LIP-STUDENTS-22-29

Investigating Flavour Anomalies

Sónia D'Azevedo[1,a] and Manuel Abreu[2,b]

[1] *Faculdade de Ciências, Universidade de Lisboa, Portugal*
[2] *Instituto Superior Técnico, Universidade de Lisboa, Portugal*

Project supervisors: Alessio Boletti and Nuno Leonardo *January 30, 2023*

Abstract. This paper deals with the $B^0 \to K^*\mu^+\mu^-$ decay, where we analyzed data collected by the CMS experiment in Run 2 of the LHC by studying the mass spectrum of the B^0 meson, using statistical methods to characterize and measure the signal and background properties, and computing the differential branching fraction as a function of the dimuon mass squared. This process is realized by the quark-level transition $b \to s\mu\mu$ that lies at the core of the current flavour anomalies.

KEYWORDS: LHC, CMS, flavor anomalies, branching fraction

1 Introduction

1.1 The CMS detector

The Large Hadron Collider (LHC) at CERN is home to countless projects associated with nuclear and particle physics research. CMS (Compact Muon Solenoid) is one of the detectors at the LHC. In this project we explore data that was collected by the CMS experiment.

Proton-proton collisions occur at the center of its cylindrical shape and the resulting particles pass through a number of sub-detectors in order to reconstruct their energy and trajectory. Right outside the collision chamber, a silicon tracker detects charged particles, and is followed by two calorimeters. Photons and electrons are identified by deposit of energy clusters in the electromagnetic calorimeter, while the hadron calorimeter, located after the EM calorimeter, helps identify hadrons. Outside these layers, there is a superconducting solenoid, which produces a magnetic field of 3.8 T, followed by the muon detectors, which consist of up to four stations of muon chambers, sandwiched between layers of iron return yoke.

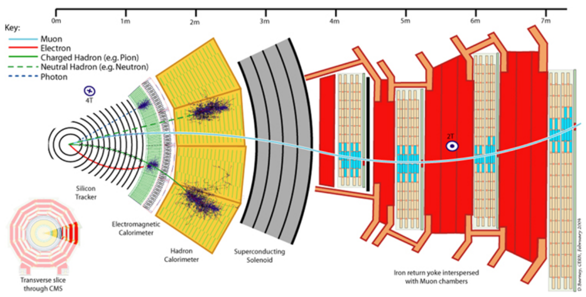

Figure 1. Schematic transverse view of the CMS detector

1.2 Lepton Flavour Universality

According to the Standard Model, the three generations of charged leptons differ from each other only in their mass, but otherwise have the same properties and interact in the same way with other particles – this principle is referred to as Lepton Flavor Universality (LFU). This would imply, for example, that the decay rates of B-mesons into electrons are the same as into muons when the effect of the leptons' mass is negligible. When significant differences in these decay rates are observed in experimental data, we are upon a flavor anomaly. And if this phenomenon is confirmed, it would be a sign of "New Physics" beyond the Standard Model as we know it.

1.3 The $B^0 \to K^*\mu^+\mu^-$ decay

In this project, we studied in particular the $B^0 \to K^*\mu^+\mu^-$ decay, with data collected by the CMS experiment in Run 2 of the LHC, in 2018. The main goal of our project was to determine the branching fraction of this decay (shown below). The measurement further uses the resonant $B^0 \to J/\Psi K^{*0}$ decay as normalization channel, since its branching ratio has been precisely measured by previous experiments. The measurement involves determining the signal yield (Y_s) and the normalization yield (Y_n), and both signal and normalized efficiencies (e_s and e_n):

$$\frac{d\mathcal{B}(B^0 \to K^*\mu^+\mu^-)}{dq_i^2} = \frac{Y_S}{Y_N}\frac{e_N}{e_S}\frac{\mathcal{B}(B^0 \to K^{*0}J/\Psi(\to \mu^+\mu^-))}{\Delta q_i^2} \quad (1)$$

Below is a list of the main variables used for our analysis. The yields can be measured by the detectors while the efficiencies are computed with Monte Carlo simulations.

- Y_S : The yield of the non-resonant channels;
- Y_N : The yield of the normalization channel;
- e_S : efficiency of the non-resonant channels;
- e_N : efficiency of the normalization channel;
- $\mathcal{B}(B^0 \to K^*J/\Psi)$: Branching Fraction of the normalization channel.

The decay in question produces 4 particles as a result - 2 leptons (which we can identify as muons) and 2 charged hadrons (which are produced by the K^* decay into $K^+\pi^-$

[a] e-mail: fc54756@alunos.fc.ul.pt
[b] e-mail: manuel.t.abreu@tecnico.ulisboa.pt

or its charge conjugate decay). It is known that one of the hadrons is negatively charged, while the other is positively charged. However, we do not known which of the 2 hadrons is the kaon or the pion, since both K^+ and K^- decays could occur. A way to resolve this issue will be discussed below.

2 Experimental procedure

2.1 Analysis in different q^2 bins regions

The RooFit package was used to fit the invariant mass spectrum of the B^0 candidate. We split the data in bins of the di-muon invariant mass squared, specified in Table 1. An independent analysis is made for each q^2 region because this way we are more sensitive to how New Physics may affect each q^2 region. Two bins (4 and 6) contain the resonant decays $B^0 \to K^* J/\Psi$ and $B^0 \to K^* \Psi'$, which are used for normalization and control channels.

Bin index	q^2 range
0	[1, 2]
1	[2, 4.2]
2	[4.2, 6]
3	[6, 8.68]
4	[8.68, 10.09]
5	[10.09, 12.86]
6	[12.86, 14.18]
7	[14.18, 16]

Table 1. Di-muon invariant mass squared q^2 ranges.

2.2 Flavour Assignment

The CMS detector detects 2 muons and 2 hadrons. For the 2 hadrons we can only detect their momentums and charges, so it is known that there is one positively charged hadron and one negatively charged, as previously mentioned. But how do we identify the hadrons for each event?

It's known that both $B^0 \to K^+\pi^-\mu^+\mu^-$ and $\bar{B}^0 \to K^-\pi^+\mu^+\mu^-$ decays may occur. The K^* mass is computed with both configurations. The configuration that results in a closer value to the K^* nominal mass, 892 MeV, is assigned to the event. Sometimes (it is not certain when), the event gets wrongly assigned. With a Monte Carlo simulation, since we generate the events, it is possible to determine the fraction of events that got wrongly assigned (12–14%, depending on the q^2 bin) as well as their distribution.

2.3 Fits

The signal yield is estimated through an unbinned maximum likelihood fit to the mass spectrum of the B^0 candidates. The shape of the signal mass distribution is estimated from MC simulated samples, and the presence of a fraction of events in the data sample with wrong flavour assignment requires that we estimate two shapes: one for correctly-tagged candidates, and one for wrongly-tagged ones. So the procedure is firstly to fit each of them separately.

2.3.1 Monte Carlo Fits

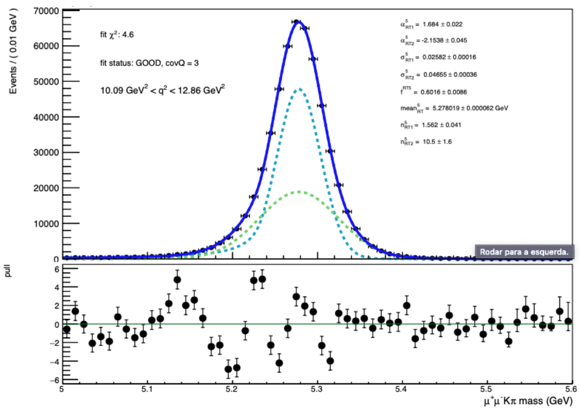

Figure 2. Fit to the B^0 mass spectrum in q^2 bin 5. This fit corresponds to the correctly tagged (RT) events. The pad below contains the pull, which is the distance between each point and the analytical function. The dashed lines are the functions used to parameterize the signal shape, in this case one Gaussian and one crystal ball, and the blue line is the sum of these two.

The function used to parameterize the signal shape is a double crystal ball for q^2 between 0 and 3, the sum of two simple crystal ball functions for q^2 bins between 4 and 6, and the sum of one simple crystal ball and a Gaussian function. The fit result for q^2 bin 5 is represented in Fig. 2, for correctly tagged events, and Fig. 3, for wrongly tagged events. The function used for the wrongly tagged events was a double crystal ball[1] for all q^2 bins.

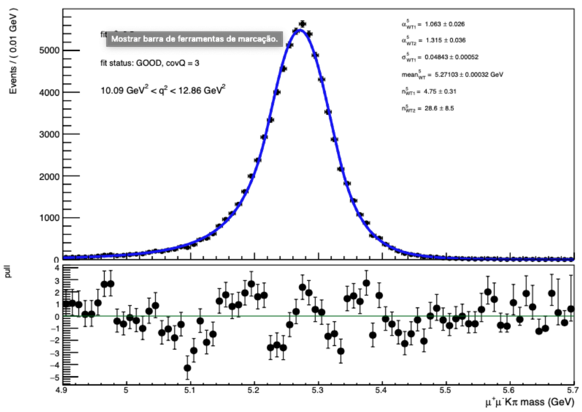

Figure 3. Fit to the B0 mass spectrum in q^2 bin 5. This fit corresponds to the wrongly tagged (WT) events. The pad below contains the pull. The blue line is the function used to parameterize the signal shape, in this case a double crystal ball.

2.3.2 Data Fits

After fitting the MC simulations, we proceeded to fit the data provided by the detector.

[1] a double crystal ball function is a function composed by a Gaussian kernel and two crystal ball tails, one on each side

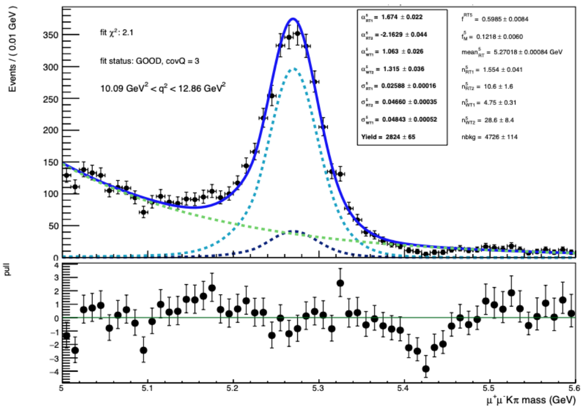

Figure 4. Fit to the B^0 mass spectrum in q^2 bin 5.

The probability density function (PDF) used for this fit has three components:

- The correctly-tagged component, the function used for the fit with Monte Carlo correctly tagged events;
- The wrongly-tagged component, the function used for the fit with Monte Carlo wrongly tagged events;
- The background component, which is an exponential function.

2.4 Yields

The Yields are obtained from the fits to the B^0 mass spectrum.

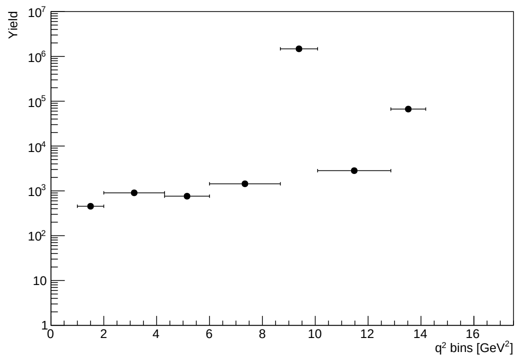

Figure 5. Yields for each q^2 bin region, including uncertainties. The y-axis is in logarithmic scale.

2.5 Efficiencies

The efficiencies are the ratio between the events that we detected and reconstructed successfully and the total of events created. The exact number of events created in reality is not known, but it is possible to estimate this value with the Monte Carlo simulations. The efficiency (ϵ) is obtained through the number of events selected by the MC simulations (N_{sel}) and the number of events generated (N_{gen}):

$$\epsilon_{nom} = \frac{N_{sel}}{N_{gen}} \tag{2}$$

However, the MC simulations may not represent the detectors perfectly, so it is important to take into account a correction factor (f_c) in the efficiency:

$$\epsilon_{corrected} = \epsilon_{nom} * f_c \tag{3}$$

This correction factor for MC not describing perfectly the detector effects can be added in the future, but the current result does not include it.

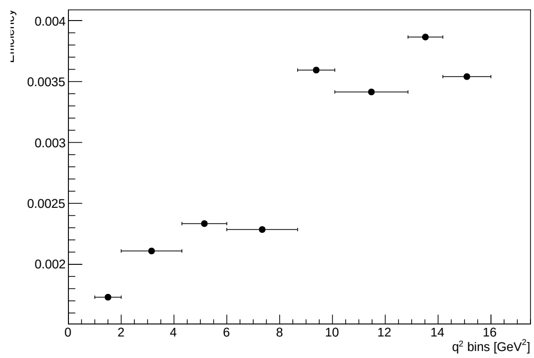

Figure 6. Efficiencies for each q^2 bin region, including uncertainties.

3 Results

3.1 Differential Branching Fraction

The Branching Fraction for the normalization channel is precisely measured by previous experiments, [2]

$$\mathcal{B}(B^0 \to K^* J/\Psi(\to \mu^+\mu^-)) = (7.5 \pm 0.3) * 10^{-5} \tag{4}$$

Once the yields and efficiencies are obtained, the Branching Fraction can be measured, in each q^2 bin region, as shown in equation (1).

[2] This result was obtained through the multiplication of 2 values: branching fraction of $J/\Psi(1S)$ decay through muons, (5.93 ± 0.06)% and branching fraction of B^0 decay to $J/\psi(1S)K^*(892)^0$, $(1.27 \pm 0.05)10^{-3}$ [5].

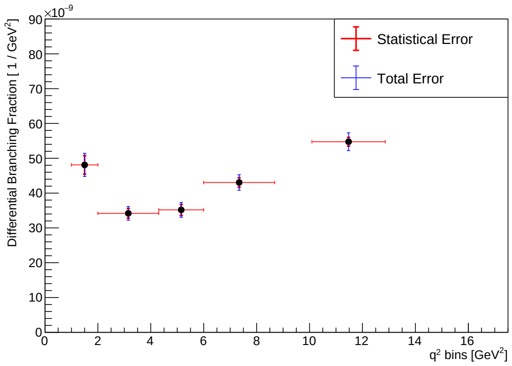

Figure 7. Differential Branching Fraction as function of q^2, including uncertainties, blue being the total uncertainty and red being the statistical uncertainty.

3.2 Uncertainties

The relative total uncertainty squared is the sum of both the statistical and systematic relative squared uncertainties.

The statistical relative squared uncertainty (σ^2_{stat}) is given by

$$\sigma^2_{stat} = \sigma^2_{SY} + \sigma^2_{NY} \tag{5}$$

- σ_{SY}: the signal yield's relative uncertainty (or the uncertainty of the signal yield divided by its value);
- σ_{NY}: the relative uncertainty for the yield of the normalization channel.

The yields' uncertainties are obtained through the fits.

The systematic relative squared uncertainty (σ^2_{syst}) is given by

$$\sigma^2_{syst} = \sigma^2_{SE} + \sigma^2_{NE} + \sigma^2_{BF} \tag{6}$$

- σ_{SE}: the signal efficiency's relative uncertainty and $\sigma_{SE} = 1/\sqrt{N_{sel}}$, where N_{sel} is the number of events selected by Monte Carlo from expression (2);
- σ_{NE}: the relative uncertainty for the efficiency of the normalization channel (computed the same way as the signal one);
- σ_{BF}: the relative uncertainty of the Branching Fraction of the normalization channel (the uncertainty of $\mathcal{B}(B^0 \to K^{*0} J/\Psi$ divided by its value, both precisely measured by previous experiments).

To obtain the total, systematic or statistical uncertainties, we multiply the respective relative uncertainty by the differential branching fraction value computed with equation (1).

4 Conclusions

We have performed a study of the $B^0 \to K^*\mu^+\mu^-$ decay, based on CMS data taken during LHC Run 2. The dimuon invariant mass squared (q^2) was divided in different regions. For each region, the signal yield is obtained from a fit to the B^0 mass spectrum. The efficiencies are computed, for each region, based on MC simulations. The differential branching fraction was thus measured as a function of q^2. The statistical uncertainties are obtained from the fit to the data while the systematic uncertainties are associated to the calculation of the efficiencies as well as the Branching Fraction of the normalization channel.

The differential branching fraction of the muonic decay is an important result to evaluate the LFU observable,

$$R_K = \frac{\mathcal{B}(B^0 \to K^*\mu^+\mu^-)}{\mathcal{B}(B^0 \to K^*e^+e^-)} \tag{7}$$

($R_K \approx 1$ in the Standard Model). A future measurement of the $B^0 \to K^*e^+e^-$ branching fraction can be combined with the result presented here to complete the test for LFU in $b \to sll$ decays.

Acknowledgements

We would like to thank Dr. Alessio Boletti and Prof. Nuno Leonardo for all the patience, support and knowledge they have provided us during this summer internship in LIP. It is thanks to them that we could deepen our understanding in this area of particle physics, a topic which we are both very interested in.

We would also like to extend our gratitude to Instituto Superior Técnico (IST) and Faculdade de Ciências da Universidade de Lisboa (FCUL), for giving us this opportunity to broaden our horizons and get a glimpse of what scientific research is like in the job market.

References

[1] M.Faria, A.Boletti, N.Leonardo, CMS meeting 7/2021
[2] M.Botas, J.Jafari, R.Pozzi, LIP-STUDENTS-21-23
[3] nLab, Flavour anomaly result compilation
[4] X.Shi, Studies of rare B decays with the CMS detector
[5] R.L. Workman et al. (Particle Data Group), Prog. Theor. Exp. Phys. 2022, 083C01 (2022).