Semi-inclusive studies of semileptonic B_s decays at Belle

C. Oswald,[3] P. Urquijo,[36] J. Dingfelder,[3] A. Abdesselam,[58] I. Adachi,[14,11] H. Aihara,[62] S. Al Said,[58,27] D. M. Asner,[49] T. Aushev,[38,22] R. Ayad,[58] V. Babu,[59] I. Badhrees,[58,26] A. M. Bakich,[57] V. Bhardwaj,[55] A. Bobrov,[4] G. Bonvicini,[67] A. Bozek,[45] M. Bračko,[34,23] T. E. Browder,[13] D. Červenkov,[5] M.-C. Chang,[8] V. Chekelian,[35] A. Chen,[42] B. G. Cheon,[12] K. Chilikin,[22] K. Cho,[28] V. Chobanova,[35] Y. Choi,[56] D. Cinabro,[67] J. Dalseno,[35,60] Z. Doležal,[5] Z. Drásal,[5] A. Drutskoy,[22,37] D. Dutta,[59] S. Eidelman,[4] H. Farhat,[67] J. E. Fast,[49] T. Ferber,[7] O. Frost,[7] B. G. Fulsom,[49] V. Gaur,[59] N. Gabyshev,[4] S. Ganguly,[67] A. Garmash,[4] D. Getzkow,[9] R. Gillard,[67] R. Glattauer,[19] Y. M. Goh,[12] P. Goldenzweig,[25] B. Golob,[32,23] O. Grzymkowska,[45] T. Hara,[14,11] J. Hasenbusch,[3] K. Hayasaka,[40] H. Hayashii,[41] X. H. He,[50] W.-S. Hou,[44] M. Huschle,[25] H. J. Hyun,[30] T. Iijima,[40,39] A. Ishikawa,[61] R. Itoh,[14,11] Y. Iwasaki,[14] I. Jaegle,[13] T. Julius,[36] K. H. Kang,[30] P. Kapusta,[45] E. Kato,[61] T. Kawasaki,[46] C. Kiesling,[35] D. Y. Kim,[54] J. B. Kim,[29] J. H. Kim,[28] K. T. Kim,[29] M. J. Kim,[30] S. H. Kim,[12] Y. J. Kim,[28] K. Kinoshita,[6] B. R. Ko,[29] P. Kodyš,[5] S. Korpar,[34,23] P. Križan,[32,23] P. Krokovny,[4] T. Kuhr,[25] T. Kumita,[64] Y.-J. Kwon,[69] J. S. Lange,[9] D. H. Lee,[29] I. S. Lee,[12] Y. Li,[66] L. Li Gioi,[35] J. Libby,[16] D. Liventsev,[66] P. Lukin,[4] D. Matvienko,[4] H. Miyata,[46] R. Mizuk,[22,37] G. B. Mohanty,[59] A. Moll,[35,60] H. K. Moon,[29] E. Nakano,[48] M. Nakao,[14,11] H. Nakazawa,[42] T. Nanut,[23] Z. Natkaniec,[45] M. Nayak,[16] S. Nishida,[14,11] T. Nozaki,[14] S. Okuno,[24] P. Pakhlov,[22,37] G. Pakhlova,[38,22] C. W. Park,[56] H. Park,[30] T. K. Pedlar,[33] L. Pesántez,[3] R. Pestotnik,[23] M. Petrič,[23] L. E. Piilonen,[66] C. Pulvermacher,[25] E. Ribežl,[23] M. Ritter,[35] A. Rostomyan,[7] M. Rozanska,[45] S. Ryu,[53] Y. Sakai,[14,11] S. Sandilya,[59] L. Santelj,[14] T. Sanuki,[61] Y. Sato,[39] V. Savinov,[51] O. Schneider,[31] G. Schnell,[1,15] C. Schwanda,[19] D. Semmler,[9] K. Senyo,[68] O. Seon,[39] M. E. Sevior,[36] M. Shapkin,[20] V. Shebalin,[4] C. P. Shen,[2] T.-A. Shibata,[63] J.-G. Shiu,[44] A. Sibidanov,[57] F. Simon,[35,60] Y.-S. Sohn,[69] E. Solovieva,[22] S. Stanič,[47] M. Starič,[23] J. Stypula,[45] M. Sumihama,[10] T. Sumiyoshi,[64] U. Tamponi,[21,65] Y. Teramoto,[48] K. Trabelsi,[14,11] M. Uchida,[63] Y. Unno,[12] S. Uno,[14,11] Y. Usov,[4] C. Van Hulse,[1] P. Vanhoefer,[35] G. Varner,[13] A. Vinokurova,[4] V. Vorobyev,[4] A. Vossen,[17] M. N. Wagner,[9] C. H. Wang,[43] M.-Z. Wang,[44] P. Wang,[18] X. L. Wang,[66] Y. Watanabe,[24] K. M. Williams,[66] E. Won,[29] H. Yamamoto,[61] S. Yashchenko,[7] Y. Yook,[69] Z. P. Zhang,[52] V. Zhilich,[4] V. Zhulanov,[4] and A. Zupanc[23]

(Belle Collaboration)

[1]*University of the Basque Country UPV/EHU, 48080 Bilbao*
[2]*Beihang University, Beijing 100191*
[3]*University of Bonn, 53115 Bonn*
[4]*Budker Institute of Nuclear Physics SB RAS and Novosibirsk State University, Novosibirsk 630090*
[5]*Faculty of Mathematics and Physics, Charles University, 121 16 Prague*
[6]*University of Cincinnati, Cincinnati, Ohio 45221*
[7]*Deutsches Elektronen–Synchrotron, 22607 Hamburg*
[8]*Department of Physics, Fu Jen Catholic University, Taipei 24205*
[9]*Justus-Liebig-Universität Gießen, 35392 Gießen*
[10]*Gifu University, Gifu 501-1193*
[11]*SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193*
[12]*Hanyang University, Seoul 133-791*
[13]*University of Hawaii, Honolulu, Hawaii 96822*
[14]*High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801*
[15]*IKERBASQUE, Basque Foundation for Science, 48013 Bilbao*
[16]*Indian Institute of Technology Madras, Chennai 600036*
[17]*Indiana University, Bloomington, Indiana 47408*
[18]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049*
[19]*Institute of High Energy Physics, Vienna 1050*
[20]*Institute for High Energy Physics, Protvino 142281*
[21]*INFN—Sezione di Torino, 10125 Torino*
[22]*Institute for Theoretical and Experimental Physics, Moscow 117218*
[23]*J. Stefan Institute, 1000 Ljubljana*
[24]*Kanagawa University, Yokohama 221-8686*
[25]*Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe*
[26]*King Abdulaziz City for Science and Technology, Riyadh 11442*
[27]*Department of Physics, Faculty of Science, King Abdulaziz University, Jeddah 21589*
[28]*Korea Institute of Science and Technology Information, Daejeon 305-806*
[29]*Korea University, Seoul 136-713*
[30]*Kyungpook National University, Daegu 702-701*
[31]*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015*

[32]*Faculty of Mathematics and Physics, University of Ljubljana, 1000 Ljubljana*
[33]*Luther College, Decorah, Iowa 52101*
[34]*University of Maribor, 2000 Maribor*
[35]*Max-Planck-Institut für Physik, 80805 München*
[36]*School of Physics, University of Melbourne, Victoria 3010*
[37]*Moscow Physical Engineering Institute, Moscow 115409*
[38]*Moscow Institute of Physics and Technology, Moscow Region 141700*
[39]*Graduate School of Science, Nagoya University, Nagoya 464-8602*
[40]*Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602*
[41]*Nara Women's University, Nara 630-8506*
[42]*National Central University, Chung-li 32054*
[43]*National United University, Miao Li 36003*
[44]*Department of Physics, National Taiwan University, Taipei 10617*
[45]*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
[46]*Niigata University, Niigata 950-2181*
[47]*University of Nova Gorica, 5000 Nova Gorica*
[48]*Osaka City University, Osaka 558-8585*
[49]*Pacific Northwest National Laboratory, Richland, Washington 99352*
[50]*Peking University, Beijing 100871*
[51]*University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
[52]*University of Science and Technology of China, Hefei 230026*
[53]*Seoul National University, Seoul 151-742*
[54]*Soongsil University, Seoul 156-743*
[55]*University of South Carolina, Columbia, South Carolina 29208*
[56]*Sungkyunkwan University, Suwon 440-746*
[57]*School of Physics, University of Sydney, NSW 2006*
[58]*Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451*
[59]*Tata Institute of Fundamental Research, Mumbai 400005*
[60]*Excellence Cluster Universe, Technische Universität München, 85748 Garching*
[61]*Tohoku University, Sendai 980-8578*
[62]*Department of Physics, University of Tokyo, Tokyo 113-0033*
[63]*Tokyo Institute of Technology, Tokyo 152-8550*
[64]*Tokyo Metropolitan University, Tokyo 192-0397*
[65]*University of Torino, 10124 Torino*
[66]*CNP, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*
[67]*Wayne State University, Detroit, Michigan 48202*
[68]*Yamagata University, Yamagata 990-8560*
[69]*Yonsei University, Seoul 120-749*

(Received 8 April 2015; published 22 October 2015)

We present an analysis of the semi-inclusive decays $B_s \to D_s^- X \ell^+ \nu$ and $B_s \to D_s^{*-} X \ell^+ \nu$, where X denotes a final state that may consist of additional hadrons or photons and ℓ is an electron or muon. The studied B_s decays are contained in the 121.4 fb^{-1} $\Upsilon(5S)$ data sample collected by the Belle detector at the KEKB asymmetric-energy e^+e^- collider. The branching fractions of the decays are measured to be $\mathcal{B}(B_s \to D_s^- X \ell^+ \nu) = [8.2 \pm 0.2(\text{stat}) \pm 0.6(\text{syst}) \pm 1.4(\text{ext})]\%$ and $\mathcal{B}(B_s \to D_s^{*-} X \ell^+ \nu) = [5.4 \pm 0.4(\text{stat}) \pm 0.4(\text{syst}) \pm 0.9(\text{ext})]\%$, where the first two uncertainties are statistical and systematic and the last is due to external parameters. The measurement also provides an estimate of the $B_s^{(*)}\bar{B}_s^{(*)}$ production cross section, $\sigma(e^+e^- \to B_s^{(*)}\bar{B}_s^{(*)}) = [53.8 \pm 1.4(\text{stat}) \pm 4.0(\text{syst}) \pm 3.4(\text{ext})]$ pb, at the center-of-mass energy $\sqrt{s} = 10.86$ GeV.

DOI: 10.1103/PhysRevD.92.072013 PACS numbers: 14.40.Nd, 13.20.He

I. INTRODUCTION

Analyses of semileptonic decays $B \to X_c \ell \nu$, where X_c denotes a hadronic final state with a charm quark, play an important role in the determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$. The extraction of $|V_{cb}|$ from the measured decay rates relies on form factors that describe the accompanying strong interaction processes. Measurements of semileptonic B_s decays provide complementary information to test and validate the QCD calculations of these form factors. Since large B_s samples have become available at Belle and the

experiments at the Large Hadron Collider, the interest in the topic of semileptonic B_s decays has intensified recently. Theoretical predictions of form factors and branching fractions are based on QCD sum rules [1–4], lattice QCD [5,6] and constituent quark models [7–15]. The predicted exclusive branching fractions vary from 1.0% to 3.2% for $B_s \to D_s \ell\nu$ decays and from 4.3% to 7.6% for $B_s \to D_s^* \ell\nu$ decays. There are also predictions for the modes with higher excitations of the D_s meson, denoted hereinafter by "D_s^{**}". The LHCb and D0 experiments have measured the semi-inclusive branching fractions of the decays $B_s \to D_{s1}(2536)X\mu^+\nu$ and $B_s \to D_{s2}^*(2573)X\mu^+\nu$, where the D_s^{**} mesons were reconstructed in $D^{(*)}K$ final states [16,17]. The inclusive semileptonic branching fraction of $B_s \to X\ell\nu$ decays was recently measured by Belle and *BABAR* [18,19] and found to be in agreement with the expectations from SU(3) flavor symmetry [20,21]. We report here the first measurements of the semi-inclusive branching fractions $\mathcal{B}(B_s \to D_s X\ell\nu)$ and $\mathcal{B}(B_s \to D_s^* X\ell\nu)$ using the Belle $\Upsilon(5S)$ data set. The number of $B_s^{(*)}\bar{B}_s^{(*)}$ pairs in the data set,

$$N_{B_s\bar{B}_s} = \sigma(e^+e^- \to B_s^{(*)}\bar{B}_s^{(*)}) \cdot \mathcal{L}_{\Upsilon(5S)}, \qquad (1)$$

where $\sigma(e^+e^- \to B_s^{(*)}\bar{B}_s^{(*)})$ is the production cross section and $\mathcal{L}_{\Upsilon(5S)}$ is the integrated luminosity, is the limiting systematic uncertainty in this measurement and other untagged B_s measurements at Belle [22]. The value $N_{B_s\bar{B}_s} = (7.1 \pm 1.3) \times 10^6$ was obtained from a measurement of the inclusive D_s yield in the data set [23]. The measured $B_s \to D_s X\ell\nu$ yield, together with an estimate for the branching fraction $\mathcal{B}(B_s \to D_s X\ell\nu)$, provides an alternative way to determine $N_{B_s\bar{B}_s}$. A similar approach was already pursued by the LEP experiments [24–26] and LHCb [27].

II. DETECTOR, DATA SAMPLE AND SIMULATION

The Belle detector located at the KEKB asymmetric-energy e^+e^- collider [28] is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector, a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) comprised of CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux-return located outside of the coil is instrumented for K_L^0 mesons and muon detection (KLM). The detector is described in detail elsewhere [29].

This analysis uses a data set with an integrated luminosity of $\mathcal{L}_{\Upsilon(5S)} = (121.4 \pm 0.8)$ fb^{-1} collected at a center-of-mass (CM) energy of $\sqrt{s} = 10.86$ GeV [30], corresponding to the mass of the $\Upsilon(5S)$ resonance. The B_s mesons are produced in pairs in the following production modes, with the respective production fractions given in parentheses: $B_s^*\bar{B}_s^*$ $[(87.8 \pm 1.5)\%]$, $B_s^*\bar{B}_s$ $[(6.7 \pm 1.2)\%]$ and $B_s\bar{B}_s$ $[(2.6 \pm 2.6)\%]$ [31]. All production modes are considered for the analysis. Moreover, we use a 62.8 fb^{-1} sample collected below the production threshold for open B production to study the continuum processes $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$).

A sample of simulated events with a size corresponding to six times the integrated data luminosity is generated using Monte Carlo (MC) techniques. The simulated data emulate the different types of events produced at the $\Upsilon(5S)$ CM energy, comprising events with B and B_s decays, bottomonium production and the $q\bar{q}$ continuum processes. The events are generated with the EvtGen package [32] and are processed through a GEANT [33] based detector simulation. Final state photon radiation is added with the PHOTOS package [34].

The branching fractions in the simulation are set to the latest averages from the Particle Data Group [31]. However, for semileptonic B_s decays only measurements of the $D_{s1}(2536)$ and $D_{s2}^*(2573)$ modes are available, so we use instead the calculations from Faustov and Galkin [8], who predict the full set of branching fractions and thus provide a self-consistent picture of the semileptonic width. The B_s semileptonic decay modes considered in this analysis, with their corresponding branching fractions given in parentheses, are: D_s (2.1%), D_s^* (5.3%), $D_{s1}(2536)$ (0.84%), $D_{s0}^*(2317)$ (0.36%), $D_{s1}(2460)$ (0.19%) and $D_{s2}^*(2573)$ (0.67%). The decays $B_s^0 \to D_s^{(*)}\ell\nu$ are described by the Caprini-Lellouch-Neubert model [35], based on heavy quark effective theory [36]. Assuming SU(3) flavor symmetry, the form factors of the semileptonic B_s decays are taken to be identical to the ones measured in the corresponding B decays [37]; we use the following values of the form factor parameters: $\rho_D = 1.186$ for $B_s \to D_s\ell\nu$ decays, and $\rho = 1.207$, $R_1 = 1.403$, $R_2 = 0.854$ for $B_s \to D_s^*\ell\nu$ decays. The $B_s^0 \to D_s^{**}\ell\nu$ decays are described by the Leibovich-Ligeti-Stewart-Wise (LLSW) model [38] originally developed for $B \to D^{**}\ell\nu$ decays. We replace in this model the B and D^{**} masses by the B_s and D_s^{**} masses, respectively. The nominal branching fractions for the D_s^{**} decays in this analysis are listed in Table I.

III. ANALYSIS OVERVIEW

This analysis is based on samples of reconstructed $D_s^-\ell^+$ and $D_s^{*-}\ell^+$ pairs [39]. Incorrectly reconstructed D_s and D_s^* candidates constitute a large background in the analysis. We therefore perform fits to the $D_s^{(*)}$ mass distributions to determine the yields of events with correctly reconstructed $D_s^{(*)}$ mesons. These events contain the following signal and background categories:

(1) $e^+e^- \to c\bar{c}$ continuum;
(2) $B \to D_s^{(*)}K\ell\nu$ decays, which have a branching fraction of $(6.1 \pm 1.0) \times 10^{-4}$ [31,40,41];

TABLE I. Nominal branching fractions of D_s^{**} decays to different final states in the MC simulation. The branching fraction of the $D_s^* \to D_s X$ decays is set to 100% and this crossfeed is included in the calculation of the branching fractions to the $D_s X$ final state.

	Branching fraction [%]		
Y	$Y \to D_s X$	$Y \to D_s^* X$	$Y \to D^{(*)} K$
$D_{s0}^*(2317)$	100	63	0
$D_{s1}(2460)$	100	3	0
$D_{s1}(2536)$	0	0	100
$D_{s2}^*(2573)$	0	0	100

(3) *opposite-$B_{(s)}$* background, where a lepton candidate is combined with a $D_s^{(*)}$ meson from the second B_s in the event; the lepton candidate can be either a primary lepton from a $B_{(s)} \to X\ell\nu$ decay, a lepton originating from a secondary decay or a misidentified hadron track;
(4) *same-$B_{(s)}$* background from secondary leptons and from hadron tracks misidentified as leptons, which stem from the decay of the same B_s meson as the reconstructed $D_s^{(*)}$ meson;
(5) *signal*: in the $D_s X\ell\nu$ channel, the signal comprises $B_s \to D_s\ell\nu$ decays and crossfeed from $B_s \to D_s^*\ell\nu$ and $B_s \to D_s^{**}\ell\nu$ decays; in the $D_s^* X\ell\nu$ channel, the dominant signal contributions are $B_s \to D_s^*\ell\nu$ decays with a small crossfeed contribution from $B_s \to D_s^{**}\ell\nu$ decays.

The continuum background is estimated using off-resonance data, and the $B \to D_s^{(*)} K\ell\nu$ background is estimated from MC simulation. We use the kinematic properties of the reconstructed decay to determine the normalizations of the other three components from data. For this, we consider the lepton momentum in the CM system of the e^+e^- collision, p_ℓ^*, and the variable

$$X_{\rm mis} = \frac{E_{B_s}^* - (E_{D_s\ell}^* + p_{D_s\ell}^*)}{p_{B_s}^*}, \tag{2}$$

where $E_{B_s}^*$ is the energy of the B_s meson in the CM system approximated by $\sqrt{s}/2$; $p_{B_s}^*$ is the momentum of the B_s meson in the CM system approximated by $\sqrt{s/4 - m_{B_s}^2}$; $E_{D_s\ell}^* = E_\ell^* + E_{D_s}^*$ is the sum of the reconstructed energies in the CM system and $p_{D_s\ell}^* = |\vec{p}_\ell^{\,*} + \vec{p}_{D_s}^{\,*}|$ is the absolute value of the sum of the reconstructed lepton and D_s momenta in the CM system. When the D_s meson and the lepton candidate stem from the decay of the same B_s meson, $X_{\rm mis}$ takes values larger than -1 because the momentum of the unreconstructed B_s decay products, $p_{\rm other}^*$, is constrained by the triangle inequality $p_{B_s}^* - p_{D_s\ell}^* \le p_{\rm other}^*$ and by $p_{\rm other}^* \le E_{\rm other}^* = E_{B_s}^* - E_{D_s\ell}^*$. We divide the data samples into three regions:

(A) $X_{\rm mis} < -1$,
(B) $X_{\rm mis} \ge -1$ and $p_\ell^* < 1.4$ GeV,
(C) $X_{\rm mis} \ge -1$ and $p_\ell^* \ge 1.4$ GeV.

As these regions are later used to determine the signal yields, we refer to them as "counting regions" in the following. Region A contains only *opposite-$B_{(s)}$* background and can be used to determine the normalization of this background. The normalization of the other two components can be extracted from the measured yields in regions B and C, which have an enhanced fraction of *same-$B_{(s)}$* background and *signal* events, respectively. The boundary $p_\ell^* = 1.4$ GeV is chosen to achieve approximately equal event yields in regions B and C. The analysis is insensitive to the modeling of the $X_{\rm mis}$ distribution for signal decays, which depends on the mass of the B_s^* meson, $m_{B_s^*}$, and is thus strongly influenced by the poor precision on $m_{B_s^*}$. The semi-inclusive branching fractions are obtained from the relation

$$\mathcal{B}(B_s \to D_s^{(*)} X\ell\nu) = \frac{N_{\rm sig}}{2 \cdot N_{B_s\bar{B}_s} \epsilon \mathcal{B}_{D_s^{(*)}}}, \tag{3}$$

where $N_{\rm sig}$ is the measured signal yield, ϵ is the average signal efficiency and $\mathcal{B}_{D_s^{(*)}}$ is the branching fraction of the reconstructed $D_s^{(*)}$ decay mode:

$$\mathcal{B}_{D_s} = \mathcal{B}(D_s^- \to \phi\pi^-; \phi \to K^+K^-), \tag{4}$$

$$\mathcal{B}_{D_s^*} = \mathcal{B}(D_s^{*-} \to D_s^-\gamma) \cdot \mathcal{B}_{D_s}. \tag{5}$$

IV. EVENT SELECTION

We select tracks originating from the interaction region by requiring $|dz| < 2.0$ cm and $dr < 0.5$ cm, where dz and dr are the impact parameters along the e^+ beam and in the transverse plane, respectively. Kaon or pion hypotheses are assigned to the tracks based on a likelihood combining the 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. The kaon (pion) identification efficiency for tracks with a typical momentum of 0.75 GeV is about 96% (92%), while the rate of pions (kaons) being misidentified as kaons (pions) is 7% (2%). The kaon and pion candidates are used to reconstruct D_s mesons in the high-purity decay channel $D_s^- \to \phi\pi^-$; $\phi \to K^+K^-$. A D_s candidate is retained in the analysis if it has a reconstructed mass, $M_{KK\pi}$, within a ± 65 MeV window around the nominal D_s mass, $m_{D_s} = 1968.5$ MeV [31], that includes large enough sidebands to determine the combinatorial background of random $KK\pi$ combinations. The reconstructed di-kaon invariant mass, M_{KK}, is required to be in the mass window between 1004 and 1034 MeV,

corresponding to three times the FWHM of the reconstructed ϕ mass peak. To suppress combinatorial background, we impose the criterion $|\cos\theta_{\rm hel}| > 0.3$ on the helicity angle, defined as the angle between the momentum of the D_s and the K^- in the rest frame of the ϕ resonance.

The D_s candidates with a reconstructed mass, $M_{KK\pi}$, within the range between 1953.5 and 1983.5 MeV, corresponding to three times the RMS of the D_s mass peak, are utilized for the reconstruction of D_s^* candidates in the dominant decay channel $D_s^* \to D_s\gamma$. Photon candidates are reconstructed from ECL clusters that are not attributed to a track candidate. The photon candidate must have a minimum energy of 125 MeV in the lab frame and the ratio of the energy deposit in the central 3×3 cells of the ECL cluster to the energy deposit in the central 5×5 cells must be at least 90%. To veto photons from π^0 decays, we combine the photon candidate with any other photon candidate in the detector and require that the invariant mass of the two photons differs from the nominal π^0 mass [31] by more than 5 MeV, corresponding to about 0.8 times the experimental resolution. The angle between the D_s meson and the photon in the lab frame is typically less than 90°, so only candidates fulfilling this requirement are retained. The D_s^* candidates whose mass difference between the reconstructed D_s^* and D_s candidates, $\Delta M = M_{KK\pi\gamma} - M_{KK\pi}$, lies between 78.8 and 208.8 MeV are retained.

Electron and muon candidates are reconstructed from tracks that are not used for the $D_s^{(*)}$ reconstruction. Electrons are selected based on the position matching between the track and the ECL cluster, the ratio of the energy measured in the ECL to the charged track momentum, the transverse ECL shower shape, specific ionization in the CDC and the ACC light yield. Muons are identified using their penetration depth and the transverse scattering in the KLM. Hadron tracks misidentified as leptons and leptons from secondary decays tend to have lower momenta than primary leptons and are suppressed by rejecting lepton candidates with a momentum in the lab frame below 900 MeV. The electron (muon) identification efficiency in the selected momentum region is better than 89% (82%) and the probability that a charged pion or kaon track is misidentified as an electron (muon) is below 1% (2%). Leptons, ℓ^+, from $J/\psi \to \ell^+\ell^-$ decays are vetoed by requiring $|M_{\ell^+h^-} - m_{J/\psi}| < 5$ MeV, where $M_{\ell^+h^-}$ is the invariant mass of the lepton and any accepted track of the opposite charge, h^-, to which we assign the ℓ mass hypothesis. Furthermore, electrons are rejected if they are likely to stem from photon conversions, $|M_{\ell^+h^-}| <$ 100 MeV, or Dalitz π^0 decays, $|M_{\ell^+h^-\gamma} - m_{\pi^0}| < 32$ MeV.

We form a signal B_s^0 candidate by pairing a $D_s^{(*)-}$ candidate with an oppositely charged lepton candidate ℓ^+. To suppress background from $c\bar{c}$ continuum, we reject events where the normalized D_s momentum, $x(D_s) = p^*(D_s)/\sqrt{s/4 - m_{D_s}^2}$, is larger than 0.5 (for explanations, see Ref. [18]). Further suppression of the $c\bar{c}$ continuum background is achieved by rejecting events with a jetlike topology characterized by $|\cos\theta_{\rm thrust}| > 0.8$, where $\theta_{\rm thrust}$ is the thrust angle defined by the two thrust axes maximizing the projection of the momenta of the tracks and photon candidates of the B_s^0 candidate and the rest of the event, respectively.

After applying the selection criteria, 7.9% (0.4%) of the events contain more than one (two) $D_s^+\ell^-$ candidate(s). We perform a χ^2 fit to the vertex of the three tracks used for D_s reconstruction and select the candidate with the best goodness-of-fit. This approach selects a correct candidate in 80% of the cases. The selected D_s candidate in an event is used for D_s^* reconstruction; in 36.2% (9.7%) of the events, more than one (two) $D_s\gamma$ combinations meet the D_s^* requirements. We choose the photon candidate with the highest energy fraction deposited in the central 3×3 cells of a 5×5 cell ECL cluster. In the case that more than one photon candidate deposits all of its energy in the central 3×3 cells of the cluster, the candidate with the higher energy in the lab frame is selected. If two or more lepton candidates pass all of these selection criteria (2.1% of all events), we choose a random lepton candidate.

V. FIT RESULTS

A. D_s fits

We determine the yields of correctly reconstructed D_s mesons with binned extended maximum likelihood fits to the reconstructed D_s mass, $M = M_{KK\pi}$, in 50 equal bins, indexed by j. The probability density function (PDF) of correctly reconstructed D_s mesons, $\mathcal{P}_{\rm sig}(M)$, is modeled by the sum of two Gaussian functions with a common mean. The PDF of the combinatorial background, $\mathcal{P}_{\rm bkg}(M)$, is a first-order Chebychev polynomial. We do not determine the shape parameters from simulation, but rather allow them to vary as free parameters in the fit. The D_s mass fits are performed simultaneously in the three counting regions ($i = A, B, C$) defined above. The width of the first Gaussian function, σ_1, the ratio of the widths of the two Gaussian functions, r_σ, and the ratio of the normalizations of the two Gaussian functions, r_N, are common fit parameters in all three regions. The means of the Gaussian functions, μ_i, and the slopes of the polynomials describing the background, b_i, are fitted in each counting region individually. The likelihood function is

$$L(\nu^{\rm sig}, \nu^{\rm bkg}, \theta) = \prod_{i=A,B,C} \frac{\nu_i^{n_i}}{n_i!} e^{-\nu_i} \prod_{\text{bins j}} \frac{\nu_{ij}^{n_{ij}}}{n_{ij}!} e^{-\nu_{ij}}, \qquad (6)$$

where $\nu^{\rm sig} = (\nu_A^{\rm sig}, \nu_B^{\rm sig}, \nu_C^{\rm sig})$ and $\nu^{\rm bkg} = (\nu_A^{\rm bkg}, \nu_B^{\rm bkg}, \nu_C^{\rm bkg})$ is the vector of signal and background yields in the three counting regions, $\theta = (\sigma_1, r_\sigma, r_N, \mu_A, \mu_B, \mu_C, b_A, b_B, b_C)$ are the shape parameters for the signal and background

PDFs, and n_{ij} and ν_{ij} are the observed and expected event yields in bin j of counting region i, respectively, with $n_i = \sum_j n_{ij}$ and $\nu_i = \sum_j \nu_{ij}$. The expected event yield, ν_{ij}, is a function of $\nu_i^{\rm sig}$,$\nu_i^{\rm bkg}$ and θ:

$$\nu_{ij} = \int_{M_{j,\min}}^{M_{j,\max}} [\nu_i^{\rm sig}\mathcal{P}_{\rm sig}(M) + \nu_i^{\rm bkg}\mathcal{P}_{\rm bkg}(M)]{\rm d}M. \quad (7)$$

Figures 1(a) and 1(b) show the $KK\pi$ mass distributions together with the fit results.

B. D_s^* fits

The D_s^* yields are determined from binned extended maximum likelihood fits to the mass difference ΔM in 25 equal bins, indexed by j. The combinatorial background is modeled by a third-order Chebyshev polynomial, $\mathcal{P}_{\rm bkg}(\Delta M)$, whose parameters are constrained to the values obtained from fits to simulated background distributions. Since the background shapes vary for the different counting regions, the shape parameters are determined for each counting region separately. The signal peak is modeled by the sum of a Gaussian function and a Crystal Ball function [42] to account for energy loss due to material in front of the calorimeter:

$$\mathcal{P}_{\rm sig}(\Delta M) \propto r_N \exp\left(-\frac{(\Delta M-\mu)^2}{(r_\sigma\cdot\sigma)^2}\right) + \begin{cases} \exp\left(-\frac{(\Delta M-\mu)^2}{2\sigma^2}\right) & \text{if } \frac{\Delta M-\mu}{\sigma} > -\alpha \\ \frac{(\frac{\beta}{\alpha})^\beta\cdot {\rm e}^{-\alpha^2/2}}{(\frac{\beta}{\alpha}-\alpha-\frac{\Delta M-\mu}{\sigma})^\beta} & \text{if } \frac{\Delta M-\mu}{\sigma} \le -\alpha. \end{cases}$$

A common mean, μ, is used for both the Gaussian and the Crystal Ball functions. We perform a fit to the simulated signal distribution and fix the parameters r_N, r_σ, α and n at the obtained values. The width σ and the mean of the signal peak μ are varied in the fit to data; the parameter σ is fitted simultaneously in all counting regions while μ is fitted individually for each counting region. The likelihood function is constructed analogous to Eqs. (6) and (7) with additional factors, to implement the constraints of the background PDF parameters taking into account their correlations. The results of the ΔM fits in the different counting regions are presented in Figs. 1(c) and 1(d).

C. Background subtraction

To estimate the $c\bar{c}$ continuum background, the D_s and D_s^* yields are measured in $D_s^-\ell^+$ and $D_s^{*-}\ell^+$ samples reconstructed in the off-resonance data. Since the size of the off-resonance data sample is not sufficient to determine the shape parameters in the fits, they are fixed to the values obtained in the fits to $\Upsilon(5S)$ data in the corresponding counting region. The CM energy, $\sqrt{s}$, in the expression for the $X_{\rm mis}$ variable in Eq. (2) is replaced by a constant value of 10.876 GeV because, otherwise, the denominator would not be defined. The $c\bar{c}$ continuum yields from the fits to off-resonance data are multiplied by the scale factor $S = (\mathcal{L}_{\Upsilon(5S)}/s_{\Upsilon(5S)})/(\mathcal{L}_{\rm off}/s_{\rm off}) = 1.81 \pm 0.02$ to account for the differences in integrated luminosities, $\mathcal{L}$, and the $1/s$ dependence of the $e^+e^- \to c\bar{c}$ cross section. Additionally, a shape correction for differences of the yields in the counting regions between off-resonance and $\Upsilon(5S)$ data is determined from MC simulation and applied. The small background from $B \to D_s^{(*)}K\ell\nu$ decays is estimated from MC simulation using a simple phase space model. The backgrounds from continuum processes and $B \to D_s^{(*)}K\ell\nu$ decays are subtracted in each counting region from the yields measured in $\Upsilon(5S)$ data.

D. Signal extraction

After subtraction of the continuum and the $B \to D_s^{(*)}K\ell\nu$ background components, the remaining yields contain three contributions: *opposite-*$B_{(s)}$ background, *same-*$B_{(s)}$ backgrounds and *signal*. The three contributions are constrained by the event yields in the three counting regions. We introduce a scale factor, a_j, for each contribution, j. The determination of the scale factors is equivalent to solving a system of three linear equations with three unknowns. In order to obtain the uncertainties on the scale factors, we minimize:

$$\chi^2 = \sum_{i=A,B,C} \frac{(D_i - \sum_j a_j N_{i,j})^2}{(\Delta D_i)^2 + \sum_j (a_j \Delta N_{i,j})^2}, \quad (8)$$

where the index i runs over the three counting regions, D_i is the event yield determined by the fits to the $M_{KK\pi}$ or ΔM distributions in data, and ΔD_i is the statistical uncertainty of these fits, $N_{i,j}$ is the MC prediction for the contribution j, and $\Delta N_{i,j}$ is its statistical uncertainty. Table II lists the scale factors, a_j, obtained from the χ^2 minimization and the signal yields,

$$N_{\rm sig} = a_{\rm sig} \cdot \sum_i N_{i,{\rm sig}}. \quad (9)$$

Figure 2 shows the $X_{\rm mis}$ and $p^*(\ell)$ distributions in the three counting regions after applying the scale factors a_j.

VI. SYSTEMATIC UNCERTAINTIES

The different sources of systematic uncertainties on the measured signal yields are described below. They comprise detector effects and the modeling of the signal and backgrounds. An overview can be found in Table V.

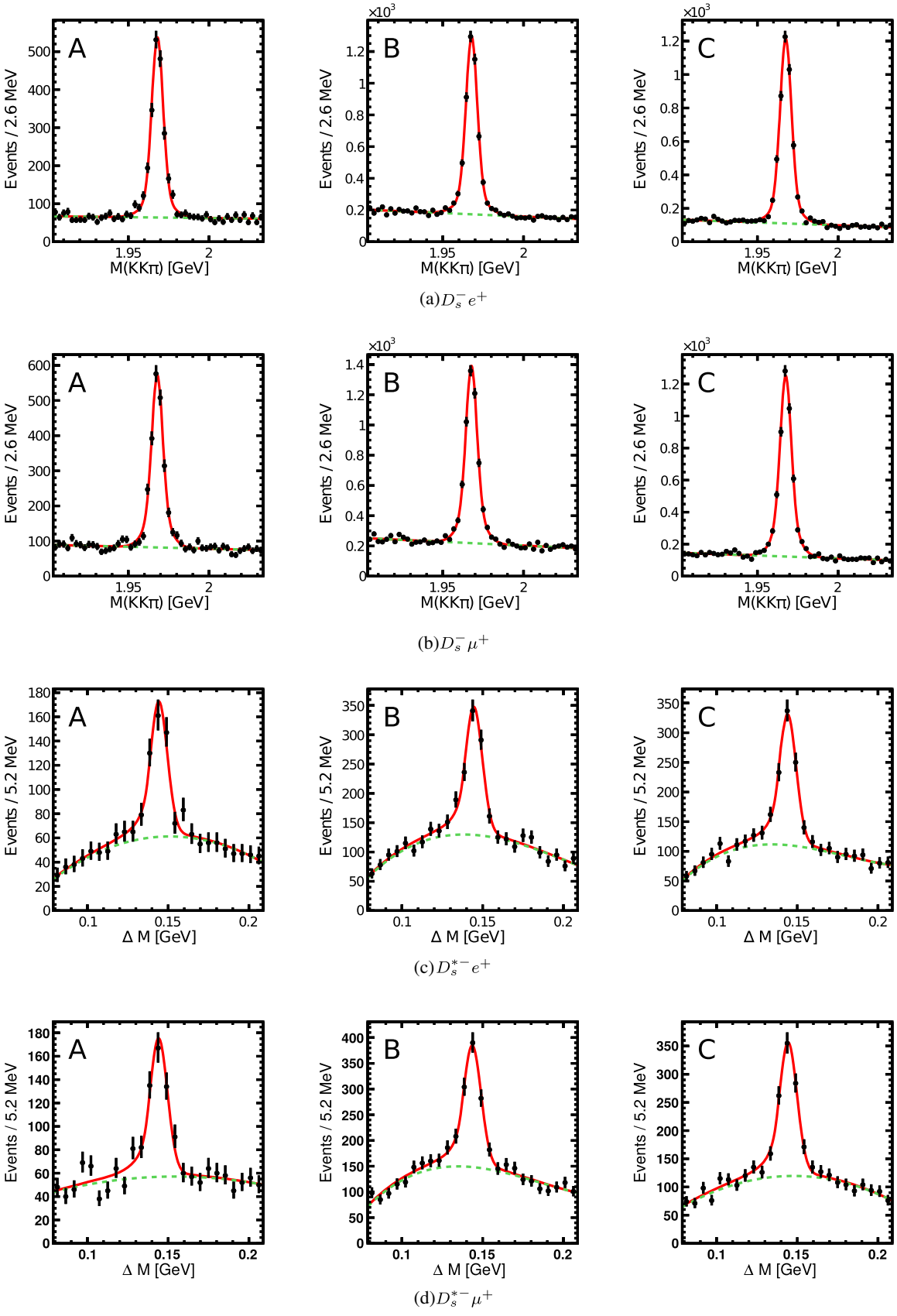

FIG. 1 (color online). The $M_{KK\pi}$ distributions for $D_s^- \ell^+$ events and ΔM distributions for $D_s^{*-} \ell^+$ events reconstructed in the $\Upsilon(5S)$ data for the three counting regions. The black points with uncertainty bars are the data, the red solid curve represents the total fit result, and the green dashed line is the fitted background component.

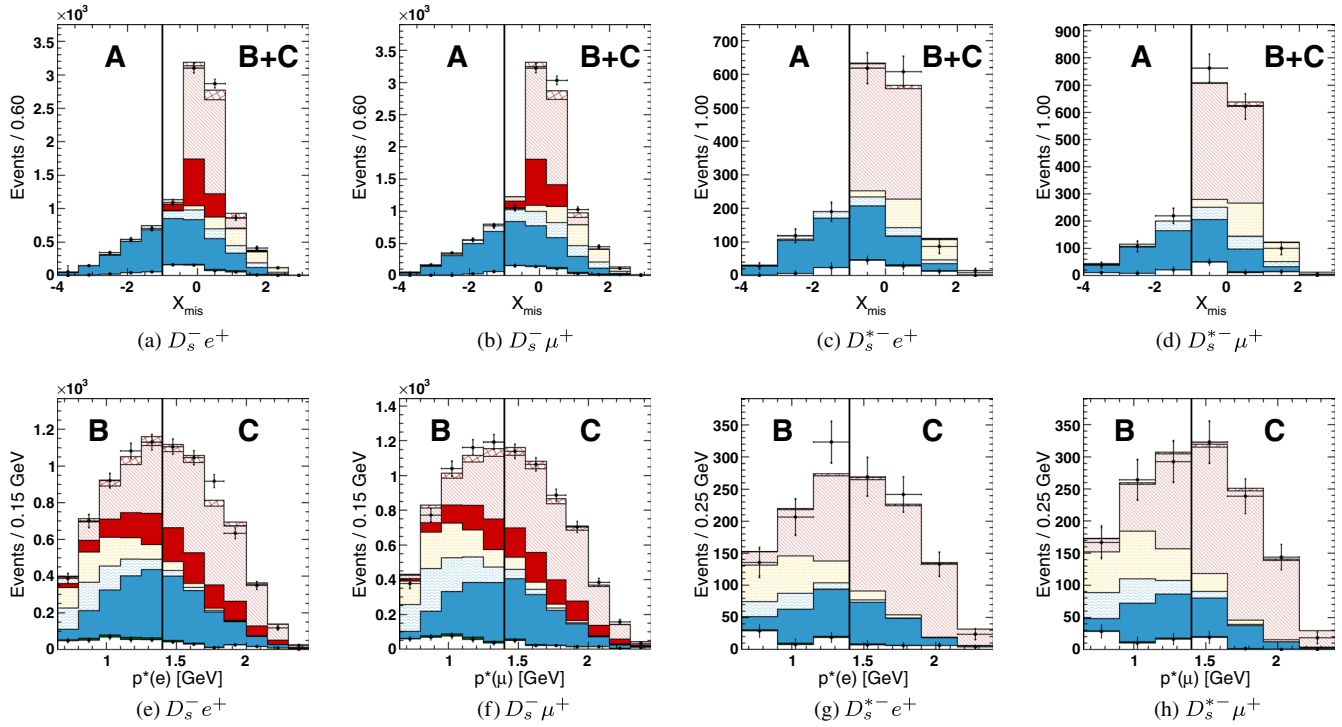

FIG. 2 (color online). Distributions of $X_{\rm mis}$ and $p^*(\ell)$ for reconstructed $D_s^{(*)-}\ell^+$ events. The black points with uncertainty bars show the $D_s^{(*)-}$ yields in the $\Upsilon(5S)$ data determined by fits to the $M_{KK\pi}$ distributions for $D_s^-\ell^+$ and the ΔM distributions for $D_s^{*-}\ell^+$. The stacked histograms represent the signal and background expectations after applying the scale factors a_j (see Table II). The components are, from bottom to top: continuum background (white), $B \to D_s^{(*)} K \ell\nu$ background (dark green), *opposite-*$B_{(s)}$ primary leptons (solid blue), *opposite-*$B_{(s)}$ secondary leptons and misidentified hadrons (hatched blue), *same-*$B_{(s)}$ background (hatched yellow), *signal* $B_s \to D_s \ell\nu$ (solid red), *signal* $B_s \to D_s^* \ell\nu$ (hatched red), *signal* $B_s \to D_s^{**} \ell\nu$ (cross-hatched red). The vertical black line illustrates the division of the counting regions. The displayed binning of the $X_{\rm mis}$ and $p^*(\ell)$ distributions is used only to illustrate the data-MC agreement; the signal yield, $N_{\rm sig}$, is extracted from the measured $D_s^{(*)-}$ yields in the three counting regions A, B and C listed in Tables III and IV.

A. Detector effects

The uncertainty on the track finding efficiency is 0.35% per track and thus 1.4% for four tracks. The photon efficiency is studied with radiative Bhabha events, from which the uncertainty is estimated to be 2%. The calibration of kaon and pion identification efficiencies is estimated from a sample of reconstructed $D^{*+} \to D^0\pi^+; D^0 \to K^-\pi^+$ decays. A variation of the obtained calibration factors within their uncertainties changes the measured signal yield by 1.4%. The efficiency of the lepton identification is estimated using the two processes $\gamma\gamma \to \ell^+\ell^-$ and $J/\psi \to \ell^+\ell^-$. The corresponding uncertainties on the measured signal yields are 1.0% and 1.6% for the electron and muon modes, respectively. The rates of hadrons being

TABLE II. The scale factors, a_j, for the MC components obtained by minimizing the χ^2 function defined in Eq. (8). The errors are the statistical uncertainties of the data and the MC sample. The signal yields are determined from Eq. (9). The yields of the other components are given in Tables III and IV. The signal efficiencies are obtained by averaging over the efficiencies for the $D_s\ell\nu$, $D_s^*\ell\nu$ and $D_s^{**}\ell\nu$ modes, taking into account the expected relative abundance of the signal components. The given errors of the efficiencies are the statistical uncertainties of the MC sample.

	Scale factors				
Channel	*Opposite-*$B_{(s)}$	*Same-*$B_{(s)}$	*Signal*	Signal yield	Efficiency [%]
$D_sXe\nu$	1.02 ± 0.04	1.00 ± 0.20	1.06 ± 0.04	4470 ± 161	16.9 ± 0.1
$D_sX\mu\nu$	1.06 ± 0.04	0.94 ± 0.16	1.09 ± 0.04	4411 ± 161	16.3 ± 0.1
$D_s^*Xe\nu$	0.89 ± 0.12	1.66 ± 0.71	1.00 ± 0.11	724 ± 79	4.6 ± 0.1
$D_s^*X\mu\nu$	0.96 ± 0.12	1.50 ± 0.58	1.13 ± 0.12	804 ± 86	4.6 ± 0.1

TABLE III. The $D_s^-\ell^+$ yields obtained from the $M_{KK\pi}$ fits to $\Upsilon(5S)$ data in the three counting regions (A, B, C) and the corresponding signal and background expectations. The scale factors from Table II obtained by minimizing Eq. (8) are applied to the MC expectations (3)–(5). The errors are the statistical uncertainties of the data and MC samples, respectively, and do not contain the scale factor uncertainties. Uncertainties are omitted if they are smaller than 0.5.

	Electrons			Muons		
	A	B	C	A	B	C
$\Upsilon(5S)$ data	1807 ± 53	4274 ± 87	4215 ± 82	1902 ± 54	4544 ± 89	4375 ± 81
(1) Continuum (scaled off-resonance data)	130 ± 34	278 ± 37	137 ± 22	102 ± 32	298 ± 40	134 ± 25
(2) $B \to D_s K\ell\nu$	0	48 ± 7	18 ± 4	0	46 ± 7	18 ± 4
(3) Opposite-$B_{(s)}$, secondary leptons, mis-ID hadrons	110 ± 4	555 ± 10	61 ± 3	205 ± 6	826 ± 12	107 ± 4
(3) Opposite-$B_{(s)}$, primary leptons	1565 ± 16	1165 ± 14	1032 ± 13	1594 ± 17	1081 ± 14	1043 ± 14
(4) Same-$B_{(s)}$ background	0	638 ± 10	89 ± 4	1	798 ± 11	158 ± 5
(5) Signal ($D_s\ell\nu$)	0	492 ± 9	669 ± 11	0	489 ± 9	693 ± 11
(5) Signal ($D_s^*\ell\nu$)	1	951 ± 13	2072 ± 19	0	872 ± 13	2072 ± 19
(5) Signal ($D_s^{**}\ell\nu; D_s^{**} \to D_s^*$)	0	28 ± 2	41 ± 3	0	26 ± 2	40 ± 3
(5) Signal ($D_s^{**}\ell\nu; D_s^{**} \nrightarrow D_s^*$)	0	117 ± 5	98 ± 4	0	109 ± 4	110 ± 4

TABLE IV. The $D_s^{*-}\ell^+$ yields obtained from the ΔM fits to $\Upsilon(5S)$ data in the three counting regions (A, B, C) and the corresponding signal and background expectations. The scale factors from Table II obtained by minimizing Eq. (8) are applied to the MC expectations (3)—(5). The errors are the statistical uncertainties of the data and MC samples, respectively, and do not contain the scale factor uncertainties.

	Electrons			Muons		
	A	B	C	A	B	C
$\Upsilon(5S)$ data	336 ± 33	656 ± 48	662 ± 46	370 ± 35	739 ± 52	741 ± 50
(1) Scaled off-resonance data	32 ± 22	61 ± 17	24 ± 11	49 ± 19	54 ± 18	20 ± 11
(2) $B \to D_s K\ell\nu$	0	6 ± 2	2 ± 1	0	4 ± 2	2 ± 1
(3) Opposite-$B_{(s)}$, secondary leptons, mis-ID hadrons	24 ± 2	60 ± 3	4 ± 1	48 ± 3	99 ± 4	13 ± 1
(3) Opposite-$B_{(s)}$, primary leptons	279 ± 6	147 ± 5	120 ± 4	273 ± 7	147 ± 5	109 ± 4
(4) Same-$B_{(s)}$ background	0	151 ± 6	20 ± 2	0	188 ± 7	39 ± 3
(5) Signal ($D_s^*\ell\nu$)	0	227 ± 6	483 ± 9	0	241 ± 7	547 ± 10
(5) Signal ($D_s^{**}\ell\nu; D_s^{**} \to D_s^*$)	0	6 ± 1	8 ± 1	0	6 ± 1	11 ± 1

misidentified as leptons are estimated with the aforementioned D^{*+} sample. The uncertainties due to this estimation are 0.1% ($D_sXe\nu$), 1.3% ($D_sX\mu\nu$), 0.1% ($D_s^*Xe\nu$) and 1.9% ($D_s^*X\mu\nu$).

B. Signal and background modeling

To study uncertainties of the PDFs in the D_s and D_s^* fits, we repeat the fits with alternative fit models and assign the resulting change of the signal yield as systematic uncertainty. Herein, we focus on the tails of the signal peaks because they can be easily assigned in the fit to the background component without deteriorating the agreement of the data with the fitted curve. The signal PDF in the D_s fits is modified by replacing the second Gaussian function by a bifurcated Gaussian function. This choice is motivated by a small asymmetry of the signal peak due to final state radiation. The normalization and the widths of the bifurcated Gaussian function are determined relative to the normalization and width of the Gaussian function from a fit to signal MC. These parameters are fixed in the fit to data. Based on the observed change of the signal yield, we assign a 3% PDF uncertainty. In the D_s^* fits, the tails of the signal peak are described by the Gaussian component of the signal PDF. When this Gaussian function is removed from the signal PDF, i.e. a Crystal Ball function only is used (cf. Ref. [43]), the signal yields decrease by 5%. Hence, we estimate the PDF uncertainty with 5%.

The uncertainty due to the continuum scale factor, S, is negligible. The uncertainty due to the shape correction for the continuum background is estimated as the full difference of the result with and without the correction applied, which is 1.2% and 0.3% for electrons and muons, respectively. To estimate the influence of the choice of the $B \to D_s^{(*)}K\ell\nu$ decay model, we replace the phase space model used in the nominal result with the ISGW2 model [44], assuming that the decay proceeds via $B \to D_0^*\ell\nu$;

TABLE V. Relative systematic uncertainties on the signal yields in %.

	$D_sXe\nu$	$D_sX\mu\nu$	$D_s^*Xe\nu$	$D_s^*X\mu\nu$
Detector				
Tracking efficiency	1.4	1.4	1.4	1.4
Photon efficiency	...	...	2.0	2.0
Kaon and pion ID	1.4	1.4	1.4	1.4
Lepton efficiency	1.0	1.6	1.0	1.6
Hadron misidentification	0.1	1.3	0.1	1.9
Signal and background modeling				
PDF for $M_{KK\pi}$ and ΔM fits	3.0	3.0	5.0	5.0
Continuum shape	1.2	0.3	1.2	0.3
$B \to D_s^{(*)}K\ell\nu$ modeling	0.3	0.3	0.1	0.1
Signal				
Composition	4.8	4.8	0.3	< 0.1
Form factors	0.9	1.0	1.0	1.0
Efficiency	3.1	3.1	3.0	3.0
Opposite-$B_{(s)}$ background				
Composition	1.6	2.2	1.0	2.5
B_s fraction	0.2	0.2	< 0.1	< 0.1
Shape	1.0	1.0	1.0	1.0
Same-$B_{(s)}$ background				
Composition and shape	0.3	0.3	0.4	0.7
B_s production mode	0.1	0.1	0.3	0.3
Beam energy	1.0	1.0	0.5	0.5
Total	7.3	7.6	6.9	7.6

$D_0^* \to D_s^{(*)}K$. The use of this alternative model increases the signal yields by 0.3% and 0.1% for the $D_sX\ell\nu$ and $D_s^*X\ell\nu$ channels, respectively. We also vary the $B \to D_s^{(*)}K\ell\nu$ branching fraction by the measured uncertainty and observe no significant change in the measured yield. We test the stability of the signal extraction when the boundary between counting region B and C is varied between $p_\ell^* = 1.3$ and 1.5 GeV. The resulting change of the signal yields is consistent with the expected change due to the increase/decrease of statistics in the respective counting regions and, therefore, no systematic uncertainty is assigned.

The systematic uncertainty on the *signal* composition in the $D_sX\ell\nu$ channels is obtained by evaluating the effect of scaling the relative amount of $B_s \to D_s^*\ell\nu$ decays up and down by 30% and adjusting the $B_s \to D_s\ell\nu$ component such that the total number of MC events is conserved. This variation covers most of the recent theory predictions and causes a 4.7% change of the signal yields. To estimate the impact of $D_s^{**} \to D_sX$ crossfeed, we double the $B_s \to D_s^{**}\ell\nu$ contribution in the signal component, which increases the signal yield by 1%. The $D_s^*X\ell\nu$ *signal* component is expected to be dominated by $B_s \to D_s^*\ell\nu$ decays and hence the uncertainty due to the amount of $D_s^{**} \to D_s^*X$ crossfeed is negligible for this channel.

The $B \to D^{(*)}\ell\nu$ form factor parameters from Ref. [37] used to simulate the $B_s \to D_s^{(*)}\ell\nu$ decays are measured with an accuracy of 2–3%. However, SU(3) flavor symmetry breaking effects may cause deviations at the order of 10% [4]. To account for these differences, we vary each form factor parameter of a given decay independently up and down by 10%. The resulting average deviation from the nominal signal yield is added linearly for each variation. The uncertainty of the LLSW model for $B_s \to D_s^{**}\ell\nu$ decays is evaluated by repeating the measurement with different sets of model parameters, as specified in Ref. [38]. The total systematic uncertainty due to form factor modeling is given by the quadratic sum of the uncertainties from all decay modes and does not exceed 1%.

The *signal* efficiencies are studied in bins of three distributions: the lepton momentum, the D_s momentum, and the angle between the reconstructed D_s meson and the lepton in the CM system. A recalculation of the average efficiencies based on the observed data yields changes the signal by at most 3.1%.

The modeling of the *opposite-$B_{(s)}$* component is studied in same-sign $D_s^+\ell^+$ control samples. The same-sign selection ensures that these samples contain only *opposite-$B_{(s)}$* combinations. Compared to the $D_s^{(*)-}\ell^+$ samples, the relative contribution of B_s decays is enhanced in this control sample. Two components of the *opposite-$B_{(s)}$* sample are distinguished: (i) primary leptons and (ii) secondary leptons and hadron tracks misidentified as leptons. Scale factors for the normalization of these two MC components are determined from fits to the $p^*(\ell)$ distributions of the $D_s^+\ell^+$ samples. The obtained scale factors are in agreement within the fit uncertainties of about 10%. A variation of the normalizations of the two components in the $D_s^{(*)-}\ell^+$ samples within this 10% uncertainty changes the signal yields between 1.0% and 2.5%, depending on the reconstructed channel. We also vary the fraction of B_s decays in the *opposite-$B_{(s)}$* component by 20%, corresponding to the uncertainty of the B_s production rate, f_s [31]. The resulting change of the signal yields is less than 0.2%. The shape uncertainty of the *opposite-$B_{(s)}$* component is evaluated in a data-driven way by using again the $D_s^+\ell^+$ samples, from which the event yields are determined in the three counting regions with the identical procedure as applied in the measurement. We calculate the ratios of data and MC yields for each counting region. These ratios range from 0.86 to 0.91 for electrons and from 0.96 to 0.97 for muons. We then modify the MC predictions for the *opposite-$B_{(s)}$* component in the $D_s^-\ell^+$ simulation accordingly and study the impact on the measurement. The results change by less than 0.4%, so an uncertainty of 1% on the modeling of the *opposite-$B_{(s)}$* component is a reasonable estimate, considering the differences between the $D_s^+\ell^+$ control samples and the $D_s^-\ell^+$ signal samples. The described approach cannot be transferred to the $D_s^{*-}\ell^+$

measurements because of the smaller sample sizes. However, the composition of the *opposite-$B_{(s)}$* background in the $D_s^{*-}\ell^+$ sample is similar to the one in the $D_s^-\ell^+$ sample and hence the same uncertainty is assigned.

The decays contributing to the *same-$B_{(s)}$* background component can be grouped into four classes with the corresponding fraction in the electron/muon channel given in parentheses: $B_{(s)} \to D_s^{(*)} X_c$ decays (70%/48%), leptons stemming from τ produced via B_s and D_s decays (21%/16%) and hadrons misidentified as leptons (9%/34%). There are no significant differences in the composition between the $D_s^-\ell^+$ and the $D_s^{*-}\ell^+$ channels. We vary the fraction of leptons from τ decays and the fraction of misidentified hadrons by ±50% and take half the difference of the resulting signal yields as the systematic uncertainty, which is below 1% for all measurements. Potential modeling uncertainties of the *same-$B_{(s)}$* component are assumed to be covered by the large variation of the composition.

We estimate the impact of the uncertainty on the different B_s production channels at the $\Upsilon(5S)$ energy by scaling the $B_s\bar{B}_s^*$ component up and down by 30% and assign half of the change in the signal yield as the systematic uncertainty of 0.1% and 0.3% for the $D_s^-\ell^+$ and $D_s^{*-}\ell^+$ modes, respectively. The beam energy is conservatively varied by ±3 MeV and signal yield variations of 1% and 0.5% are observed for the $D_s^-\ell^+$ and $D_s^{*-}\ell^+$ modes, respectively.

VII. RESULTS AND DISCUSSION

The semi-inclusive semileptonic B_s branching fractions are calculated from Eq. (3). Since the $D_s \to K^+K^-\pi^+$ reconstruction mode is also used in the determination of $N_{B_s\bar{B}_s}$ [23], the $\mathcal{B}(D_s^+ \to K^+K^-\pi^+)$ branching fraction cancels out. Using the branching fraction ratio $\mathcal{B}(D_s^+ \to \phi\pi^+)/\mathcal{B}(D_s \to K^+K^-\pi^+) = (41.6 \pm 0.8)\%$ and the branching fraction $\mathcal{B}(D_s^* \to D_s\gamma) = (94.2 \pm 0.7)\%$ [31], we obtain the semi-inclusive branching fractions:

$$D_sXe\nu\text{: } [8.1 \pm 0.3(\text{stat}) \pm 0.6(\text{syst}) \pm 1.4(\text{ext})]\%,$$
$$D_sX\mu\nu\text{: } [8.3 \pm 0.3(\text{stat}) \pm 0.6(\text{syst}) \pm 1.5(\text{ext})]\%,$$
$$D_s^*Xe\nu\text{: } [5.2 \pm 0.6(\text{stat}) \pm 0.4(\text{syst}) \pm 0.9(\text{ext})]\%,$$
$$D_s^*X\mu\nu\text{: } [5.7 \pm 0.6(\text{stat}) \pm 0.4(\text{syst}) \pm 1.0(\text{ext})]\%.$$

The first uncertainty is the statistical uncertainty of the data and MC samples, the second is the systematic uncertainty of the measurement, and the last uncertainty is due to the external measurements of $N_{B_s\bar{B}_s}$ and $\mathcal{B}_{D_s^{(*)}}$. The electron and muon samples are statistically independent because only one candidate is selected per event. Taking into account that the systematic uncertainties are all correlated except the one for lepton identification, we calculate the combination of the measurements as weighted averages:

$$D_sX\ell\nu\text{: } [8.2 \pm 0.2(\text{stat}) \pm 0.6(\text{syst}) \pm 1.4(\text{ext})]\%,$$
$$D_s^*X\ell\nu\text{: } [5.4 \pm 0.4(\text{stat}) \pm 0.4(\text{syst}) \pm 0.9(\text{ext})]\%.$$

The obtained $B_s \to D_sX\ell\nu$ branching fraction can be compared to the difference between the inclusive branching fraction, $\mathcal{B}(B_s \to X_c\ell\nu)$, and the branching fraction of the $D_s^{**}\ell\nu$ modes, where the D_s^{**} does not decay to a D_s meson. The value of $\mathcal{B}(B_s \to X_c\ell\nu)$ is estimated to be $(10.0 \pm 0.4)\%$, using the branching fraction $\mathcal{B}(B^0 \to X_c\ell\nu)$ [31,45], an estimate for the ratio of the semileptonic widths of the B_s and B^0 meson, $\Gamma_{\text{sl}}(B_s)/\Gamma_{\text{sl}}(B^0) = 0.99$ [46] and the measured lifetimes of the B^0 and B_s mesons [31].

We assume that only the semileptonic decay modes with $D_{s1}(2536)$ and $D_{s2}(2573)$ mesons do not contain D_s mesons in the final state. We obtain the estimate,

$$\begin{aligned}\mathcal{B}_{\text{est}}(B_s \to D_sX\ell\nu) &= \mathcal{B}(B_s \to X_c\ell\nu)\\ &\cdot [1 - \mathcal{B}(B_s \to D_{s2}X\ell\nu)/\mathcal{B}(B_s \to X_c\ell\nu)\\ &- \mathcal{B}(B_s \to D_{s1}X\ell\nu)/\mathcal{B}(B_s \to X_c\ell\nu)] = (9.1 \pm 0.4)\%,\end{aligned} \tag{10}$$

where the ratios $\mathcal{B}(B_s \to D_{s2}X\ell\nu)/\mathcal{B}(B_s \to X_c\ell\nu) = [3.3 \pm 1.0(\text{stat}) \pm 0.4(\text{syst})]\%$ and $\mathcal{B}(B_s \to D_{s1}X\ell\nu)/\mathcal{B}(B_s \to X_c\ell\nu) = [5.4 \pm 1.2(\text{stat}) \pm 0.5(\text{syst})]\%$ were measured at LHCb [47]. The result of our measurement is in agreement with the estimate, $\mathcal{B}_{\text{est}}(B_s \to D_sX\ell\nu)$. The rate of $B_s \to D_s^{**}\ell\nu; D_s^{**} \to D_s^*X$ decays can be constrained from the comparison between the measured semi-inclusive branching fraction, $\mathcal{B}(B_s \to D_s^*X\ell\nu)$ with the exclusive theory predictions for $\mathcal{B}(B_s \to D_s^*\ell\nu)$. For example, using the prediction from Ref. [8], one obtains $\mathcal{B}(B_s \to D_s^{**}\ell\nu; D_s^{**} \to D_s^*X) < 2.0\%$ at the 90% confidence level.

The measurement can also be used to determine $N_{B_s\bar{B}_s}$ using the estimate of the $B_s \to D_sX\ell\nu$ branching fraction from Eq. (10):

$$N_{B_s\bar{B}_s} = \frac{N_{\text{sig}}/[\epsilon\mathcal{B}(D_s \to \phi(K^+K^-)\pi^+)]}{2\mathcal{B}_{\text{est}}(B_s \to D_sX\ell\nu)}. \tag{11}$$

For N_{sig}/ϵ, we insert the weighted average of the electron and muon modes, $N_{\text{sig}}/\epsilon = [26.7 \pm 0.7(\text{stat}) \pm 2.0(\text{syst})] \times 10^3$; for $\mathcal{B}(D_s \to \phi(K^+K^-)\pi^+)$, we use the value $(2.24 \pm 0.10)\%$ [31]. We obtain $N_{B_s\bar{B}_s} = [6.53 \pm 0.17(\text{stat}) \pm 0.49(\text{syst}) \pm 0.41(\text{ext})] \times 10^6$, corresponding to the cross section $\sigma(e^+e^- \to B_s^{(*)}\bar{B}_s^{(*)}) = [53.8 \pm 1.4(\text{stat}) \pm 4.0(\text{syst}) \pm 3.4(\text{ext})]$ pb at the CM energy $\sqrt{s} = 10.86$ GeV. The first two uncertainties are the statistical and systematic uncertainties from the measurement, respectively, and the last uncertainty is due to $\mathcal{B}(D_s \to \phi(K^+K^-)\pi^+)$ and $\mathcal{B}_{\text{est}}(B_s \to D_sX\ell\nu)$. The obtained result is in agreement with $N_{B_s\bar{B}_s} = (7.1 \pm 1.3) \times 10^6$

obtained by Belle with a different technique [23] and has a significantly improved precision.

VIII. SUMMARY

We have presented the first measurements of the semi-inclusive branching fractions of $B_s \to D_s X \ell\nu$ and $B_s \to D_s^* X \ell\nu$ decays. The measured branching fractions are $\mathcal{B}(B_s \to D_s X \ell\nu) = [8.2 \pm 0.2(\text{stat}) \pm 0.6(\text{syst}) \pm 1.4(\text{ext})]\%$ and $\mathcal{B}(B_s \to D_s^* X \ell\nu) = [5.4 \pm 0.4(\text{stat}) \pm 0.4(\text{syst}) \pm 0.9(\text{ext})]\%$. In addition, the analysis of these decays provides the currently most precise estimate of the $B_s^{(*)}\bar{B}_s^{(*)}$ production cross section at the CM energy $\sqrt{s} = 10.86$ GeV: $\sigma(e^+e^- \to B_s^{(*)}\bar{B}_s^{(*)}) = [53.8 \pm 1.4(\text{stat}) \pm 4.0(\text{syst}) \pm 3.4(\text{ext})]$ pb.

ACKNOWLEDGMENTS

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, the National Institute of Informatics, and the PNNL/EMSL computing group for valuable computing and SINET4 network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS), and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council and the Australian Department of Industry, Innovation, Science and Research; Austrian Science Fund under Grants No. P 22742-N16 and No. P 26794-N20; the National Natural Science Foundation of China under Contracts No. 10575109, No. 10775142, No. 10875115, No. 11175187, and No. 11475187; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LG14034; the Carl Zeiss Foundation, the Deutsche Forschungsgemeinschaft and the VolkswagenStiftung; the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; National Research Foundation (NRF) of Korea Grants No. 2011-0029457, No. 2012-0008143, No. 2012R1A1A2008330, No. 2013R1A1A3007772, No. 2014R1A2A2A01005286, No. 2014R1A2A2A01002734, No. 2014R1A1A2006456; the Basic Research Lab program under NRF Grant No. KRF-2011-0020333, No. KRF-2011-0021196, Center for Korean J-PARC Users, No. NRF-2013K1A3A7A06056592; the Brain Korea 21-Plus program and the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information; the Polish Ministry of Science and Higher Education and the National Science Center; the Ministry of Education and Science of the Russian Federation and the Russian Foundation for Basic Research; the Slovenian Research Agency; the Basque Foundation for Science (IKERBASQUE) and the Euskal Herriko Unibertsitatea (UPV/EHU) under program UFI 11/55 (Spain); the Swiss National Science Foundation; the National Science Council and the Ministry of Education of Taiwan; and the U.S. Department of Energy and the National Science Foundation. This work is supported by a Grant-in-Aid from MEXT for Science Research in a Priority Area ("New Development of Flavor Physics") and from JSPS for Creative Scientific Research ("Evolution of Tau-lepton Physics").

[1] R. H. Li, C. D. Lu, and Y. M. Wang, Phys. Rev. D **80**, 014005 (2009).
[2] K. Azizi, Nucl. Phys. **B801**, 70 (2008).
[3] K. Azizi and M. Bayar, Phys. Rev. D **78**, 054011 (2008).
[4] P. Blasi, P. Colangelo, G. Nardulli, and N. Paver, Phys. Rev. D **49**, 238 (1994).
[5] M. Atoui, V. Mornas, D. Bečirevic, and F. Sanfilippo, Eur. Phys. J. C **74**, 2861 (2014).
[6] J. A. Bailey *et al.*, Phys. Rev. D **85**, 114502 (2012); **86**, 039904(E) (2012).
[7] Y. Y. Fan, W. F. Wang, and Z. J. Xiao, Phys. Rev. D **89**, 014030 (2014).
[8] R. N. Faustov and V. O. Galkin, Phys. Rev. D **87**, 034033 (2013).
[9] G. Li, F. L. Shao, and W. Wang, Phys. Rev. D **82**, 094031 (2010).
[10] J. M. Zhang and G. L. Wang, Chin. Phys. Lett. **27**, 051301 (2010).
[11] S. M. Zhao, X. Liu, and S. J. Li, Eur. Phys. J. C **51**, 601 (2007).
[12] X. J. Chen, H. F. Fu, C. S. Kim, and G. L. Wang, J. Phys. G **39**, 045002 (2012).
[13] J. Segovia, C. Albertus, D. R. Entem, F. Fernandez, E. Hernandez, and M. A. Perez-Garcia, Phys. Rev. D **84**, 094029 (2011).
[14] C. Albertus, Phys. Rev. D **89**, 065042 (2014).
[15] A. Bhol, Europhys. Lett. **106**, 31001 (2014).
[16] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **102**, 051801 (2009).
[17] R. Aaij *et al.* (LHCb Collaboration), Phys. Lett. B **698**, 14 (2011).
[18] C. Oswald *et al.* (Belle Collaboration), Phys. Rev. D **87**, 072008 (2013); **90**, 119901(E) (2014).
[19] J. P. Lees *et al.* (*BABAR* Collaboration), Phys. Rev. D **85**, 011101 (2012).
[20] I. I. Bigi, T. Mannel, and N. Uraltsev, J. High Energy Phys. 09 (2011) 012.
[21] M. Gronau and J. L. Rosner, Phys. Rev. D **83**, 034025 (2011).

[22] C. Oswald and T. K. Pedlar, Mod. Phys. Lett. A **28**, 1330036 (2013).
[23] The value $N_{B_s\bar{B}_s}\mathcal{B}(D_s^+ \to K^+K^-\pi^+) = (3.90 \pm 0.69) \times 10^5$ was obtained by the Belle Collaboration with the method described in A. Drutskoy *et al.* (Belle Collaboration), Phys. Rev. Lett. **98**, 052001 (2007).
[24] D. Buskulic *et al.* (ALEPH Collaboration), Phys. Lett. B **361**, 221 (1995).
[25] P. Abreu *et al.* (DELPHI Collaboration), Phys. Lett. B **289**, 199 (1992).
[26] P. D. Acton *et al.* (OPAL Collaboration), Phys. Lett. B **295**, 357 (1992).
[27] R. Aaij *et al.* (LHCb Collaboration), Phys. Rev. D **85**, 032008 (2012).
[28] 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.
[29] 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).
[30] Throughout this paper, the convention $c = 1$ is used.
[31] K. A. Olive *et al.* (Particle Data Group), Chin. Phys. C **38**, 090001 (2014).
[32] D. J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
[33] R. Brun *et al.*, GEANT 3.21 CERN Report No. DD/EE/84-1, 1984.
[34] E. Barberio and Z. Wąs, Comput. Phys. Commun. **79**, 291 (1994).
[35] I. Caprini, L. Lellouch, and M. Neubert, Nucl. Phys. **B530**, 153 (1998).
[36] M. Neubert, Phys. Rep. **245**, 259 (1994).
[37] Y. Amhis *et al.* (Heavy Flavor Averaging Group), arXiv:1207.1158.
[38] A. K. Leibovich, Z. Ligeti, I. W. Stewart, and M. B. Wise, Phys. Rev. D **57**, 308 (1998).
[39] Throughout this paper, the inclusion of the charge conjugate mode decay is implied unless otherwise stated.
[40] J. Stypula *et al.* (Belle Collaboration), Phys. Rev. D **86**, 072007 (2012).
[41] P. del Amo Sanchez *et al.* (*BABAR* Collaboration), Phys. Rev. Lett. **107**, 041804 (2011).
[42] M. J. Oreglia, Ph.D. thesis: Stanford University, 1980; Report No. SLAC-R-236, 1980; J. E. Gaiser, Ph.D. thesis: Stanford University, 1982; Report No. SLAC-R-255 1982.
[43] B. Aubert *et al.* (*BABAR* Collaboration), Phys. Rev. D **72**, 091101 (2005).
[44] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995).
[45] P. Urquijo *et al.* (Belle Collaboration), Phys. Rev. D **75**, 032001 (2007).
[46] I. I. Bigi, T. Mannel, and N. Uraltsev, J. High Energy Phys. 09 (2011) 012.
[47] The $\mathcal{B}(B_s \to D_{sj}X\ell\nu)/\mathcal{B}(B_s \to X_c\ell\nu)$ $(j = 1, 2)$ measurements are described in Ref. [17]. We consider the systematic uncertainties as fully correlated, except for the 8% shape error on the D_{s2} fit.