

**Observation of mixing-induced  $CP$  violation in the neutral  $B$  meson system**

K. Abe,<sup>9</sup> K. Abe,<sup>43</sup> R. Abe,<sup>31</sup> T. Abe,<sup>44</sup> I. Adachi,<sup>9</sup> Byoung Sup Ahn,<sup>17</sup> H. Aihara,<sup>45</sup> M. Akatsu,<sup>24</sup> Y. Asano,<sup>50</sup> T. Aso,<sup>49</sup> T. Aushev,<sup>14</sup> A. M. Bakich,<sup>40</sup> Y. Ban,<sup>35</sup> E. Banas,<sup>29</sup> S. Behari,<sup>9</sup> P. K. Behera,<sup>51</sup> A. Bondar,<sup>2</sup> A. Bozek,<sup>29</sup> M. Bračko,<sup>22,15</sup> T. E. Browder,<sup>8</sup> B. C. K. Casey,<sup>8</sup> P. Chang,<sup>28</sup> Y. Chao,<sup>28</sup> B. G. Cheon,<sup>39</sup> R. Chistov,<sup>14</sup> S.-K. Choi,<sup>7</sup> Y. Choi,<sup>39</sup> L. Y. Dong,<sup>12</sup> J. Dragic,<sup>23</sup> A. Drutskoy,<sup>14</sup> S. Eidelman,<sup>2</sup> V. Eiges,<sup>14</sup> Y. Enari,<sup>24</sup> C. W. Everton,<sup>23</sup> F. Fang,<sup>8</sup> C. Fukunaga,<sup>47</sup> M. Fukushima,<sup>11</sup> N. Gabyshev,<sup>9</sup> A. Garmash,<sup>2,9</sup> T. Gershon,<sup>9</sup> B. Golob,<sup>21,15</sup> A. Gordon,<sup>23</sup> H. Guler,<sup>8</sup> R. Guo,<sup>26</sup> J. Haba,<sup>9</sup> H. Hamasaki,<sup>9</sup> K. Hanagaki,<sup>36</sup> F. Handa,<sup>44</sup> K. Hara,<sup>33</sup> T. Hara,<sup>33</sup> N. C. Hastings,<sup>23</sup> H. Hayashii,<sup>25</sup> M. Hazumi,<sup>9</sup> E. M. Heenan,<sup>23</sup> I. Higuchi,<sup>44</sup> T. Higuchi,<sup>45</sup> T. Hojo,<sup>33</sup> T. Hokuue,<sup>24</sup> Y. Hoshi,<sup>43</sup> S. R. Hou,<sup>28</sup> W.-S. Hou,<sup>28</sup> S.-C. Hsu,<sup>28</sup> H.-C. Huang,<sup>28</sup> T. Igaki,<sup>24</sup> T. Iijima,<sup>9</sup> H. Ikeda,<sup>9</sup> K. Inami,<sup>24</sup> A. Ishikawa,<sup>24</sup> H. Ishino,<sup>46</sup> R. Itoh,<sup>9</sup> H. Iwasaki,<sup>9</sup> Y. Iwasaki,<sup>9</sup> D. J. Jackson,<sup>33</sup> H. K. Jang,<sup>38</sup> H. Kakuno,<sup>46</sup> J. H. Kang,<sup>54</sup> J. S. Kang,<sup>17</sup> P. Kapusta,<sup>29</sup> N. Katayama,<sup>9</sup> H. Kawai,<sup>3</sup> H. Kawai,<sup>45</sup> Y. Kawakami,<sup>24</sup> N. Kawamura,<sup>1</sup> T. Kawasaki,<sup>31</sup> H. Kichimi,<sup>9</sup> D. W. Kim,<sup>39</sup> Heejong Kim,<sup>54</sup> H. J. Kim,<sup>54</sup> H. O. Kim,<sup>39</sup> Hyunwoo Kim,<sup>17</sup> S. K. Kim,<sup>38</sup> T. H. Kim,<sup>54</sup> K. Kinoshita,<sup>5</sup> H. Konishi,<sup>48</sup> S. Korpar,<sup>22,15</sup> P. Krizan,<sup>21,15</sup> P. Krokovny,<sup>2</sup> R. Kulasiri,<sup>5</sup> S. Kumar,<sup>34</sup> A. Kuzmin,<sup>2</sup> Y.-J. Kwon,<sup>54</sup> J. S. Lange,<sup>6</sup> G. Leder,<sup>13</sup> S. H. Lee,<sup>38</sup> A. Limosani,<sup>23</sup> D. Liventsev,<sup>14</sup> R.-S. Lu,<sup>28</sup> J. MacNaughton,<sup>13</sup> G. Majumder,<sup>41</sup> F. Mandl,<sup>13</sup> D. Marlow,<sup>36</sup> T. Matsuishi,<sup>24</sup> S. Matsumoto,<sup>4</sup> T. Matsumoto,<sup>24</sup> Y. Mikami,<sup>44</sup> W. Mitaroff,<sup>13</sup> K. Miyabayashi,<sup>25</sup> Y. Miyabayashi,<sup>24</sup> H. Miyake,<sup>33</sup> H. Miyata,<sup>31</sup> G. R. Moloney,<sup>23</sup> S. Mori,<sup>50</sup> T. Mori,<sup>4</sup> A. Murakami,<sup>37</sup> T. Nagamine,<sup>44</sup> Y. Nagasaka,<sup>10</sup> Y. Nagashima,<sup>33</sup> T. Nakadaira,<sup>45</sup> E. Nakano,<sup>32</sup> M. Nakao,<sup>9</sup> J. W. Nam,<sup>39</sup> Z. Natkaniec,<sup>29</sup> K. Neichi,<sup>43</sup> S. Nishida,<sup>18</sup> O. Nitoh,<sup>48</sup> S. Noguchi,<sup>25</sup> T. Nozaki,<sup>9</sup> S. Ogawa,<sup>42</sup> T. Ohshima,<sup>24</sup> T. Okabe,<sup>24</sup> S. Okuno,<sup>16</sup> S. L. Olsen,<sup>8</sup> W. Ostrowicz,<sup>29</sup> H. Ozaki,<sup>9</sup> P. Pakhlov,<sup>14</sup> H. Palka,<sup>29</sup> C. S. Park,<sup>38</sup> C. W. Park,<sup>17</sup> H. Park,<sup>19</sup> K. S. Park,<sup>39</sup> L. S. Peak,<sup>40</sup> J.-P. Perroud,<sup>20</sup> M. Peters,<sup>8</sup> L. E. Piilonen,<sup>52</sup> E. Prebys,<sup>36</sup> J. L. Rodriguez,<sup>8</sup> F. Ronga,<sup>20</sup> M. Rozanska,<sup>29</sup> K. Rybicki,<sup>29</sup> H. Sagawa,<sup>9</sup> Y. Sakai,<sup>9</sup> M. Satpathy,<sup>51</sup> A. Satpathy,<sup>9,5</sup> O. Schneider,<sup>20</sup> S. Schrenk,<sup>5</sup> C. Schwanda,<sup>9,13</sup> S. Semenov,<sup>14</sup> K. Senyo,<sup>24</sup> M. E. Sevier,<sup>23</sup> H. Shibuya,<sup>42</sup> B. Schwartz,<sup>2</sup> V. Sidorov,<sup>2</sup> J. B. Singh,<sup>34</sup> S. Stanič,<sup>50</sup> A. Sugi,<sup>24</sup> A. Sugiyama,<sup>24</sup> K. Sumisawa,<sup>9</sup> T. Sumiyoshi,<sup>9</sup> K. Suzuki,<sup>9</sup> S. Suzuki,<sup>53</sup> S. Y. Suzuki,<sup>9</sup> H. Tajima,<sup>45</sup> T. Takahashi,<sup>32</sup> F. Takasaki,<sup>9</sup> M. Takita,<sup>33</sup> K. Tamai,<sup>9</sup> N. Tamura,<sup>31</sup> J. Tanaka,<sup>45</sup> M. Tanaka,<sup>9</sup> G. N. Taylor,<sup>23</sup> Y. Teramoto,<sup>32</sup> S. Tokuda,<sup>24</sup> M. Tomoto,<sup>9</sup> T. Tomura,<sup>45</sup> S. N. Tovey,<sup>23</sup> K. Trabelsi,<sup>8</sup> W. Trischuk,<sup>36,\*</sup> T. Tsuboyama,<sup>9</sup> T. Tsukamoto,<sup>9</sup> S. Uehara,<sup>9</sup> K. Ueno,<sup>28</sup> Y. Unno,<sup>3</sup> S. Uno,<sup>9</sup> Y. Ushiroda,<sup>9</sup> S. E. Vahsen,<sup>36</sup> K. E. Varvell,<sup>40</sup> C. C. Wang,<sup>28</sup> C. H. Wang,<sup>27</sup> J. G. Wang,<sup>52</sup> M.-Z. Wang,<sup>28</sup> Y. Watanabe,<sup>46</sup> E. Won,<sup>38</sup> B. D. Yabsley,<sup>9</sup> Y. Yamada,<sup>9</sup> M. Yamaga,<sup>44</sup> A. Yamaguchi,<sup>44</sup> H. Yamamoto,<sup>44</sup> T. Yamanaka,<sup>33</sup> Y. Yamashita,<sup>30</sup> M. Yamauchi,<sup>9</sup> J. Yashima,<sup>9</sup> M. Yokoyama,<sup>45</sup> Y. Yuan,<sup>12</sup> Y. Yusa,<sup>44</sup> H. Yuta,<sup>1</sup> C. C. Zhang,<sup>12</sup> J. Zhang,<sup>50</sup> Y. Zheng,<sup>8</sup> V. Zhilich,<sup>2</sup> and D. Zontar<sup>50</sup>

(Belle Collaboration)

<sup>1</sup>Aomori University, Aomori<sup>2</sup>Budker Institute of Nuclear Physics, Novosibirsk<sup>3</sup>Chiba University, Chiba<sup>4</sup>Chuo University, Tokyo<sup>5</sup>University of Cincinnati, Cincinnati, Ohio<sup>6</sup>University of Frankfurt, Frankfurt<sup>7</sup>Gyeongsang National University, Chinju<sup>8</sup>University of Hawaii, Honolulu, Hawaii<sup>9</sup>High Energy Accelerator Research Organization (KEK), Tsukuba<sup>10</sup>Hiroshima Institute of Technology, Hiroshima<sup>11</sup>Institute for Cosmic Ray Research, University of Tokyo, Tokyo<sup>12</sup>Institute of High Energy Physics, Chinese Academy of Sciences, Beijing<sup>13</sup>Institute of High Energy Physics, Vienna<sup>14</sup>Institute for Theoretical and Experimental Physics, Moscow<sup>15</sup>J. Stefan Institute, Ljubljana<sup>16</sup>Kanagawa University, Yokohama<sup>17</sup>Korea University, Seoul<sup>18</sup>Kyoto University, Kyoto<sup>19</sup>Kyungpook National University, Taegu<sup>20</sup>IPHE, University of Lausanne, Lausanne<sup>21</sup>University of Ljubljana, Ljubljana<sup>22</sup>University of Maribor, Maribor<sup>23</sup>University of Melbourne, Victoria<sup>24</sup>Nagoya University, Nagoya<sup>25</sup>Nara Women's University, Nara<sup>26</sup>National Kaohsiung Normal University, Kaohsiung<sup>27</sup>National Lien-Ho Institute of Technology, Miao Li<sup>28</sup>National Taiwan University, Taipei<sup>29</sup>H. Niewodniczanski Institute of Nuclear Physics, Krakow<sup>30</sup>Nihon Dental College, Niigata

- <sup>31</sup>Niigata University, Niigata  
<sup>32</sup>Osaka City University, Osaka  
<sup>33</sup>Osaka University, Osaka  
<sup>34</sup>Panjab University, Chandigarh  
<sup>35</sup>Peking University, Beijing  
<sup>36</sup>Princeton University, Princeton, New Jersey  
<sup>37</sup>Saga University, Saga  
<sup>38</sup>Seoul National University, Seoul  
<sup>39</sup>Sungkyunkwan University, Suwon  
<sup>40</sup>University of Sydney, Sydney NSW  
<sup>41</sup>Tata Institute of Fundamental Research, Bombay  
<sup>42</sup>Toho University, Funabashi  
<sup>43</sup>Tohoku Gakuin University, Tagajo  
<sup>44</sup>Tohoku University, Sendai  
<sup>45</sup>University of Tokyo, Tokyo  
<sup>46</sup>Tokyo Institute of Technology, Tokyo  
<sup>47</sup>Tokyo Metropolitan University, Tokyo  
<sup>48</sup>Tokyo University of Agriculture and Technology, Tokyo  
<sup>49</sup>Toyama National College of Maritime Technology, Toyama  
<sup>50</sup>University of Tsukuba, Tsukuba  
<sup>51</sup>Utkal University, Bhubaneswer  
<sup>52</sup>Virginia Polytechnic Institute and State University, Blacksburg, Virginia  
<sup>53</sup>Yokkaichi University, Yokkaichi  
<sup>54</sup>Yonsei University, Seoul

(Received 13 February 2002; published 29 August 2002)

This article describes an observation of mixing-induced  $CP$  violation and a measurement of the  $CP$  violation parameter,  $\sin 2\phi_1$ , with the Belle detector at the KEKB asymmetric  $e^+e^-$  collider. Using a data sample of  $29.1 \text{ fb}^{-1}$  recorded on the  $\Upsilon(4S)$  resonance that contains 31.3 million  $B\bar{B}$  pairs, we reconstruct decays of neutral  $B$  mesons to the following  $CP$  eigenstates:  $J/\psi K_S^0$ ,  $\psi(2S)K_S^0$ ,  $\chi_{c1}K_S^0$ ,  $\eta_c K_S^0$ ,  $J/\psi K_L^0$  and  $J/\psi K^{*0}$ . The flavor of the accompanying  $B$  meson is identified by combining information from primary and secondary leptons,  $K^\pm$  mesons,  $\Lambda$  baryons, slow and fast pions. The proper-time interval between the two  $B$  meson decays is determined from the distance between the two decay vertices measured with a silicon vertex detector. The result  $\sin 2\phi_1 = 0.99 \pm 0.14(\text{stat}) \pm 0.06(\text{syst})$  is obtained by applying a maximum likelihood fit to the 1137 candidate events. We conclude that there is large  $CP$  violation in the neutral  $B$  meson system. A zero value for  $\sin 2\phi_1$  is ruled out by more than six standard deviations.

DOI: 10.1103/PhysRevD.66.032007

PACS number(s): 11.30.Er, 12.15.Hh, 13.25.Hw

## I. INTRODUCTION

The phenomenon of  $CP$  violation is one of the major unresolved issues in our understanding of particle physics today. In 1973, Kobayashi and Maskawa (KM) proposed a model where  $CP$  violation is accommodated as an irreducible complex phase in the weak-interaction quark mixing matrix [1], which is defined as

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \quad (1)$$

where the nontrivial complex phases are conventionally assigned to the furthest off-diagonal elements  $V_{ub}$  and  $V_{td}$ . Unitarity of this CKM matrix (Cabibbo-Kobayashi-Maskawa

matrix) implies that  $\sum_i V_{ij}V_{ik}^* = \delta_{jk}$ , which gives the following relation involving  $V_{ub}$  and  $V_{td}$ :

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \quad (2)$$

This expression can be visualized as a closed triangle in the complex plane as shown in Fig. 1.

The three interior angles of the unitarity triangle originate from the nonvanishing  $CP$ -violating phase (the KM phase) and are defined as [2]

$$\phi_1 \equiv \pi - \arg(-V_{td}V_{tb}^*/-V_{cd}V_{cb}^*),$$

$$\phi_2 \equiv \arg(V_{td}V_{tb}^*/-V_{ud}V_{ub}^*), \quad (3)$$

$$\phi_3 \equiv \arg(V_{ud}V_{ub}^*/-V_{cd}V_{cb}^*).$$

\*On leave from University of Toronto, Toronto, Ontario.

The bold ansatz of Kobayashi and Maskawa required the existence of six quarks at a time when only the  $u$ ,  $d$  and  $s$  quarks were known. The subsequent discoveries of the  $c$ ,  $b$  and  $t$  quarks as well as the consistency with the  $CP$  violation observed in the neutral kaon system led to the incorporation of the KM mechanism as an essential component of the standard model (SM).

In 1980, Sanda, Bigi and Carter pointed out that the KM model contained the possibility of sizable  $CP$ -violating asymmetries in certain neutral  $B$  meson decays [3]. The subsequent observation of a long  $b$  quark lifetime [4] and a large mixing in the neutral  $B$  meson system [5] indicated that it would be feasible to measure  $CP$  violation in  $B$  meson decays at an asymmetric  $e^+e^-$  collider at the  $Y(4S)$  energy.

Until recently the only observation of  $CP$  violation was in the neutral kaon system, where the interpretation of results is complicated due to large corrections from the strong interaction. By contrast, these corrections are absent or very small for the aforementioned  $CP$  violation in the neutral  $B$  meson system. Thus its measurement can be used to over-constrain and test the consistency of the SM.

A pair of neutral  $B$  mesons created in the decay  $Y(4S) \rightarrow B^0\bar{B}^0$  is in a state with  $C = -1$  at the time of production ( $t=0$ ), where  $C$  denotes the charge conjugation. Although oscillation then starts, the state preserves the  $C$  odd configuration and is not allowed to be  $B^0B^0$  or  $\bar{B}^0\bar{B}^0$ . The time evolution of the pair is given by

$$|\Psi(t)\rangle = e^{-t/\tau_{B^0}}|\Psi(t=0)\rangle, \quad (4)$$

$$|\Psi(t=0)\rangle = \frac{1}{\sqrt{2}}[|B^0(\vec{k})\bar{B}^0(-\vec{k})\rangle - |B^0(-\vec{k})\bar{B}^0(\vec{k})\rangle],$$

where  $\vec{k}$  and  $-\vec{k}$  are the  $B$  mesons' momenta in the  $Y(4S)$  rest frame. This coherence is preserved until one  $B$  meson decays. Hence, if we can determine the flavor and the decay time  $t_{\text{tag}}$  of one of the  $B$  mesons decaying into a final state  $f_{\text{tag}}$ , we are able to determine the time-dependent decay amplitude of the other  $B$  at any time  $t$  as a function of the time difference  $t - t_{\text{tag}}$ . We consider the case where the other  $B$  meson decays at  $t = t_{CP}$  to a  $CP$  eigenstate,  $f_{CP}$ . When the decay is dominated by a single transition amplitude, the following formulas for the decay rates hold to a good approximation [3]:

$$R(B^0 \rightarrow f_{CP}; \Delta t) = e^{-|\Delta t|/\tau_{B^0}}/2\tau_{B^0}[1 + \xi_f \sin 2\phi_{CP} \sin(\Delta m_d \Delta t)], \quad (5)$$

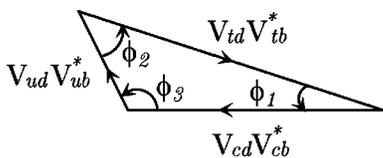


FIG. 1. The unitarity triangle relevant to  $B$  decays. Angles of the triangle arise from the KM phase.

$$R(\bar{B}^0 \rightarrow f_{CP}; \Delta t) = e^{-|\Delta t|/\tau_{B^0}}/2\tau_{B^0}[1 - \xi_f \sin 2\phi_{CP} \sin(\Delta m_d \Delta t)],$$

and the time-dependent  $CP$ -violating asymmetry is [6]

$$A(\Delta t) \equiv \frac{R(\bar{B}^0 \rightarrow f_{CP}; \Delta t) - R(B^0 \rightarrow f_{CP}; \Delta t)}{R(\bar{B}^0 \rightarrow f_{CP}; \Delta t) + R(B^0 \rightarrow f_{CP}; \Delta t)} = -\xi_f \sin 2\phi_{CP} \sin(\Delta m_d \Delta t), \quad (6)$$

where  $\xi_f$  is the  $CP$  eigenvalue of  $f_{CP}$ ,  $\Delta m_d$  is the mass difference between the two  $B^0$  meson mass eigenstates [7] and  $\Delta t = t_{CP} - t_{\text{tag}}$ . Because the asymmetry,  $A(\Delta t)$ , vanishes in the time-integrated rate, it is very important to measure the time dependence.

The angle  $\phi_{CP}$  is directly related to the interior angles of the unitarity triangle, and is the phase difference between two interfering amplitudes, one for  $B^0(\bar{B}^0) \rightarrow f_{CP}$  and the other for the mixing process  $B^0(\bar{B}^0) \rightarrow \bar{B}^0(B^0) \rightarrow f_{CP}$ . The quantity  $\phi_{CP}$  is equal to  $\phi_1$  if  $f_{CP} = J/\psi K_S^0$  or any other  $CP$  eigenstate that arises from a  $b(\bar{b}) \rightarrow c\bar{c}s(\bar{s})$  transition. The hadronic uncertainty in this case is negligibly small because the amplitude of the  $b \rightarrow s$  flavor-changing transition with associated  $c\bar{c}$  production is not only small but has the same weak phase.

As is described in the next section, the KEKB  $e^+e^-$  collider produces the  $Y(4S)$  with a Lorentz boost of  $\beta\gamma = 0.425$ . Since the  $B^0$  and  $\bar{B}^0$  are nearly at rest in the  $Y(4S)$  center of mass system (cms),  $\Delta t$  can be determined from the displacement between the two  $B$  decay vertices—i.e.

$$\Delta t \approx (z_{CP} - z_{\text{tag}})/\beta\gamma c \equiv \Delta z/\beta\gamma c, \quad (7)$$

where the  $z$  axis is defined to be anti-parallel to the positron beam direction.

Following initial experimental studies [8,9], the BaBar [10] and Belle [11] Collaborations recently reported the first clear observations of  $CP$  violation in the neutral  $B$  meson system. In this paper we describe the details of the measurement of  $\sin 2\phi_1$  with the Belle detector at the KEKB asymmetric  $e^+e^-$  collider with the same  $29.1 \text{ fb}^{-1}$  data sample reported in Ref. [11]. In Sec. II we describe the KEKB collider and the Belle detector. The measurement of  $\sin 2\phi_1$  requires the reconstruction of  $B^0 \rightarrow f_{CP}$  decays (denoted by  $B_{CP}$ ), the determination of the  $b$  flavor of the accompanying (tagging)  $B$  meson, the measurement of  $\Delta t$ , and a fit of the expected  $\Delta t$  distribution to the measured distribution using a likelihood method. The selection and tagging procedures are described in Secs. III and IV. After introducing the methods to extract  $\sin 2\phi_1$  from the  $\Delta t$  distributions in Sec. V, we present the results of the fit and discuss the interpretation of the  $CP$  violation in Sec. VI. We summarize the results in Sec. VII.

## II. EXPERIMENTAL APPARATUS

KEKB [12] is an asymmetric  $e^+e^-$  collider 3 km in circumference, which consists of 8 GeV  $e^-$  and 3.5 GeV  $e^+$  storage rings and an injection linear accelerator. It has a single interaction point (IP) where the  $e^+$  and  $e^-$  collide with a crossing angle of 22 mrad. The data used in this analysis were taken between January 2000 and July 2001. The collider was operated during this period with a peak beam current of 930 mA ( $e^+$ ) and 780 mA ( $e^-$ ), giving a peak luminosity of  $4.5 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ . Because of the energy asymmetry, the  $Y(4S)$  resonance and its daughter  $B$  mesons are produced with a Lorentz boost of  $\beta\gamma=0.425$ . On average, the  $B$  mesons decay approximately 200  $\mu\text{m}$  from the  $Y(4S)$  production point.

The Belle detector [13] is a general-purpose large solid angle magnetic spectrometer surrounding the interaction point. It consists of a barrel, forward and rear components. It is placed in such a way that the axis of the detector solenoid is parallel to the  $z$  axis. In this way we minimize the Lorentz force on the low energy positron beam.

Precision tracking and vertex measurements are provided by a central drift chamber (CDC) [14] and a silicon vertex detector (SVD) [15]. The CDC is a small-cell cylindrical drift chamber with 50 layers of anode wires, including 18 layers of stereo wires. A low- $Z$  gas mixture [He (50%) and  $\text{C}_2\text{H}_6$  (50%)] is used to minimize multiple Coulomb scattering to ensure a good momentum resolution, especially for low momentum particles. It provides three-dimensional trajectories of charged particles in the polar angle region  $17^\circ < \theta < 150^\circ$  in the laboratory frame, where  $\theta$  is measured with respect to the  $z$  axis. The SVD consists of three layers of double-sided silicon strip detectors arranged in a barrel and covers 86% of the solid angle. The three layers at radii of 3.0, 4.5 and 6.0 cm surround the beam pipe, a double-wall beryllium cylinder of 2.3 cm radius and 1 mm thickness. The strip pitches are 84  $\mu\text{m}$  for the measurement of  $z$  coordinate and 25  $\mu\text{m}$  for the measurement of azimuthal angle  $\phi$ . The impact parameter resolution for reconstructed tracks is measured as a function of the track momentum  $p$  (measured in GeV/ $c$ ) to be  $\sigma_{xy} = [19 \oplus 50/(p\beta \sin^{3/2}\theta)] \mu\text{m}$  and  $\sigma_z = [36 \oplus 42/(p\beta \sin^{5/2}\theta)] \mu\text{m}$ . The momentum resolution of the combined tracking system is  $\sigma_{p_t}/p_t = (0.30/\beta \oplus 0.19p_t)\%$ , where  $p_t$  is the transverse momentum in GeV/ $c$ .

The identification of charged pions and kaons uses three detector systems: the CDC measurements of  $dE/dx$ , a set of time-of-flight counters (TOF) [16] and a set of aerogel Cherenkov counters (ACC) [17]. The CDC measures energy loss for charged particles with a resolution of  $\sigma(dE/dx) = 6.9\%$  for minimum-ionizing pions. The TOF consists of 128 plastic scintillators viewed on both ends by fine-mesh photomultipliers that operate stably in the 1.5 T magnetic field. Their time resolution is 95 ps (rms) for minimum-ionizing particles, providing three standard deviation ( $3\sigma$ )  $K^\pm/\pi^\pm$  separation below 1.0 GeV/ $c$ , and  $2\sigma$  up to 1.5 GeV/ $c$ . The ACC consists of 1188 aerogel blocks with refractive indices between 1.01 and 1.03 depending on the polar angle. Fine-mesh photomultipliers detect the Cherenkov light. The effective number of photoelectrons is approximately 6 for  $\beta=1$

particles. Using this information,  $P(K/\pi) = \text{Prob}(K)/[\text{Prob}(K) + \text{Prob}(\pi)]$ , the probability for a particle to be a  $K^\pm$  meson, is calculated. A selection with  $P(K/\pi) > 0.6$  retains about 90% of the charged kaons with a charged pion misidentification rate of about 6%.

Photons and other neutrals are reconstructed in a CsI(Tl) crystal calorimeter (ECL) [18] consisting of 8736 crystal blocks, 16.1 radiation lengths ( $X_0$ ) thick. Their energy resolution is 1.8% for photons above 3 GeV. The ECL covers the same angular region as the CDC. Electron identification in Belle is based on a combination of  $dE/dx$  measurements in the CDC, the response of the ACC, the position and the shape of the electromagnetic shower, as well as the ratio of the cluster energy to the particle momentum [19]. The electron identification efficiency is determined from two-photon  $e^+e^- \rightarrow e^+e^-e^+e^-$  processes to be more than 90% for  $p > 1.0$  GeV/ $c$ . The hadron misidentification probability, determined using tagged pions from inclusive  $K_S^0 \rightarrow \pi^+\pi^-$  decays, is below 0.5%.

All the detectors mentioned above are inside a superconducting solenoid of 1.7 m radius that generates a 1.5 T magnetic field. The outermost spectrometer subsystem is a  $K_L^0$  and muon detector (KLM) [20], which consists of 14 layers of iron absorber (4.7 cm thick) alternating with resistive plate counters (RPC). The KLM system covers polar angles between 20 and 155 degrees. The overall muon identification efficiency, determined by using a two-photon process  $e^+e^- \rightarrow e^+e^-\mu^+\mu^-$  and simulated muons embedded in  $B\bar{B}$  candidate events, is greater than 90% for tracks with  $p > 1$  GeV/ $c$  detected in the CDC. The corresponding pion misidentification probability, determined using  $K_S^0 \rightarrow \pi^+\pi^-$  decays, is less than 2%.

In our analysis, Monte Carlo (MC) events are generated using the  $QQ$  event generator [21] and the response of the Belle detector is precisely simulated by a GEANT3-based program [22]. The simulated events are then reconstructed and analyzed with the same procedure as is used for the real data.

## III. RECONSTRUCTION OF $B^0$ DECAYS

We use a  $29.1 \text{ fb}^{-1}$  data sample, which contains 31.3 million  $B\bar{B}$  pairs, accumulated at the  $Y(4S)$  resonance between January 2000 and July 2001. The entire data sample has been analyzed and reconstructed with the same procedure.

We reconstruct  $B^0$  decays to the following  $CP$  eigenstates [23]:  $J/\psi K_S^0$ ,  $\psi(2S)K_S^0$ ,  $\chi_{c1}K_S^0$  and  $\eta_c K_S^0$  having  $\xi_f = -1$ ; and  $J/\psi K_L^0$  having  $\xi_f = +1$ . We also use the decay  $B^0 \rightarrow J/\psi K^{*0}$ ,  $K^{*0} \rightarrow K_S^0 \pi^0$ , which is a mixture of even and odd  $CP$  eigenstates. The selection of these  $B_{CP}$  candidates is described in the following sections.

### A. $B\bar{B}$ event pre-selection

To select generic  $B\bar{B}$  candidates, we require at least three tracks that satisfy  $\sqrt{x^2+y^2} < 2.0$  cm,  $|z| < 4.0$  cm, and  $p_t > 0.1$  GeV/ $c$ , where  $x, y, z$  represent the point of closest

approach of the track to the beam axis, and  $p_t$  is the momentum of the track projected onto the  $xy$  plane. We also require that more than one neutral cluster is observed and have energy greater than 0.1 GeV.

The sum of all cluster energies, boosted back to the cms assuming each cluster is generated by a massless particle, is required to be between 10% and 80% of the total cms energy. The total visible energy in the cms,  $E_{\text{vis}}^{\text{cms}}$ , is computed from the selected tracks, assuming they are pions, and the calorimeter clusters that are not associated with the tracks. We require that  $E_{\text{vis}}^{\text{cms}}$  is greater than 20% of the total cms energy. The absolute value of the  $z$  component of the cms momentum is required to be less than 50% of the cms energy. The event vertex reconstructed from the selected tracks must be within 1.5 cm and 3.5 cm of the interaction region in the directions perpendicular and parallel to the  $z$  axis, respectively. Monte Carlo simulation shows that the selection criteria described above retain more than 99% of  $B\bar{B}$  events and  $J/\psi$  inclusive events.

To suppress continuum background, which consists of  $e^+e^- \rightarrow q\bar{q}$  where  $q$  is a  $u$ ,  $d$ ,  $s$  or  $c$  quark, we also require  $R_2 \equiv H_2/H_0 \leq 0.5$ , where  $H_2$  and  $H_0$  are the second and zeroth Fox-Wolfram moments [24].

### B. $B^0 \rightarrow \text{charmonium } K_S^0(K^{*0})$ reconstruction

The candidate  $J/\psi$  and  $\psi(2S)$  mesons are reconstructed using their decays to lepton pairs, i.e.  $J/\psi \rightarrow \mu^+\mu^-$  and  $e^+e^-$ . The  $\psi(2S)$  meson is also reconstructed via its  $J/\psi\pi^+\pi^-$  decay, the  $\chi_{c1}$  meson via its  $J/\psi\gamma$  decay, and the  $\eta_c$  meson via its  $K^+K^-\pi^0$  and  $K_S^0(\pi^+\pi^-)K^\pm\pi^\mp$  decays.

For  $J/\psi$  and  $\psi(2S) \rightarrow l^+l^-$  decays, we use oppositely charged track pairs where both tracks are positively identified as leptons. For the  $B^0 \rightarrow J/\psi K_S^0(\pi^+\pi^-)$  mode, which has the smallest background fraction among the  $CP$  eigenstates that are used, the requirement for *one* of the tracks is relaxed to improve the efficiency: a track with an ECL energy deposit consistent with a minimum ionizing particle is accepted as a muon and a track that satisfies either the  $dE/dx$  or the ECL shower energy requirements as an electron. In order to remove either badly measured tracks or tracks that do not come from the interaction region, we require  $|dz| < 5$  cm for both lepton tracks. In order to account partially for final-state radiation and bremsstrahlung, the invariant mass calculation of the  $e^+e^-$  pairs is corrected by adding photons found within 50 mrad of the  $e^+$  or  $e^-$  direction. Nevertheless, a radiative tail remains and we use an asymmetric invariant mass requirement  $-150 \leq M_{e^+e^-} - M_{J/\psi[\psi(2S)]} \leq 36$  MeV/ $c^2$ . Since the  $\mu^+\mu^-$  radiative tail is smaller, we select  $-60 \leq M_{\mu^+\mu^-} - M_{J/\psi[\psi(2S)]} \leq 36$  MeV/ $c^2$  [25]. Events with a candidate  $J/\psi \rightarrow l^+l^-$  decay are accepted if the  $J/\psi$  momentum in the cms is below 2 GeV/ $c$ . Figure 2 shows the invariant mass distributions for  $J/\psi \rightarrow \mu^+\mu^-$  and  $J/\psi \rightarrow e^+e^-$  with the selection criteria applied for the  $J/\psi K_S^0$  mode.

To reconstruct  $\psi(2S) \rightarrow J/\psi\pi^+\pi^-$  decays, we select  $\pi^+\pi^-$  pairs with an invariant mass greater than 400 MeV/ $c^2$ . This requirement is based on the measured

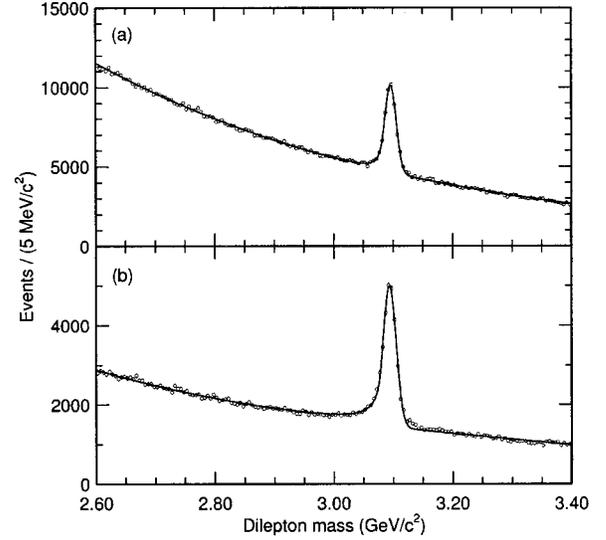


FIG. 2. The invariant mass distributions for (a)  $J/\psi \rightarrow \mu^+\mu^-$  and (b)  $J/\psi \rightarrow e^+e^-$  in the reconstruction of the  $J/\psi K_S^0$  mode where the selection for *one* of the tracks is relaxed (details are explained in the text).

mass distribution [26] and improves the signal-to-background ratio. The  $\psi(2S)$  candidates are then selected requiring the mass difference,  $M_{l^+l^-\pi^+\pi^-} - M_{l^+l^-}$ , to be between 0.58 GeV/ $c^2$  and 0.60 GeV/ $c^2$ . This corresponds to a  $\pm 3\sigma$  requirement where  $\sigma$  is the resolution on the mass difference [27].

The  $\chi_{c1} \rightarrow J/\psi\gamma$  candidates are selected by requiring the mass difference,  $M_{l^+l^-\gamma} - M_{l^+l^-}$ , to be between 0.385 GeV/ $c^2$  and 0.4305 GeV/ $c^2$ . We veto photon candidates that form a good  $\pi^0$  candidate with any other photon candidate of energy greater than 60 MeV in the event. A good  $\pi^0$  candidate is defined by an invariant mass within  $-28$  to  $+17$  MeV/ $c^2$  of the nominal  $\pi^0$  mass, and by a  $\chi^2$  of less than 10 after a mass-constrained kinematic fit [27].

For  $K_S^0 \rightarrow \pi^+\pi^-$  reconstruction, we select oppositely charged track pairs that satisfy the following requirements: (1) when both pions have associated SVD hits, the distance of the closest approach of both pion tracks in the  $z$  direction should be smaller than 1 cm; (2) when only one of the two pions has associated SVD hits, the distance of closest approach of both the pion tracks to the nominal interaction point in the  $x$ - $y$  plane should be larger than 0.25 mm; (3) when neither pion has an associated SVD hit, the  $\phi$  coordinate of the  $\pi^+\pi^-$  vertex and the  $\phi$  direction of the  $\pi^+\pi^-$  candidate's three-momentum vector should agree within 0.1 rad. The invariant mass of the candidate  $\pi^+\pi^-$  pair is required to be between 482 and 514 MeV/ $c^2$ , which retains 99.7% of the  $K_S^0$  candidates.

For the  $\chi_{c1} K_S^0$  and  $\eta_c K_S^0$  modes, more stringent track selection criteria are applied in  $K_S^0 \rightarrow \pi^+\pi^-$  reconstruction to reduce the background: (1) the flight length in the  $r$ - $\phi$  plane should be greater than 1 mm (2 mm for  $\eta_c \rightarrow K^+K^-\pi^0$ ); (2) a mismatch in the  $z$  direction at the  $K_S^0$  vertex point for two  $\pi^\pm$  tracks should be less than 2.5 cm (10 cm for  $\eta_c K_S^0$ ); (3) the angle in the  $r$ - $\phi$  plane between the  $K_S^0$  momentum vector

and the direction defined by the  $K_S^0$  and  $J/\psi$  (or  $\eta_c$ ) decay vertices should be less than 0.2 (0.1 for  $\eta_c K_S^0$ ) radian; and (4) for the  $\chi_{c1} K_S^0$  selection we also require that the distance of closest approach of the  $J/\psi$  vertex in the radial direction for each  $\pi^\pm$  track should be greater than 0.25 mm.

To reconstruct  $K_S^0 \rightarrow \pi^0 \pi^0$  candidates, we first select photons that have an energy of at least 20 MeV. For  $\pi^0 \rightarrow \gamma\gamma$  candidates, we require that the invariant mass of the two photons be between 80 MeV/ $c^2$  and 150 MeV/ $c^2$  and the momentum of  $\pi^0$  be greater than 100 MeV/ $c$ . Initially, the  $\pi^0$  decay vertex is assumed to be at the nominal interaction point. To select  $K_S^0 \rightarrow \pi^0 \pi^0$  candidates, we find the best decay vertex where the invariant masses of two  $\pi^0$  candidates are the most consistent with the nominal  $\pi^0$  mass. To this end, first the  $K_S^0$  flight direction is measured from the sum of the momenta of the four photons, then we calculate the  $\chi^2$  of the mass constrained fit for each  $\pi^0$ , varying the decay vertex along the  $K_S^0$  direction through the IP. We choose the vertex point that minimizes the sum of the  $\chi^2$  for the two  $\pi^0$  candidates. We require that the distance between the IP and the reconstructed  $K_S^0$  decay vertex be larger than  $-20$  cm where the positive direction is defined by the  $K_S^0$  momentum. To reduce the combinatorial background, the  $\chi^2$  for each  $\pi^0$  meson at the  $K_S^0$  decay vertex point determined by this method is also required to be less than 10. Using the calculated  $K_S^0$  decay vertex, we finally require that the invariant mass of  $K_S^0$  candidate lie between 470 MeV/ $c^2$  and 520 MeV/ $c^2$ .

For  $J/\psi K^{*0} (K_S^0 \pi^0)$  decays, we use  $K_S^0 \pi^0$  combinations that have an invariant mass within 75 MeV/ $c^2$  of the nominal  $K^{*0}$  mass. Here, the  $\pi^0$  candidate is reconstructed from photons with an energy greater than 40 MeV and the two photon invariant mass is required to be between 125 and 145 MeV/ $c^2$ . We reduce the background from low-momentum  $\pi^0$  mesons by requiring  $\cos \theta_{K^{*0}} < 0.8$ , where  $\theta_{K^{*0}}$  is the angle between the  $K^{*0}$  flight direction and the  $K_S^0$  momentum vector calculated in the  $K^{*0}$  rest frame.

We reconstruct  $B^0 \rightarrow \eta_c K_S^0$  candidates in two  $\eta_c$  decay modes:  $\eta_c \rightarrow K_S^0 K^\pm \pi^\mp$  and  $\eta_c \rightarrow K^+ K^- \pi^0$ . We require charged kaons to be positively identified using CDC  $dE/dx$  measurements and information from the TOF and ACC systems. For the  $\eta_c \rightarrow K_S^0 K^\pm \pi^\mp$  channel, we require an invariant mass ranging from 2.935 to 3.035 GeV/ $c^2$ . In order to suppress the continuum background, we require  $R_2 < 0.45$  and  $|\cos \theta_{\text{thr}}| < 0.85$ , where  $\theta_{\text{thr}}$  is the angle between the thrust axis of the  $B^0$  candidate and that of all remaining charged and neutral particles in the event. In the  $\eta_c \rightarrow K^+ K^- \pi^0$  mode, we reconstruct the  $\pi^0$  meson from photons having an energy larger than 50 (200) MeV in the ECL barrel (end-cap) region. The invariant mass of the  $\eta_c \rightarrow K^+ K^- \pi^0$  candidate is required to be between 2.890 and 3.040 GeV/ $c^2$ . The continuum background is suppressed by requiring  $R_2 < 0.40$  and  $|\cos \theta_{\text{thr}}| < 0.80$ .

For  $B^0$  reconstruction, we calculate the energy difference,  $\Delta E$ , and the beam-energy constrained mass,  $M_{\text{bc}}$ . The energy difference is defined as  $\Delta E \equiv E_B^{\text{cms}} - E_{\text{beam}}^{\text{cms}}$  and the beam-energy constrained mass  $M_{\text{bc}} \equiv \sqrt{(E_{\text{beam}}^{\text{cms}})^2 - (p_B^{\text{cms}})^2}$ ,

where  $E_{\text{beam}}^{\text{cms}}$  is the cms beam energy, and  $E_B^{\text{cms}}$  and  $p_B^{\text{cms}}$  are the cms energy and momentum of the  $B^0$  candidate. To improve the momentum resolution, a vertex fit and then a mass-constrained fit are performed wherever needed. The resulting fitted momenta are used in the  $\Delta E$  and  $M_{\text{bc}}$  calculations. A scatter plot of  $M_{\text{bc}}$  and  $\Delta E$  for  $J/\psi K_S^0 (\pi^+ \pi^-)$  candidates is shown in Fig. 3 along with the projections onto each axis. The  $B^0$  candidates are selected by requiring  $5.270 < M_{\text{bc}} < 5.290$  GeV/ $c^2$  ( $|M_{\text{bc}} - M_{B^0}| < 3.5\sigma$ ) and by applying the mode-dependent requirements on  $\Delta E$  listed in Table I. Figure 4 shows the  $M_{\text{bc}}$  distribution after the  $\Delta E$  selection.

The background in the signal region is estimated by simultaneously fitting the  $M_{\text{bc}}$  and  $\Delta E$  distributions with signal and background functions in the region  $5.2 < M_{\text{bc}} < 5.3$  GeV/ $c^2$  and  $-0.1 < \Delta E < 0.2$  GeV. We use a two-dimensional Gaussian for the signal. For the background, we use the ARGUS background function [28] for  $M_{\text{bc}}$  and a linear function for  $\Delta E$ . The details are described in Sec. V D. The number of candidates observed as well as the estimated background are given in Table I.

### C. $B^0 \rightarrow J/\psi K_L^0$ reconstruction

The reconstruction of  $B^0 \rightarrow J/\psi K_L^0$  is an experimental challenge but is very important because its yield is expected to be large. In addition, since this mode is a  $CP$ -even eigenstate, we should observe a time-dependent asymmetry reversed in sign compared to  $J/\psi K_S^0$ , which provides an important experimental consistency check.

While the detached vertex and invariant mass of the  $K_S^0$  provide significant background reduction for  $J/\psi K_S^0$ , the background is larger for  $J/\psi K_L^0$  as only the  $K_L^0$  direction is measured. Since the energy of the  $K_L^0$  is not measured,  $M_{\text{bc}}$  and  $\Delta E$  cannot be used as the final kinematical variables to identify  $B^0$  candidates as in other final states. Using the four-momentum of a reconstructed  $J/\psi$  candidate and the  $K_L^0$  flight direction, we calculate the momentum of the  $K_L^0$  candidate requiring  $\Delta E = 0$ . We then calculate  $p_B^{\text{cms}}$  which is used for the final selection.

The selection criteria are necessarily tighter than those used for the  $J/\psi K_S^0$  candidates. However, precise determination of the  $K_L^0$  flight direction with the KLM and ECL allows us to reconstruct  $J/\psi K_L^0$  candidates with sufficient efficiency and purity.

We use tracks which are positively identified as electrons (muons) in the identification of  $J/\psi \rightarrow e^+ e^-$  ( $\mu^+ \mu^-$ ) candidates. We require the invariant mass of the lepton pair to lie in the range  $3.05 < M_{l+l-} < 3.13$  GeV/ $c^2$ . The radiative photon correction for electron pairs is made in the same way as for the other modes. Events are rejected if one of the following decay modes are exclusively reconstructed, and satisfy  $|\Delta E| < 0.05$  GeV and  $5.27 < M_{\text{bc}} < 5.29$  GeV/ $c^2$ :  $J/\psi K^+$ ,  $J/\psi K_S^0$ ,  $J/\psi K^{*+}$  ( $K^{*+} \rightarrow K^+ \pi^0, K_S^0 \pi^+$ ), and  $J/\psi K^{*0}$  ( $K^{*0} \rightarrow K^+ \pi^-, K_S^0 \pi^0$ ).

We select  $K_L^0$  candidates based on the KLM and ECL information. There are two classes of  $K_L^0$  candidates that we refer to as KLM and ECL candidates. To select the KLM

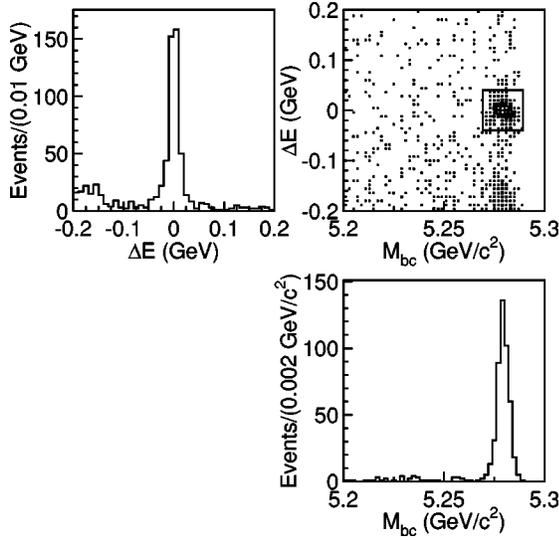


FIG. 3. The scatter plot of  $\Delta E$  versus  $M_{bc}$  for  $J/\psi K_S^0(\pi^+\pi^-)$  candidates. The box represents the signal region. The upper left figure is the  $\Delta E$  projection with  $5.270 < M_{bc} < 5.290$   $\text{GeV}/c^2$ . The lower right figure is the  $M_{bc}$  projection with  $|\Delta E| < 0.04$   $\text{GeV}$ . The enhancement in the negative  $\Delta E$  region is due to decay modes with additional pions, e.g.  $B \rightarrow J/\psi K^*$ ,  $K^* \rightarrow K_S^0 \pi$ .

candidates, a cluster of KLM hits is formed by combining the hits within a  $5^\circ$  opening angle. We require hits in two or more KLM layers and calculate the center of the KLM cluster. If there is an ECL cluster with energy greater than 0.16  $\text{GeV}$  within a  $15^\circ$  cone, we relax our criteria to allow a cluster with a hit in just one KLM layer. In this case the direction of the ECL cluster is taken as the  $K_L^0$  direction. If the cluster lies within a  $15^\circ$  cone of the extrapolation of a charged track to the first layer of the KLM, it is discarded.

ECL candidates are selected from ECL clusters using the following information: the distance between the ECL cluster and the closest charged track position; the ECL cluster energy; the ratio of energies summed in  $3 \times 3$  and  $5 \times 5$  arrays

TABLE I. Summary of  $\Delta E$  signal region, the numbers of signal candidates, and the expected background.

Decay mode	$\Delta E$ cut (MeV)	Signal candidates	Expected background
	Lower to upper		
$B^0 \rightarrow J/\psi K_S^0$			
$K_S^0 \rightarrow \pi^+\pi^-$	-40-40	457	11.9
$K_S^0 \rightarrow \pi^0\pi^0$	-150-100	76	9.4
$B^0 \rightarrow \psi(2S)K_S^0$			
$\psi(2S) \rightarrow l^+l^-$	-40-40	39	1.2
$\psi(2S) \rightarrow J/\psi\pi^+\pi^-$	-40-40	46	2.1
$B^0 \rightarrow \chi_{c1}K_S^0$	-40-40	24	2.4
$B^0 \rightarrow \eta_c K_S^0$			
$\eta_c \rightarrow K_S^0 K^\pm \pi^\mp$	-40-40	23	11.3
$\eta_c \rightarrow K^+ K^- \pi^0$	-60-40	41	13.6
$B^0 \rightarrow J/\psi K^{*0}$	-50-30	41	6.7

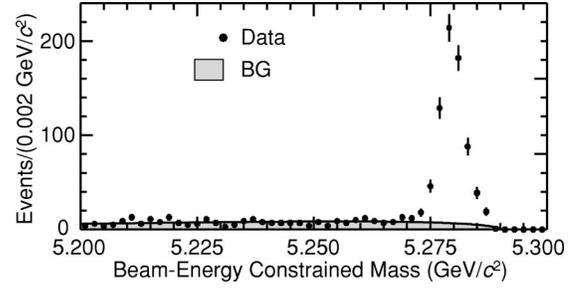


FIG. 4. The  $M_{bc}$  distribution for all exclusive decay modes combined except  $J/\psi K_L^0$ . The  $\Delta E$  cuts are imposed. The shaded area is the estimated background.

of CsI crystals surrounding the crystal at the center of the shower; the ECL shower width and the invariant mass of the shower. After a very loose pre-selection based on the above five discriminants, we calculate signal and background likelihood values for each discriminant based on a  $J/\psi$  inclusive MC. Taking the products of the above five likelihoods for each signal and background, we form the likelihood ratio  $\mathcal{L}_{K_L^0}/(\mathcal{L}_{K_L^0} + \mathcal{L}_{\text{fake}})$ . This likelihood ratio is required to be greater than 0.5.

We examine the characteristics of  $K_L^0$  candidates in both the data and the MC. We obtain consistent distributions for the number of  $K_L^0$  candidates per event, the  $K_L^0$  flight directions in the laboratory system, the total number of hit KLM layers and the number of hit first-layers in the  $K_L^0$  candidates. We investigate the momentum dependence of the KLM response by using charged pions and kaons. The MC simulation reproduces the results obtained with data well. We also use  $e^+e^- \rightarrow \gamma\phi$  followed by  $\phi \rightarrow K_L^0 K_S^0$ , where the exact  $K_L^0$  direction and momentum can be obtained by reconstructing  $\gamma$  and  $K_S^0 \rightarrow \pi^+\pi^-$ . These studies indicate that the  $K_L^0$  identification is well reproduced by the MC with an exception of an overall detection efficiency. The  $K_L^0$  detection efficiency in the data is found to be lower than the MC expectation. This, however, does not cause a difficulty in our analysis since we do not rely on the MC  $K_L^0$  detection efficiency.

For both KLM and ECL candidates, we require that the  $K_L^0$  direction should be within  $45^\circ$  of its expected direction calculated from the  $J/\psi$  candidate momentum assuming that the  $B^0$  candidate was at rest in the cms. We also require that no photon from a  $\pi^0$  decay is found near the  $K_L^0$  candidate. For this requirement, we select  $\pi^0$  candidates satisfying  $0.12 < M_{\gamma\gamma} < 0.15$   $\text{GeV}/c^2$  and with momentum above 0.8 (1.2)  $\text{GeV}/c$  for KLM (ECL) candidates.

To reconstruct  $B^0 \rightarrow J/\psi K_L^0$ , we first use the KLM candidates. If none of the KLM candidates satisfy the selection criteria below, we use the ECL candidates. Thus, the two classes of  $B^0 \rightarrow J/\psi K_L^0$  candidates are mutually exclusive. In order to suppress the background, we calculate a probability density function (PDF) for each of the following variables and then form a product of the PDF's to obtain the combined likelihood: the cms momentum of the  $J/\psi$ ,  $p_{J/\psi}^{\text{cms}}$ ; the angle between the  $K_L^0$  candidate and the closest charged track having a momentum larger than 0.7  $\text{GeV}/c$ ; the Fox-Wolfram moment ratio  $R_2$ ;  $\cos\theta_B$  where  $\theta_B$  is the polar angle of the reconstructed  $B$  in the cms; and the number of charged tracks with  $p_t > 0.1$   $\text{GeV}/c$ ,  $|dr| < 2$   $\text{cm}$ , and  $|dz| < 4$   $\text{cm}$ . In ad-

dition to these five variables, two other variables are included conditionally to reduce the background from  $B^+ \rightarrow J/\psi K^{*+}$ ,  $K^{*+} \rightarrow K_L^0 \pi^+$  decays. One such variable is the cms momentum of the  $J/\psi K_L^0 \pi^+$  system,  $p_B^{\text{cms}}(J/\psi K_L^0 \pi^+)$ , and the other is the momentum of the additional pion. We use  $p_B^{\text{cms}}(J/\psi K_L^0 \pi)$  if the invariant mass of the  $K_L^0$  and a charged track, with the nominal pion mass being assumed, is above  $0.85 \text{ GeV}/c^2$  and below  $0.93 \text{ GeV}/c^2$ , and  $p_B^{\text{cms}}(J/\psi K_L^0 \pi) < 0.8 \text{ GeV}/c$  [29]. The pion momentum is used if the addition of the extra pion results in  $0.2 < p_B^{\text{cms}}(J/\psi K_L^0 \pi) < 0.45 \text{ GeV}/c$  given the requirement above on the  $K_L^0 \pi^+$  invariant mass.

The combined likelihood is calculated for  $J/\psi K_L^0$  (signal) and inclusive  $B \rightarrow J/\psi$  decays (background) using MC samples. Background events from misidentified (fake)  $J/\psi$  mesons are not included in the likelihood construction since their contribution is small. We then form a likelihood ratio  $\mathcal{L}_{J/\psi K_L^0} / (\mathcal{L}_{J/\psi K_L^0} + \mathcal{L}_{\text{bkg}})$  that is used as a discriminant variable. Figure 5 shows the likelihood ratio distributions for the data and Monte Carlo candidates.

We require that the likelihood ratio be larger than 0.4 to identify  $J/\psi K_L^0$  candidates.

Figure 6 shows the resultant  $p_B^{\text{cms}}$  distribution for events that satisfy all the selection criteria. We define the signal region as  $0.2 \leq p_B^{\text{cms}} \leq 0.45$  (0.40)  $\text{GeV}/c$  for  $K_L^0$  candidates identified with the KLM (ECL) criteria [30], and obtain 569 candidates, out of which 397 are KLM candidates and 172 are ECL candidates.

We extract the  $J/\psi K_L^0$  signal yield by fitting the  $p_B^{\text{cms}}$  distribution of the data to a sum of four components: (1) signal; (2) background with  $K_L^0$ ; (3) background without  $K_L^0$ ; and (4) combinatorial  $J/\psi$  mesons. The shapes of the first three components are determined from the  $J/\psi$  inclusive MC sample and look-up tables are used in the fit. The normalizations of these three components are treated as free parameters in the fit to minimize the effect of the aforementioned uncertainty in the  $K_L^0$  detection efficiency in the MC simulation. The combinatorial component is evaluated using events with  $e\mu$  pairs that satisfy the requirements for  $J/\psi$  reconstruction. The shape is modeled by a second-order polynomial. An additional parameter of the fit is an offset in  $p_B^{\text{cms}}$ , allowing the signal shape to shift with respect to the background distribution.

The result of the fit to the  $p_B^{\text{cms}}$  distribution is shown in Fig. 6. By integrating each component obtained by the fit in the signal region, we find a total of  $(346.3 \pm 28.8)$   $J/\psi K_L^0$  signal events, and a signal purity of 61%.

The fitting procedure is applied to the KLM and ECL candidates separately. The  $\chi^2$  of the fit to KLM (ECL) candidates is 54.8 (31.4) for  $ndf = 35$ . Table II summarizes the fit results.

#### D. Control samples of flavor-specific $B$ decays

The reconstruction of flavor-specific  $B$  decays is also a key ingredient in this analysis, since such decays are used for various purposes: evaluation of the performance of flavor

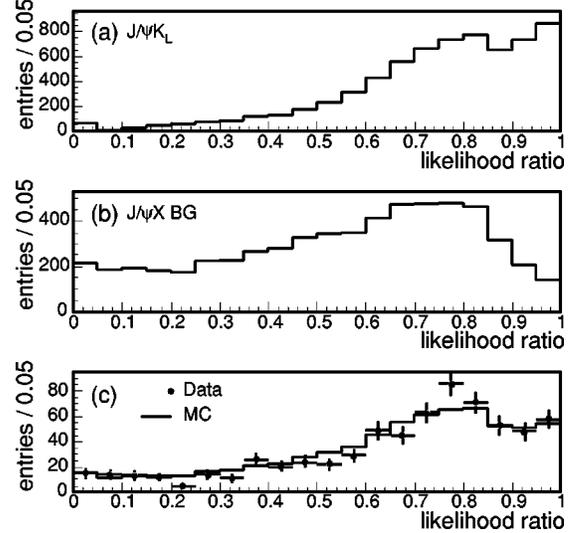


FIG. 5. The likelihood ratio distributions for (a)  $J/\psi K_L^0$  Monte Carlo sample, (b)  $J/\psi + X$  background Monte Carlo sample, and (c) the data in the signal region with Monte Carlo events overlaid.

tagging (Sec. IV B); extraction of the proper-time resolution function (Sec. V B);  $B$  lifetime measurements as a cross check (Sec. VI B 2); and the demonstration of null asymmetries in non- $CP$  final states (Sec. VI B 3). Since a large number of events with high purity are required for these purposes, we use semileptonic decays and hadronic decays from  $b \rightarrow c\bar{u}d$  transitions.

For the semileptonic decays, we use the decay chains  $B^0 \rightarrow D^{*-} l^+ \nu$  and  $D^{*-} \rightarrow \bar{D}^0 \pi^-$ , where  $\bar{D}^0 \rightarrow K^+ \pi^-$ ,  $\bar{D}^0 \rightarrow K^+ \pi^- \pi^0$  and  $\bar{D}^0 \rightarrow K^+ \pi^+ \pi^- \pi^-$ . All tracks used must have associated SVD hits and radial impact parameters  $|dr| < 0.2 \text{ cm}$ , except for the slow pion from the  $D^{*-}$  decay. Candidate  $\bar{D}^0$  decays are selected by requiring the invariant mass to be consistent with the nominal  $\bar{D}^0$  mass. The invariant mass requirement depends on the  $\bar{D}^0$  decay mode, varying from  $+9$  to  $+23 \text{ MeV}/c^2$  above the  $\bar{D}^0$  mass and from  $-9$  to  $-37 \text{ MeV}/c^2$  below the  $\bar{D}^0$  mass, respectively. For the  $D^{*-}$  reconstruction, we also impose a requirement on the mass difference,  $M_{\text{diff}}$ , between a  $D^{*-}$  candidate and the corresponding  $\bar{D}^0$  candidate. We require  $M_{\text{diff}}$  to be within

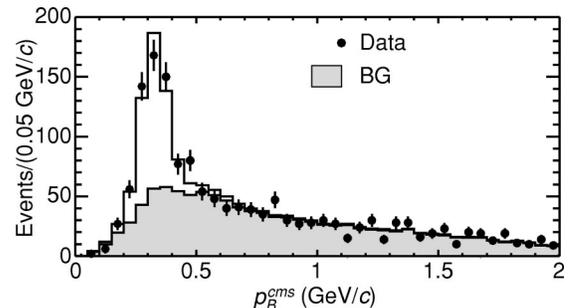


FIG. 6. The  $p_B^{\text{cms}}$  distribution of the  $J/\psi K_L^0$  candidates. The solid line is a sum of the signal and background. The shaded histogram shows the background component only.

TABLE II. Results of the fit to the  $p_B^{\text{cms}}$  distribution of the  $J/\psi K_L^0$  candidates. Yields are for the signal  $p_B^{\text{cms}}$  region.

	KLM	ECL	Sum
Signal	$244.8 \pm 25.1$	$101.5 \pm 14.2$	$346.3 \pm 28.8$
Background			
with $K_L^0$	118.1	26.5	144.6
without $K_L^0$	27.5	38.6	66.1
Combinatorial	6.6	5.4	12.0
Total yield	397	172	569

0.8 to 1.75  $\text{MeV}/c^2$  of the nominal mass difference, depending on the  $\bar{D}^0$  decay mode. In addition, we require the  $D^{*-}$  cms momentum to be less than 2.6  $\text{GeV}/c$  to suppress  $D^{*-}$  mesons from the continuum. The  $D^{*-}$  candidates are combined with  $\mu^+$  or  $e^+$  candidates having the opposite charge to the  $D^{*-}$  candidate. Lepton candidates must satisfy  $1.4 < p_l^{\text{cms}} < 2.4 \text{ GeV}/c$ , where  $p_l^{\text{cms}}$  is the cms momentum of the lepton. We exploit the massless character of the neutrino, to calculate  $M_{\text{miss}}^2$ , the effective missing mass squared in the cms defined by  $M_{\text{miss}}^2 \equiv (E_B^{\text{cms}} - E_{D^{*l}}^{\text{cms}})^2 - |\vec{p}_B^{\text{cms}}|^2 - |\vec{p}_{D^{*l}}^{\text{cms}}|^2$ , and a product of the momenta of the  $B$  and the  $D^{*l}$  system,  $C \equiv 2|\vec{p}_B^{\text{cms}}| |\vec{p}_{D^{*l}}^{\text{cms}}|$ . The cosine of the angle between  $\vec{p}_B^{\text{cms}}$  and  $\vec{p}_{D^{*l}}^{\text{cms}}$  is given by  $-M_{\text{miss}}^2/C$ . In the  $M_{\text{miss}}^2$  versus  $C$  plane, therefore, we select  $D^{*l}\nu$  candidates inside the triangle shown in Fig. 7. The triangle is defined by  $C \geq -(1.45/1.595)M_{\text{miss}}^2$ ,  $C \geq (1.02/1.1)M_{\text{miss}}^2$  and  $C \leq -(0.43/2.695)(M_{\text{miss}}^2 + 1.595) + 1.45$ .

The vertex position of the  $D^{*l}\nu$  candidate is obtained by first fitting the vertex of the  $D$  candidate from its daughter tracks, and then fitting the lepton and  $D$  candidate to obtain the  $B$  vertex. The slow pion track from the  $D^{*}$  meson is also

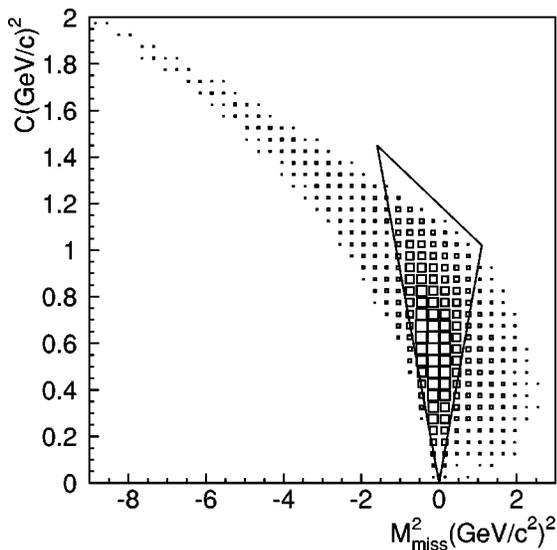


FIG. 7. The distribution of candidate  $B^0 \rightarrow D^{*-} l^+ \nu$  decays in the plane of  $C$  versus  $M_{\text{miss}}^2$ . The signal region (triangle) is also shown.

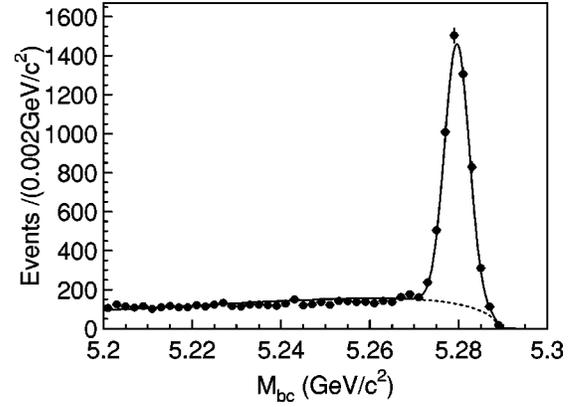


FIG. 8.  $M_{bc}$  distribution for all hadronic decay modes ( $B^0 \rightarrow D^- \pi^+$ ,  $D^{*-} \pi^+$ , and  $D^{*-} \rho^+$ ) with  $|\Delta E| < 50 \text{ MeV}$ . Fit curves for background (dashed) and signal plus background events (solid) are superimposed.

included in the fit. This is important since the re-fitted helix parameters of the slow pion are used to recalculate  $M_{\text{diff}}$  and improve its resolution. The overall signal fraction in these control samples is estimated to be 79.4%. The backgrounds consist of fake  $D^{*}$  mesons (10.4%), incorrect combinations of  $D^{*}$  with leptons that do not show an angular correlation (4.1%, called “uncorrelated background”), continuum events (2.1%) and  $B^\pm$  decays (4%). The fraction of fake leptons in the uncorrelated background is estimated with a Monte Carlo simulation to be 0.5% and is neglected. We estimate the fake  $D^{*}$  background by using events in the  $\bar{D}^0$  mass sideband as well as fake  $D^{*}$  events that are reconstructed with wrong-charge slow pions. We evaluate the uncorrelated background fraction by counting signal events in a sample wherein we flip the candidate lepton momentum vector artificially. In this case the number of uncorrelated background in the signal region remains the same level while the signal events are rejected [31]. We apply the same event selection to off-resonance data ( $2.3 \text{ fb}^{-1}$ ) and count the number of events in the signal region after subtracting the fake  $D^{*}$  background. We estimate the continuum background fraction by scaling the result with the integrated luminosity. We fit the  $C$  distribution with a range  $-10 < C < 1.1$  to estimate the signal and  $B \rightarrow D^{*l}\nu$  background fractions. We use MC information to model the signal and  $B \rightarrow D^{*l}\nu$  distribution. The  $C$  distributions and fractions of all the other backgrounds are obtained from the aforementioned special background samples and are fixed in the fit. We treat the background fraction uncertainties as a source of systematic errors in the determination of wrong tag fractions, which will be discussed in Sec. IV B.

For hadronic decays, we use the decay modes  $B^0 \rightarrow D^- \pi^+$ ,  $D^{*-} \pi^+$  and  $D^{*-} \rho^+$ , where  $D^- \rightarrow K^+ \pi^- \pi^-$ ,  $\rho^+ \rightarrow \pi^+ \pi^0$ . We reconstruct  $D^{*-}$  candidates in the same modes that were used for the  $D^{*-} l^+ \bar{\nu}$  mode. For  $D$  and  $D^{*-}$  candidates we apply mode-dependent requirements on the reconstructed  $D$  mass (ranging from  $\pm 30$  to  $\pm 60 \text{ MeV}/c^2$ ) and  $M_{\text{diff}}$  (ranging from  $\pm 3$  to  $\pm 12 \text{ MeV}/c^2$ ), in a similar way as for  $D^{*-} l^+ \bar{\nu}$  mode. We

TABLE III. Summary of the numbers of flavor-specific decay candidates and their estimated purities.

Decay mode	Candidates	Purity
$B^0 \rightarrow D^{*-} l^+ \nu$	16101	79.4%
$B^0 \rightarrow D^- \pi^+$	2241	85.5%
$B^0 \rightarrow D^{*-} \pi^+$	2126	87.2%
$B^0 \rightarrow D^{*-} \rho^+$	1620	72.2%

select  $\rho^+$  candidates by requiring the  $\pi^+ \pi^0$  invariant mass to be within 150 MeV/ $c^2$  of the nominal  $\rho^+$  mass. In order to suppress continuum background, we impose mode-dependent cuts on  $R_2$  (upper cut values ranging from 0.5 to 1.0) and  $\cos \theta_{\text{thr}}$  (upper cut values ranging from 0.92 to 1.0). The cut values are chosen to maximize the figure of merit  $S/\sqrt{S+B}$  for each mode, where  $S$  and  $B$  are numbers of signal and background, respectively. We select  $B^0$  candidates by requiring  $5.27 < M_{bc} < 5.29$  GeV/ $c^2$  and  $|\Delta E| < 50$  MeV. Background contributions are estimated by fitting the  $M_{bc}$  and  $\Delta E$  distributions in the same way as we do for  $CP$  modes. Figure 8 shows the  $M_{bc}$  distribution for the three control sample modes combined. The overall signal purity is  $\sim 82\%$ . We study background components with a MC sample that includes both  $B\bar{B}$  and continuum events. We find no significant peaking background in the signal region. Therefore we use the ARGUS function to model the background in the  $M_{bc}$  distribution. A possible deviation due to combinatorial backgrounds may introduce ambiguities in the signal fraction determination and the background  $\Delta t$  model parameters of the control sample, resulting in small uncertainties in the final determination of the  $CP$  asymmetry. This is considered as a source of systematic uncertainties.

The numbers of candidates and their purities for flavor-specific samples are summarized in Table III.

#### IV. FLAVOR TAGGING

After the exclusive reconstruction of a neutral  $B$  meson decaying into a final state,  $f_{CP}$ , all the remaining particles should belong to the final state,  $f_{\text{tag}}$ , of the decay of the other  $B$  meson. To observe time-dependent  $CP$  violation, we need to ascertain whether  $f_{\text{tag}}$  is from a  $B^0$  or  $\bar{B}^0$ . This determination is called ‘‘flavor tagging.’’ The simplest and most reliable method for flavor tagging uses the charge of high-momentum leptons in semileptonic decays, i.e.  $B^0 \rightarrow X l^+ \nu$  and  $\bar{B}^0 \rightarrow X l^- \bar{\nu}$ . The charges of final-state kaons can also be used since the decays  $B^0 \rightarrow K^+ X$  (with  $\bar{b} \rightarrow \bar{c} \rightarrow \bar{s}$ ) and  $\bar{B}^0 \rightarrow K^- X$  (with  $b \rightarrow c \rightarrow s$ ) dominate. In addition to these two leading discriminants, our algorithm includes other categories of tracks whose charges depend on the  $b$  quark’s flavor: lower momentum leptons from  $c \rightarrow s l^+ \nu$ ;  $\Lambda$  baryons from the cascade decay  $b \rightarrow c \rightarrow s$ ; high-momentum pions that originate from decays like  $B^0 \rightarrow D^{(*)-} (\pi^+, \rho^+, a_1^+, \text{etc.})$ ; and slow pions from  $D^{*-} \rightarrow \bar{D}^0 \pi^-$ . All these inputs are combined, taking their correlations into account, in a way that maximizes the flavor tagging performance. The perfor-

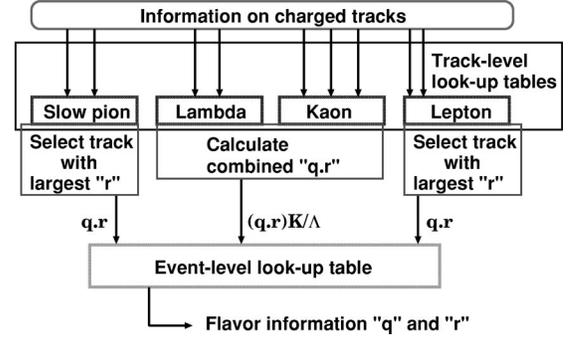


FIG. 9. A schematic diagram of the two-stage flavor tagging. See the text on the definition of the parameters ‘‘ $q$ ’’ and ‘‘ $r$ .’’

mance is characterized by two parameters,  $\epsilon$  and  $w$ . The parameter  $\epsilon$  is the raw tagging efficiency, while  $w$  is the probability that the flavor tagging is wrong (wrong tag fraction). A non-zero value of  $w$  results in a dilution of the true asymmetry. For example, if the true numbers of reconstructed  $B^0$  and  $\bar{B}^0$  are  $n_{B^0}$  and  $n_{\bar{B}^0}$ , the corresponding asymmetry is  $\mathcal{A} = (n_{B^0} - n_{\bar{B}^0}) / (n_{B^0} + n_{\bar{B}^0})$ . With realistic flavor tagging, the observed numbers are  $N_{B^0} = \epsilon[(1-w)n_{B^0} + wn_{\bar{B}^0}]$  for  $B^0$ ,  $N_{\bar{B}^0} = \epsilon[(1-w)n_{\bar{B}^0} + wn_{B^0}]$  for  $\bar{B}^0$ , and the observed asymmetry becomes  $(1-2w)\mathcal{A}$ . Since the statistical error of the measured asymmetry is proportional to  $\epsilon^{-1/2}$ , the number of events required to observe the asymmetry for a certain statistical significance is proportional to  $\epsilon_{\text{eff}} = \epsilon(1-2w)^2$ , which is called the ‘‘effective efficiency.’’ Note that an imperfect knowledge of  $w$  shifts the central value of the measurement and thus represents a potential source of systematic error.

In light of the above, our tagging algorithm has been designed to maximize  $\epsilon_{\text{eff}}$ . Moreover, since  $w$  directly affects the central value of our result, we have developed an approach wherein it can be determined from the data.

In our approach we use two parameters,  $q$  and  $r$ , to represent the tagging information. The first,  $q$ , corresponds to the sign of the  $b$  quark charge where  $q = +1$  for  $\bar{b}$  and hence  $B^0$ , and  $q = -1$  for  $b$  and  $\bar{B}^0$ . The parameter  $r$  is an event-by-event flavor-tagging dilution factor that ranges from  $r = 0$  for no flavor discrimination to  $r = 1$  for unambiguous flavor assignment. The values of  $q$  and  $r$  are determined for each event from a look-up table. Each entry of the table is prepared using a large statistics MC sample and is given by

$$q \cdot r \equiv \frac{N(B^0) - N(\bar{B}^0)}{N(B^0) + N(\bar{B}^0)}, \quad (8)$$

where  $N(B^0)$  and  $N(\bar{B}^0)$  are the numbers of  $B^0$  and  $\bar{B}^0$  in the MC sample.

In this analysis, we sort flavor-tagged events into six bins in  $r$ . For each  $r$  bin, we empirically determine  $w$  directly from data by using control samples, as described in Sec. IV B.

### A. Flavor tagging method

Flavor tagging proceeds in two stages. In the first stage, the flavor tagging information ( $q$  and  $r$ ) provided by each track in the event is calculated. In the second, the track-level results are combined to determine event-level values for  $q$  and  $r$ .

Tracks are sorted into four categories, namely those that resemble leptons, kaons,  $\Lambda$  baryons, and slow pions. For each category, we consider several tagging discriminants, such as track momentum and particle identification information. The value of  $q$  and  $r$  for each track is assigned based on MC-generated look-up tables that take the tagging discriminants as input.

In the second stage, the results from the four track categories are combined to determine the values of  $q$  and  $r$  for each event. Again a look-up table is prepared to provide  $q \cdot r$ .

Figure 9 shows a schematic diagram of the flavor tagging method. The event-level parameter  $r$  should satisfy  $r \simeq 1 - 2w$  where we measure  $w$  from control samples. Using this MC-determined dilution factor  $r$  as a measure of the tagging quality is a straightforward and powerful way of taking into account correlations among various discriminants. Using two stages, we keep the look-up tables small enough to provide sufficient MC statistics for each entry. In the following, we provide additional details about each stage of the flavor tagging.

#### 1. Track-level flavor tagging

We select tracks that do not belong to  $B_{CP}$  and that satisfy  $|dr| < 2$  cm and  $|dz| < 10$  cm. Tracks that are part of a  $K_S^0$  candidate are not used. Each selected tag-side track is examined and assigned to one of the four track categories. Tracks in the lepton category are subdivided into categories for electron-like and muon-like tracks. If the cms momentum,  $p_l^{\text{cms}}$ , of a track is larger than 0.4 GeV/ $c$  and the ratio of its electron and kaon likelihoods is larger than 0.8, the track is assigned to the electron-like category. If a track has  $p_l^{\text{cms}}$  larger than 0.8 GeV/ $c$  and the ratio of its muon and kaon likelihoods is larger than 0.95, it is passed to the muon-like category. The likelihood is calculated by combining the ACC, TOF,  $dE/dx$ , and ECL or KLM information. In the lepton category, leptons from semileptonic  $B$  decays yield the largest effective efficiency. Leptons from  $B \rightarrow D$  cascade decays and high-momentum pions from  $B^0 \rightarrow D^{(*)-} \pi^+ X$  also make a small contribution to this category. We choose the following six discriminants: the track charge; the magnitude of the momentum in the cms,  $p_l^{\text{cms}}$ ; the polar angle in the laboratory frame,  $\theta_{\text{lab}}$ ; the recoil mass,  $M_{\text{recoil}}$ , calculated using all the tag side tracks except the lepton candidate; the magnitude of the missing momentum in the cms,  $P_{\text{miss}}^{\text{cms}}$ ; and the lepton-ID quality value.

The track charge directly provides the  $b$ -flavor  $q$ . The lepton-ID quality distinguishes leptons from pions. Its performance is reinforced by variables  $p_l^{\text{cms}}$  and  $\theta_{\text{lab}}$ , which have distributions that are different for leptons and pions. The variables  $p_l^{\text{cms}}$ ,  $M_{\text{recoil}}$  and  $P_{\text{miss}}^{\text{cms}}$  discriminate between high momentum leptons from semileptonic  $B$  decays and in-

termediate momentum leptons from  $B \rightarrow D$  cascade decays where the  $D$  decays semi-leptonically.

If a track cannot be positively identified as a kaon and its momentum is less than 0.25 GeV/ $c$ , it is assigned to the slow-pion category, since low-momentum pions often come from charged  $D^* \rightarrow D \pi$  decays. Here the discriminant variables are the track charge; the momentum and polar angle in the laboratory frame,  $p_{\text{lab}}$  and  $\theta_{\text{lab}}$ ; the ratio of the electron to  $\pi$  probability from  $dE/dx$ ; and  $\cos \alpha_{\text{thr}}$ , the cosine of the angle between the slow pion candidate and the thrust axis of the tag-side particles in the cms. The main background in this category comes from other (i.e. non- $D^*$  daughter) low momentum pions, electrons from photon conversions and  $\pi^0$  Dalitz decays. To separate slow pions from those electrons we use  $dE/dx$  for this class. Since the direction of the slow pion from a  $D^*$  decay is approximately parallel to the  $D^*$  direction, it is also almost parallel to the thrust axis. The variables  $\cos \alpha_{\text{thr}}$ ,  $p_{\text{lab}}$  and  $\theta_{\text{lab}}$ , thus help to identify the slow pions originating from  $D^*$  decays.

If a track forms a  $\Lambda$  candidate with another track, it is assigned to the  $\Lambda$  category. In that category the discriminant variables are the flavor ( $\Lambda$  or  $\bar{\Lambda}$ ); the invariant mass of the reconstructed  $\Lambda$  candidate; the angle difference between the  $\Lambda$  momentum vector and the direction of the  $\Lambda$  vertex point from the nominal IP; the mismatch in the  $z$  direction of the two tracks at the  $\Lambda$  vertex point; and the proton-ID quality value.

If a track does not fall in any of the categories described above, and is not positively identified as a proton, it is classified as a kaon. The kaon category is subdivided into two parts, one for events with  $K_S^0$  decays, and the other for events without  $K_S^0$ 's. Separate treatment is necessary as events with  $K_S^0$  have a larger wrong tag fraction because of their additional strange quark content. We use the track charge,  $p^{\text{cms}}$ ,  $\theta_{\text{lab}}$  and the probability ratio of kaon to pion as the tagging discriminants. The charge of kaons is the most important discriminant. The other three variables help separate kaons from pions.

Although the discriminating power of high-momentum pions is weaker than that of charged kaons, they do provide some tagging information and are therefore included in the kaon category. Approximately half of the pions with  $p^{\text{cms}} > 1.0$  GeV/ $c$  are included in the kaon category, while the other half falls into the lepton class, mostly in the muon-like category.

As is described in Sec. IV B, we have developed a method wherein the wrong tag fractions in our flavor tagging method are evaluated from data. Thus possible discrepancies between data and MC in the distributions of discriminant variables do not affect our  $\sin 2\phi_1$  measurement, although they might result in the degradation in the effective efficiency. Nevertheless, we have made detailed comparisons between data and MC, which are described elsewhere [32], and obtain consistent distributions.

#### 2. Event-level flavor tagging

The event-level flavor tagging combines the results from each of the track categories to determine an overall  $q$  and  $r$ .

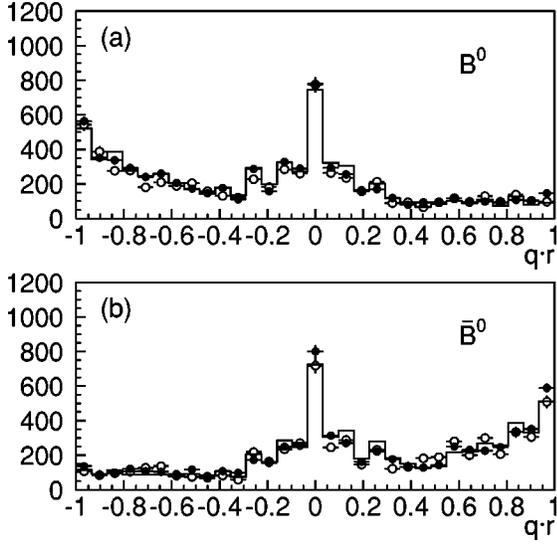


FIG. 10. The  $q \cdot r$  distribution for (a)  $B^0$  and (b)  $\bar{B}^0$  candidates. In each figure the closed (open) points with errors show the  $B^0 \rightarrow D^{*+} l^+ \nu$  (hadronic  $B^0$  decays) data with the background subtracted, while the histogram is the MC prediction.

For the lepton and slow-pion track categories, we take the  $b$ -flavor assignment from the track with the highest  $r$  value in each category. For the kaon and  $\Lambda$  categories, a combined  $b$ -flavor output is calculated as the product of likelihood values for all tracks:

$$(q \cdot r)_{K/\Lambda} = \frac{\prod_i [1 + (q \cdot r)_i] - \prod_i [1 - (q \cdot r)_i]}{\prod_i [1 + (q \cdot r)_i] + \prod_i [1 - (q \cdot r)_i]}, \quad (9)$$

where the subscript  $i$  runs over all tracks in the kaon and  $\Lambda$  categories. The product likelihood is designed to use the information from the sum of the strangeness, which provides better flavor-tagging performance than simply choosing the best candidate.

Using the three aforementioned track-level  $q \cdot r$  values, the event-level  $q$  and  $r$  values are obtained from a look-up table that is prepared with a MC sample that is independent of the sample used to obtain  $q \cdot r$  values in the track categories. The probability that we can assign a non-zero value for  $r$  is 99.6% in MC; i.e. almost all the reconstructed candidates can be used to extract  $\sin 2\phi_1$ .

We specify the following six regions:  $0 < r \leq 0.25$ ,  $0.25 < r \leq 0.5$ ,  $0.5 < r \leq 0.625$ ,  $0.625 < r \leq 0.75$ ,  $0.75 < r \leq 0.875$  and  $0.875 < r \leq 1$ . For each region we obtain the wrong tag fraction,  $w_l$ , where  $l$  is the region ID ( $l=1, \dots, 6$ ), using hadronic and semileptonic control samples, which is described in the next section. In this way, the analysis is insulated from systematic differences between the MC simulation and the data due to imperfections in the modeling of the detector response, decay branching fractions, and fragmentation in our MC simulation.

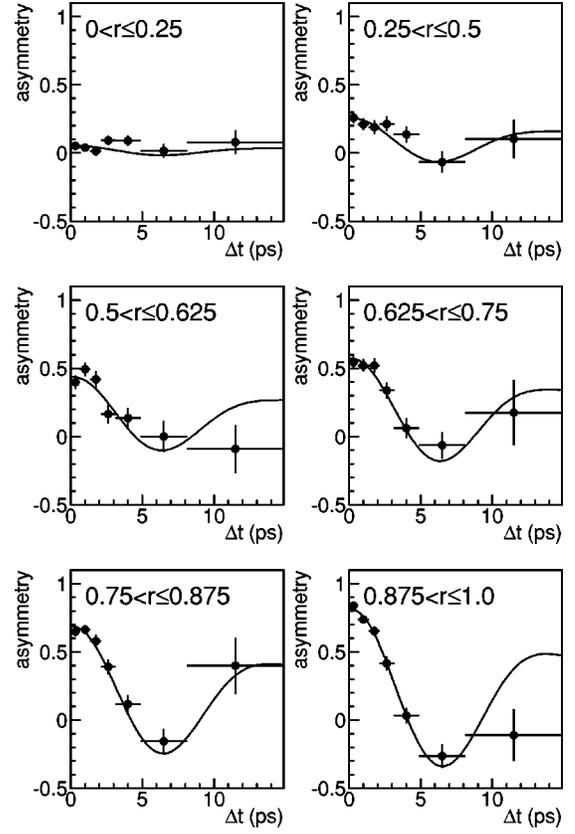


FIG. 11. Measured asymmetries between the OF events and the SF events (OF-SF asymmetries) for six regions of  $r$  obtained for the control sample  $B^0 \rightarrow D^{*+} l \nu$ . The definition of the OF and SF events is given in the text. The background is not subtracted in the asymmetry plots. Fit curves are also shown.

As a validation, we compare the distribution of  $q \cdot r$  in the  $D^{*+} l \nu$  control sample with the MC expectation. As shown in Fig. 10, the data and MC are in good agreement.

## B. Flavor tagging performance

The flavor tagging performance is evaluated by replacing the  $CP$ -eigenstate side of the event with a flavor-specific decay and tagging the  $b$  flavor for the other side using the method described above. We use the semileptonic decay  $B \rightarrow D^{*+} l \nu$  and hadronic modes  $B^0 \rightarrow D^{(*)-} \pi^+$ , and  $D^{*+} \rho^+$  for this purpose. The overall efficiency of our flavor tagging is 99.7% which is consistent with the MC expectation.

Since we know the flavors of both  $B$  mesons in this case, we can observe the time evolution of neutral  $B$ -meson pairs with opposite flavor (OF) or same flavor (SF), which is given by

$$\mathcal{P}_{\text{OF}}(\Delta t) \propto 1 + (1 - 2w) \cos(\Delta m_d \Delta t),$$

$$\mathcal{P}_{\text{SF}}(\Delta t) \propto 1 - (1 - 2w) \cos(\Delta m_d \Delta t),$$

and the OF-SF asymmetry,

$$A_{\text{mix}} \equiv \frac{\mathcal{P}_{\text{OF}} - \mathcal{P}_{\text{SF}}}{\mathcal{P}_{\text{OF}} + \mathcal{P}_{\text{SF}}} = (1 - 2w) \cos(\Delta m_d \Delta t),$$

TABLE IV. The event fractions ( $\epsilon_l$ ) and wrong tag fractions ( $w_l$ ) for each  $r$  interval. The errors include both statistical and systematic uncertainties. The event fractions are obtained from the  $J/\psi K_S^0$  simulation.

$l$	$r$	$\epsilon_l$	$w_l(D^{*}l\nu)$	$w_l(\text{hadronic})$	$w_l(\text{combined})$
1	0.000–0.250	0.405	$0.463^{+0.011}_{-0.011}$	$0.469^{+0.015}_{-0.016}$	$0.465^{+0.010}_{-0.009}$
2	0.250–0.500	0.149	$0.351^{+0.019}_{-0.017}$	$0.352^{+0.026}_{-0.026}$	$0.352^{+0.015}_{-0.014}$
3	0.500–0.625	0.081	$0.254^{+0.021}_{-0.020}$	$0.219^{+0.031}_{-0.030}$	$0.243^{+0.021}_{-0.030}$
4	0.625–0.750	0.099	$0.169^{+0.019}_{-0.018}$	$0.192^{+0.028}_{-0.027}$	$0.176^{+0.022}_{-0.017}$
5	0.750–0.875	0.123	$0.107^{+0.015}_{-0.015}$	$0.127^{+0.032}_{-0.031}$	$0.110^{+0.022}_{-0.014}$
6	0.875–1.000	0.140	$0.041^{+0.012}_{-0.011}$	$0.041^{+0.024}_{-0.023}$	$0.041^{+0.011}_{-0.010}$

where  $w$  is the wrong tag fraction. We thus obtain the value of  $w$  directly from the data by measuring the amplitude of the OF-SF asymmetry.

We obtain the wrong tag fraction by fitting the  $\Delta t$  distribution of the SF and OF events, with  $\Delta m_d$  fixed at the world average value of  $0.472 \text{ ps}^{-1}$  [6]. The procedure to form the probability density function (PDF) for the fit is quite similar to that adopted for the maximum likelihood analysis of  $CP$  eigenstates, which is described in the next section.

The resolution function for signal events, which models how the true distribution is smeared by the finite vertex resolution, is constructed by fitting the proper-time distributions without discriminating between the OF and SF events and with the lifetime fixed to the world average value. In the fit we use the background fraction estimated for each region of  $r$ , and the proper-time distribution for background obtained using events outside the signal region. For hadronic modes the sideband regions in  $M_{bc}$  and  $\Delta E$  are used. For semileptonic decays the upper sideband in  $M_{\text{diff}}$  is used for the fake  $D^*$  backgrounds. Uncorrelated backgrounds are modeled with the events that are found in the signal region after inverting the momentum of the lepton. Semileptonic decays  $D^*Xl\nu$  are treated as signal events since they approximately obey the same OF-SF asymmetry.

Figure 11 shows the measured OF-SF asymmetries as a function of  $\Delta t$  for tagged  $D^{*\mp}l^\pm\nu$  events for the six regions of  $r$ . The curves in the figure are obtained by the fit. The background is not subtracted in the plots. For hadronic modes the fits to OF and SF events are similar to those in the semileptonic case.

We also fit signal MC samples to examine the difference between the generated and reconstructed values. We apply small corrections to  $w_l$  that correspond to the difference. For hadronic modes the corrections range from 0.003 to 0.03, depending on the  $r$  region. For semileptonic decays the difference is consistent with zero within statistical errors, and we apply no correction.

To combine the results from semileptonic and hadronic decays, we calculate the weighted average and its error. We conservatively treat the difference between the weighted average and each measurement as an additional systematic error, and add this difference in quadrature with the error. The systematic errors for the semileptonic mode are dominated by the uncertainties on the background fractions and are comparable to the statistical errors. As explained in Sec. III D, the background estimation relies little on MC information since we use control samples whenever possible. One

important exception is the  $C$  distribution of the  $D^{*}l\nu$  background; we use several  $D^{*}$  components and add them with fixed fractions using MC. Since these fractions are poorly known experimentally, we conservatively assume that each component dominates the  $D^{*}l\nu$  distribution and repeat the fit procedure to obtain the systematic error. For the hadronic modes, the main contribution to the systematic error comes from the uncertainty of the fit bias obtained from the MC simulation, but the statistical errors dominate. The event fractions and wrong tag fractions are summarized in Table IV.

The total effective efficiency obtained by summing over the  $r$  regions is calculated to be

$$\begin{aligned} \epsilon_{\text{eff}} &= \sum_l \epsilon_l (1 - 2w_l)^2 \\ &= [27.0 \pm 0.8(\text{stat})_{-0.9}^{+0.6}(\text{syst})] \%, \end{aligned}$$

where  $\epsilon_l$  is the event fraction in each of the six regions.

Our simulation indicates that events with high-momentum leptons dominate the highest  $r$  region and provide the cleanest tagging information. Events with charged kaons have lower  $r$ , but are more numerous, and thus provide the largest contribution to the effective tagging efficiency. The effective efficiency using each category alone is examined with MC. We obtain 13% for the lepton category, 19% for the kaon and lambda categories combined, and 4% for the slow pion category. Note that the sum of these values exceeds  $\epsilon_{\text{eff}}$  (29.6% in MC) since an event contains tracks in different categories.

We check for possible biases in the flavor tagging by measuring the effective tagging efficiency for the  $B^0$  and  $\bar{B}^0$  control samples separately, and for the  $q = +1$  and  $-1$  samples separately. We find no statistically significant difference.

## V. MAXIMUM LIKELIHOOD FIT

We determine  $\sin 2\phi_1$  by performing an unbinned maximum-likelihood fit of a  $CP$  violating probability density function (PDF) to the observed  $\Delta t$  distributions. These PDF's come from the theoretical distributions diluted and smeared by the detector response. For modes other than  $J/\psi K^{*0}$  the PDF expected for the signal is

$$\mathcal{P}_{\text{sig}}(\Delta t, q, w_l, \xi_f) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}} \mathcal{P}_{CP}(\Delta t, q, w_l, \xi_f), \quad (10)$$

where

$$\mathcal{P}_{CP}(\Delta t, q, w_l, \xi_f) = 1 - \xi_f q (1 - 2w_l) \sin 2\phi_1 \sin(\Delta m_d \Delta t). \quad (11)$$

In order to take into account the effect of finite vertex resolution on the  $\Delta t$  distribution, this PDF is convolved with a resolution function,  $R_{\text{sig}}(\Delta t)$ . Our vertex reconstruction method is explained in Sec. V A. The parametrization and extraction of  $R_{\text{sig}}(\Delta t)$  are described in Sec. V B. We also incorporate the effect of background that dilutes the significance of  $CP$  violation in the time distribution of Eq. (10). The  $\Delta t$  distribution for background events,  $P_{\text{bkg}}(\Delta t)$ , is constructed in a similar way to the signal distribution and is described in detail in Sec. V C.

As a result, we adopt the following  $\Delta t$  distribution function for each event:

$$\begin{aligned} P(\Delta t_i; \sin 2\phi_1) &= f_{\text{sig}} \int \mathcal{P}_{\text{sig}}(\Delta t', q, w_l, \xi_f) \\ &\quad \times R_{\text{sig}}(\Delta t_i - \Delta t') d\Delta t' + (1 - f_{\text{sig}}) \\ &\quad \times P_{\text{bkg}}(\Delta t_i), \end{aligned} \quad (12)$$

where  $f_{\text{sig}}$  is the probability that the event is signal, being calculated for each candidate from  $p_B^{\text{cms}}$  for  $J/\psi K_L^0$  and a combination of  $\Delta E$  and  $M_{\text{bc}}$  for other modes. The only free parameter is  $\sin 2\phi_1$ , which is determined by maximizing the likelihood function

$$L = \prod_i P(\Delta t_i; \sin 2\phi_1), \quad (13)$$

where the product is over all candidates. We perform a blind analysis: The fitting algorithms were developed and finalized without using the flavor information  $q$ .

In the following we explain the details of  $R_{\text{sig}}(\Delta t)$ ,  $P_{\text{bkg}}(\Delta t)$  and  $f_{\text{sig}}$  in turn. The likelihood for  $J/\psi K^{*0}(\rightarrow K_S^0 \pi^0)$  candidates is described separately in Sec. V E.

### A. Vertex reconstruction

The decay vertices for the  $CP$  side that include a  $J/\psi$  candidate are reconstructed using leptons from the  $J/\psi$  and a constraint on the  $B$  decay point. The  $B$  decay point is constrained by the measured profile of the interaction point (IP profile) convolved with the finite  $B$  flight length in the plane perpendicular to the  $z$  axis (the  $r$ - $\phi$  plane). The IP profile is represented by a three-dimensional Gaussian distribution. The standard deviation of each Gaussian is determined using pre-selected  $B\bar{B}$  candidates on a run-by-run basis, while the mean is evaluated in finer subdivisions. The typical size of the IP profile is 100  $\mu\text{m}$  in  $x$ , 5  $\mu\text{m}$  in  $y$  and 3 mm in  $z$ . Since the size in the  $y$  direction is too small to be measured from the vertex distribution, it is taken from special measure-

ments by the KEKB accelerator group. For leptons, we require that there are sufficient SVD hits associated with a CDC track by a Kalman filter technique, i.e. both  $z$  and  $r$ - $\phi$  hits in at least one layer and at least one additional layer with a  $z$  hit. In order to remove events with misreconstructed tracks, we require that the reduced  $\chi^2$  ( $\chi^2/n$ ,  $n$  = number of degrees of freedom) of the vertex be less than 20. The vertex reconstruction efficiency is measured to be 95% with  $B^\pm \rightarrow J/\psi K^\pm$  and  $B^0 \rightarrow J/\psi K^{*0}(\rightarrow K^\pm \pi^\mp)$  events. This is consistent with the expectation from the SVD acceptance and cluster matching efficiency. The resolution estimated by MC is typically 75  $\mu\text{m}$  (rms).

For  $B \rightarrow \eta_c K_S^0$  candidates, the method is basically the same as for  $J/\psi$ , replacing  $l^+ l^-$  with  $K^+ K^-$  in the case of  $\eta_c \rightarrow K^+ K^- \pi^0$  and with  $K^\pm \pi^\mp$  for  $\eta_c \rightarrow K_S^0 K^\pm \pi^\mp$ . Although the resolution in these cases is worse than for candidates with a  $J/\psi$  vertex, it is still better than the tag-side vertex resolution.

The algorithm for tag-side vertex reconstruction is chosen to minimize the effect of long-lived particles, secondary vertices from charmed hadrons and a small fraction of poorly reconstructed tracks. From all the charged tracks except those used to reconstruct the  $CP$  side, we select tracks that have associated SVD hits in the same way as for the  $CP$  side. We also require that the impact parameter with respect to the  $CP$ -side vertex be less than 0.5 mm in the  $r$ - $\phi$  plane, less than 1.8 mm in  $z$ , and the vertex error in  $z$  be less than 0.5 mm. Tracks are removed if they form a  $K_S^0$  candidate satisfying  $|M_{K_S^0} - M_{\pi^+ \pi^-}| < 15 \text{ MeV}/c^2$ . Tracks satisfying these criteria are used to reconstruct the tag-side vertex where the IP constraint is also applied. If the reduced  $\chi^2$  of the vertex is good, we accept this vertex. Otherwise we remove the track that gives the largest contribution to the  $\chi^2$  and repeat the vertex reconstruction. If the track to be removed is a lepton with  $p_l^{\text{cms}} > 1.1 \text{ GeV}/c$ , however, we keep the lepton and remove the track with the second worst  $\chi^2$ . This trimming procedure is repeated until we obtain a good reduced  $\chi^2$ . The reconstruction efficiency was measured to be 93% for  $B^\pm \rightarrow J/\psi K^\pm$  and  $B^0 \rightarrow J/\psi K^{*0}(\rightarrow K^\pm \pi^\mp)$  candidates, consistent with the MC expectation (91%). The resolution estimated from the simulation is typically 140  $\mu\text{m}$  (rms).

### B. Signal resolution function

The resolution function  $R_{\text{sig}}(\Delta t - \Delta t')$  is parametrized by the sum of two Gaussians:

$$\begin{aligned} R_{\text{sig}}(\Delta t - \Delta t') &= (1 - f_{\text{tail}}) G(\Delta t - \Delta t'; \mu_{\Delta t}, \sigma_{\Delta t}) \\ &\quad + f_{\text{tail}} G(\Delta t - \Delta t'; \mu_{\Delta t}^{\text{tail}}, \sigma_{\Delta t}^{\text{tail}}), \end{aligned} \quad (14)$$

where  $G(x; \mu, \sigma)$  is a Gaussian distribution in  $x$  with mean  $\mu$  and rms  $\sigma$ . The parameter  $f_{\text{tail}}$  describes the fraction of the tail of the resolution function, and  $\sigma_{\Delta t}$ ,  $\sigma_{\Delta t}^{\text{tail}}$ ,  $\mu_{\Delta t}$  and  $\mu_{\Delta t}^{\text{tail}}$  are the proper-time difference resolutions and the mean value shifts of the proper-time difference for the main part and the tail of the resolution function, respectively. The value of  $f_{\text{tail}}$

is determined to be  $0.03 \pm 0.02$  from the lifetime analysis of hadronic samples using the same resolution function.

The proper-time difference resolutions  $\sigma_{\Delta t}$  and  $\sigma_{\Delta t}^{\text{tail}}$  are calculated on an event-by-event basis taking into account the error  $\sigma_K$  in the kinematic approximation  $\Delta t \approx \Delta z / c\beta\gamma$ ,  $\beta\gamma = p_z(Y)/m(Y)$ :

$$\sigma_{\Delta t} = \sqrt{\left(\frac{\sigma_{\Delta z}}{c\beta\gamma}\right)^2 + \sigma_K^2},$$

$$\sigma_{\Delta t}^{\text{tail}} = \sqrt{\left(\frac{\sigma_{\Delta z}^{\text{tail}}}{c\beta\gamma}\right)^2 + (\sigma_K^{\text{tail}})^2}.$$

We measure  $\sigma_K = 0.287 \pm 0.004$  ps and  $\sigma_K^{\text{tail}} = 0.32 \pm 0.19$  ps using the MC simulation. These parameters are independent of the detector performance.

The parameters  $\sigma_{\Delta z}$  and  $\sigma_{\Delta z}^{\text{tail}}$  are calculated from the event-by-event vertex errors of the two  $B$  mesons,  $\sigma_z^{CP}$  and  $\sigma_z^{\text{tag}}$ , which are computed from the track helix errors in the vertex fit. We use

$$\sigma_{\Delta z}^2 = S_{\text{det}}^2 (\sigma_z^{CP})^2 + (S_{\text{det}}^2 + S_{\text{charm}}^2) (\sigma_z^{\text{tag}})^2,$$

$$(\sigma_{\Delta z}^{\text{tail}})^2 = (S_{\text{det}}^{\text{tail}})^2 (\sigma_z^{CP})^2 + [(S_{\text{det}}^{\text{tail}})^2 + (S_{\text{charm}}^{\text{tail}})^2] (\sigma_z^{\text{tag}})^2,$$

where  $S_{\text{charm}}$  and  $S_{\text{charm}}^{\text{tail}}$  are scaling factors to account for the degradation of the vertex resolution on the tag-side due to contamination from charm daughters, and  $S_{\text{det}}$  and  $S_{\text{det}}^{\text{tail}}$  are global scaling factors that account for systematic uncertainties in the vertex errors  $\sigma_z^{CP}$  and  $\sigma_z^{\text{tag}}$ . We determine  $S_{\text{charm}} = 0.58 \pm 0.01$  and  $S_{\text{charm}}^{\text{tail}} = 2.16 \pm 0.10$  using the MC simulation. The values of  $S_{\text{det}}$  and  $S_{\text{det}}^{\text{tail}}$  are measured from the data as they depend on the detector performance. We determine  $S_{\text{det}}$  using a  $D^0 \rightarrow K^+ \pi^-$  control sample. The production point of the  $D^0$  is obtained from the primary tracks in the same hemisphere as the  $D^0$  candidate using the IP constraint. The distance between the  $D^0$  decay vertex and the production vertex in the  $z$  direction is fit with the same resolution function and the known  $D^0$  lifetime to obtain  $S_{\text{det}}$ . We measure  $S_{\text{det}} = 0.88 \pm 0.01$  from the data and  $S_{\text{det}} = 1.05 \pm 0.01$  from a MC simulation of the  $D^0$  sample. Finding  $S_{\text{det}} = 1.035 \pm 0.003$  for a  $B \rightarrow J/\psi K$  MC sample, we use  $S_{\text{det}} = (0.88 \pm 0.01) \times (1.035 \pm 0.003) / (1.05 \pm 0.01) = 0.86 \pm 0.01$  for the data. We determine  $S_{\text{det}}^{\text{tail}}$  to be  $3.51 \pm 0.88$  from the lifetime analysis of the flavor-specific hadronic samples.

For  $B \rightarrow \eta_c K_S^0$  decays, we introduce an additional scale factor to account for the difference between the  $\eta_c$  and  $J/\psi$  decays. Using the MC sample, we determine the additional scale factor for  $\eta_c$  to be  $1.07 \pm 0.02$  for both  $\eta_c \rightarrow K^+ K^- \pi^0$  and  $\eta_c \rightarrow K_S^0 K^\pm \pi^\mp$ .

A small fraction of events have a large reduced  $\chi^2$ . We have found that the vertex error computed from track helix errors in the vertex fit underestimates the vertex resolution and the vertex with larger  $\chi^2$  has worse resolution. In order to take into account this effect, we introduce effective vertex resolutions  $\tilde{\sigma}_z^{CP}$  and  $\tilde{\sigma}_z^{\text{tag}}$  when  $\chi^2/n$  is greater than 3:

$$(\tilde{\sigma}_z^{CP})^2 = \{1 + \alpha_{CP}[(\chi^2/n)_{CP} - 3]\} (\sigma_z^{CP})^2,$$

$$(\tilde{\sigma}_z^{\text{tag}})^2 = \{1 + \alpha_{\text{tag}}[(\chi^2/n)_{\text{tag}} - 3]\} (\sigma_z^{\text{tag}})^2,$$

where  $(\chi^2/n)_{CP}$  and  $(\chi^2/n)_{\text{tag}}$  are the reduced  $\chi^2$  of the vertex fits for the  $CP$  and tagging  $B$  decay vertices, respectively. The coefficients  $\alpha_{CP} = 1.02 \pm 0.03$  and  $\alpha_{\text{tag}} = 1.64 \pm 0.05$  are determined by a MC study of  $B^0 \rightarrow J/\psi K_S^0$ .

As mentioned above, the offsets  $\mu_{\Delta t}$  and  $\mu_{\Delta t}^{\text{tail}}$  originate from the mean shifts of the  $\Delta z$  measurements  $\mu_{\Delta z}$  and  $\mu_{\Delta z}^{\text{tail}}$ , respectively,

$$\mu_{\Delta t} = \frac{\mu_{\Delta z}}{c\beta\gamma}, \quad \mu_{\Delta t}^{\text{tail}} = \frac{\mu_{\Delta z}^{\text{tail}}}{c\beta\gamma}.$$

The mean value shifts,  $\mu_{\Delta z}$  and  $\mu_{\Delta z}^{\text{tail}}$ , are caused by contamination from charm daughters in the vertex reconstruction on the tag side and are correlated with  $\sigma_z^{\text{tag}}$ :

$$\mu_{\Delta z}(\sigma_z^{\text{tag}}) = \mu_0 + \alpha_\mu \sigma_z^{\text{tag}},$$

$$\mu_{\Delta z}^{\text{tail}}(\sigma_z^{\text{tag}}) = \mu_0^{\text{tail}} + \alpha_\mu^{\text{tail}} \sigma_z^{\text{tag}}.$$

The values for  $\mu_0$  and  $\mu_0^{\text{tail}}$  are determined from hadronic samples to be  $\mu_0 = (-21.4 \pm 3.7) \mu\text{m}$  and  $\mu_0^{\text{tail}} = (151 \pm 128) \mu\text{m}$ , while  $\alpha_\mu$  and  $\alpha_\mu^{\text{tail}}$  are derived from MC simulation where we obtain  $\alpha_\mu = -0.10 \pm 0.01$  and  $\alpha_\mu^{\text{tail}} = -1.42 \pm 0.17$ .

Figure 12(a) shows the  $\Delta t_{\text{rec}} - \Delta t_{\text{gen}}$  distribution for the MC  $J/\psi K_S^0(\pi^+ \pi^-)$  candidates along with the resolution function, where  $\Delta t_{\text{rec}}$  and  $\Delta t_{\text{gen}}$  are the reconstructed and true proper-time differences, respectively. The resolution function is obtained by summing event-by-event resolution functions. The distribution is well represented by the resolution function. The average resolution function obtained from the  $J/\psi K_S^0(\pi^+ \pi^-)$  data is shown in Fig. 12(b), which is represented by the sum of two Gaussian distributions with the following parameters:  $\mu_{\text{main}} = -0.24$  ps,  $\mu_{\text{tail}} = 0.18$  ps,  $\sigma_{\text{main}} = 1.49$  ps,  $\sigma_{\text{tail}} = 3.85$  ps.

### C. Background shape

The background likelihood function is defined in a similar way to the signal function,

$$P_{\text{bkg}}(\Delta t) = \int_{-\infty}^{+\infty} \mathcal{P}_{\text{bkg}}(\Delta t') \cdot R_{\text{bkg}}(\Delta t - \Delta t') d\Delta t'. \quad (16)$$

Although  $R_{\text{bkg}}(\Delta t - \Delta t')$  is treated as a resolution function for the background, it does not need to be an exact model of the vertex resolution. It is more important that  $P_{\text{bkg}}$  represents the proper-time distribution of the whole background sample with sufficient precision.

The PDF for background events is expressed as

$$P_{\text{bkg}}(\Delta t) = f_\tau \frac{e^{-|\Delta t|/\tau_{\text{bkg}}}}{2\tau_{\text{bkg}}} + (1 - f_\tau) \delta(\Delta t), \quad (17)$$

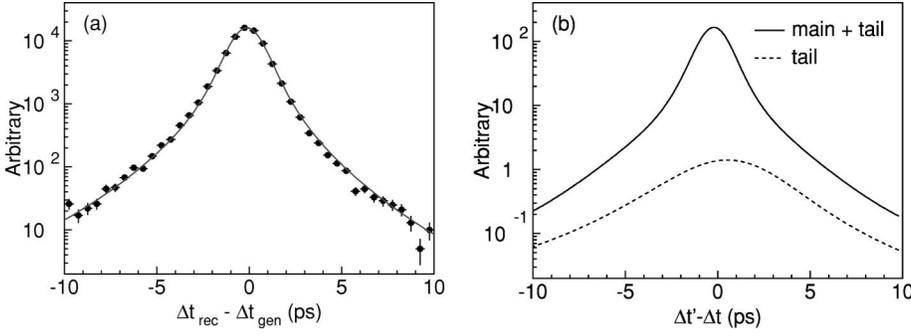


FIG. 12. (a) The  $\Delta t_{\text{rec}} - \Delta t_{\text{gen}}$  distributions for the MC  $J/\psi K_S^0(\pi^+\pi^-)$  candidates with the resolution function and (b) the average resolution function from  $J/\psi K_S^0(\pi^+\pi^-)$  data. The vertical scales are arbitrary.

where  $f_\tau$  is the fraction of the background component with an effective lifetime of  $\tau_{\text{bkg}}$ , and  $\delta$  is the Dirac delta function. We assume no asymmetry in the background  $\Delta t$  distribution. We find  $f_\tau$  to be small using background-dominated regions in the  $\Delta E$  versus  $M_{\text{bc}}$  plane of  $J/\psi K_S^0$  and  $J/\psi K^+$  candidates. We thus use  $\mathcal{P}_{\text{bkg}}(\Delta t) = \delta(\Delta t)$  for all  $f_{CP}$  modes except for  $J/\psi K_L^0$ .

The background in the  $J/\psi K_L^0$  mode is dominated by  $B \rightarrow J/\psi X$  decays, including  $CP$  eigenstates that have to be treated differently from non- $CP$  states. The  $\mathcal{P}_{\text{bkg}}$  for the  $J/\psi K_L^0$  mode is determined by a MC simulation study separately for each background component:  $J/\psi K^{*0}(K_L^0\pi^0)$ ,  $\xi_f = -1$   $CP$  modes ( $J/\psi K_S^0$ ),  $\xi_f = +1$   $CP$  modes ( $\psi(2S)K_L^0$ ,  $\chi_{c1}K_L^0$  and  $J/\psi\pi^0$ ), the other  $B^0$ ,  $B^\pm$  decays and combinatorial background. For the  $CP$ -mode backgrounds, we use the signal PDF given in Eq. (11) with the appropriate  $\xi_f$  values. For the  $J/\psi K^{*0}(K_L^0\pi^0)$  mode, which is a mixture of  $\xi_f = -1$  (about 81%) and  $\xi_f = +1$  (about 19%) states [33], we use a net  $CP$  eigenvalue of  $\xi_{\psi K^{*0}} = -0.62 \pm 0.11$ .

Accordingly, we obtain the background PDF for  $J/\psi K_L^0$ :

$$\begin{aligned} \mathcal{P}_{\text{bkg}}(\Delta t) \equiv & \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}} \\ & \times [f_{\text{bg}B^0} + f_{\text{bg}\psi K^{*0}} \mathcal{P}_{CP}(\Delta t, q, w_l, \xi_{\psi K^{*0}}) \\ & + f_{\text{bg}CP_{\text{odd}}} \mathcal{P}_{CP}(\Delta t, q, w_l, -1) \\ & + f_{\text{bg}CP_{\text{even}}} \mathcal{P}_{CP}(\Delta t, q, w_l, +1)] \\ & + f_{\text{bg}B^\pm} \frac{e^{-|\Delta t|/\tau_{\text{bg}B^\pm}}}{2\tau_{\text{bg}B^\pm}} + f_{\text{bg}Cmb} \left[ f_{\tau_{\text{Cmb}}} \frac{e^{-|\Delta t|/\tau_{\text{bg}Cmb}}}{2\tau_{\text{bg}Cmb}} \right. \\ & \left. + (1 - f_{\tau_{\text{Cmb}}}) \delta(\Delta t) \right], \end{aligned} \quad (18)$$

where  $f_{\text{bg}\psi K^{*0}}$ ,  $f_{\text{bg}CP_{\text{odd}}}$ ,  $f_{\text{bg}CP_{\text{even}}}$ ,  $f_{\text{bg}B^0}$ ,  $f_{\text{bg}B^\pm}$ , and  $f_{\text{bg}Cmb}$  are the fractions of background components from  $J/\psi K^{*0}(K_L^0\pi^0)$ ,  $\xi_f = -1$   $CP$  modes,  $\xi_f = +1$   $CP$  modes, the remaining  $B^0$ ,  $B^\pm$ , and combinatorial, respectively ( $f_{\text{bg}\psi K^{*0}} + f_{\text{bg}CP_{\text{odd}}} + f_{\text{bg}CP_{\text{even}}} + f_{\text{bg}B^0} + f_{\text{bg}B^\pm} + f_{\text{bg}Cmb} = 1$ ). The fraction of each background component is a function of  $p_B^{\text{cms}}$ . The fraction,  $f_{\text{bg}Cmb}$ , is calculated as described in Sec. III C. The combinatorial background includes a prompt component with the fraction of  $(1 - f_{\tau_{\text{Cmb}}})$  where  $f_{\tau_{\text{Cmb}}} = 0.26$

$\pm 0.08$ . The lifetime distribution of the combinatorial background is obtained from  $e-\mu$  combinations with invariant masses in the  $J/\psi$  region that satisfy our selection criteria. The effective lifetime is determined to be  $\tau_{\text{bg}Cmb} = 1.03 \pm 0.12$  ps, also from the  $e-\mu$  control sample. An MC study shows that the effective lifetime for background from  $B^\pm$ ,  $\tau_{\text{bg}B^\pm}$ , is shorter than the  $B^\pm$  lifetime due to the contamination of charged tracks from  $f_{CP}$  [mostly  $\pi^\pm$  from  $J/\psi K^{*\pm}(K_L^0\pi^\pm)$ ] into the tag-side vertex. The value of  $\tau_{\text{bg}B^\pm}$  is determined from the MC simulation to be  $(1.49 \pm 0.04)$  ps. The same MC study shows that the effective lifetime for  $B^0$  backgrounds is consistent with the nominal  $B^0$  lifetime. Thus we use the nominal  $B^0$  lifetime in our fit.

For the  $J/\psi X$  background in the  $J/\psi K_L^0$  mode, we use the signal resolution function  $R_{\text{sig}}$  to model the background since both the  $CP$ - and tag-side vertices are reconstructed with similar combinations of tracks for these backgrounds.

For the combinatorial background, we use

$$\begin{aligned} R_{\text{bkg}}(\Delta t' - \Delta t) = & (1 - f_{\text{tail,bkg}}) G(\Delta t' - \Delta t; \mu_{\Delta t}^{\text{bkg}}, \sigma_{\Delta t}^{\text{bkg}}) \\ & + f_{\text{tail,bkg}} G(\Delta t' - \Delta t; \mu_{\Delta t}^{\text{tail,bkg}}, \sigma_{\Delta t}^{\text{tail,bkg}}), \end{aligned} \quad (19)$$

where  $f_{\text{tail,bkg}}$ ,  $\mu_{\Delta t}^{\text{bkg}}$ , and  $\mu_{\Delta t}^{\text{tail,bkg}}$  are constants determined from data. The resolutions  $\sigma_{\Delta t}^{\text{bkg}}$  and  $\sigma_{\Delta t}^{\text{tail,bkg}}$  are calculated on an event-by-event basis as

$$\begin{aligned} \sigma_{\Delta t}^{\text{bkg}} &= \frac{\sigma_{\Delta z}^{\text{bkg}}}{c\beta\gamma}, \\ \sigma_{\Delta t}^{\text{tail,bkg}} &= \frac{\sigma_{\Delta z}^{\text{tail,bkg}}}{c\beta\gamma}, \end{aligned}$$

$$(\sigma_{\Delta z}^{\text{bkg}})^2 = (S_{\text{det}}^{\text{bkg}})^2 [(\tilde{\sigma}_z^{CP})^2 + (\tilde{\sigma}_z^{\text{tag}})^2],$$

$$(\sigma_{\Delta z}^{\text{tail,bkg}})^2 = (S_{\text{bkg}}^{\text{tail}})^2 [(\tilde{\sigma}_z^{CP})^2 + (\tilde{\sigma}_z^{\text{tag}})^2],$$

where  $\tilde{\sigma}_z^{CP}$  and  $\tilde{\sigma}_z^{\text{tag}}$  are calculated as shown in Eq. (15).

We use different values of  $\mu_{\Delta t}^{\text{bkg}}$  and  $\mu_{\Delta t}^{\text{tail,bkg}}$  for the finite lifetime component and the zero-lifetime component, since they come from different types of events. The background shape parameters for all modes except  $J/\psi K_L^0$  are obtained from events in the background-dominated regions of  $\Delta E$  versus  $M_{\text{bc}}$ . For  $J/\psi K_L^0$ , we use events with  $e-\mu$  pairs to deter-

mine the properties of fake  $J/\psi$  candidates, as discussed in Sec. III C. The parameters used in the fit are summarized in Table V.

#### D. Signal probability

The signal probability,  $f_{\text{sig}}$ , is calculated as a function of  $\Delta E$  and  $M_{\text{bc}}$  for each event. It is given by

$$f_{\text{sig}}(\Delta E, M_{\text{bc}}) = \frac{F_{\text{SIG}}(\Delta E, M_{\text{bc}})}{F_{\text{BG}}(\Delta E, M_{\text{bc}}) + F_{\text{SIG}}(\Delta E, M_{\text{bc}})}, \quad (20)$$

where  $F_{\text{SIG}}(\Delta E, M_{\text{bc}})$  is the signal function and  $F_{\text{BG}}(\Delta E, M_{\text{bc}})$  is the background function.

In the case of  $B^0 \rightarrow J/\psi K_S^0$  each distribution of  $\Delta E$  and  $M_{\text{bc}}$  is well modeled by a Gaussian function. For the background, we use a linear function for  $\Delta E$  and the ARGUS parametrization [28] for  $M_{\text{bc}}$ :

$$F_{\text{SIG}}(\Delta E, M_{\text{bc}}) = a \cdot G(\Delta E; \mu_{\Delta E}, \sigma_{\Delta E}) \cdot G(M_{\text{bc}}; \mu_{M_{\text{bc}}}, \sigma_{M_{\text{bc}}}),$$

$$F_{\text{BG}}(\Delta E, M_{\text{bc}}) = b \cdot (1 + c \cdot \Delta E) \cdot M_{\text{bc}} \sqrt{1 - (M_{\text{bc}}/E_{\text{beam}})^2} \times \exp\{n \cdot [1 - (M_{\text{bc}}/E_{\text{beam}})^2]\} \quad (21)$$

where  $a$  and  $b$  are normalization factors consistent with the overall signal-to-background ratios obtained from the fit to the  $M_{\text{bc}}$  distribution in the  $\Delta E$  signal region. The values  $\sigma_{M_{\text{bc}}}$ ,  $\mu_{M_{\Delta E}}$ ,  $\sigma_{M_{\Delta E}}$ ,  $\mu_{M_{\text{bc}}}$ ,  $c$  and  $n$  are determined from a fit to the data.

The  $\Delta E$  and  $M_{\text{bc}}$  distributions for  $B^0 \rightarrow \psi(2S)K_S(\psi(2S) \rightarrow l^+l^-)$ ,  $B^0 \rightarrow \psi(2S)K_S(\psi(2S) \rightarrow J/\psi \pi^+ \pi^-)$ , and  $B^0 \rightarrow \eta_c K_S^0$  ( $\eta_c \rightarrow K_S^0 K^\pm \pi^\mp$ ), are determined using the same procedure as that for  $B^0 \rightarrow J/\psi K_S$ . For  $B^0 \rightarrow \chi_{c1} K_S$ , the  $\Delta E$  and  $M_{\text{bc}}$  distributions are determined from MC simulation because the data sample for this mode is too small to estimate the parameters reliably.

The treatment for modes that include  $\pi^0$  mesons, such as  $B^0 \rightarrow J/\psi K_S(K_S^0 \rightarrow \pi^0 \pi^0)$  and  $B^0 \rightarrow \eta_c K_S^0$  ( $\eta_c \rightarrow K^+ K^- \pi^0$ ), is different. While the fit function for the  $M_{\text{bc}}$  distribution remains the same, the  $\Delta E$  distributions are better represented by the Crystal Ball function [34]:

$$\frac{1}{A} \exp\left(-\frac{(\Delta E - \mu_{\Delta E})^2}{2\sigma_{\Delta E}^2}\right) \quad \text{for } \Delta E > \mu_{\Delta E} - \alpha \sigma_{\Delta E},$$

$$\frac{1}{A} \frac{\exp(-\alpha^2/2)}{\left[1 - \frac{(\Delta E - \mu_{\Delta E})\alpha}{\sigma_{\Delta E}} - \frac{\alpha^2}{n}\right]^n} \quad \text{for } \Delta E < \mu_{\Delta E} - \alpha \sigma_{\Delta E}.$$

All the parameters for these fits were determined from MC simulation because the number of events for these modes in data is too small. The integrated background fractions in the signal region are listed in Table VI. For the  $B^0 \rightarrow J/\psi K_L^0$  fit, we define the signal probability as a function of  $p_B^{\text{cms}}$ , as described in Sec. III C.

TABLE V. Background shape parameters for the combinatorial background.

Parameters	$c\bar{c}K_S^0(K^{*0})$	$J/\psi K_L^0$
$S_{\text{det}}^{\text{bkg}}$	$1.08 \pm 0.06$	$0.88 \pm 0.07$
$S_{\text{bkg}}^{\text{tail}}$	$3.31 \pm 0.28$	$3.94 \pm 1.14$
$f_{\text{tail,bkg}}$	$0.14 \pm 0.04$	$0.05 \pm 0.02$
lifetime component		
$\mu_{\Delta t}^{\text{bkg}}(\text{ps})$	N/A	$-0.33 \pm 0.10$
$\mu_{\Delta t}^{\text{tail,bkg}}(\text{ps})$	N/A	$1.86 \pm 1.28$
prompt component		
$\mu_{\Delta t}^{\text{bkg}}(\text{ps})$	$-0.05 \pm 0.04$	$-0.22 \pm 0.15$
$\mu_{\Delta t}^{\text{tail,bkg}}(\text{ps})$	$-0.12 \pm 0.26$	$-5.00 \pm 1.10$

#### E. Likelihood for $J/\psi K^{*0} (\rightarrow K_S^0 \pi^0)$

For the  $B^0 \rightarrow J/\psi K^{*0}$  fit, the signal PDF we use is

$$\mathcal{P}_{\text{sig}}(\Delta t, \theta_{\text{tr}}, q, w_l, \xi_f) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}} \left[ (1 - f_{\text{odd}}) \frac{3}{8} (1 + \cos^2 \theta_{\text{tr}}) \mathcal{P}_{CP}(\Delta t, q, w_l, +1) + f_{\text{odd}} \frac{3}{4} (1 - \cos^2 \theta_{\text{tr}}) \mathcal{P}_{CP}(\Delta t, q, w_l, -1) \right], \quad (22)$$

where  $f_{\text{odd}}$  is the fraction of  $\xi_f = -1$  decays in the  $B^0 \rightarrow J/\psi K^{*0}(K^{*0} \rightarrow K_S^0 \pi^0)$  mode determined from a full angular analysis to be  $0.19 \pm 0.04(\text{stat}) \pm 0.04(\text{syst})$  [33]. Here  $\theta_{\text{tr}}$  is defined in the transversity basis [35] as the angle between the positive  $J/\psi$  decay lepton direction and the axis normal to the  $K^{*0}$  decay plane in the  $J/\psi$  rest frame.  $\mathcal{P}_{CP}(\Delta t, q, w_l, \xi_f)$  is defined in Eq. (11). The signal resolution function is identical to that used for the other modes. For the background shape, we also use Eq. (19) for  $R_{\text{bkg}}$  except for the  $J/\psi X$  background where we use  $R_{\text{sig}}$  in the same way as for the  $J/\psi K_L^0$  fit.

We use the following background PDF:

TABLE VI. Summary of the numbers of candidates and background fraction in the signal region for each mode. The values are obtained for events that have successful vertex reconstruction and flavor tagging.

Decay mode	Events	Background fraction
$B^0 \rightarrow J/\psi K_S^0, K_S^0 \rightarrow \pi^+ \pi^-$	387	$0.038 \pm 0.010$
$B^0 \rightarrow J/\psi K_S^0, K_S^0 \rightarrow \pi^0 \pi^0$	57	$0.272 \pm 0.054$
$B^0 \rightarrow \psi(2S)K_S^0, \psi(2S) \rightarrow l^+ l^-$	33	$0.038 \pm 0.028$
$B^0 \rightarrow \psi(2S)K_S^0, \psi(2S) \rightarrow J/\psi \pi^+ \pi^-$	32	$0.078 \pm 0.027$
$B^0 \rightarrow \chi_{c1} K_S^0, \chi_{c1} \rightarrow J/\psi \gamma$	17	$0.144 \pm 0.056$
$B^0 \rightarrow \eta_c K_S^0, \eta_c \rightarrow K_S^0 K^\pm \pi^\mp$	35	$0.242 \pm 0.045$
$B^0 \rightarrow \eta_c K_S^0, \eta_c \rightarrow K^+ K^- \pi^0$	17	$0.560 \pm 0.164$
$B^0 \rightarrow J/\psi K_L^0$	523	$0.379 \pm 0.048$
$B^0 \rightarrow J/\psi K^{*0}, K^{*0} \rightarrow K_S^0 \pi^0$	36	$0.163 \pm 0.054$

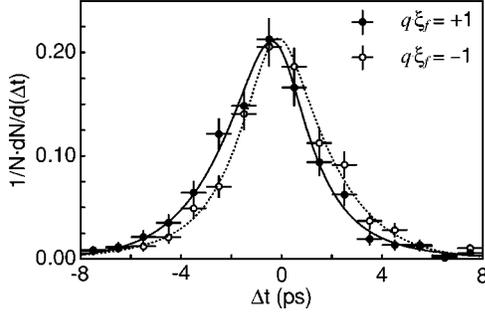


FIG. 13.  $\Delta t$  distributions for the events with  $q\xi_f = +1$  (solid points) and  $q\xi_f = -1$  (open points). The results of the global fit (with  $\sin 2\phi_1 = 0.99$ ) are shown as solid and dashed curves, respectively.

$$\mathcal{P}_{\text{bkg}}(\Delta t) \equiv f_{\text{bgFA}} \frac{e^{-|\Delta t|/\tau_{B_{\text{FA}}}}}{2\tau_{B_{\text{FA}}}} + f_{\text{bgNR}} \frac{e^{-|\Delta t|/\tau_{B_{\text{NR}}}}}{2\tau_{B_{\text{NR}}}} + f_{\text{bgCmb}} \delta(\Delta t), \quad (23)$$

where  $f_{\text{bgFA}}$ ,  $f_{\text{bgNR}}$ , and  $f_{\text{bgCmb}}$  are the fractions of background components from feed-across from other  $J/\psi K^*$  modes, non-resonant  $B^0 \rightarrow J/\psi K_S^0 \pi^0$  decays and combinatorial background. The fractions of feed-across and non-resonant decays are determined from the MC simulation and from  $K^*$  mass sideband events, respectively, and are functions of  $M_{\text{bc}}$ . The fraction of combinatorial background is determined in the same way as for the  $K_S^0$  modes. The effective lifetimes of the feed-across and non-resonant decay backgrounds,  $\tau_{B_{\text{FA}}}$  and  $\tau_{B_{\text{NR}}}$ , are fixed to the  $B^0$  lifetime in the fit.

Finally, the determination of  $f_{\text{sig}}$  follows the method for other modes that include  $\pi^0$  mesons. The  $\Delta E$  distribution is modeled by a Crystal Ball function. We consider contributions from the feed-across from other  $B^0 \rightarrow J/\psi K^* \pi^0$  modes as well as from the non-resonant  $B^0 \rightarrow J/\psi K_S^0 \pi^0$  mode, which make a peak in the signal region. These background fractions are determined from the MC simulation and  $K^* \pi^0$  mass sideband data, respectively.

## VI. FIT RESULTS

The likelihood fit is applied to the 1137 candidates where the vertex reconstruction and flavor tagging have been successful. We obtain

$$\sin 2\phi_1 = 0.99 \pm 0.14(\text{stat}).$$

The observed  $CP$  violation is large. Figure 13 shows the  $\Delta t$  distributions together with the results from the fit. Indeed, the broken  $CP$  symmetry is visually apparent from the difference between the number of events for  $q\xi_f = +1$  and  $q\xi_f = -1$  at each  $\Delta t$  bin, despite the dilution from the vertex resolution, background events and incorrect flavor tagging.

We examined the value of  $\sin 2\phi_1$  in various sub-samples. Applying the likelihood fit to  $(c\bar{c})K_S^0$  ( $\xi_f = -1$ ) and  $J/\psi K_L^0$  ( $\xi_f = +1$ ) separately, we obtain  $\sin 2\phi_1 = 0.84 \pm 0.17(\text{stat})$  and  $1.31 \pm 0.23(\text{stat})$ , respectively. Figure 14 shows the logarithmic-likelihood values as a function of  $\sin 2\phi_1$  for

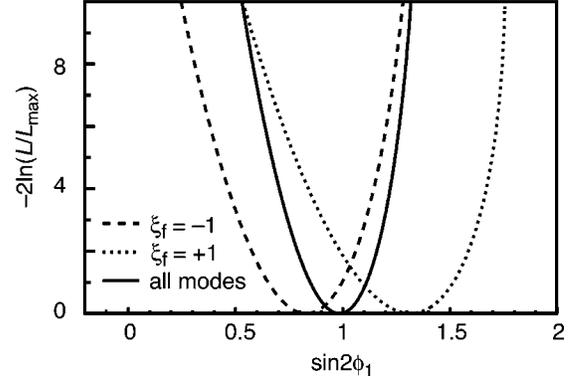


FIG. 14. Values of  $-2 \ln(L/L_{\text{max}})$  vs  $\sin 2\phi_1$  for the  $\xi_f = -1$  (dashed) and  $+1$  (dotted) modes separately and for both modes combined (solid).

$CP$ -odd,  $CP$ -even, and all decay modes. A more detailed breakdown along with separate results for  $q = +1$  and  $-1$  is given in Table VII. We find no systematic trends beyond statistical fluctuations.

Figure 15 shows the asymmetry,  $\sin 2\phi_1 \cdot \sin(\Delta m_d \Delta t)$ , obtained in each  $\Delta t$  bin for (a) all modes, (b)  $CP$ -odd modes and (c)  $CP$ -even modes. The unbinned maximum likelihood fit is performed separately for events in each  $\Delta t$  bin. The value of  $\sin 2\phi_1$  and its error are multiplied by the average value of  $\sin(\Delta m_d \Delta t)$  in each  $\Delta t$  bin of the plot. The points are plotted at the average  $\Delta t$  of each bin.

We also checked the values of  $\sin 2\phi_1$  in the different  $r$  ranges of the flavor tagging. The results are listed in Table VIII. No systematic variation is seen. Finally, we subdivided the  $CP$  sample into three data taking periods: in 2000, from January to April 2001 and the rest. The  $\sin 2\phi_1$  values we obtain are  $0.84 \pm 0.29(\text{stat})$ ,  $1.33 \pm 0.30(\text{stat})$  and  $0.94 \pm 0.20(\text{stat})$ , respectively. Again the results are consistent within the statistical fluctuations.

### A. Systematic errors

The sources of systematic error we consider are listed in Table IX [37]. Adding all the systematic errors in quadrature, we obtain

$$\sin 2\phi_1 = 0.99 \pm 0.14(\text{stat}) \pm 0.06(\text{syst}).$$

Below we explain each item in order.

#### 1. Vertex reconstruction

The largest contribution comes from vertex reconstruction. We searched for possible biases by using two different vertexing algorithms and changing the track selection criteria for the tag-side vertex. In the alternative vertexing algorithm, we first obtain a seed vertex using tracks of good quality: an impact parameter from the IP in the  $r$ - $\phi$  direction is smaller than 2.5 times the  $r$ - $\phi$  vertex error, the vertex error in  $z$  is less than 0.5 mm, and the cms momenta are larger than 0.3 GeV/c. We then repeat the vertex fit using tracks within 3 (4)  $\sigma$  in  $z$  from the seed vertex for the cms track momentum less (larger) than 1 GeV/c, where  $\sigma$  is the error of the seed vertex in  $z$ . We also estimated the effects of the vertex

TABLE VII. Summary of  $\sin 2\phi_1$  fit results. Only statistical errors are shown.

Sample	Events	$\sin 2\phi_1$
$f_{\text{tag}}=B^0$ ( $q=+1$ )	560	$0.84 \pm 0.21$
$f_{\text{tag}}=\bar{B}^0$ ( $q=-1$ )	577	$1.11 \pm 0.17$
$(c\bar{c})K_S^0$	578	$0.84 \pm 0.17$
$J/\psi K_S^0(\pi^+\pi^-)$	387	$0.81 \pm 0.20$
$(c\bar{c})K_S^0$ except $J/\psi K_S^0(\pi^+\pi^-)$	191	$1.00 \pm 0.40$
$J/\psi K_L^0$	523	$1.31 \pm 0.23$
$J/\psi K^{*0}(K_S^0\pi^0)$ [36]	36	$0.97 \pm 1.40$
All	1137	$0.99 \pm 0.14$

TABLE VIII. The  $r$  dependence of  $\sin 2\phi_1$  fit result. Only statistical errors are shown.

$r$ region	0.0–0.5	0.5–0.75	0.75–0.875	0.875–1.0
Events	613	239	119	166
$\sin 2\phi_1$	$0.60^{+0.65}_{-0.67}$	$0.40^{+0.31}_{-0.32}$	$1.56^{+0.25}_{-0.29}$	$1.05^{+0.15}_{-0.18}$

resolution tails using samples with small  $\Delta t$  ( $|\Delta t| < 6$  ps) and tighter vertex quality cuts.

## 2. Resolution function

We estimate the contribution due to the uncertainty in the resolution function by varying its parameters (given in Sec. V B) by  $\pm 1\sigma$ .

## 3. Wrong tag fraction

Systematic errors due to uncertainties in the wrong tag fractions given in Sec. IV B (Table IV) are studied by varying the wrong tag fraction individually for each  $r$  region. We added the contributions from each variation in quadrature.

## 4. Physics parameters

The  $B$  meson lifetime and mixing parameter are fixed to the world average values [6] in our fit, i.e.  $\tau_B = (1.548 \pm 0.032)$  ps and  $\Delta m_d = (0.472 \pm 0.017)$  ps $^{-1}$ . We estimate the systematic error by repeating the fit varying these parameters by their errors. Another physics-related uncertainty is the  $CP$  eigenvalue of  $J/\psi K^{*0}$  ( $\xi_{J/\psi K^{*0}}$ ) measured from the angular distribution of the decay daughters [33]. This systematic uncertainty is determined from the  $\pm 1\sigma$  uncertainty in the measurement.

## 5. Background fraction except for $J/\psi K_L^0$

The background fraction in our PDF,  $1 - f_{\text{sig}}$ , is calculated from the signal and background distribution functions of  $\Delta E$  and  $M_{\text{bc}}$  as described in Sec. V D. The distribution functions of  $\Delta E$  and  $M_{\text{bc}}$  are determined from data or the MC simulation depending on the decay mode. To estimate the systematic errors associated with the choice of parametrization, we varied the parameters obtained from the MC simulation by  $\pm 2\sigma$  and the parameters obtained from the data by  $\pm 1\sigma$ . The likelihood fit was repeated. A wider range

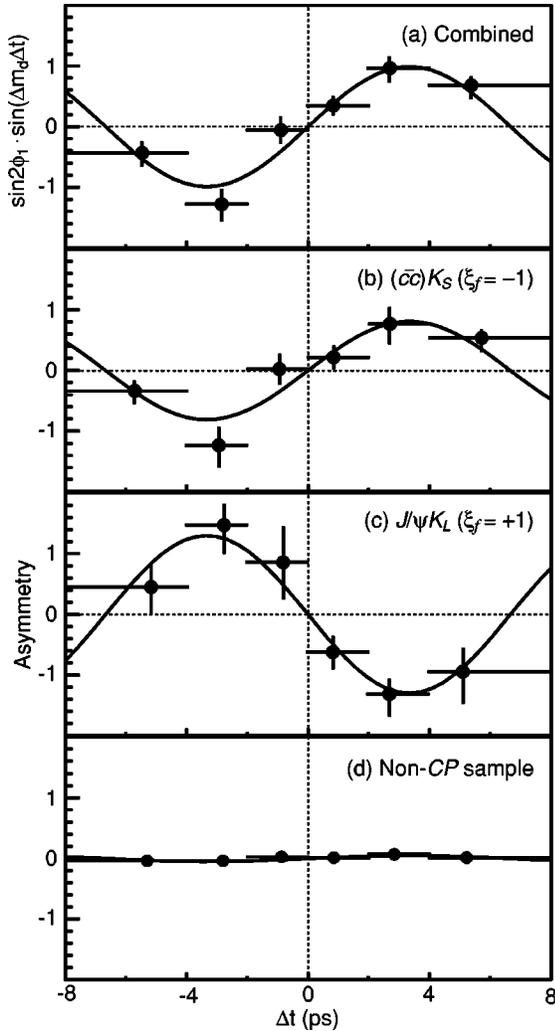


FIG. 15. (a) The asymmetry obtained from separate fits to each  $\Delta t$  bin for the full data sample; the curve is the result of the global fit. The corresponding plots for the (b)  $(c\bar{c})K_S^0$  ( $\xi_f = -1$ ), (c)  $J/\psi K_L^0$  ( $\xi_f = +1$ ) and (d) non- $CP$   $B^0$  control samples are also shown. The curves are the results of the combined fit applied separately to the individual data samples.

TABLE IX. List of systematic errors on  $\sin 2\phi_1$ .

Source	+ error	- error
vertex reconstruction	+0.040	-0.040
resolution function	+0.022	-0.032
wrong tag fraction	+0.022	-0.025
physics ( $\tau_{B^0}, \Delta m_d, J/\psi K^{*0}$ )	+0.007	-0.004
background fraction (except for $J/\psi K_L^0$ )	+0.003	-0.004
background fraction ( $J/\psi K_L^0$ )	+0.020	-0.020
background shape	+0.001	-0.001
Total	+0.06	-0.06

of uncertainty was conservatively chosen for parameters obtained from the MC simulation to take into account the possible difference between the MC simulation and data. We also estimated the systematic errors for the integrated background fractions, listed in Table VI, by varying these parameters by  $\pm 1\sigma$ . We added the results of these calculations for each decay mode in quadrature.

### 6. Background fraction for $J/\psi K_L^0$

As described in Sec. III C, the background fraction for the  $J/\psi K_L^0$  sample is obtained from a fit to  $p_B^{\text{cms}}$  distribution and is given in Table II. In this fit, the sum of components is automatically constrained to the total number of events in the signal region. Thus, the signal yield and the size of other backgrounds are strongly anti-correlated. To determine the systematic error on  $\sin 2\phi_1$  that comes from the uncertainty of the background, we need to take this anti-correlation into account. To this end, we repeat the fit to the  $p_B^{\text{cms}}$  distribution with the background fractions as free parameters but with the signal yield fixed  $+1\sigma$  or  $-1\sigma$  away from the central value obtained in the nominal fit. The resultant background yields are used to repeat the procedure to obtain  $\sin 2\phi_1$ . We regard the difference between the thus obtained value and our nominal  $\sin 2\phi_1$  value as the systematic error. We also check the systematic error due to the uncertainty in the  $CP$  content of the background. We repeated the fit varying parameters to determine the various background fractions. Since these parameters are obtained from the MC simulation, we estimate the systematic error by conservatively changing each parameter by  $\pm 2\sigma$  and adding the resulting changes in quadrature.

### 7. Background shape

The parameters that determine  $\mathcal{P}_{\text{bkg}}(\Delta t)$  and  $R_{\text{bkg}}$ , given in Sec. V C, are varied within their errors and fits are repeated.

### B. Crosschecks

We performed several crosschecks: The fitting procedure was examined using MC samples based on our likelihood functions (toy MC samples). We also measured the  $B$  meson lifetime using the same vertex reconstruction method. In addition we tested non- $CP$  control samples. These crosschecks are described below.

#### 1. Ensemble test

A thousand toy MC samples, each containing 1137 events, are generated based on our likelihood function to check the fitting procedure. Figure 16 shows the distribution of residuals [ $\sin 2\phi_1(\text{fit}) - \sin 2\phi_1(\text{input})$ ], pulls (residuals divided by fit errors), and the positive and negative errors on  $\sin 2\phi_1$  returned from the fits. All the toy MC samples have an input value of  $\sin 2\phi_1 = 0.99$ . The center of the residual in Fig. 16(a) is consistent with zero, and the standard deviation of the pull distribution in Fig. 16(b) is consistent with unity. Therefore the global fit returns the input  $\sin 2\phi_1$  value and a reasonable error. Figures 16(c) and 16(d) also show that the positive and negative errors obtained from the fit are consistent with expectations.

We also generate toy MC samples for  $J/\psi K_S^0$  and  $J/\psi K_L^0$ . The average negative errors are 0.19 and 0.28 for  $J/\psi K_S^0$  and  $J/\psi K_L^0$ , respectively. Our measurement for  $J/\psi K_S^0$  (0.20) is in good agreement with the expectation, while the result for  $J/\psi K_L^0$  (0.23) is smaller. We obtain the probability of obtaining smaller errors than this measurement to be 1.4% for  $J/\psi K_L^0$ , which is within a possible fluctuation.

### 2. $B^0$ lifetime

The  $B^0$  lifetime has been measured with the same data sample. We apply the same vertex reconstruction algorithm for fully reconstructed  $B^0$  decays as for the  $CP$  decays and the tracks on the tag side. Unbinned maximum likelihood fits are made with an exponential PDF convolved with the same  $\Delta t$  resolution function and background PDF as in the fit for  $CP$  eigenstates. For the combined  $B^0 \rightarrow J/\psi K^{*0}(K^+\pi^-)$ ,  $D^{(*)-}\pi^+$ ,  $D^{*-}\rho^+$  and  $D^{*-}l^+\nu$  decay modes, the  $B^0$  lifetime is measured to be  $\tau_{B^0} = 1.547 \pm 0.021(\text{stat})$  ps. The result is consistent with the world average value [6].

### 3. Tests on control samples

We use control samples of non- $CP$  eigenstates,  $B^0 \rightarrow J/\psi K^{*0}(K^+\pi^-)$ ,  $B^0 \rightarrow D^{(*)-}\pi^+$ ,  $B^0 \rightarrow D^{*-}\rho^+$  and  $B^0 \rightarrow D^{*-}l^+\nu$ , to check for biases in the analysis. We perform the same fit to these control samples as for the  $CP$  eigenstate modes. The results, summarized in Table X, show no systematic tendency. A combined fit to all the modes yields  $0.05 \pm 0.04$ , consistent with zero at the  $1.2\sigma$  level, as shown in Fig. 15(d).

We check for a possible bias due to  $CP$  asymmetry in the background. We fit  $J/\psi K_L^0$  candidates in the background region ( $1.0 < p_B^{\text{cms}} < 2.0$  GeV/c) treating all the events as  $J/\psi K_L^0$  candidates. Note that the fraction of events with definite  $CP$  in this region of  $p_B^{\text{cms}}$  is expected to be negligible. The result is  $\sin 2\phi_1 = 0.49 \pm 0.35$ , consistent with zero at the  $1.4\sigma$  level.

### C. Discussion

We have performed several statistical analyses of the results described in the previous sections. Using a Gaussian likelihood function based on the statistical and the systematic errors, we calculated the confidence intervals bounded by the physical region for  $\sin 2\phi_1$  using two methods: the Feldman-Cousins [38] frequentist approach and the Bayesian method with a flat prior PDF. We find a lower bound on  $\sin 2\phi_1$  of 0.70 at the 95% C.L. in both cases. We also estimated the Bayesian lower limit using the exact likelihood function, shown in Fig. 14, and obtained 0.69. We conclude that the likelihood function is Gaussian to a good approximation. Combinations of indirect measurements typically constrain  $0.50 < \sin 2\phi_1 < 0.86$  in the framework of the SM [39]. Although our measured value is large, it is consistent with the higher range of the SM prediction. We are continuing the measurement with much higher statistics in order to test the KM ansatz more precisely.

Finally we comment on the possibility of direct  $CP$  violation. The signal PDF for a neutral  $B$  meson decaying into a

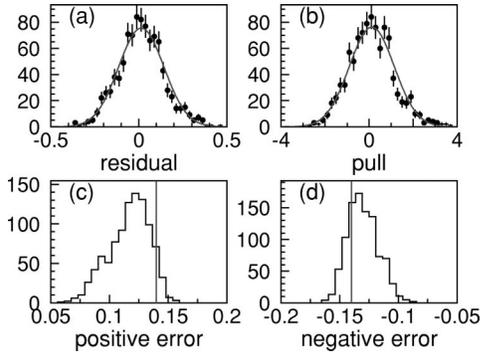


FIG. 16. The result of our toy Monte Carlo test of our fit: (a) the distribution of residuals [ $\sin 2\phi_1(\text{fit}) - \sin 2\phi_1(\text{input})$ ], (b) pull distribution [(residual)/(error of fit)], and (c) the positive and (d) negative errors of  $\sin 2\phi_1$  returned from the fits. The vertical lines in (c) and (d) indicate the errors obtained from the fit to data.

$CP$  eigenstate [Eqs. (10) and (11)] can be expressed in a more general form as

$$\mathcal{P}_{\text{sig}}(\Delta t, q, w_l, \xi_f) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{2\tau_{B^0}(1+|\lambda|^2)} \times \left\{ \frac{1+|\lambda|^2}{2} - q(1-2w_l) \right. \\ \times \left[ \xi_f \mathcal{A} |\lambda| \sin(\Delta m_d \Delta t) \right. \\ \left. \left. - \frac{|\lambda|^2 - 1}{2} \cos(\Delta m_d \Delta t) \right] \right\}, \quad (24)$$

where  $\lambda$  is a complex parameter that depends on both  $B^0\bar{B}^0$  mixing and on the amplitudes for  $B^0$  and  $\bar{B}^0$  decay to a  $CP$  eigenstate [2]. The coefficient of  $\sin(\Delta m_d \Delta t)$  is given by  $\mathcal{A}|\lambda| = -\xi_f \text{Im} \lambda$  and is equal to  $\sin 2\phi_1$  in the SM. The presence of the cosine term ( $|\lambda| \neq 1$ ) would indicate direct  $CP$  violation. Throughout our study we have assumed  $|\lambda| = 1$ , as expected in the SM [2]. In order to test this assumption, we also performed a fit using the above expression with  $\mathcal{A}$  and  $|\lambda|$  as free parameters, keeping everything else the same. We obtain  $|\lambda| = 1.09 \pm 0.14$  and  $\mathcal{A} = 0.80 \pm 0.19$  for the  $J/\psi K_S^0(\pi^+\pi^-)$  sample only, and  $|\lambda| = 1.03 \pm 0.09$  and  $\mathcal{A} = 0.99 \pm 0.14$  for all  $CP$  modes combined, where the errors are statistical only. This result confirms the assumption used in our analysis.

TABLE X.  $CP$  asymmetries for control samples.

	$J/\psi K^{*0}(K^+\pi^-)$	$D^{(*)-}\pi^+$ and $D^{*-}\rho^+$	$D^{*-}l^+\nu$
Events	816	5560	10232
$CP$ asymmetry	$0.01 \pm 0.14$	$0.12 \pm 0.06$	$0.01 \pm 0.05$

## VII. CONCLUSION

We have measured the  $CP$  violation parameter  $\sin 2\phi_1$  at the KEKB asymmetric  $e^+e^-$  collider using a data sample of  $29.1 \text{ fb}^{-1}$  recorded on the  $\Upsilon(4S)$  resonance with the Belle detector. To extract  $\sin 2\phi_1$ , we apply a maximum likelihood fit to the 1137 candidate  $B$  meson decays to  $CP$  eigenstates. We obtain

$$\sin 2\phi_1 = 0.99 \pm 0.14(\text{stat}) \pm 0.06(\text{sys}).$$

We conclude that there is large  $CP$  violation in the neutral  $B$  meson system. A zero value for  $\sin 2\phi_1$  is ruled out by more than six standard deviations.

We have calculated the confidence intervals bounded by the  $\sin 2\phi_1$  physical region for the Bayesian and Feldman-Cousins frequentist approaches. The lower bound is found to be 0.70 at the 95% C.L. in both cases. Our result is consistent with the higher range of values allowed from the constraints of the standard model as well as with our previous measurements [9].

## ACKNOWLEDGMENTS

We wish to thank the KEKB accelerator group for the excellent operation of the KEKB accelerator. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and the Japan Society for the Promotion of Science; the Australian Research Council and the Australian Department of Industry, Science and Resources; the National Science Foundation of China under contract No. 10175071; the Department of Science and Technology of India; the BK21 program of the Ministry of Education of Korea and the CHEP SRC program of the Korea Science and Engineering Foundation; the Polish State Committee for Scientific Research under contract No. 2P03B 17017; the Ministry of Science and Technology of the Russian Federation; the Ministry of Education, Science and Sport of the Republic of Slovenia; the National Science Council and the Ministry of Education of Taiwan; and the U.S. Department of Energy.

- [1] M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **49**, 652 (1973).  
 [2] H. Quinn and A.I. Sanda, *Eur. Phys. J. C* **15**, 628 (2000). Another naming convention,  $\beta$  ( $=\phi_1$ ),  $\alpha$  ( $=\phi_2$ ) and  $\gamma$  ( $=\phi_3$ ), is also used in the literature.  
 [3] A.B. Carter and A.I. Sanda, *Phys. Rev. Lett.* **45**, 952 (1980); *Phys. Rev. D* **23**, 1567 (1981); I.I. Bigi and A.I. Sanda, *Nucl. Phys.* **B193**, 85 (1981).  
 [4] MAC Collaboration, E. Fernandez *et al.*, *Phys. Rev. Lett.* **51**,

- 1022 (1983); Mark II Collaboration, N.S. Lockyer *et al.*, *ibid.* **51**, 1316 (1983).  
 [5] ARGUS Collaboration, H. Albrecht *et al.*, *Phys. Lett. B* **192**, 245 (1987).  
 [6] Particle Data Group, D.E. Groom *et al.*, *Eur. Phys. J. C* **15**, 1 (2000).  
 [7] Within the standard model the short-distance contributions guarantee that the sign of  $\Delta m_d$  is positive. A detailed discussion is found in I.I. Bigi and A.I. Sanda, *Phys. Rev. D* **60**,

- 033001 (1999).
- [8] OPAL Collaboration, K. Ackerstaff *et al.*, *Eur. Phys. J. C* **5**, 379 (1998); CDF Collaboration, T. Affolder *et al.*, *Phys. Rev. D* **61**, 072005 (2000); ALEPH Collaboration, B. Barate *et al.*, *Phys. Lett. B* **492**, 259 (2000); BaBar Collaboration, B. Aubert *et al.*, *Phys. Rev. Lett.* **86**, 2515 (2001).
- [9] Belle Collaboration, A. Abashian *et al.*, *Phys. Rev. Lett.* **86**, 2509 (2001).
- [10] BaBar Collaboration, B. Aubert *et al.*, *Phys. Rev. Lett.* **87**, 091801 (2001).
- [11] Belle Collaboration, K. Abe *et al.*, *Phys. Rev. Lett.* **87**, 091802 (2001).
- [12] S. Kurokawa *et al.*, KEK Preprint 2001-157, 2001.
- [13] Belle Collaboration, A. Abashian *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **479**, 117 (2002).
- [14] H. Hirano *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **455**, 294 (2000); M. Akatsu *et al.*, *ibid.* **454**, 322 (2000).
- [15] G. Alimonti *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **453**, 71 (2000).
- [16] H. Kichimi *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **453**, 315 (2000).
- [17] T. Iijima *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **453**, 321 (2000).
- [18] H. Ikeda *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **441**, 401 (2000).
- [19] K. Hanagaki *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **485**, 490 (2002).
- [20] A. Abashian *et al.*, *Nucl. Instrum. Methods Phys. Res. A* **449**, 112 (2000).
- [21] The QQ  $B$  meson decay event generator was developed by the CLEO Collaboration. See the following URL: <http://www.lns.cornell.edu/public/CLEO/soft/QQ>
- [22] CERN Program Library Long Writeup W5013, CERN, 1993.
- [23] Throughout this report, when a mode is quoted the inclusion of the charge conjugate mode is implied.
- [24] G.C. Fox and S. Wolfram, *Phys. Rev. Lett.* **41**, 1581 (1978).
- [25] For  $J/\psi K^*$  candidates, we adopt slightly different windows. We require  $2.95 \leq M_{e^+e^-} \leq 3.15 \text{ GeV}/c^2$  and  $3.05 \leq M_{\mu^+\mu^-} \leq 3.15 \text{ GeV}/c^2$ .
- [26] R. Mir, *Nucl. Phys. B (Proc. Suppl.)* **16**, 326 (1990).
- [27] Mass distributions for  $\psi(2S)$  and  $\chi_{c1}$ , as well as more details on the selection, will be available in a separate paper on branching fractions for exclusive charmonium decays, Belle Collaboration, K. Abe *et al.* (unpublished).
- [28] ARGUS Collaboration, H. Albrecht *et al.*, *Phys. Lett. B* **241**, 278 (1990).
- [29] If there is more than one pion candidate, we select the one with the invariant mass of  $K_L$  and the pion closest to the nominal  $K^*$  mass.
- [30] If two or more combinations of  $J/\psi$  and  $K_L$  clusters are found, the one that gives the best  $p_B^{\text{cms}}$  is selected.
- [31] For a precise estimation of the uncorrelated background, we require that a  $B^0 \rightarrow D^{*-} l^+ \nu$  candidate should not be in the signal region when we artificially flip the lepton momentum vector.
- [32] H. Kakuno *et al.* (unpublished).
- [33] Belle Collaboration, K. Abe *et al.*, “Measurements of Polarization and  $CP$  Asymmetry in  $B \rightarrow J/\psi + K^*$  decays,” BELLE-CONF-0105, 2001.
- [34] T. Skwarnicki, Ph.D. thesis, Institute for Nuclear Physics, Krakow, 1986.
- [35] I. Dunietz, H. Quinn, A. Snyder, W. Toki, and H.J. Lipkin, *Phys. Rev. D* **43**, 2193 (1991).
- [36] The result for  $J/\psi K^{*0}(K_S \pi^0)$  quoted in Ref. [11] was updated with an improved treatment of the background fraction.
- [37] The errors originating from the flavor tagging and from the resolution function differ slightly, after a further investigation, from the values quoted in Ref. [11]. The total systematic error is unchanged.
- [38] G.J. Feldman and R.D. Cousins, *Phys. Rev. D* **57**, 3873 (1998).
- [39] A. Höcker, H. Lacker, S. Laplace, and F. Le Diberder, *Eur. Phys. J. C* **21**, 225 (2001).