Evidence for B^{-}→D_{s}^{+}K^{-}ℓ^{-}ν[over ¯]_{ℓ} and search for B^{-}→D_{s}^{*+}K^{-}ℓ^{-}ν[over ¯]_{ℓ}

J. Stypula *et al.* (The Belle Collaboration)

Phys. Rev. D **86**, 072007 — Published 23 October 2012

DOI: 10.1103/PhysRevD.86.072007

Evidence for $B^- \to D_s^+ K^- \ell^- \bar{\nu}_\ell$ and search for $B^- \to D_s^{*+} K^- \ell^- \bar{\nu}_\ell$

J. Stypula,[33] M. Rozanska,[33] I. Adachi,[9] K. Adamczyk,[33] H. Aihara,[49] D. M. Asner,[38] T. Aushev,[15] A. M. Bakich,[43] V. Bhardwaj,[30] B. Bhuyan,[10] M. Bischofberger,[30] A. Bondar,[2] G. Bonvicini,[54] A. Bozek,[33] M. Bračko,[25,16] T. E. Browder,[8] M.-C. Chang,[5] P. Chang,[32] V. Chekelian,[26] A. Chen,[31] P. Chen,[32] B. G. Cheon,[7] R. Chistov,[15] I.-S. Cho,[56] K. Cho,[19] Y. Choi,[42] J. Dalseno,[26,45] M. Danilov,[15] J. Dingfelder,[1] Z. Doležal,[3] Z. Drásal,[3] A. Drutskoy,[15] S. Eidelman,[2] H. Farhat,[54] J. E. Fast,[38] V. Gaur,[44] N. Gabyshev,[2] R. Gillard,[54] Y. M. Goh,[7] B. Golob,[23,16] J. Haba,[9] K. Hayasaka,[29] H. Hayashii,[30] Y. Horii,[29] Y. Hoshi,[47] W.-S. Hou,[32] Y. B. Hsiung,[32] H. J. Hyun,[21] T. Iijima,[29,28] K. Inami,[28] A. Ishikawa,[48] R. Itoh,[9] M. Iwabuchi,[56] Y. Iwasaki,[9] T. Julius,[27] J. H. Kang,[56] P. Kapusta,[33] T. Kawasaki,[35] H. Kichimi,[9] C. Kiesling,[26] H. J. Kim,[21] J. B. Kim,[20] J. H. Kim,[19] K. T. Kim,[20] Y. J. Kim,[19] K. Kinoshita,[4] B. R. Ko,[20] P. Kodyš,[3] S. Korpar,[25,16] R. T. Kouzes,[38] P. Križan,[23,16] P. Krokovny,[2] T. Kuhr,[18] T. Kumita,[51] A. Kuzmin,[2] Y.-J. Kwon,[56] S.-H. Lee,[20] J. Li,[41] Y. Li,[53] J. Libby,[11] C. Liu,[40] Y. Liu,[4] Z. Q. Liu,[12] D. Liventsev,[15] R. Louvot,[22] K. Miyabayashi,[30] H. Miyata,[35] Y. Miyazaki,[28] R. Mizuk,[15] G. B. Mohanty,[44] A. Moll,[26,45] N. Muramatsu,[39] E. Nakano,[37] M. Nakao,[9] Z. Natkaniec,[33] C. Ng,[49] S. Nishida,[9] K. Nishimura,[8] O. Nitoh,[52] T. Nozaki,[9] S. Ogawa,[46] T. Ohshima,[28] S. Okuno,[17] S. L. Olsen,[41,8] Y. Onuki,[49] P. Pakhlov,[15] G. Pakhlova,[15] C. W. Park,[42] H. Park,[21] H. K. Park,[21] T. K. Pedlar,[24] R. Pestotnik,[16] M. Petrič,[16] L. E. Piilonen,[53] M. Ritter,[26] M. Röhrken,[18] S. Ryu,[41] H. Sahoo,[8] Y. Sakai,[9] S. Sandilya,[44] D. Santel,[4] T. Sanuki,[48] Y. Sato,[48] O. Schneider,[22] C. Schwanda,[13] K. Senyo,[55] O. Seon,[28] M. E. Sevior,[27] M. Shapkin,[14] C. P. Shen,[28] T.-A. Shibata,[50] J.-G. Shiu,[32] B. Shwartz,[2] A. Sibidanov,[43] F. Simon,[26,45] P. Smerkol,[16] Y.-S. Sohn,[56] A. Sokolov,[14] E. Solovieva,[15] S. Stanič,[36] M. Starič,[16] M. Sumihama,[6] T. Sumiyoshi,[51] Y. Teramoto,[37] M. Uchida,[50] T. Uglov,[15] Y. Unno,[7] S. Uno,[9] P. Urquijo,[1] Y. Usov,[2] P. Vanhoefer,[26] G. Varner,[8] K. E. Varvell,[43] V. Vorobyev,[2] P. Wang,[12] X. L. Wang,[12] M. Watanabe,[35] Y. Watanabe,[17] J. Wiechczynski,[33] K. M. Williams,[53] E. Won,[20] B. D. Yabsley,[43] H. Yamamoto,[48] Y. Yamashita,[34] Z. P. Zhang,[40] V. Zhilich,[2] V. Zhulanov,[2] and A. Zupanc[18]

(The Belle Collaboration)

[1]*University of Bonn, Bonn*
[2]*Budker Institute of Nuclear Physics SB RAS and Novosibirsk State University, Novosibirsk 630090*
[3]*Faculty of Mathematics and Physics, Charles University, Prague*
[4]*University of Cincinnati, Cincinnati, Ohio 45221*
[5]*Department of Physics, Fu Jen Catholic University, Taipei*
[6]*Gifu University, Gifu*
[7]*Hanyang University, Seoul*
[8]*University of Hawaii, Honolulu, Hawaii 96822*
[9]*High Energy Accelerator Research Organization (KEK), Tsukuba*
[10]*Indian Institute of Technology Guwahati, Guwahati*
[11]*Indian Institute of Technology Madras, Madras*
[12]*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing*
[13]*Institute of High Energy Physics, Vienna*
[14]*Institute of High Energy Physics, Protvino*
[15]*Institute for Theoretical and Experimental Physics, Moscow*
[16]*J. Stefan Institute, Ljubljana*
[17]*Kanagawa University, Yokohama*
[18]*Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Karlsruhe*
[19]*Korea Institute of Science and Technology Information, Daejeon*
[20]*Korea University, Seoul*
[21]*Kyungpook National University, Taegu*
[22]*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne*
[23]*Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana*
[24]*Luther College, Decorah, Iowa 52101*
[25]*University of Maribor, Maribor*
[26]*Max-Planck-Institut für Physik, München*
[27]*University of Melbourne, School of Physics, Victoria 3010*
[28]*Graduate School of Science, Nagoya University, Nagoya*
[29]*Kobayashi-Maskawa Institute, Nagoya University, Nagoya*
[30]*Nara Women's University, Nara*
[31]*National Central University, Chung-li*

[32]Department of Physics, National Taiwan University, Taipei
[33]H. Niewodniczanski Institute of Nuclear Physics, Krakow
[34]Nippon Dental University, Niigata
[35]Niigata University, Niigata
[36]University of Nova Gorica, Nova Gorica
[37]Osaka City University, Osaka
[38]Pacific Northwest National Laboratory, Richland, Washington 99352
[39]Research Center for Electron Photon Science, Tohoku University, Sendai
[40]University of Science and Technology of China, Hefei
[41]Seoul National University, Seoul
[42]Sungkyunkwan University, Suwon
[43]School of Physics, University of Sydney, NSW 2006
[44]Tata Institute of Fundamental Research, Mumbai
[45]Excellence Cluster Universe, Technische Universität München, Garching
[46]Toho University, Funabashi
[47]Tohoku Gakuin University, Tagajo
[48]Tohoku University, Sendai
[49]Department of Physics, University of Tokyo, Tokyo
[50]Tokyo Institute of Technology, Tokyo
[51]Tokyo Metropolitan University, Tokyo
[52]Tokyo University of Agriculture and Technology, Tokyo
[53]CNP, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[54]Wayne State University, Detroit, Michigan 48202
[55]Yamagata University, Yamagata
[56]Yonsei University, Seoul

We report measurements of the decays $B^- \to D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell$ in a data sample containing 657×10^6 $B\bar{B}$ pairs collected with the Belle detector at the KEKB asymmetric-energy e^+e^- collider. We observe a signal with a significance of 6σ for the combined D_s and D_s^* modes and find the first evidence of the $B^- \to D_s^+ K^- \ell^- \bar{\nu}_\ell$ decay with a significance of 3.4σ. We measure the following branching fractions: $\mathcal{B}(B^- \to D_s^+ K^- \ell^- \bar{\nu}_\ell) = (0.30 \pm 0.09(\text{stat})^{+0.11}_{-0.08}(\text{syst})) \times 10^{-3}$ and $\mathcal{B}(B^- \to D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell) = (0.59 \pm 0.12(\text{stat}) \pm 0.15(\text{syst})) \times 10^{-3}$ and set an upper limit $\mathcal{B}(B^- \to D_s^{*+} K^- \ell^- \bar{\nu}_\ell) < 0.56 \times 10^{-3}$ at the 90% confidence level. We also present the first measurement of the $D_s^+ K^-$ invariant mass distribution in these decays, which is dominated by a prominent peak around 2.6 GeV/c^2.

PACS numbers: 13.20.He, 14.40.Nd

Semileptonic B decays play a key role in testing the Standard Model (SM) and in the understanding of heavy quark dynamics. In particular, they are used to determine the weak mixing parameters $|V_{qb}|$ ($q = c, u$), complementing the measurements of CP asymmetries used to verify the Cabibbo-Kobayashi-Maskawa (CKM) mechanism of the SM [1]. The tension at the level of 2 standard deviations (σ) between the values of $|V_{qb}|$ extracted from inclusive and exclusive B decays [2], as well as some discrepancies between measurements and theoretical expectations for semileptonic B decays to excited charmed mesons, may indicate problems in the theoretical tools or in the interpretation of the experimental results.

Semileptonic B decays to final states containing a $D_s^{(*)+}\bar{K}$ system [3] provide information about the poorly known region of hadronic masses above 2.46 GeV/c^2, covering radially excited D meson states [4]. Further exploration of this region may help solving some puzzles in semileptonic B decays [5]. Recently, BaBar reported an observation of $B^- \to D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell$ (which did not distinguish between the D_s and D_s^* final states) with a branching fraction of $\mathcal{B}(B^- \to D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell) = (6.13^{+1.04}_{-1.03}(\text{stat}) \pm 0.43(\text{syst}) \pm 0.51(\mathcal{B}(D_s))) \times 10^{-4}$ [6].

In this paper, we present measurements of $B^- \to D_s^+ K^- \ell^- \bar{\nu}_\ell$ and $B^- \to D_s^{*+} K^- \ell^- \bar{\nu}_\ell$ decays using a data sample containing 657×10^6 $B\bar{B}$ pairs that were collected with the Belle detector at the KEKB asymmetric-energy e^+e^- collider [7] operating at the $\Upsilon(4S)$ resonance (center-of-mass energy $\sqrt{s} = 10.58$ GeV). The Belle detector is a large-solid-angle magnetic spectrometer consisting of a silicon vertex detector, a 50-layer central drift chamber, a system of aerogel Cherenkov counters, time-of-flight scintillation counters and an electromagnetic calorimeter 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 the coil is instrumented to identify K_L^0 mesons and muons. A detailed description of the detector can be found in Ref. [8]. We use Monte Carlo (MC) simulations to estimate signal efficiencies and background contributions. Large signal samples of $B^- \to D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell$ decays are generated with the EvtGen package [9], using a phase space model and the ISGW2 model [10] including the resonances that can decay to $D_s^{(*)}\bar{K}$. Radiative effects are modeled by PHOTOS [11]. MC samples equivalent to about ten (six) times the accumulated data are used to evaluate the background from $B\bar{B}$ (continuum $q\bar{q}$, where $q = u, d, s, c$) events.

In the analysis, we use charged tracks with impact parameters that are consistent with an origin at the beam spot and have transverse momenta above 50 MeV/c. Masses are assigned using information from particle identification subsystems. The efficiency for kaon (pion) identification ranges from 84% to 98% (92% to 94%) depending on the track momentum with a pion (kaon) misidentification probability of about 8% (16%). Electrons and muons are selected with an efficiency of about 90% and a misidentification rate below 0.2% (e) and 1.4% (μ). The momenta of particles identified as electrons are corrected for bremsstrahlung by adding photons within a 50 mrad cone around the charged particle's trajectory.

D_s^+ candidates are reconstructed in the cleanest decay chain: $D_s^+ \to \phi\pi^+$, $\phi \to K^+K^-$ (2.32 ± 0.14% product branching fraction) and subjected to a vertex fit. We accept candidates in the invariant mass range of $1.934\ \mathrm{GeV}/c^2 < M_{D_s} < 2.003\ \mathrm{GeV}/c^2$, and define the signal window within $\pm 14\ \mathrm{MeV}/c^2$ around the world average D_s mass [12]. The width of this window corresponds to 4σ of the reconstructed D_s mass, using the resolution determined from control samples in data (mentioned later). The regions outside the signal window are considered as M_{D_s} sidebands. D_s^+ candidates are combined with photons with an energy $E_\gamma > 125$ MeV to form D_s^{*+} candidates, subjected to a mass constrained vertex fit. Throughout this paper, all kinematic variables are defined in the $\Upsilon(4S)$ rest frame, unless otherwise stated. D_s^{*+} candidates with an invariant mass in the range of $2.079\ \mathrm{GeV}/c^2 < M_{D_s^*} < 2.155\ \mathrm{GeV}/c^2$ are accepted for further analysis. The signal window is defined as $2.087\ \mathrm{GeV}/c^2 < M_{D_s^*} < 2.137\ \mathrm{GeV}/c^2$ (3.7σ in $M_{D_s^*}$). Signal candidates for the decays considered here (B_{sig}) are formed by combining a negatively charged kaon and lepton (e or μ) with a $D_s^{(*)+}$ candidate. In the case of multiple B_{sig} candidates (22% of events after final selection requirements have multiple B_{sig} candidates), the one with the greatest confidence level of the vertex fit is chosen. Events with accepted $D_s^{*+}K^-\ell^-$ candidates (D_s^* sample) are removed from the set of $D_s^+K^-\ell^-$ candidates (D_s sample). Another charge configuration, $D_s^{(*)+}K^+\ell^-$, populated by decays of the type $B \to D_s^{(*)+}\bar{D}^{(*)}$, $\bar{D} \to \ell^-\bar{\nu}_\ell K^+X$, is used as a control sample.

Signal events are identified using the variable X_{mis}, introduced in Ref. [13] and defined as: $X_{\mathrm{mis}} \equiv (E_{\mathrm{beam}} - E_{D_sK\ell} - |\vec{p}_{D_sK\ell}|)/\sqrt{E^2_{\mathrm{beam}} - m^2_{B^+}}$, where E_{beam} is the beam energy, $E_{D_sK\ell}$ and $\vec{p}_{D_sK\ell}$ denote the total energy and momentum of the $D_sK\ell$ system, respectively, and m_{B^+} is the nominal B^+ mass. For decays with at most one massless invisible particle, as expected for the signal, X_{mis} takes values in the range of $[-1, 1]$, defined as the signal region, while the background has a much broader distribution. X_{mis} is calculated with the four-momentum of the D_s both in the D_s and D_s^* samples, causing a small shift of X_{mis} toward higher values for the D_s^* case due to the additional low-energy photon. With this definition, the X_{mis} distribution is more robust against imperfect modeling of photon spectra in MC and simplifies the signal extraction.

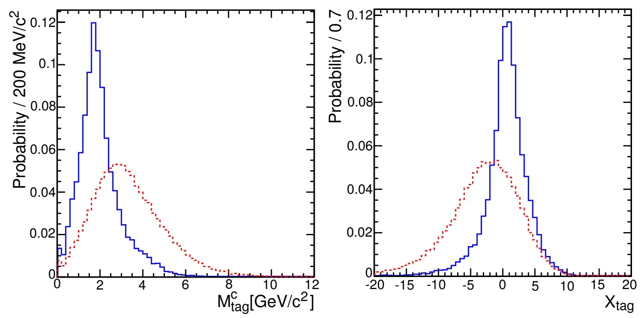

FIG. 1: M^c_{tag} (left) and X_{tag} (right) distributions for signal (blue, solid) and background (red, dashed) MC.

Particles not assigned to the B_{sig} are used to reconstruct the tagging side of the event (B_{tag}). Exploiting the information given by B_{tag} allows for background suppression without assumptions on the (unknown) signal dynamics. We require zero total event charge as well as a negatively charged lepton with a momentum above 0.5 GeV/c on the tagging side. This reduces the main background, where a D_s^+ produced in a decay of the type $B \to D_s^{(*)+}\bar{D}^{(*)}$ is combined with a lepton and a kaon from the subsequent D decay in a semileptonic decay $\bar{B} \to \ell^-\bar{\nu}_\ell D^{(*)}X$ of the accompanying $\bar{B}$ meson. Further improvement of the sensitivity is achieved with two tagging side variables $M^c_{\mathrm{tag}} \equiv \sqrt{(E_{\mathrm{tag}} - E^\ell_{\mathrm{tag}})^2 - (\vec{p}_{\mathrm{tag}} - \vec{p}^{\,\ell}_{\mathrm{tag}})^2}$ and $X_{\mathrm{tag}} \equiv (E_{\mathrm{beam}} - E_{\mathrm{tag}} - |\vec{p}_{\mathrm{tag}}|)/\sqrt{E^2_{\mathrm{beam}} - m^2_{B^+}}$, where E_{tag} and $\vec{p}_{\mathrm{tag}}$ denote the total energy and momentum of all reconstructed particles not assigned to B_{sig}, and E^ℓ_{tag} and $\vec{p}^{\,\ell}_{\mathrm{tag}}$ represent the energy and momentum of the prompt tagging lepton. Here M^c_{tag} represents the inclusively reconstructed mass of the hadronic system produced in the B_{tag} decay and X_{tag} is the tagging side equivalent of X_{mis}. The M^c_{tag} and X_{tag} distributions for signal and background are shown in Fig. 1.

In this blind analysis, the selection criteria for X_{tag} and M^c_{tag} are optimized for the D_s mode by maximizing the expected statistical significance, $N_S/\sqrt{N_S + N_B}$, where N_S (N_B) is the predicted number of signal (background) events in the (X_{mis}, M_{D_s}) signal window. This optimization is carried out for signal branching fractions $\mathcal{B}(B^- \to D_s^+K^-\ell^-\bar{\nu}_\ell)$ in the range of $(0.25 - 0.50) \times 10^{-3}$ and yields similar optimal selection criteria for the whole range, namely $-2 < X_{\mathrm{tag}} < 3$ and $M^c_{\mathrm{tag}} < 2.4\ \mathrm{GeV}/c^2$. N_B is evaluated considering two background categories in the D_s sample: "true D_s" background with correctly reconstructed D_s^+, described by the MC scaled to the integrated luminosity in data, and a "fake D_s" component, where random track combinations are misreconstructed as D_s^+, which is evaluated from the M_{D_s} sidebands. In the D_s^* sample, the background with true D_s is split into two parts: "true D_s^*" with properly reconstructed D_s^{*+} and "fake D_s^*", where a true D_s^+ is combined with a random photon candidate. The background model is tested using distributions in the sideband regions $X_{\rm mis} < -1$ and $X_{\rm mis} > 1$.

The $X_{\rm mis}$ and $M_{D_s^{(*)}}$ distributions in data are shown in Fig. 2. Figure 3 shows the invariant mass distribution of the $D_s^+K^-$ system, M_{D_sK}, for the combined D_s and D_s^* samples in the signal window and in the $X_{\rm mis}$ sidebands. Superimposed histograms represent the expected backgrounds. While the background model describes the experimental M_{D_sK} distribution well in the $X_{\rm mis}$ sidebands, a clear excess over the expected background is seen in the signal region. The M_{D_sK} distribution in the signal window is dominated by a prominent peak at ≈ 2.6 GeV/c^2, similarly to that observed in $B^- \to D_s^+K^-\pi^-$ decays [14].

The signal yields are extracted from a simultaneous, extended unbinned maximum likelihood fit to the D_s and D_s^* samples, consisting of 2175 and 396 events, respectively. The D_s and D_s^* samples are fitted in two $(X_{\rm mis}, M_{D_s})$ and three $(X_{\rm mis}, M_{D_s}, M_{D_s^*})$ dimensions, respectively. The likelihood function is constructed as follows:

$$\mathcal{L} = e^{-(\sum_k N_k + \sum_{k'} N^*_{k'})} \prod_{i=1}^{N} [\sum_k N_k \mathcal{P}_k(x_i, y_i)] \times \prod_{i'=1}^{N^*} [\sum_{k'} N^*_{k'} \mathcal{P}^*_{k'}(x_{i'}, y_{i'}, z_{i'})],$$

where x_l, y_l, z_l denote $X_{\rm mis}$, M_{D_s} and $M_{D_s^*}$ in the $l^{\rm th}$ event, and $N^{(*)}$ denotes the total number of events in the $D_s^{(*)}$ data sample. The index k (k') runs over the signal and background components in the D_s (D_s^*) sample; $N_k^{(*)}$ and $\mathcal{P}_k^{(*)}$ denote the number of events and the probability density functions (PDF) for each component, respectively. In the D_s sample, we consider two signal components coming from the decay $B^- \to D_s^+K^-\ell^-\bar{\nu}_\ell$ and from the decay $B^- \to D_s^{*+}K^-\ell^-\bar{\nu}_\ell$ if a photon from the D_s^{*+} has been missed. In the D_s^* sample, we distinguish three signal components: one coming from the $B^- \to D_s^+K^-\ell^-\bar{\nu}_\ell$ mode, where the D_s meson is associated with a random photon, and two from the $B^- \to D_s^{*+}K^-\ell^-\bar{\nu}_\ell$ mode, with true and fake D_s^* defined similarly to the background case discussed above. The coefficients $N_k^{(*)}$ for the signal components are expressed as the products $N_k^{(*)} = N_{D_s^{(*)}} f_k^{(*)}$, where $N_{D_s^{(*)}}$ denotes the total number of signal events in the $B^- \to D_s^{(*)+}K^-\ell^-\bar{\nu}_\ell$ modes. The coefficients $f_k^{(*)}$ (listed in Table I) represent the signal fraction reconstructed in each component and are evaluated from the signal MC. The coefficients $N_k^{(*)}$ for background components with fake D_s are evaluated from the M_{D_s} sidebands in data and are fixed in the fit. The two- (three-) dimensional PDF is parameterized as the product of two (three) one-dimensional PDFs for each variable. The validity of this parameterization has been checked with MC by examining the correlation between $X_{\rm mis}$ and M_{D_s}, which has been found negligible. The components with true $D_s^{(*)}$ are parameterized as a sum of two Gaussian functions in M_{D_s} or as a single Gaussian function in $M_{D_s^*}$, with means set to the world average $D_s^{(*)}$ mass values [12] and with the remaining parameters fixed from fits to control samples in data. The components with fake $D_s^{(*)}$ are parameterized as linear functions in $M_{D_s^{(*)}}$. The $X_{\rm mis}$ distribution of the signal components is modeled with two line shapes, one describing the two components of the $B^- \to D_s^+K^-\ell^-\bar{\nu}_\ell$ mode and the other one describing the three components of the $B^- \to D_s^{*+}K^-\ell^-\bar{\nu}_\ell$ decay. They are parameterized using the function $Ce^{-|(X_{\rm mis}-\mu)/\sigma|^n}e^{-\alpha(X_{\rm mis}-\mu)}$, where C is a normalization coefficient and the parameters μ, σ, α and the integer parameter n are fixed from fits to the signal MC samples. The $X_{\rm mis}$ distributions of the background components are parameterized as bifurcated Gaussian functions with parameters fixed from the simulated $B\bar{B}$ events with generic B decays (true D_s) or from the M_{D_s} sidebands in data (fake D_s). The free parameters in the fit are the two signal yields $N_{D_s^{(*)}}$, the three background yields $N_m^{(*)}$ of the components with true D_s, and the coefficients of polynomials that describe the distributions in $M_{D_s^{(*)}}$ for the fake D_s components. The range of the fit is as shown in Fig. 2. The signal yields extracted from the fit are 84 ± 24 events for the decay $B^- \to D_s^+K^-\ell^-\bar{\nu}_\ell$ and 41 ± 22 events for the decay $B^- \to D_s^{*+}K^-\ell^-\bar{\nu}_\ell$ with statistical significances of 3.9σ and 1.9σ, respectively. The significance is defined as $\Sigma = \sqrt{-2\ln(\mathcal{L}_0/\mathcal{L}_{\rm max})}$, where $\mathcal{L}_{\rm max}$ and $\mathcal{L}_0$ denote the maximum likelihood value and the likelihood value for the zero signal hypothesis, respectively. The fit results are summarized in Table II and the fit projections in $X_{\rm mis}$ and M_{D_s} are shown in Fig. 2. The fitted signal yields are used to compute the branching fractions with the formula: $\mathcal{B}(B^- \to D_s^{(*)+}K^-\ell^-\bar{\nu}_\ell) = N_s^{(*)}/(2N_{B^+B^-}\epsilon^{(*)}\mathcal{B}_{int})$, where $N_{B^+B^-}$ is the number of B^+B^- pairs in data, $\epsilon^{(*)}$ denotes the reconstruction efficiency of the signal decay chain and $\mathcal{B}_{int}$ is the product of intermediate branching fractions set to their world average values [12]. The reconstruction efficiency is expressed as $\epsilon^{(*)} = \epsilon_{PS}^{(*)}\Delta\epsilon_{cor}^{(*)}$, where $\epsilon_{PS}^{(*)}$ is the efficiency calculated from the signal MC with the phase space model and $\Delta\epsilon_{cor}^{(*)} = 1.20$ (0.57) corrects for the difference be-

TABLE I: The coefficients $f_k^{(*)}$, representing the signal fraction reconstructed in each component, evaluated from the signal MC.

Signal component k	Sample	$f_k^{(*)}$
$B^- \to D_s^+K^-\ell^-\bar{\nu}_\ell$	D_s	$(84 \pm 1)\%$
	D_s^*	$(16 \pm 1)\%$
$B^- \to D_s^{*+}K^-\ell^-\bar{\nu}_\ell$	D_s^* with true D_s^*	$(21 \pm 1)\%$
	D_s^* with fake D_s^*	$(13 \pm 1)\%$
	D_s	$(66 \pm 1)\%$

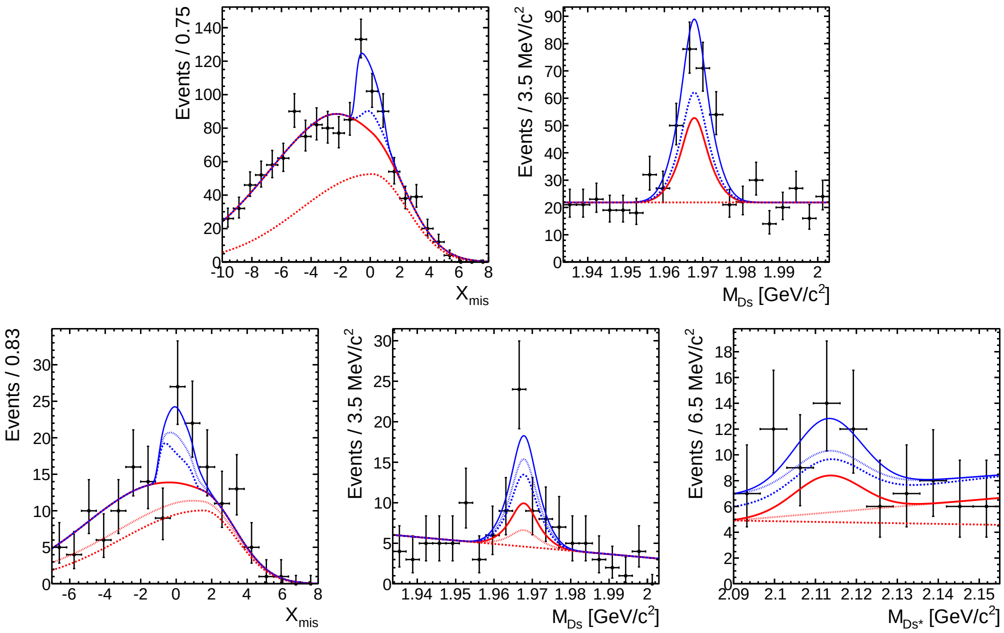

FIG. 2: Distributions (from left to right) in $X_{\rm mis}$ and M_{D_s} in the D_s sample (top), and $X_{\rm mis}$, M_{D_s} and $M_{D_s^*}$ in the D_s^* sample (bottom). Points with error bars are the data, and lines show the fit projections. Each variable is shown in the signal region of the other variable(s). For the D_s sample the lines represent (from bottom to top) the fitted background components with fake (red dashed) and true D_s (red solid), and the signal contributions from the D_s^* (blue dashed) and D_s (blue solid) modes. For the D_s^* sample the lines (from bottom to top) represent the fitted background components with fake D_s (red dashed), fake D_s^* (red dotted), true D_s^* (red solid), and the signal contributions from the D_s mode (blue dashed), the D_s^* mode with fake D_s^* (blue dotted), and with true D_s^* (blue solid). The fitted contributions are superimposed additively.

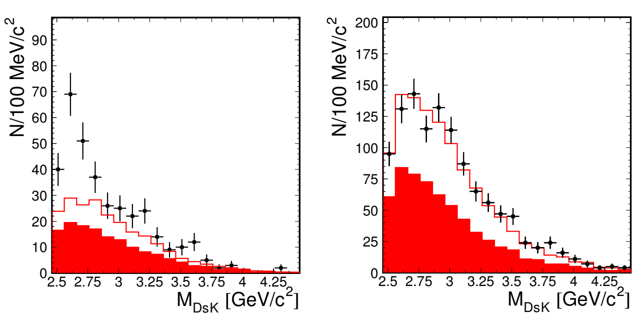

FIG. 3: The invariant mass distribution of $D_s^+K^-$ for the combined D_s and D_s^* samples in the signal window (left) and in the $X_{\rm mis}$ sidebands (right). The full (blank) histograms show the expected background contribution from fake (true) D_s. The histograms are superimposed additively.

TABLE II: Signal yields ($N_{D_s^{(*)}}$), reconstruction efficiencies ($\epsilon^{(*)}$), statistical significances (Σ) and branching fractions ($\mathcal{B}$). The errors on the signal yields are statistical, while for the branching fractions both statistical (first) and systematic (second) errors are provided. The correlation coefficient between N_{D_s} and $N_{D_s^*}$ equals -66%.

Mode	$N_{D_s^{(*)}}$	$\epsilon^{(*)}[\%]$	Σ	$\mathcal{B}\times 10^{-3}$
$B^-\to D_s^+K^-\ell^-\bar{\nu}_\ell$	84 ± 24	1.78	3.9	$0.30\pm 0.09^{+0.11}_{-0.08}$
$B^-\to D_s^{*+}K^-\ell^-\bar{\nu}_\ell$	41 ± 22	0.85	1.9	$0.29\pm 0.16^{+0.11}_{-0.10}$

tween the data and the phase space distribution. It is calculated as a function of the effective masses of the two-body subsystems $D_s^+K^-$, $D_s^+\ell^-$, and $K^-\ell^-$ and averaged using the experimentally observed distributions. We obtain $\mathcal{B}(B^-\to D_s^+K^-\ell^-\bar{\nu}_\ell)=(0.30\pm 0.09)\times 10^{-3}$ and $\mathcal{B}(B^-\to D_s^{*+}K^-\ell^-\bar{\nu}_\ell)=(0.29\pm 0.16)\times 10^{-3}$.

The dominant systematic uncertainty on the signal yield is due to the parameterization of the $X_{\rm mis}$ dependence of the signal and found to be $^{+23}_{-\ 6}(^{+7}_{-9})$ events for the $D_s(D_s^*)$ mode. It is evaluated by refitting the data with the parameters μ, σ, and α allowed to float, and by changing the integer parameter n by ± 1. Uncertainties in modeling the $X_{\rm mis}$ distributions of the background components containing true D_s are evaluated to be $^{+5}_{-7}$ $(^{+8}_{-7})$ events from fits with the background shape parameters varied by $\pm 1\sigma$, taking into account correlations between the parameters. We also repeat the fits with the parameters, whose values are determined from data (and which are fixed in the nominal fit), floating. The resulting uncertainty is $^{+4}_{-2}$ $(^{+0}_{-1})$ events. The effect of an imperfect estimation of the relative contributions of the signal components is determined to be ± 1 (± 1) from fits with the parameters $f_k^{(*)}$ varied by $\pm 1\sigma$ and taking into account a $\pm 3\%$ uncertainty on the photon reconstruction efficiency. The above uncertainties are summed in quadrature to obtain the total systematic uncertainty of the signal yield of $^{+24}_{-10}$ $(^{+12}_{-11})$ events for the D_s (D_s^*) modes. We include the effect of these uncertainties on the significance of the observed signals by convolving the likelihood function obtained in the fit with a Gaussian systematic error distribution. The significance of the signal in the $B^- \to D_s^+ K^- \ell^- \bar{\nu}_\ell$ ($B^- \to D_s^{*+} K^- \ell^- \bar{\nu}_\ell$) mode, after including systematic uncertainties, is 3.4σ (1.8σ).

In a similar way, we obtain a significance of 6σ for the combined $B^- \to D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell$ modes from the 2-dimensional ($X_{\rm mis}$, M_{D_s}) fit for the combined D_s and D_s^* samples. The much higher significance for the combined modes compared to the individual modes is due to the large cross-feed between the D_s and the D_s^* modes.

The uncertainty on the branching fractions, except for the systematic uncertainty of the signal yield, is evaluated to be 23.2% for each signal mode. It includes uncertainties in charged track reconstruction efficiency (6.6%), particle identification efficiency (3.9%), intermediate branching fractions (6.1%), number of B^+B^- pairs (1.5%) and the reconstruction efficiency correction $\Delta\epsilon_{cor}$ (21%).

The largest uncertainty, due to $\Delta\epsilon_{cor}$, is determined by calculating $\Delta\epsilon_{cor}$ in 10000 toy MC experiments. The width of a Gaussian function fitted to the obtained efficiencies is taken as systematic uncertainty. The uncertainties due to the intermediate branching fractions are taken from the errors quoted in [12]. Combining all uncertainties, we obtain $\mathcal{B}(B^- \to D_s^+ K^- \ell^- \bar{\nu}_\ell) = (0.30 \pm 0.09({\rm stat})^{+0.11}_{-0.08}({\rm syst})) \times 10^{-3}$, $\mathcal{B}(B^- \to D_s^{*+} K^- \ell^- \bar{\nu}_\ell) = (0.29 \pm 0.16({\rm stat})^{+0.11}_{-0.10}({\rm syst})) \times 10^{-3}$ and $\mathcal{B}(B^- \to D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell) = (0.59 \pm 0.12({\rm stat}) \pm 0.15({\rm syst})) \times 10^{-3}$ for the combined modes obtained in a similar way, taking correlations into account. Since the significance in the D_s^* mode does not exceed 3σ, we set an upper limit of $\mathcal{B}(B^- \to D_s^{*+} K^- \ell^- \bar{\nu}_\ell) < 0.56 \times 10^{-3}$ at the 90% confidence level, using the likelihood integration method.

In conclusion, we find evidence for the decay $B^- \to D_s^+ K^- \ell^- \bar{\nu}_\ell$ with a significance of 3.4σ and measure $\mathcal{B}(B^- \to D_s^+ K^- \ell^- \bar{\nu}_\ell) = (0.30 \pm 0.09({\rm stat})^{+0.11}_{-0.08}({\rm syst})) \times 10^{-3}$. The combined $B^- \to D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell$ decay modes are observed with a significance of 6σ to be $\mathcal{B}(B^- \to D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell) = (0.59 \pm 0.12({\rm stat}) \pm 0.15({\rm syst})) \times 10^{-3}$. The branching fraction results are consistent with the measurement of BaBar [6]. We also present the first measurement of the $D_s^+ K^-$ invariant mass distribution, which is dominated by a prominent peak around 2.6 GeV/c^2, possibly from excited D mesons decays.

We thank the KEKB group for excellent operation of the accelerator; the KEK cryogenics group for efficient solenoid operations; and the KEK computer group, the NII, and PNNL/EMSL for valuable computing and SINET4 network support. We acknowledge support from MEXT, JSPS and Nagoya's TLPRC (Japan); ARC and DIISR (Australia); NSFC (China); MSMT (Czechia); DST (India); INFN (Italy); MEST, NRF, GSDC of KISTI, and WCU (Korea); MNiSW and NCN (Poland); MES and RFAAE (Russia); ARRS (Slovenia); SNSF (Switzerland); NSC and MOE (Taiwan); and DOE and NSF (USA).

[1] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963); M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[2] D. Asner *et al.* (Heavy Flavor Averaging Group), arXiv:1010.1589v3 [hep-ex].
[3] Throughout this paper, the inclusion of the charge-conjugate decay mode is implied.
[4] P. del Amo Sanchez *et al.* (BaBar Collaboration), Phys. Rev. D **82**, 111101(R) (2010).
[5] F. U. Bernlochner, Z. Ligeti, S. Turczyk, arXiv:1202.1834 [hep-ph].
[6] P. Amo del Sanchez *et al.* (BaBar Collaboration), Phys. Rev. Lett. **107**, 041804 (2011).
[7] S. Kurokawa and E. Kikutani, Nucl. Instr. and. Meth. A **499**, 1 (2003), and other papers included in this volume.
[8] A. Abashian *et al.* (Belle Collaboration), Nucl. Instr. and Meth. A **479**, 117 (2002).
[9] D. J. Lange, Nucl. Instr. and Meth. A **462**, 152 (2001).
[10] D. Scora and N. Isgur, Phys. Rev. D **52**, 2783 (1995).
[11] E. Barberio and Z. Wąs, Comput. Phys. Commun. **79**, 291 (1994).
[12] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010) and 2011 partial update for the 2012 edition.
[13] A. Matyja *et al.* (Belle Collaboration), Phys. Rev. Lett. **99**, 191807 (2007).
[14] B. Aubert *et al.* (BaBar Collaboration), Phys. Rev. Lett. **100**, 171803 (2008); J. Wiechczynski *et al.* (Belle Collaboration), Phys. Rev. D **80**, 052005 (2009).