Measurement of branching fraction and first evidence of CP violation in $B^0 \to a_1^{\pm}(1260)\pi^{\mp}$ decays

J. Dalseno,[25,44] I. Adachi,[8] H. Aihara,[48] D. M. Asner,[37] V. Aulchenko,[2] T. Aushev,[14] A. M. Bakich,[42] A. Bay,[21] K. Belous,[13] B. Bhuyan,[9] A. Bozek,[33] M. Bračko,[24,15] O. Brovchenko,[17] T. E. Browder,[7] V. Chekelian,[25] A. Chen,[30] P. Chen,[32] B. G. Cheon,[6] K. Chilikin,[14] R. Chistov,[14] I.-S. Cho,[54] K. Cho,[18] Y. Choi,[41] Z. Doležal,[3] Z. Drásal,[3] S. Eidelman,[2] J. E. Fast,[37] V. Gaur,[43] N. Gabyshev,[2] A. Garmash,[2] Y. M. Goh,[6] H. Hayashii,[29] Y. Horii,[28] Y. Hoshi,[46] W.-S. Hou,[32] Y. B. Hsiung,[32] H. J. Hyun,[20] T. Iijima,[28,27] K. Inami,[27] A. Ishikawa,[47] R. Itoh,[8] M. Iwabuchi,[54] Y. Iwasaki,[8] T. Iwashita,[29] T. Julius,[26] J. H. Kang,[54] C. Kiesling,[25] H. O. Kim,[20] J. B. Kim,[19] Y. J. Kim,[18] K. Kinoshita,[4] B. R. Ko,[19] S. Koblitz,[25] P. Kodyš,[3] S. Korpar,[24,15] P. Križan,[22,15] P. Krokovny,[2] B. Kronenbitter,[17] T. Kuhr,[17] T. Kumita,[50] Y.-J. Kwon,[54] S.-H. Lee,[19] J. Li,[40] J. Libby,[10] C. Liu,[39] Z. Q. Liu,[11] R. Louvot,[21] J. MacNaughton,[8] D. Matvienko,[2] S. McOnie,[42] K. Miyabayashi,[29] H. Miyata,[35] Y. Miyazaki,[27] G. B. Mohanty,[43] D. Mohapatra,[37] A. Moll,[25,44] N. Muramatsu,[38] M. Nakao,[8] Z. Natkaniec,[33] E. Nedelkovska,[25] C. Ng,[48] S. Nishida,[8] K. Nishimura,[7] O. Nitoh,[51] S. Ogawa,[45] T. Ohshima,[27] S. Okuno,[16] P. Pakhlov,[14] G. Pakhlova,[14] C. W. Park,[41] H. K. Park,[20] T. K. Pedlar,[23] R. Pestotnik,[15] M. Petrič,[15] L. E. Piilonen,[52] M. Prim,[17] K. Prothmann,[25,44] M. Ritter,[25] M. Röhrken,[17] H. Sahoo,[7] Y. Sakai,[8] T. Sanuki,[47] O. Schneider,[21] C. Schwanda,[12] A. J. Schwartz,[4] K. Senyo,[53] O. Seon,[27] M. E. Sevior,[26] M. Shapkin,[13] V. Shebalin,[2] C. P. Shen,[27] T.-A. Shibata,[49] J.-G. Shiu,[32] A. Sibidanov,[42] F. Simon,[25,44] P. Smerkol,[15] Y.-S. Sohn,[54] E. Solovieva,[14] M. Starič,[15] M. Sumihama,[5] T. Sumiyoshi,[50] G. Tatishvili,[37] Y. Teramoto,[36] K. Trabelsi,[8] M. Uchida,[49] S. Uehara,[8] Y. Unno,[6] S. Uno,[8] P. Urquijo,[1] Y. Usov,[2] P. Vanhoefer,[25] G. Varner,[7] C. H. Wang,[31] P. Wang,[11] M. Watanabe,[35] Y. Watanabe,[16] K. M. Williams,[52] E. Won,[19] B. D. Yabsley,[42] Y. Yamashita,[34] Z. P. Zhang,[39] V. Zhilich,[2] V. Zhulanov,[2] and A. Zupanc[17]

(Belle Collaboration)

[1]*University of Bonn, Bonn 53115*
[2]*Budker Institute of Nuclear Physics SB RAS and Novosibirsk State University, Novosibirsk 630090*
[3]*Faculty of Mathematics and Physics, Charles University, 121 16 Prague*
[4]*University of Cincinnati, Cincinnati, Ohio 45221*
[5]*Gifu University, Gifu 501-1193*
[6]*Hanyang University, Seoul 133-791*
[7]*University of Hawaii, Honolulu, Hawaii 96822*
[8]*High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[9]*Indian Institute of Technology Guwahati, Guwahati 781039, Assam*
[10]*Indian Institute of Technology Madras, Chennai 600036*
[11]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049*
[12]*Institute of High Energy Physics, Vienna 1050*
[13]*Institute of High Energy Physics, Protvino 142281*
[14]*Institute for Theoretical and Experimental Physics, Moscow 117218*
[15]*J. Stefan Institute, 1000 Ljubljana*
[16]*Kanagawa University, Yokohama 221-8686*
[17]*Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Karlsruhe 76131*
[18]*Korea Institute of Science and Technology Information, Daejeon 305-806*
[19]*Korea University, Seoul 136-713*
[20]*Kyungpook National University, Daegu 702-701*
[21]*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015*
[22]*Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana*
[23]*Luther College, Decorah, Iowa 52101*
[24]*University of Maribor, 2000 Maribor*
[25]*Max-Planck-Institut für Physik, 80805 München*
[26]*School of Physics, University of Melbourne, Victoria 3010*
[27]*Graduate School of Science, Nagoya University, Nagoya 464-8602*
[28]*Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602*
[29]*Nara Women's University, Nara 630-8506*
[30]*National Central University, Chung-li 32054*
[31]*National United University, Miao Li 36003*
[32]*Department of Physics, National Taiwan University, Taipei 10617*

[33]*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
[34]*Nippon Dental University, Niigata 951-8580*
[35]*Niigata University, Niigata 950-2181*
[36]*Osaka City University, Osaka 558-8585*
[37]*Pacific Northwest National Laboratory, Richland, Washington 99352*
[38]*Research Center for Electron Photon Science, Tohoku University, Sendai 980-8578*
[39]*University of Science and Technology of China, Hefei 230026*
[40]*Seoul National University, Seoul 151-742*
[41]*Sungkyunkwan University, Suwon 440-746*
[42]*School of Physics, University of Sydney, New South Wales 2006*
[43]*Tata Institute of Fundamental Research, Mumbai 400005*
[44]*Excellence Cluster Universe, Technische Universität München, 85748 Garching*
[45]*Toho University, Funabashi 274-8510*
[46]*Tohoku Gakuin University, Tagajo 985-8537*
[47]*Tohoku University, Sendai 980-8578*
[48]*Department of Physics, University of Tokyo, Tokyo 113-0033*
[49]*Tokyo Institute of Technology, Tokyo 152-8550*
[50]*Tokyo Metropolitan University, Tokyo 192-0397*
[51]*Tokyo University of Agriculture and Technology, Tokyo 184-8588*
[52]*CNP, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*
[53]*Yamagata University, Yamagata 990-8560*
[54]*Yonsei University, Seoul 120-749*

(Received 31 May 2012; published 26 November 2012)

We present a measurement of the branching fraction and time-dependent CP violation parameters in $B^0 \to a_1^\pm(1260)\pi^\mp$ decays. The results are obtained from the final data sample containing 772×10^6 $B\bar{B}$ pairs collected at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. We obtain the product branching fraction $\mathcal{B}(B^0 \to a_1^\pm(1260)\pi^\mp) \times \mathcal{B}(a_1^\pm(1260) \to \pi^\pm\pi^\mp\pi^\pm) = (11.1 \pm 1.0(\text{stat}) \pm 1.4(\text{syst})) \times 10^{-6}$ and an upper limit on the product branching fraction for a possible decay with the same final state $\mathcal{B}(B^0 \to a_2^\pm(1320)\pi^\mp) \times \mathcal{B}(a_2^\pm(1320) \to \pi^\pm\pi^\mp\pi^\pm) < 2.2 \times 10^{-6}$ at 90% CL. In a time-dependent measurement to extract CP asymmetries, we obtain the CP violation parameters $\mathcal{A}_{CP} = -0.06 \pm 0.05(\text{stat}) \pm 0.07(\text{syst})$, $\mathcal{C}_{CP} = -0.01 \pm 0.11(\text{stat}) \pm 0.09(\text{syst})$, $\mathcal{S}_{CP} = -0.51 \pm 0.14(\text{stat}) \pm 0.08(\text{syst})$, representing time- and flavor-integrated direct, flavor-dependent direct and mixing-induced CP violation, respectively. Simultaneously, we also extract the CP-conserving parameters $\Delta\mathcal{C} = +0.54 \pm 0.11(\text{stat}) \pm 0.07(\text{syst})$, $\Delta\mathcal{S} = -0.09 \pm 0.14(\text{stat}) \pm 0.06(\text{syst})$, which, respectively, describe a rate difference and strong phase difference between the decay channels where the $a_1^\pm$ does not contain the spectator quark and those where it does. We find first evidence of mixing-induced CP violation in $B^0 \to a_1^\pm(1260)\pi^\mp$ decays with 3.1σ significance. The rate where the $a_1^\pm$ does not contain the spectator quark from the B meson is found to dominate the rate where it does at the 4.1σ level. However, there is no evidence for either time- and flavor-integrated direct CP violation or flavor-dependent direct CP violation.

DOI: 10.1103/PhysRevD.86.092012 PACS numbers: 11.30.Er, 12.15.Hh, 13.25.Hw

I. INTRODUCTION

CP violation in the standard model arises from a complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix [1,2]. Mixing-induced CP violation in the B sector has been clearly observed by the *BABAR* [3] and Belle [4] Collaborations in the $\bar{b} \to \bar{c}c\bar{s}$-induced decay, $B^0 \to J/\psi K_S^0$, while many other modes provide additional information on CP-violating parameters.

Decays that proceed dominantly through the $\bar{b} \to \bar{u}u\bar{d}$ transition are sensitive to the interior angle of the unitarity triangle $\phi_2(\alpha) \equiv \arg(-V_{td}V_{tb}^*)/(V_{ud}V_{ub}^*)$. The *BABAR* and Belle Collaborations have reported time-dependent CP asymmetries in these modes that include decays such as $B^0 \to \pi^+\pi^-$ [5,6], $\rho^\pm\pi^\mp$ [7,8] and $\rho^+\rho^-$ [9,10].

This paper describes the measurement of the branching fraction and time-dependent CP-violation parameters of the $\bar{b} \to \bar{u}u\bar{d}$ channel, $B^0 \to a_1^\pm\pi^\mp$, shown in Fig. 1. The left diagram shows the dominant first-order or tree process, while the right diagram shows the leading second-order loop or penguin process. This analysis can be used to test the QCD factorization framework which has been used to predict the branching fraction and CP asymmetries of this decay channel [11–13]. Similar to $\rho^\pm\pi^\mp$, the state $a_1^\pm\pi^\mp$ is not a CP eigenstate; rather, it is a flavor nonspecific state with four flavor-charge configurations that must be considered: $B^0(\bar{B}^0) \to a_1^\pm\pi^\mp$ [14]. The combined information of the B flavor and a_1 charge allows the determination of additional information compared to CP eigenstates. It allows us to separate the cases where the d quark from

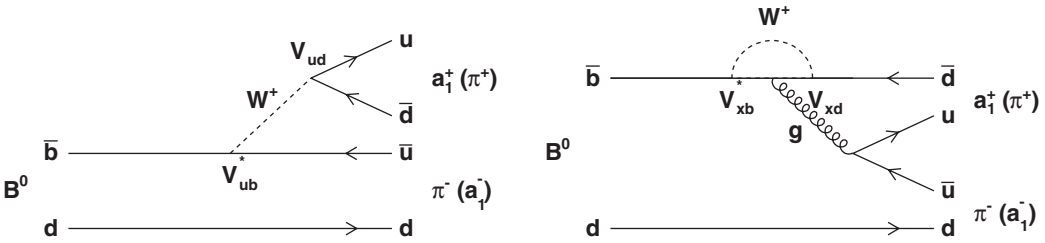

FIG. 1. Leading-order tree (left) and penguin (right) diagrams for the decay $B^0 \to a_1^\pm \pi^\mp$, where the parentheses in the figure indicate the two possible decays of the B^0. The d quark may become part of the π^- or a_1^-. In the penguin diagram, the x in V_{xb} refers to the flavor of the intermediate-state quark ($x = u, c, t$).

the B meson, which does not participate in the interaction (spectator), becomes part of the $a_1^\pm$ or $\pi^\pm$.

The decay of the $\Upsilon(4S)$ can produce a $B^0\bar{B}^0$ pair that must be coherent, of which one ($B^0_{\rm Rec}$) may be reconstructed in the $a_1^\pm\pi^\mp$ final state. This final state does not determine whether the $B^0_{\rm Rec}$ decayed as a B^0 or as a $\bar{B}^0$. The other B meson in the event ($B^0_{\rm Tag}$), however, can be reconstructed in a final state that determines its b flavor and, therefore, the flavor of the $B^0_{\rm Rec}$ at the time of the $B^0_{\rm Tag}$ decay. The proper time interval between $B^0_{\rm Rec}$ and $B^0_{\rm Tag}$, which decay at time $t_{\rm Rec}$ and $t_{\rm Tag}$, respectively, is defined as $\Delta t \equiv t_{\rm Rec} - t_{\rm Tag}$. For the case of coherent $B^0\bar{B}^0$ pairs, the time-dependent decay rate in the quasi-two-body approximation, when $B^0_{\rm Tag}$ possesses flavor q (B^0: $q = +1$; $\bar{B}^0$: $q = -1$) and the a_1 possesses charge c (a_1^+: $c = +1$; a_1^-: $c = -1$), is given by [14]

$$\mathcal{P}(\Delta t, q, c) = (1 + c\mathcal{A}_{CP})\frac{e^{-|\Delta t|/\tau_{B^0}}}{8\tau_{B^0}} \times \{1 + q[(\mathcal{S}_{CP} + c\Delta\mathcal{S})\sin\Delta m_d\Delta t - (\mathcal{C}_{CP} + c\Delta\mathcal{C})\cos\Delta m_d\Delta t]\}. \quad (1)$$

Here, τ_{B^0} is the B^0 lifetime and Δm_d is the mass difference between the two mass eigenstates of the neutral B meson. This assumes CPT invariance, no CP violation in the mixing, and a negligible difference in the decay rates between the two mass eigenstates. The parameter $\mathcal{A}_{CP}$ measures time- and flavor-integrated direct CP violation,

$$\mathcal{A}_{CP} = \frac{\Gamma(B \to a_1^+\pi^-) - \Gamma(B \to a_1^-\pi^+)}{\Gamma(B \to a_1^+\pi^-) + \Gamma(B \to a_1^-\pi^+)}. \quad (2)$$

The parameter $\mathcal{S}_{CP}$ measures mixing-induced CP violation, and $\mathcal{C}_{CP}$ measures flavor-dependent direct CP violation. The quantity $\Delta\mathcal{C}$ measures the rate asymmetry between the flavor-charge configurations, where the a_1 does not contain the spectator quark, $(\Gamma[B^0 \to a_1^+\pi^-] + \Gamma[\bar{B}^0 \to a_1^-\pi^+])$, and where it does contain the spectator quark, $(\Gamma[B^0 \to a_1^-\pi^+] + \Gamma[\bar{B}^0 \to a_1^+\pi^-])$, while $\Delta\mathcal{S}$ is related to the strong phase difference between these two processes, δ. These two parameters are not sensitive to CP violation. Sensitivity to ϕ_2 comes from this relation between four of the measured parameters,

$$\mathcal{S}_{CP} \pm \Delta\mathcal{S} = \sqrt{1 - (\mathcal{C}_{CP} \pm \Delta\mathcal{C})^2}\,\sin(2\phi_2^{\rm eff\pm} \pm \delta). \quad (3)$$

A feature common to this channel and the other $\bar{b} \to \bar{u}u\bar{d}$ modes mentioned earlier is that an effective angle, $\phi_2^{\rm eff}$, is determined, rather than ϕ_2 itself, due to the possible presence of additional loop contributions. In the limit that only the dominant tree amplitude contributes, no flavor-dependent CP violation is expected and $\phi_2^{\rm eff\pm} = \phi_2$. Fortunately, this inconvenience can be overcome with bounds on $\Delta\phi_2 \equiv \phi_2 - \phi_2^{\rm eff}$ determined using either an isospin analysis [15] or $SU(3)$ flavor symmetry [14].

From these parameters, $\phi_2^{\rm eff}$ can be determined up to a four-fold ambiguity [14],

$$\phi_2^{\rm eff} = \frac{1}{4}\left[\arcsin\left(\frac{\mathcal{S}_{CP} + \Delta\mathcal{S}}{\sqrt{1 - (\mathcal{C}_{CP} + \Delta\mathcal{C})^2}}\right) + \arcsin\left(\frac{\mathcal{S}_{CP} - \Delta\mathcal{S}}{\sqrt{1 - (\mathcal{C}_{CP} - \Delta\mathcal{C})^2}}\right)\right]. \quad (4)$$

These results can also be transformed into more physically intuitive parameters that characterize direct CP violation in decays with particular topologies [16],

$$A_{+-} = \frac{-(\mathcal{A}_{CP} + \mathcal{C}_{CP} + \mathcal{A}_{CP}\Delta\mathcal{C})}{1 + \Delta\mathcal{C} + \mathcal{A}_{CP}\mathcal{C}_{CP}}, \qquad A_{-+} = \frac{(-\mathcal{A}_{CP} + \mathcal{C}_{CP} + \mathcal{A}_{CP}\Delta\mathcal{C})}{-1 + \Delta\mathcal{C} + \mathcal{A}_{CP}\mathcal{C}_{CP}}, \quad (5)$$

which describe CP violation involving diagrams where the a_1 contains and does not contain the spectator quark, respectively.

The *BABAR* Collaboration has performed a branching fraction measurement of $B^0 \to a_1^\pm\pi^\mp$ with 218×10^6 $B\bar{B}$ pairs [17] and a time-dependent CP violation measurement with 384×10^6 $B\bar{B}$ pairs [18]. These results are collected in Table I. From a subsequent study of $B \to K_1(1270)\pi$ and $B \to K_1(1400)\pi$, the *BABAR* Collaboration has also obtained bounds on $|\Delta\phi_2|$ using $SU(3)$ flavor symmetry [19].

Two separate measurements are described in this paper. In Sec. II, we briefly describe the data set and Belle detector. The branching fraction measurement is described in Sec. III. There, we explain the selection criteria used to

TABLE I. Summary of physics parameters for $B^0 \to a_1^\pm \pi^\mp$ obtained by the *BABAR* Collaboration [17,18].

Parameter	Value
$\mathcal{B}(B^0 \to a_1^\pm \pi^\mp) \times \mathcal{B}(a_1^\pm(1260) \to \pi^\pm\pi^\mp\pi^\pm)$	$(16.6 \pm 1.9(\text{stat}) \pm 1.5(\text{syst})) \times 10^{-6}$
$\mathcal{A}_{CP}$	$-0.07 \pm 0.07(\text{stat}) \pm 0.02(\text{syst})$
$\mathcal{C}_{CP}$	$-0.10 \pm 0.15(\text{stat}) \pm 0.09(\text{syst})$
$\mathcal{S}_{CP}$	$+0.37 \pm 0.21(\text{stat}) \pm 0.07(\text{syst})$
$\Delta\mathcal{C}$	$+0.26 \pm 0.15(\text{stat}) \pm 0.07(\text{syst})$
$\Delta\mathcal{S}$	$-0.14 \pm 0.21(\text{stat}) \pm 0.06(\text{syst})$

obtain signal candidates and suppress backgrounds followed by the fit method used to extract the signal component. After this, the results of the fit are presented along with a discussion of the systematic uncertainties. In Sec. IV, these same issues are described again for the time-dependent *CP*-violation measurement followed by our conclusions in Sec. V.

II. DATA SET AND BELLE DETECTOR

This measurement of the branching fraction and time-dependent *CP* violation parameters in $B^0 \to a_1^\pm \pi^\mp$ decays is based on the final data sample containing 772×10^6 $B\bar{B}$ pairs collected with the Belle detector at the KEKB asymmetric-energy e^+e^- (3.5 on 8 GeV) collider [20]. At the $\Upsilon(4S)$ resonance ($\sqrt{s} = 10.58$ GeV), the Lorentz boost of the produced $B\bar{B}$ pairs was $\beta\gamma = 0.425$ along the z direction, which is opposite the positron beam direction.

The Belle detector is a large solid-angle magnetic spectrometer that consists of a silicon vertex detector (SVD), a 50-layer central drift chamber, an array of aerogel threshold Cherenkov counters, a barrel-like arrangement of time-of-flight scintillation counters, and an electromagnetic calorimeter comprising of CsI(Tl) crystals (ECL) located inside a superconducting solenoid coil that provides a 1.5-T magnetic field. An iron flux-return located outside of the coil is instrumented to detect K_L^0 mesons and to identify muons (KLM). The detector is described in detail elsewhere [21]. Two inner detector configurations were used. A 2.0-cm radius beam pipe and a three-layer silicon vertex detector (SVD1) was used for the first sample of 152×10^6 $B\bar{B}$ pairs, while a 1.5-cm radius beam pipe, a four-layer silicon detector (SVD2) and a small-cell inner drift chamber were used to record the remaining 620×10^6 $B\bar{B}$ pairs [22]. We use a GEANT-based Monte Carlo (MC) simulation to model the response of the detector and determine its acceptance [23].

III. BRANCHING FRACTION MEASUREMENT

A. Event selection

We reconstruct $B^0 \to a_1^\pm \pi^\mp$, where $a_1^\pm \to (\pi^+\pi^-)\pi^\pm$. The $a_1^\pm$ decay proceeds mainly through the $\rho^0\pi^\pm$ and $f_0(600)\pi^\pm$ intermediate states [24]. We assume that the $\rho^0\pi^\pm$ intermediate state gives the dominant contribution and treat the $f_0(600)\pi^\pm$ contribution separately in the systematic uncertainties. Thus, the signal MC events for establishing the selection criteria are generated as $B^0 \to a_1^\pm \pi^\mp$ decays, where $a_1^\pm \to \rho^0\pi^\pm$. The $a_1^\pm$ mass and width are taken to be $m_{a_1} = 1.23$ GeV/c^2 and $\Gamma_{a_1} = 0.40$ GeV/c^2 [24].

Charged tracks are identified using a loose requirement of distance of closest approach to the interaction point (IP) along the beam direction, $|dz| < 4.0$ cm, and in the transverse direction, $dr < 0.4$ cm. With information obtained from the central drift chamber, aerogel threshold Cherenkov counters and time-of-flight scintillation counters, particle identification is determined with the likelihood ratio $\mathcal{L}_i/(\mathcal{L}_i + \mathcal{L}_j)$. Here, $\mathcal{L}_i$ ($\mathcal{L}_j$) is the likelihood that the particle is of type i (j). A requirement of $\mathcal{L}_{K/\pi} < 0.4$ is placed on all charged pion candidates, which retains 91% of all pions from $B^0 \to a_1^\pm \pi^\mp$ but only 9% of kaons. To further suppress background from particle misidentification, vetoes are applied on particles consistent with the electron or proton hypotheses. Additional SVD requirements of two z hits and one r-ϕ hit [25] are imposed on the charged tracks so that a good quality vertex of the reconstructed B candidate can be determined.

The intermediate dipion state is reconstructed above the K_S^0 region with an invariant mass $0.52\ \text{GeV}/c^2 < m(\pi^+\pi^-) < 1.1\ \text{GeV}/c^2$. This is combined with another pion and forms an $a_1^\pm$ candidate if the invariant mass is in the window $0.85\ \text{GeV}/c^2 < m_{3\pi} < 1.75\ \text{GeV}/c^2$ below the charm threshold. Upon combination with another pion, a B candidate is formed.

Reconstructed B candidates are described with two nearly uncorrelated kinematic variables: the beam-energy-constrained mass $M_{\text{bc}} \equiv \sqrt{(E_{\text{beam}}^{\text{CMS}})^2 - (p_B^{\text{CMS}})^2}$ and the energy difference $\Delta E \equiv E_B^{\text{CMS}} - E_{\text{beam}}^{\text{CMS}}$, where $E_{\text{beam}}^{\text{CMS}}$ is the beam energy and E_B^{CMS} (p_B^{CMS}) is the energy (momentum) of the B meson, all evaluated in the center-of-mass system (CMS). The B candidates that satisfy $M_{\text{bc}} > 5.27\ \text{GeV}/c^2$ and $|\Delta E| < 0.1$ GeV are selected for further analysis.

To reduce combinatorial background in forming the $a_1^\pm$ candidate, the cosine of the angle between the prompt pion from $a_1^\pm \to \rho^0\pi^\pm$ and the B candidate in the $a_1^\pm$ rest frame

is required to be between -0.85 and $+0.85$. This distribution is roughly flat for signal while peaking at ± 1 for combinatorial background. This selection retains 80% of signal events while rejecting 43% of the combinatorial background.

The dominant background in the reconstruction of $B^0_{\rm Rec}$ is from continuum ($e^+e^- \to q\bar{q}$ where $q = u,\ d,\ s,\ c$) events. Since their topology tends to be jetlike, in contrast to the spherical $B\bar{B}$ decay, continuum events can be distinguished from $B\bar{B}$ events using event-shape variables, which we combine into a Fisher discriminant $\mathcal{F}_{b\bar{b}/q\bar{q}}$ [26]. The $B\bar{B}$ training sample is taken from signal MC, while the $q\bar{q}$ training sample comes from data taken below the $\Upsilon(4S)$ resonance. The Fisher discriminant is then constructed from the following eight variables:

(i) $|\cos\theta_{\rm TB,TO}|$, where the angle is between the $B^0_{\rm Rec}$ thrust direction and the thrust of the tag side. The thrust is defined as the vector which maximizes the sum of the longitudinal momenta of the particles. For a $B\bar{B}$ event, the pair is essentially at rest in the CMS, so the thrust axis of $B^0_{\rm Rec}$ is uncorrelated with the thrust axis of $B^0_{\rm Tag}$. In a $q\bar{q}$ event, the decay products lie in the two jets that are back to back, so the two thrust axes tend to be collinear. This variable provides the strongest discrimination against continuum.

(ii) $|\cos\theta_{B,z}|$, where the angle is between the $B^0_{\rm Rec}$ flight direction and the z direction. This is the second most important variable in discriminating against continuum. $B\bar{B}$ pairs are produced in a correlated state, for which this variable follows a sine-squared distribution, whereas for $q\bar{q}$ events, the distribution is uniform.

(iii) $|\cos\theta_{{\rm TB},z}|$, where the angle is between the $B^0_{\rm Rec}$ thrust direction and the z direction. This distribution is strongly influenced by detector acceptance at large values. As thrust and decay axes are related, this variable displays similar tendencies to the previously described variable.

(iv) $\sum p_t^{\rm CMS}$, where the transverse CMS momentum sum runs over all particles on the tag side. The continuum distribution generally has a higher mean because its decay product multiplicity is lower compared to $B\bar{B}$.

(v) $L_0^{\rm c,n} \equiv \sum_{c,n} p_{c,n}^{\rm CMS}$, where the CMS momentum sum runs over either the charged tracks (c) or neutral clusters (n) on the tag side. These distributions exhibit similar tendencies to the previous variable.

(vi) $L_2^{\rm c,n} \equiv \sum_{c,n} p_{c,n}^{\rm CMS} \cos^2\theta_{p_{c,n}^{\rm CMS},{\rm TB}}$, where the CMS momentum sum runs over either the charged tracks (c) or neutral clusters (n) on the tag side. The angle is between the particle direction and the $B^0_{\rm Rec}$ thrust direction. In addition to the factors explained for the previous variable, the mean of the continuum distribution increases all the more due to higher values of $\cos\theta_{p_{c,n}^{\rm CMS},{\rm TB}}$. For jetlike events like continuum, the momentum of a particle in $B^0_{\rm Tag}$ is closely aligned with its thrust, which itself is strongly correlated with the $B^0_{\rm Rec}$ thrust as explained for the first variable.

The distributions for each of these discriminating variables are shown for simulated (MC) signal and continuum (data) events in Fig. 2. Before training, a loose requirement of $\cos\theta_{\rm TB,TO} < 0.9$ is placed that retains 90% of the signal while rejecting 50% of the continuum background. The Fisher discriminant is also required to satisfy $-3 < \mathcal{F}_{b\bar{b}/q\bar{q}} < 2$.

The next largest background comes from charm ($b \to c$) and charmless ($b \to u,\ d,\ s$) decays of the B meson and is found to exhibit peaking structure in the signal region due to the reconstruction of particular channels with a four-track final state. Defining the decay chain $B^0 \to a_1\pi_1$, $a_1 \to \rho\pi_2$, $\rho \to \pi_3\pi_4$, we apply vetoes to remove these peaking backgrounds as summarized in Table II.

The background coming from $b \to u\bar{u}d$ channels with the same final state as signal is studied separately. We consider a possible contribution from $B^0 \to a_2^{\pm}\pi^{\mp}$ by constructing the helicity variable $\mathcal{H}_{3\pi}$, defined as the cosine of the angle between the normal to the plane of the a_1 candidate and the flight direction of the bachelor pion from the B evaluated in the 3π rest frame.

On average, 1.6 B candidates are reconstructed per signal event. Selecting the best B candidate having the nearest $M_{\rm bc}$ with respect to the nominal B meson mass, the correct B is chosen in 91% of cases where the event contains multiple candidates. The fraction of misreconstructed signal events, in general, is 17%. As this procedure introduces bias to the $M_{\rm bc}$ distribution, this variable is excluded from the fit to extract the signal yield.

The full selection used to define the event sample for the CP asymmetry measurement is also applied for the branching fraction determination. Since the $B^0_{\rm Rec}$ and $B^0_{\rm Tag}$ mesons are approximately at rest in the $\Upsilon(4S)$ CMS, the difference in decay time between the two B mesons, Δt, can be approximately determined from the displacement in z between the final state decay vertices,

$$\Delta t \simeq \frac{(z_{\rm Rec} - z_{\rm Tag})}{\beta\gamma c} \equiv \frac{\Delta z}{\beta\gamma c}. \tag{6}$$

The vertex of reconstructed B candidates is determined from the charged daughters using the known IP. The IP profile is smeared in the plane perpendicular to z to account for the finite flight length of the B meson in this plane. To obtain the Δt distribution, we reconstruct the tag side vertex from the tracks not used to reconstruct $B^0_{\rm Rec}$ [25]. The events must satisfy the requirements $|\Delta t| < 70$ ps and $h_{\rm Rec,Tag} < 500$, where $h_{\rm Rec,Tag}$ is the multitrack vertex goodness-of-fit, calculated in three-dimensional space without using the interaction-region profile's constraint. To reduce the necessity of also modeling the event-dependent observables describing the variable Δt

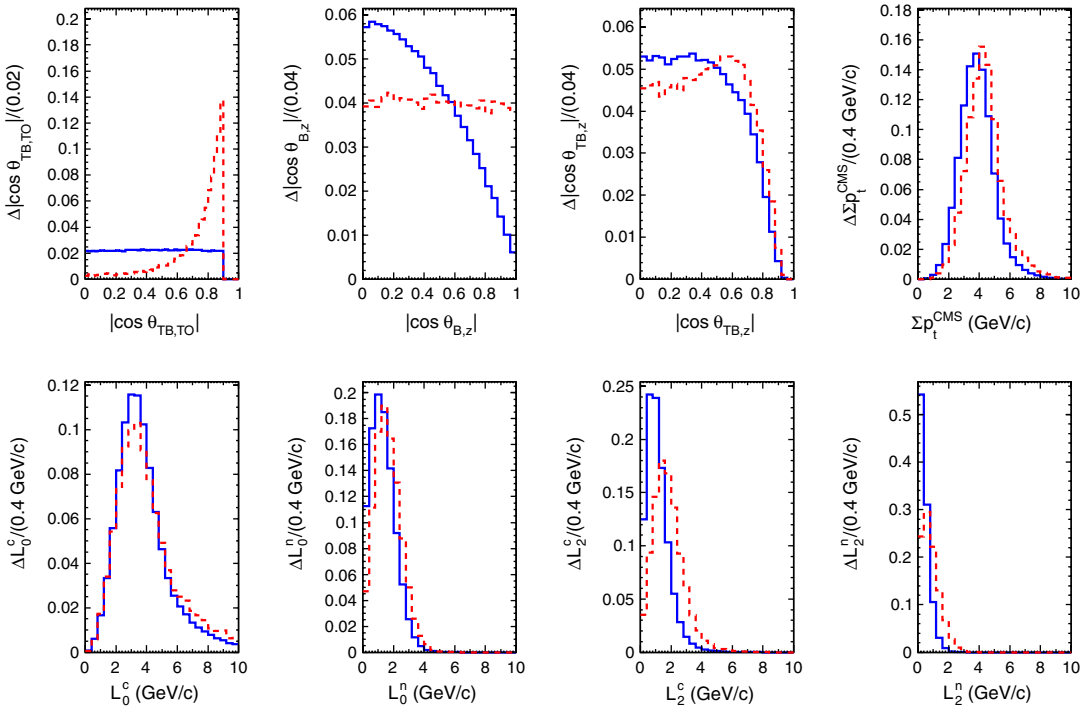

FIG. 2 (color online). Simulated (MC) and off-resonance data distributions for the quantities used to construct the Fisher discriminant $\mathcal{F}_{b\bar{b}/q\bar{q}}$, normalized to have the same area. The solid blue histograms show signal MC events, while the dashed red histograms show continuum events from data taken below the $\Upsilon(4S)$ resonance.

resolution in the fit, the vertex uncertainty is required to be $\sigma_z^{\mathrm{Rec,Tag}} < 200\ \mu$m for multitrack vertices and $\sigma_z^{\mathrm{Rec,Tag}} < 500\ \mu$m for single-track vertices.

After these selection criteria, the MC detection efficiencies are found to be

$$\begin{aligned} &\text{SVD1: } \epsilon(B^0 \to a_1^{\pm}\pi^{\mp}) = 0.1713 \pm 0.0004 \\ &\text{SVD2: } \epsilon(B^0 \to a_1^{\pm}\pi^{\mp}) = 0.2037 \pm 0.0005, \end{aligned} \tag{7}$$

where the uncertainty comes from limited MC statistics. Using independent data samples, we also determine correction factors to these efficiencies that account for the difference between data and MC. Correction factors in our reconstruction algorithm arise only from differences in particle identification and are determined from an inclusive $D^{*+} \to D^0[K^-\pi^+]\pi^+$ sample to be

$$\begin{aligned} &\text{SVD1: } \eta(B^0 \to a_1^{\pm}\pi^{\mp}) = 0.860 \pm 0.031 \\ &\text{SVD2: } \eta(B^0 \to a_1^{\pm}\pi^{\mp}) = 0.855 \pm 0.047. \end{aligned} \tag{8}$$

We employ the flavor-tagging routine described in Ref. [27]. The tagging information is represented by two parameters, the B^0_{Tag} flavor q and the purity r. The parameter r is an event-by-event, MC-determined, flavor-tagging parameter that ranges from $r = 0$ for no flavor discrimination to $r = 1$ for unambiguous flavor assignment. Due to a finite mistag probability w, the CP asymmetry is diluted by a factor $1 - 2w$. The measure of the performance of

TABLE II. Summary of peaking background vetoes. Alternate mass hypotheses have been applied to specific tracks where indicated in order to remove certain channels. The efficiency loss caused by these vetoes is also included. The X represents any charged track(s) that lead to a four-body final state.

Regions vetoed	Modes vetoed	Efficiency loss
$1.85\ \mathrm{GeV}/c^2 < m(\pi_2 K_3 \pi_4) < 1.89\ \mathrm{GeV}/c^2$ $1.85\ \mathrm{GeV}/c^2 < m(\pi_2 \pi_3 K_4) < 1.89\ \mathrm{GeV}/c^2$	$B \to D^+[K^-\pi^+\pi^+]X$	3.7%
$3.06\ \mathrm{GeV}/c^2 < m(\mu_1\mu_2) < 3.14\ \mathrm{GeV}/c^2$ $3.06\ \mathrm{GeV}/c^2 < m(\mu_1\mu_3) < 3.14\ \mathrm{GeV}/c^2$ $3.06\ \mathrm{GeV}/c^2 < m(\mu_1\mu_4) < 3.14\ \mathrm{GeV}/c^2$	$B \to J/\psi[\mu^+\mu^-]X$	6.1%
$0.480\ \mathrm{GeV}/c^2 < m(\pi_2\pi_3) < 0.516\ \mathrm{GeV}/c^2$ $0.480\ \mathrm{GeV}/c^2 < m(\pi_2\pi_4) < 0.516\ \mathrm{GeV}/c^2$	$B \to K^0_S[\pi^+\pi^-]X$	11.9%

the flavor-tagging algorithm is the total effective tagging efficiency $\epsilon_{\rm eff} = \epsilon_{\rm Tag}(1-2w)^2$, as the statistical significance of the CP parameters is proportional to $(1-2w)\times\sqrt{\epsilon_{\rm Tag}}$, where $\epsilon_{\rm Tag}$ is the raw tagging efficiency. These are determined to be $\epsilon_{\rm eff} = 0.284 \pm 0.010$ and $\epsilon_{\rm eff} = 0.301 \pm 0.004$ for SVD1 and SVD2 data, respectively.

B. Event model

The branching fraction is extracted from a four-dimensional extended unbinned maximum likelihood fit to ΔE, $\mathcal{F}_{b\bar{b}/q\bar{q}}$, $m_{3\pi}$ and $\mathcal{H}_{3\pi}$ from a data sample divided into seven bins ($l = 0\ldots6$) in the flavor-tag quality r and two SVD configurations s. We consider 12 categories in the event model: correctly reconstructed signal (referred to as truth signal hereafter), misreconstructed signal, continuum, charm neutral and charged $B\bar{B}$ decays, charmless neutral and charged $B\bar{B}$ decays, and five peaking backgrounds. In most categories, the linear correlations between fit variables are small, so the probability density function (PDF) for each category j, is taken as the product of individual PDFs for each variable $\mathcal{P}_j^{l,s}(\Delta E, \mathcal{F}_{b\bar{b}/q\bar{q}}, m_{3\pi}, \mathcal{H}_{3\pi}) = \mathcal{P}^{l,s}(\Delta E)\times\mathcal{P}^{l,s}(\mathcal{F}_{b\bar{b}/q\bar{q}})\times\mathcal{P}^{l,s}(m_{3\pi})\times\mathcal{P}^{l,s}(\mathcal{H}_{3\pi})$ in each l, s bin, unless otherwise mentioned.

The truth model shape is determined from correctly reconstructed signal MC events. The PDF for ΔE is taken to be the sum of two asymmetric-width (bifurcated) Gaussians incorporating calibration factors that correct for the difference between data and MC. These factors calibrate the mean and width of the core bifurcated Gaussian and are determined from a large-statistics control sample $B^0 \to D^-[K^+\pi^-\pi^-]\pi^+$. The PDF for $\mathcal{F}_{b\bar{b}/q\bar{q}}$, here and throughout this analysis, for all 12 categories is the sum of two bifurcated Gaussians in each flavor-tag bin l. The shape for all $B\bar{B}$ categories is fixed from the truth model except for the mean of the core distribution, incorporating calibration factors that correct for the shape difference between data and MC. The $m_{3\pi}$ distribution is modeled with an efficiency-corrected relativistic Breit-Wigner

$$\mathcal{P}_{\rm Sig}^{l,s}(m_{3\pi}) \equiv \epsilon^s(m_{3\pi})\frac{m_{a_1}\Gamma(m_{3\pi})}{(m_{3\pi}^2 - m_{a_1}^2)^2 + m_{a_1}^2\Gamma^2(m_{3\pi})}, \quad (9)$$

where ϵ^s is the mass-dependent detection efficiency and Γ is the mass-dependent width

$$\Gamma(m) = \Gamma_{a_1}\frac{\rho^{1+S}(m_{3\pi})}{\rho^{1+S}(m_{a_1})}. \quad (10)$$

Due to the finite width of the ρ^0, the phase space ρ^{1+S} of the $a_1^\pm$ decay, where the superscript represents the decay into a spin-1 meson in an S-wave configuration [28], cannot be calculated simply. Therefore, we model the phase space empirically with a sixth-order Chebyshev polynomial

$$\rho^{1+S}(m_{3\pi}) \equiv 1 + \sum_{i=1}^{6} c_i C_i(m_{3\pi}), \quad (11)$$

where C_i is a Chebyshev polynomial of order i, and c_i is the fit coefficient. The PDF for $\mathcal{H}_{3\pi}$ is a sum of symmetric Chebyshev polynomials up to sixth order.

The misreconstructed model shape is determined from incorrectly reconstructed signal MC events. For ΔE, the PDF is taken as a smoothed one-dimensional histogram, the $m_{3\pi}$ PDF is the sum of two asymmetric-width Gaussians and the PDF for $\mathcal{H}_{3\pi}$ is the sum of symmetric Chebyshev polynomials up to eighth order.

The parameterization of the continuum model is chosen based on data taken below the $\Upsilon(4S)$ resonance; however, all the shape parameters are free in the fit to extract the branching fraction. Since continuum is by far the dominant component, extra care must be taken to ensure this background shape is understood, so correlations above 2% are considered. The PDF for ΔE is taken to be a first-order Chebyshev polynomial in each flavor-tag bin with a coefficient depending linearly on $\mathcal{F}_{b\bar{b}/q\bar{q}}$,

$$\mathcal{P}_{q\bar{q}}^{l,s}(\Delta E|\mathcal{F}_{b\bar{b}/q\bar{q}}) = 1 + (p_0^{l,s} + p_1^s\mathcal{F}_{b\bar{b}/q\bar{q}})\Delta E. \quad (12)$$

The $m_{3\pi}$ shape is observed to shift quadratically with $\mathcal{F}_{b\bar{b}/q\bar{q}}$, so the PDF is a sum of Chebyshev polynomial up to fourth order that incorporates an offset. In addition, a small excess is seen above this distribution and is modeled with a Gaussian. The result is

$$\begin{aligned}\mathcal{P}_{q\bar{q}}^{l,s}&(m_{3\pi}|\mathcal{F}_{b\bar{b}/q\bar{q}}) \\ &\equiv (1-f^s)\left[1 + \sum_{k=1}^{4} c_k^s C_k(m_{3\pi} - p_1^s\mathcal{F}_{b\bar{b}/q\bar{q}} - p_2^s\mathcal{F}_{b\bar{b}/q\bar{q}}^2)\right] \\ &\quad + f^s G(m_{3\pi};\mu^s,\sigma^s). \end{aligned} \quad (13)$$

The PDF for $\mathcal{H}_{3\pi}$ is the sum of symmetric Chebyshev polynomials p to sixth order and a mirrored Gaussian,

$$\begin{aligned}\mathcal{P}_{q\bar{q}}^{l,s}(\mathcal{H}_{3\pi}) \equiv (1-2f^s)&\left[1 + \sum_{k=1}^{6} c_k^s C_k(\mathcal{H}_{3\pi})\right] \\ &+ f^s G(\mathcal{H}_{3\pi};\mu^s,\sigma^s) + f^s G(\mathcal{H}_{3\pi};-\mu^s,\sigma^s). \end{aligned} \quad (14)$$

The charm $B\bar{B}$ background shape is determined from a large sample of MC containing generic $b \to c$ transitions and is subdivided into neutral and charged B samples. For ΔE, the PDF is a smoothed one-dimensional histogram; for $m_{3\pi}$, the PDF is the sum of Chebyshev polynomials up to fourth order; for $\mathcal{H}_{3\pi}$, the PDF is the sum of symmetric Chebyshev polynomials up to sixth order plus a mirrored Gaussian.

The charmless $B\bar{B}$ background shape is determined from a large sample of MC containing generic $b \to u$, d, s transitions and is subdivided into neutral and charged B samples. A sizeable correlation is seen between ΔE and

TABLE III. List of peaking backgrounds, assumed branching fractions and their respective expected yields after all selection criteria has been applied.

Mode	$\mathcal{B}$ ($\times 10^{-6}$)	SVD1 expected events	SVD2 expected events
$B^0 \to \rho^0\rho^0$	0.73 ± 0.28	2	11
$B^0 \to b_1^{\pm}[\pi^{\pm}\pi^{\mp}\pi^{\pm}]\pi^{\mp}$	0.16 ± 0.03	5	23
$B^0 \to \rho^0\pi^+\pi^-$	4.35 ± 4.35	28	137
$B^0 \to \pi^+\pi^-\pi^+\pi^-$	9.65 ± 9.65	26	125
$B^0 \to f_0\rho^0$	<0.3	<2	<7
$B^0 \to f_0 f_0$	<0.1	<1	<1
$B^0 \to f_0\pi^+\pi^-$	<3.8	<14	<65

$m_{3\pi}$ and is taken into account with a smoothed two-dimensional histogram. The $\mathcal{H}_{3\pi}$ PDF is the sum of symmetric Chebyshev polynomials up to sixth order with a mirrored Gaussian.

Many decay channels contain the same final state as $B^0 \to a_1^{\pm}\pi^{\mp}$. In addition to the possibility of $B^0 \to a_2^{\pm}\pi^{\mp}$, we also consider those listed in Table III which includes their expected yields in 772×10^6 $B\bar{B}$ pairs. With the exception of $B^0 \to a_2^{\pm}\pi^{\mp}$, the assumed peaking background branching fractions are fixed in the nominal fit from the Heavy Flavor Averaging Group [29]. Where a mode is to be included in the fit model but only an upper limit is known, the branching fraction is taken as half the upper limit unless the mode contains an f_0, in which case it is assigned zero branching fraction and instead considered solely in the systematic uncertainties. The peaking background shapes are determined from individually generated MC samples and are fixed in the fit to data. The PDF for ΔE borrows the shape of correctly reconstructed signal events and includes a first-order Chebyshev polynomial to model the misreconstructed contribution underneath. The $m_{3\pi}$ PDF depends on the peaking background. For $B^0 \to a_2^{\pm}\pi^{\mp}$, a sum of three Gaussians is used; for $B^0 \to \rho^0\rho^0$ and $b_1^{\pm}\pi^{\mp}$, a smoothed histogram is used; otherwise, the sum of Chebyshev polynomials up to fourth order is used. The $\mathcal{H}_{3\pi}$ PDF is the sum of symmetric Chebyshev polynomials up to eighth order, except for $B^0 \to \rho^0\rho^0$, which is modeled with a smoothed symmetrized one-dimensional histogram.

The total likelihood for 208238 $B^0 \to a_1^{\pm}\pi^{\mp}$ candidates in the fit region is

$$\mathcal{L} \equiv \prod_{l,s} \frac{e^{-\sum_j N_j^s \sum_{l,s} f_j^{l,s}}}{N_{l,s}!} \prod_{i=1}^{N_{l,s}} \sum_j N_j^s f_j^{l,s} \mathcal{P}_j^{l,s} \times (\Delta E^i, \mathcal{F}^i_{b\bar{b}/q\bar{q}}, m^i_{3\pi}, \mathcal{H}^i_{3\pi}), \quad (15)$$

which iterates over i events, j categories, l flavor-tag bins and s detector configurations. Instead of two free signal yields N^s_{Sig}, the branching fraction is chosen as a single free parameter and is incorporated into the fit with

$$N^s_{\mathrm{sig}} = \mathcal{B}(B^0 \to a_1^{\pm}\pi^{\mp}) \times \mathcal{B}(a_1^{\pm} \to \pi^{\pm}\pi^{\mp}\pi^{\pm}) N^s_{B\bar{B}} \epsilon^s_{\mathrm{Sig}} \eta^s_{\mathrm{Sig}}, \quad (16)$$

where $\epsilon^s_{\mathrm{Sig}}$ and η^s_{Sig} are given in Eqs. (7) and (8), respectively. Similar conversions are done for the remaining peaking backgrounds using their expected values from Table III. The fraction of events in each flavor-tag bin l, for category j, is denoted by $f_j^{l,s}$. The fraction of signal events in each l, s bin, $f^{l,s}_{\mathrm{Sig}}$, has been calibrated with the $B^0 \to D^-[K^+\pi^-\pi^-]\pi^+$ control sample. Other free physics parameters include the $a_1^{\pm}$ width and the product-branching fraction $\mathcal{B}(B^0 \to a_2^{\pm}\pi^{\mp}) \times \mathcal{B}(a_2^{\pm} \to \pi^{\pm}\pi^{\mp}\pi^{\pm})$. Also free are the yields $N^s_{q\bar{q}}$, $N^{\mathrm{charm};s}_{B^0\bar{B}^0}$ and $N^{\mathrm{charmless};s}_{B^0\bar{B}^0}$; the remaining yields are fixed to the values given in Table IV as determined from MC. In total, there are 121 free parameters in the fit. In addition, all shape parameters of the continuum model are free in the fit to data.

TABLE IV. Summary of yields fixed relative to other yields free in the fit where the uncertainties are from limited MC statistics. The misreconstructed yield is fixed relative to the signal yield, and the charm and charmless B^+B^- background yields are fixed relative to their respective $B^0\bar{B}^0$ background yields.

Yield	SVD1	SVD2
N^s_{Mis}	$(0.167 \pm 0.001)N^{\mathrm{SVD1}}_{\mathrm{Sig}}$	$(0.166 \pm 0.001)N^{\mathrm{SVD2}}_{\mathrm{Sig}}$
$N^{\mathrm{charm};s}_{B^+B^-}$	$(1.585 \pm 0.019)N^{\mathrm{charm;SVD1}}_{B^0\bar{B}^0}$	$(1.700 \pm 0.010)N^{\mathrm{charm;SVD2}}_{B^0\bar{B}^0}$
$N^{\mathrm{charmless};s}_{B^+B^-}$	$(0.568 \pm 0.006)N^{\mathrm{charmless;SVD1}}_{B^0\bar{B}^0}$	$(0.574 \pm 0.003)N^{\mathrm{charmless;SVD2}}_{B^0\bar{B}^0}$

C. Fit result

We perform a fit to the data, with the projections shown in Fig. 3, and obtain the product branching fractions

$$\mathcal{B}(B^0 \to a_1^{\pm}(1260)\pi^{\mp}) \times \mathcal{B}(a_1^{\pm}(1260) \to \pi^{\pm}\pi^{\mp}\pi^{\pm}) = (11.1 \pm 1.0(\text{stat}) \pm 1.4(\text{syst})) \times 10^{-6}, \quad (17)$$

corresponding to a yield of 1445 ± 216 events, and

$$\mathcal{B}(B^0 \to a_2^{\pm}(1320)\pi^{\mp}) \times \mathcal{B}(a_2^{\pm}(1320) \to \pi^{\pm}\pi^{\mp}\pi^{\pm}) = (1.5 \pm 0.4(\text{stat}) \pm 0.4(\text{syst})) \times 10^{-6}, \quad (18)$$

corresponding to a yield of 282 ± 106 events. The statistical correlation coefficient between these two measurements is -0.41, and the statistical significance of the $a_1^{\pm}$ peak is 16σ, estimated by comparing the likelihood of the nominal fit result with that of a fit where the $B^0 \to a_1^{\pm}\pi^{\mp}$ branching fraction is fixed to zero. We also measure the $a_1^{\pm}$ width to be $\Gamma_{a_1} = 381 \pm 43(\text{stat})$ MeV. From this fit, the relative contributions of each component are 0.8 signal, 95.6 continuum, 3.3 $B\bar{B}$ background and 0.3% peaking background.

Our measurement of the $B^0 \to a_1^{\pm}\pi^{\mp}$ branching fraction is lower than that measured by the *BABAR* Collaboration [17] by 1.9σ though still in marginal agreement. We are also in agreement with predictions made in the QCD factorization framework given in Refs. [11,12], but not Ref. [13]. The upper limit of the branching fraction of $B^0 \to a_2^{\pm}\pi^{\mp}$ is also determined to be

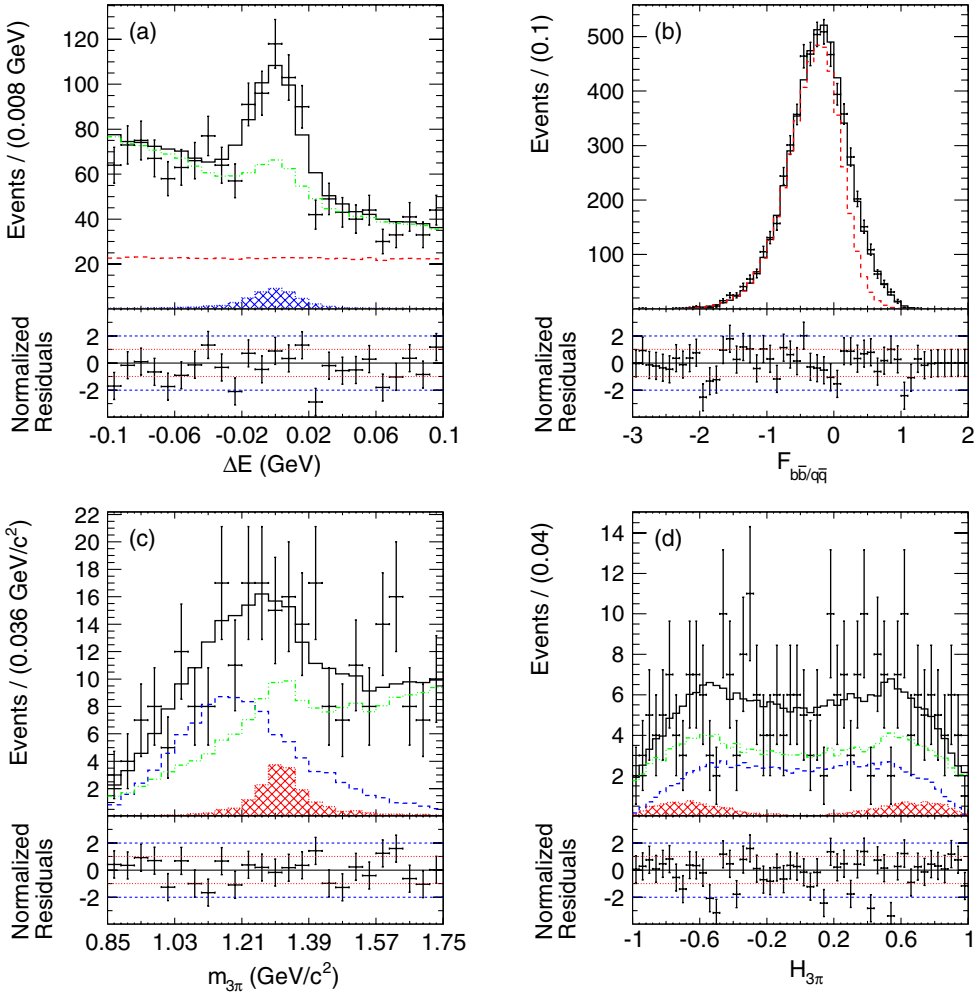

FIG. 3 (color online). Projections of the fit to the $B^0 \to a_1^{\pm}\pi^{\mp}$ data. The points with error bars represent the data and the solid black histogram represents the fit result. (a) shows the ΔE projection for $\mathcal{F}_{b\bar{b}/q\bar{q}} > 0.5$ and $r > 0.5$. The blue hatched histogram shows the peaking background, the dashed red histogram shows the continuum contribution and the dash-dotted green histogram shows the total background. (b) shows the $\mathcal{F}_{b\bar{b}/q\bar{q}}$ projection for $|\Delta E| < 0.01$ GeV and $r > 0.5$. The dashed red histogram shows the continuum background contribution. (c) and (d) show the $m_{3\pi}$ and $\mathcal{H}_{3\pi}$ projections, respectively, for $\mathcal{F}_{b\bar{b}/q\bar{q}} > 0.5$, $|\Delta E| < 0.01$ GeV and $r > 0.5$. The dashed blue histogram shows the $a_1^{\pm}$ contribution, the red hatched histogram shows the $a_2^{\pm}$ contribution and the dash-dotted green histogram shows the total background.

$$\mathcal{B}(B^0 \to a_2^{\pm}(1320)\pi^{\mp}) \times \mathcal{B}(a_2^{\pm}(1320) \to \pi^{\pm}\pi^{\mp}\pi^{\pm}) < 2.2 \times 10^{-6} \quad \text{at } 90\%\,\text{CL}, \tag{19}$$

which is improved over the current world average [24] by about two orders of magnitude.

D. Systematic uncertainties

Systematic errors from various sources are considered and estimated with independent internal studies and cross-checks. These are summarized in Table V. This includes the uncertainty on the number of produced $B\bar{B}$ events in the data sample. Contributions to the uncertainty in the selection efficiency due to particle identification and tracking are calculated by independent studies at Belle. The systematic uncertainty arising from the assumption that the $a_1^{\pm}$ decays exclusively through the dominant $\rho^0\pi^{\pm}$ intermediate state is accounted for by recalculating the detection efficiency with an exclusive $f_0(600)\pi^{\pm}$ MC.

The uncertainty in the $a_1^{\pm}$ shape is determined by varying the fixed mass within its world average uncertainty [24]. We account for a difference in the fraction of misreconstructed signal events between data and MC by varying this parameter by $\pm 5\%$ and repeating the fit. Variations in the parametric model shape due to limited statistics are accounted for by varying each parameter within their errors. The dominant contribution to this category comes from the uncertainties in the signal shape correction factors obtained from analyzing a high-statistics control sample $B^0 \to D^-[K^+\pi^-\pi^-]\pi^+$. Uncertainties in the nonparametric shapes are obtained by varying the contents of the histogram bins within $\pm 1\sigma$. The systematic uncertainty due to fixing the peaking background yields is estimated by varying the branching fraction by its world average error [24] and repeating the fit. For modes where only an upper limit is known, the variation is taken as half of the upper limit. The fit bias is determined from pseudoexperiments by searching for a difference between the generated and fitted physics parameters. As channels containing an f_0 are ignored in the nominal model, we account for a possible effect on the signal yield by embedding such events into these pseudoexperiments and determining further bias on the $a_1\pi$ and $a_2\pi$ branching fractions. Finally, the uncertainty from neglecting interference between a_1 and a_2 was estimated by constructing a four-body amplitude including detector effects and generating three relative interference configurations between a_1 and a_2: maximum constructive interference, no interference and maximal destructive interference. The largest deviation from the sample with no interference when fitting all with the nominal model gives the systematic uncertainty from interference.

TABLE V. Systematic uncertainties of branching fractions.

Category	$\delta\mathcal{B}(B^0 \to a_1^{\pm}\pi^{\mp})$ (%)	$\delta\mathcal{B}(B^0 \to a_2^{\pm}\pi^{\mp})$ (%)
$N(B\bar{B})$	1.4	1.4
Tracking	2.1	2.1
Particle identification	4.5	4.5
$f_0(600)$	3.0	N/A
a_1 shape	5.1	19.4
Misreconstruction fraction	2.1	3.3
Model shape	4.2	14.0
Histogram shape	4.8	4.6
Peaking background	5.1	11.4
Fit bias	0.4	4.1
Interference	4.1	4.1
Total	12.2	28.1

IV. TIME-DEPENDENT MEASUREMENT

A. Event selection

In addition to the event selection criteria for the branching fraction measurement, events are selected for the time-dependent CP-violation measurement if they satisfy $|\Delta E| < 0.04$ GeV. This requirement retains 97% of signal and 90% of peaking backgrounds, while rejecting 60% of the continuum background and 63% of the $B\bar{B}$ background. After this selection, the relative contributions of each component are 1.9 signal, 94.2 continuum, 3.1 $B\bar{B}$ background and 0.8% peaking background.

B. Event model

The signal PDF is given by

$$\mathcal{P}_{\text{Sig}}^{l,s}(\Delta t, q, c) \equiv (1 + c\mathcal{A}_{CP})\frac{e^{-|\Delta t|/\tau_{B^0}}}{8\tau_{B^0}} \times \{1 - q\Delta w^{l,s} + q(1 - 2w^{l,s}) \times [(\mathcal{S}_{CP} + c\Delta\mathcal{S})\sin\Delta m_d\Delta t - (\mathcal{C}_{CP} + c\Delta\mathcal{C})\cos\Delta m_d\Delta t]\} \otimes R^s_{B^0\bar{B}^0}(\Delta t), \tag{20}$$

which accounts for CP dilution from the probability of incorrect flavor tagging $w^{l,s}$ and the wrong tag difference $\Delta w^{l,s}$ between B^0 and $\bar{B}^0$, both of which are determined from flavor- specific control samples [27]. This PDF is convolved with the Δt resolution function for neutral B particles $R^s_{B^0\bar{B}^0}$, described in Ref. [25].

The reconstructed vertex position of misreconstructed events is dominated by the high- momentum bachelor pion from $B^0 \to a_1^{\pm}\pi^{\mp}$. This pion is rarely misreconstructed, and the effect of borrowing a lower momentum track from the tag side results in a slightly smaller lifetime. Thus, misreconstructed events are modeled using the truth model PDF with an effective lifetime and share CP parameters with the truth model. The effective lifetime is determined from misreconstructed signal MC events, and the effect of this choice is accounted for in the systematic uncertainties.

The continuum shape is determined from data taken below the $\Upsilon(4S)$ resonance with the model

$$P^{l,s}_{q\bar{q}}(\Delta t, q, c) \equiv \frac{1 + qc\Delta\mathcal{C}_{q\bar{q}}}{4}\Big[(1 - f_\delta)\frac{e^{-|\Delta t|/\tau_{q\bar{q}}}}{2\tau_{q\bar{q}}} + f_\delta\delta(\Delta t - \mu^s_\delta)\Big] \otimes R^s_{q\bar{q}}(\Delta t). \quad (21)$$

This model contains a lifetime and prompt component to account for the charm and charmless contributions, respectively, and is convolved with a sum of two Gaussians

$$R^s_{q\bar{q}}(\Delta t) \equiv (1 - f^s_{\rm tail})G(\Delta t; \mu^s_{\rm mean}, S^s_{\rm main}\sigma) + f^s_{\rm tail}G(\Delta t; \mu^s_{\rm mean}, S^s_{\rm main}S^s_{\rm tail}), \quad (22)$$

which uses the event-dependent Δt error constructed from the vertex resolution $\sigma \equiv (\sqrt{\sigma^2_{\rm Rec} + \sigma^2_{\rm Tag}})/\beta\gamma c$ as a scale factor. We also account for an asymmetry in the product qc, with the parameter $\Delta\mathcal{C}_{q\bar{q}}$, which is due to the jetlike topology of continuum. As a high momentum π^+ (π^-) in $B^0_{\rm Rec}$ is correlated with a high momentum π^- (π^+) on the tag side, q and c will more often have the same sign.

The $B\bar{B}$ shape is determined from a large MC sample and is divided into generic charmed and charmless, each further divided into neutral and charged $B\bar{B}$ decays, using a lifetime model

$$\mathcal{P}^{l,s}_{B\bar{B}}(\Delta t, q, c) \equiv \frac{1 + cq\Delta\mathcal{C}_{B\bar{B}}}{4}\frac{e^{-|\Delta t|/\tau_{B\bar{B}}}}{2\tau_{B\bar{B}}} \otimes R^s_{B\bar{B}}(\Delta t), \quad (23)$$

where $R_{B\bar{B}}$ is the relevant Δt resolution function for either neutral or charged B particles. Since reconstructed $B\bar{B}$ events may borrow a particle from the tag side, the average Δt lifetime tends to be smaller and is taken into account with the effective lifetime, $\tau_{B\bar{B}}$. Like continuum, $B\bar{B}$ events also exhibit a qc asymmetry; however, it is found to be more complex and is modeled with a first-order polynomial in Δt

$$\Delta\mathcal{C}_{B\bar{B}} \equiv p_0 + p_1|\Delta t|. \quad (24)$$

The shapes of the peaking backgrounds are determined from individually generated MC events. As these backgrounds may exhibit CP violation, we use a model similar to the truth model, but with an effective lifetime. We also fix all time-dependent parameters to null with the effects of this choice reflected in the systematic uncertainties.

To account for the broad underlying Δt events not yet described by either signal or background PDFs, a broad Gaussian outlier PDF is introduced

$$\mathcal{P}^{l,s}_{\rm Out}(\Delta t, q, c) \equiv \frac{1}{4}G(\Delta t; 0, \sigma^s_{\rm Out}). \quad (25)$$

The total likelihood for 83799 $B^0 \to a_1^\pm\pi^\mp$ candidates in the fit region becomes

$$\mathcal{L} \equiv \prod_{l,s}\prod_{i=1}^{N_{l,s}}\sum_j f^{l,s}_j(\Delta E^i, \mathcal{F}^i_{b\bar{b}/q\bar{q}}, m^i_{3\pi}, \mathcal{H}^i_{3\pi})\mathcal{P}^{l,s}_j(\Delta t^i, q^i, c^i), \quad (26)$$

where $f^{l,s}_j$ is the event-dependent probability of component j, in flavor-tag bin l, with detector configuration s

$$f^{l,s}_j(\Delta E^i, \mathcal{F}^i_{b\bar{b}/q\bar{q}}, m^i_{3\pi}, \mathcal{H}^i_{3\pi}) = \frac{N^s_j f^{l,s}_j \mathcal{P}^{l,s}_j(\Delta E^i, \mathcal{F}^i_{b\bar{b}/q\bar{q}}, m^i_{3\pi}, \mathcal{H}^i_{3\pi})}{\sum_j N^s_j f^{l,s}_j \mathcal{P}^{l,s}_j(\Delta E^i, \mathcal{F}^i_{b\bar{b}/q\bar{q}}, m^i_{3\pi}, \mathcal{H}^i_{3\pi})}, \quad (27)$$

constructed from the branching fraction measurement. Only the five time-dependent coefficients of the truth model are free in the fit to data.

As a consistency check, we perform a fit to data to measure the B^0 lifetime while fixing the five parameters of the truth model to zero. We obtain $\tau_{B^0} = 1.389 \pm 0.085$ ps, which is in agreement with the current world average $\tau_{B^0} = 1.519 \pm 0.007$ ps [24].

C. Fit result

We perform a fit to the data and obtain the CP-violating parameters

$$\begin{aligned}\mathcal{A}_{CP} &= -0.06 \pm 0.05(\text{stat}) \pm 0.07(\text{syst}),\\ \mathcal{C}_{CP} &= -0.01 \pm 0.11(\text{stat}) \pm 0.09(\text{syst}),\\ \mathcal{S}_{CP} &= -0.51 \pm 0.14(\text{stat}) \pm 0.08(\text{syst}),\end{aligned} \quad (28)$$

and the CP conserving parameters

$$\begin{aligned}\Delta\mathcal{C} &= +0.54 \pm 0.11(\text{stat}) \pm 0.07(\text{syst}),\\ \Delta\mathcal{S} &= -0.09 \pm 0.14(\text{stat}) \pm 0.06(\text{syst}).\end{aligned} \quad (29)$$

The background subtracted fit results are shown in Fig. 4, where the data points are the signal yields obtained by repeating the branching fraction fits in Δt, q or Δt, qc bins, accounting for the selection criteria on ΔE. Ours are the most precise measurements of these parameters to date and are in agreement with those obtained by the *BABAR* Collaboration [18]. We are also in very good agreement with the theoretical predictions made in Ref. [11]. The statistical correlation coefficients between the obtained parameters are given in Table VI.

A MC technique is employed to obtain $\phi_2^{\rm eff}$ using Eq. (4) in order to take correlations between the fitted parameters into account. We generate multiple vectors $(\mathcal{C}_{CP}, \Delta\mathcal{C}, \mathcal{S}_{CP}, \Delta\mathcal{S})$ based on a correlated multidimensional Gaussian constructed from the fit result and solve for the four solutions of $\phi_2^{\rm eff}$ each time. We take the central values and their uncertainties from the resulting distributions of $\phi_2^{\rm eff}$ and obtain the four solutions

$$\begin{aligned}\phi_2^{\rm eff} &= (-17.3 \pm 6.6(\text{stat}) \pm 4.8(\text{syst}))^\circ,\\ \phi_2^{\rm eff} &= (41.6 \pm 6.2(\text{stat}) \pm 3.4(\text{syst}))^\circ,\\ \phi_2^{\rm eff} &= (48.4 \pm 6.2(\text{stat}) \pm 3.4(\text{syst}))^\circ,\\ \phi_2^{\rm eff} &= (107.3 \pm 6.6(\text{stat}) \pm 4.8(\text{syst}))^\circ.\end{aligned} \quad (30)$$

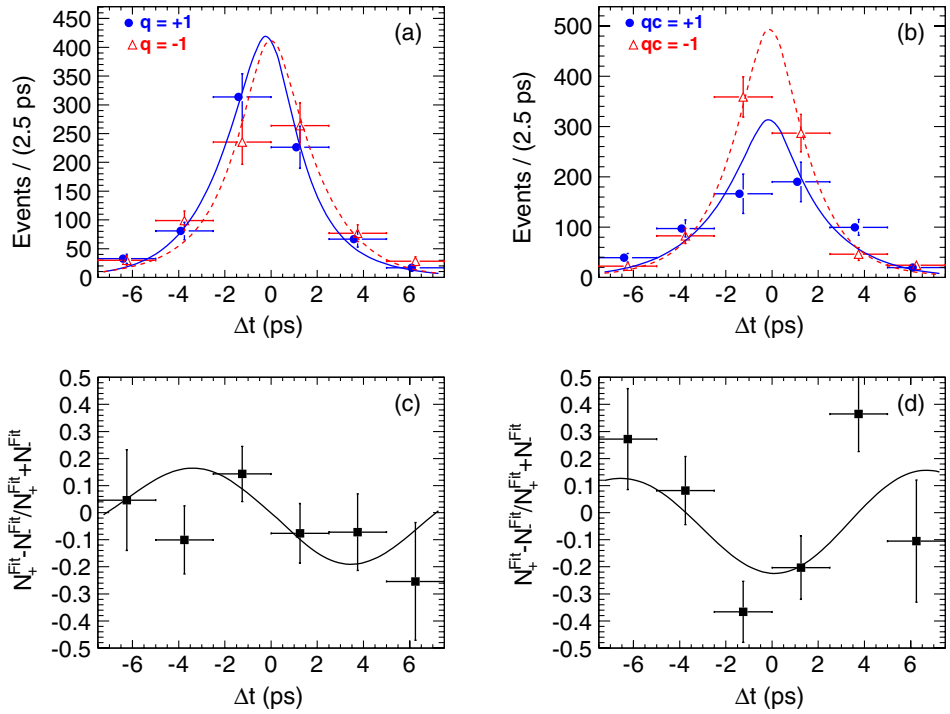

FIG. 4 (color online). Background subtracted time-dependent fit results for $B^0 \to a_1^{\pm}\pi^{\mp}$. (a) and (b) show the Δt distributions for the $B^0_{\rm Tag}$ flavor q and the product of the $B^0_{\rm Tag}$ flavor and a_1 charge qc, respectively. The left plots are useful for visualizing the effect of flavor-dependent CP violation, while the plots on the right show the effects of the CP-conserving parameters. The solid blue and dashed red curves represent the Δt distributions for positive and negative quantity, respectively. (c) and (d) show the asymmetry of the plots immediately above them, $(N_+^{\rm Fit} - N_-^{\rm Fit})/(N_+^{\rm Fit} + N_-^{\rm Fit})$, where $N_+^{\rm Fit}$ ($N_-^{\rm Fit}$) is the measured signal yield of positive (negative) quantities in bins of Δt.

Using a similar technique, we obtain the direct CP-violation parameters given in Eq. (5)

$$\begin{aligned} A_{+-} &= +0.07 \pm 0.08(\text{stat}) \pm 0.10(\text{syst}), \\ A_{-+} &= -0.04 \pm 0.26(\text{stat}) \pm 0.19(\text{syst}), \end{aligned} \tag{31}$$

where the statistical correlation coefficient between these two parameters is 0.61.

We compose Gaussian distributions from the four solutions for $\phi_2^{\rm eff}$ and construct a two-sided p-value plot for $\phi_2^{\rm eff}$ as shown in Fig. 5. We also perform a likelihood scan to estimate the significance of $\mathcal{S}_{CP}$ and $\Delta\mathcal{C}$ as shown in Fig. 6. The significance of mixing-induced CP violation is found to be 3.1σ including systematic uncertainties, while the rate where the $a_1^{\pm}$ does not contain the spectator quark is found to dominate the rate where it does at the 4.1σ level.

TABLE VI. Statistical correlation matrix for the fit result.

	$\mathcal{A}_{CP}$	$\mathcal{C}_{CP}$	$\Delta\mathcal{C}$	$\mathcal{S}_{CP}$	$\Delta\mathcal{S}$
$\mathcal{A}_{CP}$	1				
$\mathcal{C}_{CP}$	−0.20	1			
$\Delta\mathcal{C}$	+0.01	+0.03	1		
$\mathcal{S}_{CP}$	+0.02	−0.01	−0.03	1	
$\Delta\mathcal{S}$	+0.09	−0.05	−0.01	+0.02	1

D. Systematic uncertainties

The systematic uncertainties for the time-dependent parameters are summarized in Table VII. These are estimated from various sources including uncertainties in the IP profile, charged track selection based on track helix errors, helix parameter corrections, Δt and vertex

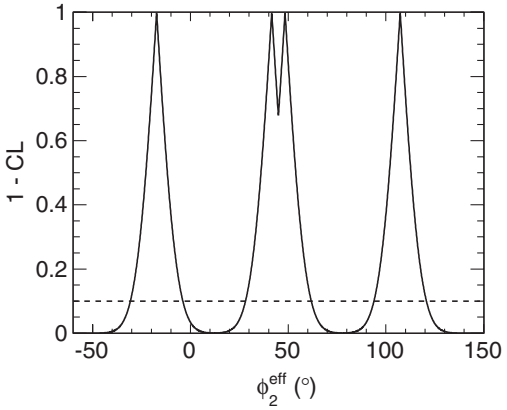

FIG. 5. Difference 1-CL plotted for a range of $\phi_2^{\rm eff}$ values as shown by the solid curve. The p value is taken as the maximum p value that can be calculated from all four solutions. The dashed line indicates CL = 90%.

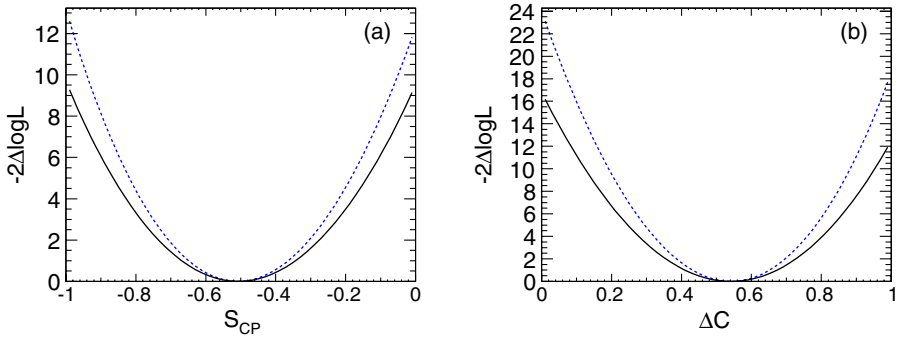

FIG. 6 (color online). Likelihood scan of $\mathcal{S}_{CP}$ (a) and $\Delta\mathcal{C}$ (b). The dotted blue curve shows the statistical likelihood, while the solid black curve includes the systematic uncertainty.

goodness-of-fit selection, Δz bias and SVD misalignment. The fixed physics parameters τ_{B^0} and Δm_d, Δt resolution function and data model shape parameters including background effective lifetimes and qc asymmetries, as well as the flavor-tagging performance parameters w and Δw, are varied by $\pm 1\sigma$. We generate MC pseudoexperiments and perform an ensemble test to obtain systematic biases from interference on the tag side arising between the CKM-favored $b\bar{d} \to (c\bar{u}d)\bar{d}$ and doubly CKM-suppressed $\bar{b}d \to (\bar{u}c\bar{d})d$ amplitudes in final states used for flavor tagging [30]. These sources should not affect $\mathcal{A}_{CP}$, as this parameter represents a time- and flavor-integrated asymmetry.

The remaining systematic categories affect all time-dependent parameters. The parameters and nonparametric shapes describing signal probability are varied in the same way as for the branching fraction measurement. The CP violation parameters of misreconstructed events are assumed to be the same as signal. To account for this, a sample of GEANT MC is produced with the nominal fit result. The systematic error is taken as the difference between the fit result to the correctly reconstructed subsample and a fit to the whole sample sharing the CP parameters between the signal and misreconstructed components. The fit bias is determined from an ensemble test by searching for a difference between the generated and fitted physics parameters.

Possible CP violation in the background is the dominant systematic uncertainty. We assume that the neutral $B\bar{B}$ background possesses a 20% CP-violating component and that peaking backgrounds display a 50% CP-violating effect. For the neutral $B\bar{B}$ component, the uncertainty is determined by fixing the CP parameters and refitting the data. For the peaking backgrounds, special GEANT MC samples are produced with the nominal fit result for signal and CP violation generated in the peaking background. We perform a fit with the peaking background

TABLE VII. Systematic uncertainties of time-dependent parameters. Categories related to Δt reconstruction and flavor tagging are not applicable to the time- and flavor-integrated $\mathcal{A}_{CP}$.

Category	$\delta\mathcal{A}_{CP}$ (10^{-2})	$\delta\mathcal{C}_{CP}$ (10^{-2})	$\delta\Delta\mathcal{C}$ (10^{-2})	$\delta\mathcal{S}_{CP}$ (10^{-2})	$\delta\Delta\mathcal{S}$ (10^{-2})
IP profile	N/A	0.2	0.2	1.0	1.0
$B^0_{\rm Tag}$ track selection	N/A	1.2	0.4	0.8	1.1
Track helix errors	N/A	0.0	0.0	0.0	0.0
Δt selection	N/A	0.1	0.0	0.1	0.0
Vertex quality selection	N/A	0.3	0.9	0.3	0.2
Δz bias	N/A	0.5	0.5	0.4	0.4
Misalignment	N/A	0.4	0.4	0.2	0.2
τ_{B^0} and Δm_d	N/A	0.3	0.3	0.2	0.2
Δt resolution function	N/A	1.3	0.9	2.8	1.7
Flavor tagging	N/A	0.3	0.2	0.2	0.1
Model shape	N/A	0.3	2.9	0.6	0.5
Tag-side interference	N/A	3.6	0.2	0.5	0.4
Signal probability	0.5	0.4	1.9	2.0	0.8
Misreconstruction	0.2	0.1	0.7	0.3	0.3
Fit bias	0.4	1.6	0.3	1.0	0.1
Background CP violation	6.6	7.6	5.7	6.8	5.1
Interference	0.8	1.1	2.0	0.2	1.0
Total	6.6	8.9	7.2	7.9	5.8

CP parameters fixed to null asymmetry and compare this with a fit where they are fixed to the generated values.

To estimate the effects of interference on the signal probability, we employ the same four-body amplitude generator described in the branching fraction measurement. MC samples are generated with the nominal fit result with the three-phase configurations: maximum constructive interference, no interference and maximal destructive interference. Firstly, signal probability is predetermined for all three samples by fitting for the branching fractions of the $a_1\pi$ and $a_2\pi$ components, assuming no interference in the fit model. Finally, a time-dependent fit is performed to the three-phase configurations and the systematic error taken as the maximum deviation from the sample with no generated interference.

V. CONCLUSION

We have presented a measurement of the product branching fraction and time-dependent parameters in $B^0 \to a_1^{\pm}(1260)\pi^{\mp}$ decays, which are in agreement with measurements performed by the *BABAR* Collaboration [17,18]. We obtain the product branching fraction

$$\mathcal{B}(B^0 \to a_1^{\pm}(1260)\pi^{\mp}) \times \mathcal{B}(a_1^{\pm}(1260) \to \pi^{\pm}\pi^{\mp}\pi^{\pm}) = (11.1 \pm 1.0(\text{stat}) \pm 1.4(\text{syst})) \times 10^{-6} \quad (32)$$

and an upper limit on the product branching fraction for a possible decay with the same final state

$$\mathcal{B}(B^0 \to a_2^{\pm}(1320)\pi^{\mp}) \times \mathcal{B}(a_2^{\pm}(1320) \to \pi^{\pm}\pi^{\mp}\pi^{\pm}) < 2.2 \times 10^{-6} \quad \text{at } 90\%\,\text{CL}. \quad (33)$$

This upper limit is an improvement over the current world's most restrictive limit by about two orders of magnitude. In a time-dependent measurement to extract *CP* asymmetries, we obtain the *CP*-violation parameters

$$\begin{aligned} \mathcal{A}_{CP} &= -0.06 \pm 0.05(\text{stat}) \pm 0.07(\text{syst}), \\ \mathcal{C}_{CP} &= -0.01 \pm 0.11(\text{stat}) \pm 0.09(\text{syst}), \\ \mathcal{S}_{CP} &= -0.51 \pm 0.14(\text{stat}) \pm 0.08(\text{syst}), \end{aligned} \quad (34)$$

representing time- and flavor-integrated direct, flavor-dependent direct and mixing-induced *CP* violation, respectively. Simultaneously, we also extract the *CP*-conserving parameters

$$\begin{aligned} \Delta\mathcal{C} &= +0.54 \pm 0.11(\text{stat}) \pm 0.07(\text{syst}), \\ \Delta\mathcal{S} &= -0.09 \pm 0.14(\text{stat}) \pm 0.06(\text{syst}), \end{aligned} \quad (35)$$

which, respectively, describe a rate difference and strong phase difference between the decay channels where the $a_1^{\pm}$ does not contain the spectator quark and those where it does. We find first evidence of mixing-induced *CP* violation in $B^0 \to a_1^{\pm}(1260)\pi^{\mp}$ decays with 3.1σ significance, and the rate where the $a_1^{\pm}$ does not contain the spectator quark is found to dominate the rate where it does at the 4.1σ level. However, there is no evidence for either time- and flavor-integrated direct *CP* violation or flavor-dependent direct *CP* violation. Our results are in good agreement with theoretical predictions given within the QCD factorization framework [11,12] and may be used in either an isospin analysis [15] or $SU(3)$ flavor symmetry [14] to extract ϕ_2.

ACKNOWLEDGMENTS

We thank the KEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; and the KEK computer group, the National Institute of Informatics and the PNNL/EMSL computing group for valuable computing and SINET4 network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS) and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council and the Australian Department of Industry, Innovation, Science and Research; the National Natural Science Foundation of China under Contracts No. 10575109, No. 10775142, No. 10875115, and No. 10825524; the Ministry of Education, Youth and Sports of the Czech Republic under Contracts No. LA10033 and No. MSM0021620859; the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; the BK21 and WCU program of the Ministry Education Science and Technology, National Research Foundation of Korea, and GSDC of the Korea Institute of Science and Technology Information; the Polish Ministry of Science and Higher Education; the Ministry of Education and Science of the Russian Federation and the Russian Federal Agency for Atomic Energy; the Slovenian Research Agency; the Swiss National Science Foundation; the National Science Council and the Ministry of Education of Taiwan; and the U.S. Department of Energy and the National Science Foundation. This work is supported by a Grant-in-Aid from MEXT for Science Research in a Priority Area ("New Development of Flavor Physics") and from JSPS for Creative Scientific Research ("Evolution of Tau-lepton Physics"). We also thank our theory colleagues at the Max-Planck-Institut für Physik, S. Borowka and W. Ochs, for helpful discussions.

[1] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).
[2] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[3] I. Adachi *et al.* (Belle Collaboration), Phys. Rev. Lett. **108**, 171802 (2012).
[4] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **79**, 072009 (2009).
[5] H. Ishino *et al.* (Belle Collaboration), Phys. Rev. Lett. **98**, 211801 (2007).
[6] B. Aubert *et al.* (*BABAR* Collaboration), arXiv:0807.4226v2.
[7] A. Kusaka *et al.* (Belle Collaboration), Phys. Rev. Lett. **98**, 221602 (2007).
[8] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **76**, 012004 (2007).
[9] A. Somov *et al.* (Belle Collaboration), Phys. Rev. D **76**, 011104 (2007).
[10] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **76**, 052007 (2007).
[11] H.-Y. Cheng and K.-C. Yang, Phys. Rev. D **76**, 114020 (2007).
[12] V. Laporta, G. Nardulli, and T. N. Pham, Phys. Rev. D **74**, 054035 (2006).
[13] G. Calderón, J. H. Muñoz, and C. E. Vera, Phys. Rev. D **76**, 094019 (2007).
[14] M. Gronau and J. Zupan, Phys. Rev. D **73**, 057502 (2006).
[15] M. Gronau and D. London, Phys. Rev. Lett. **65**, 3381 (1990).
[16] J. Charles *et al.* (CKMfitter Group), Eur. Phys. J. C **41**, 1 (2005).
[17] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **97**, 051802 (2006).
[18] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **98**, 181803 (2007).
[19] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **81**, 052009 (2010).
[20] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003), and other papers included in this volume.
[21] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002).
[22] Z. Natkaniec *et al.* (Belle SVD2 Group), Nucl. Instrum. Methods Phys. Res., Sect. A **560**, 1 (2006).
[23] R. Brun *et al.*, GEANT 3.21 Report No. CERN DD/EE/84-1, 1984.
[24] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).
[25] H. Tajima *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **533**, 370 (2004).
[26] R. A. Fisher, Ann. Hum. Genet. **7**, 179 (1936).
[27] H. Kakuno *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **533**, 516 (2004).
[28] T. A. Armstrong *et al.* (WA76 Collaboration), Z. Phys. C **48**, 213 (1990).
[29] Y. Amhis *et al.* (Heavy Flavor Averaging Group), arXiv:1207.1158v1 and online updates at http://www.slac.stanford.edu/xorg/hfag.
[30] O. Long, M. Baak, R. N. Cahn, and D. Kirkby, Phys. Rev. D **68**, 034010 (2003).