PROCEEDINGS OF SCIENCE

$B_{d/s} \to \mu^+\mu^-$ in ATLAS

Jaroslav Guenther[*†]
Innsbruck University
E-mail: jaroslav.guenther@cern.ch

The ATLAS Experiment has conducted a search for the rare decays of B_s and B_d into two muons. 25 fb^{-1} of integrated luminosity of proton-proton collisions collected during LHC Run 1 were studied and new results were obtained, as presented in this talk. An upper limit is set on the branching ratio of $\mathscr{B}(B_d \to \mu^+\mu^-) < 4.2 \times 10^{-10}$ at 95% confidence level. For B_s, ATLAS measurement yields the branching ratio $\mathscr{B}(B_s \to \mu^+\mu^-) = (0.9^{+1.1}_{-0.8}) \times 10^{-9}$. The result is consistent with the Standard Model expectation and other available measurements.

XIII International Conference on Heavy Quarks and Leptons
22- 27 May, 2016
Blacksburg, Virginia, USA

[*]Speaker.
[†]on behalf of the ATLAS Collaboration

PoS(HQL 2016)057

1. Introduction

Branching ratios of B_s and B_d meson decays into pairs of oppositely charged muons are accurately predicted by the Standard Model (SM) [1]: $\mathscr{B}(B_s \to \mu^+\mu^-) = (3.65 \pm 0.23) \times 10^{-9}$ and $\mathscr{B}(B_d \to \mu^+\mu^-) = (1.06 \pm 0.09) \times 10^{-10}$. These helicity suppressed flavour-changing neutral-current (FCNC) processes are forbidden to be mediated via tree level in the SM and proceed through loop Feynman diagrams. Direct probe of SM-unpredicted physics is possible due to the sensitivity to a presence of hypothetical (pseudo-)scalar particles in higher perturbative orders of SM extensions. Any observation of a significant enhancement or suppression of the SM-predicted branching ratios of these decays would be a clear sign of interference with a wide range of possible New Physics (NP) phenomena [11–19].

Run I of the LHC provided enough data to discover $B_s \to \mu^+\mu^-$ and observe evidences for $B_d \to \mu^+\mu^-$ decays by the CMS and LHCb collaborations. Their combined measurement yields: $\mathscr{B}(B_s \to \mu^+\mu^-) = (2.8^{+0.7}_{-0.6}) \times 10^{-9}$ and $\mathscr{B}(B_d \to \mu^+\mu^-) = (3.9^{+1.6}_{-1.4}) \times 10^{-10}$ [2].

The ATLAS Collaboration [10] has also conducted a search for the rare decays of B_s and B_d into two muons, as described further in this document. The ATLAS measurement [3] explored 25 fb^{-1} of data from proton-proton collisions at the LHC with a centre of mass energy of 7 TeV (4.7 fb^{-1}) and 8 TeV (20.3 fb^{-1}).

2. Analysis Overview

The ATLAS strategy was to perform a blind analysis. Hence, the signal region of $B_{d/s}$ invariant mass between 5166 and 5526 MeV was excluded from the data. This crucial part of the collected data was used only once the entire analysis procedure had been firmly established. Extensive use of multivariate analysis techniques [6] was necessary in order to discriminate against very large background of these rare decays. Two Boosted Decision Tree (BDT) classifiers were developed for suppression of very large continuum and peaking backgrounds.

The measurement of the $B_{d/s} \to \mu^+\mu^-$ branching ratios was performed relative to the reference signal decay $B^\pm \to J/\psi K^\pm$ which had sufficiently large number of observed events. Choosing this strategy substantially reduced systematic uncertainties of this measurement. Hence, the $\mathscr{B}(B_{d/s} \to \mu^+\mu^-)$ was measured following this prescription:

$$\mathscr{B}(B_{d/s} \to \mu^+\mu^-) = N_{\mu^+\mu^-} \times \left[\mathscr{B}(B^\pm \to J/\psi K^\pm) \times \mathscr{B}(J/\psi \to \mu^+\mu^-)\right] \times \frac{f_u}{f_{d/s}} \times \frac{1}{\mathscr{D}_{\text{norm}}}, \quad (2.1)$$

with

$$\mathscr{D}_{\text{norm}} = \sum_k N^k_{J/\psi K^\pm} \alpha_k \left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}\right)_k . \quad (2.2)$$

The $\frac{f_u}{f_{d/s}}$ factor compensates for the differences in hadronisation probabilities of b quarks to form a $B^\pm$ or a $B_{d/s}$ mesons. The denominator $\mathscr{D}_{\text{norm}}$ consists of a sum whose index k runs over four event selection categories. In the sum, the α_k parameter takes into account the different trigger prescale factors and integrated luminosities in the signal and normalisation channels. The next important factor is the $\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}$ term, which takes into account the differences in acceptance and efficiency

between the signal and the reference channel. $N_{\mu^+\mu^-}$ and $N^k_{J/\psi K^\pm}$ denote the yields of the reference and signal channels.

The input values for Equation 2.1 were acquired from the following sources. The input reference channel branching ratios were taken from the PDG [4]. The $\frac{f_s}{f_d}$ fraction of hadronisation probabilities (assuming $\frac{f_u}{f_d} = 1$) was measured by ATLAS [5] in the same fiducial kinematic region of the B meson as the one used for this study. The term $\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}$ encompassing both efficiency and acceptance was measured on simulation and calibrated on data. The data were used to measure the α_k terms as well as, eventually, the $N_{\mu^+\mu^-}$ and $N^k_{J/\psi K^\pm}$ yields which were extracted by extended unbinned likelihood fits to the invariant mass distributions.

In order to avoid efficiency losses, all events collected by di-muon triggers of the ATLAS detector were split in four categories T_i. For the 2012 data sample, three categories were defined by mutually exclusive requirements on transverse momentum (p_T) and pseudorapidity (η) of the two muon candidates (T_1, T_2, T_3). A separate category was defined for the 2011 data sample (T_4):

T_1: "higher threshold" trigger with $p_T > 6$ GeV for one muon and $p_T > 4$ GeV for the other one;

T_2: "barrel" trigger with $p_T > 4$ GeV for both muon candidates and at least one of them with $|\eta| < 1.05$ (and T_1 requirement not satisfied);

T_3: basic dimuon trigger with $p_T > 4$ GeV for both muon candidates (and T_1, T_2 requirements not satisfied);

T_4: basic dimuon trigger with $p_T > 4$ GeV for both muon candidates;

Muon candidates were required to have $p_T > 4$ GeV, $|\eta| < 2.5$ and tracks reconstructed in both the Inner Detector (ID) and the Muon Spectrometer (MS). The latter is one of many novelties in this analysis: the use of combined tracking algorithms in reconstruction improved the observed mass resolution in the forward detector regions. Kaons were required to have $p_T > 1$ GeV, $|\eta| < 2.5$ and tracks reconstructed in the ID system. B meson candidates were reconstructed imposing quality requirements on their secondary vertices (SV). Another significant improvement in this analysis is the new matching of the SV to the corresponding primary vertex (PV). The selected primary vertex is the one with minimal distance along the collision axis (z-axis) to a point S, where S is a point on the z-axis which is closest to the B candidate momentum line (in 3D). This assignment has proven to be more than 99% efficient in the simulation. Finally only B candidates in the fiducial kinematic region of $p_T > 8$ GeV, $|\eta| < 2.5$ were accepted for analysis.

3. Backgrounds

The background sources can be classified into three main categories.

1: *Combinatorial background* is the dominant background source. It is formed by two real muons produced independently in the b or $\bar{b}$ decay chain.

2: *Partially reconstructed decays* (PRDs) where only a part of the full event is reconstructed as a signal decay. In this case the muons can come from the following topologies: (a) same-side

(SS) combinatorial background from decay cascades ($b \to c\mu\nu \to s(d)\mu\mu\nu\nu$); (b) same-vertex (SV) background from B decays containing a muon pair (e.g. $B^0 \to K^{*0}\mu\mu$, $B \to J/\psi X \to \mu\mu\mu X'$); (c) B_c decays (e.g. $B_c \to J/\psi\mu\nu \to \mu\mu\mu\nu$); (d) semileptonic b-hadron decays where a final-state hadron is misidentified as a muon.

3: *Peaking background* dangerously overlaps with the main signal due to $B_{d/s} \to hh'$ decays, with both hadrons hh' (h stands for charged π or K) misidentified as muons.

The combinatorial background is several orders of magnitude bigger than the actual signal. To reduce this enormous background a BDT (c-BDT) was trained using 1.4 billion simulated events of b- or c-quark decays (anti-quarks included) containing muons in the final state. The simulated sample was validated to be in very good agreement with data candidates from the low-mass sideband of the B candidate invariant mass distribution. c-BDT used 15 input variables exploiting: B-meson kinematics, separation of primary and secondary vertices, collinearity between PV-SV direction with the reconstructed momentum of B meson candidate, properties of the muons and the rest of the event (such as the B-isolation[1]). Before applying the c-BDT classifier, several additional cuts were made to reduce the background by factor of 2.5 while losing only 5% of signal efficiency. The final c-BDT variable distribution for signal and background events is shown in Figure 1. With a final cut on the c-BDT output at 0.24, the background was reduced by a factor of thousand while keeping the remaining signal efficiency at the level of 54% (as measured in simulation).

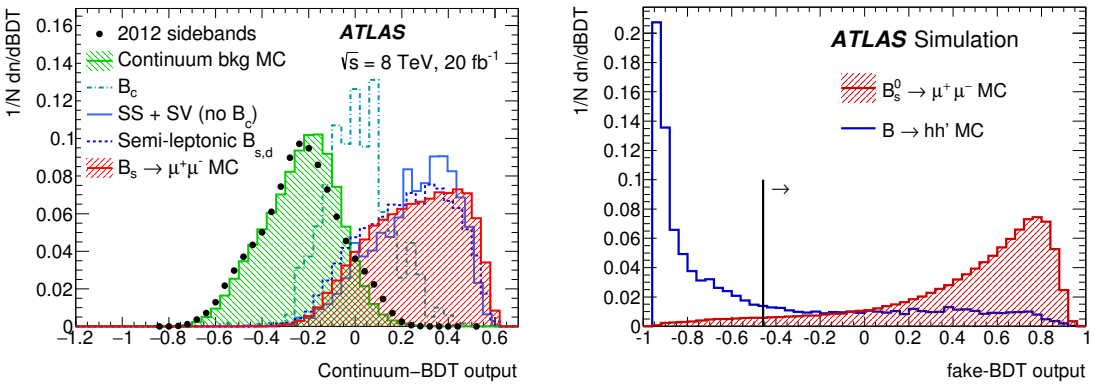

Figure 1: *Left:* Continuum-BDT distribution for the signal and background events: signal $B_{d/s}$, partially reconstructed B events (SS+SV), B_c decays and continuum. The solid histograms are obtained from simulation, while the points represent data collected in the sidebands. All distributions are normalised to unity. The distributions are shown after the preliminary selection, and before applying any reweighting to the variables used in the classifier. *Right:* MC distributions for the fake-BDT variable for real muons from signal $B_s \to \mu^+\mu^-$ sample and fake muons from peaking background $B^0_{(s)} \to hh'$ sample. The red filled histogram corresponds to real muons, while the blue histogram corresponds to fake muons. Figures from Ref. [3].

Another critical aspect of this analysis which could potentially destroy sensitivity to NP are hadrons (protons, pions, kaons) mimicking our signal as mis-identified muons. To reduce the

[1]The isolation variable is defined as the ratio of $|\overrightarrow{p_T}^B|$ to the sum of $|\overrightarrow{p_T}^B|$ and of the transverse momenta of all additional tracks contained within a cone of size $\Delta R < 0.7$ around the B direction. Only tracks with $p_T > 0.5$ GeV and matched to the same PV as the B candidate are included in the sum.

background originating from such mis-identification (*peaking background, type (d) of PRDs*), a second BDT (f-BDT) event classifier was developed using simulated events. Eight input variables of this discrimination algorithm describe muon candidate track quality as well as kinematic and calorimetric properties. The performance of f-BDT[2] was validated on several data samples ($B^\pm \to J/\psi K^\pm$, $\Phi \to K^+K^-$, signal events passing reverse f-BDT cut) and the f-BDT variable distribution for real and mis-identified muons can be seen in Figure 1. With the final signal efficiency set to 95%, f-BDT achieved a reduction 7 times better than in the previous analysis [7]. Only an estimated number of 1.0 ± 0.4 events of this background in the final B candidate selection entered the main signal fit.

4. Reference Channel Yield and Relative Efficiency Extraction

Once the background was reduced, one could concentrate on extracting the reference channel yield and computing the relative term $\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}$ of Equation 2.1.

The reference channel yield was measured by four simultaneous unbinned extended maximum likelihood fits to the invariant mass distribution of the $B^\pm$ candidates. Each of these fits corresponds to one of the event selection categories (T_1, T_2, T_3, T_4). All fits have a very similar baseline structure and consist of 4 components: partially reconstructed decays (PRDs), combinatorial background, $B^\pm \to J/\psi(\to \mu^+\mu^-)K^\pm$ signal and $B^\pm \to J/\psi(\to \mu^+\mu^-)\pi^\pm$ exclusive background. The shapes of PRDs, $B^\pm \to J/\psi K^\pm$ and $B^\pm \to J/\psi \pi^\pm$ components were constrained by including data from simulations in these simultaneous fits. Several parameters were free to vary independently such as all normalisations and the shape of the combinatorial background. Additional parameters, mass scale and mass resolution, were introduced in order to accommodate corresponding data–MC differences. The total contribution to the systematic uncertainty on the normalisation term 2.2 of Equation 2.1 was $\pm$ 0.8%.

Inclusion of a $B^\pm \to J/\psi(\to \mu^+\mu^-)\pi^\pm$ component was necessary for the fit stability and its yield was extracted simultaneously with the reference channel yield. Hence, a natural parallel measurement of the ratio of the corresponding branching fractions was performed, resulting in the following value for the error-weighted combination from the four categories to be:

$$\frac{\mathscr{B}(B^\pm \to J/\psi\pi^\pm)}{\mathscr{B}(B^\pm \to J/\psi K^\pm)} = 0.035 \pm 0.003\ (\text{stat.}) \pm 0.012\ (\text{syst.}) \tag{4.1}$$

This result is in agreement with the PDG average [4].

The $\left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}\right)_k$ ratios between signal and reference channel were measured within the fiducial region of the B meson kinematics of $p_T > 8$ GeV and $|\eta| < 2.5$. This measurement used simulated events and was performed separately for each channel in each of the four measurement categories. The simulated samples were reweighted to match the following distributions in data (control samples $B_s \to J/\psi\phi$, $B^\pm \to J/\psi K^\pm$): number of primary vertices (pile-up), $p_T(B)$, $|\eta(B)|$ and trigger efficiencies (as a function of $p_T(\mu)$, $|\eta(\mu)|$). It was taken into account that the $B_s \to \mu^+\mu^-$ decay proceeds dominantly through the heavy B_s mass eigenstate [8] and a correction to the mean lifetime in the simulated samples was applied. In addition, a correction was applied in the simulated

[2] f-BDT reduced the mis-identification probability to $0.09\%, 0.04\%, < 0.01\%$ for K, π, p

samples also for B isolation and any residual and other discrepancies between data and simulated samples in the spectra of the 15 input variables used in the c-BDT that could affect the $\left(\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}\right)_k$ ratios were accounted for as a systematic uncertainty. The total systematic uncertainty on the normalisation term 2.2 of Equation 2.1 was measured to be $\pm$ 5.9%.

5. $B_{d/s} \to \mu^+\mu^-$ Signal Yield Extraction

Both yields of $B_s \to \mu^+\mu^-$ (N_{B_s}) and $B_d \to \mu^+\mu^-$ (N_{B_d}) decays were extracted together from the data by an unbinned extended maximum likelihood fit to the di-muon invariant mass distribution. To get better control on the background shapes and in order to increase the discriminating power, the fit was performed simultaneously in 3 bins of the c-BDT classifier: 0.240-0.346, 0.346-0.446 and 0.446-1.000 (see Figure 1). The boundaries between the three c-BDT bins were set to have the same signal efficiency of 18% (measured in the simulation).

The baseline fit configuration has three background models: (1) a first order polynomial parametrises the combinatorial background, which smoothly crosses the invariant mass distribution, (2) an exponential function describes the background coming from the same-side and same vertex backgrounds described in Section 3 and (3) a signal-like model which is needed for the peaking background. This last model (3) has shape parameters set to constant values measured on simulated events and the normalisation was set to 1.0 ± 0.4 events which were constrained to be equally distributed in the 3 c-BDT bins of the fit (see Section 3). Both remaining background models have total yield parameters as well as relative fractions of these events in each c-BDT bin allowed to vary independently in the fit. The same freedom is given also to the slope of the combinatorial background model. The shape of the exponential model was constrained to be the same in each of the three bins. Two double Gaussian models describe the signals $B_s \to \mu^+\mu^-$ and $B_d \to \mu^+\mu^-$ in the fit and have shape parameters set to values obtained from fitting simulated signal samples. Two global normalisation factors N_{B_s} and N_{B_d} are introduced which can vary independently in the fit and serve to measure the final yields. Their relative contributions from each bin is set to the expected equal efficiency.

The expected fit sensitivity, assuming the SM predictions, was studied using pseudo-MC experiments and measured to be $s_{B_s\to\mu^+\mu^-} = 3.1\sigma$ and $s_{B_d\to\mu^+\mu^-} = 0.2\sigma$. The systematic uncertainties were evaluated also with the use of pseudo-MC experiments. The major contributors were coming from the relative signal efficiency of the three c-BDT bins (9%) and alternative signal and background models (6%). The evaluation of systematic uncertainties on the signal yields considered also the effects originating (first) from mass resolution and scale discrepancies between data and simulation and (second) from the constraint on the signal efficiencies to be the same across the three c-BDT bins. Eventually, the total systematic uncertainties on the fitted signal yields were: $\sigma_{syst.}(N_{B_d}) = 3$ and $\sigma_{syst.}(N_{B_s}) = \sqrt{4 \pm 0.06 \times N_{B_s}}$. These values were implemented as Gaussian constraints in the likelihood before the fit to data.

At this point, all inputs to Equation 2.1 were known except the signal yields. Assuming the validity of the SM and imposing $N_{B_{d/s}} > 0$, one would most probably expect to see a yield of $N_{B_s} = 41$ and $N_{B_d} = 5$ events in the unblinded signal region. The final fit to the di-muon invariant mass of the reconstructed B candidates selected from data can be seen in Figure 2. The number of fitted $B_{d/s} \to \mu^+\mu^-$ signal events was found to be: $N_{B_s} = 16 \pm 12$ and $N_{B_d} = -11 \pm 9$.

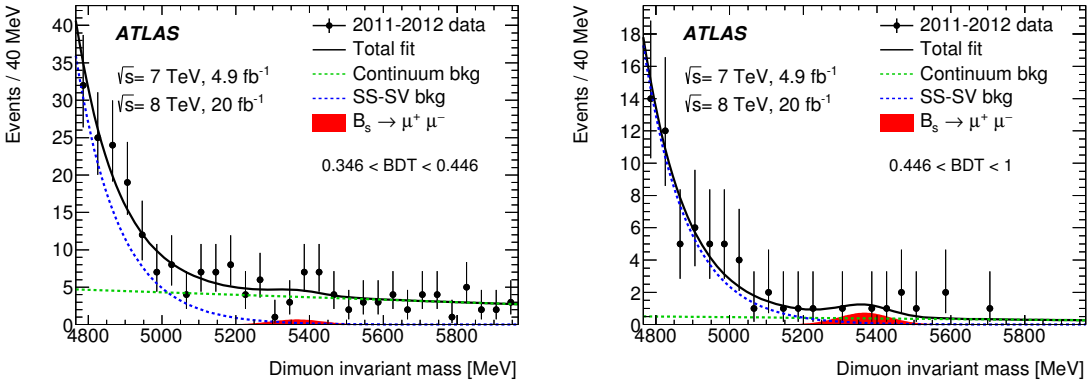

Figure 2: Di-muon invariant mass distributions in the unblinded data, for the second and the third c-BDT bin. Superimposed is the result of the maximum-likelihood fit, obtained imposing the boundary of non-negative signal contributions. The total fit is shown as a black continuous line, the filled area corresponds to the observed signal component, the blue dashed line to the SS+SV background, and the green dashed line to the combinatorial background. Figures from Ref. [3].

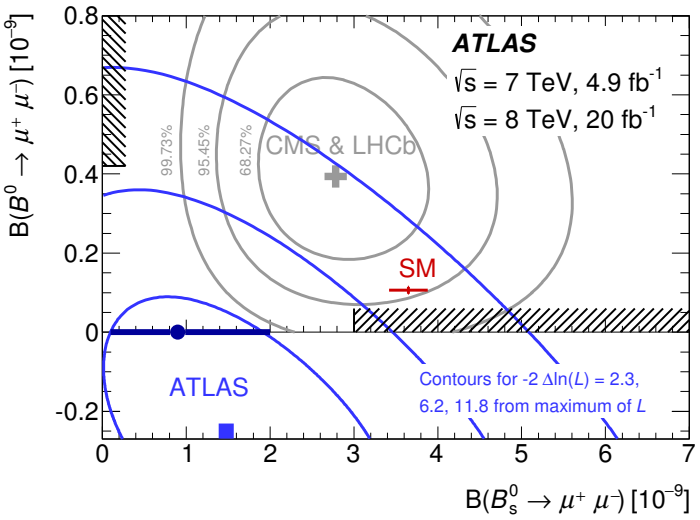

Figure 3: Contours in the plane $\mathscr{B}(B_s \to \mu^+\mu^-)$, $\mathscr{B}(B_d \to \mu^+\mu^-)$ for intervals of $-2\Delta ln(L)$ equal to 2.3, 6.2 and 11.8 relative to the absolute maximum of the likelihood, without imposing the constraint of non-negative branching fractions. The absolute maximum corresponds to $\mathscr{B}(B_s \to \mu^+\mu^-) = (1.5 \pm 1.2) \times 10^{-9}$, $\mathscr{B}(B_d \to \mu^+\mu^-) = (-0.25 \pm 0.20) \times 10^{-9}$, with the uncertainties determined by the likelihood fit and the correlation coefficient equal to -0.4. The relative maximum obtained for non-negative branching fractions is also shown, at $\mathscr{B}(B_s \to \mu^+\mu^-) = (0.9^{+1.1}_{-0.8}) \times 10^{-9}$, $\mathscr{B}(B_d \to \mu^+\mu^-) = 0$, with the 68% confidence range for $\mathscr{B}(B_s \to \mu^+\mu^-)$ obtained from pseudo-MC experiments. The hatched areas on the axes correspond to the excluded values from the extracted upper limits $\mathscr{B}(B_s \to \mu^+\mu^-) < 3.0 \times 10^{-9}$, $\mathscr{B}(B_d \to \mu^+\mu^-) < 0.42 \times 10^{-9}$ at 95% CL. Also shown are the likelihood maximum and the corresponding contours for the combined result of the CMS and LHCb experiments, as well as the SM prediction. Figure from Ref. [3].

6. Results of the $\mathscr{B}(B_{d/s} \to \mu^+\mu^-)$ Measurement

The likelihood fit of the signal yield is followed by the final extraction of the branching ratios. All input terms to the Equation 2.1 were measured as briefly described in the previous sections, taken from PDG world averages [4] or recent ATLAS measurements [5] (see Section 1). These terms were used in a profile-likelihood fit to extract the result of the $\mathscr{B}(B_{d/s} \to \mu^+\mu^-)$ measurement from data. For $B_s \to \mu^+\mu^-$, the minimum of the likelihood was located assuming both signal yields to have physically allowed values (≥ 0). The following result was found: $\mathscr{B}(B_s \to \mu^+\mu^-) = (0.9^{+1.1}_{-0.8}) \times 10^{-9}$ and $\mathscr{B}(B_d \to \mu^+\mu^-) = 0$. The observed compatibility of this result with the SM predictions was found using pseudo-MC experiments to be equivalent to 2 standard deviations (p-value of 0.048). Upper limits were set at 95% confidence level using the CL_s [9] technique: $\mathscr{B}(B_d \to \mu^+\mu^-) < 4.2 \times 10^{-10}$ and $\mathscr{B}(B_s \to \mu^+\mu^-) < 3.0 \times 10^{-9}$. The compatibility of ATLAS results [3] with the SM predictions as well as with the combined results obtained by the CMS and LHCb collaborations [2] can be seen in Figure 3. The ATLAS Run I results show an open door to the possible interference of NP with the SM suppressing the current SM predictions. More details will be surely uncovered in the next iteration of the ATLAS analysis which will benefit from the increased integrated luminosity of Run 2, improved trigger techniques and an upgraded ATLAS detector.

References

[1] C. Bobeth et al., *$B_{s,d} \to l^+l^-$ in the Standard Model with Reduced Theoretical Uncertainty*, Phys. Rev. Lett., **112**, 101801, 2014.

[2] CMS and LHCb Collaborations, *Observation of the rare $B^0_s \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data*, Nature, **522**, 2015.

[3] ATLAS Collaboration, *Study of the rare decays of B^0_s and B^0 into muon pairs from data collected during the LHC Run 1 with the ATLAS detector*, Submitted to Eur. Phys. J. C., arXiv:1604.04263 [hep-ex].

[4] Particle Data Group, *The Review of Particle Physics*, Chin. Phys. C, **38**, 090001, 2014.

[5] ATLAS Collaboration, *Determination of the Ratio of b-quark Fragmentation Fractions f_s/f_d in pp Collisions at $\sqrt{s} = 7$ TeV with the ATLAS Detector*, Phys. Rev. Lett., **115**, 262001, 2015.

[6] A. Hoecker et al., *TMVA 4, Toolkit for Multivariate Data Analysis with ROOT Users Guide*, arXiv:physics/0703039.

[7] ATLAS Collaboration, *Search for the decay $B^0_s \to \mu^+\mu^-$ with the ATLAS detector*, Phys. Lett. B **713**, 387, 2012.

[8] K. De Bruyn et al., *Probing New Physics via the $B^0_s \to \mu^+\mu^-$ Effective Lifetime*, Phys. Rev. Lett. **109**, 041801, 2012.

[9] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. G, **28**, 2693-2704, 2002.

[10] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3**, S08003 (008).

[11] C.-S. Huang, W. Liao and Q.-S. Yan, *The Promising Process to Distinguish Supersymmetric Models with Large tan β from the Standard Model: $B \to X(s)\mu^+\mu^-$*, Phys. Rev. D **59**, 011701, 1999, arXiv:hep-ph/9803460 [hep-ph].

[12] C. Hamzaoui, M. Pospelov and M. Toharia, *Higgs mediated FCNC in Supersymmetric Models with Large tan* β, Phys. Rev. D **59**, 095005, 1999, arXiv:hep-ph/9807350 [hep-ph].

[13] S. R. Choudhury and N. Gaur, *Dileptonic decay of* $B(s)$ *meson in SUSY models with large tan* β, Phys. Lett. **B 451**, 86-92, 1999, arXiv:hep-ph/9810307 [hep-ph].

[14] K. S. Babu and C. F. Kolda, *Higgs mediated* $B^0 \to \mu^+\mu^-$ *in minimal supersymmetry*, Phys. Rev. Lett. **84**, 228-231, 2000, arXiv:hep-ph/9909476 [hep-ph].

[15] S. R. Choudhury et al., *Signatures of new physics in dileptonic B-decays*, Int. J. Mod. Phys. **A 21**, 2617-2634, 2006, arXiv:hep-ph/0504193 [hep-ph].

[16] G. D'Ambrosio et al., *Minimal flavor violation: An Effective field theory approach*, Nucl. Phys. B **645**, 155-187, 2002, arXiv:hep-ph/0207036 [hep-ph].

[17] A. J. Buras, *Relations between* $\Delta M_{s,d}$ *and* $B_{s,d} \to \mu\bar{\mu}$ *in Models with Minimal Flavour Violation*, Phys. Lett. B **566**, 115-119, 2003, arXiv:hep-ph/0303060 [hep-ph].

[18] S. Davidson and S. Descotes-Genon, *Minimal Flavour Violation for Leptoquarks*, JHEP **11**, 073, 2010, arXiv:1009.1998 [hep-ph].

[19] D. Guadagnoli and G. Isidori, $B(B_s \to \mu^+\mu^-)$ *as an electroweak precision test*, Phys. Lett. B **724**, 63-67, 2013, arXiv:1302.3909 [hep-ph].

PoS(HQL 2016)057