Measurement of b-quark fragmentation fraction ratios at the CMS experiment: a key ingredient for the $B_s^0 \to \mu^+\mu^-$ rare decay analysis

Bruno Afonso Fontana Santos Alves[1, 2]

Supervisors: Nuno Leonardo[1, 2]; João Varela[1, 2]

[1] *Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal*
[2] *Laboratório de Instrumentação e Física Experimental de Partículas, Lisbon, Portugal*

The current theoretical basis for Particle Physics, the Standard Model, although extremely successful in its predictions, faces a series of limitations. Many extensions have been proposed, but no clear deviations from the Standard Model have been found. Rare decays are very sensitive to hypothetical New Physics interactions. The precise measurement of the branching fraction (BF) of the $B_s^0 \to \mu^+\mu^-$decay imposes strong constraints on Beyond the Standard Model theories and may give hints on how to explore New Physics processes. This work represents the measurement of one of the main ingredients for the mentioned branching fraction measurement: the f_s/f_d fragmentation fraction ratio. Data produced at the Large Hadron Collider and detected by the Compact Muon Solenoid at a 13 TeV centre of mass collision energy is used. The measurement of yields, efficiencies and systematic uncertainties is performed. The f_s/f_d ratio is measured to be 0.186 ± 0.002 (stat.) $\pm$ 0.021 (syst.) $\pm$ 0.016 (BF). Two other ratios are measured to confirm the robustness of the results: $f_s/f_u = 0.216 \pm 0.003$ (stat.) $\pm$ 0.029 (syst.) $\pm$ 0.017 (BF) and $f_d/f_u = 1.164 \pm 0.008$ (stat.) $\pm$ 0.120 (syst.) $\pm$ 0.057 (BF). We also conclude that no kinematic dependence is found for the ratios in the acceptance region of the measurement, defined in terms of rapidity, $|y| < 2.25$, and transverse momentum, $10 < p_T < 90$ GeV, of the B mesons.

I. INTRODUCTION

The Standard Model (SM) of particle physics is a theoretical description of the fundamental building blocks of the Universe. It encompasses all known subatomic particles, grouping them according to their quantum numbers, and explaining the way they interact. The SM predicts most experimental results obtained so far. Its last missing piece was unveiled only in 2012, with the observation of the Higgs boson [1, 2]. It is a well-tested theory in describing particle interactions up to TeV scales. Despite all of its predictive success, the SM is known to be incomplete. Several factors motivate the existence of New Physics (NP). For example, the lack of a dark matter candidate and the observed baryon asymmetry in the Universe have no explanation inside the SM.

A. B Mesons

The bottom and charm quarks are known as the "heavy flavour" quarks, since they are the most massive quarks that can comprise observable particles. That is not the case for the top quark, despite having a larger mass: it has an extremely short lifetime. After their production, the heavy flavour quarks hadronize into larger composite particles, the hadrons. Hadrons containing b quarks will be the ones with the largest mass. B mesons are hadrons composed by two quarks, being one of them an anti-b quark, $\bar{b}$, and the other a different quark (see Table I). In general, from the two quarks that form a B meson, the $\bar{b}$ quark is the one which decays, being the most massive inside the meson, while the other is referred to as the spectator quark. The dominant decay mode of a $\bar{b}$ anti-quark is $\bar{b} \to \bar{c}\,W^+$, since the otherwise CKM favoured $\bar{b} \to \bar{t}\,W^+$ is virtual only (the mass of the two particles is larger than the mass of the $\bar{b}$). Then, the W decays either into a pair of leptons $l\bar{\nu}$ (semileptonic decays), or into a pair of quarks which hadronize.

TABLE I: Properties of the lightest B meson states, where 'S' stands for spin and 'P' for parity [3].

Meson	Rest mass [MeV]	S^P	Lifetime [ps]
B^0 ($d\bar{b}$)	5279.62 ± 0.15	0^-	1.520 ± 0.004
B^+ ($u\bar{b}$)	5279.31 ± 0.15	0^-	1.638 ± 0.004
B_s^0 ($s\bar{b}$)	5366.82 ± 0.22	0^-	1.511 ± 0.014
B_c^+ ($c\bar{b}$)	6274.9 ± 0.8	0^-	0.507 ± 0.009

1. Production

Measured cross section deviations from Quantum Chromodynamics (QCD) predictions may hint at the effects of Beyond the Standard Model (BSM) physics. In order to study the cross section behaviour according to a given kinematic observable, a differential cross section is measured. The latter is given by [4]:

$$\frac{d\sigma(pp \to B\,X)}{dx} = \frac{N}{2\mathcal{A}\epsilon\mathcal{B}L\Delta x}\,, \tag{1}$$

where 'X' represents all particles a proton-proton (pp) collision can produce together with a B meson, N is the measured signal yield, $\mathcal{A}\epsilon$ is the product of acceptance and efficiency of the detector, $\mathcal{B}$ is a product of branching fractions (as an example, when studying the B_s^0 production, $\mathcal{B}$ may be the product of the $B_s^0 \to J/\psi\,\phi$,

$J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$ branching fractions), and Δx represents the bin width, with x being, for the case of this work, the transverse momentum or the rapidity. The factor of two is needed since N includes both B_s^0 and $\overline{B_s^0}$.

2. *Lifetime*

The B lifetime is long enough so that the primary and secondary vertexes can be spatially separated. Precise B lifetime measurements are key for understanding the role of the CKM matrix in CP violation, through the determination of mixing parameters and of one of its entries, V_{cb} [3]. The proper decay time t is determined from the measured decay length L as follows:

$$t = \frac{L}{\gamma v}, \qquad \text{with } \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}\ , \tag{2}$$

where γ is the Lorentz factor, c is the speed of light and v is the velocity of the particle in the lab frame. For a B meson momentum of 100 GeV, the decay length observed in the detector will be on average 0.85 cm.

3. *Mixing*

The B^0 and B_s^0 mesons ($B_{(s)}^0$) oscillate between their particle and anti-particle states, just like the K^0 ($d\bar{s}$) and D^0 ($c\bar{u}$) neutral mesons. Their mixing is determined by the mass difference between their heavy and light mass eigenstates and it happens at high frequencies [3]:

$$\Delta m(B^0) = (50.64 \pm 0.19) \times 10^{10}\ s^{-1},$$
$$\Delta m(B_s^0) = (17.757 \pm 0.021) \times 10^{12}\ s^{-1}.$$

The amount of CP violation present in $B_{(s)}^0$ mixing is very small [5]. As a consequence the mass eigenstates are very similar to CP eigenstates. The $B_s^0 \to \mu^+\mu^-$ decay, for example, occurs essentially through the heavy mass eigenstate decay only. Thus, according to the SM, the effective lifetime of the $B_s^0 \to \mu^+\mu^-$ process should be equal to the lifetime of the heavy mass eigenstate. This is not necessarily true in BSM scenarios where the light mass state can also decay into a muon pair. The $B_s^0 \to \mu^+\mu^-$ effective lifetime is therefore a good probe for BSM searches [6, 7].

B. Rare decays

The B meson decays which do not occur through the $\bar{b} \to \bar{c}\,W^+$ transition are usually denoted "rare" [8]. Specifically, the $B_{(s)}^0 \to \mu^+\mu^-$ decays have an extremely small and precisely predicted branching fraction [9]:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-)_{SM} = (3.57 \pm 0.17) \times 10^{-9},$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-)_{SM} = (1.06 \pm 0.09) \times 10^{-10}.$$

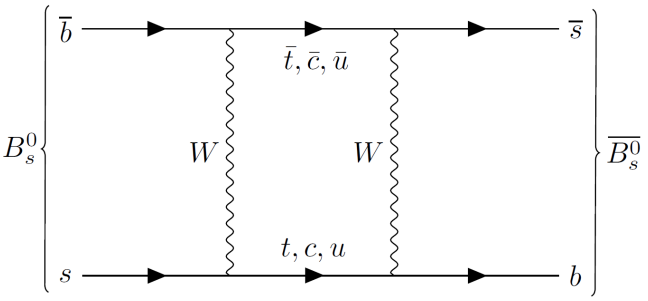

FIG. 1: Box mixing diagram contributing to the B_s^0 mixing. Another possibility is obtained by rotating this diagram 90°.

Their rarity is due to: *i*) FCNC are not allowed at tree-level in the SM and as a consequence the decays occur through higher-order diagrams (see Figure 2), *ii*) CKM suppression and *iii*) helicity suppression [10, 11]. The available experimental data suggests that tree-level contributions from NP are suppressed. Large new contributions are most likely to be present in loop-mediated processes [12]. The $B_{(s)}^0 \to \mu^+\mu^-$ decay channels represent therefore powerful discriminants to check the validity of BSM theories, in particular models containing additional Higgs bosons [12, 13].

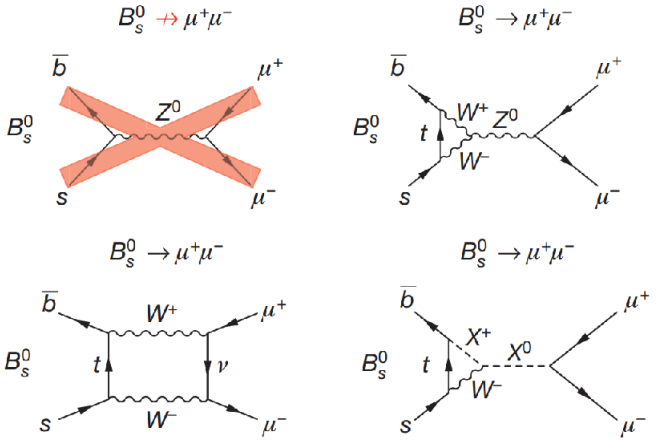

FIG. 2: The SM does not allow FCNC at tree-level (upper left). Rare decays only occur through higher-order diagrams (upper right and bottom left). BSM physics can contribute to rare processes amplitudes (bottom right) [6].

SM predicted branching fractions might be different according to certain NP scenarios. The possibility of finding an unexpected excess relative to the SM is what drives the measurement of highly suppressed branching fractions. The $B_s^0 \to \mu^+\mu^-$ decay was observed for the first time in 2015 by the CMS and LHCb Collaborations [6], and its branching fraction yielded:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (2.8^{+0.7}_{-0.8}) \times 10^{-9},$$

which is compatible with SM predictions. The CMS Collaboration is currently pursuing the precise measurement

of the same quantity using data collected at 13 TeV centre of mass collision energy. This measurement will provide strong constraints on the parameters of BSM theories in case the $B_s^0 \to \mu^+\mu^-$ branching fraction is measured to be compatible with the SM predictions. Otherwise, a direct indication of NP will be found.

C. Fragmentation Fractions

The $B_s^0 \to \mu^+\mu^-$ branching fraction is measured using the following formula:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \frac{N_{B_s^0 \to \mu^+\mu^-}}{N_{B^+ \to J/\psi K^+}} \frac{\varepsilon_{B^+}}{\varepsilon_{B_s^0}} \frac{f_u}{f_s} \mathcal{B}_{B^+}. \quad (3)$$

The symbol ε represents the overall efficiency for the $B^+ \to J/\psi K^+$ and $B_s^0 \to \mu^+\mu^-$ signals, which includes the acceptance of the detector. The 'N' refers to the observed number of signal events, and $\mathcal{B}_{B^+}$ denotes the product of the $B^+ \to J/\psi K^+$ and $J/\psi \to \mu^+\mu^-$ branching fractions. Finally, f_u/f_s denotes a ratio between "b-quark fragmentation fractions", which represent the probability for a b quark to hadronize into a certain b-flavoured hadron. Neglecting excited B decays, the fragmentation fraction is defined as $f_X \equiv N_{b\to B_X}/N_{b\to all}$, where '$B_X$' here refers to any b hadron, and not just B mesons. The sum of all fragmentation fractions must be equal to unity since they represent a probability. In addition, $f_u = f_d$ can be assumed when neglecting quark mass differences and when considering that excited B mesons decay to the B^+ and B^0 mesons equally. This work will test this assumption (see Eq. 4). To measure the quantity shown in Eq. 3, the f_s/f_d ratio will be used instead of the f_s/f_u one, since its systematic uncertainty might be lower due to cancellations present in the ratio (see Eqs. 5 and 6). The quantities in Eqs. 4, 5 and 6 have the same meaning as explained after Eq. 3. It is important to note that the fragmentation fractions at different machines might not be strictly equal, since they could depend on the kinematic regions under study [14]. It is still unknown whether the fragmentation fraction ratio f_d/f_s depends on the transverse momentum or rapidity of the B mesons. Fig. 3 shows past f_d/f_s transverse momentum dependency studies. While LHCb saw a dependency, ATLAS found the results to be compatible with no dependency.

$$\frac{f_d}{f_u} = \frac{N_{B^0 \to J/\psi K^{*0}}}{N_{B^+ \to J/\psi K^+}} \frac{\varepsilon(B^+ \to J/\psi K^+)}{\varepsilon(B^0 \to J/\psi K^{*0})} \frac{\mathcal{B}(B^+ \to J/\psi K^+)}{\mathcal{B}(B^0 \to J/\psi K^{*0})\mathcal{B}(K^{*0} \to K^+\pi^-)}, \quad (4)$$

$$\frac{f_s}{f_u} = \frac{N_{B_s^0 \to J/\psi \phi}}{N_{B^+ \to J/\psi K^+}} \frac{\varepsilon(B^+ \to J/\psi K^+)}{\varepsilon(B_s^0 \to J/\psi \phi)} \frac{\mathcal{B}(B^+ \to J/\psi K^+)}{\mathcal{B}(B_s^0 \to J/\psi \phi)\,\mathcal{B}(\phi \to K^+K^-)}, \quad (5)$$

$$\frac{f_s}{f_d} = \frac{N_{B_s^0 \to J/\psi \phi}}{N_{B^0 \to J/\psi K^{*0}}} \frac{\varepsilon(B^0 \to J/\psi K^{*0})}{\varepsilon(B_s^0 \to J/\psi \phi)} \frac{\mathcal{B}(B^0 \to J/\psi K^{*0})\mathcal{B}(K^{*0} \to K^+\pi^-)}{\mathcal{B}(B_s^0 \to J/\psi \phi)\mathcal{B}(\phi \to K^+K^-)}. \quad (6)$$

The joined CMS and LHCb analysis in Ref. [6] used the LHCb f_s/f_d measurement [15]. An additional systematic uncertainty of 5% was assigned to account for the extrapolation of the LHCb result to the CMS acceptance [6]. This is particularly problematic since we are now dealing with a precision measurement. Once CMS measures the f_s/f_d quantity, the uncertainty will be more robustly determined.

This work represents a description of the author's contribution to the measurement of the f_s/f_d fragmentation fraction ratio in the framework of the $B_s^0 \to \mu^+\mu^-$ rare decay analysis, using CMS Run2 data (2015). The f_d/f_u and f_s/f_u ratios are measured to verify the consistency of the results. The transverse momentum and rapidity dependency of the three ratios is studied.

II. EXPERIMENTAL APPARATUS

This work uses data collected by the Compact Muon Solenoid (CMS) at the Large Hadron Collider (LHC). The LHC can be found in a 27 km circular tunnel 100 m underground where two counter-rotating beams of protons are accelerated up to speeds larger than 99.99999% the speed of light, with a centre of mass collision energy of 13 TeV. The LHC has up to now delivered an integrated luminosity of about 130 fb^{-1}, while CMS has recorded close to 120 fb^{-1} [16]. The protons are squeezed together to form bunches that cross each other every 25 ns. Even though each of these bunches contains more than 10^{11} protons, only about 30 collisions take place per crossing. Still, due to the 40 MHz bunch collision frequency, the LHC produces almost 1 billion collisions per second [17]. The CMS detector is located at one of the existing interaction points in the LHC tunnel. It employs a 13 m long, 6 m diameter superconducting solenoid magnet operated at circa 3.8 T. It is composed of five major com-

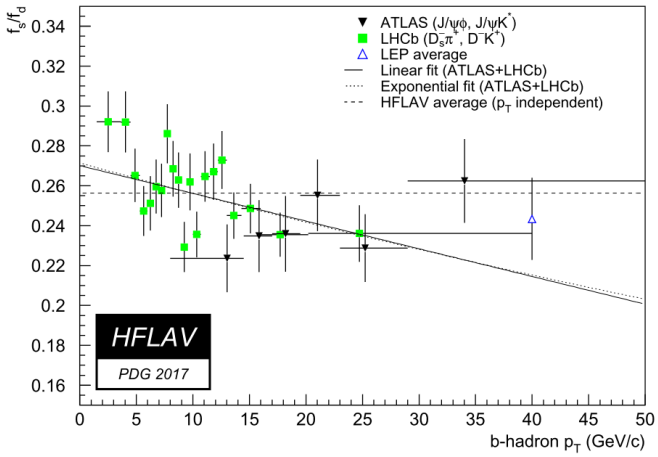

FIG. 3: Measurement of the f_s/f_d ratio as a function of p_T(B). While LHCb (green points) reported a dependence, ATLAS (black points) did not find it. This work attempts to shed some light on this inconsistency [14].

ponents: the silicon tracker, the electromagnetic and the hadronic calorimeter, the magnet itself, and the muon stations (see Fig. 4). For the purpose of this work, the silicon tracker and the muon stations represent the most important detector parts, which are responsible, respectively, for measuring the trajectories of all charged particles ("tracks") and specifically muons. In addition to this, only potentially relevant collisions are stored. The two-level trigger system plays a crucial role in the event selection: the large storage space and bandwidth required and the prohibitive costs associated impose a constraint in the maximum number of particles subjected to analysis.

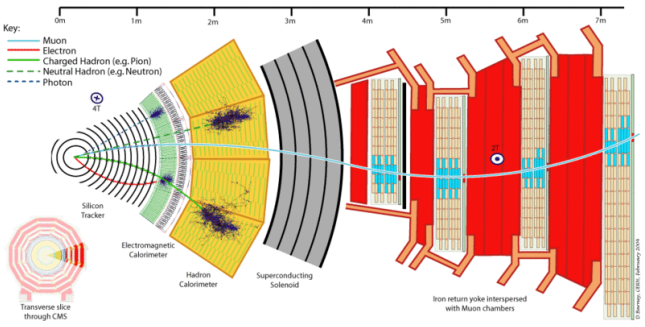

FIG. 4: Transverse section of the CMS detector, which is composed of the following (from left to right): silicon tracker, electromagnetic calorimeter, hadronic calorimeter, solenoid and four muons stations interleaved with return yoke plates. The silicon tracker measures the trajectory of charged particles. The muon stations are essential to distinguish muons (blue line) from other tracks (red and green lines).

III. DATA AND MC SAMPLES

Proton collisions at the LHC (13 TeV) collected by CMS in 2015 are studied, corresponding to a 2.54 fb^{-1} integrated luminosity. The data are processed through a dedicated trigger path, common for the three channels of interest, which requires the presence of two muons compatible with the J/ψ mass, and an additional track, originating from a common vertex displaced from the primary vertex. The muon candidates with transverse momentum $p_T(\mu) > 4.2$ GeV, pseudorapidity $|\eta(\mu)| < 2.4$ and with opposite charges are paired to form J/ψ candidates. Each J/ψ candidate must have $p_T > 8$ GeV and a χ^2 probability for a fit to the dimuon vertex larger than 10%. Candidate B mesons are reconstructed by combining a J/ψ candidate with one (B^+) or two (B^0, B_s^0) additional charged tracks with $p_T > 1$ GeV and $|\eta| < 2.4$. The track-fit χ^2 must be less than five times the number of degrees of freedom. Track candidates overlapping with muon candidates are rejected. A kinematic fit is performed to the dimuon-tracks combination, constraining the dimuon mass to the nominal J/ψ mass. The combination must be compatible with having a common vertex with a vertex-fit χ^2 probability larger than 10% and a reconstructed invariant mass in the 5–6 GeV range. The transverse decay length significance, defined as the distance between the primary and secondary vertexes in the transverse plane, divided by its uncertainty, is required to exceed 3.5. The cosine of the angle between the B candidate momentum and the vector formed by the interaction point and the secondary vertex in the transverse plane must be greater than 0.99. The B_s^0 (B^0) candidates are formed by assigning the mass of the two kaons (one kaon and one pion) to the tracks, yielding an invariant ditrack mass within a 20 MeV (100 MeV) window centered on the φ (K^{*0}) mass. The B candidates are required to lie in the acceptance range of the analysis: $10 < p_T(B) < 90$ GeV and $|y(B)| < 2.25$.

IV. DATA MC COMPARISON

The MC distributions should describe the data. To check for this, data and MC are compared; a bad MC description of the data would affect the measurement of the efficiencies, since these are MC-based. The comparison is performed with a standard sideband subtraction method, where the mass of the B candidates is the discriminating variable. The signal and background sidebands are chosen according to Table II. By fitting the mass sidebands with an exponential function, one is able to estimate the amount of background in the signal region. The signal distribution of a certain variable is obtained by subtracting the background in the signal region from the full data in the same signal region, according to:

$$\text{signal} = \text{full} - \frac{N_{\text{central}}}{N_{\text{left}} + N_{\text{right}}} \times \text{background}, \qquad (7)$$

where 'signal' and 'full' refer to the data in the signal region, 'background' refers to the data in the sidebands, and 'N_X' refers to the number of events under the exponential fit in the specified 'X' region. This is done for many of the variables described in Section III. Note that the data distributions in Eq. 7 are histograms, where information about single events is not present; that information is lost since the three N_X are determined using fits. No left sideband is specified for the $B^+ \to J/\psi\, K^+$ decay since that region is dominated not by combinatorial background but by a physics background component arising from partially reconstructed b hadron decays where one or more tracks are not detected. The right sideband is thus larger than the right sidebands of the other channels to compensate for this effect.

TABLE II: Signal and background regions used for the sideband subtraction method. All values have GeV units.

Channel	B^+	B^0	B_s^0
Signal region	5.2 - 5.35	5.2 - 5.35	5.3 - 5.45
Background left region	—	4.8 - 5.0	4.8 - 5.15
Background right region	5.5 - 5.9	5.5 - 5.75	5.55 - 5.85

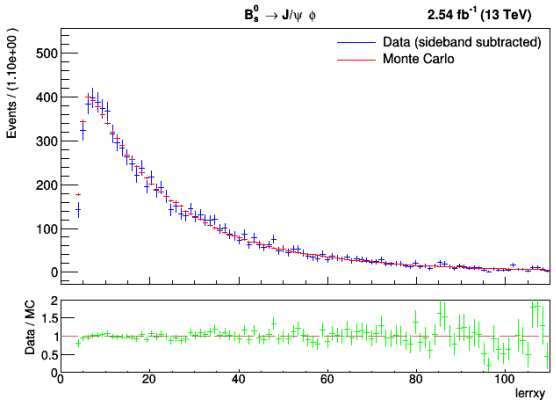

FIG. 5: L_{xy}/σ_{xy} data (blue) and MC (red) distribution for the $B_s^0 \to J/\psi\, \phi$ channel. The green weights are obtained according to Eq. 8.

A comparison between the signal in the data and in the MC is meaningful only after normalizing the latter to the area of the former since, in general, the MC samples have much larger yields than data samples. After performing the normalization, a binned distribution of weights is obtained by dividing the data histogram by the MC histogram for each variable we want to compare:

$$W_{\mathrm{Var/Bin}} = \frac{\mathrm{Data}_{\mathrm{Var/Bin}}}{\mathrm{MC}_{\mathrm{Var/Bin}}}\,, \tag{8}$$

where the variable and bin dependence is explicitly written. A large number of binned weight distributions is obtained, equal to the number of variables under study times the three channels of interest. An example can be found in Fig. 5, in the panel under the data/MC comparison plot (depicted in green). In Section VII B we discuss how the weights are used to calculate the systematic uncertainty of the efficiency.

V. YIELDS

The $B^+ \to J/\psi\, K^+$, $B^0 \to J/\psi\, K^{*0}$ and $B_s^0 \to J/\psi\, \phi$ channels of interest are fitted using the extended unbinned maximum likelihood method, considering a sum of two Gaussians for the signal and an Exponential for the background. The fit is "extended" since the probability distribution function (PDF) includes a Poisson normalization term. Other channel-specific components are considered for the B^+ and B^0 channels. The B^+ channel includes both the Cabibbo-suppressed $B^+ \to J/\psi\, \pi^+$ decays, where the pion is misidentified as a kaon, and the partially reconstructed B→J/ψ+h+X decays that appear in the mass region to the left of the signal, where the hadron 'h' is reconstructed as a kaon and the other decay products, 'X', are missed (see example in Fig. 6). The former is modelled by a triple Gaussian fixed from previous MC studies, and the latter uses an error function with magnitude and slope free to float in the fit. The B^0 decay channel also includes an additional

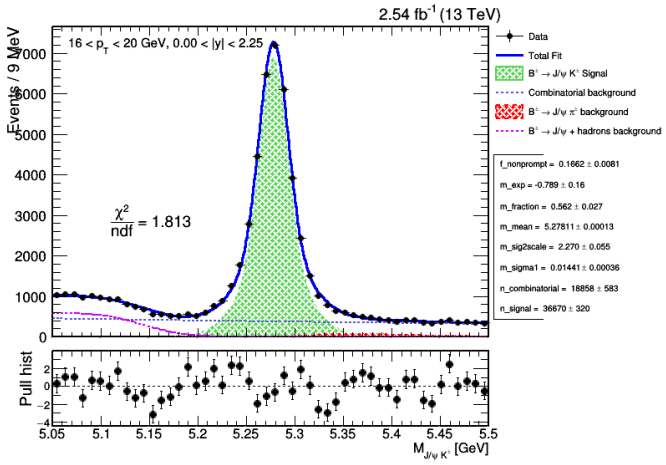

FIG. 6: Unbinned extended maximum likelihood fit for the $B^+ \to J/\psi\, K^+$ channel for the $16 < p_T < 20$ bin. The total PDF is displayed in blue, while the individual components refers to the signal (green), the combinatorial background (blue, dashed) $B^+ \to J/\psi\, \pi^+$ decay (red) and the partially reconstructed decays (pink, dashed).

component (sum of three Crystal Ball functions obtained from MC) that is used to describe the "swapped events". This terminology refers to the $K^{*0} \to K^+\pi^-$ decay: since CMS does not include a dedicated particle identification subdetector, it is unknown whether the first track is a kaon and the second one is a pion, or vice-versa. The pair combination whose invariant mass lies the closest to

the nominal mass is selected, but this does not guarantee that the selected pair is always the right one. Using MC samples with the same event selection as applied in data, we extract the fraction of candidates where an incorrect assignment is done. We refer to the candidates where this happens as "swapped", while the others are known as "true". Both swapped and true events are considered to be signal in the full fits. The fit to the swapped events is shown in Fig. 7. The obtained fraction of swapped events is $(12.91 \pm 0.17)\%$. It is fixed in the B^0 invariant mass fit. This fraction is alternatively estimated from direct counting, and it is found to be compatible with the value extracted from the fit.

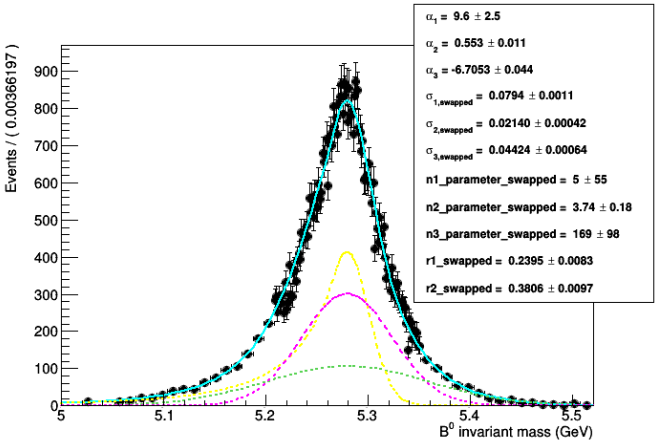

FIG. 7: Unbinned maximum likelihood fit to the B^0 swapped signal invariant mass. The cyan line describes the full PDF and the pink, yellow and green lines reflect the behaviour of each of the three crystal ball functions used. The parameters obtained from the fit are shown.

VI. EFFICIENCIES

Signal events may pass through the detector without being measured. This can happen because some particles may not lie in the detector's acceptance range (for instance, if the muons have $|\eta| > 2.4$), or because they do not pass specific online or offline selection cuts, or perhaps because some decays are not properly reconstructed. Therefore, the signal yield which appears in the master formula for this analysis (Eq. 6) does not refer to the total signal yield, but it rather refers to the observed yield, which consequently has to be corrected by an efficiency. The efficiencies are here measured using MC samples, which contain only signal.

An efficiency is simply obtained by measuring the fraction of signal that passes a certain requirement, or cut. The efficiencies are calculated using the `TEfficiency` class belonging to the `ROOT` software framework, making use of the embedded implementation for uncertainty calculations. We show in Fig. 8 an example, namely, the efficiency ratio for the f_s/f_d measurement as a function of $p_T(B)$ and $y(B)$ bins. The efficiencies for the ratios are determined from the single channel efficiencies.

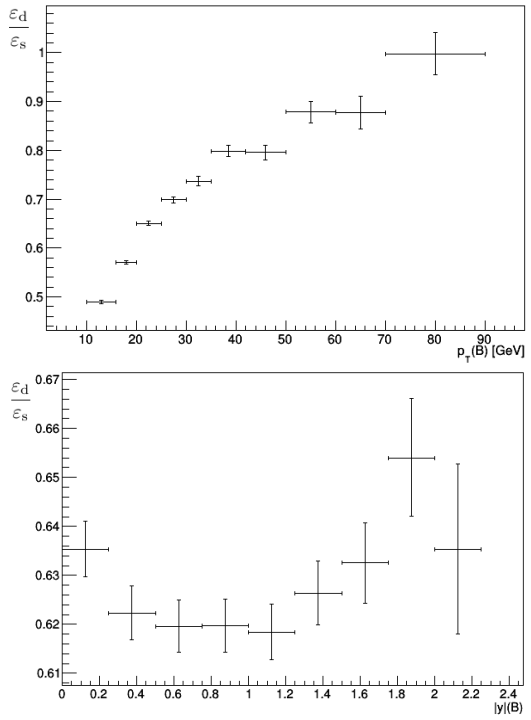

FIG. 8: Efficiencies of the f_s/f_d ratio as a function of p_T (top) and $|y(B)|$ (bottom). Other efficiency plots are as well obtained for the other two fragmentation fraction ratios and for the three channels of interest.

VII. SYSTEMATICS

One cannot reasonably consider all possible sources of systematic uncertainties. Instead, we here study the ones that could significantly alter the final measurements. It is important to find a balance between excessively conservative systematics and the opposite. While the former can affect the capability we have to detect a dependence on the fragmentation fraction ratios and can increase the overall systematic of the $B_s^0 \to \mu^+\mu^-$ branching fraction measurement, the latter is equally bad, since the error bars would give a wrong idea on where the "true" value lies.

A. Fit bias

It is important to verify whether a bias in the fits exist by guaranteeing that the signal yield obtained from the fits follows the Central Limit Theorem pattern (unitary Gaussian). Simulated samples of pseudo-data, the so-called "Toy MC", are generated according to the PDFs described in Section V. Each Toy MC sample is fitted

with the same function as its data counterpart. The procedure is independently performed for the three channels of interest, and for all the bins of the analysis. 10000 Toy MC samples are used for each bin. The distribution of the following quantity:

$$\frac{\mathrm{N}-\mu(\mathrm{N})}{\sigma(\mathrm{N})}\,, \tag{9}$$

which is called "pull", is expected to follow a Gaussian with null mean and unitary standard deviation. 'N' represents the signal yield which is repeatedly retrieved from the fits to the Toy MC samples, and μ(N) and σ(N) are, respectively, the mean and the standard deviation of the yield (N) distribution. The Toy MC samples are generated according to a Poisson distribution since the original fits to the data included an extended term. An unwanted bias, *i.e.*, a mean different than zero or a width different than unity, can occur due to a bug in the data fitting code, but it can never be related with the data itself. It exclusively depends on how the PDFs are implemented. We show, as an example, the result for the B^+ channel for p_T(B) in the 30−35 GeV range (see Fig. 9).

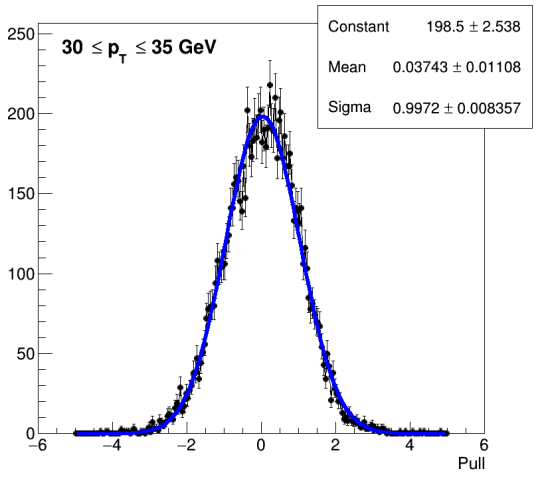

FIG. 9: Pull of the number of signal events for the B^+ channel in a representative bin: $30 \leq p_T(B) \leq 35$ GeV.

The pulls are generally well behaved, *i.e.*, unitary. Leftover deviations of the pull means from zero are accounted for as systematic uncertainties. This is done by multiplying the value of the deviation, for each channel and bin, by the statistical uncertainty of the signal yield of the fit to the corresponding bin. In order to obtain a relative uncertainty, the result is then divided by the nominal signal yield obtained from the fit to the data. The largest values are of the order of 1%.

B. MC reweighting

The weights obtained in Section IV are used to obtain a corrected version of the MC. This is known as "reweighting". It can be used to directly correct the MC in the nominal analysis or else to evaluate a systematic uncertainty that quantifies the data/MC disagreements. For the case of this work, the latter possibility is chosen, since the MC description of the data is in general good. It is important to recall that any MC problems will mostly impact the efficiency measurement, since its calculation is entirely MC-based.

We use the weights from Eq. 8 to check how would the efficiencies change if we had use the corrected MC instead of the original one. We calculated the corrected efficiencies using weights from 5 variables (p_T (track$_1$), η (track$_1$), p_T (μ_1), η (μ_1) and L_{xy}/σ_{xy}) and we used for each bin the weights that gave a larger systematic per channel. The latter is calculated by taking the difference between the nominal and corrected efficiencies. In practice, when calculating the efficiency, instead of adding '1' for each event that successfully passed the cuts, we add its respective weight. The systematic for the fragmentation fraction ratios is obtained as follows:

$$\text{Syst. uncertainty} = \frac{\left|\dfrac{\varepsilon_{\text{corrected X}}}{\varepsilon_{\text{corrected Y}}} - \dfrac{\varepsilon_{\mathrm{X}}}{\varepsilon_{\mathrm{Y}}}\right|}{\varepsilon_{\mathrm{X}}/\varepsilon_{\mathrm{Y}}}, \tag{10}$$

where both 'X' and 'Y' represent one of the three channels of interest ($\mathrm{X} \neq \mathrm{Y}$). The systematics for the full bin turned out to be 0.6%, 3.1% and 2.5%, respectively, for the f_d/f_u, f_s/f_u and f_s/f_d ratios. For individual bins these numbers stayed usually under 5% except for low yield bins, where the systematic was larger.

C. Mass window cuts

When the window mass cut is applied both to the K^{*0} coming from the B^0 meson and to the ϕ coming from the B_s^0 meson (see the selection cuts in Section IV), the correct measurement of the double track mass cuts efficiency relies on the MC samples which are used to establish those cuts. The widths of those distributions are, although similar, not equal. This represents an additional systematic uncertainty source.

Data and MC samples with all applied selection cuts except the one on the double track mass are produced. The efficiencies of the cuts are measured by fitting the double track mass distributions and later calculating the integral of the fitting function inside the used mass window. The mass distribution is fitted using a sum of three Gaussians for the peak plus a 6^{th} order Chebyshev polynomial to describe the asymmetry on the right tail of the Gaussian, which exists in both MC and in the sideband subtracted data distributions. For the case of the data some fluctuations due to low yields can be seen, which are mitigated by the fitting functions. In Fig. 10 an example of such fits is shown. The fraction of the area under the fitting function inside the mass window is taken as the double track mass cut efficiency ($\varepsilon_{\text{Fit Data}}$ and $\varepsilon_{\text{Fit MC}}$). The systematic uncertainty reached a maximum of 5.51%

for the B^0 channel and it is obtained using the following formula:

$$\text{Syst. uncertainty} = \frac{|\varepsilon_{\text{Fit Data}} - \varepsilon_{\text{Fit MC}}|}{\varepsilon_{\text{Fit MC}}} \quad (11)$$

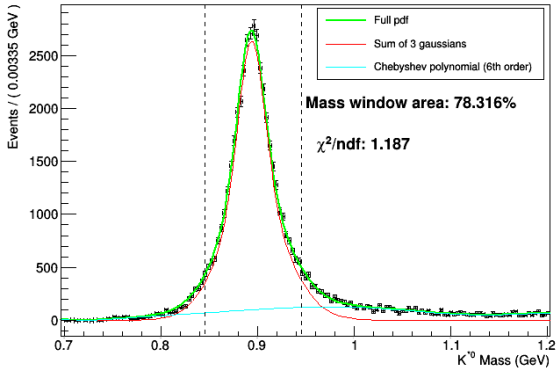

FIG. 10: Double track invariant mass spectrum for the $K^{*0} \to K^+\pi^-$ channel using signal data obtained with the sideband subtraction method. The spectrum is fitted with a maximum likelihood fit consisting of a sum of three Gaussians, plus a Chebyshev polynomial to describe some asymmetries.

D. Tag and Probe

Muons are copiously produced at the LHC. The high rates involved require the use of triggers. These only select the muons which might represent a signal process for the analysis. The sample of muons that one can then access in order to perform an efficiency measurement is, as a consequence, biased; the information provided by the muons that did not pass the trigger is lost. The efficiency measurement performed in this work is not affected by this because it is entirely based on MC samples. The overall efficiency can be calculated without introducing a bias, since all produced muons are accessible, from generation level to selection level. However, there exists a method which avoids a biased measurement of the efficiency, while still using data: the "Tag and Probe" method.

The Tag and Probe technique is used to measure the single muon efficiencies, using a very clever concept: tight selection criteria, including single muon triggers, are applied to a group of reconstructed muons. Their misidentification will then be unlikely. These particles are almost certainly real muons: the "tag" muons. The latter represent a strongly biased group of muons: an efficiency cannot be measured with them. Instead, the Tag and Probe trick is to match the tag muons with "probe" muons, being the latter muons to which no selection relative to the efficiency being measured is applied. The matching between the tag and the probe is performed by requiring the invariant mass of the two muons together, a tag and a probe, to lie close to a resonant peak, usually from the Z or J/ψ particles, since they both decay into two muons. In this way, given that the tags are very likely muons, it is almost certain that the probes will be muons as well. The latter can then be used to perform a reliable efficiency measurement; they are unbiased. The probes are then split into two groups, the "passing probes" and the "failing probes", according to their behaviour under a given cut. The passing probes pass the cut, while the failing probes do not. The passing and failing distributions are then simultaneously fitted to the resonant peak, using common floating parameters such as mean or width. The fits take into account the signal and background separately, using different PDFs for each component. The efficiency is directly extracted from the signal; it is the fraction of signal passing probes among all signal probes.

The data-driven efficiencies and their resemblance to MC single muon efficiencies were measured by the CMS Collaboration using the same muon selection as the one applied in this work. A set of weights as a function of the p_T and η of the muons was obtained. A systematic uncertainty can then be retrieved as follows. Each reconstructed B meson is inspected, and the values of p_T and η of its two corresponding muons (the ones which came from the $J/\psi \to \mu^+\mu^-$ decay) are obtained. From the available binned weights distributions the corresponding weights are extracted. For each B meson, the two weights are multiplied. This procedure is repeated for all reconstructed B mesons in the three channels of interest. When all the muon pairs have their corresponding weights assigned, the systematic uncertainty is calculated per channel and per $p_T(B)$ and $y(B)$ bins using the following expression:

$$\text{Syst. uncertainty} = \left| 1 - \frac{\sum_{\text{B mesons}} (\mathbb{W}_{\mu_1}.\mathbb{W}_{\mu_2})}{\mathcal{N}} \right| \quad (12)$$

where $\mathcal{N}$ refers to the number of reconstructed B mesons in a certain $(p_T(B), y(B))$ bin. A value of about 2.5% is obtained for the full kinematic region of the analysis, for the three channels of interest.

Since the fragmentation fraction numerator and denominator is calculated using kinematically similar muons, we here argue that the systematic uncertainty coming from the Tag and Probe method will cancel for the case of the f_s/f_d measurement. It can be seen that the uncertainties follow the same trend in different channels, when comparing the same $p_T(B)$ and $y(B)$ bins. Furthermore, this uncertainty is not dominant. The work performed in this section will be useful for future cross section measurements, because in that case no ratio involving muons is present. We decided not to include the systematic uncertainty retrieved with the Tag and Probe method in the final result.

E. Fit variations

To assess the systematic uncertainties which may exist in the determination of the yields, both the chosen fitting functions and the mass window range are changed to check whether the obtained yield remains the same or not. The difference between the nominal and the alternative yields of every variation is taken, and divided by the nominal yield in order to retrieve a relative uncertainty.

All the nominal PDFs are replaced, for all $p_T(B)$ and y(B) bins of the analysis. For the case of the signal PDF, one Gaussian only instead of the sum of two Gaussians is tried. The background is fitted with a Bernstein polynomial of 3rd degree which replaces the nominal Exponential. The $B^+ \to J/\psi\, K^+$ decay channel includes more components, and as such more systematic sources are considered: both the $B^+ \to J/\psi\, \pi^+$ and the $B \to J/\psi\, h\, X$ decays are accordingly modified, respectively, by removing the former and fitting the latter with a Gaussian rather than with an error function. We show an example of the variation of the signal fit in Figure 11.

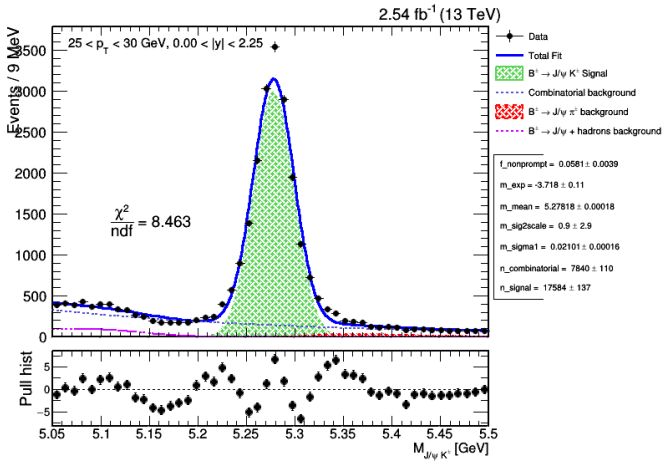

FIG. 11: Modification done to the signal fit: instead of using the sum of two Gaussians, the PDF contains one Gaussian only ($25 < p_T(B) < 30$ GeV bin).

Variations of the background yields obtained when modifying the mass window fitting range are also considered. The upper and lower bounds of the mass window are varied by 10%, and a systematic is retrieved by comparing, again, the nominal yield with the alternative yield. The variations of the PDFs gave final systematics generally under 15%, except for low yield bins, where the uncertainty occasionally rose up to 20%/25%.

F. Other systematic sources

A previous CMS study indicates that a systematic uncertainty for the track reconstruction must be considered: 2.8% for 2015 and 2.3% for 2016. The f_s/f_d ratio is not affected, since both channels have the same number of tracks. The f_d/f_u and f_s/f_u ratios include this systematic. The used branching fractions also involve uncertainties that are obtained from Ref. [3], and represent a significant uncertainty in the final measurement (see Table III). Finally, the statistical uncertainty of the efficiencies is added as a systematic. It is sub-dominant.

TABLE III: Global relative systematic uncertainties for the three fragmentation fraction ratios. The branching fraction (BF) uncertainty is displayed separately. The values are expressed in %.

Ratio	Tracking	Resolution	Pull	Total	BF
f_d/f_u	2.8	0.04	0.60	2.87	4.94
f_s/f_u	2.8	5.51	1.09	6.03	8.06
f_s/f_d	0	5.51	1.01	5.60	8.44

VIII. RESULTS AND DISCUSSION

The b-quark fragmentation fraction ratios are measured to be:

$$\begin{aligned} \frac{f_d}{f_u} &= 1.164 \pm 0.008 \pm 0.120 \pm 0.057, \\ \frac{f_s}{f_u} &= 0.216 \pm 0.003 \pm 0.029 \pm 0.017, \\ \frac{f_s}{f_d} &= 0.186 \pm 0.002 \pm 0.021 \pm 0.016. \end{aligned} \tag{13}$$

where the first uncertainty is statistical, the second is systematic and the third is due to the branching fractions. The ratio f_d/f_u yields a result compatible with unit. This serves as a corroboration for the procedures of the analysis. The f_s/f_u and f_s/f_d ratios are found compatible when considering the three shown uncertainty sources.

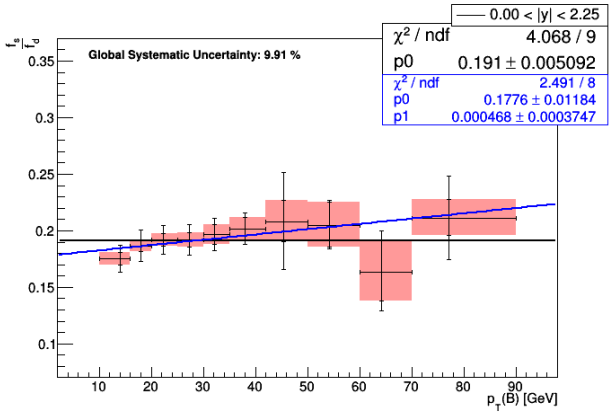

FIG. 12: Measurement of the f_s/f_d ratio as a function of $p_T(B)$. The constant fit appears in black and the linear one can be seen in blue.

The most precise result is obtained with the latter, as expected. The three ratios are measured as a function of the momentum and rapidity of the B mesons. The f_s/f_d ratio is shown in Figs. 12 and 13. The systematic and statistical uncertainties are included. The global systematic uncertainties are not displayed as vertical error bars but rather appear indicated in black in the top left corner.

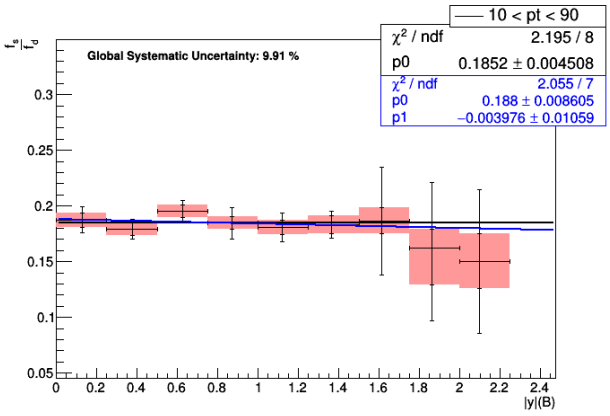

FIG. 13: Same as in Fig. 12, but as a function of |y(B)|.

The results are fitted with two alternative models: constant and linear (see Figs. 12 and 13). The slope parameters of the three ratios are close to zero. This suggests the absence of possible dependencies of these ratios on $p_T(B)$ and y(B), within the accessible precision, and in the acceptance region of the measurement: $|y(B)| < 2.25$ and $10 < p_T(B) < 90$.

This is the first fragmentation measurement performed by CMS. It is the sole result to date to have been obtained with 13 TeV data. No kinematic dependence was observed in the $|y(B)| < 2.25$ and $10 < p_T(B) < 90$ range. Both ATLAS and LHCb have measured the f_s/f_d ratio with 7 TeV data: LHCb obtained $f_s/f_d = 0.256 \pm 0.020$ [15] and ATLAS got $f_s/f_d = 0.240 \pm 0.045$ [18]. These f_s/f_d results lie within 2.70σ (LHCb) and 2.09σ (ATLAS) from the result here obtained. The dependence result here reported agrees with the ATLAS result.

Acknowledgments

We thank LIP and the CMS-LIP group for providing everything necessary to complete this work. A special thanks goes to Prof. Dr. Nuno Leonardo, for his constant dedication to the project. We further acknowledge the support of the "Collaboration in the CMS Experiment" (Ref: CERN/FIS-NUC/0029/2015) and "Probing new physics with muon pairs and heavy flavor at the LHC" (Ref: IF/01454/2013/CP1172/CT0003) grants, funded through *Fundação para a Ciência e Tecnologia* (FCT).

[1] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[2] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.

[3] C. Patrignani *et al.* (Particle Data Group), Chin. Phys. C, 40, 100001 (2016).

[4] The CMS Collaboration. Measurement of the Strange B Meson Production Cross Section with J/ψ ϕ Decays in *pp* Collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev.*, D84:052008, 2011.

[5] The CMS Collaboration. Measurement of the CP-violating weak phase ϕ_s and the decay width difference $\Delta\Gamma_s$ using the $B_s^0 \to J/\psi\phi(1020)$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV. *Phys. Lett.*, B757:97–120, 2016.

[6] The CMS and LHCb Collaborations. Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data. *Nature*, 522:68–72, 2015.

[7] The LHCb Collaboration. 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.*, 118:191801, May 2017.

[8] Thomas E. Browder and Klaus Honscheid. B mesons. *Progress in Particle and Nuclear Physics*, 35(C):81–219, 1995.

[9] Martin Beneke, Christoph Bobeth, and Robert Szafron. Enhanced Electromagnetic Corrections to the Rare Decay. *Physical Review Letters*, 120(1):011801, Jan 2018.

[10] Francis Halzen and Alan D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics.* John Wiley & Sons, 1984.

[11] David Griffiths. *Introduction to Elementary Particles.* Wiley-VCH Verlag GmbH & Co. KGaA, 2nd, revised edition, 2008.

[12] Nuno Leonardo. *Matter antimatter fluctuations, search, discovery and analysis of B_s flavor oscillations.* LAP, 2011.

[13] Patrick Koppenburg, Zdenek Dolezal, and Maria Smizanska. Rare decays of b hadrons. *Scholarpedia*, 11(6):32643, 2016.

[14] Y. Amhis. Averages of b-hadron, c-hadron, and tau-lepton properties as of summer 2016. *The European Physical Journal C*, 77(12):895, Dec 2017.

[15] The LHCb Collaboration. Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics. *Journal of High Energy Physics*, 2013(4):1–16, 2013.

[16] The CMS Collaboration. CMS Luminosity - Public Results. Accessed on April 6th 2018.

[17] CERN. LHC Machine Outreach. Accessed on April 6th 2018.

[18] The ATLAS Collaboration. 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(26):262001, 2015.