

## Search for new physics in rare and semi-rare decays of B-mesons at ATLAS

---

Iskander Ibragimov\* on behalf of the ATLAS collaboration†

Department of Physics, University of Siegen, Walter-Flex-Str. 3, 57068 Siegen, Germany

E-mail: [ibragimov@hep.physik.uni-siegen.de](mailto:ibragimov@hep.physik.uni-siegen.de)

Processes involving the Flavour Changing Neutral Current transitions in b-hadron decays are suppressed in the Standard Model and are sensitive to New Physics. We present a search for the rare decays  $B_s^0 \rightarrow \mu^+\mu^-$  and  $B^0 \rightarrow \mu^+\mu^-$  using  $25 \text{ fb}^{-1}$  of data collected by the ATLAS experiment at 7 and 8 TeV proton-proton collision energy. For the  $B^0$  dimuon decay, an upper limit on the branching fraction is set at  $\mathcal{B}(B^0 \rightarrow \mu^+\mu^-) < 4.2 \times 10^{-10}$  at 95% confidence level. For  $B_s^0$ , the branching fraction  $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-) = (0.9_{-0.8}^{+1.1}) \times 10^{-9}$  is obtained. The results are consistent with the Standard Model expectation and with the combined result of the CMS and LHCb experiments.

*38th International Conference on High Energy Physics  
3-10 August 2016  
Chicago, USA*

---

\*Speaker.

†This work was supported in part by grants of the German Federal Ministry of Education and Research (BMBF).



## 1. Introduction

The branching fractions of the decays  $B_{(s)}^0 \rightarrow \mu^+ \mu^-$  are accurately predicted in the Standard Model (SM):  $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (3.65 \pm 0.23) \times 10^{-9}$  and  $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) = (1.06 \pm 0.09) \times 10^{-10}$  [1]. Any significant deviations from these values would be a clear sign of contributions from New Physics. With the full sample of the LHC Run 1 data the CMS and LHCb collaborations have reported the observation of  $B_s^0 \rightarrow \mu^+ \mu^-$  and evidence of  $B^0 \rightarrow \mu^+ \mu^-$ , with combined values:  $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (2.8_{-0.6}^{+0.7}) \times 10^{-9}$  and  $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) = (3.9_{-1.4}^{+1.6}) \times 10^{-10}$  [2]. Results of an ATLAS analysis [3] using the full LHC Run 1 ATLAS data sample of  $25 \text{ fb}^{-1}$  are presented in this report.

## 2. Analysis Strategy

In order to be independent of the uncertainties of the luminosity and  $b\bar{b}$  production cross-section measurements, the branching fraction measurement is performed with respect to an abundant normalisation (or reference) decay  $B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-)K^+$  with a well-measured branching fraction:

$$\mathcal{B}(B_{(s)}^0 \rightarrow \mu^+ \mu^-) = N_{d(s)} \times [\mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)] \times \frac{f_u}{f_{d(s)}} \times \frac{1}{\mathcal{D}_{\text{norm}}}, \quad (1)$$

with

$$\mathcal{D}_{\text{norm}} = \sum_k N_{J/\psi K^+}^k \alpha_k \left( \frac{\epsilon_{\mu^+ \mu^-}}{\epsilon_{J/\psi K^+}} \right)_k,$$

where  $N_{d(s)}$  ( $N_{J/\psi K^+}$ ) denote the signal (reference) channel yields,  $\epsilon_{\mu^+ \mu^-}$  ( $\epsilon_{J/\psi K^+}$ ) are the corresponding values of acceptance times efficiency, and  $f_u/f_{d(s)}$  is the ratio of the hadronisation probabilities of a  $b$ -quark into  $B^+$  and  $B_{(s)}^0$ . The index  $k$  in the denominator  $\mathcal{D}_{\text{norm}}$  runs over four different data-taking periods and trigger selections. In the sum,  $\alpha_k$  takes the different trigger prescale factors and integrated luminosities in the signal and normalisation channels into account.

Signal and reference channel events are selected by similar dimuon triggers. For the 2011 data-taking period, the trigger selection of  $p_T > 4 \text{ GeV}$  for each of the two muon candidates has been used. The different prescaling factors of low-pT dimuon triggers during the 2012 data-taking period are addressed by splitting the 2012 sample into three mutually exclusive trigger categories ( $T_1$ ,  $T_2$  and  $T_3$ ) depending on the two muons' pT thresholds at 4 GeV and 6 GeV and their trajectory in the detector.

In the preliminary selection of  $B$  candidates, we require two muons with  $p_T > 4 \text{ GeV}$  and  $|\eta| < 2.5$ , reconstructed both in the inner detector and in the muon spectrometer [4]. Additional kaons used in the reconstruction of the normalisation and control channels are required to have  $p_T > 1 \text{ GeV}$ . Only  $B$  candidates inside the fiducial kinematic region of  $p_T^B > 8.0 \text{ GeV}$  and  $|\eta^B| < 2.5$  are retained. In the final selection stage the signal fraction is enhanced by two multivariate classifiers against combinatorial background (c-BDT) and against hadrons which are misidentified as muons (f-BDT). The analysis is blind, i.e. the data from a dimuon invariant mass region between 5166 MeV and 5526 MeV has been excluded from the optimisation of the signal selection procedure.

### 3. Background Suppression

The background to the  $B_{(s)}^0 \rightarrow \mu^+ \mu^-$  signal is comprised of three components: combinatorial background, partially reconstructed decays (PRD) and resonant background. The dominant combinatorial background component originates from the random combination of muon tracks created in quark-anti-quark annihilation processes dominated by  $b\bar{b} \rightarrow \mu^+ \mu^- X$  decays and has a smooth dimuon mass dependence.

The PRD component consists of several topologies: *same-side* (SS) combinatorial background from decay cascades ( $b \rightarrow c\mu\nu \rightarrow s(d)\mu\mu\nu\nu$ ), *same-vertex* (SV) background from  $B$  decays containing a muon pair like  $B^0 \rightarrow K^{*0}\mu\mu$ ,  $B_c$  decays (e.g.  $B_c \rightarrow J/\psi\mu\nu \rightarrow \mu\mu\mu\nu$ ) as well as *semileptonic b-hadron decays* where a final-state hadron is misidentified as a muon. The PRD component accumulates to the left of the signal peak with the tail of the invariant mass distribution reaching into the signal region.

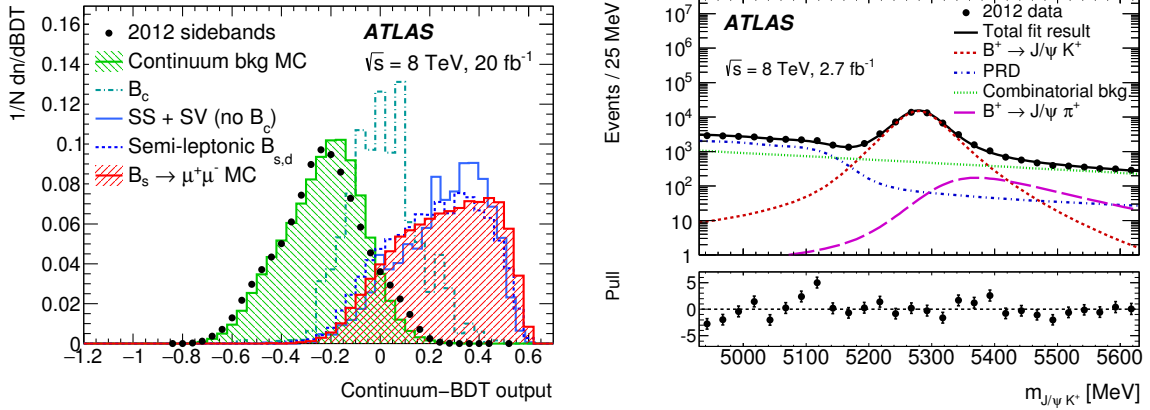
The resonant background, which peaks in the signal region, originates from the decays of neutral  $B$ -mesons with two hadrons in the final state misidentified as muons. The hadron misidentification is mainly due to in-flight decays with probabilities of 0.4% for  $K^\pm$ , 0.2% for  $\pi^\pm$  and less than 0.01% for protons. These probabilities have been further reduced by a factor of 0.4 using a dedicated BDT (f-BDT) selection, which retains 95% of true muons. After applying this selection to the resonant background, its contribution in the signal region is expected to be  $1.0 \pm 0.4$  events.

To suppress the dominant combinatorial background another dedicated BDT (c-BDT) classifier has been employed. It uses 15 input variables related to the B-decay topology and properties of its decay muons, other collision tracks and pile-up vertices in the event. The classifier has been trained on a large inclusive Monte Carlo (MC) sample of uncorrelated semileptonic decays of  $b$  and  $c$  hadrons using the TMVA package [5] and validated on the sidebands of the invariant mass distribution in data. The distributions of the BDT output variable for the sideband data, signal and different background MC samples are shown in Figure 1, left. For the final selection with a c-BDT output value larger than 0.24, 54% of signal events are retained and the combinatorial background component is suppressed by a factor of  $10^3$ .

### 4. Normalisation Channel Yield and Efficiency Ratio Extraction

The  $N_{J/\psi K^\pm}$  yield has been computed in the four data-taking and trigger categories by performing an unbinned maximum likelihood fit to the  $J/\psi K^\pm$  invariant mass distributions after all signal selections have been applied. The fit model consists of four components:  $B^+ \rightarrow J/\psi K^+$  and Cabibbo-suppressed  $B^+ \rightarrow J/\psi \pi^+$  exclusive decays, partially reconstructed  $B$  decays to the left of the  $B^+ \rightarrow J/\psi K^+$  peak and the continuum background. The shape parameters of each component are determined from a simultaneous fit to data and corresponding MC distributions. The fit result for the  $T_1$  category is shown in Figure 1, right. As a byproduct, from the yields of  $B^+ \rightarrow J/\psi K^+$  and  $B^+ \rightarrow J/\psi \pi^+$  decays the ratio  $\rho_{\pi/K}$  of the corresponding branching fractions has been obtained to  $0.035 \pm 0.003 \pm 0.012$ , which is in agreement with the most accurate results available [6, 7].

The efficiency ratios  $(\epsilon_{\mu^+\mu^-}/\epsilon_{J/\psi K^+})_k$  in the four data-taking and trigger categories are obtained from MC after applying data-driven corrections to the  $p_T^B$ ,  $|\eta^B|$  spectra and to the number



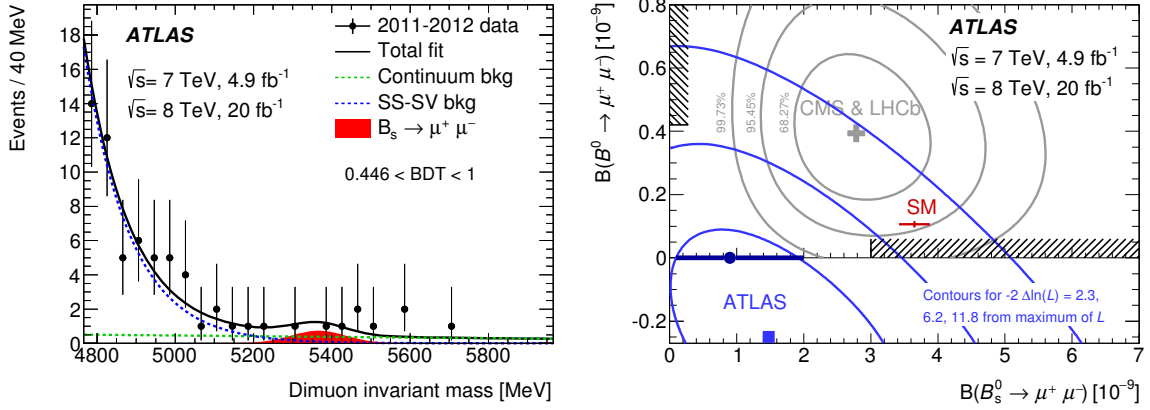
**Figure 1** – *Left*: Continuum-BDT distribution for the signal and background events [3]: signal  $B_s^0$ , partially reconstructed  $B$  events (SS+SV),  $B_c$  decays and continuum. The solid histograms are obtained from simulation, while the points represent data collected in the sidebands. All distributions are normalised to unity. *Right*:  $J/\psi K^+$  invariant mass distribution for all  $B^+$  candidates in the  $T_1$  trigger category in 2012 data [3]. The result of the fit is overlaid and the inset at the bottom of the plot shows the bin-by-bin pulls.

of reconstructed primary vertices, rescaling each category according to the equivalent integrated luminosity in data. In addition, the  $B_s^0$  efficiencies have been corrected to take into account the effective lifetime in the  $B_s^0 \rightarrow \mu^+ \mu^-$  decay, as it proceeds almost exclusively through the heavy  $B_s^0$  mass eigenstate [8]. The systematic uncertainty on the normalisation term  $\mathcal{D}_{\text{norm}}$  of Eq.1 totals  $\pm 5.9\%$ , dominated by residual data-to-MC discrepancies in the 15 c-BDT discriminating variables.

## 5. Signal Yield Extraction

The signal yields  $N_{d(s)}$  are extracted simultaneously by an unbinned maximum-likelihood fit to the dimuon invariant mass distribution after unblinding. The sample is divided into three intervals of c-BDT output with equal share of signal content (18% each) but different levels of background contamination.

The signal decays are parameterised by a superposition of two Gaussian distributions with their widths extracted from MC simulation. The shape and the relative signal efficiencies are constrained to be the same in all three c-BDT intervals. The same model is used for the resonant background, with its contribution (see Section 3) equally shared among the three c-BDT intervals. The combinatorial background is parameterised by a linear function, with independent mass slope parameter and relative fractions in all three c-BDT intervals. The SS+SV background in the low-mass sideband region is modelled by an exponential function. Its shape is constrained to be the same in all three c-BDT intervals with independent normalisations. The total systematic uncertainties on the fitted signal yields are  $\sigma_{\text{syst}}(N_s) = \sqrt{2^2 + (0.06 \times N_s)^2}$  and  $\sigma_{\text{syst}}(N_d) = 3$ . For the SM prediction the expected yield for  $N_s(N_d)$  is 41(5) events. The unblinded dimuon invariant mass distribution for the BDT interval with the largest expected signal sensitivity is shown in Figure 2, left. Applying the requirement of non-negative yields, the fit returns  $N_s = 11$  and  $N_d = 0$ , while the unconstrained fit obtains  $N_s = 16 \pm 12$  and  $N_d = -11 \pm 9$ .



**Figure 2** – *Left*: Dimuon invariant mass distribution in the unblinded data, in the continuum-BDT interval with the largest expected signal sensitivity [3]. Superimposed is the result of the maximum-likelihood fit, obtained imposing the boundary of non-negative signal contributions.

*Right*: Contours in the plane  $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ ,  $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-)$  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 [3]. The absolute maximum corresponds to  $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (1.5 \pm 1.2) \times 10^{-9}$ ,  $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) = (-0.25 \pm 0.20) \times 10^{-9}$ , 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  $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = 0.9_{-0.8}^{+1.1} \times 10^{-9}$ ,  $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) = 0$ , with the 68% confidence range for  $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$  obtained from pseudo-MC experiments. The hatched areas on the axes correspond to the excluded values from the extracted upper limits  $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) < 3.0 \times 10^{-9}$ ,  $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) < 0.42 \times 10^{-9}$  at 95% CL. Also shown are the likelihood maximum and the corresponding contours for the combined result of the CMS and LHCb experiments, and the SM prediction.

## 6. $\mathcal{B}(B_{(s)}^0 \rightarrow \mu^+ \mu^-)$ Extraction

The branching fractions  $\mathcal{B}(B_{(s)}^0 \rightarrow \mu^+ \mu^-)$  are extracted from data using a profile-likelihood fit replacing the fit parameters  $N_{s(d)}$  in Eq. (1) by the corresponding branching fractions divided by normalisation terms. The value of the  $\mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$  is obtained from world averages [9] and the hadronisation probability ratio is taken from the ATLAS measurement as  $f_s/f_d = 0.240 \pm 0.020$  [10], assuming  $f_u/f_d = 1$ . By combining these components with  $\mathcal{D}_{\text{norm}}$ , one obtains a single-event sensitivity equal to  $(8.9 \pm 1.0) \times 10^{-11}$  for  $B_s^0 \rightarrow \mu^+ \mu^-$  and  $(2.21 \pm 0.15) \times 10^{-11}$  for  $B^0 \rightarrow \mu^+ \mu^-$ .

Imposing the constraint on the signal branching fractions to be non-negative one obtains  $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (0.9_{-0.8}^{+1.1}) \times 10^{-9}$ , with a significance of  $1.4\sigma$  w.r.t. the background-only hypothesis. Alternatively, using the  $\text{CL}_s$  method [11], implemented with pseudo-MC experiments, upper limits on the  $B_s^0 \rightarrow \mu^+ \mu^-$  ( $B^0 \rightarrow \mu^+ \mu^-$ ) branching fraction of  $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) < 3.0 \times 10^{-9}$  ( $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) < 4.2 \times 10^{-10}$ ) have been placed at the 95% confidence level. The upper limit on the  $B_s^0 \rightarrow \mu^+ \mu^-$  branching fraction is lower than the SM prediction and the central value of CMS and LHCb combination, with a better agreement to the latter. For  $B^0 \rightarrow \mu^+ \mu^-$ , the limit covers both, the SM prediction and the combined CMS and LHCb measurement.

Without the constraint, the result of the simultaneous fit overlaid with the combined result of the CMS and LHCb experiments and with the SM prediction is shown in Figure 2, right. The compatibility of the result with the SM prediction is given by a  $p$ -value of 0.048, which corresponds to  $2\sigma$ .

## References

- [1] C. Bobeth *et al.*,  $B_{s,d} \rightarrow l^+l^-$  in the Standard Model with Reduced Theoretical Uncertainty, Phys. Rev. Lett. **112** (2014) 101801.
- [2] CMS and LHCb Collaborations, Observation of the rare  $B_s^0 \rightarrow \mu^+\mu^-$  decay from the combined analysis of CMS and LHCb data, Nature **522** (2015) 68.
- [3] ATLAS Collaboration, Study of the rare decays of  $B_s^0$  and  $B^0$  into muon pairs from data collected during the LHC Run 1 with the ATLAS detector, Eur. Phys. J. C **76** (2016) 513.
- [4] ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, JINST **3** (2008) S08003.
- [5] A. Hoecker *et al.*, TMVA 4, Toolkit for Multivariate Data Analysis with ROOT, Users Guide, tech. rep., CERN, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- [6] LHCb Collaboration, Measurements of the branching fractions and CP asymmetries of  $B^\pm \rightarrow J/\psi \pi^\pm$  and  $B^\pm \rightarrow \psi(2S)\pi^\pm$  decays, Phys. Rev. D **85** (2012) 091105.
- [7] BABAR Collaboration, Study of  $B^\pm \rightarrow J/\psi \pi^\pm$  and  $B^\pm \rightarrow J/\psi K^\pm$  decays: Measurement of the ratio of branching fractions and search for direct CP violation, Phys. Rev. Lett. **92** (2004) 241802.
- [8] K. De Bruyn *et al.*, Probing New Physics via the  $B_s^0 \rightarrow \mu^+\mu^-$  Effective Lifetime, Phys. Rev. Lett. **109** (2012) 041801.
- [9] Particle Data Group, Chin.Phys.C **38** (2014) 090001.
- [10] 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** (2015) 262001.
- [11] A. Read, Presentation of search results: The  $CL(s)$  technique, J. Phys. G: Nucl. Part. Phys. **28** (2002) 2693.