



Search for the lepton-flavour violating decays $B^0 \rightarrow K^{*0} \tau^\pm \mu^\mp$

LHCb collaboration

Abstract

A first search for the lepton-flavour violating decays $B^0 \rightarrow K^{*0} \tau^\pm \mu^\mp$ is presented. The analysis is performed using a sample of proton-proton collision data, collected with the LHCb detector at centre-of-mass energies of 7, 8 and 13 TeV between 2011 and 2018, corresponding to an integrated luminosity of 9 fb^{-1} . No significant signal is observed, and upper limits on the branching fractions are determined to be $\mathcal{B}(B^0 \rightarrow K^{*0} \tau^+ \mu^-) < 1.0 \text{ (1.2)} \times 10^{-5}$ and $\mathcal{B}(B^0 \rightarrow K^{*0} \tau^- \mu^+) < 8.2 \text{ (9.8)} \times 10^{-6}$ at the 90% (95%) confidence level.

Submitted to JHEP

1 Introduction

Lepton-flavour violation (LFV) has been observed in neutral leptons through the phenomenon of neutrino oscillations. In the Standard Model (SM), LFV is negligible in the charged sector [1]. As a consequence, any observation of a charged LFV decay would provide clear evidence for physics beyond the SM. The recent anomalies observed in some universality tests of lepton interactions in $b \rightarrow s\ell^+\ell^-$ transitions [2–5] have led to proposals of many extensions of the SM [6–19], predicting a significant enhancement of charged LFV decays, including $b \rightarrow s\tau\mu$ transitions.

LHCb has performed searches for lepton-flavour violating b -hadron decays into final states with an electron and a muon [20–22] and with a tau and a muon [23, 24]; however, no signal has been observed. The same holds for analogous searches performed by other experiments [25–28]. The most stringent limit on a $b \rightarrow s\tau\mu$ transition is $\mathcal{B}(B_s^0 \rightarrow \tau^\pm\mu^\mp) < 3.4 \times 10^{-5}$ at the 90% confidence level, set by the LHCb experiment [23].

In this article, the first ever search for the charged lepton-flavour violating decay $B^0 \rightarrow K^{*0}\tau^\pm\mu^\mp$, not investigated by any prior experiment, is presented. The K^{*0} meson is reconstructed through its decay into a K^+ and a π^- . The final states $B^0 \rightarrow K^{*0}\tau^-\mu^+$ with the charged kaon and tau having opposite charges and $B^0 \rightarrow K^{*0}\tau^+\mu^-$ with the charged kaon and tau having the same charge (charged conjugate processes are implied throughout) are treated independently. From a theoretical point of view, these channels could be affected differently by model extensions beyond the SM [29], and from the experimental point of view they are affected by different background contributions. The tau lepton is reconstructed through the decays $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$ or $\tau^- \rightarrow \pi^-\pi^+\pi^-\pi^0\nu_\tau$, representing approximately 14% of all tau decays.

The analysis is performed on a data set of proton-proton (pp) collisions collected with the LHCb detector at centre-of-mass energies of 7 and 8 TeV in 2011-2012 (Run 1) and 13 TeV in 2015-2018 (Run 2), corresponding to an integrated luminosity of 9 fb^{-1} . The decay $B^0 \rightarrow D^-D_s^+$ followed by $D^- \rightarrow K^+\pi^-\pi^-$ and $D_s^+ \rightarrow K^+K^-\pi^+$, with a topology similar to the signal decay and a precisely known branching fraction, is used both as a normalisation channel and a control channel to test the reliability of the simulation and evaluate some systematic uncertainties.

2 LHCb detector, trigger and simulation

The LHCb detector [30, 31] 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 charged-particle reconstruction (tracking) system consisting of a silicon-strip vertex detector surrounding the pp interaction region [32], 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 [33, 34] 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/ c . The minimum distance of a track to a primary vertex (PV), the impact parameter (IP), is measured with a resolution of $(15 + 29/p_T) \mu\text{m}$, where p_T is the component of the momentum transverse to the beam, in GeV/ c . Different types of charged hadrons are distinguished from one another

using information from two ring-imaging Cherenkov (RICH) detectors [35]. Photons, electrons and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers [36].

The online event selection is performed by a trigger [37], consisting of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction. In the hardware stage, signal candidates are required to have at least one high- p_T muon. For the normalisation channel, at least one hadron with high transverse energy is required. The software trigger requires a two-, three- or four-track secondary vertex with a significant displacement from any PV. At least one charged particle must have significant transverse momentum and be inconsistent with originating from a PV. A multivariate algorithm [38,39] based on kinematic, geometric and lepton identification criteria is used for the identification of secondary vertices consistent with the decay of a b hadron.

Simulation is used to optimise the selection, determine the signal model for the fit and obtain the selection efficiencies. In the simulation, pp collisions are generated using PYTHIA 8 [40] with a specific LHCb configuration [41]. The B^0 decay is assumed to proceed according to a uniform phase-space model. The tau decay is simulated using the TAUOLA decay library tuned with BaBar data [42]. The decays of all other unstable particles are described by EVTGEN [43], in which final-state radiation is generated using PHOTOS [44]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [45] as described in Ref. [46].

3 Event selection

The $B^0 \rightarrow K^{*0}\tau^\pm\mu^\mp$ candidates are reconstructed by combining six good quality charged tracks, with p smaller than 110 GeV/ c , p_T larger than 250 MeV/ c , η between 2 and 4.9. Two of the tracks of opposite charge, one compatible with a kaon hypothesis and the other with a pion hypothesis, are required to form a K^{*0} candidate, with mass within the range 700 to 1100 MeV/ c^2 and with a good quality vertex. A third track is identified as a muon. The other three remaining tracks, identified as pions, with momentum higher than 2 GeV/ c each, should come from another vertex. These pions form the tau candidate, with charge opposite to that of the muon and with a reconstructed mass within 0.5 to 2.0 GeV/ c^2 . The tau vertex must have a radial distance between 0.1 and 7 mm and a distance along the z -axis larger than 5 mm with respect to the best PV.

The K^{*0} , the muon and the tau candidates should have a transverse momentum greater than 1 GeV/ c each. They form a B^0 candidate, with a vertex of good quality and sufficiently displaced from any PV. Finally, the $K^{*0}\tau$ mass should be lower than 5 GeV/ c^2 and the reconstructed B^0 candidate mass be within 2 to 10 GeV/ c^2 . A sample of data, with tau and muon of the same charge (called same-sign data in the following), is also selected as a proxy for backgrounds.

Neutral particles in the tau decays, namely the neutrino and possibly π^0 s, are not explicitly reconstructed. For this reason, the invariant mass of the six tracks, $m_{K^*\tau\mu}$, does not peak at the B^0 meson mass. The corrected mass $m_{\text{corr}} = \sqrt{p_\perp^2 + m_{K^*\tau\mu}^2} + p_\perp$ is used to recover part of the missing energy, where p_\perp is the component of the missing

momentum perpendicular to the direction of flight of the B meson [47].

Backgrounds can be divided into two categories: combinatorial background, arising from random combinations of tracks; and physics background due to b -hadron decays, which are partially reconstructed and/or reconstructed with particle mis-identification, called physics background. In particular, a large component of background involves D mesons with a decay time compatible with the tau and decaying into multiple charged tracks. In order to optimise the rejection of these backgrounds of a different nature, a multi-stage selection procedure is used. The procedure was developed without looking at candidates in the region where the signal is expected. This region, called the signal region, is defined to be $4.6 < m_{\text{corr}} < 6.4 \text{ GeV}/c^2$. Consequently, the region $6.4 < m_{\text{corr}} < 18.0 \text{ GeV}/c^2$ is called the upper mass sideband, while the region $3.0 < m_{\text{corr}} < 4.6 \text{ GeV}/c^2$ is called the lower mass sideband. The former is dominated by combinatorial background while the latter contains both combinatorial and physics background.

The selection is optimised independently for Run 1 and Run 2 data-taking periods, to account for possible different background levels due to different data-taking conditions. The Punzi figure of merit [48], defined as $\varepsilon/(3/2 + \sqrt{Y_b})$, is used for identifying the optimal selection requirement, apart from the vetoes. Here ε is the efficiency of the signal selection, measured with simulation, and Y_b is the background yield in the signal region, estimated differently depending on the selection step.

The first stage of the selection is based on a multivariate discriminant exploiting the differences between the topologies of the signal decays and the combinatorial background. A boosted decision tree (BDT) [49], using the AdaBoost algorithm from the TMVA package [50], combines information from: the χ^2 of the vertices of the B^0 , K^{*0} and tau candidates; the χ^2 of the flight-distances of the B^0 and the tau candidates. The BDT is trained using simulated signal samples and the upper sideband of the same-sign data sample as a background proxy. A k -folding approach [51] is used to exploit the training samples without biasing the output of the classifiers. For the threshold optimisation, the background yield in the signal region is extrapolated from a fit with a decreasing exponential function to the candidate m_{corr} distribution in the region 11 to $18 \text{ GeV}/c^2$, dominated by combinatorial background.

The second stage is a multivariate selection dedicated to the rejection of charmed mesons mis-identified as tau leptons. The decay of a tau into three charged pions and a neutrino occurs mostly through the a_1^\pm resonance, which in turn decays into a ρ^0 and a π^\pm meson. In order to exploit the kinematic properties of this decay chain, the minimum and maximum of the momenta of the pions from the tau candidate and the masses of pion pairs with zero combined charge are combined into a BDT. The BDT is trained using a k -folding approach, with $B^0 \rightarrow K^{*0}\tau^\pm\mu^\mp$ simulation and inclusive $b\bar{b}$ simulated events reconstructed as $B^0 \rightarrow K^{*0}\tau^\pm\mu^\mp$ as a proxy for the background. Particle identification requirements and K^{*0} kinematic constraints are removed, so that particles from charmed mesons in the $b\bar{b}$ sample can form tau and K^{*0} candidates. For the threshold optimisation, the background yield is extrapolated from a fit to the B^0 candidate mass distribution in the 3 to $18 \text{ GeV}/c^2$ mass range, excluding the signal region. The fit model used is the sum of a Crystal-Ball function [52], modelling the lower mass sideband, dominated by partially reconstructed events, and an exponential function, with a slope primarily determined by events in the upper mass sideband. The reliability of this background yield estimation is verified on the same-sign data, where the prediction can be compared to the observed number of candidates in the signal region.

The third selection step consists of requirements on the muon, kaon and pion particle-identification variables. A three-dimensional scan over the possible thresholds of the three selection criteria is performed, and a configuration maximising the Punzi figure of merit is chosen. For this optimisation, the background yields evaluated in the previous step are scaled by the rejection of the particle-identification requirements, as estimated from the same-sign sample within the signal region.

The search sensitivity is further increased by exploiting the discriminating power of the mass of K^{*0} candidates, required to be within 856 to 946 MeV/ c^2 . In addition, the mass of the tau candidate is calculated using the known mass of the pion [53] and considering only the pions and neutrino momentum components orthogonal to the B^0 direction of flight. The neutrino component is measured as the difference between the momentum orthogonal components of the $K^{*0}\mu$ system and the three pions system. This mass is required to be within 789 to 1900 MeV/ c^2 . Both mass regions have been optimised using the Punzi figure of merit, as in the previous step. A particle selected from a partially reconstructed decay is oftentimes surrounded by other particles which are not used to reconstruct the candidate. To exploit this feature, a Fisher discriminant [54] is built, using the k -folding procedure and based on the following variables evaluating the particle isolation: the logarithm of the smallest variation of the B^0 , the K^{*0} or the tau vertex χ^2 when adding one or two tracks to them, independently; and the invariant mass of all the particles forming the vertex under the hypotheses of such additions; the number of tracks compatible with the aforementioned vertices (*i.e.* whose addition keeps the overall χ^2 below 9). The Fisher discriminant is chosen for its stability in the training procedure, despite the small size of the training samples available at this point of the selection. Signal simulation and, for the background, same-sign data that fall into the signal mass region, are used for the training and for the optimisation of the threshold.

A powerful variable to reject the physics backgrounds is the tau flight-distance χ^2 ; it is already used to reject the combinatorial background, which tends to have larger values than the signal. It also offers residual discriminating power against physics backgrounds from charmed mesons, for example, $B^0 \rightarrow D^{*-}\mu^+\nu$ decays for the $B^0 \rightarrow K^{*0}\tau^-\mu^+$ decay channel. Given the different nature of the physics backgrounds, the thresholds applied to the $B^0 \rightarrow K^{*0}\tau^-\mu^+$ and $B^0 \rightarrow K^{*0}\tau^+\mu^-$ channels are optimised separately.

The remaining physics backgrounds are studied with events rejected by the first BDT and within the region $4.7 < m_{\text{corr}} < 5.7$ GeV/ c^2 . Mass distributions from possible final state particle combinations are used to define vetoes that remove the observed resonant structures.

For the $B^0 \rightarrow K^{*0}\tau^-\mu^+$ case, an important physics background comes from the decay $B^0 \rightarrow D^{*-}\mu^+\nu$, with $D^{*-} \rightarrow \bar{D}^0\pi^-$ and $\bar{D}^0 \rightarrow K^+\pi^-\pi^+\pi^-$. Three of the four pions can be mis-reconstructed as a tau candidate, while the remaining pion and kaon are used to mis-reconstruct a K^{*0} candidate. To reduce the backgrounds from D^0 mesons, events with $m(K^+\tau^-)$ or $m(K^{*0}\pi^-\pi^+)$ within 60 MeV/ c^2 from the known D^0 mass are removed. This requirement additionally rejects most of the D^{*0} contributions. However, considering the minimal impact on the signal efficiency, events with $m(K^{*0}\tau^-) - m(K^{*0}\pi^-\pi^+)$ or $m(K^{*0}\tau^-) - m(K^+\tau^-)$ between 135.5 and 155.5 MeV/ c^2 are also rejected. Masses accounting for a possible mis-identification of the muon as a pion have been checked, but no significant contributions from charmed mesons have been observed.

For $B^0 \rightarrow K^{*0}\tau^+\mu^-$ decays, a significant source of physics background at this stage of the selection comes from the decay $B^0 \rightarrow D^{*-}\tau^+\nu_\tau$, with $D^{*-} \rightarrow \bar{D}^0\pi^-$, $\bar{D}^0 \rightarrow$

$K^+\mu^-\bar{\nu}_\mu$ and $\tau^+ \rightarrow \pi^+\pi^-\pi^+\bar{\nu}_\tau$. This background is rejected by selecting only events with $m(K^+\mu^-) > 1885 \text{ MeV}/c^2$.

After the selection procedure described above, there is never more than one B^0 candidate selected per event.

4 Normalisation channel

The normalisation channel $B^0 \rightarrow D^-D_s^+$, with $D^- \rightarrow K^+\pi^-\pi^-$ and $D_s^+ \rightarrow K^+K^-\pi^+$, is reconstructed using six good-quality tracks, with $2 < p < 110 \text{ GeV}/c$, $p_T > 250 \text{ MeV}/c$, $2 < \eta < 4.9$ and having associated deposits in the RICH detectors. Three of these tracks, one identified as kaon and the other two as pions, are used to form a D^- candidate with a displaced vertex of good quality and an invariant mass lying within the range 1750 to 2080 MeV/c^2 . Analogously, one track identified as pion and other two tracks identified as kaons form the D_s^+ candidate, with a displaced vertex of good quality and an invariant mass within the range 1938 to 1998 MeV/c^2 . Both the D^- and D_s^+ candidates are required to have $p_T > 1 \text{ GeV}/c$.

Combinatorial background is rejected by a BDT exploiting the same topological variables used in the selection of the signal channel, but with two flight-distance significances, one for each D meson. The BDT is trained using simulated $B^0 \rightarrow D^-D_s^+$ decays and $B^0 \rightarrow D^-D_s^+$ candidates with a mass exceeding 5400 MeV/c^2 as a background proxy. A k -folding procedure is applied. Considering that a significant signal is expected for the normalisation mode, the selection optimisation is based on the figure of merit $Y_s/\sqrt{Y_s+Y_b}$, where Y_s and Y_b are the yields of the signal and background, respectively. These yields are determined from fits in the B^0 mass range 5150 to 5400 MeV/c^2 , using a Gaussian model to parameterise the signal component and a decreasing exponential function to describe the background.

The same particle identification requirements used for selecting the signal are applied on kaons and pions. For each charmed meson, the difference between its reconstructed mass and known mass [55] must be less than 20 MeV/c^2 .

Finally, the normalisation channel yields are measured from fits of the B^0 mass distribution for each year of data taking. The global fits for Run 1 and Run 2 data, leading to yields of 1155 ± 35 and 6516 ± 84 , respectively, are shown in Fig. 1.

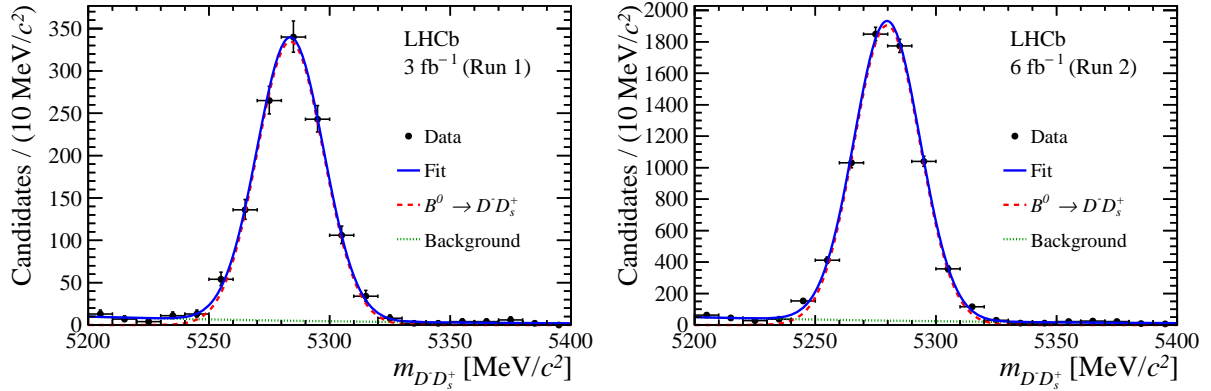


Figure 1: Mass distributions of selected $D^- D_s^+$ candidates in (left) Run 1 and (right) Run 2 data, with the fit overlaid (blue solid line). The red dashed line is signal, while the green dotted line is background.

5 Determination of efficiencies

The efficiencies for the different steps of the selection chain are evaluated separately for each year of data taking. With the exception of the particle identification efficiency, they are estimated with simulation. The reliability of the simulation is checked using B^0 , D^- and D_s^+ mesons from the normalisation channel $B^0 \rightarrow D^- D_s^+$ (with $D^- \rightarrow K^+ \pi^- \pi^-$ and $D_s^+ \rightarrow K^+ K^- \pi^+$) as proxies for two- and three-prong particle vertices, selected with the additional requirement that the B^0 candidate mass lies within $\pm 50 \text{ MeV}/c^2$ from its nominal value. All the variables used in the multivariate selections are found to be well described by simulation. Residual discrepancies are assessed as a systematic uncertainty.

The particle identification variables are not perfectly described in simulation. Therefore, the particle identification efficiency is evaluated using data calibration samples. A pure sample of pions and kaons is obtained from the decay $D^{*+} \rightarrow (D^0 \rightarrow K \pi) \pi^+$ while muons are selected from the decay $J/\psi \rightarrow \mu^+ \mu^-$, without relying on PID selection criteria [56]. The efficiencies of the PID requirements can then be computed in intervals of the kinematic variables of the final-state-particles momentum and pseudorapidity [57]. Intervals are chosen making a compromise between the stability of the efficiency within the interval and statistical uncertainty. The event occupancy, parameterised by the number of tracks per event, is also taken into account, relying on same-sign events in the case of signal and on the observed distribution for the normalisation channel.

The distribution of the selection efficiency as a function of the kinematic variables $m_{K^*0\mu^\pm}^2$ and $m_{\tau^\mp\mu^\pm}^2$, normalised to unity, is shown in Fig. 2. These distributions can be used to re-cast the results for models having $B^0 \rightarrow K^{*0} \tau^- \mu^+$ and $B^0 \rightarrow K^{*0} \tau^+ \mu^-$ decay kinematics different from the uniform distribution in the phase space used for the signal simulation in this analysis.

6 Strategy to fit the corrected mass distributions

An extended maximum likelihood fit to the m_{corr} distributions of the selected events is performed independently for the $B^0 \rightarrow K^{*0} \tau^- \mu^+$ and $B^0 \rightarrow K^{*0} \tau^+ \mu^-$ decay channels. For each channel the distribution is described by a function P_{tot} which is the sum of

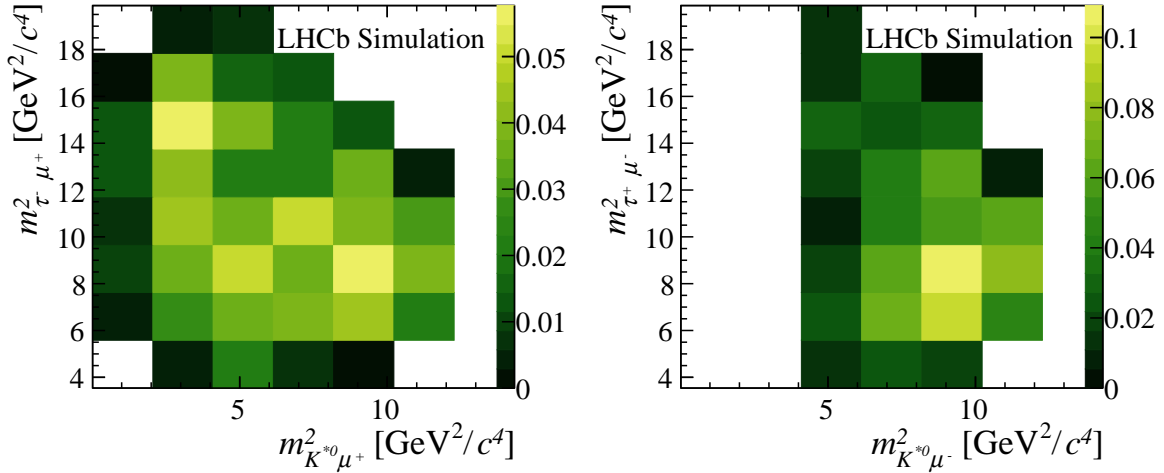


Figure 2: Efficiency as a function of $m_{K^*0\mu^\pm}^2$ and $m_{\tau^\mp\mu^\pm}^2$ for the (left) $B^0 \rightarrow K^*0\tau^-\mu^+$ and (right) $B^0 \rightarrow K^*0\tau^+\mu^-$ cases, normalised to unity. The selection requirement $m(K^+\mu^-) > 1885 \text{ MeV}/c^2$ for the $B^0 \rightarrow K^*0\tau^+\mu^-$ decay implies zero efficiency for $m_{K^*0\mu^-}^2 < 4.1 \text{ GeV}^2/c^4$ on the right plot.

three components, $P_{\text{tot}} = Y_{\tau_{3\pi}} P_{\tau_{3\pi}} + Y_{\tau_{3\pi\pi^0}} P_{\tau_{3\pi\pi^0}} + Y_{\text{bkg}} P_{\text{bkg}}$. The function $P_{\tau_{3\pi}}$ models the dominant signal contribution with $\tau^- \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$ and $Y_{\tau_{3\pi}}$ indicates the corresponding yields. Analogously, $P_{\tau_{3\pi\pi^0}}$ models the subdominant signal contribution with $\tau^- \rightarrow \pi^- \pi^+ \pi^- \pi^0 \nu_\tau$ and $Y_{\tau_{3\pi\pi^0}}$ represents the corresponding yields. Finally, P_{bkg} models the background, with yields Y_{bkg} .

The signal components $P_{\tau_{3\pi}}$ and $P_{\tau_{3\pi\pi^0}}$ are parameterised independently, using double-sided Crystal-Ball functions (DSCB), *i.e.* a crystal ball with an additional tail on the right side. The parameters are determined and fixed from fits to the simulated signal samples. The yields for the dominant signal contribution are expressed, for the $B^0 \rightarrow K^*0\tau^-\mu^+$ channel and separately for Run 1 and Run 2, as

$$Y_{\tau_{3\pi}}^{\text{run}} = \frac{\mathcal{B}(B^0 \rightarrow K^*0\tau^-\mu^+) \cdot \mathcal{B}(K^*0 \rightarrow K^+\pi^-) \cdot \mathcal{B}(\tau^- \rightarrow \pi^- \pi^+ \pi^- \nu_\tau)}{\mathcal{B}(B^0 \rightarrow D^- D_s^+) \cdot \mathcal{B}(D^- \rightarrow K^+\pi^-\pi^-) \cdot \mathcal{B}(D_s^+ \rightarrow K^+ K^- \pi^+)} \cdot \sum_{y \in \text{run}} \left(\frac{\varepsilon Y_N}{\varepsilon_N} \right)_y. \quad (1)$$

Analogous expressions hold for the $B^0 \rightarrow K^*0\tau^+\mu^-$ channel, and for the subdominant signal components. Here, ε represents the signal efficiency, while Y_N and ε_N indicate the yields and the efficiency for the normalisation channel. The ratio $\varepsilon Y_N / \varepsilon_N$ is calculated for each year in the run, as indicated by the y index. The parameter of interest $\mathcal{B}(B^0 \rightarrow K^*0\tau^-\mu^+)$ is in common among the yields of the dominant and subdominant signal contributions, as well as among runs of data taking, and is therefore fitted simultaneously. A fit bias on the measured signal branching ratio, evaluated on background-only pseudoexperiments, is at the level of 10^{-7} and is subtracted. The other branching fractions are taken from Ref. [53]. Apart from the signal branching fraction, all the parameters are constrained with Gaussian functions to account for the systematic uncertainties described in the next section.

The functional form used to parameterise the background shape is also a DSCB function. The background yields are free to vary in the fit, while the shape parameters

are constrained to the value determined on a control sample obtained by loosening the combinatorial multivariate selection. Signal contamination in the control regions is estimated to be less than 5% of the total signal, assuming a branching fraction of 10^{-5} .

7 Systematic uncertainties

For the determination of the limit, some branching fractions are used as external inputs. A systematic uncertainty is assigned imposing Gaussian constraints on their values, taken from the PDG [53].

The normalisation procedure involves ratios of signal and normalisation efficiencies, for which systematic uncertainties cancel. However, some efficiencies do not cancel in the ratio and are assessed as follows.

The uncertainty related to the limited size of the simulation samples used to determine the efficiencies is included as part of the statistical uncertainty.

The ratio between the signal and the normalisation tracking efficiencies is determined from simulation. Possible differences in data with respect to simulation will be of the same order in the numerator and denominator for five out of the six particles in the final state, and will cancel. Because the sixth hadron in the normalisation channel interacts differently with the material with respect to the muon in the signal channel, a 1.4% uncertainty is assigned to the tracking efficiency ratios [58].

The systematic uncertainty on the determination of the PID selection efficiency accounts for: the limited size of simulation and calibration samples; the choice of the p , η and number of tracks interval sizes; and the use of the *sPlot* technique [59] to extract the signal yield in the control samples.

Although the simulation of the variables used in the multivariate classifiers is validated with the normalisation channel, small residual discrepancies could affect the evaluation of the classifier efficiencies. To account for this, the $B^0 \rightarrow D^- D_s^+$ candidates with a mass within $50 \text{ MeV}/c^2$ of the known mass [53] of the B^0 meson are selected from both data and simulation, providing a high purity sample on which the classifiers used to select the $B^0 \rightarrow K^{*0} \tau \mu$ signal are applied. The variables of the classifiers relative to the tau are applied to the D^- meson and analogously those of the K^{*0} are applied to the D_s^+ meson. For each classifier, the requirement giving an efficiency on the $B^0 \rightarrow D^- D_s^+$ simulation sample equivalent to that obtained by the requirement on the simulated $B^0 \rightarrow K^{*0} \tau \mu$ sample is determined. The relative systematic uncertainty is the absolute value of the difference in efficiencies of this requirement between $B^0 \rightarrow D^- D_s^+$ simulation and data, divided by the $B^0 \rightarrow D^- D_s^+$ simulation efficiency.

Over the entire data taking period, changes were applied to the hardware muon trigger, in particular regarding the requirement on the transverse momentum of the muon. To account for these changes, the hardware muon trigger efficiency evaluated from simulation is multiplied by a correction factor derived as the ratio between the efficiency from data and simulation for the $B^\pm \rightarrow J/\psi K^\pm$ control sample. These corrections are functions of the muon transverse momentum. The difference between the efficiency resulting from this weighting procedure and the efficiency from simulation is taken as a systematic uncertainty. The overall effect is less than 1%.

The hardware hadron trigger, used in the selection of the normalisation channel, is less well reproduced in simulation. A data-driven efficiency estimation is provided by

Table 1: Relative increase (in %) of the observed upper limit when applying each of the systematic uncertainties.

Systematic effect	Limit increase [%]	
	$B^0 \rightarrow K^{*0}\tau^-\mu^+$	$B^0 \rightarrow K^{*0}\tau^+\mu^-$
Input branching fractions	4	3
Efficiencies	2	1
Normalisation yields	1	1
Background control region choice	18	26
Background analytical shape	1	1

the number of events triggered simultaneously by any muon in the event and one of the hadrons in the $B^0 \rightarrow D^- D_s^+$ decay, divided by the number of events triggered by any muon in the event. The absolute difference with the efficiency estimated on the $B^0 \rightarrow D^- D_s^+$ simulation is assumed as a systematic uncertainty.

The systematic uncertainty associated to the fit model used for measuring the normalisation channel yields (Y_n) is assessed via pseudoexperiments, generated using a kernel estimation of the $B^0 \rightarrow D^- D_s^+$ mass distribution and subsequently fit using the nominal normalisation channel fit model. The relative systematic uncertainty is the ratio between the fitted mean of the yields with respect to the mean of the generated values, and is 1.8% for Run 1 and 1.7% for Run 2.

The parameters of the DSCB function describing the background are Gaussian constrained to the values determined by the fit of the background control region, obtained by loosening the combinatorial multivariate selection. Alternative control regions are defined by choosing different requirements on the combinatorial multivariate discriminant. The mean value and width of the DSCB are Gaussian constrained to half the maximum spread of the results obtained from fits on the alternative control regions, if this is larger than the parameter uncertainty on the default control region. The systematic uncertainty for the choice of the DSCB functional form is assessed by generating samples using a kernel estimation of background control samples, and fitting with the default DSCB function. A bias term is added to each of the signal branching ratios, and its value is constrained to a Gaussian function with a mean of zero and a width set to the average value of the bias measured on pseudoexperiments, which is at the level of 6×10^{-7} .

The effect of the systematic uncertainties is summarised in Tab. 1, where the increase of the observed upper limit when applying each of the systematic uncertainties in addition to the others is shown. The dominant systematic effect comes from the uncertainty on the arbitrary choice of the control region.

8 Results

The results of the extended maximum likelihood fit to the m_{corr} distributions are shown in Figs. 3 and 4 for $B^0 \rightarrow K^{*0}\tau^+\mu^-$ and $B^0 \rightarrow K^{*0}\tau^-\mu^+$, respectively. No significant signal contribution is observed. Therefore, upper limits on the branching fractions are set via the CL_s method [60, 61], using the asymptotic approximation:

$\mathcal{B}(B^0 \rightarrow K^{*0}\tau^+\mu^-) < 1.0$ (1.2) $\times 10^{-5}$ and $\mathcal{B}(B^0 \rightarrow K^{*0}\tau^-\mu^+) < 8.2$ (9.8) $\times 10^{-6}$ at the 90% (95)% confidence level, as shown in Fig. 5.

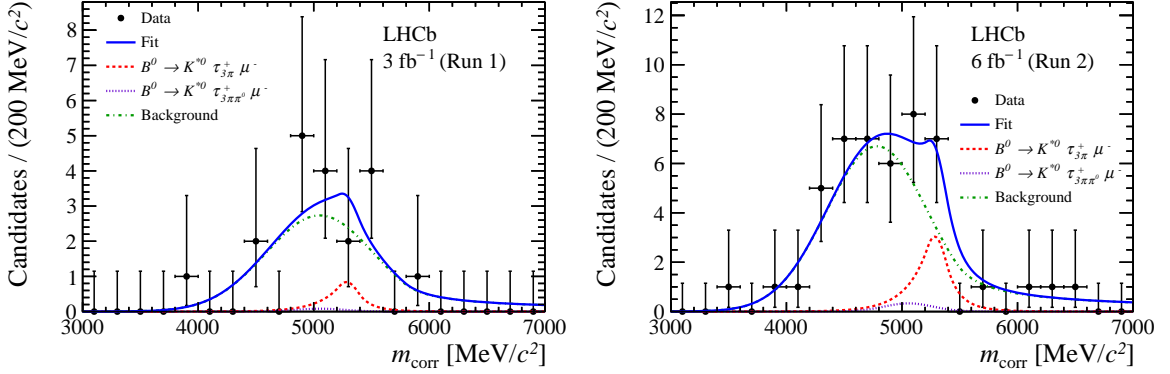


Figure 3: Distribution of the corrected mass m_{corr} of selected $B^0 \rightarrow K^{*0}\tau^+\mu^-$ candidates in (left) Run 1 and (right) Run 2 data, with the simultaneous fit overlaid (blue solid line). The red dashed line is the dominant signal component with $\tau^+ \rightarrow \pi^+\pi^-\pi^+\nu_\tau$; the violet dotted line is the subdominant signal component with $\tau^+ \rightarrow \pi^+\pi^-\pi^+\pi^0\nu_\tau$, extremely small and consequently barely visible; and the green dash-dotted line is the background.

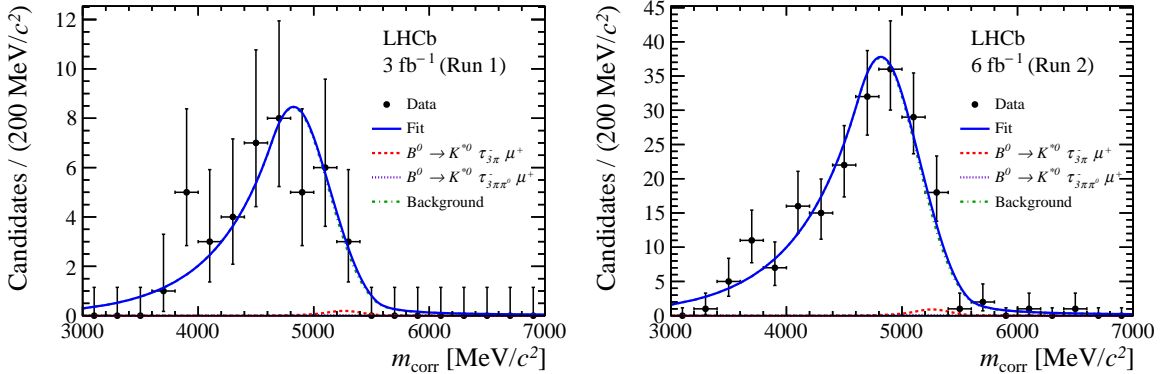


Figure 4: Distribution of the corrected mass m_{corr} of selected $B^0 \rightarrow K^{*0}\tau^-\mu^+$ candidates in (left) Run 1 and (right) Run 2 data, with the simultaneous fit overlaid (blue solid line). The red dashed line is the dominant signal component with $\tau^- \rightarrow \pi^-\pi^+\pi^-\nu_\tau$; the violet dotted line is the subdominant signal component with $\tau^- \rightarrow \pi^-\pi^+\pi^-\pi^0\nu_\tau$, extremely small and consequently barely visible; and the green dash-dotted line is the background.

9 Conclusion

The first search for the lepton-flavour violating decays $B^0 \rightarrow K^{*0}\tau^\pm\mu^\mp$ has been performed on a sample of proton-proton collision data, collected with the LHCb detector at centre-of-mass energies of 7, 8 and 13 TeV between 2011 and 2018, corresponding to an integrated luminosity of 9fb^{-1} . No significant signal is observed, and upper limits on the branching fractions are set: $\mathcal{B}(B^0 \rightarrow K^{*0}\tau^+\mu^-) < 1.0$ (1.2) $\times 10^{-5}$ and $\mathcal{B}(B^0 \rightarrow K^{*0}\tau^-\mu^+) < 8.2$ (9.8) $\times 10^{-6}$ at the 90% (95%) confidence level. These results

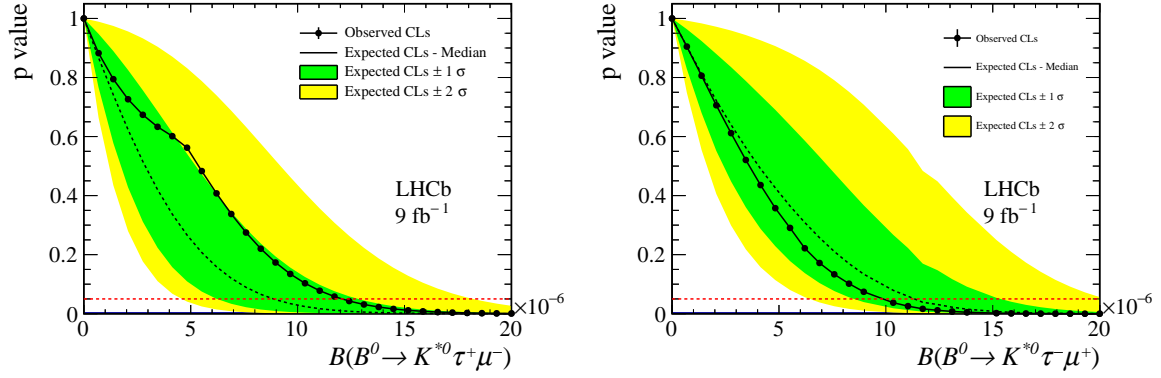


Figure 5: The expected and observed p-values derived with the CL_s method as a function of the (left) $B^0 \rightarrow K^{*0} \tau^+ \mu^-$ and (right) $B^0 \rightarrow K^{*0} \tau^- \mu^+$ branching fraction. The red line corresponds to the 95% CL.

assume a uniform distribution of the signal events within the phase space accessible to the K^{*0} , tau and muon. They are currently the most stringent upper limits on $b \rightarrow s \tau \mu$ transitions.

References

- [1] M. Raidal *et al.*, *Flavour physics of leptons and dipole moments*, Eur. Phys. J. **C57** (2008) 13, arXiv:0801.1826.
- [2] LHCb collaboration, R. Aaij *et al.*, *Tests of lepton universality using $B^0 \rightarrow K_S^0 \ell^+ \ell^-$ and $B^+ \rightarrow K^{*+} \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **128** (2022) 191802, arXiv:2110.09501.
- [3] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in beauty-quark decays*, arXiv:2103.11769, to appear in Nature Physics.
- [4] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $\Lambda_b^0 \rightarrow p K^- \ell^+ \ell^-$ decays*, JHEP **05** (2020) 040, arXiv:1912.08139.
- [5] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays*, JHEP **08** (2017) 055, arXiv:1705.05802.
- [6] S. L. Glashow, D. Guadagnoli, and K. Lane, *Lepton flavor violation in B decays?*, Phys. Rev. Lett. **114** (2015) 091801, arXiv:1411.0565.
- [7] B. Bhattacharya, A. Datta, D. London, and S. Shivashankara, *Simultaneous explanation of the R_K and $R(D^{(*)})$ puzzles*, Phys. Lett. **B742** (2015) 370, arXiv:1412.7164.
- [8] D. Bečirević, O. Sumensari, and R. Zukanovich Funchal, *Lepton flavor violation in exclusive $b \rightarrow s$ decays*, Eur. Phys. J. **C76** (2016) 134, arXiv:1602.00881.
- [9] A. Crivellin *et al.*, *Lepton-flavour violating B decays in generic Z' models*, Phys. Rev. **D92** (2015) 054013, arXiv:1504.07928.

- [10] I. de Medeiros Varzielas and G. Hiller, *Clues for flavor from rare lepton and quark decays*, JHEP **06** (2015) 072, [arXiv:1503.01084](#).
- [11] D. Bečirević, N. Košnik, O. Sumensari, and R. Zukanovich Funchal, *Palatable leptoquark scenarios for lepton flavor violation in exclusive $b \rightarrow s\ell_1\ell_2$ modes*, JHEP **11** (2016) 035, [arXiv:1608.07583](#).
- [12] A. D. Smirnov, *Vector leptoquark mass limits and branching ratios of $K_L^0, B^0, B_s \rightarrow l_i^+ l_j^-$ decays with account of fermion mixing in leptoquark currents*, Mod. Phys. Lett. **A33** (2018) 1850019, [arXiv:1801.02895](#).
- [13] R. Barbieri, G. Isidori, A. Pattori, and F. Senia, *Anomalies in B-decays and U(2) flavour symmetry*, Eur. Phys. J. **C76** (2016) 67, [arXiv:1512.01560](#).
- [14] M. Bordone, C. Cornella, J. Fuentes-Martin, and G. Isidori, *A three-site gauge model for flavor hierarchies and flavor anomalies*, Phys. Lett. **B779** (2018) 317, [arXiv:1712.01368](#).
- [15] M. Bordone, C. Cornella, J. Fuentes-Martín, and G. Isidori, *Low-energy signatures of the PS³ model: from B-physics anomalies to LFV*, JHEP **10** (2018) 148, [arXiv:1805.09328](#).
- [16] M. Duraisamy, S. Sahoo, and R. Mohanta, *Rare semileptonic $B \rightarrow K(\pi)l_i^- l_j^+$ decay in a vector leptoquark model*, Phys. Rev. **D95** (2017) 035022, [arXiv:1610.00902](#).
- [17] L. Di Luzio, A. Greljo, and M. Nardecchia, *Gauge leptoquark as the origin of B-physics anomalies*, Phys. Rev. **D96** (2017) 115011, [arXiv:1708.08450](#).
- [18] L. Di Luzio *et al.*, *Maximal Flavour Violation: a Cabibbo mechanism for leptoquarks*, JHEP **11** (2018) 081, [arXiv:1808.00942](#).
- [19] C. Cornella, J. Fuentes-Martin, and G. Isidori, *Revisiting the vector leptoquark explanation of the B-physics anomalies*, JHEP **07** (2019) 168, [arXiv:1903.11517](#).
- [20] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour violating decays $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$* , JHEP **03** (2018) 078, [arXiv:1710.04111](#).
- [21] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour violating decays $B^+ \rightarrow K^+ \mu^\pm e^\mp$* , Phys. Rev. Lett. **123** (2019) 231802, [arXiv:1909.01010](#).
- [22] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour violating decays $B^0 \rightarrow K^{*0} \mu^\pm e^\mp$ and $B_s^0 \rightarrow \phi \mu^\pm e^\mp$* , LHCb-PAPER-2022-008, in preparation.
- [23] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour-violating decays $B_s^0 \rightarrow \tau^\pm \mu^\mp$ and $B^0 \rightarrow \tau^\pm \mu^\mp$* , Phys. Rev. Lett. **123** (2019) 211801, [arXiv:1905.06614](#).
- [24] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton flavour violating decay $B^+ \rightarrow K^+ \mu^- \tau^+$ using B_s^{*0} decays*, JHEP **06** (2020) 129, [arXiv:2003.04352](#).

- [25] BELLE collaboration, S. Choudhury *et al.*, *Test of lepton flavor universality and search for lepton flavor violation in $B \rightarrow K\ell\ell$ decays*, JHEP **03** (2021) 105, [arXiv:1908.01848](#).
- [26] BaBar collaboration, B. Aubert *et al.*, *Measurements of branching fractions, rate asymmetries, and angular distributions in the rare decays $B \rightarrow K\ell^+\ell^-$ and $B \rightarrow K^*\ell^+\ell^-$* , Phys. Rev. **D73** (2006) 092001, [arXiv:hep-ex/0604007](#).
- [27] BaBar collaboration, J. P. Lees *et al.*, *Search for lepton-number violating $B^+ \rightarrow X^-\ell^+\ell^+$ decays*, Phys. Rev. **D89** (2014) 011102, [arXiv:1310.8238](#).
- [28] BaBar collaboration, J. P. Lees *et al.*, *A search for the decay modes $B^\pm \rightarrow h^\pm\tau^\pm\ell$* , Phys. Rev. **D86** (2012) 012004, [arXiv:1204.2852](#).
- [29] M. Bordone, O. Catà, and T. Feldmann, *Effective theory approach to new physics with flavour: general framework and a leptoquark example*, JHEP **01** (2020) 067, [arXiv:1910.02641](#).
- [30] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.
- [31] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, [arXiv:1412.6352](#).
- [32] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, [arXiv:1405.7808](#).
- [33] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, [arXiv:1311.3893](#).
- [34] P. d'Argent *et al.*, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, JINST **12** (2017) P11016, [arXiv:1708.00819](#).
- [35] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, [arXiv:1211.6759](#).
- [36] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, [arXiv:1211.1346](#).
- [37] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, [arXiv:1211.3055](#).
- [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. Likhomanenko *et al.*, *LHCb topological trigger reoptimization*, J. Phys. Conf. Ser. **664** (2015) 082025.
- [40] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, [arXiv:0710.3820](#).
- [41] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

- [42] I. M. Nugent *et al.*, *Resonance chiral lagrangian currents and experimental data for $\tau^- \rightarrow \pi^- \pi^- \pi^+ \nu_\tau$* , Phys. Rev. **D88** (2013) 093012.
- [43] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [44] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS interface in C++: Technical and physics documentation*, Comp. Phys. Comm. **199** (2016) 86, arXiv:1011.0937.
- [45] 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.
- [46] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.
- [47] SLD collaboration, K. Abe *et al.*, *Measurement of the b quark fragmentation function in Z^0 decays*, Phys. Rev. **D65** (2002) 092006, arXiv:hep-ex/0202031, [Erratum: Phys. Rev. **D66** (2002) 079905].
- [48] G. Punzi, *Sensitivity of searches for new signals and its optimization*, eConf **C030908** (2003) MODT002, arXiv:physics/0308063.
- [49] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.
- [50] P. Speckmayer, A. Hocker, J. Stelzer, and H. Voss, *The toolkit for multivariate data analysis, TMVA 4*, J. Phys. Conf. Ser. **219** (2010) 032057.
- [51] A. Bevan, R. G. Goñi, T. Stevenson, and T. Stevenson, *Support vector machines and generalisation in HEP*, J. Phys. Conf. Ser. **898** (2017) 072021, arXiv:1702.04686.
- [52] 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-F31-86-02.
- [53] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020) 083C01.
- [54] R. A. Fisher, *The use of multiple measurements in taxonomic problems*, Annals of Eugenics **7** (1936) 179.
- [55] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.
- [56] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, Eur. Phys. J. Tech. Instr. **6** (2019) 1, arXiv:1803.00824.
- [57] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021, 2016.
- [58] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, arXiv:1408.1251.

- [59] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0402083](#).
- [60] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. **G28** (2002) 2693.
- [61] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, [arXiv:1007.1727](#), [Erratum: Eur. Phys. J. **C73** (2013) 2501].