Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method

G. Caria,49 P. Urquijo,49 I. Adachi,16,13 H. Aihara,84 S. Al Said,79,37 D. M. Asner,3 H. Atmacan,76 T. Aushev,53 V. Babu,8 I. Badhrees,79,36 S. Bahinipati,21 A. M. Bakich,78 P. Behera,24 C. Beleño,12 J. Bennett,50 B. Bhuyan,22 T. Bilka,5 J. Biswal,32 A. Bozek,1,61 M. Bračko,47,32 T. E. Browder,15 M. Campajola,29,56 D. Červenkov,5 P. Chang,60 R. Cheaib,50 V. Chekelian,48 A. Chen,58 B. G. Cheon,14 K. Chilikin,44 H. E. Cho,14 K. Cho,39 Y. Choi,17 S. Choudhury,23 D. Cinabro,88 S. Cunliffe,8 N. Dash,21 G. De Nardo,29,56 F. Di Capua,29,56 S. Di Carlo,42 Z. Doležal,5 T. V. Dong,10 S. Eidelman,6,4,4,44 D. Epifanov,44,64 J. E. Fast,66 T. Ferber,8 D. Ferleicw,49 B. G. Fulsom,66 R. Garg,67 V. Gaur,87 N. Gabyshev,4,64 A. Garnash,4,64 A. Giri,23 P. Goldenzweig,33 D. Greenwald,31 O. Grzymkowska,61 Y. Guan,7 O. Hartbrich,6 K. Hayasaka,63 H. Hayashii,57 T. Higuchi,34 W.-S. Hou,60 C.-L. Hsu,78 T. Iijima,55,54 K. Inami,54 G. Inguglia,77 A. Ishikawa,16,13 R. Itoh,16,13 M. Iwasaki,65 Y. Iwasaki,16 W. W. Jacobs,25 H. B. Jeon,41 S. Jia,2 Y. Jin,84 D. Joffe,35 K. J. Joo,6 A. B. Kaliyar,24 K. H. Kang,41 G. Karyan,8 T. Kawasaki,38 H. Kichimi,16 C. H. Kim,14 D. Y. Kim,75 H. J. Kim,41 K. T. Kim,40 S. H. Kim,14 K. Kinoshita,7 P. Kodyš,5 S. Korpar,47,32 D. Kotchetkov,15 P. Križan,91,32 R. Kroeger,50 J.-F. Krohn,49 P. Krokovny,4,64 T. Kuhr,46 R. Kumar,70 Y.-J. Kwon,90 J. S. Lange,14 I. S. Lee,74 J. K. Lee,73 S. C. Lee,41 L. K. Li,26 Y. B. Li,68 L. Li Giot,48 J. Libby,24 K. Lieret,45 D. Liventsev,87,16 T. Luo,10 C. MacQueen,49 M. Masuda,83 T. Matsuda,51 D. Matvienko,6,4,4,44 M. Merola,29,56 F. Metzner,33 K. Miyabasahi,57 G. B. Mohanty,80 T. J. Moon,73 T. Mori,54 R. Mussa,30 K. R. Nakamura,16 M. Nakao,16,13 K. J. Nath,22 M. Nayak,88,16 N. K. Nisar,69 S. Nishida,16,13 K. Nishimura,15 K. Ogawa,63 H. Ono,62,63 Y. Onuki,84 P. Osinski,44 P. Pakhlova,44,52 G. Pakhlova,44,53 B. Pal,9 T. Pang,9 H. Park,41 S.-H. Park,90 S. Patra,72 S. Paul,91 T. K. Pedlar,46 R. Pestotnik,32 L. E. Piilonen,87 V. Popov,44,53 E. Prêncipe,18 M. T. Prim,33 A. Rabusov,81 P. K. Resmi,24 M. Ritter,45 M. Rozanska,61 G. Russo,56 D. Sahoo,80 Y. Sakai,16,13 S. Sandilya,7 L. Santelj,16 T. Sanuki,82 V. Savinov,69 O. Schneider,43 G. Schnell,1,19 J. Schueler,15 C. Schwanda,27 A. J. Schwartz,7 Y. Seino,63 K. Senyo,89 M. E. Sevior,49 V. Shebalin,15 J.-G. Shiu,66 B. Shwartz,4,64 F. Simon,48 A. Sokolov,28 E. Solovieva,44 M. Starič,92 Z. S. Strottler,87 T. Sumiyoshi,86 W. Sutcliffe,33 M. Takizawa,74,17,71 U. Tamponi,30 K. Tanida,31 F. Tencchini,8 K. Trabelsi,42 M. Uchida,85 T. Uglow,44,53 S. Uno,16,13 Y. Usov,4,64 S. E. Valsen,15 R. Van Tonder,33 G. Varner,15 K. E. Varvell,78 A. Vossen,9 E. Waheed,49 B. Wang,48 C. H. Wang,59 M.-Z. Wang,60 P. Wang,26 X. L. Wang,10 S. Watanuki,82 J. Wiechczynski,61 E. Won,40 H. Yamamoto,82 S. B. Yang,40 H. Ye,8 J. H. Yin,26 C. Z. Yuan,26 Z. P. Zhang,72 V. Zhilich,4,64 V. Zhukova,15 and V. Zhulanov,4,64

(The Belle Collaboration)

1 University of the Basque Country UPV/EHU, 48080 Bilbao
2 Beihang University, Beijing 100191
3 Brookhaven National Laboratory, Upton, New York 11973
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, Gwangju 61186
7 University of Cincinnati, Cincinnati, Ohio 45221
8 Deutsches Elektronen-Synchrotron, 22607 Hamburg
9 Duke University, Durham, North Carolina 27708
10 Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) and Institute of Modern Physics, Fudan University, Shanghai 200443
11 Justus-Liebig-Universität Gießen, 35392 Gießen
12 II. Physikalisches Institut, Georg-August-Universität Göttingen, 37073 Göttingen
13 SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193
14 Department of Physics and Institute of Natural Sciences, Hangang University, Seoul 04763
15 University of Hawaii, Honolulu, Hawaii 96822
16 High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801
17 J-PARC Branch, KEK Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801
18 Forschungszentrum Jülich, 52425 Jülich
19 Ikerbasque, Basque Foundation for Science, 48013 Bilbao
20 Indian Institute of Science Education and Research Mohali, SAS Nagar, 140306
21 Indian Institute of Technology Bhubaneswar, Satya Nagar 751007
22 Indian Institute of Technology Guwahati, Assam 781039
23 Indian Institute of Technology Hyderabad, Telangana 502285
24 Indian Institute of Technology Madras, Chennai 600036
25 Indiana University, Bloomington, Indiana 47408
The experimental results on the ratios of branching fractions \( \mathcal{R}(D) = \mathcal{B}(\bar{B} \to D^{(*)} \tau^- \bar{\nu}_\tau)/\mathcal{B}(\bar{B} \to D^{(*)} \ell^- \bar{\nu}_\ell) \) and \( \mathcal{R}(D^*) = \mathcal{B}(\bar{B} \to D^{(*)} \tau^- \bar{\nu}_\tau)/\mathcal{B}(\bar{B} \to D^{(*)} \ell^- \bar{\nu}_\ell) \), where \( \ell \) denotes an electron or a muon, show a long-standing discrepancy with the Standard Model predictions, and might hint to a violation of lepton flavor universality. We report a new simultaneous measurement of \( \mathcal{R}(D) \) and \( \mathcal{R}(D^*) \), based on a data sample containing \( 772 \times 10^6 B \bar{B} \) events recorded at the \( T(4S) \) resonance with the Belle detector at the KEKB \( e^+e^- \) collider. In this analysis the tag-side \( B \) meson is reconstructed in a semileptonic decay mode and the signal-side \( \tau \) is reconstructed in a purely leptonic decay. The measured values are \( \mathcal{R}(D) = 0.307 \pm 0.037 \pm 0.016 \) and \( \mathcal{R}(D^*) = 0.283 \pm 0.018 \pm 0.014 \), where the first uncertainties are statistical and the second are systematic. These results are in agreement with the Standard Model predictions within 0.2, 1.1 and 0.8 standard deviations for \( \mathcal{R}(D) \), \( \mathcal{R}(D^*) \) and their combination, respectively. This work constitutes the most precise measurements of \( \mathcal{R}(D) \) and \( \mathcal{R}(D^*) \) performed to date as well as the first result for \( \mathcal{R}(D) \) based on a semileptonic tagging method.

PACS numbers: 13.20.He, 14.40.Nd

Semileptonic \( B \) meson decays, involving the transition \( b \to c \tau \bar{\nu}_\tau \), are sensitive probes for physics beyond the Standard Model (SM). Any difference in the branching fraction of these processes with respect to the SM prediction would violate lepton flavor universality, which enforces equal coupling of the gauge bosons to the three lepton generations. Indeed, in many models beyond the SM, new interactions with enhanced coupling to the third family are postulated. Among such new mediators, charged Higgs bosons, which appear in supersymmetry [1] and other models with two Higgs doublets [2], may contribute measurably to the \( b \to c \tau \bar{\nu}_\tau \) decay rate due to the large masses of the \( \tau \) and the \( b \) quark. Similarly, leptoquarks [3], which carry both lepton and baryon numbers, may also contribute to this process.

The ratios of branching fractions,

\[
\mathcal{R}(D^{(*)}) = \frac{\mathcal{B}(\bar{B} \to D^{(*)} \tau^- \bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \to D^{(*)} \ell^- \bar{\nu}_\ell)}
\]

where the denominator represents the average of electron and muon modes, are typically measured instead of the absolute branching fractions of \( \bar{B} \to D^{(*)} \tau^- \bar{\nu}_\tau \) to reduce common systematic uncertainties, such as those due to the detection efficiency, the magnitude of the quark-mixing matrix element \( |V_{cb}| \), and the semileptonic decay form factors. Hereafter, \( \bar{B} \to D^{(*)} \tau^- \bar{\nu}_\tau \) [4] and \( \bar{B} \to D^{(*)} \ell^- \bar{\nu}_\ell \) will be referred to as the signal and normalization modes, respectively. The SM calculations for these ratios, performed by several groups [5][6], are averaged by HFLAV [7] to obtain \( \mathcal{R}(D) = 0.299 \pm 0.003 \) and \( \mathcal{R}(D^*) = 0.258 \pm 0.005 \).

Semileptonic \( B \) decays were first observed by Belle in 2007 [10], with subsequent studies reported by Belle [11]-[14], BaBar [15], and LHCb [16][17]. The average values of the experimental results, excluding the result presented in this Letter, are \( \mathcal{R}(D) = 0.407 \pm 0.039 \pm 0.024 \) and \( \mathcal{R}(D^*) = 0.306 \pm 0.013 \pm 0.007 \) [9], where the first uncertainty is statistical and the second is systematic. These values exceed SM predictions by 2.1\( \sigma \) and 3.0\( \sigma \), respectively, where \( \sigma \) denotes the standard deviation. A combined analysis of \( \mathcal{R}(D) \) and \( \mathcal{R}(D^*) \) taking correlations into account finds that the deviation from the SM prediction is approximately 3.8\( \sigma \) [9]. This large discrepancy must be investigated with complementary and more precise measurements.

Measurements at the \( e^+e^- \) “B-factory” experiments Belle and BaBar, are commonly performed by first reconstructing one of the \( B \) mesons in the \( T(4S) \to B \bar{B} \) decay, denoted as \( B_{\text{tag}} \), using a dedicated tagging algorithm. So far, simultaneous measurements of \( \mathcal{R}(D) \) and \( \mathcal{R}(D^*) \) at Belle and BaBar have been performed using hadronic tagging methods on both \( B^0 \) and \( B^+ \) decays [12][15], while only \( \mathcal{R}(D^{*-}) \) was measured with a semileptonic tagging method [13]. In this Letter, we report the first measurement of \( \mathcal{R}(D) \) using the semileptonic tagging method, and we update or measurement of \( \mathcal{R}(D^*) \) by combining results of \( B^0 \) and \( B^+ \) decays with a more efficient tagging algorithm. Our previous measurement of \( \mathcal{R}(D^{*-}) \) with a semileptonic tagging method is therefore superseded by this work.

We use the full \( T(4S) \) data sample containing \( 772 \times 10^6 B \bar{B} \) events recorded with the Belle detector [18] at the KEKB \( e^+e^- \) collider [19]. Belle was a general-purpose magnetic spectrometer, which consisted 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. These components were located inside a superconducting solenoid coil that provided a 1.5 T magnetic field. An iron flux-return yoke located outside the coil was instrumented to detect \( K_L^0 \) mesons and muons (KLM). The detector is described in detail elsewhere [13].
To determine the reconstruction efficiency and probability density functions (PDFs) for signal, normalization, and background modes, we use Monte Carlo (MC) simulated events generated with the EvtGen event generator [20]. The detector response is simulated with the GEANT3 package [21].

Semileptonic $B \rightarrow D^{(*)}\ell\nu$ decays are generated with the HQET2 EvtGen package, based on the CLN parametrization [22]. As the measured parameters of the model have been updated since our MC sample was generated, we apply an event-by-event correction factor obtained by taking the ratio of differential decay rates in the updated CLN parameters compared to those used in the MC. For the MC samples of $B \rightarrow D^{(*)}\ell\nu$ decays, we used the ISGW2 EvtGen package, based on the quark model described in Ref. [23]. This model has been superseded by the LLSW model [24], thus we weight events with a correction factor based on the ratio of the analytic predictions of LLSW and MC distributions generated with ISGW2. Here, $D^{**}$ denotes the orbitally excited states $D_1$, $D_2^*$, $D_1'$, and $D_0^*$. We consider $D^{**}$ decays to a $D^{(*)}$ and a pion, a $\rho$ or an $\eta$ meson, or a pair of pions, where branching fractions are based on quantum number, phase-space, and isospin arguments. The sizes of the inclusive $\Upsilon(4S) \rightarrow BB$ MC sample and the dedicated $B \rightarrow D^{(*)}\ell\nu$ MC sample correspond to about 10 times and 5 times the integrated luminosity of the $\Upsilon(4S)$ data sample, respectively.

The $B_{\text{tag}}$ is reconstructed using a hierarchical algorithm based on boosted decision trees (BDT) [25] in $D\ell\nu_\ell$ and $D^*\ell\bar{\nu}_\ell$ channels, where $\ell = e, \mu$. The BDT classifier assigns to each $B_{\text{tag}}$ candidate a probability of representing a well-reconstructed $B$ meson. The range of the BDT classifier extends from 0 to 1, with well-reconstructed candidates having the highest values. We select $B_{\text{tag}}$ candidates with a BDT classifier output greater than $10^{-1.5}$. We suppress $B \rightarrow D^*\tau^-(\rightarrow b\nu\nu)\nu$ events on the $B_{\text{tag}}$ side by applying a selection on $\cos\theta_{B,D^{(*)}\ell}$. This variable corresponds to the cosine of the angle between the momenta 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:

$$\cos\theta_{B,D^{(*)}\ell} = \frac{2E_{\text{beam}}E_{D^{(*)}\ell} - m_B^2 - m_{D^{(*)}\ell}^2}{2|p_B||p_{D^{(*)}\ell}|},$$  \hspace{1cm} (2)

Here $E_{\text{beam}}$ is the beam energy, and $E_{D^{(*)}\ell}$, $p_{D^{(*)}\ell}$, and $m_{D^{(*)}\ell}$ are the energy, momentum, and mass, of the $D^{(*)}\ell$ system, respectively. The quantities $m_B$ and $|p_B|$ are the nominal $B$ meson mass [26] and momentum, respectively. All quantities are evaluated in the $\Upsilon(4S)$ rest frame.

Correctly reconstructed $B \rightarrow D^{(*)}\ell\nu$ decays are expected to have a value of $\cos\theta_{B,D^{(*)}\ell}$ between $-1$ and $+1$. Correctly reconstructed as well as misreconstructed $B \rightarrow D^{(*)}\tau\nu$ decays generally have $\cos\theta_{B,D^{(*)}\ell}$ values below $-1$ due to the presence of additional missing particles. To account for detector resolution effects we apply the requirement $-2.0 < \cos\theta_{B,D^{(*)}\ell} < 1.0$ for the $B_{\text{tag}}$.

In each event with a selected $B_{\text{tag}}$ candidate, we search for the opposite-flavor signature $D^{(*)}\ell$ among the remaining tracks and calorimeter clusters, since we only reconstruct pure leptonic tau decays $\tau \rightarrow \ell\nu\nu$. We define four disjoint data samples, denoted $D^+\ell^-$, $D^0\ell^-$, $D^{*+}\ell^-$, and $D^{*0}\ell^-$. Charged particle tracks are reconstructed with the SVD and CDC by requiring a point of closest approach to the interaction point smaller than 5.0 cm along the direction of the $e^+$ beam and 2.0 cm in the direction perpendicular to it. These requirements do not apply to the pions from $K_S^0$ decays. 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 CDC, the response of the ACC, the cluster shape in the ECL, and the match between positions of the cluster and the track at the ECL. To recover bremsstrahlung photons from electrons, we add the four-momentum of each photon detected within a cone of 0.05 rad of the original track direction to the electron momentum. Muons are identified by the track penetration depth and hit distribution in the KLM. Charged kaons are identified by combining information from the $dE/dx$ measured in the CDC, the flight time measured with the TOF, and the response of the ACC. We do not apply any particle identification criteria for charged pion candidates.

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 $\pm15$ MeV/$c^2$ of the nominal $K_S^0$ mass [26], which corresponds to approximately seven times the reconstructed mass resolution. Further selection is performed with an algorithm based on a neural network [27].

Photons are measured as an electromagnetic cluster in the ECL with no associated charged track. Neutral pions are reconstructed in the $\pi^0 \rightarrow \gamma\gamma$ channel, and their energy resolution is improved by performing a mass-constrained fit of the two photon candidates to the nominal $\pi^0$ mass [26]. For neutral pions from $D$ decays, we require the daughter photon energies to be greater than 50 MeV and their asymmetry to be less than 0.6 in the laboratory frame, the cosine of the angle between two photons to be greater than zero, and the $\gamma\gamma$ invariant mass to be within $[-15, +15]$ MeV/$c^2$ of the nominal $\pi^0$ mass, which corresponds to approximately $\pm1.8$ times the resolution. Low-energy $\pi^0$ candidates from $D^{*}$ are reconstructed using less restrictive energy requirements: one photon must have an energy of at least 50 MeV, while the other must have a minimum energy of 20 MeV. We also require a narrower window around the diphoton invariant mass to compensate for the lower photon-energy require-
ment: within 10 MeV/$c^2$ of the nominal $\pi^0$ mass, which corresponds to approximately $\pm 1.6$ times the resolution.

Neutral $D$ mesons are reconstructed in the following decay modes: $D^{0} \rightarrow K^−\pi^+\pi^0$, $K^−\pi^+\pi^−\pi^0$, $K^+\pi^−\pi^−$, $K^0_S\pi^0$, $K^0_L\pi^0$, $K^0_LK^+$, $K^−K^+$, and $\pi^+\pi^−$. Similarly, charged $D$ mesons are reconstructed in the following modes: $D^+ \rightarrow K^−\pi^+\pi^+, K^0_S\pi^+\pi^−\pi^0$, $K^0_L\pi^0\pi^+\pi^−$, $K^0_L\pi^+\pi^+, K^+K^+\pi^0$, and $K^0_LK^+$. The combined branching fractions for reconstructed channels are 30% and 22% for $D^0$ and $D^+$, respectively. For $D$ decays without a $\pi^0$ in the final state, we require the invariant mass of the reconstructed candidates to be within 15 MeV/$c^2$ of the nominal $D^0$ or $D^+$ mass, which corresponds to a window of approximately $\pm 2.8$ times the resolution. In the case of channels with a $\pi^0$ in the final state, which have worse mass resolution, we require a wider window: from $-45$ to $+30$ MeV/$c^2$ around the nominal $D^0$ mass, and from $-36$ to $+24$ MeV/$c^2$ around the nominal $D^+$ mass. These windows correspond to approximately $[−1.1, +1.6]$ and $[−1.0, +1.4]$ times the resolution, respectively. Candidate $D^*$ mesons are reconstructed in the channels $D^0\pi^+$ and $D^+\pi^0$, and $D^{0*}$ in the channel $D^0\pi^0$. We do not consider the $D^{0*} \rightarrow D^0\gamma$ decay channel due to its higher background level.

We require the mass difference $D^*-D$ to be within 2.5 MeV/$c^2$ for the $D^{*+} \rightarrow D^{0}\pi^+$ decay mode, and within 2.0 MeV/$c^2$ for the $D^{*+} \rightarrow D^{+}\pi^0$ and $D^{*0} \rightarrow D^0\pi^0$ decay modes. These windows correspond to $\pm 3.0$ and $\pm 1.9$ times the resolution, respectively. We require a tighter mass window in the $D^*$ modes that contain a low-momentum ("slow") $\pi^0$ to suppress the large background arising from misreconstructed neutral pions.

On the signal side, we require $\cos \theta_{B,D^{(*)}\ell}$ to be less than 1.0 and the $D^{(*)}$ momentum in the $T(J/\psi)$ rest frame to be less than 2.0 GeV/$c$. Finally, we require that events contain no extra prompt charged tracks, $K^0_S$ candidates, or $\pi^0$ candidates, which are reconstructed with the same criteria as those used for the $D$ candidates. All selection criteria used for event reconstruction have been the subject of optimization studies. When multiple $B_{\text{tag}}$ or $B_{\text{sig}}$ candidates are found in an event, we first select the $B_{\text{tag}}$ candidate with the highest tagging classifier output, and then the $B_{\text{sig}}$ candidate with the highest p-value from the vertex fit of the $B$ candidate’s charm daughter.

To distinguish signal and normalization events from background processes, we use the sum of the energies of neutral clusters detected in the ECL that are not associated with any reconstructed particles, denoted as $E_{\text{ECL}}$. To mitigate the varying effects of photons related to beam background in the calculation of $E_{\text{ECL}}$, we only include clusters with energies greater than 50, 100, and 150 MeV, respectively, from the barrel, forward, and backward ECL regions [18]. Signal and normalization events peak near zero in $E_{\text{ECL}}$, while background events populate a wider range. We require that $E_{\text{ECL}}$ be less than 1.2 GeV.

To separate reconstructed signal and normalization events, we employ a BDT based on the XGBoost package [28]. The input variables to the BDT are $\cos \theta_{B,D^{(*)}\ell}$; the approximate missing mass squared $m_{\text{miss}}^2 = (E_{\text{beam}} - E_{D^{(*)}\ell} - E_i)^2 - (p_{D^{(*)}\ell} + p_i)^2$; the visible energy $E_{\text{vis}} = \sum E_i$, where $(E_i,p_i)$ is the four-momentum of particle $i$. We do not apply any selection on the BDT classifier output, denoted as $O_{\text{cls}}$; instead we use it as one of the fitting variables for the extraction of $R(D^{(*)})$. Signal events have $O_{\text{cls}}$ values near 1, while normalization events have values near 0.

We extract the yields of signal and normalization modes from a two-dimensional (2D) extended maximum-likelihood fit to the variables $O_{\text{cls}}$ and $E_{\text{ECL}}$. The fit is performed simultaneously to the four $D^{(*)}\ell$ samples, and exploits the isospin constraint $R(D^{(*)}_0) = R(D^{(*)}_+)$.

The distribution of each sample is described as the sum of several components: $D^{(*)}\tau\nu$, $D^{(*)}\ell\nu$, feed-down from $D^*(\ell)\nu$ to $D\ell(\tau)\nu$, $D^{**}(\ell)\nu$, and other backgrounds. The PDFs of these components are determined from MC simulations as 2D histogram templates. A large fraction of $B \rightarrow D^*\ell\nu$ decays from both $B^0$ and $B^+$ are reconstructed in the $D\ell$ samples (denoted feed-down). We leave these two contributions free in the fit and use their fitted yields to correct the MC estimated feed-down rate of $B \rightarrow D^*\tau\nu$ decays. The events of the $D^*\ell$ samples that appear as feed-down are treated as a component of the signal or normalization yields. As the probability of $B \rightarrow D\ell(\tau)\nu$ decays contributing to the $D^*\ell$ samples is very small, the relative rates of these contributions are fixed to the MC expected values.

The free parameters in the final fit are the yields of signal, normalization, $B \rightarrow D^{**}\ell\nu$, and feed-down from $D^*\ell$ to $D\ell$ components. The yields of other backgrounds are fixed to their MC expected values. The ratios $R(D^{(*)})$ are given by the formula:

$$R(D^{(*)}) = \frac{1}{2B(\tau^- \rightarrow \ell^- \bar{\nu}_\ell \nu_\ell)} \frac{\epsilon_{\text{norm}}}{\epsilon_{\text{sig}}} \frac{N_{\text{sig}}}{N_{\text{norm}}}, \quad (3)$$

where $\epsilon_{\text{sig(norm)}}$ and $N_{\text{sig(norm)}}$ are the detection efficiency including tagging efficiency and yields of signal (normalization) modes and $B(\tau^- \rightarrow \ell^- \bar{\nu}_\ell \nu_\ell)$ is the average of the world-average branching fractions for $\ell = e$ and $\ell = \mu$. To improve the accuracy of the MC simulation, we apply a series of correction factors determined from control sample measurements, such as those associated to lepton and hadron identification efficiencies as well as slow pion tracking efficiencies. Correction factors for the lepton efficiencies are evaluated as a function of the lepton momentum and direction using $e^+e^- \rightarrow e^+e^-\ell^-\ell^-$ and $J/\psi \rightarrow \ell^+\ell^-\gamma$ decays. Furthermore, to determine the expected yield of fake and misreconstructed $D^{(*)}$ mesons, treated as background, we use data sidebands of difference between their nominal and reconstructed mass, and we correct for differences in the reconstruction efficiency.
of the tagging algorithm between data and MC simulation.

The $E_{ECL}$ projections of the fit are shown in Fig. 1. The fit finds $R(D) = 0.307 \pm 0.037$ and $R(D^{*}) = 0.283 \pm 0.018$, where the error is statistical.

To estimate various systematic uncertainties contributing to $R(D^{(*)})$, we vary each fixed parameter 500 times, sampling from a Gaussian distribution built using the value and uncertainty of the parameter. For each variation, we repeat the fit. The associated systematic uncertainty is taken as the standard deviation of the resulting distribution of fitted results. The systematic uncertainties are listed in Table I.

In Table I the label “$D^{**}$ composition” refers to the uncertainty introduced by the branching fractions of the $B \rightarrow D^{**}\ell\nu_{\ell}$ channels and the decays of the $D^{**}$ mesons, which are not well known and hence contribute significantly to the total PDF uncertainty. The uncertainties on the branching fraction of $B \rightarrow D^{**}\ell\nu_{\ell}$ are assumed to be $\pm 6\%$ for $D_{1}$, $\pm 10\%$ for $D_{2}^{*}$, $\pm 83\%$ for $D_{1}^{*}$, and $\pm 100\%$ for $D_{0}^{*}$, while the uncertainties on each of the $D^{**}$ decay branching fractions are conservatively assumed to be $\pm 100\%$.

A large systematic uncertainty arises from the limited size of the MC samples. Firstly, this is reflected in the uncertainty of the PDF shapes. To estimate this contribution, we recalculate PDFs for signal, normalization, fake $D^{(*)}$ events, $B \rightarrow D^{**}\ell\nu_{\ell}$, feed-down, and other backgrounds by generating toy MC samples from the nominal PDFs according to Poisson statistics, and then repeating the fit with the new PDFs. Secondly, the reconstruction efficiency of feed-down events, together with the efficiency ratio of signal to normalization events, are varied within their uncertainties, which are limited by the size of the MC samples as well.

The efficiency factors for the fake $D^{(*)}$ and $B_{\text{tag}}$ reconstruction are calibrated using collision data. The uncertainties on these factors are affected by the size of the samples used in the calibration. We vary the factors within their errors and extract associated systematic uncertainties.

The effect of the lepton efficiency and fake rate, as well as that due to the slow pion efficiency, do not cancel out in the $R(D^{(*)})$ ratios. This is due to the dif-
TABLE I. Systematic uncertainties contributing to the $\mathcal{R}(D^{(*)})$ results, together with their correlation.

| Source                        | $\Delta \mathcal{R}(D)$ (%) | $\Delta \mathcal{R}(D^*)$ (%) | Correlation |
|-------------------------------|------------------------------|-------------------------------|-------------|
| $D^{(*)}$ composition          | 0.76                         | 1.41                          | −0.41       |
| PDF shapes                    | 4.39                         | 2.25                          | −0.55       |
| Feed-down factors             | 1.69                         | 0.44                          | 0.53        |
| Efficiency factors            | 1.93                         | 4.12                          | −0.57       |
| Fake $D^{(*)}$ calibration    | 0.19                         | 0.11                          | −0.76       |
| $B_{sg}$ calibration          | 0.07                         | 0.05                          | −0.76       |
| Lepton efficiency and fake rate| 0.36                         | 0.33                          | −0.83       |
| Slow pion efficiency          | 0.08                         | 0.08                          | −0.98       |
| $B$ decay form factors        | 0.55                         | 0.28                          | −0.60       |
| Luminosity, $f^{+−}, f^{00}$  | 0.10                         | 0.04                          | −0.58       |
| and $B(\Upsilon(4S))$         |                              |                               |             |
| $B(B \to D^{(*)}\ell\nu)$     | 0.05                         | 0.02                          | −0.69       |
| $B(D)$                        | 0.35                         | 0.13                          | −0.65       |
| $B(D^*)$                      | 0.04                         | 0.02                          | −0.51       |
| $B(\tau \to \ell^−\bar{\nu}_\ell\nu_\sigma)$ | 0.15                        | 0.14                          | −0.11       |
| Total                         | 5.21                         | 4.94                          | −0.52       |

Different momentum spectra of leptons and charm mesons in the normalization and signal modes. The uncertainties introduced by these factors are included in the total systematic uncertainty.

We include minor systematic contributions from other sources: one related to the parameters that are used for re-weighting the semileptonic $B \to D^{(*)}\ell\nu$ and $B \to D^{(*)}\ell\nu$ decays; and others from the integrated luminosity, the $B$ production fractions at the $\Upsilon(4S)$, $f^{+−}$ and $f^{00}$, and the branching fractions of $B \to D^{(*)}\ell\nu$, $D$, $D^*$, and $\tau^− \to \ell^−\bar{\nu}_\ell\nu_\sigma$ decays [26]. The total systematic uncertainty is estimated by summing the aforementioned contributions in quadrature.

In conclusion, we have measured the ratios $\mathcal{R}(D^{(*)}) = B(B \to D^{(*)}\tau^−\bar{\nu}_\tau)/B(B \to D^{(*)}\ell^−\bar{\nu}_\ell)$, where $\ell$ denotes an electron or a muon, using a semileptonic tagging method and a data sample containing $772 \times 10^6 BB$ events collected with the Belle detector. The results are

$$\mathcal{R}(D) = 0.307 \pm 0.037 \pm 0.016 \quad (4)$$

$$\mathcal{R}(D^*) = 0.283 \pm 0.018 \pm 0.014 \quad (5)$$

where the first uncertainties are statistical and the second are systematic. These results are in agreement with the SM predictions within $0.2\sigma$ and $1.1\sigma$, respectively. The combined result agrees with the SM predictions within $0.8\sigma$. This work constitutes the most precise measurements of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ performed to date and the first result for $\mathcal{R}(D)$ based on a semileptonic tagging method. The results of this analysis, together with the most recent Belle results on $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ [12,13] obtained using a hadronic tag, are combined to provide the Belle combination, which yields $\mathcal{R}(D) = 0.326 \pm 0.034$, $\mathcal{R}(D^*) = 0.283 \pm 0.018$ with a correlation equal to $−0.47$ between the $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ values. This combined result is in agreement with the SM predictions within $1.6$ standard deviations.

We thank the KEKB group for excellent operation of the accelerator; the KEK cryogenics group for efficient solenoid operations; and the KEK computer group, the NII, and PNBL/EMSL for valuable computing and SINET5 network support. We acknowledge support from MEXT, JSPS and Nagoya’s TLPRC (Japan); ARC (Australia); FWF (Austria); NSFC and CCEPP (China); MSMT (Czechia); CZF, DFG, EXC153, and VS (Germany); DST (India); INFN (Italy); MOE, MSIP, NRF, RSRI, FLRAS project, GSAC of KISTI and KREONET/GLORIAD (Korea); MNI SW and NCN (Poland); MSHE, Agreement 14.W03.10.0026 (Russia); ARRS (Slovenia); IKERBASQUE (Spain); SNSF (Switzerland); MOE and MOST (Taiwan); and DOE and NSF (USA). We acknowledge the support provided by the Albert Shimmins Fund for the writing of this Letter.

[1] S. P. Martin, Adv. Ser. Direct. High Energy Phys. 18, 1 (1998), arXiv:hep-ph/9708356 [hep-ph].
[2] J. F. Gunion, H. E. Haber, G. L. Kane, and S. Dawson, Front. Phys. 80, 1 (2000).
[3] W. Buchmuller, R. Ruckl, and D. Wyler, Phys. Lett. B191, 442 (1987) [Erratum: Phys. Lett. B448, 320 (1999)].
[4] Throughout this paper, the inclusion of the charge-conjugate decay mode is implied.
[5] D. Bigi and P. Gambino, Phys. Rev. D94, 094008 (2016), arXiv:1606.08030 [hep-ph].
[6] F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson, Phys. Rev. D95, 115008 (2017), [erratum: Phys. Rev. D97, 059902 (2018)], arXiv:1703.05330 [hep-ph].
[7] D. Bigi, P. Gambino, and S. Schacht, JHEP 11, 061 (2017) arXiv:1707.09509 [hep-ph].
[8] S. Jaiswal, S. Nandi, and S. K. Patra, JHEP 12, 060 (2017) arXiv:1707.09977 [hep-ph].
[9] Y. Ahmias et al. (Heavy Flavor Averaging Group), “Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018,” (2019), arXiv:1909.12524 [hep-ex].
[10] A. Matyja et al. (Belle Collaboration), Phys. Rev. Lett. 99, 191807 (2007), arXiv:0706.4429 [hep-ex].
[11] A. Bozek et al. (Belle Collaboration), Phys. Rev. D82, 072005 (2010), arXiv:1005.2302 [hep-ex].
[12] M. Huschle et al. (Belle Collaboration), Phys. Rev. D92, 072014 (2015), arXiv:1507.03233 [hep-ex].
[13] Y. Sato et al. (Belle Collaboration), Phys. Rev. D94, 072007 (2016), arXiv:1607.07923 [hep-ex].
[14] S. Hirose et al. (Belle Collaboration), Phys. Rev. Lett. 118, 211801 (2017), arXiv:1612.00529 [hep-ex], Phys. Rev. D97, 012004 (2018), arXiv:1709.00129 [hep-ex].
[15] J. P. Lees et al. (BaBar Collaboration), Phys. Rev. Lett. 109, 101802 (2012), arXiv:1205.5442 [hep-ex], Phys. Rev. D88, 072012 (2013), arXiv:1303.0571 [hep-ex].
[16] R. Aaij et al. (LHCb Collaboration), Phys. Rev. Lett. 115, 111803 (2015), arXiv:1506.08614 [hep-ex].
[17] R. Aaij et al. (LHCb Collaboration), Phys. Rev. D97,
[18] A. Abashian et al. (Belle Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A 479, 117 (2002); also see detector section in J. Brodzicka et al., Prog. Theor. Exp. Phys. 2012, 04D001 (2012).

[19] 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.

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

[21] R. Brun et al., GEANT3, CERN Report No. DD/EE/84-1, (1987).

[22] I. Caprini, L. Lellouch, and M. Neubert, Nucl. Phys. B530, 153 (1998); arXiv:hep-ph/9712417 [hep-ph].

[23] D. Scora and N. Isgur, Phys. Rev. D52, 2783 (1995).

[24] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, Phys. Rev. D57, 308 (1998).

[25] T. Keck et al., Computing and Software for Big Science 3, 6 (2019); arXiv:1807.08680 [hep-ex].

[26] M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018).

[27] M. Feindt and U. Kerzel, Nucl. Instrum. Methods Phys. Res., Sect. A 559, 190 (2006); proceedings of the X International Workshop on Advanced Computing and Analysis Techniques in Physics Research.

[28] T. Chen and C. Guestrin, (2016), proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, arXiv:1603.02754 [cs.LG].