Measurement of $\mathcal{B}(B_s \to D_s X)$ with B_s semileptonic tagging

B. Wang,[7,45] K. Kinoshita,[7] H. Aihara,[77] D. M. Asner,[3] T. Aushev,[49] R. Ayad,[72] V. Babu,[8] I. Badhrees,[72,34] A. M. Bakich,[71] P. Behera,[24] C. Beleño,[11] J. Bennett,[47] M. Bessner,[15] V. Bhardwaj,[20] T. Bilka,[5] J. Biswal,[31] A. Bobrov,[4,59] G. Bonvicini,[80] A. Bozek,[56] M. Bračko,[44,31] T. E. Browder,[15] M. Campajola,[28,51] L. Cao,[32] D. Červenkov,[5] A. Chen,[53] K. Chilikin,[40] H. E. Cho,[14] K. Cho,[35] S.-K. Choi,[13] Y. Choi,[70] D. Cinabro,[80] S. Cunliffe,[8] N. Dash,[21] F. Di Capua,[28,51] S. Di Carlo,[38] Z. Doležal,[5] S. Eidelman,[4,59,40] D. Epifanov,[4,59] J. E. Fast,[61] T. Ferber,[8] B. G. Fulsom,[61] R. Garg,[62] V. Gaur,[79] A. Garmash,[4,59] A. Giri,[23] P. Goldenzweig,[32] Y. Guan,[7] K. Hayasaka,[58] H. Hayashii,[52] W.-S. Hou,[55] K. Inami,[50] A. Ishikawa,[16] M. Iwasaki,[60] Y. Iwasaki,[16] S. Jia,[2] Y. Jin,[77] D. Joffe,[33] K. K. Joo,[6] A. B. Kaliyar,[24] K. H. Kang,[37] G. Karyan,[8] C. Kiesling,[45] D. Y. Kim,[69] K. T. Kim,[36] S. H. Kim,[14] P. Kodyš,[5] S. Korpar,[44,31] D. Kotchetkov,[15] P. Križan,[41,31] R. Kroeger,[47] P. Krokovny,[4,59] T. Kuhr,[42] A. Kuzmin,[4,59] Y.-J. Kwon,[82] J. S. Lange,[10] I. S. Lee,[14] J. Y. Lee,[67] S. C. Lee,[37] L. K. Li,[25] Y. B. Li,[63] L. Li Gioi,[45] J. Libby,[24] K. Lieret,[42] D. Liventsev,[79,16] T. Luo,[9] J. MacNaughton,[48] C. MacQueen,[46] M. Masuda,[76] T. Matsuda,[48] D. Matvienko,[4,59,40] M. Merola,[28,51] K. Miyabayashi,[52] R. Mizuk,[40,49] G. B. Mohanty,[73] R. Mussa,[29] M. Nakao,[16,12] K. J. Nath,[22] M. Nayak,[80,16] N. K. Nisar,[64] S. Nishida,[16,12] K. Nishimura,[15] S. Ogawa,[74] H. Ono,[57,58] Y. Onuki,[77] G. Pakhlova,[40,49] B. Pal,[3] T. Pang,[64] S. Pardi,[28] H. Park,[37] S.-H. Park,[82] T. K. Pedlar,[43] R. Pestotnik,[31] L. E. Piilonen,[79] V. Popov,[40,49] E. Prencipe,[18] M. Prim,[32] M. Ritter,[42] A. Rostomyan,[8] G. Russo,[51] S. Sandilya,[7] L. Santelj,[16] T. Sanuki,[75] V. Savinov,[64] O. Schneider,[39] G. Schnell,[1,19] C. Schwanda,[26] A. J. Schwartz,[7] Y. Seino,[58] K. Senyo,[81] M. E. Sevior,[46] V. Shebalin,[15] J.-G. Shiu,[55] A. Sokolov,[27] E. Solovieva,[40] M. Starič,[31] J. F. Strube,[61] T. Sumiyoshi,[78] M. Takizawa,[68,17,65] K. Tanida,[30] F. Tenchini,[8] T. Uglov,[40,49] S. Uno,[16,12] Y. Usov,[4,59] R. Van Tonder,[32] G. Varner,[15] A. Vinokurova,[4,59] E. Waheed,[46] C. H. Wang,[54] M.-Z. Wang,[55] P. Wang,[25] O. Werbycka,[56] E. Won,[36] S. B. Yang,[36] H. Ye,[8] Y. Yusa,[58] Z. P. Zhang,[66] V. Zhilich,[4,59] V. Zhukova,[40] and V. Zhulanov[4,59]

(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, 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*
[11]*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]*Hanyang 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*
[53]*National Central University, Chung-li 32054*
[54]*National United University, Miao Li 36003*
[55]*Department of Physics, National Taiwan University, Taipei 10617*
[56]*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
[57]*Nippon Dental University, Niigata 951-8580*
[58]*Niigata University, Niigata 950-2181*
[59]*Novosibirsk State University, Novosibirsk 630090*
[60]*Osaka City University, Osaka 558-8585*
[61]*Pacific Northwest National Laboratory, Richland, Washington 99352*
[62]*Panjab University, Chandigarh 160014*
[63]*Peking University, Beijing 100871*
[64]*University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
[65]*Theoretical Research Division, Nishina Center, RIKEN, Saitama 351-0198*
[66]*University of Science and Technology of China, Hefei 230026*
[67]*Seoul National University, Seoul 151-742*
[68]*Showa Pharmaceutical University, Tokyo 194-8543*
[69]*Soongsil University, Seoul 156-743*
[70]*Sungkyunkwan University, Suwon 440-746*
[71]*School of Physics, University of Sydney, New South Wales 2006*
[72]*Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451*
[73]*Tata Institute of Fundamental Research, Mumbai 400005*
[74]*Toho University, Funabashi 274-8510*
[75]*Department of Physics, Tohoku University, Sendai 980-8578*
[76]*Earthquake Research Institute, University of Tokyo, Tokyo 113-0032*
[77]*Department of Physics, University of Tokyo, Tokyo 113-0033*
[78]*Tokyo Metropolitan University, Tokyo 192-0397*
[79]*Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*
[80]*Wayne State University, Detroit, Michigan 48202*
[81]*Yamagata University, Yamagata 990-8560*
[82]*Yonsei University, Seoul 120-749*

(Received 21 June 2021; revised 3 December 2021; accepted 17 December 2021; published 11 January 2022)

We report the first direct measurement of the inclusive branching fraction $\mathcal{B}(B_s \to D_s X)$ via B_s tagging in $e^+e^- \to \Upsilon(5S)$ events. Tagging is accomplished through a partial reconstruction of semileptonic decays $B_s \to D_s X \ell \nu$, where X denotes unreconstructed additional hadrons or photons and ℓ 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 $\mathcal{B}(B_s \to D_s X) = (60.2 \pm 5.8 \pm 2.3)\%$, where the first uncertainty is statistical and the second is systematic.

DOI: 10.1103/PhysRevD.105.012004

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 [1–3]. 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 $\mathcal{B}(\bar{b} \to B_s^0) \cdot \mathcal{B}(B_s^0 \to D_s X)$. The branching fraction $\mathcal{B}(B_s^0 \to D_s X)$ was evaluated using a model-dependent value of $\mathcal{B}(\bar{b} \to B_s^0)$ 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,

$$\mathcal{B}(\Upsilon(5S) \to D_x X)/2 = f_s \cdot \mathcal{B}(B_s \to D_x X) + f_q \cdot \mathcal{B}(B \to D_x X), \quad (1)$$

where D_x 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 $\mathcal{B}(\Upsilon(5S) \to D^0 X)$ [7], 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 branching fraction $\mathcal{B}(B_s \to D_s X)$ 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_S^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 25° and 145° 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 ($\mathcal{L}_e$) 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 $\mathcal{L}_e > 0.8$ ($\mathcal{L}_\mu > 0.8$) are included as electrons (muons) in the analysis. The efficiency for electron (muon) tracks to pass this criterion is $(94.7 \pm 0.2)\%$ ($(96.7 \pm 0.2)\%$).

The neutral intermediate particles ϕ, K_S^0 and K^{*0} [13] are reconstructed from charged tracks. For $\phi \to K^+K^-$ reconstruction, any pair of oppositely charged kaons with invariant mass within 15 MeV$/c^2$ of the ϕ nominal mass [9] is considered to be a ϕ candidate. The K_S^0 candidates are reconstructed via the decay $K_S^0 \to \pi^+\pi^-$, following standard criteria [14], and are further required to have an invariant mass within 20 MeV$/c^2$ (≈4.4σ in resolution) of the nominal mass. For $K^{*0} \to K^+\pi^-$, the candidate tracks are oppositely charged K and π, with invariant mass within 50 MeV$/c^2$.

Candidates for D_s^+ are reconstructed in the final states $\phi\pi^+$, $K_S^0 K^+$, 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 1.92–2.02 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 χ^2 of the fit output is required to be less than 100. Nearly all correctly reconstructed D_s, $(98.1 \pm 0.1)\%$, are found to pass this requirement. The decays $D_s^+ \to \phi(K^+K^-)\pi^+$ and $D_s^+ \to \bar{K}^{*0}(K^-\pi^+)K^+$ 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 $|\cos\theta_{\mathrm{hel}}| > 0.5$, where the helicity angle θ_{hel} is defined as the angle between the momentum of the D_s^+ and K^+ (π^+) in the rest frame of the ϕ ($\bar{K}^{*0}$) resonance.

We tag B_s events through a "partial reconstruction" of the semileptonic decay $B_s^0 \to D_s^- X \ell^+ \nu$, with the D_s^- modes $\phi\pi^-$ and $K_S^0 K^-$, using a procedure similar to one applied at the $\Upsilon(4S)$ resonance [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 $B\bar{B}$ 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

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

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_{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 $M_{\mathrm{miss}}^2 = 0$. The thrust angle, θ_{thrust} is defined as the angle between the thrust axis [16] of the selected $D_s\ell$ system and that of the remaining tracks in the event. To suppress continuum background, we require $|\cos\theta_{\mathrm{thrust}}| < 0.8$. 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 χ^2.

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

(1) Correctly tagged candidates.

(2) Incorrect tag, where a lepton from a B_s semileptonic decay is paired with a real D_s from the other B_s. This can happen if B_s mixing has occurred.

(3) Other incorrect tags: all other sources of candidates. In addition to $B_s^{(*)}\bar{B}_s^{(*)}$ events, sources include $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$, and $B^{(*)}\bar{B}^{(*)}X$ events.

For each component, the M_{miss}^2 distribution is taken to be a histogram obtained via Monte Carlo (MC) simulation. For correctly reconstructed D_s, the distribution in $M_{D_{s\_\mathrm{tag}}}$ is represented 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, each distribution is well-represented by a linear function. Tag decays, $B_s \to D_s X \ell \nu$, are modeled as a sum of $B_s \to D_s \ell \nu$ and $B_s \to D_s^* \ell \nu$; all semileptonic B_s decays to higher excited D_s states observed to date involve DK rather than D_s in the final state, and decays including the states $D_{s0}^*(2317)$ and $D_{s1}(2460)$, which are known to decay to D_s, have not been observed [9]. The presence of higher mass excited D_s in $D_s X \ell \nu$ final states would be manifested as a knee or bump to the right side of the M_{miss}^2 peak. The data are found to be consistent with contributions from D_s and D_s^* only (Fig. 1, top). We find $N_{\mathrm{tag}}^{\phi\pi} = 6473 \pm 119$ and $N_{\mathrm{tag}}^{K_S^0 K} = 4435 \pm 126$. 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 rate of signal D_s in tagged events is determined through a binned 3D maximum-likelihood fit in the tag-side variables, M_{miss}^2 and $M_{D_{s\_\mathrm{tag}}}$, and the invariant mass of the signal-side D_s candidate, $M_{D_{s\_\mathrm{sig}}}$. Each tag + signal candidate corresponds on the tag side to one of the three components comprising the 2D fit and on the signal side with a real or combinatorial D_s. Events containing $B_s \to D_s X \ell \nu$ and inclusive $B_s \to D_s X$ may have a correctly reconstructed tag (component 1) with a signal D_s or an incorrect tag (component 2) with a D_s that is actually from the tag side. We define the first type of event as "signal" and the second as "cross-feed." Both types are included in our fit and used to determine the rate of $B_s \to D_s X$.

For signal events, where the tag-side (signal-side) D_s decays to channel i (j), the raw branching fraction ($\mathcal{B}_{\mathrm{raw}}$) is found by dividing the number observed ($N_{\mathrm{sig};ij}$) by the total number of reconstructed tags in channel i ($N_{\mathrm{tag};i}$), the branching fraction for the channel j ($\mathcal{B}_j$), and the reconstruction efficiency ($\mathcal{E}_{ij;\mathrm{tag}}$) for D_s in channel j:

$$\mathcal{B}_{\mathrm{raw}} = \frac{N_{\mathrm{sig};ij}}{N_{\mathrm{tag};i}\mathcal{B}_j\mathcal{E}_{ij;\mathrm{tag}}}. \qquad (3)$$

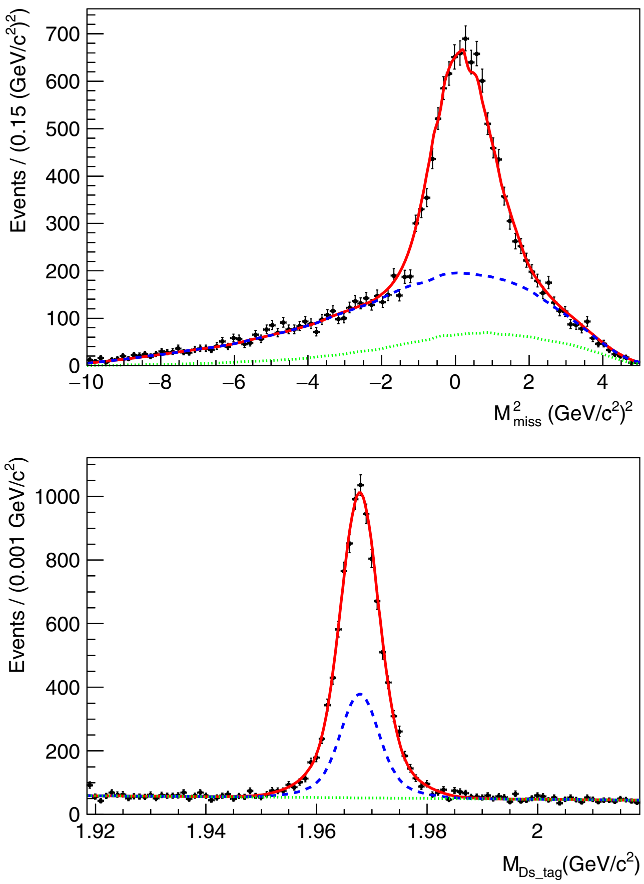

FIG. 1. Distributions in M^2_{miss} (top) and D_s candidate mass (bottom) for tag candidates with $D_s \to \phi\pi$ in data (points with error bars), overlaid with fit results (cumulative): correct tags (red, solid), incorrect tags with real D_s (blue, dashed), and other incorrect tags (green, dotted). In each plot, a signal band requirement is made on the quantity that is not displayed ($m^{PDG}_{D_s} \pm 0.015\ \text{GeV}/c^2$, $|M^2_{\text{miss}}| < 2\ (\text{GeV}/c^2)^2$)).

We evaluate $\mathcal{E}_{ij;\text{tag}}$ via MC for each pair of channels (Table I).

For cross-feed events, the raw branching fraction is obtained through the relationship of their rate to that of signal events. For both signal and cross-feed, the number of events found in a data set depends on many of the same factors: number of B_s events, branching fractions of the reconstructed D_s modes, branching fractions of $B_s \to D_s X \ell\nu$, and $B_s \to D_s X$. The reason for this is clear: the two types have a common origin, differing only in the assigning of D_s to tag- *vs* signal-side. The differences stem from the selection processes and the fact that cross-feed is sourced only from the 50% of events where $B_s \leftrightarrow \bar{B}_s$ mixing has occurred. Thus, the expected ratio, R_{ij} of observed cross-feed to signal events for each pair of D_s decay channels is the ratio of selection efficiencies times 0.5. These ratios are obtained via MC simulation. From the number of observed cross-feed ($N_{\text{cf};ij}$) we then have

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

Tag channel	Signal channel	Efficiency (%)
$\phi\pi$	$\phi\{K^+K^-\}\pi$	26.1 ± 0.5
	$K^0_S\{\pi^+\pi^-\}K$	38.5 ± 0.6
	$K^{*0}\{K^\pm\pi^\mp\}K$	24.6 ± 0.5
$K^0_S K$	$\phi\{K^+K^-\}\pi$	27.6 ± 0.5
	$K^0_S\{\pi^+\pi^-\}K$	37.8 ± 0.6
	$K^{*0}\{K^\pm\pi^\mp\}K$	24.6 ± 0.4

$$\mathcal{B}_{\text{raw}} = \frac{N_{\text{cf};ij}}{N_{\text{tag};i}\mathcal{B}_j\mathcal{E}_{ij;\text{tag}}R_{ij}}. \tag{4}$$

A fit for $\mathcal{B}_{\text{raw}}$ is performed simultaneously for the six D_s tag-signal channel combinations, using the efficiencies and efficiency ratios determined as described above. Intermediate branching fractions are fixed to PDG values.

Our fit yields $\mathcal{B}_{\text{raw}} = (58.2 \pm 5.8)\%$, which corresponds to a fitted total of 101 ± 10 signal and 36 ± 4 cross-feed events. Projections of the fit are shown in Fig. 2. To obtain $\mathcal{B}(B_s \to D_s X)$, we must make a correction to $\mathcal{B}_{\text{raw}}$, due to the fact that the signal mode, $B_s \to D_s X$, is inclusive of the tagging mode, $B_s \to D_s X \ell\nu$. We define $\mathcal{B}(B_s \to D_s X e\nu) + \mathcal{B}(B_s \to D_s X \mu\nu) \equiv \mathcal{B}_{D_s\ell}$, $\mathcal{B}(B_s \to D_s X) \equiv \mathcal{B}_{D_s}$, and the respective reconstruction efficiencies $\epsilon_{D_s\ell}$ and ϵ_{D_s}. We take $N_{B_sB_s}$ to be the number of $B_s^{(*)}\bar{B}_s^{(*)}$ events. The numbers of tags and signal are then

$$\begin{aligned} N_{\text{tag}} &= N_{B_sB_s}(2\epsilon_{D_s\ell}\mathcal{B}_{D_s\ell} - (\epsilon_{D_s\ell}\mathcal{B}_{D_s\ell})^2) \\ &= N_{B_sB_s}\epsilon_{D_s\ell}\mathcal{B}_{D_s\ell}(2 - \epsilon_{D_s\ell}\mathcal{B}_{D_s\ell}), \end{aligned} \tag{5}$$

$$\begin{aligned} N_{\text{sig}} &= N_{B_sB_s}(2\epsilon_{D_s\ell}\mathcal{B}_{D_s\ell}\epsilon_{D_s}\mathcal{B}_{D_s} - (\epsilon_{D_s\ell}\mathcal{B}_{D_s\ell})^2) \\ &= N_{B_sB_s}\epsilon_{D_s\ell}\mathcal{B}_{D_s\ell}(2\epsilon_{D_s}\mathcal{B}_{D_s} - \epsilon_{D_s\ell}\mathcal{B}_{D_s\ell}). \end{aligned} \tag{6}$$

Their ratio, corrected for efficiencies, is $\mathcal{B}_{\text{raw}}$:

$$\begin{aligned} \mathcal{B}_{\text{raw}} &= \frac{N_{\text{sig}}/\epsilon_{D_s}}{N_{\text{tag}}} \\ &= \frac{2\epsilon_{D_s}\mathcal{B}_{D_s} - \epsilon_{D_s\ell}\mathcal{B}_{D_s\ell}}{\epsilon_{D_s}(2 - \epsilon_{D_s\ell}\mathcal{B}_{D_s\ell})} \\ &= \frac{\mathcal{B}_{D_s} - \frac{\epsilon_{D_s\ell}}{2\epsilon_{D_s}}\mathcal{B}_{D_s\ell}}{1 - \epsilon_{D_s\ell}\mathcal{B}_{D_s\ell}/2}. \end{aligned} \tag{7}$$

Thus,

$$\mathcal{B}_{D_s} = \mathcal{B}_{\text{raw}}\left(1 - \frac{\epsilon_{D_s\ell}\mathcal{B}_{D_s\ell}}{2}\right) + \frac{\epsilon_{D_s\ell}}{2\epsilon_{D_s}}\mathcal{B}_{D_s\ell}. \tag{8}$$

To estimate $\epsilon_{D_s\ell}$, we use $\mathcal{B}_{D_s\ell} = (16.2 \pm 2.6(\text{sys}))\%$ [9], $2N_{B_sB_s} = (1.66 \pm 0.27(\text{sys})) \times 10^7$ [17], $N_{\text{tag}} = 10908 \pm 173(\text{stat})$ (our measurement), where errors are indicated

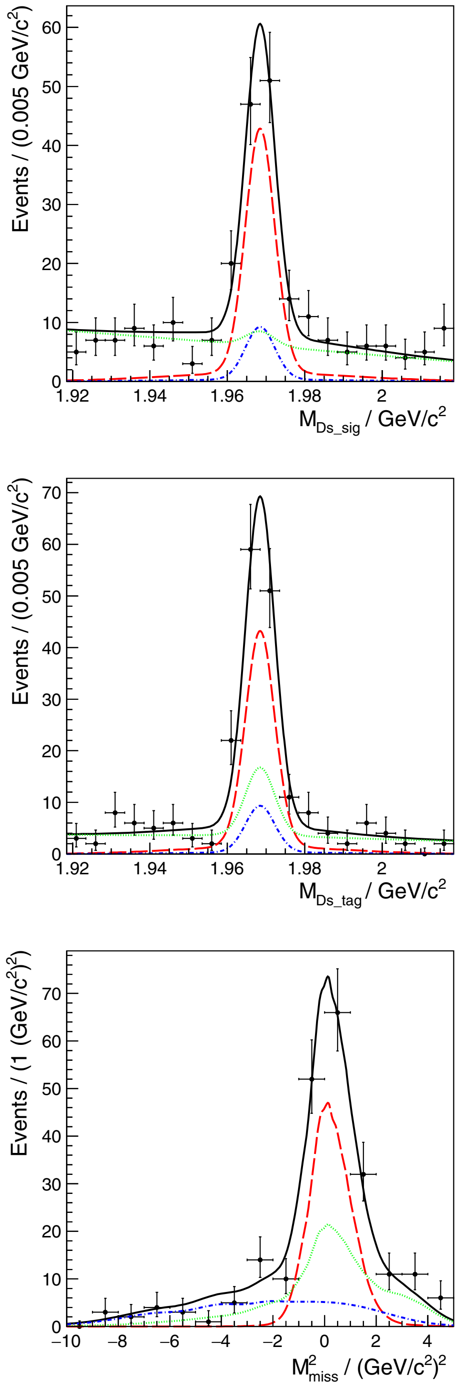

FIG. 2. 1D Projections of results from 3D fits, all D_s modes combined, for $M_{Ds}^{\rm sig}$ (top), $M_{Ds}^{\rm tag}$ (center) and $M^2_{\rm miss}$ (bottom): data (points with error bars), signal (red, dashed), cross-feed (blue, dash-dotted), background (green, dotted), and total (black, solid). For each projected variable, signal band requirements are made in the other two: $M_{Ds}^{\rm sig}, M_{Ds}^{\rm tag} \in m_{Ds}^{\rm PDG} \pm 0.02$ GeV$/c^2$, $M^2_{\rm miss} \in [-2, 2]({\rm GeV}/c^2)^2$.

as being statistical or systematic in origin. We calculate ϵ_{D_s} from Table I and branching fractions from [9]. We find

$$\epsilon_{D_s\ell} \approx \frac{N_{\rm tag}}{2N_{B_sB_s}\mathcal{B}_{D_s\ell}} = (4.1 \pm 0.1({\rm stat}) \pm 0.7({\rm sys})) \times 10^{-3}, \quad (9)$$

$$\epsilon_{D_s} = \sum_i \epsilon_i \mathcal{B}_i = (1.62 \pm 0.03({\rm sys})) \times 10^{-2}. \quad (10)$$

The first correction term is found to be negligible (less than 10^{-3}), and the second is

$$\mathcal{B}_{D_s} - \mathcal{B}_{\rm raw} = \frac{\epsilon_{D_s\ell}}{2\epsilon_{D_s}}\mathcal{B}_{D_s\ell} = (2.03 \pm 0.03({\rm stat}) \pm 0.33({\rm sys})) \times 10^{-2}. \quad (11)$$

We thus find

$$\mathcal{B}_{D_s} = (60.2 \pm 5.8 \pm 0.3)\%. \quad (12)$$

As a cross-check of our method, we fit for signal while floating the cross-feed component and find $\mathcal{B}_{\rm raw} = (64.8 \pm 8.1)\%$, which is consistent with our result.

To confirm the 3D fitting procedure and correction to $\mathcal{B}_{\rm raw}$, we generated ensembles of simulated data distributions with varied signal content. Signal and crossfeed distributions were generated by randomly selecting from our large sample of MC-generated signal events. For background we generated distributions according to those used in the fit, with parameters fixed to the results of the fit to data. Ensembles of 200 experiments were generated for each of ten branching fractions in the range 10%–100%, in 10% increments. Each distribution was fitted according to our procedure. The resulting ensemble mean branching fractions, corrected and plotted against input branching fractions, were fitted to a line. This test was repeated for each of the six D_s mode combinations, as well as the combined set. All showed consistency with a unit slope and no systematic shifts.

Our estimates of systematic uncertainties are summarized in Table II. We evaluate the effects from the considered sources by varying each and taking the resulting shift observed in $\mathcal{B}_{\rm raw}$ 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 $B_s \to D_sX\ell\nu$ on efficiency and on the $M^2_{\rm miss}$ 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[9] and by varying the HQET2 parameters in the MC generator by ±10%. For the "other incorrect tag" (type 3, above), the $M^2_{\rm miss}$ distribution

TABLE II. Systematic uncertainties on $\mathcal{B}_{\mathrm{raw}}$, in %. The total is the sum in quadrature from all sources.

	Channel						
	$\phi\pi$ Tag			K_S^0K Tag			
Source	$\phi\pi$	K_S^0K	$K^{*0}K$	$\phi\pi$	K_S^0K	$K^{*0}K$	Combined
Model, tag		1.5			1.1		1.5
Model, signal	0.1	0.1	0.3	0.1	0.1	0.1	0.3
Model, cross-feed	0.4	0.3	0.3	0.2	0.1	0.1	0.4
M^2_{miss} shape, $M_{B_s^*} - M_{B_s}$		1.2			1.2		1.2
M^2_{miss} background	0.1	0.2	0.1	0.5	0.2	0.3	0.5
$M(D_s)$ signal shape	0.2	0.2	1.2	0.1	0.1	1.0	1.2
$M(D_s)$ background shape	1.0	0.6	< 0.1	< 0.1	0.1	0.1	1.0
Cross-feed efficiency	0.5	0.3	0.6	0.3	0.1	0.3	0.6
Reconstruction efficiency	0.4	0.2	0.4	0.2	0.1	0.2	0.4
Statistics, linearity test	0.2	0.3	0.3	0.3	0.4	0.4	0.4
$B \to D_s^{(*)} K \ell \nu$		< 0.02			0.5		0.5
$\mathcal{B}(D_s \to \phi\pi)$			…				1.2
$\mathcal{B}(D_s \to K_S^0 K)$			…				0.5
$\mathcal{B}(D_s \to K^{*0} K)$			…				1.2
Single tag selection			…				1.0
Tracking			…				1.1
K-π identification			…				1.3
Total			…				3.8

in data from tags with "sideband" D_s candidates, $|M_{\mathrm{cand}} - m_{D_s} \pm 40| < 10$ 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 $\pm 1\sigma$ [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_{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_{miss}. For this analysis, the values are 45.9 MeV/c^2 for MC generation and 47.3 MeV/c^2 for M^2_{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_{miss} is fitted in both the numerator and denominator to obtain $\mathcal{B}_{\mathrm{raw}}$, 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 $\delta E/c^2$ and the value used in MC generation in the range 45.9–49.0 MeV/c^2. Changing the calculation of M^2_{miss} results in a maximum excursion in $\mathcal{B}_{\mathrm{raw}}$ 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 nonstrange B decay $\mathcal{B}(B \to D_s^{(*)} K \ell \nu)$, which is not included in our generic MC generator. We use $\mathcal{B}(B^+ \to D_s^{(*)-} K^+ \ell^+ \nu) = (6.1 \pm 1.0) \times 10^{-4}$ [9], assume that $\mathcal{B}(B^0 \to D_s^{(*)-} K^0 \ell^+ \nu)$ is the same, and multiply by a factor of two to account for both electrons and muons. Taking $\mathcal{B}(\Upsilon(5S) \to B\bar{B}X) = 76\%$, $\mathcal{B}(\Upsilon(5S) \to B_s \bar{B}_s X) = 20\%$, and $\mathcal{B}(B_s \to X \ell \nu) = 9.6\%$ [9], we estimate

$$\frac{\mathcal{B}(\Upsilon(5S) \to B\bar{B}X) \cdot \mathcal{B}(B \to D_s^{(*)} K \ell \nu)}{\mathcal{B}(\Upsilon(5S) \to B_s^{(*)} \bar{B}_s^{(*)}) \cdot \mathcal{B}(B_s \to D_s X \ell \nu)} \approx 0.048. \qquad (13)$$

As the shape in M^2_{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 $B \to D_s^{(*)} K \ell \nu$ in

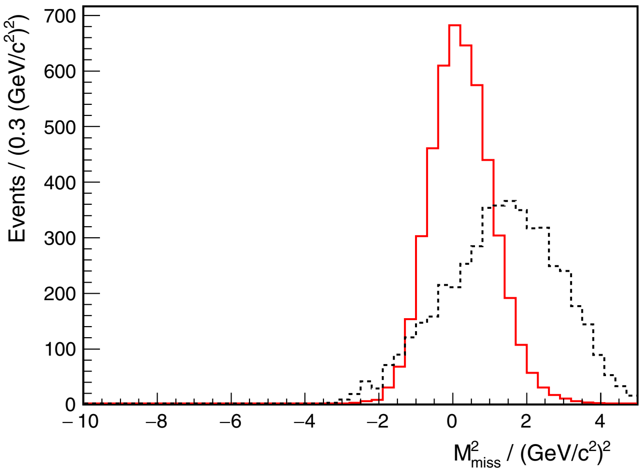

FIG. 3. The distributions in $M^2_{\rm miss}$ for $B_s \to D_s X \ell \nu$ (red) and $B \to D_s K X \ell \nu$ (black), with $D_s \to \phi\pi$.

the D_s tag modes. Figure 3 illustrates the difference. We measure the effect on our MC tag fit of including such events, and estimate a contribution to $\mathcal{B}(B_s \to D_s X \ell \nu)$ of $< 0.02\%$ (0.5%) to the $D_s \to \phi\pi$ ($D_s \to K^0_S K$) 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 $\mathcal{B}_{\rm raw}$ of 3.8%. Adding the systematic uncertainties in quadrature, we find

$$\mathcal{B}(B_s \to D_s X) = (60.2 \pm 5.8 \pm 2.3)\%. \qquad (14)$$

The central value is lower than the theoretical expectation $(86^{+8}_{-13})\%$ [18], and $\approx 1.3\sigma$ below the world average, $(93 \pm 25)\%$ [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 [7,17], 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 $f_{\rm nob} = 4.9^{+5.0}_{-0.6}\%$ [19]. Charm is highly suppressed in these decays, so we take $f_q = 1 - f_s - f_{\rm nob}$. Using $\mathcal{B}(\Upsilon(5S) \to D_s X) = (45.4 \pm 3.0)\%$ [20] and $\mathcal{B}(B \to D_s X) = (8.3 \pm 0.8)\%$ [9], we solve Eq. (1) for f_s and find

$$f_s = 0.285 \pm 0.032(\text{stat}) \pm 0.037(\text{sys}). \qquad (15)$$

This value is larger than the world average, $f_s = 0.201 \pm 0.031$ [9], which is evaluated assuming the model-based estimates $\mathcal{B}(B_s \to D_s X) = (92 \pm 11)\%$ and $\mathcal{B}(B_s \to D^0 X) = (8 \pm 7)\%$ [7]; the impact of introducing $f_{\rm nob}$ to the calculation is minor. Our result uses the same value of $\mathcal{B}(\Upsilon(5S) \to D_s X)$ from which f_s is derived in [17] and thus supersedes the value presented there. It is consistent with a recent Belle measurement of f_s by an independent method [19]. An older Belle measurement of f_s from semileptonic decays [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 $DK(X)$ [22], this value is likely an underestimate, so we do not compare it with the result reported here.

Applying Eq. (1) with $\mathcal{B}(B \to D^0/\bar{D}^0 X) = (61.5 \pm 2.9)\%$ [9], $\mathcal{B}(\Upsilon(5S) \to D^0 X) = (108 \pm 8)\%$ [9], and our result for f_s, we find $\mathcal{B}(B_s \to D^0 X) = (46 \pm 2(\text{stat}) \pm 20(\text{sys}))\%$, where the systematic uncertainties on $\mathcal{B}(\Upsilon(5S) \to D^0 X)$ and $f_{\rm nob}$ 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 $B_s \to D^0 X$ are currently included in the PDG tables [9].

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

$$\mathcal{B}(B_s \to D_s X) = (60.2 \pm 5.8(\text{stat}) \pm 2.3(\text{sys}))\%, \qquad (16)$$

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,

$$f_s = 0.285 \pm 0.032(\text{stat}) \pm 0.037(\text{sys}). \qquad (17)$$

This value supersedes that reported in [17].

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 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 No. DP180102629, No. DP170102389, No. DP170102204, No. DP150103061, No. 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. 18PJ1401000; 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 Grants No. 2016R1D1A1B01010135, No. 2016R1D1A1B02012900, No. 2018R1A2B3003643, No. 2018R1A6A1A06024970, No. 2019K1A3A7A09033840, No. 2019R1I1A3A01058933, No. 2021R1A6A1A03043957, No. 2021R1F1A1060423, No. 2021R1F1A1064008; 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 No. 14.W03.31.0026, and the HSE University Basic Research Program, Moscow; University of Tabuk research grants No. S-1440-0321, No. S-0256-1438, and No. S-0280-1439 (Saudi Arabia); the Slovenian Research Agency Grants No. J1-9124 and No. 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 Phys. Res., Sect. A **479**, 117 (2002).
[11] S. Kurokawa and E. Kikutani, Nucl. Instrum. Methods Phys. Res., Sect. 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] S. Esen *et al.* (Belle Collaboration), Phys. Rev. D **87**, 031101(R) (2013).
[18] M. Suzuki, Phys. Rev. D **31**, 1158 (1985).
[19] R. Mizuk *et al.* (Belle Collaboration), J. High Energy Phys. 06 (2021) 137.
[20] This value is obtained by Belle using 121.4 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).