EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)

LHCb-CONF-2020-002
CMS PAS BPH-20-003
ATLAS-CONF-2020-049

Combination of the ATLAS, CMS and LHCb results on the $B^0_{(s)} \to \mu^+\mu^-$ decays

The ATLAS, CMS and LHCb collaborations

ATLAS-CONF-2020-049
14 August 2020

Abstract

A combination of results on the rare $B^0_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays from the ATLAS, CMS, and LHCb experiments using data collected at the Large Hadron Collider between 2011 and 2016, is presented. The $B^0_s \to \mu^+\mu^-$ branching fraction is obtained to be $\left(2.69\,^{+\,0.37}_{-\,0.35}\right) \times 10^{-9}$ and the effective lifetime of the $B^0_s \to \mu^+\mu^-$ decay is measured to be $\tau_{B^0_s \to \mu^+\mu^-} = 1.91^{+0.37}_{-0.35}$ ps. An upper limit on the $B^0 \to \mu^+\mu^-$ branching fraction is evaluated to be $\mathcal{B}(B^0 \to \mu^+\mu^-) < 1.6\ (1.9) \times 10^{-10}$ at 90% (95%) confidence level. An upper limit on the ratio of the $B^0 \to \mu^+\mu^-$ and $B^0_s \to \mu^+\mu^-$ branching fractions is obtained to be 0.052 (0.060) at 90% (95%) confidence level.

1 Introduction

In this note a combination of the experimental searches for the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays at the ATLAS, CMS and LHCb experiments is presented. This combination is based on the analyses and data sets used for the previously published results detailed in Ref. [1] for ATLAS, Ref. [2] for CMS, and Ref. [3] for LHCb.

Measurements of low-energy processes can provide indirect constraints on new physics effects at energy scales beyond the range directly accessible at the LHC. This is particularly true for Flavour Changing Neutral Current (FCNC) processes, which are highly suppressed in the Standard Model (SM) and can only occur through higher-order Feynman diagrams. The $B_{(s)}^0 \to \mu^+\mu^-$ decays are among the most sensitive FCNC processes, owing to the precision of the calculation of their rates in the SM and their clean experimental signature [4–8]. In the SM, the decays $B_{(s)}^0 \to \mu^+\mu^-$ proceed only via loop diagrams and are also helicity suppressed, leading to very small expected branching fractions. Recently, theoretical uncertainties in the calculation of these branching fractions have been reduced as a result of progress in lattice QCD [9–13], in the calculation of electroweak effects at next-to-leading order [5], and in the calculation of QCD effects at next-to-next-to-leading order [6]. Enhanced electromagnetic contributions from virtual photon exchange have been included in the calculation, showing larger corrections than previously assumed in the theoretical uncertainties [7, 8], albeit still minor with respect to the total theoretical uncertainties. The most up-to-date SM predictions for the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ branching fractions are calculated in Ref. [8] and yield

$$\begin{array}{rcl} \mathcal{B}(B_s^0 \to \mu^+\mu^-) & = & (3.66 \pm 0.14) \times 10^{-9} \quad \text{and} \\ \mathcal{B}(B^0 \to \mu^+\mu^-) & = & (1.03 \pm 0.05) \times 10^{-10}\,. \end{array} \tag{1}$$

These branching fractions are CP-averaged and time-integrated, and account for the finite width difference measured in the B_s^0 system in the comparison between theory and experiments (see Refs. [14, 15]). They can be readily compared with experimental results. Similar predictions are obtained using the relation between $B_{(s)}^0 \to \mu^+\mu^-$ decays and $\Delta M_{d(s)}$, the mass difference of the $B_{(s)}^0$ mass eigenstates [16, 17]. Using the same input values as in Ref. [8] the ratio $\mathcal{R}$ of the $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$ branching fractions can be estimated to be: $\mathcal{R} = 0.0281 \pm 0.0016$. It is worth noting that this prediction has smaller uncertainty than the individual branching fractions, owing to the cancellation of several common factors. Contributions from new processes or new heavy particles can enhance or suppress the values of the individual branching fractions or modify their ratio. However, $\mathcal{R}$ has the same value in the SM and in all theories obeying the Minimal Flavour Violation hypothesis [18, 19], and as such it specifically probes the latter.

For the $B_s^0 \to \mu^+\mu^-$ decay, it is also interesting to measure its effective lifetime, which

is the mean lifetime of the B_s^0 meson in this decay. This can be expressed as

$$\tau_{B_s^0\to\mu^+\mu^-} \equiv \frac{\int_0^\infty t\,\langle\Gamma\,(B_s^0\to\mu^+\mu^-)\rangle\,dt}{\int_0^\infty \langle\Gamma\,(B_s^0\to\mu^+\mu^-)\rangle\,dt} \tag{2}$$

$$= \frac{\tau_{B_s^0}}{1-y_s^2}\left[\frac{1+2\mathcal{A}_{\Delta\Gamma}y_s+y_s^2}{1+\mathcal{A}_{\Delta\Gamma}y_s}\right], \tag{3}$$

where t is the proper decay time of the B_s^0 meson, Γ is its time-dependent partial-width, $\tau_{B_s^0}$ is the B_s^0 lifetime and $\Delta\Gamma_s$ is the width difference between the light and the heavy mass-eigenstates. The parameters y_s and $\mathcal{A}_{\Delta\Gamma}$ are defined as

$$y_s \equiv \frac{\Delta\Gamma_s}{2\Gamma_s}, \qquad \mathcal{A}_{\Delta\Gamma} \equiv \frac{R_H^{\mu^+\mu^-}-R_L^{\mu^+\mu^-}}{R_H^{\mu^+\mu^-}+R_L^{\mu^+\mu^-}}, \tag{4}$$

where $R_H^{\mu^+\mu^-}$ and $R_L^{\mu^+\mu^-}$ are the contributions of the heavy and light mass eigenstates of the B_s^0 system to the untagged $B_s^0\to\mu^+\mu^-$ decay rate. The $\mu^+\mu^-$ final state is CP odd. Hence in the SM $\mathcal{A}_{\Delta\Gamma}=+1$ and the effective lifetime corresponds to the lifetime of the heavy B_s^0 mass eigenstate, which is measured to be $\tau_{B_s^0\to\mu^+\mu^-}=1.609\pm0.010\,\text{ps}$ [20]. As discussed in Ref. [14], $\mathcal{A}_{\Delta\Gamma}$ can receive contributions from new physics effects, particularly from scalar and pseudoscalar operators, even in the case that the branching fractions are not modified.

The analyses of the three experiments are very similar. Starting from data triggered mostly by muons, pairs of two well-identified oppositely charged muons are combined to form a vertex displaced from the primary pp interaction. The resulting $B_{(s)}^0\to\mu^+\mu^-$ candidates are further selected with additional criteria based on kinematics and geometry, as well as isolation. The final search is performed in bins of multivariate discriminators trained to distinguish signal from combinatorial background based on a variety of observables that are not correlated with the dimuon mass. The signal yields are extracted from a simultaneous fit to the dimuon mass distributions in bins of the output of the multivariate discriminator. In addition to the residual combinatorial component, backgrounds are expected and included in the final fit from hadronic two-body decays and semileptonic decays, where the hadrons have been misidentified as muons. The yield of the $B_s^0\to\mu^+\mu^-$ candidates is converted into its branching fraction by normalising to the $B^+\to J/\psi\,K^+$ decay, as

$$\mathcal{B}(B_s^0\to\mu^+\mu^-) = \frac{f_d}{f_s}\frac{\varepsilon_{B^+\to J/\psi\,K^+}}{\varepsilon_{B_s^0\to\mu^+\mu^-}}\frac{N_{B_s^0\to\mu^+\mu^-}}{N_{B^+\to J/\psi\,K^+}}\mathcal{B}(B^+\to J/\psi\,K^+), \tag{5}$$

where N is the yield, ε the efficiency, and $\mathcal{B}$ the branching fraction of the corresponding channel. The term $\mathcal{B}(B^+\to J/\psi\,K^+)$ includes the $J/\psi\to\mu^+\mu^-$ branching fraction. The term f_d/f_s represents the ratio of fragmentation fractions of a b quark to a B^0 meson over the one to a B_s^0 meson [21]. Here and in the following it is assumed that the fragmentation fractions of B^0 and B^+ mesons are equal due to isospin symmetry to a good accuracy [22]. All experiments use the $B^+\to J/\psi\,K^+$ decay as normalisation channel, while LHCb experiment also employs the $B^0\to K^+\pi^-$ decay as additional normalisation mode. The

normalisation of the $B^0 \to \mu^+\mu^-$ proceeds similarly but without the f_d/f_s term. The efficiencies are determined using simulation and corrected using data-driven methods.

The ratio of fragmentation fractions f_d/f_s is known with limited precision and is the largest systematic uncertainty on the $B_s^0 \to \mu^+\mu^-$ branching fraction. Being common among the three experiments, a dedicated treatment is needed to take this correlation into account, as explained in Section 3.

2 Inputs to the combination

The results obtained by the three experiments and used for this combination are detailed in the following.

The ATLAS results are extracted from data samples of pp collisions corresponding to integrated luminosities of 25 fb^{-1} at a centre-of-mass energy $\sqrt{s} = 7$ and $\sqrt{s} = 8$ TeV collected in 2011 and 2012, and 26.3 fb^{-1} at $\sqrt{s} = 13$ TeV collected in 2015 and 2016. The ATLAS analysis yields [1]

$$\begin{aligned} \mathcal{B}(B_s^0 \to \mu^+\mu^-) &= \left(2.8\,^{+0.8}_{-0.7}\right) \times 10^{-9}, \\ \mathcal{B}(B^0 \to \mu^+\mu^-) &= (-1.9 \pm 1.6) \times 10^{-10}, \end{aligned} \tag{6}$$

with a significance for the $B_s^0 \to \mu^+\mu^-$ signal of 4.6 standard deviations (σ). A 95 % confidence level (CL) upper limit for the $B^0 \to \mu^+\mu^-$ signal is $\mathcal{B}(B^0 \to \mu^+\mu^-) < 2.1 \times 10^{-10}$, as obtained with the Neyman procedure [23], where the systematic uncertainties are included following the procedure described in Ref. [24]. Profiled likelihood contours are also computed using the procedure outlined in Ref. [25] and are found to be compatible with the Neyman contours. Therefore, to simplify the process, the profiled likelihood contours are used in the combination between experiments.

The CMS results are based on a data sample of pp collisions corresponding to integrated luminosities of 5 fb^{-1} at $\sqrt{s} = 7$ TeV, 20 fb^{-1} at $\sqrt{s} = 8$ TeV, and 36 fb^{-1} at $\sqrt{s} = 13$ TeV, collected during 2011, 2012, and 2016. The measured branching fractions are [2]

$$\begin{aligned} \mathcal{B}(B_s^0 \to \mu^+\mu^-) &= \left[2.9\,^{+0.7}_{-0.6}(\text{exp}) \pm 0.2(\text{frag})\right] \times 10^{-9}, \\ \mathcal{B}(B^0 \to \mu^+\mu^-) &= \left(0.8\,^{+1.4}_{-1.3}\right) \times 10^{-10}, \end{aligned} \tag{7}$$

with a signal significance of $5.6\,\sigma$ and $1.0\,\sigma$, respectively. In the quoted $B_s^0 \to \mu^+\mu^-$ branching fraction measurement, the first uncertainty combines the experimental statistical and systematic uncertainties on the measurement, while the second is due to the uncertainty in the ratio of fragmentation fractions f_d/f_s. A 95 % CL upper limit for the $B^0 \to \mu^+\mu^-$ decay is $\mathcal{B}(B^0 \to \mu^+\mu^-) < 3.6 \times 10^{-10}$, as obtained with the CLs method [25, 26]. The effective lifetime of the $B_s^0 \to \mu^+\mu^-$ decay was measured to be

$$\tau_{B_s^0 \to \mu^+\mu^-} = 1.70^{+0.60}_{-0.43} \pm 0.09\,\text{ps}, \tag{8}$$

where the first uncertainty is statistical and the second systematic.

The LHCb results are based on a data sample of pp collisions corresponding to an integrated luminosity of $1\,\text{fb}^{-1}$ recorded in 2011 at $\sqrt{s} = 7\,\text{TeV}$, $2\,\text{fb}^{-1}$ recorded in 2012 at $\sqrt{s} = 8\,\text{TeV}$ and $1.4\,\text{fb}^{-1}$ recorded in 2015 and 2016 at $\sqrt{s} = 13\,\text{TeV}$. The measured branching fractions are [3]

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \left(3.0 \pm 0.6\,^{+0.3}_{-0.2}\right) \times 10^{-9}, \quad (9)$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = \left(1.5\,^{+1.2\,+0.2}_{-1.0\,-0.1}\right) \times 10^{-10}, \quad (10)$$

where the first uncertainty is statistical and the second systematic, with signal significances of $7.8\,\sigma$ and $1.6\,\sigma$, respectively, and an upper limit $\mathcal{B}(B^0 \to \mu^+\mu^-) < 3.4 \times 10^{-10}$ at 95 % CL, as obtained with the CLs method [25, 26]. In addition, the effective lifetime of the $B_s^0 \to \mu^+\mu^-$ decay was measured to be

$$\tau_{B_s^0 \to \mu^+\mu^-} = 2.04 \pm 0.44 \pm 0.05\,\text{ps}. \quad (11)$$

where the first uncertainty is statistical and the second systematic. Results from the three experiments assume $\mathcal{A}_{\Delta\Gamma} = +1$ in the calculation of the efficiencies for the $B_s^0 \to \mu^+\mu^-$ decay and thus for its branching fractions, and they are combined in the following under this hypothesis. Corrections are given by the single experiments for two possible alternative hypotheses of $\mathcal{A}_{\Delta\Gamma} = 0, -1$. For LHCb, the increase in the branching fraction of $B_s^0 \to \mu^+\mu^-$ is of 4.6% (10.9%) for $\mathcal{A}_{\Delta\Gamma} = 0$ (-1), while for ATLAS, the increase is of 3.6% (7.8%) for $\mathcal{A}_{\Delta\Gamma} = 0$ (-1). In CMS this effect is not taken as a separate correction, but it is absorbed into the overall systematic uncertainty in the efficiency, with comparable magnitude.

3 Combination

In order to correctly take into account the correlation between $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$, and to allow the upper limit evaluation of $\mathcal{B}(B^0 \to \mu^+\mu^-)$, the combination is performed using the binned two-dimensional profile likelihoods obtained by each experiment from their fit to the dimuon invariant mass distributions. Each likelihood assumes the lifetime for the $B_s^0 \to \mu^+\mu^-$ decay to be the SM one, *i.e.* $\mathcal{A}_{\Delta\Gamma} = +1$. The uncertainties on the current measurements are dominated by the statistical component. The systematic uncertainties affecting the three measurements are considered to be independent, therefore the three likelihoods are profiled separately with respect to the nuisance parameters. The only exception is the ratio of the fragmentation fractions f_d/f_s that is a common nuisance parameter and represents the only relevant source of correlation among the three experiments. In order to take this into account, this factor was profiled separately in each likelihood, retaining its uncertainty in only one of the three experiments. In practice the f_d/f_s uncertainty was maintained in the LHCb likelihood. The effect of the correlation of f_d/f_s is tested comparing results of this combination obtained with and without the f_d/f_s uncertainty in the ATLAS and CMS results and found to be negligible, being limited by the statistical uncertainties. The effect of the dependence of f_d/f_s on the

transverse momentum was also considered and found to be within the already assigned uncertainties.

Each likelihood is fitted with an analytical function in the region of the $(\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ – $\mathcal{B}(B^0 \to \mu^+\mu^-))$ plane where both branching fractions are positive. However, for each experiment, the original likelihood is unconstrained, such that the fit to the dimuon invariant mass distribution is allowed to return all possible values (including negative values) of the two branching fractions. Since the number of observed $B^0_{(s)} \to \mu^+\mu^-$ candidates is small, the shape of the likelihood cannot be represented by a Gaussian distribution. To take into account the strong likelihood asymmetry, a variable-width Gaussian [27] has been used. This function describes asymmetric likelihoods (which means asymmetric uncertainties) and also possible correlation between the two observables which in this analysis is predominantly due to the overlap between the populations of the two mass peaks resulting from the finite dimuon mass resolution. The analytical function is found to be in good agreement with the measured binned likelihood for each experiment.

The three binned log-likelihoods are then summed and fitted using a two-dimensional variable-width Gaussian. The maximum is used to evaluate the central values and the uncertainties of the branching fractions for the two processes, yielding:

$$\begin{array}{rcl} \mathcal{B}(B_s^0 \to \mu^+\mu^-) & = & \left(2.69\,^{+\,0.37}_{-\,0.35}\right) \times 10^{-9} \text{ and} \\ \mathcal{B}(B^0 \to \mu^+\mu^-) & = & (0.6 \pm 0.7) \times 10^{-10}. \end{array} \tag{12}$$

The two-dimensional likelihood contours of the results for the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays for the three experiments together with their combination are shown in Fig. 1: likelihood contours correspond to the values of $-2\Delta\ln\mathcal{L} = 2.3$, 6.2, 11.8, 19.3, and 30.2. In a bidimensional Gaussian approximation, these contours correspond to 1 to 5 sigma levels. Since the measured significance for the $B^0 \to \mu^+\mu^-$ decay is lower than 1σ, an upper limit is evaluated from the one-dimensional negative log-likelihood. The upper limit on $\mathcal{B}(B^0 \to \mu^+\mu^-)$ is 1.6 (1.9) $\times 10^{-10}$ at 90% (95%) CL. The upper limit is computed under the hypothesis that $\mathcal{B}(B^0 \to \mu^+\mu^-)$ is positive, by renormalising the likelihood in the region where $\mathcal{B}(B^0 \to \mu^+\mu^-)$ is positive and by integrating it up to the quantiles of interest. The negative log-likelihoods from the combined analysis for both $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$ are shown in Fig. 2. The central value of the $B_s^0 \to \mu^+\mu^-$ branching fraction is lower than any single experiment's results: this is expected given the anti-correlation of the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ branching fractions present, at different levels, in the three experiments [28].

When including the theoretical uncertainties, the one-dimensional compatibility with the SM is estimated to be 2.4σ for the $B_s^0 \to \mu^+\mu^-$ and 0.64σ for the $B^0 \to \mu^+\mu^-$, while the two-dimensional compatibility with the SM point is estimated to be of 2.1σ. These values are derived from the $-2\Delta\ln\mathcal{L}$ assuming Wilks' theorem [29].

Starting from the two-dimensional likelihood, a profile likelihood of the ratio $\mathcal{R}$ of the $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$ branching fractions is obtained by computing the minimum negative log-likelihood value in the full two-dimensional plane for each $\mathcal{R}$ value. With this approach, the correlation between the two branching fractions is properly taken into

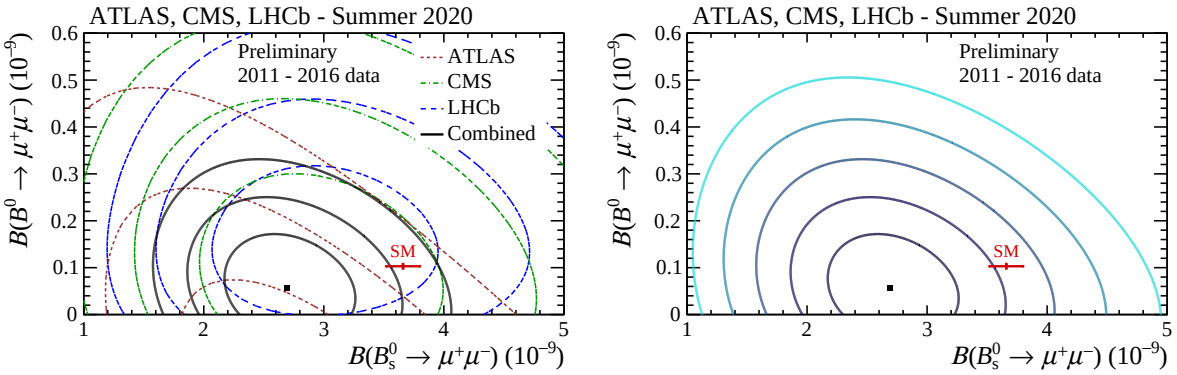

Figure 1: In the left-hand plot, the two-dimensional likelihood contours of the results for the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays for the three experiments are shown together with their combination. The dataset used was collected from 2011 to 2016. The red dashed line represents the ATLAS experiment, the green dot-dashed line the CMS experiment, the blue long-dashed line the LHCb experiment and the continuous line their combination. For each experiment and for the combination, likelihood contours correspond to the values of $-2\Delta\ln\mathcal{L} = 2.3$, 6.2, and 11.8, respectively. In the right-hand plot, the combination of the three experiments is shown with contours of different shades. Likelihood contours correspond to the values of $-2\Delta\ln\mathcal{L} = 2.3$, 6.2, 11.8, 19.3, and 30.2, represented in order by darkest to less dark colour. In both plots, the red point shows the SM predictions with their uncertainties. The published results from the three experiments are detailed in Ref. [1–3].

account. The resulting curve is shown in Fig. 3. The value of the ratio is determined to be

$$\mathcal{R} = 0.021^{+0.030}_{-0.025} \tag{13}$$

and its upper limit at 90% (95)% CL is $\mathcal{R} < 0.052$ (0.060). The upper limit is computed in the same manner as for $\mathcal{B}(B^0 \to \mu^+\mu^-)$, by integrating the likelihood only in the positive region.

The CMS and LHCb experiments also measured the effective lifetime of the observed $B_s^0 \to \mu^+\mu^-$ candidates. The LHCb $B_s^0 \to \mu^+\mu^-$ effective lifetime is measured from a fit to the background-subtracted decay-time distribution of signal candidates. The CMS measurement is determined with a two-dimensional likelihood fit to the proper decay time and dimuon invariant mass; the model introduced in the likelihood fit adopts the per-event decay time resolution as a conditional parameter in the resolution model. For both experiments, the measurement is fully dominated by its statistical uncertainty, hence the two results are uncorrelated. Two variable-width Gaussian likelihoods are used to describe the CMS and LHCb original likelihoods and the value of $-2\Delta\ln\mathcal{L}$ obtained from these functions (shown in Fig. 4) is then minimised to obtain the combined value and the

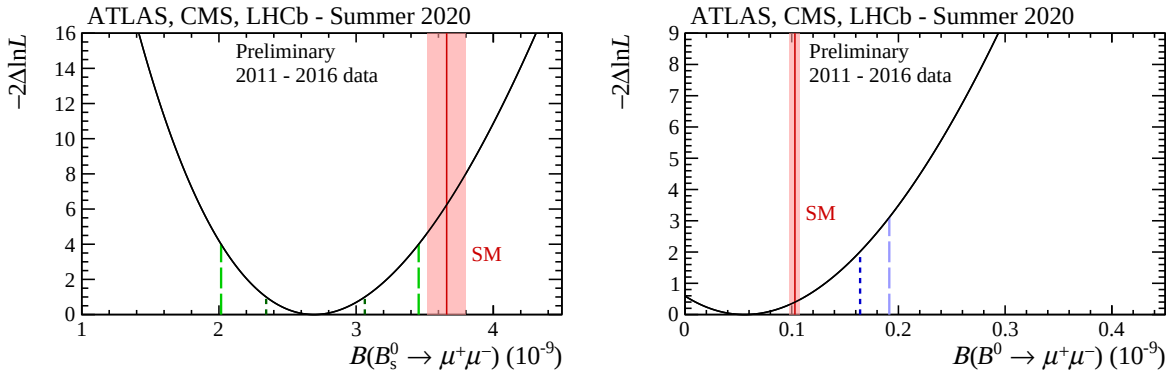

Figure 2: Value of $-2\Delta\ln\mathcal{L}$ for $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ (left) and $\mathcal{B}(B^0 \to \mu^+\mu^-)$ (right), shown in both as solid black line. In the left-hand plot, the dark (light) green dashed lines represent the 1σ (2σ) interval. In the right-hand plot, the dark (light) blue dashed lines represent the 90% (95%) CL. In both plots, the red solid band shows the SM prediction with its uncertainty. The published results from the three experiments are detailed in Ref. [1–3].

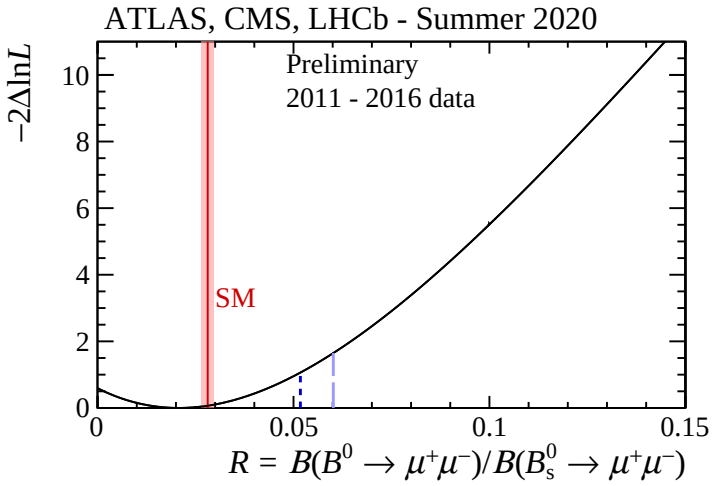

Figure 3: Value of $-2\Delta\ln\mathcal{L}$ for the ratio of the $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$ branching fractions, $\mathcal{R}$, shown as solid black line. The light (dark) blue dashed line represents the 90% (95%) CL and the red solid band shows the SM prediction with its uncertainty. The published results from the three experiments are detailed in Ref. [1–3].

68% CL interval (found where $-2\Delta\ln\mathcal{L} = 1$), which reads

$$\tau_{B_s^0 \to \mu^+\mu^-} = 1.91^{+0.37}_{-0.35}\,\text{ps}. \tag{14}$$

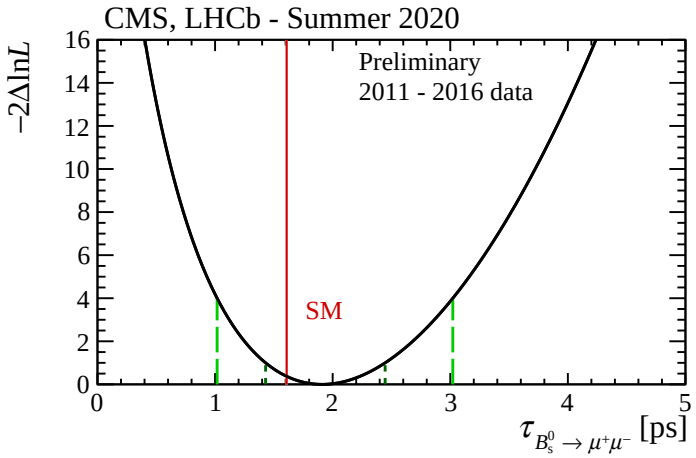

Figure 4: Value of $-2\Delta\ln\mathcal{L}$ for the combination of CMS and LHCb measurements [2, 3] of the $B_s^0 \to \mu^+\mu^-$ effective lifetime, shown as solid black line. The dark and light green dashed lines represent the intervals corresponding to $-2\Delta\ln\mathcal{L}$ =1 and 4, respectively, and the red solid band shows the SM prediction with its uncertainty.

4 Conclusions

In summary, the results of the ATLAS, CMS, and LHCb experiments on the $B^0_{(s)} \to \mu^+\mu^-$ decays obtained from the data collected between 2011 and 2016 have been combined. The $B_s^0 \to \mu^+\mu^-$ branching fraction is measured to be

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \left(2.69\,^{+0.37}_{-0.35}\right) \times 10^{-9}$$

assuming $\mathcal{A}_{\Delta\Gamma} = +1$, while an upper limit on the $B^0 \to \mu^+\mu^-$ branching fraction is set at

$$\begin{aligned} \mathcal{B}(B^0 \to \mu^+\mu^-) &< 1.6 \times 10^{-10} \text{ at } 90\% \text{ CL, and} \\ \mathcal{B}(B^0 \to \mu^+\mu^-) &< 1.9 \times 10^{-10} \text{ at } 95\% \text{ CL.} \end{aligned}$$

An upper limit on the ratio of the $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$ branching fractions is also obtained to be $\mathcal{R} < 0.052$ (0.060) at 90% (95%) confidence level.

The results are compatible with the SM predictions within 2.1 standard deviations in the two-dimensional plane of the branching fractions, when the theoretical uncertainties are included.

The effective lifetimes of the $B_s^0 \to \mu^+\mu^-$ decay, measured by the CMS and LHCb experiments, are combined yielding

$$\tau_{B_s^0 \to \mu^+\mu^-} = 1.91^{+0.37}_{-0.35}\,\text{ps}.$$

These results for the branching fractions and lifetime are the most precise to date.

References

[1] ATLAS Collaboration, M. Aaboud *et al.*, *Study of the rare decays of B_s^0 and B^0 mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector*, JHEP **04** (2019) 098, arXiv:1812.03017.

[2] CMS Collaboration, A. M. Sirunyan *et al.*, *Measurement of properties of $B_{\rm s}^0 \to \mu^+\mu^-$ decays and search for $B^0 \to \mu^+\mu^-$ with the CMS experiment*, JHEP **04** (2020) 188, arXiv:1910.12127.

[3] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays*, Phys. Rev. Lett. **118** (2017) 191801, arXiv:1703.05747.

[4] C. Bobeth *et al.*, *$B_{s,d} \to \ell^+\ell^-$ in the Standard Model with reduced theoretical uncertainty*, Phys. Rev. Lett. **112** (2014) 101801, arXiv:1311.0903.

[5] C. Bobeth, M. Gorbahn, and E. Stamou, *Electroweak Corrections to $B_{s,d} \to \ell^+\ell^-$*, Phys. Rev. D **89** (2014) 034023, arXiv:1311.1348.

[6] T. Hermann, M. Misiak, and M. Steinhauser, *Three-loop QCD corrections to $B_s \to \mu^+\mu^-$*, JHEP **12** (2013) 097, arXiv:1311.1347.

[7] M. Beneke, C. Bobeth, and R. Szafron, *Enhanced electromagnetic correction to the rare B-meson decay $B_{s,d} \to \mu^+\mu^-$*, Phys. Rev. Lett. **120** (2018) 011801, arXiv:1708.09152.

[8] M. Beneke, C. Bobeth, and R. Szafron, *Power-enhanced leading-logarithmic QED corrections to $B_q \to \mu^+\mu^-$*, JHEP **10** (2019) 232, arXiv:1908.07011.

[9] Flavour Lattice Averaging Group, S. Aoki *et al.*, *FLAG Review 2019: Flavour Lattice Averaging Group (FLAG)*, Eur. Phys. J. C **80** (2020) 113, arXiv:1902.08191.

[10] Fermilab Lattice and MILC, A. Bazavov *et al.*, *B- and D-meson leptonic decay constants from four-flavor lattice QCD*, Phys. Rev. D **98** (2018) 074512, arXiv:1712.09262.

[11] ETM Collaboration, A. Bussone *et al.*, *Mass of the b quark and B meson decay constants from N_f=2+1+1 twisted-mass lattice QCD*, Phys. Rev. D **93** (2016) 114505, arXiv:1603.04306.

[12] HPQCD Collaboration, R. J. Dowdall *et al.*, *B-meson decay constants from improved lattice nonrelativistic QCD with physical u, d, s, and c quarks*, Phys. Rev. Lett. **110** (2013) 222003, arXiv:1302.2644.

[13] C. Hughes, C. T. H. Davies, and C. J. Monahan, *New methods for B meson decay constants and form factors from lattice NRQCD*, Phys. Rev. D **97** (2018) 054509, arXiv:1711.09981.

[14] K. De Bruyn *et al.*, *Probing new physics via the $B_s^0 \to \mu^+\mu^-$ effective lifetime*, Phys. Rev. Lett. **109** (2012) 041801, arXiv:1204.1737.

[15] K. De Bruyn *et al.*, *Branching Ratio Measurements of B_s^0 Decays*, Phys. Rev. **D86** (2012) 014027, arXiv:1204.1735.

[16] A. J. Buras, *Relations between Δ M(s, d) and B(s, d) $\to \mu\bar{\mu}$ in models with minimal flavor violation*, Phys. Lett. B **566** (2003) 115, arXiv:hep-ph/0303060.

[17] D. King, A. Lenz, and T. Rauh, *B_s mixing observables and $|V_{td}/V_{ts}|$ from sum rules*, JHEP **05** (2019) 034, arXiv:1904.00940.

[18] G. D'Ambrosio, G. F. Giudice, G. Isidori, and A. Strumia, *Minimal flavor violation: An effective field theory approach*, Nucl. Phys. B **645** (2002) 155, arXiv:hep-ph/0207036.

[19] A. J. Buras, *Minimal flavor violation*, Acta Phys. Polon. B **34** (2003) 5615, arXiv:hep-ph/0310208.

[20] Heavy Flavour Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. C **77** (2017) 895, arXiv:1612.07233.

[21] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286, f_s/f_d value updated in LHCb-CONF-2013-011.

[22] Particle Data Group, M. Tanabashi *et al.*, *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001.

[23] J. Neyman, *Outline of a theory of statistical estimation based on the classical theory of probability*, Phil. Trans. R. Soc. London A, **236** (1937) 333-380.

[24] J. Conrad, O. Botner, A. Hallgren, and C. Perez de los Heros, *Including systematic uncertainties in confidence interval construction for Poisson statistics*, Phys. Rev. D **67** (2003) 012002, arXiv:hep-ex/0202013.

[25] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv:1007.1727, [Erratum: Eur.Phys.J.C 73, 2501 (2013)].

[26] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693.

[27] R. Barlow, *Asymmetric statistical errors*, arXiv:physics/0406120, in Statistical Problems in Particle Physics, Astrophysics and Cosmology.

[28] L. Lyons, D. Gibaut, and P. Clifford, *How to Combine Correlated Estimates of a Single Physical Quantity*, Nucl. Instrum. Meth. A **270** (1988) 110.

[29] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Annals Math. Statist. **9** (1938) 60.