Search for lepton-universality violation in $B^+ \to K^+ \ell^+ \ell^-$ decays

LHCb collaboration

Abstract

A measurement of the ratio of branching fractions of the decays $B^+ \to K^+ \mu^+ \mu^-$ and $B^+ \to K^+ e^+ e^-$ is presented. The proton-proton collision data used correspond to an integrated luminosity of $5.0 \text{fb}^{-1}$ recorded with the LHCb experiment at centre-of-mass energies of 7, 8 and 13 TeV. For the dilepton mass-squared range $1.1 < q^2 < 6.0 \text{GeV}^2/c^4$ the ratio of branching fractions is measured to be $R_K = 0.846^{+0.060}_{-0.054}^{+0.016}_{-0.014}$, where the first uncertainty is statistical and the second systematic. This is the most precise measurement of $R_K$ to date and is compatible with the Standard Model at the level of 2.5 standard deviations.


© 2019 CERN for the benefit of the LHCb collaboration. CC-BY-4.0 licence

†Authors are listed at the end of this paper.
Decays involving $b \to s \ell^+\ell^-$ transitions, where $\ell$ represents a lepton, are mediated by flavour-changing neutral currents. Such decays are suppressed in the Standard Model (SM), as they proceed only through amplitudes that involve electroweak loop diagrams. These processes are sensitive to virtual contributions from new particles, which could have masses that are inaccessible to direct searches for resonances, even at Large Hadron Collider experiments.

Theoretical predictions for exclusive $b \to s \ell^+\ell^-$ decays rely on the calculation of hadronic effects, and recent measurements have therefore focused on quantities where the uncertainties from such effects are reduced to some extent, such as angular observables and ratios of branching fractions. The results of the angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ [1,9] and measurements of the branching fractions of several $b \to s \ell^+\ell^-$ decays [10,13] are in some tension with SM predictions [14–19]. However, the treatment of the hadronic effects in the theoretical predictions is still the subject of considerable debate [20–29].

The electroweak couplings of all three charged leptons are identical in the SM and, consequently, the decay properties (and the hadronic effects) are expected to be the same up to corrections related to the lepton mass, regardless of the lepton flavour (referred to as lepton universality). The ratio of branching fractions for $B \to H\mu^+\mu^-$ and $B \to He^+e^-$ decays, where $H$ is a hadron, can be predicted precisely in an appropriately chosen range of the dilepton mass squared $q^2$, such ratios are predicted to be unity with $\mathcal{O}(1\%)$ precision [32]. The inclusion of charge-conjugate processes is implied throughout this Letter.

The most precise measurements of $R_K$ in the region $1.0 < q^2 < 6.0\text{ GeV}^2/c^4$ and $R_{K}\ast\phi$ in the regions $0.045 < q^2 < 1.1\text{ GeV}^2/c^4$ and $1.1 < q^2 < 6.0\text{ GeV}^2/c^4$ have been made by the LHCb collaboration and, depending on the theoretical prediction used, are 2.6 [33], 2.1–2.3 and 2.4–2.5 standard deviations [34] below their respective SM expectations [20,21,32,35,42]. These tensions and those observed in the angular and branching-fraction measurements can all be accommodated simultaneously in models with an additional heavy neutral gauge boson [43,46] or with leptoquarks [47,51].

This Letter presents the most precise measurement of the ratio $R_K$ in the range $1.1 < q^2 < 6.0\text{ GeV}^2/c^4$. The analysis is performed using 5.0 fb$^{-1}$ of proton-proton collision data collected with the LHCb detector during three data-taking periods in which the centre-of-mass energy of the collisions was 7, 8 and 13 TeV. The data were taken in the years 2011, 2012 and 2015–2016, respectively. Compared to the previous LHCb $R_K$ measurement [33], the analysis benefits from a larger data sample and an improved reconstruction; moreover the lower limit of the $q^2$ range is increased, in order to be compatible with other LHCb $b \to s \ell^+\ell^-$ analyses and to suppress further the contribution from $B^+ \to \phi(\to \ell^+\ell^-)K^+$ decays. The results supersede those of Ref. [33].

Throughout this Letter $B^+ \to K^+\ell^+\ell^-$ refers only to decays with $1.1 < q^2 < 6.0\text{ GeV}^2/c^4$, which are denoted nonresonant, whereas $B^+ \to J/\psi(\to \ell^+\ell^-)K^+$...
decays are referred to as resonant. The nonresonant $q^2$ range excludes the resonant $B^+ \to J/\psi(\to \ell^+\ell^-)K^+$ region and the high-$q^2$ region that contains contributions from excited charmonium resonances.

The analysis strategy is designed to reduce systematic uncertainties induced by the markedly different reconstruction of decays with muons in the final state compared to decays with electrons. These differences arise due to the significant bremsstrahlung emission of the electrons and the different signatures exploited in the online trigger selection. Systematic uncertainties that would otherwise affect the calculation of the efficiencies of the $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ decay modes are suppressed by measuring $R_K$ as a double ratio of branching fractions,

$$R_K = \frac{\mathcal{B}(B^+ \to K^+\mu^+\mu^-)}{\mathcal{B}(B^+ \to J/\psi(\to \mu^+\mu^-)K^+)} \times \frac{\mathcal{B}(B^+ \to K^+e^+e^-)}{\mathcal{B}(B^+ \to J/\psi(\to e^+e^-)K^+)}.$$  \hspace{1cm} (2)

The measurement requires knowledge of the observed yield and the efficiency to trigger, reconstruct and select each decay mode. The use of this double ratio exploits the fact that $J/\psi \to \ell^+\ell^-$ decays are observed to have lepton-universal branching fractions within 0.4% \cite{52,53}. Using Eq. \((2)\) then requires the nonresonant $B^+ \to K^+e^+e^-$ detection efficiency to be known only relative to that of the resonant $B^+ \to J/\psi(\to e^+e^-)K^+$ decay, rather than the $B^+ \to K^+\mu^+\mu^-$ decay. As the detector signatures of each resonant decay are similar to those of the corresponding nonresonant decay, systematic effects are reduced and the precision on $R_K$ is dominated by the statistical uncertainty.

After the application of selection criteria, which are discussed below, the four decay modes $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$, $B^+ \to J/\psi(\to e^+e^-)K^+$, $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^+e^+e^-$ are separated from background on a statistical basis, using fits to the $m(K^+\ell^+\ell^-)$ distributions. For the resonant decays, the mass $m_{J/\psi}(K^+\ell^+\ell^-)$ is computed by constraining the dilepton system to the known $J/\psi$ mass \cite{53}. This improves the electron-mode mass resolution (full width at half maximum) from 140 to 24.5 MeV/$c^2$ and the muon-mode mass resolution from 30 to 17.5 MeV/$c^2$. The $m(K^+\ell^+\ell^-)$ fit ranges and the $q^2$ selection used for the different decay modes are shown in Table 1. The selection requirements applied to the resonant and nonresonant decays are otherwise identical. The two ratios of efficiencies required to form Eq. \((2)\) are taken from simulation. The simulation is calibrated using data-derived control channels, including $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ and $B^+ \to J/\psi(\to e^+e^-)K^+$. Correlations arising from the use of these decay modes both for this calibration and in the determination of the double ratio of Eq. \((2)\) are taken into account. A further feature of the analysis strategy is that the results were not inspected until all analysis procedures were finalised.

The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, described in detail in Refs. \cite{54,55}. The detector includes a silicon-strip vertex detector surrounding the proton-proton interaction region, tracking stations either side of a dipole magnet, ring-imaging Cherenkov (RICH) detectors, calorimeters and muon chambers. The simulation used in this analysis is produced using the software described in Refs. \cite{56,61}. Final-state radiation is simulated using PHOTOS++ 3.61 in the default configuration \cite{59,62}, which is observed to agree with a full quantum electrodynamics calculation at the level of 1% \cite{32}.

Candidate events are first required to pass a hardware trigger that selects either a high transverse momentum ($p_T$) muon; or an electron, hadron or photon with high transverse energy deposited in the calorimeters. In this analysis, it is required that $B^+ \to K^+\mu^+\mu^-$.
Table 1: Resonant and nonresonant mode $q^2$ and $m(K^+\ell^+\ell^-)$ ranges. The variables $m(K^+\ell^+\ell^-)$ and $m_{J/\psi}(K^+\ell^+\ell^-)$ are used for nonresonant and resonant decays, respectively.

<table>
<thead>
<tr>
<th>Decay mode</th>
<th>$q^2$ [GeV$^2$/c$^4$]</th>
<th>$m_{J/\psi}(K^+\ell^+\ell^-)$ [GeV/c$^2$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>nonresonant $e^+e^-$</td>
<td>1.1 – 6.0</td>
<td>4.88 – 6.20</td>
</tr>
<tr>
<td>resonant $e^+e^-$</td>
<td>6.00 – 12.96</td>
<td>5.08 – 5.70</td>
</tr>
<tr>
<td>nonresonant $\mu^+\mu^-$</td>
<td>1.1 – 6.0</td>
<td>5.18 – 5.60</td>
</tr>
<tr>
<td>resonant $\mu^+\mu^-$</td>
<td>8.68 – 10.09</td>
<td>5.18 – 5.60</td>
</tr>
</tbody>
</table>

and $B^+ \to J/\psi (\to \mu^+\mu^-)K^+$ candidates are triggered by one of the muons, whereas $B^+ \to K^+e^+e^-$ and $B^+ \to J/\psi (\to e^+e^-)K^+$ candidates are required to be triggered in one of three ways: by either one of the electrons; by the kaon from the $B^+$ decay; or by particles in the event that are not part of the signal candidate. In the software trigger, the tracks of the final-state particles are required to form a vertex that is significantly displaced from any of the primary proton-proton interaction vertices (PVs) in the event. A multivariate algorithm is used for the identification of secondary vertices consistent with the decay of a $b$ hadron [63,64].

Candidates are formed from a particle identified as a charged kaon, together with a pair of well-reconstructed oppositely charged particles identified as either electrons or muons. Each particle is required to have sizeable $p_T$ and to be inconsistent with coming from a PV. The particles must originate from a common vertex with good vertex-fit quality, which is displaced significantly from all of the PVs in the event. The $B^+$ momentum vector is required to be aligned with the vector connecting one of the PVs in the event (subsequently referred to as the associated PV) and the $B^+$ decay vertex.

Kaons and muons are identified using the output of multivariate classifiers that exploit information from the tracking system, the RICH detectors, the calorimeters and the muon chambers [55,65–69]. Electrons are identified by matching tracks to electromagnetic calorimeter (ECAL) showers and adding information from the RICH detectors. The ratio of the energy detected in the ECAL to the momentum measured by the tracking system is central to this identification. If an electron radiates a photon downstream of the dipole magnet, the photon and electron deposit their energy in the same ECAL cells and the original energy of the electron is measured. However, if an electron radiates a photon upstream of the magnet, the energy of the photon will not be deposited in the same ECAL cells as the electron. For each electron track, a search is therefore made for ECAL showers around the extrapolated track direction (before the magnet) that are not associated with any other charged tracks. The energy of any such shower is added to the electron energy that is derived from the measurements made in the tracker.

Backgrounds from exclusive decays of $b$ hadrons and the so-called combinatorial background, formed from the reconstructed fragments of multiple heavy-flavour hadron decays, are reduced using selection criteria that are discussed below. The muon modes benefit from superior mass resolution so that a reduced mass range can be used (see Table 1). Consequently, the only remaining backgrounds after the application of the selection criteria are combinatorial and, for the resonant mode, from the Cabibbo-suppressed decay $B^+ \to J/\psi \pi^+$, where the pion is misidentified as a kaon. For the
electron modes, where a wider mass range is used, significant residual exclusive backgrounds also contribute. Since higher-mass $K^*$ resonances are suppressed in the mass range selected, the dominant exclusive backgrounds for the resonant and nonresonant modes are from partially reconstructed $B^{0,+} \to J/\psi(\to e^+e^-)K^*(892)(0,+)\to K^+\pi(0,0)$ and $B^{0,+} \to K^*(892) (0,+)\to K^+\pi(0,0)e^+e^-$ decays, respectively, where the pion is not included in the candidate. At the level of $O(1\%)$ of the $K^+e^+e^-$ signal, there are also exclusive background contributions from $B^+ \to D^0(\to K^+e^+\nu_e)\to e^+\nu_e$ decays and, at low $m(K^+e^+e^-)$, from the radiative tail of $B^+ \to J/\psi(\to e^+e^-)K^+$ decays. This tail is visible in the distribution of $m(K^+e^+e^-)$ versus $q^2$, which is given in the Supplemental Material to this Letter.

Cascade backgrounds of the form $H_b \to H_c(\to K^+\ell^-\nu_X)\ell^+\nu_Y$, where $H_b$ is a beauty hadron ($B^+, B^0, B^+_s$ or $L^0_b$), $H_c$ a charm hadron ($D^0, D^+, D^+_s, L^+_c$), and $X, Y$ are particles that are not reconstructed, are suppressed by requiring that the kaon-lepton invariant mass satisfies the constraint $m(K^+\ell^-) > m_{D^0}$, where $m_{D^0}$ is the known $D^0$ mass [53]. Cascade backgrounds with a misidentified particle are suppressed by applying a similar veto, but with the lepton-mass hypothesis changed to that of a pion (denoted $\ell(\to \pi)$). In the muon case, it is sufficient to reject $K\mu[\to \pi]$ combinations with a mass smaller than $m_{D^0}$. In the electron case this veto is applied without the bremsstrahlung recovery, i.e., based on only the measured track momenta, and a window around the $D^0$ mass is used to reject candidates. The vetoes retain 97% of $B^+ \to K^+\mu^+\mu^-$ and 95% of $B^+ \to K^+e^+e^-$ decays passing the full selection. The relevant mass distributions are given in the Supplemental Material [70].

Other exclusive $b$-hadron decays require at least two particles to be misidentified in order to form backgrounds. These include the decays $B^+ \to K^+\pi^+\pi^-$ and misreconstructed $B^+ \to J/\psi(\to \ell^+\ell^-)K^+$ and $B^+ \to \psi(2S)(\to \ell^+\ell^-)K^+$ decays, where the kaon is misidentified as a lepton and the lepton (of the same electric charge) as a kaon. The particle-identification criteria used in the selection render such backgrounds negligible. Backgrounds from decays with a photon converted into an $e^+e^-$ pair are also negligible.

Combinatorial background is reduced using Boosted Decision Tree (BDT) algorithms [71], which employ the gradient boosting technique [72]. For the nonresonant muon mode and for each of the three different trigger categories of the nonresonant electron mode, a single BDT is trained for the 7 and 8 TeV data, and an additional BDT is trained for the 13 TeV data. The same BDTs are used to select the resonant decays. The BDT training uses nonresonant $K^+\ell^+\ell^-$ candidates selected from the data with $m(K^+\ell^+\ell^-) > 5.4$ GeV/$c^2$ as a proxy for the background, and simulated nonresonant $K^+\ell^+\ell^-$ candidates as a proxy for the signal decays. The training and testing is performed using the $k$-folding technique with $k = 10$ [73]. The variables used as input to these BDTs are: the $p_T$ of the $B^+$, $K^+$ and dilepton candidates, and the minimum and maximum $p_T$ of the leptons; the $B^+$, dilepton and $K^+ \chi^2_{IP}$ with respect to the associated PV, where $\chi^2_{IP}$ is defined as the difference in the vertex-fit $\chi^2$ of the PV reconstructed with and without the particle being considered; the minimum and maximum $\chi^2_{IP}$ of the leptons; the $B^+$ vertex-fit quality; the significance of the $B^+$ flight distance; and the angle between the $B^+$ candidate momentum vector and the direction between the associated PV and the $B^+$ decay vertex. The selection applied to the BDT output variables is chosen to maximise the predicted significance of the nonresonant signal yield. The BDT selection reduces the combinatorial background by approximately 99%, while retaining 85% of the signal modes. The efficiency of each BDT response is independent of $m(K^+\ell^+\ell^-)$ in the regions
used to determine the event yields.

An unbinned extended maximum-likelihood fit to the $m(K^+e^+e^-)$ and $m(K^+\mu^+\mu^-)$ distributions of nonresonant candidates is used to determine $R_K$. In order to take into account the correlation between the selection efficiencies, the different trigger categories and data-taking periods are fitted simultaneously. The resonant decay mode yields are incorporated as constraints in this fit, such that the $B^+ \to K^+\mu^+\mu^-$ yield and $R_K$ are fit parameters. The resonant yields are determined from separate unbinned extended maximum-likelihood fits to the $m_{J/\psi}(K^+\ell^+\ell^-)$ distributions. For all the mass-shape models described below, the parameters are derived from simulated decays that are calibrated using data control channels.

All four signal modes are modelled by functions with multi-Gaussian cores and power-law tails on both sides of the peak [74,75]. The electron-mode signal mass shapes are described with the sum of three distributions which model whether a bremsstrahlung photon cluster was added to neither, either or both of the $e^\pm$ candidates. The fraction of signal decays in each of the bremsstrahlung categories is constrained to the value obtained from the simulation.

The shape of the $B^+ \to J/\psi\pi^+$ background is taken from simulation, while its size is constrained with respect to the $B^+ \to J/\psi K^+$ mode using the known ratio of the relevant branching fractions [53,76] and efficiencies. In each trigger category, the shape and relative fraction of the background from partially reconstructed $B^{0,+} \to K^+(892)(0,+)\to K^+\pi^{(-,0)}e^+e^-$ or $B^{0,+} \to J/\psi (\to e^+e^-)K^+(892)(0,+)\to K^+\pi^{(-,0)}$ decays are also taken from simulation. The overall yield of these partially reconstructed decays is left free to vary in the fit, in order to accommodate possible lepton-universality violation in such decays. In the fits to nonresonant $K^+e^+e^-$ candidates, the shape of the radiative tail of $B^+ \to J/\psi (\to e^+e^-)K^+$ decays is taken from simulation and its yield is constrained to the expected value within its uncertainty. In all fits, the combinatorial background is modelled with an exponential function with a freely varying yield and shape.

In order to evaluate the efficiencies accurately, weights are applied to simulated candidates to correct for the imperfect modelling of the $B^+$ production kinematics, the particle-identification performance, and the trigger response. The weights are computed sequentially, making use of control samples of $J/\psi \to \mu^+\mu^-$, $D^+ \to D^0(\to K^-\pi^+)\pi^+$ and $B^+ \to J/\psi (\to \ell^+\ell^-)K^+$ decays, and are applied to both resonant and nonresonant simulated candidates. Only subsets of the $B^+ \to J/\psi (\to \ell^+\ell^-)K^+$ samples are used to derive these corrections, which minimises the number of common candidates being used for both the determination of the corrections and the measurement. The correlations between samples are taken into account in the results and cross-checks presented below. The overall effect of the corrections on the $R_K$ measurement is at the 0.02 level, demonstrating the robustness of the double-ratio method in suppressing systematic biases that affect the resonant and nonresonant decay modes similarly.

Two classes of systematic uncertainty are considered: those that only affect the nonresonant decay yields, and those that affect the ratio of efficiencies for different trigger categories and data-taking periods in the fit for $R_K$. The uncertainty from the choice of mass-shape models falls into the former category and is estimated by fitting pseudoexperiments with alternative models that still describe the data well. The effect on $R_K$ is at the $\pm0.01$ level. Systematic uncertainties in the latter category affect the ratios of efficiencies and hence the value of $R_K$ that maximises the likelihood. These
uncertainties are accounted for through constraints on the efficiency values used in the fit to determine $R_K$, taking into account the correlations between different trigger categories and data-taking periods. The combined statistical and systematic uncertainty is then determined from a profile-likelihood scan. In order to isolate the statistical contribution to the uncertainty, the profile-likelihood scan is repeated with the efficiencies fixed to their fitted values. For the subsamples of the electron-mode data where the trigger is based on the kaon or on other particles in the event that are not part of the signal candidate, the dominant systematic uncertainties come from the (data-derived) calibration of the trigger efficiencies. For the electron trigger, there are comparable contributions from the statistical uncertainties associated with various calibration samples and the calibration of data-simulation differences.

The migration of events in $q^2$ is studied in the simulation. The effect of the differing $q^2$ resolution between data and simulation, which alters the estimate of the migration, gives a negligible uncertainty in the determination of the ratio of efficiencies. The uncertainties on parameters used in the simulation decay decay model (Wilson coefficients, form factors, other hadronic uncertainties etc.) affect the $q^2$ distribution and hence the selection efficiencies determined from simulation. The variation caused by the uncertainties on these parameters is propagated to an uncertainty on $R_K$ using predictions from the FLAVIO software package [41]. The resulting systematic effect on $R_K$ is negligible, even when non-SM values of the Wilson coefficients are considered.

Several cross-checks are used to verify the analysis procedure. The single ratio $r_{J/\psi} = B(B^+ \to J/\psi(\to \mu^+\mu^-)K^+)/B(B^+ \to J/\psi(\to e^+e^-)K^+)$ is known to be compatible with unity at the 0.4% level [52,53]. This ratio does not benefit from the cancellation of systematic effects that the double ratio used to measure $R_K$ exploits, and is therefore a stringent test of the control of the efficiencies. The corrections applied to the simulation do not force $r_{J/\psi}$ to be unity and some of the corrections shift $r_{J/\psi}$ in opposing directions. The value of $r_{J/\psi}$ is found to be $1.014 \pm 0.035$, where the uncertainty includes the statistical uncertainty and those systematic effects relevant to the $R_K$ measurement. It does not include additional subleading systematic effects that should be accounted for in a complete measurement of $r_{J/\psi}$. As a further cross-check, the double ratio of branching fractions, $R_{K}^{(2S)}$, defined by

$$R_{K}^{(2S)} = \frac{B(B^+ \to \psi(2S)(\to \mu^+\mu^-)K^+) / B(B^+ \to J/\psi(\to \mu^+\mu^-)K^+)}{B(B^+ \to J/\psi(\to e^+e^-)K^+) / B(B^+ \to J/\psi(\to e^+e^-)K^+)}$$

is determined to be $0.986 \pm 0.013$, where again the uncertainty includes the statistical uncertainty but only those systematic effects that are relevant to the $R_K$ measurement. This ratio provides an independent validation of the analysis procedure.

Leptons from $B^+ \to J/\psi K^+$ decays have a different $q^2$ value than those from the nonresonant decay modes. However, the detector efficiency depends on laboratory-frame variables rather than on $q^2$, e.g. the momenta of the final-state particles, opening angles, etc. In these laboratory variables there is significant overlap between the nonresonant and resonant modes, even if the decays do not overlap in $q^2$ (see the Supplemental Material [70]). The $r_{J/\psi}$ ratio is examined as a function of a number of reconstructed variables. Any trend would indicate an uncontrolled systematic effect that would only partially cancel in the double ratio. For each of the variables examined, no significant trend is observed. Figure 1 shows one example and others are provided in the Supplemental Material [70]. Assuming the deviations that are observed indicate genuine mismodelling of the efficiencies, rather
than fluctuations, and taking into account the spectrum of the relevant variables in the nonresonant decay modes of interest, a total shift on $R_K$ is computed for each of the variables examined. In each case, the resulting variation is within the estimated systematic uncertainty on $R_K$. The $r_{J/\psi}$ ratio is also computed in two- and three-dimensional bins of the considered variables. Again, no trend is seen and the deviations observed are consistent with the systematic uncertainties on $R_K$. An example is shown in Fig. S7 in the Supplemental Material [70]. Independent studies of the electron reconstruction efficiency using control channels selected from the data also give consistent results.

The results of the fits to the $m(K^+\ell^+\ell^-)$ and $m_{J/\psi}(K^+\ell^+\ell^-)$ distributions are shown in Fig. 2. A total of 1943 ± 49 $B^+ \to K^+\mu^+\mu^-$ decays are observed. A study of the $B^+ \to K^+\mu^+\mu^-$ differential branching fraction gives results that are consistent with previous LHCb measurements [12] but, owing to the selection criteria optimised for the precision on $R_K$, are less precise. The $B^+ \to K^+\mu^+\mu^-$ differential branching fraction observed is consistent between the 7 and 8 TeV data and the 13 TeV data.

The value of $R_K$ is measured to be

$$R_K = 0.846^{+0.060+0.016}_{-0.054-0.014},$$

where the first uncertainty is statistical and the second systematic. This is the most precise measurement to date and is consistent with the SM expectation at the level of 2.5 standard deviations [21, 32, 35, 39, 41]. The likelihood profile as a function of $R_K$ is given in the Supplemental Material [70]. The value for $R_K$ obtained is consistent across
In summary, in the dilepton mass-squared region $1.1 < q^2 < 6.0$ GeV$^2$/c$^4$, the ratio of the branching fractions for $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^+ \rightarrow K^+ e^+ e^-$ decays is measured to be $R_K = 0.846^{+0.060}_{-0.054} \pm 0.016$. This is the most precise measurement of this ratio to date and is consistent with the SM prediction at the level of 2.5 standard deviations. Further
reduction in the uncertainty on $R_K$ can be anticipated when the data collected by LHCb in 2017 and 2018, which have a statistical power approximately equal to that of the full data set used here, are included in a future analysis. In the longer term, there are good prospects for high-precision measurements as much larger samples are collected with an upgraded LHCb detector [78].

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, FAPERJ and FINEP (Brazil); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); NWO (Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MSHE (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NSF (USA). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (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); ANR, Labex P2IO and OCEVU, and Région Auvergne-Rhône-Alpes (France); Key Research Program of Frontier Sciences of CAS, CAS PIFI, and the Thousand Talents Program (China); RFBR, RSF and Yandex LLC (Russia); GVA, XuntaGal and GENCAT (Spain); the Royal Society and the Leverhulme Trust (United Kingdom); Laboratory Directed Research and Development program of LANL (USA).

References


[31] Y. Wang and D. Atwood, *Rate difference between $b \to s\mu^+\mu^-$ and $b \to s e^+ e^-$ in supersymmetry with large $\tan \beta$*, Phys. Rev. D68 (2003) 094016, arXiv:hep-ph/0304248


[34] LHCb collaboration, R. Aaij et al., *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, JHEP 08 (2017) 055, arXiv:1705.05802


[44] A. Crivellin, G. D’Ambrosio, and J. Heeck, Explaining $h \to \mu^+\tau^+$, $B \to K^*\mu^+\mu^-$ and $B \to K\mu^+\mu^-/B \to K e^+ e^-$ in a two-Higgs-doublet model with gauged $L_\mu - L_\tau$, Phys. Rev. Lett. 114 (2015) 151801, arXiv:1501.00993.


[52] BESIII collaboration, M. Ablikim et al., Precision measurements of $B(\psi(3686) \to \pi^+\pi^- J/\psi)$ and $B(J/\psi \to \ell^+\ell^-)$, Phys. Rev. D88 (2013) 032007, arXiv:1307.1189.


[70] Supplemental Material to this Letter.


The two-dimensional distributions of \([m(K^+\ell^+\ell^-), q^2]\) for muon and electron candidates are shown in Fig. S1. For the muon sample, nonresonant candidates can be seen to accumulate in a vertical band around the \(B^+\) meson mass. For the electron candidates, only some of the bremsstrahlung energy is recovered by the procedure described in the Letter and this results in a worse mass resolution and a long tail to lower \(K^+e^+e^-\) masses. The vertical band of signal candidates is then more difficult to discern. The resonant signals from \(B^+ \rightarrow J/\psi (\rightarrow \ell^+\ell^-)K^+\) and \(B^+ \rightarrow \psi(2S)(\rightarrow \ell^+\ell^-)K^+\) decays are visible as diagonal bands, where the extended tails originate from both radiative and resolution effects, which are especially marked for the electron decay modes. As the energy loss affects both \(m(K^+\ell^+\ell^-)\) and \(q^2\) measurements, the angle of these bands is fixed and it is not possible for candidates to migrate into the bulk of the signal region in \([m(K^+\ell^+\ell^-), q^2]\).

For the electron mode, the lower radiative tail of \(B^+ \rightarrow J/\psi (\rightarrow e^+e^-)K^+\) decays enters the \(1.1 < q^2 < 6.0\,\text{GeV}^2/\text{c}^4\) region only at the lower part of the \(m(K^+e^-)\) fit range around \(4.9\,\text{GeV}/\text{c}^2\) (see also the left side of the \(B^+ \rightarrow K^+e^+e^-\) fit projection in Fig. 2 of the Letter).

The reconstructed properties of simulated decays are shown in Fig. S2. The distributions for resonant and nonresonant decays are similar and consequently the determination of the efficiency of each nonresonant decay with respect to its corresponding resonant decay results in the cancellation of systematic effects.

Figure S3 shows the \(m(K^+e^-)\) mass distribution for \(B^+ \rightarrow K^+e^+e^-\) signal decays and for several cascade background decays. For the mass reconstructed taking into account the bremsstrahlung correction, signal candidates are required to satisfy \(m(K^+e^-) > m_{D^0}\), suppressing the majority of cascade backgrounds to negligible levels. However, for cascade backgrounds involving \(D^0 \rightarrow K^+\pi^-\) decays, where the \(\pi^+\) is misidentified as an electron, the bremsstrahlung correction gives rise to a long tail of candidates with \(m(K^+e^-) > m_{D^0}\).

Figure S1: Two-dimensional distributions of \([m(K^+\ell^+\ell^-), q^2]\) for (left) muon and (right) electron candidates after the application of the pre-selection and trigger requirements but not the multivariate selection.
without the bremsstrahlung correction, i.e. these projections come from independent fits to each period/category. The nonresonant periods and trigger categories are shown in Fig. S4 (Fig. S5). For the resonant modes the remainder of the selection requirements.

Such decays are suppressed by placing an additional veto on the $K^+e^-$ mass reconstructed without the bremsstrahlung correction, i.e. based on the measured track momentum alone. This veto removes background around the known $D^0$ mass, as shown in Fig. S3. After the application of both these vetoes, the cascade backgrounds are reduced to a negligible level while retaining 97% of $B^+ \to K^+\mu^+\mu^-$ and 95% of $B^+ \to K^+e^+e^-$ decays passing the remainder of the selection requirements.

The fits to the nonresonant (resonant) decay modes divided into different data-taking periods and trigger categories are shown in Fig. S4 (Fig. S5). For the resonant modes these projections come from independent fits to each period/category. The nonresonant
Table S1: Total yields of the decay modes $B^+ \to K^+ e^+ e^-$, $B^+ \to K^+ \mu^+ \mu^-$, $B^+ \to J/\psi (\to e^+ e^-) K^+$ and $B^+ \to J/\psi (\to \mu^+ \mu^-) K^+$ obtained from the fits to the data.

<table>
<thead>
<tr>
<th>Decay Mode</th>
<th>Event Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^+ \to K^+ e^+ e^-$</td>
<td>766 ± 48</td>
</tr>
<tr>
<td>$B^+ \to K^+ \mu^+ \mu^-$</td>
<td>1 943 ± 49</td>
</tr>
<tr>
<td>$B^+ \to J/\psi (\to e^+ e^-) K^+$</td>
<td>344 100 ± 610</td>
</tr>
<tr>
<td>$B^+ \to J/\psi (\to \mu^+ \mu^-) K^+$</td>
<td>1 161 800 ± 1 100</td>
</tr>
</tbody>
</table>
Figure S4: Fit to the $m(K^+\ell^+\ell^-)$ invariant-mass distribution of nonresonant candidates in the (left) 7 and 8 TeV and (right) 13 TeV data samples. The top row shows the fit to the muon modes and the subsequent rows the fits to the electron modes triggered by (second row) one of the electrons, (third row) the kaon and (last row) by other particles in the event.
Figure S5: Fit to the $m_{J/ψ(K^{+}ℓ^{+}ℓ^{-})}$ invariant-mass distribution of resonant candidates in the (left) 7 and 8 TeV and (right) 13 TeV data samples. The top row shows the fit to the muon modes and the subsequent rows the fits to the electron modes triggered by (second row) one of the electrons, (third row) the kaon and (last row) by other particles in the event. Some large pulls are observed but have a negligible impact on the yields extracted.
Figure S6: (Top) distributions of the spectra of (left) the $B^+$ transverse momentum and (right) the minimum $p_T$ of the leptons. (Bottom) the single ratio $r_{J/\psi}$ relative to its average value $\langle r_{J/\psi} \rangle$ as a function of these variables.

Figure S7: (Left) the value of $r_{J/\psi}$, relative to the average value of $r_{J/\psi}$, measured in two-dimensional bins of the maximum lepton momentum ($p(l)$) and the opening angle between the two leptons ($\alpha(l^+, l^-)$). (Right) the bin definition in this two-dimensional space together with the distribution for $B^+ \to K^+ \mu^+ \mu^-$ ($B^+ \to J/\psi (\rightarrow e^+ e^-) K^+$) decays depicted as red (blue) contours.
Figure S8: Likelihood function from the fit to the data profiled as a function of $R_K$ (solid line). The blue dashed line depicts the expected shape of the likelihood profile if the uncertainties were Gaussian.

The profile likelihood for the fit is shown in Fig. S8. The likelihood is Gaussian to a reasonable approximation in the range $0.75 < R_K < 0.95$, but non-Gaussian effects can be seen outside of this range due to the comparatively low yield in the $B^+ \rightarrow K^+ e^+ e^-$ decay.

The $R_K$ values derived from a fit to just the 7 and 8 TeV data, and a fit to just the 13 TeV data are

\[
R_{K}^{7 \text{ and } 8 \text{ TeV}} = 0.717 \pm 0.083 \pm 0.017, \\
R_{K}^{13 \text{ TeV}} = 0.928 \pm 0.089 \pm 0.020,
\]

where the first set of uncertainties are statistical and the second systematic. The combination of these values, or a combination of the latter value with the previously published LHCb result [33], requires that correlations are properly taken into account, as is done in the simultaneous fit used to derive the $R_K$ measurement given in the main body of the Letter.
School of Physics and Astronomy, University of Manchester, Manchester, United Kingdom
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
Laboratory of Mathematical and Subatomic Physics, Constantine, Algeria, associated to
Pontificia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, Brazil, associated to
South China Normal University, Guangzhou, 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
Van Swinderen Institute, University of Groningen, Groningen, Netherlands, associated to
National Research Centre Kurchatov Institute, Moscow, Russia, associated to
National University of Science and Technology “MISIS”, Moscow, Russia, associated to
National Research University Higher School of Economics, 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
University of Michigan, Ann Arbor, United States, associated to
Los Alamos National Laboratory (LANL), Los Alamos, United States, 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, Kraków, Poland
m LIFAELS, La Salle, Universitat Ramon Llull, Barcelona, Spain
n Hanoi University of Science, Hanoi, Vietnam
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 H.H. Wills Physics Laboratory, University of Bristol, Bristol, United Kingdom
w MSU - Iligan Institute of Technology (MSU-IIT), Iligan, Philippines
x Novosibirsk State University, Novosibirsk, Russia
y Sezione INFN di Trieste, Trieste, Italy
z School of Physics and Information Technology, Shaanxi Normal University (SNNU), Xi’an, China
aa Physics and Micro Electronic College, Hunan University, Changsha City, China
ab Lanzhou University, Lanzhou, China
† Deceased