Search for $B^{0}\to\pi^{-}\tau^{+}\nu_{\tau}$ with hadronic tagging at Belle

P. Hamer *et al.* (Belle Collaboration)

Phys. Rev. D **93**, 032007 — Published 22 February 2016

DOI: 10.1103/PhysRevD.93.032007

Search for $B^0 \to \pi^- \tau^+ \nu_\tau$ with hadronic tagging at Belle

P. Hamer,[9] A. Frey,[9] A. Abdesselam,[57] I. Adachi,[13,10] H. Aihara,[64] S. Al Said,[57,28] K. Arinstein,[3,49] D. M. Asner,[51] T. Aushev,[40] R. Ayad,[57] V. Babu,[58] I. Badhrees,[57,27] A. M. Bakich,[56] E. Barberio,[38] B. Bhuyan,[16] J. Biswal,[23] A. Bozek,[47] M. Bračko,[36,23] T. E. Browder,[12] D. Červenkov,[4] V. Chekelian,[37] A. Chen,[44] B. G. Cheon,[11] K. Chilikin,[39] K. Cho,[29] V. Chobanova,[37] Y. Choi,[55] D. Cinabro,[69] J. Dalseno,[37,59] M. Danilov,[39] I. Danko,[52] N. Dash,[15] J. Dingfelder,[2] Z. Doležal,[4] Z. Drásal,[4] A. Drutskoy,[39] D. Dutta,[58] S. Eidelman,[3,49] H. Farhat,[69] J. E. Fast,[51] T. Ferber,[7] B. G. Fulsom,[51] V. Gaur,[58] N. Gabyshev,[3,49] A. Garmash,[3,49] D. Getzkow,[8] R. Gillard,[69] R. Glattauer,[20] Y. M. Goh,[11] P. Goldenzweig,[25] B. Golob,[33,23] D. Greenwald,[60] J. Haba,[13,10] T. Hara,[13,10] J. Hasenbusch,[2] K. Hayasaka,[42] H. Hayashii,[43] W.-S. Hou,[46] T. Iijima,[42,41] K. Inami,[41] G. Inguglia,[7] A. Ishikawa,[62] R. Itoh,[13,10] Y. Iwasaki,[13] I. Jaegle,[12] H. B. Jeon,[31] D. Joffe,[26] K. K. Joo,[5] T. Julius,[38] K. H. Kang,[31] E. Kato,[62] T. Kawasaki,[48] C. Kiesling,[37] D. Y. Kim,[54] J. B. Kim,[30] J. H. Kim,[29] K. T. Kim,[30] M. J. Kim,[31] S. H. Kim,[11] Y. J. Kim,[29] K. Kinoshita,[6] P. Kodyš,[4] S. Korpar,[36,23] P. Križan,[33,23] P. Krokovny,[3,49] T. Kuhr,[34] A. Kuzmin,[3,49] Y.-J. Kwon,[71] J. S. Lange,[8] D. H. Lee,[30] I. S. Lee,[11] H. Li,[18] L. Li,[53] Y. Li,[68] J. Libby,[17] Y. Liu,[6] D. Liventsev,[68,13] P. Lukin,[3,49] M. Masuda,[63] D. Matvienko,[3,49] K. Miyabayashi,[43] H. Miyake,[13,10] H. Miyata,[48] R. Mizuk,[39,40] G. B. Mohanty,[58] S. Mohanty,[58,67] A. Moll,[37,59] H. K. Moon,[30] R. Mussa,[22] E. Nakano,[50] M. Nakao,[13,10] T. Nanut,[23] Z. Natkaniec,[47] M. Nayak,[17] N. K. Nisar,[58] S. Nishida,[13,10] S. Ogawa,[61] S. Okuno,[24] C. Oswald,[2] P. Pakhlov,[39] G. Pakhlova,[40] B. Pal,[6] H. Park,[31] T. K. Pedlar,[35] L. Pesántez,[2] R. Pestotnik,[23] M. Petrič,[23] L. E. Piilonen,[68] C. Pulvermacher,[25] J. Rauch,[60] E. Ribežl,[23] M. Ritter,[34] A. Rostomyan,[7] H. Sahoo,[12] Y. Sakai,[13,10] S. Sandilya,[58] L. Santelj,[13] T. Sanuki,[62] V. Savinov,[52] O. Schneider,[32] G. Schnell,[1,14] C. Schwanda,[20] Y. Seino,[48] K. Senyo,[70] M. Shapkin,[21] V. Shebalin,[3,49] T.-A. Shibata,[65] J.-G. Shiu,[46] A. Sibidanov,[56] F. Simon,[37,59] Y.-S. Sohn,[71] A. Sokolov,[21] E. Solovieva,[40] M. Starič,[23] J. Stypula,[47] T. Sumiyoshi,[66] Y. Teramoto,[50] K. Trabelsi,[13,10] V. Trusov,[25] M. Uchida,[65] S. Uehara,[13,10] Y. Unno,[11] S. Uno,[13,10] P. Urquijo,[38] Y. Usov,[3,49] C. Van Hulse,[1] P. Vanhoefer,[37] G. Varner,[12] V. Vorobyev,[3,49] A. Vossen,[18] C. H. Wang,[45] M.-Z. Wang,[46] P. Wang,[19] Y. Watanabe,[24] E. Won,[30] H. Yamamoto,[62] J. Yamaoka,[51] S. Yashchenko,[7] Y. Yook,[71] Z. P. Zhang,[53] V. Zhilich,[3,49] V. Zhulanov,[3,49] and A. Zupanc[23]

(The Belle Collaboration)

[1] *University of the Basque Country UPV/EHU, 48080 Bilbao*
[2] *University of Bonn, 53115 Bonn*
[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] *Justus-Liebig-Universität Gießen, 35392 Gießen*
[9] *II. Physikalisches Institut, Georg-August-Universität Göttingen, 37073 Göttingen*
[10] *SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193*
[11] *Hanyang University, Seoul 133-791*
[12] *University of Hawaii, Honolulu, Hawaii 96822*
[13] *High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[14] *IKERBASQUE, Basque Foundation for Science, 48013 Bilbao*
[15] *Indian Institute of Technology Bhubaneswar, Satya Nagar 751007*
[16] *Indian Institute of Technology Guwahati, Assam 781039*
[17] *Indian Institute of Technology Madras, Chennai 600036*
[18] *Indiana University, Bloomington, Indiana 47408*
[19] *Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049*
[20] *Institute of High Energy Physics, Vienna 1050*
[21] *Institute for High Energy Physics, Protvino 142281*
[22] *INFN - Sezione di Torino, 10125 Torino*
[23] *J. Stefan Institute, 1000 Ljubljana*
[24] *Kanagawa University, Yokohama 221-8686*
[25] *Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe*
[26] *Kennesaw State University, Kennesaw GA 30144*
[27] *King Abdulaziz City for Science and Technology, Riyadh 11442*
[28] *Department of Physics, Faculty of Science, King Abdulaziz University, Jeddah 21589*
[29] *Korea Institute of Science and Technology Information, Daejeon 305-806*
[30] *Korea University, Seoul 136-713*
[31] *Kyungpook National University, Daegu 702-701*

[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]University of Pittsburgh, Pittsburgh, Pennsylvania 15260
[53]University of Science and Technology of China, Hefei 230026
[54]Soongsil University, Seoul 156-743
[55]Sungkyunkwan University, Suwon 440-746
[56]School of Physics, University of Sydney, NSW 2006
[57]Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451
[58]Tata Institute of Fundamental Research, Mumbai 400005
[59]Excellence Cluster Universe, Technische Universität München, 85748 Garching
[60]Department of Physics, Technische Universität München, 85748 Garching
[61]Toho University, Funabashi 274-8510
[62]Tohoku University, Sendai 980-8578
[63]Earthquake Research Institute, University of Tokyo, Tokyo 113-0032
[64]Department of Physics, University of Tokyo, Tokyo 113-0033
[65]Tokyo Institute of Technology, Tokyo 152-8550
[66]Tokyo Metropolitan University, Tokyo 192-0397
[67]Utkal University, Bhubaneswar 751004
[68]CNP, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[69]Wayne State University, Detroit, Michigan 48202
[70]Yamagata University, Yamagata 990-8560
[71]Yonsei University, Seoul 120-749

We search for the process $B^0 \to \pi^- \tau^+ \nu_\tau$ using the full Belle data set of $711\,\mathrm{fb}^{-1}$, corresponding to $772 \times 10^6 B\bar{B}$ pairs, collected at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. We reconstruct one B meson in a hadronic decay and search for the $B^0 \to \pi^- \tau^+ \nu_\tau$ process in the remainder of the event. No significant signal is observed and an upper limit of $\mathcal{B}\left(B^0 \to \pi^- \tau^+ \nu_\tau\right) < 2.5 \times 10^{-4}$ is obtained at the 90% confidence level.

PACS numbers: 13.20.He,12.60.Fr

I. INTRODUCTION

The decay $B^0 \to \pi^- \tau^+ \nu_\tau$ [1] is mediated by the W^+ boson via the $\bar{b} \to \bar{u}$ transition. The transition amplitude is described by [2]

$$\left\langle \pi^- | u \gamma_\mu \bar{b} | B^0 \right\rangle = f^+(q^2) \left[2p_\mu + \left(\frac{1 - m_B^2 - m_\pi^2}{q^2} \right) q_\mu \right] + f^0(q^2) \frac{m_B^2 - m_\pi^2}{q^2} q_\mu, \tag{1}$$

with p and q being the momentum transfers to the pion and lepton pair, respectively.

The form factors f^+ and f^0 can be computed from QCD light-cone sum rules [2, 3] for $q^2 < 16\ \mathrm{GeV}^2/c^4$ and lattice QCD [4–6] for $q^2 > 16\ \mathrm{GeV}^2/c^4$. Various parametrizations exist to interpolate between the two regions. In this study, we use the parametrization introduced by Bourrely, Caprini, and Lellouch (BCL) [7], which can describe both form factors in $m_\tau^2 \leq q^2 \leq (m_B - m_\pi)^2$. The parameter values are taken from Ref. [2].

It has been stated [2, 8] that the differential ratio

$$\frac{\mathrm{d}\Gamma\left(B \to \pi \tau \nu_\tau\right)/\mathrm{d}q^2}{\mathrm{d}\Gamma\left(B \to \pi \ell \nu_\ell\right)/\mathrm{d}q^2}, \quad \ell = e,\ \mu \tag{2}$$

can be used as a test for the Standard Model (SM) as it depends solely on the ratio of the scalar and vector form factors f^0/f^+. The CKM matrix [9] element $|V_{ub}|$ enters both differential branching fractions and cancels in the ratio.

In new physics models like the two-Higgs-doublet model (2HDM) [10, 11], the decay $B^0 \to \pi^-\tau^+\nu_\tau$ can also be mediated by a charged Higgs boson. Possible contribution of a H^+ and other couplings in the 2HDM and MSSM [12, 13], which would affect the branching fraction and the differential ratio of branching fractions, have been evaluated in Refs. [2] and [14–17].

The decay $B^0 \to \pi^-\tau^+\nu_\tau$ has not been observed, nor has an upper limit on the branching fraction been obtained. Recent results [6] on the two form factors obtained from a joint fit to (2+1)-flavor lattice QCD calculations and $B \to \pi\ell\nu$ data from Belle [18, 19] and BaBar [20, 21] result in $\mathcal{B}\left(B^0 \to \pi^-\tau^+\nu_\tau\right)/\mathcal{B}\left(B^0 \to \pi^-\ell^+\nu_\ell\right) = 0.641(17)$ and $\mathcal{B}\left(B^0 \to \pi^-\tau^+\nu_\tau\right) = 9.35(38) \times 10^{-5}$ [22].

The signal decay is reconstructed in the four one-prong decays of the τ lepton, $\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$ with $\ell = e$ or μ, $\tau^- \to \pi^-\bar{\nu}_\tau$, and $\tau^- \to \rho^-\nu_\tau$, corresponding to 72% of all τ decays [23]. The most powerful decay modes are the two aforementioned hadronic τ decays and the $\tau^- \to e^-\bar{\nu}_e\nu_\tau$ mode, while the $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$ decay mode does not improve the final expected significance. This is mainly due to low muon momenta in the signal decay and the resulting low muon identification efficiency. The result of this analysis is based on the three most powerful τ decay modes.

II. DATA SAMPLE

The search for $B^0 \to \pi^-\tau^+\nu_\tau$ described in this paper is performed on the full data sample collected with the Belle detector at the KEKB asymmetric-energy e^+e^- (3.5 on 8.0 GeV) collider [24], operating at the $\Upsilon(4S)$ resonance. The data sample consists of an integrated luminosity of 711 fb^{-1}, which corresponds to $(771.6 \pm 10.6) \times 10^6 B\bar{B}$ pairs.

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 (TOF) scintillation counters, and an electromagnetic calorimeter (ECL) comprised of 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_L^0 mesons and to identify muons (KLM). Two inner detector configurations were used. A 2.0 cm beampipe and a 3-layer SVD were used for the first sample of $152 \times 10^6 B\bar{B}$ pairs, while a 1.5 cm beampipe, a 4-layer silicon detector and a small-cell inner drift chamber were used to record the remaining $620 \times 10^6 B\bar{B}$ pairs [25]. The detector is described in detail in Ref. [26].

The study was performed as a blind analysis based on simulated data. Monte Carlo (MC) samples were generated with EvtGen [27] and the detector simulation was performed by GEANT3 [28]. Recorded beam background was added to the MC samples. The expected non-beam background is estimated using MC samples that describe all physics processes at Belle. A resonant $\Upsilon(4S)$ event at Belle produces a $B\bar{B}$ pair. Two samples of $b \to c$ decays for $B^0\bar{B}^0$ and B^+B^- events, respectively, each contain 10 times the integrated luminosity of the data sample. Semileptonic $b \to u$ decays are simulated in a sample containing 20 times the integrated luminosity. Rare $b \to s$ and other rare decays are described in another sample corresponding to 50 times the integrated luminosity of the data. Continuum $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) was generated with PYTHIA [29] and included in the analysis in an MC sample containing five times the integrated luminosity of the data sample. Additionally, a high statistics sample of $B^0 \to X_u\tau\nu$ containing 24×10^6 events was generated with a phase-space and ISGW2 [30] model.

The signal MC sample is generated using BCL results for the vector and scalar form factors [2]. A total of 84×10^6 $B^0\bar{B}^0$ events were generated with one meson decaying into the signal final state and the other decaying generically.

No constraints on the τ decay were applied. The signal MC sample corresponds to approximately 2000 times the expected $\mathcal{B}\left(B^0 \to \pi^-\tau^+\nu_\tau\right) = 9.35 \times 10^{-5}$.

III. EVENT SELECTION

The complete reconstruction of the B meson decay into the signal final state ($B_{\rm sig}$) is not possible due to the presence of at least two neutrinos. However, since the initial state of the e^+e^- collision is completely defined by the momenta of the colliding leptons, we can constrain the signal side by fully reconstructing the other B meson ($B_{\rm tag}$) in hadronic decay modes. Tracks and clusters in the event that are not assigned to the $B_{\rm tag}$ after the successful reconstruction are assumed to originate from $B_{\rm sig}$.

A. Tag side

This analysis uses the Belle hadronic full-reconstruction algorithm [31] based on the artificial Neural Network package NeuroBayes [32]. Neural networks were trained to reconstruct B^0 and B^+ candidates from a total of 1104 decay channels.

Additional event shape variables are added to suppress continuum events. B mesons of resonant events are nearly at rest in the center-of-mass (c.m.) frame, leading to a spherical distribution of their decay products. Continuum events, on the other hand, produce back-to-back jets due the large available kinetic energy. Useful observables that differentiate between the two event types are the thrust axis of the $B_{\rm tag}$ meson [33] and modified

Fox-Wolfram moments [34]. For this analysis, the thrust axis and the second modified Fox-Wolfram moment are included in the neural network for the full hadronic reconstruction. If the algorithm does not succeed in reconstructing a B^0 candidate, the event is discarded.

Differences in the full reconstruction efficiency between MC and data, depending on the network output and $B_{\rm tag}$ reconstruction channel, are observed [19] but depend on the tag-side reconstruction only; a correction factor is determined from charmed semileptonic decays of the signal-side B meson.

The beam-energy-constrained mass,

$$M_{\rm bc} = \sqrt{E_{\rm beam}^2 - (\vec{p}_{B_{\rm tag}} c)^2} \,/c^2,$$

is required to be greater than 5.27 GeV/c^2, where $E_{\rm beam}$ and $\vec{p}_{B_{\rm tag}}$ denote the beam energy and reconstructed three-momentum, respectively, of the $B_{\rm tag}$, evaluated in the e^+e^- c.m. frame. With this requirement and the correction factor applied, we estimate a reconstruction efficiency of 0.18% from the signal MC sample, which is in very good agreement with the reconstruction efficiency of B^0 mesons in the Belle data sample [31]. The neural network output, $o_{\rm tag}^{\rm cs} \in [0, 1]$, is a continuous variable whose high (low) values correspond to candidates which are likely (unlikely) to be a true B meson. It is used at a later selection stage, as described below. The distributions of $M_{\rm bc}$ and $\ln(o_{\rm tag}^{\rm cs})$ for the three reconstruction channels are shown in Fig. 1, where the green (solid) contribution shows the dominant charmed $B \to X_c$ background decays. No further requirements are applied to the $M_{\rm bc}$ distributions, while $M_{\rm bc} > 5.27$ GeV/c^2 is required for the $\ln(o_{\rm tag}^{\rm cs})$ distributions.

B. Signal side

Only one-prong decays of the τ lepton are considered in this search. For of a correctly reconstructed $B_{\rm tag}$, there should be exactly two remaining oppositely-charged tracks in the detector. Additionally, the event should contain undetected (missing) momentum. Since the initial state of the e^+e^- collisions is given by the four-momenta of the colliding leptons, the undetected momentum can be measured. The missing momentum is defined as

$$p_{\rm miss} = 2p_{\rm beam} - p_{B_{\rm tag}} - p_{B_{\rm vis}},$$

where $2p_{\rm beam} = p_{e^+} + p_{e^-}$ is twice the beam momentum and $B_{\rm vis}$ denotes the visible part of the $B_{\rm sig}$ meson. Tracks with low transverse momentum p_t can curl in the solenoidal field and be detected as two tracks with opposite charge. Any two tracks with $p_t < 275$ MeV/c with an angle between the momentum vectors, calculated at the points closest to the interaction point, below 15° and total momentum difference less than 100 MeV/c are therefore counted as a single track. We reduce the number of poor quality tracks by requiring that $|dr| < 2.0$ cm and $|dz| < 4.0$ cm, where $|dz|$ and $|dr|$ are the distances of closest approach of a track to the interaction point along the z-axis and in the transverse plane, respectively.

Electron identification [35] is performed by calculating a likelihood ratio using the matching of charged tracks with the shower position in the ECL, the shower shape, the ratio of the energy deposited in the ECL and the measured momentum, the energy loss dE/dx in the CDC, and the Cherenkov light production in the ACC. Muon identification [36] is also done by evaluating a likelihood ratio. Clusters in the KLM are matched to charged tracks by extrapolation. For matched tracks, the difference between expected and measured penetration depth and the transverse deviation of all KLM hits associated with the track are used in this likelihood ratio. For a charged track not identified as a lepton, a kaon veto is applied using a likelihood ratio that discriminates between kaons and pions [37]. The ratio is formed from the energy loss dE/dx in the CDC, flight time information from the TOF, and photon yield in the ACC. All remaining tracks are identified as a pion. Neutral pions are reconstructed from pairs of photons. The absolute difference between the invariant mass of the π^0 candidate and the nominal π^0 mass, normalized to its uncertainty, must be below 3.0. Photons are required to have energies in the laboratory frame greater than 50 MeV for the ECL barrel and 100(150) MeV for the forward (backward) endcap. Neutral pion candidates with at least one photon being used in the tag-side full reconstruction are discarded.

Events are required to have exactly two oppositely charged particles within the allowed impact parameter range, with one additional track allowed outside the range. At least one charged pion is required. If the event contains two charged pions and neutral pion candidates, we search for $\rho^\pm$ candidates. The charged pion with the lower momentum in the c.m. frame is combined with each neutral pion candidate and a mass vertex fit is performed. A pair that can be successfully fitted with $\chi^2 < 20$ is accepted as a charged $\rho^\pm$ meson if its invariant mass is between 625 and 925 MeV/c^2. If multiple candidates are found, the candidate with a mass closest to the nominal ρ^+ mass [23] is selected. Due to the broad $\rho^\pm$ mass range, not all $\tau^- \to \rho^-\nu_\tau$ events are correctly reconstructed. These events contain two oppositely charged pions and a neutral pion in the final state and are miscategorized in the $\tau^- \to \pi^-\nu_\tau$ channel. Each event is reconstructed in one of the four reconstruction channels. In many $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$ events, the momentum of the muon is too low to reach the KLM and thus is not identified as a muon. In most of these cases, the muon is identified as a pion so that the event is placed in the $\tau^- \to \pi^-\nu_\tau$ sample.

Since K_L mesons do not completely deposit their energy in the detector, charmed B decays with subsequent decays $D \to K_L\pi$ or $D \to K_L\ell\nu_\ell$ have the signal's missing-momentum signature. A K_L candidate is identified as a cluster in the KLM without an associated

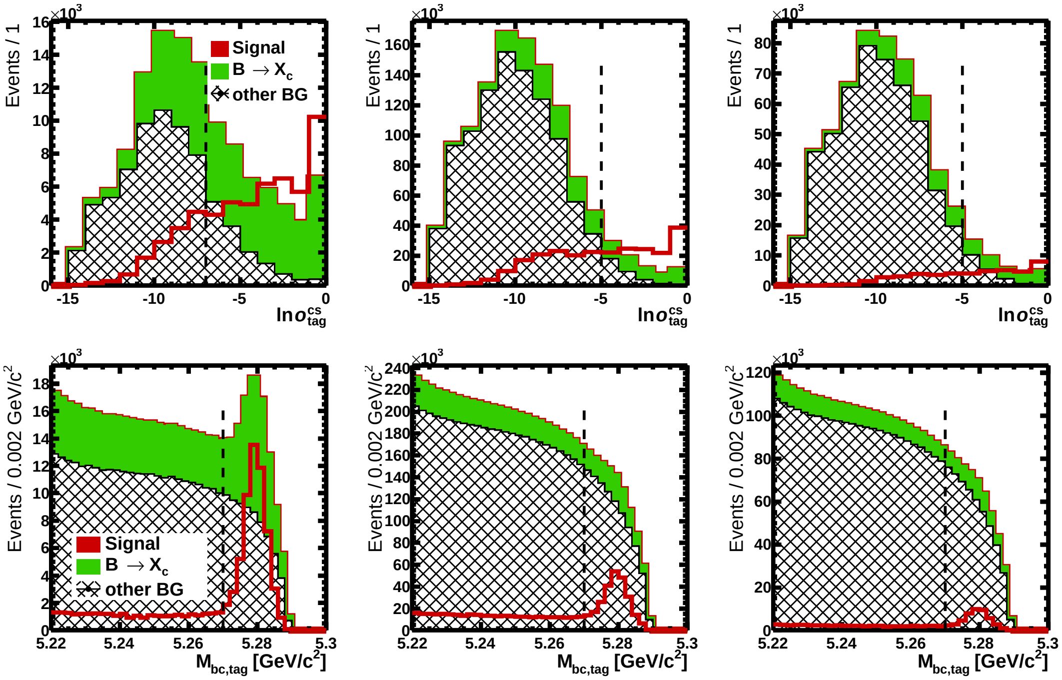


FIG. 1: Distributions of $\ln(o^{\rm cs}_{\rm tag})$(top) and $M_{\rm bc}$ (bottom) of the tag side candidate for the three reconstruction channels $\tau \to e$ (left), $\tau \to \pi$ (middle) and $\tau \to \rho$ (right) for signal and background, produced from Belle simulation. A requirement of $M_{\rm bc} > 5.27$ GeV/c^2 is applied in the left plot but no further requirements are applied other than successful reconstruction into one of the three modes. The histograms are produced from MC samples, normalized to the data sample size. The signal histogram is scaled by an arbitrary factor of 2000 for better visibility. The dashed vertical line indicates the minimal value required in the final selection; see Section III E.

charged track. An ECL cluster without an associated charged track is associated with the K_L cluster in the KLM if it lies along the flight path extrapolated from the interaction point to the KLM cluster. As described below, the extra energy in the ECL is used to determine the signal yield. Therefore, only events with a K_L without energy deposition in the ECL are vetoed.

C. Extra energy

We extract the signal yield from a fit to the distribution of the energy deposited in the ECL ($E_{\rm ECL}$) by particles not used in the full reconstruction or by the two remaining charged signal tracks. To reduce background, the aforementioned photon energy requirements are applied. For signal decays, there is no additional energy deposition, so $E_{\rm ECL}$ peaks strongly at zero. Misreconstructions of $B_{\rm tag}$ lead to a small tail towards higher energy depositions for true signal events. In contrast, most background decays exhibit non-vanishing extra energy due to the presence of additional neutral particles.

D. Boosted decision trees

Final event selection uses requirements on three variables: $\ln(o^{\rm cs}_{\rm tag})$; missing mass squared ($M^2_{\rm miss} = p^2_{\rm miss}$); and the output of the boosted decision tree (BDT). For each τ reconstruction mode, one BDT is trained using the TMVA framework [38]. All use different input variables, background training samples, and BDT growth parameters. The signal training sample consists of 3×10^7 events out of the complete signal MC sample. To improve the training, events are required to have $E_{\rm ECL} < 1$ GeV and $B_{\rm tag}$ is required to have a quality of $\ln(o^{\rm cs}_{\rm tag}) > -7$. One additional track outside the impact parameter requirement is allowed.

Another 3×10^7 signal events are used for performance tests of the BDT for receiver-operation characteristics (ROC) calculation and overtraining evaluation.

The input variables of the BDT used in the $\tau^- \to e^-\bar{\nu}_e\nu_\tau$ selection are the magnitude of the three-momenta of the pion and electron, the squared lepton-pair momentum transfer q^2, $M^2_{\rm miss}$, and different combinations of all available four-momenta. The momentum transfer can be calculated using the fact that both B mesons are at rest in the c.m. frame, which implies $p_{B_{\rm sig}} = -p_{B_{\rm tag}}$ and $q = p_{B_{\rm sig}} - p_\pi$. Due to the low efficiency of the full reconstruction, an additional signal sample is used with 2×10^7 $\tau^- \to e^-\bar{\nu}_e\nu_\tau$ events on the signal side. The background training sample consists of charmed B^0 decays and $B^0 \to X_u \ell \nu_\ell$ decays. The input variables are linearly decorrelated before their use in the training.

The background sample used in the $\tau^- \to \pi^- \nu_\tau$ selection BDT contains $b \to c$ decays and semileptonic $b \to u$ decays of B^0 mesons. Principal-component analysis (PCA) [38] is applied to the input variables. PCA is an orthogonal transformation which rotates a sample of data points such that the variability along the new axis is maximized. In this way, the variables are decorrelated. The input variables are $M^2_{\rm miss}$, the missing energy, q^2, the absolute three-momentum of $B_{\rm vis}$ in the c.m. frame, combinations of available four-momenta, and the number of unused neutral pions in the event.

The BDT training for the $\tau^- \to \rho^- \nu_\tau$ selection uses the same sample size of $b \to c$ decays, but not semileptonic $b \to u$ decays. The correlation of the input variables is again reduced by a PCA transformation. The variables used in the training are $M^2_{\rm miss}$, the missing energy, q^2, and combinations of the available four-momenta in the decay.

The performance of the three final BDTs is shown in form of the ROC curves in Fig. 2.

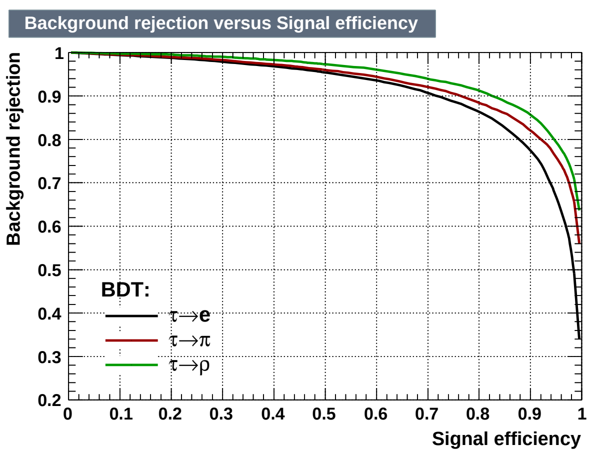


FIG. 2: Background rejection versus signal efficiency determined on the testing sample, from Belle simulation.

E. Final selection

The final selection criteria are determined from MC samples by maximizing individually the expected significance of each single τ reconstruction mode. We perform a scan over three variables simultaneously to obtain the optimal selection: $\ln(o^{\rm cs}_{\rm tag})$, $M^2_{\rm miss}$, and the BDT output. We require $\ln(o^{\rm cs}_{\rm tag}) > -7$ for the leptonic τ reconstruction and $\ln(o^{\rm cs}_{\rm tag}) > -5$ for the hadronic τ reconstruction, as shown in Fig. 1. A minimum requirement on the $M^2_{\rm miss}$ is applied to reject semileptonic $B \to \pi \ell \nu$ events, which have the same final state as the signal decay: since no energy is deposited in the ECL, decays of this type peak at zero extra energy. Also, as there is only a single neutrino in these decays, $M^2_{\rm miss}$ peaks at zero, unlike the case for signal decays, which contain at least two neutrinos. We require $M^2_{\rm miss}$ to be greater than 2.2 GeV/c^4 in the electron channel, 0.0 GeV/c^4 in the pion channel, and 0.6 GeV/c^2 in the ρ channel. The BDT output is the last variable used in the scan. The expected significance is calculated as $\sqrt{-2\ln(\mathcal{L}_0/\mathcal{L}_1)}$; the likelihood is given by $\mathcal{L}_k = \prod^i_{\rm bins} {\rm Poisson}(n_{\rm obs}|n_{\rm bg} + k \cdot n_{\rm sig})$, where $n_{\rm obs} = n_{\rm bg} + n_{\rm sig}$ is the best estimate from the MC samples.

The efficiency of the final selection is determined from MC to be 4.57×10^{-4}. The dominant reconstructed τ decay modes and their relative occurrences are listed in Table I.

TABLE I: Signal reconstruction by τ decay modes. Percentages are obtained from signal MC and sum to 100%.

τ^- decay	Relative Occurrence (%)
$\rho^- \nu_\tau$	29.54
$e^- \bar{\nu}_e \nu_\tau$	29.43
$\pi^- \nu_\tau$	16.70
$\mu^- \bar{\nu}_\mu \nu_\tau$	13.21
$a_1^- \nu_\tau$	8.72
other	2.4

The dominant background in the low-$E_{\rm ECL}$ region arises from $B^0 \to D^{(*)} \ell \nu_\ell$ and $B^0 \to D^{(*)} \rho$ decays with a subsequent decay of $D \to K_L \pi$. The K_L is undetected in these cases and the resulting decay signature resembles that of the signal. No explicit selection is available to further suppress decays of this type.

IV. SYSTEMATIC UNCERTAINTIES

In the computation of the significance level and upper limit, systematic uncertainties are included in the likelihood as nuisance parameters. The likelihood is built from probability density functions (PDFs) determined from MC predictions of each background sample, as described in Section V. All systematic uncertainties are assumed to be Gaussian-distributed and are evaluated at one standard deviation (σ).

Uncertainties of the particle identification and of the correction factor needed for the full reconstruction efficiency are included as a flat effect over all bins in $E_{\rm ECL}$. All other uncertainties are included in a bin-by-bin fashion. A constant uncertainty of 0.35% has been determined for each charged track with $p_t > 0.2$ GeV/c. Tracks below that threshold have to be treated differently depending on the track momentum [39]. The uncertainty on the number of produced B-meson pairs is 1.4%. The uncertainty due to the K_L veto is determined by varying the K_L efficiency by its uncertainty. The branching fractions of the dominant backgrounds are varied by their errors stated in Ref. [23] to determine the effect on the MC prediction. The uncertainty on the correction factor of the tagside reconstruction is determined in Ref. [19] and applied to the samples. The discrepancy between inclusive and exclusive $|V_{ub}|$ measurements has been included as a flat but asymmetric uncertainty in the $B \to X_u \ell \nu$ sample of $(^{+5}_{-15})$%. An uncertainty of $\pm 10\%$ is applied to the branching fractions in the MC sample of rare $b \to s$ and other rare B decays. Additionally, decays of type $B \to X_u \tau \nu$ are present in the final event selection. The contribution to the E_{ECL} distribution is evaluated from the MC sample assuming a $\mathcal{B}(B^0 \to \rho^+ \tau^- \nu_\tau) = 1.5 \times 10^{-4}$ and found to be small; a relative uncertainty of $\pm 50\%$ is applied. Statistical uncertainties in the PDF shape due to finite MC sample size are included in a way similar to the approach by Barlow and Beeston [40]. Instead of using one Poisson constraint per background sample per bin per τ decay channel, only one constraint term per bin per channel is used. The uncertainty introduced by this approximation is negligible for bins with non-vanishing content and reduces the amount of computation time needed. Instead of the finite MC uncertainty, the fit error is included as a systematic uncertainty for the dominant $b \to c$ contribution. The theoretical uncertainties of the signal form factors f^+ and f^0 are included by generating additional signal MC with one form factor fixed and the other varied by its 1σ uncertainty. The relative uncertainties determined in this way are combined into a single uncertainty estimate. The systematic uncertainties due to the tracking efficiency and particle identification affect only the overall efficiency and are only included in the calculation of the upper limit. The relative effect on the branching fraction is determined by repeatedly fitting modified PDFs to data. The PDFs are modified by replacing each background contribution with the respective contribution where the systematic effect is applied. For each systematic uncertainty, two fits are performed for the positive and negative deviation. The maximum, absolute deviation is quoted in Table II.

TABLE II: Effects of the systematic uncertainties on the branching fraction.

Source	Relative error (%)
Particle ID	2.4
Track efficiency	0.7
N($B\bar{B}$)	1.4
K_L veto	3.2
BG $\mathcal{B}$	2.8
Tagside	4.6
$\lvert V_{ub}\rvert$	2.8
Rare processes	2.0
$B \to X_u \tau \nu$	2.2
Background fit	0.2
Signal model	1.8
Total	8.3

V. RESULT

A binned maximum likelihood fit is performed to E_{ECL} in bins of 0.15 GeV. Due to similar shapes in the background predictions, all background contributions except for the dominant $b \to c$ transitions are fixed to the MC prediction. Possible errors introduced by this approach are accounted for as systematic uncertainties. The fit is performed simultaneously in all three reconstruction modes. The signal strength parameter μ is constrained between the three modes while the background contributions of the three reconstruction modes are floating parameters. The fit result of the $B^0 \to X_c$ background contribution agrees well with the prediction obtained from the MC sample. The signal strength has been chosen such that $\mu = 1.0$ corresponds to $\mathcal{B}(B^0 \to \pi^- \tau^+ \nu_\tau) = 1.0 \times 10^{-4}$. We obtain a best fit of $\mu = 1.52 \pm 0.72$, corresponding to 51.9 ± 24.3 signal events. The fit results by τ reconstruction mode are listed in Table III. The E_{ECL} distribution and fit results are shown in Fig. 3.

TABLE III: Fit results for signal yield. Total and split by τ reconstruction mode.

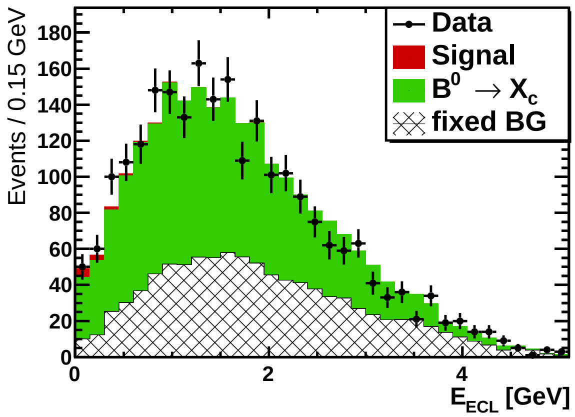
Mode	Signal Yield
e	13.2 ± 6.2
π	30.6 ± 14.3
ρ	8.1 ± 3.8
Total	51.9 ± 24.3

The significance of the measurement is obtained from a pseudo MC study. A test statistic based on the profile likelihood ratio is used. The likelihood is built in bins of 0.15 GeV in E_{ECL}. The binned likelihood is given by

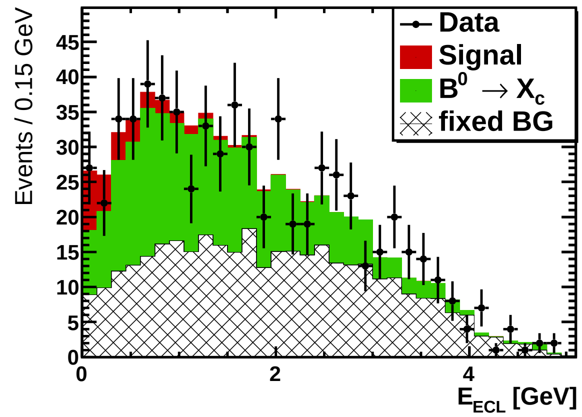
$$\mathcal{L} = \prod_c \prod_b \mathrm{Poisson}(n_{cb}|\nu_{cb}) \cdot \prod_{p \in \mathbb{P}} f_p(a_p|\alpha_p), \qquad (3)$$

where the indices c and b label the reconstruction channel and bin in E_{ECL}, respectively, and $\mathbb{P}$ denotes the set of systematic uncertainties p that are included as nuisance parameters α_p in the calculation of the number of expected events ν_{cb} per channel per bin. The nuisance parameters are parametrized as a relative effect on the nominal template prediction, assumed to be Gaussian-distributed with the nominal value being the global observable a_p. The number of events in the background-only hypothesis is determined from MC simulation and a fit to data for the dominant $b \to c$ background. The likelihood is constructed using the `HistFactory` tool in the RooStats package [41, 42].

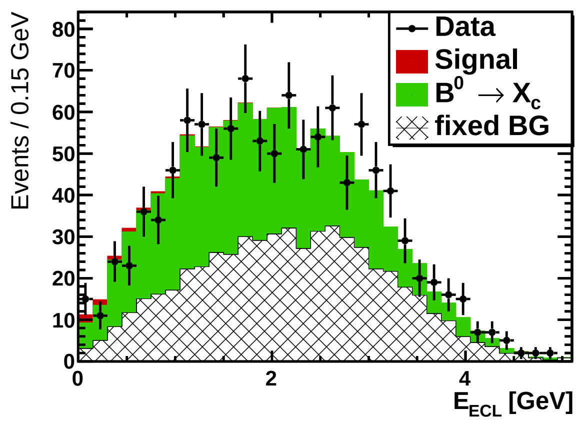
The distribution of the test statistic is obtained by pseudo-experiments. A full frequentist approach is used in both the computations of the significance level and



(a) $\tau \to e\nu\nu$



(b) $\tau \to \pi\nu$



(c) $\tau \to \rho\nu$

FIG. 3: Distributions of E_{ECL} in the three τ reconstruction modes. The signal and $b \to c$ contributions are scaled according to the fit result.

the upper limit. First, the likelihood is fitted to data to obtain the maximum likelihood estimates (MLEs) of all nuisance parameters on data. In each pseudo-experiment generation, the nuisance parameters are fixed to their respective MLE. In the subsequent maximization of the likelihood, the nuisance parameters are free parameters. The global observables are randomized in each pseudo-experiment.

Using pseudo-experiments, the p-value of the background-only hypothesis for data is determined and the significance level Z is computed in terms of standard deviations as

$$Z = \Phi^{-1}\left(1 - p\right),$$

where Φ^{-1} is the cumulative distribution function of the standard normal Gaussian.

We observe a signal significance of 2.8σ, not including systematic uncertainties in the calculation. Including all relevant systematic effects results in a significance of 2.4σ. For this result, the test statistic has been computed on 10 000 background-only pseudo-experiments.

Given the level of significance of these results, we invert the hypothesis test and compute an upper limit on the branching fraction. pseudo-experiments are generated for different signal strength parameters for both signal-plus-background and background-only hypotheses in order to obtain CL_{s+b} and CL_b, respectively. The upper limit is then computed using $CL_s = CL_{s+b}/CL_b$ [43], where a scan over reasonable signal strength parameter values is performed. At each step, 10 000 pseudo-experiments have been evaluated for both hypotheses.

At the 90% confidence level, we obtain an upper limit of $\mathcal{B}\left(B^0 \to \pi^-\tau^+\nu_\tau\right) < 2.5 \times 10^{-4}$. The upper limit at the 95% confidence level has been computed to $\mathcal{B}\left(B^0 \to \pi^-\tau^+\nu_\tau\right) < 2.8 \times 10^{-4}$. This result is the first result on $\mathcal{B}\left(B^0 \to \pi^-\tau^+\nu_\tau\right)$ and is in good agreement with the SM prediction.

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; Austrian Science Fund under Grant No. P 22742-N16 and P 26794-N20; the National Natural Science Foundation of China under Contracts No. 10575109, No. 10775142, No. 10875115, No. 11175187, and No. 11475187; the Chinese Academy of Science Center for Excellence in Particle Physics; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LG14034; the Carl Zeiss Foundation, the Deutsche Forschungsgemeinschaft and the VolkswagenStiftung; the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; the WCU program of the Ministry of Education, National Research Foundation (NRF) of Korea Grants No. 2011-0029457, No. 2012-0008143, No. 2012R1A1A2008330, No. 2013R1A1A3007772, No. 2014R1A2A2A01005286, No. 2014R1A2A2A01002734, No. 2015R1A2A2A01003280 , No. 2015H1A2A1033649; the Basic Research Lab program under NRF Grant No. KRF-2011-0020333, Center for Korean J-PARC Users, No. NRF-2013K1A3A7A06056592; the Brain Korea 21-Plus program and Radiation Science Research Institute; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Education and Science of the Russian Federation and the Russian Foundation for Basic Research; the Slovenian Research Agency; the Basque Foundation for Science (IKERBASQUE) and the Euskal Herriko Unibertsitatea (UPV/EHU) under program UFI 11/55 (Spain); 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").

---

[1] Throughout this paper, the inclusion of the charge-conjugate mode process is implied unless otherwise stated.

[2] A. Khodjamirian, T. Mannel, N. Offen, and Y.-M. Wang, Phys. Rev. D **83**, 094031 (2011), arXiv:1103.2655 [hep-ph].
[3] G. Duplancic, A. Khodjamirian, T. Mannel, B. Melic, and N. Offen, JHEP **0804**, 014 (2008), arXiv:0801.1796 [hep-ph].
[4] E. Dalgic, A. Gray, M. Wingate, C. T. Davies, G. P. Lepage, *et al.*, Phys. Rev. D **73**, 074502 (2006), arXiv:hep-lat/0601021 [hep-lat]; E. Gulez, A. Gray, M. Wingate, C. T. H. Davies, G. P. Lepage, and J. Shigemitsu, Phys. Rev. D **75**, 119906 (2007), erratum.
[5] J. A. Bailey, C. Bernard, C. E. DeTar, M. Di Pierro, A. El-Khadra, *et al.*, Phys. Rev. D **79**, 054507 (2009), arXiv:0811.3640 [hep-lat].
[6] J. A. Bailey *et al.* (Fermilab Lattice, MILC), Phys. Rev. D **92**, 014024 (2015), arXiv:1503.07839 [hep-lat].
[7] C. Bourrely, L. Lellouch, and I. Caprini, Phys. Rev. D **79**, 013008 (2009), arXiv:0807.2722 [hep-ph].
[8] A. Khodjamirian, PoS **BEAUTY2009**, 045 (2009).
[9] M. Kobayashi and T. Maskawa, PTEP **49**, 652 (1973).
[10] T. D. Lee, Phys. Rev. D **8**, 1226 (1973).
[11] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, Phys. Rept. **516**, 1 (2012), arXiv:1106.0034 [hep-ph].
[12] J. Wess and B. Zumino, Nucl. Phys. B **70**, 39 (1974).
[13] Yu. A. Golfand and E. P. Likhtman, JETP Lett. **13**, 323 (1971), [Pisma Zh. Eksp. Teor. Fiz.13,452(1971)].
[14] R. Dutta, A. Bhol, and A. K. Giri, Phys. Rev. D **88**, 114023 (2013), arXiv:1307.6653 [hep-ph].
[15] Y. Grossman and Z. Ligeti, Phys. Lett. B **332**, 373 (1994), arXiv:hep-ph/9403376 [hep-ph].
[16] C.-H. Chen and C.-Q. Geng, JHEP **10**, 053 (2006), arXiv:hep-ph/0608166 [hep-ph].
[17] C. S. Kim and R.-M. Wang, Phys. Rev. D **77**, 094006 (2008), arXiv:0712.2954 [hep-ph].
[18] H. Ha *et al.* (Belle Collaboration), Phys. Rev. D **83**, 071101 (2011), arXiv:1012.0090 [hep-ex].
[19] A. Sibidanov *et al.* (Belle Collaboration), Phys. Rev. D **88**, 032005 (2013), arXiv:1306.2781 [hep-ex].
[20] P. del Amo Sanchez *et al.* (BaBar Collaboration), Phys. Rev. D **83**, 032007 (2011), arXiv:1005.3288 [hep-ex].
[21] J. P. Lees *et al.* (BaBar Collaboration), Phys. Rev. D **86**, 092004 (2012), arXiv:1208.1253 [hep-ex].
[22] D. Daping, R. van de Water, and R. Zhou (FNAL/MILC Collaboration), (private communication).
[23] K. Olive *et al.* (Particle Data Group), Chin. Phys. C **38**, 090001 (2014).
[24] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res. Sect. A **499**, 1 (2003); T. Abe *et al.*, PTEP **2013**, 03A001 (2013) and references therein
[25] Z. Natkaniec *et al.* (Belle SVD2 Group), Nucl. Instrum. Methods Phys. Res. Sect. A **560**, 1 (2006); Y. Ushiroda (Belle SVD2 Group), Nucl. Instrum. Methods Phys. Res. Sect. A **511**, 6 (2003).
[26] 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.* (Belle Collaboration), PTEP **2012**, 04D001 (2012), arXiv:1212.5342 [hep-ex].
[27] D. J. Lange, Nucl. Instrum. Methods Phys. Res. Sect. A **462**, 152 (2001).
[28] R. Brun *et al.*, *GEANT 3.21*, CERN Report DD/EE/84-1 (1984).
[29] T. Sjostrand, S. Mrenna, and P. Z. Skands, JHEP **0605**, 026 (2006), arXiv:hep-ph/0603175 [hep-ph].
[30] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995), arXiv:hep-ph/9503486 [hep-ph].
[31] M. Feindt, F. Keller, M. Kreps, T. Kuhr, S. Neubauer, D. Zander, and A. Zupanc, Nucl. Instrum. Methods Phys. Res. Sect. A **654**, 432440 (2011).
[32] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res. Sect. A **559**, 190 (2006).
[33] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, Phys. Lett. **12**, 57 (1964).
[34] G. C. Fox and S. Wolfram, Phys. Rev. Lett. **41**, 1581 (1978); the modified version used in this paper is described in K. Abe *et al.* (Belle Collaboration), Phys. Lett. B **511**, 151 (2001), arXiv:hep-ex/0103042 [hep-ex].
[35] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, Nucl. Instrum. Methods Phys. Res. Sect. A **485**, 490 (2002), arXiv:hep-ex/0108044 [hep-ex].
[36] A. Abashian, K. Abe, K. Abe, P. Behera, F. Handa, *et al.*, Nucl. Instrum. Methods Phys. Res. Sect. A **491**, 69 (2002).
[37] E. Nakano, Nucl. Instrum. Methods Phys. Res. Sect. A **494**, 402 (2002).
[38] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, PoS **ACAT**, 040 (2007), arXiv:physics/0703039.
[39] W. Dungel *et al.* (Belle), Phys. Rev. **D82**, 112007 (2010), arXiv:1010.5620 [hep-ex].
[40] R. J. Barlow and C. Beeston, Comput. Phys. Commun. **77**, 219 (1993).
[41] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro, *et al.*, PoS **ACAT2010**, 057 (2010), arXiv:1009.1003 [physics.data-an].
[42] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke (ROOT Collaboration), *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, Tech. Rep. CERN-OPEN-2012-016 (New York U., New York, 2012).
[43] A. L. Read, J. Phys. G **28**, 2693 (2002).