

Measurement of the branching fraction and CP asymmetry of $B^0 \rightarrow \pi^0 \pi^0$ decays using $198 \times 10^6 B\bar{B}$ pairs in Belle II data

F. Abudinén[✉], I. Adachi[✉], K. Adamczyk[✉], L. Aggarwal[✉], H. Ahmed[✉], H. Aihara[✉], N. Akopov[✉], A. Aloisio[✉], N. Anh Ky[✉], D. M. Asner[✉], H. Atmacan[✉], T. Aushev[✉], V. Aushev[✉], H. Bae[✉], S. Bahinipati[✉], P. Bambade[✉], Sw. Banerjee[✉], S. Bansal[✉], M. Barrett[✉], J. Baudot[✉], M. Bauer[✉], A. Baur[✉], A. Beaubien[✉], J. Becker[✉], P. K. Behera[✉], J. V. Bennett[✉], E. Bernieri[✉], F. U. Bernlochner[✉], V. Bertacchi[✉], M. Bertemes[✉], E. Bertholet[✉], M. Bessner[✉], S. Bettarini[✉], V. Bhardwaj[✉], B. Bhuyan[✉], F. Bianchi[✉], T. Bilka[✉], S. Bilokin[✉], D. Biswas[✉], D. Bodrov[✉], A. Bolz[✉], J. Borah[✉], A. Bozek[✉], M. Bračko[✉], P. Branchini[✉], R. A. Briere[✉], T. E. Browder[✉], A. Budano[✉], S. Bussino[✉], M. Campajola[✉], L. Cao[✉], G. Casarosa[✉], C. Cecchi[✉], M.-C. Chang[✉], P. Chang[✉], R. Cheaib[✉], P. Cheema[✉], V. Chekelian[✉], C. Chen[✉], Y. Q. Chen[✉], B. G. Cheon[✉], K. Chilikin[✉], K. Chirapatpimol[✉], H.-E. Cho[✉], K. Cho[✉], S.-J. Cho[✉], S.-K. Choi[✉], S. Choudhury[✉], D. Cinabro[✉], J. Cochran[✉], L. Corona[✉], S. Cunliffe[✉], S. Das[✉], F. Dattola[✉], E. De La Cruz-Burelo[✉], S. A. De La Motte[✉], G. de Marino[✉], G. De Nardo[✉], M. De Nuccio[✉], G. De Pietro[✉], R. de Sangro[✉], M. Destefanis[✉], S. Dey[✉], A. De Yta-Hernandez[✉], R. Dhamija[✉], A. Di Canto[✉], F. Di Capua[✉], Z. Doležal[✉], I. Domínguez Jiménez[✉], T. V. Dong[✉], M. Dorigo[✉], K. Dort[✉], D. Dossett[✉], S. Dreyer[✉], S. Dubey[✉], G. Dujany[✉], P. Ecker[✉], M. Eliachevitch[✉], P. Feichtinger[✉], T. Ferber[✉], D. Ferlewicz[✉], T. Fillinger[✉], C. Finck[✉], G. Finocchiaro[✉], A. Fodor[✉], F. Forti[✉], A. Frey[✉], B. G. Fulsom[✉], A. Gabrielli[✉], E. Ganiev[✉], M. Garcia-Hernandez[✉], G. Gaudino[✉], V. Gaur[✉], A. Gellrich[✉], G. Ghevondyan[✉], R. Giordano[✉], A. Giri[✉], B. Gobbo[✉], R. Godang[✉], P. Goldenzweig[✉], W. Gradl[✉], T. Grammatico[✉], S. Granderath[✉], E. Graziani[✉], D. Greenwald[✉], Z. Gruberová[✉], T. Gu[✉], Y. Guan[✉], K. Gudkova[✉], J. Guillems[✉], S. Halder[✉], T. Hara[✉], K. Hayasaka[✉], H. Hayashii[✉], S. Hazra[✉], C. Hearty[✉], M. T. Hedges[✉], I. Heredia de la Cruz[✉], M. Hernández Villanueva[✉], A. Hershenhorn[✉], T. Higuchi[✉], E. C. Hill[✉], M. Hohmann[✉], C.-L. Hsu[✉], T. Humair[✉], T. Iijima[✉], K. Inami[✉], G. Inguglia[✉], N. Ipsita[✉], A. Ishikawa[✉], S. Ito[✉], R. Itoh[✉], M. Iwasaki[✉], P. Jackson[✉], W. W. Jacobs[✉], D. E. Jaffe[✉], E.-J. Jang[✉], Q. P. Ji[✉], S. Jia[✉], Y. Jin[✉], K. K. Joo[✉], H. Junkerkalefeld[✉], H. Kakuno[✉], M. Kaleta[✉], A. B. Kaliyar[✉], J. Kandra[✉], K. H. Kang[✉], S. Kang[✉], R. Karl[✉], G. Karyan[✉], T. Kawasaki[✉], C. Kiesling[✉], C.-H. Kim[✉], D. Y. Kim[✉], K.-H. Kim[✉], Y.-K. Kim[✉], H. Kindo[✉], K. Kinoshita[✉], P. Kodyš[✉], T. Koga[✉], S. Kohani[✉], K. Kojima[✉], T. Konno[✉], A. Korobov[✉], S. Korpar[✉], E. Kovalenko[✉], R. Kowalewski[✉], T. M. G. Kraetzschmar[✉], P. Križan[✉], P. Krokovny[✉], T. Kuhr[✉], J. Kumar[✉], R. Kumar[✉], K. Kumara[✉], T. Kunigo[✉], A. Kuzmin[✉], Y.-J. Kwon[✉], S. Lacaprra[✉], T. Lam[✉], L. Lanceri[✉], J. S. Lange[✉], M. Laurenza[✉], K. Lautenbach[✉], R. Lebourcher[✉], F. R. Le Diberder[✉], P. Leitl[✉], C. Li[✉], L. K. Li[✉], Y. B. Li[✉], J. Libby[✉], K. Lieret[✉], Z. Liptak[✉], Q. Y. Liu[✉], Z. Q. Liu[✉], D. Liventsev[✉], S. Longo[✉], A. Lozar[✉], T. Lueck[✉], C. Lyu[✉], M. Maggiora[✉], S. P. Maharana[✉], R. Maiti[✉], S. Maity[✉], R. Manfredi[✉], E. Manoni[✉], A. C. Manthei[✉], M. Mantovano[✉], S. Marcello[✉], C. Marinas[✉], L. Martel[✉], C. Martellini[✉], A. Martini[✉], T. Martinov[✉], L. Massaccesi[✉], T. Matsuda[✉], K. Matsuoka[✉], D. Matvienko[✉], S. K. Maurya[✉], J. A. McKenna[✉], F. Meier[✉], M. Merola[✉], F. Metzner[✉], M. Milesi[✉], C. Miller[✉], K. Miyabayashi[✉], H. Miyake[✉], R. Mizuk[✉], G. B. Mohanty[✉], N. Molina-Gonzalez[✉], S. Moneta[✉], H.-G. Moser[✉], M. Mrvar[✉], R. Mussa[✉], I. Nakamura[✉], K. R. Nakamura[✉], M. Nakao[✉], H. Nakayama[✉], Y. Nakazawa[✉], A. Narimani Charan[✉], M. Naruki[✉], D. Narwal[✉], Z. Natkaniec[✉], A. Natochii[✉], L. Nayak[✉], M. Nayak[✉], G. Nazaryan[✉], C. Niebuhr[✉], N. K. Nisar[✉], S. Nishida[✉], S. Ogawa[✉], H. Ono[✉], Y. Onuki[✉], P. Oskin[✉], P. Pakhlov[✉], G. Pakhlova[✉], A. Paladino[✉], A. Panta[✉], E. Paoloni[✉], S. Pardi[✉], K. Parham[✉], H. Park[✉], J. Park[✉], S.-H. Park[✉], B. Paschen[✉], A. Passeri[✉], S. Patra[✉], S. Paul[✉], T. K. Pedlar[✉], I. Peruzzi[✉], R. Peschke[✉], R. Pestotnik[✉], F. Pham[✉], M. Piccolo[✉], L. E. Piilonen[✉], G. Pinna Angioni[✉], P. L. M. Podesta-Lerma[✉], T. Podobnik[✉], S. Pokharel[✉], L. Polat[✉], C. Praz[✉], S. Prell[✉], E. Prencipe[✉], M. T. Prim[✉], H. Purwar[✉], N. Rad[✉], P. Rados[✉], G. Raeuber[✉], S. Raiz[✉], A. Ramirez Morales[✉], M. Reif[✉], S. Reiter[✉], M. Remnev[✉], I. Ripp-Baudot[✉], G. Rizzo[✉], L. B. Rizzuto[✉], S. H. Robertson[✉], D. Rodríguez Pérez[✉], J. M. Roney[✉], A. Rostomyan[✉], N. Rout[✉], G. Russo[✉], D. A. Sanders[✉], S. Sandilya[✉], A. Sangal[✉], L. Santelj[✉], Y. Sato[✉], V. Savinov[✉], B. Scavino[✉], J. Schueler[✉], C. Schwanda[✉], A. J. Schwartz[✉], Y. Seino[✉], A. Selce[✉], K. Senyo[✉], J. Serrano[✉], M. E. Seviror[✉], C. Sfienti[✉], C. Sharma[✉], C. P. Shen[✉], X. D. Shi[✉], T. Shillington[✉], J.-G. Shiu[✉], B. Shwartz[✉], A. Sibidanov[✉], F. Simon[✉], J. B. Singh[✉], J. Skorupa[✉], R. J. Sobie[✉], A. Soffer[✉], A. Sokolov[✉], E. Solovieva[✉], S. Spataro[✉], B. Spruck[✉], M. Starič[✉], S. Stefkova[✉], Z. S. Stottler[✉], R. Stroili[✉], J. Strube[✉], Y. Sue[✉], M. Sumihama[✉], K. Sumisawa[✉], W. Sutcliffe[✉], S. Y. Suzuki[✉], H. Svidras[✉], M. Takizawa[✉], U. Tamponi[✉], K. Tanida[✉], H. Tanigawa[✉], F. Tenchini[✉], A. Thaller[✉], R. Tiwary[✉], D. Tonelli[✉], E. Torassa[✉], N. Toutounji[✉], K. Trabelsi[✉], I. Tsaklidis[✉], M. Uchida[✉], I. Ueda[✉], Y. Uematsu[✉], T. Uglov[✉], K. Unger[✉], Y. Unno[✉], K. Uno[✉], S. Uno[✉], P. Urquijo[✉], Y. Ushiroda[✉], S. E. Vahsen[✉], R. van Tonder[✉], G. S. Varner[✉], K. E. Varvell[✉], A. Vinokurova[✉], V. S. Vismaya[✉], L. Vitale[✉], V. Vobbilisetti[✉], A. Vossen[✉], H. M. Wakeling[✉], S. Wallner[✉], E. Wang[✉], M.-Z. Wang[✉], X. L. Wang[✉], A. Warburton[✉], M. Watanabe[✉], S. Watanuki[✉], M. Welsch[✉], C. Wessel[✉], J. Wiechczynski[✉], E. Won[✉], X. P. Xu[✉],

B. D. Yabsley¹, S. Yamada², W. Yan³, S. B. Yang⁴, H. Ye⁵, J. Yelton⁶, J. H. Yin⁷, Y. M. Yook⁸, K. Yoshihara⁹,
C. Z. Yuan¹⁰, Y. Yusa¹¹, L. Zani¹², Y. Zhai¹³, Y. Zhang¹⁴, V. Zhilich¹⁵, Q. D. Zhou¹⁶,
X. Y. Zhou¹⁷, V. I. Zhukova¹⁸, and R. Žlebčák¹⁹

(Belle II Collaboration)

 (Received 22 March 2023; accepted 9 May 2023; published 28 June 2023; corrected 22 May 2024)

We report measurements of the branching fraction and CP asymmetry in $B^0 \rightarrow \pi^0 \pi^0$ decays reconstructed at Belle II in an electron-positron collision sample containing 198×10^6 $B\bar{B}$ pairs. We measure a branching fraction $\mathcal{B}(B^0 \rightarrow \pi^0 \pi^0) = (1.38 \pm 0.27 \pm 0.22) \times 10^{-6}$ and a CP asymmetry $\mathcal{A}_{CP}(B^0 \rightarrow \pi^0 \pi^0) = -0.14 \pm 0.46 \pm 0.07$, where the first uncertainty is statistical and the second is systematic.

DOI: [10.1103/PhysRevD.107.112009](https://doi.org/10.1103/PhysRevD.107.112009)

The study of decay-time-dependent CP asymmetries in decays dominated by the $b \rightarrow u$ transition, specifically hadronic decays of bottom mesons into charmless two-body final states, is currently the most precise way to measure the least known angle of the unitarity triangle, ϕ_2 (or α) $\equiv \arg(-V_{td}V_{tb}^*/V_{ud}V_{ub}^*)$. Here, V_{ij} are elements of the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix [1]. Improved measurements of ϕ_2 will test the unitarity of the CKM matrix and constrain possible flavor-structure extensions of the standard model (SM). One approach is to measure the time-dependent decay-rate asymmetry between $\bar{B}^0$ and B^0 mesons that decay to $\pi^+\pi^-$ final states. This asymmetry would be proportional to $\sin(2\phi_2)$ if the decay involved only tree-level $b \rightarrow u$ processes. However, the asymmetry is affected by an unknown and difficult-to-predict shift with respect to the desired ϕ_2 angle due to the presence of $b \rightarrow d$ loop (“penguin”) contributions. The tree-level and penguin amplitudes have similar magnitudes, so the shift is sizable and complicates the determination of ϕ_2 . The penguin and tree contributions can be disentangled using the $B \rightarrow \pi\pi$ isospin relations [2,3]

$$A^{+0} = \frac{1}{\sqrt{2}}A^{+-} + A^{00} \quad \text{and} \quad \bar{A}^{-0} = \frac{1}{\sqrt{2}}\bar{A}^{+-} + \bar{A}^{00}, \quad (1)$$

where A^{ij} and $\bar{A}^{ij}$ are amplitudes for the decays $B \rightarrow \pi^i \pi^j$ and $\bar{B} \rightarrow \pi^i \pi^j$, respectively. Here, B and π indicate charged or neutral bottom-mesons and pions, respectively, while i and j refer to electric charge. Taking advantage of these relations requires precise measurements of the branching fraction $\mathcal{B}$ and CP asymmetries of each $B \rightarrow \pi\pi$ decay

mode. The greatest limitation to exploiting the isospin relations lies in the uncertainty of the $B^0 \rightarrow \pi^0 \pi^0$ inputs, $\mathcal{B}$ and the time-integrated CP asymmetry,

$$\mathcal{A}_{CP} = \frac{\Gamma(\bar{B}^0 \rightarrow \pi^0 \pi^0) - \Gamma(B^0 \rightarrow \pi^0 \pi^0)}{\Gamma(\bar{B}^0 \rightarrow \pi^0 \pi^0) + \Gamma(B^0 \rightarrow \pi^0 \pi^0)}, \quad (2)$$

where Γ is the decay width. The world-average values $\mathcal{B}(B^0 \rightarrow \pi^0 \pi^0) = (1.59 \pm 0.26) \times 10^{-6}$ and $\mathcal{A}_{CP}(B^0 \rightarrow \pi^0 \pi^0) = 0.33 \pm 0.22$ [4] combine measurements reported by the BABAR [5] and Belle [6] Collaborations. Additional measurements would improve our knowledge of ϕ_2 .

Theoretical predictions for $\mathcal{B}(B^0 \rightarrow \pi^0 \pi^0)$ based on QCD factorization [7–10] and perturbative QCD [11,12] are approximately 5 times smaller than the world-average value. Furthermore, the ratio of color-suppressed to color-allowed tree amplitudes, as inferred from other charmless two-body decay modes, does not agree well with expectations [13]. This might indicate large electro-weak-penguin contributions, which are difficult to explain in the SM [14,15]. Various approaches, which predict a wide range of values for $\mathcal{B}$ and $\mathcal{A}_{CP}$, have been proposed as possible solutions to this disagreement [16–19]. More precise measurements of these quantities would help in discriminating among the various solutions proposed to address this discrepancy. In addition, a better understanding of the color-suppressed tree amplitude could help resolve the so-called $B \rightarrow K\pi$ puzzle [20–22]

In this paper, we present a measurement of $\mathcal{B}$ and $\mathcal{A}_{CP}$ for the $B^0 \rightarrow \pi^0 \pi^0$ decay using a data sample consisting of $(198.0 \pm 3.0) \times 10^6$ $B\bar{B}$ pairs [23] collected from 2019 through 2021 [24]. The sample is collected with the Belle II detector, located at the SuperKEKB asymmetric-energy e^+e^- collider [25]. A full description of the Belle II detector is given in Ref. [26]. The detector consists of several subdetectors arranged in a cylindrical structure around the beam pipe. The z axis of the lab frame is defined as the symmetry axis of a superconducting

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

solenoid, which generates a 1.5 T uniform field along the beam direction. The positive direction is given by the electron-beam direction, and the polar angle θ is defined with respect to the $+z$ axis. The detector is divided into three regions, and in increasing order of θ , they are referred to as the forward end cap, barrel, and backward end cap. The inner subdetectors are a silicon pixel detector surrounded by a four-layer double-sided silicon strip detector and a central drift chamber (CDC). These subdetectors are used to reconstruct charged particles and measure their momentum. A time-of-propagation counter [27] and an aerogel ring-imaging Cherenkov detector cover the barrel and forward end cap regions, respectively, and are used for charged particle identification. The electromagnetic calorimeter (ECL) is a segmented array of 8736 thallium-doped cesium iodide [CsI(Tl)] crystals arranged in a projective geometry toward the interaction point and covering about 90% of the solid angle in the center-of-mass (c.m.) frame. The ECL identifies electrons and photons in an energy range of 20 MeV to 4 GeV and occupies the remaining volume inside the superconducting solenoid. Resistive plate chambers and scintillating fibers to identify muons and K_L^0 mesons are installed in the flux return of the magnet.

We use Geant4-based [28] simulated samples to optimize event selection criteria, compare distributions observed in data with expectations, determine fit models, calculate signal efficiencies, and study sources of background. To study the signal, we use 10^7 $\Upsilon(4S) \rightarrow B^0 \bar{B}^0$ decays generated with EvtGen [29], where one B meson decays as $B^0 \rightarrow \pi^0 \pi^0$. To study backgrounds, we use a simulated sample approximately 5 times larger than the data sample. This sample consists of $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ processes and continuum $e^+e^- \rightarrow q\bar{q}$ background, generated with EvtGen and PYTHIA [30], where q denotes a u , d , s , or c quark. To account for a large observed $\tau^+\tau^-$ background, we use a sample of $e^+e^- \rightarrow \tau^+\tau^-$ events generated with KKMC [31] and TAUOLA [32] that is the same size as the continuum sample. To validate our analysis, we use the $B^0 \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^0$ decay as a control mode, as it contains two π^0 particles in the final state and has an order of magnitude more yield. We use a simulated sample of 5×10^6 control-mode events generated with EvtGen. To calibrate and validate our photon reconstruction, we use the $D^{*+} \rightarrow D^0(\rightarrow K_S^0(\rightarrow \pi^+\pi^-)\pi^0)\pi^+$ mode. The data are processed with the Belle II analysis software framework [33,34]. This is the first measurement of this channel at Belle II.

Measuring $B^0 \rightarrow \pi^0 \pi^0$ decay properties is challenging, as the decay is both CKM suppressed and color suppressed. As the final state consists of photons with no tracks, it is difficult to reconstruct. In addition, the large number of neutral pions produced in $e^+e^- \rightarrow q\bar{q}$ continuum events can be combined to mimic the $B^0 \rightarrow \pi^0 \pi^0$ signal. Finally, the reconstruction is susceptible to extraneous photons arising from beam interactions with the beam pipe and

residual gas; this is referred to as beam-induced background. Hence, conventional and machine-learning-based approaches, validated on data, are employed to achieve optimized selections. The signal yield and $\mathcal{A}_{CP}$ are determined by performing a maximum-likelihood fit to the data.

The online-event selection requires all events to pass criteria based on total energy and neutral-particle multiplicity. In the offline analysis we identify photon candidates by requiring that the number of crystals in an ECL energy deposition (cluster), which can be fractional as a result of energy splitting with nearby clusters, be greater than 1.5. The cluster timing is required to be within 200 ns of the offline estimated event time. We require that the cluster energy exceed 20.0 MeV in the barrel region of the ECL ($32.2^\circ < \theta < 128.7^\circ$) and 22.5 MeV in the forward end cap ($12.4^\circ < \theta < 31.4^\circ$) and backward end cap ($130.7^\circ < \theta < 155.1^\circ$) regions. Since scintillation light from CsI(Tl) crystals has a relatively long decay time, high-energy events from Bhabha processes ($e^+e^- \rightarrow e^+e^-$) can overlap with subsequent hadronic event signals. A random photon from the hadronic event can be combined with the residual energy (misreconstructed photon) in the CsI(Tl) crystals to form a π^0 candidate. To suppress nonsignal photons, and to account for the angular dependence of ECL-related variables, we employ a boosted decision tree (BDT) [35], separately for each of the three polar-angle regions of the ECL. The BDT is trained on a simulated sample of $B^0 \rightarrow \pi^0 \pi^0$ events that includes the effect of beam-induced background and uses ten input variables: the photon energy and transverse momentum, the energy recorded in the crystal having the highest signal, the distance between the ECL cluster and the nearest charged particle hitting the ECL, four variables that depend on how energy is distributed between the clusters, and two variables that depend on the fraction of cluster energy detected in the central crystal. We refer to this classifier as the “photon-BDT.” We impose a requirement on the photon-BDT output that maximizes a figure of merit $S/\sqrt{S+B}$, where S and B are the expected number of signal B^0 and background events, respectively. In simulated samples, this selection removes 68% of misreconstructed photons and retains 96% of genuine photons. Studies of the $D^{*+} \rightarrow D^0(\rightarrow K_S^0(\rightarrow \pi^+\pi^-)\pi^0)\pi^+$ decay are used to validate the photon-BDT classifier’s performance. The signal efficiency and photon-BDT output distribution are consistent between simulations and data.

Selected photons are paired to form π^0 candidates. We require that the π^0 momentum in the lab frame be greater than 1.5 GeV/ c and that the angle between the momenta of final-state photons in the lab frame be less than 0.4 rad. These requirements suppress the combinatorial background from low-energy photons. The cosine of the helicity angle, defined as the angle between the higher-energy γ direction in the π^0 rest frame and the π^0 direction in the lab frame, is required to be less than 0.99 to reject misreconstructed π^0 ’s,

which tend to peak very close to one. The diphoton mass is required to be between 0.115 and 0.150 GeV/ c^2 , which corresponds to a range of approximately $+2.0\sigma$ and -2.5σ about the known π^0 mass. The mass requirement is asymmetric as the reconstructed π^0 mass has a slight negative skew due to energy leakage from the ECL calorimeter. We improve the momentum resolution of the π^0 candidates by performing a kinematic fit that constrains their mass to the known value [4]. Signal B^0 candidates are reconstructed by combining two π^0 candidates. To select signal B^0 candidates, two kinematic variables are defined:

$$M_{bc} = \sqrt{E_{\text{beam}}^2 - |\vec{p}_B|^2} \quad \text{and} \quad \Delta E = E_B - E_{\text{beam}}, \quad (3)$$

where E_{beam} is the beam energy and $(E_B, \vec{p}_B)$ is the reconstructed four-momentum of the B candidate. All quantities are calculated in the c.m. frame of the $\Upsilon(4S)$ resonance. The M_{bc} and ΔE distributions of signal decays peak at the B mass and zero, respectively. Candidate B mesons are required to have $5.26 < M_{bc} < 5.29$ GeV/ c^2 and $-0.3 < \Delta E < 0.2$ GeV. The ΔE requirement is not centered around zero because of energy leakage from the ECL cluster.

In data, photon-energy corrections are applied to correct for ECL miscalibration. Studies of the $D^{*+} \rightarrow D^0(\rightarrow K_S^0(\pi^+\pi^-)\pi^0)\pi^+$ control mode are used to validate the corrections. The π^0 momentum is predicted using the momenta of the charged pions and energy-momentum conservation. In simulation, this predicted momentum is typically closer to the true momentum than the measured momentum as the momentum of charged pions is measured more precisely in the CDC than photon energies are measured in the ECL. When the corrections are applied, the data-simulation difference between the predicted and measured π^0 momentum decreases, and the difference between the ΔE peak position in data and simulation is approximately 1 MeV.

The sample includes a large continuum background. To reduce this background, we use topological variables that take advantage of the jetlike nature of $q\bar{q}$ events and the spherical distribution of $B\bar{B}$ events. We train a BDT classifier to analyze 28 variables comprising modified Fox-Wolfram moments [36], sphericity-related quantities [37], thrust-related quantities [38], and sets of concentric cones with various opening angles centered around the thrust axis. Variables that show correlations with ΔE and M_{bc} greater than 5% are excluded. We train the continuum classifier to identify statistically significant signal and background features using simulated signal samples and sideband data; the latter consists of events that satisfy all selection criteria but are in a signal-depleted region $5.22 < M_{bc} < 5.27$ GeV/ c^2 and $0.1 < \Delta E < 0.5$ GeV. We use these simulated samples to determine a minimum threshold C_{min} of the continuum classifier output C that

minimizes the expected statistical uncertainty of the $\mathcal{A}_{CP}$ measurement. This selection rejects 93% of the background while retaining 76.5% of the signal. The continuum classifier output is transformed into a Gaussian-like shape according to $T_c = \log[(C - C_{\text{min}})/(C_{\text{max}} - C)]$, where C_{max} is the maximum value of the continuum classifier output. Candidate B mesons are required to have $|T_c| < 3$. The T_c distributions of signal candidates and continuum are expected to peak at one and zero, respectively.

After suppression of continuum background, 1.3% of events have more than one B^0 candidate. For such events, the average multiplicity is 2.03 candidates per event. We choose the candidate with the minimum sum of the absolute deviations of the reconstructed π^0 masses from the known value [4]. This requirement is 56% efficient in selecting the correct B^0 candidate. Following all selections, 35.5% of signal events remain, of which 99.0% are correctly reconstructed. The high fraction of correctly reconstructed events is due to the low percentage of $B\bar{B}$ events in which there are three high-momentum π^0 candidates.

The resulting event sample consists of four main components: signal, continuum, background from nonsignal B decays ($B\bar{B}$ background), and $\tau^+\tau^-$ events. From studies of simulated samples, we find that 90% of $B\bar{B}$ background is from B^+B^- events in which the B^+ meson decays into a $\rho^+\pi^0$ final state and the charged pion from the subsequent $\rho^+ \rightarrow \pi^+\pi^0$ decay is not reconstructed. The remaining 10% is dominated by $B^0\bar{B}^0$ events in which the B^0 meson decays into a $K_S^0(\rightarrow \pi^0\pi^0)\pi^0$ final state, and one π^0 from the K_S^0 decay is not reconstructed. The $B\bar{B}$ background peaks at similar values of M_{bc} and T_c as the signal, but its ΔE distribution is shifted to negative values due to the energy lost from the signal-decay daughter that is not reconstructed. In addition, 2.9% of signal candidates arise from $\tau^+\tau^-$ events; these candidates are treated as part of the continuum background because their respective M_{bc} , ΔE , and T_c distributions are nearly identical.

To measure $\mathcal{A}_{CP}$, the flavor of the signal B^0 is determined by reconstructing the accompanying (tag-side) B meson in each event using the category-based algorithm described in Ref. [39]. The tagging information is encoded in two parameters: the b or $\bar{b}$ flavor of the tag-side B (q) and the purity (r). The value $q = +1$ tags a B^0 , whereas $q = -1$ tags a $\bar{B}^0$. The value of r is the algorithm's confidence for an assigned q value. It is defined as $r = 1 - 2w$, where w is the fraction of wrongly tagged events and ranges from zero, for no flavor distinction between B^0 and $\bar{B}^0$, to one for an unambiguous flavor assignment. For example, for a sample of events having $r = 0$, there would be an equal number of correctly and incorrectly tagged events; for a sample having $r = 1$, there would be no incorrectly tagged events. We divide signal candidates into seven intervals of r , with intervals having approximately equal numbers of events. The signal yield and $\mathcal{A}_{CP}$ values are determined by

performing a three-dimensional (M_{bc} , ΔE , T_c) unbinned extended maximum-likelihood fit simultaneously to events in the seven intervals of r . The likelihood function is given by

$$\mathcal{L} = \frac{e^{-\sum_j N_j}}{\prod_k N_k!} \prod_k \left[\prod_{i=1}^{N_k} \left(\sum_j f_k^j N_j P_k^j(M_{bc}^i, \Delta E^i, T_c^i, q^i) \right) \right], \quad (4)$$

where i is the number of candidates, j is the sample component in terms of signal (s), continuum (c), and $B\bar{B}$, and k indicates the r interval. Here, N_j denotes the yield for component j , N_k denotes the number of candidates in the k th bin, f_k^j is the fraction of candidates in the k th bin for the j th component, and $P_k^j(M_{bc}^i, \Delta E^i, T_c^i, q^i)$ is the probability density function (PDF) to have the i th event of the j th component in the k th bin. The values of f_k^c implicitly include a factor of one-half due to the division of the data into positive and negative q values for each r intervals. Sideband data are used to determine f_k^c , while $f_k^{B\bar{B}}$ is obtained from large simulated samples.

The PDF for the signal component is

$$\begin{aligned} P_k^s(M_{bc}, \Delta E, T_c, q) &= [1 - q\Delta w_k + q\Delta\epsilon_k(1 - 2w_k) \\ &\quad + [q(1 - 2w_k) + \Delta\epsilon_k(1 - q\Delta w_k)](1 - 2\chi_d)\mathcal{A}_{CP}] \\ &\quad \times P^s(M_{bc}, \Delta E, T_c), \end{aligned} \quad (5)$$

where w_k is the fraction of signal events incorrectly tagged (wrong-tag), Δw_k is the difference in the wrong-tag fraction between positive and negative tags, and $\Delta\epsilon_k = \Delta\epsilon_k/2\epsilon_k$ is the asymmetry of the tagging efficiency. Here, ϵ_k is the tagging efficiency and $\Delta\epsilon_k$ is the difference in the tagging efficiency between positive and negative tags. The fraction of signal events in each r interval (f_k^s), along with w_k , Δw_k , and $\Delta\epsilon_k$, are fixed to values obtained from a fit to $B^0 \rightarrow D^{(*)-}h^+$ decays, where h^+ stands for a π^+ or K^+ , following Ref. [39]. The CP asymmetry in data is diluted by a factor $(1 - 2w)$ due to incorrect tagging and by a factor of $(1 - 2\chi_d)$ due to $B^0\bar{B}^0$ mixing, where $\chi_d = 0.1875 \pm 0.0017$ is the time-integrated $B^0\bar{B}^0$ -mixing probability [4].

The T_c PDFs of the signal, continuum, and $B\bar{B}$ components are each modeled using the sum of a Gaussian and a bifurcated Gaussian function with independent mean and width parameters. The T_c PDFs for the signal and $B\bar{B}$ are modeled independently for each r bin using simulated data; this accounts for an observed dependence of T_c on r . The T_c PDF for continuum events is the same for all r bins and is taken from the data sideband.

For the (M_{bc} , ΔE) modeling, a small correlation between M_{bc} and ΔE for the signal is taken into account by using a two-dimensional kernel-density shape. To simplify the M_{bc}

and ΔE modeling of the B^+B^- and $B^0\bar{B}^0$ backgrounds, we assume that they follow the same distributions as the dominant $B^+ \rightarrow \rho^+\pi^0$ and $B^0 \rightarrow K_S^0(\rightarrow \pi^0\pi^0)\pi^0$ decays, respectively. The PDFs for the $B\bar{B}$ backgrounds are the sum of two ARGUS functions in M_{bc} and a kernel-density shape in ΔE . All signal and $B\bar{B}$ PDF parameters are fixed to those obtained from fits to large samples of simulated events.

The upper end point of the M_{bc} distribution depends on the beam energy, which varied throughout the course of data taking. To account for this, the continuum is modeled with eight ARGUS functions that have end points evenly spaced from 5.287 to 5.290 GeV/ c^2 . The contribution of each ARGUS function is fixed to the fraction of events reconstructed at each of the corresponding c.m. energies. Using eight ARGUS functions models well the variation of the M_{bc} end point and provides a good fit to the data. The ΔE distribution of the continuum is modeled with a straight line. We determine the parameters of the continuum PDF for all r bins by fitting to the data sideband region. A small dependence of ΔE on $q \cdot r$ found in simulated samples is neglected, as there are insufficient events in the higher r bins of the data sideband for a reliable fit. The $q \cdot r$ distribution for the continuum events shows an asymmetry that could bias the $\mathcal{A}_{CP}$ results. This asymmetry, defined similarly to Eq. (2), is determined to be 0.033 ± 0.002 . To account for this, we include a $q \cdot r$ asymmetry term in the continuum PDF that is equal to the $q \cdot r$ asymmetry term extracted from the sideband data. From simulated experiments, small biases of 1% in the branching fraction and 0.02 in $\mathcal{A}_{CP}$ are found; we treat these biases as systematic uncertainties.

The reconstruction and fitting procedure is further validated using $B^0 \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^0$ decays. This control sample includes a small cross-feed component in which a particle from the tag side is mistakenly included in the signal reconstruction. All photon and π^0 selections are the same, with the exception of the 1.5 GeV/ c threshold on π^0 momentum, which is removed since the π^0 from the D^0 has significantly lower momentum than the π^0 from a signal decay. We determine the branching fraction to be $\mathcal{B}(B^0 \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^0) = (3.66 \pm 0.21) \times 10^{-5}$ and the direct CP asymmetry to be $\mathcal{A}_{CP}(B^0 \rightarrow \bar{D}^0\pi^0) = 0.01 \pm 0.16$. The uncertainties for the control mode measurements are statistical only. These values agree with previous measurements [40]. Figure 1 shows signal-enhanced projections of the fits to data. The signal-enhanced region is defined as $5.275 < M_{bc} < 5.285$ GeV/ c^2 , $-0.10 < \Delta E < 0.05$ GeV, and $0 < T_c < 3$; for each plot, the selection on the plotted variable is not applied. On average, these signal-enhanced regions contain 47% of signal decays but only 11% of background. The control mode is also used to calibrate the ΔE width of the signal mode, which is taken from simulation.

We apply the fit described above to the 3177 selected $B^0 \rightarrow \pi^0\pi^0$ candidate events. The signal yield, $\mathcal{A}_{CP}$, and

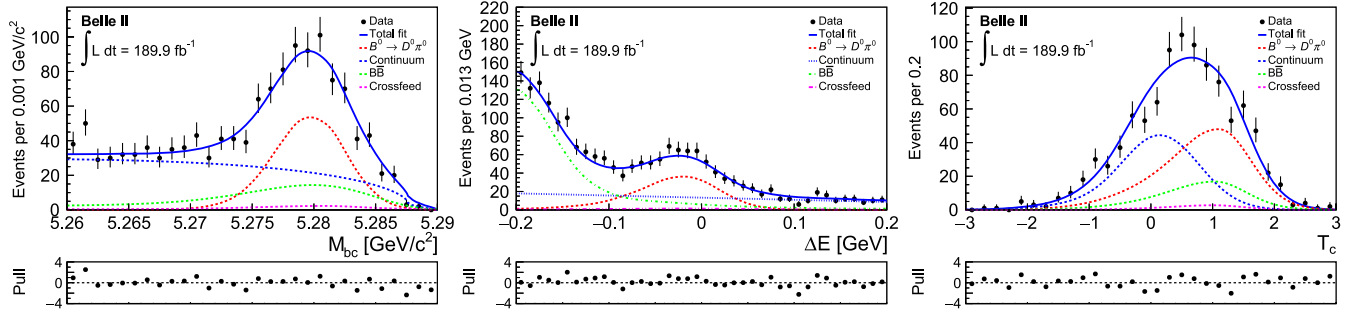


FIG. 1. Distributions of M_{bc} (left), ΔE (middle), and T_c (right) for the $B^0 \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^0$ candidates, for all seven r bins combined. The result of the fit to the data is shown as a solid blue curve. The fit components are shown as a red dashed curve (signal), blue dotted curve (continuum background), green dash-dotted curve ($B\bar{B}$ background), and magenta solid-dotted curve (cross feed). The plots are signal enhanced, which correspond to candidates with $5.275 < M_{bc} < 5.285$ GeV/c^2 , $-0.10 < \Delta E < 0.05$ GeV , and $0 < T_c < 3$. When the respective variable is displayed, the selections on that variable are not applied. The difference between observed and fit value divided by the uncertainty from the fit (pulls) is shown below each distribution.

continuum yield are free to vary, while the $B\bar{B}$ yield is fixed to the expectation from simulations. We obtain a signal yield of 93 ± 18 events. Figure 2 shows the signal-enhanced projections of the fits to data, separately for positive and negative q tags. The signal-enhanced region for the $B^0 \rightarrow \pi^0\pi^0$ signal decay is the same as that for the $B^0 \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^0$ control mode and rejects approximately 96% of the continuum background. To determine the signal significance, we convolve the statistical and additive systematic uncertainties and calculate the test

statistic $2(\log \mathcal{L}_m - \log \mathcal{L}_0) = 32.0$ with two degrees of freedom, where $\log \mathcal{L}_m$ is the log-likelihood of the measured signal yield and $\log \mathcal{L}_0$ is determined by fixing the signal yield to zero. The second degree of freedom is lost due to $\mathcal{A}_{CP} = 0$ when there is no signal. A total significance of 5.2 standard deviations is obtained.

The branching fraction is calculated using

$$\mathcal{B}(B^0 \rightarrow \pi^0\pi^0) = \frac{N_s(1 + f^{+-}/f^{00})}{2\epsilon N_{B\bar{B}}\mathcal{B}(\pi^0 \rightarrow \gamma\gamma)^2}, \quad (6)$$

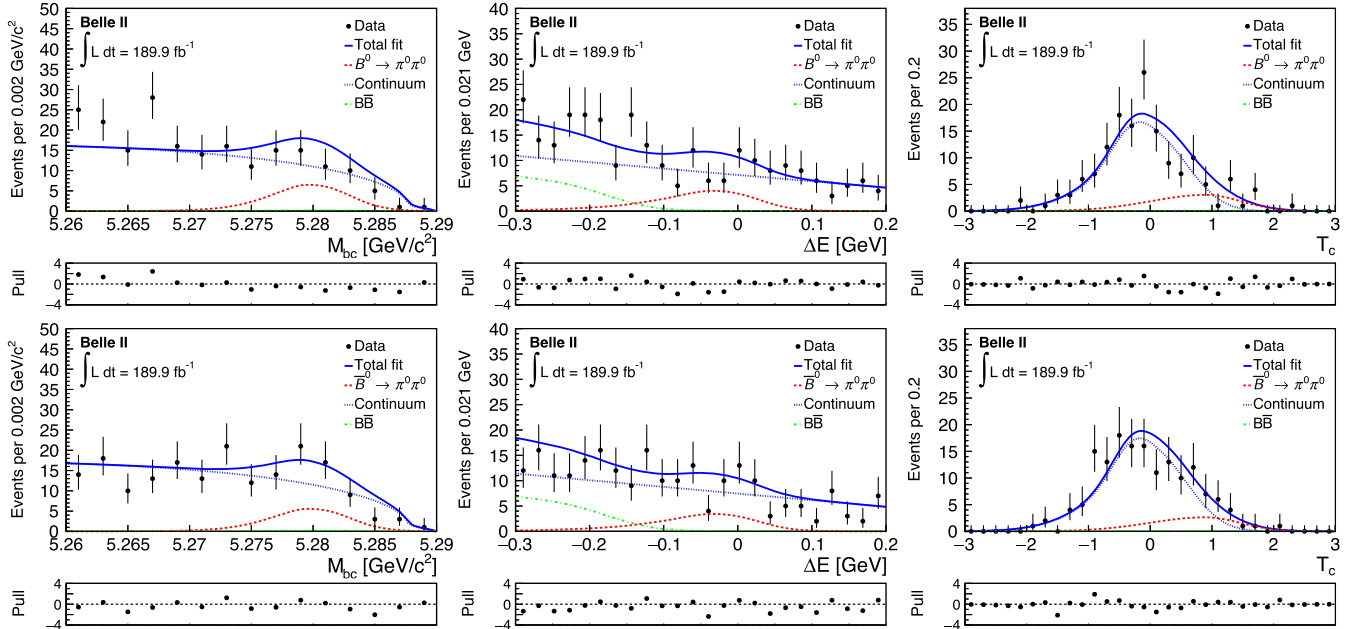


FIG. 2. Distributions of M_{bc} (left), ΔE (middle), and T_c (right) for the $B^0 \rightarrow \pi^0\pi^0$ candidates, for all seven r bins combined, with positive (top) and negative (bottom) q tags. The result of the fit to the data is shown as a solid blue curve. The fit components are shown as a red dashed curve (signal), blue dotted curve (continuum background), and green dash-dotted curve ($B\bar{B}$ background). The plots are signal enhanced, which correspond to candidates with $5.275 < M_{bc} < 5.285$ GeV/c^2 , $-0.10 < \Delta E < 0.05$ GeV , and $0 < T_c < 3$. When the respective variable is displayed, the selections on that variable are not applied. The difference between observed and fit value divided by the uncertainty from the fit (pulls) are shown below each distribution.

where N_s is the signal yield, ε is the signal reconstruction and selection efficiency, $N_{B\bar{B}}$ is the number of $B\bar{B}$ pairs produced, $\mathcal{B}(\pi^0 \rightarrow \gamma\gamma)$ [4] is the $\pi^0 \rightarrow \gamma\gamma$ branching fraction, and f^{+-}/f^{00} is the ratio of the branching fractions for the decay of $\Upsilon(4S)$ to B^+B^- and $B^0\bar{B}^0$. The ratio f^{+-}/f^{00} is determined to be $1.065 \pm 0.012 \pm 0.019 \pm 0.047$ [41], where the first and second uncertainties are statistical and systematic, respectively, and the third uncertainty is due to the assumption of isospin symmetry in $B \rightarrow J/\psi(\rightarrow \ell\ell)K$, where $\ell = e$ or μ . Inserting the values $N_s = 93 \pm 18$, $\varepsilon = (35.5 \pm 4.7)\%$, $N_{B\bar{B}} = (198.0 \pm 3.0) \times 10^6$, and $\mathcal{B}(\pi^0 \rightarrow \gamma\gamma) = (98.823 \pm 0.034)\%$ [4], we obtain

$$\mathcal{B}(B^0 \rightarrow \pi^0\pi^0) = (1.38 \pm 0.27 \pm 0.22) \times 10^{-6}, \quad (7)$$

where the first and second uncertainties are statistical and systematic (discussed below), respectively. The uncertainty in ε is due to the systematic uncertainty associated with π^0 reconstruction and continuum classifier efficiency.

The main sources of systematic uncertainties are listed in Table I and are evaluated as follows. A 3.4% systematic uncertainty associated with the π^0 reconstruction efficiency is determined from data using the decays $D^{*-} \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^-$ and $D^{*-} \rightarrow \bar{D}^0(\rightarrow K^+\pi^-)\pi^-$, where the π^0 selection is identical to that of the signal. The π^0 reconstruction efficiency as a function of momentum is also measured using $\tau^- \rightarrow 3\pi\pi^0\nu$ and $\tau^- \rightarrow 3\pi\nu$ decays. A difference of 4.7% in efficiency is observed between the measurement based on D decays and the measurement based on τ leptons. This difference increases the systematic

uncertainty for a total of 5.8% per pion. The total systematic uncertainty associated with the π^0 reconstruction efficiency is then 11.6%, as there are two pions and their errors are fully correlated.

The systematic uncertainty associated with the continuum parametrization accounts for the uncertainty in each of the data-driven continuum PDF parameters. The contribution of each parameter is determined by refitting on simulated data with the parameter used in the continuum PDF fluctuated by its one-standard-deviation uncertainties. All other continuum PDF parameters are correspondingly shifted according to their correlation with the fluctuated parameter. The systematic uncertainty is the sum in quadrature of the change in signal yield for each parameter.

The systematic uncertainty associated with the efficiency of the continuum classifier is determined using $B^0 \rightarrow \bar{D}^0(\rightarrow K^+\pi^-\pi^0)\pi^0$ decays. The efficiencies of the classifier selection in data and simulation are consistent within the statistical uncertainties. The overall statistical uncertainty is assigned as systematic uncertainty. The f^{+-}/f^{00} systematic uncertainty combines the original systematic uncertainty and the uncertainty due to the assumption of isospin symmetry. The systematic uncertainty associated with the fixed $B\bar{B}$ background yield is determined by refitting on simulated data with the generated $B\bar{B}$ yield fluctuated by its one-standard-deviation uncertainties. The systematic uncertainty associated with the fixed signal fractions for r bins is determined by refitting simulated data with the signal fractions fluctuated by their one-standard-deviation uncertainties. The systematic uncertainty is the sum in quadrature of the change in signal yield for each bin. A similar procedure to determine the systematic uncertainty associated with fixing the continuum fractions in the r bins is also performed. The systematic uncertainty associated with the photon-energy corrections is determined by refitting on data with the values of the corrections fluctuated by their uncertainties. The largest change in yield is taken as the systematic uncertainty. The systematic uncertainty associated with the assumption of independence of ΔE from r is determined by refitting on simulated data with the ΔE slope for each r bin separately estimated using large simulated samples. The procedure for estimating the uncertainty in the number of $B\bar{B}$ meson pairs is described in Ref. [24]. The systematic uncertainty associated with the choice of $(M_{bc}, \Delta E)$ signal models is determined by refitting on simulated data with two uncorrelated Crystal Ball functions [42]. A small bias in the calculated branching fraction due to the limitations of the PDFs used to model the data is included as a systematic uncertainty. The systematic uncertainty associated with a possible bias due to best candidate selection is determined by refitting on data with the best candidate randomly selected. The total systematic uncertainty is taken to be the sum in quadrature of the individual contributions (16.2%).

TABLE I. Summary of systematic uncertainties. The total is calculated by adding all systematic uncertainties in quadrature.

Source	$\mathcal{B}(\%)$	$\mathcal{A}_{CP}$
π^0 reconstruction efficiency	11.6	...
Continuum parametrization	7.4	0.02
Continuum classifier efficiency	6.5	...
$1 + f^{+-}/f^{00}$	2.5	...
Fixed $B\bar{B}$ background yield	2.3	0.01
Fixed signal r bin fractions	2.2	0.01
Knowledge of the photon-energy scale	2.0	...
Assumption of independence of ΔE from r	1.8	<0.01
Number of $B\bar{B}$ meson pairs	1.5	<0.01
Choice of $(M_{bc}, \Delta E)$ signal model	1.3	0.02
Fixed continuum r bin fraction	1.1	<0.01
Branching fraction fit bias	1.0	...
Best candidate selection	0.2	<0.01
Mistagging parameters	...	0.05
Potential nonzero $B\bar{B}$ background $\mathcal{A}_{CP}$	...	0.03
$\mathcal{A}_{CP}$ fit bias	...	0.02
Continuum $q \cdot r$ asymmetry	...	0.01
Total	16.2	0.07

The CP asymmetry of $B^0 \rightarrow \pi^0 \pi^0$ decays is measured to be

$$\mathcal{A}_{CP}(B^0 \rightarrow \pi^0 \pi^0) = -0.14 \pm 0.46 \pm 0.07. \quad (8)$$

The main sources of systematic uncertainties are listed in Table I and are evaluated as follows. The systematic uncertainties for the continuum parametrization, fixed $B\bar{B}$ background yield, fixed signal r bin fraction, choice of $(M_{bc}, \Delta E)$ signal model, and $\mathcal{A}_{CP}$ fit bias are evaluated as previously described. The systematic uncertainty due to the fixed values of w_k , Δw_k , and $\Delta \epsilon_k$ are determined by refitting simulated data with each parameter individually fluctuated by its one-standard-deviation uncertainties. The systematic uncertainty is the sum in quadrature of the change in $\mathcal{A}_{CP}$ for each parameter. The systematic uncertainty associated with bias due to potential nonzero $\mathcal{A}_{CP}$ for the $B\bar{B}$ background is determined by refitting on simulated data with the generated $\mathcal{A}_{CP}$ for the two dominant $B\bar{B}$ backgrounds fluctuated by one standard deviation from their known values. The systematic uncertainty associated with the $q \cdot r$ asymmetry of the continuum is determined by refitting the data with the $q \cdot r$ asymmetry term of the continuum PDF fluctuated by its one-standard-deviation uncertainty.

We average our results with previous measurements of $\mathcal{B}$ and $\mathcal{A}_{CP}$ for $B^0 \rightarrow \pi^0 \pi^0$ and use the isospin analysis in Ref. [2], along with previous measurements of the branching fractions and CP -asymmetry parameters for $B^0 \rightarrow \pi^+ \pi^-$ and $B^+ \rightarrow \pi^+ \pi^0$ [4], and find that the ϕ_2 exclusion interval at the 68% confidence level increases by 1.0° , corresponding to a relative increase in precision of 1.4%. Similarly, at the 95% confidence level the exclusion interval increases by 1.3° , corresponding to a relative increase in precision of 2.0%.

In conclusion, we measure the branching fraction and direct CP asymmetry to be $\mathcal{B}(B^0 \rightarrow \pi^0 \pi^0) = (1.38 \pm 0.27 \pm 0.22) \times 10^{-6}$ and $\mathcal{A}_{CP} = -0.14 \pm 0.46 \pm 0.07$, respectively. These measurements agree with previous measurements [4]. The branching fraction uncertainty is similar in size to those reported by the BABAR and Belle Collaborations, despite using a sample 2.4 and 4.0 times smaller, respectively. These improvements are due to a 60% higher signal efficiency with approximately 40% less background [6]. The higher efficiency and lower background result from improved photon timing, BDT-based photon selection, and data-driven continuum suppression.

This work, based on data collected using the Belle II detector, which was built and commissioned prior to March 2019, was supported by Science Committee of the Republic of Armenia Grant No. 20TTCG-1C010; Australian Research Council and research Grants No. DE220100462, No. DP180102629, No. DP170102389, No. DP170102204,

No. DP150103061, No. FT130100303, No. FT130100018, and No. FT120100745; Austrian Federal Ministry of Education, Science and Research, Austrian Science Fund No. P 31361-N36 and No. J4625-N, and Horizon 2020 ERC Starting Grant No. 947006 “InterLeptons”; Natural Sciences and Engineering Research Council of Canada, Compute Canada and CANARIE; Chinese Academy of Sciences and research Grant No. QYZDJ-SSW-SLH011, National Natural Science Foundation of China and research Grants No. 11521505, No. 11575017, No. 11675166, No. 11761141009, No. 11705209, and No. 11975076, LiaoNing Revitalization Talents Program under Contract No. XLYC1807135, Shanghai Pujiang Program under Grant No. 18PJ1401000, Shandong Provincial Natural Science Foundation Project No. ZR2022JQ02, and the CAS Center for Excellence in Particle Physics (CCEPP); the Ministry of Education, Youth, and Sports of the Czech Republic under Contract No. LTT17020 and Charles University Grant No. SVV 260448 and the Czech Science Foundation Grant No. 22-18469S; European Research Council, Seventh Framework PIEF-GA-2013-622527, Horizon 2020 ERC-Advanced Grants No. 267104 and No. 884719, Horizon 2020 ERC-Consolidator Grant No. 819127, Horizon 2020 Marie Skłodowska-Curie Grant Agreement No. 700525 “NIOBE” and No. 101026516, and Horizon 2020 Marie Skłodowska-Curie RISE project JENNIFER2 Grant Agreement No. 822070 (European grants); L’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3) du CNRS (France); BMBF, DFG, HGF, MPG, and AvH Foundation (Germany); Department of Atomic Energy under Project Identification No. RTI 4002 and Department of Science and Technology (India); Israel Science Foundation Grant No. 2476/17, U.S.–Israel Binational Science Foundation Grant No. 2016113, and Israel Ministry of Science Grant No. 3-16543; Istituto Nazionale di Fisica Nucleare and the research grants BELLE2; Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research Grants No. 16H03968, No. 16H03993, No. 16H06492, No. 16K05323, No. 17H01133, No. 17H05405, No. 18K03621, No. 18H03710, No. 18H05226, No. 19H00682, No. 22H00144, No. 26220706, and No. 26400255, the National Institute of Informatics, and Science Information NETwork 5 (SINET5), and the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan; National Research Foundation (NRF) of Korea Grants No. 2016R1-D1A1B-02012900, No. 2018R1-A2B-3003643, No. 2018R1-A6A1A-06024970, No. 2018R1-D1A1B-07047294, No. 2019R1-I1A3A-01058933, No. 2022R1-A2C-1003993, and No. RS-2022-00197659, Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project, the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information and KREONET/GLORIAD; Universiti Malaya

RU grant, Akademi Sains Malaysia, and Ministry of Education Malaysia; Frontiers of Science Program Contracts No. FOINS-296, No. CB-221329, No. CB-236394, No. CB-254409, and No. CB-180023, and SEP-CINVESTAV research Grant No. 237 (Mexico); the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Science and Higher Education of the Russian Federation, Agreement No. 14.W03.31.0026, and the HSE University Basic Research Program, Moscow; University of Tabuk research Grants No. S-0256-1438 and No. S-0280-1439 (Saudi Arabia); Slovenian Research Agency and research Grants No. J1-9124 and No. P1-0135; Agencia Estatal de Investigacion, Spain Grant No. RYC2020-029875-I and Generalitat Valenciana, Spain Grant No. CIDEGENT/2018/020 Ministry of Science and Technology and research Grants No. MOST106-2112-M-002-005-MY3 and No. MOST107-2119-M-002-035-MY3, and the Ministry of Education (Taiwan); Thailand Center of Excellence in Physics; TUBITAK ULAKBIM (Turkey); National Research Foundation of Ukraine, Project No. 2020.02/

0257, and Ministry of Education and Science of Ukraine; the U.S. National Science Foundation and research Grants No. PHY-1913789 and No. PHY-2111604, and the U.S. Department of Energy and research Grants No. DE-AC06-76RLO1830, No. DE-SC0007983, No. DE-SC0009824, No. DE-SC0009973, No. DE-SC0010007, No. DE-SC0010073, No. DE-SC0010118, No. DE-SC0010504, No. DE-SC0011784, No. DE-SC0012704, No. DE-SC0019230, No. DE-SC0021274, and No. DE-SC0022350; and the Vietnam Academy of Science and Technology (VAST) under Grant No. DL0000.05/21-23. These acknowledgments are not to be interpreted as an endorsement of any statement made by any of our institutes, funding agencies, governments, or their representatives. We thank the SuperKEKB team for delivering high-luminosity collisions; the KEK cryogenics group for the efficient operation of the detector solenoid magnet; the KEK computer group and the NII for on-site computing support and SINET6 network support; and the raw-data centers at BNL, DESY, GridKa, IN2P3, INFN, and the University of Victoria for off-site computing support.

-
- [1] M. Kobayashi and T. Maskawa, *Prog. Theor. Phys.* **49**, 652 (1973).
 - [2] M. Gronau and D. London, *Phys. Rev. Lett.* **65**, 3381 (1990).
 - [3] J. Charles, O. Deschamps, S. Descotes-Genon, and V. Niess, *Eur. Phys. J. C* **77**, 574 (2017).
 - [4] R. L. Workman *et al.* (Particle Data Group), *Prog. Theor. Exp. Phys.* **2022**, 083C01 (2022).
 - [5] J. P. Lees *et al.* (BABAR Collaboration), *Phys. Rev. D* **87**, 052009 (2013).
 - [6] T. Julius *et al.* (Belle Collaboration), *Phys. Rev. D* **96**, 032007 (2017).
 - [7] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, *Phys. Rev. Lett.* **83**, 1914 (1999).
 - [8] M. Beneke and M. Neubert, *Nucl. Phys.* **B675**, 333 (2003).
 - [9] M. Beneke and S. Jager, *Nucl. Phys.* **B751**, 160 (2006).
 - [10] V. Pilipp, *Nucl. Phys.* **B794**, 154 (2008).
 - [11] C.-D. Lu, K. Ukai, and M.-Z. Yang, *Phys. Rev. D* **63**, 074009 (2001).
 - [12] Y.-L. Zhang, X.-Y. Liu, Y.-Y. Fan, S. Cheng, and Z.-J. Xiao, *Phys. Rev. D* **90**, 014029 (2014).
 - [13] Y.-Y. Charng and H.-n. Li, *Phys. Rev. D* **71**, 014036 (2005).
 - [14] S. Mishima and T. Yoshikawa, *Phys. Rev. D* **70**, 094024 (2004).
 - [15] Y.-Y. Charng and H.-n. Li, *Phys. Lett. B* **594**, 185 (2004).
 - [16] C.-F. Qiao, R.-L. Zhu, X.-G. Wu, and S. J. Brodsky, *Phys. Lett. B* **748**, 422 (2015).
 - [17] H.-n. Li and S. Mishima, *Phys. Rev. D* **83**, 034023 (2011).
 - [18] Y.-F. Li and X.-Q. Yu, *Phys. Rev. D* **95**, 034023 (2017).
 - [19] H.-Y. Cheng, C.-W. Chiang, and A.-L. Kuo, *Phys. Rev. D* **91**, 014011 (2015).
 - [20] H.-n. Li, S. Mishima, and A. I. Sanda, *Phys. Rev. D* **72**, 114005 (2005).
 - [21] S. Baek, C.-W. Chiang, and D. London, *Phys. Lett. B* **675**, 59 (2009).
 - [22] V. Barger, C.-W. Chiang, P. Langacker, and H.-S. Lee, *Phys. Lett. B* **598**, 218 (2004).
 - [23] Throughout this paper, charge-conjugate modes are implicitly included unless noted otherwise.
 - [24] F. Abudinén *et al.* (Belle II Collaboration), *Chin. Phys. C* **44**, 021001 (2020).
 - [25] K. Akai, K. Furukawa, and H. Koiso (SuperKEKB Collaboration), *Nucl. Instrum. Methods* **907**, 188 (2018).
 - [26] T. Abe *et al.* (Belle II Collaboration), *arXiv:1011.0352*.
 - [27] B. Wang (Belle II iTOP Group), in *Proceedings of the Meeting of the APS Division of Particles and Fields* (2017), *arXiv:1709.09938*.
 - [28] S. Agostinelli *et al.* (GEANT4 Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **506**, 250 (2003).
 - [29] D. J. Lange, *Nucl. Instrum. Methods Phys. Res., Sect. A* **462**, 152 (2001).
 - [30] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *Comput. Phys. Commun.* **191**, 159 (2015).
 - [31] S. Jadach, B. F. L. Ward, and Z. Was, *Comput. Phys. Commun.* **130**, 260 (2000).
 - [32] S. Jadach, Z. Was, R. Decker, and J. H. Kuhn, *Comput. Phys. Commun.* **76**, 361 (1993).
 - [33] T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth, and N. Braun (Belle II Framework Software Group), *Comput. Softw. Big Sci.* **3**, 1 (2019).

- [34] Belle II Collaboration, Belle II Analysis Software Framework (basf2), [10.5281/zenodo.5574115](https://doi.org/10.5281/zenodo.5574115).
- [35] T. Keck, *Comput. Software Big Sci.* **1**, 2 (2017).
- [36] The Fox-Wolfram moments were introduced in G.C. Fox and S. Wolfram, *Phys. Rev. Lett.* **41**, 1581 (1978); The modified moments used in this paper are described in K. Abe *et al.* (Belle Collaboration), *Phys. Rev. Lett.* **87**, 101801 (2001); *Phys. Lett. B* **511**, 151 (2001).
- [37] J.D. Bjorken and S.J. Brodsky, *Phys. Rev. D* **1**, 1416 (1970).
- [38] E. Farhi, *Phys. Rev. Lett.* **39**, 1587 (1977).
- [39] F. Abudinén *et al.* (Belle II Collaboration), *Eur. Phys. J. C* **82**, 283 (2022).
- [40] T. Bloomfield *et al.* (Belle Collaboration), *Phys. Rev. D* **105**, 072007 (2022).
- [41] S. Choudhury *et al.* (Belle Collaboration), *Phys. Rev. D* **107**, L031102 (2023).
- [42] T. Skwarnicki, A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances, Ph.D. thesis, Cracow, INP, 1986.

Correction: The previously published Fig. 2 contained plots that were positioned incorrectly and has been replaced. A missing sign in an inline equation located in the abstract, Eq. (8), and the first sentence of the penultimate paragraph has been fixed.