Measurement of $\mathcal{B}(B_s \to D_s X)$ with $B_s$ Semileptonic Tagging

B. Wang, K. Kinoshita, H. Aihara, D. M. Asner, T. Aushev, R. Ayad, V. Babu, I. Badhrees, A. M. Bakich, P. Behera, C. Belénio, J. Bennett, M. Bessner, V. Bhardwaj, T. Bilka, J. Biswal, A. Bobrov, G. Bonvicini, A. Bozek, M. Bracko, T. E. Browder, M. Campajola, L. Cao, D. Červenkov, A. Chen, K. Chilikin, G. B. Mohanty, K. Lieret, R. Kroeger, B. Pal, K. Inami, K. T. Kim, J.-G. Shiu, A. Giri, T. Matsuda, G. Russo, A. Bobrov, K. Nishimura, A. Kuzmin, Y. Jin, S. Jia, Y. Guan, O. Werbycka, R. Van Tonder, J. Bennett, V. Zhulanov, →s

(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, Kwangju 660-701
(7) University of Cincinnati, Cincinnati, Ohio 45221
(8) Deutsches Elektronen–Synchrotron, 22607 Hamburg
(9) Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) and Institute of Modern Physics, Fudan University, Shanghai 200443
(10) Justus-Liebig-Universität Gießen, 35392 Gießen

Typeset by REVTeX
II. Physikalisches Institut, Georg-August-Universität Göttingen, 37073 Göttingen
12 SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193
13 Gyeongsang National University, Chinju 660-701
14 Hangyang University, Seoul 133-791
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 Institute of High Energy Physics,
Chinese Academy of Sciences, Beijing 100049
26 Institute of High Energy Physics, Vienna 1050
27 Institute for High Energy Physics, Protvino 142281
28 INFN - Sezione di Napoli, 80126 Napoli
29 INFN - Sezione di Torino, 10125 Torino
30 Advanced Science Research Center,
Japan Atomic Energy Agency, Naka 319-1195
31 J. Stefan Institute, 1000 Ljubljana
32 Institut für Experimentelle Teilchenphysik,
Karlsruher Institut für Technologie, 76131 Karlsruhe
33 Kennesaw State University, Kennesaw, Georgia 30144
34 King Abdulaziz City for Science and Technology, Riyadh 11442
35 Korea Institute of Science and Technology Information, Daejeon 305-806
36 Korea University, Seoul 136-713
37 Kyungpook National University, Daegu 702-701
38 LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay
39 École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015
40 P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow 119991
41 Faculty of Mathematics and Physics,
University of Ljubljana, 1000 Ljubljana
42 Ludwig Maximilians University, 80539 Munich
43 Luther College, Decorah, Iowa 52101
44 University of Maribor, 2000 Maribor
45 Max-Planck-Institut für Physik, 80805 München
46 School of Physics, University of Melbourne, Victoria 3010
47 University of Mississippi, University, Mississippi 38677
48 University of Miyazaki, Miyazaki 889-2192
49 Moscow Institute of Physics and Technology, Moscow Region 141700
50 Graduate School of Science, Nagoya University, Nagoya 464-8602
51 Università di Napoli Federico II, 80055 Napoli
52 Nara Women’s University, Nara 630-8506
Abstract

We report the first direct measurement of the inclusive branching fraction $\mathcal{B}(B_s \to D_sX)$ via $B_s$ tagging in $e^+e^- \to \Upsilon(5S)$ events. Tagging is accomplished through a partial reconstruction of semileptonic decays $B_s \to D_sX\ell\nu$, where $X$ denotes unreconstructed additional hadrons or photons and $\ell$ is an electron or muon. With 121.4 fb$^{-1}$ of data collected at the $\Upsilon(5S)$ resonance by the Belle detector at the KEKB asymmetric-energy $e^+e^-$ collider, we obtain $\mathcal{B}(B_s \to D_sX) = (61.6 \pm 5.3 \pm 2.1)\%$, where the first uncertainty is statistical and the second is systematic.
The study of \( B_s \)-meson properties at the \( \Upsilon(5S) \) resonance may provide important insights into the CKM matrix and hadronic structure, as well as sensitivity to new physics phenomena [13]. The branching fraction for the inclusive decay \( B_s \to D_s X \) plays an important role in the determination of the \( B_s \) production rate in \( \Upsilon(5S) \) events[4]. This rate, usually expressed as the fraction \( f_s \) of \( b \bar{b} \) events at the \( \Upsilon(5S) \), is necessary for measuring absolute rates and branching fractions. Two experiments at LEP, ALEPH [5] and OPAL [6], measured the product branching fraction \( \mathcal{B}(b \to B_s^0) \cdot \mathcal{B}(B_s^0 \to D_s X) \). The branching fraction \( \mathcal{B}(B_s^0 \to D_s X) \) was evaluated using a model-dependent value of \( \mathcal{B}(b \to B_s^0) \) and was subject to large statistical and theory uncertainties. Belle measured the branching fractions of \( \Upsilon(5S) \to D_s X \) and \( \Upsilon(5S) \to D^0 X \) [7] with 1.86 fb\(^{-1} \) of data collected at the \( \Upsilon(5S) \) energy. These are related to the inclusive \( B_s \) branching fractions to \( D_s \) and \( D^0/\bar{D}^0 \) by the following relations,

\[
\mathcal{B}(\Upsilon(5S) \to D_s X)/2 = f_s \cdot \mathcal{B}(B_s \to D_s X) + f_q \cdot \mathcal{B}(B \to D_s X),
\]

where \( D_s \) is \( D_s \) or \( D^0/\bar{D}^0 \), \( f_s \) is the fraction of \( \Upsilon(5S) \) events containing \( B_s \)-meson pairs, and \( f_q \) is the fraction containing charged or neutral \( B \) pairs. Using the measured value of \( \mathcal{B}(\Upsilon(5S) \to D^0 X) \) [4], and assuming \( f_q = 1 - f_s \) and \( \mathcal{B}(B_s \to D^0 X + \text{c.c.}) = 8 \pm 7\% \) [8], which was estimated based on phenomenological arguments, Belle found \( f_s = (18.1 \pm 3.6 \pm 7.5)\% \) [7]. This input, with the measured \( \mathcal{B}(\Upsilon(5S) \to D_s X) \) [7], was used to evaluate \( \mathcal{B}(B_s \to D_s X) = (91 \pm 18 \pm 41)\% \) [7]. The current world average, \( (93 \pm 25)\% \) [9], is based on measurements made with the methods described above, which rely on model-dependent assumptions.

In this paper, we present the first direct measurement of \( \mathcal{B}(B_s \to D_s X) \) using a \( B_s \) semileptonic tagging method with \( \Upsilon(5S) \) events. Throughout this paper, the inclusive \( B_s \) branching fractions to \( D_s \) and \( D^0/\bar{D}^0 \) is defined as the mean number of \( D_s \)-mesons per \( B_s \) decay.

We use a data sample of 121.4 fb\(^{-1} \), collected with the Belle detector [10] at the KEKB asymmetric-energy \( e^+e^- \) collider [11] operating near the \( \Upsilon(5S) \) resonance. The Belle detector is a general-purpose large-solid-angle spectrometer consisting of a silicon vertex detector (SVD), a central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. Outside the coil, an iron flux-return yoke is instrumented to detect \( K_L^0 \)-mesons and to identify muons (KLM). A detailed description of the detector can be found in Ref. [10].

All charged tracks, except those from \( K_L^0 \) decay, are required to be consistent with originating from the interaction point (IP), with the point of closest approach to the IP within 2.0 cm along the beam axis and within 0.5 cm in the plane transverse to the beam. Additionally, all tracks must have, within the SVD, at least one associated hit in the plane transverse to the beam and two hits along the beam axis. To suppress the continuum background from \( e^+e^- \to q\bar{q} \) with \( q = u, d, s, \) or \( c \), we require that the variable \( R_2 \), the ratio of second- to zeroth-order Fox-Wolfram moments [12], be less than 0.4. Kaon and pion hypotheses are assigned to the tracks based on likelihood, which is calculated using information from the Cherenkov light yield in the ACC, the time-of-flight information of the TOF, and the specific ionization \( (dE/dx) \) in the CDC. Charged kaon (pion) candidates are required to have a kaon/pion likelihood ratio \( \mathcal{L}_K/(\mathcal{L}_K + \mathcal{L}_\pi) > 0.6 \) (< 0.6). The angle between each lepton and the positron beam is required to be between 18° and 150° for electrons and between
twenty-five and fourty-five degrees for muons. Selected electrons and muons must have a minimum momentum of 1.0 GeV/c in the $e^+e^-$ center-of-mass (CM) frame. An electron/pion likelihood ratio ($L_e$) is calculated based on information from the CDC, ACC, and ECL. A muon/hadron likelihood ratio is calculated based on information from the KLM. Tracks with $L_e > 0.8$ ($L_\mu > 0.8$) are included as electrons (muons) in the analysis. The efficiency for electron (muon) tracks to pass this criterion is (94.7 ± 0.2)\% ((96.7 ± 0.2)\%).

The neutral intermediate particles $\phi, K_S^0$ and $K^0$ are reconstructed from charged tracks. For $\phi \rightarrow K^+K^-$ reconstruction, any pair of oppositely charged kaons with invariant mass within 15 MeV/c$^2$ of the $\phi$ nominal mass\cite{9} is considered to be a $\phi$ candidate. The $K_S^0$ candidates are reconstructed via the decay $K_S^0 \rightarrow \pi^+\pi^-$, following standard criteria\cite{14}, and are further required to have an invariant mass within 20 MeV/c$^2$ ($\approx$4.4 $\sigma$ in resolution) of the nominal mass. For $K^{*0} \rightarrow K^+\pi^-$, the candidate tracks are oppositely charged $K$ and $\pi$, with invariant mass within 50 MeV/c$^2$.

Candidates for $D_s^\pm$ are reconstructed in the final states $\phi\pi^+, K_S^0K^+$, and $\bar{K}^{*0}K^+$. The CM momentum of the candidate is required to be in the range 0.5 GeV/c – 3.0 GeV/c. Candidates with invariant mass in the range 1.92-2.22 GeV/c$^2$ are considered. For $\phi\pi^+$ and $\bar{K}^{*0}K^+$ modes, a vertex fit is performed for the three tracks used to reconstruct the candidate, and the $\chi^2$ of the fit output is required to be less than 100. Nearly all correctly reconstructed $D_s^0$, (98.1 ± 0.1)\%, are found to pass this requirement. The decays $D_s^+ \rightarrow \phi(K^+K^-)\pi^+$ and $D_s^+ \rightarrow \bar{K}^{*0}(K^-\pi^+)K^+$ are transitions of a pseudoscalar particle to a vector and a pseudoscalar, with the vector decaying to two pseudoscalars. To suppress combinatorial background, we require $|\cos \theta_{\text{hel}}| > 0.5$, where the helicity angle $\theta_{\text{hel}}$ is defined as the angle between the momentum of the $D_s^+$ and $K^+ (\pi^+)$ in the rest frame of the $\phi$ ($\bar{K}^{*0}$) resonance.

We tag $B_s$ events through a “partial reconstruction” of the semileptonic decay $B_s^0 \rightarrow D_s^-X\ell^+\nu$, with the $D_s^-$ modes $\phi\pi^-$ and $K_S^0K^-$, using a procedure similar to one applied at the $\Upsilon(4S)$ resonance\cite{15}, where a lepton (electron or muon) is paired with a charm meson to form a $B$ candidate. In contrast to the $\Upsilon(4S)$, where the exclusive production of $BB$ ensures that each $B$-meson’s total energy is half the CM energy, $\sqrt{s}/2$, the $B_s$’s in $\Upsilon(5S)$ events occur predominantly in $B_s^*\bar{B}_s^*$ events. In this case the energy of each $B_s$ is well approximated as $\sqrt{s}/2 - \delta E$, where $\delta E/c^2$ is the $B_s^* - B_s$ mass difference. We use $\delta E$=47.3 MeV. We thus define the “missing mass squared” of the selected $D_s^-\ell^+$ candidate as

$$M_{\text{miss}}^2 = (\sqrt{s}/2 - \delta E - E_{D\ell}^*)^2 - (p_{D\ell}^*)^2,$$

where $E_{D\ell}^*$ and $p_{D\ell}^*$ are the energy and momentum of the $D_s\ell$ system in the CM frame. The distribution in $M_{\text{miss}}^2$ for tagged $B_s$ represents the undetected neutrino plus additional low-momentum daughters of excited $D_s$, photons and pions, and is expected to peak broadly at $M_{\text{miss}}^2 = 0$. The thrust angle, $\theta_{\text{thrust}}$, is defined as the angle between the thrust axis\cite{16} of the selected $D_s\ell$ system and that of the remaining tracks in the event. To suppress continuum background, we require $|\cos \theta_{\text{thrust}}| < 0.8$. In events with more than one tag candidate, we perform a combined fit on each candidate’s three-track $D_s$ vertex, and on the vertex of the extrapolated $D_s$ trajectory with the lepton, and select the candidate having the smallest $\chi^2$.

The number of $B_s$ tags for each $D_s$ decay channel is found by a binned maximum-likelihood fit of the distribution in $M_{\text{miss}}^2$ and $M_{D_s}$ to a sum of three components, according to candidate origin:

1. Correctly tagged candidates
| Tag Channel | Signal Channel | Efficiency (%) |
|-------------|----------------|----------------|
| $\phi\pi$   | $\phi\{K^+K^-\}\pi$ | 26.1 ± 0.5     |
|             | $K_S^0\{\pi^+\pi^-\}K$ | 38.5 ± 0.6     |
|             | $K^*0\{K^\pm\pi^\mp\}K$ | 24.6 ± 0.5     |
| $K_S^0K$    | $\phi\{K^+K^-\}\pi$ | 27.6 ± 0.5     |
|             | $K_S^0\{\pi^+\pi^-\}K$ | 37.8 ± 0.6     |
|             | $K^*0\{K^\pm\pi^\mp\}K$ | 24.6 ± 0.4     |

TABLE I. Signal-side $D_s$ reconstruction efficiencies, by tag-side and signal-side $D_s$ decay channel.

2. Incorrect tag, where a lepton from a $B_s$ semileptonic decay is paired with a real $D_s$ from the other $B_s$.

3. Other incorrect tags: all other sources of candidates.

The $M_{\text{miss}}^2$ distribution for each is taken to be a histogram obtained via Monte Carlo (MC) simulation. Decays $B_s \to D_s X \ell \nu$ are modeled as a combination of $B_s \to D_s \ell \nu$ and $B_s \to D_s^* \ell \nu$, with no contributions from higher mass $D_s$ states. The data are found to be consistent with this model. For correctly reconstructed $D_s$, the invariant mass distribution is modeled by a sum of two Gaussians with a common mean. The widths of the Gaussians and their relative areas are obtained from MC simulation. For combinatorial $D_s$ background, the distribution is well-represented by a linear function. We find $N_{\phi\pi}^{\text{tag}} = 6473 \pm 119$ and $N_{K^0S}^{\text{tag}} = 4435 \pm 126$. The fit results for $D_s \to \phi\pi$ are shown in Fig. 1.

After selecting a $B_s$ candidate as the tag, we reconstruct the “signal-side” $D_s$ from the remaining tracks in the event. Candidates are reconstructed in all three modes discussed earlier, and we allow none of the tracks from the selected tag candidate to be used. The number of signal $D_s$ in tagged events is found through a binned 3D maximum-likelihood fit in $M_{\text{miss}}^2$ and the invariant masses of the tag- and signal-side $D_s$ candidates. Each signal candidate is associated on the tag side with one of the three components used for the tag fit and on the signal side with a real or combinatorial $D_s$, leading to six distinct contributions to the 3D distribution.
The signal-side $D_s$ reconstruction efficiency, $\varepsilon_{\text{sig}}$, is defined for each mode as the number, $N_{\text{sig}}$, of signal-side $D_s$ reconstructed in a sample of $N_{\text{gen}}$ MC events containing a correctly selected tag and a $D_s$ decaying in the reconstructed mode:

$$\varepsilon_{\text{sig}} = \frac{N_{\text{sig}}}{N_{\text{gen}}}. \quad (3)$$

Table I summarizes the reconstruction efficiencies, for each combination of tag-side and signal-side $D_s$ channel.

The signal, two real $D_s$ reconstructed in events with $B_s \to D_s X\ell\nu$, can appear in our distributions not only as a correctly selected tag (type 1, above) with a real signal $D_s$ but also with an incorrect tag (type 2, above) consisting of a tag-side lepton and a signal-side $D_s$, where a $B_s$ has undergone particle-antiparticle oscillation and the charges of the two $D_s$ are the same. We describe this occurrence as “cross-feed.” Because the cross-feed rate is proportional to the signal rate, the associated distribution is included as signal in the fit, with the efficiency ratio for each pair of $D_s$ channels obtained via MC simulation and fixed.

The fit is performed simultaneously for the six pairs of $D_s$ tag-signal channel combinations, constraining the relative efficiencies and intermediate branching fractions, such that the fitted parameter is the ratio, $B_{\text{raw}}$, of the number of $D_s$ in events containing a tag (corrected for reconstruction efficiencies and intermediate branching fractions) to the number of tags.

To confirm the 3D fitting procedure and correction, we performed a “linearity test” using a large set of MC-generated signal and background events. Ten ensembles of 200 independent MC signal samples, each corresponding to a branching fraction in the range 10-100% in 10% increments, were fitted and the resulting branching fractions plotted as a function of the input value. A linear fit showed consistency with a unit slope and no systematic shifts, for each of the six $D_s$ mode combinations and for the simultaneous fit.

Our fit yields $B_{\text{raw}} = (58.2 \pm 5.8)\%$, which corresponds to a fitted total of 101 $\pm 10$ signal and 66 $\pm 6$ cross-feed events. Projections of the fit are shown in Fig. 2. To confirm the 3D fitting procedure and correction, we performed a “linearity test” using a large set of MC-generated signal and background events. Ten ensembles of 200 independent MC signal samples, each corresponding to a branching fraction in the range 10-100% in 10% increments, were fitted and the resulting branching fractions plotted as a function of the input value. A linear fit showed consistency with a unit slope and no systematic shifts, for each of the six $D_s$ mode combinations and for the simultaneous fit.

Our fit yields $B_{\text{raw}} = (58.2 \pm 5.8)\%$, which corresponds to a fitted total of 101 $\pm 10$ signal and 66 $\pm 6$ cross-feed events. Projections of the fit are shown in Fig. 2. To confirm the 3D fitting procedure and correction, we performed a “linearity test” using a large set of MC-generated signal and background events. Ten ensembles of 200 independent MC signal samples, each corresponding to a branching fraction in the range 10-100% in 10% increments, were fitted and the resulting branching fractions plotted as a function of the input value. A linear fit showed consistency with a unit slope and no systematic shifts, for each of the six $D_s$ mode combinations and for the simultaneous fit.

Our fit yields $B_{\text{raw}} = (58.2 \pm 5.8)\%$, which corresponds to a fitted total of 101 $\pm 10$ signal and 66 $\pm 6$ cross-feed events. Projections of the fit are shown in Fig. 2. To confirm the 3D fitting procedure and correction, we performed a “linearity test” using a large set of MC-generated signal and background events. Ten ensembles of 200 independent MC signal samples, each corresponding to a branching fraction in the range 10-100% in 10% increments, were fitted and the resulting branching fractions plotted as a function of the input value. A linear fit showed consistency with a unit slope and no systematic shifts, for each of the six $D_s$ mode combinations and for the simultaneous fit.
in $B_{\text{raw}}$ as the uncertainty. In cases affecting the $D_s$ mode combinations separately, the maximum excursion is taken as a conservative estimate of the uncertainty on the combined result. Because this measurement involves tagging, many of the systematic uncertainties associated with tagging cancel approximately in taking the ratio of tags, with and without signal. The effect from the uncertainty due to the composition and model of $B_s \to D_s X \ell \nu$ on efficiency and on the $M^2_{\text{miss}}$ fitting shape is estimated by varying the relative rates of $B_s \to D_s \ell \nu$ and $B_s \to D_s^* \ell \nu$ within the uncertainties\cite{9} and by varying the HQET2 parameters in the MC generator by $\pm 10\%$. For the “other incorrect tag” (type 3, above), the $M^2_{\text{miss}}$ distribution in data from tags with “sideband” $D_s$ candidates, $|M_{\text{cand}} - m_{D_s} \pm 40| < 10$ MeV, is substituted in the fit. Uncertainties due to fitting of the $D_s$ mass distributions are determined by changing the signal shape from two Gaussians to three and the background from a first-order to a second-order polynomial. We vary each ratio of signal to cross-feed efficiency in the fit by $\pm 1\sigma$. The uncertainties due to branching fractions of the reconstructed $D_s$ decays are estimated by varying each by $\pm 1\sigma$\cite{9} of its value in the fitting procedure. The reconstruction efficiencies are varied by the amount of their statistical error from the MC sample. The uncertainty due to the limited statistical power of our linearity test is estimated by varying the parameters from the linear fit by $\pm 1\sigma$. To estimate effects from our selection of a single tag candidate per event, we reanalyze the data using random selection and take...
the shift in the result to be the uncertainty.

The uncertainty on the tracking efficiency affects only the three signal-side tracks comprising the \( D_s \) candidate and is estimated to be 0.35\% per track, thus, we take 1.1\% as the uncertainty from this source. The systematic uncertainty from \( K-\pi \) identification efficiencies is estimated to be 1.3\%.

The fitted shape of the \( M^{2}_{\text{miss}} \) distribution depends on the \( B_s - B_s \) mass difference, \( \delta E/c^2 \), and its uncertainty may affect the fit in two ways: in the value used to generate the MC signal events (vs the actual value) and in the value used to calculate \( M^{2}_{\text{miss}} \). For this analysis, the values are 45.9 MeV/c^2 for MC generation and 47.3 MeV/c^2 for \( M^{2}_{\text{miss}} \). The PDG presents two numbers, (46.1 ± 1.5) MeV/c^2 as a world average and a PDG fit of (48.6^{+1.8}_{-1.5}) MeV/c^2 [9]. As \( M^{2}_{\text{miss}} \) is fitted in both the numerator and denominator to obtain \( \mathcal{B}_{\text{raw}} \), effects from such differences are expected to cancel, at least in part. To estimate possible systematic shifts due to these differences, we vary separately the calculation using such differences are expected to cancel, at least in part. To estimate possible systematic shifts due to these differences, we vary separately the calculation using \( \delta E/c^2 \) and the value used in MC generation in the range 45.9-49.0 MeV/c^2. Changing the calculation of \( M^{2}_{\text{miss}} \) results in a maximum excursion in \( \mathcal{B}_{\text{raw}} \) of less than 0.1\%. Changing the value in the MC generator results in a maximum excursion of 1.2\%. We assign an uncertainty of 1.2\%.

We consider possible contributions to the tag-side sample from the non-strange \( B \) decay \( \mathcal{B}(B \rightarrow D_s^{(*)} K \ell \nu) \), which is not included in our generic MC generator. We use \( \mathcal{B}(B^+ \rightarrow D_s^{(*)} K^+ \ell^- \nu) = (6.1 \pm 1.0) \times 10^{-4} \) [9], assume that \( \mathcal{B}(B^0 \rightarrow D_s^{(*)} K^0 \ell^+ \nu) \) is the same, and multiply by a factor of two to account for both electrons and muons. Taking \( \mathcal{B}(\Upsilon(5S) \rightarrow B\bar{B}X = 76\%, \mathcal{B}(\Upsilon(5S) \rightarrow B_s B_s X = 20\%, \) and \( \mathcal{B}(B_s \rightarrow X \ell \nu) = 9.6\%\) [9], we estimate

\[
\frac{\mathcal{B}(\Upsilon(5S) \rightarrow B\bar{B}X) \cdot \mathcal{B}(B \rightarrow D_s^{(*)} K \ell \nu)}{\mathcal{B}(\Upsilon(5S) \rightarrow B_s^{(*)} \bar{B}_s^{(*)}) \cdot \mathcal{B}(B_s \rightarrow D_s X \ell \nu)} \approx 0.048.
\]

As the shape in \( M^{2}_{\text{miss}} \) includes a kaon in addition to the neutrino, it is expected to peak more broadly and at a higher value than does the \( B_s \) channel. This is confirmed in studies of MC-generated \( B\bar{B}X \) events containing \( B \rightarrow D_s^{(*)} K \ell \nu \) in the \( D_s \) tag modes. Fig. 3 illustrates the difference. We measure the effect on our MC tag fit of including such events, and estimate a contribution to \( \mathcal{B}(B_s \rightarrow D_s X \ell \nu) \) of < 0.02\% (0.5\%) to the \( D_s \rightarrow \phi \pi \) \( (D_s \rightarrow K_0^0 K) \) channel. We assign an overall systematic uncertainty of 0.5\%. The uncertainties from the above sources are summed in quadrature to arrive at the total fractional systematic uncertainty in \( \mathcal{B}_{\text{raw}} \) of 3.8\%.

After propagating the systematic error on \( \mathcal{B}_{\text{raw}} \) to the branching fraction and adding in quadrature the uncertainty from \( \mathcal{B}_{D_s \ell} \), we find

\[
\mathcal{B}(B_s \rightarrow D_s X) = (61.6 \pm 5.3 \pm 2.1\%)\.
\]

The central value is lower than the theoretical expectation (86^{+8}_{-13})\% [17], and ≈ 1.2\% below the world average (93 ± 25)\% [9]. Given the history of uncertainty on the rates and composition of charm states at higher mass in \( B \) decay, a lower value may be explained by a rate of \( c\bar{s} \) to \( D \) vs. \( D_s \) that is higher than anticipated. The implications of a lower central value are notable. Experimentally, the value affects the derived fraction \( f_s \) of \( B_s \) events among \( \Upsilon(5S) \) decays, which impacts the absolute normalization of all \( B_s \) branching fractions measured via \( \Upsilon(5S) \) decays. In the earlier Belle measurements of \( f_s \) [7, 13], Eq. 1 was used with \( f_q = 1 - f_s \). More recently, it has been found that there is a nonzero rate to bottomonia, including \( \Upsilon(1S), \Upsilon(2S), \Upsilon(3S), h_b(1P) \) and \( h_b(2P) \). We take the rate of events with “no open bottom” to be \( f_{nob} = 4.9^{+5.0}_{-0.6}\% [19] \). Charm is highly suppressed in
these decays, so we take \( f_s = 1 - f_q - f_{\text{nob}} \). Using \( B(\Upsilon(5S) \to D_sX) = (45.4 \pm 3.0)\% \) \cite{20} and \( B(B \to D_sX) = (8.3 \pm 0.8)\% \) \cite{9}, we solve Eq. 1 for \( f_s \) and find

\[
 f_s = 0.278 \pm 0.028(\text{stat}) \pm 0.035(\text{sys}) \tag{7}
\]

This value is larger than the world average, \( f_s = 0.201 \pm 0.031 \) \cite{9}, which is evaluated assuming the model-based estimates \( B(B_s \to D_sX) = (92 \pm 11)\% \) and \( B(B \to D_sX) = (8 \pm 7)\% \) \cite{7}; the impact of introducing \( f_{\text{nob}} \) to the calculation is minor. Our result uses the same value of \( B(\Upsilon(5S) \to D_sX) \) from which \( f_s \) is derived in \cite{18} and thus supersedes the value presented there. It is consistent with a recent Belle measurement of \( f_s \) by an independent method \cite{19}. An older Belle measurement of \( f_s \) from semileptonic decays \cite{21} assumed that only \( D_{s1} \) and \( D_{s2} \) contribute to non-strange charm, \( B_s \to DKX\ell\nu \). Given recently reported evidence of substantial contributions from nonresonant \( DK(X) \) \cite{22}, this value is likely an underestimate, so we do not compare it with the result reported here.

Applying Eq. 1 with \( B(B \to D^0/\bar{D}^0X) = (61.5 \pm 2.9)\% \) \cite{9}, \( B(\Upsilon(5S) \to D^0X) = (108 \pm 8)\% \) \cite{9}, and our result for \( f_s \), we find \( B(B_s \to D^0X) = (45 \pm 2(\text{stat}) \pm 19(\text{sys}))\% \), where the systematic uncertainties on \( B(\Upsilon(5S) \to D^0X) \) and \( f_{\text{nob}} \) dominate. This value is consistent with our finding of a lower rate of \( D_s \) from \( B_s \) decay, as the total charm content would need to be accounted for by an increased rate of nonstrange charm. No experimental results for \( B_s \to D^0X \) are currently included in the PDG tables \cite{9}.

We have made the first direct measurement of the \( B_s \to D_sX \) inclusive branching fraction, using a \( B_s \) semileptonic tagging method at the \( \Upsilon(5S) \) resonance. We find

\[
 B(B_s \to D_sX) = (61.6 \pm 5.3(\text{stat}) \pm 2.1(\text{sys}))\%,
\]

which is substantially lower than the world average but consistent within its large uncertainties. This result is used to recalculate the fraction \( f_s \) of \( \Upsilon(5S) \) events containing \( B_s \),

\[
 f_s = 0.278 \pm 0.028(\text{stat}) \pm 0.035(\text{sys}).
\]

This value supersedes that reported in \cite{18}.

We thank the KEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; and the KEK computer group, and the Pacific Northwest National Laboratory (PNNL) Environmental Molecular Sciences Laboratory (EMSL) computing group for strong computing support; and the National Institute
| Source                      | Channel                  | Combined |
|-----------------------------|--------------------------|----------|
| Source                      | Channel                  | Combined |
| Model, tag                  | φπ Tag                   | φπ K₀⁺K K⁺0 K |
| Model, signal               | 0.1 0.1 0.3              | 0.1 0.1 0.1 |
| Model, cross-feed           | 0.4 0.3 0.3              | 0.2 0.1 0.1 |
| M₂miss shape, M_Bs⁻ - M_Bs  | 1.2                      | 1.2      |
| M₂miss background           | 0.1 0.2 0.1              | 0.5 0.2 0.3 |
| M(D_s) signal shape         | 0.2 0.2 1.2              | 0.1 0.1 1.0 |
| M(D_s) background shape     | 1.0 0.6 <0.1             | <0.1 0.1 0.1 |
| Cross-feed efficiency       | 0.5 0.3 0.6              | 0.3 0.1 0.3 |
| Reconstruction efficiency   | 0.4 0.2 0.4              | 0.2 0.1 0.2 |
| Statistics, linearity test  | 0.2 0.3 0.3              | 0.3 0.4 0.4 |
| B → D_s⁽⁺⁻⁾Kℓν              | <0.02                    | 0.5      |
| B(D_s → φπ)                 | -                        | 1.2      |
| B(D_s → K₀⁺K)               | -                        | 0.5      |
| B(D_s → K⁺⁺K)               | -                        | 1.2      |
| Single tag selection        | -                        | 1.0      |
| Tracking                    | -                        | 1.1      |
| K-π identification          | -                        | 1.3      |
| Total                       | -                        | 3.8      |

TABLE II. Systematic uncertainties on B_{raw}, in %. The total is the sum in quadrature from all sources.

of Informatics, and Science Information NETwork 5 (SINET5) for valuable 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 including grants DP180102629, DP170102389, DP170102204, DP150103061, FT130100303; Austrian Federal Ministry of Education, Science and Research (FWF) and FWF Austrian Science Fund No. P 31361-N36; the National Natural Science Foundation of China under Contracts No. 11435013, No. 11475187, No. 11521505, No. 11575017, No. 11675166, No. 11705209; Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS), Grant No. QYZDJ-SSW-SLH011; the CAS Center for Excellence in Particle Physics (CCEPP); the Shanghai Pujiang Program under Grant No. 18PJ140100; the Shanghai Science and Technology Committee (STCSM) under Grant No. 19ZR1403000; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020; Horizon 2020 ERC Advanced Grant No. 884719 and ERC Starting Grant No. 947006 “InterLeptons” (European Union); the Carl Zeiss Foundation, the Deutsche Forschungsgemeinschaft, the Excellence Cluster Universe, and the VolkswagenStiftung; the Department of Atomic Energy (Project Identification No. RTI 4002) and the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; National Research Foundation (NRF) of Korea Grant Nos. 2016R1D1A1B01010135, 2016R1D1A1B02012900, 2018R1A2B3003643,
2018R1A6A1A06024970, 2018R1D1A1B07047294, 2019K1A3A7A09033840, 2019R1I1A3A-01058933; Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project, the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information and KREONET/GLORIAD; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Science and Higher Education of the Russian Federation, Agreement 14.W03.31.0026, and the HSE University Basic Research Program, Moscow; University of Tabuk research grants S-1440-0321, S-0256-1438, and S-0280-1439 (Saudi Arabia); the Slovenian Research Agency Grant Nos. J1-9124 and P1-0135; Ikerbasque, Basque Foundation for Science, Spain; the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology of Taiwan; and the United States Department of Energy and the National Science Foundation.

[1] A. F. Falk and A. A. Petrov, Phys. Rev. Lett. 85, 252 (2000).
[2] S. Petrak, SLAC Report No. 2001-041 (2001).
[3] D. Atwood and A. Soni, Phys. Lett. B 533, 37 (2002).
[4] Throughout this paper, “$D_s$” denotes both $D_s^+$ and $D_s^-$ and “$B_s$” denotes both $B_s^0$ and $\bar{B}_s^0$.
[5] D. Buskulic et al., Z. Phys. C 69, 585 (1996).
[6] P.D. Acton et al., Phys. Lett. B 295, 357 (1992).
[7] A. Drutskoy et al. (Belle Collaboration), Phys. Rev. Lett. 98, 052001 (2007).
[8] M. Artuso et al. (CLEO Collaboration), Phys. Rev. Lett. 95, 261801 (2005).
[9] P.A. Zyla et al., (Particle Data Group), Prog. Theor. Exp. Phys. 2020, 083C01 (2020).
[10] A. Abashian et al., (Belle Collaboration), Nucl. Instrum. Methods A 479, 117 (2002).
[11] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods A 499, 1 (2003).
[12] G.C. Fox and S. Wolfram, Phys. Rev. Lett. 41, 1581 (1978).
[13] Inclusion of charge-conjugate modes is implied throughout this paper.
[14] K.-F. Chen et al. (Belle Collaboration), Phys. Rev. D 72, 012004 (2005).
[15] D. Bortoletto et al. (CLEO collaboration), Phys. Rev. Lett. 63, 1667 (1989).
[16] S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, Phys. Lett. 12, 57 (1964).
[17] M. Suzuki, Phys. Rev. D 31, 1158 (1985).
[18] S. Esen et al. (Belle Collaboration), Phys. Rev. D 87, 031101(R) (2013).
[19] R. Mizuk et al. (Belle Collaboration), [arXiv:2104.08371 [hep-ex]]; submitted to JHEP.
[20] This value is obtained by Belle using $121.4 \text{ fb}^{-1}$ of data and the method described in Ref. [7], with $D_s^+$ reconstructed in the mode $K^+K^-\pi^+$.
[21] C. Oswald et al. (Belle Collaboration), Phys. Rev. D 92, 072013 (2015).
[22] R. Aaij et al. (LHCb Collaboration), Phys. Rev. D 100, 031102 (2019).