Test of lepton universality with $B^0 \rightarrow K^{*0}\ell^+\ell^-$ decays

The LHCb collaboration

Abstract

A test of lepton universality, performed by measuring the ratio of the branching fractions of the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ and $B^0 \rightarrow K^{*0}e^+e^-$ decays, $R_{K^{*0}}$, is presented. The $K^{*0}$ meson is reconstructed in the final state $K^+\pi^-$, which is required to have an invariant mass within 100 MeV/$c^2$ of the known $K^{*}(892)^0$ mass. The analysis is performed using proton-proton collision data, corresponding to an integrated luminosity of about 3 fb$^{-1}$, collected by the LHCb experiment at centre-of-mass energies of 7 and 8 TeV. The ratio is measured in two regions of the dilepton invariant mass squared, $q^2$, to be

$$R_{K^{*0}} = \begin{cases} 0.66^{+0.11}_{-0.07} \text{(stat)} \pm 0.03 \text{(syst)} & \text{for } 0.045 < q^2 < 1.1 \text{ GeV}^2/c^4, \\ 0.69^{+0.11}_{-0.07} \text{(stat)} \pm 0.05 \text{(syst)} & \text{for } 1.1 < q^2 < 6.0 \text{ GeV}^2/c^4. \end{cases}$$

The corresponding 95.4% confidence level intervals are [0.52, 0.89] and [0.53, 0.94]. The results, which represent the most precise measurements of $R_{K^{*0}}$ to date, are compatible with the Standard Model expectations at the level of 2.1–2.3 and 2.4–2.5 standard deviations in the two $q^2$ regions, respectively.
1 Introduction

In the Standard Model (SM) of particle physics, the electroweak couplings of leptons to
gauge bosons are independent of their flavour and the model is referred to as exhibiting
lepton universality (LU). Flavour-changing neutral-current (FCNC) processes, where a
quark changes its flavour without altering its electric charge, provide an ideal laboratory
to test LU. The SM forbids FCNCs at tree level and only allows amplitudes involving
electroweak loop (penguin and box) Feynman diagrams. The absence of a dominant
tree-level SM contribution implies that such transitions are rare, and therefore sensitive
to the existence of new particles. The presence of such particles could lead to a sizeable
increase or decrease in the rate of particular decays, or change the angular distribution
of the final-state particles. Particularly sensitive probes for such effects are ratios of the
type

\[ R_H = \frac{\int \frac{d\Gamma(B \to H\mu^+\mu^-)}{dq^2} \ dq^2}{\int \frac{d\Gamma(B \to He^+e^-)}{dq^2} \ dq^2}, \]

where \( H \) represents a hadron containing an \( s \) quark, such as a \( K \) or a \( K^* \) meson. The
decay rate, \( \Gamma \), is integrated over a range of the squared dilepton invariant mass, \( q^2 \). The
\( R_H \) ratios allow very precise tests of LU, as hadronic uncertainties in the theoretical
predictions cancel, and are expected to be close to unity in the SM [1–3]. At \( e^+e^- \) colliders operating at the \( \Upsilon(4S) \) resonance, the ratios \( R_K^*(\ell) \) have been
measured to be consistent with unity with a precision of 20 to 50\% [4,5]. More recently,
the most precise determination to date of \( R_K^{0} \) in the \( q^2 \) range between 1.0 and 6.0 GeV\(^2/c^4\) has been performed by the LHCb collaboration. The measurement has a relative precision
of 12\% [6] and is found to be 2.6 standard deviations lower than the SM expectation [1].
Hints of LU violation have been observed in \( B \to D^{(*)}\ell\nu \ell \) decays [7–9]. Tensions with
the SM have also been found in several measurements of branching fractions [10–12]
and angular observables [13,14] of rare \( b \to s \) decays. Models containing a new, neutral,
heavy gauge boson [15–20] or leptoquarks [21, 22] have been proposed to explain these
measurements.

A precise measurement of \( R_{K^0}^{0} \) can provide a deeper understanding of the nature
of the present discrepancies [23]. Some of the leading-order Feynman diagrams for the
\( B^{0} \to K^{0}\ell^+\ell^- \) decays, where \( \ell \) represents either a muon or an electron, are shown in
figure [1] for both SM and possible New Physics (NP) scenarios. If the NP particles
couple differently to electrons and muons, LU could be violated. The \( K^{*0} \) represents a
\( K^*(892)^0 \) meson, which is reconstructed in the \( K^+\pi^- \) final state by selecting candidates
within 100 MeV/c\(^2\) of the known mass [24]. No attempt is made to separate the \( K^{*0} \)
meson from S-wave or other broad contributions present in the selected \( K^+\pi^- \) region.
The S-wave fraction contribution to the \( B^0 \to K^{*0}\mu^+\mu^- \) mode has been measured by the
LHCb collaboration and found to be small [25]. Inclusion of charge-conjugate processes is
implied throughout the paper, unless stated otherwise. The analysis is performed in two
regions of \( q^2 \) that are sensitive to different NP contributions: a low-\( q^2 \) bin, between 0.045
and 1.1 GeV\(^2/c^4\), and a central-\( q^2 \) bin, between 1.1 and 6.0 GeV\(^2/c^4\). The lower boundary of
the low-\( q^2 \) region corresponds roughly to the dimuon kinematic threshold. The boundary
at 1.1 GeV\(^2/c^4\) is chosen such that \( \phi(1020) \to \ell^+\ell^- \) decays, which could potentially dilute
NP effects, are included in the low-\( q^2 \) interval. The upper boundary of the central-\( q^2 \)
Figure 1: Feynman diagrams in the SM of the $B^0 \to K^{*0} \ell^+ \ell^-$ decay for the (top left) electroweak penguin and (top right) box diagram. Possible NP contributions violating LU: (bottom left) a tree-level diagram mediated by a new gauge boson $Z'$ and (bottom right) a tree-level diagram involving a leptoquark $LQ$.

bin at $6.0 \text{ GeV}^2/c^4$ is chosen to reduce contamination from the radiative tail of the $J/\psi$ resonance.

The measurement is performed as a double ratio of the branching fractions of the $B^0 \to K^{*0} \ell^+ \ell^-$ and $B^0 \to K^{*0} J/\psi (\to \ell^+ \ell^-)$ decays

$$R_{K^{*0}} = \frac{\mathcal{B}(B^0 \to K^{*0} \mu^+ \mu^-)}{\mathcal{B}(B^0 \to K^{*0} J/\psi (\to \ell^+ \ell^-))} / \frac{\mathcal{B}(B^0 \to K^{*0} e^+ e^-)}{\mathcal{B}(B^0 \to K^{*0} J/\psi (\to e^+ e^-))},$$

where the two channels are also referred to as the “nonresonant” and the “resonant” modes, respectively. The experimental quantities relevant for the measurement are the yields and the reconstruction efficiencies of the four decays entering in the double ratio. Due to the similarity between the experimental efficiencies of the nonresonant and resonant decay modes, many sources of systematic uncertainty are substantially reduced. This helps to mitigate the significant differences in reconstruction between decays with muons or electrons in the final state, mostly due to bremsstrahlung emission and the trigger response. The decay $J/\psi \to \ell^+ \ell^-$ is measured to be consistent with LU [24]. In order to avoid experimental biases, a blind analysis was performed. The measurement is corrected for final-state radiation (FSR). Recent SM predictions for $R_{K^{*0}}$ in the two $q^2$ regions are reported in table 1. Note that possible uncertainties related to QED corrections are only included in Ref. [26], and these are found to be at the percent level. The $R_{K^{*0}}$ ratio is smaller than unity in the low-$q^2$ region due to phase-space effects.

The remainder of this paper is organised as follows: section 2 describes the LHCb detector, as well as the data and the simulation samples used; the experimental challenges in studying electrons as compared to muons are discussed in section 3; section 4 details
Table 1: Recent SM predictions for $R_{K^*0}$.

| $q^2$ range [GeV$^2$/c$^4$] | $R_{K^*0}^{\text{SM}}$ | References |
|----------------------------|-----------------------|------------|
| [0.045, 1.1]               | 0.906 ± 0.028         | BIP 26     |
|                            | 0.922 ± 0.022         | CDHMV 27–29|
|                            | 0.919 ± 0.004         | EOS 30–31  |
|                            | 0.925 ± 0.004         | flav.io 32–34 |
|                            | 0.920 ± 0.007         | JC 35      |
| [1.1, 6.0]                 | 1.000 ± 0.010         | BIP 26     |
|                            | 1.000 ± 0.006         | CDHMV 27–29|
|                            | 0.9968 ± 0.0005       | EOS 30–31  |
|                            | 0.9964 ± 0.005        | flav.io 32–34 |
|                            | 0.996 ± 0.002         | JC 35      |

how the simulation is adjusted in order to improve the modelling of the data; the selection of the candidates, rejection of the background and extraction of the yields are outlined in sections 5, 6 and 7; section 8 discusses the efficiency determination; the cross-checks performed and the systematic uncertainties associated with the measurement are summarised in sections 9 and 10, respectively; the results are presented in section 11; and section 12 presents the conclusions of the paper.

2 The LHCb detector and data set

The LHCb detector [36,37] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed to study particles containing $b$ or $c$ quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the $pp$ interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of momentum, $p$, with a relative uncertainty that varies from 0.5% at low values to 1.0% at 200 GeV/c. The minimum distance of a track to a primary vertex (PV), the impact parameter (IP), is measured with a resolution of $(15 + 29/p_T) \mu m$, where $p_T$ is the component of the momentum transverse to the beam, in GeV/c. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL). Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers.

The trigger system 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. The hardware muon trigger selects events containing at least one muon.
with significant $p_T$ (from $\sim 1.5$ to $\sim 1.8\text{ GeV} / c$, depending on the data-taking period). The hardware electron trigger requires the presence of a cluster of calorimeter cells with significant transverse energy, $E_T$, (from $\sim 2.5$ to $\sim 3.0\text{ GeV}$, depending on the data-taking period) in the ECAL. The hardware hadron trigger requires the presence of an energy deposit with $E_T$ above $\sim 3.5\text{ GeV}$ in the calorimeters. The software trigger requires a two-, three- or four-track secondary vertex, with a significant displacement from the PV. At least one charged particle must have significant $p_T$ and be inconsistent with originating from any PV. A multivariate algorithm [38] is used for the identification of secondary vertices consistent with the decay of a $b$ hadron.

The analysis is based on $pp$ collision data collected with the LHCb detector at centre-of-mass energies of 7 and 8 TeV during 2011 and 2012, and corresponding to an integrated luminosity of about $3\text{ fb}^{-1}$. Samples of simulated $B^0 \rightarrow K^{*0} \mu^+ \mu^-$, $B^0 \rightarrow K^{*0} e^+ e^-$, $B^0 \rightarrow K^{*0} J/\psi (\rightarrow \mu^+ \mu^-)$ and $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ events are used to determine the efficiency to trigger, reconstruct and select signal events, as well as to model the shapes used in the fits for signal candidates. In addition, specific simulated samples are utilised to estimate the contributions from backgrounds and to model their mass distributions. The $pp$ collisions are generated using PYTHIA [39] with a specific LHCb configuration [40]. Decays of hadronic particles are described by EvtGen [41], in which FSR is generated using Photos [42], which is observed to agree with a full QED calculation at the level of $\sim 1\%$ [26]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [43] as described in Ref. [44].

3 Electron reconstruction effects

The experimental environment in which the LHCb detector operates leads to significant differences in the treatment of decays involving muons or electrons in the final state. The two types of leptons behave differently when travelling through the detector material. Electrons emit a much larger amount of bremsstrahlung which, if not accounted for, would result in a significant degradation of the momentum resolution and consequently in a degradation of the $B$ mass resolution. If the radiation occurs downstream of the dipole magnet, the photon energy is deposited in the same calorimeter cell as that of the lepton, and the momentum of the electron is correctly measured. If the photons are emitted upstream of the magnet, the electron and photon deposit their energy in different calorimeter cells, and the electron momentum is evaluated after bremsstrahlung emission. However, for both types of emissions, the ratio of the energy detected in the ECAL to the momentum measured by the tracking system, an important variable to identify electrons, remains unbiased.

A dedicated bremsstrahlung recovery procedure is used to improve the electron momentum reconstruction. Searches are made within a region of the ECAL defined by the extrapolation of the electron track upstream of the magnet for energy deposits with $E_T > 75\text{ MeV}$ that are not associated with charged tracks. Such “bremsstrahlung clusters” are added to the measured electron momentum. If the same cluster can be associated with both the $e^+$ and the $e^-$, its energy is added to one of the two electrons at random. In $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ decays, one bremsstrahlung cluster is added to either electron of the pair in about half of the cases; the remaining half is equally split between cases when no bremsstrahlung cluster is found, or two or more clusters are added. These fractions are
reproduced well by the simulation and depend only weakly on $q^2$. The bremsstrahlung recovery procedure is limited in three ways: the energy threshold of the clusters that are added; the calorimeter acceptance and resolution; and the presence of energy deposits wrongly interpreted as bremsstrahlung clusters. These limitations degrade the resolution of the reconstructed invariant masses of both the dielectron pair and the $B$ candidate.

Since the occupancy of the calorimeters is significantly higher than that of the muon stations, the constraints on the trigger rate require that higher thresholds are imposed on the electron $E_T$ than on the muon $p_T$. In the central-$q^2$ region the higher threshold causes a loss of about half of the electron signal. The efficiency decreases slightly at lower $q^2$ values. To partially mitigate this effect, decays with electrons in the final state can also be selected through the hadron hardware trigger, using clusters associated with the $K^{*0}$ decay products, or by any hardware trigger from particles in the event that are not associated with the signal candidate.

In decays with electrons, since the mass resolution of the reconstructed $B$ candidate is worse than in final states with muons, the background contamination in the signal region is larger. The level of combinatorial background, arising from the accidental association of particles produced by different $b$- and $c$-hadron decays, is also higher in such channels, due to a larger number of electron candidates. As a result, the discriminating power of the fits to extract the signal yields is reduced (see section 7). Differences due to bremsstrahlung and the trigger response lead to a reconstruction efficiency for the $B^{0} \to K^{*0}J/\psi (\to \mu^{+}\mu^{-})$ decays that is about five times smaller than for the $B^{0} \to K^{*0}J/\psi (\to e^{+}e^{-})$ decays.

4 Corrections to the simulation

In order to optimise the selection criteria and accurately evaluate the efficiencies, a set of corrections is determined from unbiased control samples selected from the data. The procedure is applied to the simulated samples of the nonresonant and resonant modes.

The first correction accounts for differences between simulation and data in the particle identification (PID) performance [45]. The PID efficiencies are directly measured using a tag-and-probe method on high-purity data samples of pions and kaons from $D^{*+} \to D^{0}(\to K^{-}\pi^{+})\pi^{+}$ decays. Similarly, the electron and muon identification efficiencies are obtained from $B^{+} \to K^{+}J/\psi (\to \mu^{+}\mu^{-})$ decays. Corrections are determined as a function of the track momentum and pseudorapidity.

The second step of the procedure adjusts the simulation for the charged-track multiplicity in the event, which is not described well in simulation. A small correction for the $B^0$ kinematics is also applied. Resonant $B^0 \to K^{*0}J/\psi (\to \mu^{+}\mu^{-})$ decays are used since the muon triggers are observed to be well modelled in simulation.

The third step corrects the simulation of the trigger response for both the hardware and software levels using a tag-and-probe technique. Whenever possible, $B^0 \to K^{*0}J/\psi (\to \mu^{+}\mu^{-})$ decays are used as a control sample in place of $B^0 \to K^{*0}J/\psi (\to e^{+}e^{-})$ decays in order to take advantage of the larger sample size. In such cases, the two decays are compared and found to give consistent results. The tag sample is defined by events where the hardware trigger is fired by activity in the event not associated with any of the signal decay particles. Alternatively, when probing the leptonic (hadronic) hardware triggers, the tag is required to have triggered the hadronic (leptonic) hardware trigger. The corrections for the leptonic hardware triggers are parameterised...
as a function of the cluster $E_T$ or track $p_T$. The hadron hardware trigger efficiency is known to be sensitive to tracks overlapping in the HCAL, however, a good description can be obtained when the efficiency is measured as a function of the $p_T$ of the $K^+\pi^-$ pair instead of the kaon or the pion independently. Corrections are determined separately in the different calorimeter regions [36], in order to take into account potential differences due to different occupancies. When the hardware trigger is fired by activity in the event not associated with any of the signal decay particles, the correction is determined as a function of the $B^0$ $p_T$ and the charged-track multiplicity in the event in order to take into account correlations in the production between the two $b$ hadrons in the event. For the software trigger, the corrections are determined as a function of the minimum $p_T$ of the $B^0$ decay products.

Finally, residual differences between data and simulation in the reconstruction performance are accounted for using $B^0 \rightarrow K^{*0} J/\psi (\rightarrow \ell^+ \ell^-)$ candidates to which the full selection is applied, as well as additional requirements to further reduce the background contamination. The corrections are determined by matching the distribution of the $B^0$ kinematics and vertex fit quality in simulation to the data, separately for muon and electron samples.

The correction factors are determined sequentially as histograms, with the previous corrections applied before deriving the subsequent one. To avoid biases in the procedure due to common candidates being used for both the determination of the corrections and the measurement, a $k$-folding [46] approach with $k = 10$ is adopted. To dilute the dependence on the choice of the binning schemes, all corrections are linearly interpolated between adjacent bins. After all the corrections are applied to the simulation, a very good agreement with the data is obtained.

5 Selection of signal candidates

A $B^0$ candidate is formed from a pair of well-reconstructed oppositely charged particles identified as either muons or electrons, combined with two well-reconstructed oppositely charged particles, one identified as a kaon and the other as a pion. The $K^+\pi^-$ invariant mass is required to be within 100 MeV/$c^2$ of the known $K^{*0}$ mass. The kaon and pion must have $p_T$ exceeding 250 MeV/$c$, while for the muons (electrons) $p_T > 800$ ($500$) MeV/$c$ is required. Only dilepton pairs with a good-quality vertex are used to form signal candidates. The $K^{*0}$ meson and $\ell^+\ell^-$ pair are required to originate from a common vertex in order to form a $B^0$ candidate. When more than one PV is reconstructed, the one with the smallest $\chi^2_{IP}$ is selected, where $\chi^2_{IP}$ is the difference in $\chi^2$ of a given PV reconstructed with and without the considered $B^0$ candidate. With respect to this selected PV, the impact parameter of the $B^0$ candidate is required to be small, its decay vertex significantly displaced, and the momentum direction of the $B^0$ is required to be consistent with its direction of flight. This direction is given by the vector between the PV and decay vertex. The distribution of $q^2$ as a function of the four-body invariant mass for the $B^0$ candidates is shown in figure 2 for both muon and electron final states. The requirements on the neural-network classifier and $m_{corr}$ (see section 5) are not applied. In each plot, the contributions due to the charmonium resonances are clearly visible at the $J/\psi$ and $\psi(2S)$ masses. For electrons, these distributions visibly extend above the nominal mass values due to the calorimeter resolution affecting the bremsstrahlung recovery procedure.
(see section 3). The empty region in the top left corresponds to the kinematic limit of the $B^0 \to K^{*0} \ell^+ \ell^-$ decay, while the empty region in the top right corresponds to the requirement that rejects the $B^+ \to K^+ \ell^+ \ell^-$ background (see section 6).

The $B^0$ mass resolution and the contributions of signal and backgrounds depend on the way in which the event was triggered. The data sample of decay modes involving an $e^+e^-$ pair is therefore divided into three mutually exclusive categories, which in order of precedence are: candidates for which one of the electrons from the $B^0$ decay satisfies the hardware electron trigger (L0E), candidates for which one of the hadrons from the $K^{*0}$ decay meets the hardware hadron trigger (L0H) requirements, and candidates triggered by activity in the event not associated with any of the signal decay particles (L0I). For $B^0 \to K^{*0} \mu^+ \mu^-$ candidates, at least one of the two leptons must satisfy the requirements of the hardware muon trigger.

For the $B^0 \to K^{*0} J/\psi (\to \mu^+ \mu^-)$ decay mode, a dimuon mass interval within 100 MeV/$c^2$ of the known $J/\psi$ mass is selected to identify candidates. It is not possible to apply a tight $q^2$ requirement to identify the $B^0 \to K^{*0} J/\psi (\to e^+ e^-)$ mode as, despite the bremsstrahlung recovery, the $e^+ e^-$ invariant mass distribution has a long radiative tail towards low values. This tail can be seen in figure 2. The $q^2$ interval used to select $B^0 \to K^{*0} J/\psi (\to e^+ e^-)$ candidates is between 6.0 and 11.0 GeV/$c^2$, with the lower limit corresponding to the upper boundary of the central-$q^2$ bin.

The separation of the signal from the combinatorial background is based on neural-network classifiers [47]. The same classifier is used for the resonant and nonresonant modes, but muon and electron channels are treated separately. The classifiers are trained using simulated $B^0 \to K^{*0} \ell^+ \ell^-$ decays, which have been corrected for known differences between data and simulation (see section 4), to represent the signal. Data candidates with $K^+ \pi^- \ell^+ \ell^-$ invariant masses larger than 5400 MeV/$c^2$ and 5600 MeV/$c^2$ are used to represent background samples for the muon and electron channel, respectively. To best exploit the size of the available data sample for the training procedure, a $k$-folding technique [46] is adopted with $k = 10$. The variables used as input to the classifiers are: the transverse momentum, the quality of the vertex fit, the $\chi^2_P$, the $\chi^2_{VD}$ (the $\chi^2$ on the measured distance between the PV and the decay vertex), and the angle between the direction of flight and the momentum of the $B^0$ candidate, the $K^+ \pi^-$ and the dilepton.

**Figure 2:** Number of candidates for $B^0 \to K^{*0} \ell^+ \ell^-$ final states with (left) muons and (right) electrons as a function of the dilepton invariant mass squared, $q^2$, and the four-body invariant mass of the $B^0$. 

![Image of Figure 2](image-url)
pairs; the minimum and maximum of the kaon and pion $p_T$; and of their $\chi^2_{IP}$; the minimum and maximum of the lepton $p_T$ values, and of their $\chi^2_{IP}$; and finally, the most discriminating variable, the quality of the kinematic fit to the decay chain (this fit is performed with a constraint on the vertex that requires the $B^0$ candidate to originate from the PV). In each fold, only variables that significantly improve the discriminating power of the classifier are kept.

For the muon modes, a requirement on the four-body invariant mass of the $B^0$ candidate to be larger than 5150 MeV/$c^2$ excludes backgrounds due to partially reconstructed decays, $B \to K^{*0}\mu^+\mu^- X$, where one or more of the products of the $B$ decay, denoted as $X$, are not reconstructed. A kinematic fit that constrains the dielectron mass to the known $J/\psi$ mass allows the corresponding background to be separated from the $B^0 \to K^{*0}J/\psi (\to e^+e^-)$ signal by requiring the resulting four-body invariant mass to be at least 5150 MeV/$c^2$. For the nonresonant electron mode, the partially reconstructed backgrounds can be reduced by exploiting the kinematics of the decay. The ratio of the $K^{*0}$ and the dielectron momentum components transverse to the $B^0$ direction of flight is expected to be unity, unless the electrons have lost some energy due to bremsstrahlung that was not recovered (see figure 3). In the approximation that bremsstrahlung photons do not modify the dielectron direction significantly, which is particularly valid for low dilepton masses, this ratio can be used to correct the momentum of the dielectron pair. The invariant mass of the signal candidate calculated using the corrected dielectron momentum, $m_{corr}$, has a poor resolution that depends on $\chi^2_{VD}$. Nevertheless, since the missing momentum of background candidates does not originate from the dielectron pair, $m_{corr}$ still acts as a useful discriminating variable. Signal and partially reconstructed backgrounds populate different regions of the two-dimensional plane defined by $m_{corr}$ and $\chi^2_{VD}$ (see figure 4). The requirements in this plane and on the classifier response are optimised simultaneously, but separately for each $q^2$ region. The optimisation maximises a figure of merit defined as $N_S/\sqrt{N_S + N_B}$, where the expected signal yield, $N_S$, is evaluated by scaling the observed number of $B^0 \to K^{*0}J/\psi (\to \ell^+\ell^-)$ candidates by the ratio of the branching fractions of the nonresonant and resonant modes, and the expected background yield, $N_B$, is obtained by fitting the mass sidebands in data.

After the full selection, 1 to 2% of the events contain multiple candidates. This fraction is consistent between the resonant and nonresonant modes, and between final states with electrons and muons. About half of the multiple candidates are due to cases where the kaon is misidentified as the pion and vice versa. In all cases only one candidate, chosen randomly, is retained.

6 Exclusive backgrounds

Specific requirements are applied to reject backgrounds from $b$-hadron decays, while ensuring a negligible loss of signal, as verified using simulation. In the low-$q^2$ region, the size of the contamination from $B^0 \to K^{*0}V(\to \ell^+\ell^-)$ decays, where $V$ is a $\rho$, $\omega$ or $\phi$ meson, is evaluated in Refs. [48,49]. The contamination due to direct decays or interference with the signal channel is found to be smaller than 2% and similar for muons and electrons. As a consequence, the residual effect in the double ratio is expected to be very small and can therefore be safely neglected.

Misreconstructed $B^0 \to K^{*0}J/\psi (\to \mu^+\mu^-)$ and $B^0 \to K^{*0}\psi(2S)(\to \mu^+\mu^-)$ decays can
Figure 3: Sketch of the topology of a $B^0 \rightarrow K^{*0}e^+e^-$ decay. The transverse momentum lost via bremsstrahlung is evaluated as the difference between the $p_T$ of the $K^{*0}$ meson and that of the dielectron system, where both are calculated with respect to the $B^0$ meson direction of flight. Bremsstrahlung photons that are not recovered by the reconstruction are assumed to follow the dielectron momentum direction.

Figure 4: Distribution of $\chi^2_{VD}$ as a function of $m_{corr}$ for (left) $B^0 \rightarrow K^{*0}e^+e^-$ and (right) $B \rightarrow X(\rightarrow YK^{*0})e^+e^-$ simulated candidates. The distributions are normalised to the same number of candidates.

contaminate the signal region if the identities of one of the hadrons and one of the muons are swapped. To avoid this, the invariant mass of the hadron candidate (under the muon mass hypothesis) and the oppositely charged muon is required to be outside of a 60 MeV/c$^2$ interval around the known $J/\psi$ or the $\psi(2S)$ masses.

A large, nonpeaking background comes from the $B^0 \rightarrow D^-\ell^+\nu$ decay, with $D^- \rightarrow K^{*0}\ell^-\bar{\nu}_\ell$, which has a branching fraction four orders of magnitude larger than that of the signal. In the rare case where both neutrinos have low energies, the signal selection will be less effective at rejecting this background. This decay can be separated from the signal by exploiting the angular distribution of the dilepton pair. For $B^0 \rightarrow D^-\ell^+\nu$ decays, the angle $\theta_{\ell}$ between the direction of the $\ell^+$ in the dilepton rest frame and the direction of the dilepton in the $B^0$ rest frame tends to be small. This background is suppressed by requiring $|\cos\theta_{\ell}| < 0.8$.

When combined with a low-momentum $\pi^-$ meson from the rest of the event, $B^+ \rightarrow K^+\ell^+\ell^-$ decays can pass the selection and populate the upper mass sideband.
region that is used to represent the combinatorial background for the training of the neural-network classifiers. Such decays are vetoed by requiring the invariant mass of the \(K^+\ell^+\ell^-\) combination to be less than 5100 MeV/c^2. Candidates where the \(\pi^-\) from the \(K^{*0}\) is misidentified as a kaon and paired with a \(\pi^+\) are similarly rejected. To suppress background from \(B_0^+\to \phi\ell^+\ell^-\) decays, with \(\phi \to K^+K^-\) where one of the kaons is misidentified as a pion, the invariant mass of the two hadrons computed under the \(K^+K^-\) mass hypothesis is required to be larger than 1040 MeV/c^2.

7 Fits to the \(K^+\pi^-\ell^+\ell^-\) invariant mass distributions

The signal yields are determined using unbinned extended maximum likelihood fits to the four-body invariant mass, \(m(K^+\pi^-\ell^+\ell^-)\), of the selected candidates in each \(q^2\) interval and for each lepton type. The reconstructed invariant mass is calculated using a kinematic fit with a constraint on the vertex that requires the \(B^0\) candidate to originate from the PV. In order to improve the quality and stability of the results, the fits are performed simultaneously on the nonresonant and resonant modes, and some parameters are shared.

For the muon channel, the fit is performed in an invariant mass window of 5150–5850 MeV/c^2. The low edge is chosen to reject the partially reconstructed background that populates the low mass region. The probability density function (PDF) for the signal is defined by a Hypatia function [50], where the parameters are fixed from simulation. However, in order to account for possible residual discrepancies with data, the mean and width are allowed to vary freely in the fit, independently for the resonant and nonresonant modes and in each \(q^2\) region. The combinatorial background is parameterised using an exponential function, which has a different slope in the resonant and nonresonant modes, and in each \(q^2\) region, that is free to vary in the fit. For the resonant mode, two additional sources of background are included: \(\Lambda^0_b \to K^+pJ/\psi(\to \mu^+\mu^-)\) decays, where the \(p\) candidate is misidentified as a \(\pi^-\) meson, and \(B^0_s \to K^{*0}J/\psi(\to \mu^+\mu^-)\) decays. The former are described using a kernel estimation technique [51] applied to simulated events for which the \(K^+\pi^-\) invariant mass distribution has been matched to data from Ref. [52]. The latter are modelled using the same PDF as for the signal, but with the mean value shifted by the known difference between the \(B^0\) and the \(B^0_s\) masses. The equivalent backgrounds to the nonresonant mode are found to be negligible.

For the electron channel, due to the limited resolution on the \(K^+\pi^-e^+e^-\) invariant mass, a wider window of 4500–6200 MeV/c^2 is used. The resolution on the reconstructed invariant mass of the \(B^0\) and the background composition depends on the kinematics of the decay, as well as on the trigger category. For this reason, simultaneous fits to the four-body invariant mass of the \(B^0 \to K^{*0}J/\psi(\to e^+e^-)\) and \(B^0 \to K^{*0}e^+e^-\) channels are performed separately in the three trigger categories. Following the strategy of Ref. [6], the \(K^+\pi^-e^+e^-\) signal PDF is observed to depend on the number of calorimeter clusters that are added to the dielectron candidate in order to correct for the effects of bremsstrahlung. Three bremsstrahlung categories are considered, depending on whether zero, one or more clusters are recovered. The PDF is described by the sum of a Crystal Ball function [53] (CB) and a wide Gaussian function. The CB function accounts for FSR and bremsstrahlung that is not fully recovered, and corresponds to over 90% of the total signal PDF. Cases where bremsstrahlung clusters were incorrectly associated are accounted for by the Gaussian function. The shape parameters and the fraction of candidates in each bremsstrahlung
category are taken from simulation, the latter having been checked on data control channels (see figure 5). In order to account for possible data-simulation discrepancies, the mean (width) of the PDF for each trigger category is allowed to shift (scale). These shift and scale factors are common between the nonresonant and resonant PDFs. An additional scale factor is also applied to the parameter describing the tail of the CB functions. The combinatorial background is described by an exponential function with different slope parameters for the resonant and nonresonant modes, and in each trigger category and $q^2$ region, that are free to vary in the fit. The shape of the partially reconstructed hadronic background, $B \rightarrow X(\rightarrow Y K^{*0}) e^+ e^-$ (where the decay product $Y$ is not reconstructed), is obtained from simulation using a sample that includes decays of higher kaon resonances, $X$, such as $K_1^+(1270)$ and $K_2^+(1430)$. The mass distribution is modelled using a kernel estimation technique separately in each trigger category and $q^2$ region. The fraction of this background is free to vary in both $q^2$ intervals. Due to the requirement on the four-body invariant mass with a $J/\psi$ mass constraint (see section 5), there is no partially reconstructed background left to contaminate $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ candidates. Due to the long radiative tail of the dielectron invariant mass, $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ decays can contaminate the central-$q^2$ region and an additional background component is considered (see figure 2), however this contribution does not peak at the nominal $B^0$ mass. The distribution is modelled using simulated events, while the normalisation is constrained using a mixture of data and simulation. The contributions to the resonant modes from $\Lambda_b^0 \rightarrow K^+ p J/\psi (\rightarrow e^+ e^-)$ and $\bar{B}^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ decays are treated following the same procedure as for the muon channel. The normalisations are fixed to the yields returned by the muon fit after correcting for efficiency differences between the two final states.

The results of the fits to the muon channels are shown in figure 6, while figure 7 displays the fit results for the electron channels, where the three trigger categories have been combined. The distribution of the normalised fit residuals of the $B^0 \rightarrow K^{*0} J/\psi (\rightarrow \mu^+ \mu^-)$ mode shows an imperfect description of the combinatorial background at high mass values, although the effect on the signal yield is negligible. The resulting yields are listed in table 2.
Table 2: Yields obtained from the mass fits to the muon and electron (in the three trigger categories) channels. The uncertainties are statistical only.

|                  | \( B^0 \rightarrow K^{*0} \ell^+ \ell^- \) | \( B^0 \rightarrow K^{*0} J/\psi \rightarrow \ell^+ \ell^- \) |
|------------------|------------------------------------------|-------------------------------------------------|
|                  | low-\( q^2 \) | central-\( q^2 \) | central-\( q^2 \) | central-\( q^2 \) |
| \( e^+ e^- \) (L0E) | 285 ± 18/18 | 353 ± 21/21 | 274416 ± 602/654 |
| \( e^+ e^- \) (L0H) | 67 ± 10/10 | 19 ± 5/5 | 43468 ± 222/221 |
| \( e^+ e^- \) (L0I) | 25 ± 7/6 | 11505 ± 115/114 |

8 Efficiencies

The efficiency for selecting each decay mode is defined as the product of the efficiencies of the geometrical acceptance of the detector, the complete reconstruction of all tracks, the trigger requirements and the full set of kinematic, PID and background rejection requirements. All efficiencies are determined using simulation that is tuned to data, as described in section 4, and account for bin migration in \( q^2 \) due to resolution, FSR and bremsstrahlung in the detector. The net bin migration amounts to about 1% and 5% in
which is expected to be equal to unity. This quantity represents an extremely stringent
test, as it does not benefit from the large cancellation of the experimental systematic
fractions of the muon and electron resonant channels

The efficiency ratios between the nonresonant and the resonant modes, $\varepsilon_{\ell^+\ell^-}/\varepsilon_{J/\psi(\ell^+\ell^-)}$, which directly enter in the $R_{K^{*0}}$ measurement, are reported in table 3. Besides a dependence on the kinematics, the difference between the ratios in the two $q^2$ regions is almost entirely due to the different requirement on the neural-network classifier. The relative fraction of the electron trigger categories is checked using simulation to depend on $q^2$ as expected: the fraction of L0E decreases when decreasing in $q^2$, while L0H increases; on the other hand, the fraction of L0I only mildly depends on $q^2$.

9 Cross-checks

A large number of cross-checks were performed before unblinding the result. The control
of the absolute scale of the efficiencies is tested by measuring the ratio of the branching
fractions of the muon and electron resonant channels

$$r_{J/\psi} = \frac{\mathcal{B}(B^0 \rightarrow K^{*0}J/\psi(\rightarrow \mu^+\mu^-))}{\mathcal{B}(B^0 \rightarrow K^{*0}J/\psi(\rightarrow e^+e^-))},$$

which is expected to be equal to unity. This quantity represents an extremely stringent
test, as it does not benefit from the large cancellation of the experimental systematic
Table 3: Efficiency ratios between the nonresonant and resonant modes, $\varepsilon_{\ell^+\ell^-}/\varepsilon_{J/\psi(\ell^+\ell^-)}$, for the muon and electron (in the three trigger categories) channels. The uncertainties are statistical only.

|               | low-$q^2$ | central-$q^2$ |
|---------------|-----------|---------------|
| $\mu^+\mu^-$ | 0.679 ± 0.009 | 0.584 ± 0.006 |
| $e^+e^-$ (L0E) | 0.539 ± 0.013 | 0.522 ± 0.010 |
| $e^+e^-$ (L0H) | 2.252 ± 0.098 | 1.627 ± 0.066 |
| $e^+e^-$ (L0I) | 0.789 ± 0.029 | 0.595 ± 0.020 |

effects provided by the double ratio. The $r_{J/\psi}$ ratio is measured to be $1.043 ± 0.006 ± 0.045$, where the first uncertainty is statistical and the second systematic. The same sources of systematic uncertainties as in the $R_{K^{*0}}$ measurement are considered (see section 10). The result, which is in good agreement with unity, is observed to be compatible with being independent of the decay kinematics, such as $p_T$ and $\eta$ of the $B^0$ candidate and final-state particles, and the charged-track multiplicity in the event.

The extent of the cancellation of residual systematics in $R_{K^{*0}}$ is verified by measuring a double ratio, $R_{\psi(2S)}$, where $B^0 \rightarrow K^{*0}\psi(2S)(\rightarrow \ell^+\ell^-)$ decays are used in place of $B^0 \rightarrow K^{*0}\ell^+\ell^-$. The $R_{\psi(2S)}$ ratio, measured with a statistical precision of about 2%, is found to be compatible with unity within one standard deviation.

The branching fraction of the decay $B^0 \rightarrow K^{*0}\mu^+\mu^-$ is measured and found to be in good agreement with Ref. [25]. Furthermore, the branching fraction of the $B^0 \rightarrow K^{*0}\gamma$ decay, where decays with a photon conversion are used, is determined with a statistical precision of about 7% and is observed to be in agreement with the expectation within two standard deviations. The $B^0 \rightarrow K^{*0}\gamma(\rightarrow e^+e^-)$ selection and determination of the signal yield closely follows that of the $B^0 \rightarrow K^{*0}e^+e^-$ decay.

If no correction is made to the simulation, the ratio of the efficiencies changes by less than 5%. The relative population of the three bremsstrahlung categories is compared between data and simulation using both $B^0 \rightarrow K^{*0}J/\psi(\rightarrow e^+e^-)$ and $B^0 \rightarrow K^{*0}\gamma(\rightarrow e^+e^-)$ candidates to test possible $q^2$ dependence of the modelling. Good agreement is observed, as shown in figure 5.

The sPlot technique [54], where $m(K^+\pi^-\ell^+\ell^-)$ is used as the discriminating variable, is adopted to subtract statistically the background from the $B^0 \rightarrow K^{*0}\ell^+\ell^-$ selected data, and test the agreement between muons and electrons, data and simulation, using several control quantities (see figure 5): the $q^2$ distributions show good agreement in both $q^2$ regions; a clear $K^{*0}$ peak is visible in the $K^+\pi^-$ invariant mass distributions, and the muon and electron channels show good agreement; while the distribution of the opening angle between the two leptons in the central-$q^2$ region are very similar between the muon and electron channels, this is not the case at low-$q^2$ due to the difference in lepton masses; the distribution of the distance between the $K^+\pi^-$ and $\ell^+\ell^-$ vertices shows that the pairs of hadrons and leptons consistently originate from the same decay vertex.
Figure 8: (hatched) Background-subtracted distributions for (darker colour) $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and (lighter colour) $B^0 \rightarrow K^{*0} e^+ e^-$ candidates, compared to (full line) simulation. From top to bottom: $q^2$, $K^+ \pi^-$ invariant mass, $m(K^+ \pi^-)$, opening angle between the two leptons, $\theta_{\text{lepton}}$, and projection along the beam axis of the distance between the $K^+ \pi^-$ and $\ell^+ \ell^-$ vertices, $\Delta z_{\text{vertex}}$. The distributions are normalised to unity. The hatched areas correspond to the statistical uncertainties only. The data are not efficiency corrected.
Table 4: Systematic uncertainties on the $R_{K^*0}$ ratio for the three trigger categories separately (in percent). The total uncertainty is the sum in quadrature of all the contributions.

| Trigger category                  | Low-$q^2$ | Central-$q^2$ |
|----------------------------------|-----------|---------------|
| Corrections to simulation        | L0E  2.5  | L0H  3.9  | L0I  2.2  | L0E  2.2  | L0H  3.4  | L0I  3.4  |
| Trigger                          | L0E  0.1  | L0H  0.1  | L0I  0.2  | L0E  0.2  | L0H  0.8  | L0I  0.2  |
| PID                              | L0E  0.2  | L0H  0.4  | L0I  0.3  | L0E  0.2  | L0H  1.0  | L0I  0.5  |
| Kinematic selection              | L0E  2.1  | L0H  2.1  | L0I  2.1  | L0E  2.1  | L0H  2.1  | L0I  2.1  |
| Residual background              | L0E  –    | L0H  –    | L0I  –    | L0E  5.0  | L0H  5.0  | L0I  5.0  |
| Mass fits                        | L0E  1.4  | L0H  2.1  | L0I  2.5  | L0E  2.0  | L0H  0.9  | L0I  1.0  |
| Bin migration                    | L0E  1.0  | L0H  1.0  | L0I  1.0  | L0E  1.6  | L0H  1.6  | L0I  1.6  |
| $r_{J/\psi}$ ratio               | L0E  1.6  | L0H  1.4  | L0I  1.7  | L0E  0.7  | L0H  2.1  | L0I  0.7  |
| Total                            | L0E  4.0  | L0H  6.1  | L0I  5.5  | L0E  6.4  | L0H  7.5  | L0I  6.7  |

10 Systematic uncertainties

Since $R_{K^*0}$ is measured as a double ratio, many potential sources of systematic uncertainty cancel. The remaining systematics and their effects on $R_{K^*0}$ are summarised in table 4 and are described below.

**Corrections to simulation:** the uncertainty induced by the limited size of the simulated sample used to compute the efficiencies is considered; an additional systematic uncertainty is determined using binned corrections instead of interpolated ones; finally, since the data samples used to determine the corrections have a limited size, particularly for the electron hardware trigger, a systematic uncertainty is assessed with a bootstrapping technique [55].

**Trigger efficiency:** for the hardware triggers, the corrections to the simulation are determined using different control samples and the change in the result is assigned as a systematic uncertainty; for the software trigger, the corrections to the simulation do not show dependences on the kinematic of the decays, and therefore only the statistical uncertainty on the overall correction is considered as a systematic uncertainty.

**Particle identification:** the particle identification response is calibrated using data; a systematic uncertainty due to the procedure and kinematic differences between these control samples and the signal modes is included; the effects due to the identification of leptons and hadrons are considered; however, discrepancies in the description of the latter are small and further cancel in the double ratio.

**Kinematic selection:** a systematic uncertainty due to the choice of the mass fit range and to the two-dimensional requirement on $\chi^2_{VD}$ and $m_{corr}$ is determined
by comparing the efficiencies in simulation and background-subtracted samples of $B^0 \rightarrow K^{*0}J/\psi(\rightarrow \mu^+\mu^-)$ or $B^0 \rightarrow K^{*0}J/\psi(\rightarrow e^+e^-)$ decays.

**Residual background:** background due to $B^0 \rightarrow K^{*0}J/\psi(\rightarrow e^+e^-)$ decays where one of the hadrons is misidentified as an electron and vice versa is studied; using simulation that is tuned to data (see section 4) this contribution is estimated to be small; however, a few candidates with one electron of the dilepton pair having a low probability to be genuine are observed in background subtracted data; a systematic uncertainty is assigned based on the distribution of the PID information of these candidates.

**Mass fit:** the systematic uncertainty due to the parameterisation of the signal invariant mass distributions is found to be negligible for the muon channel; for the electron channel, the signal PDF is changed from the sum of a CB and a Gaussian function to the sum of two CB functions, where the mean parameter is shared and, additionally, the mass shift and the width scale factors are constrained using the $B^0 \rightarrow K^{*0}\gamma(\rightarrow e^+e^-)$ decay mode instead of $B^0 \rightarrow K^{*0}J/\psi(\rightarrow e^+e^-)$; the relative fractions of the three bremsstrahlung categories are measured in data using $B^0 \rightarrow K^{*0}J/\psi(\rightarrow e^+e^-)$ and the observed differences with respect to simulation are used in the mass fit (see figure 5); for the backgrounds, a component that describes candidates where the hadron identities are swapped is added both to the muon and electron $B^0 \rightarrow K^{*0}J/\psi(\rightarrow \ell^+\ell^-)$ modes, and constrained to the expected values observed in simulation; the kernel of the nonparametric models is also varied, as well as the mixture of the $K_{1}^{+}(1270)$ and $K_{2}^{*+}(1430)$ components that is constrained using data \[56\]; the contributions to the systematic uncertainty from these sources are evaluated using pseudoexperiments that are generated with modified parameters and fitted with the PDFs used to fit the data.

**Bin migration:** for the electron channel, the degraded $q^2$ resolution due to bremsstrahlung emission causes a nonnegligible fraction of signal candidates to migrate in and out of the given $q^2$ bin; the effect is included in the efficiency determination, but introduces a small dependence on the shape of the differential branching fraction that no longer perfectly cancels in the ratio to the muon channel; pseudoexperiments are generated, where the parameters modelling the $d\Gamma(B^0 \rightarrow K^{*0}e^+e^-)/dq^2$ distribution are varied within their uncertainties \[34\]; the maximum spread of the variation in $R_{K^{*0}}$ is taken as a systematic uncertainty; furthermore, the $q^2$ resolution is smeared for differences between data and simulation that are observed in the resonant mode.

**$r_{J/\psi}$ ratio:** the ratio of the efficiency-corrected yield of the resonant modes (see section 9) is expected to be unity to a very high precision; deviations from unity are therefore considered to be a sign of residual imperfections in the evaluation of the efficiencies; the $r_{J/\psi}$ ratio is studied as a function of various event and kinematic properties of the decay products, and the observed residual deviations from unity are used to assign a systematic uncertainty on $R_{K^{*0}}$.

For the $R_{K^{*0}}$ measurement, all the uncertainties are treated as uncorrelated among the trigger categories, except for those related to particle identification, to the kinematic
selection criteria, to the residual background, to the fit to the invariant mass and to bin migration.

11 Results

The determination of $R_{K^*0}$ exploits the log-likelihoods resulting from the fits to the invariant mass distributions of the nonresonant and resonant channels in each trigger category and $q^2$ region. Each log-likelihood is used to construct the PDF of the true number of decays, which is used as a prior to obtain the PDF of $R_{K^*0}$. The true number of decays is assumed to have a uniform prior. The three electron trigger categories are combined by summing the corresponding log-likelihoods. Uncorrelated systematic uncertainties are accounted for by convolving the yield PDFs with a Gaussian distribution of appropriate width. Correlated systematic uncertainties are treated by convolving the $R_{K^*0}$ PDF with a Gaussian distribution. The one, two and three standard deviation intervals are determined as the ranges that include 68.3%, 95.4% and 99.7% of the PDF. In each $q^2$ region, the measured values of $R_{K^*0}$ are found to be in good agreement among the three electron trigger categories (see figure 9). The results are given in table 5 and presented in figure 10 where they are compared both to the SM predictions (see table 1) and to previous measurements from the $B$ factories [4,5].

The combined $R_{K^*0}$ PDF is used to determine the compatibility with the SM expectations. The $p$-value, calculated by integrating the PDF above the expected value, is translated into a number of standard deviations. The compatibility with the SM expectations [26–35] is determined to be 2.1–2.3 and 2.4–2.5 standard deviations, for the low-$q^2$ and the central-$q^2$ regions, respectively, depending on the theory prediction used.

12 Conclusions

This paper reports a test of lepton universality performed by measuring the ratio of the branching fractions of the decays $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ and $B^0 \rightarrow K^{*0} e^+ e^-$. The $K^{*0}$ meson is reconstructed in the final state $K^+ \pi^-$, which is required to have an invariant mass within 100 MeV/$c^2$ of the known $K^*(892)^0$ mass. Data corresponding to an integrated luminosity
Table 5: Measured $R_{K^*0}$ ratios in the two $q^2$ regions. The first uncertainties are statistical and the second are systematic. About 50% of the systematic uncertainty is correlated between the two $q^2$ bins. The 95.4% and 99.7% confidence level (CL) intervals include both the statistical and systematic uncertainties.

|          | low-$q^2$     | central-$q^2$ |
|----------|---------------|---------------|
| $R_{K^*0}$ | $0.66 \pm 0.11_{-0.07}^{+0.03}$ | $0.69 \pm 0.11_{-0.07}^{+0.05}$ |
| 95.4% CL  | [0.52, 0.89]    | [0.53, 0.94]  |
| 99.7% CL  | [0.45, 1.04]    | [0.46, 1.10]  |

Figure 10: (left) Comparison of the LHCb $R_{K^*0}$ measurements with the SM theoretical predictions: BIP [26] CDHMV [27–29], EOS [30, 31], flav.io [32–34] and JC [35]. The predictions are displaced horizontally for presentation. (right) Comparison of the LHCb $R_{K^*0}$ measurements with previous experimental results from the $B$ factories [4, 5]. In the case of the $B$ factories the specific vetoes for charmonium resonances are not represented.

of 3 fb$^{-1}$ of $pp$ collisions, recorded by the LHCb experiment during 2011 and 2012, are used. The $R_{K^*0}$ ratio is measured in two regions of the dilepton invariant mass squared to be

$$R_{K^*0} = \begin{cases} 
0.66 \pm 0.11_{-0.07}^{+0.03} \text{ (stat) } \pm 0.03 \text{ (syst)} & \text{for } 0.045 < q^2 < 1.1 \text{ GeV}^2/c^4, \\
0.69 \pm 0.11_{-0.07}^{+0.05} \text{ (stat) } \pm 0.05 \text{ (syst)} & \text{for } 1.1 < q^2 < 6.0 \text{ GeV}^2/c^4.
\end{cases}$$

The corresponding 95.4% confidence level intervals are [0.52, 0.89] and [0.53, 0.94]. The results, which represent the most precise measurements of $R_{K^*0}$ to date, are compatible with the SM expectations [26–35] at 2.1–2.3 standard deviations for the low-$q^2$ region and 2.4–2.5 standard deviations for the central-$q^2$ region, depending on the theoretical prediction used.

Model-independent fits to the ensemble of FCNC data that allow for NP contributions [27–35] lead to predictions for $R_{K^*0}$ in the central-$q^2$ region that are similar to the value observed; smaller deviations are expected at low-$q^2$. The larger data set currently being accumulated by the LHCb collaboration will allow for more precise tests of these predictions.
Acknowledgements

We would like to thank F. Le Diberder for many interesting and helpful discussions on statistics. We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); NWO (The Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MinES and FASO (Russia); MinECo (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); NS (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 OCEVU, Région Auvergne (France), RFBR and Yandex LLC (Russia), GVA, XuntaGal and GENCAT (Spain), Herchel Smith Fund, The Royal Society, Royal Commission for the Exhibition of 1851 and the Leverhulme Trust (United Kingdom).

References

[1] G. Hiller and F. Krüger, More model-independent analysis of $b \rightarrow s$ processes, Phys. Rev. D69 (2004) 074020, arXiv:hep-ph/0310219.

[2] C. Bobeth, G. Hiller, and G. Piranishvili, Angular distributions of $\bar{B} \rightarrow \bar{K}\ell\ell$ decays, JHEP 12 (2007) 040, arXiv:0709.4174.

[3] C. Bouchard et al., Standard Model predictions for $B \rightarrow K\ell\ell$ with form factors from lattice QCD, Phys. Rev. Lett. 111 (2013) 162002, arXiv:1306.0434.

[4] BaBar collaboration, J. P. Lees et al., Measurement of branching fractions and rate asymmetries in the rare decays $B \rightarrow K(\ast)\ell^+\ell^-$, Phys. Rev. D86 (2012) 032012, arXiv:1204.3933.

[5] Belle collaboration, J.-T. Wei et al., Measurement of the differential branching fraction and forward-backward asymmetry for $B \rightarrow K(\ast)\ell^+\ell^-$, Phys. Rev. Lett. 103 (2009) 171801, arXiv:0904.0770.

[6] LHCb collaboration, R. Aaij et al., Test of lepton universality using $B^+ \rightarrow K^+\ell^+\ell^-$ decays, Phys. Rev. Lett. 113 (2014) 151601, arXiv:1406.6482.

[7] BaBar collaboration, B. Aubert et al., Measurements of the semileptonic decays $\bar{B} \rightarrow D\ell\bar{\nu} and \bar{B} \rightarrow D^*\ell\bar{\nu}$ using a global fit to $DX\ell\bar{\nu}$ final states, Phys. Rev. D79 (2009) 012002, arXiv:0809.0828.
[8] Belle collaboration, M. Huschle et al., Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^+\bar{\nu}_{\tau}$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_{\ell}$ decays with hadronic tagging at Belle, Phys. Rev. D92 (2015) 072014, arXiv:1507.03233

[9] LHCb collaboration, R. Aaij et al., Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_{\tau}) / \mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_{\mu})$, Phys. Rev. Lett. 115 (2015) 111803, arXiv:1506.08614.

[10] LHCb collaboration, R. Aaij et al., Differential branching fractions and isospin asymmetries of $B \to K^*\mu^+\mu^-$ decays, JHEP 06 (2014) 133, arXiv:1403.8044.

[11] LHCb collaboration, R. Aaij et al., Differential branching fraction and angular analysis of $A^0_b \to \Lambda\mu^+\mu^-$ decays, JHEP 09 (2015) 179, arXiv:1503.07138.

[12] LHCb collaboration, R. Aaij et al., Angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3\,fb^{-1}$ of integrated luminosity, JHEP 02 (2016) 104, arXiv:1512.04442.

[13] Belle collaboration, S. Wehle et al., Lepton-flavor-dependent angular analysis of $B \to K^*\ell^+\ell^-$, Phys. Rev. D88 (2013) 074002, arXiv:1307.5683.

[14] W. Altmannshofer and D. M. Straub, New physics in $B \to K^*\mu\mu$?, Eur. Phys. J. C73 (2013) 2646, arXiv:1308.1501.

[15] G. Hiller and M. Schmaltz, Diagnosing lepton-nonuniversality in $b \to s\ell\ell$, Phys. Rev. D94 (2016) 115021, arXiv:1608.03501.

[16] A. Crivellin, D. Müller, and T. Ota, Simultaneous explanation of $R(D^{(*)})$ and $b \to s\mu^+\mu^-$: the last scalar leptoquarks standing, arXiv:1703.09226.

[17] S. Descotes-Genon, J. Matias, and J. Virto, Understanding the $B \to K^*\mu^+\mu^-$ anomaly, Phys. Rev. D88 (2013) 074002, arXiv:1310.1237.

[18] R. Gauld, F. Goertz, and U. Haisch, An explicit $Z'$-boson explanation of the $B \to K^*\mu^+\mu^-$ anomaly, JHEP 01 (2014) 069, arXiv:1310.1082.

[19] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, Quark flavor transitions in $L_{\mu} - L_{\tau}$ models, Phys. Rev. D89 (2014) 095033, arXiv:1403.1269.

[20] W. Altmannshofer, S. Gori, S. Profumo, and F. S. Queiroz, Explaining dark matter and $B$ decay anomalies with an $L_{\mu} - L_{\tau}$ model, JHEP 12 (2016) 106, arXiv:1609.04026.

[21] D. Bečirević, S. Fajfer, N. Košnik, and O. Sumensari, Leptoquark model to explain the $B$-physics anomalies, $R_K$ and $R_D$, Phys. Rev. D94 (2016) 115021, arXiv:1608.03501.

[22] A. Crivellin, D. Müller, and T. Ota, Simultaneous explanation of $R(D^{(*)})$ and $b \to s\mu^+\mu^-$: the last scalar leptoquarks standing, arXiv:1703.09226.

[23] G. Hiller and M. Schmaltz, Diagnosing lepton-nonuniversality in $b \to s\ell\ell$, JHEP 02 (2015) 055, arXiv:1411.4773.
[24] Particle Data Group, C. Patrignani et al., Review of particle physics, Chin. Phys. C40 (2016) 100001.

[25] LHCb collaboration, R. Aaij et al., Measurement of the S-wave fraction in $B^0 \rightarrow K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \rightarrow K^*(892)^0\mu^+\mu^-$ differential branching fraction, JHEP 11 (2016) 047, Erratum ibid. 04 (2017) 142, arXiv:1606.04731.

[26] M. Bordone, G. Isidori, and A. Pattori, On the Standard Model predictions for $R_K$ and $R_{K^*}$, Eur. Phys. J. C76 (2016) 440, arXiv:1605.07633.

[27] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, Global analysis of $b \rightarrow s\ell\ell$ anomalies, JHEP 06 (2016) 092, arXiv:1510.04239.

[28] B. Capdevila, S. Descotes-Genon, J. Matias, and J. Virto, Assessing lepton-flavour non-universality from $B \rightarrow K^*\ell\ell$ angular analyses, JHEP 10 (2016) 075, arXiv:1605.03156.

[29] B. Capdevila, S. Descotes-Genon, L. Hofer, and J. Matias, Hadronic uncertainties in $B \rightarrow K^*\mu^+\mu^-$: a state-of-the-art analysis, JHEP 04 (2017) 016, arXiv:1701.08672.

[30] N. Serra, R. Silva Coutinho, and D. van Dyk, Measuring the breaking of lepton flavour universality in $B \rightarrow K^+\ell^+\ell^-$, Phys. Rev. D95 (2017) 035029, arXiv:1610.08761.

[31] D. van Dyk et al., EOS — A HEP program for flavor observables, https://eos.github.io; D. van Dyk et al., EOS (“delta456” release), doi: 10.5281/zenodo.159680.

[32] A. Bharucha, D. M. Straub, and R. Zwicky, $B \rightarrow V\ell^+\ell^-$ in the Standard Model from light-cone sum rules, JHEP 08 (2016) 098, arXiv:1503.05534.

[33] W. Altmannshofer, C. Niehoff, P. Stangl, and D. M. Straub, Status of the $B \rightarrow K^*\mu^+\mu^-$ anomaly after Moriond 2017, arXiv:1703.09189.

[34] D. Straub et al., flav-io/flavio v0.19, doi: 10.5281/zenodo.292991.

[35] S. Jäger and J. Martin Camalich, Reassessing the discovery potential of the $B \rightarrow K^*\ell^+\ell^-$ decays in the large-recoil region: SM challenges and BSM opportunities, Phys. Rev. D93 (2016) 014028, arXiv:1412.3183.

[36] LHCb collaboration, A. A. Alves Jr. et al., The LHCb detector at the LHC, JINST 3 (2008) S08005.

[37] LHCb collaboration, R. Aaij et al., LHCb detector performance, Int. J. Mod. Phys. A30 (2015) 1530022, arXiv:1412.6352.

[38] V. V. Gligorov and M. Williams, Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree, JINST 8 (2013) P02013, arXiv:1210.6861.

[39] T. Sjöstrand, S. Mrenna, and P. Skands, PYTHIA 6.4 physics and manual, JHEP 05 (2006) 026, arXiv:hep-ph/0603175; T. Sjöstrand, S. Mrenna, and P. Skands, A brief introduction to PYTHIA 8.1, Comput. Phys. Commun. 178 (2008) 852, arXiv:0710.3820.
I. Belyaev et al., Handling of the generation of primary events in GAUSS, the LHCb simulation framework, Nuclear Science Symposium Conference Record (NSS/MIC) IEEE (2010) 1155.

D. J. Lange, The EvtGen particle decay simulation package, Nucl. Instrum. Meth. A462 (2001) 152.

P. Golonka and Z. Was, PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays, Eur. Phys. J. C45 (2006) 97, arXiv:hep-ph/0506026.

Geant4 collaboration, J. Allison et al., Geant4 developments and applications, IEEE Trans. Nucl. Sci. 53 (2006) 270, Geant4 collaboration, S. Agostinelli et al., Geant4: a simulation toolkit, Nucl. Instrum. Meth. A506 (2003) 250.

M. Clemencic et al., The LHCb simulation application, GAUSS: design, evolution and experience, J. Phys. Conf. Ser. 331 (2011) 032023.

M. Adinolfi et al., Performance of the LHCb RICH detector at the LHC, Eur. Phys. J. C73 (2013) 2431, arXiv:1211.6759.

A. Blum, A. Kalai, and J. Langford, Beating the hold-out: Bounds for k-fold and progressive cross-validation, in Proceedings of the Twelfth Annual Conference on Computational Learning Theory, COLT ’99, (New York, NY, USA), pp. 203–208, ACM, 1999. doi: 10.1145/307400.307439.

M. Feindt and U. Kerzel, The NeuroBayes neural network package, Nucl. Instrum Meth. A559 (2006) 190.

A. Y. Korchin and V. A. Kovalchuk, Contribution of low-lying vector resonances to polarization observables in B^0 \rightarrow K^{*0}e^+e^- decay, Phys. Rev. D82 (2010) 034013, arXiv:1004.3647.

S. Jäger and J. M. Camalich, On B \rightarrow V\ell\ell at small dilepton invariant mass, power corrections, and new physics, JHEP 05 (2013) 043, arXiv:1212.2263.

D. Martínez Santos and F. Dupertuis, Mass distributions marginalized over per-event errors, Nucl. Instrum. Meth. A764 (2014) 150, arXiv:1312.5000.

K. S. Cranmer, Kernel estimation in high-energy physics, Comput. Phys. Commun 136 (2001) 198, arXiv:hep-ex/0011057.

LHCb collaboration, R. Aaij et al., Observation of J/ψp resonances consistent with pentaquark states in \Lambda_b^0 \rightarrow J/ψpK^- decays, Phys. Rev. Lett. 115 (2015) 072001, arXiv:1507.03414.

T. Skwarnicki, A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-P31-86-02.

M. Pivk and F. R. Le Diberder, sPlot: a statistical tool to unfold data distributions, Nucl. Instrum. Meth. A555 (2005) 356, arXiv:physics/0402083.
[55] B. Efron, *Bootstrap methods: Another look at the jackknife*, Ann. Stat. 7 (1979) 1.

[56] LHCb collaboration, R. Aaij *et al.*, *First observations of the rare decays $B^+ \rightarrow K^+\pi^+\pi^-\mu^+\mu^-$ and $B^+ \rightarrow \phi K^+\mu^+\mu^-$*, JHEP 10 (2014) 064, arXiv:1408.1137
associated to 38

71 Van Swinderen Institute, University of Groningen, Groningen, The Netherlands, associated to 43

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

† Deceased