Available online at www.sciencedirect.com
ScienceDirect
Nuclear Physics B Proceedings Supplement 00 (2012) 1–6
Nuclear Physics B Proceedings Supplement

Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays in CMS

Gemma Tinti for the CMS collaboration[1]

Department of Physics & Astronomy, University of Kansas, Lawrence, KS, USA

Abstract

The rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ are potentially sensitive to physics beyond the Standard Model as their branching fraction can be highly enhanced in new physics models. The CMS experiment at the LHC has already reported the studies of these rare decays in pp collisions at $\sqrt{s} = 7$ TeV, corresponding to an integrated luminosity of 5 fb^{-1}. In both decays, the number of observed events is consistent with the expectation from background plus Standard Model signal prediction. Results of the combination of the exclusion limits for two rare decays in different LHC experiments will also be presented.

Keywords: rare decays, $B_s^0 \to \mu^+\mu^-$, $B^0 \to \mu^+\mu^-$

1. Introduction

In the Standard Model (SM) of particle physics, the decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ are highly suppressed as flavor changing neutral currents are possible only through box or penguin diagrams. Furthermore, these decays are helicity suppressed by a factor m_μ^2/m_B^2 and are suppressed by the internal quark annihilation within the B meson. The SM expectations for the branching ratio ($\mathcal{B}$) of the two decays are [1] $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.2 \pm 0.2) \times 10^{-9}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.0 \pm 0.1) \times 10^{-10}$. In several extensions of the SM, the branching fractions may be different from those of the SM.

In this paper, we describe the searches for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ rare decays performed by the Compact Muon Solenoid (CMS) experiment at the Large Hadron Collider (LHC) accelerator. The event counting search is performed in the dimuon invariant mass regions around $m_{B_s^0}$ and m_{B^0} and has been published in [2]. The search has been performed blindly: the signal region ($5.2 < m_{\mu\mu} < 5.45$ GeV) in data was not analyzed until all selection criteria were established. Data were collected in 2011 from proton-proton collisions at center of mass energy $\sqrt{s} = 7$ TeV, and collecting an integrated luminosity of 5 fb^{-1}.

Section 2 describes the main features of the CMS detector that made the analysis possible. Details of the analysis are presented in Section 3. The results obtained by the CMS collaboration are presented in Section 4, while improved results given by the combination of the CMS results with other LHC experiments are given in Section 5.

2. The CMS detector

A detailed description of the multi-purpose CMS detector is available in [3]. The main subdetectors used in this analysis are the silicon tracker (composed of pixel and strip detectors inside the 3.8 T axial magnetic field) and the muon detectors (drift tubes, cathode strip chambers and resistive plate chambers embedded in the steel return yoke of the solenoid). The tracker detects particles within the pseudorapidity range of $|\eta| < 2.5$, where the pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$ and θ is the polar angle with respect to the counterclockwise proton beam direction. The silicon tracker provides an impact parameter resolution of ≈ 15 μm and

[1]Email address: gemma.tinti@psi.ch

momentum measurement with a transverse momentum (p_T) resolution of 1.5% for the charged particles in the p_T range relevant for this analysis. The muon chambers detect muons with $|\eta| < 2.4$. Muon candidates are reconstructed by combining tracks in the silicon tracker and in the muon detector. In order to ensure high purity muons, tight selections are applied on the number of track hits and segments in the muon stations and silicon tracker, on the χ^2 of the combined track and on the impact parameter with respect to the beam spot.

The dimuon candidates are selected by two levels of trigger: the first level only uses muon detector information, while the high-level trigger (HLT) uses additional input from the tracker. The HLT requirements had to adapt to the increasing LHC luminosity and became more stringent during the evolution of the run. The offline selection criteria apply tighter cuts than the HLT requirements. The muon reconstruction and trigger efficiencies in the analysis are evaluated from Monte Carlo (MC) simulation, but they are cross checked with a standard data-driven technique [4], which uses J/ψ decaying into a pair of muons. A "tag" muon, satisfying strict muon criteria, is paired with a "probe" track (a muon candidate): the single muon efficiency is determined by the number of probe tracks passing or failing the muon identification algorithm.

3. Analysis

3.1. Signal selection

An event counting experiment is performed in the dimuon mass region $4.9 < m_{\mu\mu} < 5.9$ GeV. The background to real $B^0_{s(d)}$ decaying into two muons is given by *(a)* semileptonic decays from two B mesons, *(b)* one semileptonic B decay and one misidentified hadron and *(c)* single B decays with two misidentified hadrons. The analysis is performed in two channels: the barrel (both muons are required to have $|\eta| < 1.4$), and the endcap (at least one muon with $|\eta| \geq 1.4$) and the channels are combined for the final results. The separation of the barrel and endcap channels is due to a different signal over background ratio and worse mass resolution in the endcap. MC simulation is used for the estimation of the background for the decays from rare B meson decays, while sideband data are used for the estimation of the combinatorial background. The final goal is a measurement of the decay branching ratio. To limit the effect of some uncertainties, the branching ratio measurement is performed relative to the "normalization" channel $B^\pm \to J/\psi K^\pm$ (with $J/\psi \to \mu^+\mu^-$). The branching ratio for $B^0_s \to \mu^+\mu^-$ can be expressed as a function of the normalization channel as:

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = \frac{N^{B^0_s}_{\rm obs}}{N^{B^+}_{\rm obs}} \times \frac{f_u}{f_s} \times \frac{\varepsilon^{B^+}_{\rm tot}}{\varepsilon^{B^0_s}_{\rm tot}} \times \mathcal{B}(B^\pm \to J/\psi K^\pm), \quad (1)$$

where $N^{B^0_s}_{\rm obs}$ is the background subtracted number of observed signal candidates for $B^0_s \to \mu^+\mu^-$ processes with dimuon invariant mass in the signal window, $N^{B^+}_{\rm obs}$ is the number of reconstructed $B^\pm \to J/\psi K^\pm$ decays, $\varepsilon^{B^0_s}_{\rm tot}$ is the total efficiency for B^0_s reconstruction, $\varepsilon^{B^+}_{\rm tot}$ is the total efficiency for B^+ reconstruction, f_s/f_u is the ratio of the of the B^+ and B^0_s production cross section. The value of $f_s/f_u = 0.267 \pm 0.021$ is taken from [5] and the value of $\mathcal{B}(B^\pm \to J/\psi K^\pm)$ is taken from [6]. The validation of the signal selection uses the "control" channel $B^0_s \to J/\psi\phi$, where $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$.

The two opposite charged muons are required to have an invariant mass within the overall mass window $4.9 < m_{\mu\mu} < 5.9$ GeV (the signal mass window for $B^0 \to \mu^+\mu^-$ is $5.2 < m_{\mu\mu} < 5.3$ GeV and for $B^0_s \to \mu^+\mu^-$ is $5.3 < m_{\mu\mu} < 5.45$ GeV). The two muons have to originate from a common vertex, which is the B candidate decay vertex. The cut selection presented in this analysis has been optimized to provide the best expected combined upper limit. In general, the cuts are tighter in the endcap (E) compared to the barrel (B) as the endcap channel has higher background:

- Cuts on the transverse momentum of the two muons are applied: $p_T^{\mu_1} > 4.5$ GeV (for both B and E) and $p_T^{\mu_2} > 4.0(4.2)$ GeV for the B (E). Note that the muons have been ordered in p_T, so that μ_1 is the leading muon in p_T.

- The two muons are combined to give a B candidate track: a kinematic cut on the transverse momentum of the B candidate is also applied: $p_T^B > 6.5(8.5)$ GeV in the B (E).

- The common vertex of the two muons is fit (secondary vertex) and requirements on the goodness of the fit are applied: χ^2/dof <2.2(1.8) in the B (E).

The primary vertex (PV) associated to the B candidate is chosen among all the reconstructed PVs in the event as the minimally separated PV from the B candidate along the z axis (where the z coordinate is along the beam direction). The PV is refit removing the tracks that already form the B candidate secondary vertex. Once the PV has been determined, more variables can be constructed:

- The momentum of the B candidate has to be aligned to the flight direction evaluated connecting the PV to the secondary vertex: this requires the 3D pointing angle α to be smaller than 0.05(0.03) rad in the B(E).

- The 3D impact parameter of the B candidate δ_{3D}, its uncertainty $\sigma(\delta_{3D})$, and its significance $\delta_{3D}/\sigma(\delta_{3D})$ are measured with respect to the PV and required to be: $\delta_{3D} < 0.008$ cm and $\delta_{3D}/\sigma(\delta_{3D}) < 2$ for both B and E.

- The secondary vertex has to be well separated from the PV, and this requires the 3D flight length significance to be $l_{3D}/\sigma(l_{3D}) > 13(15)$ in the B(E).

Real signal events can be selected also looking at the isolation of the primary B vertex. Isolation properties are studied in three variables:

- The isolation variable I is evaluated as:

$$I = \frac{p_T(B)}{p_T(B) + \sum_{\mathrm{trk},\Delta R<0.7,p_T>0.9} p_T} \qquad (2)$$

  where the sum over the track momenta is performed only using tracks with a minimum p_T of 0.9 GeV and for tracks contained in a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.7$ ($\Delta\eta$ and $\Delta\phi$ are the differences in pseudorapidity and azimuthal angle between a charged track and the B candidate momentum). To make the selection robust versus the beam intensity (the "pile-up" problem will be discussed in Section 3.4), the sum is performed only on tracks coming either from the same PV or not associated to any PV (but having a distance of closest approach < 0.05 cm). For good isolation, I is required to be $I > 0.80$ for both B and E.

- For good isolation, the number of tracks with $p_T > 0.5$ GeV and distance of closest approach < 0.03 cm with respect to the B candidate vertex has to be small: $N_{\mathrm{trk}}^{\mathrm{close}} < 2$ for both B and E.

- We require a minimum distance of closest approach d_{ca}^0 between tracks and the B candidate vertex to be $d_{\mathrm{ca}}^0 > 0.015$ cm for both B and E. This requirement is applied to all tracks in the event not associated to a different PV than the one in consideration.

The distributions for a subset of the variables used in the selection are shown in Figure 1 for both signal MC and sideband data (background). Each variable shown has already the cuts on all the other variables applied.

The signal selection has been validated on the normalization $B^\pm \to J/\psi K^\pm$ and control $B_s^0 \to J/\psi\phi$ sample. The reconstruction of $B^\pm \to J/\psi K^\pm$ and $B_s^0 \to J/\psi\phi$ events is very similar to the $B_s^0 \to \mu^+\mu^-$ signal. Only minimal changes to account for the slightly different decay topologies are necessary: the dimuon invariant mass is required to be $3.0 < m_{\mu\mu} < 3.2$ GeV, $p_T(\mu\mu) > 7$ GeV due to trigger conditions, $p_T(K) > 0.5$ GeV and all tracks are used in vertexing. For the $B_s^0 \to J/\psi\phi$ candidates, we also require the invariant mass of the di-kaon system to be $0.995 < m_{K^+K^-} < 1.045$ GeV and $\Delta R(K^+, K^-) < 0.25$. The sideband subtracted data have been compared to the MC simulation for the signal in the $B^\pm \to J/\psi K^\pm$ and $B_s^0 \to J/\psi\phi$ samples for all the selection variables. The data/MC difference in the selection variables is evaluated as an uncertainty on the particle yield due to the selection and it is propagated to the final result. The total difference in data and MC has been quantified by summing in quadrature the percentage difference in data and MC in each variable: this leads to 4% uncertainty for the normalization sample and to 3% for the control sample (this uncertainty on the observed particle yield will be used for the $B_s^0 \to \mu^+\mu^-$ signal uncertainty).

3.2. Measurement of $B^\pm \to J/\psi K^\pm$

The number of reconstructed B^+ mesons is estimated by applying the selection cuts explained in Section 3.1 and by performing a fit on the obtained invariant mass distribution. The signal in the fit is modeled using a double Gaussian, while the background is fit as a decaying exponential plus an error function with center 5.145 GeV. Figure 2 shows the dimuon mass distribution for the selection on the normalization sample and the relative fit. The results of the fit give $N_{\mathrm{obs}}^{B^+} = 82712 \pm 4146$ in the barrel and 23809 ± 1203 in the endcap. The total efficiency is $\varepsilon_{\mathrm{tot}}^{B^+} = 0.00110 \pm 0.00009$ in the barrel and 0.00032 ± 0.00004 in the endcap. The total systematic uncertainty on the normalization sample yield is 5%.

3.3. Background treatment for $B_s^0 \to \mu^+\mu^-$

There are different types of background events that pass the signal selection and are in the signal mass window. In addition to combinatorial background, the background is given by $B \to hh'$ rare decays (e.g. $B_s^0 \to K^+K^-$ and $B^0 \to K^+\pi^-$), where h, h' are charged hadrons misidentified as muons (note that in this case the full B mass is reconstructed into a "peaking" background, but with erroneous h, h' mass assumptions) or

Figure 1: Comparison between signal MC ($B_s^0 \to \mu^+\mu^-$) and background dimuon distributions (from sideband data) for some signal selection variables. At the top: leading muon p_T, subleading muon p_T, B candidate transverse momentum. In the middle: 3D pointing angle, 3D flight length significance and 3D impact parameter significance. At the bottom: the isolation variable I, the distance of closest approach and the number of close tracks. The MC distributions are normalized to the number of entries in the data sidebands. For each variable, the cuts on all other variables have been applied.

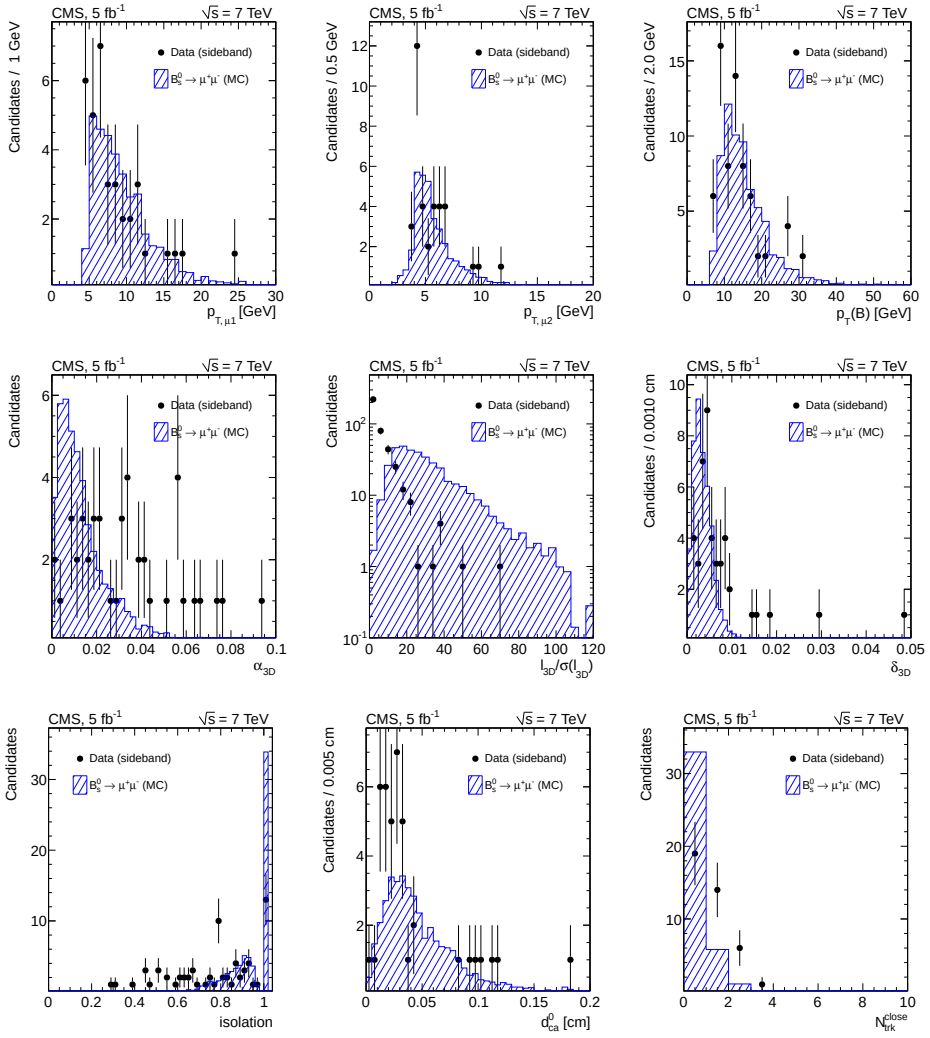
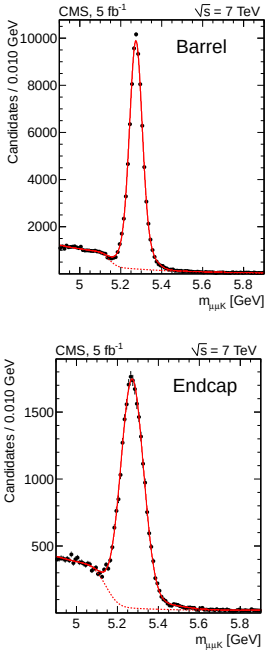


Table 1: The number of observed events in the signal window compared to the expectations. The number of events expected for the signal and peaking background is obtained from MC simulation normalized to the B^+ yield. The number of expected combinatorial background is obtained from sideband data (after subtraction of the rare background in the sidebands from MC).

Variable	$B^0 \to \mu^+\mu^-$ Barrel	$B_s^0 \to \mu^+\mu^-$ Barrel	$B^0 \to \mu^+\mu^-$ Endcap	$B_s^0 \to \mu^+\mu^-$ Endcap
$\varepsilon_{\text{tot}}^{B_s^0}$	0.0029 ± 0.0002	0.0029 ± 0.0002	0.0016 ± 0.0002	0.0016 ± 0.0002
$N_{\text{signal}}^{\text{exp}}$	0.24 ± 0.02	2.70 ± 0.41	0.10 ± 0.01	1.23 ± 0.18
$N_{\text{peak}}^{\text{exp}}$	0.33 ± 0.07	0.18 ±0.06	0.15 ± 0.03	0.08 ± 0.02
$N_{\text{comb}}^{\text{exp}}$	0.40 ± 0.34	0.59 ± 0.50	0.76 ± 0.35	1.14 ± 0.53
$N_{\text{tot}}^{\text{exp}}$	0.97 ± 0.35	3.47 ± 0.65	1.01 ± 0.35	2.45 ± 0.56
N^{obs}	2	2	0	4

Figure 2: $B^{\pm} \to J/\psi K^{\pm}$ invariant mass distributions in the barrel (top) and the endcap (bottom). The continuous line is the fit to the sum of signal and background.



rare semileptonic decays $B \to h\mu\nu$, where h is misidentified as a muon. In the case of rare semileptonic background, the full B meson mass is not reconstructed and the resulting dimuon invariant mass spectrum is a continuous distribution on the lower side of the mass spectrum. The expected number of rare events is evaluated from the MC simulation and normalized to the measured B^+ yield, taking into account the muon misidentification probability measured from data in $D^{*+} \to D^0\pi^+$ (where $D^0 \to K^-\pi^+$) and $\Lambda \to p\pi^-$. The misidentification rate is less than 0.10% (0.05%) for π and K (p). Table 1 shows the expected number of background events in the signal region: $N^{\rm exp}_{\rm peak}$ is the number of background events in the signal mass window from the peaking rare decays, $N^{\rm exp}_{\rm comb}$ is the number of expected combinatorial background events which is evaluated by interpolating according to a flat distribution the number of events from the sideband data, after subtracting the expected semileptonic background from MC. $N^{\rm exp}_{\rm signal}$ is the number of expected signal events assuming SM branching fractions and normalized to the B^+ yield.

3.4. Cross checks and systematic uncertainties

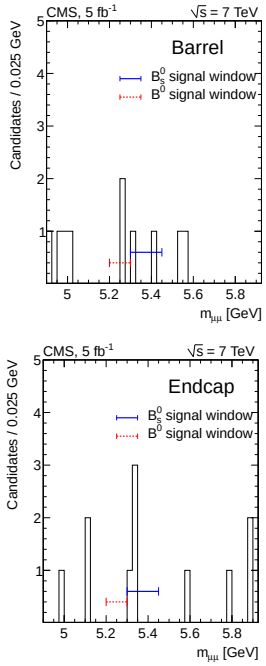
In addition to the systematic uncertainties already mentioned, other uncertainties are considered in the analysis. The errors on the trigger and muon identification efficiency ratios are between 3% and 8% depending on the muon kinematics. The $b - \overline{b}$ production process affects the b quark transverse momentum distribution and the B candidate isolation: the MC simulation has been found to describe the distributions adequately. The mass scale and resolution are known to 3%. The uncertainty on the production ratio of f_u/f_s is taken as 8%. The uncertainties on the combinatorial and rare backgrounds are evaluated to be 4% and 20%, respectively.

Many cross checks where performed. Among those, we mention the check of the stability of the selection versus the beam intensity. On average, the number of reconstructed PVs due to multiple collisions in an event was approximately 8 in the 2011 runs (with a root-mean-square equal to approximately 5-6 cm in the z direction). The selection efficiency is very robust against pile-up (i.e. multiple collisions in the same event) and is approximately flat versus the number of PVs. Another important cross check is the evaluation of the branching fraction for the control decay $B^0_s \to J/\psi\phi$, using a similar method to Equation (1). The computed branching ratio is in agreement with the world average [6] and the agreement of the results in the barrel and endcap channels within the statistical uncertainties prove the validity of the value of f_s/f_u from [5] also in the barrel.

4. Results

Figure 3 shows the measured dimuon invariant-mass distributions after unblinding. In the sidebands, the observed number of events is equal to six (seven) for the barrel (endcap) channel. Two (four) events are observed in the signal region in the barrel (endcap) for the $B^0_s \to \mu^+\mu^-$ decay channel, while two (zero) events are observed in the barrel (endcap) channel for the $B^0 \to \mu^+\mu^-$. The numbers are also reported in the last row of Table 1: the observation $N^{\rm obs}$ is consistent with the SM expectation for signal plus background. Upper limits on the $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ branching fractions are determined using the CL$_s$ method [7] by combining the endcap and barrel channels and taking into account the statistical and systematic uncertainties. The combined upper limits are: $\mathcal{B}(B^0_s \to \mu^+\mu^-) < 7.7 \times 10^{-9}$ at 95% CL and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.8 \times 10^{-9}$ at 95% CL. The SM median expected upper limits are 8.4×10^{-9} for $B^0_s \to \mu^+\mu^-$ and 1.6×10^{-9} for $B^0 \to \mu^+\mu^-$ at 95% CL. The upper limit results are shown in Figure 4. The background-only p value is 0.11 (0.24) for $B^0_s \to \mu^+\mu^-$ ($B^0 \to \mu^+\mu^-$), corresponding to 1.2 (0.7) standard deviations.

Figure 3: Final dimuon invariant mass spectrum for the barrel (top) and endcap (bottom). The spectra show the observed number of events in the sidebands and in the signal windows after unblinding.

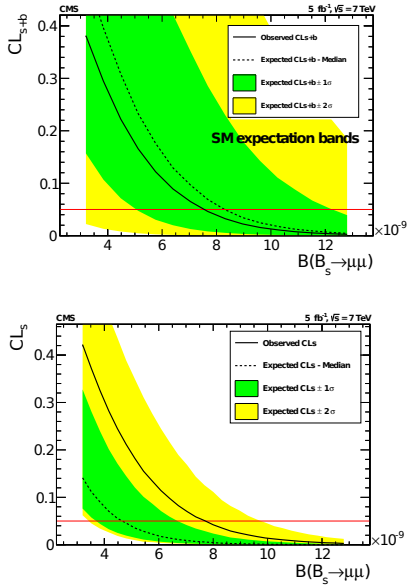


5. Combination with other experiments and summary

The search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ has been performed in pp collisions at $\sqrt{s} = 7$ TeV by the CMS experiment. In the performed analysis, the observed number of events is consistent with SM signal plus background. The combined upper limits are: $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 7.7 \times 10^{-9}$ at 95% CL and $\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.8 \times 10^{-9}$ at 95% CL.

Soon after the CMS results in [2] were published, two other LHC experiments released results on the same decay studies [8, 9, 10]). An official LHC combination has been released [11], with observed upper limit results equal to $4.2{\times}10^{-9}$ at 95% CL for $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ and $8.1{\times}10^{-10}$ at 95% CL for $\mathcal{B}(B^0 \to \mu^+\mu^-)$. The combination results are mainly driven by the results in [8]. As seen from the combination results, the upper limit measurements are getting closer and closer to the SM branching fraction expectation. The data being collected at $\sqrt{s} = 8$ TeV during the 2012 running from both the CMS and LHCb experiment have potential to allow a measurement of the SM branching fraction with a significance of three standard deviations for the $B_s^0 \to \mu^+\mu^-$ decay for each of the experiments.

Figure 4: CL_{s+b} and CL_s versus the $B_s^0 \to \mu^+\mu^-$ branching fraction assuming SM signal expectations plus background (top) or background only hypothesis (bottom).



6. Acknowledgments

Gemma Tinti is supported by the PIRE grant OISE-0730173 of the US-NSF.

References

[1] A. J. Buras, Minimal flavour violation and beyond: Towards a flavour code for short distance dynamics, Acta Phys.Polon. B41 (2010) 2487–2561.
[2] S. Chatrchyan, et al., Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays, JHEP 1204 (2012) 033.
[3] S. Chatrchyan, et al., The CMS experiment at the CERN LHC, JINST 3 (2008) S08004.
[4] Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV, CMS-PAS-MUO-10-002.
[5] R. Aaij, et al., Measurement of b hadron production fractions in 7 TeV pp collisions, Phys.Rev. D85 (2012) 032008.
[6] K. Nakamura, et al., Review of particle physics, 2010-2011. review of particle properties, J. Phys. G 37 (7A) (2010) 075021.
[7] T. Junk, Confidence level computation for combining searches with small statistics, Nucl.Instrum.Meth. A434 (1999) 435–443.
[8] R. Aaij, et al., Strong constraints on the rare decays $B_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$, Phys.Rev.Lett. 108 (2012) 231801.
[9] R. Aaij, et al., Search for the rare decays $B_s \to \mu^+\mu^-$ and $B_d \to \mu^+\mu^-$, Phys.Lett. B699 (2011) 330–340.
[10] G. Aad, et al., Search for the decay $B_s^0 \to \mu^+\mu^-$ with the ATLAS detector, Phys.Lett. B713 (2012) 387–407.
[11] Search for the rare decays $B_{(s)}^0 \to \mu^+\mu^-$ at the LHC with the ATLAS, CMS and LHCb experiments, CMS-PAS-BPH-12-009.