First observations of the rare decays $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ and $B^+ \rightarrow \phi K^+ \mu^+ \mu^-$

LHCb Collaboration; Bernet, R; Müller, K; Steinkamp, O; Straumann, U; Vollhardt, A; et al

Abstract: First observations of the rare decays $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ and $B^+ \rightarrow \phi K^+ \mu^+ \mu^-$ are presented using data corresponding to an integrated luminosity of 3.0 fb$^{-1}$, collected by the LHCb experiment at centre-of-mass energies of 7 and 8 TeV. The branching fractions of the decays are

$$B(B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-) = \text{amp}; = \text{amp}; (4.36 \pm 0.29 \text{ (stat)} \pm 0.21 \text{ (syst)} \pm 0.18 \text{ (norm)}) \times 10^{-7},$$

$$B(B^+ \rightarrow \phi K^+ \mu^+ \mu^-) = \text{amp}; = \text{amp}; (0.82 \pm 0.19 \text{ (stat)} \pm 0.17 \text{ (syst)} \pm 0.04 \text{ (norm)}) \times 10^{-7},$$

where the uncertainties are statistical, systematic, and due to the uncertainty on the branching fractions of the normalisation modes. A measurement of the differential branching fraction in bins of the invariant mass squared of the dimuon system is also presented for the decay $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$.

DOI: [https://doi.org/10.1007/JHEP10(2014)064](https://doi.org/10.1007/JHEP10(2014)064)
First observations of the rare decays

$B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ and $B^+ \rightarrow \phi K^+ \mu^+ \mu^-$

The LHCb collaboration

E-mail: shall@cern.ch

ABSTRACT: First observations of the rare decays $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ and $B^+ \rightarrow \phi K^+ \mu^+ \mu^-$ are presented using data corresponding to an integrated luminosity of $3.0 \text{fb}^{-1}$, collected by the LHCb experiment at centre-of-mass energies of 7 and 8 TeV. The branching fractions of the decays are

$$B(B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-) = (4.36^{+0.29}_{-0.27} \text{ (stat)} \pm 0.21 \text{ (syst)} \pm 0.18 \text{ (norm)}) \times 10^{-7},$$

$$B(B^+ \rightarrow \phi K^+ \mu^+ \mu^-) = (0.82^{+0.19}_{-0.17} \text{ (stat)}^{+0.10}_{-0.04} \text{ (syst)} \pm 0.27 \text{ (norm)}) \times 10^{-7},$$

where the uncertainties are statistical, systematic, and due to the uncertainty on the branching fractions of the normalisation modes. A measurement of the differential branching fraction in bins of the invariant mass squared of the dimuon system is also presented for the decay $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$.

KEYWORDS: Rare decay, Hadron-Hadron Scattering, B physics, Flavor physics

ArXiv ePrint: 1408.1137
1 Introduction

The $B^+ \rightarrow K^+ \pi^- \mu^+ \mu^-$ and $B^+ \rightarrow \phi K^+ \mu^+ \mu^-$ decays proceed via $b \rightarrow s$ flavour changing neutral currents (FCNC).\textsuperscript{1} In the Standard Model (SM), FCNC decays are forbidden at the tree level and are only allowed as higher-order electroweak loop processes. In extensions of the SM, new particles can significantly change the branching fractions and angular distributions of the observed final-state particles. Due to their sensitivity to effects beyond the SM, semileptonic $B$ decays involving FCNC transitions are currently under intense study at the LHCb experiment [1–4].

The $K^+ \pi^+ \pi^-$ system in the final state of the $B^+ \rightarrow K^+ \pi^- \mu^+ \mu^-$ decay can result from the decay of several strange resonances. Its composition was studied by the Belle collaboration for the tree-level decay $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+ \pi^+ \pi^-$ [5], where the $K_1(1270)^+$ meson was found to have a prominent contribution. The $K_1(1270)^+$ and the $K_1(1400)^+$ mesons are the mass eigenstates that result from mixing of the $P$-wave axial vector mesons $^3P_1$ ($K_{1A}$) and $^1P_1$ ($K_{1B}$) with the mixing angle $\theta_{K_1}$ [6]. The value of $\theta_{K_1}$ is either about $-33^\circ$ or $-57^\circ$ [6–11] with most recent determinations favouring the former [8–11]. The decay $B^+ \rightarrow J/\psi \phi K^+$ was first observed by the CLEO collaboration [12] and recently investigated in the search for the $X(4140)$ [13–16].

The branching fraction of the rare decay $B^+ \rightarrow K_1(1270)^+ \mu^+ \mu^-$, which is expected to contribute significantly to the $K^+ \pi^+ \pi^- \mu^+ \mu^-$ final-state, is predicted to be $\mathcal{B}(B^+ \rightarrow K_1(1270)^+ \mu^+ \mu^-) = (2.3^{+1.3}_{-1.0} + 0.0) \times 10^{-6}$ [17]. Here, the first uncertainty originates from the form-factor calculations, while the second is from the uncertainty on the

\textsuperscript{1}Charge conjugation is implied throughout this paper.
mixing angle $\theta_{K_1}$. However, due to the unknown resonance structure of the final-state hadrons, there are no inclusive theoretical predictions available for the branching fractions of the decays $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ and $B^+ \rightarrow \phi K^+ \mu^+ \mu^-$. 

This paper presents the first observations of the decays $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ and $B^+ \rightarrow \phi(1020)K^+ \mu^+ \mu^-$, using a data sample collected by the LHCb experiment, corresponding to an integrated luminosity of 3.0 fb$^{-1}$. The data were recorded in the years 2011 and 2012 at centre-of-mass energies of 7 and 8 TeV, respectively. In addition, a measurement of the differential branching fraction $d\mathcal{B}(B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-)/dq^2$, where $q^2$ is the invariant mass squared of the dimuon system, is presented.

2 The LHCb detector

The LHCb detector [18] 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. The tracking system provides a measurement of momentum, $p$, with a 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 $pp$ interaction vertex (PV), the impact parameter (IP), is measured with a resolution of $(15 + 29/p_T) \mu$m, where $p_T$ is the the component of $p$ transverse to the beam, in GeV/$c$. Charged hadrons are identified using two ring-imaging Cherenkov detectors (RICH) [19]. 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 [20]. The trigger [21] 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.

Simulated events are used to determine trigger, reconstruction and selection efficiencies. In addition, simulated samples are used to estimate possible backgrounds from $B$ meson decays that can mimic the final states of the signal decays. Simulated events are generated using PYTHIA [22, 23] with a specific LHCb configuration [24]. Decays of hadronic particles are described by EVTGEN [25], in which final-state radiation is generated using PHOTOS [26]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [27, 28] as described in ref. [29].

3 Selection of signal candidates

The $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ and $B^+ \rightarrow \phi K^+ \mu^+ \mu^-$ signal candidates are first required to pass the hardware trigger stage, which selects muons with $p_T > 1.76$ GeV/$c$. In the subsequent software trigger stage, at least one of the final-state hadrons (muons) is required to have both $p_T > 1.6$ GeV/$c$ (1.0 GeV/$c$) and IP larger than 100 $\mu$m with respect to any PV in
the event. A multivariate algorithm [30] is used to identify secondary vertices that are consistent with the decay of a $b$ hadron with muons in the final state.

Signal candidates are formed by combining two muons of opposite charge with three charged hadrons. Reconstructed signal candidate tracks must have significant displacement from any PV in the event. The signal candidate tracks are required to form a secondary vertex of good fit quality which is significantly displaced from the PV. Particle identification information from the RICH detectors (PID) is used to identify the final-state hadrons. For $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ decays, the invariant mass of the $K^+ \pi^+ \pi^-$ system is required to be below $2400 \text{MeV}/c^2$. For $B^+ \rightarrow \phi K^+ \mu^+ \mu^-$ decays with $\phi \rightarrow K^+ K^-$, the invariant mass of the $K^+ K^-$ system is required to be within $12 \text{MeV}/c^2$ of the known $\phi$ meson mass [31]. This mass region contains almost entirely $\phi \rightarrow K^+ K^-$ meson decays with negligible background.

The final states of the signal decays can be mimicked by other $B$ decays, which represent potential sources of background. Resonant decays, where the muon pair originates from either $J/\psi$ or $\psi(2S)$ meson decays, are removed by rejecting events where the invariant mass of the dimuon system is in the veto regions $2946 < m(\mu^+ \mu^-) < 3176 \text{MeV}/c^2$ or $3586 < m(\mu^+ \mu^-) < 3766 \text{MeV}/c^2$. The radiative tails of the $J/\psi (\psi(2S))$ decays are suppressed by extending the lower edge of these veto regions down by $250 \text{MeV}/c^2$ ($100 \text{MeV}/c^2$) if the reconstructed $B^+$ mass is smaller than 5230 MeV/$c^2$. In the mass region $5330 < m(B^+) < 5450 \text{MeV}/c^2$ the upper edge of the vetoes is extended up by $40 \text{MeV}/c^2$ to reject a small fraction of misreconstructed $J/\psi$ and $\psi(2S)$ meson decays. The resonant decays can also be misreconstructed as signal if a muon from the charmonium decay is misidentified as a hadron and vice versa. To remove this potential background the invariant mass of the $\mu^+ \mu^-$ or $\mu^+ K^-$ system is calculated assigning the muon mass to the hadron. If the mass falls within $50 \text{MeV}/c^2$ of the known $J/\psi$ or $\psi(2S)$ masses [31], the candidate is rejected.

Potential background from the electroweak-penguin decay $B^0 \rightarrow K^{\ast 0} \mu^+ \mu^-$, where the $K^{\ast 0} \rightarrow K^+ \pi^-$ decay is combined with a random $\pi^+$ meson, is studied and found to be negligible. Backgrounds from semileptonic $b \rightarrow c (\rightarrow s \mu^+ \nu_{\mu}) \mu^- \bar{\nu}_{\mu}$ cascade decays, as well as fully hadronic $B$ decays such as $B^+ \rightarrow \bar{D}^0 (\rightarrow K^+ \pi^+ \pi^- \pi^-) \pi^+$ where two hadrons are misidentified as muons, are also negligible.

Combinatorial background is suppressed with a boosted decision tree (BDT) [32, 33]. The BDT training uses $s$Weighted [34] candidates from the control channel $B^+ \rightarrow J/\psi K^+ \pi^+ \pi^-$ as a signal proxy and the high $B^+$ mass sideband ($5529 < m(K^+ \pi^+ \pi^- \mu^+ \mu^-) < 5780 \text{MeV}/c^2$) of $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ candidates as a background proxy. The BDT uses geometric and kinematic variables in the training, including the $p_T$ of the final state tracks and their displacement from the PV. Additionally, the $p_T$ of the reconstructed $B^+$ candidate, as well as information on the quality of the decay vertex and its displacement are used. Requirements on the BDT response and the PID criteria, which discriminate between kaons and pions for the reconstructed final-state hadrons, are optimised simultaneously using the metric $S/\sqrt{S+B}$. Here, $S$ and $B$ denote the expected signal and background yields. The value of $S$ is calculated using an estimate for the branching fraction of the decay $B^+ \rightarrow K_1(1270)^{+} \mu^+ \mu^-$. This branching fraction is determined by scaling that of the rare decay $B^0 \rightarrow K^{\ast 0} \mu^+ \mu^-$ [1] by the branching fraction ratio of the radiative decays $B^+ \rightarrow K_1(1270)^{+} \gamma$ and $B^0 \rightarrow K^{\ast 0} \gamma$ [31].
To determine the branching fractions of the signal decays, the normalisation modes $B^+ \rightarrow \psi(2S)K^+$, with the subsequent decay $\psi(2S) \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\pi^+\pi^-$, and $B^+ \rightarrow J/\psi \phi K^+$ are used. The branching fraction of the decay $B^+ \rightarrow \psi(2S)K^+$ is $(6.27 \pm 0.24) \times 10^{-4}$ [31], and the branching fraction of the decay $B^+ \rightarrow J/\psi \phi K^+$ is $(5.2 \pm 1.7) \times 10^{-5}$ [31]. The final states of the normalisation modes are identical to those of the signal decays, which is beneficial since many systematic effects are expected to cancel. Both normalisation modes are selected in analogy to the signal decays except for additional mass requirements. For the $\psi(2S)$ decay, the reconstructed $\pi^+\pi^-\mu^+\mu^-$ mass is required to be within 60 MeV/$c^2$ of the known $\psi(2S)$ mass. The reconstructed invariant mass of the dimuon system originating from the $J/\psi$ meson decay is required to be within 50 MeV/$c^2$ of the known $J/\psi$ mass.

4 Differential branching fraction of the decay $B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-$

The determination of the differential branching fraction $d\mathcal{B}(B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-)/dq^2$ is performed in bins of $q^2$, as given in table 1. Figure 1 shows the invariant mass distribution of $B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-$ candidates in each $q^2$ bin studied. Signal yields are determined using extended maximum likelihood fits to the unbinned $K^+\pi^+\pi^-\mu^+\mu^-$ mass spectra. The $m(K^+\pi^+\pi^-\mu^+\mu^-)$ distribution of the signal component is modelled using the sum of two Gaussian functions, each with a power-law tail on the low-mass side. The background component is modelled with an exponential function, where the reductions in efficiency due to the vetoes of the radiative tails of the charmonium decays are accounted for by using scale factors. The signal yield integrated over the full $q^2$ range is $N_{K\pi\pi\mu\mu} = 367^{+24}_{-23}$. The statistical significance of the signal is in excess of 20 standard deviations, according to Wilks’ theorem [35]. Figure 2a shows the fit to the mass distribution of the control channel $B^+ \rightarrow J/\psi K^+\pi^+\pi^-$ that is used to determine the parameters describing the mass distribution of the $B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-$ signal decay. To account for partially reconstructed decays at low masses, a Gaussian function is used in addition to the exponential to describe the background component. The yield of the control channel is $59335 \pm 343$. Figure 2b shows the fit for the normalisation channel $B^+ \rightarrow \psi(2S)K^+$. To describe the mass shape, the same components are used as for the fit of the control decay and all mass shape parameters are allowed to vary in the fit. The yield of the normalisation channel is $5128 \pm 67$.

The differential branching fraction $d\mathcal{B}(B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-)/dq^2$ in a $q^2$ bin of width $\Delta q^2$ is

$$\frac{d\mathcal{B}(B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-)}{dq^2} = \frac{1}{\Delta q^2} \cdot \frac{N_{\text{sig}}}{N_{\text{norm}}} \cdot \frac{\epsilon_{\text{norm}}}{\epsilon_{\text{sig}}} \cdot \mathcal{B}(B^+ \rightarrow \psi(2S)K^+) \cdot \mathcal{B}(\psi(2S) \rightarrow J/\psi(\rightarrow \mu^+\mu^-)\pi^+\pi^-). \hspace{1cm} (4.1)$$

Here, $N_{\text{sig}}$ is the yield of the signal channel in the given $q^2$ bin and $N_{\text{norm}}$ the yield of the normalisation channel. The efficiencies for the reconstruction and selection of the signal and normalisation channels are denoted by $\epsilon_{\text{sig}}$ and $\epsilon_{\text{norm}}$, respectively. The efficiency for the signal decay is determined using simulated $B^+ \rightarrow K_1(1270)^+\mu^+\mu^-$ events generated according to ref. [17]; a separate efficiency ratio is calculated for each $q^2$ bin.
Figure 1. Invariant mass of $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ candidates in bins of $q^2$ with fit projections overlaid. The signal component (shaded light blue) is modelled by the sum of two Gaussian functions, each with a power-law tail at low mass. The background component (shaded dark blue) is modelled by an exponential function. In the $q^2$ ranges $4.30 < q^2 < 8.68$ GeV$^2$/c$^4$, $10.09 < q^2 < 12.86$ GeV$^2$/c$^4$, and $14.18 < q^2 < 19.00$ GeV$^2$/c$^4$, scaling factors are applied to account for the vetoes of the radiative tails of the charmonium resonances, resulting in steps in the background mass shape. The lower right plot shows a separate fit to the signal decay integrated over all $q^2$ bins.

The branching fraction for the $\psi(2S)$ meson to decay to the final state $\pi^+ \pi^- \mu^+ \mu^-$ is $\mathcal{B}(\psi(2S) \rightarrow J/\psi (\rightarrow \mu^+ \mu^-)\pi^+ \pi^-) = (2.016 \pm 0.031) \times 10^{-2}$ [31].

The resulting differential branching fractions for the decay $B^+ \rightarrow K^+ \pi^+ \pi^- \mu^+ \mu^-$ are shown in figure 3 with numerical values given in table 1. Summation over all $q^2$ bins yields an integrated branching fraction of $(3.43^{+0.23}_{-0.21} \text{ (stat)} \pm 0.15 \text{ (syst)} \pm 0.14 \text{ (norm)}) \times 10^{-7}$, where the uncertainties are statistical, systematic, and due to the uncertainty on the normalisation channel. The fraction of signal events removed by the vetoes of the charmonium regions is determined from simulated $B^+ \rightarrow K_1(1270)^+ \mu^+ \mu^-$ events to be...
(21.3 ± 1.5)%). The uncertainty on this number is determined from a variation of the angle $\theta_{K_1}$ and the form-factor parameters within their uncertainties. Correcting for the charmonium vetoes yields a total branching fraction of

$$\mathcal{B}(B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-) = (4.36^{+0.39}_{-0.27} \text{(stat)} \pm 0.21 \text{(syst)} \pm 0.18 \text{(norm)}) \times 10^{-7}.$$ 

Since the systematic uncertainty due to the normalisation channel is significant, we also report the branching ratio of the signal channel with respect to its normalisation mode, which is determined to be

$$\frac{\mathcal{B}(B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-)}{\mathcal{B}(B^+ \rightarrow \psi(2S)K^+)} = (6.95^{+0.46}_{-0.43} \text{(stat)} \pm 0.34 \text{(syst)}) \times 10^{-4}.$$ 

Due to the low signal yield, no attempt is made to resolve the different contributions to the $K^+\pi^+\pi^-$ system in the $K^+\pi^+\pi^-\mu^+\mu^-$ final state. However, it is possible to obtain the $m(K^+\pi^+\pi^-)$ distribution using the sPlot [34] technique. Figure 4 shows this distribution for the signal decay in the full $q^2$ region, as well as for the control decay $B^+ \rightarrow J/\psi K^+\pi^+\pi^-$. 

---

Table 1. Signal yields for the decay $B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-$ and resulting differential branching fractions in bins of $q^2$. The first contribution to the uncertainty is statistical, the second systematic, where the uncertainty due to the branching fraction of the normalisation channel is included. The $q^2$ binning used corresponds to the binning used in previous analyses of $b \rightarrow s\mu^+\mu^-$ decays [1–3]. Results are also presented for the $q^2$ range from 1 to 6 GeV$^2$/c$^4$, where theory predictions are expected to be most reliable.

<table>
<thead>
<tr>
<th>$q^2$ bin [GeV$^2$/c$^4$]</th>
<th>$N_{\text{sig}}$</th>
<th>$\frac{d\mathcal{B}}{dq^2} \times 10^{-8}$ [GeV$^{-2}$c$^4$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ 0.10, 2.00]</td>
<td>134.1 $^{+12.9}_{-12.3}$</td>
<td>7.01 $^{+0.69}_{-0.65}$ ± 0.47</td>
</tr>
<tr>
<td>[ 2.00, 4.30]</td>
<td>56.5 $^{+9.7}_{-9.1}$</td>
<td>2.34 $^{+0.41}_{-0.38}$ ± 0.15</td>
</tr>
<tr>
<td>[ 4.30, 8.68]</td>
<td>119.9 $^{+14.6}_{-13.7}$</td>
<td>2.30 $^{+0.28}_{-0.26}$ ± 0.20</td>
</tr>
<tr>
<td>[10.09, 12.86]</td>
<td>54.0 $^{+10.1}_{-9.4}$</td>
<td>1.83 $^{+0.34}_{-0.32}$ ± 0.17</td>
</tr>
<tr>
<td>[14.18, 19.00]</td>
<td>3.3 $^{+2.8}_{-2.1}$</td>
<td>0.10 $^{+0.08}_{-0.06}$ ± 0.01</td>
</tr>
<tr>
<td>[ 1.00, 6.00]</td>
<td>144.8 $^{+14.9}_{-14.3}$</td>
<td>2.75 $^{+0.29}_{-0.28}$ ± 0.16</td>
</tr>
</tbody>
</table>
Figure 3. Differential branching fraction \( d\mathcal{B}(B^+ \to K^+\pi^+\pi^-\mu^+\mu^-)/dq^2 \). Errors shown include both statistical and systematic uncertainties. Shaded regions indicate the vetoed charmonium resonances.

Figure 4. Background-subtracted \( m(K^+\pi^+\pi^-) \) distributions for (a) the signal decay \( B^+ \to K^+\pi^+\pi^-\mu^+\mu^- \) and (b) the control channel \( B^+ \to J/\psi K^+\pi^+\pi^- \). The vertical lines indicate the masses of the \( K_1(1270)^+ \) and \( K_1(1400)^+ \) resonances.

For the signal decay \( B^+ \to K^+\pi^+\pi^-\mu^+\mu^- \) the data are consistent with the presence of several broad and overlapping resonances.

### 4.1 Systematic uncertainties

The dominant systematic uncertainty comes from the branching fraction of the normalisation mode \( B^+ \to \psi(2S)K^+ \), which is known to a precision of 6%. This uncertainty is fully correlated between the \( q^2 \) bins and is quoted separately.

The systematic uncertainty introduced by the choice of signal mass model is estimated by re-evaluating the signal yield using a single Gaussian function with a power-law tail. To estimate the uncertainty of the background mass model, a linear mass shape is used instead of the nominal exponential function. The total systematic uncertainty assigned due to the modelling of the mass distribution is approximately 2%.

The majority of systematic effects bias the efficiency ratio \( \epsilon_{\text{norm}}/\epsilon_{\text{sig}} \), which is determined using simulation. To account for differences between data and simulation, correc-
tions based on data are applied to simulated events. The efficiency to identify kaons is corrected by using large $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ control samples. Muon identification performance and tracking efficiency are corrected using $J/\psi \rightarrow \mu^+\mu^-$ decays. In addition, track multiplicity and vertex fit quality are weighted according to the control channel $B^+ \rightarrow J/\psi K^+\pi^+\pi^-$. The systematic uncertainties associated with these corrections are evaluated by determining the branching fraction without the correction and taking the full observed deviation as a systematic uncertainty. In total, they constitute a systematic uncertainty of around 1%. The software trigger is observed to be well described in simulation, but slight discrepancies are observed for the hardware stage. These are corrected by weighting the simulated samples according to the maximum muon $p_T$. The branching fraction is recalculated without these weights, and the observed difference of 1% is assigned as the systematic uncertainty from the trigger simulation.

Additional systematic uncertainties stem from the fact that simulated $B^+ \rightarrow K_1(1270)^+\mu^+\mu^-$ events, modelled according to ref. [17], are used to determine the efficiency ratio $\epsilon_{\text{norm}}/\epsilon_{\text{sig}}$. To account for contributions other than the $K_1(1270)^+$ to the $K^+\pi^+\pi^-$ system, events are weighted according to the $m(K^+\pi^+\pi^-)$ distribution shown in figure 4. This results in a systematic uncertainty of 1–2%, depending on the $q^2$ range considered. The effect of a potentially different $q^2$ distribution of the signal decay is evaluated by defining the efficiency ratio using $B^+ \rightarrow K_1(1270)^+\mu^+\mu^-$ events generated according to a phase-space model. The observed deviation results in a systematic uncertainty of 1–2%.

5 Branching fraction of the decay $B^+ \rightarrow \phi K^+\mu^+\mu^-$

The signal decay $B^+ \rightarrow \phi K^+\mu^+\mu^-$ is expected to be rarer than the decay $B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-$ as an $s\bar{s}$ quark pair must be created from the vacuum. Therefore, only the total branching fraction of this decay mode is determined. Figure 5a shows the $B^+ \rightarrow \phi K^+\mu^+\mu^-$ signal candidates after the full selection. The signal yield is determined to be $N_{\text{sig}} = 25.2^{+8.0}_{-5.3}$ using an extended maximum likelihood fit to the unbinned $\phi K^+\mu^+\mu^-$ mass distribution. The statistical significance of the signal, calculated using Wilks’ theorem, is 6.6$\sigma$. The signal component is modelled using the sum of two Gaussian functions with a tail described by a power law on the low-mass side. The background mass shape is modelled using a second-order Chebychev polynomial. The parameters describing the signal mass shape are fixed to those determined using the normalisation mode $B^+ \rightarrow J/\psi \phi K^+$, as shown in figure 5b. The yield of the normalisation mode is $N_{\text{norm}} = 1908 \pm 63$.

To determine the total branching fraction of the decay $B^+ \rightarrow \phi K^+\mu^+\mu^-$, the formula

$$B(B^+ \rightarrow \phi K^+\mu^+\mu^-) = \frac{N'_{\text{sig}}}{N_{\text{norm}}} \cdot B(B^+ \rightarrow J/\psi \phi K^+) \cdot B(J/\psi \rightarrow \mu^+\mu^-)$$

is used. Here, $N'_{\text{sig}}$ denotes the signal yield determined in a fit where signal candidates are weighted by the relative efficiency $\epsilon_{\text{norm}}/\epsilon_{\text{sig}}(q^2)$, according to their $q^2$ value. This is necessary since the efficiency ratio varies significantly over the full $q^2$ range. The weights are determined in bins of $q^2$, with the same choice of $q^2$ bins as in table 1. Using the branching fraction of the normalisation channel, the integrated branching fraction is determined to be
Figure 5. Invariant $m(\phi K^+\mu^+\mu^-)$ distributions for (a) $B^+ \to \phi K^+\mu^+\mu^-$ and (b) $B^+ \to J/\psi \phi K^+$ decays with fit projections overlaid.

$$(0.81^{+0.18}_{-0.16} \text{ (stat)} \pm 0.03 \text{ (syst)} \pm 0.27 \text{ (norm)}) \times 10^{-7}$$. The fraction of signal events rejected by the charmonium vetoes is $(2^{+10}_{-2})\%$. This is calculated using simulated $B^+ \to \phi K^+\mu^+\mu^-$ events generated according to a phase-space model. The uncertainty is estimated by comparison with the model given in ref. [17] for the decay $B^+ \to K_1(1270)^+\mu^+\mu^-$ and weighting to correct for the large mass of the $\phi K^+$ system. Accounting for the charmonium vetoes results in a total branching fraction of

$$B(B^+ \to \phi K^+\mu^+\mu^-) = (0.82^{+0.19}_{-0.17} \text{ (stat)}^{+0.10}_{-0.04} \text{ (syst)} \pm 0.27 \text{ (norm)}) \times 10^{-7}.$$ 

The branching fraction of the signal channel with respect to its normalisation mode is determined to be

$$\frac{B(B^+ \to \phi K^+\mu^+\mu^-)}{B(B^+ \to J/\psi \phi K^+)} = (1.58^{+0.36}_{-0.32} \text{ (stat)}^{+0.19}_{-0.07} \text{ (syst)}) \times 10^{-3}.$$ 

5.1 Systematic uncertainties

The main systematic uncertainty arises from the measurement of the branching fraction of the normalisation channel, which is known to 33% [31]. The systematic uncertainty due to the choice of signal mass model is determined by using a single Gaussian function with power-law tail on the low-mass side to determine the signal yield. For the background mass model, a first-order polynomial, instead of the nominal second-order polynomial, is used. The total systematic uncertainty from the model used to describe the $m(\phi K^+\mu^+\mu^-)$ distribution is 3%.

The majority of the systematic uncertainties affect the efficiency ratio $\epsilon_{\text{norm}}/\epsilon_{\text{sig}}(q^2)$ and arise from the corrections based on data that are applied to simulation, as described in section 4.1. The systematic uncertainties caused by these corrections are determined to be 1% in total. The limited size of the simulated samples available to calculate the efficiency ratio introduces an uncertainty of 1.5%. Imperfect modelling of the hardware trigger is corrected for in the same way as for the measurement of $B(B^+ \to K^+\pi^+\pi^-\mu^+\mu^-)$ in section 4 and results in a systematic uncertainty of 1.5%.
The efficiency ratio $\epsilon_{\text{norm}} / \epsilon_{\text{sig}}(q^2)$ is determined using simulated $B^+ \to \phi K^+ \mu^+ \mu^-$ events generated according to a phase-space model. The uncertainty due to the $q^2$ distribution in the bins is evaluated by weighting simulated events to reproduce the $q^2$ distribution of $B^+ \to K_1(1270)^+ \mu^+ \mu^-$ decays. This leads to a systematic uncertainty of 1.5%.

6 Conclusions

First observations of the rare $b \to s$ FCNC decays $B^+ \to K^+ \pi^+ \pi^- \mu^+ \mu^-$ and $B^+ \to \phi K^+ \mu^+ \mu^-$ are presented. Their branching fractions are measured to be

$$B(B^+ \to K^+ \pi^+ \pi^- \mu^+ \mu^-) = (4.36^{+0.29}_{-0.27} \text{ (stat)} \pm 0.21 \text{ (syst)} \pm 0.18 \text{ (norm)}) \times 10^{-7},$$

$$B(B^+ \to \phi K^+ \mu^+ \mu^-) = (0.82^{+0.19}_{-0.17} \text{ (stat)} \pm 0.10 \text{ (syst)} \pm 0.27 \text{ (norm)}) \times 10^{-7},$$

where the first uncertainties are statistical, the second systematic and the third due to the uncertainties on the normalisation channels. Accounting for the branching fraction $B(K_1(1270)^+ \to K^+ \pi^+ \pi^-) = (35.7 \pm 3.7\%)$ [31], the measured branching fraction for the decay $B^+ \to K^+ \pi^+ \pi^- \mu^+ \mu^-$ is lower than, but compatible with, the SM prediction of $B(B^+ \to K_1(1270)^+ \mu^+ \mu^-) = (2.3^{+1.3}_{-1.0}-0.2) \times 10^{-6}$ [17]. For the decay $B^+ \to K^+ \pi^+ \pi^- \mu^+ \mu^-$, the differential branching fraction $d\mathcal{B}(B^+ \to K^+ \pi^+ \pi^- \mu^+ \mu^-)/dq^2$ is also determined.

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 (U.S.A.). 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).

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] LHCb collaboration, Differential branching fraction and angular analysis of the decay $B^0 \to K^0\mu^+\mu^-$, JHEP 08 (2013) 131 [arXiv:1304.6325] [inSPIRE].


[3] LHCb collaboration, Differential branching fraction and angular analysis of the decay $B^0_s \to \phi\mu^+\mu^-$, JHEP 07 (2013) 084 [arXiv:1305.2168] [inSPIRE].


[5] BELLE collaboration, H. Guler et al., Study of the $K^+\pi^+\pi^-$ final state in $B^+ \to J/\psi K^+\pi^+\pi^-$ and $B^+ \to \psi' K^+\pi^+\pi^-$, Phys. Rev. D 83 (2011) 032005 [arXiv:1009.5256] [inSPIRE].


[16] CMS collaboration, Observation of a peaking structure in the $J/\psi\phi$ mass spectrum from $B^\pm \to J/\psi\phi K^\pm$ decays, Phys. Lett. B 734 (2014) 261 [arXiv:1309.6920] [inSPIRE].


29 Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
30 Petersburg Nuclear Physics Institute (PNPI), Gatchina, Russia
31 Institute of Theoretical and Experimental Physics (ITEP), Moscow, Russia
32 Institute of Nuclear Physics, Moscow State University (SINP MSU), Moscow, Russia
33 Institute for Nuclear Research of the Russian Academy of Sciences (INR RAN), Moscow, Russia
34 Budker Institute of Nuclear Physics (SB RAS) and Novosibirsk State University, Novosibirsk, Russia
35 Institute for High Energy Physics (IHEP), Protvino, Russia
36 Universitat de Barcelona, Barcelona, Spain
37 Universidad de Santiago de Compostela, Santiago de Compostela, Spain
38 European Organization for Nuclear Research (CERN), Geneva, Switzerland
39 Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
40 Physik-Institut, Universität Zürich, Zürich, Switzerland
41 Nikhef National Institute for Subatomic Physics, Amsterdam, The Netherlands
42 Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, The Netherlands
43 NSC Kharkiv Institute of Physics and Technology (NSC KIPT), Kharkiv, Ukraine
44 Institute for Nuclear Research of the National Academy of Sciences (KINR), Kyiv, Ukraine
45 University of Birmingham, Birmingham, United Kingdom
46 H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
47 Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom
48 Department of Physics, University of Warwick, Coventry, United Kingdom
49 STFC Rutherford Appleton Laboratory, Didcot, United Kingdom
50 School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom
51 School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom
52 Oliver Lodge Laboratory, University of Liverpool, Liverpool, United Kingdom
53 Imperial College London, London, United Kingdom
54 School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
55 Department of Physics, University of Oxford, Oxford, United Kingdom
56 Massachusetts Institute of Technology, Cambridge, MA, United States
57 University of Cincinnati, Cincinnati, OH, United States
58 University of Maryland, College Park, MD, United States
59 Syracuse University, Syracuse, NY, United States
60 Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, associated to
61 Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China, associated to
62 Institut für Physik, Universität Rostock, Rostock, Germany, associated to
63 National Research Centre Kurchatov Institute, Moscow, Russia, associated to
64 Instituto de Fisica Corpuscular (IFIC), Universitat de Valencia-CSIC, Valencia, Spain, associated to
65 KVI - University of Groningen, Groningen, The Netherlands, associated to
66 Celal Bayar University, Manisa, Turkey, associated to
67 Universidade Federal do Triângulo Mineiro (UFTM), Uberaba-MG, Brazil
68 P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia
69 Università di Bari, Bari, Italy
70 Università di Bologna, Bologna, Italy
71 Università di Cagliari, Cagliari, Italy
72 Università di Ferrara, Ferrara, Italy
73 Università di Firenze, Firenze, Italy
74 Università di Urbino, Urbino, Italy
1 Università di Modena e Reggio Emilia, Modena, Italy
2 Università di Genova, Genova, Italy
3 Università di Milano Bicocca, Milano, Italy
4 Università di Roma Tor Vergata, Roma, Italy
5 Università di Roma La Sapienza, Roma, Italy
6 Università della Basilicata, Potenza, Italy
7 AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland
8 LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain
9 Hanoi University of Science, Hanoi, Viet Nam
10 Università di Padova, Padova, Italy
11 Università di Pisa, Pisa, Italy
12 Scuola Normale Superiore, Pisa, Italy
13 Università degli Studi di Milano, Milano, Italy