PHYSICAL REVIEW D **93,** 031101(R) (2016)

First observation of the decay $B^0 \to \psi(2S)\pi^0$

V. Chobanova,[37] J. Dalseno,[37,63] C. Kiesling,[37] A. Abdesselam,[61] I. Adachi,[14,11] H. Aihara,[68] D. M. Asner,[51] T. Aushev,[40]
R. Ayad,[61] V. Babu,[62] I. Badhrees,[61,28] S. Bahinipati,[16] A. M. Bakich,[60] E. Barberio,[38] P. Behera,[18] V. Bhardwaj,[58]
B. Bhuyan,[17] J. Biswal,[24] A. Bobrov,[3,49] A. Bozek,[47] M. Bračko,[36,24] T. E. Browder,[13] D. Červenkov,[4] V. Chekelian,[37]
A. Chen,[44] B. G. Cheon,[12] R. Chistov,[39] K. Cho,[29] Y. Choi,[59] D. Cinabro,[74] N. Dash,[16] Z. Doležal,[4] Z. Drásal,[4]
A. Drutskoy,[39] S. Eidelman,[3,49] H. Farhat,[74] J. E. Fast,[51] T. Ferber,[7] B. G. Fulsom,[51] V. Gaur,[62] N. Gabyshev,[3,49]
A. Garmash,[3,49] R. Gillard,[74] Y. M. Goh,[12] P. Goldenzweig,[26] B. Golob,[33,24] O. Grzymkowska,[47] J. Haba,[14,11] T. Hara,[14,11]
K. Hayasaka,[42] H. Hayashii,[43] W.-S. Hou,[46] T. Iijima,[42,41] K. Inami,[41] A. Ishikawa,[66] Y. Iwasaki,[14] I. Jaegle,[13] H. B. Jeon,[31]
D. Joffe,[27] K. K. Joo,[5] T. Julius,[38] E. Kato,[66] P. Katrenko,[40] T. Kawasaki,[48] D. Y. Kim,[57] H. J. Kim,[31] J. B. Kim,[30] K.
T. Kim,[30] M. J. Kim,[31] S. H. Kim,[12] Y. J. Kim,[29] K. Kinoshita,[6] P. Kodyš,[4] S. Korpar,[36,24] P. Križan,[33,24] P. Krokovny,[3,49]
T. Kuhr,[34] R. Kumar,[54] T. Kumita,[70] A. Kuzmin,[3,49] Y.-J. Kwon,[76] I. S. Lee,[12] H. Li,[19] L. Li,[55] Y. Li,[73] L. Li Gioi,[37] J. Libby,[18]
D. Liventsev,[73,14] M. Masuda,[67] D. Matvienko,[3,49] K. Miyabayashi,[43] H. Miyata,[48] R. Mizuk,[39,40] G. B. Mohanty,[62]
S. Mohanty,[62,72] A. Moll,[37,63] H. K. Moon,[30] T. Mori,[41] R. Mussa,[23] E. Nakano,[50] M. Nakao,[14,11] T. Nanut,[24] Z. Natkaniec,[47]
M. Nayak,[18] E. Nedelkovska,[37] N. K. Nisar,[62,1] S. Nishida,[14,11] S. Ogawa,[65] P. Pakhlov,[39] G. Pakhlova,[40] B. Pal,[6] C.
W. Park,[59] H. Park,[31] S. Paul,[64] T. K. Pedlar,[35] R. Pestotnik,[24] M. Petrič,[24] L. E. Piilonen,[73] C. Pulvermacher,[26] J. Rauch,[64]
E. Ribežl,[24] M. Ritter,[34] S. Ryu,[56] H. Sahoo,[13] Y. Sakai,[14,11] S. Sandilya,[62] T. Sanuki,[66] V. Savinov,[53] T. Schlüter,[34]
O. Schneider,[32] G. Schnell,[2,15] C. Schwanda,[21] A. J. Schwartz,[6] Y. Seino,[48] K. Senyo,[75] O. Seon,[41] M. E. Sevior,[38]
V. Shebalin,[3,49] T.-A. Shibata,[69] J.-G. Shiu,[46] B. Shwartz,[3,49] F. Simon,[37,63] J. B. Singh,[52] Y.-S. Sohn,[76] A. Sokolov,[22]
E. Solovieva,[40] M. Starič,[24] J. Stypula,[47] M. Sumihama,[10] K. Sumisawa,[14,11] T. Sumiyoshi,[70] U. Tamponi,[23,71]
Y. Teramoto,[50] K. Trabelsi,[14,11] M. Uchida,[69] S. Uehara,[14,11] T. Uglov,[40] Y. Unno,[12] S. Uno,[14,11] P. Urquijo,[38] Y. Usov,[3,49]
C. Van Hulse,[2] P. Vanhoefer,[37] G. Varner,[13] A. Vinokurova,[3,49] V. Vorobyev,[3,49] M. N. Wagner,[9] C. H. Wang,[45]
M.-Z. Wang,[46] P. Wang,[20] X. L. Wang,[73] M. Watanabe,[48] Y. Watanabe,[25] K. M. Williams,[73] E. Won,[30] J. Yamaoka,[51]
S. Yashchenko,[7] H. Ye,[7] J. Yelton,[8] C. Z. Yuan,[20] Y. Yusa,[48] Z. P. Zhang,[55] V. Zhilich,[3,49] V. Zhulanov,[3,49] and A. Zupanc[33,24]
(Belle Collaboration)

[1]*Aligarh Muslim University, Aligarh 202002*
[2]*University of the Basque Country UPV/EHU, 48080 Bilbao*
[3]*Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090*
[4]*Faculty of Mathematics and Physics, Charles University, 121 16 Prague*
[5]*Chonnam National University, Kwangju 660-701*
[6]*University of Cincinnati, Cincinnati, Ohio 45221*
[7]*Deutsches Elektronen–Synchrotron, 22607 Hamburg*
[8]*University of Florida, Gainesville, Florida 32611*
[9]*Justus-Liebig-Universität Gießen, 35392 Gießen*
[10]*Gifu University, Gifu 501-1193*
[11]*SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193*
[12]*Hanyang University, Seoul 133-791*
[13]*University of Hawaii, Honolulu, Hawaii 96822*
[14]*High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[15]*IKERBASQUE, Basque Foundation for Science, 48013 Bilbao*
[16]*Indian Institute of Technology Bhubaneswar, Satya Nagar 751007*
[17]*Indian Institute of Technology Guwahati, Assam 781039*
[18]*Indian Institute of Technology Madras, Chennai 600036*
[19]*Indiana University, Bloomington, Indiana 47408*
[20]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049*
[21]*Institute of High Energy Physics, Vienna 1050*
[22]*Institute for High Energy Physics, Protvino 142281*
[23]*INFN—Sezione di Torino, 10125 Torino*
[24]*J. Stefan Institute, 1000 Ljubljana*
[25]*Kanagawa University, Yokohama 221-8686*
[26]*Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe*
[27]*Kennesaw State University, Kennesaw GA 30144*
[28]*King Abdulaziz City for Science and Technology, Riyadh 11442*
[29]*Korea Institute of Science and Technology Information, Daejeon 305-806*
[30]*Korea University, Seoul 136-713*
[31]*Kyungpook National University, Daegu 702-701*

2470-0010/2016/93(3)/031101(7)

[32]*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015*
[33]*Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana*
[34]*Ludwig Maximilians University, 80539 Munich*
[35]*Luther College, Decorah, Iowa 52101*
[36]*University of Maribor, 2000 Maribor*
[37]*Max-Planck-Institut für Physik, 80805 München*
[38]*School of Physics, University of Melbourne, Victoria 3010*
[39]*Moscow Physical Engineering Institute, Moscow 115409*
[40]*Moscow Institute of Physics and Technology, Moscow Region 141700*
[41]*Graduate School of Science, Nagoya University, Nagoya 464-8602*
[42]*Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602*
[43]*Nara Women's University, Nara 630-8506*
[44]*National Central University, Chung-li 32054*
[45]*National United University, Miao Li 36003*
[46]*Department of Physics, National Taiwan University, Taipei 10617*
[47]*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
[48]*Niigata University, Niigata 950-2181*
[49]*Novosibirsk State University, Novosibirsk 630090*
[50]*Osaka City University, Osaka 558-8585*
[51]*Pacific Northwest National Laboratory, Richland, Washington 99352*
[52]*Panjab University, Chandigarh 160014*
[53]*University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
[54]*Punjab Agricultural University, Ludhiana 141004*
[55]*University of Science and Technology of China, Hefei 230026*
[56]*Seoul National University, Seoul 151-742*
[57]*Soongsil University, Seoul 156-743*
[58]*University of South Carolina, Columbia, South Carolina 29208*
[59]*Sungkyunkwan University, Suwon 440-746*
[60]*School of Physics, University of Sydney, NSW 2006*
[61]*Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451*
[62]*Tata Institute of Fundamental Research, Mumbai 400005*
[63]*Excellence Cluster Universe, Technische Universität München, 85748 Garching*
[64]*Department of Physics, Technische Universität München, 85748 Garching*
[65]*Toho University, Funabashi 274-8510*
[66]*Department of Physics, Tohoku University, Sendai 980-8578*
[67]*Earthquake Research Institute, University of Tokyo, Tokyo 113-0032*
[68]*Department of Physics, University of Tokyo, Tokyo 113-0033*
[69]*Tokyo Institute of Technology, Tokyo 152-8550*
[70]*Tokyo Metropolitan University, Tokyo 192-0397*
[71]*University of Torino, 10124 Torino*
[72]*Utkal University, Bhubaneswar 751004*
[73]*CNP, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*
[74]*Wayne State University, Detroit, Michigan 48202*
[75]*Yamagata University, Yamagata 990-8560*
[76]*Yonsei University, Seoul 120-749*

(Received 21 December 2015; published 23 February 2016)

We report a measurement of the $B^0 \to \psi(2S)\pi^0$ branching fraction based on the full $\Upsilon(4S)$ data set of 772×10^6 $B\bar{B}$ pairs collected by the Belle detector at the KEKB asymmetric-energy e^+e^- collider. We obtain $\mathcal{B}(B^0 \to \psi(2S)\pi^0) = (1.17 \pm 0.17(\mathrm{stat}) \pm 0.08(\mathrm{syst})) \times 10^{-5}$. The result has a significance of 7.2 standard deviations and is the first observation of the decay $B^0 \to \psi(2S)\pi^0$.

DOI: 10.1103/PhysRevD.93.031101

Violation of the combined charge-parity symmetry (*CP* violation) in the Standard Model (SM) arises from a single irreducible phase in the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix [1,2]. A primary objective of the Belle experiment is to overconstrain the unitarity triangle of the CKM matrix related to $B_{u,d}$ decays. This permits a precision test of the CKM mechanism for *CP* violation as well as the search for effects beyond the SM. Mixing-induced *CP* violation in the B sector has been clearly established by the Belle [3] and *BABAR* [4] collaborations in the $b \to c\bar{c}s$-induced decays $B^0 \to (c\bar{c})^0 K^0$.

While these decays allow access to the CP-violating angle $\phi_1 \equiv \arg(-V_{cd}V^*_{cb})/(V_{td}V^*_{tb})$ at first order (tree), its value is prone to distortion from suppressed higher-order loop-induced (penguin) amplitudes containing different weak phases. Applying SU(3) symmetry arguments, the related $b \to c\bar{c}d$-induced channels $B^0 \to (c\bar{c})^0\pi^0$ can be used to quantify the shift in ϕ_1 caused by these loop contributions [5]. Thus, this $b \to c\bar{c}d$ decay is a promising place to search for new physics effects [6]. This paper establishes the $B^0 \to \psi(2S)\pi^0$ channel, which may be used to constrain the penguin contamination in $B^0 \to \psi(2S)K^0$ in a future measurement of its time-dependent CP asymmetry.

The result presented in this paper is based on the final $\Upsilon(4S)$ 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 [7]. At the $\Upsilon(4S)$ resonance, corresponding to a center-of-mass energy $\sqrt{s} = 10.58$ GeV, the $B\bar{B}$ pairs are produced with a Lorentz boost $\beta\gamma = 0.425$ nearly 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 (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) comprising CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return yoke located outside of the coil is instrumented to detect K^0_L mesons and to identify muons (KLM). The detector is described in detail elsewhere [8]. Two inner detector configurations were used: a 2.0 cm radius beampipe and a three-layer silicon vertex detector (SVD1) were used for the first sample of 152×10^6 $B\bar{B}$ pairs, while a 1.5 cm radius beampipe, a four-layer silicon vertex detector (SVD2), and a small-cell inner drift chamber were used to record the remaining 620×10^6 $B\bar{B}$ pairs [9]. Simulated B decay Monte Carlo (MC) events are generated by EvtGen[10], in which final-state radiation is described with PHOTOS [11]. We use the GEANT3 [12] toolkit to model the interaction of the generated particles with the detector and its response in order to determine the detector acceptance.

We reconstruct $\psi(2S)$ candidates in the $\ell^+\ell^-$ decay channels ($\ell = e, \mu$), referred to as leptonic hereinafter, and the $J/\psi\pi^+\pi^-$ decay channel, referred to as hadronic. All charged tracks are identified using a loose requirement on the distance of closest approach with respect to the interaction point along the beam direction of under 5.0 cm and in the transverse plane of under 1.5 cm. The J/ψ candidates are reconstructed from $\ell^+\ell^-$ pairs. Electron tracks are identified by a combination of dE/dx in the CDC, shower shape and position in the ECL, light yield in the ACC, and E/p, where E is the energy deposited in the ECL and p is the momentum measured by the SVD and the CDC. To account for radiative energy losses in the e^+e^- decays, we include the bremsstrahlung photons (γ) that are in a cone with an opening angle of 50 mrad around the e^+ (e^-) tracks [so that the reconstructed J/ψ or $\psi(2S)$ candidate is denoted as $e^+e^-(\gamma)$]. For muon tracks, the identification is based on track penetration depth and hit scatter in the KLM.

We impose asymmetric requirements on the J/ψ and $\psi(2S)$ masses due to energy leakage in the ECL and bremsstrahlung. The invariant masses of the J/ψ candidates must fulfill $M_{e^+e^-(\gamma)} - m_{J/\psi} \in (-0.150, +0.036)$ GeV$/c^2$ or $M_{\mu^+\mu^-} - m_{J/\psi} \in (-0.060, +0.036)$ GeV$/c^2$, where $m_{J/\psi}$ denotes the world-average J/ψ mass [13], and $M_{e^+e^-(\gamma)}$ and $M_{\mu^+\mu^-}$ are the reconstructed invariant masses of the $e^+e^-(\gamma)$ and $\mu^+\mu^-$ candidates, respectively. For the $\psi(2S)$, the invariant masses must fulfill $M_{e^+e^-(\gamma)} - m_{\psi(2S)} \in (-0.150, +0.036)$ GeV$/c^2$ or $M_{\mu^+\mu^-} - m_{\psi(2S)} \in (-0.060, +0.036)$ GeV$/c^2$, where $m_{\psi(2S)}$ denotes the world-average $\psi(2S)$ mass [13]. For the $\psi(2S) \to J/\psi\pi^+\pi^-$ candidates, $\Delta M \equiv M_{\ell^+\ell^-(\gamma)\pi^+\pi^-} - M_{\ell^+\ell^-(\gamma)}$ must fulfill $\Delta M \in (0.580, 0.600)$ GeV$/c^2$. To reduce background particle combinations in this channel, we select $\pi^+\pi^-$ pairs with an invariant mass above a loose threshold of 400 MeV$/c^2$. Using information obtained from the CDC, ACC, and TOF, these pion candidates are also required to be inconsistent with the kaon mass hypothesis. This requirement retains 99.8% of the pion candidates, while 5% of kaons are falsely identified as pions. To improve the B-meson mass resolution, we apply a vertex- and mass-constrained kinematic fit to the J/ψ and $\psi(2S)$ candidates. We assign each candidate its nominal mass and require that its charged daughters originate from the same vertex.

Photons are identified as isolated ECL clusters that are not matched to any charged particle track. To suppress combinatorial background, the photons are required to have energies above 50 MeV if in the ECL barrel or above 100 MeV if in the ECL endcaps, where the barrel region covers the polar angle range $32° < \theta < 130°$ and the endcap regions cover the polar angle ranges $12° < \theta < 32°$ and $130° < \theta < 157°$. Two γ candidates are combined to form a π^0 candidate that must satisfy $M_{\gamma\gamma} - m_{\pi^0} \in (-17, 15)$ MeV$/c^2$, where m_{π^0} is the world-average mass of the π^0 [13]. This corresponds to about 3 times the experimental resolution. The four-momenta of retained candidates are then adjusted in a mass-constrained fit wherein the parent mass is constrained to m_{π^0}.

We combine the $\psi(2S)$ and π^0 to form a neutral B meson. The B candidates are identified using two kinematic variables: a modified beam-energy-constrained mass,

$$M'_{\mathrm{bc}} \equiv \sqrt{(E_{\mathrm{beam}})^2 - \left|\vec{p}_{\psi(2S)} + \sqrt{(E_{\mathrm{beam}} - E_{\psi(2S)})^2 - m^2_{\pi^0}}\,\frac{\vec{p}_{\pi^0}}{|\vec{p}_{\pi^0}|}\right|^2}, \tag{1}$$

and the energy difference $\Delta E \equiv E_B - E_{\mathrm{beam}}$, where $\vec{p}$ denotes the three-momentum and E_{beam} the beam energy, all evaluated in the $\Upsilon(4S)$ center-of-mass system. This definition of M'_{bc} is preferred over the standard form used at the B factories as it exhibits a lower correlation with ΔE when π^0 is present in the final state.

A significant background arises from $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) continuum events. To suppress it, we construct the ratio of second- to zeroth-order Fox-Wolfram moments [14], $R_2 = H_2/H_0$, which ranges between zero (spherical) and one (jet-like). A loose requirement of less than 0.5 is applied. This removes around 50% of all continuum background with a negligible loss of signal efficiency.

On average, 1.13 B^0 candidates are reconstructed per event and 11.6% of all events have more than one candidate. In a multicandidate event, we choose the B^0 with the lowest $\chi^2_{\mathrm{mass}} \equiv (M_{\mathrm{Rec}} - m)^2/\sigma^2_{\mathrm{Rec}}$ per daughter particle with a reconstructed mass M_{Rec}, a nominal mass m and a mass resolution σ_{Rec}. For the leptonic channels, $\chi^2_{\mathrm{mass}} \equiv (\chi^2_{\psi(2S)} + \chi^2_{\pi^0})/2$. For the hadronic channels, $\chi^2_{\mathrm{mass}} \equiv (\chi^2_{J/\psi} + \chi^2_{\Delta m} + \chi^2_{\pi^0})/3$, where $\chi^2_{\Delta m}$ is defined similarly except that the reconstructed and nominal mass differences between $\psi(2S)$ and J/ψ are used in place of M_{Rec} and m, respectively. According to MC simulation, this procedure has a 75% success rate when more than one B candidate is reconstructed and the correct B is in the list. After this best-candidate selection, the detection efficiency, including a correction for the difference between data and MC in the particle identification and including the daughter branching fraction uncertainties and the selection criteria uncertainties, is $(0.43 \pm 0.02)\%$ for the leptonic channels and $(0.52 \pm 0.02)\%$ for the hadronic. Approximately 0.5% (10%) of the signal candidates are misreconstructed in the leptonic (hadronic) channels.

The $B^0 \to \psi(2S)\pi^0$ branching fraction, $\mathcal{B}(B^0 \to \psi(2S)\pi^0)$, is extracted from an unbinned extended maximum likelihood fit to M'_{bc} and ΔE. The following categories are considered in the event model: correctly reconstructed signal, misreconstructed signal, other $b \to (c\bar{c})q$ transitions, and combinatorial background. Unless otherwise stated, the probability density function (PDF) is the product of PDFs for each observable, $\mathcal{P}^m_c(M'_{\mathrm{bc}}, \Delta E) \equiv \mathcal{P}^m_c(M'_{\mathrm{bc}}) \times \mathcal{P}^m_c(\Delta E)$, in each $\psi(2S)$ decay mode, m, and in each category, c.

We study the distributions of both signal components—correctly reconstructed and misreconstructed—using an MC sample that contains only $B^0 \to \psi(2S)\pi^0$ events. We define a correctly reconstructed event as one in which all charged tracks are correctly associated with the signal B meson. For such events, we find the distributions of the fit observables in the $\psi(2S) \to e^+e^-$ and $\psi(2S) \to J/\psi[e^+e^-]\pi^+\pi^-$ decay channels to be similar. The distributions in the $\psi(2S) \to \mu^+\mu^-$ and $\psi(2S) \to J/\psi[\mu^+\mu^-]\pi^+\pi^-$ decay modes are also alike. Thus, we divide the signal MC into an electron and a muon component and model these separately. The M'_{bc} PDF for both modes consists of a Crystal Ball (CB) function [15], $\mathcal{C}$, combined with an ARGUS distribution [16], $\mathcal{A}$, which additionally accounts for the tail towards lower M'_{bc} values due to the photon and electron energy leakage in the ECL. Due to a correlation between M'_{bc} and ΔE, we parametrize the M'_{bc} PDF in terms of ΔE,

$$\begin{aligned}\mathcal{P}^m_{\mathrm{Sig}}(M'_{\mathrm{bc}}|\Delta E) \equiv{}& (f^m + \rho^m_1 \Delta E^2)\mathcal{C}(M'_{\mathrm{bc}}; \alpha^m_{M'_{\mathrm{bc}}}, n^m_{M'_{\mathrm{bc}}}, \mu^m_{M'_{\mathrm{bc}}} \\ &+ \mu^{\mathrm{CF}}_{M'_{\mathrm{bc}}}, \sigma^m_{M'_{\mathrm{bc}}}\sigma^{\mathrm{CF}}_{M'_{\mathrm{bc}}} + \rho^m_2 g^m(\Delta E)) \\ &+ (1 - [f^m + \rho^m_1 \Delta E^2])\mathcal{A}(M'_{\mathrm{bc}}; a^m),\end{aligned} \tag{2}$$

where $\alpha^m_{M'_{\mathrm{bc}}}$, $n^m_{M'_{\mathrm{bc}}}$, $\mu^m_{M'_{\mathrm{bc}}}$, $\sigma^m_{M'_{\mathrm{bc}}}$ and a^m are parameters obtained from MC, while $\mu^{\mathrm{CF}}_{M'_{\mathrm{bc}}}$ and $\sigma^{\mathrm{CF}}_{M'_{\mathrm{bc}}}$ are correction factors obtained from a $B^+ \to J/\psi K^{*+}$ control sample; ρ^m_1 and ρ^m_2 are correlation factors and $g^m(\Delta E)$ are functions in ΔE determined from MC: $g^{e^+e^-} = \Delta E^2$ for the electron component and $g^{\mu^+\mu^-} = |\Delta E|$ for the muon component. For both types of correctly reconstructed signal events, the ΔE PDF is the combination of a CB distribution and a sum of Chebyshev polynomials up to the first order,

$$\begin{aligned}\mathcal{P}^m_{\mathrm{Sig}}(\Delta E) \equiv{}& f^m \mathcal{C}(\Delta E; \alpha^m_{\Delta E}, n^m_{\Delta E}, \mu^m_{\Delta E} + \mu^{\mathrm{CF}}_{\Delta E}, \sigma^m_{\Delta E}\sigma^{\mathrm{CF}}_{\Delta E}) \\ &+ (1 - f^m)(1 + c^m \Delta E),\end{aligned} \tag{3}$$

where $\alpha^m_{\Delta E}$, $n^m_{\Delta E}$, $\mu^m_{\Delta E}$, $\sigma^m_{\Delta E}$ and c^m are obtained from MC, while $\mu^{\mathrm{CF}}_{\Delta E}$ and $\sigma^{\mathrm{CF}}_{\Delta E}$ are correction factors obtained from the control sample.

We omit the misreconstructed signal component in the leptonic decay modes due to its insignificant contribution. Each of the two hadronic modes is modeled with a separate two-dimensional histogram in M'_{bc}–ΔE.

The major background contribution originates from $b \to (c\bar{c})q$ decays other than the signal. We study this component from an MC sample containing all known $b \to (c\bar{c})q$ decays. Since the two leptonic channels have similar distributions, as do the two hadronic channels, we divide the $b \to (c\bar{c})q$ background events into a leptonic and a hadronic subsample. We model each of these with a two-dimensional M'_{bc}–ΔE histogram.

The rest of the background events are a mixture of $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) processes and B-meson decays into open charm and charmless final states. We refer to these as combinatorial background. We study their distributions from $\Upsilon(4S)$ data in the dilepton and ΔM sidebands. The J/ψ sideband is defined as $M_{\ell^+\ell^-} \in (2.60, 2.80)\cup(3.20, 3.40)$ GeV/c^2, the $\psi(2S)$ sideband as $M_{\ell^+\ell^-} \in (3.45, 3.53)\cup(3.80, 3.90)$ GeV/c^2, and the ΔM sideband as $\Delta M \in (0.49, 0.53)\cup(0.64, 0.68)$ GeV/c^2.

In all sidebands, the M'_{bc} PDF is an ARGUS distribution. In the leptonic sidebands, we model the ΔE combinatorial background distribution with a sum of Chebyshev polynomials up to the first order. The combinatorial ΔE PDF in the ΔM sideband is a sum of Chebyshev polynomials up to the second order. We verify that the models in the lower and upper sidebands are in agreement and thus the combined model provides a reliable description of the events in the signal region.

The total extended likelihood is given by

$$\mathcal{L} \equiv \prod_m \frac{e^{-\sum_c N_c^m} N^m}{N^m!} \prod_{i=1}^{N^m} \sum_c N_c \mathcal{P}_c^m(M'^{\,i}_{\mathrm{bc}}, \Delta E^i), \qquad (4)$$

where i indexes the events, c the categories and m the decay modes.

The $B^0 \to \psi(2S)\pi^0$ branching fraction is a free parameter in the fit to the data and is obtained by transforming the signal yields according to

$$N^m_{\mathrm{Sig}} = \mathcal{B}(B^0 \to \psi(2S)\pi^0) N_{B\bar{B}} \epsilon^m_{\mathrm{Sig}}, \qquad (5)$$

where $N_{B\bar{B}}$ is the number of $B\bar{B}$ pairs collected by the Belle detector and $\epsilon^m_{\mathrm{Sig}}$ is the detection efficiency, including daughter branching fractions for each subcategory. The misreconstructed-signal yields are fixed from MC relative to the two hadronic-mode signal yields. Only the muonic hadronic mode's yield is free in the $c\bar{c}$ background category, while the yields of the three remaining decay modes are fixed from MC relative to it. The four combinatorial-background yields are free.

We study the fit performance using pseudoexperiments in a linearity test covering the region of the expected branching fraction. There is no bias in experiments where the events are generated according to the total PDF. However, a bias at the level of 10% of the statistical error tending towards higher values is observed in experiments generated by selecting random events from the MC samples that have passed the full selection. This indicates that the bias is not due to a low signal yield but rather to imperfections in the modeling of correlations. We apply a fit correction of the full bias and consider half the correction as a systematic uncertainty.

The contribution of peaking background that originates from decays to the same final state as the signal is studied in the J/ψ, $\psi(2S)$ and Δm sidebands. We define the combinatorial background as nonpeaking in M'_{bc} and ΔE, while we assume that a potential peaking background has the same shape as the correctly reconstructed signal. Using the combinatorial background and the signal PDFs in a common fit to the sidebands, we extract two yields: one for the combinatorial background and the other for the peaking background. The peaking-background yield is consistent with zero for all modes except for the muonic signal mode in the ΔM sideband, where it has a statistical significance of 3.7σ. We extrapolate the expected peaking background yield into the signal region and subtract the obtained value from the signal yield obtained from the data.

We determine the M'_{bc} and ΔE signal model correction factors from a control sample with a similar decay topology, $B^+ \to J/\psi K^{*+}$, where the K^{*+} candidates are reconstructed from a K^+ and a π^0 candidate. To ensure a high momentum of the π^0, replicating the kinematic conditions of $B^0 \to \psi(2S)\pi^0$, we require the angle between the π^0 momentum vector and the vector opposite the B flight direction in the K^{*+} rest frame to be smaller than 1.5 rad. For the J/ψ and π^0 candidates, we use the same selection criteria as for the $B^0 \to \psi(2S)\pi^0$ mode. Only K^{*+} candidates fulfilling $M_{K^+\pi^0} \in (0.793, 0.990)$ GeV/c^2 are retained. Using a model similar to $B^0 \to \psi(2S)\pi^0$ for the control sample, we obtain a $B^+ \to J/\psi K^{*+}$ signal yield of 3681 ± 71 events and the signal correction factors from the fit to the data.

From the fit to the data containing 1090 $B^0 \to \psi(2S)\pi^0$ candidates, we obtain the bias-corrected branching fraction

$$\mathcal{B}(B^0 \to \psi(2S)\pi^0) = (1.17 \pm 0.17) \times 10^{-5}. \qquad (6)$$

The branching fraction corresponds to 85 ± 12 signal events, of which 38 ± 8 are leptonic and 47 ± 9 are hadronic, 628 ± 65 events originate from other $b \to (c\bar{c})q$ decays and 377 ± 103 events belong to the combinatorial background. All uncertainties here are statistical. Fit projections to the data are shown in Fig. 1.

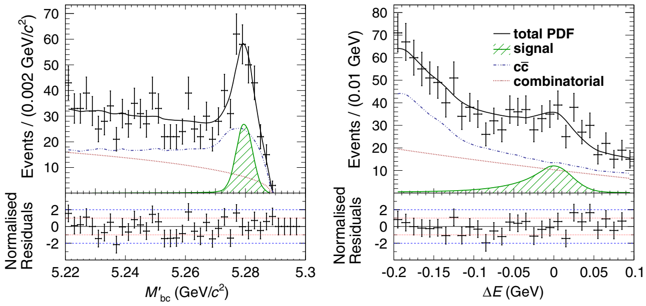

FIG. 1. Projections of the fit to the $B^0 \to \psi(2S)\pi^0$ data in the entire fit region onto M'_{bc} (left) and ΔE (right). Points with error bars represent the data and the solid black curves represent the fit results. Green hatched curves show the $B^0 \to \psi(2S)\pi^0$ signal component, blue dash-dotted curves show the $c\bar{c}$ background component, and red dotted curves indicate the combinatorial background.

TABLE I. Systematic uncertainties of the $B^0 \to \psi(2S)\pi^0$ branching fraction.

Category	$\delta\mathcal{B}(\psi(2S)\pi^0)$ [%]
$N_{B\bar{B}}$	1.4
π^0 reconstruction	4.0
$\mathcal{B}(\psi(2S) \to \ell^+\ell^-)$	3.0
$\mathcal{B}(\psi(2S) \to J/\psi\pi^+\pi^-)$	0.5
$\mathcal{B}(J/\psi \to \ell^+\ell^-)$	0.3
Electron ID	0.7
Muon ID	0.9
Hadron ID	1.3
Tracking	1.7
Misreconstruction	0.3
Parametric shape	0.9
Nonparametric shape	1.4
Peaking $b \to (c\bar{c})q$ background in $M'_{\rm bc}$	1.7
Peaking background in $M'_{\rm bc}$ and ΔE	2.2
Correction factors	0.9
Fit bias	0.6
Total	6.7

Systematic uncertainties from various sources are considered. They are estimated with both model-specific and -independent studies and cross-checks. The $\mathcal{B}(B^0 \to \psi(2S)\pi^0)$ systematic uncertainties are summarized in Table I.

The systematic uncertainty due to the error on the total number of $B\bar{B}$ pairs is calculated from the on- and off-resonance luminosity, taking into account the efficiency and luminosity scaling corrections [17]. The dominant systematic uncertainty arises from the π^0 reconstruction and is evaluated by comparing data-MC differences in the yield ratios between $\eta' \to \pi^0\pi^0\pi^0$ and $\eta' \to \pi^+\pi^-\pi^0$. We also consider the systematic uncertainties originating from the knowledge of the $\psi(2S)$ and J/ψ decay branching fractions used to calculate the efficiency. We apply the percentage error on their world averages [13] as a systematic uncertainty. The electron and muon identification efficiency uncertainties were obtained from separate Belle studies of the two-photon processes $e^+e^- \to e^+e^-\ell^+\ell^-$ and of $J/\psi \to \ell^+\ell^-$, where $\ell = e, \mu$. The uncertainty in the reconstruction efficiency due to the hadron identification is determined using $D^{*+} \to D^0[K^-\pi^+]\pi^+$ decays, where the hadron identity is unambiguously determined by its charge. The uncertainty due to the tracking efficiency is calculated by comparing data-MC differences in the reconstruction efficiencies of $D^{*\pm} \to D^0[K_S^0\{\pi^+\pi^-\}\pi^+\pi^-]\pi^\pm$. The hadron, electron and muon identification and tracking uncertainties are weighted by the reconstruction efficiencies of the corresponding B decay modes. The misreconstructed signal uncertainty is obtained by varying the misreconstructed fraction by ±20% of its value, which is a conservative estimate. The parametric and nonparametric shapes describing the background are varied within their uncertainties. For nonparametric shapes (i.e., histograms), we modify the histogram PDFs bin by bin according to a Poisson distribution and extract the branching fraction from a fit to the data. We perform 300 tests with such modified histogram PDFs and take the width of the resulting Gaussian branching-fraction distribution as a systematic uncertainty. We find that the decay $B^0 \to \psi(2S)K_S^0[\pi^0\pi^0]$ peaks in the signal region of $M'_{\rm bc}$. The $B^0 \to \psi(2S)K_S^0[\pi^0\pi^0]$ yield in the $b \to (c\bar{c})q$ background sample is varied by the uncertainty of its world-average branching fraction and the resulting difference in the $B^0 \to \psi(2S)\pi^0$ branching fraction is taken as a systematic uncertainty. The number of peaking background events obtained from the sideband study is varied by one standard deviation (σ), and the difference in the branching fraction is assigned as a systematic uncertainty. The same approach is used for the $M'_{\rm bc}$ and ΔE correction factors. Half the branching-fraction fit bias obtained from pseudoexperiments is taken as an additional systematic uncertainty. The total systematic uncertainty is 6.5% of the $B^0 \to \psi(2S)\pi^0$ branching fraction.

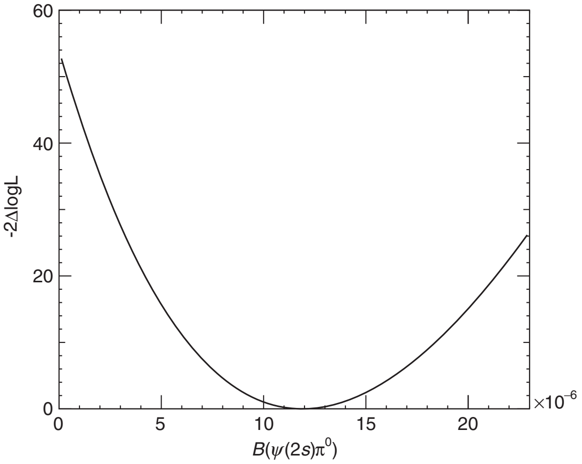

FIG. 2. $\mathcal{B}(B^0 \to \psi(2S)\pi^0)$ likelihood scan. The likelihood is convolved with an additive systematic uncertainty.

We perform a likelihood scan to obtain the statistical significance of our branching fraction measurement. We convolve the $\mathcal{L}$ distribution with a Gaussian with a zero mean and a width equal to the systematic uncertainty. The change in the $-2\log\mathcal{L}$ distribution as a function of the branching fraction is shown in Fig. 2. The statistical significance of 7.2σ is determined from $\sqrt{-2\Delta\log\mathcal{L}}$, where $\Delta\log\mathcal{L}$ is the likelihood difference between zero and the observed branching fraction. This includes the systematic uncertainties.

In summary, we report a measurement of the $B^0 \to \psi(2S)\pi^0$ branching fraction based on the full Belle data set collected at the $\Upsilon(4S)$ resonance. We obtain $\mathcal{B}(B^0 \to \psi(2S)\pi^0) = (1.17 \pm 0.17(\text{stat}) \pm 0.08(\text{syst})) \times 10^{-5}$. Our results are consistent with the naive expectation that the $B^0 \to \psi(2S)\pi^0$ to $B^0 \to \psi(2S)K_S^0$ branching fraction ratio should be similar to the $B^0 \to J/\psi\pi^0$ to $B^0 \to J/\psi K_S^0$ ratio. The $\mathcal{B}(B^0 \to \psi(2S)\pi^0)$ result has a significance of 7.2σ, which indicates the first observation of the decay $B^0 \to \psi(2S)\pi^0$.

ACKNOWLEDGMENTS

We thank the KEKB group for excellent operation of the accelerator; the KEK cryogenics group for efficient solenoid operations; and the KEK computer group, the NII, and PNNL/EMSL for valuable computing and SINET4 network support. We acknowledge support from MEXT, JSPS and Nagoya's TLPRC (Japan); ARC (Australia); FWF (Austria); NSFC and CCEPP (China); MSMT (Czechia); CZF, DFG, and VS (Germany); DST (India); INFN (Italy); MOE, MSIP, NRF, BK21Plus, WCU and RSRI (Korea); MNiSW and NCN (Poland); MES and RFAAE (Russia); ARRS (Slovenia); IKERBASQUE and UPV/EHU (Spain); SNSF (Switzerland); NSC and MOE (Taiwan); and DOE and NSF (USA).

[1] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).
[2] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[3] K. Abe *et al.* (Belle Collaboration), Phys. Rev. Lett. **87**, 091802 (2001); I. Adachi *et al.* (Belle Collaboration), Phys. Rev. Lett. **108**, 171802 (2012).
[4] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **87**, 091801 (2001); Phys. Rev. D **79**, 072009 (2009).
[5] M. Ciuchini, M. Pierini, and L. Silvestrini, Phys. Rev. Lett. **95**, 221804 (2005).
[6] Y. Grossman and M. Worah, Phys. Lett. B **395**, 241 (1997).
[7] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. A **499**, 1 (2003), and other papers included in this Volume; T. Abe *et al.*, Prog. Theor. Exp. Phys. (**2013**), 03A001. and following articles up to 03A011.
[8] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002); also see detector section in J. Brodzicka *et al.*, Prog. Theor. Exp. Phys. (**2012**) 04D001.
[9] Z. Natkaniec *et al.* (Belle SVD2 Group), Nucl. Instrum. Methods Phys. Res., Sect. A **560**, 1 (2006).
[10] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[11] P. Golonka and Z. Was, Eur. Phys. J. C **45**, 97 (2006).
[12] R. Brun *et al.*, GEANT 3.21, CERN DD/EE/84-1 (1984).
[13] K. A. Olive *et al.* (Particle Data Group), Chin. Phys. C **38**, 090001 (2014) and 2015 update.
[14] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581(1978).
[15] T. Skwarnicki, Ph.D. Thesis, Institute for Nuclear Physics, Krakow, 1986; DESY Internal Report No. DESY F31-86-02 (1986).
[16] H. Albrecht *et al.* (ARGUS Collaboration), Phys. Lett. B **241**, 278 (1990).
[17] A. J. Bevan *et al.* (*BABAR* and Belle Collaborations), Eur. Phys. J. C **74**, 3026 (2014).