EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)

Search for the lepton-flavour violating decays $B^0 \to K^{*0}\mu^{\pm}e^{\mp}$ and $B_s^0 \to \phi\mu^{\pm}e^{\mp}$

LHCb collaboration

Abstract

A search for the lepton-flavour violating decays $B^0 \to K^{*0}\mu^{\pm}e^{\mp}$ and $B_s^0 \to \phi\mu^{\pm}e^{\mp}$ is presented, using proton-proton collision data collected by the LHCb detector at the LHC, corresponding to an integrated luminosity of $9\,\text{fb}^{-1}$. No significant signals are observed and upper limits of

$$\mathcal{B}(B^0 \to K^{*0}\mu^{+}e^{-}) < 5.7 \times 10^{-9}\ (6.9 \times 10^{-9}),$$
$$\mathcal{B}(B^0 \to K^{*0}\mu^{-}e^{+}) < 6.8 \times 10^{-9}\ (7.9 \times 10^{-9}),$$
$$\mathcal{B}(B^0 \to K^{*0}\mu^{\pm}e^{\mp}) < 10.1 \times 10^{-9}\ (11.7 \times 10^{-9}),$$
$$\mathcal{B}(B_s^0 \to \phi\mu^{\pm}e^{\mp}) < 16.0 \times 10^{-9}\ (19.8 \times 10^{-9})$$

are set at 90% (95%) confidence level. These results constitute the world's most stringent limits to date, with the limit on the decay $B_s^0 \to \phi\mu^{\pm}e^{\mp}$ the first being set. In addition, limits are reported for scalar and left-handed lepton-flavour violating New Physics scenarios.

1   Introduction

Processes that are forbidden or strongly suppressed in the Standard Model (SM) are sensitive to new heavy particles beyond the SM, and can probe energy scales beyond those accessible with direct searches. Lepton-flavour violating (LFV) decays are forbidden in the SM, but the observation of neutrino oscillations shows the existence of LFV in the neutral lepton sector. An observation of LFV decays involving charged leptons would constitute a clear and unambiguous sign of New Physics (NP).

Recently, studies of rare decays of b-hadrons have received considerable attention due to the appearance of the flavour anomalies in rare $b \to s\ell^+\ell^-$ transitions [1–22]. The tensions with SM predictions seen in lepton flavour universality tests [13–22], in particular, motivate searches for LFV b-hadron decays, as lepton flavour non-universality is closely connected with LFV [23]. Specific NP scenarios that can induce LFV b-hadron decays include models with scalar or vector leptoquarks [24–26], and models with additional Z' bosons [27]. Branching fractions for $b \to s\mu^\pm e^\mp$ decays like $B^0 \to K^{*0}\mu^\pm e^\mp$ could be as large as $\mathcal{O}(10^{-7})$ [27], close to the currently best limit of $\mathcal{B}(B^0 \to K^{*0}\mu^\pm e^\mp) < 1.8 \times 10^{-7}$ at 90% confidence level (CL) obtained by the Belle collaboration [28].

The LHCb collaboration has performed searches for $b \to s\mu^\pm e^\mp$ and $b \to d\mu^\pm e^\mp$ transitions using the decays $B^0_{(s)} \to \mu^\pm e^\mp$ [29] and $B^+ \to K^+\mu^\pm e^\mp$ [30], resulting in exclusion limits of $\mathcal{B}(B^0 \to \mu^\pm e^\mp) < 1.0 \times 10^{-9}$, $\mathcal{B}(B^0_s \to \mu^\pm e^\mp) < 5.4 \times 10^{-9}$, $\mathcal{B}(B^+ \to K^+\mu^+e^-) < 6.4 \times 10^{-9}$, and $\mathcal{B}(B^+ \to K^+\mu^-e^+) < 7.0 \times 10^{-9}$ at 90% CL. Both analyses were performed using the LHCb Run 1 data sample, corresponding to an integrated luminosity of $3\,\mathrm{fb}^{-1}$.

This paper presents a search for the lepton-flavour violating decays $B^0 \to K^{*0}(\to K^+\pi^-)\mu^\pm e^\mp$ and $B^0_s \to \phi(\to K^+K^-)\mu^\pm e^\mp$. The inclusion of charge-conjugate processes is implied throughout. The symbols K^{*0} and ϕ refer to the $K^*(892)^0$ and $\phi(1020)$ vector mesons. The search uses the data sample collected by the LHCb experiment in proton-proton (pp) collisions at centre-of-mass energies of $7\,\mathrm{TeV}$ (2011), $8\,\mathrm{TeV}$ (2012), and $13\,\mathrm{TeV}$ (2015–2018), corresponding to a total integrated luminosity of $9\,\mathrm{fb}^{-1}$. As NP models can affect $b \to s\mu^+e^-$ and $b \to s\mu^-e^+$ transitions differently, limits for the decays $B^0 \to K^{*0}\mu^+e^-$ (referred to as same-sign due to muon and kaon charge being equal) and $B^0 \to K^{*0}e^+\mu^-$ (referred to as opposite-sign) are also reported separately.

The tree level decays $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$ and $B^0_s \to J/\psi(\to \mu^+\mu^-)\phi$, which exhibit large yields and have a final state similar to the signal decays, are used as normalisation channels. To avoid experimenters' bias, the candidates in the signal regions, defined using the invariant masses of the final state as $m(K^+\pi^-\mu^\pm e^\mp) \in [4900, 5600]\,\mathrm{MeV}/c^2$ and $m(K^+K^-\mu^\pm e^\mp) \in [4900, 5600]\,\mathrm{MeV}/c^2$, were not examined until the selections and limit setting procedures were finalised.

2   Detector and simulation

The LHCb detector [31, 32] 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 [33], a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about $4\,\mathrm{Tm}$, and

three stations of silicon-strip detectors and straw drift tubes [34, 35] 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 $200\,\mathrm{GeV}/c$. The minimum distance of a track to a primary pp collision vertex (PV), the impact parameter (IP), is measured with a resolution of $(15 + 29/p_\mathrm{T})\,\mu\mathrm{m}$, 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 [36]. 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 [37, 38].

The online event selection is performed by a trigger system [39]. Signal candidates first need to pass the hardware trigger (L0), which requires the final-state muon to have sizeable p_T. In the subsequent software trigger, a full event reconstruction is performed. The software trigger requires a two-, three- or four-track secondary vertex with a significant displacement from any primary pp interaction vertex. At least one charged particle must have a transverse momentum $p_\mathrm{T} > 1.6\,\mathrm{GeV}/c$ ($p_\mathrm{T} > 1.0\,\mathrm{GeV}/c$ if the particle is identified as muon) and be inconsistent with originating from a PV. A multivariate algorithm [40, 41] is used for the identification of secondary vertices consistent with the decay of a b hadron.

Simulated samples are required to determine the reconstruction and selection efficiencies, and to estimate contributions from residual backgrounds. In the simulation, pp collisions are generated using PYTHIA [42] with a specific LHCb configuration [43]. Decays of unstable particles are described by EVTGEN [44], in which final-state radiation is generated using PHOTOS [45]. The interaction of the generated particles with the detector, and its response, are implemented using the GEANT4 toolkit [46] as described in Ref. [47]. The underlying pp interaction is reused multiple times, with an independently generated signal decay for each [48]. Residual mismodelling of the particle identification and tracking performance, the p_T spectra of B^0 and B^0_s mesons, track multiplicity, and the efficiency of the L0 trigger are calibrated using high-yield control samples from data.

3 Selection of signal candidates

Candidates for $B^0 \to K^{*0}\mu^\pm e^\mp$ and $B^0_s \to \phi\mu^\pm e^\mp$ signal decays are reconstructed in the $K^+\pi^-\mu^\pm e^\mp$ and $K^+K^-\mu^\pm e^\mp$ final states, respectively. Stringent particle identification criteria are applied to the final-state hadrons and leptons, using information from the Cherenkov detectors, the muon chambers, and the calorimeter system. The final-state tracks are required to have significant χ^2_IP with respect to any PV in the event, where χ^2_IP is defined as the difference in the vertex-fit χ^2 of a given PV reconstructed with and without the track being considered. The four final-state tracks are fit to a secondary vertex (SV), which needs to have good fit quality and be significantly displaced from any PV in the event. The invariant mass of the $K^+\pi^-$ (K^+K^-) system is required to be within $100\,\mathrm{MeV}/c^2$ ($12\,\mathrm{MeV}/c^2$) of the known K^{*0} (ϕ) mass [49]. Furthermore, the reconstructed $B^0_{(s)}$ mass of signal candidates is required to be in the range $[4300, 6700]\,\mathrm{MeV}/c^2$.

Dedicated vetoes are applied to reject backgrounds originating from misidentified b-hadron decays, referred to as peaking backgrounds. The decays $B^0 \to J/\psi(\to \ell^+\ell^-)K^{*0}$ and $B^0 \to \psi(2S)(\to \ell^+\ell^-)K^{*0}$ ($B^0_s \to J/\psi(\to \ell^+\ell^-)\phi$ and $B^0_s \to \psi(2S)(\to \ell^+\ell^-)\phi$) can be

a source of background if one of the leptons is misidentified as the different lepton species. Candidates for which the reconstructed B^0 (B_s^0) mass is in the range of $[5200, 5350]$ MeV/c^2 ($[5300, 5450]$ MeV/c^2) are rejected, where the reconstructed B^0 (B_s^0) mass is determined in a fit forcing the dilepton system to have the known mass of the J/ψ or $\psi(2S)$ meson [49]. The same tree-level decays to charmonia can also mimic the signal decay if a lepton is misidentified as a hadron and vice-versa. To suppress these contributions, criteria on the invariant mass of the system comprised of a hadron and the lepton of opposite charge are applied, where the lepton is assigned the hadron mass hypothesis. Candidates are rejected if the resulting invariant mass is in the range $[800, 1050]$ MeV/c^2 ($[1000, 1040]$ MeV/c^2) and therefore consistent with the decay of a K^{*0} (ϕ) meson. Semileptonic $b \to c(\to s\ell'^+\nu_{\ell'})\ell^-\bar{\nu}_\ell$ cascade decays can result in the same charged final-state tracks in the detector as signal decays. These decays are rejected by a stringent requirement on $m(K^+\pi^-\ell^\pm) > 2$ GeV/c^2 ($m(K^+K^-\ell^\pm) > 2$ GeV/c^2), where the invariant mass is calculated using the lepton momentum without the addition of potential bremsstrahlung photons. This removes contributions from $B^0 \to D^{(*)-}\ell^+\nu_\ell$ ($B_s^0 \to D_s^{(*)-}\ell^+\nu_\ell$) decays. Semileptonic $B_{(s)}^0$ decays to higher excited $D_{(s)}^-$ resonances can escape this veto due to their higher masses; they are therefore modelled in the fit as discussed in Sec. 5. Background from several further sources are studied and found to be suppressed to negligible levels by the stringent particle identification criteria; these include rare $b \to s\ell^+\ell^-$ decays like $B^0 \to K^{*0}\ell^+\ell^-$, and $B_s^0 \to \phi\ell^+\ell^-$ (with lepton and potentially also hadron misidentification), as well as fully hadronic $B^0 \to K^{*0}\pi^+\pi^-$ and $B_s^0 \to \phi\pi^+\pi^-$ decays (with misidentification of the pions as leptons).

Background from combinations of random tracks (combinatorial background) is reduced using a boosted decision tree (BDT) [50] classifier trained with the AdaBoost algorithm [51] as implemented in the TMVA software package [52]. The BDT classifier is trained separately for the two signal decays using calibrated simulation as signal. Data from the $B_{(s)}^0$ upper mass sideband region $[5600, 6700]$ MeV/c^2 are used as a proxy for the background. The classifier is trained using a k-folding approach and its performance is verified using standard cross-validation techniques [53]. The classifier uses the (transverse) momentum of the $B_{(s)}^0$ candidate, its vertex fit quality and flight distance significance, the angle between the $B_{(s)}^0$ momentum and the vector connecting the associated PV and the $B_{(s)}^0$ decay vertex, and the χ^2_{IP} of the $B_{(s)}^0$ candidate and the final-state particles. The selection criterion on the classifier output is chosen according to the Punzi figure of merit $\varepsilon_{\text{sig}}/(3/2 + \sqrt{N_{\text{comb}}})$ [54]. Here, ε_{sig} denotes the signal efficiency and N_{comb} the expected combinatorial background yield, which is extrapolated from the upper mass sideband using the reconstructed same-sign lepton samples $K^+\pi^-\mu^\pm e^\pm$ and $K^+K^-\mu^\pm e^\pm$. Relative to the previously described selection criteria, the BDT requirement results in a signal efficiency of 55–80%, depending on the signal mode, and a rejection for combinatorial background of larger than 99%.

The normalisation modes, $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$ and $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi$, are reconstructed and selected in a way that is as similar as possible to the corresponding signal decays. In contrast to the signal modes, the muon identification criteria are applied to both final-state leptons. In addition, the invariant mass of the dimuon system is required to be within ± 60 MeV/c^2 of the known J/ψ mass [49], and the dedicated vetoes against b-hadron decays to charmonia are removed. For the $B^0 \to J/\psi K^{*0}$ normalisation mode, an additional veto on the invariant mass of the $K^+\mu^+\mu^-$ system rejects candidates

with $m(K^+\mu^+\mu^-) \in [5200, 5400]\,\mathrm{MeV}/c^2$. The decays $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi\phi$, with both $J/\psi \to e^+e^-$ and $J/\psi \to \mu^+\mu^-$, are used as control modes to validate simulation and to model the signal mass resolution as discussed in Sec. 5. The control decays $B^0 \to J/\psi(\to e^+e^-)K^{*0}$ and $B_s^0 \to J/\psi(e^+e^-)\phi$ are selected in the $m(e^+e^-)$ mass region $[2400, 3300]\,\mathrm{MeV}/c^2$.

4 Normalisation

The signal yields N_{sig} are obtained from a fit of the reconstructed $B_{(s)}^0$ mass distribution and translated to signal branching fractions $\mathcal{B}_{\mathrm{sig}}$ using the normalisation modes $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$ and $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi$ according to

$$\mathcal{B}_{\mathrm{sig}} = \underbrace{\frac{\mathcal{B}_{\mathrm{norm}}}{N_{\mathrm{norm}}} \times \frac{\varepsilon_{\mathrm{norm}}}{\varepsilon_{\mathrm{sig}}}}_{= \, \alpha} \times N_{\mathrm{sig}}, \tag{1}$$

where $\mathcal{B}_{\mathrm{norm}}$ denotes the branching fractions of the normalisation mode, given by $\mathcal{B}(B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}) = (7.57 \pm 0.30) \times 10^{-5}$ or $\mathcal{B}(B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi) = (6.07 \pm 0.29) \times 10^{-5}$ [49,55], N_{norm} the yields of the normalisation mode, and α the normalisation constant.

The efficiency ratio between normalisation and signal decays, $\varepsilon_{\mathrm{norm}}/\varepsilon_{\mathrm{sig}}$, is determined using calibrated simulation and found to be around 2.8 (2.6) for the decays $B^0 \to J/\psi K^{*0}$ and $B^0 \to K^{*0}\mu^\pm e^\mp$ ($B_s^0 \to J/\psi\phi$ and $B_s^0 \to \phi\mu^\pm e^\mp$). The efficiencies for the normalisation modes are significantly higher than for the signal decays as they exhibit higher trigger efficiencies, more efficient particle identification, and background vetoes with higher efficiencies.

The yields for the normalisation modes are determined using an unbinned extended maximum-likelihood fit to the reconstructed $B_{(s)}^0$ mass distribution, in which the invariant mass of the dimuon system is constrained to the known J/ψ mass. The fit range is limited to $[5150, 5900]\,\mathrm{MeV}/c^2$ in the B^0 mode ($[5250, 5900]\,\mathrm{MeV}/c^2$ in the B_s^0 mode) to avoid partially reconstructed b-hadron decays at low invariant masses. The signal distribution is modelled using a double-sided Hypatia function [56]. The shape parameters of the Hypatia function are determined using simulation, except for a resolution and mass shift parameter which are left to float in the fit to data to allow for mismodelling. In the fit of the decay $B^0 \to J/\psi K^{*0}$, a component for the CKM-suppressed mode $B_s^0 \to J/\psi \overline{K}^{*0}$ is also included, which is modelled identically to the normalisation mode, but shifted by the known $B_s^0 - B^0$ mass difference [49]. The $B_s^0 \to J/\psi \overline{K}^{*0}$ component is Gaussian constrained to its expectation using simulation. Similarly, in the $B_s^0 \to J/\psi\phi$ normalisation mode fit, a small component of the decay $B^0 \to J/\psi K^+ K^-$ must be accounted for, where the invariant mass of the non-resonant K^+K^- system overlaps with the ϕ selection. Again, the normalisation channel model is used to describe this component, including a mass shift by the known $B^0 - B_s^0$ mass difference. The background yield is floated in the fit. Residual backgrounds from misidentified b-hadron decays to J/ψ final states are modelled using kernel density estimates obtained from calibrated simulation. For the decay $B^0 \to J/\psi K^{*0}$, misidentified backgrounds from $\Lambda_b^0 \to J/\psi p K^-$, $B_s^0 \to J/\psi\phi$, and $B^0 \to J/\psi K^{*0}$ (with $K \leftrightarrow \pi$ misidentification), are included in the fit. For the decay $B_s^0 \to J/\psi\phi$, misidentified backgrounds from $\Lambda_b^0 \to J/\psi p K^-$ and $B^0 \to J/\psi K^{*0}$ decays are included. The yields of all backgrounds from misidentification are constrained to their expectations using
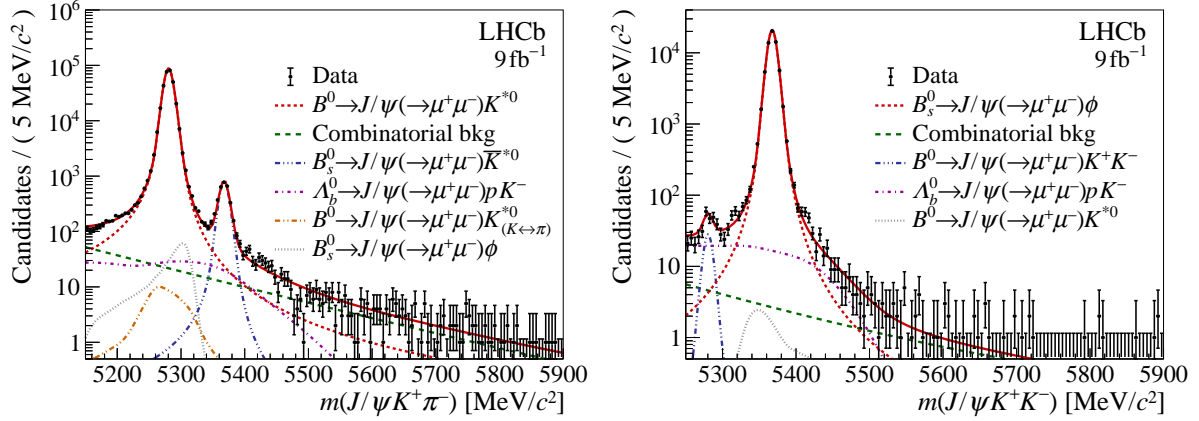
Figure 1: Mass distributions for the normalisation channels (left) $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$ and (right) $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi$ combining the different data taking periods, overlaid with the fit results.

Table 1: Normalisation mode yields $[10^3]$ for different periods of data taking.

	Yield $[10^3]$		
Mode	2011–2012	2015–2016	2017–2018
$B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$	88.88 ± 0.30	85.56 ± 0.29	139.05 ± 0.37
$B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi$	17.89 ± 0.13	16.21 ± 0.13	30.59 ± 0.18

Gaussian functions. The last component of the fit for the normalisation yields is the combinatorial background, which is modelled using an exponential function, with its slope and normalisation allowed to vary freely. The $m(J/\psi K^+\pi^-)$ and $m(J/\psi K^+K^-)$ distributions, overlaid with the fit results combined over the different data taking periods, are shown in Fig. 1. The obtained normalisation channel yields are given in Tab. 1. The resulting normalisation constants α for the different data taking periods are given in Tab. 2.

Table 2: Normalisation constant α $[10^{-9}]$ with associated statistical and systematic uncertainties, added in quadrature, for different periods of data taking. The total uncertainty is dominated by systematic effects, which are discussed in Sec. 6. The year-to-year B^0/B_s^0 ratio variation is due to different BDT criteria against combinatorial background, tuned individually for each data taking period and mode.

	$\alpha \pm (\sigma_{\text{stat}} \oplus \sigma_{\text{syst}})$ $[10^{-9}]$		
Mode	2011–2012	2015–2016	2017–2018
$B^0 \to K^{*0}\mu^+e^-$	2.47 ± 0.14	2.38 ± 0.16	1.49 ± 0.09
$B^0 \to K^{*0}\mu^-e^+$	2.50 ± 0.15	2.39 ± 0.16	1.49 ± 0.09
$B^0 \to K^{*0}\mu^\pm e^\mp$	2.48 ± 0.14	2.39 ± 0.16	1.49 ± 0.09
$B_s^0 \to \phi\mu^\pm e^\mp$	9.50 ± 0.70	9.68 ± 0.78	5.09 ± 0.39

5   Signal fit

The signal decays $B^0 \to K^{*0}\mu^{\pm}e^{\mp}$ and $B_s^0 \to \phi\mu^{\pm}e^{\mp}$ are modelled using the sum of two Crystal Ball functions [57], with power-law tails on either sides. The shape parameters are determined on simulation and fixed in the fit to data. Corrections to the mass resolution are determined using the control decays $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \phi$. Information from both $J/\psi \to \mu^+\mu^-$ and $J/\psi \to e^+e^-$ final states is combined to determine these correction factors, as no appropriate control mode with a $\mu^{\pm}e^{\mp}$ final state exists. Depending on the data taking period, the correction factors scale the core Gaussian widths by 1.04–1.08 with uncertainties at the percent level. In the nominal fit, the uncertainties on the scale factors are included as Gaussian constraints.

Semileptonic cascade decays involving higher excited $D_{(s)}^-$ resonances are modelled in the fit as they can pass the selection requirement $m(K^+\pi^-\ell^{\pm}) > 2\,\mathrm{GeV}/c^2$ ($m(K^+K^-\ell^{\pm}) > 2\,\mathrm{GeV}/c^2$) due to their high masses. Their shapes are modelled using kernel density estimates of fully simulated $B^0 \to D_2^*(2460)^-(\to \overline{D}^0(\to K^+\ell^-\bar{\nu}_\ell)\pi^-)\ell'^+\nu_{\ell'}$ and $B_s^0 \to D_{s2}^*(2573)^-(\to \overline{D}^0(\to K^+\ell^-\bar{\nu}_\ell)K^-)\ell'^+\nu_{\ell'}$ decays, as these decays are one of the dominant contributions of the remaining background from semileptonic cascades. Alternative models that include potential contributions from $D_1(2420)$ and $D_0^*(2300)$ states have also been considered. For the decay $B^0 \to K^{*0}\mu^{\pm}e^{\mp}$, an additional background contribution arises from $B^+ \to \overline{D}^0(\to K^+\ell^-\bar{\nu}_\ell)\ell'^+\nu_{\ell'}$ decays, which are combined with a random π^- from the event and thus can pass the veto against semileptonic cascade decays. This background is modelled using a kernel density estimate from simulated samples. Its yield is Gaussian constrained to the expectation from simulation. The combinatorial background is modelled with a single exponential function. Due to the low residual combinatorial background yield in the signal data samples, the exponential slope is constrained from a fit to same-sign lepton data samples with the reconstructed final states $K^+\pi^-\mu^{\pm}e^{\pm}$ and $K^+K^-\mu^{\pm}e^{\pm}$, using a relaxed BDT cut. A systematic uncertainty is assigned to this choice, as discussed in Sec. 6. The combinatorial background yields are allowed to float independently in each data taking periods.

The signal branching fractions are determined in a simultaneous fit of the data taking periods 2011–2012, 2015–2016, and 2017–2018 using Eq. 1. In the fit, the signal branching fraction and the branching fraction of the semileptonic cascade decays involving higher excited $D_{(s)}^-$ resonances, are shared between data taking periods. Figure 2 shows the reconstructed $B_{(s)}^0$ mass distributions for $B^0 \to K^{*0}\mu^{\pm}e^{\mp}$ and $B_s^0 \to \phi\mu^{\pm}e^{\mp}$ candidates, combined over data taking periods, and overlaid with the fit results for the full and the background-only model. For illustration, the signal shape, scaled to a branching fraction of 5×10^{-8} for the $B^0 \to K^{*0}\mu^{\pm}e^{\mp}$ decays and 1×10^{-7} for $B_s^0 \to \phi\mu^{\pm}e^{\mp}$, is drawn in addition. The fits for the same-sign ($B^0 \to K^{*0}(\to K^+\pi^-)\mu^+e^-$) and opposite-sign ($B^0 \to K^{*0}(\to K^+\pi^-)\mu^-e^+$) samples are also given. No significant signals are observed and limits on the signal decays are set, as detailed in Sec. 7.

6   Systematic uncertainties

Systematic effects can modify the limit setting through the normalisation constant α or through modifications of the fit model. The systematic effects on the fit model comprise the correction of the signal mass resolution and the exponential slope of the combinatorial
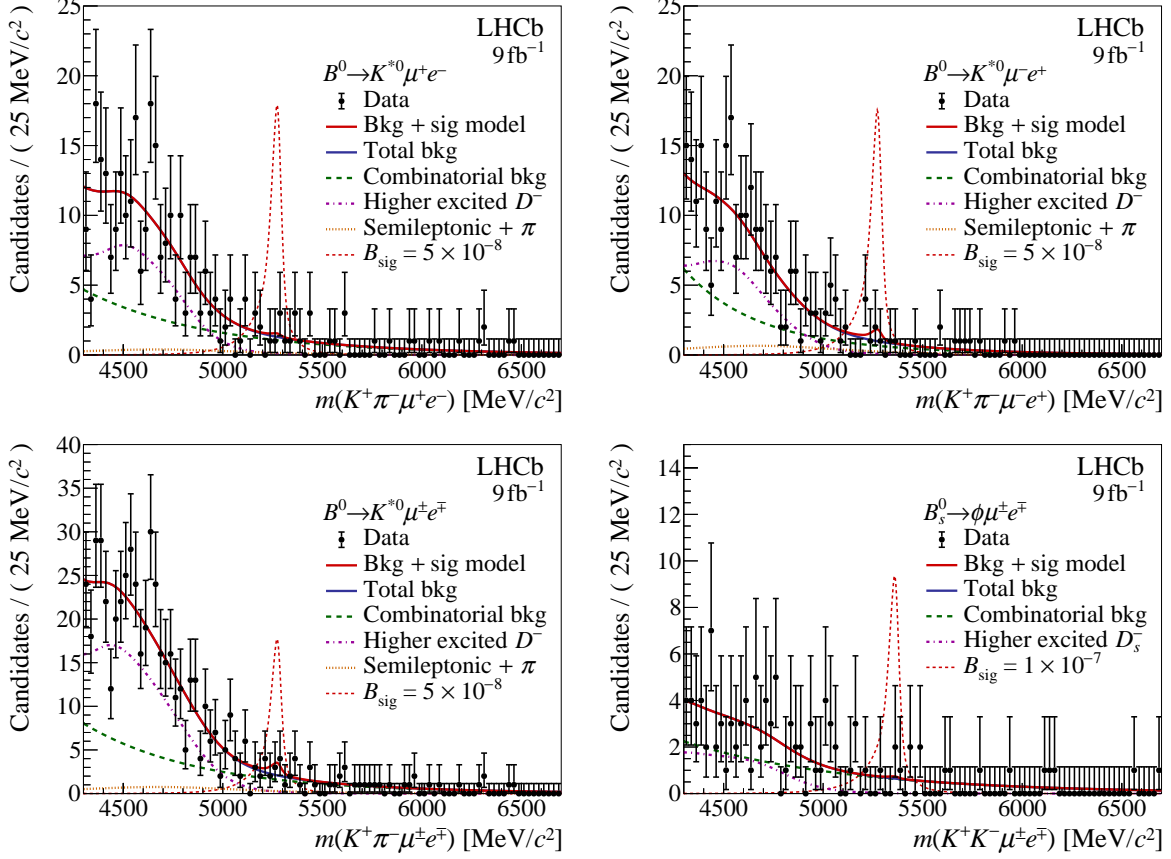
Figure 2: Mass distributions for (top left) $B^0 \to K^{*0}\mu^+e^-$, (top right) $B^0 \to K^{*0}\mu^-e^+$, (bottom left) $B^0 \to K^{*0}\mu^\pm e^\mp$, and (bottom right) $B_s^0 \to \phi\mu^\pm e^\mp$ candidates. The data are overlaid with the fit results. For illustration, the signal shape, scaled to a branching fraction of 5×10^{-8} for the $B^0 \to K^{*0}\mu^\pm e^\mp$ decays and 1×10^{-7} for $B_s^0 \to \phi\mu^\pm e^\mp$, is drawn as red dashed line.

background. They are included in the fit through Gaussian constraints. For the signal mass resolutions, the correction factors for the core Gaussian resolution are allowed to vary within their uncertainties. For the exponential slope, the difference between the fit to same-sign lepton data with a reduced BDT cut and with the nominal BDT requirement is used as an estimate for the uncertainty on this parameter. This is added in quadrature to the uncertainty of the parameter from the fit with the reduced BDT requirement, and included as Gaussian constraint in the fit.

A summary of the systematic uncertainties affecting the normalisation constant α is given in Tab. 3. The dominant source of systematic uncertainty originates from the uncertainty on the branching fraction of the normalisation channel. For the signal decay $B_s^0 \to \phi\mu^\pm e^\mp$, a systematic uncertainty of similar size arises from the significant lifetime difference of the B_s^0 mass eigenstates. The effective lifetime of the final state is a priori unknown and can affect the signal efficiency which depends on the B_s^0 decay time. The difference between the maximum and minimum lifetimes, given by those of the light and the heavy mass eigenstate, is used to determine a conservative systematic uncertainty. Further systematic uncertainties arise from the limited size of the simulation samples and the limited precision of the calibration procedures applied to simulation. These include the weighting of the B production kinematics and event multiplicity, calibration of the particle

Table 3: Sources of relative systematic uncertainties [%] on the normalisation constant α defined in Eq. 1. Where the uncertainty depends on the year of data taking, a range is provided.

Systematic source	$B^0 \to K^{*0}\mu^{\pm}e^{\mp}$	$B_s^0 \to \phi\mu^{\pm}e^{\mp}$
Normalisation $\mathcal{B}$	4.0	4.8
$\phi\mu^{\pm}e^{\mp}$ decay time distribution	–	3.8–4.5
Limited simulation sample size	0.7–1.5	0.6–1.4
Production kinematics/multiplicity	1.2–2.7	1.6–3.8
Particle identification	0.3–0.8	0.3–0.6
Muon identification	1	1
Tracking efficiency	1	1
L0 trigger calibration	1	1
HLT trigger efficiency	1	1
Residual MC differences	1.7–4.3	1.1–3.7
Sum	5.2–6.7	6.9–8.5

Table 4: Expected (background-only hypothesis) and observed limits $[10^{-9}]$ at 90% (95%) CL.

Mode	Expected	Observed
$B^0 \to K^{*0}\mu^+e^-$	4.8 (5.9)	5.7 (6.9)
$B^0 \to K^{*0}\mu^-e^+$	4.6 (5.7)	6.8 (7.9)
$B^0 \to K^{*0}\mu^{\pm}e^{\mp}$	6.1 (7.5)	10.1 (11.7)
$B_s^0 \to \phi\mu^{\pm}e^{\mp}$	14.2 (17.7)	16.0 (19.8)

identification response and the tracking efficiency, as well as the calibration of the trigger efficiencies using data [58]. In addition, systematic uncertainties originating from residual differences between data and simulation are conservatively estimated from a comparison of the BDT output distribution for background-subtracted [59] $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$ and $B_s^0 \to J/\psi(\to \mu^+\mu^-)\phi$ decays with simulation. The relative difference in the normalisation channel BDT selection efficiency in data and simulation is assigned as a systematic uncertainty.

7 Results

No significant excess of $B^0 \to K^{*0}\mu^{\pm}e^{\mp}$ or $B_s^0 \to \phi\mu^{\pm}e^{\mp}$ decays is observed and limits are set using the $\mathrm{CL_s}$ method [60]. A one-sided test statistic is used [61], as implemented in the GAMMACOMBO framework [62, 63]. The test statistic is evaluated using pseudo-experiments, which are generated using the best fit values for the nuisance parameters, and where the central values of the Gaussian constraints are varied according to their uncertainties.

The resulting $\mathrm{CL_s}$ scans are shown in Fig. 3, and upper limits at 90% and 95% CL are reported in Tab. 4; the limits given for $\mathcal{B}(B^0 \to K^{*0}\mu^{\pm}e^{\mp})$ are determined using the combined $B^0 \to K^{*0}\mu^+e^-$ and $B^0 \to K^{*0}\mu^-e^+$ data samples.

The default limits assume a uniform phase space model for the signal decays. However, NP models can result in very different decay kinematics and differential decay rates [64–66]. This is illustrated in Fig. 4 of App. A for two distinct NP scenarios: a scalar model with
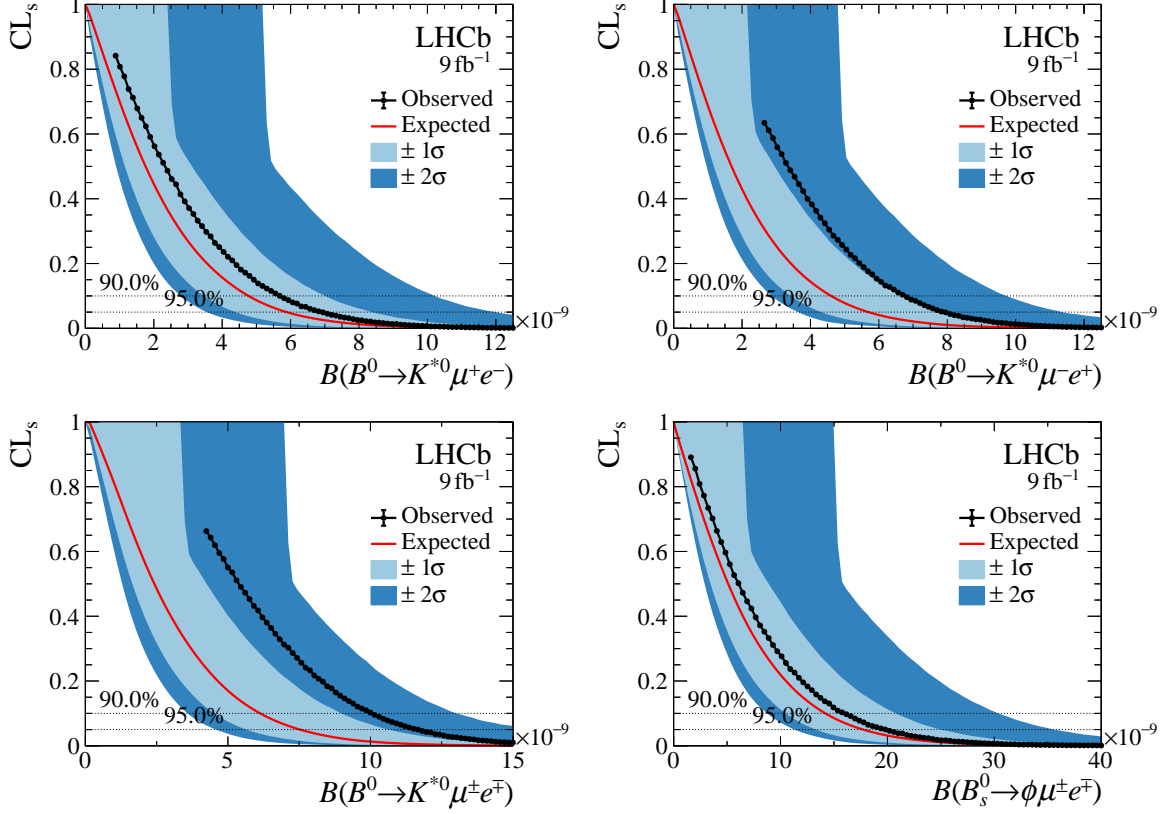
Figure 3: Observed and expected (background-only hypothesis) limits for (top left) $B^0 \to K^{*0}\mu^+e^-$, (top right) $B^0 \to K^{*0}\mu^-e^+$, (bottom left) $B^0 \to K^{*0}\mu^\pm e^\mp$, and (bottom right) $B_s^0 \to \phi\mu^\pm e^\mp$.

Table 5: Exclusion limits $[10^{-9}]$ on the $B^0_{(s)}$ branching fractions for a scalar ($C_s^{\mu e} \neq 0$) and left-handed ($C_9^{\mu e} = -C_{10}^{\mu e} \neq 0$) NP model at 90% (95%) CL.

Mode	Left-handed	Scalar
$B^0 \to K^{*0}\mu^+e^-$	6.7 (8.3)	8.4 (10.2)
$B^0 \to K^{*0}\mu^-e^+$	8.0 (9.5)	9.9 (11.5)
$B^0 \to K^{*0}\mu^\pm e^\mp$	12.0 (13.9)	14.7 (17.0)
$B_s^0 \to \phi\mu^\pm e^\mp$	16.5 (20.5)	18.8 (23.1)

$C_S^{\mu e} \neq 0$, and a left-handed model with $C_9^{\mu e} = -C_{10}^{\mu e} \neq 0$ [64,67]. Here, $C_i^{\mu e}$ denotes the lepton-flavour violating Wilson coefficients. As the reconstruction and selection efficiency shown in Fig. 5 in App. B is not flat in the decay kinematics, the total signal efficiency can differ for NP models, which needs to be accounted for in the computation of limits. For the scalar- and left-handed NP scenarios shown in App. A, the resulting limits are given in Tab. 5.

9

8   Conclusions

A search for the lepton-flavour violating decays $B^0 \to K^{*0}\mu^+e^-$, $B^0 \to K^{*0}\mu^-e^+$, $B^0 \to K^{*0}\mu^\pm e^\mp$, and $B_s^0 \to \phi\mu^\pm e^\mp$ is presented. The search uses pp collision data collected at centre-of-mass energies of 7 TeV (2011), 8 TeV (2012), and 13 TeV (2015–2018), corresponding to a total integrated luminosity of 9 fb^{-1}. No significant excesses are observed and limits on the branching fractions are obtained for a uniform phase space signal decay model, resulting in

$$
\begin{aligned}
\mathcal{B}(B^0 \to K^{*0}\mu^+e^-) &< 5.7 \times 10^{-9}\ (6.9 \times 10^{-9}), \\
\mathcal{B}(B^0 \to K^{*0}\mu^-e^+) &< 6.8 \times 10^{-9}\ (7.9 \times 10^{-9}), \\
\mathcal{B}(B^0 \to K^{*0}\mu^\pm e^\mp) &< 10.1 \times 10^{-9}\ (11.7 \times 10^{-9}), \\
\mathcal{B}(B_s^0 \to \phi\mu^\pm e^\mp) &< 16.0 \times 10^{-9}\ (19.8 \times 10^{-9})
\end{aligned}
$$

at 90% (95%) CL. In addition, limits on a scalar and left-handed NP scenario are reported. The results constitute the most stringent limits on a semileptonic LFV b-hadron decays to date. The limits on the decay $B^0 \to K^{*0}\mu^\pm e^\mp$ are improved by more than one order of magnitude in comparison to previous searches. The reported limits on the decay $B_s^0 \to \phi\mu^\pm e^\mp$ constitute the world's first constraint of a semileptonic LFV B_s^0 decay.

References

[1] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B_s^0 \to \phi\mu^+\mu^-$*, JHEP **07** (2013) 084, `arXiv:1305.2168`.

[2] 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`.

[3] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays*, JHEP **06** (2015) 115, Erratum ibid. **09** (2018) 145, `arXiv:1503.07138`.

[4] LHCb collaboration, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$*, JHEP **09** (2015) 179, `arXiv:1506.08777`.

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

[6] LHCb collaboration, R. Aaij *et al.*, *Branching fraction measurements of the rare $B_s^0 \to \phi\mu^+\mu^-$ and $B_s^0 \to f_2'(1525)\mu^+\mu^-$ decays*, Phys. Rev. Lett. **127** (2021) 151801, `arXiv:2105.14007`.

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

[8] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP-averaged observables in the $B^0 \to K^{*0}\mu^+\mu^-$ decay*, Phys. Rev. Lett. **125** (2020) 011802, `arXiv:2003.04831`.

[9] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^+ \to K^{*+}\mu^+\mu^-$ decay*, Phys. Rev. Lett. **126** (2021) 161802, `arXiv:2012.13241`.

[10] ATLAS collaboration, M. Aaboud *et al.*, *Angular analysis of $B_d^0 \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **10** (2018) 047, `arXiv:1805.04000`.

[11] CMS collaboration, V. Khachatryan *et al.*, *Angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ from pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B753** (2016) 424, `arXiv:1507.08126`.

[12] CMS collaboration, A. M. Sirunyan *et al.*, *Measurement of angular parameters from the decay $B^0 \to K^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B781** (2018) 517, `arXiv:1710.02846`.

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

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

[15] LHCb collaboration, R. Aaij *et al.*, *Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **122** (2019) 191801, `arXiv:1903.09252`.

[16] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $\Lambda_b^0 \to pK^-\ell^+\ell^-$ decays*, JHEP **05** (2020) 040, `arXiv:1912.08139`.

[17] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in beauty-quark decays*, `arXiv:2103.11769`, to appear in Nature Physics.

[18] LHCb collaboration, R. Aaij *et al.*, *Tests of lepton universality using $B^0 \to K_S^0\ell^+\ell^-$ and $B^+ \to K^{*+}\ell^+\ell^-$ decays*, Phys. Rev. Lett. **128** (2022) 191802, `arXiv:2110.09501`.

[19] Belle collaboration, S. Wehle *et al.*, *Lepton-flavor-dependent angular analysis of $B \to K^*\ell^+\ell^-$*, Phys. Rev. Lett. **118** (2017) 111801, `arXiv:1612.05014`.

[20] BaBar collaboration, J. P. Lees *et al.*, *Measurement of branching fractions and rate asymmetries in the rare decays $B \to K^{(*)}l^+l^-$*, Phys. Rev. **D86** (2012) 032012, `arXiv:1204.3933`.

[21] Belle collaboration, S. Choudhury *et al.*, *Test of lepton flavor universality and search for lepton flavor violation in $B \to K\ell\ell$ decays*, JHEP **03** (2021) 105, `arXiv:1908.01848`.

[22] Belle collaboration, A. Abdesselam *et al.*, *Test of lepton-flavor universality in $B \to K^*\ell^+\ell^-$ decays at Belle*, Phys. Rev. Lett. **126** (2021) 161801, `arXiv:1904.02440`.

[23] S. L. Glashow, D. Guadagnoli, and K. Lane, *Lepton flavor violation in B decays?*, Phys. Rev. Lett. **114** (2015) 091801, `arXiv:1411.0565`.

[24] G. Hiller, D. Loose, and K. Schönwald, *Leptoquark flavor patterns & B decay anomalies*, JHEP **12** (2016) 027, `arXiv:1609.08895`.

[25] I. de Medeiros Varzielas and G. Hiller, *Clues for flavor from rare lepton and quark decays*, JHEP **06** (2015) 072, `arXiv:1503.01084`.

[26] A. Crivellin, D. Müller, A. Signer, and Y. Ulrich, *Correlating lepton flavor universality violation in B decays with $\mu \to e\gamma$ using leptoquarks*, Phys. Rev. **D97** (2018) 015019, `arXiv:1706.08511`.

[27] A. Crivellin *et al.*, *Lepton-flavour violating B decays in generic Z' models*, Phys. Rev. **D92** (2015) 054013, `arXiv:1504.07928`.

[28] Belle collaboration, S. Sandilya *et al.*, *Search for the lepton-flavor-violating decay $B^0 \to K^{*0}\mu^\pm e^\mp$*, Phys. Rev. **D98** (2018) 071101, `arXiv:1807.03267`.

[29] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour violating decays $B^0_{(s)} \to e^\pm \mu^\mp$*, JHEP **03** (2018) 078, `arXiv:1710.04111`.

[30] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour violating decays $B^+ \to K^+ \mu^\pm e^\mp$*, Phys. Rev. Lett. **123** (2019) 231802, `arXiv:1909.01010`.

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

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

[33] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[34] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, `arXiv:1311.3893`.

[35] P. d'Argent *et al.*, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, JINST **12** (2017) P11016, `arXiv:1708.00819`.

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

[37] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, `arXiv:1211.1346`.

[38] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, `arXiv:1306.0249`.

[39] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, `arXiv:1211.3055`.

[40] 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`.

[41] T. Likhomanenko *et al.*, *LHCb topological trigger reoptimization*, J. Phys. Conf. Ser. **664** (2015) 082025.

[42] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820; T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[43] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

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

[45] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS interface in C++: Technical and physics documentation*, Comp. Phys. Comm. **199** (2016) 86, arXiv:1011.0937.

[46] 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.

[47] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[48] D. Müller, M. Clemencic, G. Corti, and M. Gersabeck, *ReDecay: A novel approach to speed up the simulation at LHCb*, Eur. Phys. J. **C78** (2018) 1009, arXiv:1810.10362.

[49] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020) 083C01.

[50] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[51] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, J. Comput. Syst. Sci. **55** (1997) 119.

[52] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, *TMVA - Toolkit for Multivariate Data Analysis with ROOT*, PoS **ACAT** (2007) 040; A. Hoecker *et al.*, *TMVA 4 — Toolkit for Multivariate Data Analysis with ROOT. Users Guide.*, arXiv:physics/0703039.

[53] 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), 203–208, ACM, 1999.

[54] G. Punzi, *Sensitivity of searches for new signals and its optimization*, eConf **C030908** (2003) MODT002, arXiv:physics/0308063.

[55] LHCb collaboration, R. Aaij *et al.*, *Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B_s^0 decay branching fractions*, Phys. Rev. **D104** (2021) 032005, arXiv:2103.06810.

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

[57] 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.

[58] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, LHCb-PUB-2014-039, 2014.

[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, Erratum ibid. **C73** (2013) 2501, arXiv:1007.1727.

[62] M. Kenzie *et al.*, *GammaCombo: A statistical analysis framework for combining measurements, fitting datasets and producing confidence intervals*, doi: 10.5281/zenodo.3371421.

[63] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle γ from a combination of LHCb results*, JHEP **12** (2016) 087, arXiv:1611.03076.

[64] D. Bečirević, O. Sumensari, and R. Zukanovich Funchal, *Lepton flavor violation in exclusive $b \to s$ decays*, Eur. Phys. J. **C76** (2016) 134, arXiv:1602.00881.

[65] W. Altmannshofer *et al.*, *Symmetries and asymmetries of $B \to K^* \mu^+ \mu^-$ decays in the Standard Model and beyond*, JHEP **01** (2009) 019, arXiv:0811.1214.

[66] C. Bobeth, G. Hiller, and G. Piranishvili, *CP asymmetries in $\bar{B} \to \bar{K}^*(\to \bar{K}\pi)\bar{\ell}\ell$ and untagged $\bar{B}_s$, $B_s \to \phi(\to K^+K^-)\bar{\ell}\ell$ decays at NLO*, JHEP **07** (2008) 106, arXiv:0805.2525.

[67] D. M. Straub, *flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond*, arXiv:1810.08132.

[68] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0} \mu^+ \mu^-$*, JHEP **08** (2013) 131, arXiv:1304.6325.

Appendices

A   New Physics models

New Physics scenarios can result in very different distributions in the three decay angles $\cos\theta_\ell$, $\cos\theta_K$ and ϕ, defined as in Ref. [68], and the four-momentum transfer $q^2 = m^2(\mu^\pm e^\mp)$ [64–66]. This is illustrated in Fig. 4.
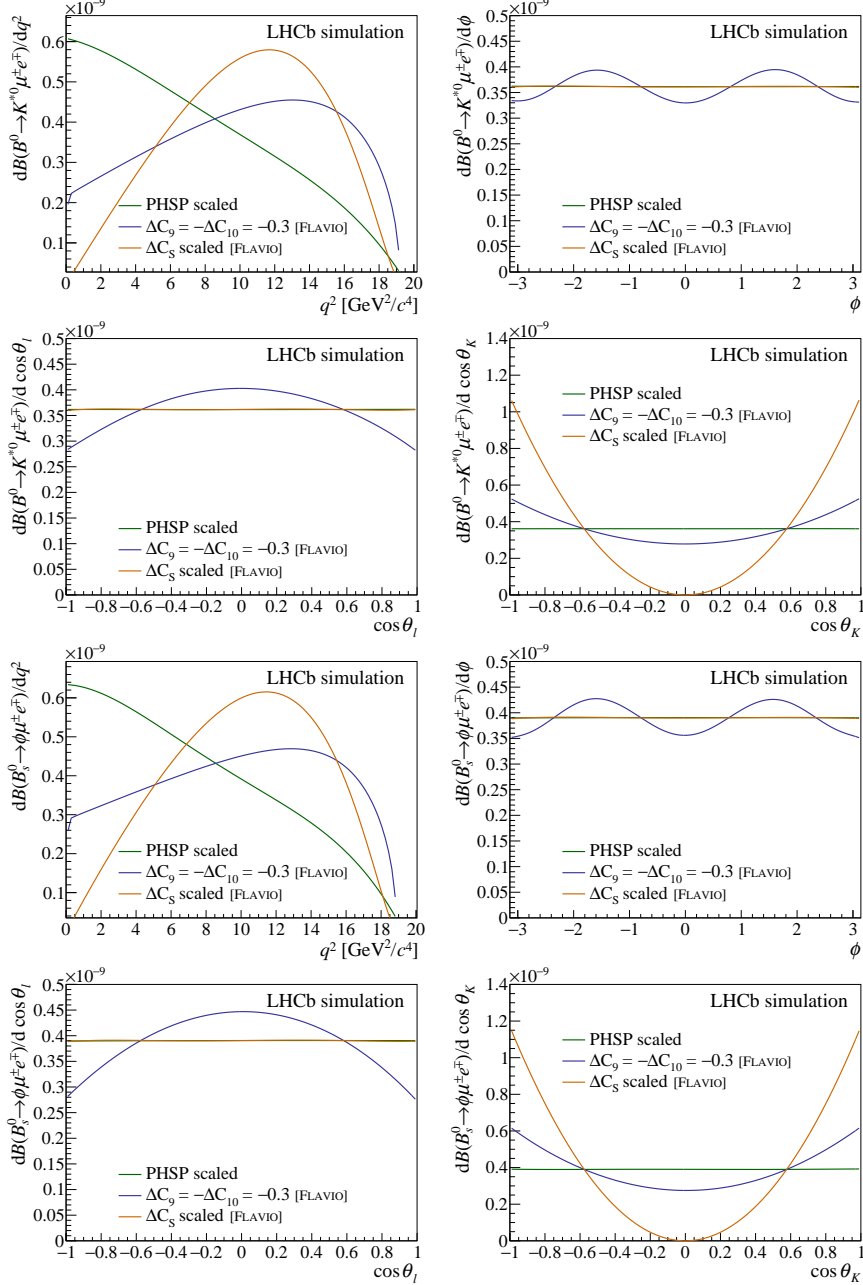

Figure 4: Differential decay rate as a function of the four-momentum transfer q^2 and the three decay angles in a left-handed ($C_9^{\mu e} = -C_{10}^{\mu e} \neq 0$) NP model for (top) the signal decay $B^0 \to K^{*0}\mu^\pm e^\mp$ and (bottom) $B_s^0 \to \phi\mu^\pm e^\mp$. The left-handed NP scenario is compared with the nominal phase space and a scalar ($C_S^{\mu e} \neq 0$) model, normalised to the same area.
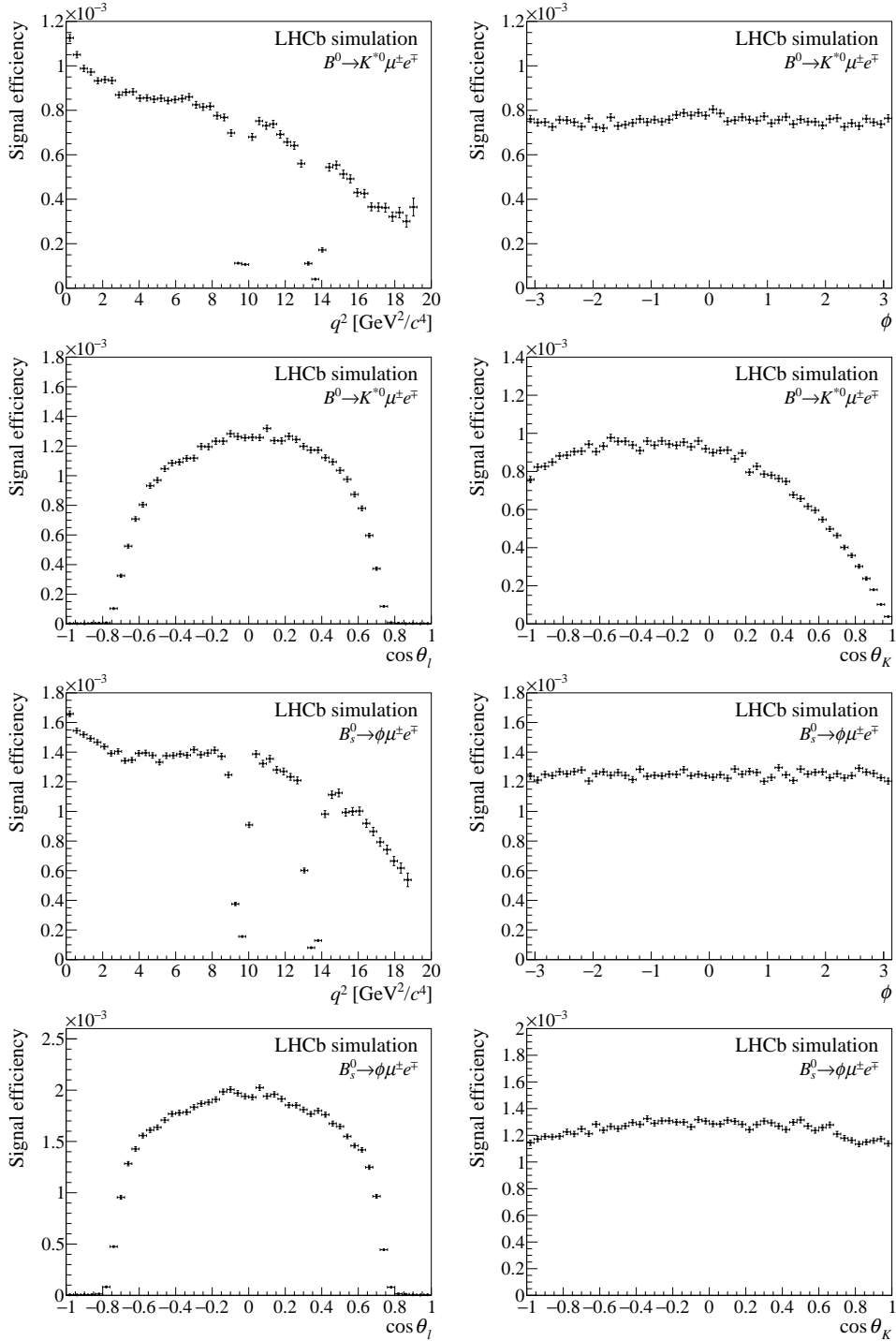
B   Efficiency projections

Figure 5: Reconstruction and selection efficiency, combined for the 2011 - 2018 samples, for (top) $B^0 \to K^{*0}\mu^\pm e^\mp$ and (bottom) $B_s^0 \to \phi\mu^\pm e^\mp$ signal decays depending on q^2 and the three decay angles $\cos\theta_\ell$, $\cos\theta_K$, and ϕ. The selection efficiency drops in the q^2 region around the J/ψ and $\psi(2S)$ masses as a result of the veto against misidentified backgrounds.