IL NUOVO CIMENTO **39 C** (2016) 231
DOI 10.1393/ncc/i2016-16231-0

Colloquia: IFAE 2015

Search for new physics in the B meson decays: $B^0_{(s)} \to \mu^+\mu^-$

L. Martini

Dipartimento di Fisica, Università di Pisa and INFN, Sezione di Pisa - Pisa, Italy

received 7 January 2016

Summary. — The rare decays $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ are among the most promising investigation channels for the search of new physics. Their tiny branching fractions and their small theoretical relative uncertainties make these decays very sensitive to many new physics scenarios. I present the first successful reconstruction of the rare decay $B^0_s \to \mu^+\mu^-$, the measurement of its branching fraction, and the most recent upper limit for the branching fraction of the $B^0 \to \mu^+\mu^-$ decay at CMS, performed with all data acquired at the LHC Run I.

1. – Theoretical and experimental overview

In the Standard Model (SM) of particle physics, the $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays proceed through penguin and box diagrams and are helicity suppressed. For these reasons the SM predicts very small Branching Fractions (BF) [1]:

$$\mathcal{B}_{\mathrm{SM}}\left(B^0_s \to \mu^+\mu^-\right) = 3.66(23) \times 10^{-9}, \tag{1}$$

$$\mathcal{B}_{\mathrm{SM}}\left(B^0 \to \mu^+\mu^-\right) = 1.06(9) \times 10^{-9}. \tag{2}$$

The largest contributions to the total uncertainty come from the CKM and decay constants.

Their small theoretical relative uncertainties (less than 10%) contribute to make these decays very sensitive to many New Physics (NP) scenarios.

In 2013, the measurement of the $B^0_s \to \mu^+\mu^-$ BF has been shown for the first time at the EPS-HEP conference, independently by the LHCb and CMS collaborations. They published soon after their results in [2] and [3], respectively, after more than 30 years of experimental efforts by many collaborations.

In this paper I describe the CMS effort, that allowed to observe the $B^0_s \to \mu^+\mu^-$ decay, with the highest single-experiment significance up to today.

2. – The CMS analysis

A detailed description of the CMS detector can be found in [4]. This analysis is performed on the proton-proton sample of data collected during 2011 and 2012 at the LHC. The 7 TeV 2011 data correspond to an integrated luminosity of $5\,\mathrm{fb}^{-1}$, while the 8 TeV 2012 data correspond to $20\,\mathrm{fb}^{-1}$.

Due to the high LHC instantaneous luminosity, CMS uses a trigger system to select interesting events. For this analysis, the trigger system selects two reconstructed muons with a dimuon invariant mass compatible with the $B^0_{(s)}$ masses, and coming from a single loose vertex.

The collected yield is normalized to the $B^\pm \to J/\psi K^\pm$ decay, to minimize uncertainties related to the $b\bar{b}$ production cross section and to remove the uncertainty on the integrated luminosity. Furthermore, many efficiency systematic uncertainties are reduced, using a decay channel with a signature similar to the signal decay. The measured $B^0_s \to \mu^+\mu^-$ branching fraction, thus, is the following:

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = \frac{N_{B^0_s}}{N_{B^\pm}} \frac{f_\mathrm{u}}{f_\mathrm{s}} \frac{\varepsilon_{B^\pm}}{\varepsilon_{B^0_s}} \mathcal{B}(B^\pm) \tag{3}$$

where N_i is the number of reconstructed signal or normalization events, ε_i is the total efficiency for the signal or the normalization decay, $\mathcal{B}(B^\pm)$ is the branching fraction for $B^\pm \to J/\psi K^\pm \to \mu^+\mu^- K^\pm$ and $f_\mathrm{u}/f_\mathrm{s}$ is the ratio of the $B^\pm$ and B^0_s fragmentation fractions.

To control the distributions of the B^0_s signal, the $B^0_s \to J/\psi\phi$ decay is used in data as a control channel to validate the Monte Carlo (MC) B^0_s sample.

To avoid biases in the analysis selections, the invariant mass signal region is kept blind until all selection criteria are established ("blind analysis").

Multivariate studies, using boosted decision trees (BDT), are used to separate dimuon combinatorial background events, coming from separate weak B decays, from signal events. Without any selections, this combinatorial background is about five orders of magnitude larger than the SM signal expectation. Variables related to the secondary vertex, such as flight length, pointing angle, impact parameter and the isolation, are included in the BDT procedure. For the combinatorial background the BDT training utilizes events from the invariant mass side-bands, while MC simulations are used for the signal. As an example of distribution, fig. 1 shows on the left the BDT outputs for signal and background for one of the fitting categories. It is possible to appreciate the separation that is achieved, mainly thanks to the CMS inner tracker performances.

A clean muon identification is needed to reduce background events from rare two-body B decays (such as $B^0_s \to K^+K^-$) that form the "peaking" background, and rare three-body B decays (such as $B^0 \to \pi^-\mu^+\nu_\mu$) that form the "semileptonic" background. All these events contain hadrons wrongly identified as muons, mainly due to punch-through or decays-in-flight. To reduce this "muon misidentification" rate, a further BDT analysis is performed on track-related quantities, such as the χ^2 of the track global fit (fig. 1, on the right) and the number of track hits, that present different distributions between true muons and misidentified muons. With this single muon identification, it has been possible to halve the fake-muon efficiency, loosing only 10% of true muons, with respect to the standard CMS reconstruction algorithm for low-p_T muons.

The training is performed with MC $B^0_s \to \mu^+\mu^-$ muons against MC $B^0_s \to K^+K^-$ kaons, wrongly identified as muons. The BDT results are tested and validated for kaons,

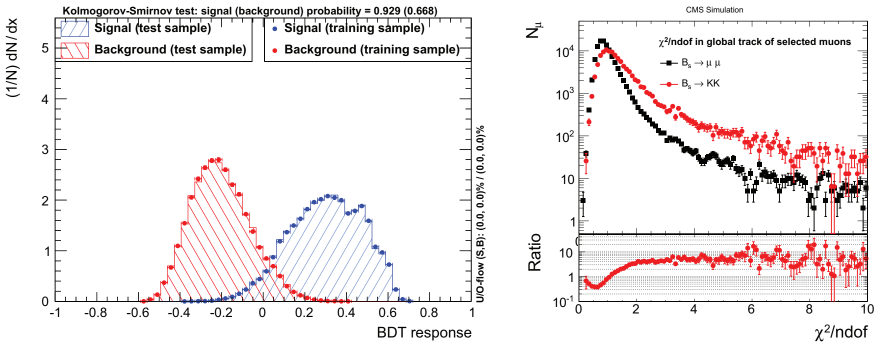

Fig. 1. – Overlay of the distributions of the BDT output for the signal and data-sidebands in the 2012 barrel category (on the left). Reduced χ^2 of the track global fit for true muons, in black and kaons, in red (on the right).

pions and protons on data through the control and normalization samples, the $D^{*+} \to D^0\pi^+ \to K^-\pi^+\pi^+$ decay and the $\Lambda^0 \to p\pi^-$ decay.

The combinatorial background arises from two reconstructed muons (real or misidentified hadrons), originating from separated particle decays. In the fit used to extract the signal yield, this background is parametrized with a linear function.

The rare background yields are normalized to the reconstructed $B^\pm \to J/\psi K^\pm$ yield in data, using the following formula:

$$N(X) = \frac{\mathcal{B}(X)}{\mathcal{B}(B^\pm \to J/\psi K^\pm) \times \mathcal{B}(J/\psi \to \mu^+\mu^-)} \frac{f_X}{f_u} \frac{\epsilon_X}{\epsilon_{B^\pm}} N^{B^\pm}_{\text{obs}}, \tag{4}$$

where X is the given rare decay, f_X is the hadronization probability and ϵ_X is the analysis acceptance and efficiency (taking into account also the muon misidentification). Figure 2 shows, as an example, the expected distributions of rare backgrounds in the barrel category for the 2012 sample.

The systematic uncertainties on the yields of these rare backgrounds take into account the uncertainties due to the muon misidentication probability and to the branching fractions.

An unbinned maximum likelihood fit on the invariant mass distribution is used to extract the signal and background yields. Data are divided into the two different data-taking periods (2011 and 2012) and two different η regions, whether any of the two muons crosses the endcap region of the detector, since resolution and background level are η-dependent. To take full information from the dimuon BDT analysis, the invariant mass distributions are sub-divided also in bins of the BDT variable, obtaining a total of twelve bins.

In each independent fit category, the likelihood is expressed as

$$L = N_{B^0_s}F_{B^0_s} + N_{B^0}F_{B^0} + N_{\text{comb}}F_{\text{comb}} + N_{\text{peak}}F_{\text{peak}} + N_{\text{semi}}F_{\text{semi}}, \tag{5}$$

where N_i is the number of events and F_i is the probability distribution function (pdf), for each contribution i. The five contributions are the two B^0_s and B^0 signals, the combinatorial background and the rare background, divided in the peaking and semileptonic

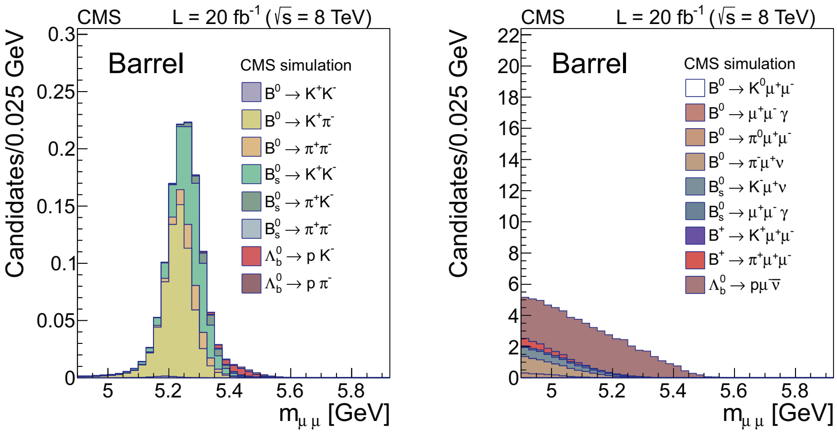

Fig. 2. – Invariant-mass distributions of rare peaking (on the left) or semileptonic (on the right) backgrounds for the barrel category in the 2012 sample.

parts. The two signals are parametrized with a Crystal Ball pdf with a per-event error width. The peaking background is described with a Crystal Ball, plus a Gaussian pdf, to take into account the enlarged width due to the diverse B peaks, and to the wrong mass hypotheses (kaons, pions and protons, instead of muons). The semileptonic background is parametrized with a Gaussian kernel distribution.

The full likelihood is therefore

$$L_{\text{tot}} = \prod_{i=1}^{12} L_i L_i^{\text{constr}}, \tag{6}$$

where L_i^{constr} represents the product of all constrained nuisance parameters per category.

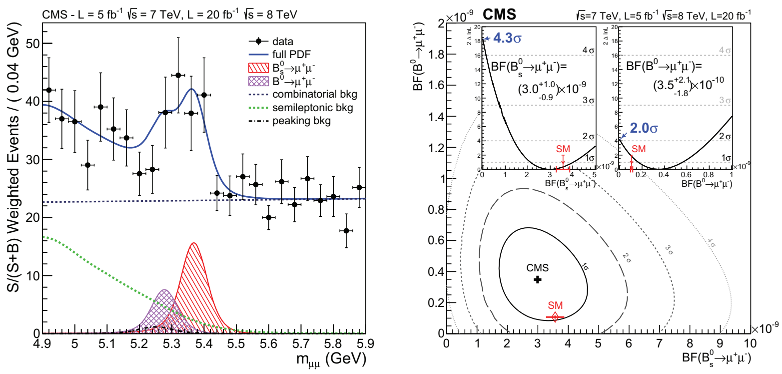

Fig. 3. – Left: Combination of all data categories for $B^0_{(s)} \to \mu^+\mu^-$. The individual categories are weighted with $S/(S+B)$, where S (B) is the signal (background) determined at the B^0_s peak position. Right: Scan of the ratio of the joint likelihood for $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$. As insets, the likelihood ratio scan for each of the branching fractions when the other is profiled together with other nuisance parameters; the significance at which the background-only hypothesis is rejected is also shown.

The final pdf has been studied with MC toy experiments, and shows no significant bias on the yield evaluation.

After all selections are fixed, the data under the $B^0_{(s)}$ yields are extracted. An excess of $B^0_s \to \mu^+\mu^-$ decays, consistent with the SM expectation, is observed above the background predictions, resulting in a measured branching fraction of $\mathcal{B}(B^0_s \to \mu^+\mu^-) = \left(3.0^{+1.0}_{-0.9}\right) \times 10^{-9}$. The significance of the excess, evaluated with the Wilk's theorem, is $4.3\,\sigma$. No significant excess is observed for the $B^0 \to \mu^+\mu^-$ decay, and an upper limit of $\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.1 \times 10^{-9}$ at 95% confidence level is determined with the $\mathrm{CL_S}$ method.

Figure 3 on the left shows, as an illustrative purpose, the fit results on the combined data. On the right, the ratio of the joint likelihood for $\mathcal{B}(B^0_{(s)} \to \mu^+\mu^-)$ is shown.

A full combination of the likelihood has been performed recently between the CMS and the LHCb results [5]:

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = \left(2.8^{+0.7}_{-0.6}\right) \times 10^{-9}, \tag{7}$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = \left(3.9^{+1.6}_{-1.4}\right) \times 10^{-10}. \tag{8}$$

The ratio is SM-compatible at the $2.3\,\sigma$ level. The significance of the B^0_s signal over the null hypothesis results $6.2\,\sigma$, while the B^0 significance sets to $3.2\,\sigma$. Thus, the empirical observation of the $B^0_s \to \mu^+\mu^-$ decay is claimed. This is also the first time of a full combined analysis between the two LHC collaborations.

REFERENCES

[1] Bobeth C., Gorbahn M., Hermann T., Misiak M., Stamou E. *et al.*, *Phys. Rev. Lett.*, **112** (2014) 101801.
[2] Aaij R. *et al.*, *Phys. Rev. Lett.*, **111** (2013) 101805.
[3] Chatrchyan S. *et al.*, *Phys. Rev. Lett.*, **111** (2013) 101805.
[4] Chatrchyan S. *et al.*, *JINST*, **3** (2008) S08004.
[5] Khachatryan V. *et al.*, *Nature*, **10** (2015) 14474.