

$B_{(s)}^0 \rightarrow \mu^+ \mu^-$ at CMS

FRANCO LIGABUE

*On behalf of the CMS Experiment,
Scuola Normale Superiore and INFN, Pisa*

ABSTRACT

The search for the rare $B^0 \rightarrow \mu^+ \mu^-$ and $B_s^0 \rightarrow \mu^+ \mu^-$ decays in pp collisions at $\sqrt{s} = 7$ GeV and $\sqrt{s} = 8$ GeV, collected at the LHC in 2011 and 2012, is briefly reviewed. The data analyzed by CMS correspond to a total integrated luminosity of 5 and 20 fb^{-1} , respectively. The time-integrated average branching fraction $\langle \mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) \rangle$ has been measured to be $(3.0_{-0.9}^{+1.0}) \times 10^{-9}$ in accordance with the Standard Model predictions, while an upper limit $\langle \mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) \rangle < 1.1 \times 10^{-9}$ has been placed on the other investigated decay at 95% CL. A preliminary combination with the results from LHCb is also presented for both channels, and prospects for the future are briefly discussed.

PRESENTED AT

The Second Annual Conference
on Large Hadron Collider Physics
Columbia University, New York, U.S.A
June 2-7, 2014

1 Introduction

The decays $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ are strongly suppressed in the Standard Model. They are Flavour Changing Neutral Current processes that are forbidden at tree level and can only proceed through higher-order (“box” and “penguin”) diagrams. Since they involve a scalar meson decaying weakly to two ultrarelativistic leptons, the decays are also helicity-suppressed. The purely leptonic final state makes it a relatively clean process to compute in the Standard Model, which predicts rather precise values for the branching ratios, namely $(3.23 \pm 0.27) \times 10^{-9}$ for $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$, and $(1.07 \pm 0.10) \times 10^{-10}$ for $\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)$. These values are expected to be significantly altered in several extensions of the Standard Model, which makes $B_{(s)}^0 \rightarrow \mu^+\mu^-$ a particularly clean probe for New Physics [1]. In some supersymmetric models, for instance, the sensitivity of $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$ to the fundamental parameter $\tan\beta$ can become as high as $\tan^6\beta$. The neutral-meson $B_{(s)}^0$ - $\bar{B}_{(s)}^0$ system is also subject to flavour oscillation, which changes the observable time-integrated flavour average branching fraction with respect to the flavour-eigenstate decay probability computed at production time ($t = 0$):

$$\langle \mathcal{B}(B_s^0 \rightarrow \mu\mu) \rangle = \frac{1 + \mathcal{A}_{\Delta\Gamma}^{\mu\mu} y_s}{1 - y_s^2} \mathcal{B}(B_s^0 \rightarrow \mu\mu)|_{t=0} \quad (1)$$

where $y_s \equiv \frac{\Delta\Gamma}{2\Gamma}$ is the relative width difference between the two mass eigenstates, and where the parameter $\mathcal{A}_{\Delta\Gamma}^{\mu\mu}$, which is unity in the Standard Model, can also be sensitive to New Physics.

2 CMS detector and muon identification

The CMS detector has been described in detail in several publications (see for instance [2]). Its relevant features here are basically its tracking, vertexing, and muon identification capabilities. The central silicon tracker, which is immersed in a 3.8 T uniform magnetic field parallel to the beam direction, can reconstruct charged tracks within a pseudo-rapidity range $|\eta| < 2.5$, with a p_T resolution around 1% and an impact parameter resolution of about 15 μm in the kinematical range relevant to this analysis. Three different types of muon-specific detectors (drift tubes, cathode strip chambers and resistive plate chambers) are located in the return yoke of the magnetic field.

Muons are very efficiently reconstructed and identified by combining silicon tracker and muon detector information. The reconstruction efficiency is checked on data with tag-and-probe techniques and is above 99% for a wide portion of the relevant phase space. In order to reduce the contamination from misidentified hadrons – a crucial issue for this analysis – a dedicated multivariate selection has been specially developed and applied on top of the standard CMS muon reconstruction. The variables used for this additional selection are either purely kinematic, related to silicon tracker or muon detector information, or related to tracking fit quality. Adding this supplemental selection reduces the hadron misidentification rate by roughly 50%, at the moderate expense of a 10% loss in efficiency. The final hadron misidentification probability ranges roughly from 0.5×10^{-3} to 2.2×10^{-3} .

3 Analysis strategy and data selection

The measurement strategy for detecting $B_{(s)}^0 \rightarrow \mu^+\mu^-$ decays in pp collisions is in principle straightforward: one must find a pair of oppositely charged muon tracks, compatible with originating from a common vertex, possibly displaced from the beam line, and whose invariant mass lies within a pre-defined interval centred around the $B_{(s)}^0$ mass. The signal can then be extracted by event counting or by fitting the mass spectrum, after suitable background subtraction. In order to derive a branching fraction from the measured yield, though, one needs to know very precisely the signal selection efficiency, as well as the $B_{(s)}^0$ production cross section and the absolute luminosity \mathcal{L} . A way of reducing the large systematics involved is to refer the observed yield to that of a “normalization” channel whose branching ratio is rather precisely known, namely $B^\pm \rightarrow J/\Psi K^\pm \rightarrow (\mu^+\mu^-)K^\pm$, by means of the following luminosity-independent formula:

$$\langle \mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) \rangle = \frac{N_{\text{obs}}^{B_s^0}}{N_{\text{obs}}^{B^\pm}} \times \frac{\varepsilon_{B^\pm}}{\varepsilon_{B_s^0}} \times \frac{f_u}{f_s} \times \mathcal{B}(B^\pm \rightarrow J/\Psi K^\pm) \times \mathcal{B}(J/\Psi \rightarrow \mu^+ \mu^-) \quad (2)$$

where the first factor in the multiplication is the ratio of the observed number of events for the signal and for the normalization channel, and the second is the ratio of detection efficiency, where many systematics uncertainties cancel to first order. The third factor (f_u/f_s) is the production ratio of B^\pm to B_s^0 mesons at LHC, and is taken from experiment [6].

The analyzed pp collision samples correspond to an integrated luminosity of 5 fb^{-1} collected by CMS in 2011 at $\sqrt{s} = 7 \text{ TeV}$, and 20 fb^{-1} collected in 2012 at $\sqrt{s} = 8 \text{ TeV}$. Events are selected via a two-level (online and offline) trigger, requiring the presence of two muons with an invariant mass within a specified window and with a minimum common vertex fit probability. After applying the multivariate selection for enhancing the muon purity, two adjacent signal regions and two sideband windows are defined within the overall accepted invariant mass region between 4.9 and 5.9 GeV/c^2 . The data falling in the signal region, between 5.20 and 5.45 GeV/c^2 are then never looked at (“blind” analysis) except in the final stages of signal yield extraction.

Exclusively reconstructed $B^\pm \rightarrow J/\Psi K^\pm$ and $B_s^0 \rightarrow (J/\Psi)\phi \rightarrow (\mu^+ \mu^-)(K^+ K^-)$, selected with kinematic cuts similar to those used for the signal, provide the so-called “normalization” and “control” samples. The former is used for branching ratio extraction according to eq. 2, both – but especially the latter – are used for data-simulation comparison and cross checks.

Most of the background events mimicking the signal involve the actual presence of b hadrons. Events where two real muons coming from semileptonic decays of separate b hadrons appear by chance to form the right mass (“combinatorial background”) are more or less flatly distributed in the invariant mass window, whereas events where the two selected tracks (either two real muons or a real muon and a misidentified hadron) come from a three-body semileptonic decay of a single b -hadron (“semileptonic background”) exhibit a monotonically sloping distribution (see left plot in fig. 1). Events where both the decay products of two-body b -hadrons decays (pions, kaons or protons) have been misidentified as muons are much more dangerous since they peak at the B mass (“peaking background”, see right plot in fig. 1) and their contribution cannot be interpolated from the data sidebands like for the previous category.

A multivariate (Boosted Decision Tree or BDT) technique [4] is used to build a discriminating variable providing a good separation of the signal from the combinatorial background. The most effective variables used in the combination turn out to be either related to the secondary vertex separation (the impact parameter and its significance, or the angle between the dimuon momentum and the vector pointing from the primary to the secondary vertex), or else “isolation” variables, such as the number of tracks close to the dimuon system, shown in the right plot of fig. 2. The BDT has been optimized using simulated events, as representative of the signal, and real data from the mass sideband windows, as representative of the combinatorial background.

4 Signal extraction and results

The events selected in the $B^0 \rightarrow \mu^+ \mu^-$ signal mass window have been analyzed with a simple counting experiment after a cut on the value of the BDT discriminating variable optimizing the sensitivity $S/\sqrt{S+B}$. The events were split into four samples according to the year of data taking and the pseudorapidity of the detected muons (“barrel” and “endcap” events). The signal select efficiency is evaluated to range roughly from 1 to 3 permille.

No significant excess is observed for this channel, and an upper limit is placed on the branching ratio at 95% confidence level using the CL_s approach, obtaining

$$\langle \mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) \rangle < 1.1 \times 10^{-9}$$

which is compatible with the Standard Model predicted limit at the 2σ level, though on the higher side, as shown in the left plot of fig. 4.

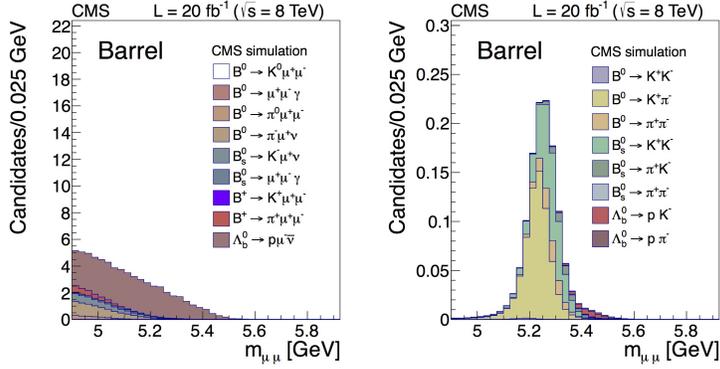


Figure 1: Dimuon invariant mass distribution for background events from rare 3-body semileptonic (left) and 2-body hadronic decays single b-hadrons in simulation.

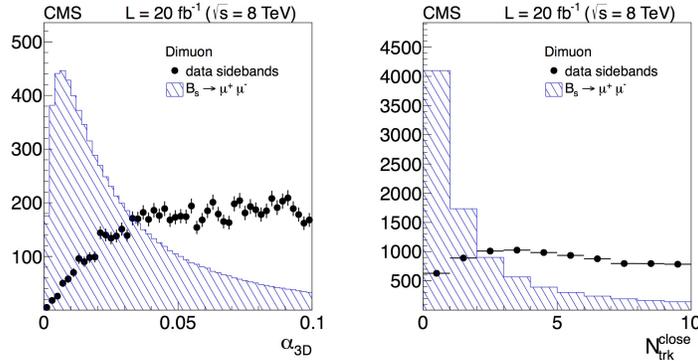


Figure 2: Distribution of two of the most effective variables used in the BDT technique, for simulated signal events (histogram) and data non-signal events from mass sidebands (data points). Left: pointing angle between B candidate momentum and primary-to-secondary vertex direction. Right: number of tracks in the vicinity of the B candidate.

The $B_s^0 \rightarrow \mu^+ \mu^-$ signal is extracted with an unbinned maximum likelihood fit to the invariant mass spectrum applied after splitting the data into twelve subsamples or “categories” according to the value of the BDT discriminating variable and, again, the year of data taking and the pseudorapidity of the detected muons. The size and boundaries the BDT intervals have been chosen so as to have roughly the same number of expected signal events in each category. For each category the signal shape is described by a Crystal Ball function with a variable (per-event) mass-resolution parameter. The combinatorial background is described by a straight line, while the shapes of the backgrounds from rare b-hadron decays are taken from simulation. The estimated systematic uncertainties (for instance due to uncertainties on hadron misidentification probability, or on the peaking background normalization) are included in the fit as further Gaussian constraints. A signal excess is observed, as is visible in the illustrative plot of fig. 3, corresponding to a measured value for the time-integrated branching fraction of

$$\langle \mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) \rangle = (3.0_{-0.9}^{+1.0}) \times 10^{-9}$$

where the quoted uncertainty includes both the (dominating) statistical component and the systematic contributions. The systematics is dominated by the uncertainty on the value of the f_u/f_s factor in eq. 2, which is taken from a published LHCb measurement [6].

The right plot in fig. 4 shows the contours of the joint likelihood for the two branching ratios, along with the Standard Model prediction, which appears to be compatible at the 1σ level with the results. The top

right inset in the same plot shows that the observed $B^0 \rightarrow \mu^+ \mu^-$ excess is at the level of 2.0σ with respect to the background-only hypothesis.

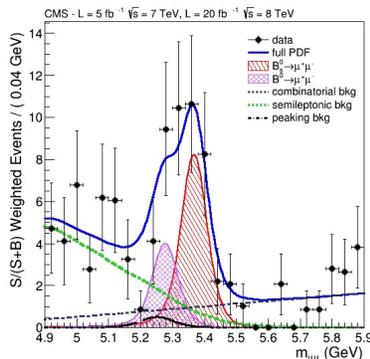


Figure 3: Plot illustrating the weighted average of the fit results on the invariant mass spectrum in the 12 data categories, for $B_s^0 \rightarrow \mu^+ \mu^-$ yield extraction. The weights are proportional to the signal purity $S/(S+B)$ determined at the B_s^0 peak position.

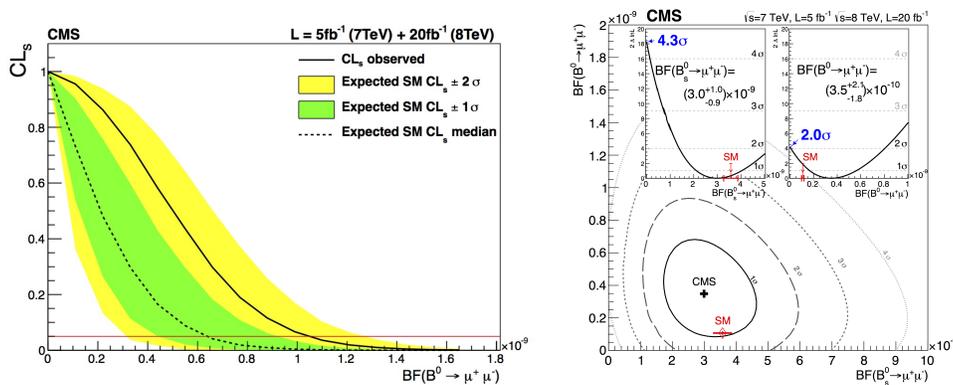


Figure 4: Left: Observed and expected 95% CL limit as a function of the $B^0 \rightarrow \mu^+ \mu^-$ assumed branching fraction. Right: Scan of the joint likelihood ratio for the two branching fractions. As insets, the likelihood ratio scans for each of the branching fraction when the other is taken as a nuisance parameter.

5 Combination with LHCb and prospects

The CMS measurement followed analogous results published by the LHCb collaboration [5]. While a proper combination, which requires building a joint likelihood, is being worked at, preliminary combined results have been obtained by the two collaborations for both branching fractions, by simply taking into account correlated and uncorrelated uncertainties:

$$\langle \mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) \rangle = (2.9 \pm 0.7) \times 10^{-9}$$

$$\langle \mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) \rangle = (3.6_{-1.4}^{+1.6}) \times 10^{-10}$$

The dominating correlated uncertainty is of course the value of f_u/f_s , which is the same for the two experiments. Although not final, the combined value for $\langle \mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) \rangle$ has a significance above 5.0σ .

Both values are compatible with the SM expectation, which can help to reduce considerably the allowed parameter space in many New Physics scenarios.

The future prospects for the measurements have been investigated by CMS [8], also in view of the foreseen (tracking) detector upgrade. The future LHC runs will be characterized by higher centre-of-mass energy, bunch-crossing frequency, and luminosity. This will lead to an increase of the signal production, partly due to the increased cross-section, which will have to meet with several serious challenges, especially concerning trigger rates and high pile-up level (which might affect isolation and therefore potentially decrease the selection efficiency). Improvements in the analysis are being studied, particularly as far as fake muon rejection is concerned (crucial for reducing the peaking background).

By rescaling the present results to the future foreseen luminosities and centre-of-mass energies, assuming mostly unchanged selection efficiencies, the significance of the $B_s^0 \rightarrow \mu^+\mu^-$ branching fraction measurement is expected to double by the end of 2017. The significance of the $B^0 \rightarrow \mu^+\mu^-$ measurement is expected to reach 5σ by the end of the HL-LHC phase.

References

- [1] J. R. Ellis *et al.*, JHEP **05** 063 (2006)
J. Parry, Nucl. Phys. **B 760**, 38 (2007)
D. Straub, <http://arxiv.org/abs/1205.6094>
- [2] S. Chatrchyan *et al.* [CMS Collaboration], “The CMS Experiment at the LHC,” JINST **3**, S08004 (2008)
- [3] S. Chatrchyan *et al.* [CMS Collaboration], “Measurement of the $B_s^0 \rightarrow \mu^+\mu^-$ Branching Fraction and Search for $B^0 \rightarrow \mu^+\mu^-$ with the CMS Experiment,” Phys. Rev. Lett. **111**, 101804 (2013)
- [4] A. Hoecker *et al.*. “TMVA - toolkit for multivariate data analysis”, (2007), [arXiv:physics/0703039 [physics.data-an]]
- [5] R. Aaij *et al.* [LHCb Collaboration], “First evidence for the decay $B_s^0 \rightarrow \mu^+\mu^-$,” Phys. Rev. Lett. **110**, 021801 (2013)
- [6] R. Aaij *et al.* [LHCb Collaboration], “Measurement of the fragmentation function f_s/f_d and its dependence on B meson kinematics,” J. High Energy Phys. **04**, 001 (2013)
- [7] The CMS and LHCb Collaborations, CMS PAS BPH-13-007, LHCb-CONF-2013-012
- [8] The CMS Collaboration “CMS reach in $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ branching fractions for the new LHC runs,” CMS PAS FTR-13-022