Investigating the flavour anomalies through the rare beauty decay $B^0 \to K^{*0}\mu^+\mu^-$

Maria Carolina Feliciano Faria
maria.faria@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

October 2021

Abstract

We measure the $B^0 \to K^{*0}\mu^+\mu^-$ decay branching fraction as a function of the di-muon invariant mass squared, using the $B^0 \to J/\psi K^{*0}$ decay as normalisation channel. We use proton-proton collision data at $\sqrt{s} = 13$ TeV, collected by the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) during the years 2016, 2017 and 2018. The analysis is validated by comparing the ratio between the branching fractions of the two resonant channels, $B^0 \to J/\psi K^{*0}$ and $B^0 \to \psi(2S)K^{*0}$, with the current world average. Our result is consistent with this value within $0.2\,\sigma$. The $B^0 \to K^{*0}\mu^+\mu^-$ branching fraction results are also consistent with previous LHC Run 1 measurements and with the Standard Model (SM) predictions. These preliminary results are the most precise to date.

The $B^0 \to K^{*0}\mu^+\mu^-$ decay is a rare beauty process which can only proceed at loop order in the SM, making it very sensitive to new physics (NP). Tensions with the SM have already been reported in one of the parameters appearing in its decay rate, P_5', belonging to the class of discrepancies in the flavour sector commonly referred to as flavour anomalies. The $B^0 \to K^{*0}\mu^+\mu^-$ branching fraction can enter in global fits to help constraining NP scenarios and investigating the source of these alluring flavour anomalies.

Keywords: Flavour anomalies, rare decays, CMS experiment, new physics, B physics, LHC

1. Introduction

The Standard Model (SM), albeit a very successful theory, cannot be the final description of nature since it fails to provide explanation to some observed phenomena, namely the existence of dark matter and the origin of the current matter-antimatter asymmetry in the universe. As a consequence, searches for physics beyond the SM are well motivated and are one of the main driving forces for the development of particle accelerators, such as the Large Hadron Collider (LHC).

Thus far no clear sign of New Physics (NP) has emerged from collider data. Nevertheless, a class of discrepancies with the flavour sector of the SM has been slowly making its stand. These are the so-called flavour anomalies and have been detected in two sets of quark-level transitions: $b \to sll$ (beauty to strange quark plus pair of charged leptons) and $b \to cl\bar{\nu}$ (beauty to charm quark plus charged lepton and neutrino).

In this work, we study a decay which is a particular realisation of the $b \to sll$ transitions, the $B^0 \to K^{*0}\mu^+\mu^-$. Tensions with the SM have already been reported in one of the parameters that enters in its decay rate, P_5'. In Fig. 1, the latest LHCb result [16] on this parameter can be seen, where a tension with the SM of 2.5 and $2.9\,\sigma$ has been reported in the di-muon invariant mass squared (q^2) ranges $4.0 < q^2 < 6.0$ and $6.0 < q^2 < 8.0\,\text{GeV}^2$, respectively.
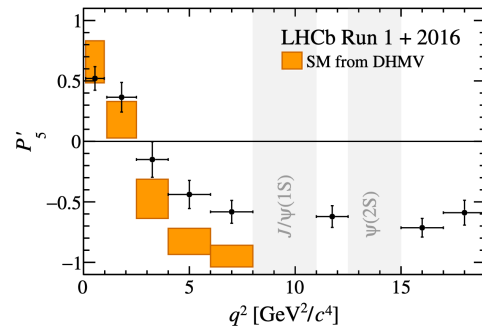

Figure 1: Latest LHCb result [16] on the P_5' parameter, where a tension with the SM of 2.5 and $2.9\,\sigma$ is reported in the ranges $4.0 < q^2 < 6.0$ and $6.0 < q^2 < 8.0\,\text{GeV}^2$, respectively (left).

Even though the individual discrepancies detected in these measurements do not yet display sufficient levels of statistical significance, when taken together in global fits [2] they do and they can be used to constrain NP models. Therefore, a clarification of the flavour anomalies is a current priority in High Energy Physics (HEP).

In this work, we measure the differential branching fraction of the $B^0 \to K^{*0}\mu^+\mu^-$ decay as a function of the di-muon invariant mass squared, using the $B^0 \to J/\psi K^{*0}$ as normalisation channel. We use proton-proton (pp) collision data at $\sqrt{s} = 13$ Tev, which was collected by the Compact Muon Solenoid (CMS) detector at the LHC, during the years 2016, 2017 and 2018. This observable can enter in the mentioned global fits, to help constraining NP scenarios and investigating the source of the alluring flavour anomalies.

The analysis is performed in bins of the di-muon invariant mass squared (q^2), as defined in Tab. 1. The q^2 ranges $8.68 < q^2 < 10.09$ GeV^2 and $12.86 < q^2 < 14.18$ GeV^2 correspond, respectively, to the two resonant channels $B^0 \to J/\psi K^{*0}$ and $B^0 \to \psi(2S)K^{*0}$, which occur much more frequently in the collisions. They are both used as control channels and the former is also used to normalise the branching fraction measurement

$$\frac{d\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)}{dq^2} = \frac{Y_S}{Y_N}\frac{\epsilon_N}{\epsilon_S}\frac{\mathcal{B}(B^0 \to J/\psi K^{*0})}{\Delta q_i^2},$$
(1)

where Y_S and Y_N are the yields of signal and normalisation channels, respectively, ϵ_S and ϵ_N are the detector efficiencies for signal and normalisation channels, respectively, and i runs over the signal q^2 bins. $\mathcal{B}(B^0 \to J/\psi K^{*0})$ is the branching fraction of the normalisation channel [17].

Table 1: q^2 bins used in the analysis.

bin index	q^2 range [GeV^2]
0	1-2
1	2-4.3
2	4.3-6
3	6-8.68
4	8.68 - 10.09 (J/ψ region)
5	10.09-12.86
6	12.86-14.18 ($\psi(2S)$ region)
7	14.18-16

The structure of this work is as follows. We start in Sec. 2 by giving a theoretical description of the $b \to sll$ transitions and how they are studied in the context of Effective Field Theories (EFTs). Within this formalism, the $B^0 \to K^{*0}\mu^+\mu^-$ decay rate is obtained. We continue on Sec. 3 with a brief description of the CMS detector, presenting the trigger (online) selections used in the analysis. In Sec. 4, we present the datasets and offline selections used. The method used to extract the yields Y_S and Y_N in Eq. (1) is described in Sec. 5, whereas the method used to compute the efficiencies ϵ_S and ϵ_N can be found in Sec. 6. The Monte Carlo (MC) simulations are validated in Sec. 7 and a study of the systematic uncertainties is presented in Sec. 8. Finally, we present the main results in

Sec. 9 and the conclusions in Sec. 10.

2. Theoretical framework

The $b \to sll$ transitions are highly suppressed in SM. On the one hand, because they correspond to a Flavour Changing Neutral Current (FCNC) transition, which is forbidden at tree level in the SM. On the other hand, because the leading order loop diagram is a penguin loop with a t-quark mediator (see Fig. 2 a)) which is Cabibbo suppressed, since it is proportional to the terms of the Cabibbo–Kobayashi–Maskawa (CKM) matrix $|V_{ts}V_{tb}| \approx 10^{-2}$. Other SM loop contributions proceed via box diagrams, as can be seen in Fig. 2 b).
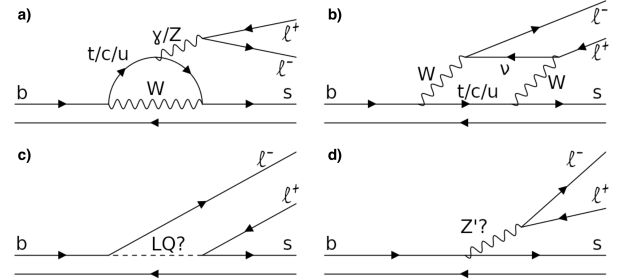

Figure 2: Loop diagrams contributing to the $b \to s\, l^+l^-$ transitions in the SM (top) and in possible NP scenarios (bottom). a) and b) correspond to a penguin loop and a box diagram, respectively. c) and d) correspond to tree level diagrams mediated by leptoquarks and a heavy gauge boson (Z'), respectively.

This fact makes decays realising them very rare and sensitive to NP, which could appear as yet undiscovered fundamental particles, such as leptoquarks (LQs) [13, 3] or a heavier gauge boson (Z') [1, 11], which would allow the existence of these FCNC transitions at tree level. These NP contributions could produce sizeable effects in the decay rates and branching fractions which would appear as flavour anomalies.

The energy scale of the $b \to sll$ transitions (the b-quark mass $m_b \approx 5$ GeV) is much lower than the electroweak scale (≈ 100 GeV) and possible NP scales (1-100 TeV). In cases where very disparate mass scales appear, it is advantageous to construct an EFT where the degrees of freedom that become relevant at much higher energy scales do not appear explicitly. In this description, the Feynman diagrams depicted in Fig. 2 reduce to point-like interactions, whose coupling constants are sensitive to the high energy degrees of freedom. In the presence of the high energy scale Λ, the effective Hamiltonian $\mathcal{H}_{eff}$, denoting a transition between an initial state $|i\rangle$ and a final state $|f\rangle$, can then be written as

$$\langle f | \mathcal{H}_{eff} | i \rangle = \sum_k C_k(\Lambda) \langle f | \mathcal{O}_k | i \rangle, \qquad (2)$$

where $C_k(\Lambda)$ are the WCs, which are the coupling constants of the local interactions and con-

tain physics information above Λ, being sensitive to NP, and $\langle f | \mathcal{O}_k | i \rangle$ are the matrix elements of the local operators $\mathcal{O}_k$, which contain physics information below Λ. NP can either change the values of the WCs or contribute with new operators $\mathcal{O}_k$.

2.1. The $B^0 \to K^{*0}\mu^+\mu^-$ decay rate

The $B^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$ has a fully charged final state composed of a kaon, a pion and two opposite charge muons, where the kaon and pion result from the decay of the K^{*0} meson. The four-particle final state can be described by a set of four kinematic variables: q^2, θ_l, θ_K and ϕ; with the three angles depicted in Fig. 3.
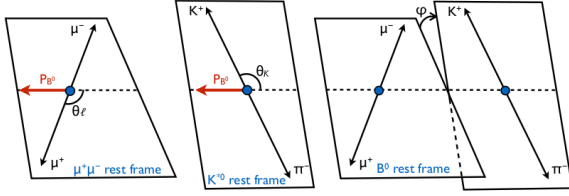

Figure 3: Graphical illustration of the three angular variables θ_l, $\theta_K \in [0, \pi]$ and $\phi \in [-\pi, \pi]$ used to describe the decay $B^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$. Taken from Ref. [6].

Since the $B^0 \to K^{*0}\mu^+\mu^-$ is a realisation of the $b \to sll$ transitions, its effective Hamiltonian as the form described in Eq. (2). In order to compute the decay rate, the amplitude $\mathcal{M} = \langle K\pi | \mathcal{H}_{eff} | \bar{B} \rangle$ needs to be obtained. The matrix elements of some operators can be expressed in terms of form factors, which below the charmonium resonances ($6\,\text{GeV}^2$) are calculated using Light Cone Sum Rules (LCSRs) [5] and above $6\,\text{GeV}^2$ are calculated using lattice Quantum chromodynamics (QCD) calculations. The matrix elements of the remaining operators are calculated using QCD factorisation techniques (QCDF) [4], below $6\,\text{GeV}^2$.

Since there are large theoretical uncertainties associated with the aforementioned form factor calculations, a set of parameters $P_i^{(')}$, which are form-factor independent at leading order, was constructed [12] in order to reduce them. In terms of these parameters, the $B^0 \to K^{*0}\mu^+\mu^-$ decay rate can be written as

$$
\frac{1}{d\Gamma/dq^2} \frac{d^4\Gamma}{dq^2 d\cos\theta_l d\cos\theta_K d\phi} = \frac{9}{32\pi} \Bigg[\frac{3}{4} F_T \sin\theta_K{}^2 + F_L \cos\theta_K{}^2
$$
$$
+ (\frac{1}{4} F_T \sin\theta_K{}^2 - F_L \cos\theta_K{}^2)\cos 2\theta_l
$$
$$
+ \frac{1}{2} P_1 F_T \sin\theta_K{}^2 \sin\theta_l{}^2 \cos 2\phi
$$
$$
+ \sqrt{F_T F_L}(\frac{1}{2} P_4' \sin 2\theta_K \sin 2\theta_l \cos\phi + P_5' \sin 2\theta_K \sin\theta_l \cos\phi)
$$
$$
\sqrt{F_T F_L}(P_6' \sin 2\theta_K \sin\theta_l \sin\phi - \frac{1}{2} P_8' \sin 2\theta_K \sin 2\theta_l \sin\phi)
$$
$$
+ 2 P_2 F_T \sin\theta_K{}^2 \cos\theta_l - P_3 F_T \sin\theta_K{}^2 \sin\theta_l{}^2 \sin 2\phi \Bigg],
$$
(3)

where F_T and F_L are the transverse and longitudinal polarisations of the K^{*0} meson, respectively. This decay rate is used to compute the detector efficiency in Sec. 6.

3. CMS detector and online selections

CMS is a general purpose detector situated at the CERN LHC. It has a cylindrical shape and is designed around a superconducting solenoid, which provides a magnetic field $B = 3.8\,\text{T}$. A schematic representation of the detector can be seen in Fig. 4 and a detailed description can be found in Ref. [7].

The main detector components used in this analysis are the silicon tracker and the muon system. The silicon traker measures charged particles within the pseudorapidity range $|\eta| < 2.5$, where

$\eta = -\ln[\tan(\theta/2)]$ and θ is the polar angle of the particle relative to the beam direction. It consists on a pixel and a miscrostrip detector and is located inside the superconducting solenoid. Muons are measured within the pseudorapidity range $|\eta| < 2.4$, with four muons stations consisting of several layers of drift tubes and cathode strip chambers in the regions $|\eta| < 1.2$ and $0.9 < |\eta| < 2.4$, respectively, complemented by resistive plate chambers convering the region $|\eta| < 1.6$. The muon chambers are interleaved with the iron return yoke (in red in Fig. 4), where the magnetic field is $B \approx 2\,\text{T}$.

As they traverse the silicon tracker, charged particles leave a series of hits in different layers of the material. Tracks are reconstructed by combining hits from several layers, allowing an estimate

of the parameters of their trajectories to be performed. From the knowledge of the magnetic field (B) and the radius of curvature of their trajectory (R), their transverse momentum $p_T = qBR$ can be obtained, where q is the electric charge. The reconstructed tracks have an impact parameter resolution of around $15\,\mu$m.

The primary and secondary vertices of the B meson correspond, respectively, to the region of space where the meson is created as a result of a pp collision and where the B meson decays. The vertices are reconstructed in the silicon tracker by grouping several tracks into vertex candidates (vertex finding) and determining the best estimates for the position of the vertices (vertex fitting). The probability of the tracks coming from the same interaction vertex is quantified by means of the χ^2 test. Since b hadrons are relatively long-lived, they travel a distance of a few millimeters in the detector. As a consequence, their secondary vertex is displaced from their primary vertex, which facilitates their reconstruction.

Muons are reconstructed independently in the silicon tracker (muon tracker track) and in the muon stations (standalone muon), making use of the residual magnetic field in the iron return yoke ($B = 2$ T). They are subsequently combined with tracks found in the silicon tracker to form global muons.
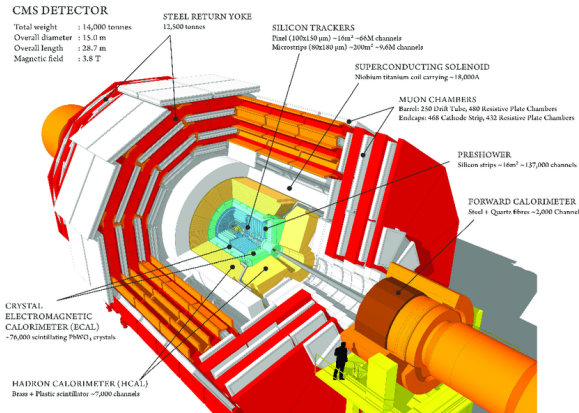

Figure 4: Schematic representation of the CMS detector identifying dimensions and different constituents.

Events are selected using a two-level trigger system composed by a Level 1 (L1) trigger and a High Level Trigger (HLT), which reduce the event rate to around $100\,$kHz and $1\,$kHz, respectively, before data storage. The L1 trigger is made from custom electronics and uses information from the calorimeters and the muon detectors to select the most interesting events. The HLT relies on $\approx$ 100 commercial processors and has access to the complete silicon tracker information, allowing more precise selections to be made online.

Events used in the analysis were collected with triggers requiring that two opposite charge muons

and an additional track are found in the event and form a 3-body displaced vertex. Some of the requirements applied at HLT are: single muon $p_T > 4\,$GeV and $|\eta| < 2.5$; di-muon $p_T > 6.9\,$GeV; di-muon invariant mass in the ranges $[2.9\text{-}3.3]\,$GeV (J/ψ); $[3.3\text{-}4.05]\,$GeV ($\psi(2S)$) and $[1\text{-}2.9] \cup [4.0\text{-}4.8]\,$GeV (non-resonant channel) and additional track with $p_T > 0.8\,$GeV and $|\eta| < 2.5$.

4. Datasets and offline selections

In this analysis, we use pp collision data at $\sqrt{s} = 13\,$TeV, collected by the CMS detector during the years 2016, 2017 and 2018, corresponding to a total integrated luminosity of $L = 139.5\,$fb^{-1}. Besides data samples, the analysis also uses Monte Carlo (MC) simulation samples. Pileup events are added to represent the number of multiple vertices per event as seen on data (pileup weight). The simulated samples are used to estimate background contributions, to compute the detector efficiency and perform studies of systematic uncertainties.

The fiducial region of the analysis is defined by the presence of a B^0 candidate with pseudorapidity $|\eta_B| < 3$. For $|\eta_B| > 3$, the probability of detecting the four final state particles resulting from the B^0 decay is very low and therefore a filter to these events is used in the MC simulations, to save computing time and disk space. This cut does not introduce any q^2 dependence on the efficiencies.

The detector acceptance is defined by the region of phase space within which the final state particles are potentially detectable. A set of requirements is applied on the generated MC in order to reject events falling outside the acceptance of the CMS detector: $|\eta(\mu^+, \mu^-, K^\pm, \pi^\mp)| < 2.5$, $p_T(\mu^+, \mu^-) > 2.5\,$GeV and $p_T(K^\pm, \pi^\mp) > 0.4\,$GeV. Events passing these selections are referred to as accepted events.

The signal and control channels $B^0 \rightarrow K^{*0}\mu^+\mu^-$, $B^0 \rightarrow J/\psi K^{*0}$ and $B^0 \rightarrow \psi(2S)K^{*0}$ are reconstructed through their decay into the fully charged final state $K^+\pi^-\mu^+\mu^-$. Only events passing the triggers described in Sec. 3 are considered. B^0 candidates are formed by combining two opposite charge muons with a K^{*0} meson. The two muons and at least one track should be matched to the objects firing the HLT. The selections applied at HLT are re-applied in the offline analysis.

Some of the requirements applied to muons are the following. Single muon $p_T > 4\,$GeV and $|\eta| < 2.5$. Number of track layers with measurements > 5. Number of pixel layers with measurements > 0. Muon track passing the high-purity flag [8] in order to reject tracks that do not correspond to a charged particle (fake tracks). Some of the requirements applied to the two hadron tracks are the following.

$p_T > 0.8\,\text{GeV}$ and $|\eta| < 2.4$. In 2017 and 2018, the p_T of the offline track matched to the one firing the trigger is required to be $> 1.2\,\text{GeV}$. Pass the high-purity flag [8]. The two tracks must successfully fit to a common vertex to form a K^{*0} meson. The cut $m_{KK} > 1.035\,\text{GeV}$, where m_{KK} is the invariant mass of the two hadron tracks with kaon mass assigned is applied in order to reject $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ events. The muon and the hadron candidates are fitted to a common vertex and a minimum χ^2 probability of 1 % is required.

Since both the $B^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$ and $\bar{B} \to \bar{K}^{*0}(K^-\pi^+)\mu^+\mu^-$ decays occur in the collisions, leaving the same signature in the CMS detector: two opposite charge muons and two opposite charge tracks; a flavour-tag is necessary to distinguish between the two. The K^{*0} invariant mass is computed for both $K^+\pi^-$ and $K^-\pi^+$ mass hypotheses, and at least one of the two combinations is required to lie in a $3\,\sigma$ mass window from the K^{*0} nominal mass [17], where σ is its natural width (50 MeV). The flavour-tag assignment consists on selecting the combination whose invariant mass is closest to the K^{*0} nominal mass [17]. This assignment does not always result in the correct flavour for the tracks and, consequently, there is a fraction of mis-tagged events in our signal. Henceforth, events with the correct flavour will be referred to as right-tagged (RT) and the ones with mis-tagged flavour will be referred to as wrong-tagged (WT). The mis-tag fraction (f^M) is defined as the ratio between the number of WT events and the total number of signal events and is determined from simulation. The mis-tag fractions are around 12-13%, depending on the q^2 bin and year.

The analysis also contains background events which contaminate our signal. A multivariate analysis was performed [10] in order to reduce the amount of background. A Boosted Decision Tree (BDT) discriminator was used to separate signal from background events. The training was performed on signal MC and on background events from the data sidebands. The BDT cut used in the analysis was chosen by maximising the expected $S/\sqrt{S+B}$, where S is the signal yield extracted from a fit to the signal MC and B is the background yield extracted from a fit to the left and right data sidebands, extrapolated to the signal region. The working points giving the best $S/\sqrt{S+B}$ values are: BDT score > 0.99 (2016), > 0.97 (2017) and > 0.975 (2018).

Additional cuts are performed in the analysis in order to reject peaking and partially reconstructed backgrounds. Peaking backgrounds are decays that do not correspond to our signal and can result from e.g. other b-hadron decays, giving rise to structure in the mass spectrum. Partially reconstructed backgrounds are formed by decays similar to our signal where an additional particle in the final state is missed in the reconstruction. The presence of these decays is studied in the J/ψ channel. $B^+ \to J/\psi K^+$ peaking background events and $B^+ \to \psi(2S)K^+$ partially reconstructed events, with $\psi(2S) \to J/\psi\pi^+\pi^-$, in which the softest pion is missed, are rejected. Furthermore, a cut based on the di-muon invariant mass was performed in order to reject events from the resonant $B^0 \to J/\psi K^{*0}$ and $B^0 \to \psi(2S)K^{*0}$ channels that can leak into the adjacent signal q^2 bins.

The combinatorial background is present is every physics analysis and describes events in which the four particles in the final state do not all come from the same interaction point. Consequently, the 4-body invariant mass can have a random value. This constitutes the main source of background in our analysis and nees to be modelled accordingly, as explained in the next section.

5. Yields extraction

The yields Y_S and Y_N in Eq. (1) are extracted from an extended unbinned maximum likelihood fit to the $K^+\pi^-\mu^+\mu^-$ invariant mass distribution. The signal shape, containing both RT and WT components, is described by a combination of crystal-ball and Gaussian functions, whereas the combinatorial background is described by an exponential function. In the case of the J/ψ channel, there is an additional background component describing unremoved partially reconstructed backgrounds. The shape of this component, taken from MC simulations, is given by a complementary error function. For all years, the fraction of partially reconstructed events with respect to signal is fixed to the value obtained from a fit to the 2018 data distribution $f_{erf} = 0.0672 \pm 0.0018$, where this value was left free to float.

Since in data there is not enough information to separate RT from WT events, a fit to the RT and WT MC is firstly performed. Subsequently, the fit to data is done applying Gaussian constraints on all parameters, except on the RT mean and yields, using the values and uncertainties of the parameters extracted from the MC fits.

The data fit results in year 2018 are shown in Fig. 5. The total fit (in black) can be seen superimposed with the data points. The RT and WT components are painted in green and red, respectively. The exponential curve of the combinatorial background is shown in blue and the component describing the partially reconstructed events in the J/ψ channel is shown in cyan. The q^2 ranges covered by each q^2 bin can be seen in the top left region of each plot, alongside the signal yields.
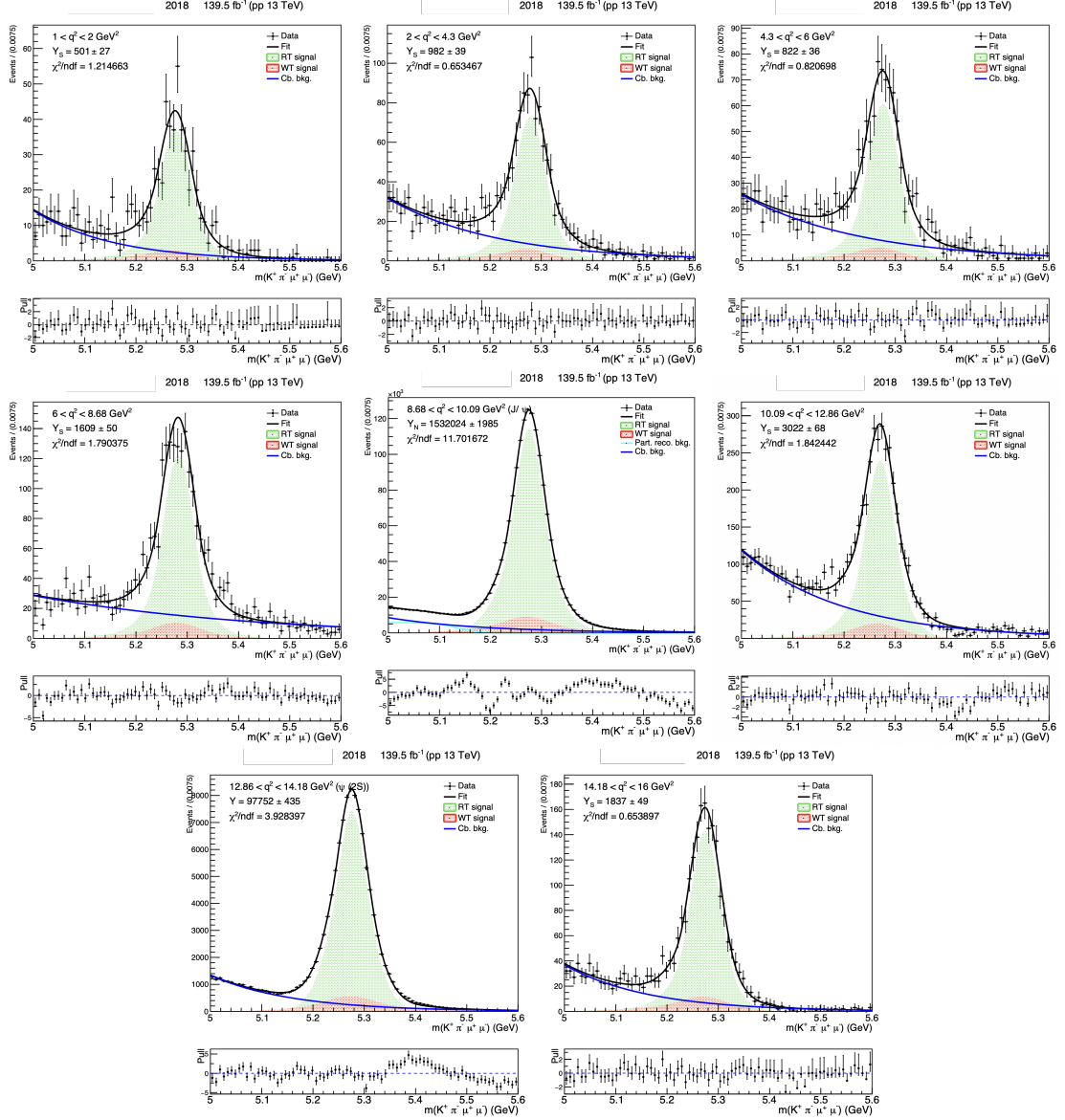
Figure 5: Data fit results in year 2018. The q^2 ranges covered by each q^2 bin can be seen in the top left region of each plot. The RT and WT components are shown in green and red, respectively. The combinatorial background is shown in blue and the partially reconstructed background in the J/ψ channel in cyan.

A toy MC study was performed in order to check whether the estimates for the signal yields are unbiased and whether we can trust their statistical uncertainties. 5000 toy samples were generated and fitted with the same model used to fit the data. We obtained the pull distribution

$$\text{Pull}(Y) = \frac{Y_i^{toy} - Y^{data}}{\sigma_i^{toy}}, \qquad (4)$$

where Y_i^{toy} is the signal yield obtained from the fit to the toy MC sample with index i, σ_i^{toy} is the corresponding statistical error and Y^{data} is the signal yield obtained from the fit to data.

If the estimator for the signal yield Y^{data} is unbiased, we expect to obtain statistically compatible estimates in each repeated measurement. The

distribution of pulls for an unbiased fit should follow a Gaussian distribution with unit width, centred upon zero. Shifts in the mean can suggest a bias in the fitting procedure, while deviations from a unit width can suggest that the statistical uncertainty returned by the fit is not reliable. The fit bias is accounted for in a systematic uncertainty, as explained in Sec. 8.

6. Detector efficiency

The detector efficiency is defined as the ratio between the number of signal candidates that are accepted, reconstructed and selected over the total number of signal events produced in the collisions, within the fiducial region of our analysis. The efficiencies ϵ_S and ϵ_N of both signal and normalisation channels, are used to correct the measured signal

6

yields Y_S and Y_N obtained in Sec. 5, so that we can compare our experimental results with the theory predictions.

In order to take into account the effect of the pileup re-weighting mentioned in Sec. 4, the detector efficiency is described by the product between two terms: the selection efficiency (N_{sel}/D_{sel}) and the acceptance (N_{acc}/D_{acc}) as

$$\epsilon = \frac{N_{acc}}{D_{acc}} \times \frac{N_{sel}}{D_{sel}}. \qquad (5)$$

The acceptance is the fraction between the number of accepted signal candidates (N_{acc}) over the total number of signal candidates within the fiducial region of the analysis (D_{acc}). The selection efficiency is the fraction between the number of reconstructed signal candidates passing all the analysis selections (N_{sel}) over the number of accepted signal events (D_{sel}). N_{sel} and D_{sel} have the pileup weight, whereas N_{acc} and D_{acc} do not.

The efficiencies are computed using MC simulations, which are generated with values for the angular parameters appearing in the $B^0 \to K^{*0}\mu^+\mu^-$ decay rate in Eq. (3) that are close to the SM predictions. Since probing the flavour anomalies is one of the goals of our study, the efficiencies are integrated over the angular variables weighted by the $B^0 \to K^{*0}\mu^+\mu^-$ decay rate. The angular parameters are taken from a 4-dimensional likelihood fit to the $K^+\pi^-\mu^+\mu^-$ invariant mass and the three angles θ_K, θ_l and ϕ describing the decay. Both the angular parameters obtained from the fits to data and MC are used. The data signal angular parameters are taken from Ref. [16] while the remaining ones are from Ref. [10].

The detector efficiency for year 2018 can be seen in Fig. 6, for each q^2 bin. The acceptance is the same for the three years.
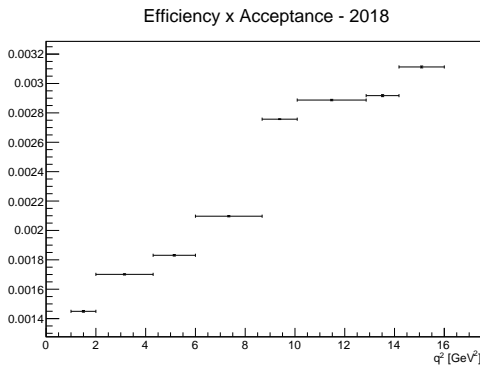

Figure 6: Averaged efficiencies for each q^2 bin in year 2018.

7. Monte Carlo validation

Since the efficiencies obtained in the last section rely on the MC simulations, we need to validate the MC by comparing it with data. Since the MC used contains only signal events, before this comparison can be made, the signal and background distributions in data need to be separated. This has been achieved by making use of the $_s\mathcal{P}lot$ method [18]. The data and MC signal distributions were then compared for a set of control variables. The largest discrepancies between data and MC were found for the pseudorapidity of the B^0 meson ("bEta").

In Fig. 7, the comparison between data and MC can be seen for the B^0 meson pseudorapidity in year 2018, for the J/ψ channel. The signal data and MC distributions are shown in red and green, respectively, and they are both normalised to the "bEta" range. In the bottom panel, the bin by bin ratio between the two distributions can be seen. These data/MC ratios are used to quantify the data vs MC discrepancies in a systematic uncertainty, as will be explained in Sec. 8.
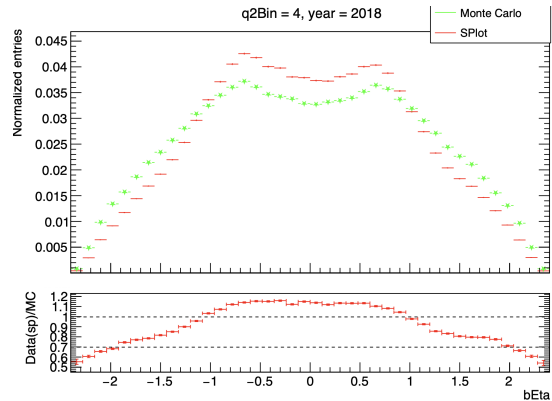

Figure 7: Comparison between data signal distribution (in red) obtained with he $_s\mathcal{P}lot$ method and the MC (in green), for the B meson pseudorapidity in year 2018 for the J/ψ channel.

8. Systematic uncertainties

Systematic uncertainties originate from the assumptions and methods used to extract the yields (Sec. 5) and to compute the efficiencies (Sec. 6) necessary to obtain the $B^0 \to K^{*0}\mu^+\mu^-$ branching fraction defined in Eq.(1). The uncertainties of the input branching fractions $\mathcal{B}(B^0 \to J/\psi K^{*0})$ and $\mathcal{B}(J/\psi \to \mu^+\mu^-)$ are also considered as part of the systematic error, being 3.94 % and 0.55 %, respectively.

The systematic uncertainties associated with the yields have two sources: fit variations and fit bias. The former is estimated by using different fit configurations, other than the nominal presented in Sec. 5, to extract the yields and it quantifies the effect the assumptions made in the nominal fit has in the yield results. The latter quantifies how reliable the yields obtained with the nominal fit are.

We used the following fit variations

- **Scale factor**: all the fit parameters, except the RT mean and the yields, are fixed to the values

found in MC. A "scale factor" parameter is multiplied by all the widths to allow for possible differences in the mass resolution between data and MC. This parameter is left free to float in the two resonant q^2 bins and is fixed to the values obtained in the J/ψ channel for the signal q^2 bins. The largest value for the scale factor is found in the J/ψ channel, in year 2017, where the data prefers a shape 10% wider than in MC.

- **Mistag fraction**: nominal fit in which the fraction of WT events is fixed to the values obtained from a fit to the MC containing both RT and WT events, applying the same Gaussian constraints used to fit the data. It allows for possible differences in the amount of WT events in data and MC.

- **No erf**: the component describing the partially reconstructed events in the J/ψ channel is removed.

- **Largest/lowest** f_{erf}: nominal fit in which the fraction of partially reconstructed events with respect to signal (f_{erf}) is fixed to its estimated largest and lowest limits, respectively. In order to estimate these limits, several fit configurations were performed in which f_{erf} was left free to float. The variation giving the largest values for f_{erf} is the one that has the largest number of degrees of freedom: the RT parameters are unconstrained. The variation giving the lowest values for f_{erf} is the one that has the lowest number of degrees of freedom: the scale factor variation.

The yield uncertainties associated with the fit variations are estimated in the following way. The difference between the signal yields obtained with each variation and the nominal fit is computed. Only the maximum difference between the variations "no erf", "largest f_{erf}" and "lowest f_{erf}" and the nominal fit is considered in the J/ψ channel. The systematic error is then given by the quadratic sum of each independent variation.

In the signal q^2 bins, the largest contribution to the systematic error comes from the scale factor variation, whereas in the normalisation channel, it comes from the "no erf" variation, in which only a single exponential is used to describe the background. The largest error of 3.99 % is found in q^2 bin 3 in year 2017 and the lowest error of 0.62 % is found in q^2 bin 6 ($\psi(2S)$) in year 2016.

The total yield systematic uncertainty is given by the quadratic sum of the two considered sources: fit variations and fit bias. The fit bias error is given by the product between the mean of the signal yield pull distributions obtained with the toy MC study

(see Sec. 5) and the yield statistical uncertainty. The fit bias errors do not change significantly the yield systematic uncertainties. The largest bias of 0.53% is found in q^2 bin 0, in year 2016, where the statistical error is the largest. The yield statistical and systematic uncertainties are propagated to the branching fraction.

The ratios between the B^0 pseudorapidity signal distribution from data ($_s\mathcal{P}lot$) and MC in the J/ψ channel are used to compute a weighted efficiency ϵ^{wei}, in which the selection efficiency term (N_{sel}/D_{sel}) in Eq. (5) is re-weighted with the values of the data/MC ratios. The branching fraction is then computed with ϵ and ϵ^{wei} and a systematic error is estimated as the difference between the two. This error quantifies the discrepancies between data and MC presented in Sec. 7. The largest error of 1.4% is found for q^2 bin 0, in year 2018, whereas the lowest error of 0.03% is found for q^2 bin 7, in year 2016.

Since the method used to compute the efficiencies in Sec. 6 makes use of the angular parameters from Refs. [10, 16], their statistical uncertainties need to be accounted for in a systematic uncertainty. For this purpose, two additional "efficiencies" ϵ^+ and ϵ^- are computed. The former, uses the nominal values plus statistical uncertainties of all the angular parameters simultaneously, whereas the latter uses their nominal values minus statistical uncertainties. Two additional branching fractions are calculated using ϵ^+ and ϵ^-. The difference between the two and the nominal branching fraction is computed and the systematic error is defined as the maximum difference with respect to the nominal value. The largest error of 1.26% is found for q^2 bin 5 in year 2017, whereas the lowest error of 0.11% is found for q^2 bin 1 in year 2017.

Finally, the finite size of the MC is also considered for the final systematic errors of the results. It corresponds to the efficiency statistical uncertainties as given by the `TEfficiency` class. The largest error of 0.42% is found for q^2 bin 0 in year 2016, whereas the lowest error of 0.16% is found for q^2 bin 5 in year 2018.

9. Results and discussion

The analysis is validated by comparing the ratio between the branching fractions of the two resonant channels, $B^0 \rightarrow \psi(2S)K^{*0}$ and $B^0 \rightarrow J/\psi K^{*0}$, with the current world average [17]. The PDG value is given by

$$R_{PDG} = \frac{\mathcal{B}(B^0 \rightarrow \psi(2S)K^{*0}) \times \mathcal{B}(\psi(2S) \rightarrow \mu^+\mu^-)}{\mathcal{B}(B^0 \rightarrow J/\psi K^{*0}) \times \mathcal{B}(J/\psi \rightarrow \mu^+\mu^-)},$$
(6)

where $\mathcal{B}(B^0 \rightarrow \psi(2S)K^{*0}) = (5.9 \pm 0.4) \times 10^{-4}$, $\mathcal{B}(\psi(2S) \rightarrow \mu^+\mu^-) = (80 \pm 6) \times 10^{-4}$, $\mathcal{B}(B^0 \rightarrow J/\psi K^{*0}) = (1.27 \pm 0.05) \times 10^{-3}$ and $\mathcal{B}(J/\psi \rightarrow$

$\mu^+\mu^-) = (5.961 \pm 0.033) \times 10^{-2}$. We compute this ratio as

$$R = \frac{Y}{Y_N}\frac{\epsilon_N}{\epsilon}, \qquad (7)$$

where Y, Y_N are the $B^0 \to \psi(2S)K^{*0}$ and $B^0 \to J/\psi K^{*0}$ signal yields, respectively, obtained as explained in Ch. 5 and ϵ, ϵ_N are the $B^0 \to \psi(2S)K^{*0}$ and $B^0 \to J/\psi K^{*0}$ averaged efficiencies, respectively, computed as described in Sec. 6.

In Fig. 8, the obtained results can be seen with the respective statistical (black) and systematic (red) uncertainties. The world average value [17] is shown in the bottom and the dashed blue line indicates its nominal value for reference. The single year results as well as the three years combined result in the top can be seen. The three years combined result is in agreement with the world average value within $0.2\,\sigma$, therefore validating the analysis.
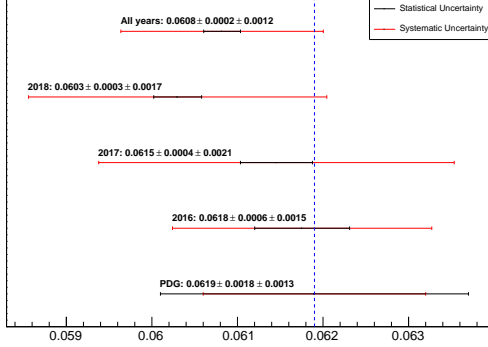
Figure 8: Single year and three year average (top) results for the ratio between the branching fractions of the two resonant channels, $B^0 \to \psi(2S)K^{*0}$ and $B^0 \to J/\psi K^{*0}$, with the respective statistical (black) and systematic (red) uncertainties. On the bottom, the world average value can be seen with the dashed blue line indicating its nominal value for reference.

The combination of the branching fraction results from the three years is performed using the weighted average

$$\bar{B} = \frac{\sum_{i=1}^{3} w_i B_i}{\sum_{i=1}^{3} w_i}, \qquad (8)$$

where B_i denotes generically the branching fraction of each of the three years, the weights w_i are defined as $w_i = 1/\sigma_i$, where σ_i is the statistical uncertainty of each measurement.

In Fig. 9, the three years averaged result can be seen in black. Both the statistical and systematic error bars are shown in the plot. The former has a horizontal tick in the end, whereas the latter does not. The two grey bands indicate the q^2 ranges of the two resonant channels, $B^0 \to J/\psi K^{*0}$ and $B^0 \to \psi(2S)K^{*0}$. Superimposed in the plot are the previous CMS [9] (lilac) and LHCb [15] (grey)

Run 1 results as well as the SM predictions [5, 14] (pink shaded region). At low q^2, below the charmonium resonances, the $B^0 \to K^{*0}$ form factor calculations come from LCSRs [5] whereas at high q^2, they come from lattice QCD [14].
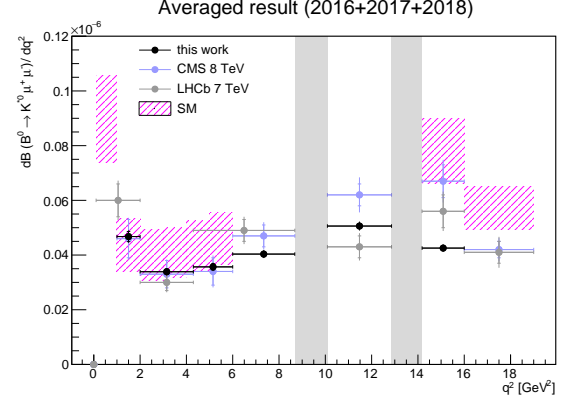
Figure 9: Three years average result (black) superimposed with the previous CMS [9] (lilac) and LHCb [15] (grey) Run 1 results as well as the SM predictions [5, 14] (pink shaded region). The two grey bands represent the q^2 ranges of the $B^0 \to J/\psi K^{*0}$ and $B^0 \to \psi(2S)K^{*0}$ resonant channels.

From Fig. 9, we can see that our results are consistent with the previous CMS [9] and LHCb [15] Run 1 measurements while having an increased precision. The factor of increase in the precision of the results in comparison with the previous CMS analysis can be seen in Tab. 2. Our results are also consistent with the SM, where the largest discrepancy in q^2 bin 7 is $2.9\,\sigma$ away from the theory prediction. In Tab. 3, the averaged branching fraction results for all q^2 bins can be seen with the respective relative statistical and systematic uncertainties. The systematic uncertainties are, in general, comparable with the statistical uncertainties.

Table 2: Factor of increase in the precision of the branching fraction measurement (three years averaged result) in comparison with the previous CMS Run 1 calculation [9].

q^2bin	0	1	2	3	5	7
Factor	3.39	3.96	3.80	2.80	3.06	4.77

Table 3: Differential branching fraction - weighted average

q^2	dB/dq^2 [10^{-7} GeV^{-2}]	Stat. error (%)	Syst. error (%)
0	$0.47 \pm 0.02 \pm 0.02$	3.99	3.71
1	$0.34 \pm 0.01 \pm 0.01$	2.95	3.75
2	$0.36 \pm 0.01 \pm 0.01$	3.28	3.80
3	$0.40 \pm 0.01 \pm 0.02$	2.25	3.81
5	$0.51 \pm 0.01 \pm 0.02$	1.75	3.82
7	$0.43 \pm 0.01 \pm 0.02$	2.10	3.56

10. Conclusions

In this thesis, we measured the differential branching fraction of the $B^0 \to K^{*0}\mu^+\mu^-$ decay as a function of the di-muon invariant mass squared (q^2),

using the $B^0 \to J/\psi K^{*0}$ resonance as normalisation channel. We used pp collision data at $\sqrt{s} = 13$ TeV collected by the CMS detector in Run 2 during the years 2016, 2017 and 2018. The analysis was performed independently for each data-taking year and q^2 bin.

The analysis was validated by computing the ratio between the resonant $B^0 \to J/\psi K^{*0}$ and $B^0 \to \psi(2S)K^{*0}$ branching fractions and comparing it with the current world average [17]. Our value was consistent with it within $0.2\,\sigma$, therefore validating our analysis The $B^0 \to K^{*0}\mu^+\mu^-$ branching fraction results were also found to be consistent with the previous CMS [9] and LHCb [15] measurements in Run 1 and with the SM predictions [5, 14], being the most precise to date.

The $B^0 \to K^{*0}\mu^+\mu^-$ is a rare beauty decay which can only proceed at loop order in the SM, which makes it very sensitive to NP. Tensions with the SM have been reported in one of the angular parameters appearing in its decay rate, P_5'. Other tensions in the flavour sector have also been reported in other observables / decays, being commonly referred to as flavour anomalies. The $B^0 \to K^{*0}\mu^+\mu^-$ branching fraction can enter in global fits which help constraining NP scenarios and investigating the source of these alluring flavour anomalies.

With the future efforts of both experimental and theoretical communities, with the former trying to reduce the experimental statistical errors, by increasing the luminosity of the experiments and the latter trying to reduce the large theoretical uncertainties of form-factor dependent observables, if the statistical significance of the flavour anomalies passes over the $5\,\sigma$ barrier, they will be the first established NP seen at the LHC.

References

[1] W. Altmannshofer, J. Davighi, and M. Nardecchia. Gauging the accidental symmetries of the standard model, and implications for the flavor anomalies. 2020. arxiv:1909.02021 [hep-ph].

[2] W. Altmannshofer and P. Stangl. New physics in rare B decays after Moriond 2021. 2021. arxiv:2103.13370 [hep-ph].

[3] A. Angelescu, D. Bečirević, D. A. Faroughy, F. Jaffredo, and O. Sumensari. On the single leptoquark solutions to the B-physics anomalies. 2021. arxiv:2103.12504 [hep-ph].

[4] M. Beneke and M. Neubert. QCD factorization for B —> PP and B —> PV decays. 2003. arxiv:0308039 [hep-ph].

[5] A. Bharucha, D. M. Straub, and R. Zwicky. $B \to V\ell^+\ell^-$ in the standard model from light-cone sum rules. 2016. arxiv:1503.05534 [hep-ph].

[6] A. Boletti. *Measurement of angular parameters from the decay $B^0 \to K^{*0}\mu^+\mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*. PhD thesis, Università degli Studi di Padova, 2018.

[7] CMS. The CMS experiment at the CERN LHC. 2008.

[8] CMS. Description and performance of track and primary-vertex reconstruction with the CMS tracker. 2014. arxiv:1405.6569 [physics.ins-det].

[9] CMS. Angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ from pp collisions at $\sqrt{s} = 8$ TeV. 2016. arxiv:1507.08126 [hep-ex].

[10] CMS. Full angular analysis of decay $B^0 \to K^{*0}\mu^+\mu^-$ at 13 Tev. 2021. AN-18-138.

[11] J. Davighi. Anomalous Z' bosons for anomalous B decays. 2021. arxiv:2105.06918 [hep-ph].

[12] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto. Optimizing the basis of $B \to K^*ll$ observables in the full kinematic range. 2013. arxiv:1303.5794 [hep-ph].

[13] G. Hiller, D. Loose, and I. Nišandžić. Flavorful leptoquarks at the LHC and beyond: spin 1. 2021. arxiv:2103.12724 [hep-ph].

[14] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate. Rare B decays using lattice QCD form factors. 2015. arxiv:1501.00367 [hep-lat].

[15] LHCb. Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$. 2013. arxiv:1304.6325 [hep-ex].

[16] LHCb. Measurement of CP-averaged observables in the $B^0 \to K^{*0}\mu^+\mu^-$ decay. 2020. arxiv:2003.04831 [hep-ex].

[17] PDG. Review of particle physics. *Progress of Theoretical and Experimental Physics*, 2020(8), 2020. 083C01.

[18] M. Pivk and F. R. Le Diberder. SPlot: a statistical tool to unfold data distributions. 2005. arxiv:physics/0402083.