STUDY OF THE $B^- \to D_s^{(\star)+} K^- \ell^- \overline{\nu}_\ell$ DECAYS AT BELLE*

K. Adamczyk

The Henryk Niewodniczański Institute of Nuclear Physics
Polish Academy od Sciences
Radzikowskiego 152, 31-342 Kraków, Poland

(Received May 22, 2012)

Preliminary results containing measurements of the decays $B^- \to D_s^{(\star)+} K^- \ell^- \overline{\nu}_\ell$ are presented (the inclusion of the charge-conjugate modes is implied). Analysis is based on data sample of 656 $B\overline{B}$ collected at Belle detector in the clean environment of KEKB asymmetric-energy e^+e^- collider. Combined $B \to D_s^{(*)} K \ell \nu_\ell$ modes were observed with significance of 6σ. First time these modes were measured separately: $\mathcal{B}(B \to D_s K \ell \nu_\ell) = [3.0 \pm 1.2(\mathrm{stat})^{+1.1}_{-0.8}(\mathrm{syst})] \times 10^{-4}$ (3.4σ, first evidence), $\mathcal{B}(B \to D_s^* K \ell \nu_\ell) < 5.4 \times 10^{-4}$ C.L. $= 90\%$. Due to the fact that analysis is model-independent it allows for first measurement of the $m(D_s K)$ spectrum, which is dominated by a pronounced peak around $2.6\,\mathrm{GeV}/c^2$.

DOI:10.5506/APhysPolB.43.1391
PACS numbers: 13.20.He, 14.40.Nd

1. Introduction

Searching for new exclusive semileptonic B decays with $b \to c$ transition is motivated by several open questions in this field. It can be noticed that known exclusive decays do not sum up to total inclusive branching fraction $B \to X_c \ell \nu$. Furthermore, there are discrepancies between measurements and theoretical expectations for semileptonic B decays to excited charmed resonances. These both issues affect the accuracy of $|V_{ub}|$ and $|V_{cb}|$ determination.

B decays to $D_s K$ system are interesting candidates for the missing semileptonic modes. $D_s K$ system allows for exploration of masses $m(D_s K) > 2.46\,\mathrm{GeV}/c^2$, where resonant and non-resonant contributions are expected. Observation of such decays has an impact on background description for many important processes, *e.g.* $B_s \to D_s X \ell \nu$.

* Presented at the Cracow Epiphany Conference on Present and Future of B Physics, Cracow, Poland, January 9–11, 2012.

Recently, BaBar reported first observation of combined D_s and $D_s^\star$ modes with a branching fraction of $\mathcal{B}(B^- \to D_s^{(\star)+} K^- \ell^- \overline{\nu}_\ell) = [6.13^{+1.04}_{-1.03}(\mathrm{stat}) \pm 0.43(\mathrm{syst}) \pm 0.51(\mathcal{B}(D_s))] \times 10^{-4}$ [1].

2. Event reconstruction

The analysis is based on a data sample consisting of 657×10^6 $B\bar{B}$ pairs that were collected with the Belle detector [2] at the KEKB asymmetric e^+e^- collider operating at the $\Upsilon(4S)$ resonance [3]. D_s^+ candidates are reconstructed in the cleanest decay chain: $D_s^+ \to \phi\pi^+$, $\phi \to K^+K^-$ ($2.32 \pm 0.14\%$ of D_s width). D_s^+ candidates are combined with photons with an energy $E_\gamma > 125$ MeV to find $D_s^{\star+}$ candidates. $D_s(D_s^\star)$ candidates with an invariant mass in the range $1.934 < m_{D_s} < 2.003\,\mathrm{GeV}/c^2$ ($2.079 < m_{D_s^\star} < 2.155\,\mathrm{GeV}/c^2$) are accepted for further analysis. The signal windows are defined as: $1.954 < m_{D_s} < 1.982\,\mathrm{GeV}/c^2$ and $2.079 < m_{D_s^\star} < 2.255\,\mathrm{GeV}/c^2$. 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^+. In the case of multiple B_{sig} candidates, the one with the greatest confidence level of the vertex fit is chosen. Events with accepted $D_s^{\star+}K^-\ell^-$ candidates ($D_s^\star$ sample) are removed from the set of $D_s^+K^-\ell^-$ candidates (D_s sample).

Signal events are identified using the variable X_{mis} [4] defined as: $X_{\mathrm{mis}} \equiv (E_{\mathrm{beam}} - E_{\mathrm{vis}} - |\overrightarrow{p}_{\mathrm{vis}}|)/\sqrt{E_{\mathrm{beam}}^2 - m_{B^\pm}^2}$, where E_{beam} is the beam energy, E_{vis} and $\overrightarrow{p}_{\mathrm{vis}}$ denote the total energy and momentum of the $D_sK\ell$ system, respectively, and $m_{B^\pm}$ is the $B^\pm$ mass. All variables are calculated in CM of $\Upsilon(4S)$. 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.

2.1. Background suppression

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 select events with a negatively charged lepton with a momentum above $0.5\,\mathrm{GeV}/c$ on the tagging side. This reduces the main background, where a D_s^+ produced in a decay of the type $B \to D_s^{(\star)+}\overline{D}^{(\star)}$ is combined with a lepton and kaon from the subsequent D decay from a semileptonic decay $\overline{B} \to \ell^-\overline{\nu}_\ell D^{(\star)}X$ of the accompanying $\overline{B}$ meson. Further improvement of the sensitivity is achieved with two tagging side variables:

$$M^c_{\rm tag} \equiv \sqrt{\left(E_{\rm tag} - E^\ell_{\rm tag}\right)^2 - \left(\overrightarrow{p}_{\rm tag} - \overrightarrow{p}^\ell_{\rm tag}\right)^2}\,,$$
$$X_{\rm tag} \equiv \left(E_{\rm beam} - E_{\rm tag} - |\overrightarrow{p}_{\rm tag}|\right)/\sqrt{E^2_{\rm beam} - M^2_B}\,,$$

where $E_{\rm tag}$ and $\overrightarrow{p}_{\rm tag}$ denote the total energy and momentum of reconstructed particles not assigned to $B_{\rm sig}$. $E^\ell_{\rm tag}$ and $\overrightarrow{p}^\ell_{\rm tag}$ represent the energy and momentum of prompt tagging lepton. $M^c_{\rm tag}$ represents the inclusively reconstructed mass of the hadronic system produced in the $B_{\rm tag}$ decay and $X_{\rm tag}$ is the tagging side equivalent of $X_{\rm mis}$. The $M^c_{\rm tag}$ and $X_{\rm tag}$ distributions for signal and background are shown in Fig. 1.

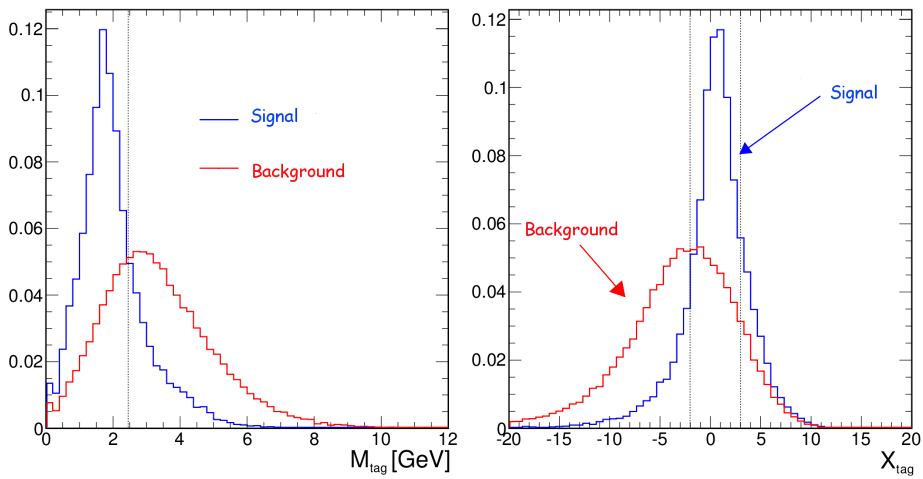

Fig. 1. $M^c_{\rm tag}$ (left) and $X_{\rm tag}$ (right) distributions for signal and background MC.

In the analysis we require $-2 < X_{\rm tag} < 3$, $M^c_{\rm tag} < 2.4\,{\rm GeV}/c^2$, and zero total event charge. The selection criteria are optimized for the D_s mode by maximizing the expected statistical significance estimated as $N_{\rm S}/\sqrt{N_{\rm S} + N_{\rm B}}$, where $N_{\rm S}(N_{\rm B})$ is the predicted number of signal (background) events in the $(X_{\rm mis},\, m_{D_s})$ signal window. This optimization was carried out for signal branching fractions in the range $\mathcal{B}(B \to D_s^{(*)} K \ell \nu) = 2.5$–$5.0 \times 10^{-4}$.

2.2. Background model

$N_{\rm B}$ 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 $D_s^\star$ sample, the background with true D_s is split into two parts: "true $D_s^\star$" with properly reconstructed $D_s^{\star +}$ and "fake $D_s^\star$", 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 $X_{\rm mis} > 1$. At this stage of analysis, signal window is blinded.

2.3. Signal in data

After defining the selection criteria, the signal region is unblinded. The resulting $X_{\rm mis}$, $m_{D_s^{(\star)}}$, $M_{D_s^+ K^-}$ distributions in data are shown in Fig. 2. $M_{D_s^+ K^-}$ is the invariant mass distribution of the $D_s^+ K^-$ system for the combined D_s and $D_s^\star$ samples in the signal window. While the background model describes the experimental distributions well in the $X_{\rm mis}$ sidebands, a clear excess over the expected background is seen in the signal region. The $M_{D_s^+ K^-}$ distribution in the signal window is dominated by a prominent peak at $\approx 2.6\,{\rm GeV}/c^2$, similarly to that observed in $B^- \to D_s^+ K^- \pi^-$ decays [5].

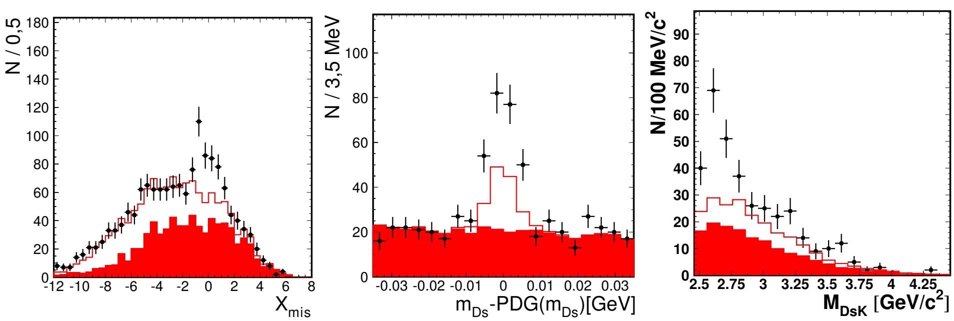

Fig. 2. $X_{\rm mis}$, $m_{D_s} - {\rm PDG}(m_{D_s})$, $M_{D_s K}$ distributions in data.

2.4. Signal extraction

The signal yield are extracted from a simultaneous, extended unbinned maximum likelihood fit to the D_s and $D_s^\star$ samples. The D_s and $D_s^\star$ samples are fitted in two ($X_{\rm mis}$, m_{D_s}) and three ($X_{\rm mis}$, m_{D_s}, $m_{D_s^\star}$) dimensions, respectively. In the D_s sample, we consider two signal components coming from the decay $B^- \to D_s^+ K^- \ell^- \overline{\nu}_\ell$ and from the decay $B^- \to D_s^{\star +} K^- \ell^- \overline{\nu}_\ell$ if a photon from the $D_s^{\star +}$ has been missed. In the $D_s^\star$ sample, we distinguish three signal components: one coming from the $B^- \to D_s^+ K^- \ell^- \overline{\nu}_\ell$ channel, where D_s meson is associated with a random photon, and two from the $B^- \to D_s^{\star +} K^- \ell^- \overline{\nu}_\ell$ mode, with true and fake $D_s^\star$ defined analogous to the background case discussed above.

The two (three)-dimensional PDFs are parameterized as the product of two (three) one-dimensional PDFs for each variable. The components with true $D_s^{(\star)}$ are parameterized as a sum of two Gaussian functions in m_{D_s} or as a single Gaussian function in $m_{D_s^\star}$, with means set to the average $D_s^{(\star)}$ mass values [6] and with the remaining parameters fixed from fits to control samples in data. The components with fake $D_s^{(\star)}$ are parameterized as first degree polynomials in $m_{D_s^{(\star)}}$. The $X_{\rm mis}$ distribution of the signal components is modeled with two line shapes, one describing the two components of the $B^- \to D_s^+ K^- \ell^- \overline{\nu}_\ell$ mode and the other one describing the three components of the $B^- \to D_s^{\star +} K^- \ell^- \overline{\nu}_\ell$ 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 $n \in N$ are fixed from fits to the signal MC samples. The $X_{\rm mis}$ distributions of the background components are parametrized as bifurcated Gaussian functions with parameters fixed from the simulated $B\overline{B}$ 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 signal and background yields, and the coefficients of the polynomials describing the $m_{D_s^{(\star)}}$ dependence of the fake $D_s^{(\star)}$.

3. Results

The signal yields extracted from the fit are presented in Table I. 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. Fit projections for each variable plotted in signal windows of the other variables are presented in Fig. 3. The fitted signal yields are used to compute the branching fractions with the formula: $\mathcal{B}(B^- \to D_s^{(\star)+} K^- \ell^- \overline{\nu}_\ell) = N_{D_s^{(\star)}}/(2N_{B^+B^-}\epsilon^{(\star)}\mathcal{B}_{\rm int})$, where $N_{B^+B^-}$ is the number of B^+B^- pairs in data, $\epsilon^{(\star)}$ denotes the reconstruction efficiency of the signal decay chain and $\mathcal{B}_{\rm int}$ is the product of intermediate branching fractions.

TABLE I

Signal yields ($N_{D_s^{(\star)}}$), branching fractions ($\mathcal{B}$), statistical significances (Σ).

Decay channel	$N_{D_s^{(*)}}$	$\mathcal{B}$	Σ
$D_s K \ell \nu$	84 ± 24	$[3.0 \pm 1.2({\rm stat})^{+1.1}_{-0.8}({\rm syst})] \times 10^{-4}$	3.4σ
$D_s^* K \ell \nu$	41 ± 22	$[2.9 \pm 1.6({\rm stat})^{+1.1}_{-1.0}({\rm syst})] \times 10^{-4}$ $< 5.4 \times 10^{-4}$ C.L. $= 90\%$	1.8σ
combined $D_s^{(*)} K \ell \nu$			6σ

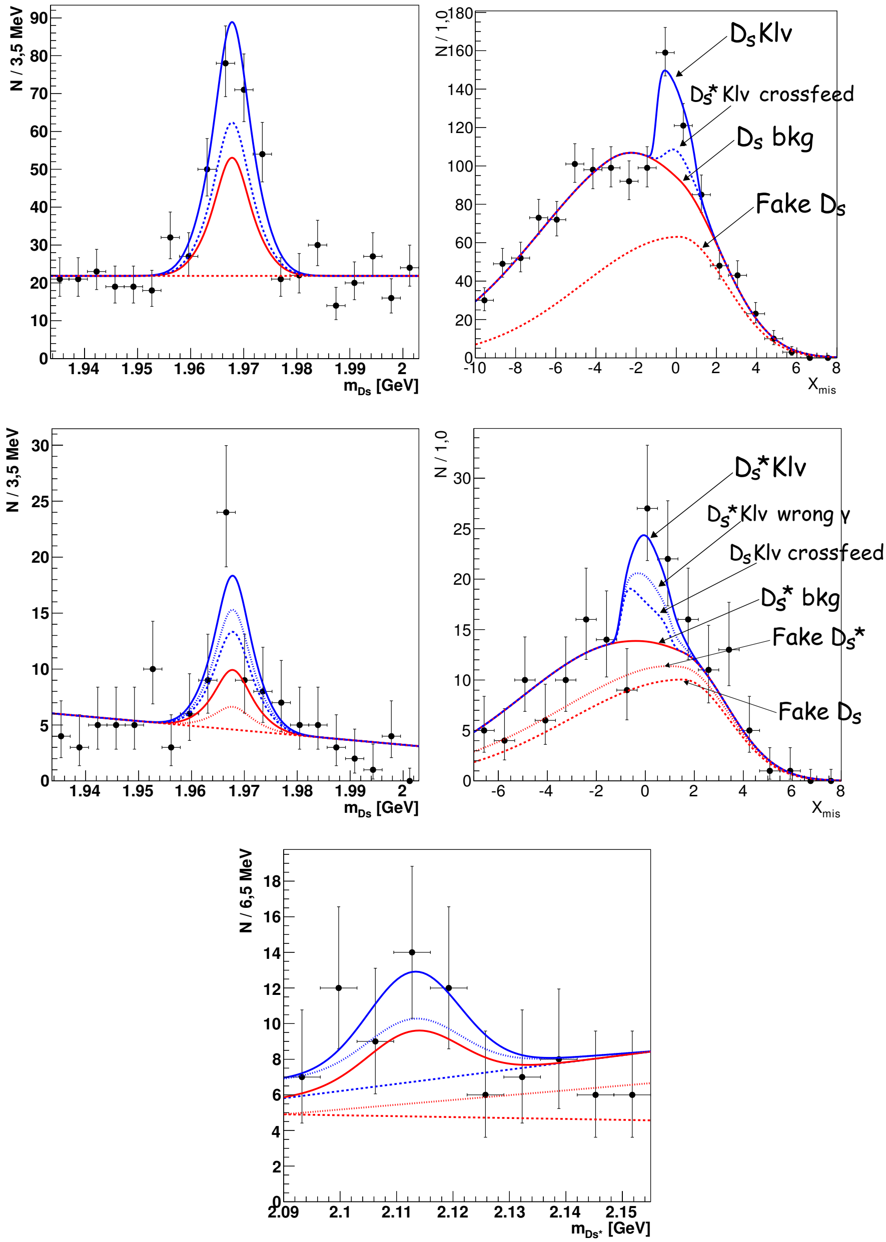

Fig. 3. $X_{\rm mis}$, m_{D_s} distributions in the $m_{D_s^{(\star)}}$ and $X_{\rm mis}$ signal window for D_s (first row) and $D_s^\star$ (second row) sample. In the third row, there is $m_{D_s^\star}$ distribution in the m_{D_s} and $X_{\rm mis}$ signal window for $D_s^\star$ sample.

Systematic uncertainties are presented in Table II. The dominant systematic uncertainty on the signal yield is due to the parametrization of the $X_{\rm mis}$ dependence of the signal and found to be $^{+27}_{-7}(^{+17}_{-22})$ events for the $D_s(D_s^\star)$. Systematic uncertainty of signal reconstruction efficiency also gives significant contribution (21%) to the total uncertainty. The reconstruction efficiency is expressed as $\epsilon^{(\star)} = \epsilon^{(\star)}_{\rm PS}\Delta\epsilon^{(\star)}_{\rm cor}$, where $\epsilon^{(\star)}_{\rm PS}$ is the efficiency calculated from the signal MC with the phase space model and $\Delta\epsilon^{(\star)}_{\rm cor} = 1.20(0.57)$ corrects for the difference between the data and the phase space distribution. It is calculated as a function of the effective masses of two-body subsystems $D_s^{(\star)+}K^-$, $D_s^{(\star)+}\ell^-$, $K^-\ell^-$, and averaged using the experimentally observed distributions.

TABLE II

Systematic uncertainties.

Source	$\Delta\mathcal{B}(D_s)$%	$\Delta\mathcal{B}(D_s^*)$%
Tracking, KID, LeptID	8	
$\mathcal{B}(D_s \to \phi\pi)$	6	
Signal efficiency	21	
$N(B^+B^-)$	2	
Signal PDF (MC)	+27, −7	+17, −22
BKG PDF (MC)	+6, −8	+20, −17
BKG PDF (Data)	+5, −1	3
Cross feed	1	2

REFERENCES

[1] P. Amo del Sanchez *et al.* [BaBar Collaboration], *Phys. Rev. Lett.* **107**, 041804 (2011).

[2] A. Abashian *et al.* [Belle Collaboration], *Nucl. Instrum. Methods* **A479**, 117 (2002).

[3] S. Kurokawa, E. Kikutani, *Nucl. Instrum. Methods* **A499**, 1 (2003) and other papers included in this volume.

[4] A. Matyja *et al.* [Belle Collaboration], *Phys. Rev. Lett.* **99**, 191807 (2007).

[5] B. Aubert *et al.* [BaBar Collaboration], *Phys. Rev. Lett.* **100**, 171803 (2008); J. Wiechczynski *et al.* [Belle Collaboration], *Phys. Rev.* **D80**, 052005 (2009).

[6] K. Nakamura *et al.* [Particle Data Group], *J. Phys. G* **37**, 075021 (2010) and 2011 partial update for the 2012 edition.