Measurement of the branching ratio of $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ relative to $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with a semileptonic tagging method

Y. Sato,[44] T. Iijima,[45,44] K. Adamczyk,[50] H. Aihara,[74] D. M. Asner,[56] H. Atmacan,[40] T. Aushev,[43] R. Ayad,[67] T. Aziz,[68] V. Babu,[68] I. Badhrees,[67,28] A. M. Bakich,[66] V. Bansal,[56] P. Behera,[20] V. Bhardwaj,[17] B. Bhuyan,[19] J. Biswal,[24] G. Bonvicini,[79] A. Bozek,[50] M. Bračko,[37,24] D. Červenkov,[6] P. Chang,[49] V. Chekelian,[38] A. Chen,[47] B. G. Cheon,[12] K. Chilikin,[33,42] R. Chistov,[33,42] K. Cho,[29] V. Chobanova,[38] Y. Choi,[65] D. Cinabro,[79] M. Danilov,[42,33] N. Dash,[18] S. Di Carlo,[79] Z. Doležal,[6] D. Dutta,[68] S. Eidelman,[5,54] D. Epifanov,[74] H. Farhat,[79] J. E. Fast,[56] T. Ferber,[8] B. G. Fulsom,[56] V. Gaur,[68] N. Gabyshev,[5,54] A. Garmash,[5,54] P. Goldenzweig,[26] B. Golob,[34,24] D. Greenwald,[70] K. Hara,[14] T. Hara,[14,11] J. Hasenbusch,[4] K. Hayasaka,[52] H. Hayashii,[46] S. Hirose,[44] T. Horiguchi,[72] W.-S. Hou,[49] K. Inami,[44] A. Ishikawa,[72] R. Itoh,[14,11] Y. Iwasaki,[14] I. Jaegle,[13] H. B. Jeon,[31] D. Joffe,[27] T. Julius,[39] K. H. Kang,[31] Y. Kato,[44] P. Katrenko,[43,33] T. Kawasaki,[52] D. Y. Kim,[63] J. B. Kim,[30] K. T. Kim,[30] M. J. Kim,[31] S. H. Kim,[12] Y. J. Kim,[29] K. Kinoshita,[7] P. Kodyš,[6] S. Korpar,[37,24] D. Kotchetkov,[13] P. Krokovny,[5,54] T. Kuhr,[35] R. Kumar,[58] Y.-J. Kwon,[81] J. S. Lange,[10] C. H. Li,[39] L. Li,[60] Y. Li,[78] L. Li Gioi,[38] J. Libby,[20] D. Liventsev,[78,14] T. Luo,[57] M. Masuda,[73] T. Matsuda,[41] D. Matvienko,[5,54] K. Miyabayashi,[46] H. Miyata,[52] R. Mizuk,[33,42,43] G. B. Mohanty,[68] A. Moll,[38,69] H. K. Moon,[30] K. R. Nakamura,[14] E. Nakano,[55] M. Nakao,[14,11] T. Nanut,[24] K. J. Nath,[19] Z. Natkaniec,[50] M. Nayak,[79,14] K. Negishi,[72] N. K. Nisar,[68,1] S. Nishida,[14,11] S. Ogawa,[71] S. Okuno,[25] S. L. Olsen,[61] Y. Onuki,[74] P. Pakhlov,[33,42] G. Pakhlova,[33,43] B. Pal,[7] C.-S. Park,[81] S. Paul,[70] T. K. Pedlar,[36] L. Pesántez,[4] R. Pestotnik,[24] M. Petrič,[24] L. E. Piilonen,[78] M. V. Purohit,[64] J. Rauch,[70] A. Rostomyan,[8] M. Rozanska,[50] Y. Sakai,[14,11] S. Sandilya,[7] L. Santelj,[14] V. Savinov,[57] T. Schlüter,[35] O. Schneider,[32] G. Schnell,[2,16] C. Schwanda,[22] A. J. Schwartz,[7] Y. Seino,[52] K. Senyo,[80] O. Seon,[44] M. E. Sevior,[39] V. Shebalin,[5,54] C. P. Shen,[3] T.-A. Shibata,[75] J.-G. Shiu,[49] B. Shwartz,[5,54] F. Simon,[38,69] E. Solovieva,[33,43] S. Stanič,[53] M. Starič,[24] J. F. Strube,[56] T. Sumiyoshi,[76] M. Takizawa,[62,15,59] U. Tamponi,[23,77] F. Tenchini,[39] K. Trabelsi,[14,11] M. Uchida,[75] S. Uno,[14,11] P. Urquijo,[39] Y. Ushiroda,[14,11] Y. Usov,[5,54] C. Van Hulse,[2] G. Varner,[13] A. Vinokurova,[5,54] V. Vorobyev,[5,54] C. H. Wang,[48] M.-Z. Wang,[49] P. Wang,[21] Y. Watanabe,[25] K. M. Williams,[78] E. Won,[30] H. Yamamoto,[72] J. Yamaoka,[56] Y. Yamashita,[51] J. Yelton,[9] Y. Yook,[81] C. Z. Yuan,[21] Y. Yusa,[52] Z. P. Zhang,[60] V. Zhilich,[5,54] V. Zhukova,[42] V. Zhulanov,[5,54] and A. Zupanc[34,24]

(The Belle Collaboration)

[1] *Aligarh Muslim University, Aligarh 202002*
[2] *University of the Basque Country UPV/EHU, 48080 Bilbao*
[3] *Beihang University, Beijing 100191*
[4] *University of Bonn, 53115 Bonn*
[5] *Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090*
[6] *Faculty of Mathematics and Physics, Charles University, 121 16 Prague*
[7] *University of Cincinnati, Cincinnati, Ohio 45221*
[8] *Deutsches Elektronen–Synchrotron, 22607 Hamburg*
[9] *University of Florida, Gainesville, Florida 32611*
[10] *Justus-Liebig-Universität Gießen, 35392 Gießen*
[11] *SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193*
[12] *Hanyang University, Seoul 133-791*
[13] *University of Hawaii, Honolulu, Hawaii 96822*
[14] *High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[15] *J-PARC Branch, KEK Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[16] *IKERBASQUE, Basque Foundation for Science, 48013 Bilbao*
[17] *Indian Institute of Science Education and Research Mohali, SAS Nagar, 140306*
[18] *Indian Institute of Technology Bhubaneswar, Satya Nagar 751007*
[19] *Indian Institute of Technology Guwahati, Assam 781039*
[20] *Indian Institute of Technology Madras, Chennai 600036*
[21] *Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049*
[22] *Institute of High Energy Physics, Vienna 1050*
[23] *INFN - Sezione di Torino, 10125 Torino*
[24] *J. Stefan Institute, 1000 Ljubljana*
[25] *Kanagawa University, Yokohama 221-8686*
[26] *Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe*
[27] *Kennesaw State University, Kennesaw, Georgia 30144*
[28] *King Abdulaziz City for Science and Technology, Riyadh 11442*
[29] *Korea Institute of Science and Technology Information, Daejeon 305-806*
[30] *Korea University, Seoul 136-713*

[31] Kyungpook National University, Daegu 702-701
[32] École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015
[33] P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow 119991
[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] Middle East Technical University, 06531 Ankara
[41] University of Miyazaki, Miyazaki 889-2192
[42] Moscow Physical Engineering Institute, Moscow 115409
[43] Moscow Institute of Physics and Technology, Moscow Region 141700
[44] Graduate School of Science, Nagoya University, Nagoya 464-8602
[45] Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602
[46] Nara Women's University, Nara 630-8506
[47] National Central University, Chung-li 32054
[48] National United University, Miao Li 36003
[49] Department of Physics, National Taiwan University, Taipei 10617
[50] H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342
[51] Nippon Dental University, Niigata 951-8580
[52] Niigata University, Niigata 950-2181
[53] University of Nova Gorica, 5000 Nova Gorica
[54] Novosibirsk State University, Novosibirsk 630090
[55] Osaka City University, Osaka 558-8585
[56] Pacific Northwest National Laboratory, Richland, Washington 99352
[57] University of Pittsburgh, Pittsburgh, Pennsylvania 15260
[58] Punjab Agricultural University, Ludhiana 141004
[59] Theoretical Research Division, Nishina Center, RIKEN, Saitama 351-0198
[60] University of Science and Technology of China, Hefei 230026
[61] Seoul National University, Seoul 151-742
[62] Showa Pharmaceutical University, Tokyo 194-8543
[63] Soongsil University, Seoul 156-743
[64] University of South Carolina, Columbia, South Carolina 29208
[65] Sungkyunkwan University, Suwon 440-746
[66] School of Physics, University of Sydney, New South Wales 2006
[67] Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451
[68] Tata Institute of Fundamental Research, Mumbai 400005
[69] Excellence Cluster Universe, Technische Universität München, 85748 Garching
[70] Department of Physics, Technische Universität München, 85748 Garching
[71] Toho University, Funabashi 274-8510
[72] Department of Physics, Tohoku University, Sendai 980-8578
[73] Earthquake Research Institute, University of Tokyo, Tokyo 113-0032
[74] Department of Physics, University of Tokyo, Tokyo 113-0033
[75] Tokyo Institute of Technology, Tokyo 152-8550
[76] Tokyo Metropolitan University, Tokyo 192-0397
[77] University of Torino, 10124 Torino
[78] Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[79] Wayne State University, Detroit, Michigan 48202
[80] Yamagata University, Yamagata 990-8560
[81] Yonsei University, Seoul 120-749

(Dated: July 28, 2016)

We report a measurement of the ratio $\mathcal{R}(D^*) = \mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell)$, where ℓ denotes an electron or a muon. The results are based on a data sample containing 772×10^6 $B\bar{B}$ pairs recorded at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB e^+e^- collider. We select a sample of $B^0\bar{B}^0$ pairs by reconstructing both B mesons in semileptonic decays to $D^{*\mp}\ell^\pm$. We measure $\mathcal{R}(D^*) = 0.302 \pm 0.030(\text{stat}) \pm 0.011(\text{syst})$, which is within 1.6σ of the Standard Model theoretical expectation, where the standard deviation σ includes systematic uncertainties. We use this measurement to constrain several scenarios of new physics in a model-independent approach.

PACS numbers: 13.20.He, 14.40.Nd, 14.80.Da

I. INTRODUCTION

Semitauonic B meson decays of the type $b \to c\tau\nu_\tau$ [1] are sensitive probes to search for physics beyond the Standard Model (SM). Charged Higgs bosons, which ap-

pear in supersymmetry [2] and other models with at least two Higgs doublets [3], may contribute measurably to the decays due to the large mass of the τ. Similarly, leptoquarks [4], which carry both baryon number and lepton number, may also contribute to this process. The ratio of branching fractions

$$\mathcal{R}(D^{(*)}) = \frac{\mathcal{B}(\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell)} \quad (\ell = e, \mu), \qquad (1)$$

is typically used instead of the absolute branching fraction of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ to reduce several systematic uncertainties, such as those on the experimental efficiency, the magnitude of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$ and the semileptonic decay form factors. The SM calculations on these ratios predict $\mathcal{R}(D^*) = 0.252 \pm 0.003$ [5] and $\mathcal{R}(D) = 0.297 \pm 0.017$ [6, 7] with a precision of better than 2% and 6% for $\mathcal{R}(D^*)$ and $\mathcal{R}(D)$, respectively. Exclusive semitauonic B decays were first observed by Belle [8], with subsequent studies reported by Belle [9, 10], *BABAR* [7], and LHCb [11]. All the experimental results are consistent with each other, and the average values of Refs. [7, 10, 11] are $\mathcal{R}(D^*) = 0.322 \pm 0.018 \pm 0.012$ and $\mathcal{R}(D) = 0.391 \pm 0.041 \pm 0.028$ [12], which exceed the SM predictions by 3.0σ and 1.7σ, respectively. The combined analysis of $\mathcal{R}(D^*)$ and $\mathcal{R}(D)$, taking into account measurement correlations, finds that the deviation is 3.9σ from the SM prediction [12].

So far, measurements of $\mathcal{R}(D^{(*)})$ at the B factories have been performed using hadronic [7, 10] or inclusive tagging methods [8, 9]. Semileptonic tagging methods have been employed in studies of $B^- \to \tau^-\bar{\nu}_\tau$ decays and have been shown to have similar experimental precision to that of the hadronic tagging method [13, 14]. In this paper, we report the first measurement of $\mathcal{R}(D^*)$ using the semileptonic tagging method. We reconstruct signal $B^0\bar{B}^0$ events in modes where one B decays semitauonically, $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ followed by $\tau^- \to \ell^-\bar{\nu}_\ell\nu_\tau$ (referred to hereinafter as $B_{\rm sig}$), and the other B decays in a semileptonic channel $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ (referred to hereinafter as $B_{\rm tag}$). In order to form $\mathcal{R}(D^*)$, we also reconstruct normalization $B^0\bar{B}^0$ events in modes where both B mesons decay to $D^{*+}\ell^-\bar{\nu}_\ell$.

II. DETECTOR AND MC SIMULATION

We use the full $\Upsilon(4S)$ data sample containing 772×10^6 $B\bar{B}$ pairs recorded with the Belle detector [15] at the KEKB e^+e^- collider [16]. The Belle detector is a general-purpose magnetic spectrometer, which consists of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) comprising CsI(Tl) crystals. The devices are located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return yoke located outside the coil is instrumented to detect K_L^0 mesons and to identify muons (KLM). The detector is described in detail elsewhere [15].

To determine the reconstruction efficiency and probability density functions (PDF) for signal, normalization, and background processes, we use Monte Carlo (MC) simulated events, which are generated with the EvtGen event generator [17] and simulated with the GEANT3 package [18]. The MC samples for signal events are generated using the decay model based on the heavy quark effective theory (HQET) in Ref. [19]. The normalization mode is simulated using HQET, and reweighted according to the current world-average form factor values: $\rho^2 = 1.207 \pm 0.015 \pm 0.021$, $R_1 = 1.403 \pm 0.033$, and $R_2 = 0.854 \pm 0.020$ [12]. Background $B \to D^{**}\ell\nu_\ell$ events are simulated with the ISGW [20] model and reweighted to match the kinematics predicted by the LLSW model [21]. Here, D^{**} denotes the orbitally excited states D_1, D_2^*, D_1', and D_0^*. Radially excited states are neglected. We consider D^{**} decays to a $D^{(*)}$ and a pion, a ρ or an η meson, or a pair of pions, where branching fractions are assumed based on quantum-number, phase-space, and isospin arguments. The sample sizes of the signal, $B\bar{B}$, and continuum $q\bar{q}$ ($q = u, d, s, c$) production processes correspond to about 40, 10, and 6 times the integrated luminosity of the on-resonance $\Upsilon(4S)$ data sample, respectively.

III. EVENT SELECTION

Charged particle tracks are reconstructed with the SVD and CDC. All tracks other than $K_S^0 \to \pi^+\pi^-$ decay daughters are required to originate from near the interaction point (IP). Electrons are identified by a combination of the specific ionization (dE/dx) in the CDC, the ratio of the cluster energy in the ECL to the track momentum measured with the SVD and CDC, the response of the ACC, the shower shape in the ECL, and the match between the positions of the shower and the track at the ECL [22]. To recover bremsstrahlung photons from electrons, we add the four-momentum of each photon detected within 0.05 radians of the original track direction to the electron momentum. Muons are identified by the track penetration depth and hit distribution in the KLM [23]. Charged kaons are identified by combining information from the dE/dx in the CDC, the flight time measured with the TOF, and the response of the ACC [24]. We do not apply any particle identification criteria on charged pions.

Candidate K_S^0 mesons are formed by combining two oppositely charged tracks with pion mass hypotheses. We require their invariant mass to lie within 15 MeV/c^2 of the nominal K^0 mass [25], which corresponds to approximately 7σ, where σ denotes the resolution of the $\pi^+\pi^-$ invariant mass. We then impose the following additional requirements: both pion tracks must have a large distance of closest approach to the IP in the plane perpen-

dicular to the electron beam line; the pion tracks must intersect at a common vertex that is displaced from the IP; and the momentum vector of the K^0_S candidate should originate from the IP.

Neutral pion candidates are formed from pairs of photons with further criteria specific to whether the π^0 is from a D^{*+} decay or D decay. For neutral pions from D decays, we require the photon daughter energies to be greater than 50 MeV, the cosine of the angle in the laboratory frame between the two photons to be greater than zero, and the $\gamma\gamma$ invariant mass to be within -15 and $+10$ MeV/c^2 of the nominal π^0 mass [25], which corresponds to approximately $\pm 1.8\sigma$. Photons are measured as an energy cluster in the ECL with no associated charged tracks. A mass-constrained fit is then performed to obtain the π^0 momentum. For neutral pions from D^{*+} decays, which have lower energies, we require one photon to have an energy of at least 50 MeV and the other to have an energy of at least 20 MeV. We also require a narrow window around the di-photon invariant mass to compensate for the lower photon-energy requirement: within 10 MeV/c^2 of the nominal π^0 mass, which corresponds to approximately $\pm 1.6\sigma$.

Neutral D mesons are reconstructed in the following decay modes: $D^0 \to K^-\pi^+$, $K^0_S\pi^0$, K^+K^-, $\pi^+\pi^-$, $K^0_S\pi^+\pi^-$, $K^-\pi^+\pi^0$, $\pi^+\pi^-\pi^0$, $K^0_SK^+K^-$, $K^-\pi^+\pi^+\pi^-$, and $K^0_S\pi^+\pi^-\pi^0$. Charged D mesons are reconstructed in the following modes: $D^+ \to K^0_S\pi^+$, $K^-\pi^+\pi^+$, $K^0_S\pi^+\pi^0$, $K^+K^-\pi^+$, and $K^0_S\pi^+\pi^+\pi^-$. The combined reconstructed branching fractions are 37% and 22% for D^0 and D^+, respectively. For D decay modes without a π^0 in the final state, we require the invariant mass of the D candidates to be within 15 MeV/c^2 of the D^0 or D^+ mass, which corresponds to a window of approximately $\pm 3\sigma$. For modes with a π^0 in the final state, we require a wider invariant mass window: from -45 to $+30$ MeV/c^2 around the nominal D^0 mass for D^0 candidates, and from -36 to $+24$ MeV/c^2 around the nominal D^+ mass for D^+ candidates. These windows correspond to approximately $[-1.2\sigma, +1.8\sigma]$ and $[-1.0\sigma, +1.5\sigma]$, respectively, in resolution. Candidate D^{*+} mesons are formed by combining D^0 and π^+ candidates or D^+ and π^0 candidates. To improve the resolution of the D^*-D mass difference, ΔM, for the $D^{*+} \to D^0\pi^+$ decay mode, the charged pion track from the D^{*+} is refitted to the D^0 decay vertex. We require ΔM to be within 2.5 MeV/c^2 and 2.0 MeV/c^2, respectively, around the value of the nominal D^*-D mass difference for the $D^{*+} \to D^0\pi^+$ and $D^{*+} \to D^+\pi^0$ decay modes. These windows correspond to $\pm 3.2\sigma$ and $\pm 2.0\sigma$, respectively, in resolution. We apply a tighter window in the $D^{*+} \to D^+\pi^0$ decay mode to suppress a large contribution to the background arising from falsely reconstructed neutral pions.

To tag semileptonic B decays, we combine D^{*+} and lepton candidates of opposite electric charge and calculate the cosine of the angle between the momentum of the B meson and the $D^*\ell$ system in the $\Upsilon(4S)$ rest frame, under the assumption that only one massless particle is not reconstructed:

$$\cos\theta_{B\text{-}D^*\ell} \equiv \frac{2E_{\rm beam}E_{D^*\ell} - m_B^2c^4 - M_{D^*\ell}^2c^4}{2|\vec{p}_B| \cdot |\vec{p}_{D^*\ell}|c^2}, \quad (2)$$

where $E_{\rm beam}$ is the energy of the beam, and $E_{D^*\ell}$, $\vec{p}_{D^*\ell}$, and $M_{D^*\ell}$ are the energy, momentum, and mass, respectively, of the $D^*\ell$ system. The variable m_B is the nominal B meson mass [25], and $\vec{p}_B$ is the nominal B meson momentum. All variables are defined in the $\Upsilon(4S)$ rest frame. Figure 1 shows the $\cos\theta_{B\text{-}D^*\ell}$ distribution for signal and normalization decay modes in MC samples. Correctly reconstructed B candidates in the normalization decay mode are expected to have a value of $\cos\theta_{B\text{-}D^*\ell}$ between -1 and $+1$. Correctly reconstructed B candidates in the signal decay mode and falsely reconstructed B candidates tend to have values of $\cos\theta_{B\text{-}D^*\ell}$ below the physical region due to contributions from additional particles.

In each event we require two tagged B candidates that are opposite in flavor. Signal events may have the same flavor due to $B\bar{B}$ mixing; however, we veto such events as they lead to ambiguous $D^*\ell$ pair assignment and larger combinatorial background. We require that at most one B meson be reconstructed from a D^+ mode to avoid large background from fake neutral pions when forming D^* candidates. In each signal event, we assign the candidate with the lower value of $\cos\theta_{B\text{-}D^*\ell}$ (referred to hereinafter as $\cos\theta^{\rm sig}_{B\text{-}D^*\ell}$) as $B_{\rm sig}$. The probability of falsely assigning the $B_{\rm sig}$ as the $B_{\rm tag}$ for signal events is about 3%, according to MC simulation.

After the identification of the $B_{\rm sig}$ and $B_{\rm tag}$ candidates, we apply further background suppression criteria. On the tag side, we require $-2.0 < \cos\theta^{\rm tag}_{B\text{-}D^*\ell} < +1.5$ to select $B \to D^*\ell\nu_\ell$. On the signal side, we require the D^* momentum in the $\Upsilon(4S)$ rest frame to be less than 2.0 GeV/c, whlie, on the tag side, we require it to be less than 2.5 GeV/c, which accounts for the lepton mass difference. Finally, we require that events contain no extra charged tracks, K^0_S candidates, or π^0 candidates, which are reconstructed with the same criteria as those used for the D candidates.

At this stage, the probability of finding multiple candidates is 7%. When multiple candidates are found in an event, we select the most signal-like candidates based on the quality of vertex-constrained fits for the D mesons.

IV. SIGNAL, NORMALIZATION AND BACKGROUND SEPARATION

To separate reconstructed signal and normalization events, we employ a neural network using the NeuroBayes software package [26]. The variables used as inputs to the network are $\cos\theta^{\rm sig}_{B\text{-}D^*\ell}$, the missing mass squared $M^2_{\rm miss} = (2E_{\rm beam} - \sum_i E_i)^2/c^4 - |\sum_i \vec{p}_i|^2/c^2$, and the visible energy $E_{\rm vis} = \sum_i E_i$, where $(E_i, \vec{p}_i)$ is the four-momentum of particle i in the $\Upsilon(4S)$ rest frame. The

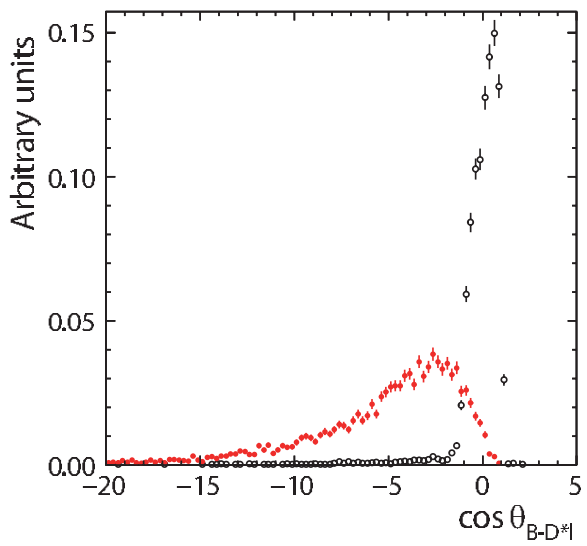


FIG. 1. The $\cos\theta_{B\text{-}D^*\ell}$ distributions for $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ (solid red circles) and $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ (open black circles) taken from MC simulation.

most powerful observable in separating signal and normalization is $\cos\theta^{\rm sig}_{B\text{-}D^*\ell}$. The neural network is trained using MC samples of signal and normalization events.

The dominant background contributions arise from events with misreconstructed $D^{(*)}$ mesons (denoted fakes). The sub-dominant contributions arise from two sources in which D^* mesons from both $B_{\rm sig}$ and $B_{\rm tag}$ are correctly reconstructed. One source is $B \to D^{**}\ell\nu_\ell$, where the D^{**} meson decays to $D^{(*)}$ and other particles. The other source is $B \to X_c D^*$ events, where one D^* meson is correctly reconstructed and the other charmed meson X_c decays semileptonically. If the hadrons in the semileptonic X_c decay are not identified, such events can mimic signal. Similarly, events in which X_c is a D_s^+ meson decaying into $\tau^+\nu_\tau$ can also mimic signal.

To separate signal and normalization events from background processes, we place a criterion on the sum of the energies of neutral clusters detected in the ECL that are not associated with reconstructed particles, denoted as $E_{\rm ECL}$. To mitigate the effects of photons related to beam background in the energy sum, we only include clusters with energies greater than 50, 100, and 150 MeV, respectively, from the barrel, forward, and backward calorimeter regions, defined in Ref. [15]. Signal and normalization events peak near zero in $E_{\rm ECL}$, while background events can populate a wider range. We require $E_{\rm ECL}$ to be less than 1.2 GeV.

V. MC CALIBRATION

To improve the accuracy of the MC simulation, we apply a series of calibration factors determined from control sample measurements. The lepton identification efficiencies are separately corrected for electrons and for muons to account for differences between the detector responses in data and MC. Correction factors for lepton identification efficiencies are evaluated as a functions of the momentum and direction of the lepton using $e^+e^- \to e^+e^-\ell^+\ell^-$ and $J/\psi \to \ell^+\ell^-$ decays. We reweight events to account for differing $D^{(*)}$ yields between data and MC.

The differing yields of correctly reconstructed $D^{(*)}$ mesons in data and MC affect the $\mathcal{R}(D^*)$ measurement, as it biases the determination of the background contribution. Calibration factors for events with both correctly- and falsely-reconstructed D mesons are estimated for each D meson decay mode using a two-dimensional fit to M_D. For this calibration, we use samples with all selection criteria other than D mass and ΔM applied. A two-dimensional PDF is constructed by taking the product of the one-dimensional functions for M_D. The PDF in each dimension is the sum of a signal component and a background component modelled with a first-order polynomial. The signal component is a triple Gaussian for D^0 decay modes without a π^0 and a Crystal Ball function [27] plus a Gaussian for D^0 decay modes with a π^0 and D^+ decay modes. In this calibration, we do not distinguish signal and tag sides. To estimate calibration factors for specific D decay modes, we fit samples in which one D meson is reconstructed in a specific mode while the other is reconstructed in any signal mode. From the ratios of data to MC samples in signal and background yields, we derive calibration factors of the specific decay mode for events with correctly and falsely reconstructed D mesons. We cannot independently determine calibration factors for all D meson decay modes as we use other decay modes when we calibrate each specific decay mode of a given D meson. To estimate all the calibration factors correctly, we first perform the two-dimensional fit for each decay mode without weighting factors, and then iterate the fits using resultant weighting factors until all calibration factors converge.

Similarly, we estimate calibration factors for events with correctly and falsely reconstructed D^* mesons from a two-dimensional fit to ΔM. Calibration factors for events with correctly and falsely reconstructed D^* mesons are separately estimated for subsequent decay to D^0 and D^+ mesons.

VI. MAXIMUM LIKELIHOOD FIT

We extract the yields of the signal and normalization processes from a two-dimensional extended maximum-likelihood fit to neural network classifier output $\mathcal{O}_{NB}$ and $E_{\rm ECL}$. The likelihood function consists of five components: signal, normalization, fake $D^{(*)}$ events, $B \to D^{**}\ell\nu_\ell$, and other backgrounds (predominantly from $B \to X_c D^*$). The PDFs of all components are determined from MC simulation. There are significant

correlations between $\mathcal{O}_{NB}$ and $E_{\rm ECL}$ in the normalization and background components, but not for the signal. We therefore construct the normalization and background PDFs using two-dimensional histograms and apply a smoothing procedure to account for its limited statistical power [28]. The signal PDF is the product of one-dimensional histograms in $\mathcal{O}_{NB}$ and $E_{\rm ECL}$.

Three parameters are floated in the final fit: the yields of the signal, normalization, and $B \to D^{**}\ell\nu_\ell$ components. The yield of fake $D^{(*)}$ events is fixed to the value estimated from sidebands in the ΔM distribution. Since the PDF shape of fake $D^{(*)}$ events depends on the composition of signal, normalization, $B \to D^{**}\ell\nu_\ell$, and other backgrounds, the relative contributions of these processes to the fake $D^{(*)}$ component are described as a function of the three fit parameters. The yields of other backgrounds are fixed to the values expected from MC simulation. The ratio $\mathcal{R}(D^*)$ is given by the formula:

$$\mathcal{R}(D^*) = \frac{1}{2\mathcal{B}(\tau^- \to \ell^- \bar{\nu}_\ell \nu_\tau)} \cdot \frac{\varepsilon_{\rm norm}}{\varepsilon_{\rm sig}} \cdot \frac{N_{\rm sig}}{N_{\rm norm}}, \quad (3)$$

where $\varepsilon_{\rm sig(norm)}$ and $N_{\rm sig(norm)}$ are the reconstruction efficiency and yields of signal (normalization) events. We use $\mathcal{B}(\tau^- \to \ell^- \bar{\nu}_\ell \nu_\tau) = 0.176 \pm 0.003$ as the average of the world averages for $\ell = e$ and $\ell = \mu$ [25]. The ratio of efficiencies, $\varepsilon_{\rm norm}/\varepsilon_{\rm sig}$, is estimated to be 1.289 ± 0.015 from MC simulation. The difference between reconstruction efficiencies of signal and normalization events arises from their distinct lepton momentum distributions, and the different event criteria on the D^* momenta.

VII. PDF VALIDATION

We validate the PDFs used in the fitting procedure by analysing various control samples. For fake $D^{(*)}$ events, we study the ΔM sidebands, where we find good agreement in both $\mathcal{O}_{NB}$ and $E_{\rm ECL}$. For $B \to D^*\ell\nu_\ell$ decays, we require one B meson to be reconstructed with the hadronic tagging method [29] and the other B meson reconstructed with the nominal criteria of this analysis. We find good agreement between data and MC in the $E_{\rm ECL}$, $M^2_{\rm miss}$, and $E_{\rm vis}$ distributions, but small discrepancies in the $\cos\theta_{B\text{-}D^*\ell}$ distributions, which we incorporate as a systematic uncertainty.

VIII. SYSTEMATIC UNCERTAINTIES

To estimate the systematic uncertainties on $\mathcal{R}(D^*)$, we vary every fixed parameter in turn by one standard deviation and repeat the fit. The systematic uncertainties are summarized in Table I. The dominant systematic uncertainty arises from the limited size of the MC samples: to estimate this uncertainty, we recalculate PDFs for signal, normalization, fake $D^{(*)}$ events, $B \to D^{**}\ell\nu_\ell$, and other backgrounds by generating toy MC samples from the nominal PDFs according to Poisson statistics and repeat the fit with the new PDFs.

Small discrepancies between the data and MC are found in the $\cos\theta_{B\text{-}D^*\ell}$ distributions in the hadronic tagged samples. We estimate it as "PDF shape of the normalization in $\cos\theta_{B\text{-}D^*\ell}$" in Table I by correcting the $\cos\theta_{B\text{-}D^*\ell}$ distribution in MC samples according to the observed discrepancies, and repeating the fit.

The branching fractions of the $B \to D^{**}\ell\nu_\ell$ decay modes and the decays of the D^{**} mesons are not well known and therefore contribute significantly to the total PDF uncertainty for $B \to D^{**}\ell\nu_\ell$ decays. The branching fraction of each $B \to D^{**}\ell\nu_\ell$ decay is varied within its uncertainty. The uncertainties are assumed to be $\pm 6\%$ for D_1, $\pm 10\%$ for D_2^*, $\pm 83\%$ for D_1', and $\pm 100\%$ for D_0^*, including the limited knowledge of the D^{**} decays. We also consider the impact of contributions from radially excited $D(2S)$ and $D^*(2S)$, where we assume the branching fractions of $B \to D^{(*)}(2S)\ell\nu_\ell$ to be as large as 0.5%.

The yield of fake D^* events is fixed to the value estimated from sidebands in the ΔM distribution. We vary this yield within its uncertainties. We also vary the calibration factors for D meson decay modes within their uncertainties for events with falsely reconstructed $D^{(*)}$ events.

The yields of other background processes, predominantly from $B \to X_c D^*$ events, are fixed to the values estimated from MC simulation. We consider variations on the yield and shape of the PDF of these background processes within their measured uncertainties. The uncertainties for the $B \to X_c D^*$ channels are assumed to be $\pm 8\%$ for $B \to D_s^* D^{*-}$, $\pm 14\%$ for $B \to D_s D^{*-}$, $\pm 8\%$ for $B \to D^{*+} D^{*-}$, and $\pm 10\%$ for $B \to D^+ D^{*-}$. Furthermore, we add an uncertainty of $\pm 4\%$ due to the size of the MC sample. We determine the uncertainty from the branching fraction of $D_s \to \tau\nu_\tau$ decay (which may peak near the signal in the $E_{\rm ECL}$ distribution) to be negligible.

The reconstruction efficiency ratio of signal to normalization events is varied within its uncertainty, which is limited by the size of the MC samples for signal events.

We include other minor systematic uncertainties from two sources: one related to the parameters that are used for the reweighting of the semileptonic $B \to D^{(*(*))}\ell\nu_\ell$ decays from the ISGW model to the LLSW model; and the other from the branching fraction of $\tau^- \to \ell^- \bar{\nu}_\ell \nu_\tau$ decay [25]. The total systematic uncertainty is estimated by summing the above uncertainties in quadrature.

IX. RESULTS

The $\mathcal{O}_{NB}$ and $E_{\rm ECL}$ projections of the fitted distributions are shown in Figure 2. The yields of signal and normalization events are measured to be 231 ± 23(stat) and 2800 ± 57(stat), respectively. The ratio $\mathcal{R}(D^*)$ is found to be

$$\mathcal{R}(D^*) = 0.302 \pm 0.030 \pm 0.011, \quad (4)$$

TABLE I. Summary of the systematic uncertainties on $\mathcal{R}(D^*)$ for electron and muon modes combined and separated. The uncertainties are relative and are given in percent.

	$\mathcal{R}(D^*)$ [%]		
Sources	$\ell^{\rm sig} = e, \mu$	$\ell^{\rm sig} = e$	$\ell^{\rm sig} = \mu$
MC size for each PDF shape	2.2	2.5	3.9
PDF shape of the normalization in $\cos\theta_{B\text{-}D^*\ell}$	$^{+1.1}_{-0.0}$	$^{+2.1}_{-0.0}$	$^{+2.8}_{-0.0}$
PDF shape of $B \to D^{**}\ell\nu_\ell$	$^{+1.0}_{-1.7}$	$^{+0.7}_{-1.3}$	$^{+2.2}_{-3.3}$
PDF shape and yields of fake $D^{(*)}$	1.4	1.6	1.6
PDF shape and yields of $B \to X_c D^*$	1.1	1.2	1.1
Reconstruction efficiency ratio $\varepsilon_{\rm norm}/\varepsilon_{\rm sig}$	1.2	1.5	1.9
Modeling of semileptonic decay	0.2	0.2	0.3
$\mathcal{B}(\tau^- \to \ell^- \bar{\nu}_\ell \nu_\tau)$	0.2	0.2	0.2
Total systematic uncertainty	$^{+3.4}_{-3.5}$	$^{+4.1}_{-3.7}$	$^{+5.9}_{-5.8}$

where the first uncertainty is statistical and the second systematic (and likewise for all following results).

We calculate the statistical significance of the signal as $\sqrt{-2\ln(\mathcal{L}_0/\mathcal{L}_{\rm max})}$, where $\mathcal{L}_{\rm max}$ and $\mathcal{L}_0$ are the maximum likelihood and the likelihood obtained assuming zero signal yield, respectively. We obtain a statistical significance of 13.8σ. We also estimate the compatibility of the measured value of $\mathcal{R}(D^*)$ and the SM prediction. The effect of systematic uncertainties is included by convolving the likelihood function with a Gaussian distribution. We conclude that our result is larger than the SM prediction by 1.6σ.

X. CROSS-CHECKS

To determine the consistency of the measured value of $\mathcal{R}(D^*)$ among τ final states, we divide the data samples by lepton flavor on the signal side and fit them separately. All PDFs for electron and muon channels are separately constructed from the MC samples. The efficiency ratios $\varepsilon_{\rm norm}/\varepsilon_{\rm sig}$ are estimated to be 1.107 ± 0.016 and 1.591 ± 0.030 for electron and muon channels of the tau decays, respectively. We obtain

$$\mathcal{R}(D^*) = 0.311 \pm 0.038 \pm 0.013 \ \ (\ell^{\rm sig} = e), \quad (5)$$
$$\mathcal{R}(D^*) = 0.304 \pm 0.051 \pm 0.018 \ \ (\ell^{\rm sig} = \mu). \quad (6)$$

The systematic uncertainties are summarized in Table I. These two results are consistent with each other.

To study $B \to D^{**}\ell\nu_\ell$ background contributions, we require an additional π^0 with respect to the nominal event selection. In this control sample, we calculate $E'_{\rm ECL}$, which is defined as the remaining energy after the energy deposit from the additional π^0 is removed from $E_{\rm ECL}$. The $B \to D^{**}\ell\nu_\ell$ background contributions are extracted from the control samples using the nominal fitting method, replacing $E_{\rm ECL}$ with $E'_{\rm ECL}$, which is defined as $E_{\rm ECL}$ without the energy deposit from the additional π^0 [30]. We find consistent results for the branching fractions of $B \to D^{**}\ell\nu_\ell$ in the control and signal samples.

XI. NEW PHYSICS COMPATIBILITY TESTS

Assuming all neutrinos are left-handed, the effective Hamiltonian that contains all possible four-fermion operators for the $b \to c\tau\nu_\tau$ decay can be described as follows [19]:

$$\mathcal{H}_{\rm eff} = \frac{4G_F}{\sqrt{2}} V_{cb} \left[\mathcal{O}_{V_1} + \sum_{X=S_1,S_2,V_1,V_2,T} C_X \mathcal{O}_X \right], \quad (7)$$

where the four-Fermi operators, $\mathcal{O}_X$, are defined as

$$\mathcal{O}_{S_1} = (\bar{c}_L b_R)(\bar{\tau}_R \nu_{\tau L}), \quad (8)$$
$$\mathcal{O}_{S_2} = (\bar{c}_R b_L)(\bar{\tau}_R \nu_{\tau L}), \quad (9)$$
$$\mathcal{O}_{V_1} = (\bar{c}_L \gamma^\mu b_L)(\bar{\tau}_L \gamma_\mu \nu_{\tau L}), \quad (10)$$
$$\mathcal{O}_{V_2} = (\bar{c}_R \gamma^\mu b_R)(\bar{\tau}_L \gamma_\mu \nu_{\tau L}), \quad (11)$$
$$\mathcal{O}_T = (\bar{c}_R \sigma^{\mu\nu} b_L)(\bar{\tau}_R \sigma_{\mu\nu} \nu_{\tau L}), \quad (12)$$

and the C_X parameters are the Wilson coefficients of $\mathcal{O}_X$. We investigate the compatibility of the data samples with new physics using a model-independent approach, separately examining the impact of each operator. In each new-physics scenario, we take into account changes in the efficiency and fit PDF shapes using dedicated signal simulation. We set the Wilson coefficients to be real in all cases. Since $\mathcal{O}_{V_1}$ is just the SM operator, it would change only $\mathcal{R}(D^*)$, but not the kinematic distributions. In the type-II two-Higgs doublet model (2HDM), the relevant Wilson coefficients are given as $C_{S_1} = -m_b m_\tau \tan^2\beta / m^2_{H^+}$ and $C_{S_2} = -m_c m_\tau / m^2_{H^+}$, where $\tan\beta$ is the ratio of the vacuum expectation values of the two Higgs doublets, and m_b, m_c, m_τ, and m_{H^+} are the masses of the b quark, c quark, τ lepton, and charged Higgs boson. Since the contribution from C_{S_2} is almost negligibly small except for the light charged Higgs boson, we neglect the contribution from C_{S_2} in the type-II 2HDM.

Various leptoquark models have been presented to explain anomalies in $\mathcal{R}(D^{(*)})$ in Ref. [4]. In addition to the model-independent study, we study two representative models: R_2 and S_1. Model R_2 contains scalar

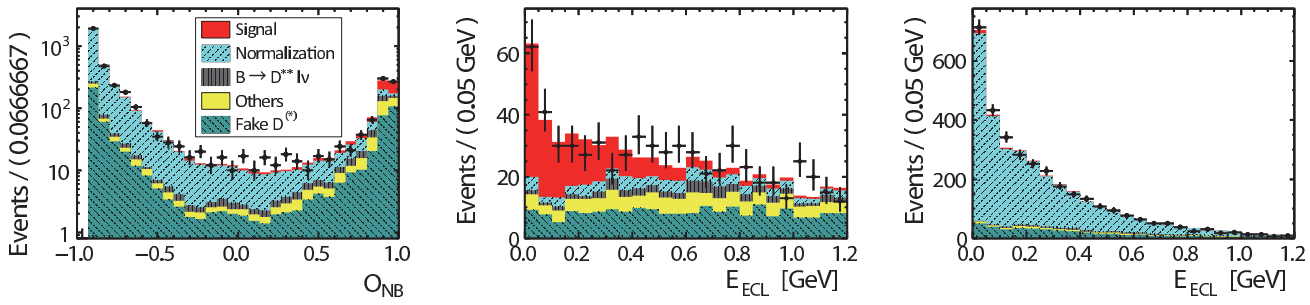


FIG. 2. Projections of the fit results with data points overlaid for (left) the neural network classifier output, $\mathcal{O}_{NB}$, and the $E_{\rm ECL}$ distribution in (center) the signal-enhanced region, $\mathcal{O}_{NB} > 0.8$, and (right) the normalization-enhanced region, $\mathcal{O}_{NB} < 0.8$. The background categories are described in detail in the text, where "others" refers to predominantly $B \to X_c D^*$ decays.

leptoquarks of the type $(3,2)_{7/6}$ using the notation $(SU(3)_c, SU(2)_L)_Y$, where $SU(3)_c$ is the representation under the generators of QCD, $SU(2)_L$ is the representation under the generators of weak isospin, and Y is the weak hypercharge. Model S_1 contains leptoquarks of the type $(3^*,1)_{1/3}$. In these leptoquark models, the relevant Wilson coefficients are related by $C_{S_2} = +7.8 C_T$ for the R_2-type leptoquark model and $C_{S_2} = -7.8 C_T$ for the S_1-type leptoquark model at the b quark mass scale, assuming a leptoquark mass scale of 1 TeV/c^2. Although the V_1 operator can appear independently of the S_2 and T operators in the S_1-type leptoquark model, we assume no contribution from the V_1 operator in this study.

Figure 3 shows the dependence of the efficiency and measured value of $\mathcal{R}(D^*)$ as a function of the values of the respective parameters in the type-II 2HDM and the R_2-type leptoquark model. Efficiency variations for other scenarios are shown in Ref. [30]. We find that efficiencies increase by up to 17% for $\mathcal{O}_{V_2}$ and $\mathcal{O}_T$, mainly due to the variation of the D^* momentum distribution. Similarly, the efficiencies increase by up to 16% and 11% in R_2- and S_1-type leptoquark models, respectively, which include contributions from $\mathcal{O}_T$. In other scenarios, the efficiency variation is 6% or less. Figure 4 shows the dependency of the measured values of $\mathcal{R}(D^*)$ on the values of the respective parameters in the type-II 2HDM and the R_2-type leptoquark model. The allowed regions with 68% confidence level (C.L.) of the respective parameters are summarized in Table II.

In Refs. [7] and [10], the $q^2 \equiv (p_B - p_{D^*})^2$ spectra are examined in order to study the effects of new physics beyond the SM. Since q^2 cannot be calculated here due to the neutrino in the decay of the $B_{\rm tag}$, we use instead the momenta of the D^* and the ℓ in $B_{\rm sig}$ at the $\Upsilon(4S)$ rest frame. Figure 5 shows the momentum distributions of the background-subtracted data in the region of $\mathcal{O}_{NB} > 0.8$ and $E_{\rm ECL} < 0.5$ GeV for the SM, type-II 2HDM with $\tan\beta/m_{H^+} = 0.7$ GeV^{-1}, and the R_2-type leptoquark model with $C_T = +0.36$. The PDF shapes of background events are taken from MC simulation and normalized to the yields obtained by the fitting. Table III shows p values for all scenarios, where we include only the statistical uncertainty. We find our data are compatible with the SM and additional contributions from scalar and vector operators; large additional contributions from tensor operator or the R_2- and S_1-type leptoquark models are disfavored.

XII. CONCLUSION

In conclusion, we report the first measurement of $\mathcal{R}(D^*)$ with a semileptonic tagging method using a data sample containing $772 \times 10^6 B\bar{B}$ pairs collected with the Belle detector. The result is

$$\mathcal{R}(D^*) = 0.302 \pm 0.030 \pm 0.011, \qquad (13)$$

which is within 1.6σ of the SM prediction including systematic uncertainties, and is in good agreement with other measurements by Belle [8–10], *BABAR* [7], and LHCb [11]. The result is statistically independent of earlier Belle measurements. We investigate the compatibility of the data samples with new physics in a model-independent method by adding the operators one by one. We also study two types of leptoquark models. We find our data allow the additional contributions from scalar and vector operators while disfavoring large additional contributions from a tensor operator with $+0.34 < C_T < +0.39$, an R_2-type leptoquark model with $+0.34 < C_T < +0.38$, or an S_1-type leptoquark model with $+0.22 < C_T < +0.28$, when considering the impact on the decay kinematics.

XIII. ACKNOWLEDGEMENTS

We thank Y. Sakaki, R. Watanabe, and M. Tanaka for their invaluable suggestions. This work was supported in part by a Grant-in-Aid for JSPS Fellows (No.13J03438) and a Grant-in-Aid for Scientific Research (S) "Probing New Physics with Tau-Lepton" (No.26220706). We thank the KEKB group for the excellent operation of the

TABLE II. Allowed regions with 68% C.L. of Wilson coefficients [19]. $-4.25 < C_{S_1} < -3.09$ corresponds to 0.65 GeV^{-1} $<$ $\tan\beta/m_{H^+} < 0.76$ GeV^{-1} in type-II 2HDM, where $m_b = 4.20$ GeV/c^2, $m_c = 0.901$ GeV/c^2 [31] and $m_\tau = 1.77682$ GeV/c^2 [25] are used.

Models or operators	Parameters	Allowed regions (68% C.L.)
$\mathcal{O}_{S_1}$	C_{S_1}	$[-4.25, -3.09]$, $[+0.44, +1.57]$
$\mathcal{O}_{S_2}$	C_{S_2}	$[-1.56, -0.43]$, $[+3.12, +4.28]$
$\mathcal{O}_{V_1}$	C_{V_1}	$[-2.15, -2.03]$, $[+0.05, +0.15]$
$\mathcal{O}_{V_2}$	C_{V_2}	$[-0.17, 0.00]$, $[+1.83, +1.96]$
$\mathcal{O}_T$	C_T	$[-0.06, -0.01]$, $[+0.34, +0.39]$
R_2-type leptoquark	$C_T(= +C_{S_2}/7.8)$	$[-0.05, -0.01]$, $[+0.34, +0.38]$
S_1-type leptoquark	$C_T(= -C_{S_2}/7.8)$	$[-0.07, -0.01]$, $[+0.22, +0.28]$

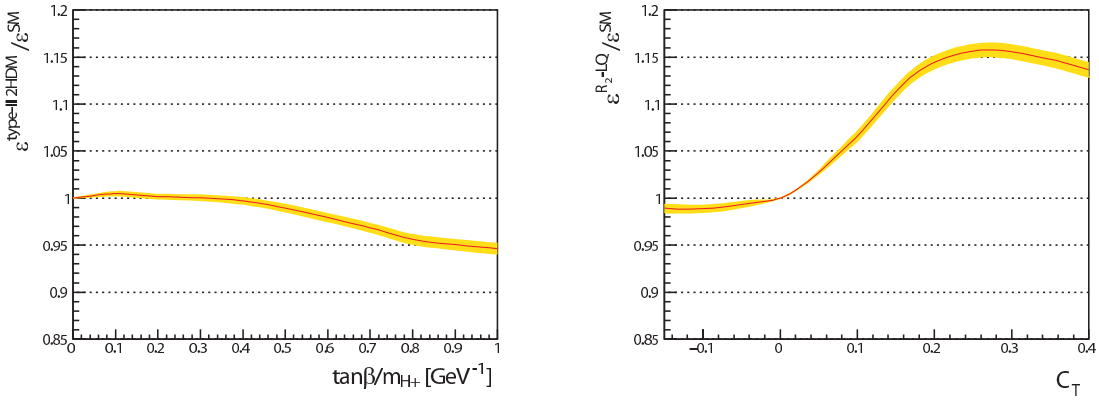


FIG. 3. The efficiencies for (left) the type-II 2HDM and (right) R_2-type leptoquark model with respect to the SM value.

TABLE III. p values for each scenario from the momentum distributions of the D^* or the lepton on the signal side in the $\Upsilon(4S)$ rest frame, where we include only the statistical uncertainty.

Model or operator	Parameter	p values [%]	
		p_{D^*}	p_ℓ
SM		37.6	25.8
Type-II 2HDM	$\frac{\tan\beta}{m_{H^+}} = 0.7$ GeV^{-1}	37.9	22.5
$\mathcal{O}_{V_2}$	$C_{V_2} = +1.88$	24.1	18.6
$\mathcal{O}_T$	$C_T = +0.36$	0.9	19.2
R_2-type leptoquark model	$C_T = +0.36$	1.4	16.2
S_1-type leptoquark model	$C_T = +0.26$	1.1	15.4

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, No. 11475187 and No. 11575017; 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, the Excellence Cluster Universe, 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 Re-

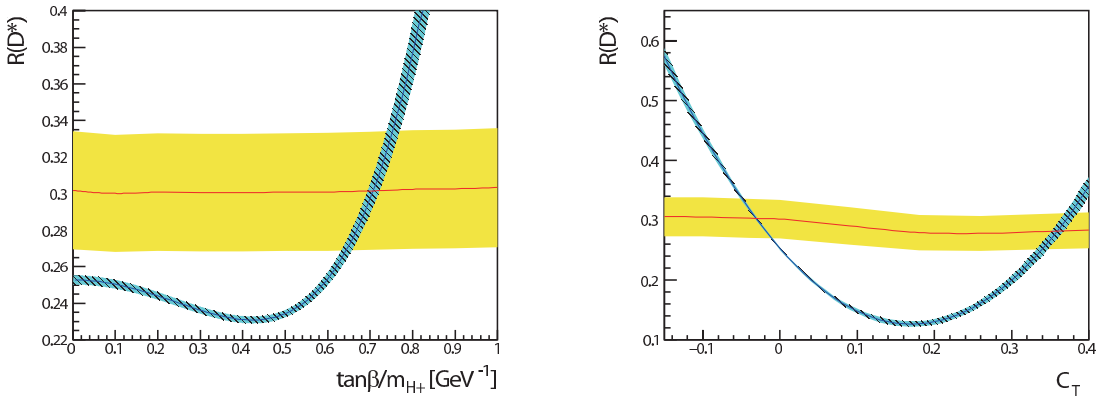


FIG. 4. The measured values of $\mathcal{R}(D^*)$ for (left) the type-II 2HDM and (right) R_2-type leptoquark models, where central values are given as the solid (red) curves and the 1σ uncertainties are given as the shaded (yellow) regions. The theoretical predictions and their 1σ uncertainties are shown as solid (blue) curves and hatched (light blue) regions, respectively [19].

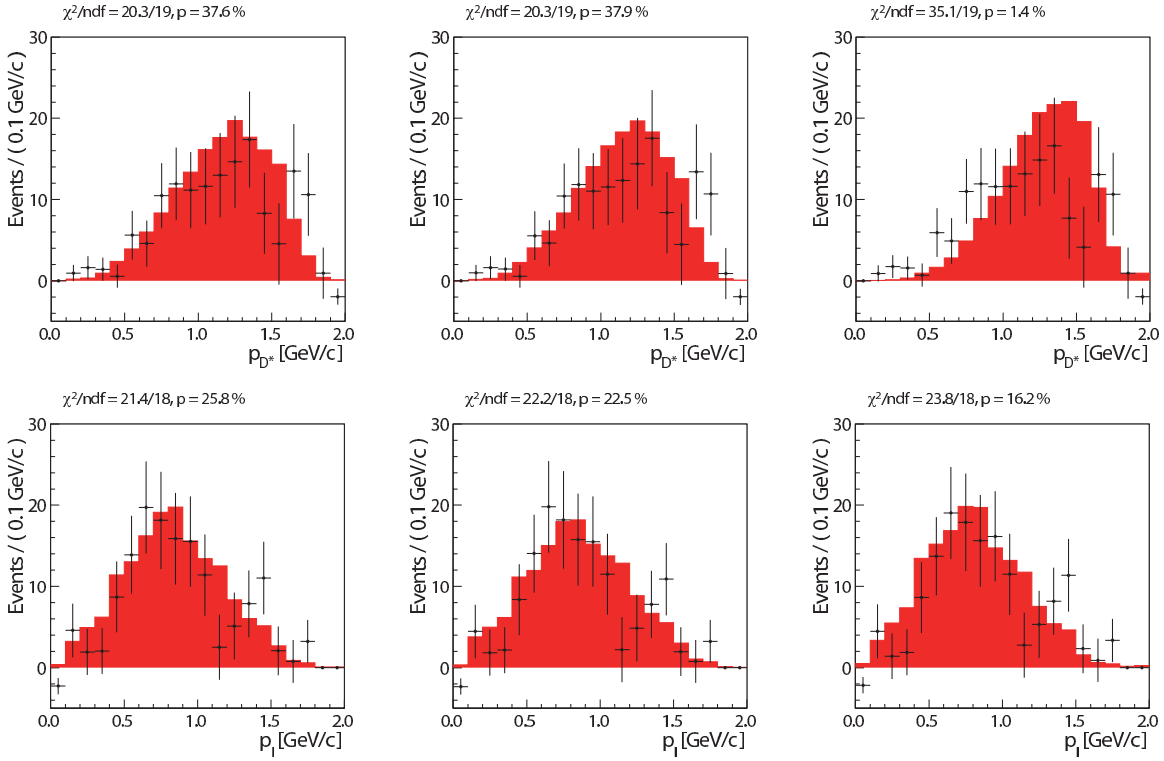


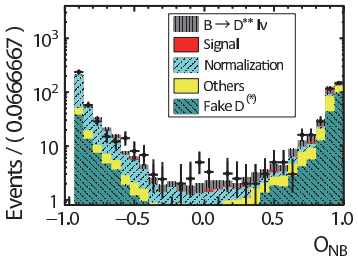
FIG. 5. Background-subtracted momentum distributions of D^* (top) and ℓ (bottom) in the region of $\mathcal{O}_{NB} > 0.8$ and $E_{\rm ECL} < 0.5$ GeV for (left) the SM, (center) the type-II 2HDM with $\tan\beta/m_{H^+} = 0.7$ GeV^{-1}, and (right) R_2-type leptoquark model with $C_T = +0.36$. The points and the shaded histograms correspond to the measured and expected distributions, respectively. The expected distributions are normalized to the number of detected events.

search 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; Ikerbasque, Basque Foundation for Science and the Euskal Herriko Unibertsitatea (UPV/EHU) under program UFI 11/55 (Spain); the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology 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

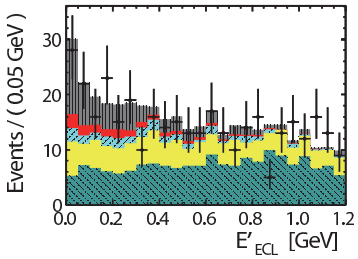
Research ("Evolution of Tau-lepton Physics").

[1] Charge-conjugate decays are implied throughout this paper, unless otherwise stated.
[2] H. Itoh, S. Komine, and Y. Okada, Prog. Theor. Phys. **114**, 179 (2005).
[3] M. Tanaka, Z. Phys. C **67**, 321 (1995).
[4] Y. Sakaki, R. Watanabe, M. Tanaka, and A. Tayduganov, Phys. Rev. D **88**, 094012 (2013).
[5] S. Fajfer, J.F. Kamenik, and I. Nisandzic, Phys. Rev. D **85**, 094025 (2012).
[6] J.F. Kamenik, and F. Mescia, Phys. Rev. D **78**, 014003 (2008).
[7] 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).
[8] A. Matyja *et al.* (Belle Collaboration), Phys. Rev. Lett. **99**, 191807 (2007).
[9] A. Bozek *et al.* (Belle Collaboration), Phys. Rev. D **82**, 072005 (2010).
[10] M. Huschle *et al.* (Belle Collaboration), Phys. Rev. D **92**, 072014 (2015).
[11] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. Lett. **115**, 111803 (2015).
[12] Y. Amhis *et al.* (Heavy Flavor Averaging Group), arXiv:1412.7515 and online update at http://www.slac.stanford.edu/xorg/hfag/.
[13] B. Kronenbitter *et al.* (Belle Collaboration), Phys. Rev. D **92**, 051102(R) (2015).
[14] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **81**, 051101(R) (2010).
[15] A. Abashian *et al.* (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **479**, 117 (2002); also see the detector section in J. Brodzicka *et al.*, Prog. Theor. Exp. Phys. (2012) 04D001.
[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 and following articles up to 03A011.
[17] D.J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[18] R. Brun *et al.*, GEANT3.21, CERN Report No. DD/EE/84-1, (1984) (unpublished).
[19] M. Tanaka and R. Watanabe, Phys. Rev. D **87**, 034028 (2013).
[20] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995).
[21] A.K. Leibovich, Z. Ligeti, I.W. Stewart, and M.B. Wise, Phys. Rev. D **57**, 308 (1998).
[22] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, Nucl. Instrum. Methods Phys. Res., Sect. A **485**, 490 (2002).
[23] A. Abashian *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **491**, 69 (2002).
[24] E. Nakano, Nucl. Instrum. Methods Phys. Res., Sect. A **494**, 402 (2002).
[25] K.A. Olive *et al* (Particle Data Group), Chin. Phys. C **38**, 090001 (2014).
[26] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res., Sect. A **559**, 190 (2006).
[27] T. Skwarnicki, Ph.D. Thesis, Institute for Nuclear Physics, Krakow 1986; DESY Internal Report, DESY F31-86-02 (1986).
[28] J.H. Friedman, Data Analysis Techniques for High Energy Particle Physics, in: Proc. 1974 CERN School of Computing, CERN 74-23 (1974).
[29] M. Feindt, F. Keller, M. Kreps, T. Kuhr, S. Neubauer, D. Zander, and A. Zunpac, Nucl. Instrum. Methods Phys. Res., Sect. A **654**, 432 (2011).
[30] See supplemental Material for fit figures of the controls samples, efficiency variation, measured value of $\mathcal{R}(D^*)$, and background-subtracted momenta distributions of D^* and lepton.
[31] Z. z. Xing, H. Zhang and S. Zhou, Phys. Rev. D **77**, 113016 (2008).

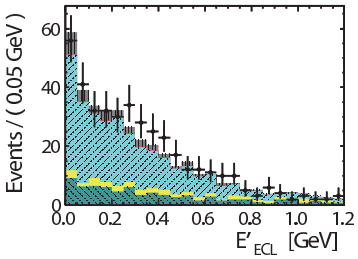
XIV. SUPPLEMENTAL MATERIAL



(a)$\mathcal{O}_{NB}$ distribution.

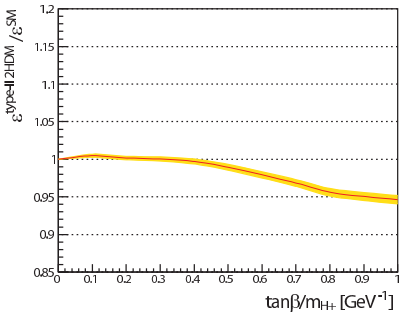


(b)$E'_{\rm ECL}$ distribution with $B \to D^{**}\ell\nu_\ell$-enhanced $\mathcal{O}_{NB}$ region ($\mathcal{O}_{NB} > 0.0$).

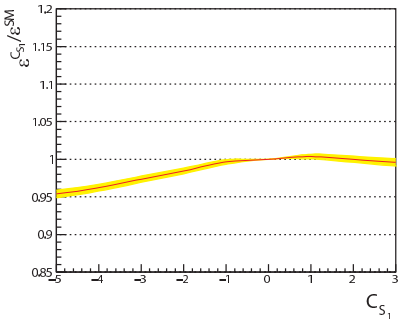


(c)$E'_{\rm ECL}$ distribution with normalization-enhanced $\mathcal{O}_{NB}$ region ($\mathcal{O}_{NB} < 0.0$).

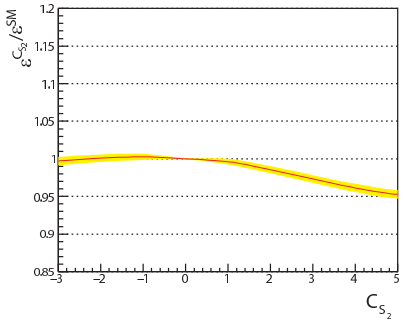
FIG. 6. Projection of the fit results of control samples with data points overlaid. In the control samples, we require π^0 in addition to the nominal event selection. $E'_{\rm ECL}$ is defined as $E_{\rm ECL}$ excluded from energy deposit from additional π^0.



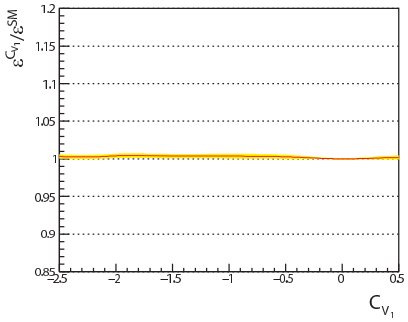
(a)Type-II 2HDM.



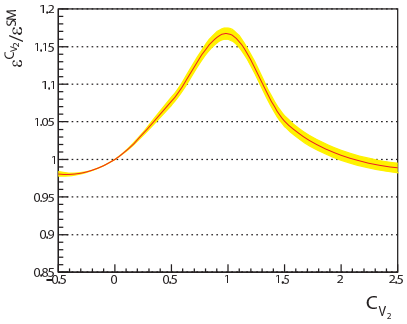
(b)SM with adding contribution from $\mathcal{O}_{S_1}$.



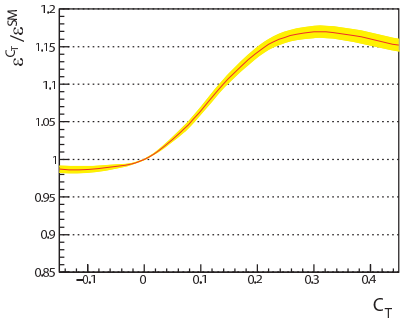
(c)SM with adding contribution from $\mathcal{O}_{S_2}$.



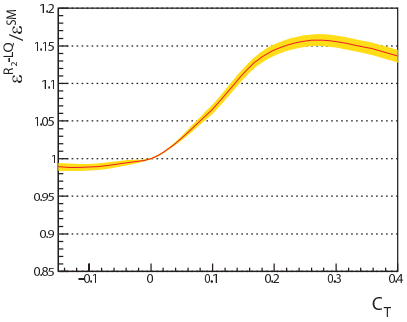
(d)SM with adding contribution from $\mathcal{O}_{V_1}$.



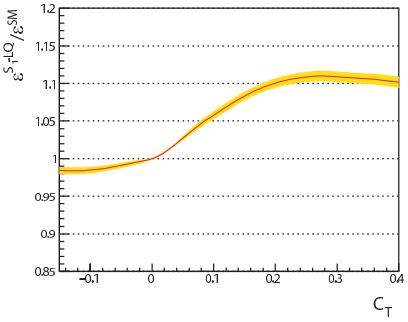
(e)SM with adding contribution from $\mathcal{O}_{V_2}$.



(f)SM with adding contribution from $\mathcal{O}_T$.

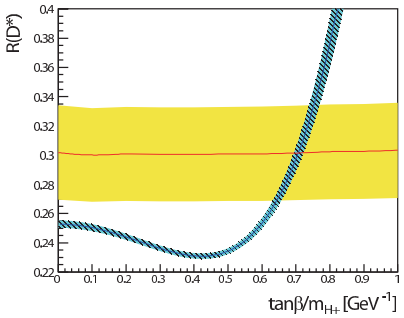


(g)R_2-type leptoquark model.

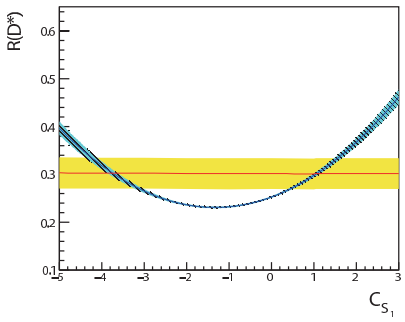


(h)S_1-type leptoquark model.

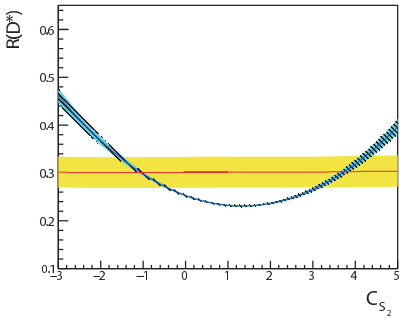
FIG. 7. Efficiency with respect to the SM value.



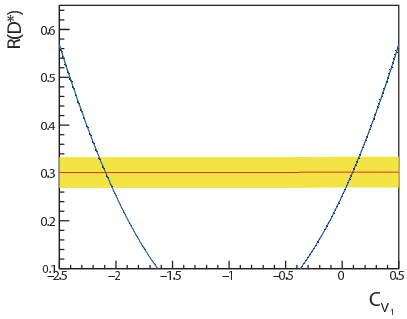
(a)Type-II 2HDM.



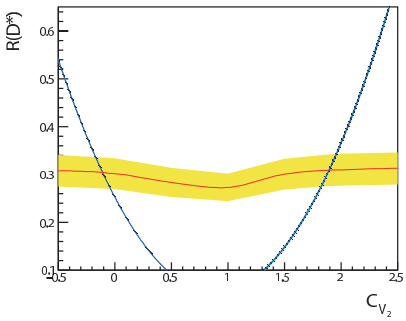
(b)SM with adding contribution from $\mathcal{O}_{S_1}$.



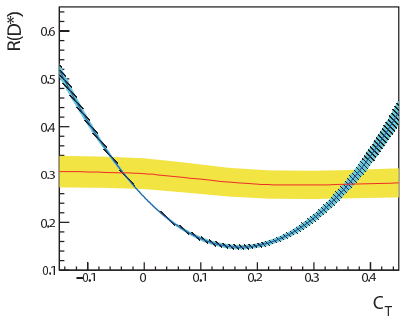
(c)SM with adding contribution from $\mathcal{O}_{S_2}$.



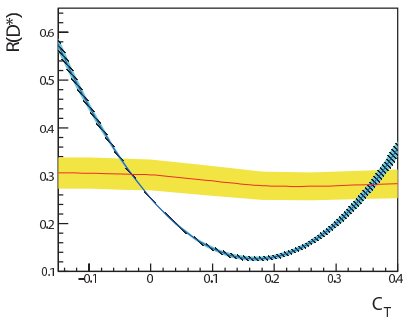
(d)SM with adding contribution from $\mathcal{O}_{V_1}$.



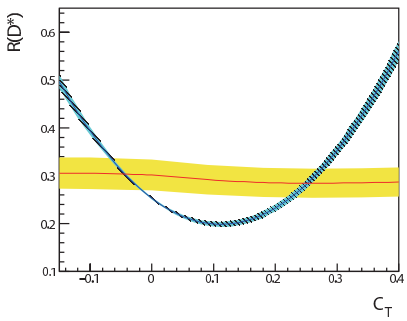
(e)SM with adding contribution from $\mathcal{O}_{V_2}$.



(f)SM with adding contribution from $\mathcal{O}_T$.

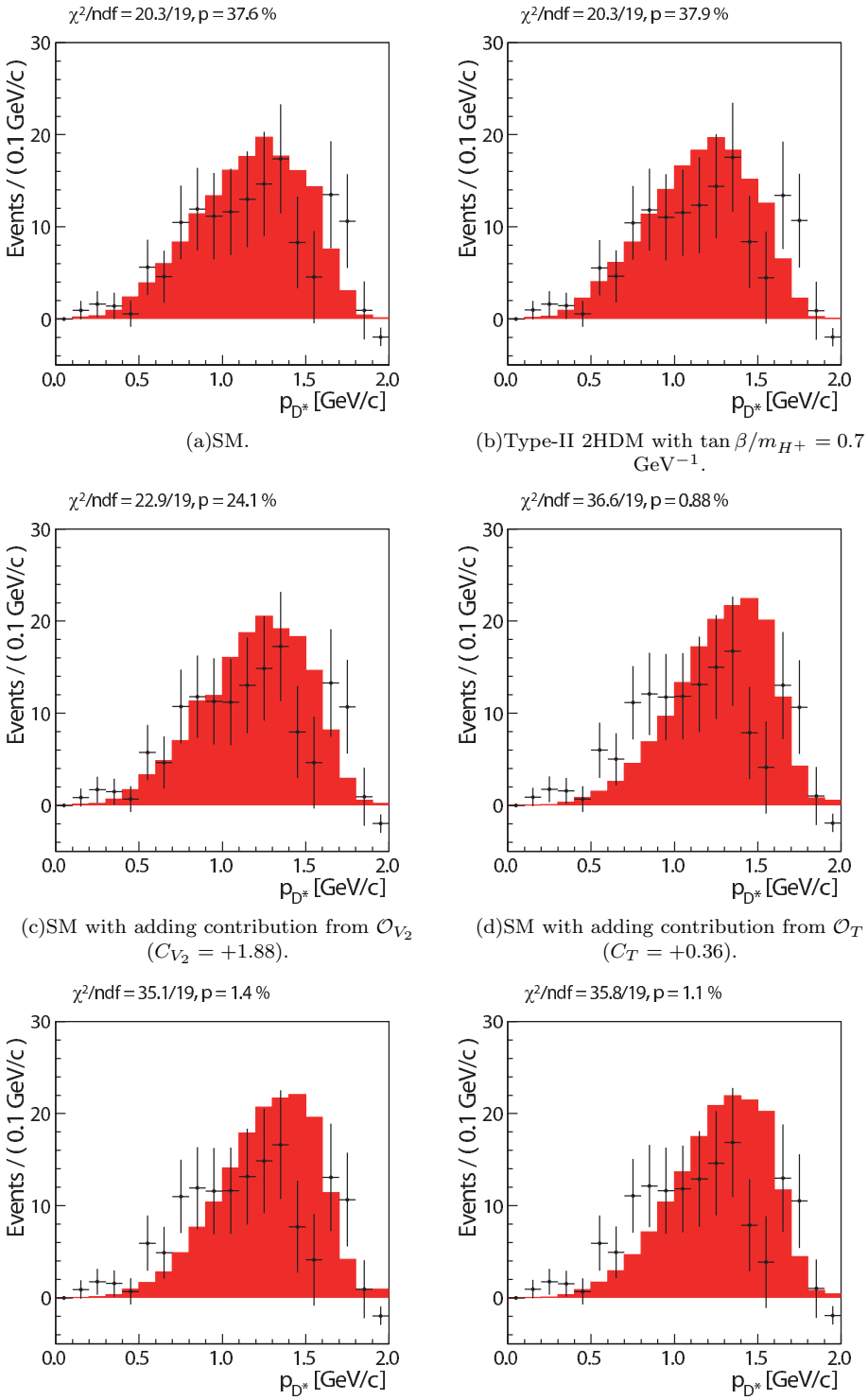


(g)R_2-type leptoquark model.



(h)S_1-type leptoquark model.

FIG. 8. Measured values of $\mathcal{R}(D^*)$ and their (1σ) uncertainties are shown by solid (red) curve and shaded (yellow) region. Theoretical predictions and their (1σ) uncertainties are shown by solid (blue) curve and hatched (light blue) region.



(a)SM.

(b)Type-II 2HDM with $\tan\beta/m_{H^+} = 0.7$ GeV^{-1}.

(c)SM with adding contribution from $\mathcal{O}_{V_2}$ ($C_{V_2} = +1.88$).

(d)SM with adding contribution from $\mathcal{O}_T$ ($C_T = +0.36$).

(e)R_2-type leptoquark model with $C_T = +0.36$.

(f)S_1-type leptoquark model with $C_T = +0.26$.

FIG. 9. Background-subtracted D^* momentum distributions in the region of $\mathcal{O}_{NB} > 0.8$ and $E_{\rm ECL} < 0.5$ GeV. The points and the shaded histograms correspond to the measured and expected distributions, respectively. The expected distributions are normalized to the number of detected events.

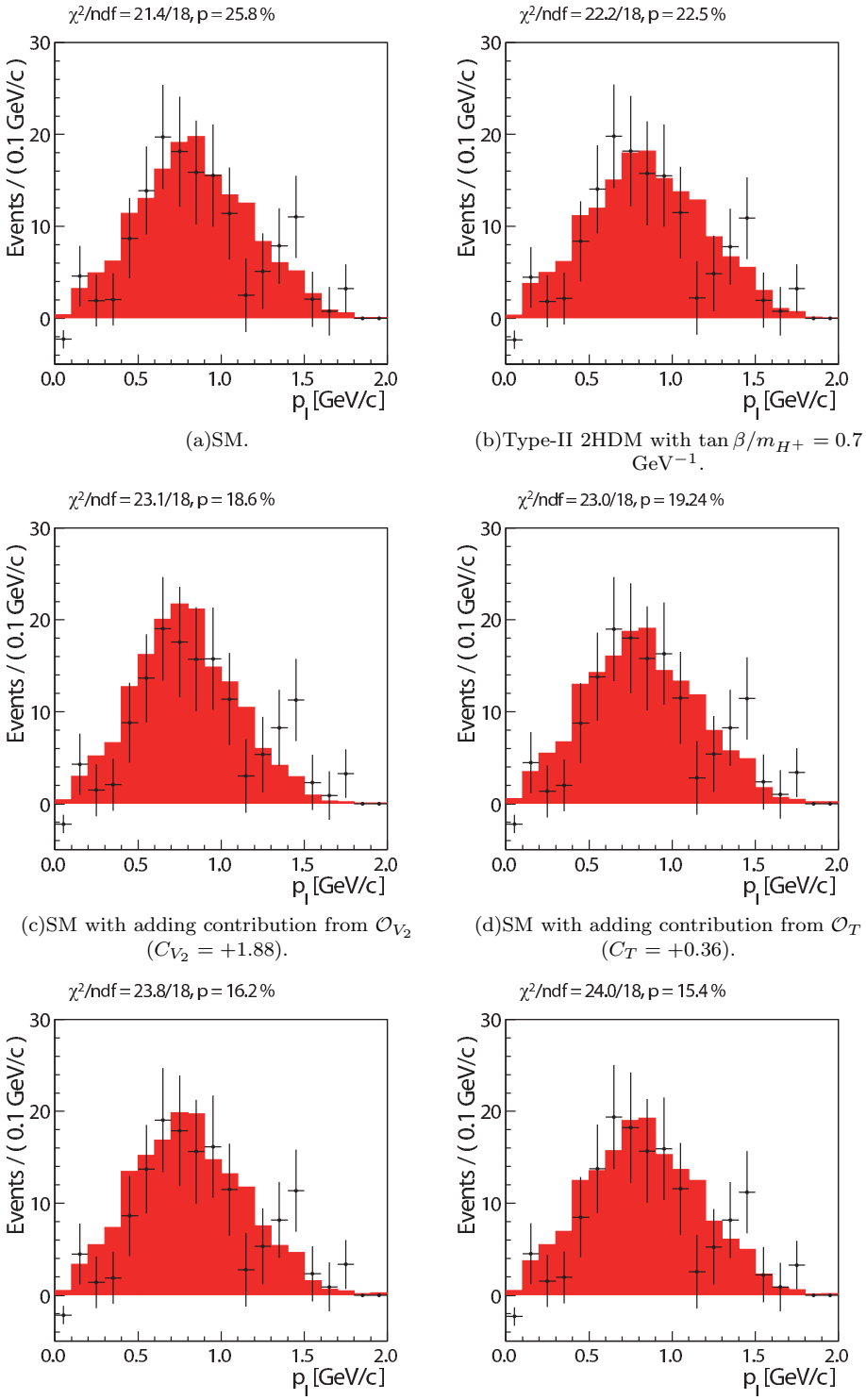


FIG. 10. Background-subtracted lepton momentum distributions in the region of $\mathcal{O}_{NB} > 0.8$ and $E_{\rm ECL} < 0.5$ GeV. The points and the shaded histograms correspond to the measured and expected distributions, respectively. The expected distributions are normalized to the number of detected events.