Measurement of the phase difference between short- and long-distance amplitudes in the $B^{++}K^{++}\overline{K}^{0}$ Decay

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

| As Published | http://dx.doi.org/10.1140/epjc/s10052-017-4703-2 |
| Publisher |  |
| Version | Final published version |
| Accessed | Tue Mar 12 22:43:47 EDT 2019 |
| Citable Link | http://hdl.handle.net/1721.1/119614 |
| Terms of Use | Creative Commons Attribution 4.0 International License |
| Detailed Terms | http://creativecommons.org/licenses/by/4.0/ |
Measurement of the phase difference between short- and long-distance amplitudes in the $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay

LHCb collaboration

CERN, 1211 Geneva 23, Switzerland

Received: 21 December 2016 / Accepted: 16 February 2017 / Published online: 16 March 2017

Abstract  A measurement of the phase difference between the short- and long-distance contributions to the $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay is performed by analysing the dimuon mass distribution. The analysis is based on $pp$ collision data corresponding to an integrated luminosity of 3 fb$^{-1}$ collected by the LHCb experiment in 2011 and 2012. The long-distance contribution to the $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay is modelled as a sum of relativistic Breit–Wigner amplitudes representing different vector meson resonances. In the Standard Model (SM), FCNC transitions are forbidden at tree level and intermediate hadronic resonances. In the Standard Model (FCNC) transitions and long-distance contributions from non-SM particles [1–5,12,16] or could indicate a problem with existing SM predictions [13,15,17]. To explain the observed tensions, long-distance effects would need to be sizeable in dimuon mass regions far from the pole masses of the resonances. Therefore, it is important to understand how well these long-distance effects are modelled in the SM and how they interfere with the short-distance contributions.

Recent global analyses of measurements involving $b \rightarrow s \ell^+ \ell^-$ flavour-changing neutral-current (FCNC) transitions and long-distance contributions from intermediate hadronic resonances in the Standard Model (SM), FCNC transitions are forbidden at tree level and must occur via a loop-level process. In many extensions of the SM, new particles can contribute to the amplitude of the $b \rightarrow s \ell^+ \ell^-$ process changing the rate of the decay or the distribution of the final-state particles. Decays like $B^+ \rightarrow K^+ \mu^+ \mu^-$ are therefore sensitive probes of physics beyond the SM.

Recent global analyses of measurements involving $b \rightarrow s \ell^+ \ell^-$ processes report deviations from SM predictions at the level of four standard deviations [1–15]. These differences could be explained by new short-distance contributions from non-SM particles [1–5,12,16] or could indicate a problem with existing SM predictions [13,15,17]. To explain the observed tensions, long-distance effects would need to be sizeable in dimuon mass regions far from the pole masses of the resonances. Therefore, it is important to understand how well these long-distance effects are modelled in the SM and how they interfere with the short-distance contributions.

Previous measurements of $b \rightarrow s \ell^+ \ell^-$ processes [18–23] excluded regions of dimuon mass around the $\phi$, $J/\psi$ and $\psi(2S)$ resonances. The amplitude in these mass regions is dominated by the narrow vector resonances and has a large theoretical uncertainty. These dimuon regions are therefore considered insensitive to new physics effects.

In this paper, a first measurement of the phase difference between the contributions to the short-distance and the narrow-resonance amplitudes in the $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay is presented.\textsuperscript{1} For the first time, the branching fraction of the short-distance component is determined without interpolation across the $J/\psi$ and $\psi(2S)$ regions. The measurement is performed through a fit to the full dimuon mass spectrum, $m_{\mu^+\mu^-}$, using a model describing the vector resonances as a sum of relativistic Breit–Wigner amplitudes. This approach is similar to that of Refs. [13,24], with the difference that the magnitudes and phases of the resonant amplitudes are determined using the LHCb data rather than using the external information on the cross-section for $e^+e^- \rightarrow$ hadrons from the BES collaboration [25]. The model includes the $\rho, \omega, \phi$, $J/\psi$ and $\psi(2S)$ resonances, as well as broad charmonium states ($\psi(3770), \psi(4040), \psi(4160)$ and $\psi(4415)$) above the open charm threshold. Evidence for the $\psi(4160)$ resonance in the dimuon spectrum of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays has been previously reported by LHCb in Ref. [26]. The continuum of broad states with pole masses above the maximum $m_{\mu^+\mu^-}$ value allowed in the decay is neglected.

The measurement presented in this paper is performed using a data set corresponding to 3 fb$^{-1}$ of integrated luminosity collected by the LHCb experiment in $pp$ collisions.

\textsuperscript{1} The inclusion of charge-conjugate processes is implied throughout.
during 2011 and 2012 at √s = 7 TeV and 8 TeV. The paper is organised as follows: Section 2 describes the LHCb detector and the procedure used to generate simulated events; the reconstruction and selection of \( B^+ \to K^+\mu^+\mu^- \) decays are described in Sect. 3; Section 4 describes the \( m_{\mu\mu} \) distribution of \( B^+ \to K^+\mu^+\mu^- \) decays, including the model for the various resonances appearing in the dimuon mass spectrum; the fit procedure to the dimuon mass spectrum, including the methods to correct for the detection and selection biases, is discussed in Sect. 5. The results and associated systematic uncertainties are discussed in Sects. 6 and 7. Finally, conclusions are presented in Sect. 8.

2 Detector and simulation

The LHCb detector [27,28] is a single-arm forward spectrometer, covering the pseudorapidity range \( 2 < \eta < 5 \), designed to study the production and decay of particles containing \( b \) or \( c \) quarks. The detector includes a high-precision tracking system divided into three subsystems: a silicon-strip vertex detector surrounding the \( pp \) interaction region, a large-area silicon-strip detector that is 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 situated downstream of the magnet. The tracking system provides a measurement of the 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 momentum scale of tracks in the data is calibrated using the \( B^+ \) and \( J/\psi K^+ \) masses measured in \( B^+ \to J/\psi K^+ \) decays [29]. The minimum distance of a track to a primary vertex (PV), the impact parameter (IP), is measured with a resolution of \( (15 + 29/\sqrt{p_T}) \) µm, 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 (RICH). 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 [30], 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.

A large sample of simulated events is used to determine the effect of the detector geometry, trigger, and selection criteria on the dimuon mass distribution of the \( B^+ \to K^+\mu^+\mu^- \) decay. In the simulation, \( pp \) collisions are generated using PYTHIA 8 [31,32] with a specific LHCb configuration [33]. The decay of the \( B^+ \) meson is described by EVTGEN [34], which generates final-state radiation using PHOTOS [35]. As described in Ref. [36], the GEANT4 toolkit [37,38] is used to implement the interaction of the generated particles with the detector and its response. Data-driven corrections are applied to the simulation following the procedure of Ref. [23]. These corrections account for the small level of mismodelling of the detector occupancy, the \( B^+ \) momentum and vertex quality, and the particle identification (PID) performance. The momentum of every reconstructed track in the simulation is also smeared by a small amount in order to better match the mass resolution of the data.

3 Selection of signal candidates

In the trigger for the 7 TeV (8 TeV) data, at least one of the muons is required to have \( p_T > 1.48 \) GeV/c (\( p_T > 1.76 \) GeV/c) and one of the final-state particles is required to have both \( p_T > 1.4 \) GeV/c (\( p_T > 1.6 \) GeV/c) and an IP > 100 µm with respect to all PVs in the event; if this final-state particle is identified as a muon, \( p_T > 1.0 \) GeV/c is required instead. Finally, the tracks of two or more of the final-state particles are required to form a vertex that is significantly displaced from all PVs.

In the offline selection, signal candidates are built from a pair of oppositely tracks that are identified as muons. The muon pair is then combined with a charged track that is identified as a kaon by the RICH detectors. The signal candidates are required to pass a set of loose preselection requirements that are identical to those described in Ref. [26]. These requirements exploit the decay topology of \( B^+ \to K^+\mu^+\mu^- \) transitions and restrict the data sample to candidates with good-quality vertex and track fits. Candidates are required to have a reconstructed \( K^+\mu^+\mu^- \) mass, \( m_{K\mu\mu} \), in the range 5100 < \( m_{K\mu\mu} < 6500 \) MeV/c².

Combinatorial background, where particles from different decays are mistakenly combined, is further suppressed with the use of a Boosted Decision Tree (BDT) [39,40] using kinematic and geometric information. The BDT is identical to that described in Ref. [26] and uses the same working point. The efficiency of the BDT for signal is uniform with respect to \( m_{K\mu\mu} \).

Specific background processes can mimic the signal if their final states are misidentified or partially reconstructed. The requirements described in Ref. [26] reduce the overall contribution of the background from such decay processes to a level of less than 1% of the expected signal yield in the full mass region. The largest remaining specific background contribution comes from \( B^+ \to \pi^+\mu^+\mu^- \) decays (including \( B^+ \to J/\psi \pi^+ \) and \( B^+ \to \psi(2S)\pi^+ \)), where the pion is mistakenly identified as a kaon.

The \( K^+\mu^+\mu^- \) mass of the selected candidates is shown in Fig. 1. The signal is modelled by the sum of two Gaussian functions and a Gaussian function with power-law tails on
both sides of the peak; these all share a common peak position. A Gaussian function is used to describe a small contribution from $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays around the known $B^+$ mass [41]. Combinatorial background is described by an exponential function with a negative gradient. At low $m_{K\mu\mu}$, the background is dominated by partially reconstructed $b$-hadron decays, e.g. from $B^{\pm}(0) \rightarrow K^{\pm}(0) \mu^+ \mu^-$ decays in which the pion from the $K^{\pm}(0)$ is not reconstructed. This background component is modelled using the upper tail of a Gaussian function. The shape of the background from $B^+ \rightarrow \pi^+ \mu^+ \mu^-$ decays is taken from a sample of simulated events. Integrating the signal component in a $\pm 40$ MeV/c$^2$ window around the known $B^+$ mass [41] yields 980 000 $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays.

When computing $m_{\mu\mu}$, a kinematic fit is performed to the selected candidates. In the fit, the $K_{\mu\mu}$ mass is constrained to the known $B^+$ mass and the candidate is required to originate from one of the PVs in the event. For simulated $B^+ \rightarrow J/\psi K^+$ decays, this improves the resolution in $m_{\mu\mu}$ by about a factor of two.

### 4 Differential decay rate

Following the notation of Ref. [42], the $CP$-averaged differential decay rate of $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays as a function of the dimuon mass squared, $q^2 \equiv m_{\mu\mu}^2$, is given by

$$
\frac{d\Gamma}{dq^2} = \frac{G_F^2 |V_{tb} V_{ts}^*|^2 |k| |\beta|}{128\pi^5} \left[ 2 |k|^2 \beta^2 \left| c_{10} f_0(q^2) \right|^2 ight] \\
+ \frac{4m_{\mu}^2 (m_{\mu}^2 - m_K^2)^2}{q^2 m_{\mu}^2} \left| c_{10} f_0(q^2) \right|^2 \\
+ |k|^2 \left[ 1 - \frac{1}{3} \beta^2 \right] \left| c_9 f_0(q^2) + 2c_{10} m_B + m_s \frac{m_B + m_K}{f}\left( |q^2| \right) \right|^2.
$$

(1)

where $|k|$ is the kaon momentum in the $B^+$ meson rest frame. Here $m_K$ and $m_B$ are the masses of the $K^+$ and $B^+$ mesons while $m_s$ and $m_b$ refer to the $s$ and $b$ quark masses as defined in Ref. [42], $\mu$ is the muon mass and $\beta^2 = 1 - 4m_{\mu}^2/q^2$. The constants $G_F, \alpha$, and $V_{tb}$ are the Fermi constant, the QED fine structure constant, and CKM matrix elements, respectively. The parameters $f_0, f$ denote the scalar, vector and tensor $B \rightarrow K$ form factors. The $C_i$ are the Wilson coefficients in an effective field theory description of the decay. The coefficient $C_0$ corresponds to the coupling strength of the vector current operator, $C_{10}$ to the axial-vector current operator and $C_7$ to the electromagnetic dipole operator. The operator definitions and the numerical values of the Wilson coefficients in the SM can be found in Ref. [43]. Right-handed Wilson coefficients, conventionally denoted $C_i'$, are suppressed in the SM and are ignored in this analysis. The Wilson coefficients $C_0$ and $C_{10}$ are assumed to be real. This implicitly assumes that there is no weak phase associated with the short-distance contribution.

In general, $CP$-violating effects are expected to be small across the $m_{\mu\mu}$ range with the exception of the region around the $\rho$ and $\omega$ resonances, which enter with different strong and weak phases [44]. The small size of the $CP$ asymmetry between $B^+$ and $B^-$ decays is confirmed in Ref. [45]. In the present analysis, there is no sensitivity to $CP$-violating effects at low masses and therefore the phases of the resonances are taken to be the same for $B^+$ and $B^-$ decays throughout.

Vector resonances, which produce dimuon pairs via a virtual photon, mimic a contribution to $C_9$. These long-distance hadronic contributions to the $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay are taken into account by introducing an effective Wilson coefficient in place of $C_9$ in Eq. 1.

$$
C_9^{\text{eff}} = C_9 + Y(q^2),
$$

(2)

where the term $Y(q^2)$ describes the sum of resonant and continuum hadronic states appearing in the dimuon mass spectrum. In this analysis $Y(q^2)$ is replaced by the sum of vector meson resonances $j$ such that

$$
C_9^{\text{eff}} = C_9 + \sum_j \eta_j e^{i\delta_j} A_j^{\text{res}}(q^2),
$$

(3)

where $\eta_j$ is the magnitude of the resonance amplitude and $\delta_j$ is its phase relative to $C_9$. These phase differences are one of the main results of this paper. The $q^2$ dependence of the magnitude and phase of the resonance is parameterised by $A_j^{\text{res}}(q^2)$. The resonances included are the $\omega, \rho^0, \phi, J/\psi, \psi(2S), \psi(3770), \psi(4040), \psi(4160)$ and $\psi(4415)$. Contributions from other broad resonances and hadronic continuum states are ignored, as are contributions from weak annihilation [46–48]. No systematic uncertainties are attributed to
these assumptions, which are part of the model that defines the analysis framework of this paper.

The function $A_{10}^q(q^2)$ is taken to have the form of a relativistic Breit–Wigner function for the $\omega, \rho^0, \phi, J/\psi, \psi(2S)$ and $\psi(4040), \psi(4160)$ and $\psi(4415)$ resonances,

$$A_{10}^{\psi}(q^2) = \frac{m_{0j} \Gamma_{0j}}{(m_{0j}^2 - q^2) - i m_{0j} \Gamma_{0j}(q^2)},$$

where $m_{0j}$ is the pole mass of the $j$th resonance and $\Gamma_{0j}$ its natural width. The running width $\Gamma_j(q^2)$ is given by

$$\Gamma_j(q^2) = \frac{p}{p_0j} \frac{m_{0j}}{\sqrt{q^2}} \Gamma_{0j},$$

where $p$ is the momentum of the muons in the rest frame of the dimuon system evaluated at $q$, and $p_0j$ is the momentum evaluated at the mass of the resonance. To account for the open charm threshold, the lineshape of the $\psi(3770)$ resonance is described by a Flatté function [49] with a width defined as

$$\Gamma_{\psi(3770)}(q^2) = \frac{p}{p_0j} \frac{m_{0j}}{\sqrt{q^2}} \left[ \Gamma_1 + \Gamma_2 \sqrt{\frac{1 - (4m_D^2/q^2)^2}{1 - (4m_D^2/q_0^2)}}, \right]$$

where $m_D$ is the mass of the $D^0$ meson and $q_0^2$ is the $q^2$ value at the pole mass of the $\psi(3770)$. The coefficients $\Gamma_1 = 0.3 \text{ MeV}/c^2$ and $\Gamma_2 = 27 \text{ MeV}/c^2$ are taken from Ref. [41] and correspond to the sum of the partial widths of the $\psi(3770)$ to states below and above the open charm threshold. For $q^2 < 4m_D^2$, the phase-space factor accompanying $\Gamma_j$ in Eq. 6 becomes complex.

The form factors are parameterised according to Ref. [50] as

$$f_0(q^2) = \frac{1}{1 - q^2/m_{B^0}^2} \sum_{i=0}^{N-1} b_i^0 z^i,$$

$$f_{+,T}(q^2) = \frac{1}{1 - q^2/m_{B^*}^2} \sum_{i=0}^{N-1} b_{i+T} z^i - (-1)^{i-N} \left( \frac{i}{N} \right) z^N,$$

with, for this analysis, $N = 3$. Here $m_{B^+}(m_{B^*})$ is the mass of the lowest-lying excited $B^+$ meson with $J^P = 1^- (0^+)$. The coefficients $b_i^0$ are allowed to vary in the fit to the data subject to constraints from Ref. [42], whereas the coefficients $b_i^0$ and $b_i^T$ are fixed to their central values. The function $z$ is defined by the mapping

$$z(q^2) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}$$

with

$$t_+ \equiv (m_B - m_K)^2$$

and

$$t_0 \equiv (m_B + m_K)(\sqrt{m_B} - \sqrt{m_K})^2.$$

### 5 Fit to the $m_{\mu\mu}$ distribution

In order to determine the magnitudes and phases of the different resonant contributions, a maximum likelihood fit in 538 bins is performed to the distribution of the reconstructed dimuon mass, $m_{\mu\mu}^{\text{rec}}$, of candidates with $m_{K\mu\mu}$ in a $\pm 40 \text{ MeV}/c^2$ window about the known $B^+$ mass. The $m_{\mu\mu}^{\text{rec}}$ distribution of the $B^+ \to K^+ \mu^+ \mu^-$ decay is described by

$$R(m_{\mu\mu}^{\text{rec}}, m_{\mu\mu}) \propto \left( \epsilon(m_{\mu\mu}) \frac{d\Gamma}{dq^2} \frac{dq^2}{dm_{\mu\mu}} \right),$$

i.e. by Eq. 1, multiplied by the detector efficiency, $\epsilon$, as a function of the true dimuon mass, $m_{\mu\mu}$, and convolved with the experimental mass resolution $R$ discussed in Sect. 5.2.

#### 5.1 Signal model

The magnitudes and phases of the resonances are allowed to vary in the fit, as are the Wilson coefficients $C_9$ and $C_{10}$. As the contribution of $C_7$ to the total decay rate is small, it is fixed to its SM value of $C_7^{\text{SM}} = -0.304 \pm 0.006$ [43].

The form factor $f_0(q^2)$ is constrained in the fit according to its value and uncertainty from Ref. [42]. The form factors $f_0(q^2)$ and $f_T(q^2)$ have a limited impact on the normalisation and shape of Eq. 1, and are fixed to their values from Ref. [42]. The masses and widths of the broad resonances above the open charm threshold are constrained according to their values in Ref. [51]. The masses and widths of the $J/\psi$ and $\psi(2S)$ mesons are fixed to their known values [41]. The large magnitude of the $J/\psi$ and $\psi(2S)$ amplitudes makes the fit very sensitive to the position of the pole mass of these resonances. Due to some residual uncertainty on the momentum scale in the data, the pole masses of the $J/\psi$ and $\psi(2S)$ mesons are allowed to vary in the fit.

The short-distance component is normalised to the branching fraction of $B^+ \to J/\psi K^+$ measured by the $B$-factory experiments [41]. After correcting for isospin asymmetries in the production of the $B^+$ mesons at the $\Upsilon(4S)$, the branching fraction is $B(B^+ \to J/\psi K^+) = (9.95 \pm 0.32) \times 10^{-4}$ [52]. This is further multiplied by $B(J/\psi \to \mu^+ \mu^-) = (5.96 \pm 0.03) \times 10^{-2}$ [41] to account for the decay of the $J/\psi$ meson.
Table 1 Resolution parameters of the different convolution regions in units of MeV/c². The \(\alpha_l\) and \(\alpha_u\) parameters are shared between the \(J/\psi\) and \(\psi(2S)\) regions. The parameters without uncertainties are fixed from fits to the simulated events

<table>
<thead>
<tr>
<th>Region (MeV/c²)</th>
<th>(\sigma_G)</th>
<th>(\sigma_C)</th>
<th>(\alpha_l)</th>
<th>(n_l)</th>
<th>(\alpha_u)</th>
<th>(n_u)</th>
<th>(f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[300, 1800]</td>
<td>3.53</td>
<td>2.98</td>
<td>-1.15</td>
<td>20.0</td>
<td>1.15</td>
<td>20.0</td>
<td>0.39</td>
</tr>
<tr>
<td>[1800, 3400]</td>
<td>6.71 ± 0.04</td>
<td>5.67 ± 0.02</td>
<td>-1.21 ± 0.02</td>
<td>9.1 ± 1.0</td>
<td>1.21 ± 0.02</td>
<td>20.0</td>
<td>0.41 ± 0.01</td>
</tr>
<tr>
<td>[3400, 4700]</td>
<td>5.63 ± 0.04</td>
<td>4.76 ± 0.02</td>
<td>-1.21 ± 0.02</td>
<td>8.5 ± 0.5</td>
<td>1.21 ± 0.02</td>
<td>7.3 ± 1.2</td>
<td>0.41 ± 0.01</td>
</tr>
</tbody>
</table>

The branching fraction of the decay \(B^+ \rightarrow K^+ \mu^+ \mu^-\) via an intermediate resonance \(J\) is computed from the fit as

\[
\tau_B \frac{G_F^2 \alpha_s^2 |V_{tb}|^2 |V_{tb}|^2}{128\pi^3} \int_{4m^2_\mu} (m - m_K)^2 |k|^3 \left[ \beta - \frac{1}{3} \beta^3 \right] \times \left| f_+(q^2)^2 \right| |r_j|^2 \left| A^{res}_J (q^2) \right|^2 dq^2,
\]

where \(\tau_B\) is the lifetime of the \(B^+\) meson. The branching fractions of \(B^+ \rightarrow \rho K^+, B^+ \rightarrow \omega K^+, B^+ \rightarrow \phi K^+\) and \(B^+ \rightarrow \psi(3770)K^+\) are also constrained assuming factorisation between the \(B\) decay and the subsequent decay of the intermediate resonance to a muon pair. These branching fractions are taken from Ref. [41].

5.2 Mass resolution

The convolution of the resolution function with the signal model is implemented using a fast Fourier transform technique [53,54]. The fit to the data is performed in three separate regions of dimuon mass: 300 \(\leq m^{rec}_{\mu\mu} \leq 1800\) MeV/c², 1800 \(< m^{rec}_{\mu\mu} \leq 3400\) MeV/c² and 3400 \(< m^{rec}_{\mu\mu} \leq 4700\) MeV/c².

To increase the speed of the fit, the resolution is treated as constant within these regions using the resolution at the \(\phi\), \(J/\psi\) and \(\psi(2S)\) pole masses. The impact of this assumption on the measured phases of the \(J/\psi\) and \(\psi(2S)\) resonances has been tested using pseudoexperiments and found to be negligible. This is to be expected as the spectra in all other regions vary slowly in comparison to the resolution function. The resolution is modelled using the sum of a Gaussian function, \(G\), and a Gaussian function with power-law tails on the lower and upper side of the peak, \(C\),

\[
R (m^{rec}_{\mu\mu}, m_{\mu\mu}) = f \ G (m^{rec}_{\mu\mu}, m_{\mu\mu}, \sigma_G) + (1 - f) \ C (m^{rec}_{\mu\mu}, m_{\mu\mu}, \sigma_C, n_l, n_u, \alpha_l, \alpha_u).
\]

The component with power-law tails is defined as

\[
C (m^{rec}_{\mu\mu}, m_{\mu\mu}, \sigma_C, n_l, n_u, \alpha_l, \alpha_u) = \begin{cases} A_1 \ (B_1 - \delta)^{-n_l} & \text{if } \delta < \alpha_l \\ \exp(-\delta^2/2) & \text{if } \alpha_l < \delta < \alpha_u \\ A_u \ (B_u + \delta)^{-n_u} & \text{if } \delta > \alpha_u \end{cases}
\]

with

\[
\delta = (m^{rec}_{\mu\mu} - m_{\mu\mu}) / \sigma_C
\]

\[
A_{1,u} = \left( \frac{m_{1,u}}{|\alpha_{1,u}|} \right)^{n_{1,u}} e^{-|\alpha_{1,u}|^2/2}
\]

\[
B_{1,u} = \left( \frac{m_{1,u}}{|\alpha_{1,u}|} \right) - |\alpha_{1,u}|
\]

and is normalised to unity.

The parameters describing the resolution model for the \(J/\psi\) and \(\psi(2S)\) regions \((f, \sigma, \sigma_G, n_l, n_u, \alpha_l, \alpha_u)\) are allowed to vary in the fit to the data. The parameters \(\alpha_l\), \(\alpha_u\) and \(f\) are shared between the \(J/\psi\) and \(\psi(2S)\) regions. The resolution parameters for the \(\phi\) region can not be determined from the data in this way and are instead fixed to their values in the simulation. The resulting values of the resolution parameters are summarised in Table 1. As a cross-check, a second fit to the \(m^{rec}_{\mu\mu}\) distribution is performed using the full \(m_{\mu\mu}\) dependence of the resolution model in Eq. 12 and a numerical implementation of the convolution. In this fit to the data, the parameters of the resolution model are taken from simulated \(B^+ \rightarrow K^+ \mu^+ \mu^-\) events and fixed up to an overall scaling of the width of the resolution function. The two fits to \(m^{rec}_{\mu\mu}\) yield compatible results.

5.3 Efficiency correction

The measured dimuon mass distribution is biased by the trigger, selection and detector geometry. The dominant sources of bias are the geometrical acceptance of the detector, the impact parameter requirements on the muons and the kaon and the \(p_T\) dependence of the trigger. Figure 2 shows the efficiency to trigger, reconstruct and select candidates as a function of \(m_{\mu\mu}\) in a sample of simulated \(B^+ \rightarrow K^+ \mu^+ \mu^-\) candidates. The rise in efficiency with increasing dimuon mass originates from the requirement that one of the muons has \(p_T > 1.48\) GeV/c \((p_T > 1.76\) GeV/c\) in the 2011 (2012) trigger. The drop in efficiency at large dimuon mass (small hadronic recoil) originates from the impact parameter requirement on the kaon. The efficiency is normalised to the efficiency at the \(J/\psi\) meson mass and is parameterised as a function of \(m_{\mu\mu}\) by the sum of Legendre polynomials, \(P_l(x)\), up to sixth order.
The reconstructed dimuon mass distribution of the combinatorial background candidates is taken from the \( m_{K\mu\mu} \) upper mass sideband, \( 5620 < m_{K\mu\mu} < 5700 \text{ MeV}/c^2 \). When evaluating \( m_{\mu\mu}^{\text{rec}} \), \( m_{K\mu\mu} \) is constrained to the centre of the sideband rather than to the known \( B^+ \) mass. Combinatorial background comprising a genuine \( J/\psi \) or \( \psi(2S) \) meson is described by the sum of two Gaussian functions. After applying the mass constraint, the means of the Gaussians do not correspond exactly to the known \( J/\psi \) and \( \psi(2S) \) masses. Combinatorial background comprising a dimuon pair that does not originate from a \( J/\psi \) or \( \psi(2S) \) meson is described by an ARGUS function [55]. The lineshape of the background from \( B^+ \to \pi^+ \mu^+ \mu^- \) decays, where the pion is mistakenly identified as a kaon, is taken from simulated events.

6 Results

The dimuon mass distributions and the projections of the fit to the data are shown in Fig. 3. Four solutions are obtained with almost equal likelihood values, which correspond to ambiguities in the signs of the \( J/\psi \) and \( \psi(2S) \) phases. The values of the phases and branching fractions of the vector meson resonances are listed in Table 3. The posterior values for the \( f_\pm \) form factor are reported in Table 4. A \( \chi^2 \) test between the data and the model, with the binning scheme used in Fig. 3, results in a \( \chi^2 \) of 110 with 78 degrees of freedom. The largest disagreements between the data and the model are localised in the \( m_{K\mu} \) region close to the \( J/\psi \) pole mass and around 1.8 GeV/c\(^2\). The latter is discussed in Sect. 7.

The branching fraction of the short-distance component of the \( B^+ \to K^+ \mu^+ \mu^- \) decay can be calculated by integrating Eq. 1 after setting the amplitudes of the resonances to zero. This gives

\[
B(B^+ \to K^+ \mu^+ \mu^-) = (4.37 \pm 0.15 \text{ (stat)} \pm 0.23 \text{ (syst)}) \times 10^{-7},
\]

where the statistical uncertainty includes the uncertainty on the form-factor predictions. The systematic uncertainty on the branching fraction is discussed in Sect. 7. This measurement is compatible with the branching fraction reported in Ref. [22]. The two results are based on the same data and therefore should not be used together in global fits. The branching fraction reported in Ref. [22] is based on a binned measurement in \( q^2 \) regions away from the narrow resonances (\( \phi, J/\psi \) and \( \psi(2S) \)) and then extrapolated to the full \( q^2 \) range. The contribution from the broad resonances was thus included in that result.

### Table 2

Parameters describing the efficiency to trigger, reconstruct and select simulated \( B^+ \to K^+ \mu^+ \mu^- \) decays as a function of \( m_{\mu\mu} \).

<table>
<thead>
<tr>
<th>( \varepsilon_0 )</th>
<th>( \varepsilon_1 )</th>
<th>( \varepsilon_2 )</th>
<th>( \varepsilon_3 )</th>
<th>( \varepsilon_4 )</th>
<th>( \varepsilon_5 )</th>
<th>( \varepsilon_6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.9262</td>
<td>0.1279</td>
<td>-0.0532</td>
<td>-0.1857</td>
<td>-0.1269</td>
<td>-0.0205</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>0.0036</td>
<td>0.0080</td>
<td>0.0116</td>
<td>0.0131</td>
<td>0.0155</td>
<td>0.0138</td>
</tr>
<tr>
<td>Correlation</td>
<td>( \varepsilon_0 )</td>
<td>( \varepsilon_1 )</td>
<td>( \varepsilon_2 )</td>
<td>( \varepsilon_3 )</td>
<td>( \varepsilon_4 )</td>
<td>( \varepsilon_5 )</td>
</tr>
<tr>
<td>( \varepsilon_0 )</td>
<td>1.000</td>
<td>-0.340</td>
<td>0.605</td>
<td>-0.208</td>
<td>0.432</td>
<td>-0.132</td>
</tr>
<tr>
<td>( \varepsilon_1 )</td>
<td>1.000</td>
<td>-0.345</td>
<td>0.635</td>
<td>-0.207</td>
<td>0.411</td>
<td>-0.094</td>
</tr>
<tr>
<td>( \varepsilon_2 )</td>
<td>1.000</td>
<td>-0.352</td>
<td>0.684</td>
<td>-0.224</td>
<td>0.455</td>
<td></td>
</tr>
<tr>
<td>( \varepsilon_3 )</td>
<td>1.000</td>
<td>-0.344</td>
<td>0.608</td>
<td>-0.154</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \varepsilon_4 )</td>
<td>1.000</td>
<td>-0.344</td>
<td>0.619</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \varepsilon_5 )</td>
<td>1.000</td>
<td>-0.259</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \varepsilon_6 )</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
positive (left between the short- and phase is negative (top left); the $J/\psi$ phase is positive and the $\psi(2S)$ phase is negative (top right); the $J/\psi$ phase is negative and the $\psi(2S)$ phase is positive (bottom left); and both phases are positive (bottom right). The component labelled interference refers to the interference between the short- and long-distance contributions to the decay. The $\chi^2$ value of the four solutions is almost identical, with a value of 110 for 78 degrees of freedom

**Table 3** Branching fractions and phases for each resonance in the fit for the four solutions of the $J/\psi$ and $\psi(2S)$ phases. Both statistical and systematic contributions are included in the uncertainties. There is a common systematic uncertainty of 4.5%, dominated by the uncertainty on the $B^+ \rightarrow J/\psi K^+$ branching fraction, which provides the normalisation for all measurements

<table>
<thead>
<tr>
<th>Resonance</th>
<th>$J/\psi$ negative/$\psi(2S)$ negative</th>
<th>$J/\psi$ negative/$\psi(2S)$ positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase [rad]</td>
<td>Branching fraction</td>
<td>Phase [rad]</td>
</tr>
<tr>
<td>$\rho(770)$</td>
<td>$-0.35 \pm 0.54$ ($1.71 \pm 0.25) \times 10^{-10}$</td>
<td>$-0.30 \pm 0.54$ ($1.71 \pm 0.25) \times 10^{-10}$</td>
</tr>
<tr>
<td>$\omega(782)$</td>
<td>$0.26 \pm 0.39$ ($4.93 \pm 0.59) \times 10^{-10}$</td>
<td>$0.30 \pm 0.38$ ($4.93 \pm 0.58) \times 10^{-10}$</td>
</tr>
<tr>
<td>$\phi(1020)$</td>
<td>$0.47 \pm 0.39$ ($2.53 \pm 0.26) \times 10^{-9}$</td>
<td>$0.51 \pm 0.37$ ($2.53 \pm 0.26) \times 10^{-9}$</td>
</tr>
<tr>
<td>$J/\psi$</td>
<td>$-1.66 \pm 0.05$ –</td>
<td>$-1.50 \pm 0.05$ –</td>
</tr>
<tr>
<td>$\psi(2S)$</td>
<td>$-1.93 \pm 0.10$ ($4.64 \pm 0.20) \times 10^{-6}$</td>
<td>$2.08 \pm 0.11$ ($4.69 \pm 0.20) \times 10^{-6}$</td>
</tr>
<tr>
<td>$\psi(3770)$</td>
<td>$-2.13 \pm 0.42$ ($1.38 \pm 0.54) \times 10^{-9}$</td>
<td>$-2.89 \pm 0.19$ ($1.67 \pm 0.61) \times 10^{-9}$</td>
</tr>
<tr>
<td>$\psi(4040)$</td>
<td>$-2.52 \pm 0.66$ ($4.17 \pm 2.72) \times 10^{-10}$</td>
<td>$-2.69 \pm 0.52$ ($4.25 \pm 2.83) \times 10^{-10}$</td>
</tr>
<tr>
<td>$\psi(4160)$</td>
<td>$-1.90 \pm 0.64$ ($2.61 \pm 0.84) \times 10^{-9}$</td>
<td>$-2.13 \pm 0.33$ ($2.67 \pm 0.85) \times 10^{-9}$</td>
</tr>
<tr>
<td>$\psi(4415)$</td>
<td>$-2.52 \pm 0.36$ ($6.04 \pm 3.93) \times 10^{-10}$</td>
<td>$-2.43 \pm 0.43$ ($7.10 \pm 4.48) \times 10^{-10}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Resonance</th>
<th>$J/\psi$ positive/$\psi(2S)$ negative</th>
<th>$J/\psi$ positive/$\psi(2S)$ positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase [rad]</td>
<td>Branching fraction</td>
<td>Phase [rad]</td>
</tr>
<tr>
<td>$\rho(770)$</td>
<td>$-0.26 \pm 0.54$ ($1.71 \pm 0.25) \times 10^{-10}$</td>
<td>$-0.22 \pm 0.54$ ($1.71 \pm 0.25) \times 10^{-10}$</td>
</tr>
<tr>
<td>$\omega(782)$</td>
<td>$0.35 \pm 0.39$ ($4.93 \pm 0.58) \times 10^{-10}$</td>
<td>$0.38 \pm 0.38$ ($4.93 \pm 0.58) \times 10^{-10}$</td>
</tr>
<tr>
<td>$\phi(1020)$</td>
<td>$0.58 \pm 0.38$ ($2.53 \pm 0.26) \times 10^{-9}$</td>
<td>$0.62 \pm 0.37$ ($2.52 \pm 0.26) \times 10^{-9}$</td>
</tr>
<tr>
<td>$J/\psi$</td>
<td>$1.47 \pm 0.05$ –</td>
<td>$1.63 \pm 0.05$ –</td>
</tr>
<tr>
<td>$\psi(2S)$</td>
<td>$-2.21 \pm 0.11$ ($4.63 \pm 0.20) \times 10^{-6}$</td>
<td>$1.80 \pm 0.10$ ($4.68 \pm 0.20) \times 10^{-6}$</td>
</tr>
<tr>
<td>$\psi(3770)$</td>
<td>$-2.40 \pm 0.39$ ($1.39 \pm 0.54) \times 10^{-9}$</td>
<td>$-2.95 \pm 0.14$ ($1.68 \pm 0.61) \times 10^{-9}$</td>
</tr>
<tr>
<td>$\psi(4040)$</td>
<td>$-2.64 \pm 0.50$ ($4.05 \pm 2.76) \times 10^{-10}$</td>
<td>$-2.75 \pm 0.48$ ($4.30 \pm 2.86) \times 10^{-10}$</td>
</tr>
<tr>
<td>$\psi(4160)$</td>
<td>$-2.11 \pm 0.38$ ($2.62 \pm 0.82) \times 10^{-9}$</td>
<td>$-2.28 \pm 0.24$ ($2.68 \pm 0.81) \times 10^{-9}$</td>
</tr>
<tr>
<td>$\psi(4415)$</td>
<td>$-2.42 \pm 0.46$ ($6.13 \pm 3.98) \times 10^{-10}$</td>
<td>$-2.31 \pm 0.48$ ($7.12 \pm 4.94) \times 10^{-10}$</td>
</tr>
</tbody>
</table>
A two-dimensional likelihood profile of \( C_9 \) and \( C_{10} \) is also obtained as shown in Fig. 4. The intervals correspond to \( \chi^2 \) probabilities assuming two degrees of freedom. Only the quadrant with \( C_9 \) and \( C_{10} \) values around the SM prediction is shown. The other quadrants can be obtained by mirroring in the axes. The branching fraction of the short-distance component provides a good constraint on the sum of \( |C_9|^2 \) and \( |C_{10}|^2 \) (see Eq. 1). This gives rise to the annular shape in the likelihood profile in Fig. 4. In addition, there is a modest ability for the fit to differentiate between \( C_9 \) and \( C_{10} \) through the interference of the \( C_9 \) component with the resonances. The visible interference pattern excludes very small values of \( |C_9| \). Overall, the correlation between \( C_9 \) and \( C_{10} \) is approximately 90\%. The best-fit point for the Wilson coefficients (in a given quadrant of the \( C_9 \) and \( C_{10} \) plane) and the corresponding \( B^+ \rightarrow K^+\mu^+\mu^- \) branching fraction are the same for the four combinations of the \( J/\psi \) and \( \psi(2S) \) phases. Including statistical and systematic uncertainties, the fit results deviate from the SM prediction at the level of 3.0 standard deviations. The uncertainty is dominated by the precision of the form factors. The best-fit point prefers a value of \(|C_{10}|\) that is smaller than \(|C_{10}^{SM}|\) and a value of \(|C_9|\) that is larger than \(|C_9^{SM}|\). However, if \( C_{10} \) is fixed to its SM value, the fit prefers \(|C_9| < |C_{10}^{SM}|\). This is consistent with the results of global fits to \( b \rightarrow s\ell^+\ell^- \) processes. Given the model assumptions in this paper, the interference with the \( J/\psi \) meson is not able to explain the low value of the branching fraction of the \( B^+ \rightarrow K^+\mu^+\mu^- \) decay while keeping the values of \( C_9 \) and \( C_{10} \) at their SM predictions.

### 7 Systematic uncertainties

Sources of systematic uncertainty are considered separately for the phase and branching fraction measurements. In both cases, the largest systematic uncertainties are accounted for in the statistical uncertainty as they are included as nuisance parameters in the fit. For smaller sources of uncertainty, the fit is repeated with variations of the inputs and the difference is assigned as a systematic uncertainty. A summary of the remaining systematic uncertainties can be found in Table 5.

The parameters governing the behaviour of the tails of the resolution function are particularly correlated with the phases. The systematic uncertainty on the resolution model is included in the statistical uncertainty by allowing the resolution parameter values to vary in the fit. If the tail parameters are fixed to their central values, the statistical uncertainties on the phase measurements decrease by approximately 20\%. The choice of parameterisation for the resolution model is validated using a large sample of simulated events and no additional uncertainty is assigned for the choice of model. For the branching fraction measurement, the uncertainty arising from the resolution model is negligible compared to other sources of systematic uncertainty.

Similarly to the resolution model, the systematic uncertainty associated with the knowledge of the \( f_+(q^2) \) form factor is included in the statistical uncertainty. If the form-factor parameters are fixed to their best-fit values, the statistical uncertainties on the phases decrease by 4\% (1\%) for the \( J/\psi (\psi(2S)) \) measurements. For the branching fraction, the uncertainty is 2\%, which is of similar size as the statistical uncertainty.

### Table 4

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Ref. [42]</th>
<th>Fit result</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b_0^+ )</td>
<td>0.466 ± 0.014</td>
<td>0.465 ± 0.013</td>
</tr>
<tr>
<td>( b_1^+ )</td>
<td>−0.89 ± 0.13</td>
<td>−0.81 ± 0.05</td>
</tr>
<tr>
<td>( b_2^+ )</td>
<td>−0.21 ± 0.55</td>
<td>0.03 ± 0.32</td>
</tr>
</tbody>
</table>

### Table 5

<table>
<thead>
<tr>
<th>Source</th>
<th>( J/\psi ) phase</th>
<th>( \psi(2S) ) phase</th>
<th>Branching fraction</th>
<th>( C_9, C_{10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad components</td>
<td>20 mrad</td>
<td>10 mrad</td>
<td>1.0%</td>
<td>0.05</td>
</tr>
<tr>
<td>Efficiency model</td>
<td>10 mrad</td>
<td>10 mrad</td>
<td>1.0%</td>
<td>0.05</td>
</tr>
<tr>
<td>( B(J/\psi \rightarrow J/\psi K^+) )</td>
<td>3 mrad</td>
<td>10 mrad</td>
<td>4.2%</td>
<td>0.19</td>
</tr>
</tbody>
</table>
At around $m_{\mu\mu} = 1.8 \text{ GeV}/c^2$ there is a small discrepancy between the data and the model (see Fig. 3). This is interpreted as a possible contribution from excited $\rho$, $\omega$ or $\phi$ resonances. Given the limited knowledge of the masses and widths of the states in this region, these broad states are neglected in the nominal fit. They are, however, visible in $e^+e^- \rightarrow \text{hadrons}$ vacuum polarisation data [41]. To test the effect of such states on the phases of the $J/\psi$ and $\psi(2S)$ mesons, an additional relativistic Breit–Wigner amplitude is included with a width and mass that are allowed to vary in the fit. The inclusion of this Breit–Wigner amplitude marginally improves the fit quality around $m_{\mu\mu} = 1.8 \text{ GeV}/c^2$ and changes the $J/\psi$ ($\psi(2S))$ phase by 40% (20%) of its statistical uncertainty, which is added as a systematic effect. The magnitude of the amplitude is not statistically significant and its mean and width do not correspond to a known state. The phases of the other resonances in the fit have larger statistical uncertainties and the inclusion of this additional amplitude has a negligible effect on their fit values. Given that the contribution of this amplitude is small compared to the short-distance component, its effect on the branching fraction is only around 1%.

Other, smaller systematic uncertainties include modelling of the combinatorial background, calculation of the efficiency as a function of $q^2$ and the uncertainty on the $B^+ \rightarrow J/\psi K^+$ branching fraction. The latter affects the branching fraction measurement and is obtained from Ref. [52], which results in a 4% uncertainty.

8 Conclusions

This paper presents the first measurement of the phase difference between the short- and long-distance contributions to the $B^+ \rightarrow K^+\mu^+\mu^-$ decay. The measurement is performed using a binned maximum likelihood fit to the dimuon mass distribution of the decays. The long-distance contributions are modelled as the sum of relativistic Breit–Wigner amplitudes representing different vector meson resonances decaying to muon pairs, each with their own magnitude and phase. The short-distance contribution is expressed in terms of an effective field theory description of the decay with the Wilson coefficients $C_0$ and $C_{10}$, which are taken to be real. These are left free in the fit and all other components set to their corresponding SM values. The $B \rightarrow K$ hadronic form factors are constrained in the fit to the predictions from Ref. [42].

The fit results in four approximately degenerate solutions corresponding to ambiguities in the signs of the $J/\psi$ and $\psi(2S)$ phases. The values of the $J/\psi$ phases are compatible with $\pm \frac{\pi}{2}$, which means that the interference with the short-distance component in dimuon mass regions far from their pole masses is small. The negative solution of the $J/\psi$ phase agrees qualitatively with the prediction of Ref. [47], where long-distance contributions are calculated at negative $q^2$ and extrapolated to the $q^2$ region below the $J/\psi$ pole-mass using a hadronic dispersion relation. The fit model, which includes the conventional $J^{PC} = 1^{-+} \rightarrow c\bar{c}$ resonances, is found to describe the data well, with no significant evidence for the decays $B^+ \rightarrow \psi(4040) K^+$ or $B^+ \rightarrow \psi(4415) K^+$. The values of the $\psi(3770)$ and $\psi(4160)$ phases are compatible with those reported in Ref. [13].

The measurement of the Wilson coefficients prefers a value of $|C_{10}| < |C_{10}^{\text{SM}}|$ and a value of $|C_0| > |C_0^{\text{SM}}|$. If the value of $C_{10}$ is set to that of $C_{10}^{\text{SM}}$, the measurement favours the region $|C_0| < |C_0^{\text{SM}}|$. These results are similar to those reported previously in global analyses. The interference between the short- and long-distance contributions in the regions around the $\rho$, $\omega$ and the $\phi$, and in the region $q^2 > m_{\psi(2S)}^2$, results in the exclusion of the hypothesis that $C_0 = 0$ at more than 5 standard deviations. The dominant uncertainty on the measurements of $C_0$ and $C_{10}$ arises from the knowledge of the $B \rightarrow K$ hadronic form factors. The current data set allows the uncertainties on these hadronic parameters to be reduced. Improved inputs on the form factors from lattice QCD calculations and the larger data set that will be available at the end of the LHC Run 2 are needed to further improve the measurement of the Wilson coefficients.

A similar strategy to the one applied in this paper can be extended to other $b \rightarrow s^+\ell^-$ decay processes to understand the influence of hadronic resonances on global fits for $C_0$ and $C_{10}$. However, the situation is more complicated in decays where the strange hadron is not a pseudoscalar meson as the amplitudes corresponding to different helicity states of the hadron can have different relative phases.

Finally, a measurement of the branching fraction of the short-distance component of $B^+ \rightarrow K^+\mu^+\mu^-$ decays is also reported and is found to be

$$B(B^+ \rightarrow K^+\mu^+\mu^-) = (4.37 \pm 0.15 \text{ (stat)} \pm 0.23 \text{ (syst)}) \times 10^{-7},$$

where the first uncertainty is statistical and second is systematic. In contrast to previous analyses, the measurement is performed across the full $q^2$ region accounting for the interference with the long-distance contributions and without any veto of resonance-dominated regions of the phase space. The value of the branching fraction is found to be compatible with previous measurements [22], but smaller than the SM prediction [42].

Acknowledgements 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 and MPG (Germany); INFN (Italy); FOM and NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FASO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASA (Ukraine); STFC (United King-
dom); NSF (USA). 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 (USA). 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), Conseil Général de Haute-Savoie, Labex ENIGMASS and OCEUVU, Région Auvergne (France), RFBR and Yandex LLC (Russia), GVA, Xunta-Gal and GENCAT (Spain), Herchel Smith Fund, The Royal Society, Royal Commission for the Exhibition of 1851 and the Leverhulme Trust (United Kingdom).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Funded by SCOAP3.

References

5. A. Crivellin, G. D’Ambrosio, J. Heeck, Explaining $h \to \mu^+\mu^-$ as $B \to K^\ast \mu^+\mu^-$ and $B \to K \mu^+\mu^-$. Eur. Phys. J. C 73, 2646 (2013). arXiv:1308.1501
12. J. Lyon, R. Zwicky, Resonances gone topsy turvy - the charm of QCD or new physics in $b \to s\ell^+\ell^-$? arXiv:1406.0566
19. CMS collaboration, V. Kharaticyan et al., Angular analysis of the decay $B^0 \to K^*0 \mu^+\mu^-$ from $pp$ collisions at $\sqrt{s}=7$ TeV. Phys. Lett. B 753, 424 (2016). arXiv:1507.08126
20. LHCb collaboration, R. Aaij et al., Differential branching fractions and isospin asymmetries of $B \to K^\ast \mu^+\mu^-$ decays. JHEP 06, 133 (2014). arXiv:1403.8044
21. LHCb collaboration, R. Aaij et al., Angular analysis of the $B^0 \to K^\ast \mu^+\mu^-$ decay using $3$ fb$^{-1}$ of integrated luminosity. JHEP 02, 104 (2016). arXiv:1512.04442
25. LHCb collaboration, A. Alves Jr. et al., The LHCb detector at the LHC. JINST 3, S08005 (2008)
42. J.A. Bailey et al., $B \to K^{\mp}\ell^{\pm}\ell'^{\mp}$ decay form factors from three-flavor lattice QCD. Phys. Rev. D 93, 025026 (2016). arXiv:1509.06235
45. LHCb collaboration, R. Aaij et al., Measurement of CP asymmetries in the decays $B^0 \to K^{0}\mu^+\mu^-$ and $B^0 \to K^{+}\mu^+\mu^-$. JHEP 09, 177 (2014). arXiv:1408.0978
47. A. Khodjamirian, T. Mannel, Y.M. Wang, $B \to K^{\mp}\ell^{\pm}\ell'^{\mp}$ decay at large hadronic recoil. JHEP 02, 010 (2013). arXiv:1211.0234
48. J. Lyon, R. Zwicky, Isospin asymmetries in $B \to (K^*, \nu\gamma)/\ell^-\ell'^{\mp}$ and $B \to \ell^-\ell'^{\mp}$ in and beyond the Standard Model. Phys. Rev. D 88, 094004 (2013). arXiv:1305.4797
49. S.M. Flatté, Coupled-channel analysis of the $\pi\eta$ and $K\bar{K}$ systems near $K\bar{K}$ threshold. Phys. Lett. B 63, 224 (1976)

LHCb collaboration

Università di Urbino, Urbino, Italy
Università della Basilicata, Potenza, Italy
Scuola Normale Superiore, Pisa, Italy
Università di Modena e Reggio Emilia, Modena, Italy
Iligan Institute of Technology (IIT), Iligan, Philippines
Novosibirsk State University, Novosibirsk, Russia
† Deceased