Search and measurement of the $B \to \mu\mu$ rare processes with LHC Run I data

Nuno T. Leonardo

Laboratório de Investigação e Física Experimental de Partículas, Portugal, and Purdue University, Department of Physics, USA
On behalf of the CMS Collaboration

Abstract

The decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ are highly suppressed in the standard model constituting sensitive probes of new physics. We present the results of a search for these rare decays in proton-proton collisions using the full data sample collected by the CMS experiment during LHC Run I. Through a fit to the dimuon invariant-mass spectrum, an excess of events with respect to the background is observed, compatible with the $B_s^0(B^0)$ signal with a significance of 4.3 (2.0) standard deviations (σ). The measured branching fractions are $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.0^{+1.0}_{-0.9}) \times 10^{-9}$ and $\mathcal{B}(B^0 \to \mu^+\mu^-) = (3.5^{+2.1}_{-1.8}) \times 10^{-10}$. The combination of the CMS and LHCb results, using their full Run I data samples, yield an event excess compatible with the B_s^0 (B^0) signal with a significance in excess of $6\,\sigma$ ($3\,\sigma$). The ongoing and projected accelerator and detector upgrades will allow to establish and to carry out precision measurements of both rare decays, and explore novel observables with further sensitivity to new physics effects.

Keywords: LHC, CMS, rare decays, heavy flavor, B physics, FCNC, searches for physics beyond the standard model

Contents

1. Introduction and motivation

The rare process $B_s^0 \to \mu^+\mu^-$ is one of the milestones of the LHC physics program. The expected probability for it to occur is very small and is precisely predicted in the standard model (SM). Any observed deviation has therefore the potential to significantly rule out the SM. Indeed, numerous scenarios beyond the SM can induce sizable modifications. The observation and measurement of the decay can be used to exclude and constrain classes of new physics scenarios. The process $B^0 \to \mu^+\mu^-$ is predicted to be even rarer and more challenging experimentally. Sensitivity to the B_s^0 decay has been attained for the first time with Run I of CERN's LHC while sensitivity to the B^0 decay will be achieved in the coming LHC runs.

In this article we present the results of a simultaneous search for the $B_{(s)}^0 \to \mu^+\mu^-$ decays at CMS. The outcome of the Run I LHC combination, involving the analysis of the CMS and LHCb data, is also reported, along with prospects for the future higher luminosity and energy LHC runs.

1.1. Theory expectations

The B_s^0 meson plays a key role in testing the quark-flavor sector of the SM. The exploration of its weak decays, after pioneering work at the Tevatron, has shifted to the LHC. Among the properties which render the B_s^0 system particularly interesting and sensitive to new phenomena are its extremely high-rate flavor oscillations between particle and antiparticle [1, 2] and its extremely low-rate decays to dileptons [3].

The exclusive decays of the neutral B mesons into muon pairs are highly suppressed – in the standard model. These processes involve effective flavor-changing neutral current (FCNC) $b \rightarrow c(d)$ transitions that can only proceed through higher-order diagrams in the SM, through box and penguin topologies. They are suppressed further by the required helicities of the final-state leptons in the decay of the spin-zero B mesons. The predicted SM values corresponding to the decay-time integrated branching fractions are $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)_{SM} = (3.57 \pm 0.30) \times 10^{-9}$ and $\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)_{SM} = (1.07 \pm 0.10) \times 10^{-10}$ [3, 4], where the uncertainties are fully dominated by nonpertubative QCD effects determined through lattice studies.

Several scenarios of physics beyond the standard model (BSM) predict significant modifications of the decay rates. Together with the precise SM expecations, this renders the study of these processes an excellent probe of BSM physics. This is the case, in particular, of theories with extended Higgs sectors. For example, in the context of the minimal supersymmetric SM, an enhancement of the branching fractions proportional to $\tan^6 \beta$ is predicted [5, 6, 7], where $\tan \beta$ is the ratio of the vacuum expectation values of the two Higgs fields. For large values of $\tan \beta$, this search belongs to the most sensitive probes for BSM physics which can be performed at collider experiments.

1.2. Pre-LHC searches and experimental status

The search for these decays has been pursued for over three decades. Significant progress in sensitivity has been made, with exclusion limits on the branching fractions improving by five orders of magnitude, as illustrated in Fig. 1. The ARGUS [8], UA1 [9], CLEO [10], Belle [11], BaBar [12], CDF [13], D0 [14], ATLAS [15], CMS [16], and LHCb [17] experiments have all published limits on these decays.

Prior to the start of the LHC, the most stringent limits were obtained by the Tevatron experiments. In the last couple of years, Tevatron's CDF [13] and LHC's LHCb [18] have both reported double-sided intervals for $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$. Notwithstanding, sufficient statistical sensitivity for the B_s^0 decay needed to wait for the
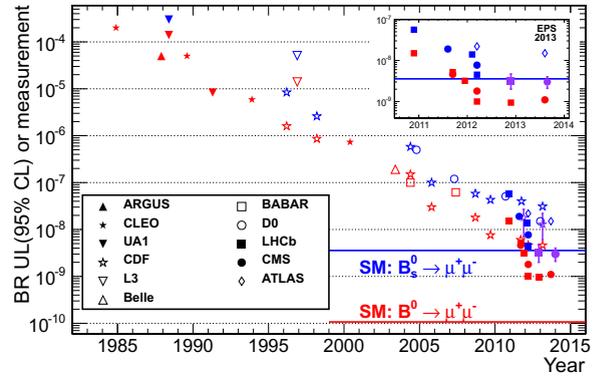
Figure 1: Summary of searches over time, illustrating the steady increase in sensitivity achieved.

full LHC Run I dataset, as is reported herein, while for B^0 it will be attained only with the coming LHC runs.

2. Analysis strategy and results

The signal is searched for in the dimuon invariant-mass spectrum in a window around the B-mesons mass. Owing to their long lifetime, B mesons tend to travel a measurable distance inside the detector before decaying. As such, the experimental signature consists of two isolated muons fro the exclusive B decay, originating from a common displaced vertex, with the dimuon momentum aligned with the flight direction between the primary and B-decay vertices.

2.1. Event selection

A dedicated trigger algorithm has been deployed to collect the data samples. It is based on the presence of two muons, with transverse momentum $p_{T,\mu}$=3–4 GeV, and invariant mass $4.8 < m_{\mu\mu} < 6.0$ GeV, forming a displaced vertex.

Vertex and isolation properties are further used offline to discriminate against combinatorial backgrounds. These arise from combinations of muons from uncorrelated sources, most importantly from two separate semileptonic $b \rightarrow c\mu\nu$ decays or from one such decay in combination with a misidentified hadron. Figure 2 illustrates the discriminating power of some of the variables used in the selection.

Instead of applying single thresholds to each of the selection variables, they are used to build boosted decision tree (BDT) multivariate discriminators. The BDTs are trained using signal events drawn from $B_s^0 \rightarrow \mu^+\mu^-$ Monte Carlo (MC) simulation and background events
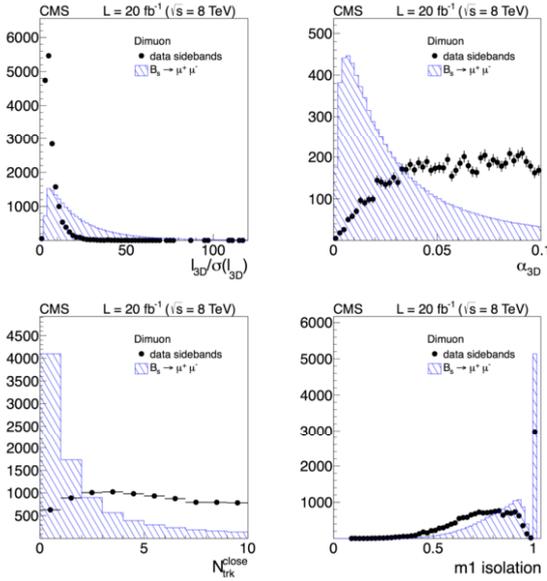
Figure 2: Subset of selection variables: (a) fligth-distance significance; (b) pointing angle between dimuon momentum and flight direction; (c) number of tracks close to the B vertex; (d) single-muon isolation. Signal distributions are from $B_s^0 \rightarrow \mu^+\mu^-$ MC simulation and background from mass sidebands in data.

from mass sidebands in data. The signal region, $5.20 < m_{\mu\mu} < 5.45$ GeV, was kept blind until all selection criteria were established. A comparison of the selection variable distributions in data and simulation reveal good agreement for the reference $B^+ \rightarrow J/\psi K^+$ and $B_s^0 \rightarrow J/\psi\phi$ samples. No pileup dependence was detected for the individual selected variables nor the combined BDT output discriminant b.

The combinatorial background described above has a monotonic depedence on the dimuon invariant mass, which allows it to be extrapolated to the signal region from the mass sidebands. There are, however, physics background sources for which this is not the case. Rare single B decays contribute a non-peaking component (e.g. $\Lambda_b \rightarrow p\mu\nu$) and a peaking component from two-body hadronic decays (e.g. $B_s^0 \rightarrow K^+K^-$), in the signal search window. These are estimated from MC simulation of the relevant processes, as indicated in Fig. 3.

These physics backgrounds arise due to the misidentification of charged hadrons as muons. A dedicated muon selection algorithm is employed to minimize the hadron-to-muon misidentification probability. It is based on a BDT that explores kinematic as well as silicon-tracker and muon-chamber combined-track information, which aims at descriminating genuine muons from those arising from hadron decay-in-flight or detec-
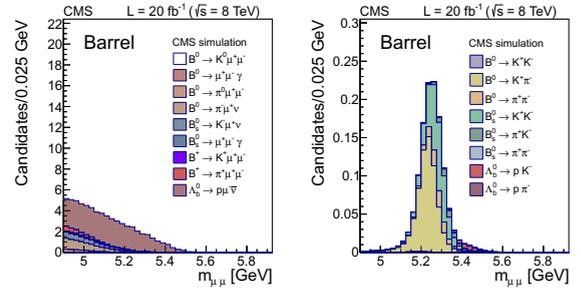
Figure 3: Processes contributing to the (left) non-peaking and (right) peaking physics backgrounds, following from hadron misidentification. Estimated from MC simulation.
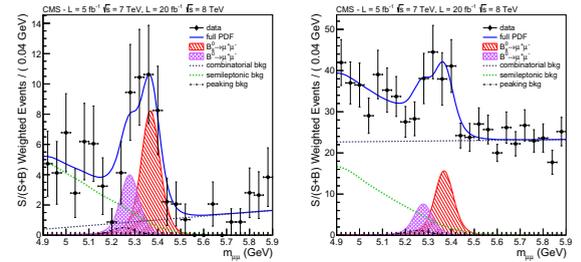
Figure 4: Illustration of the simultaneous fit to the mass spectrum, weighted over categories, for the (left) 1D-BDT and (right) categorized-BDT methods.

tor punchthrough. A misidentification rate at the permill level, $(0.4 - 2.2) \times 10^{-3}$, is achieved.

2.2. Likelihood fit and significance

The analysis is optimized separately in four separate *channels*, based on the $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV datasets and according to whether both muons lie within the region $|\eta_\mu| < 1.4$ (barrel) or not (endcap). Two methodologies are employed to determine the final results. (1) In the 1D-BDT method, a single threshold of the selection-BDT output b is optimized, per channel, for best significance. In the categorized-BDT method, the discriminant b is used to define 12 event categories, with different signal-to-background ratios but similar expected signal yield. The results are extracted via a simultaneous unbinned maximum likelihood fit to the resulting mass distributions. The fit projections are illustrated in Fig. 4 for both approaches.

An excess of events in the signal region is observed above the background predictions. The excess is found to be compatible with the $B_s^0 \rightarrow \mu^+\mu^-$ signal with an observed (expected median) significance of 4.3(4.8) standard deviations (σ) in the categorized-BDT approach, and $4.8\,\sigma(4.7\,\sigma)$ in the 1D-BDT approach. The excess in the B^0 region has a significance of $2.0\,\sigma$. The $B_s^0(B^0)$
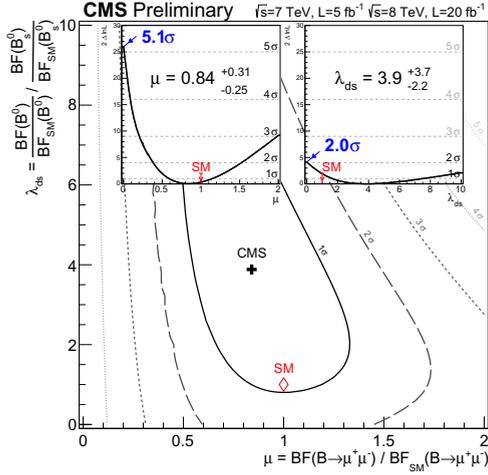
Figure 5: Scan of the ratio of the joint likelihood for the parameters μ_s and λ_{ds} defined in the text, for the categorized-BDT approach. The insets show the likelihood ratio scan for each of the parameters when the other is profiled together with other nuisance parameters; the significance at which the hypothesis wherein the parameter is zero is rejected is also shown.

significance is determined by evaluating the ratio of likelihood values for the hypothesis of no $B_s^0(B^0)$ signal and when allowing the $B_s^0(B^0)$ to float, while treating the $B^0(B_s^0)$ as a nuisance parameter that is allowed to float in the fit.

2.3. Branching fractions

Under the hypothesis that the observed excess in the data corresponds to a genuine signal, the decay branching fractions may be extracted fom the fit to the data. The fit likelihood is built having the branching fractions as the parameters of interest. For minimizing uncertainties associated to the B production cross section and reconstruction efficiency, and for constraining the normalization of the physics backgrounds, a sample of $B^+ \rightarrow J/\psi K^+$ (with $J/\psi \rightarrow \mu^+\mu^-$) decays, obtained with identical dimuon selection as for the signal, is employed. The branching fractions are measured using

$$\mathcal{B}(B_{(s)}^0 \rightarrow \mu^+\mu^-) = \frac{N^{B_{(s)}^0}}{N^{B^+}} \frac{f_u}{f_s} \frac{\varepsilon^{B^+}}{\varepsilon^{B_{(s)}^0}} \mathcal{B}(B^+), \qquad (1)$$

where N denotes the extracted signal yields, ε the total trigger and reconstruction efficiencies, $\mathcal{B}(B^+)= (6.0 \pm 0.2) \times 10^{-5}$ [19] the branching fraction for $B^+ \rightarrow \mu^+\mu^- K^+$, and $f_u/f_s = 0.256 \pm 0.020$ [20] the ratio of the B^+ and B_s^0 fragmentation fractions. The branching fractions, determined from the categorized-BDT fit, are:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-) = (3.0^{+1.0}_{-0.9}) \times 10^{-9},$$
$$\mathcal{B}(B^0 \rightarrow \mu^+\mu^-) = (3.5^{+2.1}_{-1.8}) \times 10^{-10}. \qquad (2)$$
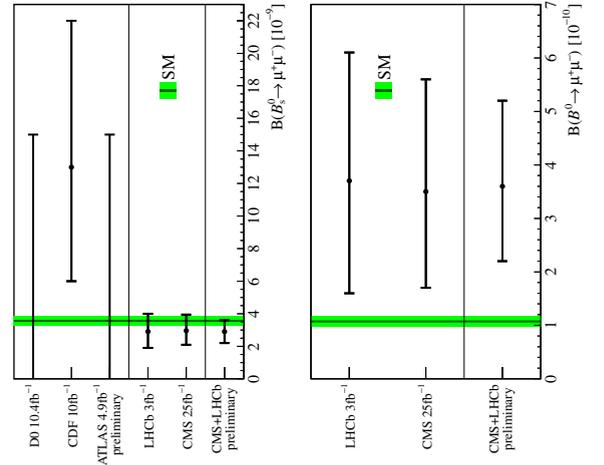
Figure 6: Illustration of the Tevatron and LHC results for (left) $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$ and (right) $\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)$, and combination [21] of CMS and LHCb measurements.

The uncertainty is statistics dominated. The main systematic uncertainties arises from f_u/f_s and the description of the physics backgrounds.

The results in Eq. 2 are compatible with the SM expectations. To verify and further quanitfy the level of agreement of the measuremens with the SM expectations, the likelihood is expressed in terms of the signal strength parameters μ_s and λ_{ds},

$$\mu_s \equiv \frac{\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)}{\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)_{SM}} = 0.84^{+0.31}_{-0.25},$$

$$\lambda_{ds} \equiv \frac{\mu_d}{\mu_s} = \frac{\frac{\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)}{\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)_{SM}}}{\frac{\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)}{\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)_{SM}}} = 3.9^{+3.7}_{-2.2}. \qquad (3)$$

The scan of the joint likelihood for μ_s and λ_{ds} is shown in Fig. 5. The compatibility with the SM point, $\mu_s = \lambda_{ds} = 1$, is further illustrated therein.

3. LHC combination

The CMS and LHCb collaborations at the LHC have both reported measurements of the branching fractions $\mathcal{B}(B_{(s)}^0 \rightarrow \mu^+\mu^-)$ based on the full Run I dataset. The ATLAS results based on the full Run I dataset have not been yet reported. A combination of the CMS and LHCb measurements is here presented.

3.1. Combined CMS and LHCb results

The results reported above in Sect. 2 are based on the analysis of pp-collision data corresponding to $5\,\text{fb}^{-1}$
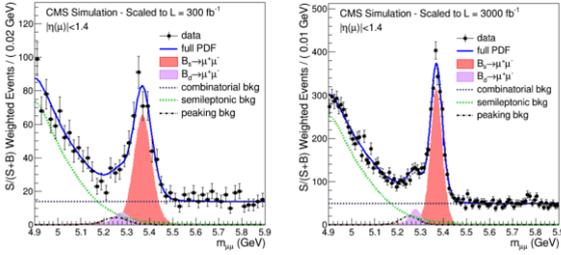
Figure 7: Simulated CMS projections indicating expected resolutions and yields for (left) 300 fb^{-1} and (right) 3000 fb^{-1}.

Table 1: Projected yields and sensitivity for $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ branching fraction measurements by CMS for future LHC runs.

$\int \mathcal{L}\, dt$ [fb^{-1}]	$B^0 \rightarrow \mu^+\mu^-$ stat. signif.	$\delta\mathcal{B}/\mathcal{B}$	$B_s^0 \rightarrow \mu^+\mu^-$ $\delta\mathcal{B}/\mathcal{B}$	$\delta\frac{\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)}{\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)}$
25	$2\,\sigma$	60%	33%	70%
300	$\approx 2.2\,\sigma$	48%	13%	50%
3000	$\approx 6.8\,\sigma$	18%	11%	21%

at $\sqrt{s}$ = 7 TeV collected in 2011 and 20 fb^{-1} at $\sqrt{s}$ = 8 TeV collected in 2012 by the CMS experiment.

LHCb has reported [22], simultaneously with CMS, the results of a search based on their full LHC Run I dataset, corresponding to 1 fb^{-1} at $\sqrt{s}$ = 7 TeV collected in 2011 and to 2 fb^{-1} at $\sqrt{s}$ = 8 TeV in 2012. LHCb equally reports an excess of events over background expectation with a significance of 4.0 σ for $B_s^0 \rightarrow \mu^+\mu^-$ and 2.0 σ for $B^0 \rightarrow \mu^+\mu^-$.

The CMS and LHCb measurements are compatible with each other. In order to increase the precision of the measured branching fractions, their combination is pursued. In the combination, known correlations and common external inputs need to be synchronized. Associated systematic uncertainties result for example from the imperfect knowledge of the ratio of fragmentation fractions f_u/f_s, that arises as the decay $B^+ \rightarrow J/\psi K^+$ is used as normalization channel.

In a preliminary combination of the results from both experiments [21], the asymmetric uncertainties are incorporated by using ensembles of simplified pseudo-experiments. The results are illustrated in Fig. 6. As for the individual analyses, and despite small upper fluctuations detected in both experiments for $\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)$, also the combined results are compatible with the SM expectations.

A more thorough approach explores the combination of the likelihood functions, through a simultaneous fit to the two datasets [23]. This allows in particular a more reliable evaluation of the statistical significance of the combined analyses. The excess of events over background observed is compatible with the $B_s^0 \rightarrow \mu^+\mu^- (B^0 \rightarrow \mu^+\mu^-)$ signals at a significance level in excess of 6 σ (3 σ).

4. Prospects

Following LHC's very sucessful Run I, both accelerator and detectors are undergoing and will further undergo in the coming years major upgrades. These will result in dramatic increases in sensitivity for rare decay searches such as those discussed in this article.

4.1. Further measurements

After having detected for the first time the $B_s^0 \rightarrow \mu^+\mu^-$ decay during LHC Run I, as is reported herein, the focus of the coming LHC runs will be on pursuing the search for the $B^0 \rightarrow \mu^+\mu^-$ decay and further establishing and carrying out precision measurements of both channels.

The analysis reported in Sect. 2 was optimized primarily for searching for the $B_s^0 \rightarrow \mu^+\mu^-$ decay. Further analysis refinements may be pursued, for each channel; for example at the level of candidate selection, fitting, and normalization channels.

It will be possible to explore additional observables, that are also theoretically clean and provide complementary sensitivity to BSM effects. A precise determination of the ratio of branching fractions allow to probe non-minimal flavor violation scenarios. B_s^0 mixing induced, effective-lifetime [4] and CP asymmetry observables are also particularly interesting.

4.2. Accelerator and detector upgrades

The LHC will increase the centre-of-mass collision energy, to 14 TeV, and the delivered luminosity in the coming years. From 2023, a major upgrade is projected (HL-LHC); the instantaneous luminosity for CMS will be leveled at a maximum of 5×10^{34} cm^{-2}s^{-1}, with the goal of delivering a total integrated luminosity of 3ab^{-1}. The (staged) increase in instantaneous luminosities, leading up to an average of 140 primary interactions per collision (pileup), demand in turn equally dramatic upgrades of the detectors.

Starting with Run II (2015-2017) of the LHC, the CMS detector will have improved muon trigger capabilities, providing sharper momentum resolution and ability to cope with increased pileup conditions. With Run III (2019-2021), it will have an improved pixel detector, with an additional layer, improving vertex resolution by 50%. For the HL-LHC phase, from 2023 onwards, the

detector will have been refurbished with an enhanced muon system and redesigned inner-tracker. The new tracker is expected to have hardware trigger level capabilities. The smaller silicon sensors pitch, combined with the reduced (x2) material budget, will improve the mass resolution by about a factor of 1.5 in the barrel region ($|\eta| < 1.4$). This will help separate the B^0 signal from the tail of the B_s^0 signal, which now becomes a background to the B^0 search and measurement.

4.3. Sensitivity projections

Projections are made by scaling the results obtained in Run I, and by parameterizing, under estimated assumptions [24], the expected effect of improvements in the detector and harsher collision environment. Same analysis and trigger performance as in Run I is assumed, along with a 35% loss in detection efficiency. The SM ratio of the two branching fractions is also assumed. Projected reconstruction yields and resolutions are illustrated in Fig. 7. Branching fraction sensitivities are summarized in Table 1. The scenarios considered for these projections include LHC Run III and HL-LHC, with correpsonding CMS upgrades [25].

5. Summary

The study of the decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ is among the highest priorities in heavy flavour physics, due to their exceptional sensitivity to sources of physics beyond the standard model. Based on the analysis of the full LHC Run I dataset collected by CMS, a first observation of the $B_s^0 \to \mu^+\mu^-$ with more the 4 standard deviations is reported. The combination of the CMS and LHCb results yields the first observation of the $B_s^0 \to \mu^+\mu^-$ channel with more than $6\,\sigma$ and first evidence of the $B^0 \to \mu^+\mu^-$ channel with $3\,\sigma$. The exploration of the $B_{(s)}^0 \to \mu^+\mu^-$ rare decays remains a highest priority of the heavy-flavor physics program into the coming LHC runs.

References

[1] A. Abulencia, et al., Observation of B_s^0–anti-B_s^0 Oscillations, Phys.Rev.Lett. 97 (2006) 242003. arXiv:hep-ex/0609040, doi:10.1103/PhysRevLett.97.242003.

[2] N. T. Leonardo, Analysis of B_s flavor oscillations at CDF, Ph.D. thesis, MIT, FERMILAB-THESIS-2006-18, ISBN:384337693X (2006).

[3] A. J. Buras, et al., On the Standard Model prediction for $B_s^0 \to \mu^+\mu^-$, Eur. Phys. J. C 72 (2012) 2172. arXiv:1208.0934, doi:10.1140/epjc/s10052-012-2172-1.

[4] K. De Bruyn, et al., Probing new physics via the $B_s^0 \to \mu^+\mu^-$ effective lifetime, Phys. Rev. Lett. 109 (2012) 041801. arXiv:1204.1737, doi:10.1103/PhysRevLett.109.041801.

[5] K. Babu, C. F. Kolda, Higgs mediated $B^0 \to \mu^+\mu^-$ in minimal supersymmetry, Phys.Rev.Lett. 84 (2000) 228–231. arXiv:hep-ph/9909476, doi:10.1103/PhysRevLett.84.228.

[6] C. Hamzaoui, M. Pospelov, M. Toharia, Higgs mediated FCNC in supersymmetric models with large tan Beta, Phys.Rev. D59 (1999) 095005. arXiv:hep-ph/9807350, doi:10.1103/PhysRevD.59.095005.

[7] C.-S. Huang, W. Liao, Q.-S. Yan, S.-H. Zhu, $B_s \to l^+l^-$ in a general 2 HDM and MSSM, Phys.Rev. D63 (2001) 114021. arXiv:hep-ph/0006250, doi:10.1103/PhysRevD.64.059902, 10.1103/PhysRevD.63.114021.

[8] H. Albrecht, et al., B meson decays into charmonium states, Phys. Lett. B 199 (1987) 451. doi:10.1016/0370-2693(87)90952-X.

[9] C. Albajar, et al., A search for rare B meson decays at the CERN SPPS collider, Phys. Lett. B 262 (1991) 163. doi:10.1016/0370-2693(91)90660-I.

[10] T. Bergfeld, et al., Search for decays of B^0 mesons into pairs of leptons: $B^0 \to e^+e^-$, $B^0 \to \mu^+\mu^-$ and $B^0 \to e^\pm\mu^\mp$, Phys. Rev. D 62 (2000) 091102. arXiv:hep-ex/0007042, doi:10.1103/PhysRevD.62.091102.

[11] M. C. Chang, et al., Search for $B^0 \to l^+l^-$ at Belle, Phys. Rev. D 68 (2003) 111101. arXiv:hep-ex/0309069, doi:10.1103/PhysRevD.68.111101.

[12] B. Aubert, et al., Search for decays of B_s^0 mesons into e^+e^-, $\mu^+\mu^-$, and $e^\pm\mu^\mp$ final states, Phys. Rev. D 77 (2008) 032007. arXiv:0712.1516, doi:10.1103/PhysRevD.77.032007.

[13] T. Aaltonen, et al., Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays with the full CDF Run II data set, Phys. Rev. D 87 (2013) 07203. arXiv:1301.7048, doi:10.1103/PhysRevD.87.072003.

[14] V. M. Abazov, et al., Search for the rare decay $B_s^0 \to \mu^+\mu^-$, Phys. Rev. D 87 (2013) 072006. arXiv:1301.4507, doi:10.1103/PhysRevD.87.072006.

[15] G. Aad, et al., Search for the decay $B_s^0 \to \mu^+\mu^-$ with the ATLAS detector, Phys. Lett. B 713 (2012) 387. arXiv:1204.0735, doi:10.1016/j.physletb.2012.06.013.

[16] S. Chatrchyan, et al., Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays, JHEP 04 (2012) 033. arXiv:1203.3976, doi:10.1007/JHEP04(2012)033.

[17] R. Aaij, et al., Search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$, Phys. Lett. B 699 (2011) 330. arXiv:1103.2465, doi:10.1016/j.physletb.2011.04.031.

[18] R. Aaij, et al., First evidence for the decay $B_s^0 \to \mu^+\mu^-$, Phys. Rev. Lett. 110 (2013) 021801. arXiv:1211.2674, doi:10.1103/PhysRevLett.110.021801.

[19] Particle Data Group, Review of particle physics, Phys. Rev. D 86 (2012) 010001. doi:10.1103/PhysRevD.86.010001.

[20] R. Aaij, et al., Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics, JHEP 04 (2013) 001. arXiv:1301.5286, doi:10.1007/JHEP04(2013)001.

[21] Combination of results on the rare decays $B_s^0 \to \mu^+\mu^-$ from the CMS and LHCb experiments, CMS Physics Analysis Summary CMS-PAS-BPH-13-007 (2014).

[22] R. Aaij, et al., Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and search for $B^0 \to \mu^+\mu^-$ decays at the LHCb experiment, Phys.Rev.Lett. 111 (2013) 101805. arXiv:1307.5024, doi:10.1103/PhysRevLett.111.101805.

[23] Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data, Nature (to appear).

[24] CMS reach in $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ branching fractions for the new LHC runs, CMS Physics Analysis Summary CMS-PAS-FTR-13-022 (2013).

[25] CMS Collaboration, CMS B Physics Results page, https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsBPH.