Updated branching fraction measurements of $B[s(s)][KShh$ decays

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Updated branching fraction measurements of 
$B_{(s)}^{0} \rightarrow K_{S}^{0} h^{+} h'^{-}$ decays

The LHCb collaboration

E-mail: louis.henry@cern.ch

ABSTRACT: The charmless three-body decays $B_{(s)}^{0} \rightarrow K_{S}^{0} h^{+} h'^{-}$ (where $h'^{-} = \pi, K$) are analysed using a sample of $pp$ collision data recorded by the LHCb experiment, corresponding to an integrated luminosity of $3 \text{ fb}^{-1}$. The branching fractions are measured relative to that of the $B^{0} \rightarrow K_{S}^{0} \pi^{+} \pi^{-}$ decay, and are determined to be:

$$
\frac{\mathcal{B}(B^{0} \rightarrow K_{S}^{0} K^{+} K^{-})}{\mathcal{B}(B^{0} \rightarrow K_{S}^{0} \pi^{+} \pi^{-})} = 0.549 \pm 0.018 \text{ (stat)} \pm 0.033 \text{ (syst)},
$$

$$
\frac{\mathcal{B}(B_{s}^{0} \rightarrow K_{S}^{0} \pi^{+} \pi^{-})}{\mathcal{B}(B^{0} \rightarrow K_{S}^{0} \pi^{+} \pi^{-})} = 0.191 \pm 0.027 \text{ (stat)} \pm 0.031 \text{ (syst)} \pm 0.011 \left(\frac{f_{s}}{f_{d}}\right),
$$

$$
\frac{\mathcal{B}(B_{s}^{0} \rightarrow K_{S}^{0} K^{+} K^{-})}{\mathcal{B}(B^{0} \rightarrow K_{S}^{0} \pi^{+} \pi^{-})} = 1.70 \pm 0.07 \text{ (stat)} \pm 0.11 \text{ (syst)} \pm 0.10 \left(\frac{f_{s}}{f_{d}}\right),
$$

where $f_{s}/f_{d}$ represents the ratio of hadronisation fractions of the $B_{s}^{0}$ and $B^{0}$ mesons.

KEYWORDS: B physics, Branching fraction, Flavor physics, Hadron-Hadron scattering (experiments)

ArXiv ePrint: 1707.01665
1 Introduction

The measurement of $CP$-violation observables in the decays $B^0 \rightarrow K^0_s \pi^+ \pi^-$ and $B^0 \rightarrow K^0_s K^+ K^-$, which are dominated by $b \rightarrow q\bar{q}s$ ($q = u, d, s$) loop transitions, are of great theoretical interest. In particular, the mixing-induced $CP$ asymmetries in these decays are predicted by the Standard Model (SM) Cabibbo-Kobayashi-Maskawa mechanism [1, 2] to be approximately equal to those governed by $b \rightarrow c\bar{c}s$ transitions, such as $B^0 \rightarrow J/\psi K^0_s$. Within the SM the weak phase measurements in $b \rightarrow q\bar{q}s$ decays are expected to deviate from the values determined in $b \rightarrow c\bar{c}s$ decays but for certain contributions to these decays, such as $B^0 \rightarrow \phi K^0_s$ and $B^0 \rightarrow \rho^0 K^0_s$, this deviation is either expected to be small or can be controlled using flavour symmetries [3–5]. The existence of new particles predicted in several extensions of the SM could introduce additional weak phases that contribute along with the SM mixing phase to the amplitudes of these loop-dominated charmless decays.

---

1Unless stated otherwise, charge conjugated modes are implicitly included throughout this article.
potentially leading to much greater deviations from the $b \rightarrow c \bar{c}s$ values $^{[6-8]}$. The mixing-induced CP-violating phase can be measured by means of a flavour-tagged time-dependent analysis of the three-body Dalitz plot of these decays $^{[9-12]}$. The current experimental measurements of this phase in $b \rightarrow q\bar{q}s$ decays $^{[13]}$ show a generally good agreement with the results for the weak phase $\beta$ from $b \rightarrow c \bar{c}s$ decays for each of the CP eigenstates studied. The experimental uncertainties are, however, currently rather larger than the size of the expected deviations, both in the SM and beyond-the-SM scenarios, and so there is a need for more precise measurements of these quantities. A similar determination of the mixing-induced CP-violating phase in the $B_0^*$ system is possible with, among others, the $B_0^s \rightarrow K_0^* K^{\pm}\pi^\mp$ decays $^{[14]}$.

It is also possible to determine the CKM angle $\gamma$ by combining information from several $B \rightarrow K_h h'$ decays, using either the methods originally proposed in refs. $^{[15, 16]}$ and recently developed further in ref. $^{[17]}$, or those proposed in refs. $^{[18-20]}$. The existing experimental results, which come from the BaBar collaboration $^{[21, 22]}$, demonstrate the feasibility of the measurement, albeit with large statistical uncertainties. The decay $B_0^s \rightarrow K_0^* \pi^+\pi^-$ is dominated by tree-level processes and as such is of particular interest for this effort, with the potential to yield a theoretically clean determination of $\gamma$ $^{[23]}$.

The measurements of the branching fractions themselves are of great importance in order to confront theoretical predictions. These predictions are based on various approaches to modelling the hadronisation processes, such as QCD factorisation or PQCD, see for example refs. $^{[24-29]}$. Comparison of the different approaches with the experimental data will allow further refinement of the theoretical models, which in turn will yield improved predictions of branching fractions and CP asymmetries of these and many other charmless decay modes. In addition, these results can be used to test the level of breaking of the flavour symmetries: isospin, U-spin and SU(3), see for example ref. $^{[30]}$.

Of the decays of neutral $B$ mesons to $K_0^* \pi^+\pi^-$, $K_0^* K^{\pm}\pi^\mp$ and $K_0^* K^+ K^-$ final states, only the decay $B_0^s \rightarrow K_0^* K^+ K^-$ remains to be observed $^{[10, 12, 31-34]}$. Most recently, a search for the three $B_0^s$ decays was reported by the LHCb experiment using the 1 fb$^{-1}$ data sample recorded in 2011 $^{[34]}$. While first observations were made for the $B_0^s \rightarrow K_0^* \pi^+\pi^-$ and $B_0^s \rightarrow K_0^* K^{\pm}\pi^\mp$ modes, no evidence for the decay $B_0^s \rightarrow K_0^* K^+ K^-$ was found. In this work, all the aforementioned charmless three-body decays of the $B_0^s$ and $B_0^0$ mesons are studied using the $pp$ collision data recorded by the LHCb detector, corresponding to an integrated luminosity of 1.0 fb$^{-1}$ at a centre-of-mass energy of 7 TeV in 2011 and 2.0 fb$^{-1}$ at a centre-of-mass energy of 8 TeV in 2012. This sample is three times larger than that used in ref. $^{[34]}$. The measurements of the time-integrated branching fractions $^{[35]}$ relative to that of $B_0^s \rightarrow K_0^* \pi^+\pi^-$ are presented. The notation $B(B_0^s \rightarrow K_0^* K^{\pm}\pi^\mp)$ is used throughout the document to indicate the sum of the branching fractions $B(B_0^s \rightarrow K_0^* K^+\pi^-)$ and $B(B_0^s \rightarrow K_0^* K^-\pi^+)$, and similarly for the corresponding $B_0^0$ decays.

## 2 Detector and simulation

The LHCb detector $^{[36, 37]}$ is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing $b$ or $c$
quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector (VELO) surrounding the $pp$ interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of momentum, $p$, of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 $\text{GeV}/c$. The minimum distance of a track to a primary vertex (PV), the impact parameter (IP), is measured with a resolution of $(15 + 29/p_T) \mu m$, where $p_T$ is the component of the momentum transverse to the beam, in $\text{GeV}/c$. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers.

Simulated data samples are used to investigate backgrounds from other $b$-hadron decays and also to study the detection and reconstruction efficiency of the signal. In the simulation, $pp$ collisions are generated using PYTHIA $^{[38, 39]}$ with a specific LHCb configuration $^{[40]}$. Decays of hadronic particles are described by EVTGEN $^{[41]}$, in which final-state radiation is generated using PHOTOS $^{[42]}$. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit $^{[43, 44]}$ as described in ref. $^{[45]}$.

3 Trigger and event selection

The online event selection is performed by a trigger $^{[46]}$, which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, in which all charged particles with $p_T > 500 (300) \text{MeV}/c$ are reconstructed for data collected in 2011 (2012). At the hardware trigger stage, events are required to have a muon with high $p_T$ or a hadron, photon or electron with high transverse energy in the calorimeters. The software trigger requires a two-, three- or four-track secondary vertex with a significant displacement from all primary $pp$ interaction vertices. At least one charged particle must have transverse momentum $p_T > 1.7 (1.6) \text{GeV}/c$ in the 2011 (2012) data and be inconsistent with originating from a PV. A multivariate algorithm $^{[47]}$ is used for the identification of secondary vertices consistent with the decay of a $b$ hadron. It is required that the software trigger decision must have been caused entirely by tracks from the decay of the signal $B$ candidate.

To suppress the ‘combinatorial’ background formed by combinations of unrelated tracks, the events satisfying the trigger requirements are filtered in two stages: a pre-selection based on loose requirements, followed by a multivariate selection. In order to minimise the variation of the selection efficiency over the Dalitz plot, the selection procedure uses only loose requirements on the momenta of the $B$-meson decay products and relies mainly on topological features such as the flight distance of the $B$ candidate. These features depend on whether the $B$ candidate or the $K^0_L, h^+, h^-$ candidates are consistent with having originated from a particular PV. It is therefore necessary to ‘associate’ each candidate with a single PV — that from which it is most consistent with having originated.
The association is defined in terms of the $\chi^2_{\text{IP}}$ quantity, which is the difference in fit $\chi^2$ of the given PV reconstructed with and without the track or tracks from the particle in question. In events that contain more than one PV, each candidate is associated with the PV that has the smallest $\chi^2_{\text{IP}}$.

Decays of $K^0_S \to \pi^+\pi^-$ are reconstructed in two different categories: the first involving $K^0_S$ mesons that decay early enough for the resulting pions to be reconstructed in the VELO; and the second containing those $K^0_S$ mesons that decay later, such that track segments of the pions cannot be formed in the VELO. These $K^0_S$ reconstruction categories are referred to as *long* and *downstream*, respectively. The long category has better mass, momentum and vertex resolution than the downstream category. There are however approximately twice as many $K^0_S$ candidates reconstructed as downstream than as long, simply due to the lifetime of the $K^0_S$ meson and the geometry of the detector. In the following, $B$ candidates reconstructed from either a long or downstream $K^0_S$ candidate, in addition to two oppositely charged tracks, are also referred to using these category names. During the 2012 data taking, a significant improvement of the trigger efficiency for long-lived particles, specifically for downstream candidates, was obtained following an update of the software trigger algorithms. To take into account the differences in trigger efficiencies and the different data-taking conditions, the data sample is divided into 2011, 2012a, and 2012b data-taking periods, and each period is divided in two sub-samples according to the $K^0_S$ reconstruction category. The 2012b sample is the largest, corresponding to 1.4 fb$^{-1}$, and also has the highest trigger efficiency.

The two charged pions that form the $K^0_S$ candidates are both required to have momentum $p > 2\text{ GeV}/c$ and have $\chi^2_{\text{IP}}$ with respect to their associated PV greater than 9 (4) for long (downstream) candidates. They are then required to form a vertex with good fit quality (quantified by the fit $\chi^2$, $\chi^2_{\text{vtx}} < 12$) and to have invariant mass within 20 MeV/c$^2$ (30 MeV/c$^2$) of the nominal $K^0_S$ mass [48] for long (downstream) candidates. A requirement on the square of the ratio of the separation of the $K^0_S$ vertex from its associated PV and the corresponding uncertainty, $\chi^2_{\text{VS}} > 80$ (50) for long (downstream) candidates, ensures a significant vertex separation. Downstream $K^0_S$ candidates are required in addition to have a momentum $p > 6\text{ GeV}/c$.

The $B$ candidates are formed from a $K^0_S$ candidate and two oppositely charged tracks (initially reconstructed under the pion mass hypothesis). Each of these two tracks is required to have $p < 100\text{ GeV}/c$, a value beyond which there is little pion-kaon discrimination. The scalar sum of the transverse momenta of the $K^0_S$ and the two $h^+h'^-$ candidates must be greater than 3.0 GeV/c (4.2 GeV/c), for long (downstream) candidates, and at least two of the three decay products must have $p_T > 0.8\text{ GeV}/c$. The IP of the $B$-meson decay product with the largest $p_T$ is required to be greater than 0.05 mm relative to the PV associated to the $B$ candidate. The $B$ candidate decay products are then required to form a vertex that has $\chi^2_{\text{vtx}} < 12$ and which is separated from any PV by at least 1.7 mm. The difference in $\chi^2_{\text{vtx}}$ when adding another track must be greater than 4. The $B$ candidates must have $p_T > 1.5\text{ GeV}/c$ and invariant mass within the range $4000 < m_{K^0_S\pi^+\pi^-} < 6200\text{ MeV}/c^2$. They are further obliged to be consistent with originating from a PV, quantified by requiring, for long (downstream) candidates, both that $\chi^2_{\text{IP}} < 8$ (6) and that the cosine of
the angle $\theta_{\text{DIR}}$ between the reconstructed momentum of the $B$ candidate and the vector between the associated PV and the decay vertex be greater than 0.9999 (0.999). Finally, the decay vertex of the $K^0_S$ candidate is required to be at least 30 mm downstream, along the beam direction, from that of the $B$ candidate.

Multivariate discriminants based on a boosted decision tree (BDT) algorithm [49, 50] are used to further reduce combinatorial backgrounds. Simulated $B^0 \rightarrow K^0_S \pi^+ \pi^-$ events and data from upper mass sidebands, $5425 < m_{K^0_S\pi^+\pi^-} < 6200$ MeV/c$^2$, are used as the signal and background training samples, respectively. Contributions from muons and protons are removed from these samples using particle identification (PID) variables. Each of the six samples (resulting from the division by the three data-taking periods and the two $K^0_S$ reconstruction categories) is further subdivided into two equally-sized subsamples. Each subsample is then used to train an independent discriminant. In the subsequent analysis the BDT trained on one subsample of a given category is used to select events from the other subsample, in order to avoid bias. The input quantities for the BDTs are: the $p_T$, $\eta$, $\chi^2_{\text{IP}}$, $\chi^2_{\text{VS}}$, $\cos \theta_{\text{DIR}}$ and $\chi^2_{\text{vtx}}$ values of the $B$ candidate; the smallest change in the $B$-candidate $\chi^2_{\text{vtx}}$ value when adding another track from the event; the sum of the $\chi^2_{\text{IP}}$ values of the $h^+$ and $h^-$ candidates; the $\chi^2_{\text{IP}}$, $\chi^2_{\text{VS}}$ and $\chi^2_{\text{vtx}}$ values of the $K^0_S$ candidate; and the $p_T$ asymmetry

$$p_T^{\text{asym}} = \frac{p_T^B - p_T^{\text{cone}}}{p_T^B + p_T^{\text{cone}}},$$

(3.1)

where $p_T^{\text{cone}}$ is the transverse component of the sum of all particle momenta inside a cone around the $B$-candidate direction, of radius $R \equiv \sqrt{\Delta \eta^2 + \Delta \phi^2} = 1.5$, where $\Delta \eta$ and $\Delta \phi$ are the difference in pseudorapidity and azimuthal angle (in radians) around the beam direction, between the momentum vector of the track under consideration and that of the $B$ candidate.

The selection requirement placed on the output of the BDTs is independently optimised for each data sample. For all signal decay modes that have previously been observed, the following figure of merit is used

$$Q_1 \equiv \frac{N_{\text{sig}}}{\sqrt{N_{\text{sig}} + N_{\text{bg}}}},$$

(3.2)

where $N_{\text{sig}}$ ($N_{\text{bg}}$) represents the number of expected signal (combinatorial background) events for a given selection. The value of $N_{\text{sig}}$ is estimated based on the known branching fractions and efficiencies, while $N_{\text{bg}}$ is calculated by fitting the sideband above the signal region and extrapolating into the signal region, defined as the invariant-mass window of five times the typical resolution around the $B^0$ and the $B^0_s$ masses.

For the yet unobserved $B^0_s \rightarrow K^0_S K^+ K^-$ mode, an alternative figure of merit [51] is used

$$Q_2 \equiv \frac{\varepsilon_{\text{sig}}}{1 + \sqrt{N_{\text{bg}}}},$$

(3.3)

where the signal efficiency ($\varepsilon_{\text{sig}}$) is estimated from the signal simulation. The optimisation is performed separately for each of the six categories. As each final state contains both $B^0$ and $B^0_s$ signals, one of which is favoured and the other suppressed, this procedure results in applying two differently optimised selections on each final state.
Particle identification requirements are subsequently applied in order to reduce backgrounds from decays such as $A_{L}^{0} \rightarrow K_{S}^{0}p\pi^{-}$ and $B_{(s)}^{0} \rightarrow J/\psi(\rightarrow \mu^{+}\mu^{-})K_{S}^{0}$ where, respectively, the proton and muons are misidentified as pions or kaons. PID information is also used to assign each candidate exclusively to one out of four possible final states: $K_{S}^{0}\pi^{+}\pi^{-}$, $K_{S}^{0}K^{+}\pi^{-}$, $K_{S}^{0}\pi^{+}\bar{K}^{-}$, and $K_{S}^{0}K^{+}\bar{K}^{-}$. The PID requirements are optimised to reduce the cross-feed between the different signal decay modes using the same figures of merit introduced for the BDT optimisation.

Fully reconstructed $B$-meson decays into two-body $Dh$ or $(c\bar{s})K_{S}^{0}$ combinations, where $(c\bar{s})$ indicates a charmonium resonance, may result in a $K_{S}^{0}h^{+}h^{-}$ final state that satisfies the selection criteria and has the same $B$-candidate invariant mass distribution as the signal candidates. The decays of $A_{L}^{0}$ baryons to $\Lambda_{c}^{+}h$ with $\Lambda_{c}^{+} \rightarrow pK_{S}^{0}$ also peak under the signal when the proton is misidentified. Therefore, the following $D$ and $\Lambda_{c}^{+}$ decays are explicitly reconstructed under the relevant particle hypotheses and vetoed in all the spectra: $D^{0} \rightarrow K^{-}\pi^{+}$, $D^{+} \rightarrow K_{S}^{0}K^{+}$, $D^{+} \rightarrow K_{S}^{0}\pi^{+}$, $D_{s}^{+} \rightarrow K_{S}^{0}K^{+}$, $D_{s}^{+} \rightarrow K_{S}^{0}\pi^{+}$, and $\Lambda_{c}^{+} \rightarrow pK_{S}^{0}$. Additional vetoes on charmonium resonances, $J/\psi \rightarrow \pi^{+}\pi^{-}$, $K^{+}K^{-}$ and $\chi_{c0} \rightarrow \pi^{+}\pi^{-}$, $K^{+}K^{-}$, are applied to remove the small number of fully reconstructed and well identified peaking $B_{(s)}^{0} \rightarrow (J/\psi,\chi_{c0})K_{S}^{0}$ decays. The vetoed region for each reconstructed charm (charmonium) state is an invariant-mass window of 30 (48) MeV/$c^2$ around the world average mass value of that state [48]. This range reflects the typical mass resolution obtained at LHCb.

The fraction of selected events containing more than one $B$ candidate is at the percent level. The candidate to be retained in each event is chosen randomly, but reproducibly.

4 Fit model

The signal yields corresponding to each of the BDT optimisations are determined by means of a simultaneous unbinned extended maximum likelihood fit to the $B$-candidate invariant mass distributions of all final states in the six categories. Four types of components contribute to each invariant mass distribution: signal decays, backgrounds resulting from cross-feeds, partially reconstructed decays, and random combinations of unrelated tracks.

Signal $B_{(s)}^{0} \rightarrow K_{S}^{0}h^{+}h^{-}$ decays with correct identification of the final-state particles are modelled with the sum of two Crystal Ball (CB) functions [52] that share common values for the peak position and width but have independent power law tails on opposite sides of the peak. The $B^{0}$ and $B_{s}^{0}$ masses (peak positions of the CB functions) are free parameters in the fit and are allowed to take different values in the different data-taking periods in order to allow for small differences in momentum calibration. Seven parameters related to the widths of the CB functions are also free parameters of the fit: the width of the downstream $B^{0} \rightarrow K_{S}^{0}\pi^{+}\pi^{-}$ signal in each of the three data-taking periods; the ratio of the widths of the $B_{s}^{0}$ and $B^{0}$ decay modes; the relative widths of $K_{S}^{0}K^{+}\pi^{-}$ and $K_{S}^{0}K^{+}\bar{K}^{-}$ to $K_{S}^{0}\pi^{+}\pi^{-}$; and the ratio of the widths in the long and downstream categories. The dependence of the width on each of these divisions is assumed to factorise; for example the width $\sigma$ of the long $B_{s}^{0} \rightarrow K_{S}^{0}K^{+}\bar{K}^{-}$ signal in the 2011 data-taking period is related to
that of the downstream \( B^0 \to K_S^0 \pi^+ \pi^- \) signal in the same data-taking period by

\[
\sigma_{B^0 \to K^0_S K^+ K^-}^{2011 \text{long}} = \sigma_{B^0 \to K^0_S K^+ K^-}^{2011 \text{downstream}} \times r_{B^0/B^0} \times r_{K^0_S K^+ K^-/K_S^0 \pi^+ \pi^-} \times r_{\text{long/downstream}},
\]

where \( r_{x/y} \) indicates the ratio of the widths of categories \( x \) and \( y \). These assumptions are made necessary by the otherwise poor determination of the width of the suppressed mode in each spectrum. The other parameters of the CB components are obtained by a simultaneous fit to simulated samples.

Cross-feed contributions from misidentified signal decays are modelled empirically by the sum of two CB functions using simulated events. Only contributions from the decays \( B^0 \to K^0_S \pi^+ \pi^- \) and \( B^0 \to K^0_S K^+ K^- \) reconstructed and selected as \( K^0_S K^\pm \pi^\mp \), or the decays \( B_S^0 \to K^0_S K^\pm \pi^\mp \) and \( B^0 \to K^0_S K^\pm \pi^\mp \) reconstructed and selected as either \( K^0_S K^+ K^- \) or \( K^0_S \pi^+ \pi^- \) are considered. Other potential misidentified decays are neglected, as their contributions have been checked to be below one event. The relative yield of each misidentified decay is constrained with respect to the yield of the corresponding correctly identified decay. The constraints are implemented using Gaussian prior probability distributions included in the likelihood. The mean values are obtained from the ratio of selection efficiencies and the widths include uncertainties originating from the finite size of the simulated event samples and the systematic uncertainties related to the determination of the PID efficiencies.

Backgrounds from partially reconstructed decays such as \( B_S^0 \to K^{\ast 0} (\to K^0_S \pi^0) R^{\ast 0} (K^- \pi^+) \), where the neutral pion is not reconstructed, are also modelled. Four categories are included in each of the final state spectra, where the background results from either charmed or charmless decays of \( B^{0,+} \) or \( B^0_S \) mesons. These decays are modelled by means of generalised ARGUS functions \([53]\) convolved with a Gaussian resolution function. Their parameters are determined from simulated samples of the expected dominant decays in each category. Radiative decays and those from \( B^0 \to \eta' K^0_S \) are considered separately and included only in the \( K^0_S \pi^+ \pi^- \) final state. The normalisation of all such contributions is constrained with respect to the signal in the relevant final state using Gaussian prior probability distributions based on the ratio of efficiencies and the ratio of branching fractions from world averages \([48]\). The relative uncertainties on these ratios vary between 20% and 100%.

The combinatorial background is modelled by a linear function. The variations of the slope parameter between data-taking periods, \( K^0_S \) reconstruction categories and the different final states are assumed to factorise (in an analogous way to the widths of the signal distributions), leaving six free parameters. This assumption, as well as the choice of the linear model, are considered as sources of systematic uncertainties.

The fit results for each BDT optimisation, combining all data-taking periods, are displayed in figures 1 and 2. The separate plots for the individual data-taking periods are shown in figures 3–14 in appendix A. Table 1 shows the signal yields for each mode summed over all data-taking periods and \( K^0_S \) reconstruction categories, along with a weighted sum of efficiencies. The fitted yields of each decay mode for each of the three data-taking periods and two \( K^0_S \) reconstruction categories are given in appendix A. Statistical correlations between the signal yields are below 10% in all cases and are neglected. For the suppressed modes, the combinatorial background is negligible in the high invariant-mass region for the \( K^0_S \pi^+ \pi^- \) and \( K^0_S K^+ K^- \) final states, leading to a small systematic uncertainty related
<table>
<thead>
<tr>
<th>Decay</th>
<th>downstream</th>
<th></th>
<th>long</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yield</td>
<td>Efficiency (%)</td>
<td>Yield</td>
<td>Efficiency (%)</td>
</tr>
<tr>
<td>$B^0 \to K_{S}^{0}\pi^+\pi^-$</td>
<td>2766 ± 66</td>
<td>0.0447 ± 0.0039</td>
<td>1411 ± 45</td>
<td>0.0168 ± 0.0015</td>
</tr>
<tr>
<td>$B^0 \to K_{S}^{0}\bar{K}^{0}+\pi^+$</td>
<td>261 ± 24</td>
<td>0.0340 ± 0.0031</td>
<td>160 ± 17</td>
<td>0.0120 ± 0.0012</td>
</tr>
<tr>
<td>$B^0 \to K_{S}^{0}\bar{K}^{0}+K^-$</td>
<td>1133 ± 39</td>
<td>0.0300 ± 0.0035</td>
<td>685 ± 29</td>
<td>0.0142 ± 0.0017</td>
</tr>
<tr>
<td>$B^0_s \to K_{S}^{0}\pi^+$</td>
<td>146 ± 19</td>
<td>0.0359 ± 0.0030</td>
<td>74 ± 11</td>
<td>0.0127 ± 0.0011</td>
</tr>
<tr>
<td>$B^0_s \to K_{S}^{0}\bar{K}^{0}+\pi^+$</td>
<td>1100 ± 41</td>
<td>0.0387 ± 0.0035</td>
<td>568 ± 28</td>
<td>0.0146 ± 0.0013</td>
</tr>
<tr>
<td>$B^0_s \to K_{S}^{0}\bar{K}^{0}+K^-$</td>
<td>12 ± 6</td>
<td>0.0282 ± 0.0023</td>
<td>7 ± 4</td>
<td>0.0094 ± 0.0013</td>
</tr>
</tbody>
</table>

Table 1. Signal yields obtained from the simultaneous fit to the data. The yields are the sum of those obtained in the three data-taking periods when fitting the data sample selected using the BDT optimisation chosen for the given decay mode. The uncertainties are statistical only. The average selection efficiencies, described in section 5, are also shown for each decay mode together with the corresponding total uncertainty due to the limited simulation sample size and systematic effects in their determination.

to the assumptions used to fit this component. In order to determine the significance of the $B^0 \to K_{S}^{0}\bar{K}^{0}+K^-$ signal, likelihood profiles are constructed for the $B^0_s \to K_{S}^{0}\bar{K}^{0}+K^-$ yield in each fit category, taking into account systematic uncertainties. The profiles are constructed from fits where the shape parameters of the $B^0 \to K_{S}^{0}\bar{K}^{0}+K^-$ signal are fixed to the values obtained from the nominal fit, which allows the change in the fit likelihood to be interpreted using Wilks’ theorem [54]. Combining these profiles yields a significance of 2.5 $\sigma$.

5 Determination of the efficiencies

The measurements of the branching fractions of the $B^0_{(s)} \to K_{S}^{0}h^{+}h'^{+}$ decays relative to the well established $B^0 \to K_{S}^{0}\pi^+\pi^-$ decay mode proceed according to

\[
\frac{\mathcal{B}(B^0_{(s)} \to K_{S}^{0}h^{+}h'^{+})}{\mathcal{B}(B^0 \to K_{S}^{0}\pi^+\pi^-)} = \frac{\varepsilon_{B^0 \to K_{S}^{0}h^{+}h'^{+}}^\text{sel}}{\varepsilon_{B^0 \to K_{S}^{0}\pi^+\pi^-}} \frac{N_{B^0_{(s)} \to K_{S}^{0}h^{+}h'^{+}}}{N_{B^0 \to K_{S}^{0}\pi^+\pi^-}} \frac{f_d}{f_d,s},
\]

where $\varepsilon^\text{sel}$ is the selection efficiency (which includes geometrical acceptance, reconstruction, selection, trigger and particle identification components), $N$ is the fitted signal yield, and $f_d$ and $f_d,s$ are the hadronisation fractions of a $b$ quark into a $B^0$ and $B^0_s$ meson, respectively. The ratio $f_s/f_d$ has been precisely determined by the LHCb experiment from hadronic and semileptonic measurements to be $f_s/f_d = 0.259 ± 0.015$ [55, 56]. Since the $CP$ content of the three $B^0_s$ decays is currently unknown, the calculation of the corresponding efficiencies assumes an effective lifetime of $1/\Gamma_s$, where $\Gamma_s$ is the average width of the two $CP$-eigenstates of the $B^0_s$ meson. The effect of varying the decay width by $\pm \Delta \Gamma_s/2$, where $\Delta \Gamma_s$ is the width difference between the two $B^0_s$ $CP$-eigenstates, results in relative changes to the average efficiency of $± 4\%$. 

---
Figure 1. Invariant mass distributions of, from top to bottom, $K^0 S K^+$, $K^0 S K^0$ and $K^0 S \pi^+ \pi^-$ candidates, summing the three periods of data taking, with the selection optimisation chosen for the favoured decay modes for (left) downstream and (right) long $K^0$ reconstruction categories. In each plot, data are the black points with error bars and the total fit model is overlaid (solid blue line). The $B^0$ ($B^0_s$) signal components are the pink (orange) short-dashed (dotted) lines, while fully reconstructed misidentified decays are the green and dark blue dashed lines close to the $B^0$ and $B^0_s$ peaks. The sum of the partially reconstructed contributions from $B$ to open charm decays, charmless hadronic decays, $B^0 \rightarrow \eta' K^0$ and charmless radiative decays are the red dash triple-dotted lines. The combinatorial background contribution is the gray long-dash dotted line.

Three-body decays are, in general, composed of several quasi-two-body decays and nonresonant contributions, all of them possibly interfering. The signal reconstruction, selection and trigger efficiencies also vary over the phase space. Hence, both the distribution of the signal events and the variation of the efficiency over the Dalitz plot [57] must be
Figure 2. Invariant mass distributions of, from top to bottom, $K^0 S K^+$, $K^0 S K^+$ and $K^0 S K^+$ candidates, summing the three periods of data taking, with the selection optimisation chosen for the suppressed decay modes for (left) downstream and (right) long $K^0 S$ reconstruction categories. Colours and line styles follow the same conventions as in figure 1.

determined in order to calculate the efficiency-corrected yield. In this analysis, efficiencies are determined for each decay mode from simulated signal samples in bins of the “square Dalitz plot” [58], where the usual Dalitz-plot coordinates have been transformed in order to map a rectangular space. The edges of the phase space are spread out in the square Dalitz plot, which permits a more precise modelling of the efficiency in the regions where it varies the most and where most of the signal candidates are expected. The square Dalitz-plot distribution of each signal mode is determined from the data using the $s$Plot technique [59], using the $K^0 S K^+$ invariant mass as the discriminating variable. The distributions of signal
events on the Dalitz plot, as obtained using this technique, are shown in appendix C. The efficiency-corrected yields are calculated as:

$$N_{corr}^{B_0 \to K_S^0 h^\pm h'^\mp} = \sum_{i} w_i \varepsilon_i,$$

(5.2)

where $w_i$ is the sPlot weight and $\varepsilon_i$ is the efficiency for event $i$, and the sum includes all events in the fitted data sample. The average efficiency for each decay mode:

$$\bar{\varepsilon} = \frac{N_{B_0 \to K_S^0 h^\pm h'^\mp}}{N_{corr}^{B_0 \to K_S^0 h^\pm h'^\mp}},$$

(5.3)

where $N_{B_0 \to K_S^0 h^\pm h'^\mp}$ is the fitted signal yield, is given for each signal decay in table 1. They are presented for each data-taking period and $K_S^0$ reconstruction category in tables 3–5 in appendix A. Their relative uncertainties due to the finite size of the simulated event samples vary from 1% to 20%.

The imperfections of the simulation are corrected for in several respects. Inaccuracies of the tracking efficiency in the simulation are mitigated by weighting the $h^+$ and $h'^-$ tracks by a correction factor obtained from a data calibration sample [60]. An analogous correction is applied for the $K_S^0$ tracking and vertex reconstruction efficiency. A control data sample of $D^{*+} \to D^0 (\to K^- \pi^+) \pi^+_s$ decays, where $\pi^+_s$ indicates a slow pion, is used to quantify the differences between the data and simulation hardware trigger stage for pions and kaons, independently for positive and negative hadron charges, as a function of $p_T$ [46]. Corrections to account for differences between data and simulation related to tracking efficiency are $O(1\%)$, while those related to trigger efficiency can be $O(10\%)$, depending on the position on the Dalitz plot. The uncertainties attached to these various corrections are propagated to the branching fraction measurements as systematic uncertainties and are further discussed in section 6.

The particle identification efficiencies and misidentification rates are determined from simulated signal decays on an event-by-event basis using a data-driven calibration to match the kinematic properties of the tracks in the decay of interest. A weighting procedure is performed in bins of $p$, $\eta$ and event multiplicity, accounting for kinematic correlations between the tracks. Calibration tracks are taken from $D^{*+} \to D^0 \pi^+_s$ decays where the $D^0$ decays to the Cabibbo-favoured $K^- \pi^+$ final state. In this case, the charge of the soft pion $\pi^+_s$ provides the kaon or pion identity of the tracks. The dependence of the PID efficiency over the Dalitz plot induced by the variations of PID performance with the track kinematics is included in the procedure described above. This calibration is performed using samples from the same data-taking period, accounting for the variation in the performance of the Cherenkov detectors over time.

6 Systematic uncertainties

Most of the systematic uncertainties are eliminated or greatly reduced by normalising the branching fraction measurements to the $B^0 \to K_S^0 \pi^+ \pi^-$ mode. A summary of the contributions, expressed as relative uncertainties, is given in table 2, including the uncertainty
Table 2. Summary of the systematic uncertainties on the ratio of branching fractions. A weighted average of the two $K^0_S$ reconstruction categories and three data-taking periods is performed. The values quoted for the individual contributions, which are illustrative of the hierarchy between sources of systematic uncertainty, each result from a weighted average in which the other systematic uncertainty contributions are disregarded. The total uncertainty is the weighted average including all contributions. All uncertainties are relative and are quoted as percentages.

in the measurement of $f_s/f_d$ [55, 56]. A detailed breakdown of systematic uncertainties per data-taking period and $K^0_S$ reconstruction category is provided in tables 6 and 7 in appendix B. The dominant contributions arise from the modelling of the combinatorial background shape in the invariant mass fit and from the determination of the efficiency of the hardware trigger.

6.1 Fit model

The fit model relies on a number of assumptions, both in terms of the values of parameters being taken from simulation and in terms of the choice of the functional forms describing the various contributions. In both cases, the uncertainties are evaluated using pseudoexperiments that are generated from the alternative parameterisation and are fitted using both the nominal and the alternative fit models. The distribution of the difference in the value of a given parameter determined in the two fit model is subsequently fitted with a Gaussian function and the corresponding systematic uncertainty is assigned as the sum in quadrature of the mean and the resolution of the Gaussian.

This procedure is employed for the fixed parameters of the signal, partially-reconstructed and cross-feed backgrounds and for the functional forms used for the signal and combinatorial background. Due to the limited sizes of the simulated event samples used to parameterise both the partially-reconstructed and cross-feed background shapes, the uncertainty related to the fixed parameters also covers any reasonable variation of the shape. For the combinatorial background, the ratios of the slopes in different $K^0_S$ reconstruction categories and in different data-taking periods are constrained to be the same for all final states. Two alternative models are considered: allowing independent ratios for each of the final states (testing the assumption of the factorisation of the slope ratios) and using an exponential model instead of the nominal linear one (testing the functional form of the combinatorial shape).
Finally, in order to evaluate the impact of residual contributions from $A_0^0$ decays that survive the proton PID veto described in section 3, fits to data are performed including a model for this contribution. As these fits show negligible difference to the nominal model, no systematic uncertainty is assigned.

The total fit model systematic uncertainty is given by the sum in quadrature of all the contributions and is mostly dominated by the combinatorial background model uncertainty. Some uncertainties are fully correlated among the different data samples that are averaged and are treated as such in the uncertainty propagation. The correlated fit model systematics include uncertainties due to the fit biases and combinatorial and signal shapes. The combination of these contributions is shown in table 2 as “Fit model”, while they are referred to in tables 6 and 7 as “Fit model (corr.)” and “Fit model (uncorr.)”, respectively.

6.2 Selection and trigger efficiencies

The accuracy of the efficiency determination is limited in most cases by the finite size of the samples of simulated signal events, duly propagated as a systematic uncertainty. In addition, the effect related to the choice of binning for the square Dalitz plot is estimated from the spread of the average efficiencies determined from several alternative binning schemes. These two sources of uncertainties are detailed in tables 6 and 7, and are labelled “Selection (statistics)” and “Selection (binning)”, respectively.

As introduced in section 5, the sources of uncertainties related to the imperfections of the tracking simulation are two-fold: the reconstruction of both long and downstream tracks and the reconstruction of the $K_S^{0}$ decay vertex (in particular for the downstream category). In both cases, the reconstruction efficiencies are determined from data calibration samples and the simulated events are weighted to match the performance measured in the data. The uncertainties arising from the finite size of the calibration samples are propagated to assign a systematic uncertainty.

Possible sources of systematic uncertainty related to the efficiency estimation of the hardware trigger have been studied. Two additional systematic uncertainties are assigned: one related to the imperfect simulation of the rate of overlapping tracks in the hadron calorimeter forming a single hadron trigger candidate and one related to the choice of the data calibration sample itself. These two sources of uncertainties are detailed in tables 6 and 7 in appendix B, labelled “Trigger (overlap)” and “Trigger (calib. sample)”, respectively. For the first source, the systematic uncertainty is estimated as the difference between the trigger efficiency correction with and without the overlapping cluster corrections. For the latter source of uncertainty, the correction factors have been determined from a sample of reconstructed $B^0 \rightarrow J/\psi K^{+} \pi^-$ events. Twice the difference between the correction factors determined from the two calibration samples is taken as the estimate of the associated systematic uncertainty.

The uncertainties due to the choice of the binning of the square Dalitz plot used to produce the efficiency maps and those related to the hardware trigger calibration samples are treated as correlated among the different data samples split by year of data taking.
6.3 Particle identification efficiencies

The procedure to evaluate the efficiencies of the PID selections uses calibration tracks that differ from the signal tracks in terms of their kinematic distributions. While the binning procedure attempts to mitigate these differences, there could be some remaining systematic effect. This is addressed by considering different ensembles of kinematical binnings to determine the efficiency. An overall 1% systematic uncertainty is assigned to quantify any bias due to the procedure. The statistical uncertainties originating from the finite sample sizes are added in quadrature.

7 Results and conclusion

The decay modes $B^0_s \to K^0_S h^\pm h^\mp$ have been analysed using a dataset, corresponding to an integrated luminosity of 3.0 fb$^{-1}$ recorded by the LHCb detector at a centre-of-mass energy of 7 TeV and 8 TeV. The branching fraction of each decay is measured relative to that of $B^0 \to K^0_S \pi^+ \pi^-$. The ratios of branching fractions are determined independently for the two $K^0_S$ reconstruction categories and three data-taking periods and then combined by performing a $\chi^2$ fit. The corresponding covariance matrix includes the statistical and systematic uncertainties. A 100% linear correlation factor is assumed for the correlated systematic uncertainties. Good agreement is obtained among all determinations from each data-taking period and $K^0_S$ reconstruction category. The results obtained from the combination are

$$
\frac{B(B^0 \to K^0_S K^\pm \pi^\mp)}{B(B^0 \to K^0_S \pi^+ \pi^-)} = 0.123 \pm 0.009 \text{ (stat)} \pm 0.015 \text{ (syst)}, \\
\frac{B(B^0 \to K^0_S K^+ K^-)}{B(B^0 \to K^0_S \pi^+ \pi^-)} = 0.549 \pm 0.018 \text{ (stat)} \pm 0.033 \text{ (syst)}, \\
\frac{B(B^0_s \to K^0_S \pi^+ \pi^-)}{B(B^0 \to K^0_S \pi^+ \pi^-)} = 0.191 \pm 0.027 \text{ (stat)} \pm 0.031 \text{ (syst)} \pm 0.011 \text{ (fs/fd)}, \\
\frac{B(B^0 \to K^0_S K^\pm \pi^\mp)}{B(B^0 \to K^0_S \pi^+ \pi^-)} = 1.70 \pm 0.07 \text{ (stat)} \pm 0.11 \text{ (syst)} \pm 0.10 \text{ (fs/fd)}, \\
\frac{B(B^0_s \to K^0_S K^+ K^-)}{B(B^0 \to K^0_S \pi^+ \pi^-)} = 0.026 \pm 0.011 \text{ (stat)} \pm 0.007 \text{ (syst)} \pm 0.002 \text{ (fs/fd)}.
$$

All measurements of branching fractions are in good agreement with the earlier LHCb determinations [34], which they supersede. The measurement of the relative branching fractions of $B^0 \to K^0_S K^\pm \pi^\mp$ and $B^0 \to K^0_S K^+ K^-$ are consistent with the world average results [13, 48].

The significance of the measured signal yield for the decay $B^0_s \to K^0_S K^+ K^-$, including systematic uncertainties, is 2.5 standard deviations. A 90% confidence level (C.L.) interval for the corresponding branching fraction relative to that of $B^0 \to K^0_S \pi^+ \pi^-$ is derived, following the approach of Feldman-Cousins [61]

$$
\frac{B(B^0 \to K^0_S K^+ K^-)}{B(B^0 \to K^0_S \pi^+ \pi^-)} \in [0.008 - 0.051] \text{ at 90\% C.L.}
$$
Using the world average value omitting the previous LHCb result, $B(B^0 \to K^0\pi^+\pi^-) = (4.96 \pm 0.20) \times 10^{-5}$ \cite{13, 48}, the measured time-integrated branching fractions are

\begin{align*}
B(B^0 \to \bar{K}^0 K^\mp \pi^\mp) &= (6.1 \pm 0.5 \pm 0.7 \pm 0.3) \times 10^{-6}, \\
B(B^0 \to K^0 K^+ K^-) &= (27.2 \pm 0.9 \pm 1.6 \pm 1.1) \times 10^{-6}, \\
B(B^0_s \to K^0\pi^+\pi^-) &= (9.5 \pm 1.3 \pm 1.5 \pm 0.4) \times 10^{-6}, \\
B(B^0_s \to \bar{K}^0 K^\mp \pi^\pm) &= (84.3 \pm 3.5 \pm 7.4 \pm 3.4) \times 10^{-6}, \\
B(B^0_s \to K^0 K^+ K^-) &\in [0.4 - 2.5] \times 10^{-6} \text{ at 90\% C.L.,}
\end{align*}

where the first uncertainty is statistical, the second systematic and the last due to the uncertainty on $B(B^0 \to K^0\pi^+\pi^-)$. These results are in agreement with the available predictions for these channels \cite{24-27}.

The first Dalitz-plot analyses by the LHCb experiment of the dominant decays ($B^0 \to K^0\pi^+\pi^-$, $B^0_s \to K^0\pi^+\pi^-$, and $B^0 \to K^0 K^+ K^-$) are the next step of the physics programme introduced in this work. These studies will follow and benefit from the selection methods developed for this analysis.

Acknowledgments

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FASO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (U.S.A.). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and OSC (U.S.A.). We are indebted to the communities behind the multiple open source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany), EPLANET, Marie Sklodowska-Curie Actions and ERC (European Union), Conseil Général de Haute-Savoie, Labex ENIGMASS and OCEUV, Région Auvergne (France), RFBR and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), Herchel Smith Fund, The Royal Society, Royal Commission for the Exhibition of 1851 and the Leverhulme Trust (United Kingdom).
A Fit results by category

Signal yields and efficiencies for the different decays, data-taking periods and $K_S^0$ reconstruction categories are shown for each of the two BDT optimisation points in tables 3–5. Fit results for the different data-taking periods and $K_S^0$ reconstruction categories are shown for each of the two BDT optimisation points in figures 3–14.

<table>
<thead>
<tr>
<th>Decay</th>
<th>downstream</th>
<th>long</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0 \rightarrow K^0_S\pi^+\pi^-$</td>
<td>Yield</td>
<td>Efficiency (%)</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0_SK^+K^-$</td>
<td>333 ± 23</td>
<td>0.0361 ± 0.0064</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0_SK^+\pi^+$</td>
<td>76 ± 13</td>
<td>0.0322 ± 0.0063</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0_S\pi^+$</td>
<td>43 ± 10</td>
<td>0.0316 ± 0.0051</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0_SK^+K^-$</td>
<td>5 ± 3</td>
<td>0.0244 ± 0.0052</td>
</tr>
</tbody>
</table>

Table 3. Signal yields obtained for the 2011 category from the simultaneous fit to the data. The yields shown are those obtained when fitting the data sample selected using the BDT optimisation chosen for the given decay mode. The uncertainties are statistical only. The average selection efficiencies, described in section 5, are also shown for each decay mode together with the corresponding total uncertainty due to the limited simulation sample size and systematic effects in their determination.

<table>
<thead>
<tr>
<th>Decay</th>
<th>downstream</th>
<th>long</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0 \rightarrow K^0_S\pi^+\pi^-$</td>
<td>Yield</td>
<td>Efficiency (%)</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0_SK^+K^-$</td>
<td>553 ± 30</td>
<td>0.0423 ± 0.0059</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0_SK^+\pi^+$</td>
<td>181 ± 15</td>
<td>0.0263 ± 0.0052</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0_SK^+\pi^+$</td>
<td>205 ± 18</td>
<td>0.0395 ± 0.0060</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0.SK^+\pi^+$</td>
<td>63 ± 11</td>
<td>0.0306 ± 0.0047</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0_S\pi^+$</td>
<td>17 ± 8</td>
<td>0.0493 ± 0.0068</td>
</tr>
<tr>
<td>$B^0 \rightarrow K^0_SK^+K^-$</td>
<td>2 ± 3</td>
<td>0.0290 ± 0.0039</td>
</tr>
</tbody>
</table>

Table 4. Signal yields obtained for the 2011 category from the simultaneous fit to the data. The yields shown are those obtained when fitting the data sample selected using the BDT optimisation chosen for the given decay mode. The uncertainties are statistical only. The average selection efficiencies, described in section 5, are also shown for each decay mode together with the corresponding total uncertainty due to the limited simulation sample size and systematic effects in their determination.
<table>
<thead>
<tr>
<th>Decay</th>
<th>downstream Yield</th>
<th>Efficiency (%)</th>
<th>long Yield</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0 \to K^0 S_{\pi^+\pi^-}$</td>
<td>1410 ± 46</td>
<td>0.0455 ± 0.0063</td>
<td>654 ± 30</td>
<td>0.0161 ± 0.0022</td>
</tr>
<tr>
<td>$B^0 \to K^0 S_{K^+K^-}$</td>
<td>671 ± 30</td>
<td>0.0395 ± 0.0076</td>
<td>344 ± 20</td>
<td>0.0128 ± 0.0025</td>
</tr>
<tr>
<td>$B^0_s \to K^0 S_{K^\pm\pi^\mp}$</td>
<td>562 ± 29</td>
<td>0.0401 ± 0.0059</td>
<td>262 ± 19</td>
<td>0.0138 ± 0.0020</td>
</tr>
<tr>
<td>$B^0 \to K^0 S_{K^\pm\pi^\mp}$</td>
<td>122 ± 17</td>
<td>0.0402 ± 0.0056</td>
<td>65 ± 10</td>
<td>0.0100 ± 0.0015</td>
</tr>
<tr>
<td>$B^0_s \to K^0 S_{\pi^+\pi^-}$</td>
<td>86 ± 14</td>
<td>0.0335 ± 0.0045</td>
<td>38 ± 5</td>
<td>0.0108 ± 0.0015</td>
</tr>
<tr>
<td>$B^0_s \to K^0 S_{K^+K^-}$</td>
<td>5 ± 4</td>
<td>0.0291 ± 0.0034</td>
<td>2 ± 2</td>
<td>0.0083 ± 0.0017</td>
</tr>
</tbody>
</table>

**Table 5.** Signal yields obtained for the 2011 category from the simultaneous fit to the data. The yields shown are those obtained when fitting the data sample selected using the BDT optimisation chosen for the given decay mode. The uncertainties are statistical only. The average selection efficiencies, described in section 5, are also shown for each decay mode together with the corresponding total uncertainty due to the limited simulation sample size and systematic effects in their determination.
Figure 3. Results of the simultaneous fit to data (downstream, 2011) with the BDT optimisation corresponding to the favoured decay modes. The modes $K^0_S K^+ K^-$, $K^0_S K^+ \pi^-$, $K^0_S \pi^+ K^-$ and $K^0_S \pi^+ \pi^-$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 4. Results of the simultaneous fit to data (downstream, 2012a) with the BDT optimisation corresponding to the favoured decay modes. The modes $K_0^0 K^+ K^−$, $K_0^0 K^− π^+$, $K_0^0 π^+ K^−$ and $K_0^0 π^+ π^−$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 5. Results of the simultaneous fit to data (downstream, 2012b) with the BDT optimisation corresponding to the favoured decay modes. The modes $K_{S}^{0}K^{+}K^{-}$, $K_{S}^{0}K^{+}\pi^{-}$, $K_{S}^{0}\pi^{+}K^{-}$ and $K_{S}^{0}\pi^{+}\pi^{-}$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 6. Results of the simultaneous fit to data (long, 2011) with the BDT optimisation corresponding to the favoured decay modes. The modes $K^0_SK^0S$, $K^0_SK^0S$, $K^0_SK^0S$ and $K^0_SK^0S$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 7. Results of the simultaneous fit to data (long, 2012a) with the BDT optimisation corresponding to the favoured decay modes. The modes $K_0^0 K^+ K^-$, $K_0^0 K^+ \pi^-$, $K_0^0 \pi^+ K^-$ and $K_0^0 \pi^+ \pi^-$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 8. Results of the simultaneous fit to data (long, 2012b) with the BDT optimisation corresponding to the favoured decay modes. The modes \(K^0_S K^+ K^-\), \(K^0_S K^+ \pi^-\), \(K^0_S \pi^+ K^-\) and \(K^0_S \pi^+ \pi^-\) are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 9. Results of the simultaneous fit to data (downstream, 2011) with the BDT optimisation corresponding to the suppressed modes. The modes $K^0_{S}K^-\ K^+$, $K^0_S K^+\pi^-$, $K^0_S\pi^+K^-$ and $K^0_S\pi^+\pi^-$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 10. Results of the simultaneous fit to data (downstream, 2012a) with the BDT optimisation corresponding to the suppressed modes. The modes $K_S^0 K^+ K^−$, $K_S^0 K^+ π^−$, $K_S^0 π^+ K^−$ and $K_S^0 π^+ π^−$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 11. Results of the simultaneous fit to data (downstream, 2012b) with the BDT optimisation corresponding to the suppressed modes. The modes $K^0S K^+$, $K^0 S K^+$, $K^0 S K^-,$ and $K^0 S K^+$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 12. Results of the simultaneous fit to data (long, 2011) with the BDT optimisation corresponding to the suppressed modes. The panels $K^0_S K^0$, $K^0_S K^-$, $K^0_S K^+$ and $K^0_S \pi^-$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 13. Results of the simultaneous fit to data (long, 2012a) with the BDT optimisation corresponding to the suppressed modes. The modes $K^0 K^- K^+$, $K^0 K^0 K^+$, $K^0 K^+ K^-$ and $K^0 K^+ \pi^-$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
Figure 14. Results of the simultaneous fit to data (long, 2012b) with the BDT optimisation corresponding to the suppressed modes. The modes $K_0^0 K^+ K^-$, $K_0^0 K^+ \pi^-$, $K_0^0 \pi^+ K^-$ and $K_0^0 \pi^+ \pi^-$ are shown from top to bottom. The left-hand side plots show the results with a logarithmic scale and the right-hand side with a linear scale. Legend is similar to that of plots shown in figure 1.
B Breakdown of systematic uncertainties

The full breakdown of the systematic uncertainties for each data-taking period is given in tables 6 and 7 for, respectively, the downstream and long $K_S^0$ reconstruction categories.

<table>
<thead>
<tr>
<th>Relative $B$ (2011 sample)</th>
<th>$B(B^0\to K_S^0K^+\pi^-)$</th>
<th>$B(B^0\to K_S^0K^+\pi^0)$</th>
<th>$B(B^0\to K_S^0K^+\pi^0)$</th>
<th>$B(B^0\to K_S^0K^+\pi^0)$</th>
<th>$B(B^0\to K_S^0K^+\pi^-)$</th>
<th>$B(B^0\to K_S^0K^+\pi^-)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit model (uncorrelated) [%]</td>
<td>12.3</td>
<td>3.2</td>
<td>7.4</td>
<td>4.4</td>
<td>12.2</td>
<td>12.2</td>
</tr>
<tr>
<td>Fit model (correlated) [%]</td>
<td>11.5</td>
<td>2.4</td>
<td>11.6</td>
<td>5.8</td>
<td>4.8</td>
<td>4.8</td>
</tr>
<tr>
<td>Selection (statistics) [%]</td>
<td>4.1</td>
<td>1.1</td>
<td>3.7</td>
<td>4.0</td>
<td>6.4</td>
<td>6.4</td>
</tr>
<tr>
<td>Selection (binning) [%]</td>
<td>2.5</td>
<td>2.6</td>
<td>4.3</td>
<td>2.6</td>
<td>2.9</td>
<td>2.9</td>
</tr>
<tr>
<td>Tracking [%]</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Trigger (overlap) [%]</td>
<td>0.1</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Trigger (calibration sample) [%]</td>
<td>3.7</td>
<td>3.7</td>
<td>6.1</td>
<td>4.2</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>PID [%]</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>$f_s/f_d$ [%]</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relative $B$ (2012a sample)</th>
<th>$B(B^0\to K_S^0K^+\pi^-)$</th>
<th>$B(B^0\to K_S^0K^+\pi^0)$</th>
<th>$B(B^0\to K_S^0K^+\pi^0)$</th>
<th>$B(B^0\to K_S^0K^+\pi^0)$</th>
<th>$B(B^0\to K_S^0K^+\pi^-)$</th>
<th>$B(B^0\to K_S^0K^+\pi^-)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit model (uncorrelated) [%]</td>
<td>4.2</td>
<td>0.8</td>
<td>4.5</td>
<td>4.6</td>
<td>58.1</td>
<td>58.1</td>
</tr>
<tr>
<td>Fit model (correlated) [%]</td>
<td>7.8</td>
<td>1.6</td>
<td>15.8</td>
<td>4.8</td>
<td>16.7</td>
<td>16.7</td>
</tr>
<tr>
<td>Selection (statistics) [%]</td>
<td>4.1</td>
<td>1.8</td>
<td>2.3</td>
<td>1.3</td>
<td>6.1</td>
<td>6.1</td>
</tr>
<tr>
<td>Selection (binning) [%]</td>
<td>4.1</td>
<td>3.7</td>
<td>2.7</td>
<td>3.4</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Tracking [%]</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Trigger (overlap) [%]</td>
<td>0.0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Trigger (calibration sample) [%]</td>
<td>2.8</td>
<td>5.9</td>
<td>6.4</td>
<td>2.6</td>
<td>12.8</td>
<td>12.8</td>
</tr>
<tr>
<td>PID [%]</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>$f_s/f_d$ [%]</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relative $B$ (2012b sample)</th>
<th>$B(B^0\to K_S^0K^+\pi^-)$</th>
<th>$B(B^0\to K_S^0K^+\pi^0)$</th>
<th>$B(B^0\to K_S^0K^+\pi^0)$</th>
<th>$B(B^0\to K_S^0K^+\pi^0)$</th>
<th>$B(B^0\to K_S^0K^+\pi^-)$</th>
<th>$B(B^0\to K_S^0K^+\pi^-)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit model (uncorrelated) [%]</td>
<td>9.8</td>
<td>1.8</td>
<td>10.1</td>
<td>2.1</td>
<td>26.4</td>
<td>26.4</td>
</tr>
<tr>
<td>Fit model (correlated) [%]</td>
<td>10.9</td>
<td>1.9</td>
<td>11.2</td>
<td>4.6</td>
<td>28.0</td>
<td>28.0</td>
</tr>
<tr>
<td>Selection (statistics) [%]</td>
<td>1.2</td>
<td>0.5</td>
<td>1.1</td>
<td>0.7</td>
<td>3.4</td>
<td>3.4</td>
</tr>
<tr>
<td>Selection (binning) [%]</td>
<td>1.6</td>
<td>1.3</td>
<td>1.3</td>
<td>3.3</td>
<td>3.8</td>
<td>3.8</td>
</tr>
<tr>
<td>Tracking [%]</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Trigger (overlap) [%]</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Trigger (calibration sample) [%]</td>
<td>2.8</td>
<td>5.9</td>
<td>6.4</td>
<td>2.6</td>
<td>12.8</td>
<td>12.8</td>
</tr>
<tr>
<td>PID [%]</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>1.1</td>
</tr>
<tr>
<td>$f_s/f_d$ [%]</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Table 6. Systematic uncertainties on the ratios of branching fractions for downstream $K_S^0$ reconstruction. All uncertainties are relative and are quoted as percentages.
C Dalitz-plot distributions of signal events

Dalitz-plot distributions of signal events as extracted using the sPlot technique are shown for each signal mode are shown in figure 15.
Figure 15. Distribution of \( s \)Plot weights in data. The \( K_0^0K^+K^- \), \( K_0^0K^+\pi^- \), \( K_0^0\pi^+K^- \), and \( K_0^0\pi^+\pi^- \) final states are shown from top to bottom, with a \( B^0 \) parent on the left, and a \( B^0_s \) on the right. All data-taking periods and \( K_0^0 \) reconstruction categories are added.
Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

[3] M. Beneke, Corrections to \( \sin(2\beta) \) from CP asymmetries in \( B^0 \to (\pi^0, \rho^0, \eta, \eta', \omega, \phi)K^0 \) decays, Phys. Lett. B 620 (2005) 143 [hep-ph/0505076] [inSPIRE].
[9] Belle collaboration, J. Dalseno et al., Time-dependent Dalitz plot measurement of CP parameters in \( B^0 \to K^{0}_s\pi^+\pi^- \) decays, Phys. Rev. D 79 (2009) 072004 [arxiv:0811.3665] [inSPIRE].
[10] BABAR collaboration, B. Aubert et al., Time-dependent amplitude analysis of \( B^0 \to K^{0}_s\pi^+\pi^- \), Phys. Rev. D 80 (2009) 112001 [arxiv:0905.3615] [inSPIRE].
[12] BABAR collaboration, J.P. Lees et al., Study of CP-violation in Dalitz-plot analyses of \( B^0 \to K^{0}_sK^+K^- \), \( B^+ \to K^+K^-K^+ \), and \( B^+ \to K^{0}_sK^0K^+ \), Phys. Rev. D 85 (2012) 112010 [arxiv:1201.5897] [inSPIRE].


[22] BaBar collaboration, J.P. Lees et al., Evidence for CP violation in $B^+ \to K^*(892)^+\pi^0$ from a Dalitz plot analysis of $B^+ \to K^0_S\pi^+\pi^0$ decays, *Phys. Rev. D* **96** (2017) 072001 [arXiv:1505.00705] [SPIRE].


[27] Y. Li, Comprehensive study of $B^0 \to K^0(\bar{K}^0)K^+\pi^\pm$ decays in the factorization approach, *Phys. Rev. D* **89** (2014) 094007 [arXiv:1402.6052] [SPIRE].


[29] Y. Li, A.-J. Ma, W.-F. Wang and Z.-J. Xiao, Quasi-two-body decays $B^0(\bar{K}^0) \to P\rho'(1450), P\rho'(1700) \to P\pi\pi$ in the perturbative QCD approach, *Phys. Rev. D* **96** (2017) 036014 [arXiv:1704.07566] [SPIRE].


[33] BaBar collaboration, P. del Amo Sanchez et al., Observation of the rare decay $B^0 \to K^0_SK^\pm\pi^\mp$, *Phys. Rev. D* **82** (2010) 031101 [arXiv:1003.0640] [SPIRE].
[34] LHCb collaboration, Study of $B^0_{(s)} \rightarrow K^0_S h^+ h^-$ decays with first observation of $B^0 \rightarrow K^0_S K^+ \pi^-$ and $B^0 \rightarrow K^0_S \pi^+ \pi^-$, *JHEP* **10** (2013) 143 [arXiv:1307.7648] [inSPIRE].


[37] LHCb collaboration, The LHCb Detector at the LHC, 2008 *JINST* **3** S08005 [inSPIRE].


[58] BABAR collaboration, B. Aubert et al., *An amplitude analysis of the decay $B^\pm \to \pi^\pm \pi^\pm \pi^\mp$*, Phys. Rev. D 72 (2005) 052002 [hep-ex/0507025] [inSPIRE].


University of Chinese Academy of Sciences, Beijing, China, associated to
School of Physics and Technology, Wuhan University, Wuhan, China, associated to
Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China, associated to
Departamento de Fisica, Universidad Nacional de Colombia, Bogota, Colombia, associated to
Institut für Physik, Universität Rostock, Rostock, Germany, associated to
National Research Centre Kurchatov Institute, Moscow, Russia, associated to
Instituto de Fisica Corpuscular, Centro Mixto Universidad de Valencia - CSIC, Valencia, Spain, associated to
Van Swinderen Institute, University of Groningen, Groningen, The Netherlands, associated to

Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil
Laboratoire Leprince-Ringuet, Palaiseau, France
P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia
Università di Bari, Bari, Italy
Università di Bologna, Bologna, Italy
Università di Cagliari, Cagliari, Italy
Università di Ferrara, Ferrara, Italy
Università di Genova, Genova, Italy
Università di Milano Bicocca, Milano, Italy
Università di Roma Tor Vergata, Roma, Italy
Università di Roma La Sapienza, Roma, Italy
AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland
LIFAEELS, La Salle, Universitat Ramon Llull, Barcelona, Spain
Hanoi University of Science, Hanoi, Viet Nam
Università di Padova, Padova, Italy
Università di Pisa, Pisa, Italy
Università degli Studi di Milano, Milano, Italy
Università di Urbino, Urbino, Italy
Università della Basilicata, Potenza, Italy
Scuola Normale Superiore, Pisa, Italy
Università di Modena e Reggio Emilia, Modena, Italy
Iligan Institute of Technology (IIT), Iligan, Philippines
Novosibirsk State University, Novosibirsk, Russia

Deceased