

ATLAS $B_s \rightarrow \mu^+ \mu^-$

Takuya Nobe*

on behalf of the ATLAS Collaboration

Tokyo Institute of Technology

E-mail: Takuya.Nobe@cern.ch

A search for $B_s \rightarrow \mu^+ \mu^-$ in pp collisions at a centre-of-mass energy of 7 TeV has been performed using data collected with the ATLAS detector at the LHC in 2011 corresponding to 2.4 fb^{-1} of the integrated luminosity. The decay $B_s \rightarrow \mu^+ \mu^-$ is a Flavor Changing Neutral Current process, which is suppressed strongly in the Standard Model (SM), while contributions of unknown particles may enhance its branching ratio. Therefore it is a very sensitive channel in the search for physics beyond the SM. The signal region was blinded until the background in signal-free region was well understood. After unblinding, no significant excess was observed in the signal region and the upper limit on the branching ratio was determined to be 2.2×10^{-8} at 95% C.L.

*14th International Conference on B-Physics at Hadron Machines,
April 8-12, 2013
Bologna, Italy*

*Speaker.



1. Introduction

The Flavor Changing Neutral Current (FCNC) process $B_s \rightarrow \mu^+ \mu^-$ is strongly suppressed in the Standard Model (SM). The branching ratio of $B_s \rightarrow \mu^+ \mu^-$, $\mathcal{B}(B_s \rightarrow \mu^+ \mu^-)$ is expected to be extremely small: $\mathcal{B}(B_s \rightarrow \mu^+ \mu^-)_{\text{SM}} = (3.56 \pm 0.18) \times 10^{-9}$ [1]. On the other hand, contributions of unknown particles, such as those predicted by Minimal Supersymmetric Standard Model and other extensions may enhance its branching ratio. Therefore it is a very sensitive channel in the search for physics beyond the SM. The first measurement of its branching ratio was performed by LHCb, yielding $\mathcal{B}(B_s \rightarrow \mu^+ \mu^-) = 3.2_{-1.2}^{+1.5} \times 10^{-9}$ [2]. This paper reports the result of a search performed with pp collisions corresponding to an integrated luminosity of 2.4 fb^{-1} , collected with the ATLAS detector at the LHC. More details are described in Ref. [3].

The branching ratio, $\mathcal{B}(B_s \rightarrow \mu^+ \mu^-)$ is measured with respect to a reference decay $B^\pm \rightarrow J/\psi K^\pm (\rightarrow \mu^+ \mu^- K^\pm)$ in order to suppress systematic uncertainties in the evaluation of the efficiencies and acceptances. It is obtained using the following formula:

$$\mathcal{B}(B_s \rightarrow \mu^+ \mu^-) = \mathcal{B}(B^\pm \rightarrow J/\psi K^\pm) \times \frac{f_u}{f_s} \times \frac{N_{\mu^+ \mu^-}}{N_{J/\psi K^\pm}} \times \frac{A_{J/\psi K^\pm} \varepsilon_{J/\psi K^\pm}}{A_{\mu^+ \mu^-} \varepsilon_{\mu^+ \mu^-}}, \quad (1.1)$$

where $\mathcal{B}(B^\pm \rightarrow J/\psi K^\pm)$ is the branching ratio for the reference channel, f_u/f_s is the relative production probability of B^\pm and B_s , and N , A and ε are the number of events observed after background subtraction, total selection acceptance, and efficiency in the signal ($\mu^+ \mu^-$) and the reference ($J/\psi K^\pm$) channels. The value $\mathcal{B}(B^\pm \rightarrow J/\psi K^\pm) = (6.01 \pm 0.21) \times 10^{-5}$ was taken from the current PDG average [4] and $f_u/f_s = 0.267 \pm 0.021$ was taken from LHCb result [5]. In order to avoid an analysis-level bias, a blind analysis was performed in which the di-muon mass region $5066 \text{ MeV} < m_{\mu\mu} < 5666 \text{ MeV}$ was removed from the analysis until the procedures for event selection, signal and limit extractions were fully defined.

2. Experimental setup

2.1 The ATLAS detector

The ATLAS detector¹ is a multi-purpose physics apparatus with a forward-backward symmetric cylindrical geometry, which covers nearly 4π in the solid angle. It consists of an inner tracking detector (ID), electromagnetic and hadronic calorimeters, and a muon spectrometer (MS). The transverse momentum (p_T) resolution in the ID is $\sigma(p_T)/p_T = (0.05 p_T \oplus 1)\%$ and the transverse impact parameter resolution is about $10 \mu\text{m}$. A detailed description of the ATLAS detector can be found elsewhere [6].

ATLAS has a multi-level muon trigger system, with a Level 1 (L1) and a High Level Trigger (HLT) of increasing complexity and selectivity.

¹ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point. The z-axis is along the beam pipe, the x-axis points to the centre of the LHC ring and the y-axis points upward. Cylindrical coordinates (r , ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity η is defined as $\eta = -\ln[\tan(\theta/2)]$ where θ is the polar angle

2.2 Monte Carlo simulation

Signal and reference channel Monte Carlo (MC) samples were used for the extraction of the acceptance times efficiency ratios: $R_{A\epsilon} = A_{J/\psi K} \epsilon_{J/\psi K} / A_{\mu^+ \mu^-} \epsilon_{\mu^+ \mu^-}$. B decay products were required to satisfy $|\eta| < 2.5$ and $p_T > 2.5(0.5)$ GeV for muons (kaons) at the generator level. Generator level re-weighting was performed using sample without pre-selection and over a wider range in the b -quark kinematics. After that, data-driven weights were derived in bins of the reconstructed p_T and η of the B meson, p_T^B and η^B , by comparing $B^\pm \rightarrow J/\psi K^\pm$ MC and observed data and applied to the signal MC. The pileup effect due to additional pp interactions in the same and other bunch crossings during the same read-out window of the detector was considered in the simulation.

3. Event selection

3.1 Pre-selection

Di-muon candidates were selected at the trigger level to satisfy $p_T > 4$ GeV for both muon candidates. A full track reconstruction of muon was performed in the HLT and loose cuts on $m_{\mu\mu}$ were applied compatible with a J/ψ (2500 to 4300 MeV) or B_s (4000 to 8500 MeV). At the offline level, several pre-selection cuts based on kinematics were applied: For both identified muons, the corresponding ID tracks were required to satisfy $|\eta| < 2.5$ and $p_T > 4$ GeV; The χ^2/dof of the fitted B decay vertex had to be less than 2.0 (6.0) for the signal (reference) channel; Reconstructed B candidate had to satisfy $p_T^B > 8.0$ GeV and $|\eta^B| < 2.5$. All tracks were required to have good quality hits at ID and be matched to a reconstructed MS track. In addition, for the reference channel only, the kaon track is required to have $|\eta| < 2.5$ and $p_T > 2.5$ GeV, and the J/ψ candidate was required to have an invariant mass between 2915 and 3175 MeV. Signal regions were defined in the invariant mass of B , m_B : 5066 to 5666 MeV for $B_s \rightarrow \mu^+ \mu^-$ and 5180 to 5380 MeV for $B^\pm \rightarrow J/\psi K^\pm$. Sideband regions were defined as 300 (200) MeV outside of the signal regions for signal (reference) channel, both higher and lower sides. The primary vertex position was obtained from a fit of all charged tracks associated with the same primary vertex as the B candidate, and excluding those tracks from the B decay.

After the pre-selection, $B_s \rightarrow \mu^+ \mu^-$ candidates were selected by multivariate analysis.

3.2 Background categorization

Two categories of background for $B_s \rightarrow \mu^+ \mu^-$ were considered: a continuum background with a smooth dependence on $m_{\mu\mu}$ and a peaking background from mis-reconstructed resonance decays. The continuum background was estimated by interpolation from the sideband data. The distribution was compared with simulated data from $b\bar{b} \rightarrow \mu^+ \mu^- X$ after the same selections, and there was a reasonable agreement. Half of the sideband data (with even event numbers) were used for this procedure while the other half (odd numbers) were used to optimize the selection cut. The peaking background is due to B decay candidates containing either one or two hadrons mis-identified as muons. The main component in this class of background had two fake muons, e.g. $B_s \rightarrow K^+ K^-$, $B^0 \rightarrow \pi^+ \pi^-$ and $B^0 \rightarrow K^\pm \pi^\pm$. Decay-in-flight muon carries most of the hadron momentum, so the reconstructed di-muon mass peaks in the signal region. The peaking background kinematics is similar to the signal, so even the multivariate discrimination is expected to be poor. The expected

background rate was obtained with Monte Carlo, by multiplying the integrated luminosity, acceptance, efficiency, and fake probability. The probability of misidentifying a hadron as a muon is 2 (4) % for π^\pm (K^\pm), with a relative uncertainty of 20%.

3.3 Multivariate analysis

The TMVA package [7] implementation of Boosted Decision Trees (BDT) was used to discriminate the signal from the continuum background. Fourteen discriminating variables were selected as BDT inputs. For example, Fig.1 shows (a) L_{xy} , (b) α_{2D} and (c) $I_{0.7}$ distributions for the signal and sideband data. $L_{xy} = (\Delta\vec{x} \cdot \vec{p}^B) / p_T^B$ is the distance between the primary vertex and the decay vertex in the plane transverse to the beam; α_{2D} is the angle between the B -candidate transverse momentum vector and the line between the primary and secondary vertex, in the plane transverse to the beam. $I_{0.7}$ is the ratio of the p_T of the B candidate to the scalar sum of its p_T with the p_T all other tracks with $p_T > 0.5$ GeV and within a cone $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.7$ around the B candidate direction. In order to suppress the pile-up dependency, only the tracks originating from the primary vertex associated with the B decay are accounted for the definition of $I_{0.7}$.

The upper plot of Fig. 2 shows the BDT output for signal Monte Carlo and sideband data. The cut values of BDT output and $m_{\mu\mu}$ were optimized to maximize the sensitivity, P defined as $\varepsilon_{\text{sig}} / (1 + \sqrt{N_{\text{bkg}}})$, where $\varepsilon_{\text{sig}} = A_{\mu^+\mu^-} \cdot \varepsilon_{\mu^+\mu^-}$ and N_{bkg} are the signal acceptance times efficiency and the background held for a given set of cuts, respectively. The invariant mass resolution depends on the η of the muons due to the increase in multiple scattering and the decrease of the magnetic field at large values of $|\eta|$. Three categories were defined according to $|\eta|_{\text{max}}$, defined as the largest $|\eta|$ of the two muons for each candidate: Barrel (0-1.0), Transition (1.0-1.5) and Endcap (1.5-2.5). The result of the optimization in each $|\eta|_{\text{max}}$ region is summarized in Tab. 1. To avoid bias, the BDT training and cut optimization were performed on odd-numbered events and even-numbered events were used for the measurement of the background and for interpolation to the signal region. BDT output was confirmed to be independent of $m_{\mu\mu}$ over the sideband regions and signal region.

The bottom plot of Fig. 2 shows the BDT output for reference channel $B^\pm \rightarrow J/\psi K^\pm$, comparing data and Monte Carlo. The data and Monte Carlo agree well after the application of similar kinematic cuts as for the signal. The same classification based on $|\eta|_{\text{max}}$ was used for the reference channel.

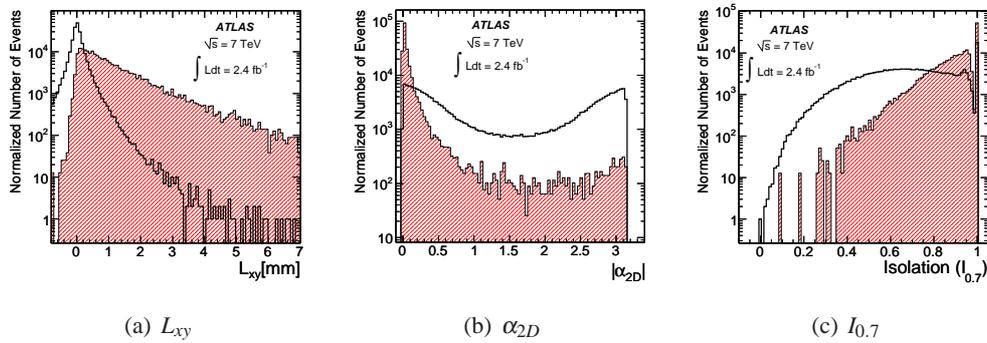


Figure 1: (a) L_{xy} , (b) α_{2D} and (c) $I_{0.7}$ distributions for signal Monte Carlo (shaded) and sideband data (empty) [3].

$ \eta _{\max}$ range	Barrel	Transition	Endcap
invariant mass window [MeV]	± 116	± 133	± 171
BDT output threshold	0.234	0.245	0.270
sideband count	5	0	2
bkg. scaling factor	1.29	1.14	0.88
expected resonant bkg.	0.10	0.06	0.08
signal region count	2	1	0

Table 1: Thresholds of BDT output and the invariant mass and the relative acceptance-times-efficiency ratio $R_{A\epsilon}$ for each mass-resolution category. Sideband count, bkg. scaling factor, expected resonant bkg., and signal region count in each $|\eta|_{\max}$ region are also represented. Bkg. scaling factor represents the expected ratio of the sidebands (even event number) to those in the search window.

4. Reference channel yield

The reference channel yield $N_{J/\psi K^\pm}$ was determined from a binned likelihood fit to the invariant mass distribution of the $\mu^+ \mu^- K^\pm$, performed in the mass range of 4930-5630 MeV. Only even-numbered events were used in the extraction of $B^\pm \rightarrow J/\psi K^\pm$ yield. The B^\pm signal was modeled with two Gaussian distributions of equal mean values. The background was modeled with the sum of: (a) an exponential function for the continuum background; (b) an exponential function multiplied by a complementary error function describing the low-mass contribution for partially reconstructed decays; and (c) a Gaussian function for the background from $B^\pm \rightarrow J/\psi \pi^\pm$. The reference channel yields were estimated to be: 4300 ± 142 (Barrel), 1410 ± 112 (Transition), and 1130 ± 163 (Endcap). Both statistical and systematic uncertainties were considered.

The acceptance times efficiency ratio $R_{A\epsilon}$ was estimated by the Monte Carlo to be 0.274 ± 0.012 in Barrel, 0.202 ± 0.015 in Transition, and 0.143 ± 0.011 in Endcap. The effect of small differences between the data and Monte Carlo are reduced by the data-driven reweighting of the Monte Carlo described in Sec. 2.2. Furthermore, only differences between signal and reference channel were considered.

5. Results

Finally, the signal regions were unblinded and no significant signal was observed in each $|\eta|_{\max}$ region. The num-

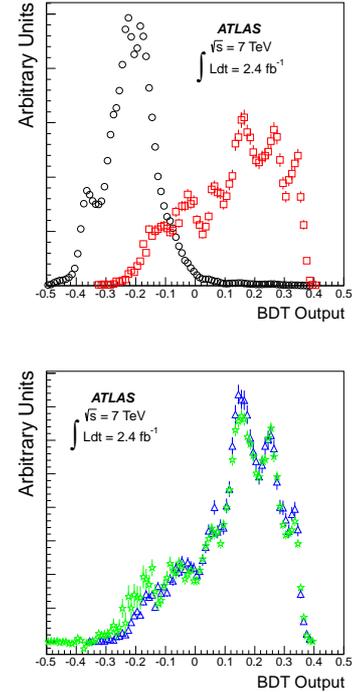


Figure 2: Distributions of the response of the BDT classifier. Top: $B_s \rightarrow \mu^+ \mu^-$ Monte Carlo sample (squares) and data sidebands (circles); bottom: $B^\pm \rightarrow J/\psi K^\pm$ events from Monte Carlo samples (triangles) and sideband-subtracted data (stars) [3].

bers of observed events and the expected background rates are summarized in Tab. 1. The upper limit on the branching ratio was obtained by means of an implementation [8] of the CLs method [9].

The extraction of limits on branching ratio was based on the likelihood considering the Poisson probability to observe events in the signal and sideband regions with signal and background expectation. The mean value of the Poisson distribution in the signal region is equal to the sum of the B_s branching ratio, the continuum background and the resonant background. In the sideband region, it is equal to the background scaled by transfer factor from the total background to even-numbered events. The continuum backgrounds in both regions are estimated simultaneously. Systematic uncertainties were convoluted into the likelihood as Gaussian distributions.

The expected limits were obtained by setting the counts in the signal region equal to the interpolated background plus the small resonant background. A median expected limit of $2.3_{-0.5}^{+1.0} \times 10^{-8}$ at 95% CL was obtained, where the range encloses 68% of the background-only pseudo-experiments. The behavior of the observed CLs for different tested values of the $\mathcal{B}(B_s \rightarrow \mu^+ \mu^-)$ is shown in Fig. 3. The observed upper limit is $2.2(1.9) \times 10^{-8}$ at 95 (90) % CL.

6. Conclusion

A limit on the branching fraction $\mathcal{B}(B_s \rightarrow \mu^+ \mu^-)$ is determined using 2.4 fb^{-1} of integrated luminosity collected in 2011 by the ATLAS detector. $B_s \rightarrow \mu^+ \mu^-$ candidates were selected by a multivariate analysis performed on three categories of events with different mass resolution, yielding a limit of $\mathcal{B}(B_s \rightarrow \mu^+ \mu^-) < 2.2(1.9) \times 10^{-8}$ at 95% (90%) CL.

Acknowledgement We thank CERN for the successful operation of the LHC and the support of the ATLAS experiment. T. N. acknowledges Tokyo Institute of Technology Global Center of Excellence Program ‘Nano science and Quantum Physics’ for the assistance of the work.

References

- [1] A. J. Buras, R. Fleischer, J. Girrbachz, and R. Knegjens, arXiv:1303.3820 [hep-ph].
- [2] LHCb Collaboration, R. Aaij et al., Phys.Rev.Lett. **110** (2013) 021801.
- [3] ATLAS Collaboration, G. Aad et al., Phys.Lett. **B713** (2012) 387–407.
- [4] K. Particle Data Group, Nakamura et al., J.Phys. **G37** (2010) 075021.
- [5] LHCb Collaboration, R. Aaji et al., Phys. Rev. D **85** (2012) 032008.
- [6] ATLAS Collaboration, G. Aad et al., JINST **3** (2008) S08003.
- [7] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, et al., PoS ACAT (2007) 040.
- [8] T. Junk, Nucl.Instrum.Meth. **A434** (1999) 435–443.
- [9] A. L. Read, J.Phys. **G28** (2002) 2693–2704.

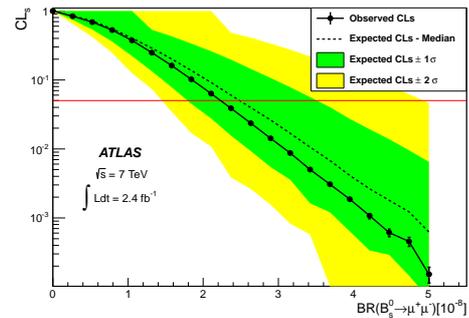


Figure 3: Observed CLs (circles) as a function of $\mathcal{B}(B_s \rightarrow \mu^+ \mu^-)$. The 95% CL limit is indicated by the horizontal line. The dark and light bands correspond to $\pm 1\sigma$ and $\pm 2\sigma$ fluctuations on the expectation, based on the number of observed events in the signal and sideband regions [3].