Observation of $B[0 \over S] \rightarrow K^*[\text{superscript} \pm] K[\text{superscript} \#]$ and evidence for $B[0 \over S] \rightarrow K^*[\text{superscript} -] \pi [\text{superscript} +]$ decays

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

As Published
http://dx.doi.org/10.1088/1367-2630/16/12/123001

Publisher
IOP Publishing

Version
Final published version

Citable link
http://hdl.handle.net/1721.1/93110

Terms of Use
Creative Commons Attribution

Detailed Terms
http://creativecommons.org/licenses/by/4.0/
Observation of $B_s^0 \rightarrow K^{*\pm}K^{\mp}$ and evidence for $B_s^0 \rightarrow K^{*+}\pi^-$ decays

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 New J. Phys. 16 123001

(http://iopscience.iop.org/1367-2630/16/12/123001)

View the table of contents for this issue, or go to the journal homepage for more.

Download details:

IP Address: 18.51.1.3
This content was downloaded on 13/01/2015 at 15:41

Please note that terms and conditions apply.
Observation of $B_s^0 \to K^{*\pm} K^\mp$ and evidence for $B_s^0 \to K^{*-\pi^+}$ decays

The LHCb Collaboration

Received 31 July 2014, revised 24 September 2014
Accepted for publication 10 October 2014
Published 2 December 2014

New Journal of Physics 16 (2014) 123001
doi:10.1088/1367-2630/16/12/123001

Abstract
Measurements of the branching fractions of $B_s^0 \to K^{*\pm} K^\mp$ and $B_s^0 \to K^{*-\pi^+}$ decays are performed using a data sample corresponding to 1.0 fb$^{-1}$ of proton-proton collision data collected with the LHCb detector at a centre-of-mass energy of 7 TeV, where the $K^{*\pm}$ mesons are reconstructed in the $K_S^0 \pi^\pm$ final state. The first observation of the $B_s^0 \to K^{*\pm} K^\mp$ decay and the first evidence for the $B_s^0 \to K^{*-\pi^+}$ decay are reported with branching fractions

$$B(B_s^0 \to K^{*\pm} K^\mp) = (12.7 \pm 1.9 \pm 1.9) \times 10^{-6},$$

$$B(B_s^0 \to K^{*-\pi^+}) = (3.3 \pm 1.1 \pm 0.5) \times 10^{-6},$$

where the first uncertainties are statistical and the second are systematic. In addition, an upper limit of $B(B^0 \to K^{*\pm} K^\mp) < 0.4 (0.5) \times 10^{-6}$ is set at 90% (95%) confidence level.

Keywords: flavour physics, B physics, branching fraction

1. Introduction

The Standard Model (SM) of particle physics predicts that all manifestations of $CP$ violation, i.e. violation of symmetry under the combined charge conjugation and parity operation, arise due to the single complex phase that appears in the Cabibbo–Kobayashi–Maskawa (CKM) quark mixing matrix [1, 2]. Since this source is not sufficient to account for the level of the

1 Authors are listed at the end of the paper.
baryon asymmetry of the Universe [3], one of the key goals of contemporary particle physics is
to search for signatures of CP violation that are not consistent with the CKM paradigm.

Among the most important areas being explored in quark flavour physics is the study of B meson decays to hadronic final states that do not contain charm quarks or antiquarks (hereafter referred to as ‘charmless’). As shown in figure 1, such decays have, in general, amplitudes that contain contributions from both ‘tree’ and ‘loop’ diagrams (see, e.g., [4]). The phase differences between the two amplitudes can lead to CP violation and, since particles hypothesized in extensions to the SM may affect the loop diagrams, deviations from the SM predictions may occur. Large CP violation effects, i.e. asymmetries of $\mathcal{O}(10\%)$ or more between the rates of $\bar{B}$ and B meson decays to CP conjugate final states, have been seen in $B^0 \rightarrow K^+\pi^-$ [5–8], $B^0_s \rightarrow K^-\pi^+$ [7, 8], and $B^+ \rightarrow \pi^+\pi^-K^+$, $K^+K^-K^+$, $\pi^+\pi^-\pi^+$ and $K^+\pi^-K^+$ decays [9–11]. However, it is hard to be certain whether these measurements are consistent with the SM predictions due to the presence of parameters describing the hadronic interactions that are difficult to determine either theoretically or from data.

An interesting approach to control the hadronic uncertainties is to exploit amplitude analysis techniques. For example, by studying the distribution of kinematic configurations of $B^0 \rightarrow K^0_S\pi^+\pi^-$ decays across the Dalitz plot [12], the relative phase between the $K^{*+}\pi^-$ and $K^0_S\rho^0$ amplitudes can be determined. This information is not accessible in studies either of two-body decays, or of the inclusive properties of three-body decays. Consequently, it may be possible to make more sensitive tests of the SM by studying decays to final states having contributions from intermediate states with one vector and one pseudoscalar meson (VP), rather than in those with two pseudoscalars.

Several methods to test the SM with B meson decays to charmless VP ($K^{*}\pi$ and $K\rho$) states have been proposed [13–18]. The experimental inputs needed for these methods are the magnitudes and relative phases of the decay amplitudes. Although the phases can only be obtained from Dalitz plot analyses of B meson decays to final states containing one kaon and two pions, the magnitudes can be obtained from simplified approaches. Dalitz plot analyses have been performed for the decays $B^+ \rightarrow K^{*+}\pi^+$ [19, 20], $B^0 \rightarrow K^0_S\pi^+\pi^-$ [21, 22] and $B^0 \rightarrow K^+\pi^-\pi^0$ [23]. Decays of B mesons to $K^{*}K$ final states can in principle be studied with similar methods, but the existing experimental results are less precise [24–29]. No previous measurements of $B_s^0$ meson decays to charmless VP final states exist. First results from the LHCb collaboration on inclusive three-body charmless $B_s^0$ decays have recently become available [30], but no attempt has previously been made to separate the different resonant and nonresonant contributions to their Dalitz plots.

In this paper, the first measurements of $B_s^0$ meson decays to $K^{*}\pi^+$ and $K^{*}\pm K^\mp$ final states and of the $B^0 \rightarrow K^{*}\pm K^\mp$ rate are reported. Throughout the remainder of the paper the symbol $K^*$ is used to denote the $K^*(892)$ resonance. Unique charge assignments of the final state
particles are specified in the expression $B^0 \rightarrow K^\ast^- \pi^+$ because the amplitude for $B^0 \rightarrow K^\ast^+ \pi^-$ is expected to be negligibly small; however, the inclusion of charge-conjugate processes is implied throughout the paper. The branching fractions are measured relative to that of the $B^0 \rightarrow K^\ast^+ \pi^-$ decay, which is known from previous measurements, $B(B^0 \rightarrow K^\ast^+ \pi^-) = (8.5 \pm 0.7) \times 10^{-6}$ [31]. Each of the relative branching fractions for $B^0 \rightarrow K^\ast^\pm h^\mp$, where $h$ refers either to a pion or kaon, are determined as

$$\frac{B(B^0 \rightarrow K^\ast^\pm h^\mp)}{B(B^0 \rightarrow K^\ast^\pm \pi^-)} = \frac{f_d \epsilon(B^0 \rightarrow K^\ast^\pm \pi^-)}{f_\pi \epsilon(B^0 \rightarrow K^\ast^\pm h^\mp)} N(B^0 \rightarrow K^\ast^\pm h^\mp),$$

while that for $B^0 \rightarrow K^\ast^\pm K^\pm$ is determined as

$$\frac{B(B^0 \rightarrow K^\ast^\pm K^\mp)}{B(B^0 \rightarrow K^\ast^\pm \pi^-)} = \frac{\epsilon(B^0 \rightarrow K^\ast^\pm \pi^-)}{\epsilon(B^0 \rightarrow K^\ast^\pm K^\mp)} N(B^0 \rightarrow K^\ast^\pm K^\mp),$$

where $N$ are signal yields obtained from data, $\epsilon$ are efficiencies obtained from simulation and corrected for known discrepancies between data and simulation, and the ratio of fragmentation fractions $f_\pi/f_d = 0.259 \pm 0.015$ [32–34]. With this approach, several potentially large systematic uncertainties cancel in the ratios. The $K^{\ast^\pm}$ mesons are reconstructed in their decays to $K_S^0 \pi^\pm$ with $K_S^0 \rightarrow \pi^+\pi^-\pi^0$ and therefore the final states $K_S^0 \pi^\pm h^\mp$, as well as the data sample, are identical to those studied in [30].

Although the analysis shares several common features to that of the previous publication [30], the selection is optimized independently based on the expected level of background within the allowed $K_S^0 \pi^\pm$ mass window. The data sample used is too small for a detailed Dalitz plot analysis, and therefore only branching fractions are measured. The fit used to distinguish signal from background is an unbinned maximum likelihood fit in the two dimensions of $B$ candidate and $K^\ast$ candidate invariant masses. This approach allows the resonant $B \rightarrow K^\ast^\pm h^\mp$ decay to be separated from other $B$ meson decays to the $K_S^0 \pi^\pm h^\mp$ final state. It does not, however, account for interference effects between the $K^\ast^\pm h^\mp$ component and other amplitudes contributing to the Dalitz plot; possible biases due to interference are considered as a source of systematic uncertainty.

2. The LHCb detector

The analysis is based on a data sample corresponding to an integrated luminosity of 1.0 fb$^{-1}$ of $pp$ collisions at a centre-of-mass energy of 7 TeV recorded with the LHCb detector at CERN. The LHCb detector [35] 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) [36] 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 [37] placed downstream. The tracking system provides a momentum measurement with relative uncertainty that varies from 0.4% at low momentum to 0.6% at 100 GeV/c. The minimum distance of a track to a primary vertex, the impact parameter, is measured with resolution of 20 $\mu$m for tracks with large momentum transverse to the beamline ($p_T$). Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors [38]. Photon, electron and hadron candidates 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 [39].

The trigger [40] consists of hardware and software stages. The hadron trigger at the hardware stage requires that there is at least one particle with transverse energy $E_T > 3.5\text{GeV}$. Events containing candidate signal decays are required to have been triggered at the hardware level in one of two ways. Events in the first category are triggered by particles from candidate signal decays that have an associated calorimeter energy deposit above the threshold, while those in the second category are triggered independently of the particles associated with the signal decay. Events that do not fall into either of these categories are not used in the subsequent analysis. The software trigger requires a two-, three- or four-track secondary vertex with a large sum of the $p_T$ of the tracks and a significant displacement from the primary $pp$ interaction vertices (PVs). A multivariate algorithm [41] is used for the identification of secondary vertices consistent with the decay of a $b$ hadron.

Simulated events are used to study the detector response to signal decays and to investigate potential sources of background. In the simulation, $pp$ collisions are generated using PYTHIA [42] with a specific LHCb configuration [43]. Decays of hadronic particles are described by EVTGEN [44], in which final state radiation is generated using PHOTOS [45]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [46] as described in [47].

3. Selection requirements

The trigger and preselection requirements are identical to those in [30]. As in that analysis, and those of other final states containing $K_S^0$ mesons [48–52], candidate signal decays, i.e. combinations of tracks that are consistent with the signal hypothesis, are separated into two categories: ‘long’, where both tracks from the $K_S^0 \rightarrow \pi^+\pi^-$ decay contain hits in the VELO, and ‘downstream’, where neither does. Both categories have associated hits in the tracking detectors downstream of the magnet. Since long candidates have better mass, momentum and vertex resolution, different selection requirements are imposed for the two categories.

The two tracks originating from the $B$ decay vertex, referred to hereafter as ‘bachelor’ tracks, are required not to have associated hits in the muon system. Backgrounds from decays with charm or charmonia in the intermediate state are vetoed by removing candidates with two-body invariant mass under the appropriate final state hypothesis within $30\text{MeV}/c$ of the known masses [53]. Vetoes are applied for $J/\psi \rightarrow \pi^+\pi^-$ or $K^+K^-$, $\Upsilon_c(0) \rightarrow \pi^+\pi^-$ or $K^+K^-\pi^+$, $D^0 \rightarrow K^-\pi^+$, $\pi^+\pi^-$ or $K^+K^-$, $D^+ \rightarrow K_S^0\pi^+$ or $K_S^0K^+$, $D^+_s \rightarrow K_S^0\pi^+$ or $K^0_SK^+$ and $\Lambda_c^+ \rightarrow K_S^0p$ decays.

The largest source of potential background is from random combinations of final state particles, hereafter referred to as combinatorial background. Signal candidates are separated from this source of background with the output of a neural network [54] that is trained and optimized separately for long and downstream candidates. In the training, simulated $B^0_s \rightarrow K^{*\pm}K^{\mp}$ decays are used to represent signal, and data from the high mass sideband of $K_S^0\pi^+\pi^-$ candidates are used as a background sample (the sideband is $40 < m(K_S^0\pi^+\pi^-) - m_{B^0} < 150\text{MeV}/c$, where $m_{B^0}$ is the known value of the $B^0$ mass [53]). The variables used are: the values of the impact parameter $\chi^2$, defined as the difference in $\chi^2$ of the associated PV with and without the considered particle, for the bachelor tracks and the $K_S^0$ and $B$ candidates; the vertex fit $\chi^2$ for the
\(K_S^0\) and \(B\) candidates; the angle between the \(B\) candidate flight direction and the line between the associated PV and the decay vertex; the separation between the PV and the decay vertex divided by its uncertainty; and the \(B\) candidate \(p_T\). Some of these variables are transformed into their logarithms or other forms that are more appropriate for numerical handling. The consistency of the distributions of these variables between data and simulation is confirmed for \(B^0 \to K_S^0 \pi^+\pi^-\) decays using the \(sPlot\) technique [55] with the \(B\) candidate mass as discriminating variable.

The criteria on the outputs of the neural network are chosen to optimize the probability to observe the \(B_s^0 \to K^{*\pm}K^{\mp}\) decay with significance exceeding five standard deviations (\(\sigma\)) [56]. For the optimization, an additional requirement on the \(K_S^0\pi^\pm\) invariant mass, \(\left|m(K_S^0\pi^\pm) - m_{K^{*\pm}}\right| < 100\) MeV/c with \(m_{K^{*\pm}}\) the known \(K^{*\pm}\) mass, calculated with the \(B\) and \(K_S^0\) candidates constrained to their known masses, is imposed to select the \(K^{*\pm}\) dominated region of the phase space. The requirements on the neural network output give signal efficiencies exceeding 90\% for candidates containing long \(K_S^0\) candidates and exceeding 80\% for candidates containing downstream \(K_S^0\) candidates, while approximately 95\% and 92\% of the background is removed from the two categories, respectively.

Requirements are imposed on particle identification information, primarily from the ring-imaging Cherenkov detectors [38], to separate \(K^{*\pm}K^{\mp}\) and \(K^{*\pm}\pi^{\mp}\) decays. The criteria are chosen based on optimization of a similar figure of merit to that used to obtain the requirement on the neural network output, and retain about 70\% of \(K^{*\pm}K^{\mp}\) and about 75\% of \(K^{*\pm}\pi^{\mp}\) decays. Candidates with tracks that are likely to be protons are rejected. After all selection requirements are applied, below 1\% of events containing one candidate also contain a second candidate; all such candidates are retained.

4. Determination of signal yields

Candidates with masses inside the fit windows of \(5000 < m(K_S^0\pi^\pm h^\mp) < 5500\) MeV/c and \(650 < m(K_S^0\pi^\pm) < 1200\) MeV/c are used to perform extended unbinned maximum likelihood fits to determine the signal yields. In these fits, signal decays are separated from several categories of background by exploiting their distributions in both \(m(K_S^0\pi^\pm h^\mp)\) and \(m(K_S^0\pi^\pm)\). The mass of the \(K_S^0\pi^\pm h^\mp\) combination is calculated assigning either the kaon or pion mass to \(h^\mp\) according to the outcome of the particle identification requirement. A single simultaneous fit to both long and downstream candidates is performed. Separate fits are performed for \(K^{*\pm}K^{\mp}\) and \(K^{*\pm}\pi^{\mp}\) candidates.

In addition to the signal components and combinatorial background, candidates can originate from several other \(b\) hadron decays. Potential sources include: decays of \(B^0\) and \(B_s^0\) mesons to \(K_S^0\pi^\pm h^\mp\) final states without an intermediate \(K^*\) state (referred to as ‘nonresonant’); misidentified \(B^0\to K^{*\pm}h^\mp\) (referred to as ‘cross-feed’) and \(A_0^0\to K^{*\pm}p\) decays; decays of \(B\) mesons to charmless final states with an additional unreconstructed pion; and \(B^+\to D^0h^+, D^0\to K_S^0\pi^+\pi^-\) decays where the additional pion is not reconstructed. Where branching fraction measurements exist [31, 50, 53], the yields of the background sources, except that for nonresonant \(B^0\to K_S^0\pi^+\pi^-\) decays, are expected to be less than 10\% of those for \(B_s^0\to K^{*\pm}K^{\mp}\). The branching fractions of the other nonresonant decays have not been previously determined.

The fit includes components for both \(B^0\) and \(B_s^0\) signal and nonresonant components, and the sources of background listed above. The signal components are parametrized by a Crystal
Ball (CB) function \[57\] in \(B\) candidate mass and a relativistic Breit–Wigner (RBW) function in \(K^*\) candidate mass. The peak positions and widths of the functions for the dominant contribution (\(B_s^0\) for \(K^{*\pm}K^\mp\), \(B^0\) for \(K^{*\pm}\pi^\mp\)) are allowed to vary freely in the fit. The relative positions of the \(B^0\) and \(B_s^0\) peaks in the \(B\) candidate mass distribution are fixed according to the known \(B^0–B_s^0\) mass difference \[53\]. The tail parameters of the CB function are fixed to the values found in fits to simulated signal events, as are the relative widths of the \(B^0\) and \(B_s^0\) shapes. Cross-feed contributions are also described by the product of CB and RBW functions with parameters determined from simulation. The misidentification causes a shift and a smearing of the \(B\) candidate mass distribution and only small changes to the shape in the \(K^*\) candidate mass.

The \(B\) candidate mass distributions for the nonresonant components are also parametrized by a CB function, with peak positions and widths identical to those of the signal components, but with different tail parameters that are fixed to values obtained from simulation. Within the \(K^*\) mass window considered in the fit, the nonresonant shape can be approximated with a linear function. All linear functions used in the fit are parametrized by their yield and the abscissa value at which they cross zero, and are set to zero beyond this threshold, \(m_0\). The relative yields of nonresonant and signal components are constrained to have the same value in the samples with long and downstream candidates, but this ratio is allowed to be different for \(B^0\) and \(B_s^0\) decays.

Backgrounds from other \(b\) hadron decays are described nonparametrically by kernel functions \[58\] in the \(B\) candidate mass and either RBW or linear functions in the \(K^*\) candidate mass, depending on whether or not the decay involves a \(K^*\) resonance. All these background shapes are determined from simulation. To reduce the number of free parameters in the fit to the \(K^{*\pm}K^\mp\) sample, the yields of the backgrounds from charmless hadronic \(B\) meson decays with missing particles are fixed relative to the yield for the \(B^0\rightarrow K^{*\pm}\pi^\mp\) cross-feed component according to expectation. The yield of the \(B^+\rightarrow D^0h^+, D^0\rightarrow K_s^0\pi^+\pi^\mp\) component is determined from the fit to data. The yield for the \(\Lambda_{b}^0\rightarrow K^{*\pm}\pi^\mp\) contribution is also a free parameter in the fit to \(K^{*\pm}K^\mp\) candidates, but is fixed to zero in the fit to \(K^{*\pm}\pi^\mp\) candidates.

The combinatorial background is modelled with linear functions in both \(B\) and \(K^*\) candidate mass distributions, with parameters freely varied in the fit to data except for the \(m_0\) threshold in \(B\) candidate mass, which is fixed from fits to sideband data. For all components, the factorization of the two-dimensional probability density functions into the product of one-dimensional functions is verified to be a good approximation using simulation and sideband data. In total there are 20 free parameters in the fit to the \(K^{*\pm}K^\mp\) sample: yields for \(B^0\) and \(B_s^0\) signals, cross-feed, \(\Lambda_{b}^0\), \(B\rightarrow Dh\) and combinatorial backgrounds (all for both long and downstream categories); ratios of yields for the \(B^0\) and \(B_s^0\) nonresonant components; peak position and width parameters for the signal in both \(B\) candidate and \(K^*\) candidate mass distributions; and parameters of the linear functions describing the combinatorial background in \(K^*\) candidate mass for both long and downstream categories. The fit to the \(K^{*\pm}\pi^\mp\) sample has the same number of free parameters, with the \(\Lambda_{b}^0\) background yields replaced by charmless background yields. The stability of both fits is confirmed using simulated pseudoexperiments.

The results of the fits are shown in figures 2 and 3 for the \(K^{*\pm}K^\mp\) and \(K^{*\pm}\pi^\mp\) final states, respectively, and the signal yields are given in table 1. All other fit results are consistent with expectations.
5. Systematic uncertainties

Systematic uncertainties occur due to possible imperfections in the fit model used to determine the signal yields, and due to imperfect knowledge of the efficiencies used to convert the yields to branching fraction results. A summary of the systematic uncertainties is given in table 2.

The fixed parameters in the functions describing the signal and background components are varied within their uncertainties, and the changes in the fitted yields are assigned as systematic uncertainties. Studies with simulated pseudoexperiments cannot exclude biases on
Figure 3. Results of the fit to $K^{*\pm}\pi^\mp$ candidates projected onto (a), (b) $B$ candidate and (c), (d) $K^*$ candidate mass distributions, for (a), (c) long and (b), (d) downstream candidates. The total fit result (black solid line) is shown together with the data points. Components for the $B^0$ (red dash dotted line) and $B_s^0$ (pink dash double-dotted line) signals are shown together with $B^0$ (dark red falling-hatched area) and $B_s^0$ (purple rising-hatched area) nonresonant components, partially reconstructed and cross-feed background (blue long-dashed line), and combinatorial background (green long-dash-dotted line) components.

Table 1. Yields and relative yields obtained from the fits to $K^{*\pm}\pi^\mp$ and $K^{*\pm}\pi^\mp$ candidates. The relative yields of nonresonant (NR) $B_{(s)}^0$ decays are constrained to be identical in long and downstream categories. Only statistical uncertainties are given.

<table>
<thead>
<tr>
<th>Yield</th>
<th>$B^0$</th>
<th>$B_s^0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Long</td>
<td>Downstream</td>
</tr>
<tr>
<td>$N (K^{*\pm}\pi^\mp)$</td>
<td>0 ± 4</td>
<td>4 ± 3</td>
</tr>
<tr>
<td>$N (K^{*\pm}\pi^\mp)$</td>
<td>80 ± 10</td>
<td>165 ± 16</td>
</tr>
<tr>
<td>$N (K_S^{0}\pi^{\pm}\pi^\mp \text{NR})/N (K^{*\pm}\pi^\mp)$</td>
<td>0.0 ± 1.0</td>
<td>0.41 ± 0.16</td>
</tr>
<tr>
<td>$N (K_S^{0}\pi^{\pm}\pi^\mp \text{NR})/N (K^{*\pm}\pi^\mp)$</td>
<td>0.79 ± 0.14</td>
<td>0.6 ± 0.4</td>
</tr>
</tbody>
</table>
the yields at the level of a few decays. An uncertainty corresponding to the size of the possible bias is assigned. The linear approximation for the shape of the nonresonant component in the $K^*$ candidate mass can only be valid over a restricted range. Therefore the mass window is varied and the change in the fitted results taken as an estimate of the corresponding uncertainty.

The largest source of systematic uncertainty arises due to imperfect cancellation of interference effects between the P-wave $K^*$ signal and the nonresonant component, in which the $K_S^0\pi^\pm$ system is predominantly S-wave. Since the efficiency is not uniform as a function of the cosine of the decay angle, $\cos \theta_{K^+}$, defined as the angle between the $B$ and $K_S^0$ candidate momenta in the rest frame of the $K_S^0\pi^\pm$ system, a residual interference effect may bias the results. The size of this uncertainty is evaluated by fitting the distribution of $\cos \theta_{K^+}$ [59]. The distribution is reconstructed from the signal $s$Weights [55] obtained from the default fit. Only the region where $\cos \theta_{K^+}$ is positive is considered, since the efficiency variation is highly nontrivial in the negative region. This ensures that the assigned uncertainty is conservative since any cancellation of the interference effects between the two sides of the distribution is neglected. In the absence of interference, the distribution will be parabolic and pass through the origin. The bias on the signal yield due to interference can therefore be evaluated from the constant and linear components resulting from a fit of the distribution to a second-order polynomial. Such fits are shown for $B^0 \rightarrow K^{*+}\pi^-$ and $B^0_s \rightarrow K^{*+}\pi^-$ signals in figure 4. The measured yields of the $B_s^0 \rightarrow K^{*-}\pi^+$ and $B^0 \rightarrow K^{*}\pi^+$ signals are too small to allow this method to be used. Therefore the same relative uncertainties are assigned to these decays as in the corresponding $B^0$ or $B^0_s$ decay.

Systematic uncertainties on the ratio of efficiencies arise due to limited sizes of the simulation samples used to determine the acceptance and selection efficiencies, and due to possible mismodelling of the detector response. Two potential sources of mismodelling are the trigger and particle identification efficiencies. These are determined from control samples and systematic uncertainties assigned using the same procedures as described in [30]. The imperfect

<table>
<thead>
<tr>
<th>Source</th>
<th>$B(B^0 \rightarrow K^{<em>+}\pi^-)/B(B^0 \rightarrow K^{</em>}\pi)$</th>
<th>$B(B^0 \rightarrow K^{<em>-}\pi^+)/B(B^0 \rightarrow K^{</em>}\pi)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Long Downstream</td>
<td>Long Downstream</td>
</tr>
<tr>
<td>Fit</td>
<td>0.14</td>
<td>0.07</td>
</tr>
<tr>
<td>S-Wave interference</td>
<td>0.32</td>
<td>0.14</td>
</tr>
<tr>
<td>Acceptance</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Selection</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>Trigger</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>Particle identification</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>$f_s/f_d$</td>
<td>0.10</td>
<td>0.08</td>
</tr>
<tr>
<td>Total</td>
<td>0.37</td>
<td>0.19</td>
</tr>
</tbody>
</table>
knowledge of the ratio of fragmentation fractions, $f_s/f_d = 0.259 \pm 0.015$ [32–34], is another source of uncertainty.

6. Results and conclusion

The significance of the signal strengths is determined from $\sqrt{-2\Delta \ln L}$, where $\Delta \ln L$ is the change in the log likelihood between the default fit result and that obtained when the relevant component is fixed to zero. This calculation is performed both with only the statistical uncertainty included, and after the likelihood function is convolved with a Gaussian function with width corresponding to the systematic uncertainty on the fitted yield. Combining the likelihoods from long and downstream categories, the statistical significances for $B^0 \rightarrow K^{*\pm}\pi^\mp$ and $B_s^0 \rightarrow K^{*\pm}\pi^\mp$ decays are $12.5 \sigma$ and $3.9 \sigma$ while the corresponding values for the total significance are $7.8 \sigma$ and $3.4 \sigma$, respectively. The significance of the $B^0 \rightarrow K^{*\pm}\pi^\mp$ signal is below $2 \sigma$. 

Figure 4. Background-subtracted distribution of $\cos \theta_{K^*}$ for (a), (b) $B^0 \rightarrow K^{*\pm}\pi^\mp$ and (c), (d) $B_s^0 \rightarrow K^{*\pm}\pi^\mp$ signals from the samples with (a), (c) long and (b), (d) downstream candidates. Results of fits with second-order polynomial functions are shown as the solid lines.
The ratios of branching fractions of equations (1) and (2) are obtained by correcting the ratios of yields by the ratios of efficiencies and, where appropriate, by the ratio of the fragmentation fractions. The particle identification efficiencies are determined from data, using samples of kaons and pions from $D^{*+} \to D^0 \pi^+$, $D^0 \to K^-\pi^+$ decays reweighted according to the kinematic distributions of the bachelor tracks in $B_s^0 \to K^{*}\pi^\mp$ decays. The relative efficiencies of the acceptance and all other selection requirements are determined from simulation. The relative efficiencies are within 10% of unity.

Since the signal for $B^0 \to K^{*}\pi^\mp$ decays is not significant, upper limits at 90% and 95% confidence level (CL) are obtained by integrating the profile likelihood function in the region of positive branching fraction. All results from the samples with long and downstream candidates are consistent and the combined results are

$$
\frac{B(B_s^0 \to K^{*}\pi^\mp)}{B(B^0 \to K^{*}\pi^\mp)} = 1.49 \pm 0.22 \text{ (stat)} \pm 0.18 \text{ (syst)},
$$

$$
\frac{B(B_s^0 \to K^{*}\pi^\mp)}{B(B^0 \to K^{*}\pi^\mp)} = 0.02 \pm 0.02 \text{ (stat)} \pm 0.01 \text{ (syst)},
$$

$$
< 0.05 \text{ (0.06) at 90\% (95\%) CL},
$$

$$
\frac{B(B_s^0 \to K^{*}\pi^\mp)}{B(B^0 \to K^{*}\pi^\mp)} = 0.39 \pm 0.13 \text{ (stat)} \pm 0.05 \text{ (syst)}.
$$

Multiplying the relative branching fractions by $B(B^0 \to K^{*}\pi^\mp) = (8.5 \pm 0.7) \times 10^{-6}$ [31] gives

$$
B(B_s^0 \to K^{*}\pi^\mp) = (12.7 \pm 1.9 \text{ (stat)} \pm 1.9 \text{ (syst)}) \times 10^{-6},
$$

$$
B(B_s^0 \to K^{*}\pi^\mp) = (0.17 \pm 0.15 \text{ (stat)} \pm 0.05 \text{ (syst)}) \times 10^{-6},
$$

$$
< 0.4 \text{ (0.5) } \times 10^{-6} \text{ at 90\% (95\%) CL},
$$

$$
B(B_s^0 \to K^{*}\pi^\mp) = (3.3 \pm 1.1 \text{ (stat)} \pm 0.5 \text{ (syst)}) \times 10^{-6}.
$$

In summary, $B_s^0 \to K^{*}h^\mp$ decays have been studied using a data sample corresponding to 1.0 fb$^{-1}$ of $pp$ collision data at a centre-of-mass energy of 7 TeV collected with the LHCb detector. The first observation of the $B_s^0 \to K^{*}K^\mp$ decay and the first evidence for the $B_s^0 \to K^{*}\pi^\mp$ decay are obtained, and their branching fractions measured. An upper limit is set on the branching fraction of the $B^0 \to K^{*}\pi^\mp$ decay. The results are consistent with several independent theoretical predictions [60–62] and represent an important step towards searches for physics beyond the Standard Model in decays of $B$ mesons to charmless final states containing one pseudoscalar and one vector meson. Dalitz plot analyses of larger samples will allow the reduction of both statistical and systematic uncertainties on these results. The additional sensitivity to relative phases provided by such analyses will also permit searches for sources of $CP$ violation beyond the Standard Model.
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); NSFC (China); CNRS/IN2P3 (France); BMBF, DFG, HGF and MPG (Germany); SFI (Ireland); INFN (Italy); FOM and NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FANO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (USA). The Tier1 computing centres are supported by IN2P3 (France), KIT and BMBF (Germany), INFN (Italy), NWO and SURF (The Netherlands), PIC (Spain), GridPP (United Kingdom). We are indebted to the communities behind the multiple open source software packages on which we depend. We are also thankful for the computing resources and the access to software R&D tools provided by Yandex LLC (Russia). Individual groups or members have received support from EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union), Conseil général de Haute-Savoie, Labex ENIGMASS and OCEVU, Région Auvergne (France), RFBR (Russia), XuntaGal and GENCAT (Spain), Royal Society and Royal Commission for the Exhibition of 1851 (United Kingdom).

The LHCb Collaboration

References

[17] Bediaga I, Guerrero G and de Miranda J M 2007 Extracting the quark mixing phase $\gamma$ from $B^+ \rightarrow K^+\pi^+\pi^-$, $B^0 \rightarrow K^0\pi^+\pi^-$, and $B^0 \rightarrow K^0\pi^+\pi^-$ Phys. Rev. D 76 073001
[19] Garmash A et al (Belle collaboration) 2006 Evidence for large direct CP violation in $B^\pm \rightarrow \rho(770)^0K^\pm$ from analysis of the three-body charmless $B^\pm \rightarrow K^\pm\pi^\pm\pi^\mp$ decays Phys. Rev. Lett. 96 251803
[20] Aubert B et al (BaBar collaboration) 2008 Evidence for direct CP violation from Dalitz-plot analysis of $B^\pm \rightarrow K^\pm\pi^\mp\pi^\mp$ Phys. Rev. D 78 012004
[21] Dalseno J et al (Belle collaboration) 2009 Time-dependent Dalitz plot measurement of $CP$ parameters in $B^0 \rightarrow K^0\pi^+\pi^-$ decays Phys. Rev. D 79 072004
[22] Aubert B et al (BaBar collaboration) 2009 Time-dependent amplitude analysis of $B^0 \rightarrow K^0\pi^+\pi^-$ Phys. Rev. D 80 112001
[23] Lees J P et al (BaBar collaboration) 2011 Amplitude analysis of $B^0 \rightarrow K^+\pi^-\pi^0$ and evidence of direct $CP$ violation in $B \rightarrow K^*\pi$ decays Phys. Rev. D 83 112010
[24] Aubert B et al (BaBar collaboration) 2006 Search for the decay of a $B^0$ or $B^0\overline{B}^0$ meson to $K^{*0}\overline{K}^{*0}$ or $K^{*0}\overline{K}^{*0}$ Phys. Rev. D 74 072008
[27] Aubert B et al (BaBar collaboration) 2009 Search for the decay $B^+ \rightarrow K^0S\pi^+$ Phys. Rev. D 79 051101
[28] del Amo Sanchez P et al (BaBar collaboration) 2010 Observation of the rare decay $B^0 \rightarrow K^0S\pi^+\pi^-$ Phys. Rev. D 82 031101
[29] Gaur V et al (Belle collaboration) 2013 Evidence for the decay $B^0 \rightarrow K^+K^-\pi^0$ Phys. Rev. D 87 091101
[30] Aaij R et al (LHCb collaboration) 2013 Study of $B^0_\gamma \rightarrow K^0_\gamma h^+h^-$ decays with first observation of $B^0_\gamma \rightarrow K^0_\gamma K^0\pi^+\pi^-$ and $B^0_\gamma \rightarrow K^0_\gamma \pi^+\pi^-$. J. High Energy Phys. JHEP10(2013)143
[34] LHCb collaboration. Updated average $f_s/f_d$ $b$-hadron production fraction ratio for 7 TeV $pp$ collisions LHCb-CONF-2013–011 https://cds.cern.ch/record/1559262
[37] Arink R et al 2014 Performance of the LHCb Outer Tracker JINST 9 P01002
[40] Aaij R et al 2013 The LHCb trigger and its performance JINST 8 P04022
[41] Gligorov V V and Williams M 2013 Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree JINST 8 P02013
[43] Belyaev I et al 2010 Handling of the generation of primary events in GAUSS, the LHCb simulation framework Nuclear Science Symposium Conference Record (NSS/MIC) IEEE p 1155
[50] Aaij R et al (LHCb collaboration) 2014 Searches for $\Lambda_b^0$ and $\Xi_b^0$ decays to $K_0^0 p\pi^-$ and $K_0^0 pK^-$ final states with first observation of the $\Lambda_b^0 \to K_0^0 p\pi^-$ decay J. High Energy Phys. JHEP04(2014)087
[59] Aaij R et al (LHCb collaboration) 2014 Measurement of the charge asymmetry in $B^\pm \to \phi K^\mp$ and search for $B^\pm \to \phi\pi^\mp$ decays Phys. Lett. B 728 85