THE BRANCHING FRACTION AND EFFECTIVE LIFETIME OF $B^0_{(s)} \to \mu^+\mu^-$ AT LHCb WITH RUN 1 AND RUN 2 DATA

M. MULDER
on behalf of the LHCb Collaboration
Nikhef, Science Park 105,
1098 XG Amsterdam, Netherlands

After Run 1 of the LHC, global fits to $b \to s\ell\ell$ observables show a deviation from the Standard Model (SM) with a significance of ~ 4 standard devations. An example of a $b \to s\ell\ell$ process is the decay of a B^0_s meson into two muons ($B^0_s \to \mu^+\mu^-$). The latest analysis of $B^0_{(s)} \to \mu^+\mu^-$ decays by LHCb with Run 1 and Run 2 data is presented. The $B^0_s \to \mu^+\mu^-$ decay is observed for the first time by a single experiment. In addition, the first measurement of the $B^0_s \to \mu^+\mu^-$ effective lifetime is performed. No significant excess of $B^0 \to \mu^+\mu^-$ decays is observed. All results are consistent with the SM and constrain New Physics in $b \to s\ell\ell$ processes.

1 Introduction

In the SM, Flavour Changing Neutral Currents (FCNCs) occur at loop level and are suppressed by the GIM mechanism, and sometimes helicity suppression. As New Physics (NP) is not necessarily suppressed, FCNCs probe physics at energies beyond the LHC centre-of-mass energy.

One such FCNC is the $b \to s\ell\ell$ transition. Several theory groups have performed global fits to $b \to s\ell\ell$ observables in the Effective Field Theory (EFT) framework, and find that the data deviate by ~ 4 standard deviations with respect to the SM [1,2,3].

One example of a $b \to s\ell\ell$ transition is the decay of a B^0_s meson into two muons ($B^0_s \to \mu^+\mu^-$). For $B^0_{(s)} \to \mu^+\mu^-$ decays, the decay amplitude can be written as

$$A(B^0_{(s)} \to \mu^+\mu^-) = \langle \mu^+\mu^- | \mathcal{H}_{eff} | B^0_{(s)} \rangle = \frac{G_F}{\sqrt{2}} V^*_{\mathrm{t(d,s)}} V_{\mathrm{tb}} \sum_{i \in 10,S,P} \mathcal{C}^{(\prime)}_i \langle \mu^+\mu^- | \mathcal{O}^{(\prime)}_i | B^0_{(s)} \rangle \,, \quad (1)$$

where G_F is the Fermi constant, $V^*_{\mathrm{t(d,s)}}$ and V_{tb} are CKM elements, $\mathcal{C}^{(\prime)}_i$ are perturbative Wilson coefficients and $\mathcal{O}^{(\prime)}_i$ are non-perturbative Wilson operators. These operators describe axial vector ($\mathcal{O}^{(\prime)}_{10}$), scalar ($\mathcal{O}^{(\prime)}_S$) and pseudo-scalar ($\mathcal{O}^{(\prime)}_P$) contributions respectively, and a prime indicates a right-handed contribution. In the SM, only $\mathcal{C}_{10}$ contributes. Interestingly, contributions from $\mathcal{C}^{(\prime)}_{10}$ are helicity suppressed, while (pseudo-)scalar contributions are not. Therefore, $B^0_{(s)} \to \mu^+\mu^-$ observables are very sensitive to (pseudo-)scalar NP.

As $B^0_{(s)} \to \mu^+\mu^-$ decays are purely leptonic, the hadronic part of the amplitude reduces to the decay constant, which is well determined using lattice QCD [4]. The branching fractions of $B^0 \to \mu^+\mu^-$ and $B^0_s \to \mu^+\mu^-$ are consequently well predicted in the SM [5,6]:

$$\mathcal{B}(B^0_s \to \mu^+\mu^-) = (3.57 \pm 0.18) \times 10^{-9} \quad (2)$$

$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.06 \pm 0.06) \times 10^{-10} \quad (3)$$

A second observable in the $B^0_{(s)} \to \mu^+\mu^-$ system is the effective lifetime of $B^0_s \to \mu^+\mu^-$ [7]. $B^0_{(s)}$ mesons mix with their anti-particles and form mass (and lifetime) eigenstates. The mass-eigenstate rate asymmetry is given by

$$\mathcal{A}_{\Delta\Gamma} = \frac{\Gamma(B^0_{(s),H} \to \mu^+\mu^-) - \Gamma(B^0_{(s),L} \to \mu^+\mu^-)}{\Gamma(B^0_{(s),H} \to \mu^+\mu^-) + \Gamma(B^0_{(s),L} \to \mu^+\mu^-)} \tag{4}$$

In the SM, $\mathcal{A}_{\Delta\Gamma} = 1$, which means that only the heavy $B^0_{(s)}$ eigenstate decays to two muons. For B^0_s mesons, the lifetime difference between their two eigenstates is measured to be $\Delta\Gamma/\Gamma = 0.12$ [a] By measuring the effective lifetime from the untagged decay distribution, the mass-eigenstate rate asymmetry is probed. The effective lifetime of $B^0_s \to \mu^+\mu^-$ is an independent and complementary observable to the branching fraction.

Here, the new LHCb analysis of $B^0_{(s)} \to \mu^+\mu^-$ observables is presented [8]. It is based on data corresponding to an integrated luminosity of $1\,\mathrm{fb}^{-1}$ of pp collisions at a centre-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$, $2\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 8\,\mathrm{TeV}$ and $1.4\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 13\,\mathrm{TeV}$. The first two datasets are referred to as Run 1, the latter as Run 2.

In Section 2, the branching fraction measurement is presented. In Section 3, the effective lifetime measurement is presented. As there is a large overlap in selection and analysis strategy, only differences with the branching fraction measurement are mentioned in the latter Section.

2 $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ branching fraction measurement

2.1 Backgrounds and selection

The analysis strategy is to search for muon pairs with opposite charges, $m_{\mu^+\mu^-} \in [4900, 6000]\,\mathrm{MeV}$ [b] and a well reconstructed, clearly displaced decay vertex.

Four background categories exist for the branching fraction measurement. First, $\mathrm{b\overline{b}} \to \mu\mu X$ decays, where both muons combine to fake a signal. This is called combinatorial background, and all other backgrounds are hereafter referred to as peaking backgrounds. Second, $B^0_{(s)} \to h^+h^{(\prime)-}$ decays, where $h \in [K, \pi]$ and both hadrons are misidentified as muons. Third, $B^0 \to \pi^-\mu^+\nu_\mu$, $B^0_s \to K^-\mu^+\nu_\mu$, and $\Lambda^0_b \to p\mu^-\nu$, where one hadron is misidentified as a muon and the neutrino is not reconstructed. Fourth, $B^+_c \to J/\psi(\to \mu^+\mu^-)\mu^+\nu_\mu$ and $B^{0(+)} \to \pi^{0(+)}\mu^+\mu^-$ decays, where two real muons are combined and one particle is not reconstructed.

To reject both combinatorial and peaking backgrounds, two multivariate classifiers are used.

For the separation of $B^0_{(s)} \to \mu^+\mu^-$ from combinatorial background, a Boosted Decision Tree (BDT) is trained. This BDT is used to divide the sample based on signal purity. The main discriminating variables in this BDT are based on track isolation, as $\mathrm{b\overline{b}} \to \mu\mu X$ decays have extra tracks close to the muon tracks.

To reject peaking backgrounds, neural networks that discriminate between muons and hadrons are trained. The neural network selection is optimised for the $\mathcal{B}(B^0 \to \mu^+\mu^-)$ measurement, as due to its lower mass it is more affected by peaking backgrounds.

Simulation is used for each peaking background to estimate the number of events after selection, as well as the mass and BDT shape. The main backgrounds ($B^0_{(s)} \to h^+h^{(\prime)-}$, $B^0 \to \pi^-\mu^+\nu_\mu$, $B^0_s \to K^-\mu^+\nu_\mu$) are also controlled using a fit to data with the $h\mu$ mass hypothesis.

2.2 Signal calibration and normalisation

To convert a yield after selection into a branching fraction, other B decays with known branching fractions are used for normalisation. In addition, the mass and BDT shape of the signal have

[a] For B^0 mesons, no lifetime difference has been observed. It is expected to be $\mathcal{O}(10^{-3})$. The effective lifetime of the $B^0 \to \mu^+\mu^-$ decay is not measured due to the small lifetime difference and its small branching fraction.

[b] For reference, the B^0_s mass is 5367 MeV/c^2 and the B^0 mass is 5280 MeV/c^2.

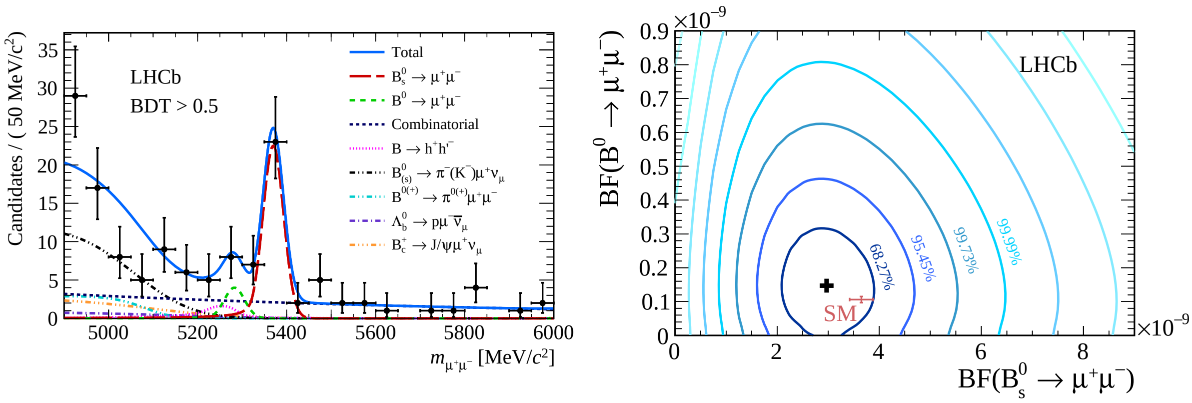

Figure 1 – (left) Mass distribution of selected $B^0_{(s)} \to \mu^+\mu^-$ candidates in the four most sensitive BDT bins. The result of the fit is overlaid. (right) The 2D confidence interval in $\mathcal{B}(B^0_s \to \mu^+\mu^-), \mathcal{B}(B^0 \to \mu^+\mu^-)$ from this measurement.

to be calibrated.

The normalisation is performed with $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ and $B^0 \to K^+\pi^-$ decays. The two muons in the final state for the $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ decay are detected similarly to the muons from $B^0_{(s)} \to \mu^+\mu^-$. The efficiency to detect the extra kaon track in the final state is corrected for using simulation. For $B^0 \to K^+\pi^-$, the kinematics are very similar to $B^0_{(s)} \to \mu^+\mu^-$, which means the same BDT can be applied. However, the trigger and PID selection for hadrons and muons are very different, and are corrected for using a data-driven approach.

For the BDT calibration, fits to $B^0 \to K^+\pi^-$ decays per BDT bin are used to determine the $B^0_{(s)} \to \mu^+\mu^-$ BDT distributions. The same corrections as for the normalisation are applied, and the BDT distribution is found to be consistent with simulation.

2.3 Branching fraction fit and results

To measure both branching fractions, a simultaneous maximum likelihood fit for $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$ is performed to $m_{\mu^+\mu^-}$ in bins of BDT per dataset (Run 1 and Run 2). The only free nuisance parameters are the combinatorial background shape parameters and yields. The shape is the same for all BDT bins in a dataset, the yields are free for each BDT bin.

From the fit, the $B^0_s \to \mu^+\mu^-$ decay is observed with a statistical significance of 7.8 standard deviations, and its branching fraction is measured to be $\mathcal{B}(B^0_s \to \mu^+\mu^-) = (3.0 \pm 0.6^{+0.3}_{-0.2}) \times 10^{-9}$, where the first uncertainty is statistical and the second systematic. The statistical uncertainty is obtained by repeating the fit with all nuisance parameters fixed. No significant excess of $B^0 \to \mu^+\mu^-$ decays is observed, and a 95% confidence level upper limit is determined, $\mathcal{B}(B^0 \to \mu^+\mu^-) < 3.4 \times 10^{-10}$.

3 Effective lifetime measurement of $B^0_s \to \mu^+\mu^-$

In contrast to the branching fraction, the effective lifetime is measured for the $B^0_s \to \mu^+\mu^-$ decay only. Therefore, the selection and fitting procedure are optimised independently.

First, events with $m_{\mu^+\mu^-} < 5320$ MeV, and thus all peaking backgrounds, as well as $B^0 \to \mu^+\mu^-$ events, are cut away. Second, the PID selection is loosened, as the mass cut makes the measurement less sensitive to peaking backgrounds. Third, a cut is applied on the BDT, instead of using the BDT to divide the sample in bins.

With the *sPlot* method [9], a mass fit is used to background-subtract the decay time distribution. The mass fit contains just two shapes ($B^0_s \to \mu^+\mu^-$ and combinatorial background). The

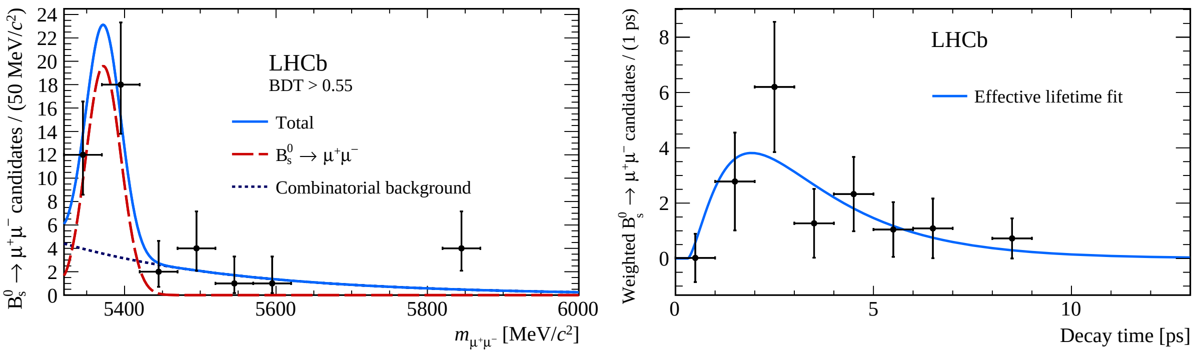

Figure 2 – (left) Mass distribution of $B_s^0 \to \mu^+\mu^-$ candidates. The result of the fit is overlaid. (right) Background-subtracted decay-time distribution with the fit result superimposed.

background-subtracted decay time distribution is fitted with an exponential, convoluted with an acceptance function.

The effective lifetime is measured for the first time and found to be $\tau(B_s^0 \to \mu^+\mu^-) = 2.04 \pm 0.44 \pm 0.05$ ps. This measurement is consistent with the $\mathcal{A}_{\Delta\Gamma} = +1(-1)$ hypothesis at 1.0 (1.4) standard deviations. The current experimental precision does not set a strong constraint, but shows the potential LHCb will have to constrain NP using the effective lifetime.

4 Conclusions

The $B_s^0 \to \mu^+\mu^-$ decay is observed with a significance of 7.8 standard deviations, and its branching fraction is measured to be $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.0 \pm 0.6^{+0.3}_{-0.2}) \times 10^{-9}$, where the first uncertainty is statistical and the second systematic. In addition, the first measurement of the $B_s^0 \to \mu^+\mu^-$ effective lifetime is performed: $\tau(B_s^0 \to \mu^+\mu^-) = 2.04 \pm 0.44 \pm 0.05$ ps. No significant excess of $B^0 \to \mu^+\mu^-$ decays is observed, and a 95% confidence level upper limit is determined, $\mathcal{B}(B^0 \to \mu^+\mu^-) < 3.4 \times 10^{-10}$. All results are consistent with the SM and constrain New Physics in $b \to s\ell\ell$ processes. Already, these results have sparked interesting discussions [6,10,11,12].

Acknowledgments

The author wishes to acknowledge the financial support from NWO and thank the organisers of Moriond EW 2017 for an interesting and captivating conference.

References

1. W. Altmannshofer and D. Straub, *Eur. Phys. J.* **C 75**, 382, (2015).
2. T. Hurth *et al.*, *Nucl. Phys.* **B 909**, 737-777, (2016).
3. S. Descotes-Genon *et al.*, *JHEP* **06**, 092, (2016).
4. S. Aoki *et al.*, *Eur. Phys. J.* **C 77**, 112, (2017).
5. C. Bobeth *et al.*, *Phys. Rev. Lett.* **112**, 101801, (2014).
6. R. Fleischer *et al.*, arXiv:1703.10160 (2017).
7. K. De Bruyn *et al.*, *Phys. Rev. Lett.* **109**, 041801, (2012).
8. R. Aaij *et al.*, arXiv:1703.05747 (2017).
9. M. Pivk and F. Le Diberder, *Nucl. Instrum. Methods* **A55**, 356-369,(2005).
10. W. Altmannshofer *et al.*, arXiv:1702.05498 (2017).
11. C. Bobeth *et al.*, arXiv:1703.04753 (2017).
12. C. Chiang *et al.*, arXiv:1703.06289 (2017).