Measurement of the branching fractions of $\overline{B} \to D^{(*)} K^- K_{(S)}^{(*)0}$ and $\overline{B} \to D^{(*)} D_s^-$ decays at Belle II

The Belle II collaboration

I. Adachi, L. Aggarwal, H. Aihara, N. Akopov, A. Aloisio, N. Althubiti,
N. Anh Ky, D. M. Asner, H. Atmacan, T. Aushev, V. Aushev,
M. Aversano, R. Ayad, V. Babu, H. Bae, S. Bahinipati, P. Bambade,
Sw. Banerjee, S. Bansal, M. Barrett, J. Baudot, A. Baur, A. Beaubien,
F. Becherer, J. Becker, J. V. Bennett, F. U. Bernlochner, V. Bertacchi,
M. Bertemes, E. Bertholet, M. Bessner, S. Bettarini, B. Bhuyan,
F. Bianchi, L. Bierwirth, T. Bilka, D. Biswas, A. Bobrov, D. Bodrov,
A. Bolz, A. Boschetti, A. Bozek, M. Bračko, P. Branchini, R. A. Briere,
T. E. Browder, A. Budano, S. Bussino, Q. Campagna, M. Campajola,
L. Cao, G. Casarosa, C. Cecchi, J. Cerasoli, M.-C. Chang, P. Chang,
P. Cheema, B. G. Cheon, K. Chilikin, K. Chirapatpimol, H.-E. Cho,
K. Cho, S.-J. Cho, S.-K. Choi, S. Choudhury, L. Corona, J. X. Cui,
F. Dattola, E. De La Cruz-Burelo, S. A. De La Motte, G. de Marino,
G. De Nardo, M. De Nuccio, G. De Pietro, R. de Sangro, M. Destefanis,
S. Dey, R. Dhamija, A. Di Canto, F. Di Capua, J. Dingfelder, Z. Doležal,
I. Domínguez Jiménez, T. V. Dong, M. Dorigo, D. Dorner, K. Dort,
D. Dossett, S. Dreyer, S. Dubey, K. Dugic, G. Dujany, P. Ecker,
M. Eliachevitch, D. Epifanov, P. Feichtinger, T. Ferber, T. Fillinger,
C. Finck, G. Finocchiaro, A. Fodor, F. Forti, A. Frey, B. G. Fulsom,
M. Garcia-Hernandez, R. Garg, G. Gaudino, V. Gaur, A. Gaz, A. Gellrich,
G. Ghevondyan, D. Ghosh, H. Ghumaryan, G. Giakoustidis, R. Giordano,
A. Giri, A. Glazov, B. Gobbo, R. Godang, O. Gogota, P. Goldenzweig,
W. Gradl, E. Graziani, D. Greenwald, Z. Gruberová, T. Gu, K. Gudkova,
I. Haide, S. Halder, Y. Han, T. Hara, C. Harris, K. Hayasaka,
H. Hayashii, S. Hazra, C. Hearty, M. T. Hedges, A. Heidelbach,
I. Heredia de la Cruz, M. Hernández Villanueva, T. Higuchi, M. Hoek,

M. Hohmann, P. Horak, C.-L. Hsu, T. Humair, T. Iijima, K. Inami,
N. Ipsita, A. Ishikawa, R. Itoh, M. Iwasaki, W. W. Jacobs, D. E. Jaffe,
E.-J. Jang, S. Jia, Y. Jin, A. Johnson, K. K. Joo, H. Junkerkalefeld,
A. B. Kaliyar, J. Kandra, K. H. Kang, S. Kang, G. Karyan, T. Kawasaki,
F. Keil, C. Kiesling, C.-H. Kim, D. Y. Kim, K.-H. Kim, Y.-K. Kim,
H. Kindo, K. Kinoshita, P. Kodyš, T. Koga, S. Kohani, K. Kojima,
T. Konno, A. Korobov, S. Korpar, E. Kovalenko, R. Kowalewski, P. Križan,
P. Krokovny, T. Kuhr, Y. Kulii, J. Kumar, M. Kumar, R. Kumar,
K. Kumara, T. Kunigo, A. Kuzmin, Y.-J. Kwon, S. Lacaprara, K. Lalwani,
T. Lam, J. S. Lange, M. Laurenza, K. Lautenbach, R. Leboucher,
F. R. Le Diberder, M. J. Lee, C. Lemettais, P. Leo, D. Levit, P. M. Lewis,
L. K. Li, S. X. Li, Y. Li, Y. B. Li, J. Libby, Z. Liptak, M. H. Liu,
Q. Y. Liu, Z. Q. Liu, D. Liventsev, S. Longo, T. Lueck, C. Lyu, Y. Ma,
M. Maggiora, S. P. Maharana, R. Maiti, S. Maity, G. Mancinelli,
R. Manfredi, E. Manoni, M. Mantovano, D. Marcantonio, S. Marcello,
C. Marinas, C. Martellini, A. Martens, A. Martini, T. Martinov,
L. Massaccesi, M. Masuda, K. Matsuoka, D. Matvienko, S. K. Maurya,
J. A. McKenna, F. Meier, M. Merola, C. Miller, M. Mirra, S. Mitra,
K. Miyabayashi, R. Mizuk, G. B. Mohanty, S. Mondal, S. Moneta,
H.-G. Moser, M. Mrvar, R. Mussa, I. Nakamura, M. Nakao, Y. Nakazawa,
M. Naruki, D. Narwal, Z. Natkaniec, A. Natochii, L. Nayak, M. Nayak,
G. Nazaryan, M. Neu, M. Niiyama, S. Nishida, S. Ogawa, Y. Onishchuk,
H. Ono, G. Pakhlova, S. Pardi, K. Parham, H. Park, J. Park,
S.-H. Park, B. Paschen, A. Passeri, S. Patra, S. Paul, T. K. Pedlar,
R. Peschke, R. Pestotnik, M. Piccolo, L. E. Piilonen, G. Pinna Angioni,
P. L. M. Podesta-Lerma, T. Podobnik, S. Pokharel, C. Praz, S. Prell,
E. Prencipe, M. T. Prim, H. Purwar, P. Rados, G. Raeuber, S. Raiz,
N. Rauls, M. Reif, S. Reiter, M. Remnev, L. Reuter, I. Ripp-Baudot,
G. Rizzo, M. Roehrken, J. M. Roney, A. Rostomyan, N. Rout,
S. Sandilya, L. Santelj, Y. Sato, V. Savinov, B. Scavino, C. Schmitt,
S. Schneider, M. Schnepf, C. Schwanda, Y. Seino, A. Selce, K. Senyo,
J. Serrano, M. E. Sevior, C. Sfienti, W. Shan, C. Sharma, C. P. Shen,
X. D. Shi, T. Shillington, T. Shimasaki, J.-G. Shiu, D. Shtol,
A. Sibidanov, F. Simon, J. B. Singh, J. Skorupa, R. J. Sobie, M. Sobotzik,
A. Soffer, A. Sokolov, E. Solovieva, S. Spataro, B. Spruck, M. Starič,
P. Stavroulakis, S. Stefkova, R. Stroili, M. Sumihama, H. Svidras,
M. Takizawa, U. Tamponi, S. Tanaka, K. Tanida, F. Tenchini, A. Thaller,
O. Tittel, R. Tiwary, D. Tonelli, E. Torassa, K. Trabelsi, I. Ueda,
T. Uglov, K. Unger, Y. Unno, K. Uno, S. Uno, Y. Ushiroda,
S. E. Vahsen, R. van Tonder, K. E. Varvell, M. Veronesi, A. Vinokurova,
V. S. Vismaya, L. Vitale, V. Vobbilisetti, R. Volpe, A. Vossen, B. Wach,
M. Wakai, S. Wallner, E. Wang, M.-Z. Wang, Z. Wang, A. Warburton,
M. Watanabe, S. Watanuki, C. Wessel, J. Wiechczynski, E. Won,

X. P. Xu [ID], **B. D. Yabsley** [ID], **S. Yamada** [ID], **S. B. Yang** [ID], **J. Yelton** [ID], **J. H. Yin** [ID],
Y. M. Yook [ID], **K. Yoshihara** [ID], **C. Z. Yuan** [ID], **L. Zani** [ID], **F. Zeng** [ID], **B. Zhang** [ID],
V. Zhilich [ID], **J. S. Zhou** [ID], **Q. D. Zhou** [ID], **V. I. Zhukova** [ID] **and R. Žlebčík** [ID]

https://belle.kek.jp/

E-mail: coll-publications@belle2.org

ABSTRACT: We present measurements of the branching fractions of eight $\overline{B}^0 \to D^{(*)+}K^-K^{(*)0}_{(S)}$, $B^- \to D^{(*)0}K^-K^{(*)0}_{(S)}$ decay channels. The results are based on data from SuperKEKB electron-positron collisions at the $\Upsilon(4S)$ resonance collected with the Belle II detector, corresponding to an integrated luminosity of 362 fb^{-1}. The event yields are extracted from fits to the distributions of the difference between expected and observed B meson energy, and are efficiency-corrected as a function of $m(K^-K^{(*)0}_{(S)})$ and $m(D^{(*)}K^{(*)0}_{(S)})$ in order to avoid dependence on the decay model. These results include the first observation of $\overline{B}^0 \to D^+K^-K^0_S$, $B^- \to D^{*0}K^-K^0_S$, and $\overline{B}^0 \to D^{*+}K^-K^0_S$ decays and a significant improvement in the precision of the other channels compared to previous measurements. The helicity-angle distributions and the invariant mass distributions of the $K^-K^{(*)0}_{(S)}$ systems are compatible with quasi-two-body decays via a resonant transition with spin-parity $J^P = 1^-$ for the $K^-K^0_S$ systems and $J^P = 1^+$ for the K^-K^{*0} systems. We also present measurements of the branching fractions of four $\overline{B}^0 \to D^{(*)+}D^-_s$, $B^- \to D^{(*)0}D^-_s$ decay channels with a precision compatible to the current world averages.

KEYWORDS: B Physics, Branching fraction, e^+-e^- Experiments, Particle and Resonance Production

Contents

1 Introduction

Knowledge of B meson hadronic decays is limited: about 40% of the total B width is not measured in terms of exclusive branching fractions ($\mathcal{B}$). Hence, unmeasured decays are usually simulated using the PYTHIA fragmentation model [1], which does not properly describe B decays. The inclusive branching fraction for B meson decays into a $D^{(*)}$ meson accompanied by a kaon pair and possibly pions could account for 6% of the B width, according to PYTHIA, However, only a small fraction of the exclusive components has been measured [2]. Improving the knowledge of these decay channels can be of significant benefit to understanding the background contamination of many analyses of hadronic B decays. The unknown fraction of the total B width is spread across many exclusive channels, therefore improvements are not expected from single branching fraction measurements, but require the systematic exploration of the large unmeasured or poorly measured channels. This paper is part of a program of measuring the missing components in the B decay width.

Assuming factorization [3], three-body $B \to DK^- K$ decays proceed via a tree-level amplitude with an external W boson emission and the production of an $s\bar{s}$ pair. The symbol B indicates a neutral or charged B meson, the symbol K indicates a K^0_S or a K^{*0} meson, while the symbol D indicates a $D^{(*)+,0}$ meson, from now on.[1] However, the quasi-two-body

---

[1]We use natural units $\hbar = c = 1$ and charge conjugation is implied throughout.

mechanism, $B \to DX^-(\to K^-K)$, with an intermediate resonance X^- that decays strongly, is also possible. Several theoretical studies have been performed to interpret these decays in term of conventional [4, 5] or exotic [6, 7] resonances. The factorization assumption forbids the production of resonances with spin larger than one. Assuming factorization and exact isospin symmetry, a spin-parity assignment $J^P = 1^-$ is expected for the $K^-K_S^0$ system [8]. If isopin symmetry (or factorization) is not assumed $J^P = 0^+$ states are also allowed. On the other hand, for the K^-K^{*0} system the spin-parity assignments $J^P = 0^-$, 1^-, or 1^+ are allowed.

A search for $B \to DK^-K$ decays was previously performed by the Belle experiment, on a sample with an integrated luminosity of 29 fb^{-1} [2]. Four $B \to DK^-K^{*0}$ channels and the $B^- \to D^0K^-K_S^0$ mode were observed, while excesses in the remaining K_S^0 channels were seen with significances of about 2.5 standard deviations.[2] A study of the $m(K^-K^{*0})$ invariant mass distribution was also reported. A dominant transition via a $J^P = 1^+$ state was claimed for $B \to DK^-K^{*0}$ decays, and the intermediate state was interpreted as a $a_1(1260)^-$ resonance. Due to the small sample size, no conclusive claims were made for $B \to DK^-K_S^0$ decays. However, the $m(K^-K_S^0)$ invariant mass and angular distributions suggested the presence of a resonant component with $J^P = 1^-$ in $B \to DK^-K_S^0$ decays. This was interpreted as a ρ-like resonance, but the $\rho(770)$ was disfavored given the limited available phase-space.

A potential application of an improvement in understanding the $B \to DK^-K$ sector is an amelioration of the B-tagging algorithms used at B-factories. At an energy-asymmetric e^+e^- collider such as SuperKEKB, $B^0\overline{B}^0$ and B^+B^- meson pairs ($B\overline{B}$ pairs) are produced at threshold from $\Upsilon(4S)$ meson decays. A large part of the Belle II physics program [9] relies on identifying the partner B meson produced in association with the signal B meson to infer the properties of the signal (B-tagging). In particular, in hadronic B-tagging the partner B meson is fully reconstructed to infer the kinematic properties of the signal using initial-state constraints. The Belle II B-tagging algorithm, Full Event Interpretation (FEI), is based on a set of multivariate classifiers trained on the Belle II Monte Carlo (MC) simulation [10]. A mismodeling in the simulation may introduce biases and uncertainties in the FEI efficiencies, and leads to suboptimal FEI performance, degrading Belle II's physics reach. Three-body $B \to DK^-K$ decay channels are currently not used by the FEI algorithm. However, the high purity of these decays [2] makes them ideal candidates to improve the B-tagging efficiency. Moreover, information on the branching fractions alone is insufficient, and knowledge of the final-state kinematic properties and the intermediate states is essential for an accurate description of these decays during FEI training. In conclusion, additional measurements of $B \to DK^-K$ decays and their dynamics can lead to an improvement of the FEI efficiency and its background rejection.

A more precise measurement of eight $B \to DK^-K$ absolute branching fractions including observation of the three unobserved $B \to DK^-K_S^0$ channels is reported in this work, using the $362\,\text{fb}^{-1}$ Belle II data sample collected at the $\Upsilon(4S)$ resonance. This analysis aims to better understand the intermediate states in $B \to DK^-K$ decays by investigating their Dalitz plots [11]. In particular, the $m(K^-K_S^0)$ and $m(K^-K^{*0})$ distributions and the angular distributions of these systems are studied.

---

[2]We use "K_S^0 channels" to refer to all the channels where a K_S^0 is present; similarly we use "K^{*0}", "D^0", "D^+", "D^{*0}" and "D^{*+} channels".

Measurements of the branching fractions of the four $B \to DD_s^-$ decays, which are reconstructed in the same final states as the $B \to DK^-K$ channels, are also included in this work. These channels are often adopted as control or normalization channels in several measurements [12–15] and improvements in $\mathcal{B}(B \to DD_s^-)$ precision will reduce their uncertainties.

The branching fractions are extracted independently for each channel. After the event reconstruction and selection, a fit to the distribution of the difference between expected and observed B-meson energy is performed to separate the signal from the backgrounds. In $B \to DK^-K$ channels, the sPlot procedure [16] is used to obtain a background-subtracted $(m(K^-K), m(DK))$ two-dimensional invariant mass distribution. The signal efficiency is evaluated using a simulated signal sample, as a function of $(m(K^-K), m(DK))$. The branching fraction is then obtained from the efficiency-corrected integral of the $(m(K^-K), m(DK))$ distribution. The $B \to DD_s^-(\to K^-K)$ branching fractions are extracted using the same strategy, but the efficiency correction is applied directly to the yield obtained from the ΔE fit, since there is no model dependence in these channels. We use simulation and data control samples to optimize the analysis. In particular, the $B \to DD_s^-(\to K^-K)$ channels, along with the $B^- \to D^{*0}\pi^-$ control channel, are also used to validate the $B \to DK^-K$ analysis and assess some of the systematic uncertainties. To reduce experimenter's bias, data containing signal candidates are only analyzed after the entire analysis procedure is finalized.

The paper is organized as follows. Section 2 gives a brief description of the Belle II detector and analysis sample. Section 3 discusses the reconstruction and event selection. Section 4 and section 5 describe the signal yield extraction and efficiency estimation, respectively. Section 6 and section 7 present the measured branching fractions and the related systematic uncertainties, respectively. In section 8 discusses the $m(K^-K)$ invariant mass and angular distributions.

2 The Belle II detector and samples

The Belle II experiment [17] is located at SuperKEKB, which collides electrons and positrons at and near the $\Upsilon(4S)$ resonance [18]. The Belle II detector has a cylindrical geometry with the symmetry axis, defined as the z axis, almost coincident with the direction of the electron beam. The detector includes a two-layer silicon-pixel detector surrounded by a four-layer double-sided silicon-strip detector [19] and a 56-layer central drift chamber (CDC). These detectors reconstruct trajectories of charged particles (tracks). Only one sixth of the second layer of the pixel detector was installed for the data analyzed here. Surrounding the CDC, which also provides dE/dx energy-loss measurements, is a time-of-propagation counter [20] in the central region and an aerogel-based ring-imaging Cherenkov counter in the forward region. These detectors provide charged-particle identification. Surrounding these subsystems is an electromagnetic calorimeter based on CsI(Tl) crystals that primarily provides energy and timing measurements for photons and electrons. Outside of the calorimeter is a superconducting solenoid magnet. The solenoid magnet provides a 1.5 T magnetic field that is parallel to the z axis. Its flux return is instrumented with resistive-plate chambers and plastic scintillator modules to detect muons, K_L^0 mesons, and neutrons.

Belle II integrates a hardware trigger and an online software event selection to suppress background processes and to select $b\bar{b}$, $c\bar{c}$, and $\tau^+\tau^-$ events with high efficiency.

The data sample of Belle II collected from 2019 to 2022 at the energy of the $\Upsilon(4S)$ is used in this analysis. The total integrated luminosity of the sample is $\mathcal{L}_{\text{int}} = (362 \pm 2)\,\text{fb}^{-1}$, which corresponds to $(387 \pm 6) \times 10^6$ $B\overline{B}$ pairs.

Simulated samples of exclusive decay channels are used to determine signal efficiencies, to define fit models, and to evaluate the systematic uncertainties. These samples have an integrated luminosity equivalent to at least 50 times the data. A simulated sample that reproduces the composition of Belle II events, including $B\overline{B}$ and $e^+e^- \to q\bar{q}$ continuum (where q indicates an u, d, s or c quark) backgrounds, and equivalent to an integrated luminosity of $1\,\text{ab}^{-1}$, is used to investigate the sample composition and validate the analysis before examining the signal region in data.

Simulated events are generated using KKMC generator for quark-antiquark production from e^+e^- collisions [21], PYTHIA8 generator for hadronization [1], EvtGen software package and PYTHIA8 generator for the decay of the generated hadrons [22, 23], and Geant4 software package for the detector response [24]. The simulation includes simulated beam-induced backgrounds [25]. The data and the MC simulations are processed using the Belle II analysis software [26, 27].

3 Reconstruction and event selection

Events are selected by the trigger based on the charged-particle multiplicity and total energy, to suppress low-multiplicity events.

Candidate charged pions and kaons are identified from their tracks, which are required to have transverse and longitudinal impact parameter calculated with respect to the interaction point, (IP) $d_r < 2$ cm and $|d_z| < 4$ cm, respectively, polar angle θ within the acceptance of the CDC ($17° \le \theta \le 150°$), and at least 20 measurement points (hits) in the CDC. The candidates are identified as pions or kaons based on the ratio between the particle-identification (PID) likelihood for the test hypothesis and the sum of the likelihoods for all other hypotheses. The PID selection efficiency is between 70% and 80% on average, with misidentification probability varying between 3% and 10%, depending on the transverse momentum p_T and $\cos\theta$ of the charged particle, both for kaon- and pion-enriched samples.

Photon candidates are selected by requiring calorimeter energy deposits (clusters) contain more than one crystal, have θ within the CDC acceptance, and a polar-angle-dependent energy threshold of $80\,\text{MeV}$, $30\,\text{MeV}$, and $60\,\text{MeV}$ in the forward endcap, barrel, and backward endcap, respectively. Clusters are required to be detected within 200 ns of the beam crossing to suppress energy deposits from beam background. The misreconstructed photons from hadronic clusters without a matched track, are suppressed with a dedicated multivariate classifier, which uses cluster-shape variables, calorimeter crystal pulse-shape, and track-to-cluster distance.

Neutral pions are reconstructed from photon candidate pairs, requiring the diphoton mass to be between 120 MeV and 145 MeV and the convergence of a mass-constrained fit. Candidate K_S^0 mesons are reconstructed from pairs of oppositely charged pions. We require the dipion mass to be within 10 MeV of the known value, a successful vertex fit, the transverse distance between the interaction point and the dipion vertex to exceed 0.4 cm,

and the cosine of the angle between the K_S^0 flight direction and its reconstructed momentum evaluated in the laboratory frame to exceed 0.8. Candidate K^{*0} mesons are reconstructed in the $K^{*0} \to K^+\pi^-$ decay channel, requiring their invariant mass to be within $50\,\text{MeV}$ of the known K^{*0} mass value.

Candidate $D^{(*)}$ mesons are reconstructed using the $D^0 \to K^-\pi^+$, $D^+ \to K^-\pi^+\pi^+$, $D^{*0} \to D^0\pi^0$, and $D^{*+} \to D^0\pi^+$ decay channels. The invariant masses of the D^0 and D^+ candidates are required to be within 15 MeV, corresponding to three units of resolution (σ), of the known values, and a mass- and vertex-constrained fit is performed. Similarly, the invariant mass of the D^{*0} and D^{*+} candidates is required to be within 3 MeV and 1.5 MeV of the known value (3σ).

Candidate B mesons are reconstructed in eight $B \to DK^-K$ channels and four $B \to DD_s^-$ channels. We require the beam-constrained mass $M_{\text{bc}} = \sqrt{E_{\text{beam}}^{*2} - p_B^{*2}} > 5.272$ GeV and -0.12 GeV $< \Delta E = E_B^* - E_{\text{beam}}^* < 0.3$ GeV, where p_B and E_B are the momentum and the energy of the B meson, respectively, and $E_{\text{beam}}^* = \sqrt{s}/2$ is the calibrated value of the beam energy. The symbol $*$ indicates that the variable is evaluated in the $\Upsilon(4S)$ center-of-mass frame. The asymmetric requirement on ΔE suppresses cross-feed from other channels in which a pion is not reconstructed. To reduce continuum background, we require that the ratio of the second- to zeroth Fox-Wolfram moments satisfies $R_2 < 0.5$ [28]; the absolute value of the cosine of the angle between the thrust axis of the B and the thrust axis of the rest of the event be smaller than 0.85, where the thrust axis is the direction that maximizes the total projection of the momenta of all particles in the event [29]; and the absolute value of the cosine of the angle between the B momentum and the beam direction satisfies $|\cos\theta_{p_B^* p_{\text{beam}}^*}| < 0.9$. The rest of the event includes all the remaining tracks and clusters selected with the following requirements. The tracks are required to be within the CDC polar angle acceptance, to have transverse momenta greater than 100 MeV, and to have $d_r < 0.5$ cm, and $|d_z| < 3$ cm. The clusters are required to be in the CDC polar acceptance, to have energies above 50 MeV, and to be detected within 200 ns of the beam crossing.

In the $B \to DK^-K$ channels, we also require $m(K^-K)$ to differ by more than $20\,\text{MeV}$ (4σ) from the known D_s^- mass, to suppress resonant $B \to DD_s^-(\to K^-K)$ decays, which has the same final state as the signal. To identify $B \to DD_s^-(\to K^-K)$ decays, the selection procedure for $B \to DK^-K$ channels is used; however, the D_s^- mass veto is inverted.

The average B candidate multiplicity per event as determined from signal MC simulation is between 1.05 and 1.08 for the D^0, D^+, and D^{*+} channels, while it is about 1.4 for the D^{*0} channels. For each event, the B candidate with M_{bc} closest to the known B mass is selected. The efficiency of this candidate selection, defined as the ratio of the number of correct candidates chosen in multiple candidate events to the total number of candidates for events with multiple candidates in which the correct candidate is reconstructed, is 67% for the $\overline{B}^0 \to D^{*0}K^-K_S^0$ channel, while it is between 80% and 92% for all the other channels, based on simulation. After the candidate selection, the purity of the sample (defined as the ratio of correctly reconstructed events to the total number of reconstructed events) is between 92% and 96% for all channels, except for the D^{*0} channels where it is 71%, according to the signal MC simulation.

The $B^- \to D^{*0}\pi^-$ control sample is reconstructed using the D^{*0}, π^- and B^- selections used in $B^- \to D^{*0}K^-K$ channels. In addition the D^{*0} momentum is required to be lower than 2.5 GeV, to obtain a momentum range similar to that of the $B^- \to D^{*0}K^-K$ signal.

We apply corrections for the track-momentum scale and photon-energy scale on data, which are derived from large data control samples. These corrections range from 10^{-4} to the 10^{-3} level.

4 Signal yield extraction

The composition of the selected samples is investigated using the ΔE distribution from MC simulation. A small, smooth $e^+e^- \to q\bar{q}$ continuum background is expected in all channels, while the signal $B \to DK^-K$ peaks at $\Delta E \approx 0$. In the $\overline{B}^0 \to D^+K^-K$ channels a cross-feed component from $\overline{B}^0 \to D^{*+}K^-K$ channels is expected, in which the π^0 from the $D^{*+} \to D^+\pi^0$ decay is not reconstructed. This component peaks at $\Delta E \approx -0.15$ GeV. A similar background is present in the $B^- \to D^0K^-K$ channels, which has a cross-feed component from the $B^- \to D^{*0}K^-K$ and $\overline{B}^0 \to D^{*+}K^-K$ channels, that both mimic the signal when the π^0 or the π^+ from the D^* decay is not reconstructed. In the $B^- \to D^{*0}K^-K$ channels, a cross-feed component from the $B^- \to D^0K^-K$ channel is reconstructed if an incorrectly reconstructed π^0 is associated with the D^0. This background component peaks around $\Delta E \approx 0.15$ GeV. In the $B^- \to D^{*0}K^-K$ channels a cross-feed component from the $\overline{B}^0 \to D^{*+}K^-K$ channels is also expected, when the π^+ from the D^{*+} decay is not reconstructed and an incorrect π^0 is included. This background component peaks under the signal, and thus it requires special treatment. The $\overline{B}^0 \to D^{*+}K^-K$ channels have no peaking backgrounds. The four $B \to DK^-K^{*0}$ channels have backgrounds from $B \to DK^-K^+\pi^-$, i.e. the non-K^{*0}-resonant component. The $B \to DD_s^-$ channels have the same composition as the $B \to DK^-K$ channels; however, they have an additional background from $B \to DK^-K$.

The RooFit package [30] is used to perform extended maximum likelihood fits to the unbinned ΔE distribution to extract the signal yield. The fit is performed in the ΔE range $[-0.12\,\text{GeV}, 0.3\,\text{GeV}]$ for all channels. The upper boundary is chosen to better constrain the fit for the $B^- \to D^0K^-K$, $\overline{B}^0 \to D^+K^-K$, and $\overline{B}^0 \to D^{*+}K^-K$ channels. In the $B^- \to D^{*0}K^-K$ channel an extended range is used to constrain the background peaking under the signal.

The ΔE distribution is described with the following model:

$$p(\Delta E) = p_{\text{sig}}(\Delta E) + p_{\text{bkg}}(\Delta E) + p_{\text{cross-feed}}(\Delta E) + p_{DKK\pi\text{-bkg}}(\Delta E) + p_{DKK\text{-bkg}}(\Delta E), \quad (4.1)$$

where the details of the components are described in table 1, and each individual component (signal, combinatorial background, cross-feed in $B \to D^{*0}K^-K$ channels, $B \to DK^+K^-\pi$ background, $B \to DK^-K$ background in $B \to DD_s$ channels, respectively) is included only in the specific B decay channel indicated in the last column of the table.

4.1 $B \to DK^-K_S^0$ channels

In the $B^- \to D^0K^-K_S^0$, $\overline{B}^0 \to D^+K^-K_S^0$, and $\overline{B}^0 \to D^{*+}K^-K_S^0$ channels, the signal is described by the sum of one symmetric (*core*) and one asymmetric (*tail*) Gaussian distributions sharing the same mean parameter. The widths of the Gaussians are fixed to the values obtained in a fit to the simulated signal sample. However, a scale factor r is assigned as a multiplier to the core Gaussian width to compensate for any difference in resolution between

Component	Description	B signal channels
p_{sig}	$C\left[(1-t)\mathcal{G}(\Delta E, \mu, r\cdot\sigma) + t\mathcal{G}_A(\Delta E, \mu, \sigma_L, \sigma_R)\right]$	All
p_{bkg}	$Be^{d\Delta E} + K$	All
$p_{\mathrm{cross\text{-}feed}}$	$N\left[\mathcal{G}_A(\Delta E, \mu+\Delta\mu, \sigma_{L0}, \sigma_{R0}) + \frac{\mathcal{B}(D^{*+})}{\mathcal{B}(D^0)}\mathcal{G}_A(\Delta E, 0, \sigma_{L+}, \sigma_{R+})\right]$	D^{*0} channels
$p_{DKK\pi\text{-}\mathrm{bkg}}$	$fC\mathcal{G}(\Delta E, \mu, r\cdot\sigma)$	$B\to DK^-K^{*0}$ channels
$p_{DKK\text{-}\mathrm{bkg}}$	$D\mathcal{G}(\Delta E, \mu, r\cdot\sigma)$	$B\to DD_s^-$ channels

Table 1. Details of the model for the ΔE fit. The symbol $\mathcal{G}(x, \mu, \sigma)$ stands for a Gaussian function with mean μ and width σ; $\mathcal{G}_A(x, \mu, \sigma_L, \sigma_R)$ stands for an asymmetric Gaussian function with mean μ, left width σ_L, and right width σ_R; $\mathcal{B}(D^{*+})/\mathcal{B}(D^0)$ is the ratio of branching fractions from eq. (4.5); f stands for the fraction of the $DKK\pi$ component as described in the text; C, B, K, N, and D are yield parameters; t is the tail-to-core Gaussian fraction; r is the resolution scale factor; d is the exponential decay constant; and $\Delta\mu$ is the shift of the mean for the cross-feed component.

simulation and data. The resolution scale factor is a free parameter in the D^0 channel fit to data. In the D^+ and D^{*+} channels, the resolution scale factors are fixed to the value from the D^0 fit ($r = 1.1 \pm 0.1$), since the expected yields are too small to constrain an additional free parameter. The fractions of the core and the tail Gaussian are fixed to the values obtained in a fit to the signal simulation sample. The background is described by the sum of a falling exponential distribution and a constant. The fit has two or three free shape parameters (mean, exponential constant, and resolution scale factor for D^0 channel), along the signal yield, and two background yields.

The $B^- \to D^{*0}K^-K_S^0$ channel parametrization requires a different approach, because of the background from the $\overline{B}^0 \to D^{*+}K^-K_S^0$ channel peaking under the signal. The yield of this cross-feed component can be constrained directly from data using the cross-feed from the $B^- \to D^0K^-K_S^0$ channel, which is shifted in ΔE by 0.15 GeV. This ΔE fitting region is almost free from continuum background, allowing an accurate determination of the cross-feed despite the low yield. The yield $N_{D^0}^{\mathrm{bkg}}$ of cross-feed from the $B^- \to D^0K^-K_S^0$ channel can be determined from data as

$$N_{D^0}^{\mathrm{bkg}} \propto \mathcal{B}(B^- \to D^0K^-K_S^0)f_{\pi^0}\varepsilon_{D^0}, \tag{4.2}$$

where ε_{D^0} is the efficiency of $B^- \to D^0K^-K_S^0$ reconstruction and f_{π^0} is the probability of incorrect π^0 association. For the $\overline{B}^0 \to D^{*+}K^-K_S^0$ cross-feed, assuming that the efficiency of a multiparticle final state factorizes into the products of single-particle efficiencies, we have

$$N_{D^{*+}}^{\mathrm{bkg}} \propto \mathcal{B}(\overline{B}^0 \to D^{*+}K^-K_S^0)f_{\pi^0}\varepsilon_{D^0\text{-}\mathrm{in}\text{-}D^{*+}}, \tag{4.3}$$

where $\varepsilon_{D^0\text{-}\mathrm{in}\text{-}D^{*+}}$ is the efficiency of $B^- \to D^0K^-K_S^0$ reconstruction from $B^- \to D^{*+}K^-K_S^0$ events. The $\overline{B}^0 \to D^{*+}K^-K_S^0$ decay chain is identical to $\overline{B}^- \to D^0K^-K_S^0$, except for the additional low-momentum π^+, and shows no relevant kinematic differences in the simulation of the $D^0K^-K_S^0$ channel that could lead to a significant difference in the efficiency. Therefore we assume

$$\varepsilon_{D^0\text{-}\mathrm{in}\text{-}D^{*+}} = \varepsilon_{D^0}, \tag{4.4}$$

to obtain

$$N_{D^{*+}}^{\text{bkg}} = \frac{\mathcal{B}(\overline{B}^0 \to D^{*+}K^-K_S^0)}{\mathcal{B}(B^- \to D^0K^-K_S^0)} N_{D^0}^{\text{bkg}}. \tag{4.5}$$

Thus, by measuring the yield of the ΔE-shifted cross-feed from the $B^- \to D^0K^-K_S^0$ channel, and adding as inputs the branching fractions of the two cross-feed backgrounds measured in this analysis, we constrain the yield of the cross-feed from $\overline{B}^0 \to D^{*+}K^-K_S^0$ decays peaking under the signal. A systematic uncertainty is assigned due to the assumptions in this procedure (as described in section 7).

The $B^- \to D^{*0}K^-K_S^0$ channel is fitted using functional forms described in table 1. The resolution scale factor is fixed to unity, as estimated from the $B^- \to D^{*0}\pi^-$ control channel. The cross-feed component from the $B^- \to D^0K^-K_S^0$ channel is parametrized as an asymmetric Gaussian distribution with widths fixed from the signal simulation and mean equal to the signal mean with a fixed shift evaluated from the signal simulation. The $B^- \to D^0K^-K_S^0$ cross-feed yield is free in the fit. The cross-feed component from the $\overline{B}^0 \to D^{*+}K^-K_S^0$ channel is fitted as an asymmetric Gaussian distribution with mean fixed to zero and width fixed from the signal simulation. The yield is fixed based on the relation in eq. (4.5), where $N_{D^0}^{\text{bkg}}$ is the yield of the $B^- \to D^0K^-K_S^0$ cross-feed, which is fitted simultaneously, and the D^{*+} and D^0 channel branching fractions are fixed. The fits have two shape parameters (mean and exponential constant), the signal yield, the yields of two backgrounds, and the yield of the $B^- \to D^0K^-K_S^0$ cross-feed component as free parameters.

Data with fit projections overlaid are shown in figure 1 for the four K_S^0 channels. The backgrounds are smooth and small as expected, with a signal-to-background ratio at $\Delta E \approx 0$ between 3 and 15. The cross-feed components in the D^{*0} channel are well modeled. The four signals have statistical significances well above five standard deviations, calculated as $\sqrt{-2\ln(\mathcal{L}_0^{\max}/\mathcal{L}^{\max})}$, where $\mathcal{L}_0^{\max}$ and $\mathcal{L}^{\max}$ are the maximized likelihood values for the background-only and the signal-plus-background hypothesis, respectively. The pulls between the data distribution and the fit are also shown, defined as the difference between the data yield and the fit value divided by the data uncertainty. The agreement is reasonable. The yields are summarized in table 2.

4.2 $B \to DK^-K^{*0}$ channels

The $B \to DK^-K^{*0}$ channels are fitted in ΔE with the same function as used for the K_S^0 channels, however the resolution scale factors are free parameters in the four channels. An additional Gaussian function is added to each mode to describe the four body $B \to DK^-K^+\pi^-$ component. The mean and the width of the additional Gaussian function are fixed to that of the core Gaussian function. The $B \to DK^-K^+\pi^-$ yield is constrained to be a fixed fraction of the signal yield. The fraction is estimated from a fit to the $m(K^+\pi^-)$ distribution. The latter is obtained by relaxing the K^{*0} invariant mass selection, performing a fit to the resulting ΔE distribution with the same model used to fit the K_S^0 channels, and using the sPlot technique to obtain the continuum-background-subtracted $m(K^+\pi^-)$ distribution. The $m(K^+\pi^-)$ distribution is fitted in the range 0.65 GeV $< m(K^+\pi^-) <$ 2.5 GeV, with the sum of a histogram template from the simulation for the signal component and a third-order
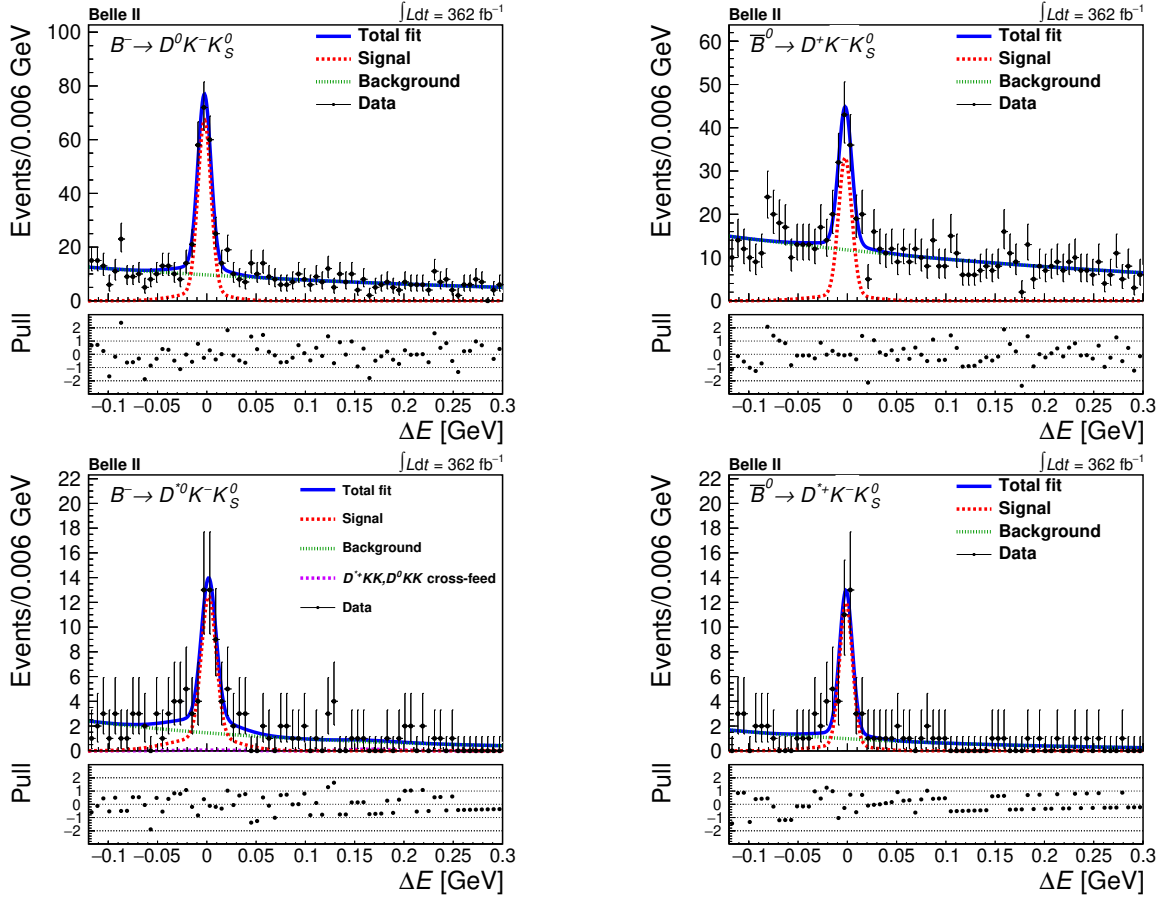
Figure 1. Distribution of ΔE for the $B^- \to D^0 K^- K_S^0$ (top left), $\overline{B}^0 \to D^+ K^- K_S^0$ (top right), $B^- \to D^{*0} K^- K_S^0$ (bottom left), and $\overline{B}^0 \to D^{*+} K^- K_S^0$ (bottom right) channels. The projections of the fits are overlaid, the fit components are highlighted, and the pulls between the fit and the data are shown in the bottom panel of each plot.

Chebyshev polynomial for the $B \to DK^- K^+ \pi^-$ component with parameters fixed from simulation. The $[1.25\,\text{GeV}, 1.60\,\text{GeV}]$ and $[1.85\,\text{GeV}, 1.87\,\text{GeV}]$ invariant mass regions are vetoed to exclude higher-mass kaon resonances and $B \to D^{(*)} D^{(*)} K$ decays, respectively. The $m(K^+ \pi^-)$ distributions with the projection of the fits overlaid are shown in figure 11 for the four channels. The non-K^{*0}-resonant fraction is obtained from the ratio between the $B \to DK^- K^+ \pi^-$ and the $B \to DK^- K$ yields within the K^{*0} signal region. The fractions are between 0.7% and 3.1% of the signal, depending on the channel. The fractions are assumed to be as uniform in the $(m(K^- K), m(DK))$ plane, given their small values and large systematic uncertainties. Data with fit projections overlaid for K^{*0} channels and the pulls between the data distribution and the fit are shown in figure 2. As in the K_S^0 channels, the remaining backgrounds are small as expected, with signal-to-background ratios at $\Delta E \approx 0$ between 8 and 100. All the four signals have statistical significances well above five standard deviations. The yields are summarized in table 2.
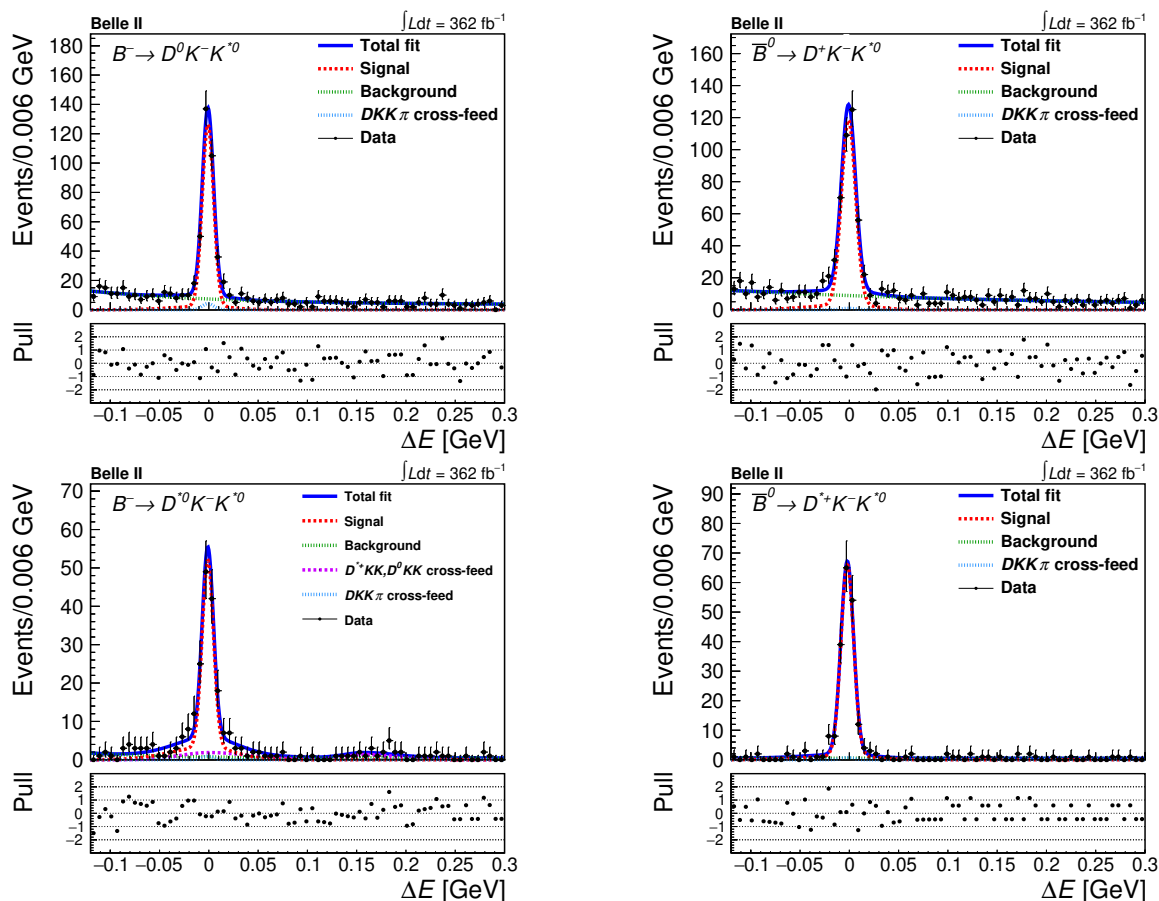
Figure 2. Distribution of ΔE for the $B^- \to D^0 K^- K^{*0}$ (top left), $\overline{B}^0 \to D^+ K^- K^{*0}$ (top right), $B^- \to D^{*0} K^- K^{*0}$ (bottom left), and $\overline{B}^0 \to D^{*+} K^- K^{*0}$ (bottom right) channels. The projections of the fits are overlaid, the fit components are highlighted, and the pulls between the fit and the data are shown in the bottom panel of each plot.

4.3 $B \to DD_s^-$ channels

The $B \to DD_s^-$ channels are fitted with the same strategy as the $B \to DK^- K$ channels. The fits are performed independently for the $D_s^- \to K^- K_S^0$ and $D_s^- \to K^- K^{*0}$ final states. The resolution scale factors are fixed in the $\overline{B}^0 \to D^{*+} D_s^-$ and $\overline{B}^- \to D^{*0} D_s^-$ fits using the average of the resolution scale factors obtained from the D^+ and D^0 channels. The $D_s^- \to K^- K^+ \pi^-$ yield is assumed to be zero in the $D_s^- \to K^- K^{*0}$ channels, given the previous measurements of D_s^- branching fractions [31]. The $B \to DK^- K$ background is described with an additional Gaussian component, with the same mean and width of the core Gaussian. The yield of the latter is fixed in the fit and is estimated from the $m(K^- K)$ distribution of the $B \to DK^- K$ channels measured in this work, assuming it uniform within the bin of interest. Data with fit projections overlaid and the pulls between data and the fit are shown in figure 3 and figure 4 for the four K_S^0 channels, and the four K^{*0} channels, respectively. The eight channels are almost free from background, except for the cross-feed in the D^{*0} channels and the $B \to DK^- K$ cross-feed, which are between 0.1% and 13.0% of the signal.
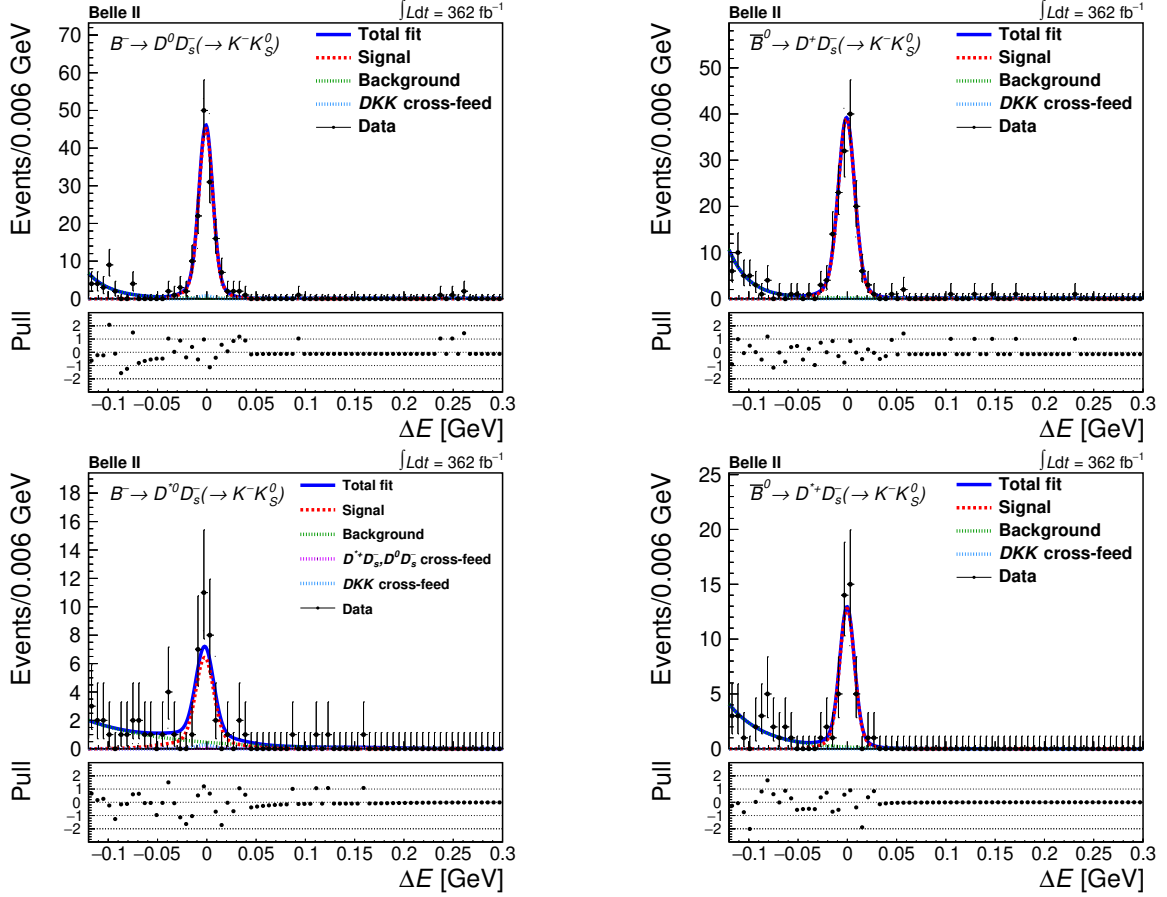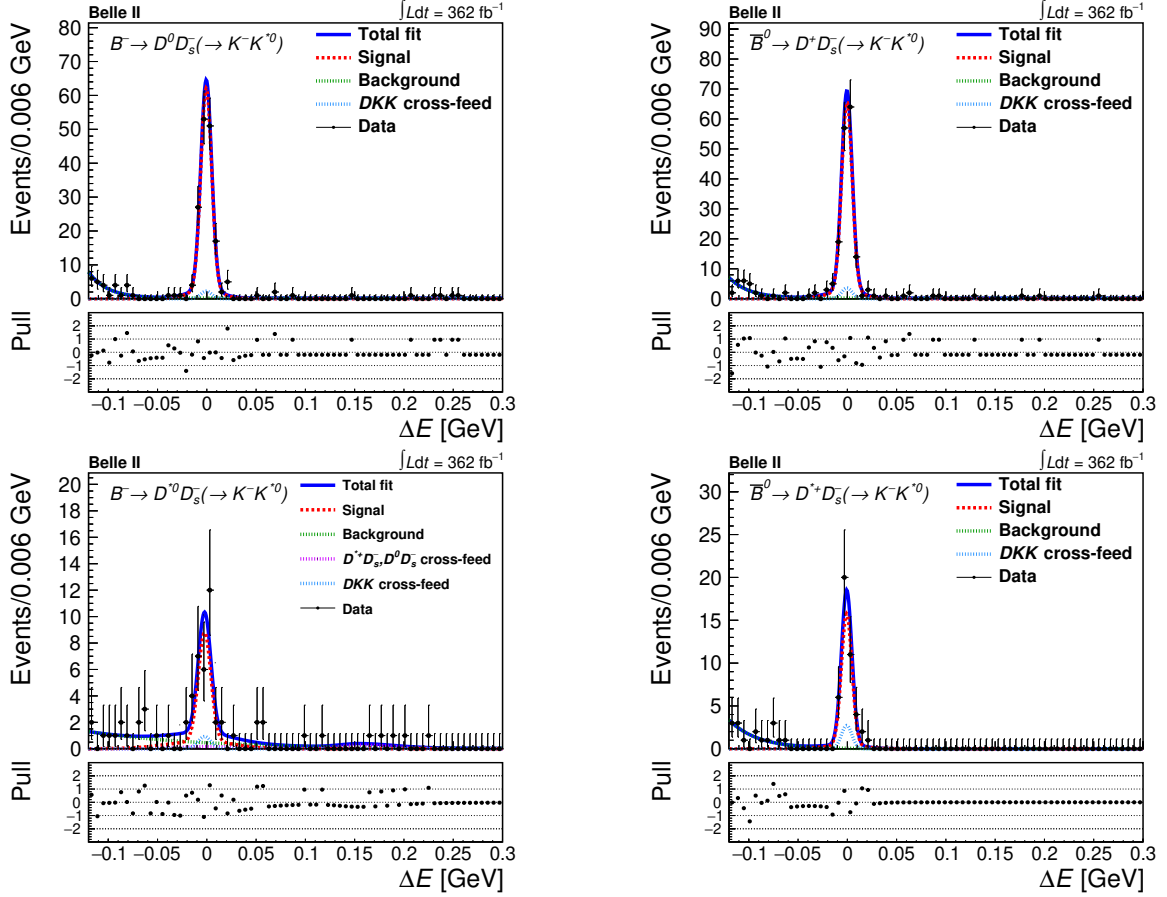
Figure 3. Distribution of ΔE for the $B^- \to D^0 D_s^-(\to K^- K_S^0)$ (top left), $\overline{B}^0 \to D^+ D_s^-(\to K^- K_S^0)$ (top right), $B^- \to D^{*0} K^- K_S^0$ (bottom left), and $\overline{B}^0 \to D^{*+} D_s^-(\to K^- K_S^0)$ (bottom right) channels. The projections of the fits are overlaid, the fit components are highlighted, and the pulls between the fit and the data are shown in the bottom panel of each plot.

The fits are validated by repeating the analysis on 10^3 Toy MC experiments, i.e. simplified simulated pseudo-experiments based on sampling the likelihood, and show no bias. The $B \to DK^-K$ fit is validated on data using the $B \to DD_s^-(\to K^-K)$ decays as control channels, which shares the same final state as the signal, and a very similar decay topology. Thus, the efficiency for the control channels is similar to that of $B \to DK^-K$ decays. In addition, the control channels are well measured [31]. The resulting control-channel branching fractions are in agreement with the world averages, validating the signal extraction on data.

5 Efficiency determination

The efficiency is obtained for each channel separately using signal simulation samples. An admixture of three-body $B \to DK^-K$ decays produced with a uniform distribution in phase-space, and $B \to DX^-$ events (where $X = a_1(1260)^-$ or $\rho(1450)^-$ for the K^{*0} or K_S^0 channels, respectively) is used. We divide the $(m(K^-K), m(DK))$ distribution into 20×20 equally-spaced intervals (bins) with $0.9 \text{ GeV} < m(K^-K) < 3.5 \text{ GeV}$ and $2 \text{ GeV} < m(DK) < 5 \text{ GeV}$. The efficiency $\varepsilon(m(K^-K), m(DK))$ is defined as the fraction of generated events that are

– 11 –

Figure 4. Distribution of ΔE for the $B^- \to D^0 D_s^-(\to K^- K^{*0})$ (top left), $\overline{B}^0 \to D^+ D_s^-(\to K^- K^{*0})$, (top right), $B^- \to D^{*0} D_s^-(\to K^- K^{*0})$ (bottom left), and $\overline{B}^0 \to D^{*+} D_s^-(\to K^- K^{*0})$ (bottom right) channels. The projections of the fits are overlaid, the fit components are highlighted, and the pulls between the fit and the data are shown in the bottom panel of each plot.

reconstructed and selected in each bin of reconstructed $(m(K^- K), m(DK))$. This allows the efficiency to be much less dependent on the $(m(K^- K), m(DK))$ distribution of the data, which is a priori unknown and possibly different from the simulation. The reconstructed $(m(K^- K), m(DK))$ spectrum is obtained by fitting the ΔE distribution of signal simulation, using the functional forms described in section 4. An sPlot is then used to obtain the $(m(K^- K), m(DK))$ distribution.

The efficiency is corrected for known data-simulation mismodeling. In particular, the PID selection efficiency is calibrated with a scale factor as a function of momentum and polar angle for each track; the K_S^0 reconstruction and selection efficiency is calibrated with a scale factor as a function of the vertex separation from the IP, and the momentum and polar angle of the K_S^0; the reconstruction and selection efficiencies of the low-momentum pions from the D^{*+} and D^{*0} decays are calibrated with two scale factors as functions of the pion momentum. The overall correction factor ranges between 1% and 10%.

The resulting efficiencies $\varepsilon(m(K^- K), m(DK))$ are shown in figure 5 and figure 6 for the four K_S^0 channels and the four K^{*0} channels, respectively. All channels show a drop
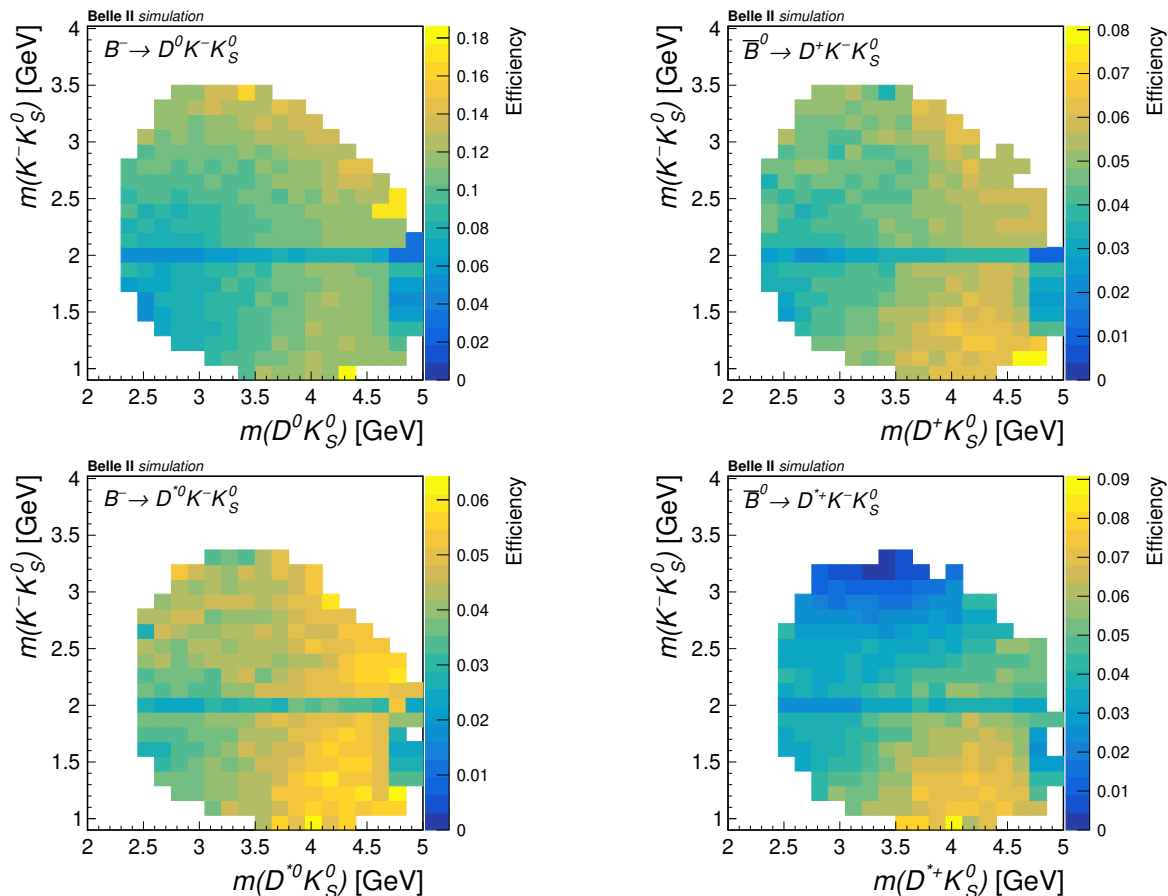
Figure 5. Efficiency as a function of $\big(m(K^-K), m(DK)\big)$ for the $B^- \to D^0 K^- K_S^0$ (top left), $\overline{B}^0 \to D^+ K^- K_S^0$ (top right), $B^- \to D^{*0} K^- K_S^0$ (bottom left), and $\overline{B}^0 \to D^{*+} K^- K_S^0$ (bottom right) channels.

in efficiency at $m(K^-K)$ close to the D_s^- mass due to the $B \to DD_s^-$ veto. For the D^{*+} channels, $\varepsilon\big(m(K^-K), m(DK)\big)$ decreases at high $m(K^-K)$ due to the anticorrelation between $m(K^-K_S^0)$ and the momentum of the pion from the D^{*+} decay, for which the efficiency decreases at low momentum. These efficiencies have statistical uncertainties between 0.5% and 5%, and increase up to 15% at the edge of the phase-space. The allowed phase-space is properly populated using the admixture of simulated samples. The invariant mass resolutions are sufficiently small compared to the efficiency bin width to allow neighbour-bin migration effects to be neglected.

The efficiencies for the $B \to DD_s$ channels are defined as the fraction of the generated events that are reconstructed and selected from the $B \to DD_s(\to K^-K)$ signal simulation. The efficiencies, listed in table 2, are estimated separately for the $D_s^- \to K^- K_S^0$ and $D_s^- \to K^- K^{*0}$ final states.
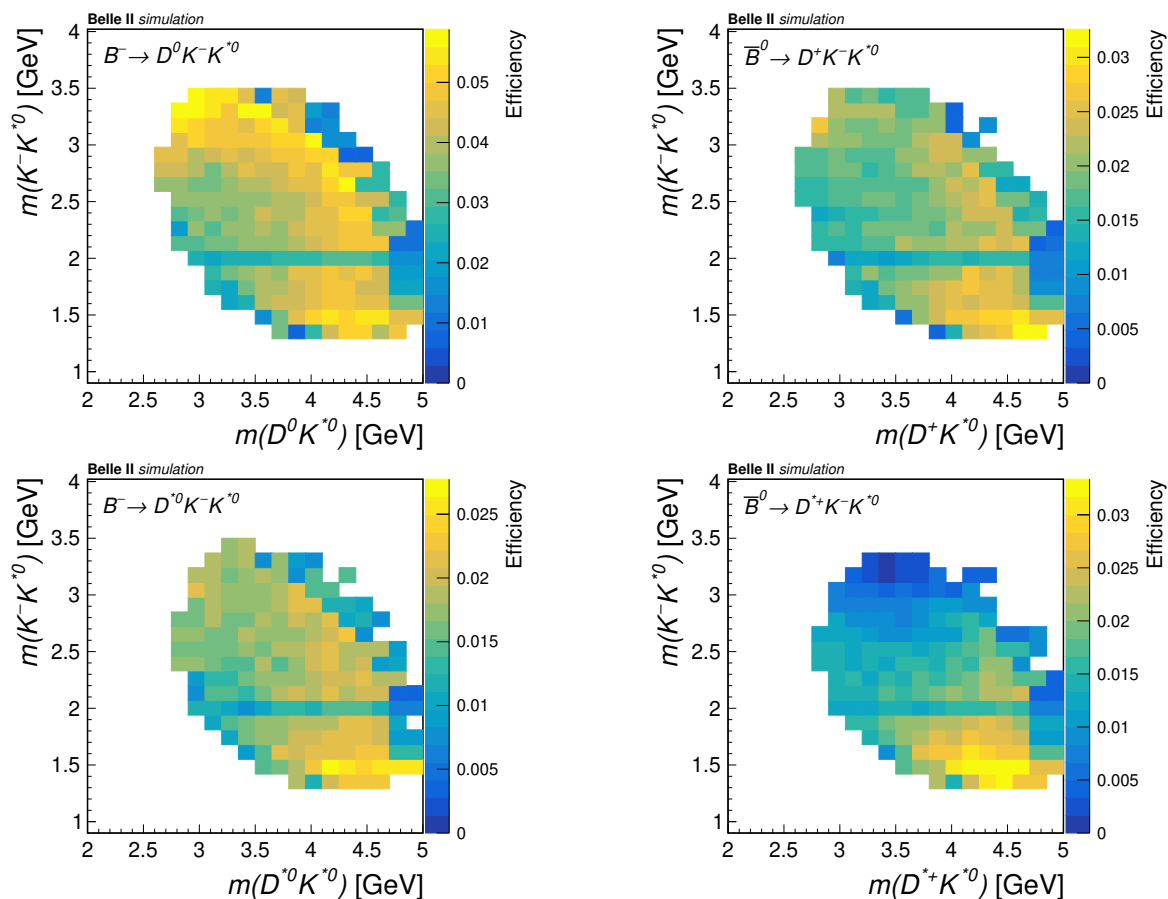
Figure 6. Efficiency as a function of $\left(m(K^-K), m(DK)\right)$ for the $B^- \to D^0 K^- K^{*0}$ (top, left), $\overline{B}^0 \to D^+ K^- K^{*0}$ (top right), $B^- \to D^{*0} K^- K^{*0}$ (bottom left), and $\overline{B}^0 \to D^{*+} K^- K^{*0}$ (bottom right) channels.

6 Branching fraction extraction

For the $B \to DK^-K$ channels, the branching fraction can be expressed as

$$\mathcal{B} = \frac{N_{\text{reco}}^{\varepsilon \text{ corr}}}{2 f_{+-,00} N_{B\overline{B}} \mathcal{B}_{D^{(*)}} \mathcal{B}_{K_{(S)}^{(*)0}}} \tag{6.1}$$

where $N_{\text{reco}}^{\varepsilon \text{ corr}}$ is the background-subtracted and efficiency-corrected signal yield, $\mathcal{B}_{D^{(*)}} \mathcal{B}_{K_{(S)}^{(*)0}}$ is the product of the branching fractions of the relevant intermediate $D^{(*)}$ and $K_{(S)}^{(*)0}$ decays in the reconstructed decay chain, $N_{B\overline{B}}$ is the total number of $B\overline{B}$ pairs, and $f_{+-,00}$ is the fraction of charged or neutral $B\overline{B}$ pairs.

The signal yield as a function of $\left(m(K^-K), m(DK)\right)$ is extracted using an *s*Plot technique. The ΔE distribution, fitted as described in section 4, is used as a discriminating variable to obtain the *s*Weights. The *s*Weights for the signal are then used to obtain the $\left(m(K^-K), m(DK)\right)$ distribution for pure signal. The prerequisites for *s*Plot validity are satisfied as the ΔE distribution is independent of $m(K^-K)$ and $m(DK)$. The efficiency

Channel	Yield	Average ε	$\mathcal{B}$ [10^{-4}]	Stat. significance [σ]
$B^- \to D^0 K^- K_S^0$	209 ± 17	0.098	$1.82 \pm 0.16 \pm 0.08$	> 10
$\overline{B}^0 \to D^+ K^- K_S^0$	105 ± 14	0.048	$0.82 \pm 0.12 \pm 0.05$	10
$B^- \to D^{*0} K^- K_S^0$	51 ± 9	0.044	$1.47 \pm 0.27 \pm 0.10$	8
$\overline{B}^0 \to D^{*+} K^- K_S^0$	36 ± 7	0.046	$0.91 \pm 0.19 \pm 0.05$	9
$B^- \to D^0 K^- K^{*0}$	325 ± 19	0.043	$7.19 \pm 0.45 \pm 0.33$	> 10
$\overline{B}^0 \to D^+ K^- K^{*0}$	385 ± 22	0.021	$7.56 \pm 0.45 \pm 0.38$	> 10
$B^- \to D^{*0} K^- K^{*0}$	160 ± 15	0.019	$11.93 \pm 1.14 \pm 0.93$	> 10
$\overline{B}^0 \to D^{*+} K^- K^{*0}$	193 ± 14	0.020	$13.12 \pm 1.21 \pm 0.71$	> 10
$B^- \to D^0 D_s^-$	144 ± 12 / 153 ± 13	0.09 / 0.04	$95 \pm 6 \pm 5$	> 10 / > 10
$\overline{B}^0 \to D^+ D_s^-$	145 ± 12 / 159 ± 13	0.05 / 0.02	$89 \pm 5 \pm 5$	> 10 / > 10
$B^- \to D^{*0} D_s^-$	30 ± 6 / 29 ± 7	0.04 / 0.02	$65 \pm 10 \pm 6$	7 / 8
$\overline{B}^0 \to D^{*+} D_s^-$	43 ± 7 / 37 ± 7	0.04 / 0.02	$83 \pm 10 \pm 6$	> 10 / > 10

Table 2. Observed yields for the eight $B \to DK^-K$ channels and the four $B \to DD_s^-$ channels and their statistical uncertainty; average efficiency with the correction described in Sec. 5; measured branching fractions with statistical (first) and systematic (second) uncertainties; statistical significances. The first/second value for yield, efficiency, and significance of $B \to DD_s^-$ channels refer to K_S^0/K^{*0} sub-channels.

correction is obtained by applying a $1/\varepsilon\big(m(K^-K), m(DK)\big)$ weight event-by-event, according to the efficiency map described in section 5.

The $f_{+-,00}$ values used in eq. (6.1) are evaluated from the ratio f_{+-}/f_{00} from ref. [32], while $N_{B\overline{B}}$ is reported in section 2. We obtain $2f_\pm N_{B\overline{B}} = (399 \pm 11) \times 10^6$ and $2f_0 N_{B\overline{B}} = (375 \pm 10) \times 10^6$. The branching fractions of the intermediate $D^{(*)}$ and K_S^0 decays are taken from the world-average values in ref. [31].

The $B \to DD_s^-$ branching fractions are first determined independently for the $D_s^- \to K^- K_S^0$ and $D_s^- \to K^- K^{*0}$ sub-channels using eq. (6.1), where the product of the branching fractions is replaced by $\mathcal{B}_{D^{(*)}} \mathcal{B}_{D_s^-} \mathcal{B}_{K^{(*)0}_{(S)}}$ and $N_{\rm reco}^{\varepsilon\,{\rm corr}} = N_{\rm reco}^K/\varepsilon_K$, where $N_{\rm reco}^K$ and ε_K are the reconstructed yield and the efficiency in the specific sub-channel. For each of the four channels, the two D_s^- sub-channels give compatible results. The $B \to DD_s^-$ branching fractions are then obtained from the weighted average of the results from the two D_s^- sub-channels, including the statistical uncertainty only in the weights.

The branching fractions are given in table 2 for each decay channel.

7 Systematic uncertainties

The contributions to the systematic uncertainties for each $B \to DK^-K$ channel are summarized in table 3, expressed as relative uncertainties to the corresponding branching fractions.

The first group of systematic uncertainties affects the efficiency estimation. They are related to the differences between efficiencies in data and in simulation, and are derived from the uncertainties on the corresponding corrections. An uncertainty associated with the limited size of the MC simulation sample is included ("Eff - MC sample size" in table 3),

which gives the statistical uncertainty in the efficiency determination. The uncertainty related to the tracking efficiency ("Eff - tracking") is estimated using a $e^+e^- \to \tau^+\tau^-$ data control sample. No correction is applied, but a per-track uncertainty of 0.24% is included assuming full correlation between the tracks. For the uncertainty associated with the K_S^0 efficiency ("Eff - K_S^0"), the nominal scale factor is varied by its uncertainty. The uncertainty is assumed to be fully correlated in the vertex separation from the IP, momentum and polar angle space. The scale factors and the uncertainties are determined in $D^{*+} \to D^0(\to K_S^0\pi^+\pi^-)\pi^+$ decays. A similar approach is adopted for the uncertainty associated with the PID-efficiency correction ("Eff - PID"). The uncertainty is assumed to be fully correlated between tracks and as a function of momentum, and polar angle of the tracks. The scale factors and the associated uncertainty are evaluated using $D^{*+} \to \overline{D}^0(\to K^-\pi^+)\pi^+$ and $K_S^0 \to \pi^+\pi^-$ data control samples. For the uncertainty associated with the efficiency for reconstructing the low-momentum charged pion ("Eff - π^+ from D^{*+}"), the scale factors are evaluated with statistical (partially correlated in momentum) and systematic (correlated) uncertainties estimated on a $B^0 \to D^{*-}\pi^+$ data control sample. A fully correlated 2.7% uncertainty is used for the scale factors, equal to the sum in quadrature of both uncertainties. This uncertainty affects only the D^{*+} channels. For the uncertainty associated with the π^0-efficiency correction ("Eff - π^0") in the D^{*0} channels, the scale factor is varied according to the uncertainties evaluated on a $B^+ \to D^{*0}\pi^+$ data control sample.

A systematic uncertainty is assigned to the efficiency ("Eff - modeling") due to the possible mismodeling of the efficiency distribution in the $(m(K^-K), m(DK))$ plane and the integration over additional degrees of freedom beyond the two adopted dimensions, due to polarization of vector particles (D^*, K^*, and possible vector resonances). The systematic uncertainty is determined by validating the efficiency map correcting distributions generated with different resonant and polarization schemes ($B \to D\rho'$, $B \to Da_1^-$, ...).

The signal model used in the ΔE fit is modified to check its robustness. Alternative branching fractions are evaluated, and a systematic uncertainty is quoted as the difference between the nominal and the alternative branching fraction ("Signal model"). A variation is considered for the channels in which the resolution scale factors are fixed using auxiliary channels. The corresponding systematic uncertainty is evaluated by shifting the resolution scale factors within their uncertainties ($\delta r = 0.1$ for $D^{(*)+}$ and $\delta r = 0.03$ for D^{*0}) and repeating the fit. This variation is applied to the D^+, D^{*0}, and D^{*+} channels. A second variation is obtained from a fit to the $B \to DD_s^-$ data sample allowing the widths of the tail Gaussian to float in the fit. The difference between the fixed values (i.e., those fixed from signal simulation samples) and the values obtained from the fit are used as variations of the fixed values of the widths in the $B \to DK^-K_S^0$ fit, to obtain an alternative value of the branching fraction. Given their small signal yields, the latter approach is not feasible for the $\overline{B}^0 \to D^+D_s^-$ and $B \to D^*D_s^-$ channels. Therefore, the variation observed in the $B^- \to D^0D_s^-$ mode is included for the D^+ and D^{*+} channels, but not for the D^{*0} channels, for which a different approach is used as described below.

The $B^- \to D^{*0}K^-K$ signal includes a 29% self-cross-feed component i.e., misreconstructed signal events, mostly due to incorrect π^0 associations. Since the self-cross-feed component is treated as signal, it does not artificially increase the branching ratio of the

signal channels. However, it does degrade the resolution in ΔE. Given the possible disagreement in the description of the π^0 misreconstruction between data and simulation, a dedicated systematic uncertainty is assigned (included in "Signal model"). A fit to the $B^- \to D^{*0}\pi^-$ data control channel is performed allowing the widths of the tail Gaussian σ_L and σ_R to vary in the fit. A second fit is performed on the simulation fixing the values of the widths. The differences in widths from the two fits is used to obtain an alternative value of the branching fractions. The systematic uncertainty is the difference between the nominal and the alternative value. The self-cross-feed is small in the other six channels and the fit-model systematic uncertainty covers possible mismodelings.

A systematic uncertainty related to the specific choice of background model is assigned ("Bkg model"). In the nominal fit, the background for all channels is described by the sum of an exponential function and a constant. Two alternative fits are performed by adding a linear term or removing the constant term. The differences between the nominal branching fractions and the results obtained with the alternative background models are quoted as systematic uncertainties.

A systematic uncertainty related to the possible mismodeling of the non-K^{*0}-resonant fraction is assigned to the four K^{*0} channels ("$DKK\pi$ bkg"). Four alternative non-K^{*0}-resonant background fractions compared to the nominal one are assumed, and the branching fraction is estimated again with the varied fractions. First, two alternative non-K^{*0}-resonant fractions are obtained by varying the fractions according to the statistical uncertainties of the fractions propagated from the fit components. Second, an alternative non-K^{*0}-resonant fraction is determined by replacing the signal template with a relativistic Breit-Wigner distribution corrected for the two-body phase-space factor. Third, the parameters of the Chebyshev polynomial are fixed from a simulation produced with `PYTHIA` instead of `EvtGen`. The total systematic uncertainty is the sum in quadrature of the four variations.

A systematic uncertainty is assigned ("D^{*0} peaking bkg") to account for the uncertainties in eq. (4.5), used to assess the yield of the cross-feed from $\overline{B}^0 \to D^{*+}K^-K$ channels in the D^{*0} channels. Equation (4.4) is verified with signal simulation, and holds with a maximum deviation of 10%. The yield of the peaking background $N_{D^{*+}}^{\text{bkg}}$ is scaled to 110% and 90% and two resulting signal yields are extracted. In addition, the branching fractions of D^{*+} and D^0 channels are scaled by their uncertainties, to take into account the correlation between the branching fractions and the $N_{D^{*+}}$ yield, producing additional variations of $\mathcal{B}(\overline{B}^0 \to D^{*0}K^-K)$. This systematic uncertainty also addresses the correlation between the D^{*0} channel and the D^{*+} and D^0 channels. The systematic uncertainty is the sum in quadrature of the differences between the resulting alternative branching fractions and the nominal one. This systematic uncertainty affects only the D^{*0} channels.

A systematic uncertainty related to the knowledge of the total number of $B\overline{B}$ pairs $\delta N_{B\overline{B}} = 5.6 \times 10^6$, which enters in eq. (6.1), is included ("$N_{B\overline{B}}$"). Similarly, a the systematic uncertainty related to $f_{+-,00}$ from ref. [32] ("$f_{+-,00}$") is included.

A systematic uncertainty related to the intermediate branching fraction used in eq. (6.1) is also included ("Intermediate $\mathcal{B}$s") by propagating the uncertainties of the known intermediate branching fractions to the final results.

Source	$D^0 K^- K^0_S$	$D^+ K^- K^0_S$	$D^{*0} K^- K^0_S$	$D^{*+} K^- K^0_S$	$D^0 K^- K^{*0}$	$D^+ K^- K^{*0}$	$D^{*0} K^- K^{*0}$	$D^{*+} K^- K^{*0}$
Eff. - MC sample size	0.5	0.8	1.1	0.9	0.5	0.7	0.9	1.2
Eff. - tracking	0.7	1.0	0.7	1.0	1.0	1.2	1.0	1.2
Eff. - π^+ from D^{*+}	-	-	-	2.7	-	-	-	2.7
Eff. - K^0_S	2.4	2.7	2.3	2.3	-	-	-	-
Eff. - PID	1.3	1.7	0.5	0.6	2.5	2.6	1.6	1.7
Eff. - π^0	-	-	5.1	-	-	-	5.1	-
Eff. - modeling	0.2	0.3	0.6	0.7	1.3	2.0	3.1	2.4
Signal model	1.5	3.6	2.3	2.7	0.8	1.0	2.5	0.6
Bkg model	0.8	1.1	0.8	0.8	1.1	0.4	0.2	0.1
$DKK\pi$ bkg	-	-	-	-	1.4	0.7	0.7	0.8
D^{*0} peaking bkg	-	-	< 0.1	-	-	-	2.0	-
$N_{B\bar{B}}$	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
$f_{+-,00}$	2.4	2.5	2.4	2.5	2.4	2.5	2.4	2.5
Intermediate $\mathcal{B}$s	0.8	1.7	1.6	1.1	0.8	1.7	0.6	1.1
Total systematic	4.4	6.1	7.1	5.7	4.6	5.1	7.8	5.4
Statistical	8.8	14.4	18.1	20.5	6.2	6.0	9.6	9.2

Table 3. Relative statistical uncertainties and breakdown of the relative systematic uncertainties to the corresponding branching fractions for the eight channels. All values are in percent. A dash is shown when the uncertainty is not applicable.

The contributions to the systematic uncertainties for each $B \to DD_s^-$ channel are summarized in table 4, expressed as relative uncertainties to the corresponding branching fractions. When feasible, the same procedure used for the $B \to DK^-K$ channels is applied. An additional systematic uncertainty is assigned for the $B \to DK^-K$ background estimation ("DKK bkg"). This uncertainty is obtained by varying the background yield by the statistical uncertainty of the relevant bin in the measured $m(K^-K)$ distribution. The systematic uncertainties are evaluated separately for the $D_s^- \to K^- K^0_S$ and $D_s^- \to K^- K^{*0}_S$ sub-channels and then a weighted average is provided.

The "Eff-MC sample size" systematic uncertainty is not correlated between the channels. The "Intermediate $\mathcal{B}$s" are partially correlated between the channels, according to the shared branching fractions. The uncertainty in the factor $f_{+-,00}$ is fully correlated between B^0 channels, and fully anti-correlated between B^- and B^0 channels. All other systematic uncertainties are fully correlated between all the relevant channels.

8 Invariant mass and helicity angles analysis for $B \to DK^-K$ channels

The signal yield as a function of $m(K^-K)$ is extracted using the signal sWeights as described in section 6 and applying the event-by-event efficiency correction using the map described in section 5.

The distributions of the θ_{KK} and θ_{K^*} helicity angles are extracted using the same approach. The former is defined as the angle between the K^{*0} momentum and the direction opposite to the B momentum, both calculated in the K^-K^{*0} system rest frame. The latter

Source	$B^- \to D^0 D_s^-$	$\overline{B}^0 \to D^+ D_s^-$	$B^- \to D^{*0} D_s^-$	$\overline{B}^0 \to D^{*+} D_s^-$
Eff. - MC sample size	< 0.1	< 0.1	< 0.1	< 0.1
Eff. - tracking	0.8	1.0	0.8	1.0
Eff. - π^+ from D^{*+}	-	-	-	2.7
Eff. - K_S^0	1.2	1.2	1.2	1.2
Eff. - PID	1.9	2.1	1.1	1.3
Eff. - π^0	-	-	5.1	-
Signal model	< 0.1	< 0.1	1.1	0.3
Bkg model	0.7	0.7	1.6	0.1
DKK bkg	1.7	2.1	6.1	4.5
D^{*0} peaking bkg	-	-	0.6	-
$N_{B\overline{B}}$	1.4	1.4	1.4	1.4
$f_{+-,00}$	2.4	2.5	2.4	2.5
Intermediate $\mathcal{B}$s	2.5	2.9	2.8	2.9
Total systematic	5.0	5.6	9.5	7.0
Statistical	6.0	5.9	15.2	11.9

Table 4. Relative statistical uncertainties and breakdown of the relative systematic uncertainties to the corresponding branching fractions for the four $B \to DD_s^-$ channels. All values are in percent. A dash is shown when the uncertainty is not applicable.

is defined as the angle between the momentum of the K^+ from K^{*0} decay and the direction opposite to the $K^- K^{*0}$ system momentum, both in the K^{*0} rest frame. The angles θ_{KK} and θ_{K_S} are defined analogously for the K_S^0 channels. There is no significant correlation between these angles and ΔE, as required for the sPlot technique.

The expected angular distributions $dN/d\theta$ are shown in table 5 for different hypotheses for the spin-parity J^P of the $K^- K_{(S)}^{(*)0}$ system, assuming a single J^P for the transition. Only the J^P states allowed by the factorization hypothesis and assuming exact isospin symmetry are shown. In the case of $J^P = 1^-$, and scalar to vector-vector decay (D^* channels), the distribution depends on the fraction of the longitudinal polarization in the B decay, so it is a mixture of different contributions with unknown strengths. Hence, the $dN/d\theta$ distribution is not known a priori for this hypothesis. In the case of $J^P = 1^+$, the D-wave contribution is assumed to be negligible compared to the S-wave [33], otherwise the angular distribution will be a mixture of different contributions as well. The θ_{K_S} distribution is not sensitive to the spin-parity of the $K^- K_S^0$ system and is expected to be uniform; thus, it is used to verify the absence of bias.

The various spin-parity hypotheses are tested by fitting the $dN/\cos\theta_{KK}$ and $dN/\cos\theta_{K_{(S)}^{(*)}}$ distributions, for each channel, with the functional forms corresponding to individual J^P hypotheses and comparing the χ^2 of the fits. The significance of the difference between the fit hypotheses is also estimated using the Akaike information criterion to compare non-nested models [34]. For the K_S^0 channels $dN/\theta_{KK} \propto \cos^2\theta$ and $dN/\theta_{K_S} \propto$ const are preferred, suggesting $J^P = 1^-$. For the K^{*0} channels $dN/\theta_{KK} \propto$ const and $dN/\theta_{K^*} \propto$ const are preferred, suggesting $J^P = 1^+$ or a non-resonant decay. The distributions of $dN/\cos\theta_{K_{(S)}^{(*)}}$

| | $K^- K^{*0}$ channels | | | | $K^- K^0_S$ channels | | | |
| | D^0, D^+ channels | | D^{*0}, D^{*+} channels | | D^0, D^+ channels | | D^{*0}, D^{*+} channels | |
J^P	$dN/d\theta_{KK}$	$dN/d\theta_{K*}$	$dN/d\theta_{KK}$	$dN/d\theta_{K*}$	$dN/d\theta_{KK}$	$dN/d\theta_{K_S}$	$dN/d\theta_{KK}$	$dN/d\theta_{K_S}$
Three-body	const	const	const	const	const	const	const	const
0^-	const	$\cos^2\theta$	const	$\cos^2\theta$	-	-	-	-
1^-	$\sin^2\theta$	$\sin^2\theta$	mix	$\sin^2\theta$	$\cos^2\theta$	const	mix	const
1^+	const†	const†	const†	const†	-	-	-	-

Table 5. Possible angular distributions given a specific spin-parity state of the $K^- K^{(*)0}_{(S)}$ system, subdividing between pseudoscalar channels (D^0, D^+) and vector channels (D^{*0}, D^{*+}). The hyphen (-) stands for a forbidden spin-parity assuming factorization and exact isospin symmetry; mix stands for a polarization dependent distribution; const stands for a uniform distribution; the † symbol indicates that the uniform distribution requires S-wave dominance.

and $dN/\cos\theta_{KK}$ for the eight channels are shown in figure 7 and figure 8; the fits projections are overlaid. Additional details about the fits are provided in the appendix A.

The $m(K^- K)$ distributions for data are shown in figure 9 and figure 10. The $m(K^- K)$ distribution from the signal simulation is also overlaid to show a comparison with phase-space and specific resonant distributions. In all channels, the bulk of the observed $m(K^- K^0_S)$ distributions is located at low $m(K^- K^0_S)$ values, and show structures different from phase-space distributions. The distributions vanish above 2.0–2.5 GeV, disfavoring the presence of a significant phase-space component. The overlay with the $\rho(1450)^-$ meson lineshape shows partial agreement of the peak position. Therefore, combining the observed $m(K^- K)$ distribution and the helicity angle constraint, we conclude that the four K^0_S channels proceed predominantly via $B \to D\rho'^-(\to K^- K^0_S)$, where ρ' indicates one or more ρ−like resonance. This interpretation is supported by ref. [35], where the $m(K^- K^0_S)$ distribution is predominantly described as the combination of the $\rho(1450)^-$ and $\rho(770)^-$ contributions. Moreover, we cannot exclude the interference of even-spin states, which are allowed if factorization is not assumed. In this case, the transition can proceed via a color-suppressed diagram with internal W boson emission, and could result in a $J^P = 0^+$ state. However, from ref. [35] this contribution is expected to be small or negligible. For the K^{*0} channels, the observed $m(K^- K^{*0})$ distribution also differs from phase-space, with the peak position and the high mass tail in good agreement with the simulated $a_1(1640)^-$ meson lineshape, thus strongly disfavoring a non-resonant transition. Given the imperfect agreement with a pure $a_1(1640)^-$ lineshape, we cannot exclude the possibility that the observed lineshape is due to the superposition of multiple a_1^- resonances. However, the $a_1(1260)^-$ is unlikely to be the dominant contribution given the observed peak position, despite the large uncertainties related to the knowledge of the $a_1(1260)^-$ width. Combining the observed $m(K^- K^{*0})$ distribution and the information from the helicity angles, we conclude that the four K^{*0} channels proceed via $B \to D a_1'^-(\to K^- K^{*0})$, where a_1' stands for one or multiple a_1-like resonances.
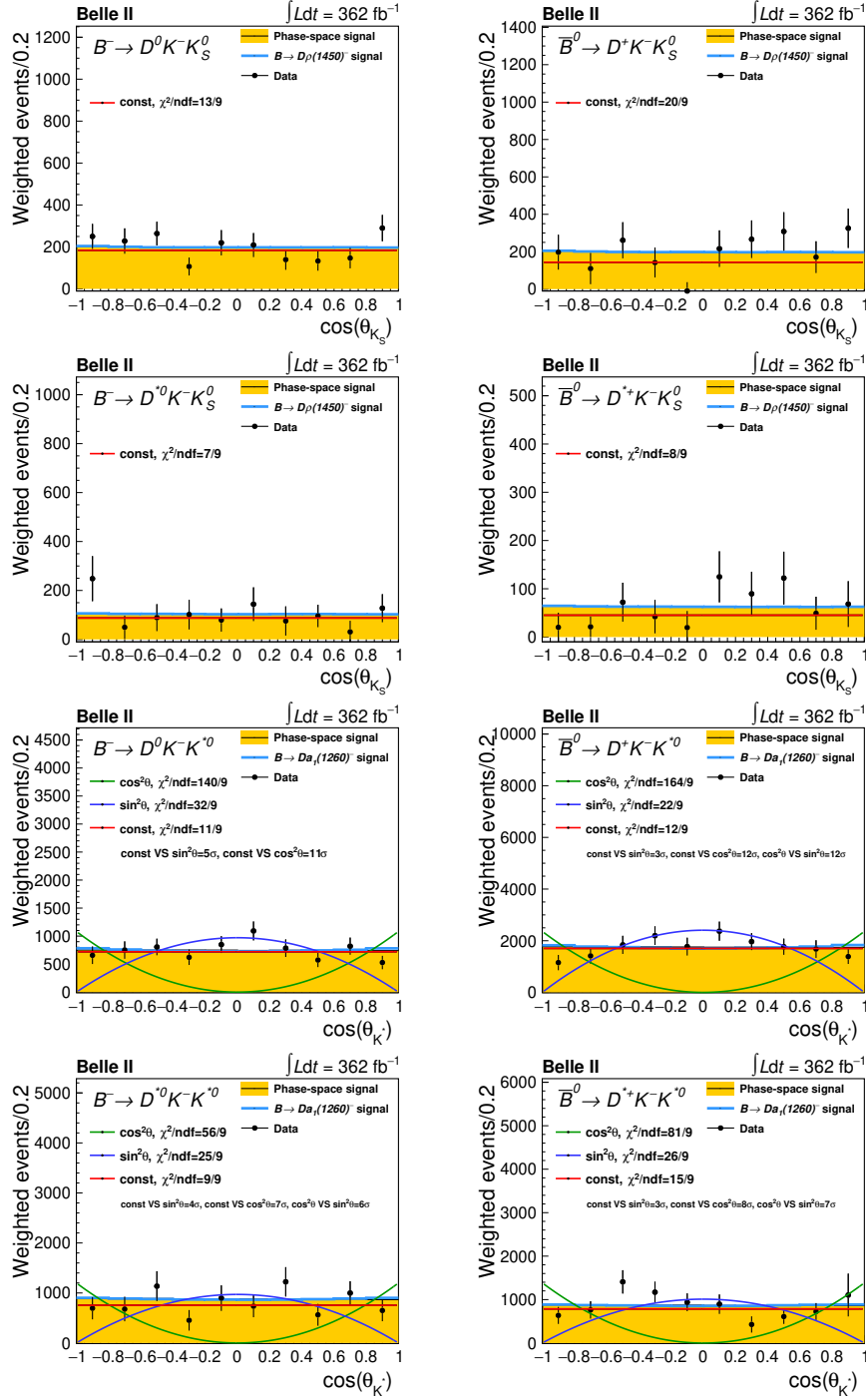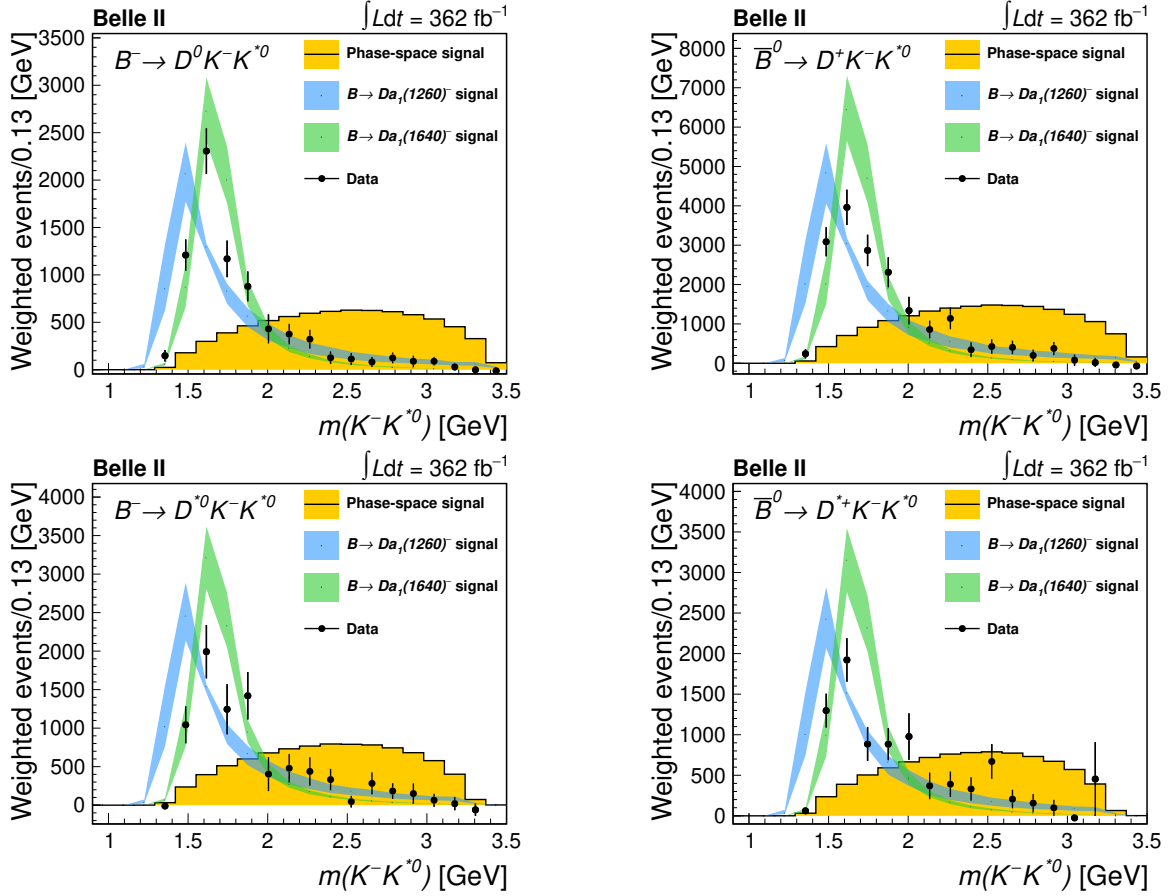
Figure 7. Background-subtracted and efficiency-corrected distribution of $dN/d\cos\theta_{K_{(S)}^{(*)}}$ for the $B^- \to D^0 K^- K_S^0$ (first line, left), $\overline{B}^0 \to D^+ K^- K_S^0$ (first line, right), $B^- \to D^{*0} K^- K_S^0$ (second line, left), $\overline{B}^0 \to D^{*+} K^- K_S^0$ (second line, right), $B^- \to D^0 K^- K^{*0}$ (third line, left), $\overline{B}^0 \to D^+ K^- K^{*0}$ (third line, right), $B^- \to D^{*0} K^- K^{*0}$ (fourth line, left), and $\overline{B}^0 \to D^{*+} K^- K^{*0}$ (fourth line, right) channels. The error bars represent the statistical uncertainty. A phase-space MC simulation and a resonant MC simulation at generator level, rescaled to the integral of the data distribution, are also shown for comparison. The three fit-hypotheses and the respective χ^2 are overlaid, together with the significance of the difference between each fit and the other two.

Figure 8. Background-subtracted and efficiency-corrected distribution of $dN/d\cos\theta_{KK}$ for the $B^- \to D^0 K^- K_S^0$ (first line, left), $\overline{B}^0 \to D^+ K^- K_S^0$ (first line, right), $B^- \to D^{*0} K^- K_S^0$ (second line, left), $\overline{B}^0 \to D^{*+} K^- K_S^0$ (second line, right), $B^- \to D^0 K^- K^{*0}$ (third line, left), $\overline{B}^0 \to D^+ K^- K^{*0}$ (third line, right), $B^- \to D^{*0} K^- K^{*0}$ (fourth line, left), and $\overline{B}^0 \to D^{*+} K^- K^{*0}$ (fourth line, right) channel. The error bars represent the statistical uncertainty. A phase-space MC simulation and a resonant MC simulation at generator level, rescaled to the integral of the data distribution, are also shown for comparison. The three fit-hypotheses and the respective χ^2 are overlaid, together with the significance of the difference between each fit and the other two.
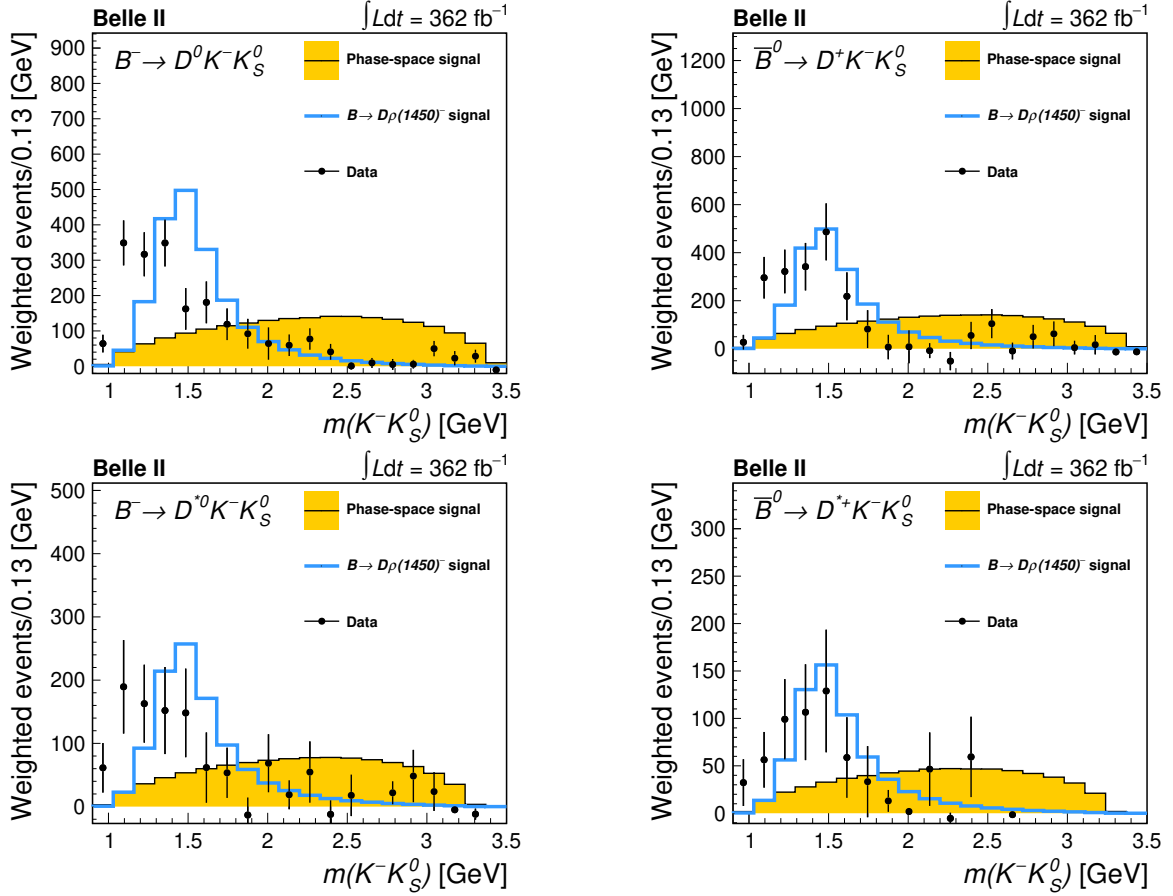
Figure 9. Background-subtracted and efficiency-corrected distribution of $m(K^-K_S^0)$ for the $B^- \to D^0 K^- K_S^0$ (top left), $\overline{B}^0 \to D^+ K^- K_S^0$ (top right), $B^- \to D^{*0} K^- K_S^0$ (bottom left), and $\overline{B}^0 \to D^{*+} K^- K_S^0$ (bottom right) channels. The error bars represent the statistical uncertainties. A phase-space MC simulation and a resonant MC simulation at generator level, rescaled to the integral of the data distribution, are also shown for comparison.

Figure 10. Background-subtracted and efficiency-corrected distribution of $m(K^- K^{*0})$ for the $B^- \to D^0 K^- K^{*0}$ (top, left), $\overline{B}^0 \to D^+ K^- K^{*0}$ (top right), $B^- \to D^{*0} K^- K^{*0}$ (bottom left), and $\overline{B}^0 \to D^{*+} K^- K^{*0}$ (bottom right) channels. The error bars represent the statistical uncertainties. A phase-space MC simulation and two resonant MC simulation at generator level, rescaled to the integral of the data distribution, are also shown for comparison. The uncertainties on the resonance parameters simulations are shown as shaded areas.

9 Conclusions

Using an electron-positron data sample collected by Belle II on the $\Upsilon(4S)$ resonance with an integrated luminosity of 362 fb^{-1}, we report the measurement of the eight branching fractions

$$\mathcal{B}(B^- \to D^0 K^- K_S^0) = (1.82 \pm 0.16 \pm 0.08) \times 10^{-4},$$
$$\mathcal{B}(\overline{B}^0 \to D^+ K^- K_S^0) = (0.82 \pm 0.12 \pm 0.05) \times 10^{-4},$$
$$\mathcal{B}(B^- \to D^{*0} K^- K_S^0) = (1.47 \pm 0.27 \pm 0.10) \times 10^{-4},$$
$$\mathcal{B}(\overline{B}^0 \to D^{*+} K^- K_S^0) = (0.91 \pm 0.19 \pm 0.05) \times 10^{-4}$$
$$\mathcal{B}(B^- \to D^0 K^- K^{*0}) = (7.19 \pm 0.45 \pm 0.33) \times 10^{-4},$$
$$\mathcal{B}(\overline{B}^0 \to D^+ K^- K^{*0}) = (7.56 \pm 0.45 \pm 0.38) \times 10^{-4},$$
$$\mathcal{B}(B^- \to D^{*0} K^- K^{*0}) = (11.93 \pm 1.14 \pm 0.93) \times 10^{-4},$$
$$\mathcal{B}(\overline{B}^0 \to D^{*+} K^- K^{*0}) = (13.12 \pm 1.21 \pm 0.71) \times 10^{-4}$$

where the first uncertainty is statistical and the second uncertainty is systematic. The decays $\overline{B}^0 \to D^+ K^- K_S^0$, $B^- \to D^{*0} K^- K_S^0$, and $\overline{B}^0 \to D^{*+} K^- K_S^0$ are observed for the first time. The precision of the eight measurements is between 5% and 20% and is limited by statistical uncertainties. The precision of $\mathcal{B}(B^- \to D^0 K^- K_S^0)$ and the four $\mathcal{B}(\overline{B} \to DK^- K^{*0})$ are improved by more than a factor of three compared to previous measurements [2].

The observed $m(K^- K_{(S)}^{(*)0})$ invariant mass and helicity angles distributions suggest the presence of a dominant resonant component in the $K^- K_{(S)}^{(*)0}$ system as has been previously reported by Belle [2]. In $B \to DK^- K_S^0$ decays, a $B \to D\rho'^-(\to K^- K_S^0)$ transition is favored, where ρ' stands for a $J^P = 1^-$ state. The transitions are likely to proceed via multiple resonant states, and we cannot exclude the presence of an even-spin intermediate state. In $B \to DK^- K^{*0}$ decays a $B \to Da_1'^-(\to K^- K^{*0})$ transition is favored, where a_1' stands for a $J^P = 1^+$ state. The observed $m(KK^{*0})$ distributions favor the contribution of not only the $a_1(1260)^-$, but also excited a_1^- states such as the $a_1(1640)^-$, in contrast to the earlier Belle measurement [2].

These eight channels can be exploited in the FEI B-tagging algorithm, given their high purity. They can provide a few percent improvement in efficiency for hadronic FEI.

We also provide measurements of the branching fractions of the four channels $B \to DD_s^-$, obtained from the same data sample. They are found to be

$$\mathcal{B}(B^- \to D^0 D_s^-) = (95 \pm 6 \pm 5) \times 10^{-4},$$
$$\mathcal{B}(\overline{B}^0 \to D^+ D_s^-) = (89 \pm 5 \pm 5) \times 10^{-4},$$
$$\mathcal{B}(B^- \to D^{*0} D_s^-) = (65 \pm 10 \pm 6) \times 10^{-4},$$
$$\mathcal{B}(\overline{B}^0 \to D^{*+} D_s^-) = (83 \pm 10 \pm 6) \times 10^{-4}$$

where the first uncertainty is statistical and the second one systematic. These measurements have a precision similar to or competitive with the current world averages [31].

A   Additional material

In figure 11 the distribution of $m(K^+\pi^-)$ invariant mass for the four K^{*0} channels is shown. The projection of the fit, as described in section 4, is overlaid. The non-K^{*0}-resonant fraction is listed in table 6 for the four K^{*0} channels. The systematic uncertainties are estimated as described in section 7.

In figure 7 and figure 8 the distributions of $dN/d\cos\theta_{K^{(*)}_{(S)}}$ and $dN/d\cos\theta_{KK}$ respectively, are shown for the eight $B \to DK^-K$ channels. The tested fit hypotheses, according to table 5, are also shown. These fits models have only the total yield as a free parameter and they consider statistical uncertainty only. A phase-space MC simulation and a specific resonant MC are overlaid as a reference. For the K^0_S channels, the four $\cos(\theta_{K_S})$ fit show good agreement with the uniform distribution, as expected. The $\cos(\theta_{KK})$ distributions for the $\overline{B}^0 \to D^+K^-K^0_S$ and $\overline{B}^0 \to D^{*+}K^-K^0_S$ channels follow a $\cos^2\theta$ distribution, in agreement with $J^P = 1^-$. On the other hand, the $\cos(\theta_{KK})$ distributions for the $B^- \to D^0K^-K^0_S$ and $B^- \to D^{*0}K^-K^0_S$ channels disagree with all the tested hypotheses. The unknown polarization of the D^{*0} justifies the observed distribution for the $B^- \to D^{*0}K^-K^0_S$ channel. The observed asymmetric $\cos(\theta_{KK})$ distribution for the $B^- \to D^0K^-K^0_S$ channel supports the hypothesis of interference between the $J^P = 1^-$ state and spin-even states. For the K^{*0} channels, both the $\cos(\theta_{K^*})$ and $\cos(\theta_{KK})$ distributions are in good agreement with the uniform distribution, as expected for $J^P = 1^+$, but for the $\cos(\theta_{KK})$ of the $B^0 \to D^{*+}K^-K^{*0}$ channel.

The $(m(K^-K), m(DK))$ distributions are shown in figure 12 for the eight $B \to DK^-K$ channels. The selection is applied together with the efficiency correction; sWeights are applied to subtract the background component. The distributions also show some structures, which are compatible with the resonant transition discussed in section 8. The $m(DK)$ and $m(DK^-)$ projections show that the reflection of the $m(K^-K)$ resonant lineshape may produce such structures. These distributions are shown in figure 13 and figure 14.

Channel	non-K^{*0}-resonant fraction [%]
$B^- \to D^0 K^- K^{*0}$	$3.1 \pm 0.5 \pm 1.2$
$\overline{B}^0 \to D^+ K^- K^{*0}$	$0.7 \pm 0.5 \pm 0.3$
$B^- \to D^{*0} K^- K^{*0}$	$1.4 \pm 0.6 \pm 0.3$
$\overline{B}^0 \to D^{*+} K^- K^{*0}$	$1.0 \pm 0.5 \pm 0.5$

Table 6. Non-K^{*0}-resonant fraction, in percent of the signal, estimated for each of the four K^{*0} channels. The first uncertainty is statistical, the second systematic.
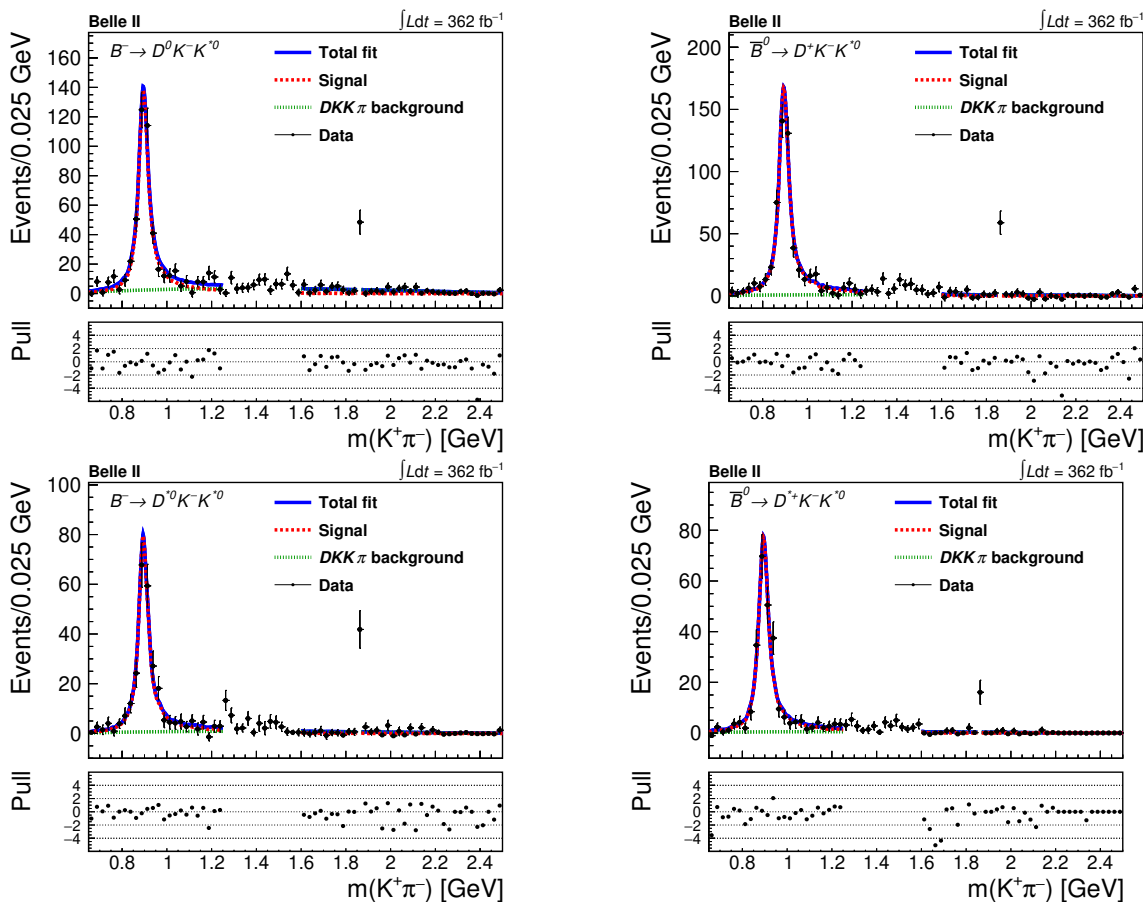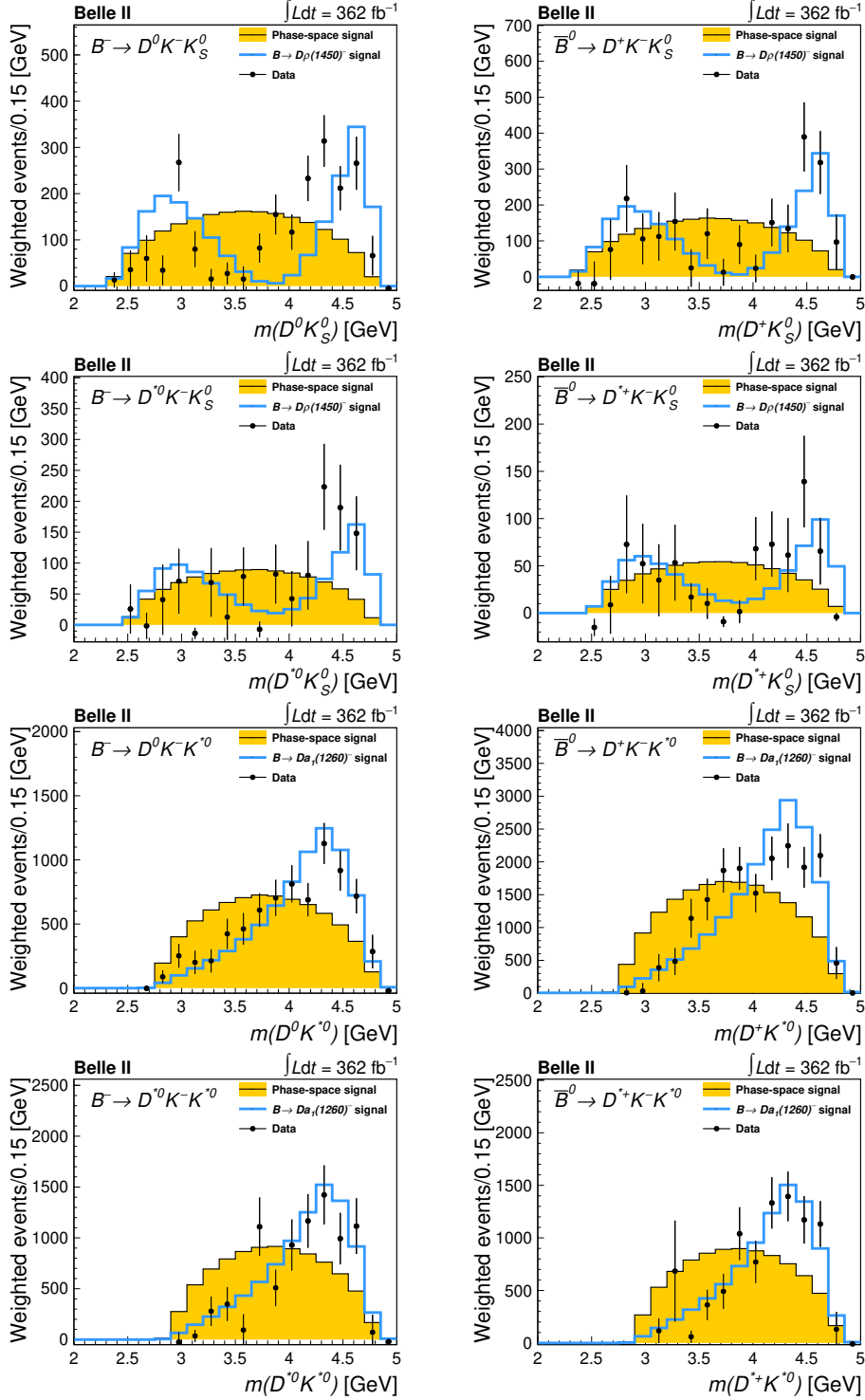

Figure 11. Distribution of $m(K^+\pi^-)$ for the $B^- \to D^0 K^- K^{*0}$ (top left), $\overline{B}^0 \to D^+ K^- K^{*0}$ (top right), $B^- \to D^{*0} K^- K^{*0}$ (bottom left), and $\overline{B}^0 \to D^{*+} K^- K^{*0}$ (bottom right) channels. The projections of the fits are overlaid, the fit components are highlighted, and the pulls between the fit and the data are shown in the bottom panel of each plot. The background $B\overline{B}$ and continuum background are subtracted by applying the signal sWeights.
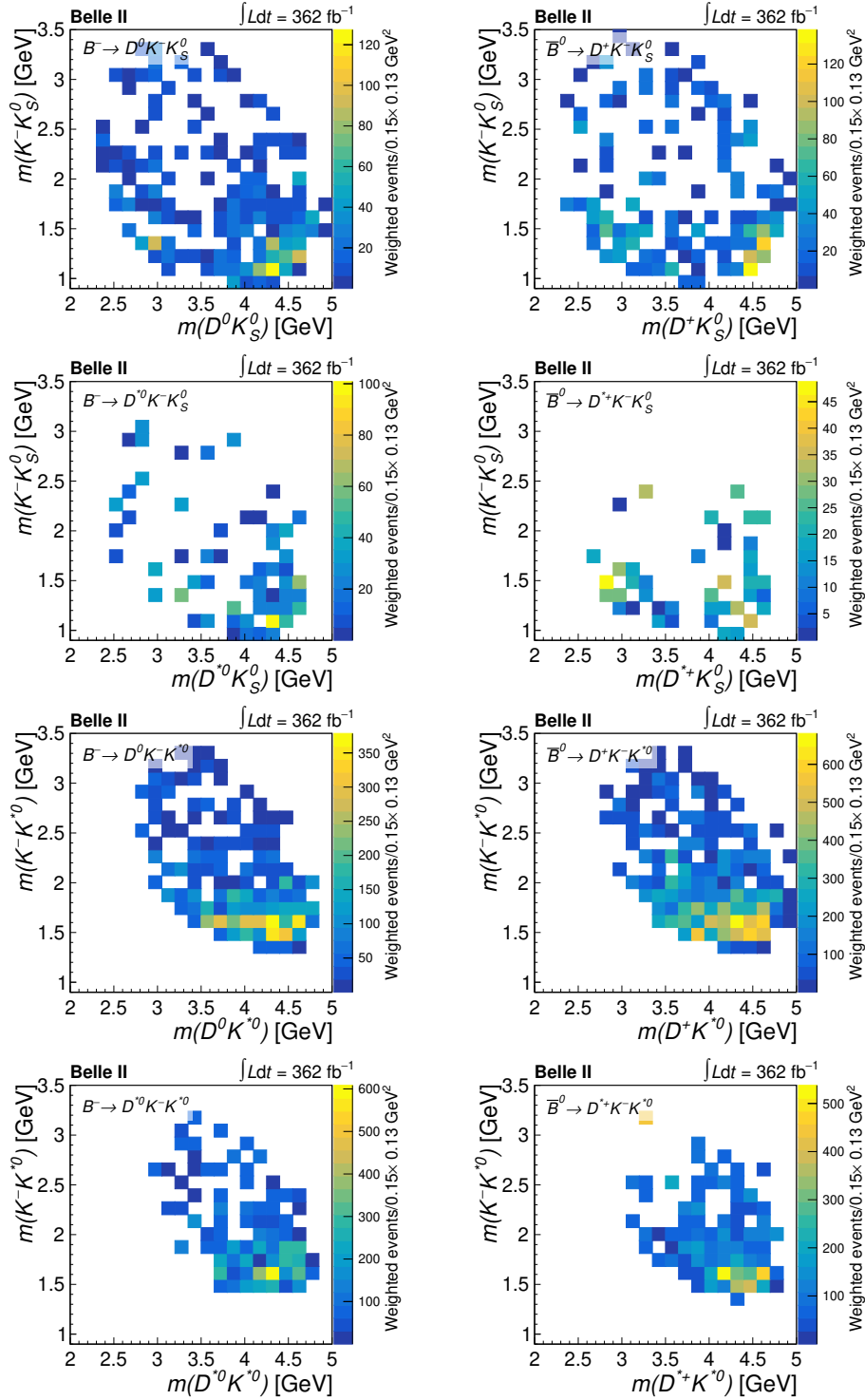
Figure 12. Background-subtracted and efficiency-corrected distribution of $\left(m(K^-K), m(DK)\right)$ for the $B^- \to D^0 K^- K_S^0$ (first line, left), $\overline{B}^0 \to D^+ K^- K_S^0$ (first line, right), $B^- \to D^{*0} K^- K_S^0$ (second line, left), $\overline{B}^0 \to D^{*+} K^- K_S^0$ (second line, right), $B^- \to D^0 K^- K^{*0}$ (third line, left), $\overline{B}^0 \to D^+ K^- K^{*0}$ (third line, right), $B^- \to D^{*0} K^- K^{*0}$ (fourth line, left), and $\overline{B}^0 \to D^{*+} K^- K^{*0}$ (fourth line, right) channels.
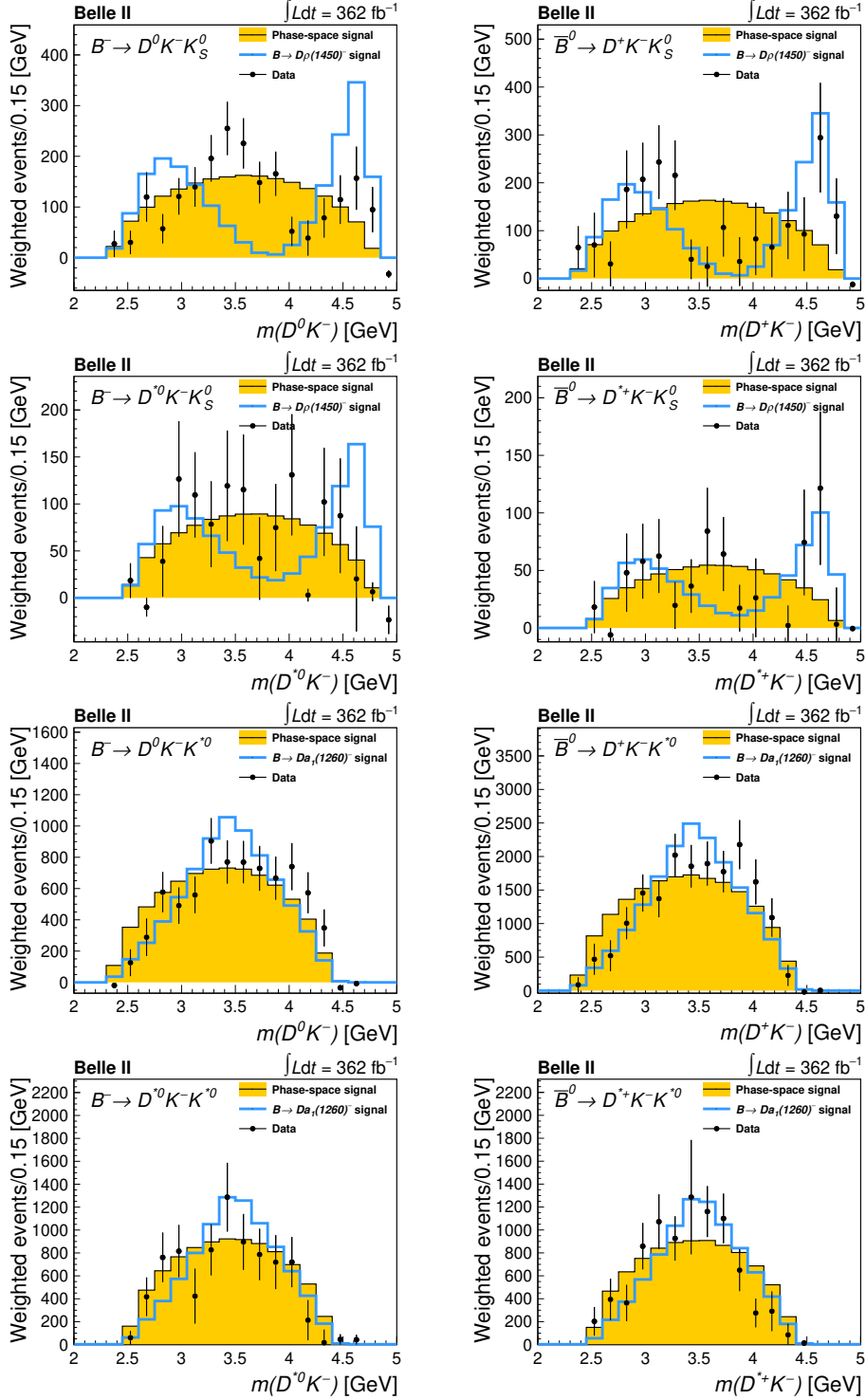
Figure 13. Background-subtracted and efficiency-corrected distribution of $m(DK)$ for the $B^- \to D^0 K^- K_S^0$ (first line, left), $\overline{B}^0 \to D^+ K^- K_S^0$ (first line, right), $B^- \to D^{*0} K^- K_S^0$ (second line, left), $\overline{B}^0 \to D^{*+} K^- K_S^0$ (second line, right), $B^- \to D^0 K^- K^{*0}$ (third line, left), $\overline{B}^0 \to D^+ K^- K^{*0}$ (third line, right), $B^- \to D^{*0} K^- K^{*0}$ (fourth line, left), and $\overline{B}^0 \to D^{*+} K^- K^{*0}$ (fourth line, right) channels. The error bars represent the statistical uncertainty. A phase-space MC simulation and a resonant MC simulation at generator level, rescaled to the integral of the data distribution, are also shown for comparison.

Figure 14. Background-subtracted and efficiency-corrected distribution of $m(DK^-)$ for the $B^- \to D^0K^-K_S^0$ (first line, left), $\overline{B}^0 \to D^+K^-K_S^0$ (first line, right), $B^- \to D^{*0}K^-K_S^0$ (second line, left), $\overline{B}^0 \to D^{*+}K^-K_S^0$ (second line, right), $B^- \to D^0K^-K^{*0}$ (third line, left), $\overline{B}^0 \to D^+K^-K^{*0}$ (third line, right), $B^- \to D^{*0}K^-K^{*0}$ (fourth line, left), and $\overline{B}^0 \to D^{*+}K^-K^{*0}$ (fourth line, right) channels. The error bars represent the statistical uncertainty. A phase-space MC simulation and a resonant MC simulation at generator level, rescaled to the integral of the data distribution, are also shown for comparison.

References

[1] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [`arXiv:1410.3012`] [INSPIRE].

[2] BELLE collaboration, *Observation of $B \to D^{(*)}K^-K^{0(*)}$ decays*, *Phys. Lett. B* **542** (2002) 171 [`hep-ex/0207041`] [INSPIRE].

[3] M. Bauer, B. Stech and M. Wirbel, *Exclusive nonleptonic decays of D, D_s, and B mesons*, *Z. Phys. C* **34** (1987) 103 [INSPIRE].

[4] A.-J. Ma and W.-F. Wang, *Contributions of the kaon pair from $\rho(770)$ for the three-body decays $B \to DK\bar{K}$*, *Phys. Rev. D* **103** (2021) 016002 [`arXiv:2010.12906`] [INSPIRE].

[5] A.-J. Ma and W.-F. Wang, *Resonances $\rho(1450)^+$ and $\rho(1700)^+$ in $B \to DKK$ decays*, *Chin. Phys. C* **46** (2022) 053104 [`arXiv:2201.06881`] [INSPIRE].

[6] S. Clymton and H.-C. Kim, *Molecular nature of the $a_1(1260)$ axial-vector meson*, *Phys. Rev. D* **106** (2022) 114015 [`arXiv:2208.04124`] [INSPIRE].

[7] L.R. Dai, R. Molina and E. Oset, *Looking for the exotic $X^0(2866)$ and its $J^P = 1^+$ partner in the $B^0 \to D^{(*)+}K^-K^{(*)0}$ reactions*, *Phys. Rev. D* **105** (2022) 096022 [`arXiv:2202.11973`] [INSPIRE].

[8] M. Diehl, *Designer mesons for exploring factorization in b decays*, *PoS* **HEP2001** (2001) 104 [`hep-ph/0111009`] [INSPIRE].

[9] BELLE II collaboration, *The Belle II physics book*, *PTEP* **2019** (2019) 123C01 [*Erratum ibid.* **2020** (2020) 029201] [`arXiv:1808.10567`] [INSPIRE].

[10] T. Keck et al., *The full event interpretation: an exclusive tagging algorithm for the Belle II experiment*, *Comput. Softw. Big Sci.* **3** (2019) 6 [`arXiv:1807.08680`] [INSPIRE].

[11] R.H. Dalitz, *On the analysis of tau-meson data and the nature of the tau-meson*, *Phil. Mag. Ser. 7* **44** (1953) 1068 [INSPIRE].

[12] LHCb collaboration, *First observations of $\bar{B}_s^0 \to D^+D^-$, $D_s^+D^-$ and $D^0\bar{D}^0$ decays*, *Phys. Rev. D* **87** (2013) 092007 [`arXiv:1302.5854`] [INSPIRE].

[13] LHCb collaboration, *Study of beauty hadron decays into pairs of charm hadrons*, *Phys. Rev. Lett.* **112** (2014) 202001 [`arXiv:1403.3606`] [INSPIRE].

[14] LHCb collaboration, *Search for the decays $B_s^0 \to \tau^+\tau^-$ and $B^0 \to \tau^+\tau^-$*, *Phys. Rev. Lett.* **118** (2017) 251802 [`arXiv:1703.02508`] [INSPIRE].

[15] LHCb collaboration, *Measurement of CP asymmetries and branching fraction ratios of B^- decays to two charm mesons*, *JHEP* **09** (2023) 202 [`arXiv:2306.09945`] [INSPIRE].

[16] M. Pivk and F.R. Le Diberder, *SPlot: a statistical tool to unfold data distributions*, *Nucl. Instrum. Meth. A* **555** (2005) 356 [`physics/0402083`] [INSPIRE].

[17] BELLE II collaboration, *Belle II technical design report*, `arXiv:1011.0352` [INSPIRE].

[18] SUPERKEKB collaboration, *SuperKEKB collider*, *Nucl. Instrum. Meth. A* **907** (2018) 188 [`arXiv:1809.01958`] [INSPIRE].

[19] BELLE II SVD collaboration, *The design, construction, operation and performance of the Belle II silicon vertex detector*, 2022 *JINST* **17** P11042 [`arXiv:2201.09824`] [INSPIRE].

[20] D. Kotchetkov et al., *Front-end electronic readout system for the Belle II imaging time-of-propagation detector*, *Nucl. Instrum. Meth. A* **941** (2019) 162342 [`arXiv:1804.10782`] [INSPIRE].

[21] S. Jadach, B.F.L. Ward and Z. Was, *The precision Monte Carlo event generator KK for two fermion final states in e^+e^- collisions*, *Comput. Phys. Commun.* **130** (2000) 260 [`hep-ph/9912214`] [ɪɴSPIRE].

[22] A. Ryd et al., *EvtGen: a Monte Carlo generator for B-physics*, EVTGEN-V00-11-07 (2005) [ɪɴSPIRE].

[23] D.J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152 [ɪɴSPIRE].

[24] GEANT4 collaboration, *GEANT4 — a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250 [ɪɴSPIRE].

[25] Z.J. Liptak et al., *Measurements of beam backgrounds in SuperKEKB phase 2*, *Nucl. Instrum. Meth. A* **1040** (2022) 167168 [`arXiv:2112.14537`] [ɪɴSPIRE].

[26] BELLE II FRAMEWORK SOFTWARE GROUP collaboration, *The Belle II core software*, *Comput. Softw. Big Sci.* **3** (2019) 1 [`arXiv:1809.04299`] [ɪɴSPIRE].

[27] BELLE II collaboration, *Belle II Analysis Software Framework (BASF2)*, Zenodo, (2022).

[28] G.C. Fox and S. Wolfram, *Event shapes in e^+e^- annihilation*, *Nucl. Phys. B* **149** (1979) 413 [*Erratum ibid.* **157** (1979) 543] [ɪɴSPIRE].

[29] E. Farhi, *A QCD test for jets*, *Phys. Rev. Lett.* **39** (1977) 1587 [ɪɴSPIRE].

[30] W. Verkerke and D.P. Kirkby, *The RooFit toolkit for data modeling*, *eConf C* **0303241** (2003) MOLT007 [`physics/0306116`] [ɪɴSPIRE].

[31] PARTICLE DATA GROUP collaboration, *Review of particle physics*, *PTEP* **2022** (2022) 083C01 [ɪɴSPIRE].

[32] BELLE collaboration, *Measurement of the B^+/B^0 production ratio in e^+e^- collisions at the $\Upsilon(4S)$ resonance using $B \to J/\psi(\ell\ell)K$ decays at Belle*, *Phys. Rev. D* **107** (2023) L031102 [`arXiv:2207.01194`] [ɪɴSPIRE].

[33] PARTICLE DATA GROUP collaboration, *Review of particle physics*, *PTEP* **2022** (2022) 083C01 [ɪɴSPIRE].

[34] E.-J. Wagenmakers and S. Farrell, *AIC model selection using Akaike weights*, *Psychon. Bull. Rev.* **11** (2004) 192.

[35] W.-F. Wang, L.-F. Yang, A.-J. Ma and À. Ramos, *The low-mass enhancement of kaon pairs in $B^+ \to \bar{D}^{(*)0}K^+\bar{K}^0$ and $B^0 \to D^{(*)-}K^+\bar{K}^0$ decays*, *Phys. Rev. D* **109** (2024) 116009 [`arXiv:2403.07499`] [ɪɴSPIRE].