Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle

M. Huschle,[26] T. Kuhr,[35] M. Heck,[26] P. Goldenzweig,[26] A. Abdesselam,[61] I. Adachi,[15,12] K. Adamczyk,[48] H. Aihara,[67]
S. Al Said,[61,29] K. Arinstein,[4,51] D. M. Asner,[53] T. Aushev,[41,23] R. Ayad,[61] T. Aziz,[62] I. Badhrees,[61,28] A. M. Bakich,[60]
V. Bansal,[53] E. Barberio,[39] V. Bhardwaj,[58] B. Bhuyan,[17] J. Biswal,[24] A. Bobrov,[4,51] A. Bozek,[48] M. Bračko,[37,24]
T. E. Browder,[14] D. Červenkov,[5] P. Chang,[47] V. Chekelian,[38] A. Chen,[45] B. G. Cheon,[13] K. Chilikin,[23] R. Chistov,[23]
K. Cho,[30] V. Chobanova,[38] Y. Choi,[59] D. Cinabro,[73] J. Dalseno,[38,63] M. Danilov,[23,40] Z. Doležal,[5] Z. Drásal,[5] D. Dutta,[62]
S. Eidelman,[4,51] D. Epifanov,[67] H. Farhat,[73] J. E. Fast,[53] T. Ferber,[8] A. Frey,[11] B. G. Fulsom,[53] V. Gaur,[62] N. Gabyshev,[4,51]
A. Garmash,[4,51] R. Gillard,[73] R. Glattauer,[20] Y. M. Goh,[13] B. Golob,[34,24] J. Grygier,[26] P. Hamer,[11] K. Hara,[15] T. Hara,[15,12]
J. Hasenbusch,[3] K. Hayasaka,[43] H. Hayashii,[44] X. H. He,[54] M. Heider,[26] A. Heller,[26] T. Horiguchi,[65] W.-S. Hou,[47]
C.-L. Hsu,[39] T. Iijima,[43,42] K. Inami,[42] G. Inguglia,[8] A. Ishikawa,[65] R. Itoh,[15,12] Y. Iwasaki,[15] I. Jaegle,[14] D. Joffe,[27]
K. K. Joo,[6] T. Julius,[39] K. H. Kang,[32] E. Kato,[65] P. Katrenko,[23] T. Kawasaki,[49] T. Keck,[26] C. Kiesling,[38] D. Y. Kim,[57]
H. J. Kim,[32] J. B. Kim,[31] J. H. Kim,[30] K. T. Kim,[31] M. J. Kim,[32] S. H. Kim,[13] Y. J. Kim,[30] K. Kinoshita,[7] B. R. Ko,[31]
N. Kobayashi,[68] P. Kodyš,[5] S. Korpar,[37,24] P. Križan,[34,24] P. Krokovny,[4,51] T. Kumita,[69] A. Kuzmin,[4,51] Y.-J. Kwon,[75]
I. S. Lee,[13] C. Li,[39] Y. Li,[72] L. Li Gioi,[38] J. Libby,[18] D. Liventsev,[72,15] P. Lukin,[4,51] M. Masuda,[66] D. Matvienko,[4,51]
K. Miyabayashi,[44] H. Miyake,[15,12] H. Miyata,[49] R. Mizuk,[23,40] G. B. Mohanty,[62] S. Mohanty,[62,71] A. Moll,[38,63]
H. K. Moon,[31] R. Mussa,[22] K. R. Nakamura,[15] E. Nakano,[52] M. Nakao,[15,12] T. Nanut,[24] M. Nayak,[18] N. K. Nisar,[62]
S. Nishida,[15,12] S. Ogawa,[64] S. Okuno,[25] C. Oswald,[3] G. Pakhlova,[41,23] B. Pal,[7] C. W. Park,[59] H. Park,[32] T. K. Pedlar,[36]
L. Pesántez,[3] R. Pestotnik,[24] M. Petrič,[24] L. E. Piilonen,[72] C. Pulvermacher,[26] E. Ribežl,[24] M. Ritter,[38] A. Rostomyan,[8]
Y. Sakai,[15,12] S. Sandilya,[62] L. Santelj,[15] T. Sanuki,[65] Y. Sato,[42] V. Savinov,[55] O. Schneider,[33] G. Schnell,[1,16] C. Schwanda,[20]
A. J. Schwartz,[7] D. Semmler,[9] K. Senyo,[74] O. Seon,[42] M. E. Sevior,[39] V. Shebalin,[4,51] C. P. Shen,[2] T.-A. Shibata,[68]
J.-G. Shiu,[47] F. Simon,[38,63] Y.-S. Sohn,[75] A. Sokolov,[21] E. Solovieva,[23] S. Stanič,[50] M. Starič,[24] M. Steder,[8]
J. Stypula,[48] M. Sumihama,[10] T. Sumiyoshi,[69] U. Tamponi,[22,70] Y. Teramoto,[52] K. Trabelsi,[15,12] V. Trusov,[26] M. Uchida,[68]
T. Uglov,[23,41] S. Uno,[15,12] Y. Usov,[4,51] C. Van Hulse,[1] P. Vanhoefer,[38] G. Varner,[14] A. Vinokurova,[4,51] V. Vorobyev,[4,51]
M. N. Wagner,[9] C. H. Wang,[46] M.-Z. Wang,[47] P. Wang,[19] X. L. Wang,[72] M. Watanabe,[49] Y. Watanabe,[25]
E. Won,[31] H. Yamamoto,[65] J. Yamaoka,[53] S. Yashchenko,[8] H. Ye,[8] Y. Yook,[75] C. Z. Yuan,[19] Y. Yusa,[49] Z. P. Zhang,[56]
V. Zhilich,[4,51] V. Zhulanov,[4,51] and A. Zupanc[24]

(Belle Collaboration)

[1]University of the Basque Country UPV/EHU, 48080 Bilbao
[2]Beihang University, Beijing 100191
[3]University of Bonn, 53115 Bonn
[4]Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090
[5]Faculty of Mathematics and Physics, Charles University, 121 16 Prague
[6]Chonnam National University, Kwangju 660-701
[7]University of Cincinnati, Cincinnati, Ohio 45221
[8]Deutsches Elektronen–Synchrotron, 22607 Hamburg
[9]Justus-Liebig-Universität Gießen, 35392 Gießen
[10]Gifu University, Gifu 501-1193
[11]II. Physikalisches Institut, Georg-August-Universität Göttingen, 37073 Göttingen
[12]SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193
[13]Hanyang University, Seoul 133-791
[14]University of Hawaii, Honolulu, Hawaii 96822
[15]High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801
[16]IKERBASQUE, Basque Foundation for Science, 48013 Bilbao
[17]Indian Institute of Technology Guwahati, Assam 781039
[18]Indian Institute of Technology Madras, Chennai 600036
[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]Institute for Theoretical and Experimental Physics, Moscow 117218
[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, Georgia 30144

[28]*King Abdulaziz City for Science and Technology, Riyadh 11442*
[29]*Department of Physics, Faculty of Science, King Abdulaziz University, Jeddah 21589*
[30]*Korea Institute of Science and Technology Information, Daejeon 305-806*
[31]*Korea University, Seoul 136-713*
[32]*Kyungpook National University, Daegu 702-701*
[33]*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015*
[34]*Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana*
[35]*Ludwig Maximilians University, 80539 Munich*
[36]*Luther College, Decorah, Iowa 52101*
[37]*University of Maribor, 2000 Maribor*
[38]*Max-Planck-Institut für Physik, 80805 München*
[39]*School of Physics, University of Melbourne, Victoria 3010*
[40]*Moscow Physical Engineering Institute, Moscow 115409*
[41]*Moscow Institute of Physics and Technology, Moscow Region 141700*
[42]*Graduate School of Science, Nagoya University, Nagoya 464-8602*
[43]*Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602*
[44]*Nara Women's University, Nara 630-8506*
[45]*National Central University, Chung-li 32054*
[46]*National United University, Miao Li 36003*
[47]*Department of Physics, National Taiwan University, Taipei 10617*
[48]*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
[49]*Niigata University, Niigata 950-2181*
[50]*University of Nova Gorica, 5000 Nova Gorica*
[51]*Novosibirsk State University, Novosibirsk 630090*
[52]*Osaka City University, Osaka 558-8585*
[53]*Pacific Northwest National Laboratory, Richland, Washington 99352*
[54]*Peking University, Beijing 100871*
[55]*University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
[56]*University of Science and Technology of China, Hefei 230026*
[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]*Toho University, Funabashi 274-8510*
[65]*Tohoku University, Sendai 980-8578*
[66]*Earthquake Research Institute, University of Tokyo, Tokyo 113-0032*
[67]*Department of Physics, University of Tokyo, Tokyo 113-0033*
[68]*Tokyo Institute of Technology, Tokyo 152-8550*
[69]*Tokyo Metropolitan University, Tokyo 192-0397*
[70]*University of Torino, 10124 Torino*
[71]*Utkal University, Bhubaneswar 751004*
[72]*CNP, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*
[73]*Wayne State University, Detroit, Michigan 48202*
[74]*Yamagata University, Yamagata 990-8560*
[75]*Yonsei University, Seoul 120-749*

We report a measurement of the branching fraction ratios $R(D^{(*)})$ of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ (where $\ell = e$ or μ) using the full Belle data sample of $772 \times 10^6 B\bar{B}$ pairs collected at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. The measured values are $R(D) = 0.375 \pm 0.064(\text{stat}) \pm 0.026(\text{syst})$ and $R(D^*) = 0.293 \pm 0.038(\text{stat}) \pm 0.015(\text{syst})$. The analysis uses hadronic reconstruction of the tag-side B meson and purely leptonic τ decays. The results are consistent with earlier measurements and do not show a significant deviation from the standard model prediction.

I. INTRODUCTION

Semileptonic $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays [1], where $\ell = e$ or μ, have been studied in detail, experimentally [2] and theoretically [3], and are used, for example, to extract the standard model (SM) parameter $|V_{cb}|$ [4]. The replacement of the light lepton by the higher-mass τ leads to an increased sensitivity to new physics (NP) effects. In particular, models with charged Higgs bosons [5,6], whose couplings are proportional to mass and thus more pronounced for τ leptons, predict measurable deviations of the branching fraction and kinematic distributions from SM expectations. The measurement of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ is challenging because the τ must be reconstructed from its decay products that include one or more neutrinos.

The first observation of an exclusive semitauonic B decay was reported by the Belle Collaboration in 2007 in the channel $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ [7]. Subsequent measurements by *BABAR* and Belle [8–10] reported branching fractions above—yet consistent with—the SM predictions. In 2012, a significant excess over the SM expectation was reported by *BABAR* [11] that suggested the presence of NP; this called for an independent confirmation. Interestingly, the two-Higgs doublet model (2HDM) of type II, which might explain a deviation from the SM expectation in a (semi) tauonic B decay [5], is incompatible with this result. A recent LHCb measurement of $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ [12] also shows a 2.1σ deviation from the SM prediction.

Measurements and predictions are usually quoted as branching fraction ratios,

$$R(D) = \frac{\mathcal{B}(\bar{B} \to D\tau^-\bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \to D\ell^-\bar{\nu}_\ell)}, \tag{1}$$

and

$$R(D^*) = \frac{\mathcal{B}(\bar{B} \to D^*\tau^-\bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \to D^*\ell^-\bar{\nu}_\ell)}, \tag{2}$$

to reduce experimental systematic uncertainties and theory uncertainties from form factors, where $\mathcal{B}(\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell) = [\mathcal{B}(\bar{B} \to D^{(*)}e^-\bar{\nu}_e) + \mathcal{B}(\bar{B} \to D^{(*)}\mu^-\bar{\nu}_\mu)]/2$. In Ref. [11] the calculations in Ref. [13] are used with updated form factor measurements to obtain the standard model predictions $R(D)_{\rm SM} = 0.297 \pm 0.017$ and $R(D^*)_{\rm SM} = 0.252 \pm 0.003$. More recent predictions of $R(D)_{\rm SM}$ are 0.299 ± 0.011 [14] and 0.300 ± 0.008 [15].

In this paper, we report new measurements of $R(D)$ and $R(D^*)$ with the full Belle $\Upsilon(4S) \to B\bar{B}$ data set of 711 fb^{-1}. The τ lepton is reconstructed in the leptonic decays $\tau^- \to e^-\bar{\nu}_e\nu_\tau$ and $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$ so that the signal and normalization modes have the same detectable final state particles. This reduces the systematic uncertainty in R but requires a method to distinguish the modes experimentally. For this purpose, we exploit the kinematics of $e^+e^- \to \Upsilon(4S) \to B\bar{B}$ by reconstructing the accompanying B meson, $B_{\rm tag}$, in a hadronic decay mode and extracting the invariant mass squared,

$$M^2_{\rm miss} = (p_{e^+e^-} - p_{\rm tag} - p_{D^{(*)}} - p_\ell)^2/c^2, \tag{3}$$

of all undetected signal-B meson daughters, where $p_{e^+e^-}$, $p_{\rm tag}$, $p_{D^{(*)}}$, and p_ℓ are the four-momenta of the colliding beam particles, the $B_{\rm tag}$ candidate, and the reconstructed signal-B daughters, respectively.

The $M^2_{\rm miss}$ distribution peaks at (above) zero for the normalization (signal) mode with one neutrino (three neutrinos) in the final state. The separation power is weaker for backgrounds where multiple final-state particles are not reconstructed. We improve the rejection of such backgrounds by training a neural network to distinguish them from the signal in the high-$M^2_{\rm miss}$ region. Since the low- and high-$M^2_{\rm miss}$ regions are dominated by different backgrounds, the data sample is split at $M^2_{\rm miss} = 0.85$ GeV$^2/c^4$ and the subsamples are fit simultaneously. In the low-$M^2_{\rm miss}$ region, which is dominated by the normalization mode, we fit the $M^2_{\rm miss}$ distribution; in the high $M^2_{\rm miss}$ region, where the background with multiple missing particles contributes, we fit the neural-network output distribution. The analysis procedure is developed and optimized with simulated data before applying it to the experimental data.

II. BELLE EXPERIMENT

This measurement is based on a data sample that contains $772 \times 10^6 B\bar{B}$ pairs, collected with the Belle detector at the KEKB asymmetric-energy e^+e^- (3.5 on 8 GeV) collider [16] operating at the $\Upsilon(4S)$ resonance. 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 comprised of CsI(Tl) crystals (ECL) located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the coil is instrumented to detect K_L^0 mesons and to identify muons (KLM). The detector is described in detail in Ref. [17]. Two inner-detector configurations were used. A 2.0-cm beampipe and a three-layer silicon vertex detector was used for the first sample of $152 \times 10^6 B\bar{B}$ pairs, while a 1.5-cm beampipe, a four-layer silicon detector, and a small-cell inner drift chamber were used for the remaining $620 \times 10^6 B\bar{B}$ pairs [18].

III. RECONSTRUCTION

We reconstruct $B_{\rm tag}$ candidates using the hierarchical hadronic full reconstruction algorithm [19], which includes 1149 B final states. The efficiency of the $B_{\rm tag}$ reconstruction is 0.3% for B^+ and 0.2% for B^0 mesons [19]. Requirements

on three observables are applied to enhance the sample's purity: the beam energy-constrained mass $M_{bc} \equiv \sqrt{E_{beam}^2 - (\mathbf{p}_{tag}c)^2}/c^2$ must lie between 5.274 and 5.286 GeV/c^2, where E_{beam} is the colliding-beam energy and $\mathbf{p}_{tag}$ is the B_{tag} momentum, both measured in the center-of-mass system (CMS); the absolute value of the energy difference $\Delta E \equiv E_{tag} - E_{beam}$ must be smaller than 50 MeV, where E_{tag} is the B_{tag} CMS energy; and the full-reconstruction neural-network quality estimator for B_{tag} (which incorporates modified Fox-Wolfram moments [20] to suppress $e^+e^- \to q\bar{q}$ continuum events) must exceed a channel-dependent threshold that preserves $\approx 85\%$ of the $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ events.

In each event with a selected B_{tag} candidate, we search for the signature $D^{(*)}\ell$, with $\ell = e$ or μ, among the remaining tracks and calorimeter clusters. The four disjoint data samples are denoted $D^+\ell^-$, $D^0\ell^-$, $D^{*+}\ell^-$, and $D^{*0}\ell^-$. We reconstruct D^+ mesons in the decays to $K^-\pi^+\pi^+$, $K_S^0\pi^+$, $K_S^0\pi^+\pi^0$, and $K_S^0\pi^+\pi^+\pi^-$; D^0 mesons to $K^-\pi^+$, $K^-\pi^+\pi^+\pi^-$, $K^-\pi^+\pi^0$, $K_S^0\pi^0$, and $K_S^0\pi^+\pi^-$; D^{*+} mesons to $D^0\pi^+$ and $D^+\pi^0$; and D^{*0} mesons to $D^0\pi^0$ and $D^0\gamma$.

Charged-particle candidates are selected from tracks that originate from within 4.0 (2.0) cm along (perpendicular to) the beam direction of the interaction point (IP). Selections on the particle-identification likelihood ratio of the electron (muon) vs. the hadron hypothesis for the candidate lepton track retain 95% (92%) of signal events. We veto a $D^{(*)}$ candidate if a charged daughter is lepton-like, with a signal efficiency of 97%. K_S^0 candidates are reconstructed from pairs of oppositely charged tracks, treated as pions, and must satisfy standard quality requirements [21].

Clusters in the ECL with an energy of at least 50 MeV and no matching track are identified as photons. Candidate π^0's are reconstructed from pairs of photons. For end-cap photons used in a π^0 candidate, the energy must be greater than 80 MeV. The momentum of π^0 candidates not originating from a D^* decay must exceed 200 MeV/c in the signal-B rest frame. The absolute value of the difference $S_{\gamma\gamma}$ between the invariant mass of the π^0 candidate and the nominal π^0 mass, normalized to its uncertainty, must be below 3.0.

We select $D^{+/0}$ meson candidates with a CMS momentum below 3.0 GeV/c. For both D and D^* candidates, the candidate D mass and $D^* - D$ mass difference must be within 1.5 standard deviations of the nominal D mass and $D^* - D$ mass difference, respectively. The resolution is asymmetric and is taken from simulated data.

The missing mass squared, M_{miss}^2, must lie between -0.2 and 8.0 GeV$^2/c^4$. The momentum transfer $q^2 \equiv (p_B - p_{D^{(*)}})^2$ on the signal side is required to be greater than 4.0 GeV$^2/c^2$, which suppresses the otherwise overwhelming contribution from semileptonic B meson decays to light leptons. Events with a remaining π^0 candidate are rejected

if the energy of either daughter photon exceeds 50/100/150 MeV in the barrel/forward/backward region. The overall charge of the event must be zero, with no additional charged tracks allowed.

If there are several B_{tag} candidates, the one with the most signal like neural-network quality estimator is selected. Then, on average, we have 1.23 signal or normalization candidates per event, with most ambiguities arising from D^* meson decays to a D meson and a neutral pion or photon in $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decays. In a multicandidate event, we select one at random.

IV. SIMULATION

We use samples of simulated (MC) events to study backgrounds, to optimize the selection criteria, and to determine the probability density function (PDF) shapes of the fit components. The decay chains in all simulated data are generated with the EVTGEN [22] package; the GEANT3 [23] framework is used to simulate the detector response. A luminosity-weighted run-dependent sample of 10^7 events for each of the four signatures is generated for the signal mode $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ using the decay model described in Ref. [24]. To investigate possible new physics effects, we produce a sample of simulated $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ signal events for the scenario of a two-Higgs-doublet model of type II with $\tan\beta/m_{H^+} = 0.5c^2/$GeV [24]. A sample that corresponds to 5 times the amount of recorded data and contains $B\bar{B}$ events with B mesons decaying generically via $b \to c$ transitions as well as $q\bar{q}$ events with $q \in \{u, d, s, c\}$ is used for the background.

Several corrections are applied to the background MC sample to improve its agreement with measured data. We first reweight the MC events to account for the imperfect estimate of the proportions of correctly reconstructed B_{tag} candidates and to better estimate the yields of background processes with good tags. (The reweighting cancels to first order in the efficiency ratio used to extract $R(D^{(*)})$.) The weights are given by the ratios of yields in simulation and data, determined from fits to the distributions of M_{bc} and M_{miss}^2 for events with a B_{tag} and a semileptonic decay on the signal side [25]. The correction factors are in the range 0.35 to 1.1, with an overall factor of approximately 0.75. To extract correction factors for the number of incorrectly reconstructed B_{tag} candidates, we compare yields of simulated and reconstructed data in a sideband of M_{bc}, requiring 5.23 GeV/c^2 < M_{bc} < 5.25 GeV/c^2. This is done separately for the four signal modes, and we exclude events with fake $D^{(*)}$ mesons or fake leptons on the signal side as these are corrected by other measures. The ratios of the yields, whose values are between 0.99 and 1.14, are then applied as weights.

Second, we apply a correction for the signal-lepton candidates to account for differing misidentification rates in simulated and recorded data. Correction factors for the

lepton candidate are provided in eight (eleven) bins in polar angle (momentum). (Lepton-identification efficiencies are compatible, within uncertainties, between simulated and recorded data.)

Third, we reweight the events to account for $D^{(*)}$ yield differences in MC and data. While the yield of candidates with a fake D meson will be estimated from sidebands and therefore does not need to be corrected in simulated data, differences in correctly reconstructed D yields can affect the determination of R. We determine the yield ratios of simulated and reconstructed data by fitting the invariant mass (mass difference) distributions of the D (D^*) mesons in a wider window than used for the nominal selection and apply the ratios as weights. This is done individually for each $D^{(*)}$ meson reconstruction channel and yields correction factors between 0.75 and 1.09. Background MC events with $D_s^- \to \ell^- \bar{\nu}_\ell$ decays are reweighted to adopt the latest branching fraction measurements [2].

Fourth, semileptonic decays of B mesons to higher excitations of D mesons, hereinafter labelled D^{**}, comprise one of the most challenging backgrounds. Our background MC sample contains semileptonic—including semitauonic—B decays to D_2^*, D_0^*, D_1, D_1', and the radial excitations $D(2S)$ and $D^*(2S)$, each in the charged and neutral variety. The decays are generated initially according to the ISGW model [26] and reweighted to reproduce the distributions in q^2 and p_ℓ^* (the lepton momentum in the signal-B frame) of the LLSW model [27]. Parameter uncertainties in this model are treated as systematic uncertainties. We consider D^{**} decays to a $D^{(*)}$ and one or two pions, a ρ, or an η meson, with branching ratio assumptions based on quantum-number, phase-space, and isospin arguments. Similar weights are applied to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ events in the background MC according to the most recent measurements of the form factors $\rho^2 = 1.207 \pm 0.015 \pm 0.021$, $R_1 = 1.403 \pm 0.033$, and $R_2 = 0.854 \pm 0.020$ for $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ and $\rho^2 = 1.186 \pm 0.036 \pm 0.041$ for $\bar{B} \to D\ell^-\bar{\nu}_\ell$ [4,28].

V. SAMPLE COMPOSITION

We identify the following components in the data samples:

(i) *lepton normalization* This originates from $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays and has both visible (i.e., non-neutrino) daughters of the B meson correctly reconstructed with a distinctive M_{miss}^2 distribution that peaks around zero. Its yield is a free parameter of the fit.

(ii) *lepton cross-feed* This arises from the misclassification of a $\bar{B} \to D^*\ell^-\bar{\nu}_\ell$ decay into the $D\ell^-$ sample with same D-meson charge due to the loss of a low-energetic π^0 or γ daughter of the D^*. The broad M_{miss}^2 distribution peaks at positive values up to roughly 1.0 GeV^2/c^4. Its yield is allowed to float in the fit.

(iii) *tau signal* This component, arising from $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays, has a correctly reconstructed $D^{(*)}$ daughter and a correctly identified τ-decay daughter lepton. With three final-state neutrinos, its broad M_{miss}^2 distribution is most prominent in regions above 1.0 GeV^2/c^4. The yield $Y_{\tau\mathrm{signal}}^{D\ell^-}$ of the tau signal in each $D\ell^-$ sample is determined by the branching-fraction ratio $R(D)$, which is a free parameter in the fit, the corresponding lepton normalization yield $Y_{\ell\mathrm{norm}}^{D\ell^-}$, and the efficiency ratio f^D for the lepton normalization and tau signal components:

$$Y_{\tau\mathrm{signal}}^{D^{+,0}\ell^-} = R(D)Y_{\ell\mathrm{norm}}^{D^{+,0}\ell^-}/(2f^{D^{+,0}}). \qquad (4)$$

The factor of 2 accounts for the inclusion of both electrons and muons in the lepton normalization component. The efficiency ratios, which include the $\tau^- \to \ell^- \nu_\tau \bar{\nu}_\ell$ branching fractions [2], are determined from simulation to be $f^{D^+} = 1.69 \pm 0.09$ and $f^{D^0} = 1.91 \pm 0.06$, where the uncertainties are statistical. In a similar way, the tau signal yield in the $D^*\ell^-$ samples is given by the floating fit parameter $R(D^*)$ and the corresponding lepton normalization yield. However, to encompass larger yields and thus obtain smaller statistical uncertainties, the cross-feeds are added to the tau signal and lepton normalization with the concomitant use of an effective efficiency ratio $f_{\mathrm{eff}}^{D^*}$, defined by

$$\frac{1}{f_{\mathrm{eff}}^{D^{*+,0}}} = \frac{1 - x_{\mathrm{CF}}}{f^{D^{*+,0}}} + \frac{x_{\mathrm{CF}}}{f_{\mathrm{CF}}^{D^{*+,0}}}, \qquad (5)$$

where x_{CF} is the fraction of lepton cross-feed events relative to the sum of lepton normalization and lepton cross-feed yields, determined from simulation, and f^{D^*} ($f_{\mathrm{CF}}^{D^*}$) is the efficiency ratio for the lepton normalization (lepton cross-feed) and tau signal (tau cross-feed) components. The values of the effective efficiency ratios are $f_{\mathrm{eff}}^{D^{*+}} = 3.11 \pm 0.13$ and $f_{\mathrm{eff}}^{D^{*0}} = 3.63 \pm 0.09$.

(iv) *tau cross-feed* This component is the analogue to the lepton cross-feed but originating from $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$ decays. Its yield and shape in M_{miss}^2 are quite similar to those of the tau signal component. It appears only in the $D\ell^-$ samples and its yield is constrained by the τ signal yield $Y_{\tau\mathrm{signal}}^{D^*\ell}$ in the respective $D^*\ell^-$ samples of same charge, assuming a π^0 or γ from the D^* decay is not reconstructed. The constraining factor is taken from the appropriate lepton normalization and lepton cross-feed yields and is calibrated by a factor g that represents the cross-feed ratio for light-lepton and τ modes; MC gives $g^+ = 0.83 \pm 0.08$ for the $D^+\ell^-$ sample and $g^0 = 0.69 \pm 0.04$ for

the $D^0\ell^-$ sample. The tau cross-feed yield $Y_{\tau\mathrm{CF}}^{D\ell}$ is given by

$$Y_{\tau\mathrm{CF}}^{D^{+,0}\ell^-} = Y_{\tau\mathrm{signal}}^{D^{*+,0}\ell^-} \frac{Y_{\ell\mathrm{CF}}^{D^{+,0}\ell^-}}{Y_{\ell\mathrm{norm}}^{D^{*+,0}\ell^-}} \frac{1}{g^{+,0}}. \quad (6)$$

(v) *wrong-charge lepton cross-feed* This component is similar to lepton cross-feed but arises from the loss of the charged pion in $D^{*+} \to D^0\pi^+$. To preserve the overall neutral charge of the event, the lost pion is absorbed into the now-misreconstructed B_{tag} meson. (Since D^{*0} mesons do not decay to charged pions, this component appears only in the $D^0\ell^-$ sample.) Its M_{miss}^2 distribution resembles that of lepton cross-feed. Its smaller yield is constrained in the fit relative to the lepton normalization yield in the D^{*+} sample with a factor $f_{\mathrm{wc}} = 0.107 \pm 0.004$, taken from simulation.

(vi) *fake* $\boldsymbol{D^{(*)}}$ This component is dominated by random combinations of final-state particles that form a fake D or D^* meson. This can happen by either missing particles in the event or misassigning particles to the wrong B meson. This background occurs in all samples and, in the $D^*\ell^-$ samples, includes combinations of a correctly reconstructed D meson and an incorrect D^* primary daughter. The M_{miss}^2 distribution is very broad and extends to the highest values.

The fake $D^{(*)}$ yield is estimated separately for each $D^{(*)}$ decay mode. A sideband region is defined in the distribution of the invariant mass M_D (the D^*–D mass difference ΔM_{D^*D}) by excluding twice the signal-region width on both sides of the nominal mass (mass difference) and a $\pm 60\ \mathrm{MeV}/c^2$ window around the D^* peak position for the $D^+ \to K_S^0\pi^+\pi^0$ channel. Multiplying the sideband yield in the real data by the yield ratio in MC of the fake $D^{(*)}$ component and sideband provides an estimate for the fake $D^{(*)}$ yield in each $D^{(*)}$ decay mode; these are summed to obtain the total yield in each of the four data samples.

(vii) $\boldsymbol{D^{**}}$ *background* This component contains candidates that originate from $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell(\nu_\tau\bar{\nu}_\tau)$ decays. The higher-excitation D states decay typically to a D or D^* meson plus one pion (although more pions are possible) so the final state here has a properly identified lepton, a properly reconstructed $D^{(*)}$ meson, and (at least) one pion that might be lost or absorbed into B_{tag}. If the pion is missed, this process mimics the tau signal and exhibits a similar M_{miss}^2 distribution. The yields of this background and the tau signal are comparable. In contrast to the other background components, it is not possible to constrain the yield from MC since the properties of

TABLE I. Fit components in each data sample. For the yield source, "fit" indicates a free parameter in the fit; "constrained" reflects a dependence on other parameters; "MC" denotes a fixed yield taken from simulation; and "SB" identifies a fixed yield derived from the corresponding sideband. The constraints are described in the text.

Component	$D^+\ell^-$	$D^0\ell^-$	$D^{*+}\ell^-$	$D^{*0}\ell^-$	Yield source
ℓ normalization	✓	✓	✓	✓	Fit
ℓ CF	✓	✓	-	-	Fit
τ signal	✓	✓	✓	✓	Fit
τ CF	✓	✓	-	-	Constrained
Wrong charge ℓ CF	-	✓	-	-	Constrained
Fake D	✓	✓	-	-	M_D SB
Fake D^*	-	-	✓	✓	ΔM_{D^*D} SB
D^{**} background	✓	✓	✓	✓	Fit
Fake ℓ	✓	✓	✓	✓	MC
D_s decay	✓	✓	✓	✓	MC
Rest	✓	✓	✓	✓	MC

$\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ and D^{**} decays are not known reliably. Thus, its yield is a free parameter in the fit.

(viii) *fake lepton* This component contains events with a misidentified lepton candidate; the track is usually a kaon or pion from the tag side or from $\bar{B} \to DK$ or $D\pi$ decays. This component also includes $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ events with a misidentified pion from a hadronic τ decay. Since lepton misidentification is far less probable than $D^{(*)}$ misreconstruction, a $D^{(*)}\ell^-$ event in which both are misidentified is classified in the fake $D^{(*)}$ component. The fake lepton background is a broad structure in M_{miss}^2 that appears in all four data samples; the fixed and relatively low yield is estimated accurately from MC.

(ix) $\boldsymbol{D_s}$ *decay* This component arises from the decay chain $\bar{B} \to D^{(*)}D_s^-$ with $D_s^- \to \ell^-\bar{\nu}_\ell(\nu_\tau\bar{\nu}_\tau)$ and so has a final state that mimics the tau signal. The decay $D_s^- \to \ell^-\bar{\nu}_\ell$ is helicity suppressed and only the tauonic D_s^- decays provide a non-negligible contribution. Its M_{miss}^2 distribution resembles that of the tau signal; its low yield in MC is confirmed by experiment, with the most precise determination provided by Belle [29]. Consequently, this component's yield is fixed in the fit to the MC value.

(x) *rest* This component encompasses all background candidates that are not captured by the other listed components. It contains candidates with well-identified final state particles that do not originate from one of the previously covered sources and may be random combinations of tag- and signal-side particles. Its yield is quite low in all four samples and is fixed in the fit to the MC value.

Table I itemizes each component in the fit for each signature. The yields of the fixed components are listed in Table II.

TABLE II. Yields for the fixed components in the four data samples.

	$D^+\ell^-$	$D^0\ell^-$	$D^{*+}\ell^-$	$D^{*0}\ell^-$
Fake $D^{(*)}$	350	1330	180	2220
Fake ℓ	20.9	69	13.7	12.9
D_s decay	22.0	112	21.0	20.7
Rest	23.6	77	4.3	4.2

VI. FIT PROCEDURE

As explained above, the low-M^2_{miss} region is dominated by the lepton normalization and has essentially no sensitivity to the tau signal; in contrast, the high-M^2_{miss} region, where the tau signal is concentrated, exhibits little discrimination power in M^2_{miss} between the tau signal and the other backgrounds—in particular, the D^{**} background. Therefore, we fit simultaneously the M^2_{miss} distribution below 0.85 GeV$^2/c^4$ to constrain the lepton normalization and lepton cross-feed yields and a neural-network output o_{NB} above 0.85 GeV$^2/c^4$ to constrain the yields of the other components. (In fact, *all* components are fit in both regions.) The partition at $M^2_{\text{miss}} = 0.85$ GeV$^2/c^4$ minimizes the expected uncertainty on $R(D)$ and $R(D^*)$.

The aforementioned neural network is trained for each of the four data samples with simulated events to distinguish the tau signal from the backgrounds in the high-M^2_{miss} region: mainly D^{**} background but also the wrong-charge cross-feed, fake lepton, D_s decay, and rest components. The neural network incorporates M^2_{miss} and several other observables that provide the desired signal-to-background separation. The most powerful observable is E_{ECL}, the unassociated energy in the ECL that aggregates all clusters that are not associated with reconstructed particles (including bremsstrahlung). A nonzero E_{ECL} value indicates a missing physical process in the event, such as a decay mode with a π^0 in which only a single daughter photon is reconstructed. Two additional network inputs are q^2 and p^*_ℓ; their additional discriminating power is limited by their strong correlation with M^2_{miss}. Other input variables, which provide marginally more discrimination, are the number of unassigned π^0 candidates with $|S_{\gamma\gamma}| < 5.0$; the cosine of the angle between the momentum and vertex displacement of the $D^{(*)}$ meson; and the decay-channel identifiers of the B and $D^{(*)}$ mesons.

For use in the fit, the neural-network output o_{NB} is transformed into
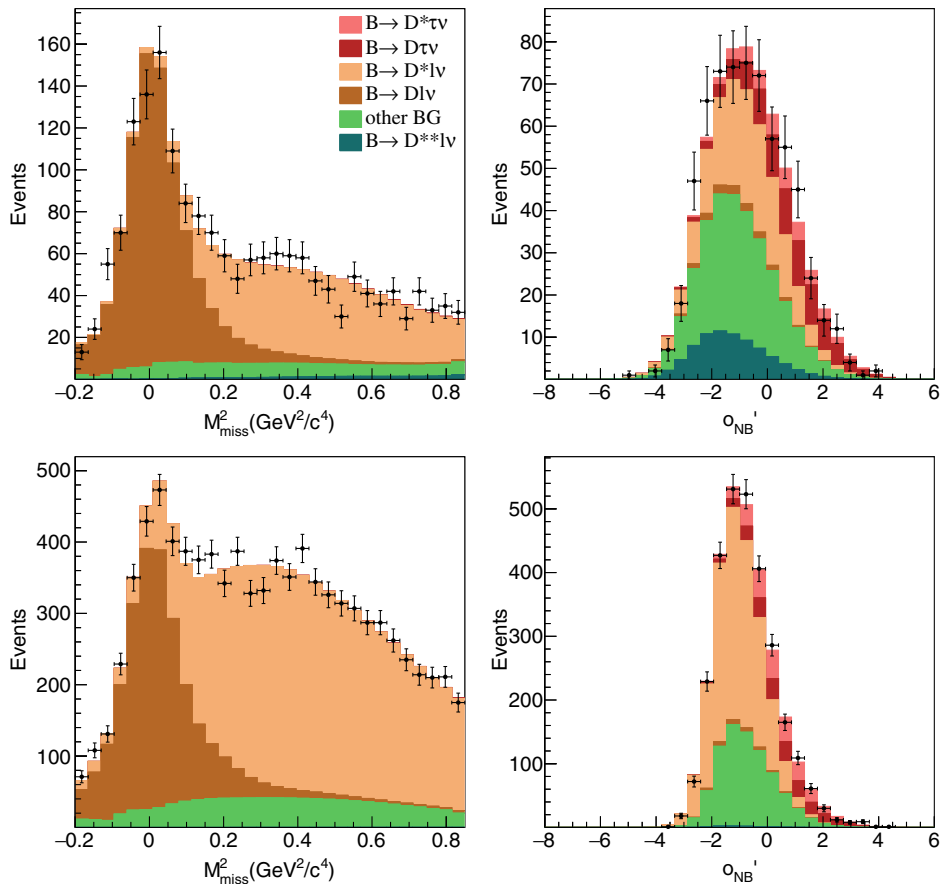

FIG. 1 (color online). Fit projections and data points with statistical uncertainties in the $D^+\ell^-$ (top) and $D^0\ell^-$ (bottom) data samples. Left: M^2_{miss} distribution for $M^2_{\text{miss}} < 0.85$ GeV$^2/c^4$; right: o'_{NB} distribution for $M^2_{\text{miss}} > 0.85$ GeV$^2/c^4$.

$$o'_{\text{NB}} \equiv \log \frac{o_{\text{NB}} - o_{\text{min}}}{o_{\text{max}} - o_{\text{NB}}}, \qquad (7)$$

where the parameters o_{min} and o_{max} are the minimum and maximum network output values, respectively, in the elected data sample. The o'_{NB} distributions have smoother shapes and can be described well with bifurcated Gaussian functions, which makes their parameterizations more robust.

For each fit component within a selected data sample, two PDFs are determined: in M^2_{miss} for $M^2_{\text{miss}} < 0.85 \text{ GeV}^2/c^4$ and in o'_{NB} for $M^2_{\text{miss}} > 0.85 \text{ GeV}^2/c^4$. The PDFs of M^2_{miss} are represented by smoothed histograms and are constructed by applying a smoothing algorithm [30] to the respective MC distributions. Each bifurcated-Gaussian PDF in o'_{NB} is parameterized by the mean, left width and right width, which are determined by an unbinned maximum likelihood fit to the MC distribution. In the fit, each component has a total yield, defined in Table I, with partial yields in the lower- and upper-M^2_{miss} regions that are fixed MC-determined fractions of the total yield.

We maximize the extended likelihood function

$$\mathcal{L} = \prod_i \left[\mathcal{Q}(N_i, K_i) \prod_{k_i=1}^{K_i} \mathcal{P}_i(\mathbf{x}_{k_i}) \right], \qquad (8)$$

where $i \in \{D^+\ell^-, D^0\ell^-, D^{*+}\ell^-, D^{*0}\ell^-\}$ is the data-sample index, $\mathcal{Q}(N_i, K_i)$ is the Poisson probability to observe K_i events for an expectation value of $N_i = \sum_j Y_{i,j}$ events (with $Y_{i,j}$ being the yield of component j in data sample i), and the vector $\mathbf{x}_{k_i}$ holds the values for M^2_{miss} and o'_{NB} of candidate k_i. The PDF $\mathcal{P}_i$ of data sample i is given by

$$\mathcal{P}_i(M^2_{\text{miss}}, o'_{\text{NB}}) = \frac{1}{N_i} \cdot \sum_j Y_{i,j} [f_{i,j,\text{low}} \mathcal{P}_{i,j,\text{low}}(M^2_{\text{miss}})$$
$$+ (1 - f_{i,j,\text{low}}) \mathcal{P}_{i,j,\text{high}}(o'_{\text{NB}})]. \qquad (9)$$

The index j runs over the components and $f_{i,j,\text{low}}$ is the fraction of events of the component j that are in the lower M^2_{miss} range. The one-dimensional probability density function $\mathcal{P}_{i,j,\text{low}}$ ($\mathcal{P}_{i,j,\text{high}}$) represents the M^2_{miss} (o'_{NB}) distribution in the low- (high-)M^2_{miss} region.
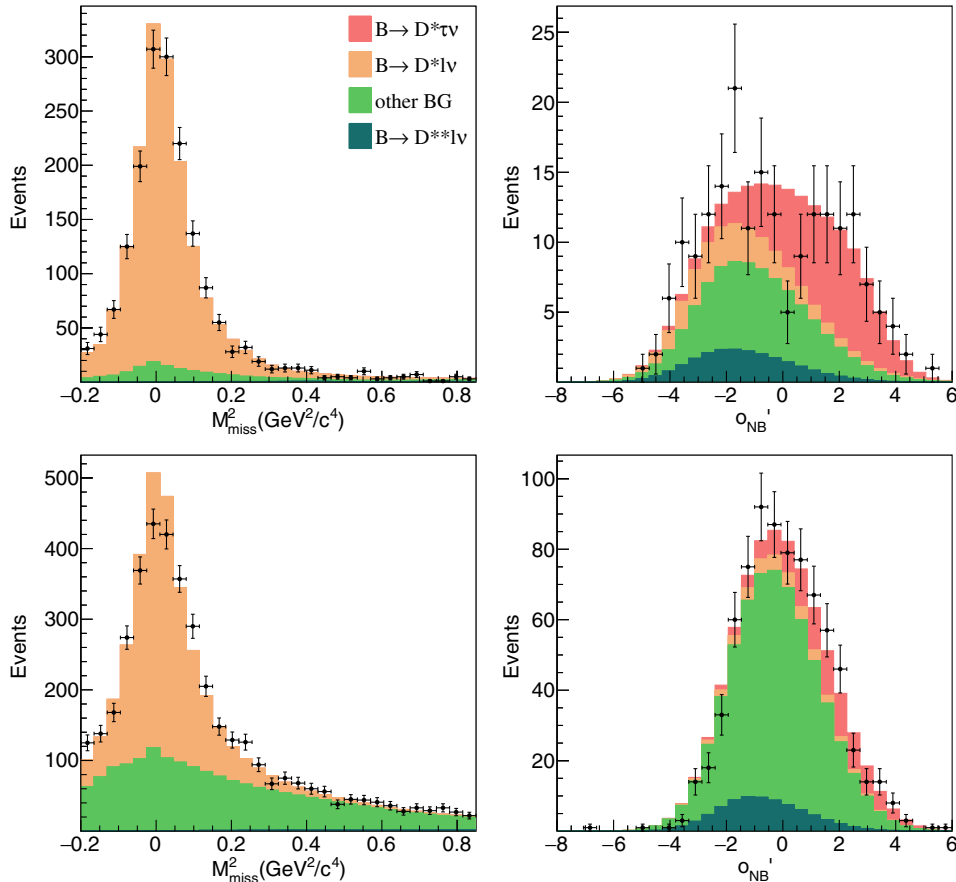

FIG. 2 (color online). Fit projections and data points with statistical uncertainties in the $D^{*+}\ell^-$ (top) and $D^{*0}\ell^-$ (bottom) data samples. Left: M^2_{miss} distribution for $M^2_{\text{miss}} < 0.85 \text{ GeV}^2/c^4$; right: o'_{NB} distribution for $M^2_{\text{miss}} > 0.85 \text{ GeV}^2/c^4$.

The simultaneous fit over all four data samples has twelve free parameters: the lepton normalization yield per sample, the lepton cross-feed yield per $D\ell^-$ sample, the D^{**} background yield per sample, and the branching-fraction ratios $R(D)$ and $R(D^*)$. Here, we assume isospin symmetry and use the same $R(D)$ and $R(D^*)$ parameters for the $\bar{B}^0$ and B^- samples.

VII. CROSS-CHECKS

The implementation of the fit procedure is tested by applying the same procedure to multiple subsets of the available simulated data. The fit accuracies are evaluated using sets of 500 pseudoexperiments and show no significant bias in any measured quantity. These are used also to test the influence on the fit result of the value of $M^2_{\text{miss}} = 0.85 \text{ GeV}^2/c^4$ that is used to partition the samples: variation of this value reduces the precision of the fit result but does not introduce any bias.

Further tests address the compatibility of the simulated and recorded data. To test resolution modelling, we use a sample of events with $q^2 < 3.5 \text{ GeV}^2/c^2$, dominated by $\bar{B} \rightarrow D^{(*)}\ell^-\bar{\nu}_\ell$ decays. As the D^{**} background is one of the most important components—with a large potential for

flaws in its modeling—we evaluate its distributions in more depth by reconstructing a data sample with enriched $\bar{B} \rightarrow D^{**}\ell^-\bar{\nu}_\ell$ content by requiring a signal-like event but with an additional π^0. The background-enriched data samples are fit individually in four dimensions separately: M^2_{miss}, $M^2_{\text{miss,no }\pi^0}$, E_{ECL}, and p^*_ℓ, where $M^2_{\text{miss,no }\pi^0}$ is the missing mass of the candidate, calculated without the additional π^0. The shapes of the components are extracted from simulated data. In each of the four $D^{(*)}\ell^-\pi^0$ samples, consistent yields are obtained from the fits to all four variables, indicating that the simulation describes faithfully the distribution in all tested dimensions.

VIII. RESULTS

The fit to the entire data sample gives

$$R(D) = 0.375 \pm 0.064 \qquad (10)$$

$$R(D^*) = 0.293 \pm 0.038, \qquad (11)$$

corresponding to a yield of 320 $\bar{B} \rightarrow D\tau^-\bar{\nu}_\tau$ and 503 $\bar{B} \rightarrow D^*\tau^-\bar{\nu}_\tau$ events; the errors are statistical. Projections of the fit are shown in Figs. 1 and 2. The high-M^2_{miss} distributions
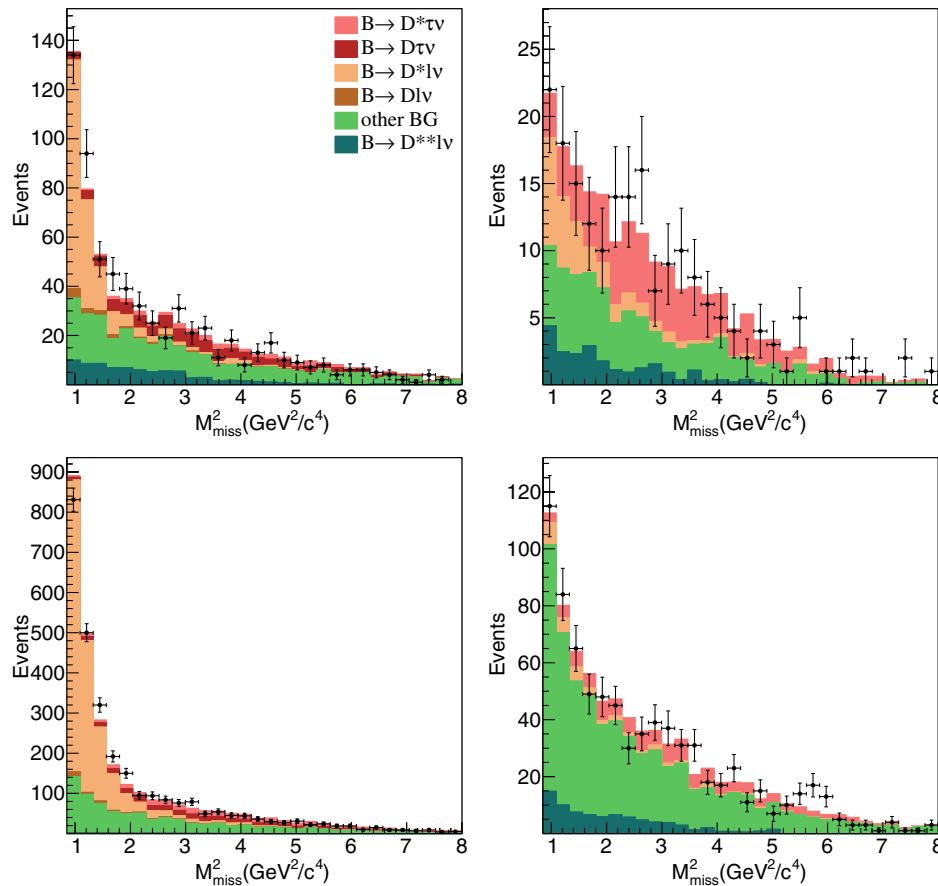

FIG. 3 (color online).   Projections of the fit results and data points with statistical uncertainties for the high M^2_{miss} region. Top left: $D^+\ell^-$; top right: $D^{*+}\ell^-$; bottom left: $D^0\ell^-$; bottom right: $D^{*0}\ell^-$.
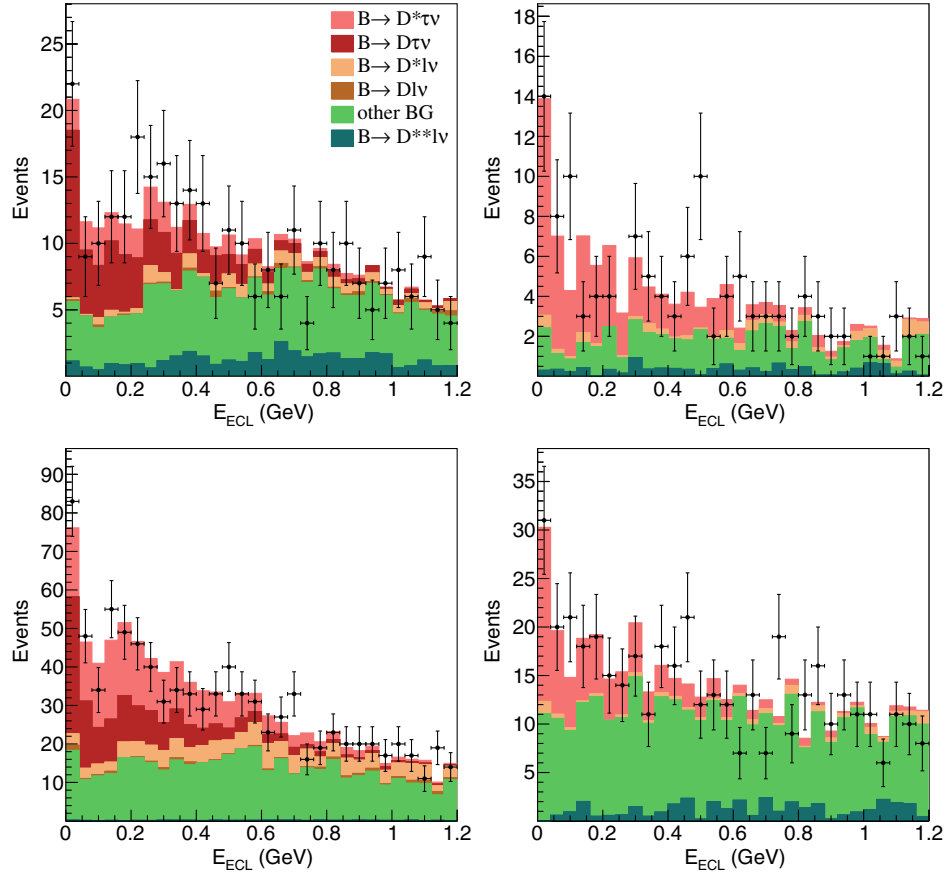
FIG. 4 (color online).   Projections of the fit results and data points with statistical uncertainties in a signal-enhanced region of $M^2_{\text{miss}} > 2.0 \text{ GeV}^2/c^4$ in the E_{ECL} dimension. Top left: $D^+\ell^-$; top right: $D^{*+}\ell^-$; bottom left: $D^0\ell^-$; bottom right: $D^{*0}\ell^-$.

and the fit projections are shown in Fig. 3. Figures 4 and 5 show the signal-enhanced ($M^2_{\text{miss}} > 2.0 \text{ GeV}^2/c^4$) fit projections in E_{ECL} (the most powerful classifier in the neural network) and p^*_ℓ, respectively. In these figures, all background components except D^{**} background are combined into the other-BG component for clarity. The best-fit yields are given in Table III.

From the fit, the correlation between R and R^* is -0.56; each, in turn, is most strongly correlated with the D^{**} background yields, with 0.1 to 0.2 for R and ≈ 0.3 for R^*.

IX. SYSTEMATIC UNCERTAINTIES

The dominant systematic uncertainties arise from our limited understanding of the D^{**} background and from uncertainties in the fixed factors used in the fit. They are summarized in Table IV and itemized below.

In the table, "$D^{(*(*))}\ell\nu$ shapes" refers to uncertainties in the parameters that are used for the shape reweighting of semileptonic decays. The effect on the result is extracted by creating different sets of weights according to shape hypotheses from varying individual production parameters within their 1σ limits.

The D^{**} background has a strong influence on the extracted yield of the tau signal because the two components overlap in the M^2_{miss} spectrum. In addition to the shape uncertainties, there are uncertainties related to the poorly determined branching fractions to the different D^{**} states. The fit is therefore repeated several times: twice for each D^{**} state, with its branching fractions varied within its uncertainties. We use the following uncertainties: 42.3% for D^*_2, 34.6% for D^*_0, 14.9% for D_1, 36.2% for D'_1, and 100.0% for the radially excited $D(2S)$ and $D^*(2S)$. The best-fit variations in R are used as systematic uncertainties. They are combined quadratically and quoted in Table IV as "D^{**} composition."

All fixed factors used in the fit are varied by their uncertainty (arising from the MC sample size). The influence of the uncertainty of these factors is shown individually in Table IV. Most factors—especially the fixed yields—have little influence on the overall uncertainty; the efficiency ratios $f^{D^{+,0}}$ and $f^{D^{*+,0}}_{\text{eff}}$ and the cross-feed probability ratios $g^{+,0}$ give the largest contributions, comparable to the D^{**} composition and $D^{(*(*))}\ell\nu$ shape uncertainties.

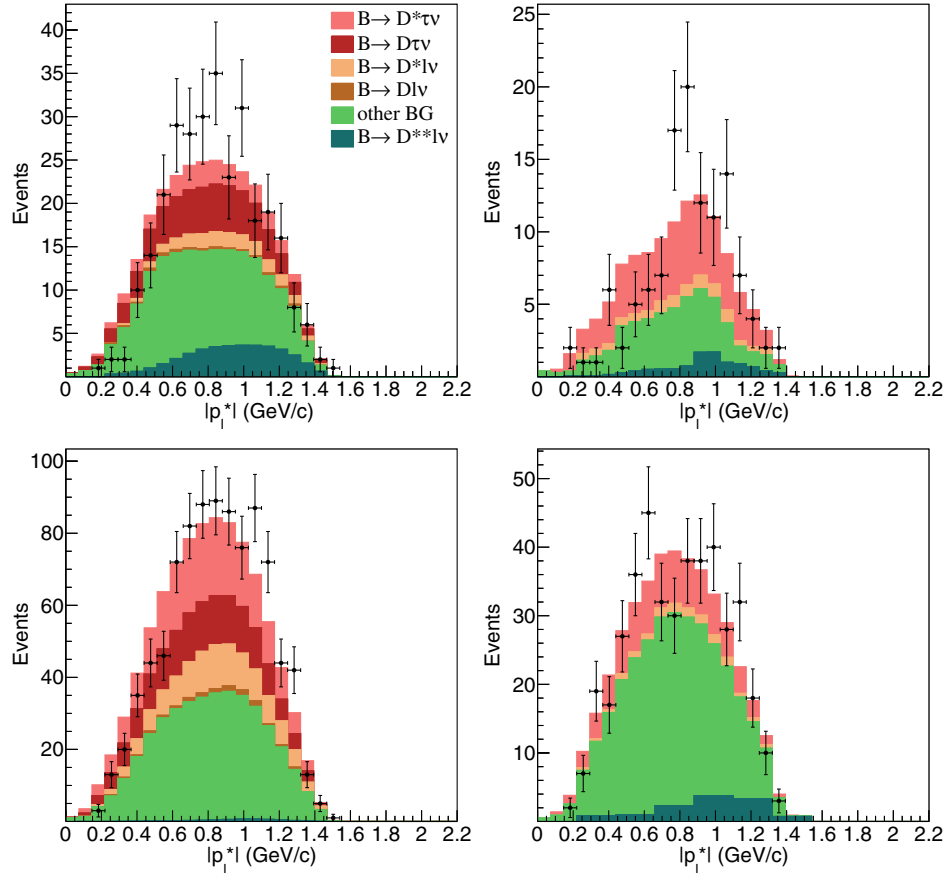To evaluate the effect of PDF uncertainties, the shapes of all components are modified and the fit is repeated. The

FIG. 5 (color online). Projections of the fit results and data points with statistical uncertainties in a signal-enhanced region of $M_{\mathrm{miss}}^2 > 2.0~\mathrm{GeV}^2/c^4$ in the p_ℓ^* dimension. Top left: $D^+\ell^-$; top right: $D^{*+}\ell^-$; bottom left: $D^0\ell^-$; bottom right: $D^{*0}\ell^-$.

nominal fit uses smoothed-histogram PDFs in M_{miss}^2; here, these are replaced by unsmoothed-histogram PDFs. The variation of the best-fit R is taken as the symmetric systematic uncertainty for "M_{miss}^2 shape" in Table IV. For the o'_{NB} alternate model, we replace the bifurcated Gaussians by kernel-estimator functions with adaptive bandwidth. Again, the deviation from the nominal fit value is taken as the symmetric systematic uncertainty for

TABLE III. Fit results and expected yields as derived from simulated data.

Sample	Component	Yield	Expected yield
$D^+\ell^-$	ℓ normalization	844 ± 34	870
$D^+\ell^-$	ℓ CF	924 ± 47	970
$D^+\ell^-$	D^{**} BG	108 ± 38	133
$D^0\ell^-$	ℓ normalization	2303 ± 64	2290
$D^0\ell^-$	ℓ CF	7324 ± 122	7440
$D^0\ell^-$	D^{**} BG	131 ± 81	210
$D^{*+}\ell^-$	ℓ normalization	1609 ± 43	1680
$D^{*+}\ell^-$	D^{**} BG	36 ± 18	76
$D^{*0}\ell^-$	ℓ normalization	2188 ± 60	2280
$D^{*0}\ell^-$	D^{**} BG	117 ± 39	40

TABLE IV. Overview of relative systematic uncertainties in percent. The last column gives the correlation between $R(D)$ and $R(D^*)$.

	$R(D)[\%]$	$R(D^*)[\%]$	Correlation
$D^{(*(*))}\ell\nu$ shapes	4.2	1.5	0.04
D^{**} composition	1.3	3.0	-0.63
Fake D yield	0.5	0.3	0.13
Fake ℓ yield	0.5	0.6	-0.66
D_s yield	0.1	0.1	-0.85
Rest yield	0.1	0.0	-0.70
Efficiency ratio f^{D^+}	2.5	0.7	-0.98
Efficiency ratio f^{D^0}	1.8	0.4	0.86
Efficiency ratio $f_{\mathrm{eff}}^{D^{*+}}$	1.3	2.5	-0.99
Efficiency ratio $f_{\mathrm{eff}}^{D^{*0}}$	0.7	1.1	0.94
CF double ratio g^+	2.2	2.0	-1.00
CF double ratio g^0	1.7	1.0	-1.00
Efficiency ratio f_{wc}	0.0	0.0	0.84
M_{miss}^2 shape	0.6	1.0	0.00
o'_{NB} shape	3.2	0.8	0.00
Lepton PID efficiency	0.5	0.5	1.00
Total	7.1	5.2	-0.32

"o'_{NB} shape" in Table IV. It is among the dominant systematic uncertainties.

The identification efficiencies for primary and secondary leptons are slightly different between simulated and real data. This difference affects the measurement by modifying the efficiency ratios. It has been calibrated for different lepton kinematics and run conditions using $J/\psi \to \ell^+\ell^-$ decays, leading to a 0.5% relative uncertainty in $R(D)$ and $R(D^*)$.

The correlations of $R(D)$ and $R(D^*)$ for each itemized systematic-uncertainty contribution are given in the last column of Table IV. These are calculated using 500 pseudoexperiments, with two exceptions: the shape uncertainties are assumed to be uncorrelated while the lepton ID efficiencies are assumed to be 100% correlated between $R(D)$ and $R(D^*)$. The total correlation of the systematic uncertainties is -0.32.

X. RESULTS AND DISCUSSION

The best-fit results, including systematic uncertainties, are

$$R(D) = 0.375 \pm 0.064 \pm 0.026 \quad (12)$$

$$R(D^*) = 0.293 \pm 0.038 \pm 0.015. \quad (13)$$

Figure 6 shows the exclusion level in the $R(D)$–$R(D^*)$ plane, based on the likelihood distribution that is convoluted with a correlated two-dimensional normal distribution according to the systematic uncertainties. The exclusions of the central values of the *BABAR* measurement [11] and the SM prediction as determined in Ref. [11] are comparably low at 1.4σ and 1.8σ, respectively. While our measurement does not favor one over the other, both measurements deviate in the same direction from the SM expectation.

We also use our fit procedure to test the compatibility of the data samples with the two-Higgs-doublet model of type II. For this purpose, we perform the analysis with the 2HDM MC sample with $\tan\beta/m_{H^+} = 0.5c^2/\mathrm{GeV}$ to extract probability density distributions. The best-fit values in this alternate model are

$$R(D) = 0.329 \pm 0.060(\mathrm{stat}) \pm 0.022(\mathrm{syst}) \quad (14)$$

$$R(D^*) = 0.301 \pm 0.039(\mathrm{stat}) \pm 0.015(\mathrm{syst}). \quad (15)$$

The effect on the measured $R(D^*)$ value is very small, but the measured value for $R(D)$ is significantly lower. For the prediction in the 2HDM of type II, we use formula (20) in Ref. [11]; the expected values are

$$R(D)_{\mathrm{2HDM}} = 0.590 \pm 0.125 \quad (16)$$

$$R(D^*)_{\mathrm{2HDM}} = 0.241 \pm 0.007. \quad (17)$$

Figure 7 shows the predictions of $R(D)$ and $R(D^*)$ as a function of $\tan\beta/m_{H^+}$ for the type II 2HDM, together with our results for the two studied values of 0 (SM) and $0.5c^2/\mathrm{GeV}$. In contrast to *BABAR*'s measurements, our results are compatible with the type II 2DHM in the $\tan\beta/m_{H^+}$ regions around $0.45c^2/\mathrm{GeV}$ and zero.

The observable most sensitive to NP extensions of the SM with a scalar charged Higgs is q^2. We estimate the signal q^2 distributions by subtracting the background, using the distributions from simulated data and the yields from the fit procedure, and correcting the distributions using efficiency estimations from simulated data. The $D^+\ell^-$ and
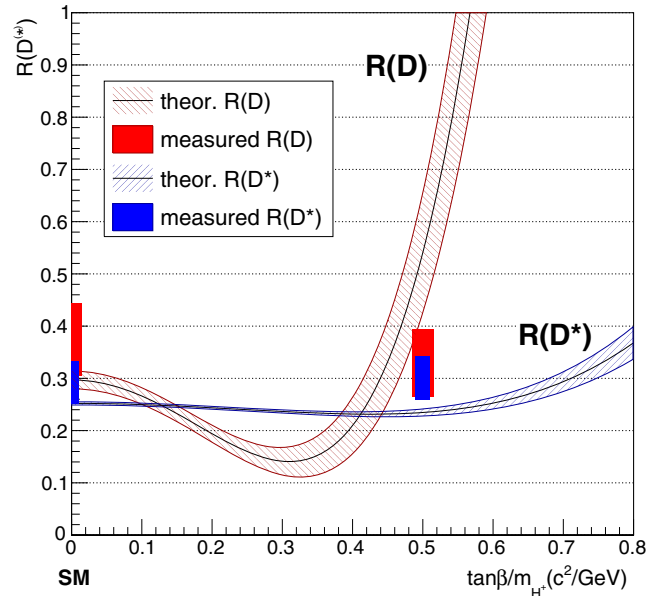
FIG. 6 (color online). Exclusion level of $R(D) - R(D^*)$ value assumptions in standard deviations, systematic uncertainties included.

FIG. 7 (color online). Theoretical predictions with 1σ error ranges for $R(D)$ (red) and $R(D^*)$ (blue) for different values of $\tan\beta/m_{H^+}$ in the 2HDM of type II. The fit results for $\tan\beta/m_{H^+} = 0.5c^2/\mathrm{GeV}$ and SM are shown with their 1σ ranges as red and blue bars with arbitrary width for better visibility.
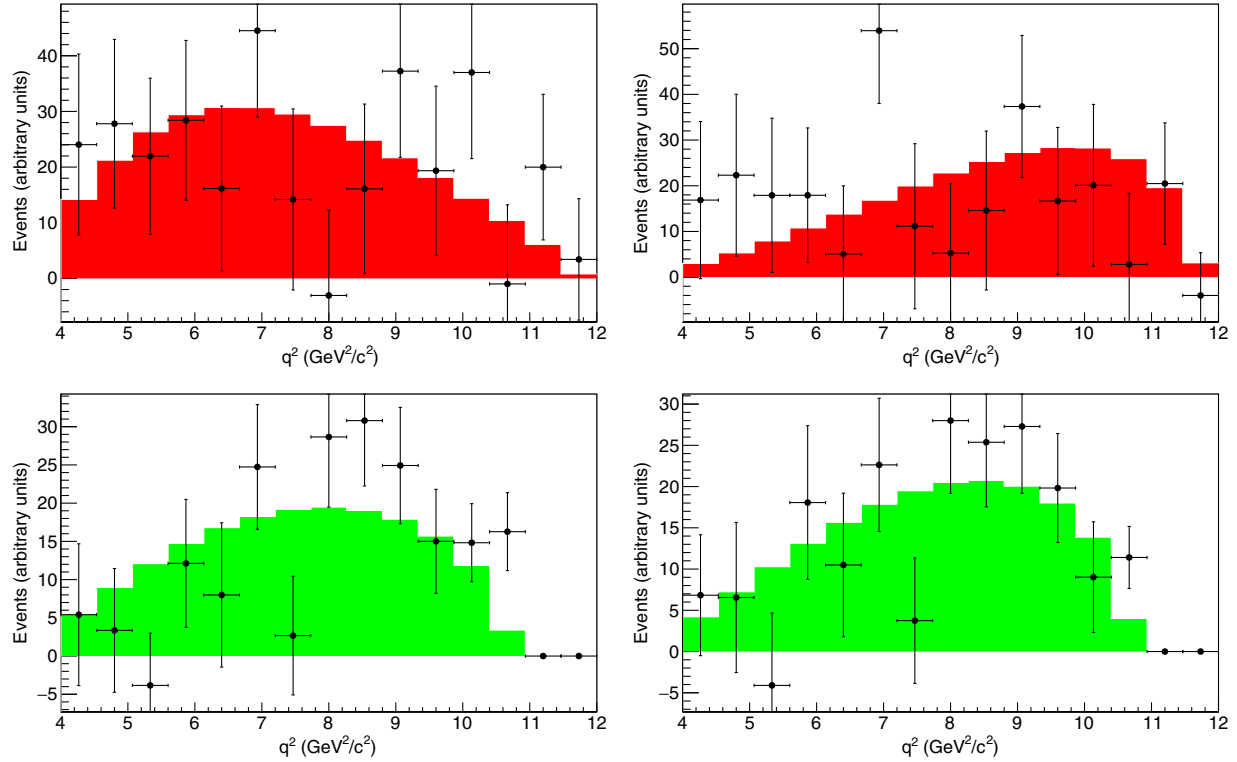
FIG. 8 (color online).   Background-subtracted q^2 distributions of the τ signal in the region of $M_{miss}^2 > 0.85$ GeV$^2/c^4$. The distributions are efficiency corrected and normalized to the fitted yield. The error bars show the statistical uncertainties. The histogram is the respective expected distribution from signal MC. Left: Standard Model result, right: Type-II 2HDM result with $\tan\beta/m_{H^+} = 0.5c^2$/GeV, top: $\bar{B} \to D\tau^-\bar{\nu}_\tau$, bottom: $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$

$D^0\ell^-$ samples and the $D^{*+}\ell^-$ and $D^{*0}\ell^-$ samples are combined to increase the available statistics, then the full procedure is repeated using the assumptions for the τ signal in a type II 2HDM model with $\tan\beta/m_{H^+} = 0.5c^2$/GeV. Figure 8 shows the measured background-subtracted and efficiency-corrected q^2 distributions for the SM and the NP point. As the signal yields are not extracted from fits to individual q^2 bins, the data distribution depends slightly on the signal model; the signal model can affect the background yields in the fit to uncorrected data, which are then subtracted. A χ^2 test shows that both hypotheses are compatible with our data with p-values for the SM distribution of 64% ($D\tau^-\bar{\nu}_\tau$) and 11% ($D^*\tau^-\bar{\nu}_\tau$), and for the NP distribution of 53% ($D\tau^-\bar{\nu}_\tau$) and 49% ($D^*\tau^-\bar{\nu}_\tau$).

XI. CONCLUSION

We present a measurement of the relative branching ratios $R(D^{(*)})$ of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ using the full $\Upsilon(4S)$ data recorded with the Belle detector. The results are

$$R(D) = 0.375 \pm 0.064(\text{stat}) \pm 0.026(\text{syst})$$

$$R(D^*) = 0.293 \pm 0.038(\text{stat}) \pm 0.015(\text{syst}).$$

In comparison to our previous preliminary results [9], which are superseded by this measurement, we utilize a more sophisticated fit strategy with an improved handling of the background from $\bar{B} \to D^{**}\ell^-\bar{\nu}_\ell$ events, impose an isospin constraint, and exploit a much higher tagging efficiency. By these methods, we reduce the statistical uncertainties by about a third and the systematic uncertainties by more than a half.

Our result lies between the SM expectation and the most recent measurement from the *BABAR* collaboration [11] and is compatible with both. It is also compatible with a 2HDM of type II in the region around $\tan\beta/m_{H^+} = 0.5c^2$/GeV, as illustrated in Figs. 7 and 8.

[1] Throughout this paper, the inclusion of the charge-conjugate decay mode is implied.

[2] K. A. Olive *et al.* (Particle Data Group Collaboration), Chin. Phys. C **38,** 090001 (2014).

[3] J. Chay, H. Georgi, and B. Grinstein, Phys. Lett. B **247,** 399 (1990); I. I. Y. Bigi, M. A. Shifman, N. G. Uraltsev, and A. I. Vainshtein, Phys. Rev. Lett. **71,** 496 (1993); C. W. Bauer, Z. Ligeti, M. Luke, A. V. Manohar, and M. Trott, Phys. Rev. D **70,** 094017 (2004).

[4] Y. Amhis *et al.* (Heavy Flavor Averaging Group (HFAG) Collaboration), arXiv:1412.7515.

[5] M. Tanaka, Z. Phys. C **67,** 321 (1995).

[6] H. Itoh, S. Komine, and Y. Okada, Prog. Theor. Phys. **114,** 179 (2005); U. Nierste, S. Trine, and S. Westhoff, Phys. Rev. D **78** (2008) 015006; M. Tanaka and R. Watanabe, Phys. Rev. D **82,** 034027 (2010); S. Fajfer, J. F. Kamenik, and I. Nisandzic, Phys. Rev. D **85,** 094025 (2012).

[7] A. Matyja *et al.* (Belle Collaboration), Phys. Rev. Lett. **99,** 191807 (2007).

[8] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **100,** 021801 (2008).

[9] I. Adachi *et al.* (Belle Collaboration), arXiv:0910.4301.

[10] A. Bozek *et al.* (Belle Collaboration), Phys. Rev. D **82,** 072005 (2010).

[11] J. P. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **109,** 101802 (2012); J. P. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. D **88,** 072012 (2013).

[12] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **115,** 111803 (2015).

[13] S. Fajfer, J. F. Kamenik, and I. Nisandzic, Phys. Rev. D **85,** 094025 (2012); J. F. Kamenik and F. Mescia, Phys. Rev. D **78,** 014003 (2008).

[14] J. A. Bailey *et al.* (Fermilab Lattice and MILC Collaborations), Phys. Rev. D **92,** 034506 (2015).

[15] H. Na, C. M. Bouchard, G. P. Lepage, C. Monahan, and J. Shigemitsu, Phys. Rev. D **92,** 054510 (2015).

[16] 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 (2013) and references therein.

[17] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479** (2002) 117; also see detector section in J. Brodzicka *et al.*, Prog. Theor. Exp. Phys. **2012,** 04D001 (2012).

[18] Z. Natkaniec *et al.* (Belle SVD2 Group), Nucl. Instrum. Methods Phys. Res., Sect. A **560,** 1 (2006).

[19] M. Feindt, F. Keller, M. Kreps, T. Kuhr, S. Neubauer, D. Zander, and A. Zupanc, Nucl. Instrum. Methods Phys. Res., Sect. A **654,** 432 (2011).

[20] S. H. Lee *et al.* (Belle Collaboration), Phys. Rev. Lett. **91,** 261801 (2003).

[21] K. Sumisawa *et al.* (Belle Collaboration), Phys. Rev. Lett. **95,** 061801 (2005).

[22] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462,** 152 (2001).

[23] R. Brun *et al.*, Report No. CERN DD/EE/84-1 (1987).

[24] M. Tanaka and R. Watanabe, Phys. Rev. D **87,** 034028 (2013).

[25] A. Sibidanov *et al.* (Belle Collaboration), Phys. Rev. D **88,** 032005 (2013).

[26] D. Scora and N. Isgur, Phys. Rev. D **52,** 2783 (1995).

[27] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, Phys. Rev. D **57,** 308 (1998).

[28] I. Caprini, L. Lellouch, and M. Neubert, Nucl. Phys. **B530,** 153 (1998).

[29] A. Zupanc *et al.* (Belle Collaboration), J. High Energy Phys. 09 (2013) 139.

[30] V. Blobel, histogram smoothing code, http://www.desy.de/˜blobel/splft.f.