LHCb-INT-2013-012
March 15, 2013

Towards model independent results for $\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-$

M. Perrin-Terrin[1], G. Mancinelli[1], F. Dettori[2], J. Serrano[1].

[1] *CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France.*
[2] *Nikhef, National Institute for Subatomic Physics, Amsterdam, Netherlands.*

Abstract

We clarify the definition of quantities theoretically computed and experimentally measured for $\mathcal{B}(\mathsf{B}^0_{(\mathsf{s})} \to \mu^+\mu^-)$. We claim that, in order to compare these quantities, the analysis needs to include model dependent corrections. We provide these corrections for the analysis of the 2012 data, where they have been neglected. For $\mathsf{B}^0 \to \mu^+\mu^-$ the shift is $(-1.46 \pm 0.01)\%$ for the SM value. For $\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-$ SM result, under the assumption that all analysis sensitivity come from the last BDT bin, for which the correction is the largest, the shift is $(-8.85 \pm 0.06)\%$. Including the additional correction required to translate the $\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-$ proper time integrated branching ratio into the initial CP-average one, the total shift is $(-15.14 \pm 0.06)\%$. Therefore the SM theoretical predictions for the branching ratios $\mathcal{B}(\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-) = (3.23 \pm 0.27) \times 10^{-9}$ and $\mathcal{B}(\mathsf{B}^0 \to \mu^+\mu^-) = (1.07 \pm 0.10) \times 10^{-10}$ have now to be compared to the following experimental values:

$$\begin{aligned} \mathcal{B}(\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-) &= (2.7^{+1.3}_{-1.0}) \times 10^{-9}, \\ \mathcal{B}(\mathsf{B}^0 \to \mu^+\mu^-) &< 9.3 \times 10^{-10} \text{ at } 95\% \text{ C.L.} \end{aligned}$$

Contents

1 Introduction

As experiment are progressing towards the measurement of the branching ratio of the $B^0_s \to \mu^+\mu^-$decay, an effort was produced by the phenomenologist community [1, 2, 3] to clarify the definition used to compute theoretical predictions. The outcome of their works is the distinction between the quantities generally computed in the literature from the time integrated branching ratio measured experimentally. These two quantities are in fact identical for all the b-mesons decays but the B^0_s ones, due to the unique large decay width difference of its mass eigenstates. In this case, the difference between the two quantities is model dependent and for $B^0_s \to \mu^+\mu^-$ in the standard model (SM) is of the order of 9%.

The same effort needs to be done from the experimental side in order to clarify the definition of the measured quantity which is reported by [4, 5, 6, 7]. This note raises two points on the treatment of life-time dependent efficiencies which are essential to the branching ratio extraction. First we argue that the $B^0_s \to \mu^+\mu^-$ signal efficiency is model dependent. Since the factor computed by [1] to translate the CP-averaged branching ratio prediction at $t = 0$ into the time integrated one is also model dependent we propose to define a new single event sensitivity (also called normalisation factor) which will be model dependent and absorb these two factors. Second, we claim that the distribution of the expected number of signal events in each BDT bin in the LHCb analysis, obtained with the calibration method, needs model dependent corrections to account for the decay rate difference between the signal and control channels which has been so far neglected. Both corrections are provided in a model dependent way for the analysis of the 2012 data.

2 Branching Ratio Definitions

2.1 Proper Time Dependent Decay Rate

The decay time dependent untagged decay rate for $B^0_s \to \mu^+\mu^-$ is defined as in [1, 8], by:

$$\begin{aligned}
\Gamma\left(B_s \to \mu^+\mu^-\right) &= \Gamma\left(B^0_s(t) \to \mu^+\mu^-\right) + \Gamma\left(\overline{B}^0_s(t) \to \mu^+\mu^-\right) \\
&= R_H e^{-\Gamma_H t} + R_L e^{-\Gamma_L t} \\
&= \left(R_H + R_L\right) e^{-\Gamma_s t}\left[\cosh\frac{y_s t}{\tau_{B^0_s}} + \mathcal{A}_{\Delta\Gamma}\sinh\frac{y_s t}{\tau_{B^0_s}}\right],
\end{aligned}$$

where B_s stands for either B^0_s or $\overline{B}^0_s$, Γ_H and Γ_L are the decay widths of the heavy and light mass eigenstates, Γ_s is the average of Γ_H and Γ_L, y_s is defined as $\left(\Gamma_L - \Gamma_H\right)/\left(\Gamma_L + \Gamma_H\right)$, $\tau_{B^0_s}$ is the B^0_s mean life-time and $\mathcal{A}_{\Delta\Gamma} = \left(R_H - R_L\right)/\left(R_H + R_L\right)$.

Considering the CP-odd structure of the $B^0_s \to \mu^+\mu^-$ decay in the SM, $R_L = 0$ or equivalently $\mathcal{A}_{\Delta\Gamma}$=1.

2.2 Initial CP-averaged and Proper Time Integrated Branching Ratio

The theoretical predictions computed in the literature (e.g. [3] and references therein) for the branching fraction of $\mathsf{B_s} \to \mu^+\mu^-$ use a definition based on the CP-averaged decay rates in the flavour-eigenstate basis at $t = 0$:

$$\begin{aligned} \mathcal{B}^0(\mathsf{B_s} \to \mu^+\mu^-) &= \frac{\tau_{\mathsf{B_s^0}}}{2}\Gamma\left(\mathsf{B_s} \to \mu^+\mu^-\right)|_{t=0} \\ &\overset{\text{SM}}{=} \frac{\tau_{\mathsf{B_s^0}}}{2} R_H. \end{aligned} \tag{1}$$

In the SM, with $\tau_{\mathsf{B_s^0}} = (1.466 \pm 0.031)\,\text{ps}$ the prediction is [3]:

$$\mathcal{B}^0(\mathsf{B_s} \to \mu^+\mu^-) \overset{\text{SM}}{=} (3.23 \pm 0.27) \times 10^{-9}. \tag{2}$$

In the following, all corrections are computed with the latest world average experimental values [9] (see Appendix A). For $\tau_{\mathsf{B_s^0}}$ this value is 2.5% larger than the one used to compute the prediction. Therefore to be completely rigorous, the branching ratio prediction to which the experimental result is compared should be shifted by +2.5%.

In [1], a decay time integrated branching fraction is also defined by:

$$\begin{aligned} \langle\mathcal{B}(\mathsf{B_s} \to \mu^+\mu^-)\rangle &= \frac{1}{2}\int_0^\infty \Gamma\left(\mathsf{B_s} \to \mu^+\mu^-\right) \mathrm{dt} \\ &= \mathcal{B}^0(\mathsf{B_s} \to \mu^+\mu^-) \times \frac{1 + \mathcal{A}_{\Delta\Gamma} y_s}{1 - y_s^2} \\ &\overset{\text{SM}}{=} \frac{\mathcal{B}^0(\mathsf{B_s} \to \mu^+\mu^-)}{1 - y_s} \\ &\overset{\text{SM}}{=} 1.09 \times \mathcal{B}^0(\mathsf{B_s} \to \mu^+\mu^-)\ [2]. \end{aligned}$$

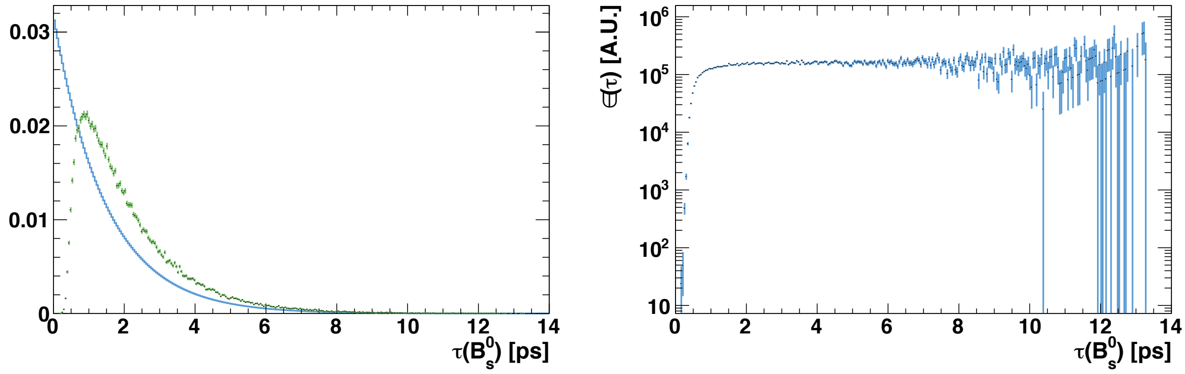
For the control channels, $\mathsf{B^+} \to \mathsf{J/\psi K^+}$ and $\mathsf{B^0} \to \mathsf{K^+}\pi^-$, these two definitions refer in fact to the same quantity since the width difference between $\mathsf{B_H^0}$ and $\mathsf{B_L^0}$ are null or negligible. In the following, the notation $\mathsf{B_s}$ is abandoned and the $\mathsf{B_s^0}$ notation refers to both $\mathsf{B_s^0}$ and $\overline{\mathsf{B}}{}_\mathsf{s}^0$.

3 Normalisation

This section describes how from the measured numbers of events the branching ratio is recovered. This conversion relies on the correct computation of efficiencies.

3.1 Efficiencies

The efficiencies depend on the decay time as the selection performed in the analysis uses variables correlated to the candidate decay time (impact parameters of the candidate



(a) (b)

Figure 1: (a) Proper time distribution before (blue) and after (green) the selection and (b) the resulting decay time dependent acceptance shape (arbitrary units).

tracks for instance). The decay time integrated efficiency is defined as:

$$\epsilon = \frac{\int_0^\infty \Gamma(t)\epsilon(t)\mathrm{d}t}{\int_0^\infty \Gamma(t)\mathrm{d}t}. \tag{3}$$

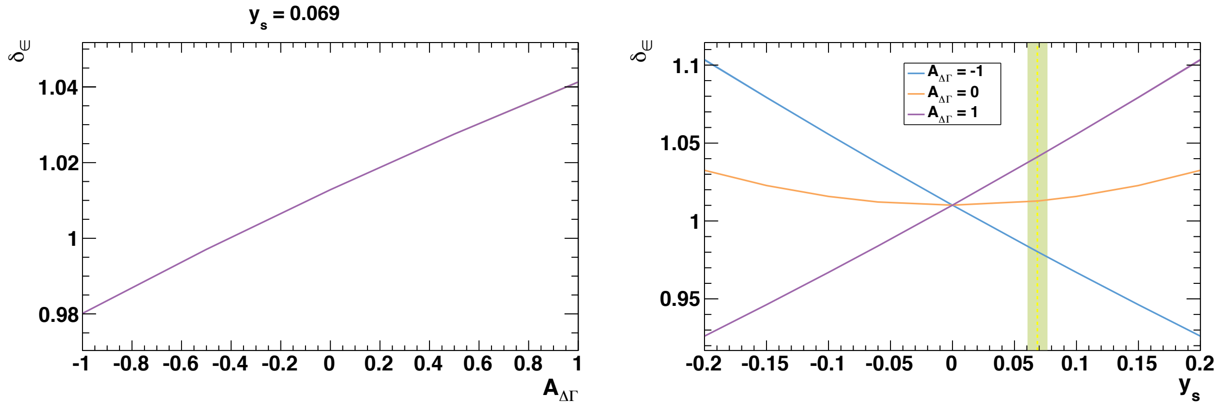
In the analysis all the efficiencies are computed using the simulated data sample either directly or indirectly to rescale some variable distributions correlated to the decay time. Therefore the decay rate assumptions (single exponential decay time distribution with $\tau = 1.469141\,\mathrm{ps}$) made to generate the simulated data induce a bias in all the signal efficiencies. The control channels are free of these biases as the assumptions made in the simulation are up to date with the latest experimental results of the decay rate (see Appendix A). For the signal the bias is expressed as:

$$\begin{aligned} \delta_\epsilon &= \frac{\epsilon^{\mathcal{A}_{\Delta\Gamma},y_s}}{\epsilon^{MC}} \\ &= \frac{\int_0^\infty (R_H e^{-\Gamma_H t} + R_L e^{-\Gamma_L t})\epsilon(t)\mathrm{d}t}{\int_0^\infty (R_H e^{-\Gamma_H t} + R_L e^{-\Gamma_L t})\mathrm{d}t} \times \frac{\int_0^\infty e^{-\Gamma_{MC} t}\mathrm{d}t}{\int_0^\infty e^{-\Gamma_{MC} t}\epsilon(t)\mathrm{d}t}. \end{aligned} \tag{4}$$

Figure 1 shows the decay time distributions before and after the selection and the resulting decay time dependent acceptance shape. From this shape the bias is computed and shown as a function of $\mathcal{A}_{\Delta\Gamma}$ and y_s in Figure 2.

For the SM ($\mathcal{A}_{\Delta\Gamma}$=1 and y_s=0.069 [9]) the bias is of $\delta_\epsilon - 1 = (+4.10 \pm 0.03)\%$. The uncertainties on this correction are only due to the statistics of the simulated data sample used to compute the decay time dependent efficiency, and they largely cancel since the numerator and denominator in Equation 5 are highly correlated.

The efficiency for $\mathrm{B}^0 \to \mu^+\mu^-$ is assumed, in the published analyses, to be identical to the one for $\mathrm{B}^0_s \to \mu^+\mu^-$. Hence the $\mathrm{B}^0 \to \mu^+\mu^-$ efficiency should also be corrected



(a) (b)

Figure 2: (a) Efficencies bias as a function of $\mathcal{A}_{\Delta\Gamma}$ for the experimental y_s value [9] and (b) as a function of y_s for different $\mathcal{A}_{\Delta\Gamma}$ values. The vertical green band is the experimental averaged y_s value and its uncertainty.

to account for the difference between the assumptions made in the simulation of the $\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-$ sample and the actual $\mathsf{B}^0 \to \mu^+\mu^-$ decay time distribution. This correction however is not model dependent since the width difference in the B^0 mode is negligible. The efficiency correction is estimated to be $\delta_\epsilon - 1 = (+1.48 \pm 0.01)\%$.

3.2 Normalisation Factor

The observables experimentally measured are the numbers of signal and control channels candidates, $N^{exp}(\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-)$, $N^{exp}(\mathsf{B}^+ \to \mathsf{J}/\psi\mathsf{K}^+)$ and $N^{exp}(\mathsf{B}^0 \to \mathsf{K}^+\pi^-)$. These numbers can be expressed as a function of the branching ratios defined in Equation 1:

$$
\begin{aligned}
N^{exp}_{sig} &= \frac{N^{tot}_{\mathsf{B}^0_\mathsf{s}}}{2} \int_0^\infty \Gamma(t)\epsilon^{sig}(t)\mathrm{d}t \\
&= N^{tot}_{\mathsf{B}^0_\mathsf{s}} \times \epsilon_{sig} \times \mathcal{B}^0(\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-) \times \frac{1 + \mathcal{A}_{\Delta\Gamma} y_s}{1 - y_s^2} \qquad (5) \\
N^{exp}_{cc} &= N^{tot}_{\mathsf{B}_{cc}} \times \epsilon_{cc} \times \mathcal{B}^0_{cc} \qquad (6)
\end{aligned}
$$

where $\epsilon(t)$ is the trigger, reconstruction, and selection efficiency as a function of the decay time, ϵ is the decay time integrated efficiency defined in Equation 3, $N^{tot}_{\mathsf{B}^0_\mathsf{s}}$ ($N^{tot}_{\mathsf{B}_{cc}}$) is the number of B^0_s (B_{cc}) and $\overline{\mathsf{B}}^0_\mathsf{s}$ ($\overline{\mathsf{B}}_{cc}$) mesons produced at t=0.[1] The sub-script *sig* stands for signal and *cc* for control channel.

From Equation 5 and Equation 6, the branching ratio of $\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-$ can be expressed from experimentally observable quantities and with the probability ratio that a b-quark

[1]Meson and anti-meson are supposed to be equally produced.

hadronises into a B_s^0 instead of the control channed B-meson f_s/f_d:

$$\begin{aligned} \mathcal{B}^0(\mathsf{B}_s^0 \to \mu^+\mu^-) &= \mathcal{B}_{cc}^0 \times \frac{N_{sig}^{exp}}{N_{cc}^{exp}} \times \frac{f_d}{f_s} \times \frac{\epsilon_{cc}}{\epsilon_{sig}(\mathcal{A}_{\Delta\Gamma}, y_s)} \times \frac{1-y_s^2}{1+\mathcal{A}_{\Delta\Gamma} y_s} \\ &\stackrel{\text{SM}}{=} \mathcal{B}_{cc}^0 \times \frac{N_{sig}^{exp}}{N_{cc}^{exp}} \times \frac{f_d}{f_s} \times \frac{\epsilon_{cc}}{\epsilon_{sig}^{SM}} \times (1-y_s) \end{aligned} \tag{7}$$

The term $(1-y_s^2)\,/\,(1+\mathcal{A}_{\Delta\Gamma} y_s)$ is the one derived by [1] to translate the theoretical prediction at $t=0$ into the decay time integrated one and simplifies in the SM to $(1-y_s)$ giving the 9% correction. This terms, as well as the signal efficiency, depends on $\mathcal{A}_{\Delta\Gamma}$ and y_s.

In all the published analysis the normalisation used was defined as:

$$\mathcal{B}^0(\mathsf{B}_s^0 \to \mu^+\mu^-) = \mathcal{B}_{cc}^0 \times \frac{N_{sig}^{exp}}{N_{cc}^{exp}} \times \frac{f_d}{f_s} \times \frac{\epsilon_{cc}}{\epsilon_{sig}^{MC}}, \tag{8}$$

and results were compared to the SM decay time integrated branching ratio. We advocate here a redefinition of the normalisation factor which would account for the efficiency correction and would convert the number of signal events directly to the branching ratio at $t=0$:

$$\alpha = \mathcal{B}_{cc}^0 \times \frac{1}{N_{cc}^{exp}} \times \frac{f_d}{f_s} \times \frac{\epsilon_{cc}}{\epsilon_{sig}(\mathcal{A}_{\Delta\Gamma}, y_s)} \times \frac{1-y_s^2}{1+\mathcal{A}_{\Delta\Gamma} y_s}. \tag{9}$$

Figure 3 gives the new normalisation factor for the 2012 analysis [7] as a function of $\mathcal{A}_{\Delta\Gamma}$ and y_s. Assuming the SM $\mathcal{A}_{\Delta\Gamma}$ value and taking y_s from [9], the fitted branching ratio should be shifted by $(-10.60 \pm 0.02)\%$ with respect to the published one [7]. The budget of this shift assign $(-3.94 \pm 0.03)\%$ to correct for the signal efficiency and -6.9% to translate to the initial CP-averaged branching ratio at. The SM prediction $\mathcal{B}(\mathsf{B}_s^0 \to \mu^+\mu^-) = (3.23 \pm 0.27) \times 10^{-9}$ should now be compared to the following experimental value:

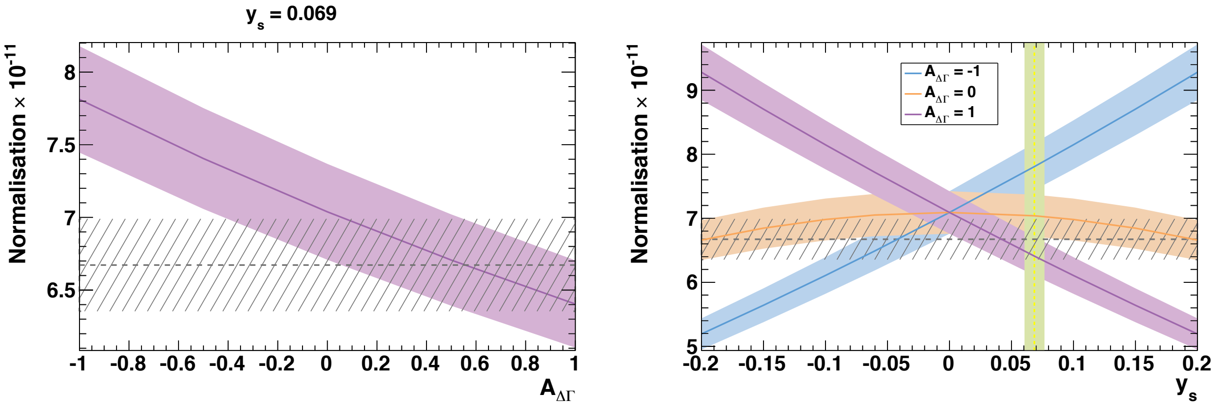
$$\mathcal{B}(\mathsf{B}_s^0 \to \mu^+\mu^-) = (2.9^{+1.3}_{-1.1}) \times 10^{-9}.$$

For $\mathsf{B}^0 \to \mu^+\mu^-$, the normalisation does not need to be redefined as long as the signal efficiency is corrected. For the 2012 analysis, this correction translates into a shift of $(-1.46 \pm 0.01)\%$ for $\mathcal{B}(\mathsf{B}^0 \to \mu^+\mu^-)$ and the upper limits. The latter becomes:

$$\mathcal{B}(\mathsf{B}^0 \to \mu^+\mu^-) < 9.3 \times 10^{-10}$$

4 BDT Calibration

This section describes how the signal BDT PDF is obtained from the control channels with the calibration method. We claim that some correction due to efficiencies computation must be accounted for in the calibration process.



(a) (b)

Figure 3: (a) New normalisation as a function of $\mathcal{A}_{\Delta\Gamma}$ for the experimental y_s value [9] and (b) as a function of y_s for different $\mathcal{A}_{\Delta\Gamma}$ values. The dashed area represents the normalisation of the 2012 analysis [7] corrected by $(1 - y_s)$.

4.1 Efficiencies

The decay time dependent efficiency in each BDT bin is obtained by:

$$\epsilon^i(t) = \frac{N^i(t)}{N^{gen}(t)}, \tag{10}$$

where $N^i(t)$ is the number of events with decay time t in BDT bin i and $N^{gen}(t)$ is the number of all events generated with decay time t.

This decay time dependent efficiency can be reasonably assumed to be identical between signal and control channels. In Figure 4, the decay time distributions in each and all the BDT bins is shown and Figure 5 exhibits the ratio of the decay time dependent efficiency in each BDT bin by the total decay time dependent selection efficiency.

From this decay time dependent efficiency, the decay time integrated efficiency ϵ^i, can be computed with Equation 3. For the signal it depends on $\mathcal{A}_{\Delta\Gamma}$ and y_s and is therefore written $\epsilon^i_{sig}(\mathcal{A}_{\Delta\Gamma}, y_s)$.

4.2 PDF Calibration

Following Equation 7, the number of signal events in each BDT bin can be expressed using $\mathrm{B}^0 \to \mathrm{h}^+\mathrm{h}'^-$ as control channel as:

$$N^{exp,i}_{sig} = N^{exp,i}_{cc} \times \frac{f_s}{f_d} \times \frac{\epsilon^i_{sig}(\mathcal{A}_{\Delta\Gamma}, y_s)}{\epsilon^i_{cc}} \times \frac{\mathcal{B}^0_{sig}}{\mathcal{B}^0_{cc}} \times \frac{1 + \mathcal{A}_{\Delta\Gamma} y_s}{1 - y_s^2}. \tag{11}$$

The ratio of Equation 11 by the total number of signal event obtained from Equation 7 with $\mathrm{B}^0 \to \mathrm{h}^+\mathrm{h}'^-$ as control channel gives for each BDT bin the shape of the PDF:

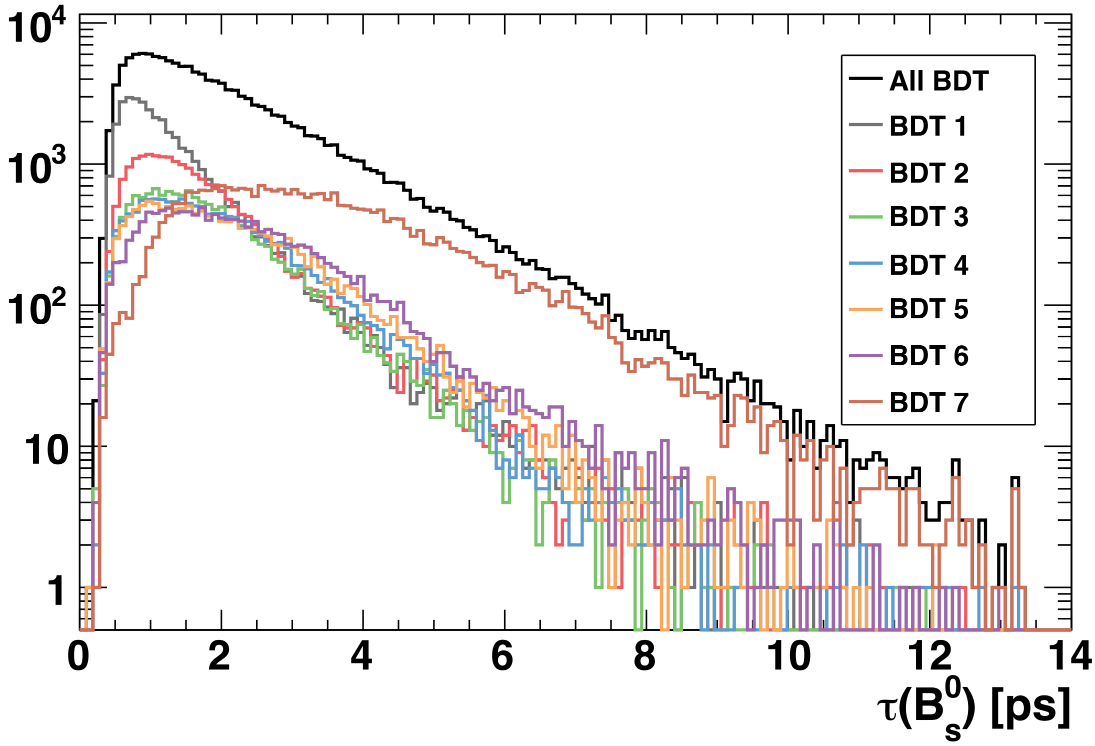


Figure 4: Proper time distributions in each BDT bin and in all BDT bins.

$$\frac{N_{sig}^{exp,i}}{N_{sig}^{exp}} = \frac{N_{cc}^{exp,i}}{N_{cc}^{exp}} \times \frac{\epsilon_{sig}^{i}(\mathcal{A}_{\Delta\Gamma}, y_s)}{\epsilon_{cc}^{i}} \times \frac{\epsilon_{cc}}{\epsilon_{sig}(\mathcal{A}_{\Delta\Gamma}, y_s)} \tag{12}$$

The decay time integrated efficiency terms are obtained by convolving the decay time dependent efficiency by the decay rate (c.f. Equation 3). The decay time dependent efficiency can be reasonably assumed to be identical between signal and control channels but their decay rates have a different dependence on the decay time and the $\mathsf{B}_\mathsf{s}^0 \to \mu^+\mu^-$ one is model dependent. As a consequence, the ratio of signal to control channels efficiencies does not cancel and is model dependent. However, in all the published analyses the PDF was obtained for each bin, neglecting these efficiencies ratio, by:

$$\frac{N_{sig}^{exp,i}}{N_{sig}^{exp}} = \frac{N_{cc}^{exp,i}}{N_{cc}^{exp}}. \tag{13}$$

Therefore the BDT PDF is biased by:

$$\delta^i_{PDF} = \frac{\epsilon^i_{sig}(\mathcal{A}_{\Delta\Gamma}, y_s)}{\epsilon^i_{cc}} \times \frac{\epsilon_{cc}}{\epsilon_{sig}(\mathcal{A}_{\Delta\Gamma}, y_s)}.$$

The correction to be applied is model dependent, parametrised by $\mathcal{A}_{\Delta\Gamma}$ and y_s. In addition each BDT bin has a different decay time dependent efficiency, hence a different correction.

The correction to be applied to BDT PDF is shown in Figure 6 as a function of $\mathcal{A}_{\Delta\Gamma}$ and y_s. The correction obtained with the SM value of $\mathcal{A}_{\Delta\Gamma}$ and the latest experimental average value of $y_s = 0.069$ [9] are reported in Table 1 together with correction to the $\mathcal{B}(\mathrm{B^0_s} \to \mu^+\mu^-)$ in each bin, if bins were fitted individually. The last BDT bin (bin 7) is the most sensitive one and drives to a large extent the analysis results. If the combined fitted result were completely determined by this bin, the combined 2012 fitted branching ratio [7] would be (assuming the SM $\mathcal{A}_{\Delta\Gamma}$ and the y_s world average values) biased by a further $-5.11 \pm 0.06\%$ yielding to:

$$\mathcal{B}(\mathrm{B^0_s} \to \mu^+\mu^-) = (2.7^{+1.3}_{-1.0}) \times 10^{-9},$$

to be compared with $(3.23 \pm 0.27) \times 10^{-9}$, the SM prediction.

An estimate of the bias on the combined $\mathcal{B}(\mathrm{B^0_s} \to \mu^+\mu^-)$ results can be obtained by averaging the correction in each bin by its relative contribution to the total analysis sensitivity. These contributions were estimated in [10] based on a former analysis [5] and are reported in Table 2. Assuming that the bins contributions to the sensitivity are similar to the latest analysis ones [7] the weighted averaged correction is: -3.5%. We also evaluate the bias on the combined $\mathcal{B}(\mathrm{B^0_s} \to \mu^+\mu^-)$ results by a maximum likelihood unbinned fit to a sample of data events with 10 signal Monte Carlo candidates embedded with and without the correction on the BDT PDF and without any uncertainty on the BDT PDF that could absorb the central value shift. A correction of the same level as the one estimated with the previous method is found.

For $\mathrm{B^0} \to \mu^+\mu^-$ no correction has to be applied since the control channel is also a $\mathrm{B^0}$ decay and has therefore the same decay rate as the signal one. The calibration performed in the analysis also uses $\mathrm{B^0_s} \to \mathrm{h^+h'^-}$ events as well. We neglect here these channels for two reasons: first, they are three times less abundant than the $\mathrm{B^0} \to \mathrm{h^+h'^-}$; second, the decay-rate of these $\mathrm{B^0_s} \to \mathrm{h^+h'^-}$ channels are required to compute the corrections and, apart for $\mathrm{B^0_s} \to \mathrm{K^+K^-}$ for which the effective life time and so $\mathcal{A}^{KK}_{\Delta\Gamma}$ has just been measured [11], the others are still unknown. We recommend to decide whether or not to include these channels in the calibration process of future analyses, by comparing the gain these additional events bring to the statistical uncertainties to the additional systematics they add.

5 Conclusions

After the dramatic experimental progress in the measurement of the branching ratio of the $\mathrm{B^0_s} \to \mu^+\mu^-$ decay, it is of prime interest to clarify the quantity which is actually measured

Table 1: Correction to the BDT PDF and to the $\mathcal{B}(\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-)$ in each bin if bins were fitted individually. The corrections assume $\mathcal{A}_{\Delta\Gamma} = 1$ and $y_s = 0.069$ [9].

Bin	PDF Correction $\delta^i_{PDF} - 1$ (%)	$\mathcal{B}(\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-)$ Correction $1/\delta^i_{PDF} - 1$ (%)
1	-3.17 ± 0.03	$+3.27 \pm 0.03$
2	-1.86 ± 0.03	$+1.90 \pm 0.04$
3	-1.16 ± 0.04	$+1.17 \pm 0.04$
4	-0.50 ± 0.05	$+0.51 \pm 0.05$
5	$+0.07 \pm 0.05$	-0.07 ± 0.05
6	$+0.95 \pm 0.06$	-0.94 ± 0.06
7	$+5.39 \pm 0.06$	-5.11 ± 0.06

Table 2: Sensitivity contribution of each BDT bin to the total analysis sensitivity.

Bin	Sensitivity Contribution (%)
1	0.0981
2	1.09
3	2.03
4	4.33
5	8.70
6	16.1
7	67.6

and the way it has to be compared with theoretical predictions. This note cure the analysis from two biases resulting from an incorrect treatment of the time-dependent efficiencies.

For $\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-$, the corrections to be applied are model dependent and for the SM, under the assumption that all analysis sensitivity come from the last BDT bin, for which the correction is maximum, the bias found is $(-8.85 \pm 0.06)\%$. Including the additional -6.9% correction required to translate the $\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-$ proper time integrated branching ratio into the initial CP-average one, the total shift is $(-15.14 \pm 0.06)\%$. Therefore the branching ratio measured in [7] which was $\mathcal{B}(\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-) = (3.2^{+1.5}_{-1.2}) \times 10^{-9}$, and compared to $(3.54 \pm 0.27) \times 10^{-9}$ is now:

$$\mathcal{B}(\mathsf{B}^0_\mathsf{s} \to \mu^+\mu^-) = (2.7^{+1.3}_{-1.0}) \times 10^{-9},$$

and is to be compared to $(3.23 \pm 0.27) \times 10^{-9}$.

For $\mathsf{B}^0 \to \mu^+\mu^-$ the correction is smaller and accounts for a $(-1.46 \pm 0.01)\%$ shift in the 95% confidence level upper limits of the latest analysis [7] from $\mathcal{B}(\mathsf{B}^0 \to \mu^+\mu^-) < 9.4 \times 10^{-9}$ to:

$$\mathcal{B}(\mathsf{B}^0 \to \mu^+\mu^-) < 9.3 \times 10^{-10}.$$

To account for the model dependency of these corrections for $B_s^0 \to \mu^+\mu^-$ in future analyses, we define a model dependent normalisation factor which corrects for the bias due to the signal efficiency computation and translates the number of observed $B_s^0 \to \mu^+\mu^-$ signal events directly to the CP-averaged branching ratio used by theoreticians to compute predictions. We provide, for the latest analysis [7], the curves of this new normalisation factor as a function of y_s and $\mathcal{A}_{\Delta\Gamma}$ the two parameters describing the model dependency. These curves allow to translate a result (limit or branching fraction) obtained for specific y_s and $\mathcal{A}_{\Delta\Gamma}$ values (e.g. SM) to any other couple of values. The same curves are provided for the correction due to the calibration. However, these corrections must be applied before fitting or extracting limits as they modify the BDT PDF.

A Inputs

LHCb inputs used for the 2011 and 2012 simulations (will change soon):

$$\begin{aligned} \tau_{B_s^0} &= 1.469141\,\mathrm{ps} \\ \tau_{B^0} &= 1.525000\,\mathrm{ps} \\ \tau_{B^+} &= 1.638000\,\mathrm{ps} \end{aligned}$$

Experimental values, taken from HFAG [9]:

$$\begin{aligned} y_s &= 0.0685 \pm 0.008 \\ \tau_{B_s^0} &= 1.503 \pm 0.010\,\mathrm{ps} \\ \tau_{B^0} &= 1.519 \pm 0.007\,\mathrm{ps} \\ \tau_{B^+} &= 1.642 \pm 0.008\,\mathrm{ps} \end{aligned}$$

References

[1] K. De Bruyn, R. Fleischer, R. Knegjens, P. Koppenburg, M. Merk, et al. Branching Ratio Measurements of B_s^0 Decays. *Phys.Rev.*, D86:014027, 2012.

[2] Kristof De Bruyn, Robert Fleischer, Robert Knegjens, Patrick Koppenburg, Marcel Merk, et al. Probing New Physics via the $B_s^0 \to \mu^+\mu^-$ Effective Lifetime. *Phys.Rev.Lett.*, 109:041801, 2012.

[3] A. Buras, J. Girrbach, D. Guadagnoli, and G. Isidori. On the Standard Model prediction for $B_s^0 \to \mu^+\mu^-$. *Eur.Phys.J.*, C72:2172, 2012.

[4] R. Aaij et al. Search for the rare decays $B_s^0 \to \mu^+\mu^-$and $B^0 \to \mu^+\mu^-$. *Phys.Lett.*, B699:330–340, 2011.

[5] R. Aaij et al. Search for the rare decays $B_s^0 \to \mu^+\mu^-$and $B^0 \to \mu^+\mu^-$. *Phys.Lett.*, B708:55–67, 2012.

[6] R. Aaij et al. Strong constraints on the rare decays $\mathsf{B_s^0} \to \mu^+\mu^-$and $\mathsf{B^0} \to \mu^+\mu^-$. *Phys.Rev.Lett.*, 108:231801, 2012.

[7] R Aaij et al. First evidence for the decay $\mathsf{B_s^0} \to \mu^+\mu^-$. *Phys.Rev.Lett.*, 110:021801, 2013.

[8] I. Dunietz, R. Fleischer, and U. Nierste. In pursuit of new physics with $\mathsf{B_s^0}$decays. *Phys.Rev.*, D63:114015, 2001.

[9] Y. Amhis et al. Averages of b-hadron, c-hadron, and tau-lepton properties as of early 2012. *arXiv hep-ex:*, 1207:1158, 2012. online update at http://www.slac.stanford.edu/xorg/hfag.

[10] M. Perrin-Terrin and G. Mancinelli. Optimisation of the binning of the discriminating variables used in the computation of $\mathsf{B_s^0} \to \mu^+\mu^-$ upper limits with the modified frequentist approach. Technical Report LHCb-INT-2012-003, CERN, Geneva, Feb 2012.

[11] R Aaij et al. Measurement of the effective $\mathsf{B_s^0} \to \mathsf{K^+K^-}$ lifetime. *Phys. Lett. B*, 716:393–400., 2012.

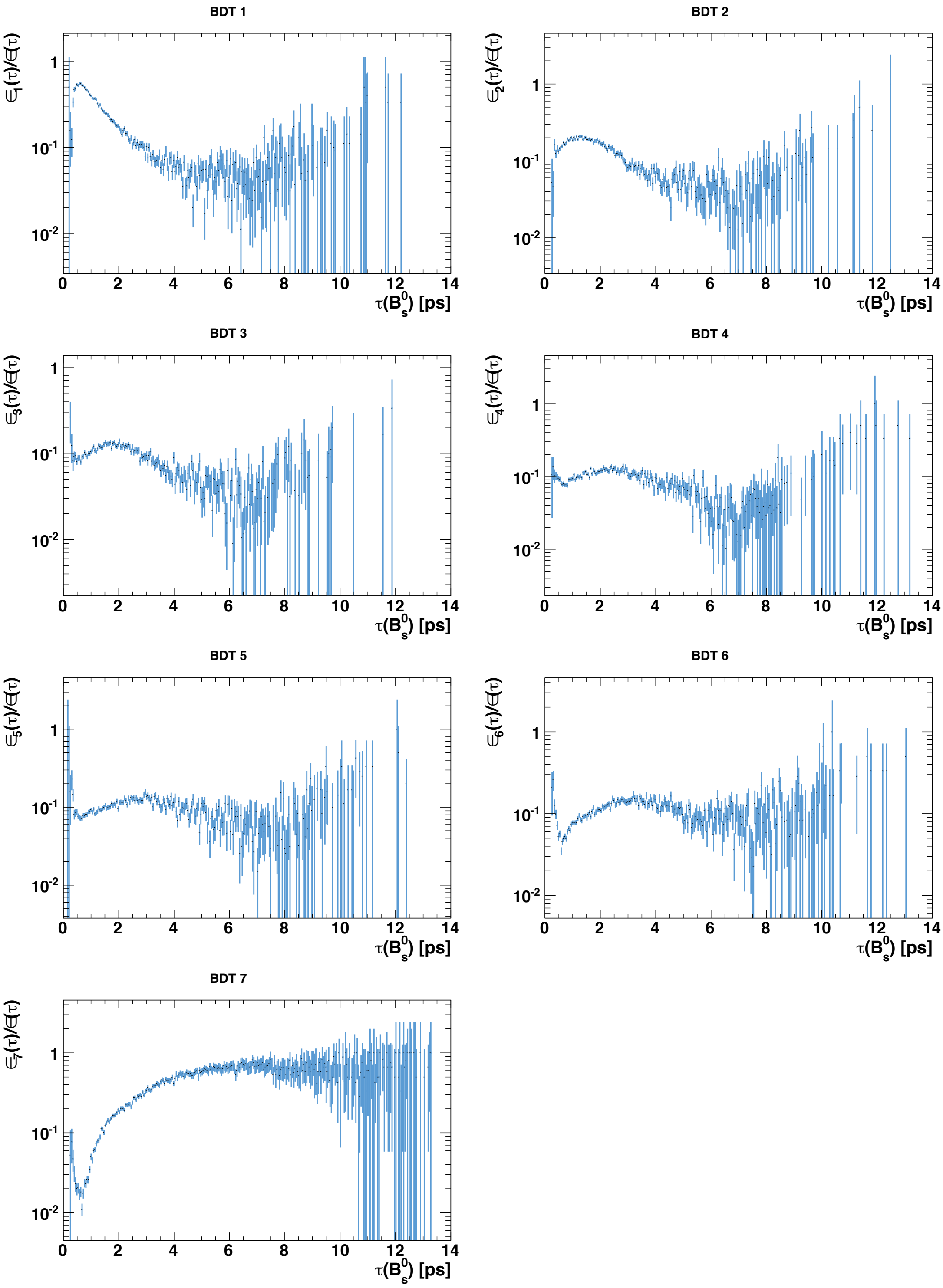
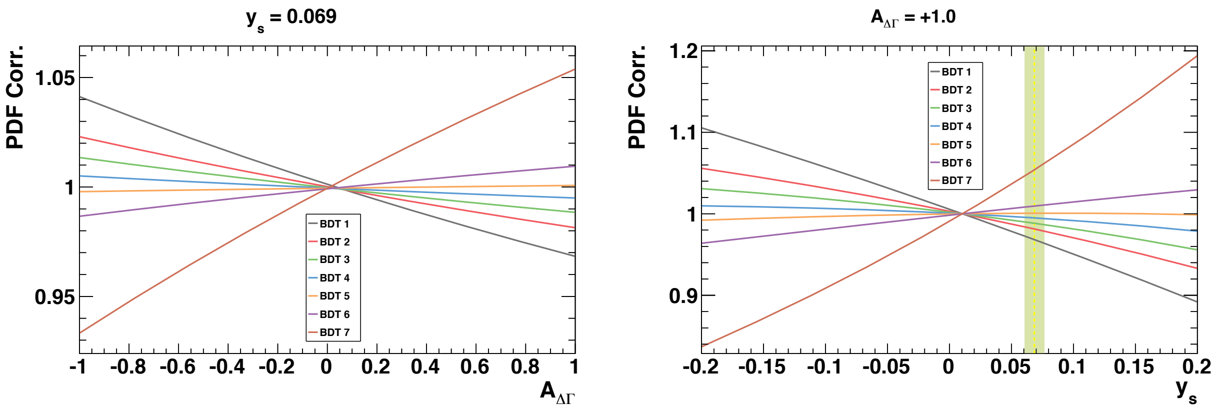


Figure 5: Ratio for each BDT bin of the decay time dependent efficiency by the total selection decay time dependent efficiency shown in Figure 1.



(a) (b)

Figure 6: (a) Correction to be applied to BDT PDF as a function of $\mathcal{A}_{\Delta\Gamma}$ for the experimental average y_s value [9] and (b) as a function of y_s for the SM $\mathcal{A}_{\Delta\Gamma}$ value (+1). The vertical green band is the experimental average y_s value and its uncertainty.