First observation of $B^+ \rightarrow D_s^+ K^+ K^-$ decays and a search for $B^+ \rightarrow D_s^+ \phi$ decays

The LHCb collaboration

E-mail: tom.hadavizadeh@cern.ch

ABSTRACT: A search for $B^+ \rightarrow D_s^+ K^+ K^-$ decays is performed using $pp$ collision data corresponding to an integrated luminosity of 4.8 fb$^{-1}$, collected at centre-of-mass energies of 7, 8 and 13 TeV with the LHCb experiment. A significant signal is observed for the first time and the branching fraction is determined to be

$$B(B^+ \rightarrow D_s^+ K^+ K^-) = (7.1 \pm 0.5 \pm 0.6 \pm 0.7) \times 10^{-6},$$

where the first uncertainty is statistical, the second systematic and the third due to the uncertainty on the branching fraction of the normalisation mode $B^+ \rightarrow D_s^+ D_s^0$. A search is also performed for the pure annihilation decay $B^+ \rightarrow D_s^+ \phi$. No significant signal is observed and a limit of

$$B(B^+ \rightarrow D_s^+ \phi) < 4.9 \times 10^{-7} \ (4.2 \times 10^{-7})$$

is set on the branching fraction at 95% (90%) confidence level.

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

ArXiv ePrint: 1711.05637
The LHCb collaboration

1 Introduction

The decay $B^+ \rightarrow D_s^+ K^+ K^-$ is mediated by a $\bar{b} \rightarrow \pi$ transition shown in figure 1 and is therefore suppressed in the Standard Model (SM) due to the small size of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element $V_{ub}$. The branching fraction for this decay is currently not measured, however a similar decay, $B^+ \rightarrow D_s^+ \pi^0$, has been observed with a branching fraction of $\mathcal{B}(B^+ \rightarrow D_s^+ \pi^0) = (1.5 \pm 0.5) \times 10^{-5}$ [1].

In the SM, the decay $B^+ \rightarrow D_s^+ \phi$ proceeds dominantly via the annihilation diagram shown in figure 1. This suppressed topology requires the wave functions of the incoming quarks to overlap sufficiently to annihilate into a virtual $W^+$ boson. The decay is further suppressed by the small magnitude of the CKM matrix element $V_{ub}$ associated with the annihilation vertex. In addition, unlike many rare hadronic decays including $B^+ \rightarrow D_s^+ K^+ K^-$, possible contributions from rescattering effects are expected to be small, for example contributions from intermediate states such as $B^+ \rightarrow D_s^+ \omega$ [2]. Several SM predictions have been made for the branching fraction of the $B^+ \rightarrow D_s^+ \phi$ decay [3–6], using input from lattice calculations [7–9]. These predictions are in the range $(1 - 7) \times 10^{-7}$, where the limit on the precision is dominated by hadronic uncertainties. However, additional diagrams contributing to this decay can arise in some extensions of the SM, such as
supersymmetric models with R-parity violation. They could enhance the branching fraction and/or produce large CP asymmetries [4, 5], which makes the $B^+ \rightarrow D_s^+ \phi$ decay a promising place to search for new physics beyond the SM.\footnote{Charge conjugation is implied throughout this paper. Furthermore, $\phi$ denotes the $\phi(1020)$ resonance.}

The LHCb experiment reported evidence for the decay $B^+ \rightarrow D_s^+ \phi$ using $pp$ collision data corresponding to an integrated luminosity of 1 fb$^{-1}$ taken during 2011, at a centre-of-mass energy of 7 TeV [10]. A total of $6.7^{+4.5}_{-2.8}$ candidates was observed. The branching fraction was determined to be

$$B(B^+ \rightarrow D_s^+ \phi) = (1.87^{+1.25}_{-0.73} \pm 0.19 \pm 0.32) \times 10^{-6},$$  \hspace{1cm} (1.1)$$

where the first uncertainty is statistical, the second is systematic and the third is due to the uncertainty on the branching fraction of the decay $B^+ \rightarrow D_s^+ \overline{D}^0$, which was used as normalisation. Given the large uncertainties on both the theoretical and experimental values, the previously measured value is consistent with the range of SM values given above.

The measurements presented in this paper reanalyse the data collected in 2011, whilst adding data corresponding to an integrated luminosity of 2 fb$^{-1}$ collected at a centre-of-mass energy 8 TeV in 2012, along with 0.3 fb$^{-1}$ from 2015 and 1.5 fb$^{-1}$ from 2016, both at 13 TeV. They supersede the previous measurement [10].

This analysis is performed in two parts: firstly $B^+ \rightarrow D_s^+ K^+ K^-$ decays are reconstructed across the entire phase space and then a dedicated search for $B^+ \rightarrow D_s^+ \phi$ decays is performed in a narrow region of $K^+ K^-$ invariant mass around the $\phi$ meson. The branching fractions are determined using the decay $B^+ \rightarrow D_s^+ \overline{D}^0$, with $\overline{D}^0 \rightarrow K^+ K^-$, as a normalisation channel. Although this $\overline{D}^0$ decay has a smaller branching fraction than $\overline{D}^0 \rightarrow K^+ \pi^-$ (0.4% vs. 3.9% [11]), sharing the same final state between the signal and normalisation channel reduces systematic uncertainties in the ratio of detection efficiencies.

2 Detector and data sample

The LHCb detector [12, 13] 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 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 GeV/$c$. The minimum distance of a track to a primary $pp$ interaction 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 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. The online event selection is performed by a trigger, which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction.

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. Two different algorithms are used in the software trigger to select candidates for this analysis. The first uses a multivariate algorithm [14] to identify the presence of a secondary vertex that has two, three or four tracks and is displaced from any PV. At least one of these charged particles must have a transverse momentum $p_T > 1.7$ GeV/$c$ and be inconsistent with originating from a PV. The second algorithm selects $\phi$ candidates decaying to two charged kaons. Each kaon must have a transverse momentum $p_T > 0.8$ GeV/$c$ and be inconsistent with originating from a PV. The invariant mass of the kaon pair must be within 20 MeV/$c^2$ of the known $\phi$ mass [11]. This algorithm is used in both the search for $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ decays.

Simulated events are used to determine the relative efficiencies of the signal and normalisation channels. The samples are generated for each of the running periods. In these simulations, $pp$ collisions are generated using PYTHIA [15, 16] with a specific LHCb configuration [17]. Decays of hadronic particles are described by EVTGEN [18], in which final-state radiation is generated using PHOTOS [19]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [20, 21] as described in ref. [22].

3 Candidate selection

Candidate $B^+ \rightarrow D_s^+ \phi$ and $B^+ \rightarrow D_s^+ K^+ K^-$ decays are selected using similar requirements. The $\phi$ mesons in $B^+ \rightarrow D_s^+ \phi$ candidates are reconstructed with $\phi \rightarrow K^+ K^-$. Both modes are reconstructed using the $D_s^+ \rightarrow K^+ K^- \pi^+$ decay, whilst $B^+ \rightarrow D_s^+ \phi$ candidates are additionally reconstructed with the decays $D_s^+ \rightarrow K^+ \pi^- \pi^+$ and $D_s^+ \rightarrow \pi^+ \pi^- \pi^+$ to increase the sensitivity of the search. The $D_s^+$ ($\phi$) candidates are required to have an invariant mass within 25 MeV/$c^2$ (40 MeV/$c^2$) of the known $D_s^+$ ($\phi$) mass [11]. In the search for $B^+ \rightarrow D_s^+ K^+ K^-$ decays, the veto $|m(K^+ K^-) - m(D^0)| > 25$ MeV/$c^2$ is applied to explicitly remove the normalisation channel from the signal mode.
The $B^+$ meson candidates are formed from well reconstructed tracks with $\chi^2_{IP} > 4.0$, where $\chi^2_{IP}$ is defined as the difference in the vertex-fit $\chi^2$ of the best PV reconstructed with and without the particle being considered. The best PV is the PV that has the smallest $\chi^2_{IP}$ value. For kaons from the $\phi$ or $B^+$ decay the momentum requirement is $p > 2$ GeV/c. At least one track of each $B^+$ meson candidate must have $p_T > 0.5$ GeV/c and $p > 5$ GeV/c.

Loose requirements are made on particle identification (PID) to reduce background from other $b$-hadron decays with misidentified hadrons. For the signal, the overall efficiency of the PID requirements varies from 80% to 90%, depending on the $D_s^+$ mode. Background from decays of $B^+$ mesons to the same final state that did not proceed via a $D_s^+$ meson (referred to as charmless decays) are suppressed by applying a requirement on the significance of the $B^+$ and $D_s^+$ vertex separation, $\chi^2_{IP}$.

The residual yields of charmless decays are estimated by determining the $B^+$ yield in candidates that are in the invariant mass range $25 < |m(h^+h^{'-}\pi^+) - m(D_s^+)| < 50$ MeV/$c^2$, where $m(h^+h^{'-}\pi^+)$ is the $D_s^+$ candidate mass and $h,h' = K, \pi$. This background estimation is performed separately for the $B^+ \rightarrow D_s^+\phi$ and $B^+ \rightarrow D_s^+K^+K^-$ searches. For the $B^+ \rightarrow D_s^+\bar{D}^0$ normalisation channel, a two-dimensional optimisation is performed to calculate the contribution from decays without a $D_s^+$ meson, $D^0$ meson or both. The optimal selection requirements are chosen such that the maximum signal efficiency is achieved for a residual charmless contribution of 2% of the normalisation yield.

For the decay $D_s^+ \rightarrow K^+K^-\pi^+$, candidates are rejected if they are consistent with $D^+ \rightarrow K^-\pi^+\pi^+$ or $\Lambda_c^+ \rightarrow pK^-\pi^+$ decays, where either a pion or a proton has been misidentified as a kaon. The candidates are reconstructed using the alternative mass hypothesis and, for those falling within 25 MeV/$c^2$ of the $D^+$ or $\Lambda_c^+$ mass, particle identification requirements are tightened on the misidentified track.

Another set of vetoes rejects decays where the tracks forming the $D_s^+$ candidate originate from an excited charged charm meson decay, for example $D^{*+} \rightarrow (D^0 \rightarrow h^++h'^-)\pi^+$. By requiring $\Delta m = m(h^+h'^-\pi^+) - m(h^+h'^-) > 150$ MeV/$c^2$ decays of this type are efficiently removed. Other specific backgrounds are removed by mass vetoes. These vetoes remove $B_s^0 \rightarrow \phi\phi$ decays in which one of the $\phi$ mesons is combined with an unrelated pion to form the $D_s^+$ candidate. Any candidates within 50 MeV/$c^2$ of the known $B_s^0$ mass [11] in the four-body invariant mass $m(K^+K^-K^+K^-)$ are removed to ensure a smooth combinatorial background distribution.

In addition, a veto is applied to the invariant mass of the kaons from the $\phi$ meson or $B^+$ candidate combined with any pion from the $D_s^+$ candidate, removing candidates within 25 MeV/$c^2$ of the known $D_s^+$ mass. This removes decays that include incorrectly reconstructed $D_s^+ \rightarrow \phi\pi^+$ or $D_s^+ \rightarrow K^+K^-\pi^+$ decays, where the $\phi$ or $K^+K^-$ pair are incorrectly assigned to have originated from the $B^+$ meson rather than the $D_s^+$ meson. For example, this incorrect assignment could lead to $B^+ \rightarrow (D_s^+ \rightarrow \phi\pi^+)K^+K^-$ decays being reconstructed as $B^+ \rightarrow (D_s^+ \rightarrow K^+K^-\pi^+)\phi$ decays. The $B^+$ ($D_s^+$) candidates are required to have $\chi^2_{IP} < 10$ ($\chi^2_{IP} > 10$), to ensure they are consistent (inconsistent) with being produced at the best PV.

Multivariate analyses (MVA) are used to separate genuine $\phi$ and $D_s^+$ candidates from random combinations of tracks [23]. The $\phi$ and $D_s^+$ MVAs use data samples of $B_s^0 \rightarrow J/\psi\phi$
and $B_s^0 \to D_s^+ \pi^-$ decays, respectively, where the background is statistically subtracted using the sPlot method [24]. The training uses the $\phi$ or $D_s^+$ sidebands as a background sample. A total of eight MVAs are trained to target the decays $\phi \to K^+ K^-$, $D_s^+ \to K^+ K^- \pi^+$, $D_s^+ \to K^+ \pi^- \pi^+$ and $D_s^+ \to \pi^+ \pi^- \pi^+$, separately in the Run 1 (2011 and 2012) and Run 2 (2015 and 2016) data. A preselection including the trigger, vetoes and PID requirements previously discussed is applied to the training samples, ensuring they are representative of the target signal decays. The samples are split into two subsamples in a random but reproducible way. One is used to train the corresponding MVA, the other to test its response. The MVA method used in this analysis is a gradient Boosted Decision Tree (BDTG) [25]. The selection criteria for each of the BDTG classifiers are determined by optimising the figure of merit $\epsilon_s/(a^2 + \sqrt{N_{BKG}})$ [26], with $a = 5$, where $\epsilon_s$ is the signal efficiency and $N_{BKG}$ is the number of background candidates determined from fits to data, calculated in the signal region.

The efficiencies of the MVAs are obtained from the test samples of $B_s^0 \to J/\psi \phi$ and $B_s^0 \to D_s^+ \pi^-$ decays. Additionally, a sample of $B^+ \to D^0 \pi^+$ decays is used to calculate the efficiency of $D_s^+ \to K^+ K^-$ decays in the normalisation channel. The efficiency calculation takes into account the kinematic differences between the training and signal samples, as well as any possible correlations between the $D_s^+$ and $\phi$ kinematics, by using input from simulation samples. Any further correlations between the $\phi$ and $D_s^+$ MVA efficiencies are found to be negligible. In the search for $B^+ \to D_s^+ K^+ K^-$ decays, calibration samples are used to correct for the imperfect modelling of the PID in simulation. These corrected samples are then used to obtain the variations in the MVA efficiencies as a function of the phase-space position, in particular of the $m(K^+ K^-)$ invariant mass.

The invariant mass of the $B^+$ meson candidates is determined from fits in which the $D_s^+$ candidate mass (and $D^0$ candidate mass for the normalisation channel) is constrained to the known value [27]. Additionally, the momentum vector of the $B^+$ meson is constrained to be parallel to the vector connecting the PV and the $B^+$ meson decay vertex.

4 Invariant mass fits

The branching fractions of the $B^+ \to D_s^+ \phi$ and $B^+ \to D_s^+ K^+ K^-$ decays are determined from unbinned maximum likelihood fits to the invariant mass of the $B^+$ candidates. However, separate fit strategies are used for the $B^+ \to D_s^+ \phi$ and $B^+ \to D_s^+ K^+ K^-$ searches.

The search for $B^+ \to D_s^+ K^+ K^-$ involves two independent fits for the signal and normalisation channels. The $B^+ \to D_s^+ K^+ K^-$ yield is corrected on a per-candidate basis to account for the phase-space dependence of the signal efficiencies in this three-body decay.

In contrast, the $B^+ \to D_s^+ \phi$ candidates are treated as quasi-two-body decays in which all signal candidates are corrected with the same efficiency. The $B^+ \to D_s^+ \phi$ signal and normalisation channels are fitted simultaneously in different categories, as are the three $D_s^+$ decay modes, with the $D_s^+ \to K^+ K^-\pi^+$ mode split further into $D_s^+ \to \phi \pi^+$ and non-$\phi$ submodes. This exploits the high purity of the $D_s^+ \to \phi \pi^+$ decay. As the $B^+ \to D_s^+ \phi$ decay involves the decay of a pseudoscalar particle to a pseudoscalar and vector particle, the $\phi$ vector meson ($J^P = 1^-$) must be produced longitudinally polarised. For a longitudinally
polarised $\phi$ meson decaying to $K^+K^-$, the distribution of the angle $\theta_K$, defined as the angle that the kaon meson forms with the $B$ momentum in the $\phi$ rest frame, is proportional to $\cos^2\theta_K$. The distribution of $\cos\theta_K$ for $B^+ \rightarrow D_s^+\phi$ as determined from simulated events is shown in figure 2. In the simultaneous fit for $B^+ \rightarrow D_s^+\phi$ candidates the candidates are split into two helicity categories: $|\cos\theta_K| > 0.4$ and $|\cos\theta_K| < 0.4$. In simulated events, 93% of $B^+ \rightarrow D_s^+\phi$ decays are found in the first category, whereas for the normalisation decay and background modes, as the distributions in $\cos\theta_K$ are approximately flat, only 60% of candidates fall into this category. Additionally, the fit further assigns candidates into two $m(K^+K^-)$ invariant mass categories, $|m(K^+K^-) - m_\phi| < 10\text{ MeV}/c^2$ and $10 < |m(K^+K^-) - m_\phi| < 40\text{ MeV}/c^2$ (figure 2), to help constrain the contribution from the different backgrounds in the signal region. Background modes involving two kaons that did not originate from a $\phi$ meson (for example $B^0_s \rightarrow D_s^{(*)+}K^-K^{(*)0}$) have different fractions in these two categories, helping to distinguish them from those decays with a real $\phi$ meson. The fractions of $B^+ \rightarrow D_s^+\phi$ candidates in each of the categories, as determined from simulated events, are listed in table 1.

**Figure 2.** Distributions of (left) $\cos\theta_K$ and (right) $m(K^+K^-)$ in $B^+ \rightarrow D_s^+\phi$ decays, as determined from simulated events. The vertical lines represent the limits of the two categories used for each variable. In the $m(K^+K^-)$ distribution, the area within the dashed red lines represents the $\phi$ signal region, and the two areas between the dashed red and blue lines represent the $\phi$ sideband region. The $B^+ \rightarrow D_s^+\phi$ signal decays are seen to primarily contribute to the $\phi$ signal region and the $|\cos\theta_K| > 0.4$ category.

**Table 1.** Fractions of $B^+ \rightarrow D_s^+\phi$ candidates expected in the helicity and $m(K^+K^-)$ invariant mass categories of the simultaneous fit.

| $|m(K^+K^-) - m_\phi| (\text{MeV}/c^2)$ | Helicity Category | $|\cos\theta_K| > 0.4$ | $|\cos\theta_K| < 0.4$ |
|-------------------------------------|-----------------|-----------------|-----------------|
| $< 10$                              | 82%             | 6%              |
| $(10, 40)$                          | 11%             | 1%              |
Table 2. Fractions of $B^+ \to D_s^+ K^+ K^-$ candidates assumed to contribute to each helicity and $m(K^+K^-)$ invariant mass categories of the simultaneous fit. The uncertainties shown are calculated from the range of fractions obtained by assuming different contributing resonances, as detailed in section 4.1.

4.1 Signal and normalisation probability density functions

The normalisation and signal components in the $B^+ \to D_s^+ \bar{D}^0$ and $B^+ \to D_s^+ K^+ K^-$ or $B^+ \to D_s^+ \phi$ invariant mass distributions are each modelled using the sum of two Crystal Ball (CB) [28] probability density functions (PDFs) with tails at lower invariant mass. The tail parameters, the ratio of the two CB widths, and the relative fraction of each CB function are determined from simulated events. The resolution parameter of the narrow CB component in each $D_s^+$ decay mode category is a free parameter in the fit, but the ratios of signal and normalisation widths are fixed to values determined from simulated events. For the normalisation mode, the fraction of $B^+ \to D_s^+ \bar{D}^0$ candidates in the two helicity bins is a free parameter in the fit, whereas for the signal the fraction in each helicity and $m(K^+K^-)$ invariant mass category of the fit is fixed to that determined from simulated events, as reported in table 1.

The search for $B^+ \to D_s^+ \phi$ decays includes a component for $B^+ \to D_s^+ K^+ K^-$ decays that did not proceed via a $\phi$ meson. The fraction of $B^+ \to D_s^+ K^+ K^-$ decays expected in each helicity angle and $m(K^+K^-)$ mass category, shown in table 2, are calculated from the average of different $K^+K^-$ resonances that could contribute to $B^+ \to D_s^+ K^+ K^-$ decays. These resonances include possible contributions from the $f_0(980)$ and $a_0(980)$ resonances. The resulting fractions are sufficiently different from those for the $B^+ \to D_s^+ \phi$ signal such that the two contributions can be distinguished. The range of fractions obtained by considering the different resonances are included as uncertainties in table 2. A systematic uncertainty is assigned to account for the fixed fractions assumed in the fit. No attempt is made to separate any of the contributing resonances in the search for $B^+ \to D_s^+ K^+ K^-$ candidates.

4.2 Background PDFs

A number of background components are included in the fit model. The dominant source of background under the signal is due to combinations of unrelated tracks. An exponential function is used to parametrise this component. The same slope parameter is used in the simultaneous fit to the signal and normalisation modes. Partially reconstructed $B^+ \to D_s^+ \bar{D}^0$ and $B^+ \to D_s^+ \bar{D}^{*0}$ decays are concentrated in the lower part of the $D_s^+ \bar{D}^0$ spectrum. They are parametrised using analytical shapes that account for the nonreconstructed neutral pion or photon from the excited $D$-meson decays. These shapes are
constructed from Gaussian distributions convolved with second-order polynomials, and are analogous to those used in similar analyses [29]. An additional component is used to model $B^+ \to D_{s}^{+}D^{*0}$ decays where one particle from each of the excited $D$ mesons is missed. Partially reconstructed $B^+ \to D_{s}^{*+}\phi$ decays can contribute to the lower part of the $D_{s}^{*+}\phi$ spectrum. These, similarly, are fitted with analytical shapes that account for the missing neutral particle from the $D_{s}^{*+}$ decay, as well as the different helicity states for the decay of a pseudoscalar meson to two vector particles. They are parametrised in an analogous way to similar analyses [30]. This background component is only included in the search for $B_{s}^{0} \to D_{s}^{*+}\phi$ decays. The modes $B_{s}^{0} \to D_{s}^{+}K^{-}K^{*0}$ and $B_{s}^{0} \to D_{s}^{*+}K^{-}K^{*0}$ form a background to $B^+ \to D_{s}^{*+}\phi$ decays when a low-momentum pion from the $K^{*0}$ decay is not reconstructed. Additionally, a neutral pion or photon can be missed from the excited $D_{s}^{+}$ meson decay in the case of $B_{s}^{0} \to D_{s}^{*+}K^{-}K^{*0}$. The PDFs are determined from simulated events. The expected fractions in each category of the $B^+ \to D_{s}^{*+}\phi$ fit are fixed using simulated events. The decays $B_{s}^{0} \to D_{s}^{+}D_{s}^{-}$, $B_{s}^{0} \to D_{s}^{*+}D_{s}^{-}$ and $B^0 \to D_{s}^{*+}D_{s}^{-}$ can form a background when a pion is not reconstructed from a $D_{s}^{*}$ or $D^+$ decaying to $K^{+}K^{-}\pi^{+}$. The PDFs are also determined from simulated events, with the fractions in each $B^+ \to D_{s}^{*+}\phi$ fit category fixed. The result of the fit to $B^+ \to D_{s}^{*+}K^+K^-$ candidates, including all the relevant background components is shown in figure 3. The result of the simultaneous fit to $B^+ \to D_{s}^{*+}\phi$ candidates in the different helicity angle and $m(K^+K^-)$ mass categories is shown in figure 4. The three contributing $D_{s}^{*}$ meson decay modes are merged.

Figure 3. Mass distribution of $B^+ \to D_{s}^{*+}K^+K^-$ candidates. The result of the fit to the data using the model described in section 4.1 is overlaid, with the PDF components given in the legend.
Figure 4. Mass distribution of \( B^+ \to D_s^+ \phi \) candidates in (top) the \( \phi \) mass region, and (bottom) the \( \phi \) mass sideband. The plots on the left are in the helicity bin \( |\cos \theta_K| > 0.4 \) and the right are in \( |\cos \theta_K| < 0.4 \). The result of the fit to the data using the model described in section 4.1 is overlaid, with the PDF components given in the legend. The \( B^+ \to D_s^+ \phi \) decays (black) are expected to primarily contribute to the \( \phi \) region with \( |\cos \theta_K| > 0.4 \).

5 Systematic uncertainties

A number of different sources of systematic uncertainty are considered. The contribution from each source is detailed in table 3.

Relative efficiencies. The calculation of the branching fractions requires a correction to the ratio of signal and normalisation yields to account for the difference in the selection efficiency of the two modes. All relative selection efficiencies except the PID and MVA efficiencies are determined from simulated events and the effect of
having a limited simulation sample size is included as a systematic uncertainty. The relative efficiency for the PID and MVA requirements are determined from data control modes, including the samples of $B^0_s \rightarrow J/\psi \phi$ and $B^0_s \rightarrow D^*_s \pi^-$ decays used to test the MVA responses. Systematic uncertainties are assigned to account for the limited sizes of the control mode samples, kinematic differences between the control modes and the signal modes and differences between the data and simulation distributions that might affect the relative efficiency.

**Table 3.** Systematic uncertainties contributing to the measurements of $\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$ and $\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-)$. The systematic uncertainty from the normalisation branching fraction is also included.

<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>$\mathcal{B}(B^+ \rightarrow D_s^+ \phi)$ (x10^{-7})</th>
<th>$\mathcal{B}(B^+ \rightarrow D_s^+ K^+ K^-)$ (x10^{-6})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative efficiencies</td>
<td>0.08</td>
<td>0.59</td>
</tr>
<tr>
<td>Signal and normalisation PDFs</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Background PDFs</td>
<td>0.69</td>
<td>0.02</td>
</tr>
<tr>
<td>Charmless contribution</td>
<td>0.02</td>
<td>0.05</td>
</tr>
<tr>
<td>$B^+ \rightarrow D_s^+ K^+ K^-$ model</td>
<td>0.38</td>
<td>–</td>
</tr>
<tr>
<td>Normalisation</td>
<td>0.12</td>
<td>0.72</td>
</tr>
</tbody>
</table>

**Signal and normalisation PDFs.** Some parameters in the signal and normalisation PDFs are fixed to values obtained from simulation. These include the tail parameters, relative widths, and fractional amounts of the two CB functions that make up the PDFs. The values obtained from simulation have associated uncertainties arising from the limited simulation sample sizes. The nominal fits are repeated with the fixed parameters modified to values sampled from Gaussian distributions, with a width given by the parameter uncertainties. All parameters are changed simultaneously. For the fit to $B^+ \rightarrow D_s^+ \phi$ candidates, the fractions of events expected in each category of the fit are also included in the procedure. The resulting variation is assigned as the systematic uncertainty.

**Background PDFs.** Some of the PDFs for the background modes are taken directly from simulated events using one-dimensional kernel estimations [31]. In the nominal fit, these are smeared to account for the differences in the mass resolution between data and simulation. To account for any systematic uncertainty arising from the choice of resolution difference, the fit is repeated, randomly varying the smearing resolution each time. The resulting variation in the branching fraction is assigned as a systematic uncertainty. Additionally, each partially reconstructed background PDF has fixed fractions in the different categories of the signal fit. To determine the effect on the branching fraction, these fractions are repeatedly sampled from Gaussian distributions with widths given by the statistical uncertainty on the fractions. For
the combinatorial background shape, the choice of parametrisation is varied and the effect included in the systematic uncertainty.

**Charmless contribution.** Residual charmless and single-charm backgrounds are expected to remain in the final selection. These contributions are neglected in the calculation of the branching fractions. However, the shift in the branching fraction caused by numerically including the charmless yields is assigned as a systematic uncertainty.

**B**$^+ \to D_s^+ K^+ K^-$ **model assumption.** The fit to $B^+ \to D_s^+ \phi$ candidates includes a shape for $B^+ \to D_s^+ K^+ K^-$ decays that do not proceed via a $\phi$ meson. In order to distinguish this component from the signal, the different fractions of candidates in the four fit categories are exploited. This requires making assumptions as to which resonances contribute to the full $B^+ \to D_s^+ K^+ K^-$ decay model. The shape is assumed to be dominated by $f_0(980)$ and $a_0(980)$ resonances. Estimates of the uncertainties on the fractions are determined by considering the range in each fraction for the models considered. The variation in the branching fraction that results from varying these fractions within the uncertainties is assigned as the systematic uncertainty.

### 6 Results

#### 6.1 Search for $B^+ \to D_s^+ K^+ K^-$ candidates

The fit to $B^+ \to D_s^+ K^+ K^-$ candidates finds a total yield of $N(B^+ \to D_s^+ K^+ K^-) = 443 \pm 29$ candidates. This constitutes the first observation of this decay mode. The branching fraction is calculated as

$$B(B^+ \to D_s^+ K^+ K^-) = \frac{N_{\text{corr}}(B^+ \to D_s^+ K^+ K^-)}{N(B^+ \to D_s^0)} \times B(B^+ \to D_s^0) \times B(D^0 \to K^+ K^-)$$

where $N(B^+ \to D_s^0)$ is the yield of normalisation decays, and $N_{\text{corr}}(B^+ \to D_s^+ K^+ K^-)$ is defined to be

$$N_{\text{corr}}(B^+ \to D_s^+ K^+ K^-) = \sum_i \frac{W_i}{\epsilon_i^{\text{ratio}}}$$

where $W_i$ is the per-candidate weight, as determined by the $sPlot$ technique for candidate $i$; and $\epsilon_i^{\text{ratio}}$ represents the relative efficiency of the signal and normalisation modes $\epsilon_i(B^+ \to D_s^+ K^+ K^-)/\epsilon_i(B^+ \to D_s^0)$ in the relevant bin of the $B^+ \to D_s^+ K^+ K^-$ Dalitz plot. The corrected yield ratio can be expressed as the ratio of signal and normalisation branching fractions using eq. (6.1). The value is measured to be

$$\frac{N_{\text{corr}}(B^+ \to D_s^+ K^+ K^-)}{N(B^+ \to D_s^0)} = \frac{B(B^+ \to D_s^+ K^+ K^-)}{B(B^+ \to D_s^0)} \times \frac{B(D^0 \to K^+ K^-)}{B(D^0 \to K^+ K^-)} = 0.197 \pm 0.015 \pm 0.017,$$

where the first uncertainty is statistical, and the second is systematic. The branching fraction for $B^+ \to D_s^+ K^+ K^-$ decays is determined to be

$$B(B^+ \to D_s^+ K^+ K^-) = (7.1 \pm 0.5 \pm 0.6 \pm 0.7) \times 10^{-6},$$
where the first uncertainty is statistical, the second is systematic and the third from the branching fractions of $D^0 \to K^+ K^-$ and of the normalisation mode $B^+ \to D_s^+ D^0$. The values used for the branching fractions are $\mathcal{B}(D^0 \to K^+ K^-) = (4.01 \pm 0.07) \times 10^{-3}$ and $\mathcal{B}(B^+ \to D_s^+ D^0) = (9.0 \pm 0.9) \times 10^{-3}$ [11]. The two-body projections $m(D_s^+ K^-)$ and $m(K^+ K^-)$ are obtained for the signal component using the $s$Plot technique, shown in figure 5. No significant peak is observed in the $\phi$ region of the $m(K^+ K^-)$ plot; rather a broad distribution of candidates is found in the region up to $m(K^+ K^-) \simeq 1900$ MeV/$c^2$.

### 6.2 Search for $B^+ \to D_s^+ \phi$ candidates

The fit to $B^+ \to D_s^+ \phi$ candidates finds a total yield of $N(B^+ \to D_s^+ \phi) = 5.3 \pm 6.7$, summed across all categories and $D_s^+$ meson decay modes. A yield of $N(B^+ \to D_s^+ K^- K^+) = 65 \pm 10$ is found, consistent with the yield obtained from the full $B^+ \to D_s^+ K^- K^-$ measurement. The branching fraction for $B^+ \to D_s^+ \phi$ decays is calculated as

$$\mathcal{B}(B^+ \to D_s^+ \phi) = R \times \frac{\mathcal{B}(D^0 \to K^+ K^-)}{\mathcal{B}(\phi \to K^+ K^-)} \times \mathcal{B}(B^+ \to D_s^+ D^0),$$  

(6.3)

where the branching fraction $\mathcal{B}(\phi \to K^+ K^-) = 0.489 \pm 0.005$ has been used [11].

The free variable $R$ is defined to be the ratio of the signal and normalisation yields, corrected for the selection efficiencies. The yield of signal candidates in each $D_s^+$ mode is constructed from $R$ and the normalisation yield for the given $D_s^+$ decay mode, $N(B^+ \to D_s^+ D^0)$. The product of these two quantities is corrected by the ratio of selection efficiencies

$$N(B^+ \to D_s^+ \phi) = R \times N(B^+ \to D_s^+ D^0) \times \frac{\epsilon(B^+ \to D_s^+ \phi)}{\epsilon(B^+ \to D_s^+ D^0)},$$  

(6.4)

The simultaneous fit measures a single value of $R$ for all $D_s^+$ decay mode categories. From an ensemble of pseudoexperiments, $R$ is distributed normally. It can be written as the ratio of signal and normalisation branching fractions using eq. (6.3). The value is

---

Figure 5. Projections of the background-subtracted two-body invariant masses (left) $m(D_s^+ K^-)$ and (right) $m(K^+ K^-)$ for $B^+ \to D_s^+ K^- K^-$ decays. These plots are additionally weighted by a factor $1/\epsilon_{\text{ratio}}$ to correct for the efficiency variation across the phase space. An expansion of the $\phi$ region of $m(K^+ K^-)$ is inset where the same $\phi$ signal region and $\phi$ sideband region have been represented as in figure 2.
determined to be
\[
R = \frac{\mathcal{B}(B^+ \to D_s^+ \phi)}{\mathcal{B}(B^+ \to D_s^0 \bar{D}_s^0)} \times \frac{\mathcal{B}(\phi \to K^+ K^-)}{\mathcal{B}(D_s^0 \to K^+ K^-)} = (1.6^{+2.2}_{-1.9} \pm 1.1) \times 10^{-3},
\]
where the first uncertainty is statistical and the second systematic. This corresponds to a branching fraction for \(B^+ \to D_s^+ \phi\) decays of
\[
\mathcal{B}(B^+ \to D_s^+ \phi) = (1.2^{+1.6}_{-1.4} \pm 0.8 \pm 0.1) \times 10^{-7},
\]
where the first uncertainty is statistical, the second systematic, and the third results from the uncertainty on the branching fractions \(\mathcal{B}(B^+ \to D_s^+ \bar{D}_s^0)\), \(\mathcal{B}(\phi \to K^+ K^-)\) and \(\mathcal{B} (D_s^0 \to K^+ K^-)\). Considering only the statistical uncertainty, the significance of the \(B^+ \to D_s^+ \phi\) signal is 0.8 standard deviations (\(\sigma\)).

Upper limits at 95% and 90% confidence levels (CL) are determined using the Feldman-Cousins approach [32]. An ensemble of pseudoexperiments is generated for different values of the branching fraction \(\mathcal{B}(B^+ \to D_s^+ \phi)\). These generated pseudoexperiments are then fitted with the nominal fit model to calculate the fitted branching fraction and associated statistical uncertainty, \(\sigma_{\text{stat}}\). This method constructs confidence bands based on a likelihood ratio method, calculating the probability of fitting a branching fraction for a given generated branching fraction. This probability is assumed to follow a Gaussian distribution with width \(\sigma = \sqrt{\sigma_{\text{stat}}^2 + \sigma_{\text{syst}}^2}\), where \(\sigma_{\text{stat}}\) and \(\sigma_{\text{syst}}\) are the statistical and systematic uncertainties. The dominant source of systematic uncertainty in this measurement is from the background PDFs. As the size of this uncertainty is not expected to vary as a function of the generated branching fraction, \(\sigma_{\text{syst}}\) is assumed to be constant. Nuisance parameters are accounted for using the plug-in method [33]. The generated confidence bands are shown in figure 6, where the statistical-only 90% CL and 95% CL bands are shown, along with the 95% CL band with systematic uncertainty included. This corresponds to a statistical-only 95% (90%) CL limit of \(\mathcal{B}(B^+ \to D_s^+ \phi) < 4.4 \times 10^{-7} (3.9 \times 10^{-7})\), and a 95% (90%) CL limit including systematic uncertainties of
\[
\mathcal{B}(B^+ \to D_s^+ \phi) < 4.9 \times 10^{-7} (4.2 \times 10^{-7}).
\]

7 Conclusions

A search for \(B^+ \to D_s^+ K^+ K^-\) decays is performed. The branching fraction is determined to be
\[
\mathcal{B}(B^+ \to D_s^+ K^+ K^-) = (7.1 \pm 0.5 \pm 0.6 \pm 0.7) \times 10^{-6},
\]
where the first uncertainty is statistical, the second systematic and the third is due to the uncertainty on the branching fraction of the normalisation mode \(B^+ \to D_s^+ \bar{D}_s^0\). This is the first observation of this decay mode. A search is also performed for the pure annihilation decay \(B^+ \to D_s^+ \phi\), but no significant signal is observed and a limit of
\[
\mathcal{B}(B^+ \to D_s^+ \phi) < 4.9 \times 10^{-7} (4.2 \times 10^{-7})
\]
is set on the branching fraction at 95% (90%) confidence level. The limit on $B(B^+ \rightarrow D_s^+ \phi)$ presented here supersedes the previous result from LHCb [10].

This updated analysis benefits from the significantly larger data sample now available at LHCb to increase the reach of these searches. The previous measurement performed by LHCb reported evidence for the decay $B^+ \rightarrow D_s^+ \phi$ with a significance greater than 3$\sigma$. This update determines that there is a sizeable contribution from $B^+ \rightarrow D_s^+ K^+ K^-$ decays that contribute within the $\phi$-meson mass window that was previously not considered. The result is consistent with the prediction that rescattering contributions to $B^+ \rightarrow D_s^+ \phi$ decays are small.

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 Skłodowska-Curie Actions and ERC (European Union), ANR, Labex P2IO, ENIGMASS and OCEVU, and Région Auvergne-Rhône-Alpes (France), RFBR and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), Herchel Smith Fund, the Royal Society, the English-Speaking Union and the Leverhulme Trust (United Kingdom).

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

[1] BABAR collaboration, B. Aubert et al., Evidence for the Rare Decay $B^+ \to D_s^+ \pi^0$, Phys. Rev. Lett. 98 (2007) 171801 [hep-ex/0611030] [INSPIRE].


[8] ETM collaboration, A. Bussone et al., Mass of the $b$ quark and $B$-meson decay constants from $N_f = 2 + 1 + 1$ twisted-mass lattice QCD, Phys. Rev. D 93 (2016) 114505 [arXiv:1603.04306] [INSPIRE].


[10] LHCb collaboration, First evidence for the annihilation decay mode $B^+ \to D^+_s \phi$, JHEP 02 (2013) 043 [arXiv:1210.1089] [INSPIRE].


[12] LHCb collaboration, The LHCb Detector at the LHC, 2008 JINST 3 S08005 [INSPIRE].


[21] LHCb collaboration, *First observations of $B_0^s \to D^+ D^-$, $D_s^+ D^-$ and $D^0 \bar{D}^0$ decays*, *Phys. Rev. D* **87** (2013) 092007 [arXiv:1302.5854] [INSPIRE].


[27] LHCb collaboration, *Measurement of CP observables in $B^\pm \to D^{(*)} K^\pm$ and $B^\pm \to D^{(*)} \pi^\pm$ decays*, *Phys. Lett. B* **777** (2018) 16 [arXiv:1708.06370] [INSPIRE].

[28] LHCb collaboration, *Model-independent measurement of the CKM angle $\gamma$ using $B^0 \to D K^{*0}$ decays with $D \to K_S^0 \pi^+ \pi^-$ and $K_S^0 K^+ K^-$*, *JHEP* **06** (2016) 131 [arXiv:1604.01525] [INSPIRE].


10 Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany
11 Max-Planck-Institut für Kernphysik (MPIK), Heidelberg, Germany
12 Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
13 School of Physics, University College Dublin, Dublin, Ireland
14 Sezione INFN di Bari, Bari, Italy
15 Sezione INFN di Bologna, Bologna, Italy
16 Sezione INFN di Cagliari, Cagliari, Italy
17 Università e INFN, Ferrara, Ferrara, Italy
18 Sezione INFN di Firenze, Firenze, Italy
19 Laboratori Nazionali dell’INFN di Frascati, Frascati, Italy
20 Sezione INFN di Genova, Genova, Italy
21 Università e INFN, Milano-Bicocca, Milano, Italy
22 Sezione di Milano, Milano, Italy
23 Sezione INFN di Padova, Padova, Italy
24 Sezione INFN di Pisa, Pisa, Italy
25 Sezione INFN di Roma Tor Vergata, Roma, Italy
26 Sezione INFN di Roma La Sapienza, Roma, Italy
27 Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
28 AGH - University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland
29 National Center for Nuclear Research (NCBJ), Warsaw, Poland
30 Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
31 Petersburg Nuclear Physics Institute (PNPI), Gatchina, Russia
32 Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia
33 Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia
34 Institute for Nuclear Research of the Russian Academy of Sciences (INR RAN), Moscow, Russia
35 Yandex School of Data Analysis, Moscow, Russia
36 Budker Institute of Nuclear Physics (SB RAS), Novosibirsk, Russia
37 Institute for High Energy Physics (IHEP), Protvino, Russia
38 ICCUB, Universitat de Barcelona, Barcelona, Spain
39 Instituto Galego de Física de Altas Enerxías (IGFAE), Universidade de Santiago de Compostela, Santiago de Compostela, Spain
40 European Organization for Nuclear Research (CERN), Geneva, Switzerland
41 Institute of Physics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
42 Physik-Institut, Universität Zürich, Zürich, Switzerland
43 Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands
44 Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, The Netherlands
45 NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine
46 Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine
47 University of Birmingham, Birmingham, United Kingdom
48 H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
49 Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
50 Department of Physics, University of Warwick, Coventry, United Kingdom
51 STFC Rutherford Appleton Laboratory, Didcot, United Kingdom
52 School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
53 School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
54 Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
55 Imperial College London, London, United Kingdom
56 School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
57 Department of Physics, University of Oxford, Oxford, United Kingdom
Massachusetts Institute of Technology, Cambridge, MA, United States
University of Cincinnati, Cincinnati, OH, United States
University of Maryland, College Park, MD, United States
Syracuse University, Syracuse, NY, United States
Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, associated to
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 Física, Universidad Nacional de Colombia, Bogotá, Colombia, associated to
Institut für Physik, Universität Rostock, Rostock, Germany, associated to
National Research Centre Kurchatov Institute, Moscow, Russia, associated to
National Research Tomsk Polytechnic University, Tomsk, Russia, associated to
Instituto de Física Corpuscular, Centro Mixto Universidad de Valencia - CSIC, Valencia, Spain, associated to
Van Swinderen Institute, University of Groningen, Groningen, The Netherlands, associated to

a Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil
b Laboratoire Leprince-Ringuet, Palaiseau, France
c P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia
d Università di Bari, Bari, Italy
e Università di Bologna, Bologna, Italy
f Università di Cagliari, Cagliari, Italy
g Università di Ferrara, Ferrara, Italy
h Università di Genova, Genova, Italy
i Università di Milano Bicocca, Milano, Italy
j Università di Roma Tor Vergata, Roma, Italy
k Università di Roma La Sapienza, Roma, Italy
l AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland
m LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain
n Hanoi University of Science, Hanoi, Viet Nam
o Università di Padova, Padova, Italy
p Università di Pisa, Pisa, Italy
q Università degli Studi di Milano, Milano, Italy
r Università di Urbino, Urbino, Italy
s Università della Basilicata, Potenza, Italy
t Scuola Normale Superiore, Pisa, Italy
u Università di Modena e Reggio Emilia, Modena, Italy
v Iligan Institute of Technology (IIT), Iligan, Philippines
w Novosibirsk State University, Novosibirsk, Russia
x Deceased