$ `NT1` (core)	-0.07 ± 0.15	-0.45 ± 0.21	-0.00	0.00
$b_1(\Delta t)$ `NT2` (core)	-0.46 ± 0.12	-0.20 ± 0.16	0.01	0.03
$b_2(\Delta t)$ (tail)	-5.0 ± 4.2	-7.5 ± 2.4	0.04	0.06
f_2(tail)	0.014 ± 0.020	0.015 ± 0.010	0.06	0.07
f_3(outlier)	0.008 ± 0.004	0.000 ± 0.014	-0.09	0.01
Signal dilutions				
$\langle D \rangle$, `lepton`	0.842 ± 0.028		0.24	
$\langle D \rangle$, `kaon`	0.669 ± 0.023		0.30	
$\langle D \rangle$, `NT1`	0.563 ± 0.044		0.11	
$\langle D \rangle$, `NT2`	0.313 ± 0.041		0.11	
ΔD, `lepton`	-0.006 ± 0.045		0.02	
ΔD, `kaon`	0.024 ± 0.033		0.01	
ΔD, `NT1`	-0.086 ± 0.068		0.00	
ΔD, `NT2`	0.100 ± 0.060		0.00	
Background properties				
τ, mixing bkgd [ps]	0.853 ± 0.036		-0.01	
$f(\tau=0)$, mixing bkgd, `lepton`	0.05 ± 0.10		0.01	
$f(\tau=0)$, mixing bkgd, `kaon`	0.42 ± 0.05		0.01	
$f(\tau=0)$, mixing bkgd, `NT1`	0.33 ± 0.08		0.01	
$f(\tau=0)$, mixing bkgd, `NT2`	0.32 ± 0.08		0.01	
Background resolution function				
S_1 (core)	1.211 ± 0.043	1.131 ± 0.046	-0.00	0.00
$b_1(\Delta t)$ (core)	-0.135 ± 0.031	-0.015 ± 0.038	-0.00	-0.00
f_3 (outlier)	0.022 ± 0.004	0.036 ± 0.007	-0.01	0.02
Background dilutions				
$\langle D \rangle$, `lepton`, $\tau=0$	0.0 ± 2.9		-0.02	
$\langle D \rangle$, `kaon`, $\tau=0$	0.52 ± 0.08		-0.03	
$\langle D \rangle$, `NT1`, $\tau=0$	0.67 ± 0.27		-0.01	
$\langle D \rangle$, `NT2`, $\tau=0$	-0.05 ± 0.13		-0.00	
$\langle D \rangle$, `lepton`, $\tau>0$	0.34 ± 0.13		0.02	
$\langle D \rangle$, `kaon`, $\tau>0$	0.26 ± 0.06		0.04	
$\langle D \rangle$, `NT1`, $\tau>0$	-0.13 ± 0.11		0.01	
$\langle D \rangle$, `NT2`, $\tau>0$	0.12 ± 0.031		0.01	

periods. As can be seen from Table XIV, the values obtained from the subsample fits are all consistent with the global result for Δm_d.

B. Systematic error estimation

Systematic errors can be grouped into four categories: signal properties and description, background properties and description, fixed external parameters and statistical limitations of Monte Carlo validation tests of the fitting procedure. A summary of these sources for the hadronic B^0 sample is shown in Table XV. In the following, the individual contributions are referenced by the lettered lines in this table.

1. Signal properties and description

For the signal events, the use of a double Gaussian plus outlier model for rescaling the event-by-event Δt errors as part of the likelihood fit means that uncertainties in the vertex resolution are incorporated into the statistical error on

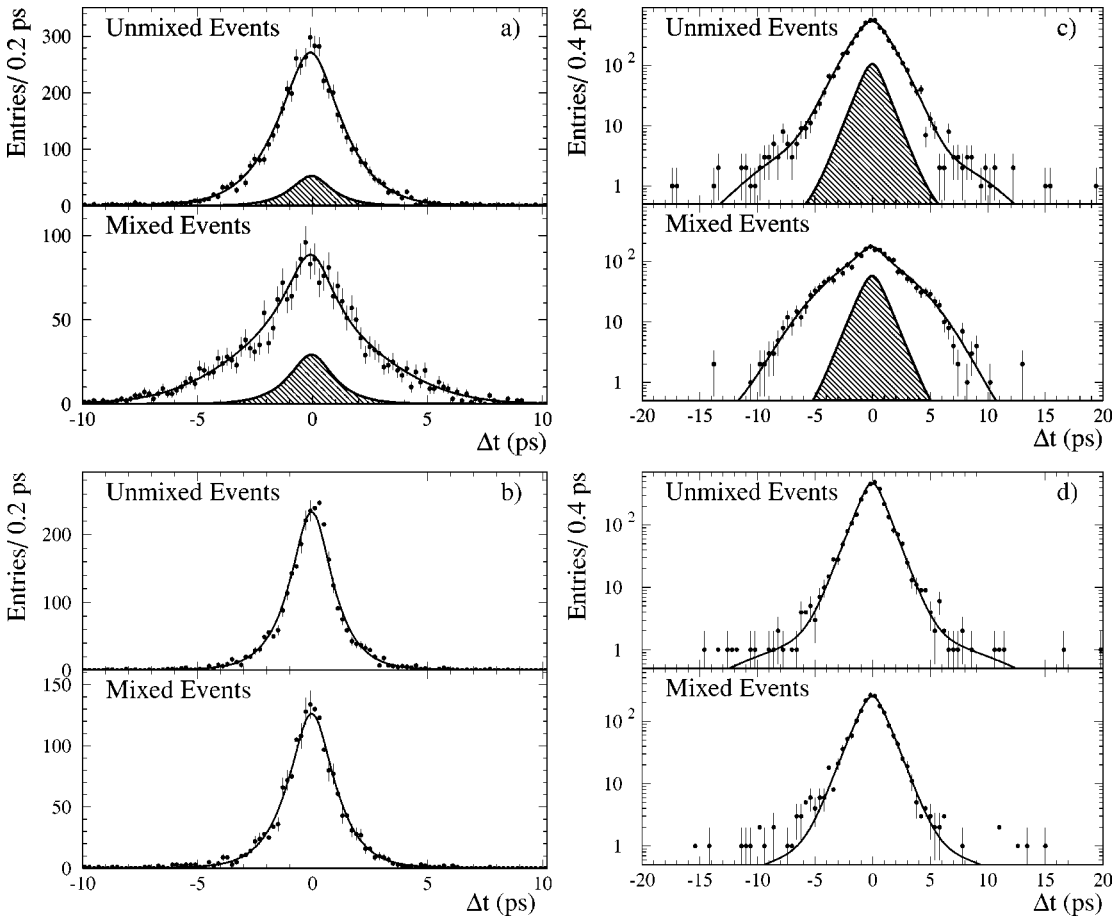

FIG. 24. Distributions of Δt for unmixed (upper panel) and mixed (lower panel) events in the hadronic B sample, divided into a signal region $m_{\mathrm{ES}}>5.27\ \mathrm{GeV}/c^2$ with a (a) linear and (c) logarithmic scale, and a sideband region $m_{\mathrm{ES}}<5.27\ \mathrm{GeV}/c^2$ with a (b) linear and (d) logarithmic scale. In all cases, the data points are overlaid with the result from the global unbinned likelihood fit, projected on the basis of the individual signal and background probabilities, and event-by-event Δt resolutions, for candidates in the respective samples. In (a) and (c), the Δt distributions obtained from the likelihood fit to the full sample are overlaid, along with the simultaneously determined background distribution shown as the curve in (b) and (d).

Δm_d, including proper treatment of all correlations. Assuming that this model is sufficiently flexible to accommodate the observed distribution in data, no additional systematic error need be assigned. The contribution to the total statistical error due to the vertex resolution can be extracted by fitting the data twice: once holding all parameters except Δm_d fixed, and once allowing the resolution function parameters to vary in addition to Δm_d. Subtracting in quadrature the respective errors on Δm_d from the two fits shows that $\pm 0.005\ \mathrm{ps}^{-1}$ of the statistical error can be attributed to the resolution parameters.

To determine the systematic error due to the assumed parametrization of the resolution model, we apply a number of possible misalignment scenarios to a sample of simulated

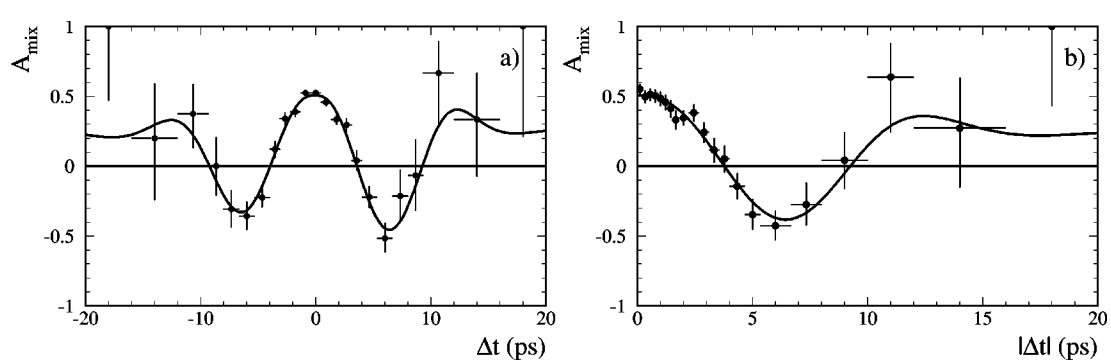

FIG. 25. Time-dependent asymmetry $\mathcal{A}(\Delta t)$ between unmixed and mixed events for hadronic B candidates with $m_{\mathrm{ES}}>5.27\ \mathrm{GeV}/c^2$; (a) as a function of Δt and (b) folded as a function of $|\Delta t|$. The asymmetry in (a) is due to the fitted bias in the Δt resolution function.

TABLE XIV. Result of fitting for Δm_d in the entire $B_{\rm flav}$ sample and in various subsamples. The difference in the fitted value of Δm_d versus the result from the fit to the full $B_{\rm flav}$ sample are reported.

Sample	$\Delta m_d - \Delta m_d(all)$
Decay mode	
$D^{*-}\pi^+$	-0.029 ± 0.030
$D^{*-}\rho^+$	$+0.017 \pm 0.039$
$D^{*-}a_1^+$	$+0.066 \pm 0.063$
$D^-\pi^+$	$+0.022 \pm 0.030$
$D^-\rho^+$	-0.031 ± 0.038
$D^-a_1^+$	-0.033 ± 0.041
$D^{*-}X$	$+0.000 \pm 0.025$
D^-X	-0.005 ± 0.023
Tagging category	
Lepton	$+0.005 \pm 0.026$
Kaon	$+0.002 \pm 0.023$
NT1	-0.032 ± 0.044
NT2	$+0.12 \pm 0.10$
$B_{\rm rec}$ state	
$B_{\rm rec} = \overline{B}^0$	$+0.015 \pm 0.023$
$B_{\rm rec} = B^0$	-0.003 ± 0.023
$B_{\rm tag}$ state	
$B_{\rm tag} = \overline{B}^0$	$+0.019 \pm 0.023$
$B_{\rm tag} = B^0$	-0.007 ± 0.022
Data sample	
Run 1	-0.012 ± 0.022
Run 2	$+0.019 \pm 0.025$

events. By comparing the value of Δm_d derived from these misaligned samples to the case of perfect alignment, we derive a systematic uncertainty of ± 0.004 ps^{-1} (a).

An additional systematic error is attributed to uncertainties in the treatment of the small fraction of Δt outliers that are the result of misreconstructed vertices. The stability of the Δm_d result is examined under variation of the width of the third Gaussian component in the resolution function between 6 and 18 ps, and through replacement of the third Gaussian with a uniform distribution and varying the width between 8 and 40 ps. On this basis, we attribute a systematic uncertainty of ± 0.002 ps^{-1} to the outlier treatment (b).

As described in detail in Sec. V A, the beam spot position is an integral part of the determination of Δt. Increasing its vertical size by up to 80 μm, and systematically biasing its vertical position by up to 80 μm, results in a corresponding variation of Δm_d by less than 0.001 ps^{-1} (c).

The requirement on the maximum allowed value of $\sigma_{\Delta t}$ is varied between 1 and 2.4 ps, and the observed variation of 0.003 ps^{-1} in Δm_d is assigned as a systematic uncertainty (d). The observed dependence is mainly due to correlations between tagging and vertexing, as described in Sec. VII B 4.

TABLE XV. Systematic uncertainties and contributions to statistical errors for Δm_d obtained with the likelihood fit to the hadronic B^0 sample.

Source	$\sigma(\Delta m_d)$ [ps^{-1}]
Signal properties	
(a) SVT alignment	0.004
(b) Δt outlier description	0.002
(c) Beam spot position or size	0.001
(d) $\sigma_{\Delta t}$ requirement	0.003
Background properties	
(e) Background fraction	0.002
(f) Background Δt structure	0.001
(g) Background Δt resolution	0.001
(h) Sideband extrapolation	0.002
(i) Peaking B^+ background	0.002
External parameters	
(j) z scale	<0.002
(k) z boost (parameters)	0.001
(l) z boost (method)	0.001
(m) B^0 lifetime	0.006
Monte Carlo studies	
(n) Signal MC statistics	0.003
(o) Tag-side D composition and lifetime	0.001
(p) Right or wrong tag resolution differences	0.001
Total systematic error	0.010
Statistical error	0.016
Contribution due to resolution function	0.005
Contribution due to mistag rate	0.005
Total error	0.019

2. Background properties

A systematic uncertainty in Δm_d arises from our ability to separate signal from background as a function of $m_{\rm ES}$. We estimate this uncertainty by varying the width and height of the fitted Gaussian peak in $m_{\rm ES}$, the slope parameter of the ARGUS background shape, and the normalizations of the signal and backgrounds by one standard deviation around their central values, resulting in an uncertainty of ± 0.002 ps^{-1} in Δm_d (e).

As discussed in Sec. VI B 1, the Δt distribution of the background is described by the combination of a prompt component and a lifetime component. To estimate the systematic uncertainty due to this choice, we add an additional component, with its own separate lifetime, that is allowed to mix; the observed value of Δm_d changes by 0.001 ps^{-1} (f). Similarly, adding an additional Gaussian distribution to the Δt background resolution function changes Δm_d by no more than 0.001 ps^{-1} (g).

Finally, the composition of the background changes slightly as a function of $m_{\rm ES}$, since the fraction of background due to continuum production slowly decreases to-

wards the B mass. As a result, the Δt structure of the background could change as well. To study this dependence, we split the $m_{\rm ES}$ sideband region into seven mutually exclusive, 10 MeV/c^2-wide intervals, and repeat the Δm_d fit with each of these slices in turn. The variation of Δm_d is then extrapolated as a function of the position of the sideband slice relative to the B mass. We correct the value of Δm_d by -0.002 ps^{-1} obtained from this extrapolation, and assign the statistical uncertainty of 0.002 ps^{-1} of this procedure as a systematic error on Δm_d (h).

A small fraction (about 1.5%) of the events attributed to the B^0 signal by the fit to the $m_{\rm ES}$ distribution consists of B^+ events, mainly due to the swapping of a soft π^0 with a charged pion as described in Sec. III C 1. The uncertainty on this peaking fraction is propagated to Δm_d, yielding a systematic error of 0.002 ps^{-1} (i).

3. External parameters

An error in the boost of the $\Upsilon(4S)$ system (0.1%) or in the knowledge of the z scale of the detector, as described in Sec. V A, could bias the Δm_d measurement because these parameters are used to reconstruct the decay length difference Δz and to convert it to the decay time difference Δt. The uncertainties on these quantities are propagated to Δm_d and lead to systematic uncertainties of 0.001 ps^{-1} (1) and less than 0.002 ps^{-1} (j), respectively. In addition to these, we assign the difference of 0.001 ps^{-1} (k) in the value of Δm_d obtained by using the Δz to Δt conversion described in Eq. (21) instead of Eq. (22) as a systematic error. Finally, in the likelihood fit, we fix the B^0 lifetime to the PDG value [11]. The present uncertainty on this value of ± 0.032 ps leads to a systematic error of ∓ 0.006 ps^{-1} (m).

4. Monte Carlo validation of measurement technique

Candidate selection criteria, or the analysis and fitting procedure, could potentially cause systematic biases in the measurement of Δm_d. These potential biases are estimated by repeating the analysis with a large sample of Monte Carlo events, which are generated with the full GEANT3 [27] detector simulation. In the Monte Carlo sample, the fitted result for Δm_d is shifted by $+0.007 \pm 0.003$ ps^{-1} from the input value. A corresponding correction with this central value is applied to the fitted result with data, and the uncertainty is assigned as a systematic error (n).

The main cause of this bias is a small correlation between the mistag rate and the Δt resolution that is not modeled in the likelihood function. This correlation is seen most readily in data for Kaon tags and is shown for simulation in Fig. 26a. We find that both the mistag rate for kaon tags and the event-by-event error $\sigma_{\Delta t}$ 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 $\sigma_{\Delta t}$, as can be seen in Fig. 26b. The mistag rate dependence originates from the kinematics of the physics sources for wrong-charge kaons. The three major sources of mistags are wrong-sign D^0 mesons from B decays to double charm, wrong-sign kaons from D^+ decays, and kaons

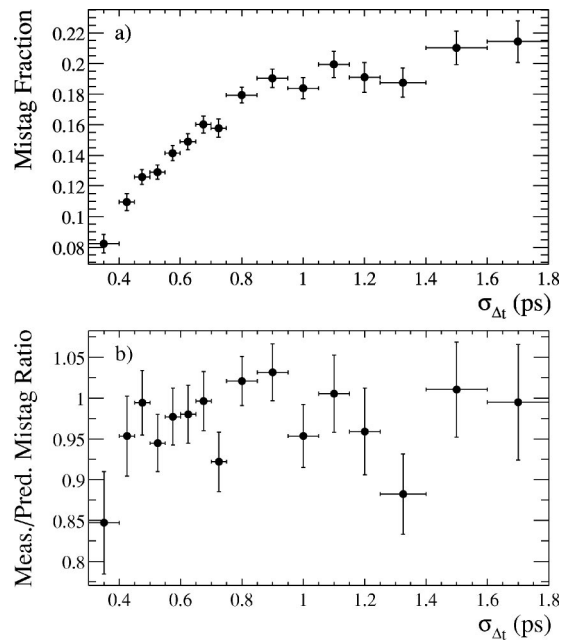

FIG. 26. (a) Correlation between the event-by-event error on Δt ($\sigma_{\Delta t}$) and the mistag rate in the Kaon category from Monte Carlo simulation; (b) dependence of the mistag rate on $\sigma_{\Delta t}$ after scaling the mistag rate by $\sqrt{\Sigma p_t^2}$.

produced directly in B decays. All these sources produce a spectrum of charged tracks that have smaller $\sqrt{\Sigma p_t^2}$ than B decays that produce a correct tag. The Δt resolution dependence originates from the $1/p_t^2$ dependence of σ_z for the individual contributing tracks.

Since the effect is small and well described by the Monte Carlo simulation, we have chosen to treat the impact of this correlation as a correction, rather than building the effect into the likelihood function. We include additional systematic errors related to the tag-side properties that could affect the accuracy of the description of this correlation in the simulation. In particular, the D^0, D^+, and D_S^+ meson branching fractions, the D meson lifetimes, and the wrong-sign kaon production rates in B meson decays are all varied. These studies lead to an assigned systematic error of ± 0.001 ps^{-1} (o).

In addition, we consider the possibility that correctly and incorrectly tagged events could have different resolution functions. Based on Monte Carlo studies of the variation in the fitted value for Δm_d with and without allowing for independent resolution functions for correctly and incorrectly tagged events, an uncertainty of ± 0.001 ps^{-1} is assigned to this source (p).

C. Validation studies and cross checks

1. Monte Carlo studies

A high-precision test of the fitting procedure was performed with fast parametrized Monte Carlo simulations, where 2000 experiments were generated with sample size and composition corresponding to that obtained from the ac-

tual data. The mistag rates and Δt distributions were generated according to the model used in the likelihood fit. The full fit was then performed on each of these experiments. The resulting distribution of pulls (defined as the difference between the fitted and generated value of a parameter divided by the statistical error as obtained from the likelihood fit) has a mean -0.038 ± 0.022 and standard deviation 1.012 ± 0.023, consistent with no measurement bias in either the value of Δm_d or its estimated error.

2. Simple counting experiment

If the mistag rate is known, the time-integrated fraction $\chi_d=\frac{1}{2}x_d/(1+x_d^2)$ of mixed events can be determined from the $B_{\rm flav}$ sample by counting mixed and unmixed events. The value for χ_d obtained by this means, after correcting for the mistag rates obtained from the full time-dependent fit and assuming the Particle Data Group (PDG) value for the B^0 lifetime, leads to a value of $\Delta m_d=x_d/\tau_{B^0}$ that differs from the likelihood-fit result by $-0.003\pm0.013\ \mathrm{ps}^{-1}$, where the quoted error is the difference in quadrature of the statistical errors of both measurements.

Because of the choice for normalization of the likelihood $\mathcal{L}_{\rm mix}$, the time-integrated ratio of the number of mixed to unmixed events contributes to our measurement of Δm_d. Alternatively, it is possible to normalize the likelihoods of mixed and unmixed events individually, in which case Δm_d is determined solely from the shape of Δt distributions. The value of Δm_d determined by the Δt distributions alone differs from the full measurement by $0.003\pm0.015\ \mathrm{ps}^{-1}$, where the quoted error is given by the difference in quadrature of the statistical errors of the two measurements.

3. Cross-check with τ_{B^0}

If we allow the value of τ_{B^0} to float in the Δm_d fit the value of Δm_d increases by $0.008\pm0.007\ \mathrm{ps}^{-1}$, and the lifetime is found to be 1.51 ± 0.03 ps, consistent with our recent measurement [17]. We have also performed a series of fits with fixed values for Δm_d and τ_{B^0} in order to determine the dependence of Δm_d on τ_{B^0}, which is found to be

$$\Delta m_d=\left[0.516-0.279\left(\frac{\tau_{B^0}}{1.548\ \mathrm{ps}}-1\right)\right]\mathrm{ps}^{-1}. \qquad (43)$$

VIII. *CP* VIOLATION IN NEUTRAL *B* DECAYS

A. Likelihood fit results for $\sin 2\beta$

The value of $\sin 2\beta$, the dilution factors $\mathcal{D}_i$, the Δt resolution parameters $\hat{a}_i$, and the background fractions and time distribution parameters are extracted with an unbinned maximum-likelihood fit to the flavor-eigenstate $B_{\rm flav}$ and B_{CP} samples as described in Sec. VI. We also demand a valid tag and Δt determination for the event, based on the algorithms described in Secs. IV and V. The looser requirements $|\Delta t|<20$ ps and $\sigma_{\Delta t}<2.4$ ps are applied to the proper time difference measurement. The fit results are summarized in Table XVI together with the correlation of the parameters with $\sin 2\beta$. The mistag fractions and vertex parameters are predominantly determined by the $B_{\rm flav}$ sample. The *CP* asymmetry and parameters describing the background for the *CP* events are determined by the *CP* sample. The value of $\sin 2\beta$ obtained from the combined $\eta_{CP}=-1$, $\eta_{CP}=+1$, and $J/\psi K^{*0}$ *CP* samples is

$$\sin 2\beta=0.59\pm0.14\pm0.05,$$

where the first error is statistical and the second systematic.

The $m_{\rm ES}$ distribution for events in $\eta_{CP}=-1$ modes, separated into tagging categories, is shown in Fig. 27. The signal probability $f_{i,\rm sig}^{CP}(m_{\rm ES})$ for the $\eta_{CP}=-1$ sample is determined from these fits as described in Sec. VI B 1.

Table XVII summarizes the event yields and $\sin 2\beta$ values determined for the full *CP* sample and various subsamples. Results are provided by *CP* channel, tagging category, B^0 versus $\overline{B}^0$ tag, J/ψ decay mode and data-taking period. The consistency between the six *CP* modes is satisfactory, the probability of finding a worse agreement being 8%. The large observed asymmetry in $B^0\to\chi_{c1}K_S^0$ causes the likelihood for this channel to become negative in certain regions of Δt. The likelihood of each of the selected candidates is of course positive. Fast parametrized Monte Carlo studies show that $\sin 2\beta$ is unbiased if the likelihood is not required to be positive for all values of Δt and that the probability to measure such a large asymmetry is about 1%. The observed asymmetry in the number of B^0 (160) and $\overline{B}^0$ (113) tags in the $J/\psi K_L^0$ sample has no impact on the $\sin 2\beta$ measurement. The results obtained with the full B_{CP} samples for run 1 and run 2 are consistent at the 1.8 sigma level. The yields and fitted values for $\sin 2\beta$ are also listed in Table XVII for the high purity $\eta_{CP}=-1$ sample alone, along with a similar breakdown into subsamples; again, no significant variation is seen.

The distribution of events as a function of Δt for B^0 and $\overline{B}^0$ tags is shown in Fig. 28a for the full *CP* sample. For this purpose, only those events with $m_{\rm ES}>5.27\ \mathrm{GeV/c^2}$ in the $\eta_{CP}=-1$ and $J/\psi K^{*0}$ samples or $\Delta E<10$ MeV in the $\eta_{CP}=+1$ sample are included. Overlaid on the data are the projections of the signal and background Δt distributions obtained from the fit, where the latter is normalized to the projected background level. Figures 28b–28d show the corresponding Δt distributions for the $\eta_{CP}=-1$, $\eta_{CP}=+1$ samples and $B^0\to J/\psi K^{*0}$ ($K^{*0}\to K_S^0\pi^0, K_S^0\to\pi^+\pi^-$). The superimposed likelihood curves show the quality of the fit for each subsample. The value of $\sin 2\beta$ obtained by fixing all other parameters to results obtained with the full *CP* sample and then fitting for $\sin 2\beta$ in bins of Δt is shown in Fig. 29a. The values obtained for $\sin 2\beta$ are all consistent, demonstrating that the oscillation as a function of Δt has the expected behavior. The observed asymmetry $\mathcal{A}(\Delta t)$ is shown in Figs. 29b and 29c for the $\eta_{CP}=-1$ and $\eta_{CP}=+1$ samples, respectively, along with the projections from the fit results.

The average dilutions and dilution differences for B^0 and $\overline{B}^0$ tags obtained from the fit to the B^0 flavor eigenstate and full *CP* sample, and the corresponding tagging efficiencies, are summarized in Table XVIII. We find a total tagging effi-

TABLE XVI. Parameters for the combined likelihood fit to the B_{CP} and $B_{\rm flav}$ samples. The first major column contains the fit results, while the second major column contains the correlation coefficients with respect to $\sin 2\beta$ for each fit parameter.

	Fit result		Correlation	
Parameter	Run 1	Run 2	Run 1	Run 2
$\sin 2\beta$	0.59 ± 0.14			
Signal resolution function				
S_1 (core)	1.2 ± 0.1	1.1 ± 0.1	0.018	0.020
$b_1(\Delta t)$ `lepton` (core)	0.07 ± 0.12	0.04 ± 0.16	0.008	0.045
$b_1(\Delta t)$ `kaon` (core)	-0.26 ± 0.08	-0.18 ± 0.09	0.002	0.021
$b_1(\Delta t)$ `NT1` (core)	-0.21 ± 0.15	-0.33 ± 0.21	0.004	0.001
$b_1(\Delta t)$ `NT2` (core)	-0.31 ± 0.11	-0.17 ± 0.15	−0.001	−0.002
$b_2(\Delta t)$ (tail)	-1.7 ± 1.5	-3.3 ± 2.8	0.001	0.006
f_2 (tail)	0.08 ± 0.06	0.04 ± 0.04	0.009	0.005
f_3 (outlier)	0.005 ± 0.003	0.000 ± 0.001	−0.001	0.000
Signal dilutions				
$\langle D\rangle$, `lepton`	0.82 ± 0.03		−0.042	
$\langle D\rangle$, `kaon`	0.65 ± 0.02		−0.083	
$\langle D\rangle$, `NT1`	0.56 ± 0.04		−0.015	
$\langle D\rangle$, `NT2`	0.30 ± 0.04		−0.032	
ΔD, `lepton`	-0.02 ± 0.04		0.010	
ΔD, `kaon`	0.04 ± 0.03		0.005	
ΔD, `NT1`	-0.11 ± 0.06		0.014	
ΔD, `NT2`	0.12 ± 0.05		−0.008	
Background properties				
τ, mixing bkgd [ps]	1.3 ± 0.1		−0.001	
$f(\tau=0)$, CP bkgd	0.60 ± 0.12		−0.011	
$f(\tau=0)$, mixing bkgd, `lepton`	0.31 ± 0.10		−0.001	
$f(\tau=0)$, mixing bkgd, `kaon`	0.65 ± 0.04		−0.001	
$f(\tau=0)$, mixing bkgd, `NT1`	0.62 ± 0.06		−0.001	
$f(\tau=0)$, mixing bkgd, `NT2`	0.64 ± 0.04		−0.001	
Background resolution function				
S_1 (core)	1.5 ± 0.1	1.3 ± 0.1	0.004	−0.003
$b_1(\Delta t)$ core [ps]	-0.16 ± 0.03	0.02 ± 0.04	0.000	−0.001
f_2(outlier)	0.016 ± 0.004	0.017 ± 0.005	−0.001	0.000
Background dilutions				
$\langle D\rangle$, `lepton`, $\tau=0$	0.33 ± 0.27		0.003	
$\langle D\rangle$, `kaon`, $\tau=0$	0.45 ± 0.03		0.008	
$\langle D\rangle$, `NT1`, $\tau=0$	0.25 ± 0.10		0.002	
$\langle D\rangle$, `NT2`, $\tau=0$	0.11 ± 0.06		0.003	
$\langle D\rangle$, `lepton`, $\tau>0$	0.33 ± 0.14		0.000	
$\langle D\rangle$, `kaon`, $\tau>0$	0.24 ± 0.06		0.000	
$\langle D\rangle$, `NT1`, $\tau>0$	0.05 ± 0.14		−0.001	
$\langle D\rangle$, `NT2`, $\tau>0$	0.09 ± 0.09		0.000	

ciency of $(68.4\pm0.7)\%$ (statistical error only). The lepton categories have the lowest mistag fractions, but also low efficiency. The Kaon category, despite having a larger mistag fraction (17.6%), has a higher effective tagging efficiency; one-third of events are assigned to this category. Altogether, lepton and kaon categories have an effective tagging efficiency $Q\approx22.4\%$. The neural network categories increase the effective tagging efficiency by ~4% to an overall $Q=(26.1\pm1.2)\%$ (statistical error only). These mistag fractions are very similar to the mistag fractions that are obtained

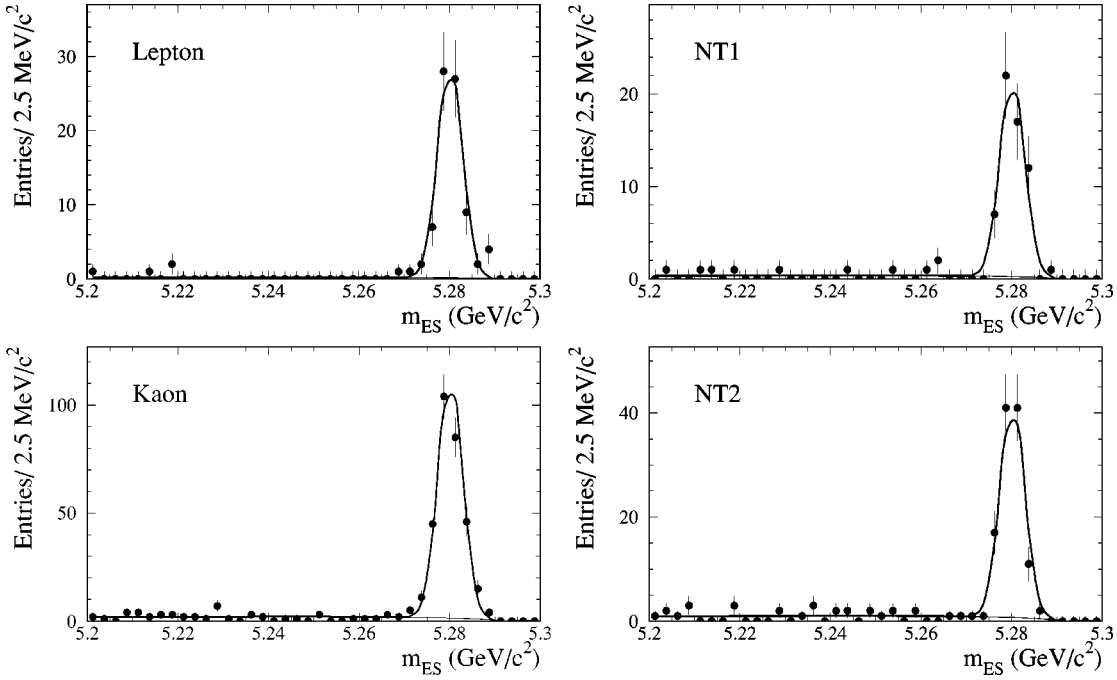

FIG. 27. Distribution of $m_{\rm ES}$ for $\eta_{CP}=-1$ candidates in separate tagging categories (`Lepton`, `Kaon`, `NT1` and `NT2`), overlaid with the result of a fit with a Gaussian distribution for the signal and an ARGUS function for the background.

from the $B_{\rm flav}$ sample alone (see Table XIII). The small differences are due to the correlation between the mistag fractions and the Δt resolution function parameters.

Based on a large number of fast parametrized Monte Carlo experiments with the same number of events as our full B_{CP} and $B_{\rm flav}$ data samples, we estimate a probability of 27% for finding a value of the maximum likelihood lower than that observed. These same studies, based on samples with the size and composition of the data, show that the expected statistical error is 0.132 with a spread of 0.005, in very good agreement with the observed error of 0.137.

B. Systematic error estimation

Just as for the Δm_d measurement, systematic errors can usefully be grouped into signal description, including detector reconstruction effects, background description, fixed external parameters, and statistical limitations of Monte Carlo validation tests for the fitting procedure (discussed in Sec. VIII C 1). A summary of these sources of systematic error is shown in Table XIX for the various *CP* samples. In the following, the individual contributions are referenced by the lettered lines in this table.

1. Signal properties and description

The parameters of the Δt resolution function, the dilutions and dilution differences are determined from the data sample itself with the likelihood fit. Thus, they do not contribute to the systematic error, but rather are incorporated into the statistical uncertainty at a level determined by the size of the data sample itself. Their overall contribution to the total error on $\sin 2\beta$ is 0.02, as determined from the difference in quadrature between the statistical error on $\sin 2\beta$ from the full likelihood fit and from a fit with only $\sin 2\beta$ allowed to vary.

While the bulk of the uncertainties from these sources is thus incorporated into the statistical error, we assign additional systematic uncertainties due to the fixed form of the parametrization for the Δt resolution function. This form may not be flexible enough to account for all possible effects. In addition, tests of the assumption that the resolution function and dilution parameters are the same for the $B_{\rm flav}$ and B_{CP} samples are limited in precision by the size of the available Monte Carlo samples.

The resolution function, described in Sec. V D, is one of several possible functional forms. In order to test possible biases induced by this particular choice, an alternative model has been considered where a Gaussian distribution is convolved with an exponential, with the effective lifetime in the exponential depending on the tagging category. No difference between the fit results with the two models is observed in Monte Carlo simulation. We assign as a systematic uncertainty the difference in the fit results observed in the data (Table XIX, line a). The largest systematic uncertainties from the Δt behavior arise from possible effects that our model of the resolution function cannot accommodate or completely parametrize. These include residual uncertainties in the SVT alignment (b) and possible differences in the Δt determination for correctly and incorrectly tagged events (c). An additional uncertainty is assigned due to the treatment of the Δt outliers (d). Fits with Monte Carlo samples of $B_{\rm flav}$ and B_{CP} signal events show no significant difference between resolution function parameters for the two samples. We assign a systematic uncertainty of ± 0.003 due to the residual shift in

TABLE XVII. Result of fitting for CP asymmetries in the entire CP sample and in various subsamples. The tagged number of events $N_{\rm tag}$ includes signal and background, while the purities are obtained from separate fits to $m_{\rm ES}$ or ΔE distributions.

Sample	$N_{\rm tag}$	Purity (%)	$\sin 2\beta$	$\mathrm{Im}\,\lambda/\lvert\lambda\rvert$	$\lvert\lambda\rvert$
CP sample	803	80	**0.59±0.14**		
Decay mode					
$J/\psi K^0_S$ ($K^0_S\to\pi^+\pi^-$)	316	98	0.45 ± 0.18	0.45 ± 0.18	0.91 ± 0.11
$J/\psi K^0_S$ ($K^0_S\to\pi^0\pi^0$)	64	94	0.70 ± 0.50	0.71 ± 0.50	0.95 ± 0.27
$\psi(2S)K^0_S$ ($K^0_S\to\pi^+\pi^-$)	67	98	0.47 ± 0.42	0.48 ± 0.45	1.22 ± 0.33
$\chi_{c1}K^0_S$ ($K^0_S\to\pi^+\pi^-$)	33	97	$2.59\pm^{0.55}_{0.67}$	2.67 ± 0.59	0.71 ± 0.23
$J/\psi K^0_L$	273	51	0.70 ± 0.34		
$J/\psi K^{*0}$ ($K^{*0}\to K^0_S\pi^0$, $K^0_S\to\pi^+\pi^-$)	50	74	0.82 ± 1.00		
Tagging category					
`Lepton`	130	82	0.54 ± 0.26		
`Kaon`	438	79	0.58 ± 0.18		
`NT1`	79	74	0.89 ± 0.30		
`NT2`	156	80	0.40 ± 0.65		
$B_{\rm tag}$ state					
B^0	420	79	0.54 ± 0.19		
$\overline{B}^0$	383	78	0.64 ± 0.20		
J/ψ mode					
$J/\psi\to e^+e^-$	385	78	0.45 ± 0.18		
$J/\psi\to\mu^+\mu^-$	418	84	0.70 ± 0.18		
Data sample					
Run 1	533	80	0.49 ± 0.20		
Run 2	270	84	0.82 ± 0.22		
$\eta_{CP}=-1$ sample	480	96	0.56 ± 0.15	0.56 ± 0.15	0.93 ± 0.09
Tagging category					
`Lepton`	74	100	0.54 ± 0.29	0.57 ± 0.29	0.77 ± 0.14
`Kaon`	271	98	0.59 ± 0.20	0.59 ± 0.20	0.98 ± 0.12
`NT1`	46	97	0.67 ± 0.45	0.57 ± 0.46	0.73 ± 0.29
`NT2`	89	95	0.10 ± 0.74	0.28 ± 1.29	2.95 ± 3.83
$B_{\rm tag}$ state					
B^0	234	98	0.50 ± 0.22		
$\overline{B}^0$	246	97	0.61 ± 0.22		
J/ψ mode					
$J/\psi\to e^+e^-$	219	94	0.54 ± 0.22	0.52 ± 0.22	1.00 ± 0.15
$J/\psi\to\mu^+\mu^-$	261	98	0.60 ± 0.21	0.63 ± 0.21	0.87 ± 0.11
Data sample					
Run 1	310	95	0.37 ± 0.20	0.37 ± 0.20	1.16 ± 0.15
Run 2	170	98	0.86 ± 0.24	0.96 ± 0.26	$0.66^{-0.11}_{+0.12}$
Control samples					
$B^0\to D^{(*)-}\pi^+/\rho^+/a_1^+$	7579	84	0.00 ± 0.04		
$B^+\to\overline{D}^{(*)0}\pi^+$	6800	86	-0.02 ± 0.04		
$B^0\to J/\psi K^{*0}$ ($K^{*0}\to K^+\pi^-$)	705	95	0.12 ± 0.12		
$B^+\to J/\psi K^{(*)+}$, $\psi(2S)K^+$	2031	94	0.07 ± 0.07		

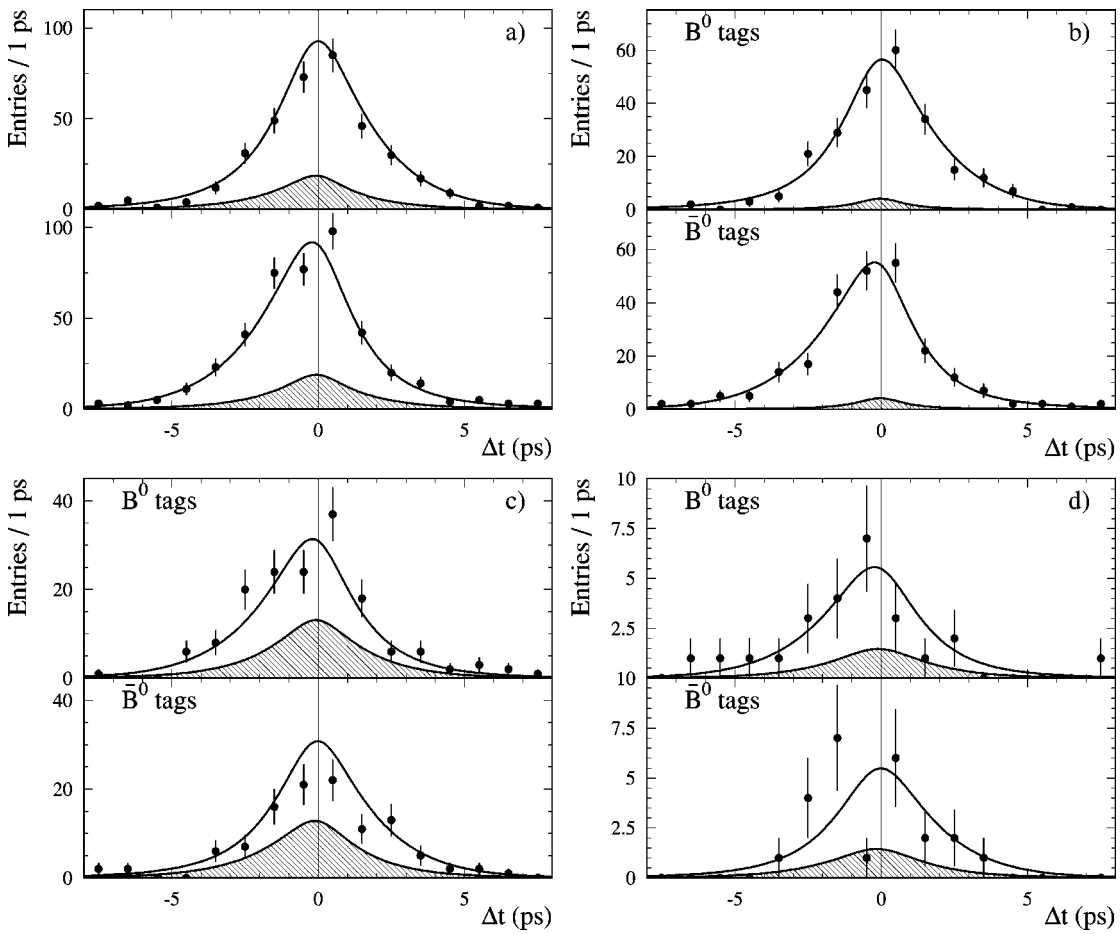

FIG. 28. (a) Distribution of Δt for tagged events in the full *CP* sample. The upper (lower) panel is the sum of B^0 ($\overline{B}^0$)-tagged events in the $\eta_{CP}=-1$ and $J/\psi K^{*0}$ samples, combined with the $\overline{B}^0$ (B^0)-tagged events in the $J/\psi K_L^0$ sample. Corresponding B^0- (lower panel) and $\overline{B}^0$-tagged (upper panel) distributions for the (b) $\eta_{CP}=-1$, (c) $J/\psi K_L^0$, and (d) $J/\psi K^{*0}$ samples are also shown. In all cases, the data points are overlaid with the result from the global unbinned likelihood fit, projected on the basis of the individual signal and background probabilities, and event-by-event Δt resolutions, for candidates in the respective samples. Therefore, the curves correspond to $\sin 2\beta = 0.59$, rather than the fitted value obtained with the individual subsample. The probability-weighted Δt spectra of the background candidates obtained from the fit are indicated by the shaded areas.

$\sin 2\beta$ between the two sets of fitted Δt resolution parameters (e).

An underlying assumption of the global fit is that dilutions and dilution differences are the same for the $B_{\rm flav}$ and B_{CP} samples. We assign the full difference as seen in Monte Carlo simulation as systematic error, ± 0.027 (f). Monte Carlo studies of lepton-tagged $J/\psi K_L^0$ events show that the resolution for the missing transverse momentum is broadened by the neutrino from the semileptonic $B_{\rm tag}$ decays. This leads to a small loss of efficiency but no significant bias in the mistag rates relative to the $B_{\rm flav}$ sample. In addition, the $B^{\pm}$ data sample was used to study any possible dependence of the dilutions on Δt. No significant effect was observed. However, a dependence of the dilutions on $\sigma_{\Delta t}$ has been seen, both in data and the Monte Carlo simulation (see Sec. VII B 4). Finally, it is possible that tagging efficiencies could be different for B^0 and $\overline{B}^0$ mesons. A separate study of the relative tagging efficiencies is described in Sec. VIII D, since the relative efficiencies form an important part of the direct *CP* violation search. The systematic error on $\sin 2\beta$ due to this effect is estimated to be ± 0.003 (g).

2. Background properties

The fraction of background events in the $\eta_{CP}=-1$ sample is estimated from fits to the $m_{\rm ES}$ distribution. Varying this fraction within the stated errors and changing the signal probability as a function of $m_{\rm ES}$ results in a systematic error of ± 0.005 on $\sin 2\beta$ (h). The uncertainty on the fraction of peaking background contributes a systematic error of ± 0.003 (i). Varying the effective $\sin 2\beta$ assumed for the ARGUS ($\mathcal{A}$ in Sec. VI B 1) and peaking ($\delta_{\rm peak}$ in Sec. VI B 1) backgrounds in the *CP* sample from -1 to $+1$ contributes a systematic error of ± 0.015 (j) and ± 0.004 (k), respectively. In addition, the contributions due to the uncertainty of the Δt

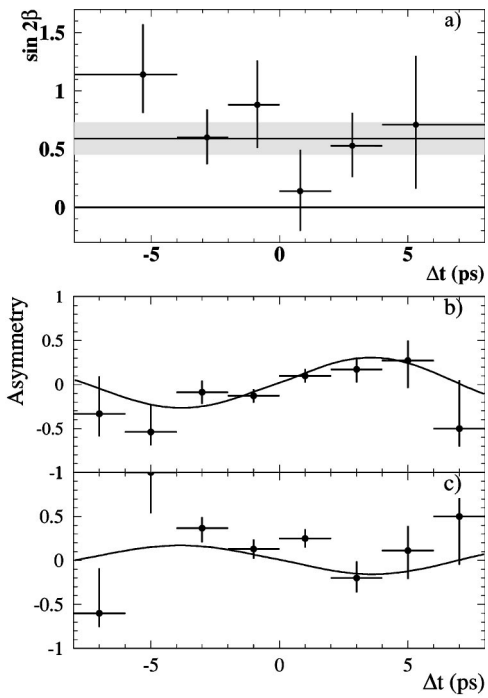

FIG. 29. (a) Fitted value of $\sin 2\beta$ obtained in bins of Δt by fixing all other parameters to the values obtained with the full *CP* sample; raw asymmetry in the number of B^0 and $\overline{B}^0$ tags in the signal region, $(N_{B^0}-N_{\overline{B}^0})/(N_{B^0}+N_{\overline{B}^0})$, with asymmetric binomial errors, as a function of Δt for (b) $\eta_{CP}=-1$ and (c) $J/\psi K_L^0$ samples. The data points are overlaid with the separate fit results for the two samples.

resolution model (± 0.002), and the effective lifetime (negligible) of the *CP* background, have been evaluated (l,m).

For the $B^0 \to J/\psi K_L^0$ channel, the signal and non-J/ψ background fractions are varied within their statistical uncertainties ($\pm 1\sigma$) as obtained with the fit to the ΔE distribution of the sample. This contributes a systematic error of ± 0.075 to the $B^0 \to J/\psi K_L^0 \sin 2\beta$ result and ± 0.01 to the final result (n). We also vary background parameters for the $B^0 \to J/\psi K_L^0$ sample, including the $J/\psi X$ branching fractions according to Table VI, the assumed η_{CP}, the mistag rates and efficiencies, the Δt resolution function, and ΔE shape (o–r). The total $B^0 \to J/\psi K_L^0$ background systematic error, summing these contributions in quadrature, is ± 0.09 for the $\sin 2\beta$ fit to the $B^0 \to J/\psi K_L^0$ sample alone and ± 0.013 for the full sample.

For the $B^0 \to J/\psi K^{*0}$ $(K_S^0 \pi^0)$ sample, the value of $R_\perp$ as well as the sample composition are varied (s,t) according to Table V.

The effect of the uncertainty on background component in the $B_{\rm flav}$ sample on $\sin 2\beta$ has also been evaluated. The only significant sources of uncertainty are the fraction of background that mixes (v) and the signal probability distribution as a function of $m_{\rm ES}$ (u,w).

3. External parameters

The residual uncertainty on the physical z scale (x) and the boost parameters of the $\Upsilon(4S)$ center of mass (y) contribute systematic uncertainties. We fix the B^0 lifetime to the current world average values $\tau_{B^0}=1.548$ ps and $\Delta m_d = 0.472\ \mathrm{ps}^{-1}$ [11]. The errors on $\sin 2\beta$ due to uncertainties in τ_{B^0} and Δm_d are ± 0.009 and ± 0.013, respectively (z–aa).

4. Monte Carlo validation of measurement technique

The analysis method has been studied with a high-statistics Monte Carlo sample. A fit result that is consistent with the generated value for $\sin 2\beta$ was found. We assign a ± 0.012 systematic error due to the statistical limitation of the Monte Carlo sample size (bb). Section VIII C 1 describes this study in more detail.

C. Validation studies and cross checks

We have used data and Monte Carlo samples to perform validation studies of the analysis technique. These tests include studies with parameterized Monte Carlo samples, full GEANT3 [27] simulation samples, as well as data samples where no *CP* asymmetry is expected.

1. Monte Carlo studies

The highest precision test of the fitting procedure was performed with fast parametrized Monte Carlo simulation,

TABLE XVIII. Average mistag fractions w_i and mistag differences $\Delta w_i = w_i(B^0) - \psi_i(\overline{B}^0)$ extracted for each tagging category i from the maximum-likelihood fit to the time distribution for the fully reconstructed B^0 sample $(B_{\rm flav}+B_{CP})$. The figure of merit for tagging is the effective tagging efficiency $Q_i = \varepsilon_i(1-2w_i)^2$, where ε_i is the fraction of events with a reconstructed tag vertex that are assigned to the ith category. ε_i is computed for the $\eta_{CP}=\pm 1$ samples as well as the combined B_{CP} and $B_{\rm flav}$ samples. Uncertainties are statistical only. The statistical error on $\sin 2\beta$ is proportional to $1/\sqrt{Q}$, where $Q=\Sigma Q_i$.

Category	$\eta_{CP}=-1$ ε [%]	$\eta_{CP}=+1$ ε [%]	$B_{\rm flav}+B_{CP}$ ε [%]	w [%]	Δw [%]	Q [%]
Lepton	11.0 ± 1.2	10.4 ± 3.0	10.9 ± 0.3	9.0 ± 1.4	0.9 ± 2.2	7.4 ± 0.5
Kaon	38.9 ± 1.9	28.3 ± 4.5	35.8 ± 1.0	17.6 ± 1.0	-1.9 ± 1.5	15.0 ± 0.9
NT1	6.9 ± 0.9	4.8 ± 2.3	7.8 ± 0.3	22.0 ± 2.1	5.6 ± 3.2	2.5 ± 0.4
NT2	13.0 ± 0.4	13.9 ± 3.3	13.8 ± 0.3	35.1 ± 1.9	-5.9 ± 2.7	1.2 ± 0.3
All	69.8 ± 2.7	57.4 ± 6.7	68.4 ± 0.7			26.1 ± 1.2

TABLE XIX. Summary of contributions to the systematic error on $\sin 2\beta$, $\mathrm{Im}\,\lambda/|\lambda|$ and $|\lambda|$. Note that the last two measurements use only the $\eta_{CP}=-1$ sample.

	CP sample					
Source	$\eta_{CP}=-1$	$J/\psi K_L^0$	$J/\psi K^{*0}$	Full	$\mathrm{Im}\,\lambda/\lvert\lambda\rvert$	$\lvert\lambda\rvert$
Signal properties						
(a) Δt signal resolution model	±0.009	±0.01	±0.07	±0.009	±0.003	±0.003
(b) SVT alignment	±0.027	±0.027	±0.027	±0.027	±0.027	±0.012
(c) Δt for right or wrong tagged events	±0.012	±0.012	±0.012	±0.012	±0.011	±0.003
(d) Δt signal resolution outliers	±0.002	±0.018	±0.03	±0.002	±0.003	±0.002
(e) Δt signal resolution	±0.003	±0.003	±0.003	±0.003	±0.003	±0.009
(f) Signal dilutions for *CP* vs B_{flav}	±0.027	±0.027	±0.027	±0.027	±0.027	±0.011
(g) Tagging efficiencies	±0.003	±0.003	±0.003	±0.003	±0.004	±0.012
Background properties: $\eta_{CP}=-1$						
(h) Background fraction	±0.006	—	—	±0.005	±0.006	±0.004
(i) *CP* bkgd peaking component	±0.004	—	—	±0.003	±0.005	±0.001
(j) *CP* bkgd *CP* content (ARGUS)	±0.015	—	—	±0.015	±0.015	±0.001
(k) *CP* bkgd *CP* content (Peak)	±0.004	—	—	±0.004	±0.004	±0.001
(l) *CP* bkgd effective lifetime	0	—	—	0	0	0
(m) *CP* bkgd resolution	±0.002	—	—	±0.002	±0.002	±0.001
Background properties: $J/\psi K_L^0$						
(n) Background fraction	—	±0.075	—	±0.01	—	—
(o) ΔE distribution	—	±0.04	—	±0.007	—	—
(p) Effective *CP* of backgrounds	—	±0.020	—	±0.001	—	—
(q) Background composition	—	±0.014	—	±0.002	—	—
(r) Background Δt and dilution	—	±0.023	—	±0.003	—	—
Background properties: $J/\psi K^{*0}$						
(s) Sample composition	—	—	±0.08	±0.001	—	—
(t) $R_\perp$	—	—	±0.08	±0.001	—	—
Background properties: B_{flav}						
(u) Background fraction	±0.001	±0.008	±0.003	±0.002	±0.002	±0.001
(v) B_{flav} bkgd mixing contrib.	±0.001	±0.002	±0.001	±0.002	±0.001	0
(w) B_{flav} bkgd peaking component	0	±0.001	±0.001	0	0	0
External parameters						
(x) z scale and boost	±0.003	±0.003	±0.003	±0.003	±0.003	±0.001
(y) Beam spot	±0.002	±0.002	±0.002	±0.002	±0.002	±0.006
(z) B^0 lifetime	±0.008	±0.011	±0.022	±0.009	±0.009	±0.012
(aa) Δm_d	±0.015	±0.012	±0.082	±0.013	±0.015	±0.001
Monte Carlo studies						
(bb) Monte Carlo statistics	±0.012	±0.012	±0.012	±0.012	±0.012	±0.007
Total systematic error	±0.05	±0.10	±0.16	±0.05	±0.05	±0.02
Statistical error	±0.15	±0.34	±1.01	±0.14	±0.15	±0.09

where 1000 experiments were generated with sample sizes corresponding to the observed B_{flav} and B_{CP} events in data, including mistag rates, Δt resolutions, and background fractions and time dependence. The full fit is performed on each of these experiments. The resulting pull distribution (defined as the difference between the fitted and generated value of a parameter divided by the statistical error as obtained from the likelihood fit) has a mean -0.029 ± 0.032 and standard deviation 1.007 ± 0.022, consistent with no measurement bias in either the value of $\sin 2\beta$ or its estimated error.

In addition, large samples of signal and background Monte Carlo events generated with a GEANT3 [27] detector simulation are used to validate the measurement. For these tests, we obtained the resolution function parameters as well as the dilutions from a Monte Carlo sample of B_{flav} events. Using these parameters, we fit for $\sin 2\beta$ in

TABLE XX. Results when Δm_d and (or) τ_{B^0} are floated in the $\sin 2\beta$ fit to the full CP sample and the $\eta_{CP}=-1$ subsample alone.

Fit	$\sin 2\beta$	Δm_d (ps^{-1})	τ_{B^0} (ps)
All CP modes			
Nominal fit	0.59 ± 0.14	0.472	1.548
Float Δm_d	0.55 ± 0.13	0.533 ± 0.015	1.548
Float τ_{B^0}	0.60 ± 0.14	0.472	1.53 ± 0.03
Float Δm_d and τ_{B^0}	0.56 ± 0.13	0.542 ± 0.016	1.50 ± 0.03
$\eta_{CP}=-1$ modes			
Nominal fit	0.56 ± 0.15	0.472	1.548
Float Δm_d	0.51 ± 0.15	0.531 ± 0.015	1.548
Float τ_{B^0}	0.57 ± 0.15	0.472	1.53 ± 0.03
Float Δm_d and τ_{B^0}	0.52 ± 0.15	0.540 ± 0.016	1.50 ± 0.03

Monte Carlo samples of CP signal events that correspond in number to the reconstructed data sample. These Monte Carlo events are generated with various values of $\sin 2\beta$ (0.1 to 0.9) and different CP-eigenstate modes, corresponding to those used in the measurement with data. The mean and spread of the pull distribution for these Monte Carlo samples can be used to check for any measurement bias and to confirm the validity of the reported error. We find that the mean pull is consistent with zero and the spread is consistent with the reported error. A systematic error of ±0.012 is assigned to $\sin 2\beta$ due to the limited Monte Carlo statistics for this test.

The effect of background has been evaluated by adding an appropriate fraction of background events to our signal Monte Carlo sample and performing the likelihood fit. The background samples are obtained either from simulated $B\to J/\psi X$ events or ΔE sidebands in data ($|\Delta E|<120$ MeV but outside the signal region). We find no significant bias for $\sin 2\beta$ with the addition of either source of background.

2. Cross checks with τ_{B^0} and Δm_d

Table XX shows results for $\sin 2\beta$ if Δm_d and τ_{B^0} are

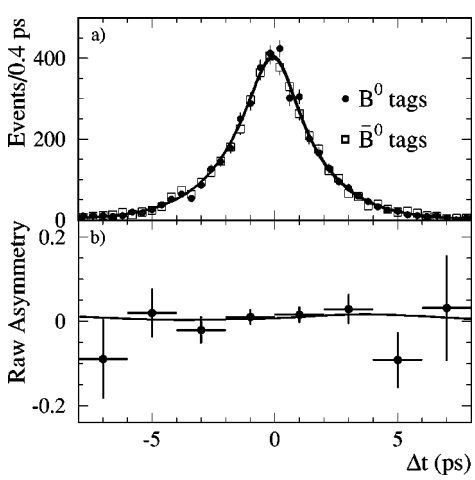

FIG. 30. (a) Distribution in Δt for B^0- and $\overline{B}^0$-tagged samples and (b) observed asymmetry for events in the flavor-eigenstate B^0 sample. The projections of the likelihood fit for the B^0- and $\overline{B}^0$-tagged samples are shown in (a) as the overlapping solid lines.

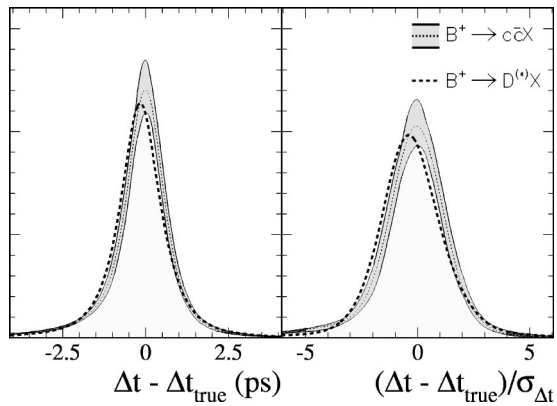

FIG. 31. Comparison of the fitted Δt resolution obtained with the data control samples $B^+\to D^{(*)}X$ and $B^+\to c\bar{c}X$, showing the fitted distribution for (a) $\delta_t=\Delta t-\Delta t_{\rm true}$ and (b) the normalized difference $\delta_t/\sigma_{\Delta t}$. The one sigma error envelope from the fit to the $B^+\to c\bar{c}X$ sample (shaded region), overlaps the central value for the five times larger $B^+\to D^{(*)}X$ sample (dashed line).

allowed to float in the combined fit to the CP and $B_{\rm flav}$ samples. The fitted value of Δm_d is somewhat larger than that reported in Sec. VII. However, with no kaon veto applied to the tagging vertex, the correction for the bias introduced by known correlations between mistag rates and the Δt resolution is also larger. Taking this into account, the two results are consistent within the independent statistical errors. Likewise the lifetime is found to be consistent with our recent measurement [17]. We have also performed fits with Δm_d and τ_{B^0} fixed to a series of values around the world average in order to determine the dependence of $\sin 2\beta$ on these two parameters, thereby finding that

$$\sin 2\beta=\left[0.59-0.35\left(\frac{\Delta m_d}{0.472\ \text{ps}^{-1}}-1\right)-0.45\left(\frac{\tau_{B^0}}{1.548\ \text{ps}}-1\right)\right]. \tag{44}$$

3. Asymmetries in data control samples

Control samples in data where the reconstructed B^0 and B^+ meson decays to a flavor-eigenstate mode with a $D^{(*)}$ or charmonium meson in the final state can be used to validate the $\sin 2\beta$ measurement, since the asymmetry is expected to be zero in this case. For these samples, the Δt resolution function parameters and the dilutions are fixed to the values obtained with the $B_{\rm flav}$ sample. The CP asymmetry and the fraction of prompt background (identical for each tagging category, as is the case for the fit to the CP data sample) are allowed to float. The measured asymmetries are all consistent with zero, as shown in Table XVII. The observed Δt distributions for the B^0- and $\overline{B}^0$-tagged events in the $B_{\rm flav}$ sample is shown in Fig. 30a, where good agreement is clearly visible. Figure 30b demonstrates that there is no visible asymmetry as a function of Δt.

Control samples are also used to check the assumption that the Δt resolution function, which is primarily deter-

TABLE XXI. Yields, efficiencies, mistag rates w, and tagging separation $Q=\epsilon_{tag}(1-2w)^2$ as measured by the single-bin method in the hadronic $B_{\rm flav}$ event sample. A comparison of the mistag rates measured in the same sample with the single-bin method, $w_{\rm sb}$, and the likelihood fit, $w_{\rm like}$ (Table XVIII), are reported as the differences $\Delta_{\rm like}=w_{\rm sb}-w_{\rm like}$ between the two extraction techniques, normalized to the uncorrelated statistical and systematic errors.

Category	Yield	Efficiency [%]	Mistag rate w [%]	Q [%]	$\Delta_{\rm like}$ [σ]
Lepton	1128	11.0 ± 0.3	$9.5\pm1.5\pm0.6$	7.2	-0.8
Kaon	3687	35.8 ± 0.5	$17.8\pm1.0\pm0.7$	14.8	-0.4
NT1	819	7.9 ± 0.3	$22.0\pm2.2\pm0.9$	2.5	$+0.0$
NT2	1428	13.9 ± 0.3	$34.3\pm1.9\pm1.1$	1.4	$+0.8$

mined by the $B_{\rm flav}$ sample, can be applied to the charmonium decay modes in the *CP* sample. Figure 31 graphically compares the fitted Δt resolution function for the $B^+\to D(*)X$ control sample with that of the $B^+\to c\bar{c}X$ control sample. A 1σ error envelope encompasses the fit to the $B^+\to c\bar{c}X$ sample, which has five times fewer events. The level of agreement is acceptable. The same comparison between the $B^0\to D(*)X$ and $B^0\to c\bar{c}X$ samples was inconclusive due to the low statistics of the $B^0\to c\bar{c}X$ sample.

4. Time-integrated measurement of mistag rates

As described in Sec. I, a time-integrated technique can also be used to measure the mistag fractions in data, thereby providing a simple check of the likelihood fit method. The statistical precision of the time-integrated measurement is enhanced by restricting the sample to events in a single optimized Δt interval. Taking into account detector vertex resolution, the optimal interval is found to be $|\Delta t|<2.5$ ps. Events with $|\Delta t|>2.5$ ps have, on average, equal numbers of mixed and unmixed events due to flavor oscillations, and therefore contribute nothing to the determination of the mistag rate. We refer to this time-integrated technique using a single optimized Δt interval as the "single-bin" method and apply it to both the $B_{\rm flav}$ sample described in Sec. III C 1 and the semileptonic B^0 sample described in Sec. III E.

To correct for the presence of backgrounds, a term is added to Eq. (15) to account for the contribution of each background source to the fraction of mixed events in the sample:

$$\chi_{\rm obs}=f_{\rm sig}(\chi_d+(1-2\chi_d)w)+\sum_\beta f_\beta\chi_\beta, \tag{45}$$

where $f_{\rm sig}$ and f_β are the fraction of signal and background source β, respectively, χ_β is the fraction of mixed events in each background source, and $\chi_{\rm obs}$ is the observed fraction of mixed events. In this expression, χ_d must also be modified to represent the integrated mixing probability for $|\Delta t|<2.5$ ps. Using the world-average values for Δm_d and τ_{B^0} [11], and taking into account the Δt resolution function $\mathcal{R}(\Delta t)$, we find

$$\chi_d'=\tfrac{1}{2}\left[1-\int e^{-|\Delta t|/\tau}\cos(\Delta m_d\Delta t)\otimes\mathcal{R}(\Delta t)d(\Delta t)\Big/\int e^{-|\Delta t|/\tau}\otimes\mathcal{R}(\Delta t)d(\Delta t)\right]=0.079, \tag{46}$$

where the integral is performed over the range $|\Delta t|<2.5$ ps and $\mathcal{R}(\Delta t)$ is modeled by a double-Gaussian distribution with five parameters (one fraction, two biases and two widths) determined directly from data using the hadronic sample. Solving Eq. (45) for w, and using the calculated value for χ_d', the mistag rates are obtained:

$$w=\frac{\chi_{\rm obs}-f_{\rm sig}\chi_d'-\Sigma_\beta f_\beta\chi_\beta}{f_{\rm sig}(1-2\chi_d')}. \tag{47}$$

TABLE XXII. Yields, efficiencies, mistag rate w, and tagging separation $Q=\epsilon_{tag}(1-2w)^2$ as measured by the single-bin method in the semileptonic B^0 event sample. A comparison of the mistag rates measured with the single-bin method are reported as the differences $\Delta_{\rm sample}=w_{\rm flav}-w_{\rm sl}$ between the mistag rates in the $B_{\rm flav}$ sample, $w_{\rm flav}$ (Table XXI), and semileptonic B^0 samples, $w_{\rm sl}$, normalized to the quadratic sum of statistical and uncorrelated systematic errors.

Category	Yield	Efficiency [%]	Mistag rate w [%]	Q [%]	$\Delta_{\rm sample}$ [σ]
Lepton	3046	11.9 ± 0.4	$8.7\pm0.9\pm1.4$	8.1	$+0.4$
Kaon	10270	36.2 ± 1.9	$19.5\pm0.7\pm1.2$	13.5	-1.1
NT1	2127	8.1 ± 0.4	$22.3\pm1.4\pm1.2$	2.5	-0.1
NT2	3967	13.5 ± 0.9	$36.0\pm1.2\pm1.3$	1.1	-0.7

TABLE XXIII. Sources of systematic error for the mistag measurement on the $B_{\rm flav}$ sample in the single-bin method.

Type	Variation	Lepton	Kaon	NT1	NT2
$\tau(B^0)$, Δm_d	$\pm 1\sigma$	0.005	0.004	0.003	0.002
Resolution	see text	0.002	0.002	0.001	0.001
Wrong-tag resolution	see text	0.003	0.006	0.007	0.009
Combinatorial bkgd	$\pm 1\sigma$	0.002	0.002	0.005	0.004
$B^{\pm}$ peaking bkgd	$\pm 1\sigma$	0.001	0.001	0.000	0.000
Total		0.006	0.007	0.009	0.011

All tagged events in the $B_{\rm flav}$ sample with $|\Delta t|<2.5$ ps are used for a single-bin study. The combinatorial background fraction in the signal sample is determined from a fit to the $m_{\rm ES}$ distribution as described in Sec. VII. The signal region is defined as events with $m_{\rm ES}>5.27\ {\rm GeV}/c^2$. The B^+ peaking background in this signal is estimated to be $(1.3 \pm 0.8)\%$. The fraction of mixed events in the combinatorial background is determined by tagging category with the sideband control sample, $5.20<m_{\rm ES}<5.27\ {\rm GeV}/c^2$, and the mistag fraction associated with the B^+ peaking background has been measured directly in data. The number of tagged events in each category is summarized in Table XXI.

A separate single-bin analysis is also performed with the sample of $B^0\to D^{*-}l^+\overline{\nu}$ events described in Sec. III E. We use tagged events with $|\Delta t|<2.5$ ps and evaluate the backgrounds for events in this time interval. The backgrounds and mixed-event fractions are evaluated separately for each tagging category. Backgrounds are larger for the semileptonic modes than for the hadronic modes and originate from a variety of sources. In the case of the combinatorial background, the estimate is obtained from the $m(\overline{D}^0\pi^-)-m(\overline{D}^0)$ sideband. For the continuum background, off-resonance data is used after correction for the combinatorial component. The mixed-event fraction for $B\overline{B}$ background is estimated with generic $B\overline{B}$ Monte Carlo simulation. The mistag fraction of the last background component, the decay $B^+\to D^*Xl\overline{\nu}$, has been determined with data. The estimates of the contributions of the various backgrounds are described in Sec. III E. The number of tagged events in each category are summarized in Table XXII.

TABLE XXIV. Sources of systematic error for the mistag measurement from the semileptonic B^0 sample in the single-bin method.

Type	Variation	Lepton	Kaon	NT1	NT2
$\tau(B^0)$, Δm_d	$\pm 1\sigma$	0.006	0.004	0.004	0.002
Resolution	see text	0.001	0.001	0.001	0.001
Wrong-tag resolution	see text	0.003	0.006	0.007	0.009
Combinatorial bkgd	$\pm 1\sigma$	0.001	0.006	0.004	0.004
Continuum bkgd	$\pm 1\sigma$	0.001	0.003	0.006	0.007
$B\overline{B}$ bkgd	$\pm 1\sigma$	0.011	0.005	0.005	0.004
$B^+\to D^*Xl\overline{\nu}$ bkgd	$\pm 1\sigma$	0.003	0.004	0.002	0.001
Total		0.014	0.012	0.012	0.013

We use Eq. (47) to obtain the mistag rates in each tagging category shown in Table XXI for the $B_{\rm flav}$ sample and Table XXII for the B^0 semileptonic sample. The sources of systematic error on these results are summarized in Tables XXIII and XXIV respectively.

Three sources of systematic uncertainties are common to both the hadronic and semileptonic samples. The first is the uncertainty due to the errors on the world-average values for the B^0 lifetime and Δm_d values. The second is due to the Δt resolution function, whose fit parameters in data are varied within errors. The third common uncertainty is related to the possibility that wrong tags have worse Δt resolution than correct tags. This effect has been studied with Monte Carlo simulation, where we observe a slightly larger rms width for events with wrong-sign tags. From this study, scale factors comparing the right and wrong-tag resolution functions have been extracted and then applied to the resolution function for wrong tags.

The systematic uncertainties unique to each sample are due to the background components. These are estimated by varying both the background fractions f_β and the fraction of mixed events associated with each background source, χ_β, by one standard deviation in their uncertainty.

For the semileptonic sample, the systematic error due to backgrounds is the dominant source. The characterization of these various backgrounds is described in Sec. III E. For the combinatorial background fraction, a relative systematic uncertainty of 20% is added in quadrature to the statistical error to cover the range of results obtained with various $m(\overline{D}^0\pi^-)-m(\overline{D}^0)$ fitting functions.

The systematic error due to the continuum background is determined by varying both the background level and the mixed fractions. The $B\overline{B}$ background fraction uncertainty is obtained by combining the statistical uncertainty and the systematic error given in Sec. III E. The systematic errors introduced by uncertainties on the background from the decay $B^+\to D^*Xl\overline{\nu}$ are obtained by varying the fraction described in Sec. III E as well as the mistag fraction of B^+ mesons measured on data. Studies with Monte Carlo simulation have been performed to verify that the mistag fractions are not affected by the presence of the extra pions in the decay $B^+\to D^*Xl\overline{\nu}$. An additional uncertainty due to the statistical precision of the Monte Carlo study has been added to the charged B mistag fractions measured with data.

Table XXI shows the difference $\Delta_{\rm like}=w_{\rm sb}-w_{\rm like}$ between the mistag rates measured with the single-bin method in the $B_{\rm flav}$ sample, $w_{\rm sb}$, and the likelihood fit result, $w_{\rm like}$ (Table XVIII). The difference is reported in terms of the uncorrelated statistical and systematic errors for the two methods, when applied to the same data sample. The component of uncorrelated statistical error is estimated with a fast parametrized Monte Carlo simulation. It varies with category due to different event yields. The differences $\Delta_{\rm sample}=w_{\rm flav}-w_{\rm sl}$ in the mistag rates measured with the single-bin method in the $B_{\rm flav}$ sample, $w_{\rm flav}$, and in the semileptonic B^0 sample, $w_{\rm sl}$, are reported in Table XXII. The quadratic sum of the statistical and uncorrelated systematic errors is used to estimate the consistency of the measurements.

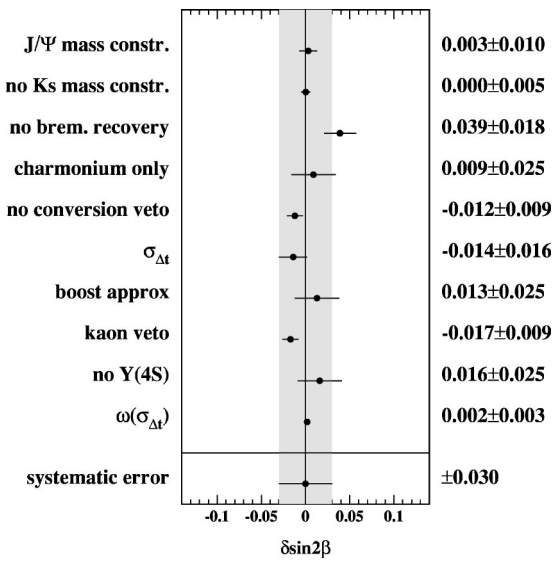

FIG. 32. Results obtained with several alternative variations on the vertexing algorithm that impact the Δt measurement. The shaded band is the systematic error assigned to the parametrization of the resolution function. The full range corresponds to one statistical standard deviation.

5. Vertexing algorithm checks

In order to verify that the results are stable under variation of the vertexing algorithm that is used for the measurement of Δt, several less powerful alternatives to the default method have been considered:

(1) *Charmonium mass constraint for vertex fit*: The mass constraint on the charmonium daughter, used in the selection of the events, is also applied in the determination of the vertex.
(2) *No K_S^0 mass constraint*: The mass constraint on the K_S^0 candidate is not applied during the vertex reconstruction.
(3) *No bremsstrahlung recovery*: Only events without an associated bremsstrahlung photon for the J/ψ daughter electrons are considered in the likelihood fit.
(4) *Charmonium daughters only*: The vertex of the fully reconstructed B meson is reconstructed only with the tracks from the charmonium daughter.
(5) *No converted photon veto*: Pairs of tracks from gamma conversions are retained in the vertex fit.
(6) *$\sigma_{\Delta t}$ requirement*: Only events with $\sigma_{\Delta t}<1.4$ ps are retained, as is required in the mixing analysis.
(7) *Boost approximation*: The boost approximation [Eq. (21)] is used to convert the Δz measurement into Δt.
(8) *Kaon veto*: The more restrictive requirement from the mixing analysis that no kaons participate in the tagging vertex is applied.
(9) *No $\Upsilon(4S)$ constraint*: The algorithm described in Sec. V A is simplified by dropping the $\Upsilon(4S)$ momentum constraint.
(10) *Dilution dependence on $\sigma_{\Delta t}$*: Dilutions for the kaon category are parametrized as a function of the error on Δt.

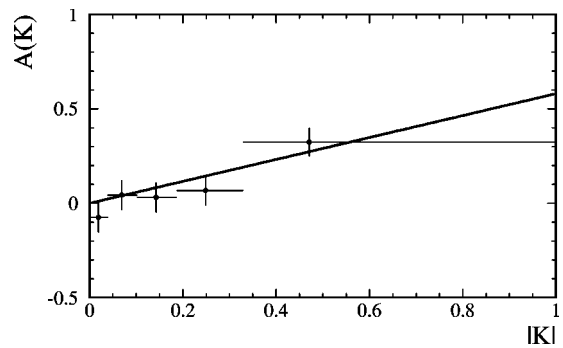

FIG. 33. Observed asymmetry $\mathcal{A}(\mathcal{K})$ as a function of $|\mathcal{K}|$, with an overlay of the expected linear dependence superimposed.

A summary of the results obtained with these different configurations for the Δt determination is provided in Fig. 32. In all cases the variation of the measured asymmetry is consistent with the error assigned to the parametrization of the resolution function.

6. $J/\psi K_L^0$ background cross-checks

As a cross-check, a likelihood fit was performed to $J/\psi K_L^0$ candidates in a ΔE sideband region ($20<\Delta E<80$ MeV) treated entirely as signal events. This sample is actually a mixture of B decay modes with an expected average η_{CP} of $+0.04$. The true value for $\sin 2\beta$ in the Monte Carlo simulation is 0.7 and consequently the expected result from the likelihood fit to the Δt distribution of the control sample is 0.03. The actual fits to sideband regions in data and Monte Carlo simulation find $\sin 2\beta=0.16\pm0.18$ and -0.03 ± 0.10 respectively, both of which are consistent with expectations.

As another cross-check, a sample of $J/\psi K_S^0$ events was selected in the data, where only the K_S^0 direction information was used, thereby emulating the K_L^0 selection. The purity and background composition of this control sample is very similar to that of the $J/\psi K_L^0$ sample. However, in this case, the subset of true $J/\psi K_S^0$ events can be identified with the normal $J/\psi K_S^0$ selection criteria. A fit to the ΔE distribution of the full control sample finds $(49\pm3)\%$ signal, which is in good agreement with the fraction, 47%, observed for the cleanly identified $J/\psi K_S^0$ subset. Likewise, a likelihood fit to the Δt distribution of the full control sample agrees well with the value of $\sin 2\beta$ obtained with the true $J/\psi K_S^0$ subsample.

7. Graphical display of the asymmetry

An elegant display of the *CP* asymmetry in the data can be obtained with the use of the so-called *Kin* variable, hereafter denoted as $\mathcal{K}$. It is also possible to verify directly the fitted value for $\sin 2\beta$ from the ratio of appropriate weighted averages for $\mathcal{K}$. In particular, $\mathcal{K}$ has a PDF with an asymmetry known to be linearly dependent with a slope given by $\sin 2\beta$ regardless of the details of the analysis.

Writing the PDF $\mathcal{F}_+(\mathcal{F}_-)$ for events with a B^0 ($\overline{B}^0$) tag in terms of the general functions $F_1(\Delta t)$ and $F_2(\Delta t)$

$$\mathcal{F}_\pm(\Delta t)=F_1(\Delta t)\pm\sin 2\beta\ F_2(\Delta t) \tag{48}$$

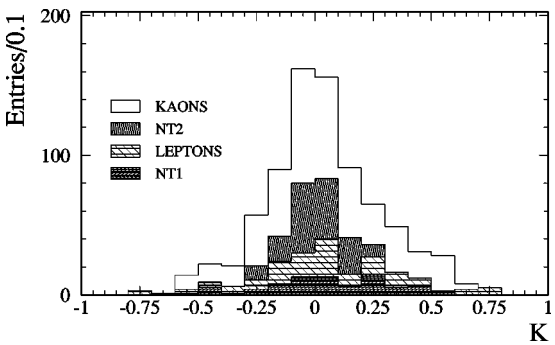

FIG. 34. Distribution of the observable $\mathcal{K}$ for the individual tagging categories.

allows us to introduce

$$\mathcal{K}(\Delta t) = \pm F_2(\Delta t)/F_1(\Delta t), \tag{49}$$

where + applies to events with a B^0 tag and − with a $\overline{B^0}$ tag. Ignoring resolution effects, dilutions and background, the standard model expectation for the Δt distribution of tagged B^0 decays into CP modes [Eq. (10)] gives $\mathcal{K}(\Delta t) = -\eta_{CP} \sin \Delta m_d \Delta t$. When these effects are included, we can still write

$$\mathcal{F}_{\pm}(\Delta t) = F_1(\Delta t)(1 + \mathcal{K}(\Delta t) \sin 2\beta), \tag{50}$$

although $\mathcal{K}$ will be a more complicated function of Δt and could depend on kinematic variables as well.

The distribution of events as a function of $\mathcal{K}$ is

$$\begin{aligned} \frac{dN}{d\mathcal{K}} &= \int d\Delta t \left[\mathcal{F}_+ \delta\left(\mathcal{K} - \frac{F_2}{F_1}\right) + \mathcal{F}_- \delta\left(\mathcal{K} + \frac{F_2}{F_1}\right) \right] \\ &= (1 + \mathcal{K} \sin 2\beta) \int d\Delta t \\ &\quad \times F_1 \left[\delta\left(\mathcal{K} - \frac{F_2}{F_1}\right) + \delta\left(\mathcal{K} + \frac{F_2}{F_1}\right) \right] \\ &= (1 + \mathcal{K} \sin 2\beta) \Psi(\mathcal{K}), \end{aligned} \tag{51}$$

where $\Psi(\mathcal{K})$ is an even function of $\mathcal{K}$. It follows that the ratio of the odd to the even part of the distribution for $\mathcal{K}$ is a linear function of $\mathcal{K}$ with coefficient $\sin 2\beta$. Thus the distribution of $\mathcal{K}$ can be used to test for the effect of an CP violation simply by examining the dependence of the asymmetry

$$\mathcal{A}(\mathcal{K}) = \frac{dN_{\mathcal{K}>0}/d\mathcal{K} - dN_{\mathcal{K}<0}/d\mathcal{K}}{dN_{\mathcal{K}>0}/d\mathcal{K} + dN_{\mathcal{K}<0}/d\mathcal{K}}. \tag{52}$$

The observed asymmetry for the CP sample is shown in Fig. 33 as a function of $\mathcal{K}$, along with an overlay of the expected linear dependence. The data agree with this hypothesis at the 55% C.L.

From the expression in Eq. (51) we find

$$\sin 2\beta = \frac{\Sigma_i \mathcal{K}_i}{\Sigma_i \mathcal{K}_i^2} \pm \frac{1}{\sqrt{\Sigma_i \mathcal{K}_i^2}} \sqrt{1 - (\sin 2\beta)^2 \frac{\Sigma_i \mathcal{K}_i^4}{\Sigma_i \mathcal{K}_i^2}}. \tag{53}$$

In averaging $\mathcal{K}$ the even component of the $\Psi(\mathcal{K})$ cancels out, while the odd component cancels in averaging $\mathcal{K}^2$. This offers a method of measuring $\sin 2\beta$ that is mathematically equivalent to the result with the global likelihood fit. However, it can only be applied when $\sin 2\beta$ is the one remaining free parameter. The moments of $\mathcal{K}$ for the full CP sample give results that are numerically identical to the likelihood fit, thereby confirming the minimization procedure used for the fit.

The fact that the mean value of $\mathcal{K}$ is proportional to $\sin 2\beta$ also allows a visual representation of the CP asymmetry. Figure 34 shows the distribution of $\mathcal{K}$ in data, with events in the individual tagging categories indicated as well. The larger the value of $\mathcal{K}$ for a given event, the larger the weight that this event carries in the measurement of $\sin 2\beta$. Again, the CP asymmetry in the data is clearly evident in the distribution of $\mathcal{K}$.

D. Fits results without assuming $|\lambda| = 1$

A more general description of the time evolution of neutral B decays to CP eigenstates contains a term proportional to $\cos \Delta m_d \Delta t$ [Eq. (8)]. The coefficient of the cosine term is expected to be negligible in the standard model, where $|\lambda| = 1$. In order to search for a non-standard model effect, we fit the $\eta_{CP} = -1$ sample for $|\lambda|$ and $\mathrm{Im}\,\lambda/|\lambda|$. The latter is equal to $\sin 2\beta$ if $|\lambda| = 1$. The $\eta_{CP} = -1$ sample has the advantage of having very little background, while the other CP modes have backgrounds that are both significantly larger and dominated by other B decay modes with possible direct CP contributions.

The fitted values for $|\lambda|$ and $\mathrm{Im}\,\lambda/|\lambda|$ with the $CP = -1$ sample and various subsamples are listed in Table XVII. The two CP parameters are almost uncorrelated, with the coefficient between $\mathrm{Im}\,\lambda/|\lambda|$ and $|\lambda|$ of −1.7%. The same systematic error studies as described in detail in Sec. VIII B were repeated for the fit to the $\eta_{CP} = -1$ sample for $\mathrm{Im}\,\lambda/|\lambda|$ and $|\lambda|$. The estimated uncertainties from these sources are listed in Table XIX.

We have also performed detailed cross-checks, similar to those described in Sec. VIII C. In particular, large samples of parametrized simulation, as well as full Monte Carlo samples, have been used to verify the fitting procedure. The B_{flav} sample has also been used to demonstrate that no bias is introduced in the measurement. The relative normalization of the tagged events in the two flavors is in fact sensitive to the coefficient of the cosine term in Eq. (8), and therefore $|\lambda|$. The systematic error introduced by the uncertainty on the parameters $\langle \epsilon_{\mathrm{tag}} \rangle_i$ and μ_i listed in Table X are uncorrelated between tagging categories. Therefore, they are added in quadrature to obtain the systematic error contribution listed in Table XIX(g).

The final result of the fit with the $\eta_{CP} = -1$ sample is

$$\begin{aligned} |\lambda| &= 0.93 \pm 0.09 \pm 0.02, \\ \mathrm{Im}\,\lambda/|\lambda| &= 0.56 \pm 0.15 \pm 0.05. \end{aligned} \tag{54}$$

Thus, we find no evidence for direct CP violation in the

$\eta_{CP}=-1$ sample and the value of $\mathrm{Im}\,\lambda/|\lambda|$ is consistent with the result from the nominal *CP* fit with $|\lambda|=1$.

IX. CONCLUSIONS AND PROSPECTS

In 29.7 fb^{-1} of e^+e^- annihilation data collected near the $\Upsilon(4S)$ resonance, we have obtained a new measurement of the time-dependent B^0-$\overline{B}^0$ oscillation frequency with a sample of 6350 tagged flavor-eigenstate B^0 meson decays that are fully reconstructed in hadronic final states:

$$\Delta m_d = 0.516 \pm 0.016(\mathrm{stat}) \pm 0.010(\mathrm{syst})\ \mathrm{ps}^{-1}.$$

This result is at a level of precision comparable to the most recent world average for Δm_d and lies about 1.7σ above the combined value of $0.472 \pm 0.017\ \mathrm{ps}^{-1}$ [11]. It is also quite compatible with our own recent measurement [28] with a dilepton sample. The Δm_d study reported here confirms our understanding of B reconstruction, flavor tagging, and Δt resolution in our data sample. Our measurement contributes significantly to the precision of the determined value for Δm_d, one of the fundamental parameters constraining our knowledge of the CKM matrix, and remains dominated by statistical errors that will improve with more data.

We have presented a measurement of the *CP*-violating asymmetry parameter $\sin 2\beta$ in the neutral B meson system:

$$\sin 2\beta = 0.59 \pm 0.14(\mathrm{stat}) \pm 0.05(\mathrm{syst}), \tag{55}$$

which establishes *CP* violation in the B^0 system at the 4.1σ level. This significance is computed from the sum in quadrature of the statistical and additive systematic errors. The probability of obtaining the observed value or higher in the absence of *CP* violation is less than 3×10^{-5}. The corresponding probability for the $\eta_{CP}=-1$ sample alone is 2×10^{-4}. Our measurement is consistent at the 1.9σ level with the recently reported result from Belle of $\sin 2\beta = 0.99 \pm 0.14(\mathrm{stat}) \pm 0.06(\mathrm{syst})$ [29], and with previous measurements [30–33]. The observed value for $\sin 2\beta$ is currently limited by the size of the *CP* sample, allowing for substantial improvement as more data are recorded in the next few years.

We have also used the $\eta_{CP}=-1$ sample to search for possible direct *CP* violation through interference of decay amplitudes. The direct *CP* parameter λ is found to be

$$|\lambda| = 0.93 \pm 0.09(\mathrm{stat}) \pm 0.02(\mathrm{syst}).$$

This result is consistent with the standard model expectation, where $|\lambda|=1$ and no significant direct *CP* violation should exist in charmonium decays.

As already noted in Sec. I, measurements of *CP* asymmetries in B decays to charmonium can be used to constrain, with little theoretical ambiguity, the parameters of the CKM matrix. In the standard model with three families, the CKM matrix V [3] incorporates three real parameters and one phase δ generating *CP* violation if $\delta \neq 0$ or π. The Wolfenstein parametrization [14] of V takes advantage of the observed hierarchy in the matrix elements in terms of the expansion parameter $\lambda_{CKM}=|V_{us}|$. The remaining parameters in this representation are denoted A, ρ, and η, where *CP* violation requires $\eta \neq 0$.

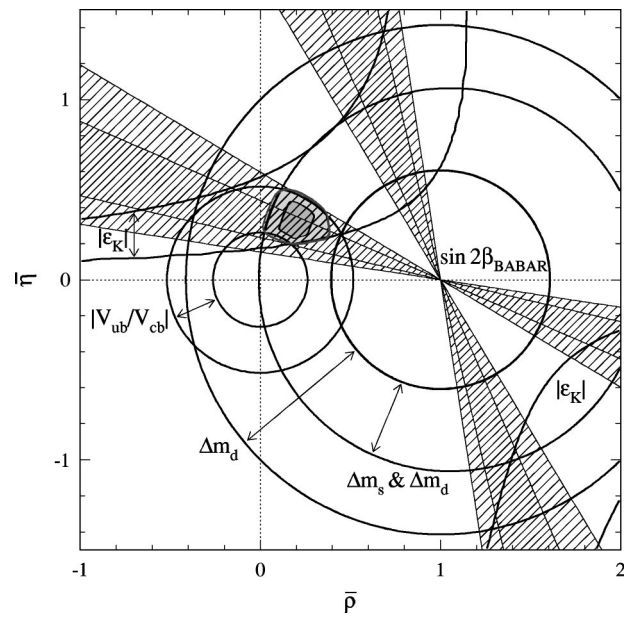

FIG. 35. Present indirect constraints on the position of the apex of the unitarity triangle in the $(\overline{\rho},\overline{\eta})$ plane, not including our measurement of $\sin 2\beta$. The fitting procedure is described in Ref. [34]. Our result $\sin 2\beta = 0.59 \pm 0.14(\mathrm{stat}) \pm 0.05(\mathrm{syst})$ is represented by diagonally hatched regions, corresponding to one and two statistical standard deviations. The individual indirect constraints lie between the pairs of solid lines that are connected by the double-ended arrows with labels.

The parameter λ_{CKM} is determined from semileptonic kaon decays and nuclear β decays. Semileptonic B meson decays to charm are used to determine the parameter A. Constraints on ρ and η are obtained from *CP* violation in mixing in the kaon sector $|\epsilon_K|$, the ratio $|V_{ub}/V_{cb}|$, and the oscillation frequency Δm_d for B^0-$\overline{B}^0$ mixing. The oscillation frequency Δm_s has not been measured, since $B_S^0-\overline{B}_S^0$ mixing has not been observed yet. However, the observed amplitude spectrum $\mathcal{A}(\Delta m_s)$ improves the constraints on ρ and η. Together, these measurements provide indirect constraints on $\sin 2\beta$.

Our overall knowledge of the CKM parameters is limited by the relatively large uncertainties in some of the theoretical quantities, mainly due to nonperturbative QCD effects. In particular, the constraints on ρ and η suffer from theoretical and systematic uncertainties in the determination of $|V_{ub}/V_{cb}|$ and from theoretical uncertainties in QCD parameters entering the prediction of $|\epsilon_K|$, Δm_d, and Δm_s. Recent analyses constraining the CKM matrix have been performed with different statistical approaches [34–40]. They mainly differ in the treatment of theoretical uncertainties and also in the choice of the input values and their errors.

Due to the fourfold ambiguity in the value of β obtained from the $\sin 2\beta$ measurement, there are four allowed regions in the ρ-η plane. One of these regions is found to be in agreement with the allowed ρ-η region obtained from CKM

fits within the standard model. Figure 35, taken from Ref. [34], shows our direct measurement and the indirect constraints in the $\bar{\rho}-\bar{\eta}$ plane in terms of the renormalized parameters $\bar{\rho}=\rho(1-\lambda_{CKM}^2/2)$ and $\bar{\eta}=\eta(1-\lambda_{CKM}^2/2)$. The contributions of the individual measurements $|\epsilon_K|$, $|V_{ub}/V_{cb}|$, Δm_d, and Δm_s [34] are indicated, as well as the allowed region if all the constraints are considered simultaneously. Overlaid as the diagonally-hatched area are the regions corresponding to one and two times the one-standard-deviation experimental uncertainty on our $\sin 2\beta$ measurement.

It should be emphasized that, beyond being a direct constraint on β, the measurement of $\sin 2\beta$ differs qualitatively in its interpretation from the indirect constraints on β obtained from $|\epsilon_K|$, $|V_{ub}/V_{cb}|$, Δm_d, and eventually Δm_s. For $\sin 2\beta$, the size of the allowed domain is determined by well-defined experimental uncertainties that are predominantly statistical in origin, while in contrast the region allowed by the indirect measurements is mostly defined by theoretical uncertainties, which makes a statistical interpretation difficult.

The current experimental uncertainty on $\sin 2\beta$ has now reached a level of precision that offers significant constraint on the standard model. Over the next few years there will continue to be substantial improvements in precision of the $\sin 2\beta$ determination, including measurements for other final states in which CP-violating asymmetries are proportional to $\sin 2\beta$. Beyond this, studies of time-dependent asymmetries in modes involving $b\to u$ transitions have already begun [41] and may provide additional constraints, although here the interpretation in terms of $\sin 2\alpha$ from the unitarity triangle is likely to be made difficult due to significant penguin contributions. Nevertheless, these measurements will be able to directly test the validity of the CKM picture as the origin for the observed CP violation in neutral B decays.

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 US 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 (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] J. H. Christenson *et al.*, Phys. Rev. Lett. **13**, 138 (1964); NA31 Collaboration, G. D. Barr *et al.*, Phys. Lett. B **317**, 233 (1993); E731 Collaboration, L. K. Gibbons *et al.*, Phys. Rev. Lett. **70**, 1203 (1993).

[2] A. D. Sakhorov, Pis'ma Zh. Eksp. Teor. Fiz. **5**, 32 (1967) [JETP Lett. **5**, 24 (1967)].

[3] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963); M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).

[4] See, for instance, "The *BABAR* Physics Book," edited by P. H. Harrison and H. R. Quinn, SLAC-R-504, 1998, Sec. 14 and references therein.

[5] For an introduction to *CP* violation, see, for instance, "The *BABAR* Physics Book," [4] Sec. 1, and references therein.

[6] M. B. Gavela *et al.*, Mod. Phys. Lett. A **9**, 795 (1994); Nucl. Phys. **B340**, 382 (1994); P. Huet and E. Sather, Phys. Rev. D **51**, 379 (1995).

[7] C. Jarlskog, in *CP Violation*, edited by C. Jarlskog (World Scientific, Singapore, 1988).

[8] "PEP-II: An Asymmetric *B* Factory," Conceptual Design Report, SLAC-418, LBL-5379, 1993.

[9] UA1 Collaboration, C. Albajar *et al.*, Phys. Lett. B **186**, 247 (1987).

[10] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **192**, 245 (1987).

[11] Particle Data Group, D. E. Groom *et al.*, Eur. Phys. J. C **15**, 1 (2000).

[12] Throughout this paper, charged and neutral flavor-eigenstate decay modes imply also their charge conjugate.

[13] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **87**, 241801 (2001).

[14] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).

[15] *BABAR* Collaboration, B. Aubert *et al.* Phys. Rev. Lett. **88**, 221802 (2002).

[16] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **87**, 091801 (2001); **86**, 2515 (2001).

[17] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **87**, 201803 (2001).

[18] *BABAR* Collaboration, B. Aubert *et al.*, Nucl. Instrum. Methods Phys. Res. A **479**, 1 (2002).

[19] P. Billoir, Nucl. Instrum. Methods Phys. Res. A **225**, 225 (1984).

[20] A. Drescher *et al.*, Nucl. Instrum. Methods Phys. Res. A **237**, 464 (1985).

[21] R. Sinkus and T. Voss, Nucl. Instrum. Methods Phys. Res. A **391**, 360 (1997).

[22] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978).

[23] ARGUS Collaboration, H. Albrecht *et al.*, Phys. Lett. B **241**, 278 (1990).

[24] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. D **65**, 032001 (2002).

[25] The Crystal Ball function is given by

$$f(x)=N\begin{cases}\exp\left(-\dfrac{(x-\overline{x})^2}{2\sigma^2}\right), & (x-\overline{x})/\sigma>\alpha,\\ A\times\left(B-\dfrac{x-\overline{x}}{\sigma}\right)^{-n}, & (x-\overline{x})/\sigma\leqslant\alpha,\end{cases}$$

where $A\equiv(n/|\alpha|)^n\times\exp(-|\alpha|^2/2)$ and $B\equiv n/|\alpha|-|\alpha|$. N is a normalization factor, $\overline{x}$ and σ are the fitted peak position and width of the Gaussian portion of the function, and α and n are the fitted point at which the function transitions to the power function and the exponent of the power function, respectively. D. Antreasyan, Crystal Ball Note 321, 1983.

[26] E687 Collaboration, P. L. Frabetti *et al.*, Phys. Lett. B **331**, 217 (1994).

[27] ''GEANT, Detector Description and Simulation Tool,'' CERN program library long writeup W5013, 1994.

[28] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. Lett. **88**, 221803 (2002). Both are somewhat higher, but consistent with, a recent dilepton result from the Belle Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **86**, 3228 (2001).

[29] Belle Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **87**, 091802 (2001).

[30] Other results have been reported by OPAL [31] ($\sin 2\beta=3.2^{+1.8}_{-2.0}\pm 0.5$) and the Collider Detecter at Fermilab (CDF) [32] ($\sin 2\beta=0.79^{+0.41}_{-0.44}$). See also ALEPH's preliminary result [33] $\sin 2\beta=0.93^{+0.64+0.36}_{-0.88-0.24}$.

[31] OPAL Collaboration, K. Ackerstaff *et al.*, Eur. Phys. J. C **5**, 379 (1998).

[32] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. D **61**, 072005 (2000).

[33] ALEPH Collaboration, ALEPH 99-099 CONF 99-54, 1999.

[34] A. Höcker, H. Lacker, S. Laplace, and F. Le Diberder, Eur. Phys. J. C **21**, 225 (2001).

[35] S. Plaszczynski and M.-H. Schune, ''Overall determination of the CKM matrix,'' LAL-99-67, hep-ph/9911280; ''The BABAR Physics Book,'' Chap. 14, SLAC-R-504, 1998.

[36] M. Ciuchini *et al.*, J. High Energy Phys. **107**, 013 (2001).

[37] M. Bargiotti *et al.*, Riv. Nuovo Cimento **23N3**, 1 (2000).

[38] P. Faccioli, Nucl. Instrum. Methods Phys. Res. A **462**, 313 (2001).

[39] D. Atwood and A. Soni, Phys. Lett. B **516**, 39 (2001).

[40] A. Ali and D. London, Eur. Phys. J. C **18**, 665 (2001).

[41] *BABAR* Collaboration, B. Aubert *et al.*, Phys. Rev. D **65**, 051502 (2002).