

# Rare and semi-rare decays at ATLAS

U De Sanctis *on behalf of the ATLAS Collaboration*

University of Sussex, Brighton BN1 9RH, United Kingdom

E-mail: [umberto.de.sanctis@cern.ch](mailto:umberto.de.sanctis@cern.ch)

**Abstract.** The measurements of the rare  $B^0$ -meson decay processes performed by the ATLAS experiment at the LHC are reviewed. Particular attention is given to the measurement of the branching ratios of the  $B_s^0$  and  $B_d^0$  meson decays into a pair of muons with the full Run 1 dataset corresponding to an integrated luminosity of  $25 \text{ fb}^{-1}$ .

## 1. Introduction

The measurements of rare and semi-rare decays of  $B^0$  mesons are extremely interesting as tests of the predictions of the Standard Model (SM). In fact, they are forbidden at lowest perturbative order in the SM due to flavour-changing neutral-current (FCNC) suppression and can only proceed through loop diagrams. New heavy particles in extensions of the SM can appear in competing Feynman diagrams and significantly affect both branching fractions and the angular distributions of final state particles. Therefore these processes are sensitive to a wide range of new physics phenomena which are not predicted by the SM.

Among these processes, decays of a  $B_s$  and  $B_d$  meson into a pair of oppositely charged muons (referred to in the following as  $B_{(s)}^0 \rightarrow \mu\mu$ ) are of particular interest. In fact, the  $B_{(s)}^0 \rightarrow \mu\mu$  branching fractions are accurately predicted in the SM:  $BR(B_s^0 \rightarrow \mu\mu) = (3.65 \pm 0.23) \times 10^{-9}$  and  $BR(B_d^0 \rightarrow \mu\mu) = (1.06 \pm 0.09) \times 10^{-10}$  [1]. Hence, any significant deviation from these predictions would be a hint of New Physics.

Recently, both CMS and LHCb collaborations have reported an observation of  $B_s^0 \rightarrow \mu\mu$  and an evidence of  $B_d^0 \rightarrow \mu\mu$  decays, leading to a combined measurement of  $BR(B_s^0 \rightarrow \mu\mu) = 2.8_{-0.8}^{+0.7} \times 10^{-9}$  and  $BR(B_d^0 \rightarrow \mu\mu) = 3.9_{-1.4}^{+1.6} \times 10^{-10}$  [2].

In this report, the measurement of the branching ratios of the  $B_s^0 \rightarrow \mu\mu$  and  $B_d^0 \rightarrow \mu\mu$  decays performed by the ATLAS Collaboration is described. The results are extracted from the entire dataset collected by the experiment during the LHC Run 1 campaign, corresponding to an integrated luminosity of  $25 \text{ fb}^{-1}$  collected at a center-of-mass energy of 7 TeV ( $4.7 \text{ fb}^{-1}$ ) and 8 TeV ( $20.3 \text{ fb}^{-1}$ ) [3].

## 2. Analysis outline

A blind analysis is performed in the di-muon invariant mass region between 5166 and 5526 MeV. All the events belonging to this region are removed until the procedures for event selection and the details of signal yield extraction are completely defined. Data used in this analysis have been collected by the ATLAS detector during stable beams conditions. Di-muon triggers have been used to select the events from the  $pp$  collisions. The data sample has been split into



four mutually exclusive trigger categories  $T_i$  depending on the requirements on the transverse momentum and the pseudorapidity of the two muons:

- $T_1$ : “higher threshold” trigger with  $p_T > 6$  GeV and  $> 4$  GeV for the two muons in the 2012 data sample;
- $T_2$ : “barrel” trigger with  $p_T > 4$  GeV for both muon candidates and at least one of them with  $|\eta| < 1.05$  in the 2012 data sample (and  $T_1$  requirement not satisfied);
- $T_3$ : basic di-muon trigger with  $p_T > 4$  GeV for both muon candidates in the 2012 data sample (and  $T_1, T_2$  requirements not satisfied).
- $T_4$ : basic di-muon trigger with  $p_T > 4$  GeV for both muon candidates in the 2011 data sample.

In order to minimise the systematic uncertainties, both branching ratios are extracted relative to the  $B^\pm \rightarrow J/\psi(\rightarrow \mu\mu)K^\pm$  reference channel, which has a well-measured branching ratio. Therefore, the procedure to extract the  $\text{BR}(B_{(s)}^0 \rightarrow \mu\mu)$  has to be corrected taking into account the difference in the hadronisation probability  $f_{u,s,d}$  to form, respectively, a  $B^+$ ,  $B_s$  or  $B_d$  meson giving a  $b$  quark. Also the different acceptances and efficiencies between the signal and the reference channel need to be taken into account. Summarising, the  $\text{BR}(B_{(s)}^0 \rightarrow \mu\mu)$ , can be expressed as:

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

with

$$\mathcal{D}_{\text{norm}} = \sum_k N_{J/\psi K^\pm}^k \alpha_k \left( \frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}} \right)_k. \quad (2)$$

Here the index  $k$  runs over the four different trigger categories used in the analysis,  $\alpha_k$  takes into account the different trigger prescale factors and integrated luminosities in the signal and reference channels.  $N_{d(s)}$  and  $N_{J/\psi K^\pm}^k$  are the event yields, while  $\varepsilon_{\mu^+\mu^-}$  and  $\varepsilon_{J/\psi K^\pm}$  represent the detector acceptance and analysis efficiencies of the signal and the reference channels, respectively.  $B_s \rightarrow J/\psi(\rightarrow \mu\mu)\phi(\rightarrow KK)$  events are also used in the analysis as control sample for the signal kinematic variables used in the analysis.

Muons are required to have  $p_T > 4$  GeV and  $|\eta| < 2.5$  and to be reconstructed in both the Muon Spectrometer and the Inner Detector (ID). Kaons are reconstructed in the ID system and are required to have  $p_T > 1$  GeV and  $|\eta| < 2.5$ .  $B$  mesons kinematic observables are reconstructed imposing quality requirements on the di-muon vertex for signal, or on the di-muon combined with one track vertex for the reference channel. The matching of a  $B$  candidate to a primary vertex (PV) is made by projecting the direction of the candidate to the point of closest approach to the collision axis ( $z$ ) and choosing the PV with the smallest separation in  $z$ . This association procedure has an efficiency of 99% evaluated using simulated signal samples.

The reconstructed  $B$  candidates are finally required to satisfy the following criteria:  $p_T(B) > 8$  GeV and  $|\eta(B)| < 2.5$ . The invariant mass regions used in the analysis for the signal, the reference and the control channels are reported in Table 1.

### 3. Backgrounds

The main backgrounds for this analysis can be classified into three categories: combinatorial background, partially reconstructed  $B$ -decays (PRD) and peaking background.

The combinatorial background consists mainly of real muons produced in the  $b$  or  $\bar{b}$  decay chains. It is the dominant background for the analysis and it is several orders of magnitude larger than the signal. Therefore, a multivariate classifier, namely a Boosted Decision Tree [4] (c-BDT),

is trained in order to reduce its contribution. The c-BDT is based on 15 kinematics variables discriminating the signal from this background. These variables describe: the kinematics of the  $B$  meson candidate, the secondary vertex displacement, the properties of the muons and the rest of the event (such as the  $B$ -isolation  $I^1$ ). The c-BDT is trained on a Monte-Carlo (MC) sample simulating decays of  $b$  and  $c$  quarks (and anti-quarks correspondingly) containing muons. This simulated sample is validated against data candidates from the mass sideband regions. A cut on the c-BDT output at 0.24 is applied, corresponding to an efficiency of 54% on simulated signal and a rejection of about  $10^3$  of the combinatorial background.

The PRD background consists of decays where the two muons in the final state come from one of the following topologies: (a) originating from a common ancestor in 'cascade' transitions (e.g.  $b \rightarrow c\mu\nu \rightarrow s\mu\mu\nu\nu$ ) referred to as *same-side* muons (SS); (b) the same decay diagram (e.g.  $B \rightarrow J/\psi X$  or  $B_d \rightarrow \mu\mu K^*$ ) referred to as *same vertex* muons (SV); (c) the  $B_c \rightarrow J/\psi\mu\nu$  decay; (d) semileptonic  $B$  decays where a hadron  $h$  ( $\pi$  or  $K$ ) is misidentified as a muon (e.g.  $B \rightarrow \mu h\nu$ ). All these backgrounds populate the low-mass sideband and in some cases may have tails extending into the signal region.

The *peaking background* consists of events where the hadronic decays  $B_{(s)}^0 \rightarrow hh'$  ( $h$  being  $\pi$  or  $K$ ) are reconstructed as signal events due to hadron misidentification. This background component peaks in the signal region and needs to be reduced since it overcomes the signal by several orders of magnitude. In addition to the reduction given by the muon identification algorithm (which gives a rejection factor of 250 and 500 for pions and kaons, respectively), a different BDT (called f-BDT) is setup based on calorimetric and tracking variables that are able to distinguish real muons from hadrons. A cut on f-BDT output is tuned to have a 95% efficiency on the signal while reducing the peaking background to  $1.0 \pm 0.4$  events in the signal region. The f-BDT selection also effectively reduces any other background containing hadrons misidentified as muons (such as the semileptonic decays  $B \rightarrow \mu h\nu$ ). Applying this f-BDT cut, an improvement of a factor of seven in the peaking background rejection is achieved with respect to the previous result [5].

#### 4. Reference channel yield extraction and efficiencies computation

In order to extract the  $\text{BR}(B_{(s)}^0 \rightarrow \mu\mu)$  using Eq. 1 the yield of the reference channel and the relative efficiency and acceptance ratios between the two channels need to be computed.

The  $B^\pm \rightarrow J/\psi(\rightarrow \mu\mu)K^\pm$  and the  $B^\pm \rightarrow J/\psi(\rightarrow \mu\mu)\pi^\pm$  yields are simultaneously extracted, using an unbinned extended maximum likelihood (ML) fit to the  $\mu\mu K^\pm$  invariant mass distribution. The functional forms used in the fit are extracted from simulation, while the shape parameters are fitted simultaneously to data and to simulation. As a byproduct, the ratio between  $B^\pm \rightarrow J/\psi(\rightarrow \mu\mu)\pi^\pm$  and  $B^\pm \rightarrow J/\psi(\rightarrow \mu\mu)K^\pm$  was also measured to be  $0.035 \pm 0.003 \pm 0.012$ , in agreement with the PDG average [6].

<sup>1</sup>  $I = \frac{|\vec{p}_T(B)|}{|\vec{p}_T(B)| + \sum_k |\vec{p}_{T,k}|}$  where  $k$  runs over all additional tracks contained within a cone of size  $\Delta R$  that are matched to the same PV as the B candidate

Channel	Low-mass sideband (MeV)	Signal region (MeV)	High-mass sideband (MeV)
$B_{(s)}^0 \rightarrow \mu\mu$	4766–5166	5166–5526	5526–5966
$B^\pm \rightarrow J/\psi K^\pm$	4930–5180	5180–5380	5380–5630
$B_s^0 \rightarrow J/\psi\phi$	5050–5297	5297–5437	5437–5650

**Table 1.** List of the invariant mass regions used in the analysis for signal, reference and control channels.

The efficiency ratio  $\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}$  between signal and reference channels is computed using simulated samples within the fiducial volume of  $p_T(B) > 8$  GeV and  $|\eta(B)| < 2.5$ . The simulated samples are reweighted in order to match the distributions of number of primary vertices (and therefore pile-up),  $p_T(B)$ ,  $|\eta(B)|$  and trigger efficiencies (as a function of  $p_T(\mu)$  and  $|\eta(\mu)|$ ) in data. In addition, a correction to the  $B_s$  lifetime in the simulated signal sample is applied to take into account that the  $B_s \rightarrow \mu\mu$  decay proceeds almost exclusively through the heavy  $B_s$  mass eigenstate [7]. A comparison between the spectra of the 15 variables used in the c-BDT in simulated samples and in data is made to look for possible discrepancies that could affect  $\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}$ . For all variables, but the  $B$  isolation  $I$ , the agreement is found to be good. The simulated spectra are reweighted, one at a time, to match those in data. For each variable, the variation in  $\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}$ , after applying the reweighting procedure, is found to be below 2% and considered as a systematic uncertainty on  $\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}$ . Signal and reference MC samples are instead reweighted to match the  $B$  isolation in data ( $B_s \rightarrow J/\psi\phi$  data for the signal and  $B^\pm \rightarrow J/\psi K^\pm$  data for the reference channel) and used throughout the analysis. The total systematic uncertainty on  $\frac{\varepsilon_{\mu^+\mu^-}}{\varepsilon_{J/\psi K^\pm}}$  is found to be 5.9%.

## 5. Signal yield extraction

The numbers of signal events  $N_{B_s}$  and  $N_{B_d}$  are extracted simultaneously using an unbinned extended ML fit to the di-muon invariant mass  $m_{\mu\mu}$  in the three invariant mass regions shown in Table 1. In order to enhance the discriminating power of the analysis, three bins in the c-BDT output are defined: 0.24–0.346, 0.346–0.446 and 0.446–1. The signal efficiency, as evaluated in simulations, is 18% in each bin. The fit is performed simultaneously in the three c-BDT bins.

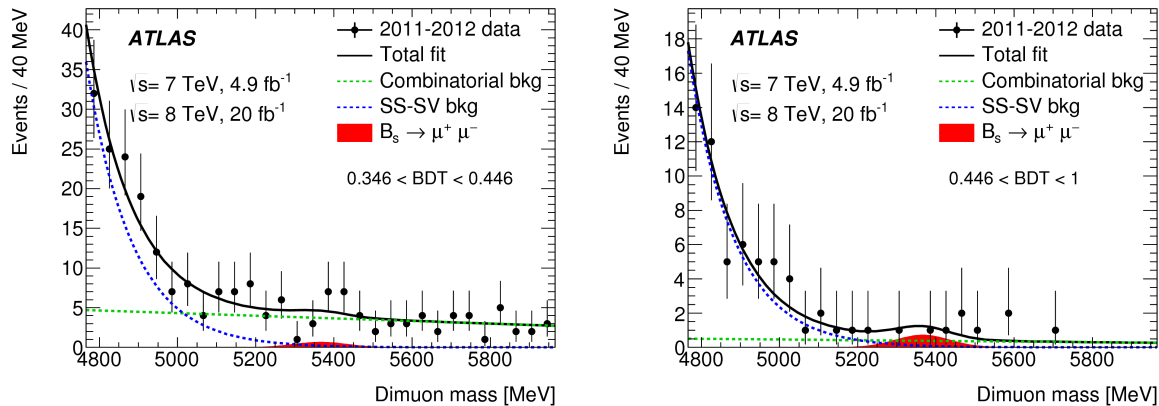
The signals are parameterised by a double Gaussian function to take into account the different resolution of the di-muon invariant mass pair in the four trigger categories. The shape and the relative signal efficiencies are constrained to be the same in all three c-BDT bins. The combinatorial background is parameterised with a first order polynomial, with mass slope and relative fractions in the three c-BDT bins as independent parameters. The background coming from SS and SV events (see Section 3) is fitted with an exponential function. The background invariant mass shape is constrained to be the same in all three c-BDT bins with independent normalisations. The peaking background is fitted with the same model as the signal and normalised to a total of  $1.0 \pm 0.4$  events (see Section 3) equally distributed in the three c-BDT bins. The resulting fit is shown in Fig. 1 for the second and the third c-BDT bin. All backgrounds are normalised independently in each c-BDT bin.

The main systematic uncertainties affecting  $N_s$  and  $N_d$  as extracted from the fit are evaluated using a toy MC. In particular, different fit PDFs are chosen for the various backgrounds. Systematic uncertainties related to mass resolution and scale and to the constraint on the signal efficiencies to be the same across the three c-BDT bins are also considered. The total systematic uncertainties on the fitted signal yields are:  $\sigma_{syst}(N_{B_d}) = 3$  and  $\sigma_{syst}(N_{B_s}) = \sqrt{4 \pm 0.06 \times N_{B_s}}$  events, respectively. These systematic uncertainties are included in the likelihood as Gaussian constraints before the fit to data.

The expected number of signal events, based on the SM calculations and taking the relevant input parameters from PDG [6] and recent ATLAS results [8], is:  $N_{B_s} = 41$  and  $N_{B_d} = 5$ . The number of fitted signal events is found to be:  $N_{B_s} = 16 \pm 12$  and  $N_{B_d} = -11 \pm 9$ .

## 6. $BR(B_{(s)}^0 \rightarrow \mu\mu)$ extraction

The  $BR(B_{(s)}^0 \rightarrow \mu\mu)$  are extracted from data through a profile-likelihood fit using the terms written in Eq. 1. The  $BR(B^\pm \rightarrow J/\psi(\rightarrow \mu\mu)K^\pm) \times BR(J/\psi \rightarrow \mu\mu)$  term is taken from PDG [6],



**Figure 1.** 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.

while the hadronisation probability ratio  $\frac{f_u}{f_{d(s)}}$  from a measurement made by ATLAS [8].

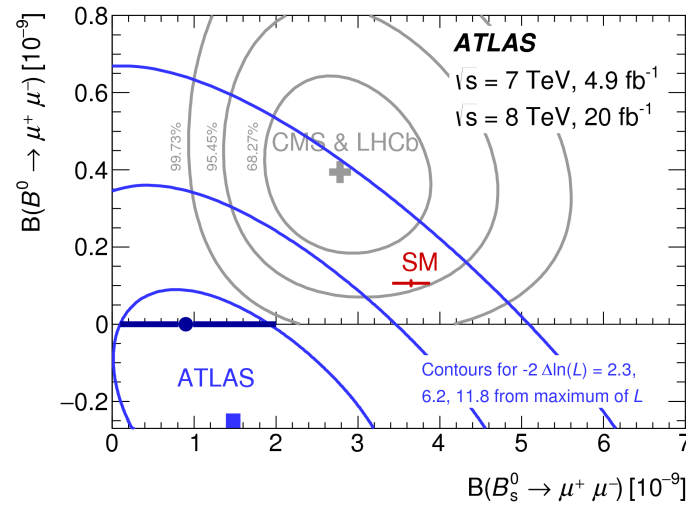
The single-event-sensitivities are found to be  $(8.9 \pm 1.0) \times 10^{-11}$  and  $(2.21 \pm 0.15) \times 10^{-11}$  for  $B_s$  and  $B_d$ , respectively. The values maximising the likelihood with the constraints on  $N_{B_s}$  and  $N_{B_d}$  to be positive are found to be:  $\text{BR}(B_s^0) = 0.9_{-0.8}^{+1.1} \times 10^{-9}$  and  $\text{BR}(B_d^0) = 0$ . All the results in the following are obtained applying those constraints if not otherwise stated. The compatibility of the results with the SM prediction is evaluated with toy-MC experiments using a hypothesis test based on the likelihood ratio  $-\ln[L(\text{SM})/L(\text{max})]$  for the simultaneous fit to  $\text{BR}(B_s^0)$  and  $\text{BR}(B_d^0)$ . The result is a  $p$ -value of 0.048, corresponding to 2.0 standard deviations.

Upper limits on both BR are also computed at the 95%CL, leading to the following results:  $\text{BR}(B_s^0) < 3.0 \times 10^{-9}$  and  $\text{BR}(B_d^0) < 4.2 \times 10^{-10}$ . The  $\text{BR}(B_s^0)$  upper limit is obtained with toy MC experiments under the background only hypothesis leaving the  $\text{BR}(B_d^0)$  free to be determined by the fit. Similarly, the  $\text{BR}(B_d^0)$  upper limit is set leaving the  $\text{BR}(B_s^0)$  free to be determined by the fit.

Figure 2 summarises the outcomes of the measurement in comparison with the SM predictions as well as previous combined results published by the CMS and LHCb collaborations [2]. Blue curves represent intervals of  $-2 \Delta \ln(L)$  equal to 2.3, 6.2 and 11.8 that are obtained with respect to the absolute maximum of the likelihood (i.e. without constraining both BR to be positive), while gray curves represent the CL intervals obtained by CMS and LHCb experiments.

## 7. Conclusions

This report summarised the measurement of the  $\text{BR}(B_{(s)}^0 \rightarrow \mu\mu)$  performed by the ATLAS Collaboration analysing the full Run 1 dataset corresponding to an integrated luminosity of  $25 \text{ fb}^{-1}$ . The results found are compatible with the SM prediction within 2.0 standard deviations and are in agreement with the measurements performed by the CMS and LHCb collaborations. Other measurements of semi-rare decay processes, such as  $B_d \rightarrow \mu\mu K^*$  are still under scrutiny by the ATLAS Collaboration. These decays, together with other rare and semi-rare processes, will be an essential part of the B-physics programme of the ATLAS Collaboration for the ongoing Run 2 campaign, where the ATLAS experiment will benefit from improved detector and trigger performance and from the increased integrated luminosity that will be available.



**Figure 2.** Contours in the plane  $\text{BR}(B_s^0 \rightarrow \mu\mu)$ ,  $\text{BR}(B_d^0 \rightarrow \mu\mu)$  for intervals of  $-2 \Delta \ln(L)$  equal to 2.3, 6.2 and 11.8 relative to the absolute maximum of the likelihood (blue square), without imposing the constraint of non-negative branching fractions. Also shown are the corresponding contours for the combined result of the CMS and LHCb experiments, the SM prediction, and the maximum of the likelihood within the boundary of non-negative branching fractions (blue dot), with the error bars covering the 68.3% confidence range for  $\text{BR}(B_s^0 \rightarrow \mu\mu)$  under the conditions explained in the text.

## References

- [1] Bobeth C et al, 2014  $B_{s,d} \rightarrow l^+ l^-$  in the Standard Model with reduced theoretical uncertainty, *Phys. Rev. Lett.*, **112**, 101801
- [2] CMS and LHCb Collaborations, 2015 Observation of the rare  $B_s^0 \rightarrow \mu^+ \mu^-$  decay from the combined analysis of CMS and LHCb data, *Nature*, **522**
- [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, 2016 *Preprint* arXiv:1604.04263 [hep-ex] *Submitted to Eur.Phys. J. C*.
- [4] Hoecker A et al, 2007 TMVA 4 Toolkit for Multivariate Data Analysis with ROOT users guide, *Preprint* arXiv:physics/0703039
- [5] ATLAS Collaboration, 2012 Search for the decay  $B_s^0 \rightarrow \mu^+ \mu^-$  with the ATLAS detector, *Phys. Lett. B* **713** 387
- [6] Particle Data Group, 2014 The review of particle physics, *Chin.Phys.C*, **38**, 090001
- [7] De Bruyn K et al, 2012 Probing New Physics via the  $B_s^0 \rightarrow \mu\mu$  effective lifetime, *Phys. Rev. Lett.* **109** 041801
- [8] ATLAS Collaboration, 2015 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