Search for the lepton flavour violating decay
\( B^+ \rightarrow K^+ \mu^- \tau^+ \) using \( B_{s2}^0 \) decays

The LHCb collaboration

E-mail: msrudolp@syr.edu

Abstract: A search is presented for the lepton flavour violating decay \( B^+ \rightarrow K^+ \mu^- \tau^+ \) using a sample of proton-proton collisions at centre-of-mass energies of 7, 8, and 13 TeV, collected with the LHCb detector and corresponding to a total integrated luminosity of 9 fb\(^{-1}\). The \( \tau \) leptons are selected inclusively, primarily via decays with a single charged particle. The four-momentum of the \( \tau \) lepton is determined by using \( B^+ \) mesons from \( B_{s2}^0 \rightarrow B^+ K^- \) decays. No significant excess is observed, and an upper limit is set on the branching fraction

\[
B(B^+ \rightarrow K^+ \mu^- \tau^+) < 3.9 \times 10^{-5} \text{ at 90\% confidence level.}
\]

The obtained limit is comparable to the world-best limit.

Keywords: B physics, Beyond Standard Model, Flavor physics, Hadron-Hadron scattering (experiments), Rare decay

ArXiv ePrint: 2003.04352
1 Introduction

A number of experimental hints of lepton flavour universality violation in the semileptonic transitions $b \to s \ell^+\ell^-$ [1-3] and $b \to c \ell^-\pi_\ell$ [4-9] have recently been found.\footnote{1}{The inclusion of charge-conjugate processes is implied throughout.} In general, physics beyond the Standard Model that generates lepton flavour non-universality is likely to also produce direct lepton flavour violation [10]. Theoretical models seeking to simultaneously explain all these anomalies, for example with a vector leptoquark, often lead to relatively large branching fractions for the decays $B \to K\mu^\pm\tau^\mp$ [11-16].

The branching fractions for the two $\mu\tau$ charge combinations are not in general the same, as they depend on the details of the physics mechanism producing the decay. In this paper, we present a search for the decay $B^+ \to K^+\mu^-\tau^+$. From an experimental point of view, this combination is preferred over $B^+ \to K^+\mu^+\tau^-$ as it has a lower background from semileptonic $B \to D X \mu^+\nu_\mu$ decays, where $X$ refers to any number of additional particles, because Cabibbo-favoured decays of the charm meson are likely to lead to kaons of the same charge as the muon. An upper limit on the branching fraction for the signal decay has been previously set by the BaBar collaboration [17] $B(B^+ \to K^+\mu^-\tau^+) < 2.8 \times 10^{-5}$ at 90% confidence level (CL).

We reconstruct the full four-momentum of the $\tau$ lepton using $B^+$ mesons from the decay $B_{s2}^0 \to B^+K^-$, which amounts to about 1% of $B^+$ production. By reconstructing the decay vertex of the $B^+$ meson from the $K^+\mu^-$ pair and the momentum of the $K^-$
meson, it is possible to determine the momentum of the $B^+$ meson up to a quadratic ambiguity by imposing mass constraints on the $B_{s2}^0$ and $B^+$ mesons [18]. This technique was first used to study relative branching fractions in $B^+ \rightarrow D^0 \chi \mu^+\nu$ decays [19]. We then search for a peak in the missing-mass squared distribution corresponding to the $\tau$ mass squared, $m_{\tau}^2$. Even signal $B^+$ mesons not coming from a $B_{s2}^0$ decay show a peak at $m_{\tau}^2$. We account for the contribution of these non-$B_{s2}^0$ candidates in the analysis. The leptons are selected inclusively, as we only require one additional charged track near the $K^+\mu^-$ pair to help discriminate against background. To normalise the branching fraction, we use the decay $B^+ \rightarrow J/\psi K^+$, with $J/\psi \rightarrow \mu^+\mu^-$. The normalisation channel is also used to quantify the contributions from $B_{s2}^{*0}$ decays, as well as non-$B_{s2}^{*0}$ candidates with nearby kaons.

In addition to providing the missing-mass discriminating variable, this method allows us to study the control sample composed of same-sign $B^+K^+$ decays, which does not include any $B_{s2}^{*0}$ component. We use this sample to optimise the signal selection, and motivate our description of the background missing-mass shape.

2 Detector, data samples, and simulation

The LHCb detector [20, 21] is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing $b$ or $c$ quarks. The detector includes a high-precision tracking system consisting of a silicon-strip vertex detector surrounding the $pp$ interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of the momentum, $p$, of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV.\footnote{Natural units with $c = 1$ are used throughout.}

The minimum distance of a track to a primary $pp$ interaction vertex (PV), the impact parameter, 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. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov detectors. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers. The online event selection is performed by a trigger, which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction. At the hardware trigger stage, events are required to have a muon with high $p_T$ or a hadron, photon or electron with high transverse energy deposited in the calorimeters. The software trigger requires a two-, three- or four-track secondary vertex with a significant displacement from any primary vertex.

We use data samples collected from 2011 to 2018, at centre-of-mass energies of 7, 8, and 13 TeV, corresponding to an integrated luminosity of 9 fb$^{-1}$. We model signal and normalisation decays using simulation. In the simulation, $pp$ collisions are generated using PYTHIA [22, 23] with a specific LHCb configuration [24]. Decays of hadrons and $\tau$ leptons are described by EVTGEN [25]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [26, 27] as described in ref. [28].
For the signal, we consider both a phase space model and variations of the decay kinematics with effective operators for the $b \to s \mu^+\tau^-$ interaction and their corresponding Wilson coefficients using the distributions from ref. [29] (see also ref. [30]) and the form factors from ref. [31]. The branching fraction limit is determined for various hypotheses: for the phase-space decay, for a decay via the vector or axial-vector operators $O^{(0)}_9$ or $O^{(0)}_{10}$, and for a decay using the scalar or pseudoscalar operators $O^{(0)}_S$ or $O^{(0)}_P$ [29].

3 Selection and missing mass calculation

The selection of $B^+$ candidates begins with a $K^+\mu^-$ pair with an invariant mass $m_{K^+\mu^-} > 1800$ MeV to reduce background from semileptonic charm decays. The $K^+$ and $\mu^-$ candidates are formed from high-quality tracks consistent with kaon and muon hypotheses and inconsistent with being produced at any PV in the event. The $K^+\mu^-$ vertex must be of high quality and well separated from any PV. The $K^+\mu^-$ pair is associated with a single PV by choosing the vertex that minimizes the quantity $2\Delta\chi^2_{ip}$, defined as the change in the $\chi^2$ of the vertex fit when including or excluding the $K^+\mu^-$ pair in the fit.

To better separate signal candidates with $\tau$ leptons from background, we require an additional track, labelled $t^+$, with charge opposite to that of the muon. This track must also be of high quality and inconsistent with being produced at any PV in the event. By adding this third track, we also fully reconstruct the normalisation mode $B^+ \to J/\psi K^+$, with $J/\psi \to \mu^+\mu^-$. Many background candidates are expected to come from $B$-meson decays of the form $B \to \bar{D}(\to K^+X\mu^-)K^+Y$, where $X$ and $Y$ refer to any number of additional particles. In these cases the kaon originating from the $\bar{D}$ meson is assigned as the additional track. Since only approximately 2% of $\tau$ decays contain a charged kaon, we apply particle identification requirements so that the track is unlikely to be a charged kaon. Events in which multiple candidates are found, or events in which a candidate $B^+\to J/\psi K^+$ decay is reconstructed, are not used in this search. This removes backgrounds and avoids overlap with ongoing searches at LHCb exclusively using this decay channel. The overall signal loss is less than 3%.

We split the data samples into signal and normalisation regions based on the invariant mass of the $K^+\mu^-t^+$ triple, using the muon hypothesis for the third track. Candidates with $m_{K^+\mu^-} < 4800$ MeV fall into the signal region, while candidates with $5180 < m_{K^+\mu^-} < 5380$ MeV and $|m_{\mu\mu} - m_{J/\psi}| < 40$ MeV fall into the normalisation region.

The $B^+$ candidate direction is estimated using the associated PV and $K^+\mu^-$ vertex positions. We next consider prompt tracks, i.e. those that are consistent with being produced at that PV. Those tracks identified as kaons, with a charge opposite to that of the kaon in the $K^+\mu^-$ pair and a small perpendicular momentum relative to the $B^+$ candidate direction, are combined with the $B^+$ candidates to form $B_{22}^0$ candidates. We refer to this sample as the opposite-sign kaon (OSK) sample. Additionally, we select a control sample, referred to as same-sign kaon (SSK) sample, by adding prompt kaons of the same sign as the kaon in the $K^+\mu^-$ pair.

From ref. [19], the two $B$-meson energy solutions are

$$E_B = \frac{\Delta^2}{2E_K} \frac{1}{1 - (p_K/E_K)^2 \cos^2 \theta} [1 \pm \sqrt{d}], \quad \text{where} \quad (3.1)$$
\[ d = \frac{p_K^2}{E_K} \cos^2 \theta - \frac{4m_{BK}^2 p_K^2 \cos^2 \theta}{\Delta^4} \left( 1 - \frac{p_K^2}{E_K^2} \cos^2 \theta \right), \quad (3.2) \]

\[ \Delta^2 = m_{BK}^2 - m_B^2 - m_K^2, \quad (3.3) \]

where \( m_{BK} = m_{B_{s2}^0} \) is the assumed \( B^+ K^- \) mass, \( p_K \) and \( E_K \) are the reconstructed prompt kaon momentum and energy, and \( \theta \) is the laboratory frame angle between the prompt kaon and \( B \)-meson directions. The missing four-momentum of the \( \tau \) lepton, \( P_{\text{miss}} \), is then reconstructed as \( P_B - P_{K+\mu^-} \), where \( P_B \) and \( P_{K+\mu^-} \) are the four-momenta of the \( B \) meson and \( K^+ \mu^- \) pair. The missing mass squared is calculated using the lowest energy, real solution for which the resulting missing energy is greater than the reconstructed energy of the third track under a pion mass hypothesis. With this choice, we correctly reconstruct the energy of signal decays in simulation in more than 75% of cases. About 9% of all signal decays have no such solution and are lost. Both signal and normalisation candidates, as well as the SSK control-sample candidates, are required to pass this procedure. Candidates in the signal region are additionally required to have the residual missing mass squared, defined as the four-momentum difference of the \( B \) meson and \( K^+ \mu^- t^+ \) triple, \( (P_B - P_{K+\mu^-} - P_t)^2 \), greater than \(-0.5 \text{ GeV}^2\). This requirement removes background and only poorly reconstructed signal candidates which do not peak at the \( \tau \) mass squared. The minimum mass difference, defined in ref. [19] as

\[ \Delta m_{\text{min}} = \sqrt{m_B^2 + m_K^2 + 2m_B \sqrt{p_K^2 \sin^2 \theta + m_K^2} - m_B - m_K}, \quad (3.4) \]

is required to be greater than 30 MeV. This removes contributions from \( B_{s1}^0 \) and \( B_{s2}^0 \to B^+ K^- \) decays, as well as background in which a kaon from the \( B \) decay is wrongly associated to the primary vertex.

Missing-mass distributions for the signal simulation and the full data sample after the above selection are shown in figure 1. All signal decays, whether they come from a \( B_{s2}^0 \) meson or not, peak at the known \( m_{BK}^2 \), however the non-\( B_{s2}^0 \) candidates have a much wider peak than the \( B_{s2}^0 \) ones. The data distributions are shown for both the OS\( K \) and SSK samples. They have similar shapes with a broad hump centred near 5 GeV\(^2\). We note that the OS\( K \) sample has a higher yield than the SSK; this excess has been observed in both fully and partially reconstructed decays [19, 32].

### 4 Normalisation

We determine the yield of the normalisation decay, as well as the relative efficiency of the signal modes with respect to the normalisation mode, separately for each data-taking year. For the normalisation mode, we determine the inclusive yield of \( B^+ \to J/\psi K^+ \) decays, whether or not they originate from a \( B_{s2}^0 \) meson, by a binned maximum-likelihood fit to the \( K^+ \mu^- t^+ \) mass distribution, where we assign the muon mass hypothesis to the third track. The signal is described with a Gaussian distribution, and the background with a linear model.
We determine the fraction of the normalisation candidates coming from $B_{s}^{0}$ decays using a $K^{+}\mu^{-}\tau^{+}$ mass fit for the combined-years data sample using the same model as the separated-years samples, along with a binned maximum-likelihood fit to the measured mass-difference distribution $m_{B^{+}K^{-}} - m_{B^{+}} - m_{K^{-}}$ around the $B_{s}^{0}$ peak. For the latter fit, we describe the signal peak with a Gaussian core that transitions to an exponential tail on each side, and we model the background with a third-degree polynomial. The results of these fits are shown in figure 2. The total data sample contains $4240 \pm 70$ $B^{+} \rightarrow J/\psi K^{+}$ decays; the fraction originating from $B_{s}^{0}$ decays is $f_{B_{s}^{0}} = (25.4 \pm 1.8)\%$, where the uncertainty combines the statistical and systematic uncertainties from the choice of fit function. The year-to-year variation is not found to be statistically significant, so we use the value obtained from the combined dataset for all years.

The relative efficiency of the signal and normalisation modes is determined using simulation with corrections from data. For $B_{s}^{0}$ decays the relative efficiencies in different years average around 30%, with an absolute year-to-year variation of less than 3%. Different signal decay models change the relative efficiency by approximately 10%, with the decays via scalar and pseudoscalar operators having a lower overall efficiency. Signal events in
which the $B^+$ meson does not originate from a $B_{s2}^{*0}$ decay have a lower selection efficiency, primarily because fewer of these candidates pass the residual missing-mass requirement and fall into the missing-mass fit range. Using simulation, we derive an additional factor for this signal component of $r_{\text{non-}B_{s2}^{*0}} = 0.849 \pm 0.007$, which gives the relative efficiency with respect to the $B_{s2}^{*0}$ mode.

5 Multivariate signal selection

We further improve the signal selection using a Boosted Decision Tree (BDT) classification with the Adaboost algorithm [33]. The BDT inputs are primarily chosen to distinguish additional tracks coming from signal $\tau$ lepton decays from various sources of background. Some examples are semileptonic $b$-hadron decays to charm where the charm hadron produces a kaon with charge opposite that of the muon, or $b$-hadron decays where the muon is produced in the semileptonic decay of a child charm hadron. The background training sample is taken from the SSK sample in the $m_{\text{miss}}^2$ region around $m_{\tau}^2$. This focuses the training on the sources of background which fall near the signal peak. We describe the signal with simulation samples that include only $B_{s2}^{*0}$ decays; the effect of the BDT on non-$B_{s2}^{*0}$ signal simulation is then estimated separately.

The training makes use of different topological reconstructions of the $K^+\mu^-\tau^+$ triple: in addition to the signal selection, we also first combine either the kaon and the track or the muon and the track into a pair before adding the third particle. The pair masses and the flight distance of the pair in each topology help to distinguish the signal from background, for instance when the pair comes from a charm hadron decay. We also include the flight distance of the $\tau$, which we reconstruct as the distance along the $\tau$ trajectory found in the missing-mass calculation from the $K^+\mu^-$ vertex to the point of closest approach of the third track. The result of a separate isolation discriminant is included to reduce background with additional charged tracks; this discriminant is trained to distinguish additional tracks belonging to the same $b$-hadron decay from other tracks in the event. It uses the kinematics of the additional tracks, their distances to and angles with the signal candidate tracks, and topological information from vertices formed by the additional tracks and signal candidate tracks.

A loose requirement on the signal optimisation BDT output is applied, keeping about 70% of all simulated $B_{s2}^{*0}$ signal candidates and about 40% of non-$B_{s2}^{*0}$ signal candidates. We perform the final fit to the $m_{\text{miss}}^2$ distribution in four bins of the BDT output. The bins are chosen by optimising the expected upper limit using a number of background events derived from the yields in the OSK and SSK $m_{\text{miss}}^2$ sidebands from 1 to 2 GeV$^2$ and 4 to 6 GeV$^2$.

6 Background studies

The background in this analysis is composed of a large number of different partially reconstructed $b$-hadron decays. None of them, however, produce a narrow peak in $m_{\text{miss}}^2$. Only
$B^+$ mesons produced from $B_{s2}^{\pm0}$ decays have a resolution comparable to the signal. Furthermore, if there is more than one missing particle then the true missing-mass distribution will be much wider than the expected signal peak. Charm hadrons have masses close to the $\tau$ mass, however there is no Standard Model decay $B^+ \rightarrow K^+\mu^-D^+$. We are not sensitive to decays such as $B^+ \rightarrow K^+\pi^-D^+$, where the pion is misidentified as a muon, because of their low branching fraction and the rate at which hadrons are misidentified as muons.

We expect that the missing-mass distribution, summed over many different background components, is smooth, and we model it as a polynomial.

These assumptions are tested using simulation and data. We produce fast simulation samples with RapidSim [34] of a number of potential exclusive background sources from $B^+$, $B^0$, $B_{s2}^{0}$, and $A_{0}$ hadrons; the true missing-mass distributions for these decays are smeared to estimate their shapes in data. No sign of any sharply peaking component is found. In data we consider a number of different control samples, namely all possible $K\mu\tau$ charge combinations in both OS and SS $K$ samples, excluding the signal selection of $K^+\mu^-\tau^+$ in the OS $K$ sample. There is no sign of any narrow peak in any of the distributions, even after applying a tight requirement on the BDT output.

Maximum-likelihood fits to the SS $K$ sample using polynomials of different degrees in the restricted $m_{\text{miss}}^2$ range from 1 to 6 GeV$^2$ are used to study the background shape in more detail. The optimal number of free polynomial parameters in the most signal-like BDT output bin, based on the best-fit value of $-2\log L$, penalised by one for each additional parameter, is four. We further study the effect of background modelling by performing a large number of pseudoexperiments, both background-only and with injected signal at branching fractions of $1 \times 10^{-5}$ and $2 \times 10^{-5}$. In these studies, we first fit a background model of some polynomial degree to one of the control samples to determine a base background model. We generate many pseudodatasets from this background model, and then fit them both with polynomials of the same degree as well as different degrees. Based on these studies, we take into account the systematic uncertainty due to the background modelling by reporting the weakest limit using background descriptions of third, fourth, or fifth degree polynomials, all of which well describe the background shapes in the pseudoexperiments.

7 Fit description

We search for the $K^+\mu^-\tau^+$ missing-mass peak with an unbinned maximum-likelihood fit simultaneously in four bins of BDT output in the OS $K$ $K^+\mu^-\tau^+$ signal channel. The fit is performed in the missing-mass range $1 < m_{\text{miss}}^2 < 6$ GeV$^2$. The parameter of interest is the branching fraction $\mathcal{B}(B^+ \rightarrow K^+\mu^-\tau^+)$. We describe the $m_{\text{miss}}^2$ shape for the signal component with a generalized hyperbolic distribution [35] with shape parameters obtained from simulation. Two signal shapes are used: one for $B_{s2}^{\pm0}$ decays, and one for the wider non-$B_{s2}^{\pm0}$ contribution. We determine the shapes separately in each bin of BDT response. The signal decay model does not significantly affect the signal missing-mass shape. The background is described by polynomial functions which vary independently in each BDT output bin.
We base the normalisation of the signal components on the yields of the $B^+ \to J/\psi K^+$ decays; the yield is determined in data year-by-year to account for different efficiencies between years. We combine this together with the relative efficiencies, $\varepsilon_{\text{rel}}$; the known $B^+ \to J/\psi K^+$ with $J/\psi \to \mu^+ \mu^-$ combined branching fraction, abbreviated as $\mathcal{B}(J/\psi K^+)$; and the parameter of interest to derive a total number of $B^+ \to K^+ \mu^- \tau^+$ signal decays. This total is divided between $B_{s2}^{0 \ast}$ and non-$B_{s2}^{0 \ast}$ decays based on the observed fraction in the normalization channel, and then distributed across the four BDT bins. This gives yields in each BDT bin $j$ of

$$N_j(B^+ \to K^+ \mu^- \tau^+|B_{s2}^{0 \ast}) = \varepsilon_{B_{s2}^{0 \ast}j} \frac{\mathcal{B}(K^+ \mu^- \tau^+)}{\mathcal{B}(J/\psi K^+)} f_{B_{s2}^{0 \ast}}$$

$$\times \sum_{i \in \text{years}} \varepsilon_{\text{rel},i} f_{i} N_i(J/\psi K^+),$$

(7.1)

$$N_j(B^+ \to K^+ \mu^- \tau^+|\text{non-}{B_{s2}^{0 \ast}}) = \varepsilon_{\text{non-}{B_{s2}^{0 \ast}j}} \frac{\mathcal{B}(K^+ \mu^- \tau^+)}{\mathcal{B}(J/\psi K^+)} (1 - f_{B_{s2}^{0 \ast}})$$

$$\times \sum_{i \in \text{years}} \varepsilon_{\text{rel},i} f_{\text{non-}{B_{s2}^{0 \ast}}} N_i(J/\psi K^+),$$

(7.2)

where $\varepsilon_{B_{s2}^{0 \ast}j}$ and $\varepsilon_{\text{non-}{B_{s2}^{0 \ast}j}}$ are the separate efficiencies for each signal component to be found in BDT bin $j$. The main parameters of the fit are thus the $B^+ \to K^+ \mu^- \tau^+$ branching fraction, four parameters for the background normalisation in each BDT bin, and up to five parameters describing the polynomial background shapes in each BDT bin.

The largest systematic uncertainty comes from the choice of background model. The fifth degree background description obtains the weakest limit among the tested background models. We include the effects of other systematic uncertainties using Gaussian-constrained nuisance parameters. These nuisance parameters modify the normalisation yield, the relative efficiency of the signal and normalisation channels, the signal yield in each BDT bin, and the signal shapes. The largest effects come from the modelling of the kinematics of $B_{s2}^{0 \ast}$ decays in simulation, which results in 5% changes in the relative efficiency and in the signal fractions in each bin of BDT response. The relative statistical uncertainty of the $B_{s2}^{0 \ast}$ fraction taken from the normalisation channel is approximately 7%. Altogether, the total effect of these systematic uncertainties on the final limit is small, at the $10^{-6}$ level.

### 8 Results and conclusion

The result at the best fit point is shown in figure 3. The obtained value for the signal branching fraction from the maximum-likelihood fit is $(1.9 \pm 1.5) \times 10^{-5}$. No significant excess is observed, and we set upper limits on the branching fraction using the CLs method [36]. We perform a scan in the signal branching fraction, obtaining the signal and background $p$-values from the distributions of a one-sided profile-likelihood-ratio test statistic obtained with pseudoexperiments in which we vary the constraints on the systematic uncertainties. The scan used to determine the observed limits, compared to the expected one, is shown in figure 4. The expected upper limit at 90% CL is $2.3 \times 10^{-5}$. The observed 90% and
95% CL limits, assuming a phase space signal decay model, are:

\[ B(B^+ \rightarrow K^+ \mu^- \tau^+) < 3.9 \times 10^{-5} \text{ at 90\% CL,} \]

\[ < 4.5 \times 10^{-5} \text{ at 95\% CL.} \]

An identical limit is obtained when the decay is generated from the effective operators \( \mathcal{O}^{(i)}_9 \) or \( \mathcal{O}^{(i)}_{10} \). If instead it is produced from \( \mathcal{O}^{(i)}_9 \) or \( \mathcal{O}^{(i)}_{10} \), the obtained limit is

\[ B(B^+ \rightarrow K^+ \mu^- \tau^+) < 4.4 \times 10^{-5} \text{ at 90\% CL and < 5.0 \times 10^{-5} at 95\% CL.} \]

This is the first result from the LHCb experiment for the lepton-flavour violating decay \( B^+ \rightarrow K^+ \mu^- \tau^+ \). By studying \( B^+ \) mesons from \( B_{s2}^{0} \) decays, we are able to make the first analysis at LHCb of a \( B \) hadron decay using inclusive \( \tau \) decays. This provides complementary information to searches for lepton-flavour violation at LHCb with three-prong \( \tau \) decays, for example \( B_{(s)}^{0} \rightarrow \tau^\pm \mu^\mp \) decays \cite{37}. We observe no significant signal, and set an upper limit slightly above that obtained by the BaBar collaboration \cite{17}.

**Acknowledgments**

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies:
Figure 4. Scan of the $p$-value in the signal branching fraction used to determine the CLs upper limits, compared to the expected one. Scan points below the best-fit branching fraction are not shown. The horizontal red line shows a $p$-value of 0.1, used to define the 90% CL upper limit.

CAPES, CNPq, 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); DOE NP and NSF (U.S.A.). 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 (U.S.A.). We are indebted to the communities behind the multiple open-source software packages on which we depend. Individual groups or members have received support from AvH Foundation (Germany); EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union); ANR, Labex P2IO 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).

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

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


[6] LHCb collaboration, Measurement of the ratio of branching fractions $B(B^+_c \to J/\psi\tau^+\nu_\tau)/B(B^+_c \to J/\psi\mu^+\nu_\mu)$, *Phys. Rev. Lett.* **120** (2018) 121801 [arXiv:1711.05623] [SPIRE].


[8] LHCb collaboration, Measurement of the ratio of the $B^0 \to D^{*-\tau^+\nu_\tau}$ and $B^0 \to D^{*-\mu^+\nu_\mu}$ branching fractions using three-prong $\tau$-lepton decays, *Phys. Rev. Lett.* **120** (2018) 171802 [arXiv:1708.08856] [SPIRE].


[17] BABAR collaboration, A search for the decay modes $B^\pm \to h^\pm \tau^\pm l$, *Phys. Rev. D* **86** (2012) 012004 [arXiv:1204.2852] [SPIRE].


1 Centro Brasileiro de Pesquisas Físicas (CBPF), Rio de Janeiro, Brazil
2 Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil
3 Center for High Energy Physics, Tsinghua University, Beijing, China
4 School of Physics State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
5 University of Chinese Academy of Sciences, Beijing, China
6 Institute Of High Energy Physics (IHEP), Beijing, China
7 Institute of Particle Physics, Central China Normal University, Wuhan, Hubei, China
8 Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IN2P3-LAPP, Annecy, France
9 Université Clermont Auvergne, CNRS/IN2P3, LPC, Clermont-Ferrand, France
10 Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France
11 Université Paris-Saclay, CNRS/IN2P3, IJCLab, 91405 Orsay, France, Orsay, France
12 LPNHE, Sorbonne Universités, Paris Diderot Sorbonne Paris Cité, CNRS/IN2P3, Paris, France
13 I. Physikalisches Instut, RWTH Aachen University, Aachen, Germany
14 Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany
15 Max-Planck-Institut für Kernphysik (MPIK), Heidelberg, Germany
16 Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany
17 School of Physics, University College Dublin, Dublin, Ireland
18 INFN Sezione di Bari, Bari, Italy
19 INFN Sezione di Bologna, Bologna, Italy
20 INFN Sezione di Ferrara, Ferrara, Italy
21 INFN Sezione di Firenze, Firenze, Italy
22 INFN Laboratori Nazionali di Frascati, Frascati, Italy
23 INFN Sezione di Genova, Genova, Italy
24 INFN Sezione di Milano-Bicocca, Milano, Italy
25 INFN Sezione di Milano, Milano, Italy
26 INFN Sezione di Cagliari, Monserrato, Italy
27 INFN Sezione di Padova, Padova, Italy
28 INFN Sezione di Pisa, Pisa, Italy
29 INFN Sezione di Roma Tor Vergata, Roma, Italy
30 INFN Sezione di Roma La Sapienza, Roma, Italy
31 Nikhef National Institute for Subatomic Physics, Amsterdam, Netherlands
32 Nikhef National Institute for Subatomic Physics and VU University Amsterdam, Amsterdam, Netherlands
33 Henryk Niewodniczanski Institute of Nuclear Physics Polish Academy of Sciences, Kraków, Poland
34 AGH — University of Science and Technology, Faculty of Physics and Applied Computer Science, Kraków, Poland
35 National Center for Nuclear Research (NCBJ), Warsaw, Poland
36 Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania
37 Petersburg Nuclear Physics Institute NRC Kurchatov Institute (PNPI NRC KI), Gatchina, Russia
38 Institute of Theoretical and Experimental Physics NRC Kurchatov Institute (ITEP NRC KI), Moscow, Russia, Moscow, Russia