

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

Bruno Afonso Fontana Santos Alves

Thesis to obtain the Master of Science Degree in

Engineering Physics

Supervisors: Prof. Dr. Nuno Teotónio Viegas Guerreiro Leonardo
Prof. Dr. João Manuel Coelho dos Santos Varela

Examination Committee

Chairperson: Prof. Dr. Jorge Manuel Rodrigues Crispim Romão

Supervisor: Prof. Dr. Nuno Teotónio Viegas Guerreiro Leonardo

Members of the Committee: Prof. Dr. Patricia Conde Muíño

Prof. Dr. Nuno Filipe da Silva Fernandes de Castro

May 2018

Acknowledgments

I acknowledge LIP (*Laboratório de Instrumentação e Física Experimental de Partículas*) and the CMS-LIP group for providing everything necessary to complete this work. A special thanks goes to Prof. Dr. Nuno Leonardo, for his guidance and constant dedication to the project. The present document considerably improved thanks to his careful review. The computational framework used for the analysis here reported was based on previous code by Dr. Kai-Feng Chen (National Taiwan University), which was later modified by Bruno Galinhas (former PhD student at LIP). Bruno was responsible for automating the framework, which turned its usage into something simpler; a very significant number of plots here shown were obtained using this framework. I thank him for sharing the code and for explaining some of its details. I further thank the ones who helped me with the more technical parts of the work: Dr. Cristóvão da Cruz e Silva (LIP Postdoc), who gave lots of precious input regarding the software used by CMS, and Oleksii Toldaiev (LIP PhD student). I finally thank all CMS-LIP members who gave relevant input during the many meetings and discussions that regularly take place at LIP. This work would also not have been possible without the contribution of the technical and administrative staff at many other CMS institutes, and of the personnel working at the LHC Computing Grid.

Other people were also responsible for improving the work here presented. The first weeks were very enjoyable thanks to Giorgio Ghillardi (Master's student at *Università degli Studi di Torino*), who gave very important input regarding Monte Carlo processing and that discussed the main framework with me. I thank Dr. Kai-Feng Chen for his availability in discussing a yield-related problem that arose during the work. The 2016 Monte Carlo processing benefited from Dr. Silvio Donati (*Universität Zürich, Physik-Institut*) technical help and suggestions, while the Tag and Probe section of this work was only possible to the weights kindly shared by Dr. Leonardo Cristella (Postdoc at *Istituto Nazionale di Fisica Nucleare*).

I 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, both funded through *Fundação para a Ciência e Tecnologia* (FCT).

I conclude by thanking the ones who did not review my thesis, did not help me configuring the software and did have a hard time understanding why Particle Physics is so interesting, and which still supported me the most: my venerable and memorable parents, my curious and delirious brother, and my lively and lovely Ana, which stay with me wherever I go.

Resumo

O Modelo Padrão, pilar da Física de Partículas actual, apesar de extremamente bem sucedido nas suas previsões, possui limitações importantes. Vários modelos alternativos foram propostos; nenhuma incompatibilidade face às previsões do Modelo Padrão foi no entanto encontrada.

Decaimentos raros são extremamente sensíveis a hipotéticas interacções de Nova Física. A medição da fracção de vezes na qual o decaimento $B_s^0 \rightarrow \mu^+ \mu^-$ ocorre (*branching fraction* – BF) permite limitar os valores de certos parâmetros previstos por modelos teóricos alternativos e, simultaneamente, indicar qual a melhor forma de estudar processos de Nova Física.

Este trabalho, usando dados produzidos pelo acelerador *Large Hadron Collider* a uma energia do centro de massa de 13 TeV e medidos pelo detector *Compact Muon Solenoid*, visa a obtenção de um dos ingredientes necessários ao cálculo da fracção mencionada, nomeadamente, um rácio de fracções de fragmentação: f_s/f_d . A medição do número de eventos observados em três processos, de eficiências e de vários erros sistemáticos é feita. O valor final encontrado para o rácio é o seguinte: $f_s/f_d = 0.186 \pm 0.002$ (stat.) ± 0.021 (syst.) ± 0.016 (BF). Outros dois rácios foram medidos para garantir a consistência dos resultados: $f_s/f_u = 0.216 \pm 0.003$ (stat.) ± 0.029 (syst.) ± 0.017 (BF) e $f_d/f_u = 1.164 \pm 0.008$ (stat.) ± 0.120 (syst.) ± 0.057 (BF). Não foi encontrada qualquer dependência dos rácios na região de aceitação do detector, que é definida em termos de rapidez, $|y| < 2.25$, e momento transversal, $10 < p_T < 90$ GeV, dos mesões B.

Palavras-chave: *Large Hadron Collider*, *Compact Muon Solenoid*, Física dos mesões B, Fracção de Fragmentação, Teorias para além do Modelo Padrão, Decaimentos Raros.

Abstract

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 $B_s^0 \rightarrow \mu^+ \mu^-$ 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.) ± 0.021 (syst.) ± 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.) ± 0.029 (syst.) ± 0.017 (BF) and $f_d/f_u = 1.164 \pm 0.008$ (stat.) ± 0.120 (syst.) ± 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.

Keywords: Large Hadron Collider, Compact Muon Solenoid, B Physics, Fragmentation Fraction, Beyond the Standard Model theories, Rare Decays.

Contents

Acknowledgments	iii
Resumo	v
Abstract	vii
List of Tables	xi
List of Figures	xiii
Glossary	xxi
1 Introduction	1
1.1 Theoretical framework	1
1.1.1 Standard Model Overview	1
1.1.2 The Flavour Sector	2
1.2 B Mesons	3
1.2.1 Production	4
1.2.2 Lifetime	5
1.2.3 Mixing	6
1.3 Rare decays	8
1.3.1 The $B_s^0 \rightarrow \mu^+ \mu^-$ Branching Fraction	10
1.3.2 Fragmentation Fractions	11
1.4 Thesis Outline	17
2 Experimental Apparatus	19
2.1 CMS Coordinate System and Parametrization	19
2.2 Tracker	20
2.3 Muon Stations	22
2.4 Trigger	22
3 Data and MC Samples	25
3.1 Processing	26
3.2 Selection	28
4 Data MC Comparison	31
4.1 The sideband subtraction method	31

4.2	Mass resolution	35
5	Yields	39
5.1	Unbinned Extended Maximum Likelihood method	39
5.2	$K \pi$ swap	40
5.3	Invariant mass fits	42
6	Efficiencies	45
7	Systematics	51
7.1	Fit bias	51
7.2	MC reweighting	52
7.3	Mass window cuts	56
7.4	Tag and Probe	58
7.5	Fit variations	63
7.6	Other systematic sources	64
7.7	Summary	65
8	Results and Discussion	69
8.1	Kinematic dependence	69
8.2	Comparison with other results	71
8.3	Future Prospects	72
	Bibliography	74
A	Data and MC Comparison	80
A.1	B^+ channel	81
A.2	B^0 channel	83
A.3	B_s^0 channel	85
B	MC Reweighting	87
B.1	$B^0 \rightarrow J/\psi K^{*0}$ channel	87
B.2	$B_s^0 \rightarrow J/\psi \phi$ channel	92
C	Invariant mass fits	98
C.1	B^+ channel	98
C.2	B^0 channel	99
C.3	B_s^0 channel	100

List of Tables

1.1	Measured masses of the six known quarks [5].	3
1.2	Properties of B meson ground states. ‘S’ stands for spin and ‘P’ for parity. Values taken from [5].	4
1.3	Values of the branching fractions used in this work. The values were obtained from [5]. . .	14
1.4	Statistical plus systematic uncertainty sources for the CMS $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction measurement [36] (added quadratically). The 5% extra uncertainty present in the ratio refers to the CMS extrapolation of the LHCb measurement [35] already explained in the main text and also explained in Ref. [21]. For the sources which include two numbers, the first refers to the “Barrel” and the second to the “Endcap” part of the CMS detector (see Chapter 2).	16
3.1	AOD data samples used in this work. The dates refers to their online processing.	26
3.2	AODSIM Monte Carlo samples used in this work. The suspension points refer to the following strings: RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9_ext1-v1, for 2015, and RunIISummer16DR80Premix-PUMoriond17_80X_mcRun2_asymptotic_2016_TracheIV_v6_xt1-v2, for 2016. The substring TuneCUEP8M1_13TeV-pythia8-evtgen was abridged to XX. The specified size refers to the number of B mesons generated in each sample.	27
3.3	Values of the luminosities for the 2015 available data eras processed. The CMS Collaboration measured a value of 2.3% for the systematic uncertainty of 2015 luminosity measurements.	28
3.4	Values of the luminosities for the 2016 available data eras processed. The CMS Collaboration measured a value of 2.5% for the systematic uncertainty of 2016 luminosity measurements.	28
3.5	Summary of the selection criteria used in this work.	29
4.1	Signal and background regions used for the Sideband Subtraction Method. All values have GeV units.	32
4.2	Resolution values for B and double tracks invariant mass distributions. Values are in MeV.	36
6.1	Cuts imposed to the second MC sample used for the determination of the efficiency for each channel of interest. There is an exception: for the $B^+ \rightarrow J/\psi K^+$ channel no cut was imposed to the single kaon in the 2016 MC sample.	46

7.1	Global fit bias systematic uncertainty for the three channels of interest and for the three ratios.	52
7.2	Systematic double track window mass cut uncertainty, for the B^0 and B_s^0 channels.	56
7.3	Global fit bias systematic uncertainty for the measured fragmentation fraction ratios.	57
7.4	Soft Muon ID trigger systematic uncertainties in $p_T(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.	61
7.5	L1.L2 trigger systematic uncertainties in $p_T(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.	61
7.6	L3 trigger systematic uncertainties in $p_T(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.	62
7.7	Soft Muon ID trigger systematic uncertainties in $y(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.	62
7.8	L1.L2 trigger systematic uncertainties in $y(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.	62
7.9	L3 trigger systematic uncertainties in $y(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.	62
7.10	Global systematic uncertainties measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.	62
7.11	Branching fraction systematics for single channels and respective ratios.	65
7.12	Global relative systematic uncertainties for the three channels of interest and for the three measured fragmentation fraction ratios. The values are expressed in %.	66
7.13	Bin-dependent relative systematic uncertainties for the f_s/f_d ratio. The values shown are expressed in %.	66
7.14	Bin-dependent relative systematic uncertainties for the f_s/f_u ratio. The values shown are expressed in %.	67
7.15	Bin-dependent relative systematic uncertainties for the f_d/f_u ratio. The values shown are expressed in %.	68
8.1	Constant and linear fit parameters, for the three ratios under study, for $p_T(B)$ and $y(B)$ bins.	70

List of Figures

- 1.1 Feynman diagrams for the $B^+ \rightarrow J/\psi K^+$ and $B_s^0 \rightarrow J/\psi \phi$ decays (used in this work). Both diagrams include a spectator quark (the ‘up’ quark on the left and the ‘strange’ quark on the right). The red symbol ‘ V_{cb} ’ refers to the CKM matrix (see Eq. 1.4). 4
- 1.2 Box mixing diagram contributing to the $B_{(s)}^0 - \overline{B}_{(s)}^0$ mixing. Another possibility is obtained by rotating this diagram 90° 8
- 1.3 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) [21]. 9
- 1.4 Helicity suppression in the $B_s^0 \rightarrow \mu^+ \mu^-$ decay. Spin is represented by the thick brown arrows, while momenta by the thin red arrows. Two opposite spin orientations are possible. If the muons were massless, both possibilities would be forbidden, since the upper situation implies a left-handed anti-muon, and the bottom one implies a right-handed muon. This suppression mechanism exists for the $B^0 \rightarrow \mu^+ \mu^-$ decay too. 9
- 1.5 Search for the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays during the last 35 years. Markers without error bars denote upper limits on the branching fractions at 90% confidence level, while measurements are denoted with error bars delimiting 68% confidence intervals. The solid horizontal lines represent the SM predictions for the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ branching fractions; the blue (red) lines and markers are related to the $B_s^0 \rightarrow \mu^+ \mu^-$ ($B^0 \rightarrow \mu^+ \mu^-$) decay [21]. 11
- 1.6 Dimuon invariant mass $m_{\mu^+ \mu^-}$ distribution for B_s^0 and B^0 meson candidates. The data points (black) were superimposed by a combined unbinned maximum extended likelihood fit (solid blue). The vertical error bars denote a 68% confidence interval, while the horizontal error bars show the chosen binning. The fit includes several components: for the signal, we see the B^0 (blue area) and B_s^0 (yellow area) peaks, while for the background the semi-leptonic background (dotted salmon line), the combinatorial background (dash-dotted green line) and the peaking background (dashed violet line) were considered. The combinatorial background arises due to random combinations of two muons. The semi-leptonic background comes in turn from B_s^0 semi-leptonic decays in which some non-interacting particles (e.g.: neutrinos, neutral pions) were not detected. Finally, the peaking background appears due to two-body hadronic decays, such as $B^0 \rightarrow K^+ \pi^-$ [21]. 12

- 1.7 Likelihood contours for the ratios of the B_s^0 and B^0 branching fractions with respect to their SM prediction. The black cross marks the central value obtained. The SM point is shown as the red square located, by construction, at $S_{SM}^{B_s^0} = S_{SM}^{B^0} = 1$. Each contour encloses a region approximately corresponding to the reported confidence level. The SM branching fractions are assumed uncorrelated to each other, and the likelihood contours take the branching fraction uncertainties into account [21]. 12
- 1.8 Example of the constraints that can be imposed to BSM theories by studying the $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ rare decays. The correlation between these decays' branching fractions is shown considering Minimal Flavour Violation (MFV) models, a SM-like model with a fourth generation of elementary particles (SM4), the Randall-Sundrum model with custodial protection (RSc) and four SUSY flavour models. The grey area is ruled out experimentally. The SM point is marked by a star. This plot was retrieved from Ref. [31], which does not include the joined LHCb and CMS measurement. The current constraints are thus a bit tighter. That can be readily seen since a null value for the $B_s^0 \rightarrow \mu^+\mu^-$ branching fraction is no longer allowed by experiment (see Eq. 1.16). 13
- 1.9 Current status of f_s/f_d measurements. LEP, CDF (Collider Detector at Fermilab), LHCb, and ATLAS measurements are shown. The error bars are depicted with thin black lines. The theory error is represented by the thick red line for the first two measurements. The black dashed line with its uncertainty (green-shaded region) shows the HFAG average obtained using the blue points (the last three measurements) [34]. 15
- 1.10 Measurement of the f_s/f_d (left) and f_{Λ_b}/f_d (right) ratios as a function of $p_T(B)$. While the LHCb and LEP (Large Electron-Positron Collider) measurements see a clear p_T dependence on the f_{Λ_b}/f_d ratio, the same cannot be said regarding the f_s/f_d quantity, which ATLAS also measured. While LHCb (left plot, green points) reported a dependence, ATLAS (left plot, black points) did not find it. This work tries to shed some light on this inconsistency, also making use of the more recent datasets which allow to reach higher p_T values, and thus testing the dependence in a wider kinematic range [33]. 16
- 2.1 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 from other tracks (blue line). 20

2.2	$z-R$ transverse section of the CMS detector. The interaction point is at the lower left corner. The pseudorapidity η is indicated for different θ angles. The locations of both the muon stations and the steel disk are shown (dark grey areas). The 4 drift tube (DT, in light orange) stations are labelled MB (“muon barrel”) and the cathode strip chambers (CSC, in green) are labelled ME (“muon endcap”). Resistive plate chambers (RPC, in blue) are in both the barrel and the endcaps of CMS, where they are labelled RB and RE, respectively [41].	21
2.3	Dimuon mass distribution collected with various dimuon triggers at 13 TeV in 2016 with 13.1 fb^{-1} . The coloured paths correspond to dedicated dimuon triggers with low p_T thresholds, in specific mass windows, while the light grey continuous distribution represents events collected with a dimuon trigger with high p_T thresholds. Note the logarithmic scales. The J/ψ peak and B_s^0 window can be clearly identified.	23
4.1	B candidates invariant mass spectra for the three channels of interest: B^+ (top left), B^0 (top right) and B_s (bottom). These events passed the selection cuts of the analysis displayed in Table 3.5. The dashed lines indicate the regions considered for the SSM: the signal region is depicted in red, while the sidebands are shown in green.	33
4.2	Background exponential fits for the B^+ , B^0 and B_s channels (from left to right), in both logarithmic (top) and linear (bottom) scales. The λ is the parameter of the exponential, which was left free to oscillate during the fit. The whole dataset is shown to facilitate the understanding of the plot; the fit was performed to the sidebands only (blue).	33
4.3	L_{xy} significance (defined in the text) distribution for the $B_s^0 \rightarrow J/\psi \phi$ channel, for the nominal (left plot) and an alternative (right plot) sidebands. The SSM does not depend on the chosen sidebands, as long as the exponential fit to the background converges. The optional region of this example uses a range between 4.6 and 5.0 GeV on the left sideband, and between 5.75 and 6.15 GeV on the right sideband.	34
4.4	Data / MC comparison for some of the variables that did not show a very good agreement. In the first row the $p_T(B)$ and the $y(B)$ are shown for the $B^+ \rightarrow J/\psi K^+$ channel, although a similar behaviour can be seen for the other two channels. The second row refers to the $B^0 \rightarrow J/\psi K^{*0}$ channel only, since the behaviour of the same variables is quite good for the other channels (see Appendix A).	36
4.5	B meson invariant mass spectra for the three channels of interest: B^+ (top), B^0 (middle) and B_s^0 (bottom). The left column displays data while the right one refers to MC. The spectra were fitted with a maximum likelihood fit consisting of a sum of three Gaussians plus a Chebyshev polynomial to describe some asymmetries. The uncertainties are displayed in Table 4.2.	37

4.6	Double track invariant mass spectra: $K^{*0} \rightarrow K^+\pi^-$ channel (top) and $\phi \rightarrow K^+K^-$ (bottom). The data distributions were obtained after applying the SSM technique. The spectra were fitted in the central region with a maximum likelihood fit consisting of a sum of three Gaussians. The dashed red lines represent the mass window applied during the selection stage. The uncertainties are displayed in Table 4.2.	38
5.1	Unbinned maximum likelihood fit to the B^0 MC true (top row), swapped (middle row) and full (bottom row) signal invariant mass, both for linear (left column) and logarithmic (right column) scales. The full dataset corresponds to the sum of the true with swapped events. The cyan line describes the full PDF in each case. For the first two rows, the pink, yellow and green lines reflect the behaviour of each of the three Crystal Ball functions used. For the last row, the yellow line represents the true signal and the green line represents the swapped signal. For the case of the left column, the parameters obtained from the fit are shown.	41
5.2	Unbinned extended maximum likelihood fit for the B^+ , B^0 and B_s^0 channels, for the $16 < p_T < 20$ GeV bin.	42
6.1	Overall efficiencies of the three channels of interest as a function of $p_T(B)$ (left) and $y(B)$ (right) bins. From top to bottom, the B^+ , B^0 and B_s^0 channels are shown.	47
6.2	Efficiencies of the three fragmentation fraction ratios studied as a function of $p_T(B)$ (left) and $y(B)$ (right). From top to bottom, the $\varepsilon_u/\varepsilon_s$, $\varepsilon_u/\varepsilon_d$ and $\varepsilon_d/\varepsilon_s$ efficiency ratios are shown.	49
7.1	Pull distributions for the number of signal events for the B^+ channel, for all considered $p_T(B)$ bins. A similar analysis was performed for the B^0 and B_s^0 channels.	53
7.2	Example of the reweighting technique. On the left plot the original data/MC comparison for the L_{xy} significance variable is shown. On the right plot the same variable is corrected using the L_{xy} significance as the reweighting variable. By construction, the two distributions now perfectly match.	55
7.3	Double track invariant mass spectra: $K^{*0} \rightarrow K^+\pi^-$ channel (top) and $\phi \rightarrow K^+K^-$ channel (bottom); the left column refers to data and the right column to MC. The data distributions are obtained after applying the SSM technique. The spectra are fitted with a maximum likelihood fit consisting of a sum of three Gaussians plus a Chebyshev polynomial to describe some asymmetries.	57
7.4	Example of Tag and Probe study performed using early 2015 CMS data. The signal and background were described by different fit functions, and were simultaneously fitted to the J/ψ resonance peak. The upper-left plot shows the passing probes distribution, while the upper-right plots shows the failing probes distribution. The bottom plot depicts the distributions of the two groups combined. These plots were created by Leonardo Cristella.	60

7.5	Example of a subset of Tag and Probe weights which are used in this work. Here the weights refer to L3 efficiencies. The bins of rapidity shown are, from top left to bottom right: $0 < \eta(\mu) < 0.2$ (a), $0.2 < \eta(\mu) < 0.3$ (b), $0.3 < \eta(\mu) < 0.9$ (c) and $0.9 < \eta(\mu) < 1.2$ (d). The points were connected by a line. The weights were obtained by Leonardo Cristella.	60
7.6	Examples of modifications done to the nominal B^+ PDF in order to obtain systematic uncertainties for the measurement of the yields. In the top left plot the signal was fitted with a Gaussian. In the top right plot no $J/\psi \pi^+$ component was included. In the bottom left plot the $J/\psi \rightarrow h X$ component was fitted with a Gaussian. In the bottom right plot the mass window was increased to the left by 10%. The three plots contain data from the $25 \leq p_T(B) \leq 30$ GeV bin.	65
8.1	Fragmentation fraction ratios f_d/f_u (top), f_s/f_u (middle), f_s/f_d (bottom) as a function of $p_T(B)$ (left) and $ \eta(B) $ (right). The pink boxes denote statistical uncertainties only. The vertical error bars show the quadratic sum of statistical and systematic uncertainties. The global systematic uncertainties are not displayed as vertical error bars but rather appear indicated in black in the top left corner. The fit results for the constant and linear fits are displayed in the top right corner and repeated in Table 8.1.	70
8.2	Ratio of fragmentation fractions f_s/f_d as functions of $p_T(B)$ (left) and $\eta(B)$ (right). The errors on the data points are the statistical and uncorrelated systematic uncertainties added in quadrature. The solid line is the result of a linear fit, and the dashed line corresponds to the fit for the no-dependence hypothesis. The average value of $p_T(B)$ or $\eta(B)$ is determined for each bin and used as the centre of the bin. The horizontal error bars indicate the bin size. Note that the scale is zero suppressed [35].	71
8.3	Projected LHC performance until 2035, showing preliminary dates for long shutdowns and projected luminosities. The integrated luminosity is displayed according to a logarithmic scale [56].	72
8.4	CMS simulations for the expected invariant mass distributions before (left) and during (right) the High Luminosity LHC running. The plot on the left considers an integrated luminosity of 300 fb^{-1} , while the plot on the right reaches 3000 fb^{-1} . Both simulations only take the barrel region ($ \eta < 1.4$) into account. The $B^0 \rightarrow \mu^+ \mu^-$ peak will probably be disentangled from the $B_s^0 \rightarrow \mu^+ \mu^-$ peak, making its experimental observation possible [56].	73
A.1	Data / MC comparison for the $B^+ \rightarrow J/\psi K^+$ channel. The distributions correspond, from top left to bottom right, to the following variables: $p_T(B)$, $y(B)$, $p_T(\mu_1)$, $p_T(\mu_2)$, $\eta(\mu_1)$ and $\eta(\mu_2)$	81
A.2	Data / MC comparison for the $B^+ \rightarrow J/\psi K^+$ channel. The distributions correspond, from top left to bottom right, to the following variables: $p_T(\text{track}_1)$, $\eta(\text{track}_1)$, L_{xy} , σ_{xy} , L_{xy}/σ_{xy} , the vertex probability, $\cos(\alpha_{2D})$ and the proper time.	82

A.3	Data / MC comparison for the $B^0 \rightarrow J/\psi K^{*0}$ channel. The distributions correspond, from top left to bottom right, to the following variables: $p_T(B)$, $y(B)$, $p_T(\mu_1)$, $p_T(\mu_2)$, $\eta(\mu_1)$, $\eta(\mu_2)$, $p_T(\text{track}_1)$ and $p_T(\text{track}_2)$	83
A.4	Data / MC comparison for the $B^0 \rightarrow J/\psi K^{*0}$ channel. The distributions correspond, from top left to bottom right, to the following variables: $\eta(\text{track}_1)$, $\eta(\text{track}_2)$, L_{xy} , σ_{xy} , L_{xy}/σ_{xy} , the vertex probability, $\cos(\alpha_{2D})$ and the proper time.	84
A.5	Data / MC comparison for the $B_s^0 \rightarrow J/\psi \phi$ channel. The distributions correspond, from top left to bottom right, to the following variables: $p_T(B)$, $y(B)$, $p_T(\mu_1)$, $p_T(\mu_2)$, $\eta(\mu_1)$, $\eta(\mu_2)$, $p_T(\text{track}_1)$ and $p_T(\text{track}_2)$	85
A.6	Data / MC comparison for the $B_s^0 \rightarrow J/\psi \phi$ channel. The distributions correspond, from top left to bottom right, to the following variables: $\eta(\text{track}_1)$, $\eta(\text{track}_2)$, L_{xy} , σ_{xy} , L_{xy}/σ_{xy} , the vertex probability, $\cos(\alpha_{2D})$ and the proper time.	86
B.1	MC reweighting applied to the $p_T(B)$ (top group) and $y(B)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	87
B.2	MC reweighting applied to the $p_T(\mu_1)$ (top group), $p_T(\mu_2)$ (middle group) and $\eta(\mu_1)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	88
B.3	MC reweighting applied to the $p_T(\mu_2)$ (top group), $p_T(\text{track}_1)$ (middle group) and $p_T(\text{track}_2)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	89
B.4	MC reweighting applied to the $p_T(\mu_2)$ (top group), $p_T(\text{track}_1)$ (middle group) and $p_T(\text{track}_2)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	90
B.5	MC reweighting applied to the σ_{xy} (top group), L_{xy}/σ_{xy} (middle group) and vertex probability (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	91
B.6	MC reweighting applied to the $\cos(\alpha_{2D})$ variable. The top left plot shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	92

B.7	MC reweighting applied to the $p_T(B)$ variable. The top left plot shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance. . . .	92
B.8	MC reweighting applied to the $y(B)$ (top group), $p_T(\mu_1)$ (middle group) and $p_T(\mu_2)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	93
B.9	MC reweighting applied to the $\eta(\mu_1)$ (top group), $\eta(\mu_1)$ (middle group) and $\eta(\text{track}_1)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	94
B.10	MC reweighting applied to the $p_T(\text{track}_2)$ (top group), $\eta(\text{track}_1)$ (middle group) and $\eta(\text{track}_2)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	95
B.11	MC reweighting applied to the L_{xy} (top group), σ_{xy} (middle group) and L_{xy}/σ_{xy} (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	96
B.12	MC reweighting applied to the vertex probability (top group) and $\cos(\alpha_{2D})$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.	97
C.1	Unbinned extended maximum likelihood fit for the B^+ channel, using bins of p_T . The same kind of fits were performed for $y(B)$ bins.	98
C.2	Unbinned extended maximum likelihood fits for the B^0 channel, using bins of p_T . The same kind of fits were performed for $y(B)$ bins.	99
C.3	Unbinned extended maximum likelihood fits for the B_s^0 channel, using bins of p_T . The same kind of fits were performed for $y(B)$ bins.	100

Glossary

AOD	Analysis Object Data
BSM	Beyond Standard Model
CB function	Crystal Ball function
CKM matrix	Cabibbo–Kobayashi–Maskawa matrix
CMSSW	Compact Muon Solenoid SoftWare
CMS	Compact Muon Solenoid experiment
CSC	Cathode Strip Chamber
DT	Drift Tube
FCNC	Flavour Changing Neutral Currents
HLT	High Level Trigger
LHC	Large Hadron Collider
LIP	<i>Laboratório de Instrumentação e Física Experimental de Partículas</i>
MIP	Minimally ionizing particle
NP	New Physics
PDF	Probability density function
PD	Primary dataset
QCD	Quantum Chromodynamics
RPC	Resistive Plate Chamber
SM	Standard Model
SSM	Sideband Subtraction Method
VEV	Vacuum Expectation Value

Chapter 1

Introduction

1.1 Theoretical framework

1.1.1 Standard Model Overview

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] at the Large Hadron Collider (LHC). The SM is a well tested theory in describing particle interactions up to TeV scales.

The SM formalism is built around a $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge invariant Lagrangian which includes several Lorentz invariant terms, where ‘Y’ stands for hypercharge, ‘L’ refers to the left chirality component of a fermion field and ‘C’ is the colour quantum number [3]. The electroweak $SU(2)_L \otimes U(1)_Y$ group is spontaneously broken into a $U(1)_Q$ group by the scalar Higgs field which acquires a non-zero vacuum expectation value (VEV). The latter group has the electric charge ‘Q’ as generator. Out of the four gauge bosons contained in the electroweak group, two of them linearly combine to form the Z_μ neutral massive boson and the massless photon (A_μ) using the so-called weak mixing angle as a rotation parameter [4]. The two remaining bosons form the massive W_μ^+ and W_μ^- charged gauge bosons. The mass of these bosons is experimentally measured to be $m_W = 80.385 \pm 0.015$ GeV and $m_Z = 91.1876 \pm 0.0021$ GeV [5]. Fermions are organised according to their chirality: while left eigenvectors of the chirality are grouped into isospin doublets, the right-handed fermions are $SU(2)_L$ singlets. For the specific case of neutrinos, only the left eigenvectors exist in the SM; their right counterpart was never observed [6]. As a consequence, according to the SM, they are massless, despite the fact that the observed neutrino oscillations show that neutrinos possess a tiny mass (< 1 eV) [7]. Leptons are $SU(3)_C$ colour singlets, which means they can be described by the electroweak group only. This is not the case for quarks, as they interact strongly with other quarks: Quantum Chromodynamics (QCD) is required. Besides, quarks remain confined; they always group together to form hadrons, *i.e.*, groups of two (bound quark-antiquark pairs with integer spin called mesons), three (the baryons, with half-integer spin) or even more quarks,

the so-called exotic states [3]. As an example, the J/ψ and ϕ mesons, which will be often mentioned in this work, are bound states of, respectively, $c\bar{c}$ (Charmonium) and $s\bar{s}$ pairs of quarks.

Despite all of its predictive success, the SM is known to be incomplete. Several factors strongly motivate the existence of New Physics (NP). The SM does not account for gravity, it does not provide a dark matter candidate and it cannot explain the observed baryon asymmetry in the Universe (the CP violating phase in the SM cannot justify the magnitude of the asymmetry [6]). In addition, all of its mass scales and corresponding hierarchies are not predicted and need to be fixed by experiment. Other hierarchies problems exist, namely in the computation of the Higgs boson mass: an extremely unlikely fine-tuning would be necessary in order for the Higgs boson to obtain a 125 GeV mass, since loop corrections to the mass of the Higgs boson are quadratically divergent. Furthermore, the SM obeys to a specific gauge group for no apparent reason, it is not left-right symmetric and imposes no definite constraint on the number of generations [8, 9, 10].

1.1.2 The Flavour Sector

Fermionic interactions are encapsulated into the Lagrangian of the SM and arise from its kinetic term after spontaneous symmetry breaking takes place [11]. When the electric charges of the two interacting fermions differs by either $+1$ or -1 (in units of the electron electric charge), the interaction is known as charged:

$$\mathcal{L}_W = -\frac{g}{\sqrt{2}}(W_\mu^+ \bar{\nu}_L \gamma^\mu l_L + W_\mu^- \bar{l}_L \gamma^\mu \nu_L) + \text{h.c.}, \quad (1.1)$$

where the first term refers to leptons and the second one to quarks, g is the $SU(2)$ gauge coupling and ‘h.c.’ is a short-hand notation for ‘hermitian conjugate’. The $\bar{\nu}_L$, l_L , $\bar{l}_L$ and n_L symbols represent three-dimensional vectors that live in the family space, composed by the three known generations. Each of their components is a Dirac spinor. The p_L and n_L symbols refer to the positively and negatively charged components of the left-handed weak isospin quark doublet. In the same way, the ν_L and l_L symbols refer to the neutrino-like and electron-like component of the $SU(2)_L$ lepton doublet. In the case where the charges of the interacting particles are equal, the interaction is neutral, and is mediated both by the Z_μ and by the A_μ . Except for the multiplicative constant, the term in the Lagrangian is similar to (1.1), but now interacting leptons and quarks have equal electric charges.

Besides generating masses for the W and Z bosons, the scalar Higgs field can in addition be used to write $SU(2)_L \otimes U(1)_Y$ invariant mass terms which give mass to leptons and quarks after spontaneous symmetry breaking takes place [4], therefore solving the problem that an explicit Dirac mass term is not invariant under the aforementioned group [11]:

$$\mathcal{L}_{\text{mass}} = -\bar{n}_L M_n n_R - \bar{p}_L M_p p_R - \bar{l}_L M_l l_R + \text{h.c.}, \quad (1.2)$$

where a neutrino mass term cannot be written without right-chirality neutrinos. The M_n , M_p and M_l mass matrices are in general not diagonal, which means that the quarks in (1.2) are not mass eigenstates. It is nonetheless always possible to perform a unitary bidiagonalization of the left-handed and right-

handed quark family vectors such that the mass matrices become diagonal [11]. By performing the same transformation on the quark term of the Lagrangian in (1.1), a mixing matrix appears:

$$\mathcal{L}_W^{(q)} = \frac{g}{\sqrt{2}}(\mathbf{W}_\mu^+ \bar{u}_L \gamma^\mu V_{CKM} d_L + \mathbf{W}_\mu^- \bar{d}_L \gamma^\mu V_{CKM}^\dagger u_L), \quad (1.3)$$

where V_{CKM} is the unitary quark flavour mixing Cabibbo–Kobayashi–Maskawa (CKM) matrix, and u and d are mass eigenstates representing, respectively, the “up-type” quarks (u, c, t) and the “down-type” quarks (d, s, b). Their measured masses are shown in Table 1.1. The neutral quark current is not affected by the mass bidiagonalization. This implies that there is no flavour mixing in the Lagrangian term related to neutral interactions; in the SM no flavour changing neutral currents (FCNC) exist at tree level. As a consequence, FCNC can only be induced by higher-order diagrams making them sensitive to Beyond Standard Model (BSM) physics [12]. Since there are no right chirality neutrinos in the SM, it is always possible to rephase the left-handed neutrinos in such a way as to avoid any mixing in the leptonic sector. Thus, there are no lepton-flavour changing interactions in the SM, both charged or neutral. In reality these interactions may happen, but they will be highly suppressed, due to the small neutrino masses.

The CKM matrix acts on the family space, and its entries have the following experimentally determined magnitudes [5]:

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97417 \pm 0.00021 & 0.2248 \pm 0.0006 & (4.09 \pm 0.39) \times 10^{-3} \\ 0.220 \pm 0.005 & 0.995 \pm 0.016 & (40.5 \pm 1.5) \times 10^{-3} \\ (8.2 \pm 0.6) \times 10^{-3} & (40.0 \pm 2.7) \times 10^{-3} & 1.009 \pm 0.031 \end{pmatrix} \quad (1.4)$$

where each V_{ij} suggests how much a charged interaction between the i and j quarks will be suppressed. The off-diagonal entries of the CKM matrix are significantly smaller than the diagonal entries, except for the V_{cd} and V_{us} entries. The CKM matrix, being a three by three unitary matrix, must include a complex phase which is the source of CP violation in the Standard Model [13].

Table 1.1: Measured masses of the six known quarks [5].

Quark	Mass	Quark	Mass	Quark	Mass
u	$2.2_{-0.4}^{+0.6} \text{ MeV}$	c	$1.28 \pm 0.03 \text{ GeV}$	t	$173.1 \pm 0.6 \text{ GeV}$
d	$4.7_{-0.4}^{+0.5} \text{ MeV}$	s	96_{-4}^{+8} MeV	b	$4.18_{-0.03}^{+0.04} \text{ GeV}$

1.2 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 too. B mesons are hadrons composed by two quarks, being one of them an anti- b quark, $\bar{b}$, and the other a different quark¹ (Table 1.2 shows some of their properties). 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 $\bar{b} \rightarrow \bar{c}W^+$, where the W boson is virtual (the CKM otherwise favoured $\bar{b} \rightarrow \bar{t}W^+$ is forbidden due to the large mass of the top²). Then, the W materializes either into a pair of leptons (semileptonic decays), or into a pair of quarks which hadronize. In Figure 1.1 we show two of the latter mentioned decays:

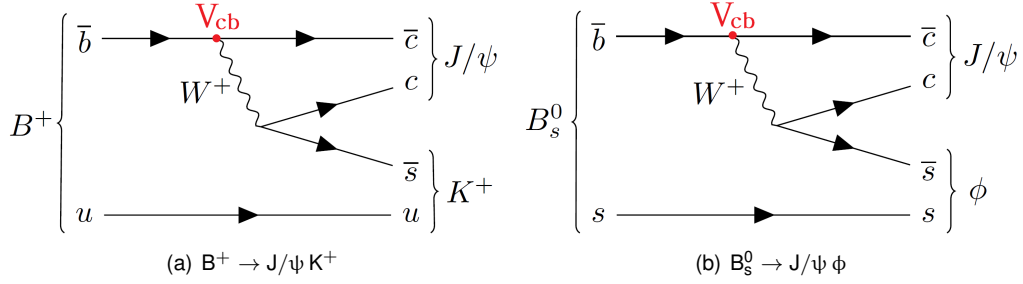


Figure 1.1: Feynman diagrams for the $B^+ \rightarrow J/\psi K^+$ and $B_s^0 \rightarrow J/\psi \phi$ decays (used in this work). Both diagrams include a spectator quark (the ‘up’ quark on the left and the ‘strange’ quark on the right). The red symbol ‘ V_{cb} ’ refers to the CKM matrix (see Eq. 1.4).

B meson decay rates are used to measure the couplings between different generation of quarks ($|V_{cb}|$, $|V_{ub}|$, $|V_{ts}|$ and $|V_{td}|$ CKM elements) and to test the CKM matrix unitarity. Besides, these mesons represent an exceptional tool to measure CP violation: their decays are used to measure CP-violating quantities. The SM predicts large CP asymmetries in B meson systems [14].

Table 1.2: Properties of B meson ground states. ‘S’ stands for spin and ‘P’ for parity. Values taken from [5].

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

1.2.1 Production

The b quark production cross section in hadron collisions has been up to now computed at next-to-leading order in perturbative QCD. Although the heavy-quark masses are larger than the intrinsic QCD energy scale, *i.e.*, the typical scale at which non-perturbative effects become important, the strong coupling constant is still small, and perturbative computations of short-distance effects are possible.

¹When the other quark is a b the resulting neutral bound state belongs to the Bottomonium family (the $\eta_b(1S)$ or the $\Upsilon(1S)$, for example), and it is not considered to be a B meson. This is just a definition.

²The so-called “Penguin” diagrams include this possibility, but the t becomes a virtual particle.

The latter have to be separated from (non-perturbative) long-distance effects. The study of heavy flavour production therefore provides an increased understanding of QCD in different dynamical regimes [12, 15, 16]. Strong interaction dynamics can be analysed, and deviations from QCD predictions may hint at NP effects.

Production is quantitatively determined through cross section measurements. In order to study its behaviour according to a given kinematic observable, one usually determines the differential cross section, *i.e.*, the cross section as a function of a certain variable. The transverse momentum, denoted p_T , or the rapidity (see Section 2.1) are commonly used. The differential cross section involving B mesons is given by [17]:

$$\frac{d\sigma(pp \rightarrow B X)}{dx} = \frac{N}{2\mathcal{A}\epsilon\mathcal{B}L\Delta x}, \quad (1.5)$$

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 $B_s^0 \rightarrow J/\psi \phi$, $J/\psi \rightarrow \mu^+\mu^-$ and $\phi \rightarrow K^+K^-$ 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 $\overline{B}_s^0$.

1.2.2 Lifetime

The CKM matrix (V_{cb} and V_{ub}) suppresses the available decay channels for hadrons containing a b quark. The B lifetime is therefore long enough so that the primary vertex, *i.e.*, the point where the two protons collide and where b -hadrons are produced, and the secondary vertex, *i.e.*, the point where the b -hadron decays, can be spatially separated³. B lifetimes measurements are key in extracting the weak parameters that are important for understanding the role of the CKM matrix in CP violation, such as the determination of V_{cb} and mixing parameters [5].

The relation between proper decay time and decay length is given by:

$$t = \frac{L}{\gamma v}, \quad \text{with } \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}, \quad (1.6)$$

where t is the proper decay time, L is the decay length, γ is the Lorentz factor, c is the speed of light and v is the velocity of the particle in the laboratory frame. The proper decay time follows an exponential distribution reproducing the exponential decay law; the lifetime is the inverse of the decay constant of the distribution. For a B meson momentum of 100 GeV, the decay length observed in the detector will be of around 0.85 cm. Despite travelling a relatively large distance, the vast majority of B mesons will decay before reaching the first layer of the Compact Muon Solenoid (CMS) detector, which lies at 4.4 cm from the collision point (see Section 2)⁴.

³This is the basis for the “ b -tagging” technique, which must be used in order to identify jets coming from a b , when its individual decay products cannot be identified.

⁴It would be really difficult to track a particle without any hits in the first silicon layers of the detector!

1.2.3 Mixing

Neutral mesons are known to oscillate between their particle and anti-particle states. A notable exception is the neutral pion (bound state of $u\bar{u}$ or $d\bar{d}$), since it is its own anti-particle. On the contrary, due to the conservation of electric charge, charged mesons do not oscillate. The neutral meson oscillations are mediated by the weak force through box diagrams (see Figure 1.2). In this section P^0 is a flavour state and may refer to either K^0 ($d\bar{s}$), D^0 ($c\bar{u}$), B^0 or B_s^0 mesons. Under certain conditions [11], the wave function of a neutral meson can be written as a superposition of particle and anti-particle states,

$$|\psi(t)\rangle = \psi_1(t) |P^0\rangle + \psi_2(t) |\overline{P^0}\rangle, \quad (1.7)$$

while its time evolution is given by the Schrödinger equation:

$$i \frac{d}{dt} |\psi(t)\rangle = \mathcal{H} |\psi(t)\rangle, \quad (1.8)$$

where the Hamiltonian $\mathcal{H}$ of the system is in general not diagonal and cannot be hermitian, otherwise the mesons would not decay, but merely oscillate [11]. The Hamiltonian is broken into two pieces, $\mathcal{H} = \mathbf{M} - i\mathbf{\Gamma}/2$, where both $\mathbf{M}$ and $\mathbf{\Gamma}$ are hermitian. The time evolution can be expressed using a mass diagonal basis. With this in mind we introduce the following basis transformation [9, 18]:

$$\begin{pmatrix} |P_H\rangle \\ |P_L\rangle \end{pmatrix} = \begin{pmatrix} p & -q \\ p & q \end{pmatrix} \begin{pmatrix} |P^0\rangle \\ |\overline{P^0}\rangle \end{pmatrix}. \quad (1.9)$$

where the complex p and q coefficients obey the $|p|^2 + |q|^2 = 1$ normalization relation. H and L refer to heavy and light mass states. It should be noted that the mass states in Eq. 1.9, as opposed to the flavour states, are not necessarily orthogonal, *i.e.*, $\langle P_H | P_L \rangle$ is zero only when $|q/p| = 1$. The eigenvalues of the diagonalized Hamiltonian matrix are defined to be $\mu_{H,L} = m_{H,L} - i\Gamma_{H,L}/2$. We also define the following quantities:

$$\begin{aligned} m &= (m_H + m_L)/2, & \Delta m &= m_H - m_L, \\ \Gamma &= (\Gamma_H + \Gamma_L)/2, & \Delta \Gamma &= \Gamma_H - \Gamma_L, \\ \Delta \mu &= \mu_H - \mu_L. \end{aligned} \quad (1.10)$$

Using Eq. 1.8 and converting the system back to the flavour basis, the time evolution can be calculated with:

$$\begin{aligned} |P^0(t)\rangle &= g_+(t) |P^0\rangle + \frac{q}{p} g_-(t) |\overline{P^0}\rangle, \\ |\overline{P^0}(t)\rangle &= \frac{p}{q} g_-(t) |P^0\rangle + g_+(t) |\overline{P^0}\rangle, \end{aligned} \quad (1.11)$$

where

$$g_{\pm}(t) = e^{-imt} e^{-\Gamma t/2} \begin{cases} \sin\left(\frac{\Delta \mu t}{2}\right) \\ \cos\left(\frac{\Delta \mu t}{2}\right) \end{cases}. \quad (1.12)$$

The probabilities for the transitions to happen are obtained taking the moduli squared of the transition amplitudes [9]:

$$\begin{aligned} \left| \langle P^0 | \mathcal{H} | \overline{P^0} \rangle \right|^2 &= \left| \frac{p}{q} \right|^2 |g_-(t)|^2, \\ \left| \langle \overline{P^0} | \mathcal{H} | P^0 \rangle \right|^2 &= \left| \frac{q}{p} \right|^2 |g_-(t)|^2, \\ \left| \langle P^0 | \mathcal{H} | P^0 \rangle \right|^2 &= \left| \langle \overline{P^0} | \mathcal{H} | \overline{P^0} \rangle \right|^2 = |g_+(t)|^2, \end{aligned} \quad (1.13)$$

where

$$|g_{\pm}(t)|^2 = \frac{e^{-\Gamma t}}{2} \left[\cosh \left(\frac{\Delta\Gamma}{2} t \right) \pm \cos(\Delta m t) \right]. \quad (1.14)$$

It is now clear that the $P^0 \rightarrow \overline{P^0}$ and $\overline{P^0} \rightarrow P^0$ rates are not the same, as long as $|q/p| \neq 1$. This indicates CP violation.

The mixing occurring in $B_{(s)}^0$ mesons⁵ is determined by the mass difference between the heavy and light mass eigenstates. The frequencies are calculated assuming that no CP violation is present ($|q/p| = 1$) and that $\Delta\Gamma/\Gamma$ is negligible [9]. The following approximated transition probabilities are fitted:

$$\begin{aligned} \mathcal{P}(P^0 \rightarrow \overline{P^0}) &= \mathcal{P}(\overline{P^0} \rightarrow P^0) = \frac{\Gamma}{2} e^{-\Gamma t} [1 - \cos(\Delta m t)], \\ \mathcal{P}(P^0 \rightarrow P^0) &= \mathcal{P}(\overline{P^0} \rightarrow \overline{P^0}) = \frac{\Gamma}{2} e^{-\Gamma t} [1 + \cos(\Delta m t)], \end{aligned} \quad (1.15)$$

yielding the below frequencies [5]:

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

The mixing can have an impact on branching fraction and lifetime measurements. Indeed, while the relative difference between decay widths of the two B^0 mass eigenstates is truly negligible, $\Delta\Gamma_d/\Gamma_d = -0.003 \pm 0.015$, the width difference for strange flavoured B mesons is important: $\Delta\Gamma_s/\Gamma_s = 0.124 \pm 0.011$ [5]. In experiments the branching fraction is usually extracted from the total event yield, which does not contain information regarding the mixing which is taken into account in theoretical calculations. One has therefore to be careful when comparing branching fraction values [19].

The amount of CP violation present in $B_{(s)}^0 - \overline{B_{(s)}^0}$ mixing, given by the CP-violating weak phase ϕ_s , is very small ($|q/p|$ is close to unity as assumed in Eq. 1.15) [20]. As a consequence, the mass and CP eigenstates almost coincide. As such, there are certain decays that essentially come from one mass eigenstate only. A relevant example is having a muon pair as final state, since only the heavy mass eigenstate can decay into it [21]. Thus, according to the SM, the effective lifetime⁶ $\tau = 1/\Gamma$ of the $B_s^0 \rightarrow \mu^+ \mu^-$ process should be equal to the B_H lifetime. This is not necessarily true in BSM scenarios, where the light mass state can also decay into a muon pair. The lifetime of the $B_s^0 \rightarrow \mu^+ \mu^-$ decay is therefore a good probe for BSM searches [21, 23].

⁵The $B_{(s)}^0$ notation refers to both B^0 and B_s^0 .

⁶The “effective lifetime” here refers to a weighted average of the lifetimes of the two mass eigenstates. See, for example, Ref. [22].

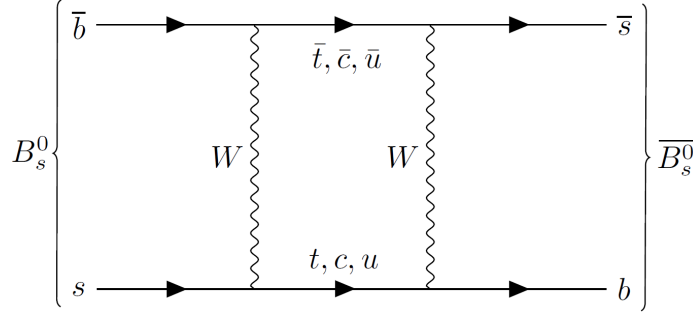


Figure 1.2: Box mixing diagram contributing to the $B_{(s)}^0 - \overline{B}_{(s)}^0$ mixing. Another possibility is obtained by rotating this diagram 90° .

1.3 Rare decays

The ground states of $B_{(s)}^0$ mesons, like the ground states of the other b -flavoured hadrons, are known to be unstable particles which decay through the weak interaction [5]. Among all the existing B meson decay channels, the ones which do not occur through the $\bar{b} \rightarrow \bar{c} W^+$ transition are denoted “rare” [14]. The $b \rightarrow s \gamma$ and $b \rightarrow s l^+ l^-$ quark decays are very good examples, where $l^\pm$ refers to a charged lepton. The inclusive branching fraction of the former B decay to a photon plus anything is in good agreement with SM predictions, imposing strong constraints on NP, in particular supersymmetry. The latter decay is currently drawing attention from the scientific community since the lepton universality ratio, $\mathcal{B}(B^+ \rightarrow K^+ \mu^+ \mu^-) / \mathcal{B}(B^+ \rightarrow K^+ e^+ e^-)$, was measured to be $0.745^{+0.090}_{-0.074} \text{ (stat.)} \pm 0.036 \text{ (syst.)}$ by LHCb [24], which corresponds to a 2.6σ distance from unity [10].

This work focuses on another interesting possibility: the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays. These decay channels have an extremely small and precisely predicted branching fraction [25, 26]:

$$\begin{aligned} \mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)_{\text{SM}} &= (3.57 \pm 0.17) \times 10^{-9}, \\ \mathcal{B}(B^0 \rightarrow \mu^+ \mu^-)_{\text{SM}} &= (1.06 \pm 0.09) \times 10^{-10}, \end{aligned}$$

where the uncertainties are dominated by the decay constants of the B mesons and by the CKM matrix elements [10]. The two aforementioned decays occur through loop processes, since FCNCs are not allowed at tree-level in the SM (see Figure 1.3). Besides, both decays are CKM suppressed. There is an additional suppression factor due to the helicity, which is a good quantum number for all particles described by the wave-function ψ which satisfies the Dirac equation $(i\gamma^\mu \partial_\mu - m)\psi = 0$. It is defined as the projection of the spin along the momentum direction, and it coincides with the chirality in the massless regime. All interactions mediated by the weak interaction deal with left chirality particles (or right chirality anti-particles) only, since parity is maximally violated by the weak force (see Eq. 1.1). For the specific case of the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays there is always either a right-handed muon or a left-handed anti-muon (see Figure 1.4). The handedness refers to the helicity and not to the chirality, and as such there is no law that forbids the decay to happen. Still, would the muon be massless, and the helicity and chirality would coincide, turning the decay into something impossible to occur. Since the mass of the muon is not zero, the decay exists, but its branching fraction is nevertheless suppressed by a

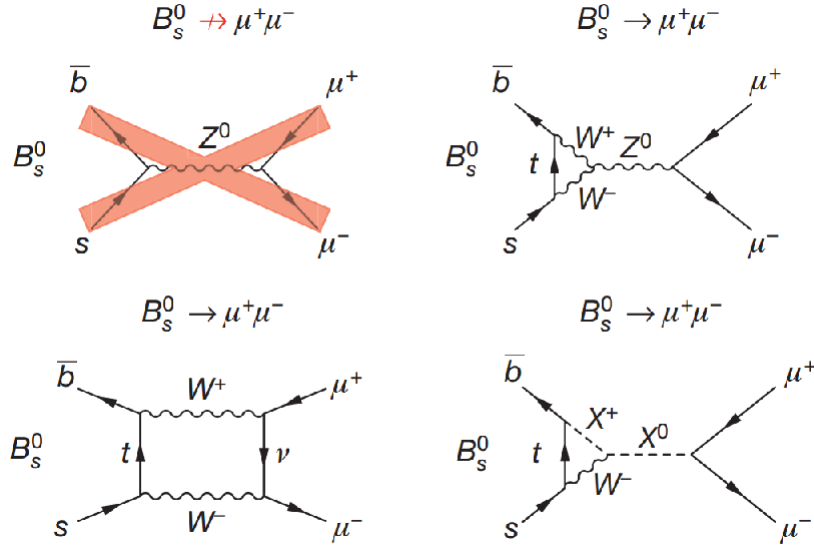


Figure 1.3: The SM does not allow FCNC at tree-level (upper left). Rare decays only occur through higher-order diagrams (upper right and bottom left). BSM physics can contribute to rare processes amplitudes (bottom right) [21].

$(m_\mu/m_{B_s^0})^3$ factor [4, 6, 10]; the $B_s^0 \rightarrow e^+e^-$ decay is for this reason out of reach in the foreseeable future⁷. 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. The search for NP is therefore facilitated if the SM processes under study are also suppressed, which is the case in rare decays. The $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decay channels represent powerful discriminants to check the existence of BSM theories, in particular models containing additional Higgs bosons [9, 10]. As an example, the $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction imposes constraints on the two-Higgs-doublet model, both in the mass of the two predicted charged Higgs bosons and in the ratio of the two VEVs [27].

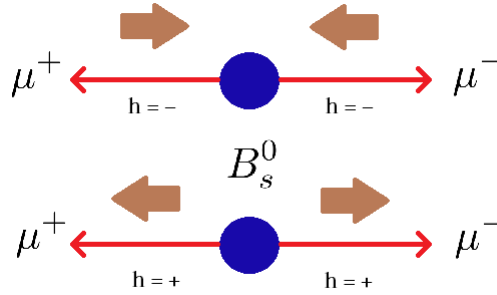


Figure 1.4: Helicity suppression in the $B_s^0 \rightarrow \mu^+ \mu^-$ decay. Spin is represented by the thick brown arrows, while momenta by the thin red arrows. Two opposite spin orientations are possible. If the muons were massless, both possibilities would be forbidden, since the upper situation implies a left-handed anti-muon, and the bottom one implies a right-handed muon. This suppression mechanism exists for the $B^0 \rightarrow \mu^+ \mu^-$ decay too.

SM predicted branching fractions might be different according to certain NP scenarios (see bottom right plot in Fig. 1.3). The possibility of finding an unexpected excess relative to the SM is what drives the measurement of highly suppressed branching fractions. For example, under certain conditions, the

⁷The reconstruction and the mass resolution of the electrons is also not as good as for the muons.

$B_s^0 \rightarrow \mu^+ \mu^-$ decay rate can be enhanced by over two orders of magnitude in the minimal supersymmetric model [28]. In the last thirty-five years, eleven collaborations at different colliders have searched for the $B_{(s)}^0 \rightarrow \mu^+ \mu^-$ decays (see Figure 1.5). An excess was not found⁸. As the amount of available data increased, the upper limits imposed to the decays became tighter and tighter. In 2015 the CMS and the LHCb Collaborations joined their efforts to observe the $B_s^0 \rightarrow \mu^+ \mu^-$ decay for the first time using data collected at a 7 and 8 TeV centre of mass collision energy (2011 and 2012), with a statistical significance larger than 5σ , which is the required threshold to claim an observation (see Figure 1.6) [21]. The $B^0 \rightarrow \mu^+ \mu^-$ mass peak did not achieve the same required significance; it was not observed so far. The final $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction measurement yielded:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (2.8_{-0.8}^{+0.7}) \times 10^{-9}. \quad (1.16)$$

For completeness, we report the most up-to-date values for the $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction measured by LHCb and ATLAS, respectively, $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = [3.0 \pm 0.6 \text{ (stat.) }_{-0.2}^{+0.3} \text{ (syst.)}] \times 10^{-9}$ [23] and $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (0.9_{-0.8}^{+1.1}) \times 10^{-9}$ [30], where the uncertainties of the last value already include the statistical and systematic uncertainties. The same collaborations imposed upper limits on the $B^0 \rightarrow \mu^+ \mu^-$ branching fraction, respectively, $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) < 3.4 \times 10^{-10}$ and $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) < 4.2 \times 10^{-10}$ at a 95% confidence level. The measured branching fractions are compatible with SM predictions (see Figure 1.7).

The CMS Collaboration is currently involved in the precise measurement of the same quantity⁹ using data collected at a 13 TeV centre of mass collision energy. This measurement will provide strong constraints on the parameters of BSM theories in case the $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction is measured to be compatible with the SM predictions (see Fig. 1.8). If not, a direct indication of NP will be found.

1.3.1 The $B_s^0 \rightarrow \mu^+ \mu^-$ Branching Fraction

Defining $P \rightarrow Y$ as the decay of a certain particle ‘P’ to a given final state ‘Y’, and considering that ‘X’ represents all possible final states to which ‘P’ can decay to, the branching fraction is defined as:

$$\mathcal{B}(P \rightarrow Y) \equiv \frac{N_{P \rightarrow Y}}{N_{P \rightarrow X}}, \quad (1.17)$$

where ‘N’ is the yield. The quantity in the denominator cannot be measured in practice. A normalization channel is therefore needed. The normalization channel is usually topologically similar to the one studied, since some systematic uncertainties cancel in the branching fraction calculation. To allow a precise measurement, the normalization channel should also have high yields, *i.e.*, the decay products of that particular channel should be detected often. The $B^+ \rightarrow J/\psi K^+$ channel is used, with $J/\psi \rightarrow \mu^+ \mu^-$ (see Table 1.3). We note that this channel, besides the two muons that also appear in the decay under

⁸Interestingly, the $\tau_{B_s^0 \rightarrow \mu^+ \mu^-}$ previously mentioned may still reveal New Physics effects even if the corresponding branching fraction is close to the prediction of the SM [29].

⁹LHCb covers a different pseudorapidity region: $2 < |\eta| < 5$. CMS studies the $0 < |\eta| < 2.5$ region. Pseudorapidity is defined in Section 2.1.

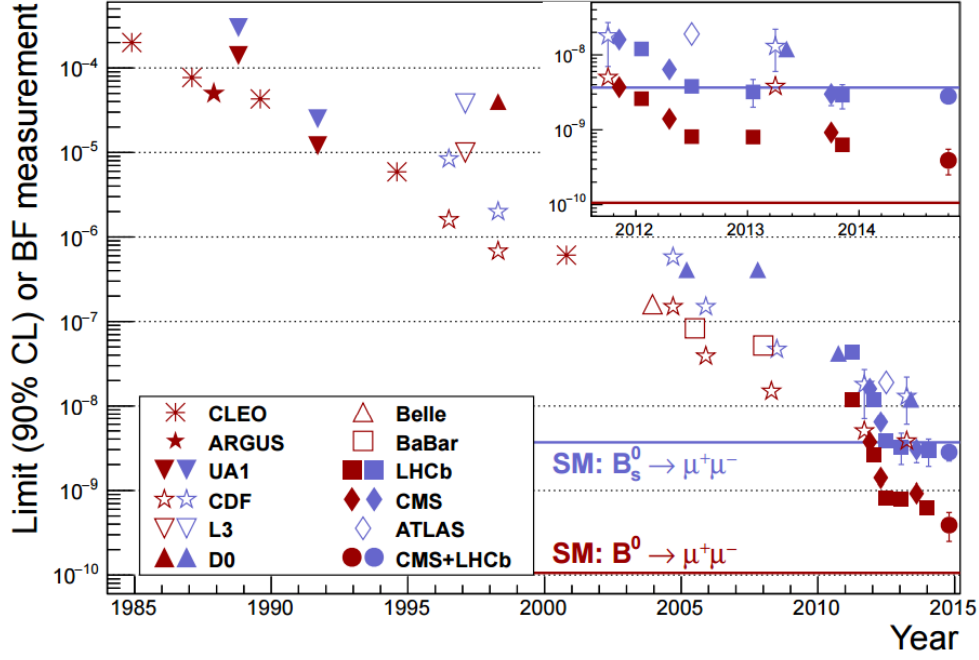


Figure 1.5: Search for the $B_{(s)}^0 \rightarrow \mu^+\mu^-$ decays during the last 35 years. Markers without error bars denote upper limits on the branching fractions at 90% confidence level, while measurements are denoted with error bars delimiting 68% confidence intervals. The solid horizontal lines represent the SM predictions for the $B_{(s)}^0 \rightarrow \mu^+\mu^-$ branching fractions; the blue (red) lines and markers are related to the $B_s^0 \rightarrow \mu^+\mu^-$ ($B^0 \rightarrow \mu^+\mu^-$) decay [21].

study, includes only one track. It is one of the most common B^+ decays which includes a J/ψ , and the uncertainty of its measurement is of 3% only. The relevant formula is then:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-) = \frac{N_{B_s^0 \rightarrow \mu^+\mu^-}}{N_{B^+ \rightarrow J/\psi K^+}} \frac{\varepsilon(B^+ \rightarrow J/\psi K^+)}{\varepsilon(B_s^0 \rightarrow \mu^+\mu^-)} \frac{f_u}{f_s} \mathcal{B}(B^+ \rightarrow J/\psi K^+) \mathcal{B}(J/\psi \rightarrow \mu^+\mu^-), \quad (1.18)$$

where ε represents the overall efficiency of the analysis for a certain decay channel, which also includes the acceptance of the detector. The ‘N’ refers to the observed number of signal events, and not to the total number of particles which follow the specified decay; the efficiencies serve as a correction. Finally, f_u/f_s denotes a ratio between the fragmentation fractions of the B^+ and B_s^0 mesons. Their meaning will be explained in the next section.

1.3.2 Fragmentation Fractions

Due to colour confinement, when a quark is produced it must hadronize. For the case of b quarks, the probability of doing it into a certain b -flavoured hadron is denoted “ b fragmentation fraction”. It is defined as $f_X \equiv N_{b \rightarrow B_X} / N_{b \rightarrow \text{all}}$, where ‘X’ may refer to any b hadron. For example, f_s is the probability for a

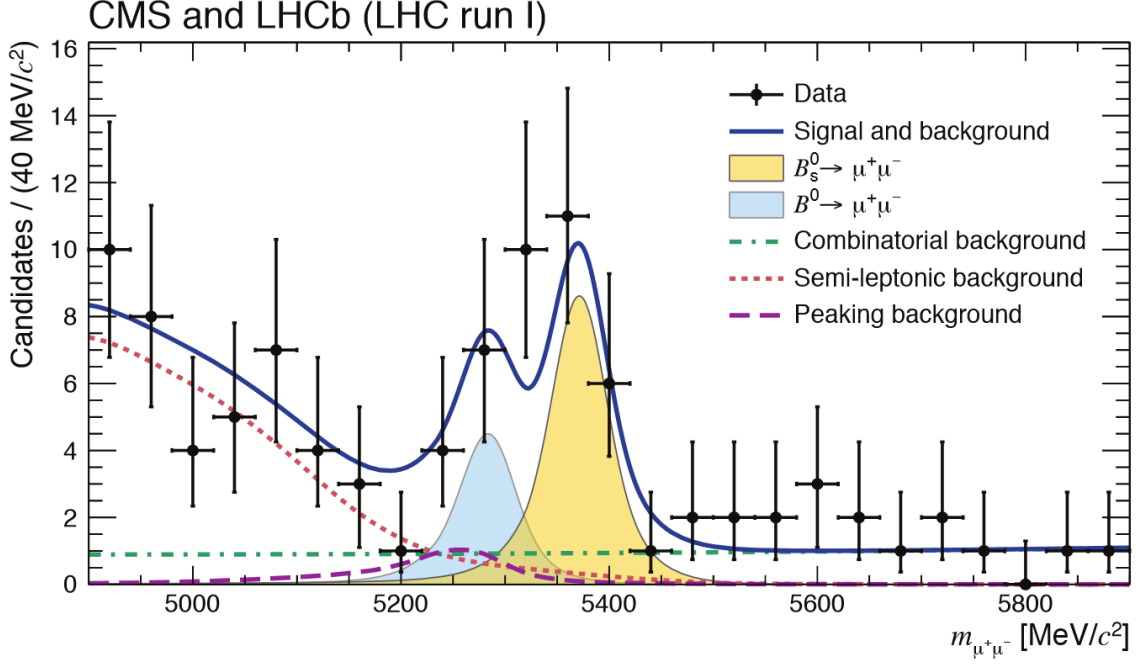


Figure 1.6: Dimuon invariant mass $m_{\mu^+\mu^-}$ distribution for B_s^0 and B^0 meson candidates. The data points (black) were superimposed by a combined unbinned maximum extended likelihood fit (solid blue). The vertical error bars denote a 68% confidence interval, while the horizontal error bars show the chosen binning. The fit includes several components: for the signal, we see the B^0 (blue area) and B_s^0 (yellow area) peaks, while for the background the semi-leptonic background (dotted salmon line), the combinatorial background (dash-dotted green line) and the peaking background (dashed violet line) were considered. The combinatorial background arises due to random combinations of two muons. The semi-leptonic background comes in turn from B_s^0 semi-leptonic decays in which some non-interacting particles (e.g.: neutrinos, neutral pions) were not detected. Finally, the peaking background appears due to two-body hadronic decays, such as $B^0 \rightarrow K^+\pi^-$ [21].

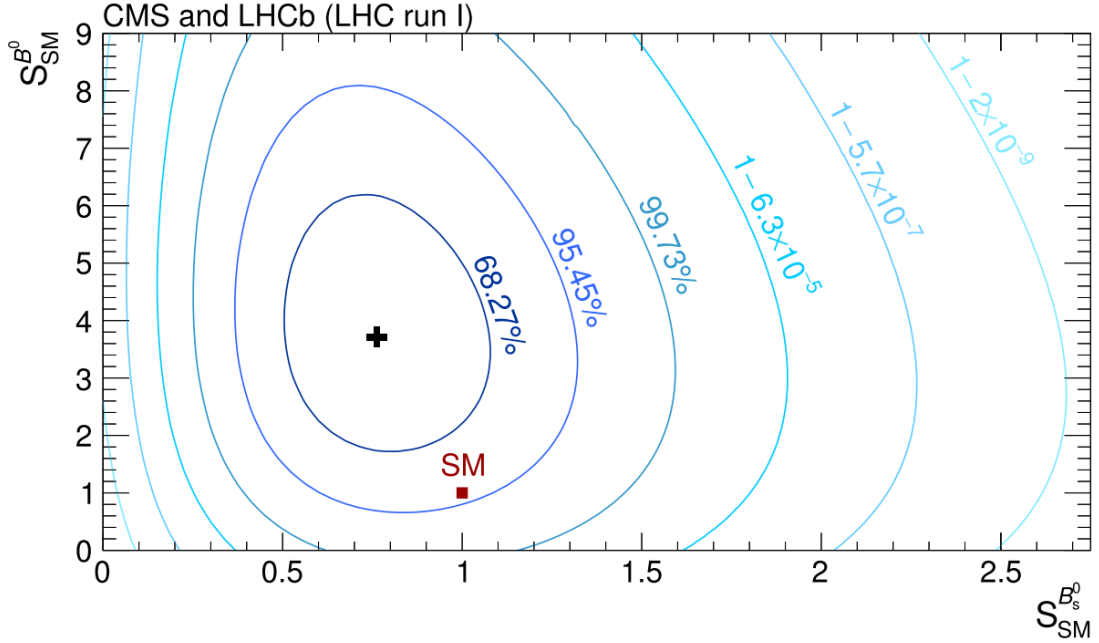


Figure 1.7: Likelihood contours for the ratios of the B_s^0 and B^0 branching fractions with respect to their SM prediction. The black cross marks the central value obtained. The SM point is shown as the red square located, by construction, at $S_{SM}^{B_s^0} = S_{SM}^{B^0} = 1$. Each contour encloses a region approximately corresponding to the reported confidence level. The SM branching fractions are assumed uncorrelated to each other, and the likelihood contours take the branching fraction uncertainties into account [21].

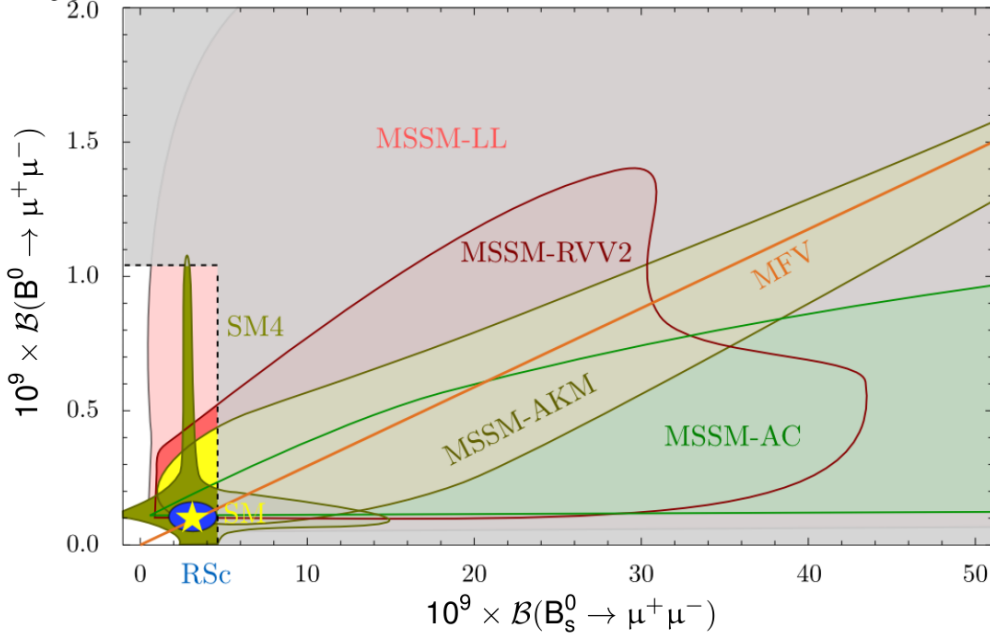


Figure 1.8: Example of the constraints that can be imposed to BSM theories by studying the $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ rare decays. The correlation between these decays' branching fractions is shown considering Minimal Flavour Violation (MFV) models, a SM-like model with a fourth generation of elementary particles (SM4), the Randall-Sundrum model with custodial protection (RSc) and four SUSY flavour models. The grey area is ruled out experimentally. The SM point is marked by a star. This plot was retrieved from Ref. [31], which does not include the joined LHCb and CMS measurement. The current constraints are thus a bit tighter. That can be readily seen since a null value for the $B_s^0 \rightarrow \mu^+\mu^-$ branching fraction is no longer allowed by experiment (see Eq. 1.16).

b quark to hadronize into a B_s^0 meson¹⁰. Fragmentation fractions are constrained by the fact that they represent a probability:

$$f_u + f_d + f_s + f_c + f_{\text{baryon}} = 1, \quad (1.19)$$

where

$$f_{\text{baryon}} = f_{\Lambda_b^0} + f_{\Xi_b^0} + f_{\Xi_b^-} + f_{\Omega_b^-} = f_{\Lambda_b^0} \left(1 + 2 \frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} + \frac{f_{\Omega_b^-}}{f_{\Lambda_b^0}} \right), \quad (1.20)$$

in which $f_{\Xi_b^-} = f_{\Xi_b^0}$ was assumed, *i.e.*, isospin invariance is expected to hold in the production of these two baryons. All the other b baryons are expected to decay to the listed ones through electromagnetic or strong interactions [33].

The fragmentation fraction ratio present in Eq. 1.18 can be calculated using channels which somehow involve the B^+ and B_s^0 mesons. Here the $B^+ \rightarrow J/\psi K^+$ and $B_s^0 \rightarrow J/\psi \phi$ channels are considered ($J/\psi \rightarrow \mu^+\mu^-$ and $\phi \rightarrow K^+K^-$):

$$\frac{f_s}{f_u} = \frac{N_{B_s^0 \rightarrow J/\psi \phi}}{N_{B^+ \rightarrow J/\psi K^+}} \frac{\varepsilon(B^+ \rightarrow J/\psi K^+)}{\varepsilon(B_s^0 \rightarrow J/\psi \phi)} \frac{\mathcal{B}(B^+ \rightarrow J/\psi K^+)}{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) \mathcal{B}(\phi \rightarrow K^+K^-)}, \quad (1.21)$$

¹⁰The meson to which the b quark hadronizes is not necessarily the one which is detected. Decays like $B_s^{**} \rightarrow B^+ K^-$ or $B_s^{**} \rightarrow B^0 \bar{K}^0$ (where ** represents a double excited state) exist and may have a significant effect. For example, it might happen that not all the measured B^+ come from the hadronization of a b quark into a B^+ meson. A more appropriate definition of f_X might be the following: the fraction of B_X mesons in an unbiased sample of weakly-decaying hadrons containing a bottom quark. An extra item is worth mentioning: a decay of the type $B_c^+ \rightarrow B_s^0 X$ may contribute to either f_c or f_s . Since the B_c^+ meson is the only ground state meson consisting of two heavy quarks of different flavour ($\bar{b}$ and c), its production is suppressed relative to other lighter B mesons, and as such it is not expected that the $B_c^+ \rightarrow B_s^0 X$ decay will sizeably affect the f_s fragmentation fraction [32].

where the variables have the same meaning as explained after Eq. 1.18. The chosen channels have high yields to allow a precise measurement, and include a dimuon in the final state, since muons leave a clear signature in the detector (see Chapter 2). Like the B^+ channel here considered, the $B_s^0 \rightarrow J/\psi \phi$ decay is one of the B_s^0 most common decay channels which includes a J/ψ . Its branching fraction is displayed in Table 1.3.

Table 1.3: Values of the branching fractions used in this work. The values were obtained from [5].

Decay channel	Branching fraction
$B^+ \rightarrow J/\psi K^+$	$(1.026 \pm 0.031) \times 10^{-3}$
$B^0 \rightarrow J/\psi K^{*0}$	$(1.28 \pm 0.05) \times 10^{-3}$
$K^{*0} \rightarrow K^+ \pi^-$	$(6.650 \pm 0.001) \times 10^{-1}$
$B_s^0 \rightarrow J/\psi \phi$	$(1.08 \pm 0.08) \times 10^{-3}$
$\phi \rightarrow K^+ K^-$	$(4.89 \pm 0.5) \times 10^{-1}$

If mass effects are neglected, decays of the type $B_s^{0**} \rightarrow B^* K$ contribute equally to B^+ and B^0 rates, where ** denotes a double excited state. In addition, if one assumes the probabilities for a b quark to hadronize into a B^+ or a B^0 mesons to be equal (isospin symmetry), and also that electromagnetic decays of B^{+*} and B^{0*} and strong decays of B^{+**} , B^{0**} and B_s^{0**} do not affect the relative amounts of the B^+ and B^0 mesons, one can use that the production of these two mesons is equal: $f_u = f_d$. This work will test this assumption, by using the following formula (which includes the $B^0 \rightarrow J/\psi K^{*0}$ decay), which is expected to return a unitary value:

$$\frac{f_d}{f_u} = \frac{N_{B^0 \rightarrow J/\psi K^{*0}}}{N_{B^+ \rightarrow J/\psi K^+}} \frac{\varepsilon(B^+ \rightarrow J/\psi K^+)}{\varepsilon(B^0 \rightarrow J/\psi K^{*0})} \frac{\mathcal{B}(B^+ \rightarrow J/\psi K^+)}{\mathcal{B}(B^0 \rightarrow J/\psi K^{*0}) \mathcal{B}(K^{*0} \rightarrow K^+ \pi^-)}. \quad (1.22)$$

The f_s/f_d ratio will also be measured, since its uncertainty might be lower than the one coming from the f_s/f_u ratio due to the cancellation of systematic uncertainties. Indeed, looking at its formula:

$$\frac{f_s}{f_d} = \frac{N_{B_s^0 \rightarrow J/\psi \phi}}{N_{B^0 \rightarrow J/\psi K^{*0}}} \frac{\varepsilon(B^0 \rightarrow J/\psi K^{*0})}{\varepsilon(B_s^0 \rightarrow J/\psi \phi)} \frac{\mathcal{B}(B^0 \rightarrow J/\psi K^{*0}) \mathcal{B}(K^{*0} \rightarrow K^+ \pi^-)}{\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) \mathcal{B}(\phi \rightarrow K^+ K^-)}, \quad (1.23)$$

we observe that the two channels here used have the same number of tracks ($K^{*0} \rightarrow K^+ \pi^-$ versus $\phi \rightarrow K^+ K^-$), and so the corresponding tracking uncertainty of the detector cancels at first order. In Fig. 1.9 we show a comparison plot of the f_s/f_d measurements done so far¹¹. The idea of systematic uncertainties cancellation also helps to understand why the fragmentation fractions are not calculated separately: they would contribute with a larger systematic contribution.

The joined LHCb and CMS analysis in Ref. [21] used the LHCb f_s/f_d measurement [35]. An additional systematic uncertainty of 5% was assigned to account for the extrapolation of the LHCb result to the CMS acceptance [21]. 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 deter-

¹¹For completeness, we mention here two concerns raised by the LHCb Collaboration regarding the decay channels used by ATLAS [34] and by this work. The first one concerns the possible existence of NP in Penguin diagrams in the $B^0 \rightarrow J/\psi K^{*0}$ and $B_s^0 \rightarrow J/\psi \phi$ decays. These contributions were still not experimentally excluded. Secondly, the $B \rightarrow J/\psi X$ decays involve a large theoretical uncertainty when compared to $B \rightarrow D h$ decays (the ones used by LHCb [35]), being 'h' a hadron.

mined.

It is also important to note that fragmentation fractions cannot be assumed to be constant. One can only plausibly do it when the square of the momentum transfer to the produced b quark is large when compared with the square of the hadronization energy scale. Moreover, there is no strong argument to claim that the fractions at different colliders should be strictly equal, since they could depend on the accessible kinematic regions¹² [33]. Still, for completeness, we provide a rough value for them:

$$\begin{aligned} f_u &\sim 40\%, f_d \sim 40\%, f_s \sim 10\%, \\ f_c &\sim 1\%, f_{\text{baryon}} \sim 8\%. \end{aligned} \quad (1.24)$$

It is still unknown whether the fragmentation fraction ratio f_d/f_s depends on the kinematics of B mesons, *e.g.*, its transverse momentum. Fig. 1.10 shows previous f_d/f_s and f_{Λ_b}/f_s transverse momentum dependency studies, which for the case of the f_d/f_s ratio were inconclusive. While LHCb saw a dependency, ATLAS found the results to be compatible with no dependency. Since these results were obtained with a 7 TeV centre of mass collision energy, we expect to probe higher p_T regions, since the data used in this work takes advantage of the 13 TeV centre of mass collision energy. Part of this analysis will be dedicated to study the fragmentation fraction's kinematic dependence on both the p_T and rapidity. The rapidity can be thought of as an angular variable (its exact definition is included in Section 2.1). Ta-

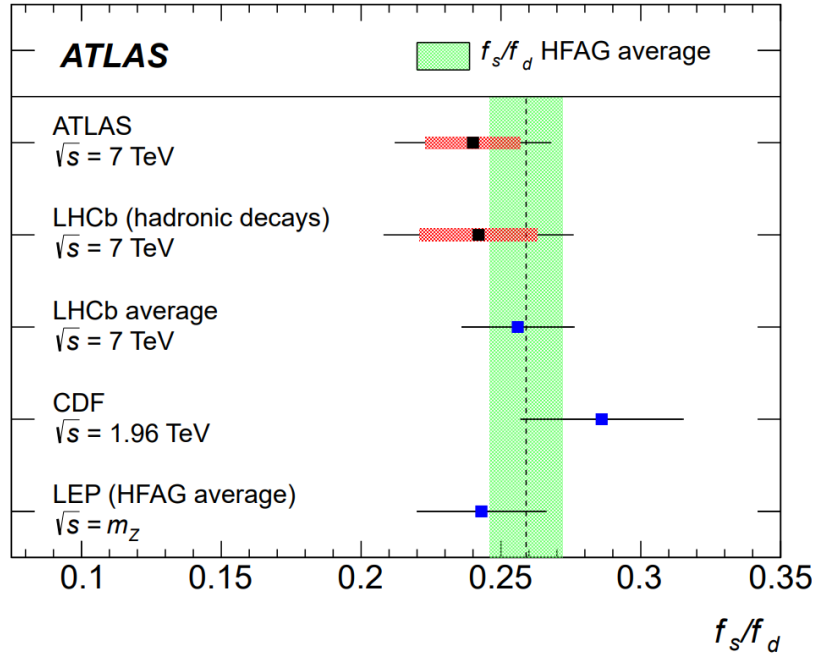


Figure 1.9: Current status of f_s/f_d measurements. LEP, CDF (Collider Detector at Fermilab), LHCb, and ATLAS measurements are shown. The error bars are depicted with thin black lines. The theory error is represented by the thick red line for the first two measurements. The black dashed line with its uncertainty (green-shaded region) shows the HFAG average obtained using the blue points (the last three measurements) [34].

ble 1.4 shows the main systematic sources which contributed for the CMS measurement leading to the

¹²It has been observed that the fragmentation fractions of b quarks produced in $Z \rightarrow b\bar{b}$ decays significantly differ from the same fractions when the b quarks are produced in $p\bar{p}$ collisions [33].

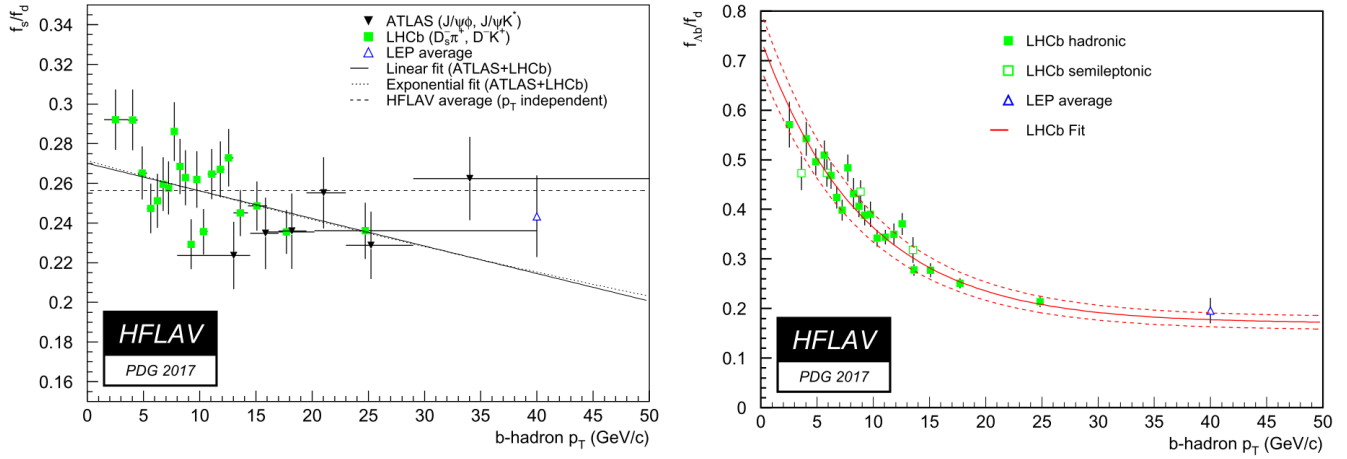


Figure 1.10: Measurement of the f_s/f_d (left) and f_{Λ_b}/f_d (right) ratios as a function of $p_T(B)$. While the LHCb and LEP (Large Electron-Positron Collider) measurements see a clear p_T dependence on the f_{Λ_b}/f_d ratio, the same cannot be said regarding the f_s/f_d quantity, which ATLAS also measured. While LHCb (left plot, green points) reported a dependence, ATLAS (left plot, black points) did not find it. This work tries to shed some light on this inconsistency, also making use of the more recent datasets which allow to reach higher p_T values, and thus testing the dependence in a wider kinematic range [33].

$B_s^0 \rightarrow \mu^+ \mu^-$ observation¹³ [36]. It can be seen that the f_s/f_d quantity contributes with one of the largest uncertainties. As a consequence, its precise measurement is crucial, since the chance to observe NP in the $B_s^0 \rightarrow \mu^+ \mu^-$ decay may depend on how small the systematic uncertainties of the final measurement are.

Table 1.4: Statistical plus systematic uncertainty sources for the CMS $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction measurement [36] (added quadratically). The 5% extra uncertainty present in the ratio refers to the CMS extrapolation of the LHCb measurement [35] already explained in the main text and also explained in Ref. [21]. For the sources which include two numbers, the first refers to the “Barrel” and the second to the “Endcap” part of the CMS detector (see Chapter 2).

Source	Uncertainty
f_d/f_s	(7.81 + 5.00) %
$\mathcal{B}(B^+ \rightarrow J/\psi K^+)$	3.33%
Total efficiency (7 TeV)	8.16% - 11.11%
Total efficiency (8 TeV)	8.54% - 14.29%
Norm. channel yield (7 TeV)	5.76% - 5.14%
Norm. channel yield (8 TeV)	5.18% - 5.05%

¹³There are many systematic uncertainty sources that consist in background events which behave in the same fashion as the channel under study. For example, in CMS a hadron to muon misidentification happens with a 10^{-3} probability. When this occurs, a $B_s^0 \rightarrow hh'$ decay (where h and h' are hadrons), can be misreconstructed as $B_s^0 \rightarrow \mu^+ \mu^-$. These events are modelled with MC simulations, but then a 50% uncertainty has to be assigned to the probability, due to differences between data and MC. A different issue affects the $\Lambda_b \rightarrow p \mu^- \bar{\nu}$ decay, which can also fake the signal, since the neutrino $\bar{\nu}$ is not seen and the proton may be confused with a muon. Its branching fraction is known with a 100% uncertainty! These systematic sources affect the yield determination, but do not directly enter Eq. 1.18.

1.4 Thesis Outline

This work represents a detailed description of the author's contribution to the measurement of the f_s/f_d fragmentation fraction ratio in the framework of the $B_s^0 \rightarrow \mu^+\mu^-$ rare decay branching fraction measurement, using CMS Run2 data (2015). The work was done at the *Laboratório de Instrumentação e Física Experimental de Partículas* (LIP). For completeness, the f_d/f_u and f_s/f_u ratios are measured to test the consistency of the results. The transverse momentum and rapidity dependency of the three ratios is studied. CMS has up to now not measured any of these quantities.

Chapter 2 contains a brief description of the CMS detector, where the silicon tracker, muon detector and trigger are described, due to their relevance for the measurements here performed. Chapter 3 explains which data and MC samples were used, how they were processed, and which selection was applied. The following chapter serves to explain the data/MC comparison techniques, expressed both in the form of the mass sideband subtraction method and in fits to the peaks of the B meson mass distributions. The rest of the work is structured according to the ingredients needed to obtain the fragmentation fraction ratios. It contains chapters which discuss the way the yields, efficiencies and systematic uncertainties were obtained (Chapters 5, 6 and 7, respectively). The final results and related discussion are presented in Chapter 8. Although not all produced plots could be included in the present work, a very representative sample can be found in the Appendixes.

Chapter 2

Experimental Apparatus

This work uses data collected by the Compact Muon Solenoid (CMS) at the 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% of 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} [37]. 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 per crossing take place. Still, due to the 40 MHz bunch collision frequency, the LHC produces almost 1 billion collisions per second [38]. The CMS detector is located at one of the existing interaction points at 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 components: the silicon tracker, the electromagnetic and the hadronic calorimeter, the magnet itself, and the muon stations (from the centre to the outside; see Fig. 2.1). For the purpose of this work, the silicon tracker and the muon stations represent the most important parts of the detector.

2.1 CMS Coordinate System and Parametrization

The chosen coordinate system at CMS adopts the centre of the detector as its origin. The y axis points vertically upwards, while the x axis points radially inwards toward the centre of the LHC. As a consequence, the z axis points along the counter-clockwise beam direction. Due to the absence of detectors close to the z axis¹, the transverse component ' p_T ' of the momentum magnitude ' p ' of a particle can be measured more precisely than the total three-momentum. These quantities are given by:

$$p = \sqrt{p_T^2 + p_z^2}, \quad p_T = \sqrt{p_x^2 + p_y^2}. \quad (2.1)$$

The p_T is also invariant with respect to Lorentz boosts along z . Furthermore, the protons that follow a trajectory on the z axis tend to be those who did not interact, so that the interesting phenomena will usually have a significant p_T . For all these reasons, the transverse momentum is widely used. The

¹The beam line has to fit somewhere!

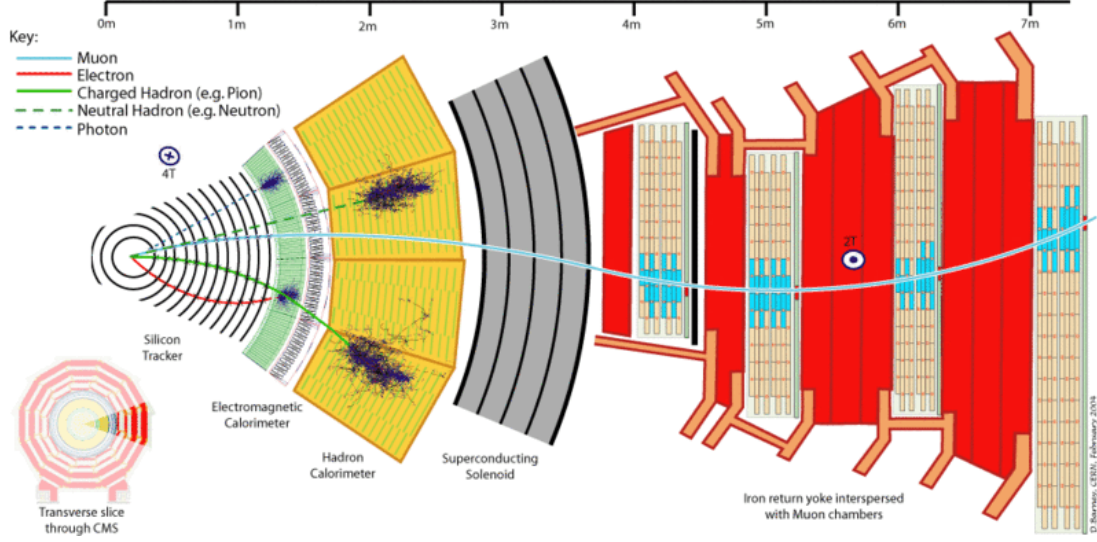


Figure 2.1: Transverse section of the CMS detector, which is composed of the following (from left to right): silicon tracker, electromagnetic calorimeter, hadronic calorimeter, solenoid and four muons stations interleaved with return yoke plates. The silicon tracker measures the trajectory of charged particles. The muon stations are essential to distinguish muons from other tracks (blue line).

azimuthal angle $\phi \in [-\pi; \pi]$ is measured from the x axis in the xy plane and the radial coordinate in this plane is denoted by $R = \sqrt{x^2 + y^2}$. The polar angle $\theta \in [0; \pi]$ is measured from the z axis. Another very common coordinate at particle colliders is the rapidity 'y', defined as:

$$y = \frac{1}{2} \ln \left(\frac{E - p_z}{E + p_z} \right). \quad (2.2)$$

Rapidity differences (but not the rapidity itself) are also Lorentz invariant under boosts along z . Unfortunately, in order to measure this quantity, one would *a priori* need to know the mass of the particle. In addition to this, the z component of the momentum cannot be precisely measured. As an alternative, a quantity named pseudorapidity is defined in the following way:

$$\eta = \frac{1}{2} \ln \left(\frac{|\mathbf{p}| - p_z}{|\mathbf{p}| + p_z} \right) = -\ln \left(\tan \frac{\theta}{2} \right). \quad (2.3)$$

In the relativistic limit $p \gg m$, being 'm' the mass of the particle, the pseudorapidity and rapidity definitions are equivalent [39, 40]. A transverse section of the CMS detector with the θ , η , R and z coordinates is shown in Fig. 2.2.

2.2 Tracker

The inner layer of CMS is occupied by a silicon tracker. The goal of the tracker is to efficiently and precisely measure the trajectories of charged particles with $|\eta| < 2.5$. It can also precisely measure

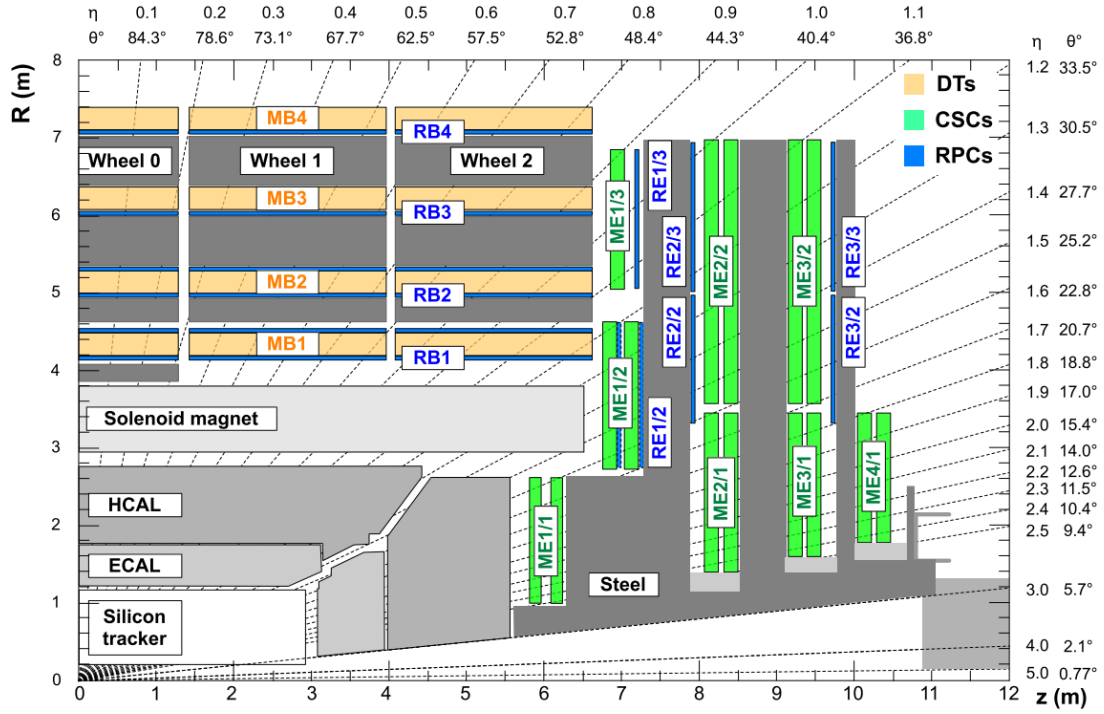


Figure 2.2: $z-R$ transverse section of the CMS detector. The interaction point is at the lower left corner. The pseudorapidity η is indicated for different θ angles. The locations of both the muon stations and the steel disk are shown (dark grey areas). The 4 drift tube (DT, in light orange) stations are labelled MB (“muon barrel”) and the cathode strip chambers (CSC, in green) are labelled ME (“muon endcap”). Resistive plate chambers (RPC, in blue) are in both the barrel and the endcaps of CMS, where they are labelled RB and RE, respectively [41].

secondary vertexes² and impact parameters, which is a fundamental feature for any B physics analysis. The tracker should extend as closely as possible towards the primary vertex, to better reconstruct the momentum of the particles. The radiation that passes through the tracker is so large that from 2015 onwards it was coupled to a cooling system which reaches -20°C temperatures. The tracker is divided into silicon pixels (closer to the beam) and silicon strips, reaching a total of 75 million of electronic read-out channels. They measure the path of a track with a spatial resolution of $10\text{-}12\ \mu\text{m}$, and can reconstruct high p_T isolated electrons and muons with an efficiency of about 99% and with a fake track³ rate lower than 1% [39, 42, 43]. It must be pointed out that CMS does not possess a dedicated particle identification detector, such as a Cherenkov detector. Distinguishing among different charged particles (proton, kaons and pions) is therefore not straightforward. Muons instead, may be easily identified since they reach the muon stations.

²The terminology is quite intuitive: the primary vertex is the point where two protons collide, and the secondary vertex is the point in which some particles produced at the pp collision further decay: in our case this refers to the B mesons. Depending on the decay channel under study, one may even observe tertiary vertexes. This is the case for some heavy flavour decays, such as the $B \rightarrow DX$ decays, being ‘D’ a charmed meson and ‘X’ anything else, since the D mesons travel a bit before decaying (its lifetime is comparable to the lifetimes of the B mesons).

³A fake track represents the detector measurement of a track that never existed. For example, some random noise in consecutive layers of the silicon tracker could simulate a charged particle.

2.3 Muon Stations

As its name might suggest, CMS is characterized by a good muon identification and momentum resolution over a wide range of momenta and angles, by a good dimuon mass resolution ($\approx 1\%$ at 100 GeV), and by the ability to unambiguously determine the charge of muons with three-momentum smaller than 1 TeV. Muons have the ability to penetrate through multiple iron layers without interacting. Its detection system is therefore placed at the edge of the detector, and can be reached by muons and neutrinos only (the latter are not directly detected). It comprises a series of muon stations interleaved with steel “return yoke” plates. The path of a muon is measured by fitting a curve to the hits among the four muon stations, together with the hits in the silicon tracker (see Figure 2.1). Interestingly enough, the magnetic field located outside the solenoid is also used for muon tracking (it has approximately a 2T magnitude). The muon system is used for muon identification, momentum measurement, and triggering. The return yoke is used as a hadron absorber for the identification of muons, and it must be very strong in order to withstand the magnetic field forces.

A total of 1400 muon chambers measure the path through which muons travel. They consist of three types of gas⁴ detectors: drift tubes (DTs), cathode strip chambers (CSCs) and resistive plate chambers (RPCs). When a muon enters into one of these detectors, it knocks out some electrons which in turn produce an electric signal, making the tracking of the muon’s original position possible. The electrons move due to an electric field created by charged wires (DTs and CSCs) or charged plates (RPCs). The DTs are located in the “barrel” part of the detector ($0 < |\eta| < 1.2$), while the CSCs mostly lie in the “endcap” region of CMS ($0.9 < |\eta| < 2.4$). They are used both for measurement and triggering. Their spatial resolution was measured to be of around 100 μm for CSCs and DTs. Their time resolution is of a few nanoseconds. The RPCs are used for triggering and they are located in both the barrel and the endcap regions. The endcap is subjected to higher particle flow rates than the barrel, since it lies closer to the beamline; the three types of muon chambers were designed accordingly. For example, a DT would not work in the endcap region [39].

2.4 Trigger

CMS does not store all collisions that occur at the interaction point, due to the large storage space required, high bandwidth (maximum rate of data transfer) needed as well as prohibitive costs. To solve these issues, a key ingredient is introduced: particles produced at pp collisions are selected with a two-level trigger system, which is designed to reduce the amount of data by a factor of 10^6 . The first trigger, L1 (“level one”), is a hardware system which decides in 4 μs if the data relative to a certain collision should be accepted or rejected. It uses information coming from the calorimeters and muons detectors only. The Calorimeter Trigger obtains information regarding muon isolation and muon compatibility with

⁴Which gas should one use? Since the detector lies in the underground, in a confined space, the mixture cannot be flammable. Organic compounds have to be avoided since they promote ageing. It should also be cheap! Mixtures of Ar/CO₂ satisfy these and other technical requirements [44].

minimally ionizing particles (MIPs)⁵. Regarding the muon trigger system, both the DT and CSC trigger on the p_T of muons with good efficiency and high background rejection. The RPC consists in a fast and independent redundant trigger system: it can identify muon tracks and unambiguously assign them to the correct bunch crossing with high efficiency (it identifies the events in much less than 25 ns, which is the time interval between consecutive bunch crossings), while estimating the track's transverse momentum [44]. Then, the Global Muon Trigger gathers all the information from the DTs, CSCs and RPCs trigger systems to improve the efficiency, reduce the trigger rates and suppress background. These goals are achieved by comparing the spatial coordinates of DTs and CSCs candidates with the ones of RPCs candidates, and check if they match. The muons are also back-extrapolated to the calorimeter region to further add the information there measured. Finally, the Global Trigger decides whether an event at L1 should be accepted or rejected, based on the information of the Calorimeter Trigger and of the Global Muon Trigger.

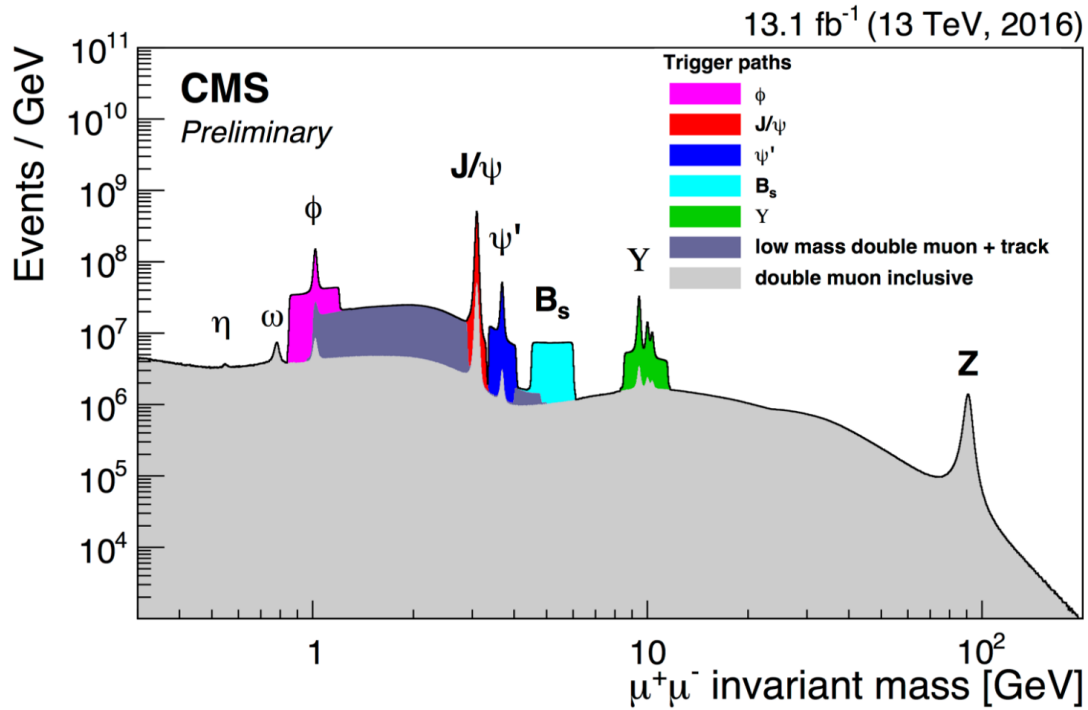


Figure 2.3: Dimuon mass distribution collected with various dimuon triggers at 13 TeV in 2016 with 13.1 fb⁻¹. The coloured paths correspond to dedicated dimuon triggers with low p_T thresholds, in specific mass windows, while the light grey continuous distribution represents events collected with a dimuon trigger with high p_T thresholds. Note the logarithmic scales. The J/ψ peak and B_s^0 window can be clearly identified.

At the next stage a software trigger is applied: the High Level Trigger (HLT). It rejects the surviving L1 events by an extra factor of 1000. It was designed to operate at a maximum of 50 kHz L1 output rate. It makes use of a large computing cluster which uses all the available information coming from the detector, including the tracking. It is sometimes divided into two separate sublevels⁶: the L2 trigger, which obtains information regarding the muon chambers and the calorimeter with a much higher precision than L1, and

⁵A muon is considered isolated if its energy deposit in the calorimeter region from which it emerged is below a defined threshold. A MIP is a particle whose average energy loss rate approaches the minimum allowed by theory. Relativistic muons are MIPs.

⁶This separation does not exist in practice, since the L2 and L3 triggers implementation relies on software. It must nevertheless be introduced, since it will be needed when discussing the "Tag and Probe" technique (Section 7.4).

the L3 trigger, which uses information from all detector elements including the tracker. A large number of “HLT paths” is available: they select events according to specific criteria, like certain p_T thresholds, mass windows (where a given particle with an invariant mass laying outside the window is excluded), or others. Each analysis must carefully design its own paths: one would like to retain as much signal as possible, *i.e.*, to keep all the relevant events for the analysis, and to reject everything else. We now draw the reader’s attention to an interesting detail. If a certain analysis is looking for particles which always decay to three tracks, the relevant HLT path will probably not be the one which selects three tracks, but rather a different one that selects one or two tracks. In the latter case more collisions are kept for further analysis; the matching of one or two extra tracks to the ones selected by the HLT can be later done offline. Indeed, when using a restrictive trigger, there is the possibility that some signal is rejected by mistake. In short: one would like to maximize the stored signal given the existing hardware and software restrictions [21, 39, 45].

In this work the used path includes a double muon trigger (see Section 3.2). Due to the good muon trigger system in the CMS detector, channels which involve muons in the final state are commonly preferred: in this work, all considered channels include muons. We point out that low p_T muons are much harder to trigger, due to their higher production rate. The CMS muon trigger capabilities are depicted in Fig. 2.3. The J/ψ peak can be clearly seen, as well as many others⁷. The B_s^0 trigger path is also shown (cyan) but the peak cannot be seen. This is not surprising: we know that the $B_s^0 \rightarrow \mu^+\mu^-$ decay is incredibly rare, and physicists cannot observe it by just applying a trigger! Other techniques, described in Ref. [36], are required.

⁷The ϕ peak (pink) refers to the ϕ decay to two muons, and not to the $\phi \rightarrow K^+K^-$ decay (which is part of the $B_s^0 \rightarrow J/\psi \phi$ decay channel used in this work). The former decay channel has a lower probability to happen than the latter (the $\phi \rightarrow \mu^+\mu^-$ branching fraction is of the order of 10^{-4} [5]).

Chapter 3

Data and MC Samples

The data are collected by the CMS detector and are processed through the HLT. Each HLT path is designated to live inside a specific “Primary dataset” (PD). Different HLT trigger paths are contained inside the same PD, since some analysis may need to use collision information retrieved with more than one HLT path. The PDs are defined according to some standard criteria: no analysis should use more than one or two PDs to avoid bookkeeping duplicate information (since the same collision may appear in several PDs); the PDs should be small to avoid the time consumption caused by running analysis on large samples; prescaled triggers, *i.e.*, triggers which randomly select only a fraction of all accepted collisions, should be kept in different PDs. For example, a prescale of 10 means that the HLT trigger randomly selects one out of 10 events [46, 47].

This work uses data which fired the `HLT_DoubleMu4_JpsiTrk_Displaced` HLT path, which lives in the so-called “Charmonium” PD. This path requires two reconstructed muons each with a transverse momentum larger than 4 GeV, that the pair is found within a 2.9 to 3.3 GeV J/ψ invariant mass window, that the dimuon satisfies $L_{xy}/\sigma_{xy} > 3$ before decaying, being L_{xy} the transverse distance between the primary and secondary vertexes and σ_{xy} its uncertainty, *i.e.*, the J/ψ is required to be displaced (the B from where the J/ψ came travelled a bit in the detector)¹ and that the dimuon is matched to a track².

It would have been possible to use the `HLT_DoubleMu4_JpsiTrkTrk_Displaced` HLT path for the B^0 and B_s^0 channels instead, in order to immediately account for the extra track. Since at the end ratios between different fragmentation fractions will be measured, it is better to use the same trigger for all channels of interest. In this way we know that the mesons were selected in similar ways and that therefore no bias is introduced. Would the `HLT_DoubleMu4_JpsiTrkTrk_Displaced` path miss some double track events, and a bias would be introduced when calculating a yield ratio involving the B^+ channel³. On the other hand, it is important to store as much data as possible. This leads to the idea that we could have also used a looser trigger, since the `HLT_DoubleMu4_Jpsi_Displaced` path is available. This time a different issue would be faced: the trigger rate would be too high. Indeed, the limited trigger band-

¹All prompt-produced J/ψ are thus removed (J/ψ that were produced at the primary vertex).

²The full description of all the triggers, including the one used in this analysis, is available at <https://github.com/cms-sw/cmssw/blob/master/HLTrigger/Configuration/python/HLT/textunderscoreFULL/textunderscorecff.py>.

³Of course that this effect can be accounted for in the efficiency of each channel, but that would imply a more difficult efficiency calculation (different for different channels) without any advantage.

width would require using a prescale, which selects only some collisions among all events that fired the trigger. A fraction of potentially interesting collisions would be lost. The selected path for this analysis is therefore unprescaled; it represents a balance between getting as many events as possible while being able to handle such large amounts of data.

The data which passed the L1 and HLT triggers can be accessed through the Data Aggregation System of CERN. In this repository one can find all kinds of reconstructed (RECO) and Analysis Object Data (AOD) information. The AOD datasets represent a sub sample of the RECO information and provide everything a physics analysis usually needs in a compact way. In extreme cases, for example when some tracks need to be refitted, the RECO datasets have to be used. In this analysis AOD “Charmonium” data samples were used (2015 and 2016 – see Table 3.1).

For the case of the MC samples used in this work (see Table 3.2), additional formats are available. While the GEN samples include only the generator information, the SIM ones considers also the simulation of the CMS detector using Geant4, which simulates the passage of particles through matter. Other formats are available but are not here discussed⁴. AODSIM samples were used: they are similar to the AOD format of the data, but include the generation level information as well⁵. The datasets used the Pythia8 event generator, and the decay tool EvtGen forced the particles to decay into certain final states. As an example, the J/ψ was always forced to decay into a pair of muons, despite doing that only circa 6% of times. The reason why two 2015 datasets were used for each channel is explained in Section 6.

Table 3.1: AOD data samples used in this work. The dates refers to their online processing.

Year	Dataset
2015	/Charmonium/Run2015C_25ns-16Dec2015-v1/AOD
	/Charmonium/Run2015D-16Dec2015-v1/AOD
2016	/Charmonium/Run2016C-23Sep2016/AOD
	/Charmonium/Run2016D-23Sep2016/AOD
	/Charmonium/Run2016E-23Sep2016/AOD
	/Charmonium/Run2016F-23Sep2016/AOD
	/Charmonium/Run2016G-23Sep2016/AOD

3.1 Processing

The chosen datasets have to be processed to reconstruct the final states under study. The muons used are selected according to “Soft Muon ID” quality cuts, which are widely used in Charmonium related analysis. They consist of the following items (especially relevant for Section 7.4):

- “high purity” tracks, as defined in Ref. [42]. This item provides stringent requirements, increasing the “purity”, *i.e.*, decreasing the number of misidentified tracks in the signal region, but reducing

⁴The CMS data formats are explained in Ref. [48].

⁵This is something that does not exist in data, since one has only access to the decay products of the particles that were produced in the pp collisions. Real data cannot be generated!

Table 3.2: AODSIM Monte Carlo samples used in this work. The suspension points refer to the following strings: RunIISpring15DR74-Asympt25ns_MCRUN2_74_V9_ext1-v1, for 2015, and RunIISummer16DR80Premix-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6_xt1-v2, for 2016. The substring TuneCUEP8M1_13TeV-pythia8-evtgen was abridged to XX. The specified size refers to the number of B mesons generated in each sample.

Year	Dataset	Size
2015	/BuToJpsiKV2_BFilter_XX/.../AODSIM	5M
	/BuToJpsiKV2_BMuonFilter_XX/.../AODSIM	5M
	/BdToJpsiKstarV2_BFilter_XX/.../AODSIM	1M
	/BdToJpsiKstarV2_BMuonFilter_XX/.../AODSIM	1M
	/BsToJpsiPhiV2_BFilter_XX/.../AODSIM	1M
	/BsToJpsiPhi_BMuonFilter_XX/.../AODSIM	1M
2016	/BuToJpsiK_BMuonFilter_SoftQCDnonD_XX/.../AODSIM	20M
	/BdToJpsiKstar_BMuonFilter_SoftQCDnonD_XX/.../AODSIM	75M
	/BsToJpsiPhi_BMuonFilter_SoftQCDnonD_XX/.../AODSIM	80M

the respective efficiency;

- matching a track measured in the silicon tracker with a track measured in a muon station;
- more than five hits on the silicon strips and at least one hit in the silicon pixels;
- transverse and longitudinal impact parameter⁶ cuts, respectively $d_{xy} < 0.3$ cm and $d_z < 20$ cm, with respect to the pp collision vertex (primary vertex).

The muons are paired to form a J/ψ meson and, for the B^0 and B_s^0 channels, the track candidates are paired to reconstruct, respectively, K^{*0} and ϕ mesons. Basic reconstruction requirements are imposed, such as checking the charge of the particle or making sure that a specific track is well reconstructed. Track candidates overlapping with muon candidates are rejected. Some basic selection is also applied: for example, the χ^2 secondary vertex probability⁷ has to be larger than 1%. A very broad mass window is also imposed to the channels under study⁸.

The processing organizes the data in a “ROOT way”, through the `TTree` ROOT class. ROOT is a C++ software framework designed to deal with data processing, statistical analysis, visualisation and storage, and which was often used for the analysis here discussed (many plots were done by using a specific ROOT package, RooFit). The CMS software (CMSSW) is employed. It is composed of one executable and several plug-in modules, which are responsible for the collision information processing. The executable is configured using a configuration file written in Python which indicates which data, modules and parameters should be used, as well as the order in which they should be run. The 2015 and 2016 data and MC samples were processed for the purpose of this analysis, and the corresponding

⁶Distance of closest approach between the line of motion of a particle and the centre of a set of scattering particles. Here the centre refers to the pp collision vertex.

⁷Each time a track is fitted the corresponding χ^2/dof (per degree of freedom) is retrieved. The mentioned probability tells us how likely it is to observe the measured χ^2/dof (or larger) even for a correct model.

⁸For example, for the B^+ channel, the mass distribution is limited from 4.6 to 6.4 GeV. By looking at the $B^+ \rightarrow J/\psi K^+$ mass fits in Section 5.3 we indeed see that the range is very broad, when compared to the width of the peak.

luminosities were obtained using the `brilcalc` command⁹. The luminosities are shown in Tables 3.3 and 3.4 (the LHC data-taking periods are split into “Eras”).

Table 3.3: Values of the luminosities for the 2015 available data eras processed. The CMS Collaboration measured a value of 2.3% for the systematic uncertainty of 2015 luminosity measurements.

Era	Luminosity (fb ⁻¹)
C	0.022 ± 0.001
D	2.515 ± 0.058
Total	2.537 ± 0.058

Table 3.4: Values of the luminosities for the 2016 available data eras processed. The CMS Collaboration measured a value of 2.5% for the systematic uncertainty of 2016 luminosity measurements.

Era	Luminosity (fb ⁻¹)
C	2.611 ± 0.067
D	4.276 ± 0.107
E	2.752 ± 0.069
F	3.109 ± 0.078
G	5.000 ± 0.125
Total	17.748 ± 0.206

3.2 Selection

Once the processing is completed, the analysis selection can be applied. The selection is not performed during the processing since re-running the entire processing stage in case something has to be changed in the selection is extremely time-consuming (the processing of the data used for this analysis took several days).

A series of kinematic constraints is imposed to muons, tracks, mesons and Charmonia candidates, including a displacement cut, $L_{xy}/\sigma_{xy} > 3.5$, being L_{xy} the transverse distance travelled by the meson before its decay, and σ_{xy} the respective uncertainty. This cut takes advantage of the relative large mean lifetime of B mesons. A cut on $\cos(\alpha_{2D})$ was also required; the α_{2D} angle represents the angle between the vector that points from the primary vertex to the secondary vertex and the average of the three-momentum of the B meson decay particles, which is expected to be zero. The already described “SoftMuonID” quality cuts were used. The tracks are selected using recommended “highPurity” cuts: they reject badly reconstructed tracks. The “veto cuts” are also used, to avoid cross-contamination between different channels. Since both the B^0 and the B_s^0 decay into a J/ψ plus two tracks, it sometimes happens that one of the channels is mistakenly reconstructed as the other channel. This may occur when a kaon is misreconstructed as a pion¹⁰. By rejecting ditracks lying in a specific mass window

⁹The `brilcalc` documentation can be found here: <https://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html#brilcalcLumi>.

¹⁰As mentioned in Section 2.2, CMS does not possess a dedicated particle identification detector.

range, the “veto cuts” make sure most contamination is removed (see the last two rows in Table 3.5). For example, if one tries to reconstruct a B_s^0 meson, it will reconstruct two tracks as being kaons, as long as they meet the other selection criteria. It is then important to make sure that the two chosen tracks, when reconstructed as a kaon plus a pion, do not have an invariant mass close to the mass of the B^0 , since in that case they could have been produced by a B^0 meson instead. All the applied selection cuts are summarized in Table 3.5. During the offline selection stage the processed data and MC datasets are further required to match the HLT_DoubleMu4_JpsiTrk_Displaced trigger path. Finally, the B candidates are required to lie in the acceptance range of the analysis: $10 < p_T(B) < 90$ GeV and $|y(B)| < 2.25$.

Table 3.5: Summary of the selection criteria used in this work.

Objects	Selection criteria
Muons	SoftMuonID $p_T > 4.2$ GeV $ \eta < 2.4$
J/ ψ	Mass: $ m_{\mu\mu} - m_{J/\psi} < 150$ MeV $p_T > 8.0$ GeV Dimuon Vertex Probability > 10 %
Tracks	highPurity $p_T > 1.0$ GeV $ \eta < 2.4$ $\chi^2/\text{ndf} < 5.0$ strip hits + pixel hits > 5
Any B meson	Vertex Probability $> 10\%$ $(L_{xy}/\sigma_{xy}) > 3.5$ $\cos(\alpha_{2D}) > 0.99$
B^0	$ m_{K\pi} - m_{K^{*0}} < 50$ MeV $ m_{KK} - m_\phi > 10$ MeV (ϕ veto)
B_s^0	$ m_{KK} - m_\phi < 10$ MeV $ m_{K\pi} - m_{K^{*0}} > 50$ MeV (K^{0*} veto)

Chapter 4

Data MC Comparison

In this work, the MC samples mentioned in Chapter 3 are mainly used to measure the efficiencies which enter in the fragmentation fraction ratios measurement (see Chapter 6). Other studies, such as the estimation of the fraction of misassigned (swapped) tracks in the B^0 channel, or the understanding of how much can the ϕ particle coming from the B_s meson contaminate the double track mass spectrum of the B^0 channel, also rely on MC samples. As a consequence, any deviation of the simulation with respect to the data can have a strong impact on the final results. There is no reason to make one think that the used MC samples flawlessly describe the data. Many factors could cause possible deviations, such as the non-optimal description of a certain physics process or a slightly different detector resolution, calibration and/or alignment¹, among others. A comparison between data and MC is therefore needed.

4.1 The sideband subtraction method

While the MC samples here used contain only the signal corresponding to each of the three channels of interest, the data distributions include background events as well, which are defined as being those that are not produced due to the processes under study. The estimate of the amount of background present in the data samples is then required in order to obtain signal distributions for data variables. For the purpose of the present work, the Sideband Subtraction Method (SSM) is the chosen technique. The data samples are split into one or more background regions (the sidebands), and one signal region, using a discriminating variable which depends on the analysis undertaken. The discriminating variable is expected to have no correlation with the variables under study. The method assumes that there is no signal outside the signal region, *i.e.*, all the events which lie in the defined background regions are indeed background-like. As a consequence, it is reasonable to assume the behaviour of the events in the sidebands to be similar to the behaviour of the background events inside the signal region. By performing an interpolation of the events in the sideband regions, one is then able to estimate the amount

¹The calibration is intended to eliminate or reduce the bias in the reading of a certain instrument over a range for all continuous values. For this purpose, reference standards with known values for selected points covering the range of interest are measured with the instrument in question. The goal of the detector's alignment is to reduce the χ^2 of the fits to the tracks and to reduce the bias and uncertainty of the fitted track parameters. Correlated displacements of sensors which introduce a track parameter bias, but do not change the mean χ^2 , are the core problem of alignment/calibration [49].

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. This has to be done for each variable one wants to study, in order to compare all of them with the MC.

Table 4.1: Signal and background regions used for the Sideband Subtraction Method. All values have GeV units.

Channel	B^+	B^0	B_s
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

In this analysis, we use the mass of the involved B meson candidates as the discriminating variable. The mass spectra can be found in Fig. 4.1. The background is modelled by an exponential function, $Ne^{-\lambda}$, where ‘N’ and ‘ λ ’ are parameters left free to oscillate during the unbinned maximum likelihood fit (see Section 5.1). The background fits for the three channel of interest can be found in Fig. 4.2, for both linear and logarithmic scale. The signal mass window covers the regions defined in Table 4.1. For the B^0 and B_s channels two background regions are defined, on the left and on the right of the peak region. The $B^+ \rightarrow J/\psi K^+$ decay is, however, an exception: the background region to the left of the signal peak 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, as discussed in Section 5.3. As such, for the case of the B^+ channel, the background region is defined only to the right of the signal peak. The right sideband is thus larger than the right sidebands of the other channels to compensate for this effect.

By performing an integral² over the exponential function, the number of background events under the signal peak is obtained. The data signal distributions are obtained using the following formula, which is applied to all studied variables:

$$\text{signal} = \text{full} - \alpha \times \text{background}, \quad (4.1)$$

where ‘signal’ and ‘full’ refer to the data in the signal region, ‘background’ refers to the data in the sidebands, and α is given by:

$$\alpha = \frac{N_{\text{central}}}{N_{\text{left}} + N_{\text{right}}}, \quad (4.2)$$

where ‘ N_x ’ refers to the number of events under the exponential fit in the specified ‘X’ region. Note that in Eq. 4.2, for the case of the B^+ channel, N_{left} will be zero. Also note that the data distributions in Eq. 4.2 are described by histograms, where information about single events is not present; that information is lost since α is determined using fits.

This method has some potential shortcomings that one must take into account. First of all, there is always the possibility of defining a background region which still contains some signal events. This problem can be solved by avoiding the definition of background regions which are adjacent to the signal

²ROOT uses by default the QAGS routine from the FORTRAN90 QUADPACK library. This routine employs an adaptative method, and it is also capable of dealing with integrable singularities.

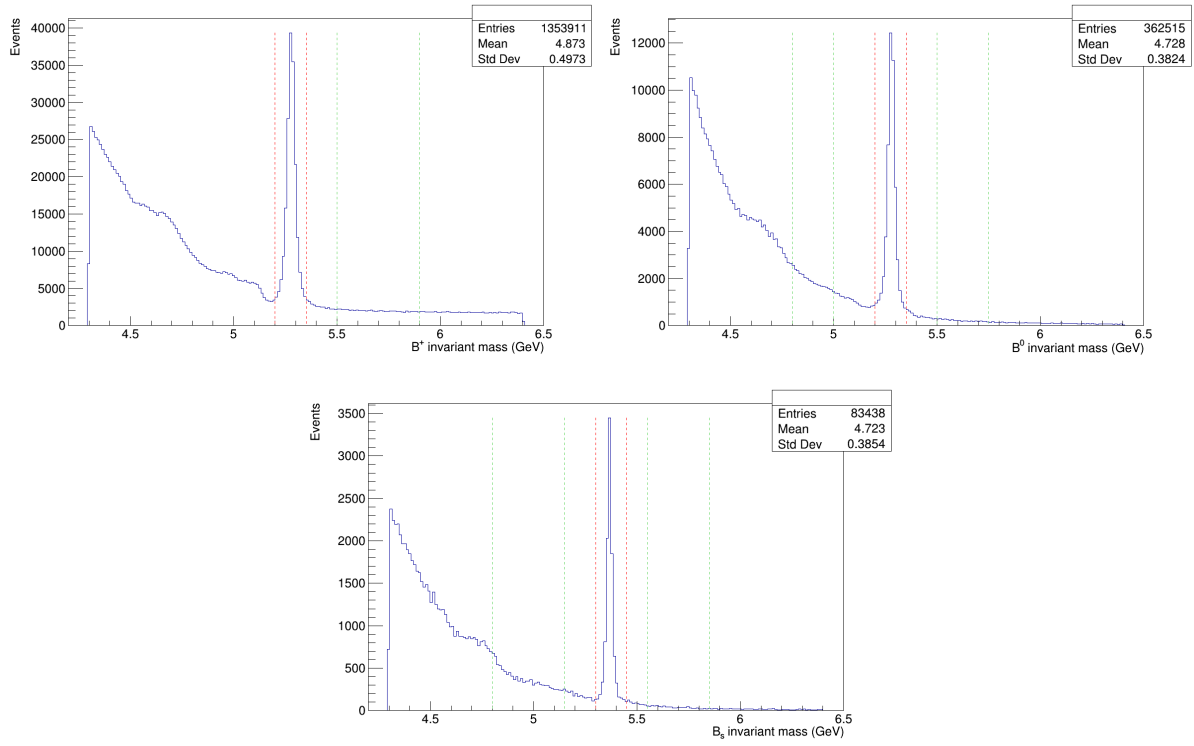


Figure 4.1: B candidates invariant mass spectra for the three channels of interest: B^+ (top left), B^0 (top right) and B_s (bottom). These events passed the selection cuts of the analysis displayed in Table 3.5. The dashed lines indicate the regions considered for the SSM: the signal region is depicted in red, while the sidebands are shown in green.

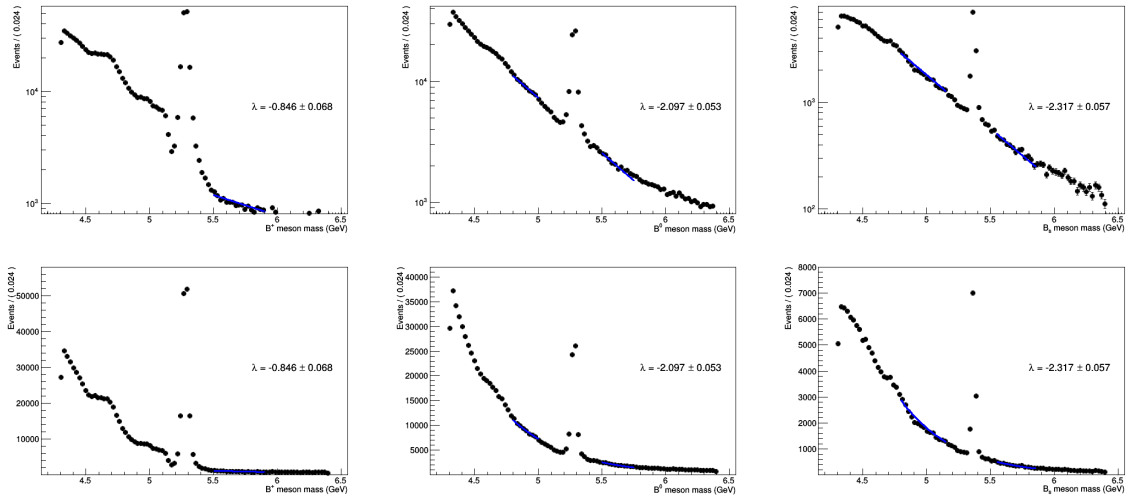


Figure 4.2: Background exponential fits for the B^+ , B^0 and B_s channels (from left to right), in both logarithmic (top) and linear (bottom) scales. The λ is the parameter of the exponential, which was left free to oscillate during the fit. The whole dataset is shown to facilitate the understanding of the plot; the fit was performed to the sidebands only (blue).

region. Still, some precaution is needed. The definition of sidebands in regions far from the peak could also imply a negative outcome: the background could cease to be of combinatorial nature only, due to the existence of some other processes. In that case, the exponential function used to perform the fit would no longer rightly describe the background events under the peak, where only combinatorial background events are expected (that is why the left sideband was not considered for the case of the B^+ channel). Something that also occurs is the existence of bins with a negative number of events, after the method is applied. This happens when performing the subtraction in Eq. 4.1: the fit to the background is subjected to statistical fluctuations and, as a consequence, in the regions of a certain variable where no signal is present the method subtracts two bins where only background-like events are present. These bins may have a slightly different numbers of events (α is obtained from a fit). This unphysical result is mentioned for completeness only; it has a very minor effect on the results. An additional source of concern would be the existence of a variation according to the used sidebands. Different sideband regions were tested, and the results did not change in any significant way, which adds to the robustness of the SSM. The sidebands had nevertheless to be large enough for the exponential fit to converge. In Figure 4.3 we show an example of one distribution both for the chosen nominal sidebands (see Table 4.1) and for an alternative choice.

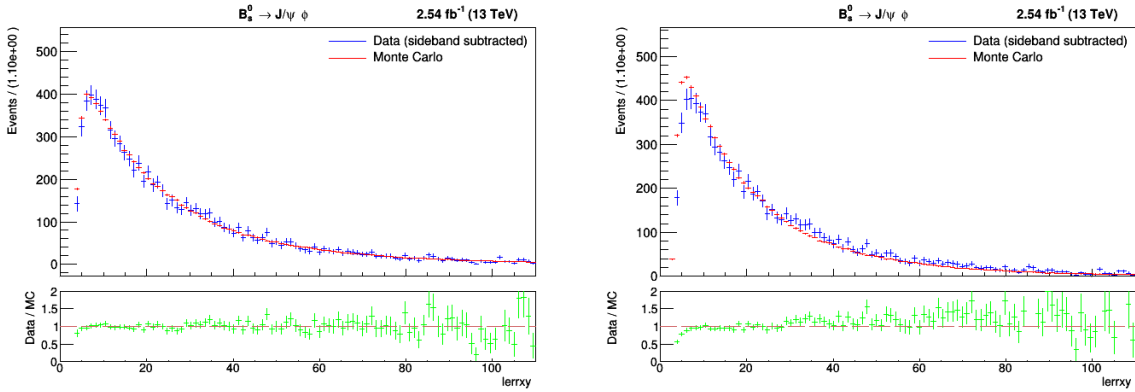


Figure 4.3: L_{xy} significance (defined in the text) distribution for the $B_s^0 \rightarrow J/\psi \phi$ channel, for the nominal (left plot) and an alternative (right plot) sidebands. The SSM does not depend on the chosen sidebands, as long as the exponential fit to the background converges. The optional region of this example uses a range between 4.6 and 5.0 GeV on the left sideband, and between 5.75 and 6.15 GeV on the right sideband.

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 (data is for example constrained by the physical processes under study, by the luminosity provided by the LHC and by the CMS trigger, while MC production constraints are related with time and storage space). The distributions of the following variables are compared:

- “pt”: $p_T(B)$;
- “y”: rapidity of the B mesons;
- “mu1pt” / “mu2pt”: $p_T(\mu_1) / p_T(\mu_2)$;
- “mu1eta” / “mu2eta”: $\eta(\mu_1) / \eta(\mu_2)$ (muon pseudorapidity);

- “tk1pt” / “tk2pt”: $p_T(\text{track}_1)$ and $p_T(\text{track}_2)$ (for the B^+ case there is $p_T(\text{track}_1)$ only);
- “tk1eta” / “tk2eta”: pseudorapidity of the hadron tracks (for the B^+ case there is $\eta(\text{track}_1)$ only);
- “lxy”: distance between the primary and secondary vertexes in the transverse plane, L_{xy} ;
- “errxy”: L_{xy} uncertainty, σ_{xy} ;
- “lerrxy”: L_{xy} significance (L_{xy}/σ_{xy});
- “vtxprob”: decay vertex probability;
- “cosalpha2d”: cosine of the pointing angle between the vector that joins the primary and secondary vertexes and the vector sum of the three-momenta of the decay products of the B meson decay; it is here denoted as $\cos(\alpha_{2D})$.

The most striking features of the comparison are worth mentioning. They follow a similar pattern in the three considered channels. The $p_T(B)$ variable is found not to have a very good agreement with its respective simulation. However, it does not create a big concern, since the analysis selection does not include a cut on the $p_T(B)$. The L_{xy} is found to have a poor agreement for the case of the $B^0 \rightarrow J/\psi K^{*0}$ decay channel, but not for the other two channels. As a consequence, the L_{xy}/σ_{xy} variable also shows a bad agreement in the same channel (see bottom plots in Fig. 4.4). This will influence the measurement of the efficiency, since a $L_{xy}/\sigma_{xy} > 3.5$ cut is imposed at selection level. Additionally, the rapidity and pseudorapidity data distributions are found to be asymmetric. This asymmetry is observed for both muons and tracks and, as a consequence, it is also observed in the $y(B)$ distributions. An example of the asymmetry is shown in the top right plot in Fig. 4.4. Nevertheless, considering all compared variables, the MC describes the data well enough so that no *a priori* correction is performed. All the comparison plots can be found in Appendix A.

After performing the normalization, a binned distribution of weights, $W_{\text{Var/Bin}}$, is obtained by dividing the data histogram containing only signal by the MC histogram for each variable we want to compare:

$$W_{\text{Var/Bin}} = \frac{\text{Data}_{\text{Var/Bin}}}{\text{MC}_{\text{Var/Bin}}}, \quad (4.3)$$

where the variable and bin dependence is explicitly written. A large number of binned weights distributions is obtained, equal to the number of variables under study times the three channels of interest. The weights can be found in Appendix A and also in Fig. 4.4, in the panels under the data/MC comparison plots (depicted in green). In Section 7.2 we discuss how the weights were used to calculate the systematic uncertainty of the efficiency.

4.2 Mass resolution

To compare the mass resolution between data and MC, the B meson and the ditrack mass distributions are inspected. For the latter, the data signal is extracted using the sideband subtraction method as described in Section 4.1. For the B invariant mass distributions the signal is fitted with a sum of three Gaussian functions while the background is described by a Chebyshev polynomial of 6th order. The fits can be seen in Figs. 4.5 and 4.6. The estimated resolutions are also summarized in Table 4.2.

Table 4.2: Resolution values for B and double tracks invariant mass distributions. Values are in MeV.

	B^+	B^0	B_s^0
Data			
B mass distribution	25.46 ± 0.91	15.18 ± 0.52	9.42 ± 1.01
Double track mass distribution	—	29.06 ± 0.57	4.24 ± 0.25
MC			
B mass distribution	21.60 ± 0.32	19.95 ± 0.06	9.71 ± 0.01
Double track mass distribution	—	32.36 ± 0.31	4.00 ± 0.03

The analysis selection includes cuts in both the ϕ and K^{*0} double track mass. The differences in mass resolution between data and simulation can therefore have an impact in the analysis: in Section 7.3 a systematic uncertainty is assigned to this source. The differences in the resolution of the B meson distributions do not affect this work, since no selection cut was applied on the mass of the B mesons.

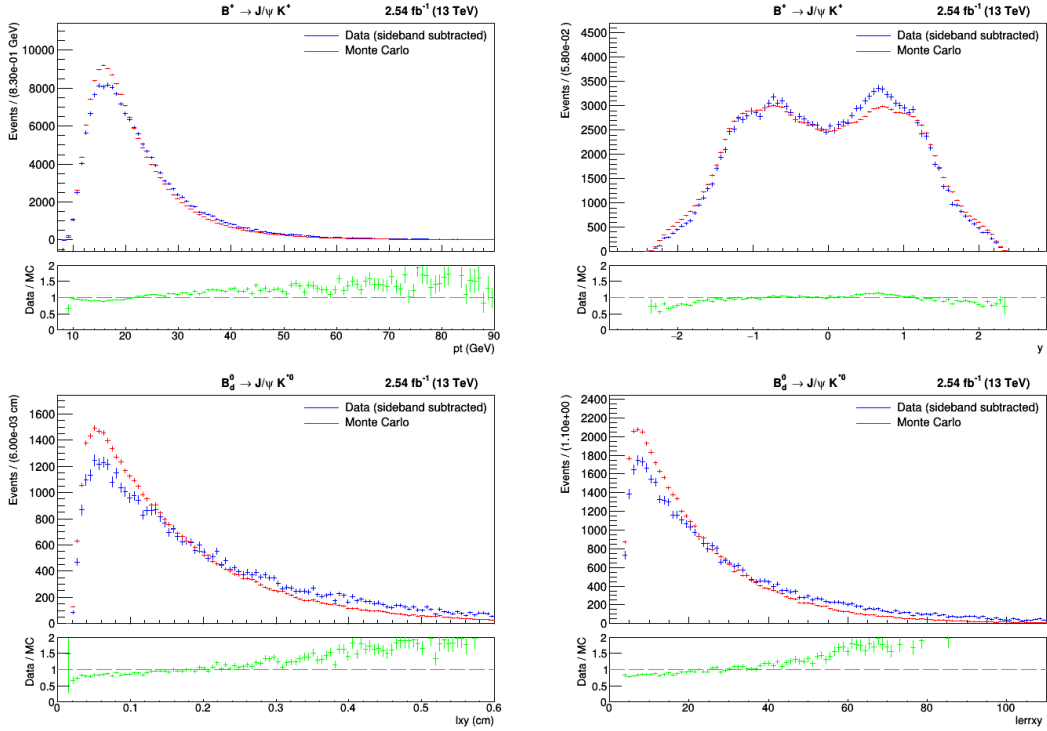


Figure 4.4: Data / MC comparison for some of the variables that did not show a very good agreement. In the first row the $p_T(B)$ and the $y(B)$ are shown for the $B^+ \rightarrow J/\psi K^+$ channel, although a similar behaviour can be seen for the other two channels. The second row refers to the $B^0 \rightarrow J/\psi K^{*0}$ channel only, since the behaviour of the same variables is quite good for the other channels (see Appendix A).

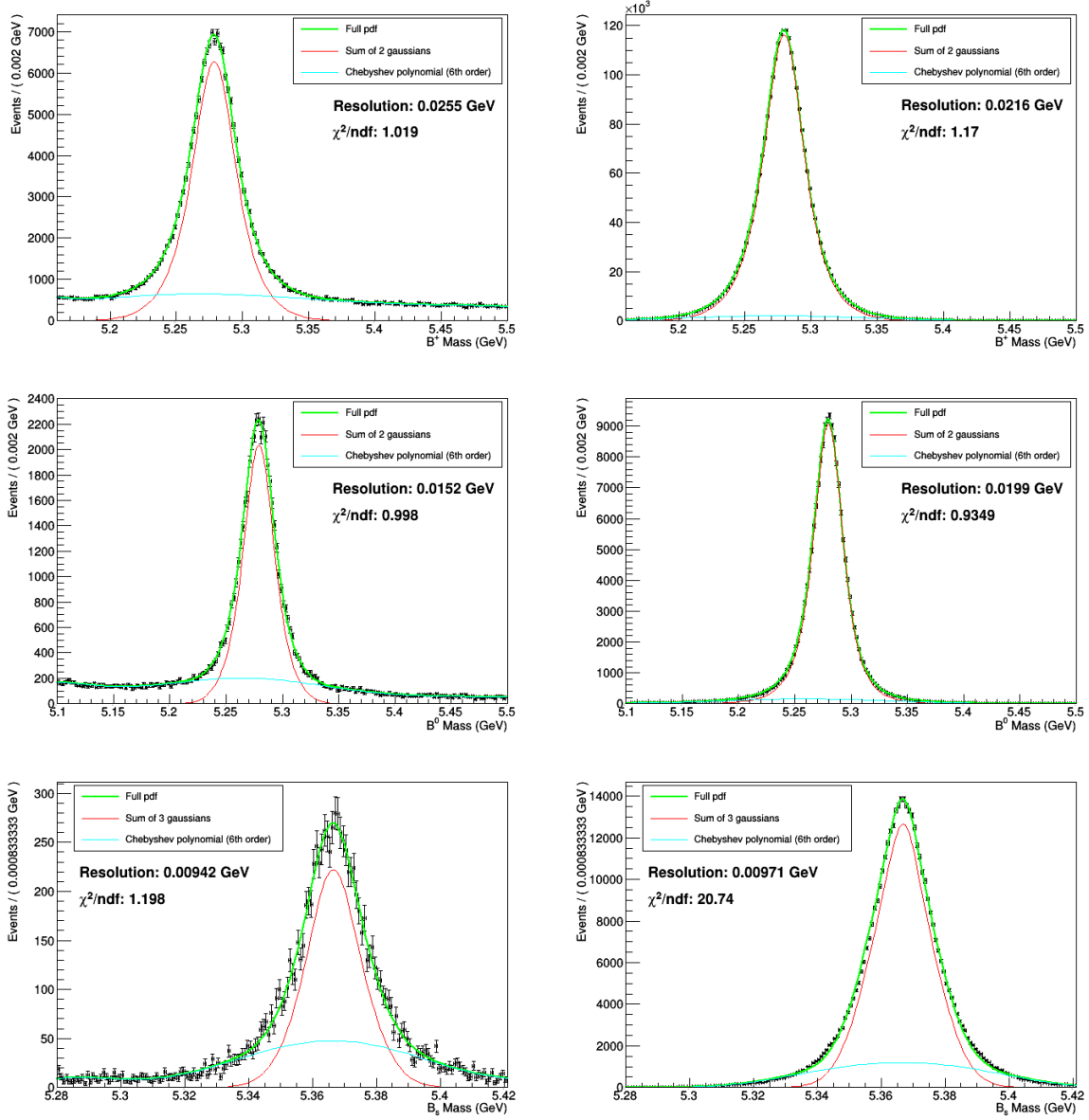


Figure 4.5: B meson invariant mass spectra for the three channels of interest: B^+ (top), B^0 (middle) and B_s^0 (bottom). The left column displays data while the right one refers to MC. The spectra were fitted with a maximum likelihood fit consisting of a sum of three Gaussians plus a Chebyshev polynomial to describe some asymmetries. The uncertainties are displayed in Table 4.2.

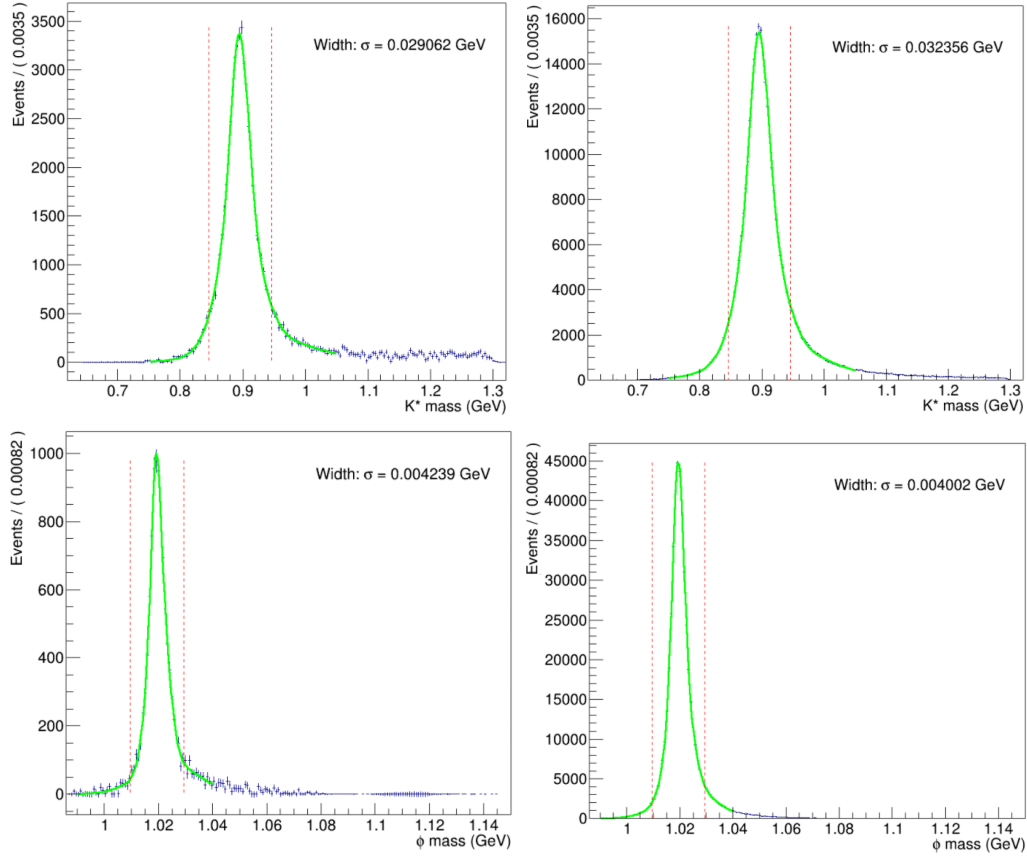


Figure 4.6: Double track invariant mass spectra: $K^{*0} \rightarrow K^+\pi^-$ channel (top) and $\phi \rightarrow K^+K^-$ (bottom). The data distributions were obtained after applying the SSM technique. The spectra were fitted in the central region with a maximum likelihood fit consisting of a sum of three Gaussians. The dashed red lines represent the mass window applied during the selection stage. The uncertainties are displayed in Table 4.2.

Chapter 5

Yields

The determination of the signal yield for the three channels of interest constitutes one of the major ingredients of the analysis performed in this work. We start this section by first giving a brief description of the unbinned maximum likelihood method, a statistical tool which was intensively used to determine the yields later mentioned. Next, the “K π swap” effect is explained, and the invariant mass fits are shown.

5.1 Unbinned Extended Maximum Likelihood method

The maximum likelihood method is a statistical method used to obtain the most likely values for a given set of unknown parameters, given the available data sample. The method is characterized by the likelihood function, which for the continuous case with n measurements is given by:

$$\mathcal{L}(\underline{\theta}|\underline{x}) = f_{\underline{X}}(\underline{x}|\underline{\theta}) = \prod_{i=1}^n f_{X_i}(x_i|\underline{\theta}), \quad \underline{\theta} \in \Theta, \quad (5.1)$$

where $\underline{X} = (X_1, X_2, \dots, X_n)$ represents a n -dimensional random sample of independent and identically distributed random variables, $\underline{x} = (x_1, x_2, \dots, x_n)$ is the observed value of $\underline{X}$, f_{X_i} is the probability density function (PDF) for the X_i random variable, $\underline{\theta}$ is the set of values of the unknown parameters whose likelihood is being evaluated, and Θ is $\underline{\theta}$'s domain, *i.e.*, the values that can be given to the parameters contained in $\underline{\theta}$ [50]. The most likely values of the unknown parameters are estimated by maximizing the likelihood function. One usually takes advantage of the fact that it is analytically much easier to deal with the logarithmic of the likelihood rather than with the likelihood defined in Eq. 5.1. The maximum of the logarithmic likelihood, considering p parameters, is then obtained by satisfying the following requirement:

$$\frac{\partial \ln \mathcal{L}[(\theta_1, \theta_2, \dots, \theta_p)|\underline{x}]}{\partial \theta_j} = 0, \quad j = 1, \dots, p. \quad (5.2)$$

In addition, the Hessian matrix $\nabla^2 \ln \mathcal{L}[(\theta_1, \theta_2, \dots, \theta_p)|\underline{x}]$ has to be negative-definite (all its eigenvalues must be negative). The requirements can be better understood when one considers only one parameter. Indeed, its maximum will be obtained by finding the point which has a null first-derivative and a negative

second-derivative. Note that the method is unbinned. Finally, it is worth mentioning that the likelihood method here used has an additional feature: in its “extended” version, the normalization is allowed to vary according to a Poisson distribution centered about some true mean value, $A(\underline{\theta})$, which we require $f_{\underline{X}}(x|\underline{\theta})$ to be normalized to: $\int f_{\underline{X}}(x|\underline{\theta})dx \equiv A(\underline{\theta})$. The likelihood then looks like [51]:

$$\mathcal{L}_{\text{extended}}(\underline{\theta}|x) = \frac{A(\underline{\theta})^n e^{-A(\underline{\theta})}}{n!} \prod_{i=1}^n \frac{f_{X_i}(x_i|\underline{\theta})}{A(\underline{\theta})}, \quad \underline{\theta} \in \Theta, \quad (5.3)$$

The extended likelihood therefore describes both the shape and the size of the distribution [52].

The unbinned maximum likelihood fits of the next section are performed using `ROOT`, which is a package belonging to the `ROOT` modular scientific software framework. The ranges included inside Θ are defined by the user, as well as the PDFs, which are empirically selected to be able to accurately model the data. The numerical program responsible to find the maximum likelihood value is called `MINUIT`.

5.2 $K \pi$ swap

Given the absence of dedicated particle identification subdetector components for protons, kaons and pions (see Chapter 2), it is not trivial to distinguish them in the detector. As a result, the $K^{*0} \rightarrow K^+ \pi^-$ decay may be reconstructed with swapped hadron mass assignments: it is unknown whether the first track is a kaon and the second one is a pion, or the other way around. The pair which has an invariant mass closer 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 swapped component can be seen in the B^0 invariant mass fits from Section 5.3).

The existing MC signal dataset is split into those two categories, and each of them is separately fitted. Both the true and swapped event shapes are extracted using three Crystal Ball (CB) functions. A CB function is made out of a Gaussian function in its core plus a power-law tail defined after a certain threshold. In both cases two of the CB functions have a tail in the low-mass region, while the third CB has a tail in the right side of the mass distribution. The shape of the swapped component is then combined with the shape of the true pairs and fitted to the B^0 MC mass distribution, by fixing all the parameters except the relative fraction of swapped events, which is obtained. This parameter represents the relative amount of swapped events with respect to the sum of both swapped and non-swapped events. This fraction is also estimated from direct counting (using MC), and it is found to be compatible with the value extracted from the fit. The fits are shown in Fig. 5.1 in linear and logarithmic scales. The fraction of swapped events obtained is $(12.91 \pm 0.17)\%$ and it is fixed in the B^0 mass fit (see the yellow component of all B^0 mass fits in Appendix C).

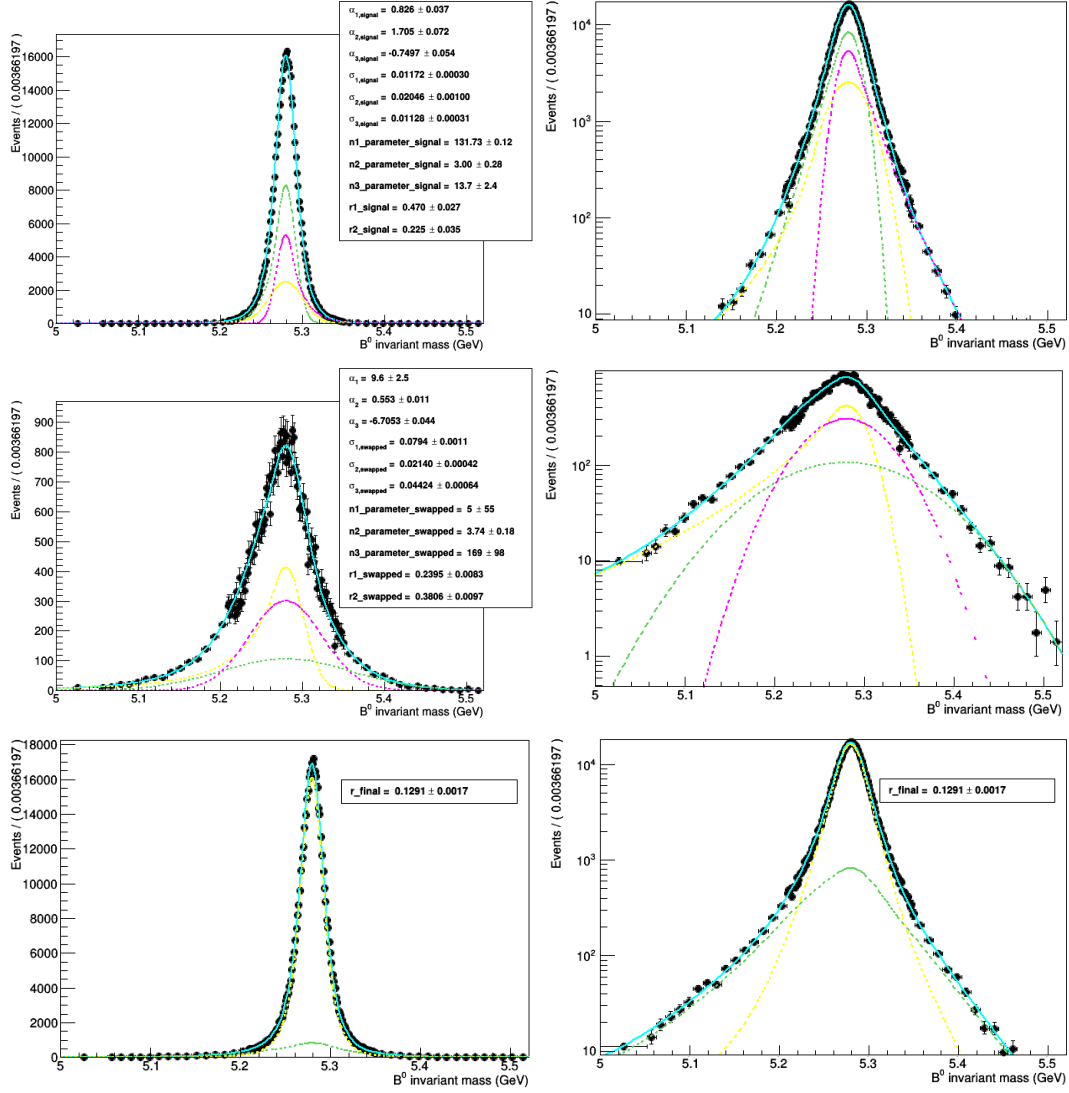


Figure 5.1: Unbinned maximum likelihood fit to the B^0 MC true (top row), swapped (middle row) and full (bottom row) signal invariant mass, both for linear (left column) and logarithmic (right column) scales. The full dataset corresponds to the sum of the true with swapped events. The cyan line describes the full PDF in each case. For the first two rows, the pink, yellow and green lines reflect the behaviour of each of the three Crystal Ball functions used. For the last row, the yellow line represents the true signal and the green line represents the swapped signal. For the case of the left column, the parameters obtained from the fit are shown.

5.3 Invariant mass fits

The three channels of interest ($B^+ \rightarrow J/\psi K^+$, $B^0 \rightarrow J/\psi K^{*0}$ and $B_s^0 \rightarrow J/\psi \phi$) are fitted using the extended unbinned maximum likelihood method, considering the following general PDF:

$$f_{X_i}(m; N_S) = N_S \cdot [\alpha G(m; M, \sigma_1) + (1 - \alpha) G(m; M, \sigma_2)] + N_B \cdot E(m; \tau_m), \quad (5.4)$$

where ‘m’ is the candidate’s mass, while ‘M’ and ‘ $\sigma_{1/2}$ ’ are the signal mass mean and widths (resolution). The Gaussian and Exponential functions are represented by the ‘G’ and ‘E’ letters, and are normalized to the fitting mass window. The signal and background yields are represented as, respectively, ‘ N_S ’ (the parameter of interest) and ‘ N_B ’. The ‘ α ’ and ‘ τ_m ’ parameters describe, respectively, the signal fractions and exponential decay slope.

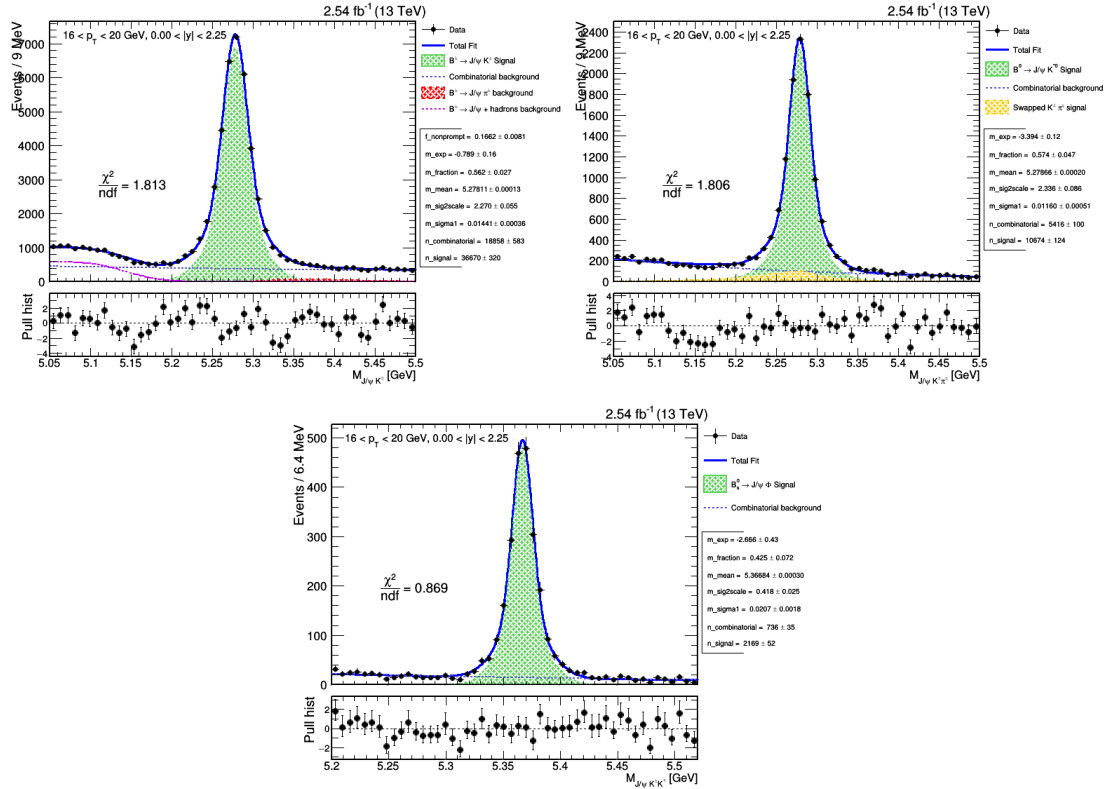


Figure 5.2: Unbinned extended maximum likelihood fit for the B^+ , B^0 and B_s^0 channels, for the $16 < p_T < 20$ GeV bin.

Other channel-specific components are considered for the $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow J/\psi K^{*0}$ channels and are added to Eq. 5.4. For the case of the latter, the signal PDF has an additional component that is used to describe the signal candidates that were misreconstructed as described in Sec. 5.2; both its fraction, $(12.91 \pm 0.17)\%$, and its shape are constrained in the fit. For the $B^+ \rightarrow J/\psi K^+$ channel, some physics backgrounds are considered: the Cabibbo-suppressed $B^+ \rightarrow J/\psi \pi^+$ decay, where the pion is misidentified as a kaon, and the $B \rightarrow J/\psi + h + X$ partially reconstructed b hadrons decay, where the hadron ‘h’ is reconstructed as a kaon and the other decay products, ‘X’, are missed. While the former physics background appears as a small wide bump to the right of the main B^+ peak, the latter dominates

the left region of the invariant mass spectrum.

The $J/\psi \pi^+$ background component is fixed in the fit: the shape (which is modelled with three Gaussians) is taken from a previous study using simulations, and the normalization is taken from the relative 4% branching fraction [5]. The partially reconstructed B^+ background component is modelled with a complementary error function, with magnitude and slope free to float in the fit.

The datasets are split into $p_T(B)$ and $y(B)$ bins and each of them is separately fitted. As an example, the fits for the $16 < p_T < 20$ GeV bin are shown in Fig. 5.2. The full set of fits is provided in Appendix C. The fit results, displaying the fitted parameter values and the reduced chi-square, are overlaid on the figures. The fits in this section rely on a computational framework that was already set¹, such that only minor adjustments were needed before running the code.

¹The framework was developed by Bruno Galinhas, a former PhD student at LIP.

Chapter 6

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 events are not properly reconstructed. Therefore, the signal yield which appears in the master formula for this analysis, Eq. 1.23, 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. The data/MC disagreements seen in Chapter 4 will then have an impact in the efficiency measurement, since the efficiency should describe the data. This effect will be accounted for when determining its respective systematic uncertainty in Section 7.2. A data-driven efficiency measurement method is discussed in Section 7.4.

An efficiency is simply obtained by measuring the fraction of signal that passes a certain requirement, or cut. We start by taking a MC sample containing generator level information, *i.e.*, a sample that gives access to the properties of the generated particles before being reconstructed. As shown in Table 3.2, two MC samples per channel are used for 2015. The first one¹ has no cuts whatsoever, while the second sample² has *a priori* cuts applied before the generation of the events (see Table 6.1). The second sample is needed since p_T distributions usually peak at relatively low values and decay exponentially (see the data/MC p_T comparison plots in Appendix A). Therefore, in order to get reasonable yields in the higher $p_T(B)$ bins, the generation of events with low $p_T(B)$ has to be avoided. The cuts in Table 3.2 are then used such that the measurement of the efficiency in high $p_T(B)$ bins is possible. We first measure the efficiency of those unphysical cuts by using the first MC samples. We call this efficiency the “prefilter efficiency”:

$$\varepsilon_{\text{pre}} = \frac{N(B \mid p_T(\mu)^{\text{gen}} > 2.8, |\eta(\mu)^{\text{gen}}| < 2.4, 10 \text{ GeV} < p_T(B)^{\text{gen}} < 90 \text{ GeV}, |y(B)^{\text{gen}}| < 2.25)}{N(B \mid 10 \text{ GeV} < p_T(B)^{\text{gen}} < 90 \text{ GeV}, |y(B)^{\text{gen}}| < 2.25)} . \quad (6.1)$$

Then, the second MC sample is used to measure the efficiency of the reconstruction and of all selection

¹The one which contains the substring “BFilter” (see Table 3.2).

²The one which contains the substring “BMuonFilter” (see Table 3.2).

cuts. We call it the “reconstruction efficiency”:

$$\varepsilon_{\text{reco}} = \frac{N(B \mid \text{Analysis cuts})}{N(B \mid p_T(\mu)^{\text{gen}} > 2.8, |\eta(\mu)^{\text{gen}}| < 2.4, 10 \text{ GeV} < p_T(B)^{\text{gen}} < 90 \text{ GeV}, |y(B)^{\text{gen}}| < 2.25)} \cdot \quad (6.2)$$

By multiplying these two efficiencies we obtain the overall efficiency, given by:

$$\varepsilon = \frac{N(B \mid \text{Analysis cuts})}{N(B \mid 10 \text{ GeV} < p_T(B)^{\text{gen}} < 90 \text{ GeV}, |y(B)^{\text{gen}}| < 2.25)} \cdot \quad (6.3)$$

Note that the detector’s acceptance is already included in this definition, and that the total efficiency for each B meson channel is denoted as ε_u , ε_d or ε_s , as defined in Eqs. 1.21, 1.22 and 1.23. For the 2016 MC samples the calculation of the efficiencies is very similar except for the unphysical cuts applied (see Table 6.1). The definitions in Eqs 6.1, 6.2 and 6.3 change accordingly. The 2015 MC samples without cuts were also used for the 2016 efficiency measurement since at generator level nothing changed between 2015 and 2016. This explains why only one 2016 sample per channel appears in Section 3. The efficiencies are calculated thanks to a set of C++ macros³ which use the `TEfficiency` class belonging to the ROOT software framework, making use of the embedded implementation for uncertainty calculations. The statistical uncertainties of the prefilter and reconstruction efficiencies are added in quadrature in order to obtain the statistical uncertainty of the total efficiency. For the fragmentation fraction ratios, the

Table 6.1: Cuts imposed to the second MC sample used for the determination of the efficiency for each channel of interest. There is an exception: for the $B^+ \rightarrow J/\psi K^+$ channel no cut was imposed to the single kaon in the 2016 MC sample.

Year	Cuts
2015	$p_T(\mu) > 2.8, \eta(\mu) < 2.4$
2016	$p_T(\mu) > 2.5, \eta(\mu) < 2.5$ $p_T(\text{track}) > 0.4, \eta(\text{track}) < 2.5$

corresponding efficiency ratios are determined from the single channel efficiencies. In Fig. 6.1 and 6.2 we show the overall efficiencies for the considered channels and ratios.

The interpretation of the shapes of the single channel efficiencies is straightforward: CMS is more efficient in detecting muons with high p_T . In the limit where the p_T is very low, muons curve so much under the influence of the magnetic field that they cannot reach the muon stations. For the shapes as a function of the $y(B)$, the efficiency goes down as we approach the regions that are not covered by CMS, namely, 2.5 for tracks and 2.4 for muons in terms of pseudorapidity. A drop at low values of rapidity is surprisingly seen and not clearly understood. As an educated guess, we can assume that the lower efficiency arises due to gaps between the muon detectors, that exist at those values of $y(B)$ [41]. It is difficult to think about a reason for this drop that is related with the selection, since no cuts were applied in the $y(B)$; a correlation between this variable and other variables to which we are applying a cut would hardly explain the phenomenon, since it would have to affect only part of the full rapidity range.

The ratios are somewhat more difficult to interpret. Still, we try to do so by considering the number of

³The framework was developed together with Bruno Galinhas, a former PhD student at LIP.

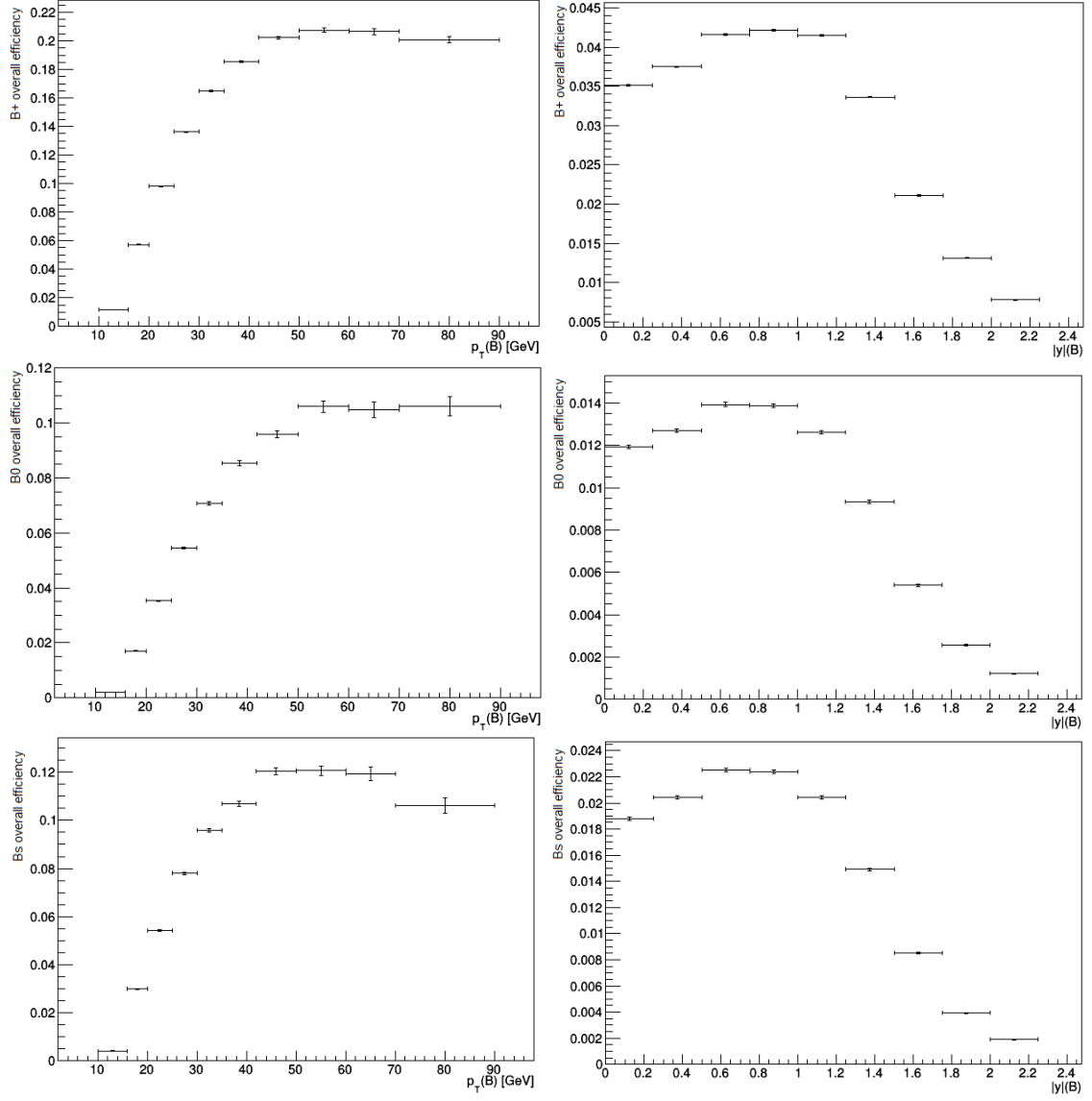


Figure 6.1: Overall efficiencies of the three channels of interest as a function of $p_T(B)$ (left) and $y(B)$ (right) bins. From top to bottom, the B^+ , B^0 and B_s^0 channels are shown.

tracks for each decay channel⁴, as well as the mass of each B meson. It is harder to efficiently measure two tracks when compared to just one. As such, the efficiency of the $B^+ \rightarrow J/\psi K^+$ channel is expected to be higher, especially for low values of p_T . As we approach of p_T , the differences will tend to become smaller, since it is easy to “see” the tracks, even if in larger numbers. A similar reasoning can be applied when looking to the plots of the distributions as a function of $y(B)$; this time, however, the region where it is harder to detect the tracks is the one with high rapidity values. What about the $\varepsilon_d/\varepsilon_s$ plots? For those, since the number of tracks is the same, a more subtle approach is required. The mass of the B_s^0 meson is higher than the mass of the B^0 meson, and as a consequence its decay products will have on average higher $p_T(B)$. That is why $\varepsilon_s > \varepsilon_d$ for all bins. This difference is more notorious for low values of $p_T(B)$, since for those values it is more difficult to detect a particle. As a result, for low $p_T(B)$ bins the B mass difference represents an important factor. For higher values of $p_T(B)$, most tracks are measured. The mass difference becomes negligible. Indeed, for the last bin the ratio is approximately equal to unity. In terms of rapidity, no clear trend is expected. The efficiency of the $B_s^0 \rightarrow J/\psi \phi$ channel is always higher than the $B^0 \rightarrow J/\psi K^{*0}$ channel efficiency, quite likely due to the mass difference mentioned previously. If one looks at the vertical scale, one sees that the efficiency ratio is approximately constant, especially when compared with the other two ratios.

⁴The muons are here not considered since all channels include the $J/\psi \rightarrow \mu^+ \mu^-$ decay.

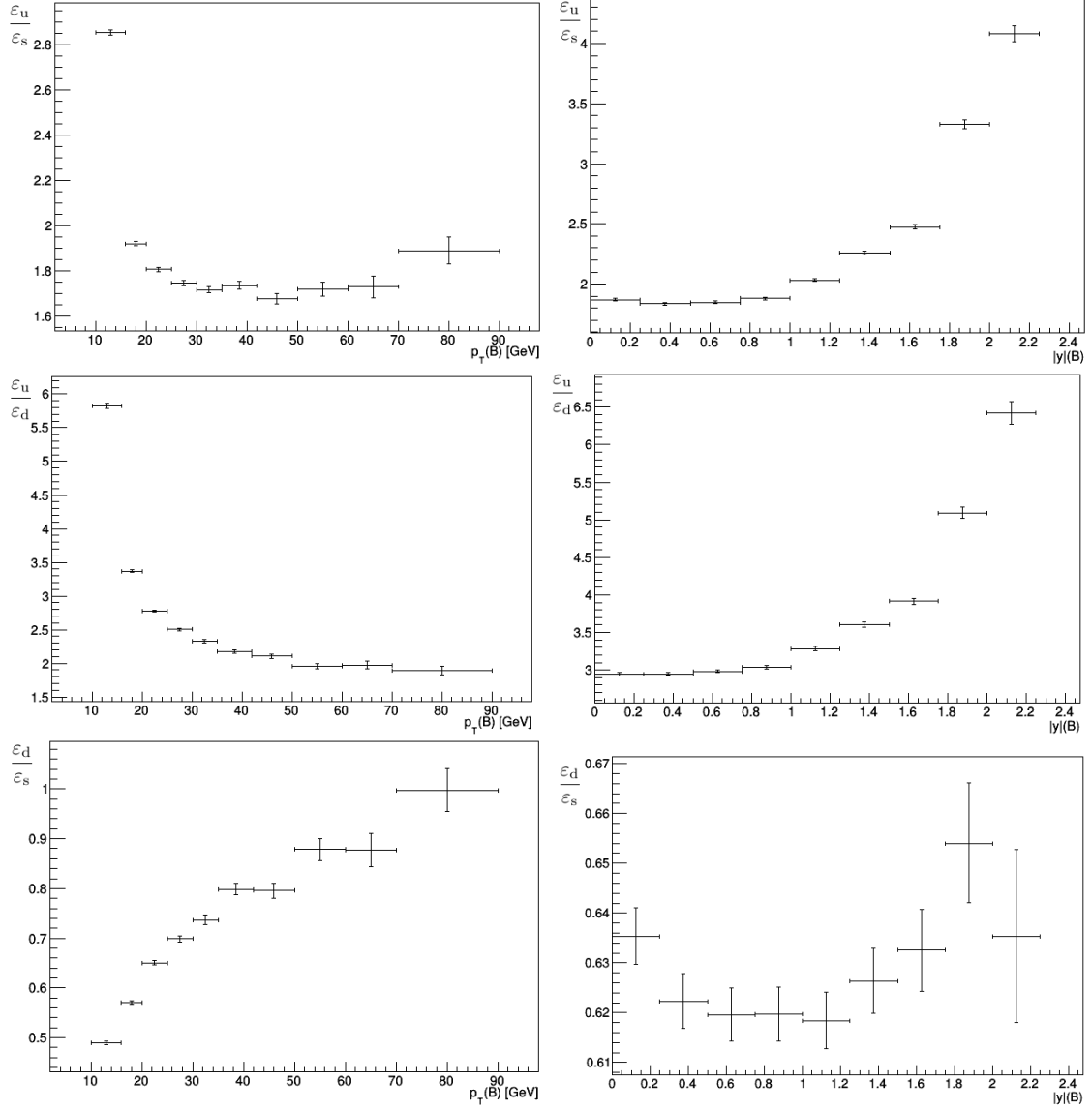


Figure 6.2: Efficiencies of the three fragmentation fraction ratios studied as a function of $p_T(B)$ (left) and $y(B)$ (right). From top to bottom, the $\varepsilon_u/\varepsilon_s$, $\varepsilon_u/\varepsilon_d$ and $\varepsilon_d/\varepsilon_s$ efficiency ratios are shown.

Chapter 7

Systematics

This chapter serves the purpose of explaining how the systematic uncertainties of the fragmentation fraction ratio measurements are calculated. One cannot reasonably consider all possible sources of systematic uncertainties. Instead, we here study the ones that can 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 $B_s^0 \rightarrow \mu^+ \mu^-$ branching fraction measurement, the latter is equally bad, since the vertical error bars would give a wrong idea on where the “true” value lies.

7.1 Fit bias

For this section we use of one of the most important principles in probability theory, the Central Limit Theorem (CLT), to quantitatively measure some bias that may have unintentionally been introduced to the analysis. The CLT theorem establishes that, when dealing with n independent and identically distributed random variables $X_1, X_2, \dots, X_n$, with $n \geq 30$ (in general), the probability distribution function of both the sum and of the arithmetic mean of those variables can be approximated by a normal distribution [50]. If one further performs the following normalization:

$$Z \equiv \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}, \quad (7.1)$$

being $\bar{X}$ the sample mean which follows a normal distribution with mean μ and variance σ^2/n , then Z will follow a Gaussian with null mean and unitary standard deviation.

For the case of our analysis, we want to know if the parameter of interest obtained from the fit, the signal yield, follows the CLT predicted pattern. The yield is the parameter of interest since it is directly used to obtain the fragmentation fraction ratios (Eqs. 1.21, 1.22 and 1.23). Simulated samples of pseudo-data, the so-called “Toy MC”, are generated according to the PDFs described in Section 5.3. 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:

$$\frac{N - \mu(N)}{\sigma(N)}, \quad (7.2)$$

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 $\mu(N)$ and $\sigma(N)$ are, respectively, the mean and the standard deviation of the yield distribution. The Toy MC samples are generated according to a Poisson distribution since the original fits to the data included an extended term (see Section 5.1). 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 the way the PDFs are implemented. We show the pull distributions for the B^+ channel in Fig. 7.1.

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. Given that the deviations of the pull means from zero are found to be small, the resulting systematic uncertainty is only a fraction of the statistical uncertainty. For each channel, we consider as systematic uncertainty the worst value among all studied bins; this is summarized in Table 7.1. The same table shows the uncertainty of the fragmentation fraction ratios, which is obtained by adding in quadrature the contributions from the three involved channels.

Table 7.1: Global fit bias systematic uncertainty for the three channels of interest and for the three ratios.

Channel	Syst. uncertainty (%)	Ratio	Syst. uncertainty (%)
B^+	0.518	f_d/f_u	0.603
B^0	0.309	f_s/f_u	1.090
B_s^0	0.959	f_s/f_d	1.008

7.2 MC reweighting

For all the variables specified in Chapter 4, the data and MC signal distributions are compared and their agreement is quantified by a set of weights obtained according to Eq. 4.3, that we here repeat:

$$W_{\text{Var/Bin}} = \frac{\text{Data}_{\text{Var/Bin}}}{\text{MC}_{\text{Var/Bin}}}, \quad (7.3)$$

where “Var” refers to the variable being inspected and “Bin” to the bin where the event to be corrected lies. We note that the data distributions here considered correspond to signal only: the mass sideband method was already applied, as described in Section 4.1. The weights follow binned distributions since the mass sideband subtraction method involves the subtraction of two histograms. We recall that the MC distribution was normalized to the area of the data distribution. Once the weights are used to correct

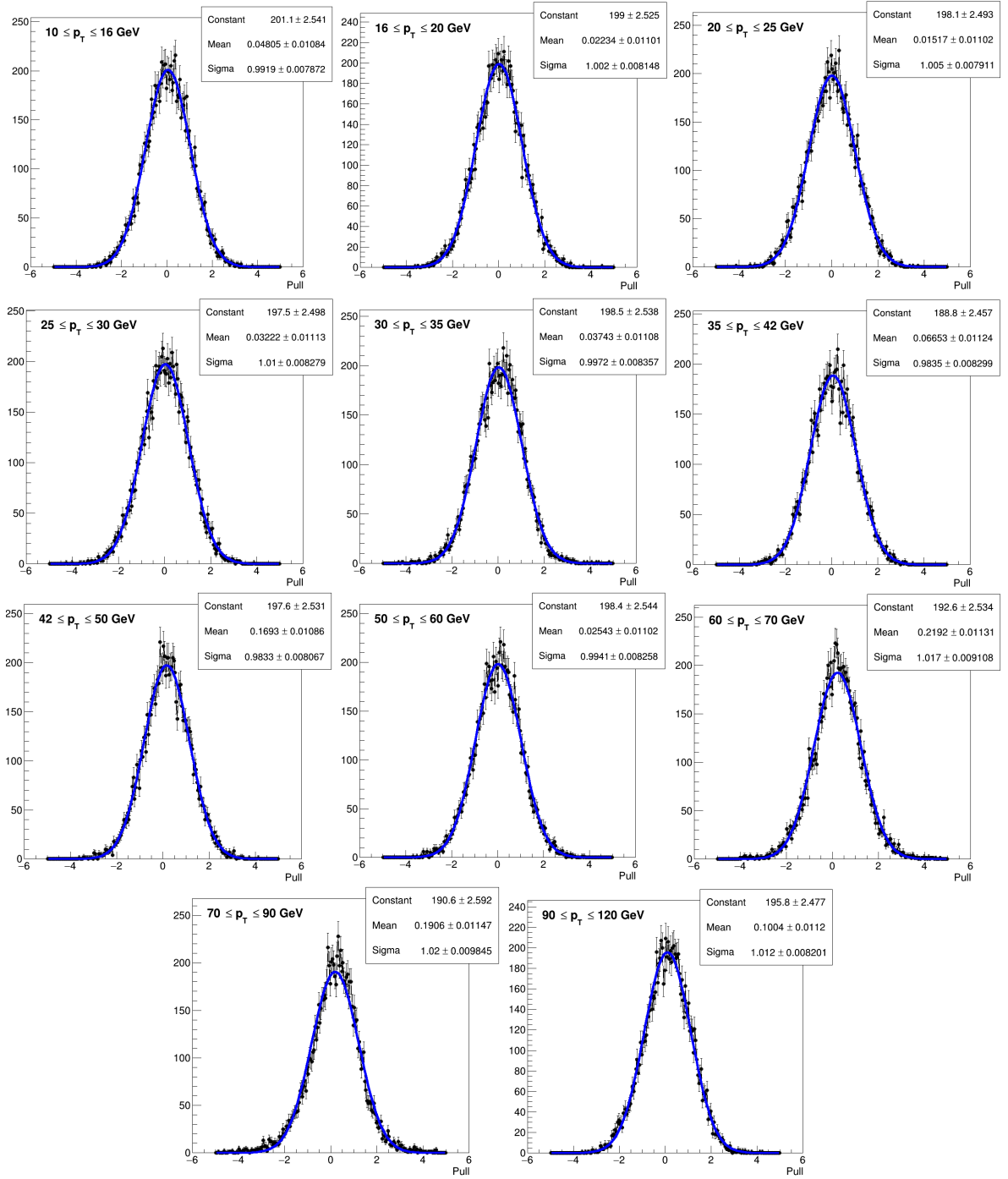


Figure 7.1: Pull distributions for the number of signal events for the B^+ channel, for all considered $p_T(B)$ bins. A similar analysis was performed for the B^0 and B_s^0 channels.

the MC distribution from which they were calculated a perfect description of the data for the variable being corrected is obtained. The procedure of correcting the MC to match the observed data is called “reweighting”. The reweighting technique can be used to directly correct a MC distribution, whenever the MC poorly describes the data, or else for evaluating the systematic uncertainties related to the quality of the MC agreement. 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. The worse the data/MC agreement

is, the less accurate will the efficiency determination be.

Ideally, the weights calculated from a particular variable would be able to correct not only the same variable, which gets corrected by construction, but also all the other variables under study. In reality, however, that is not the case. Some weights correct some variables better than others, such that it is always necessary to combine different distributions of weights. This can be done in different ways. One possibility would be to determine multi-dimensional sets of weights, instead of the one-dimensional weights described by Eq. 7.3. Just as an example, it would be possible to divide two-dimensional histograms of the L_{xy}/σ_{xy} and $\cos(\alpha_{2D})$ variables to obtain two-dimensional weights that would then by construction correct those two variables instead of just one. This procedure can be extended to more dimensions, eventually matching the number of variables under study, such that all used variables are perfectly described by the MC. In this work we decided to pursue a different approach. Still using the one-dimensional weights, we considered more than one set and saw which of them represented the largest systematic source in each of the three studied fragmentation fraction ratios.

We decided to use the binned weight distributions coming from the following variables (the “reweighting variables”) to correct the nominal MC distribution:

- $p_T(\text{track}_1)$,
- $\eta(\text{track}_1)$,
- $p_T(\mu_1)$,
- $\eta(\mu_1)$,
- L_{xy} significance.

We further note that the $p_T(B)$ and $y(B)$ variables should not be used as reweighting variables, since no selection cuts are performed on these variables, and as such the overall efficiency cannot depend on them. The weights were set to 1 whenever the lack of reasonably large yields produced a weight which was not trustworthy. The distributions of all the variables considered for this analysis (see Section 4.1) before and after being reweighted by the five chosen reweighting variables are shown in Appendix B. Fig. 7.2 shows an example in which the weights are used to correct the same variable.

The systematic uncertainty is calculated event by event, although the weights themselves follow a binned distribution, *i.e.*, each event was corrected using weights which are common to other events as well. The systematics are calculated as follows. The nominal efficiency of the analysis is measured by calculating the fraction of the number of events that passed all stages of selection, as explained in Chapter 6. The corrected reconstruction efficiency will be instead measured using the following formula:

$$\varepsilon_{\text{corrected reco}} = \frac{\sum_{\text{Events}} w_{\text{Var/Bin}}}{N(B|p_T(\mu)^{\text{gen}} > 2.8, |\eta(\mu)^{\text{gen}}| < 2.4, 10 \text{ GeV} < p_T(B)^{\text{gen}} < 90 \text{ GeV}, |y(B)^{\text{gen}}| < 2.25)}, \quad (7.4)$$

where the “Events” mentioned in the numerator refer to the same events that appear in the numerator of the nominal reconstruction efficiency in Eq. 6.2, and $w_{\text{Var/Bin}}$ refers to the single weights present in the $W_{\text{Var/Bin}}$ sets. In the limiting case where the MC perfectly agrees with the data, $w_{\text{Var/Bin}} = 1$ and Eqs. 6.2

and 7.4 are equivalent. The systematic uncertainty is taken as the largest variation obtained among all reweighting variables used and it is obtained with the following expression for each of the analysis bins:

$$\text{Syst. uncertainty} = \frac{\left| \frac{\varepsilon_{\text{corrected reco X}}}{\varepsilon_{\text{corrected reco Y}}} - \frac{\varepsilon_{\text{reco X}}}{\varepsilon_{\text{reco Y}}} \right|}{\varepsilon_{\text{reco X}} / \varepsilon_{\text{reco Y}}}, \quad (7.5)$$

where both 'X' and 'Y' represent one of the three channels of interest ($X \neq Y$). The uncertainty could have been calculated using the single efficiencies instead of using the ratios, but it was verified that the final systematic would be more conservative for the three ratios under study. It is an expected effect, since this work deals with ratios of channels with similar topologies. Finally, it is necessary to explain the reason why the prefilter efficiency (see Eq. 6.1) is not considered. Indeed, the measured efficiency is the overall efficiency and not just the reconstructed one. This happens because the weights obtained with Eq. 7.3 only refer to events that already passed the reconstruction phase. It is impossible to find an equivalent for generated events, which are the kind of events the prefilter efficiency deals with, since that would imply finding data at generator level, which makes no sense whatsoever. As a consequence, the systematic uncertainty of the overall efficiency can be measured by studying the reconstruction efficiency alone. We finally stress that the weights were used for the systematic uncertainty determination only, and not for correcting the nominal MC.

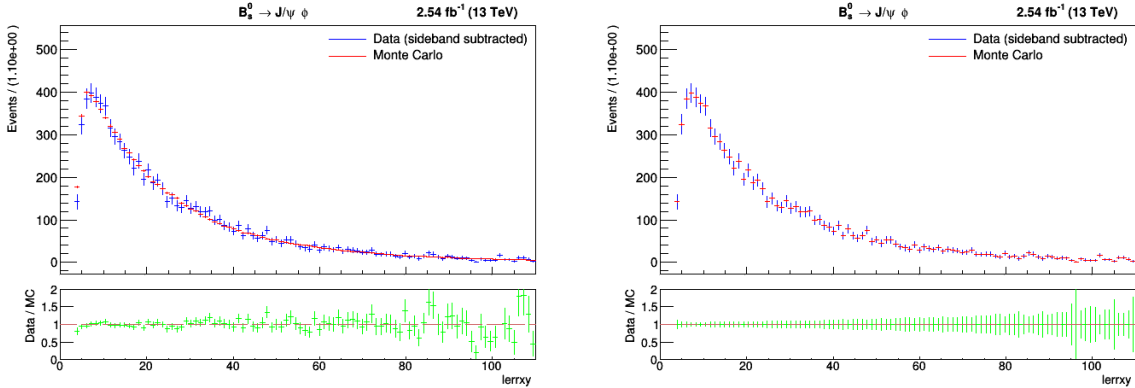


Figure 7.2: Example of the reweighting technique. On the left plot the original data/MC comparison for the L_{xy} significance variable is shown. On the right plot the same variable is corrected using the L_{xy} significance as the reweighting variable. By construction, the two distributions now perfectly match.

To give an example, let us focus on a particular event. We consider the five mentioned reweighting variables, one at a time. We choose one, and we check its value for this particular event. Once we have that value we can extract the corresponding weight, which came from the data/MC comparison of the reweighting variable we are considering. As long as the event does not have a value for the reweighting variable which lies outside the region where the comparison was done (and the regions are large enough so that virtually no event misses it), a weight will always be found. If we are considering $p_T(\text{track}_1)$ as the reweighting variable, and its value is found to be 15 GeV for this particular event, we can find the corresponding weight looking at the weight existing at $p_T(\text{track}_1) = 15 \text{ GeV}$, which was obtained by comparing the data and MC distributions of the $p_T(\text{track}_1)$ variable. Then, when calculating the corrected efficiency, instead of considering one event, we will consider its weight. We recall that the

efficiency is calculated simply by summing all events that pass a certain set of cuts, and then dividing this number by the total number of events. Instead of adding ‘1’ for each event that successfully passed the cuts (numerator), we add ‘ $w_{\text{Var/Bin}}$ ’ in order to obtain the systematic uncertainty (check Eq. 7.4). The process is repeated for all reweighting variables. By comparing all the corrected efficiency ratios with the nominal efficiency ratios, a systematic uncertainty can be extracted. The largest deviation among these comparisons with different reweighting variables is stored as the systematic uncertainty. This must be done per bin and per fragmentation fraction ratio under study. The resulting systematic uncertainties are shown in Tables 7.13, 7.14 and 7.15.

7.3 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 3.2), the correct measurement of the double track mass cuts efficiency relies on the MC samples which are used to establish those cuts. As we saw in Section 4.2, the widths of the mass distributions of both the K^{*0} and ϕ were, 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 were produced. The efficiencies of the cuts were measured by fitting the double track mass distributions and later calculating the integral of the fitting function inside the used mass window (see Fig. 7.3). The mass distribution was fitted using a sum of three Gaussians for the peak plus a 6th 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. 7.3 the fits are shown, where the mentioned areas refer to a comparison between the fit function integral inside the mass window with the integral over the full fit range. The former is taken as the double track mass cut efficiency, after dividing by the latter. The systematic uncertainty was obtained by using the following formula:

$$\text{Syst. uncertainty} = \frac{|\varepsilon_{\text{Data}} - \varepsilon_{\text{MC}}|}{\varepsilon_{\text{MC}}} \quad (7.6)$$

The MC efficiency appears in the denominator because it is the nominal efficiency in this work. The obtained systematic uncertainties are depicted in Table 7.2. The ratio uncertainties are obtained by adding the single channel efficiency uncertainties in quadrature, and are displayed in Table 7.3.

Table 7.2: Systematic double track window mass cut uncertainty, for the B^0 and B_s^0 channels.

	B^0	B_s^0
Data efficiency	0.783 ± 0.001	0.849 ± 0.001
MC efficiency	0.786 ± 0.001	0.899 ± 0.001
Systematic (%)	0.04	5.51

Table 7.3: Global fit bias systematic uncertainty for the measured fragmentation fraction ratios.

Ratio	Syst. uncertainty (%)
f_d/f_u	0.04
f_s/f_u	5.51
f_s/f_d	5.51

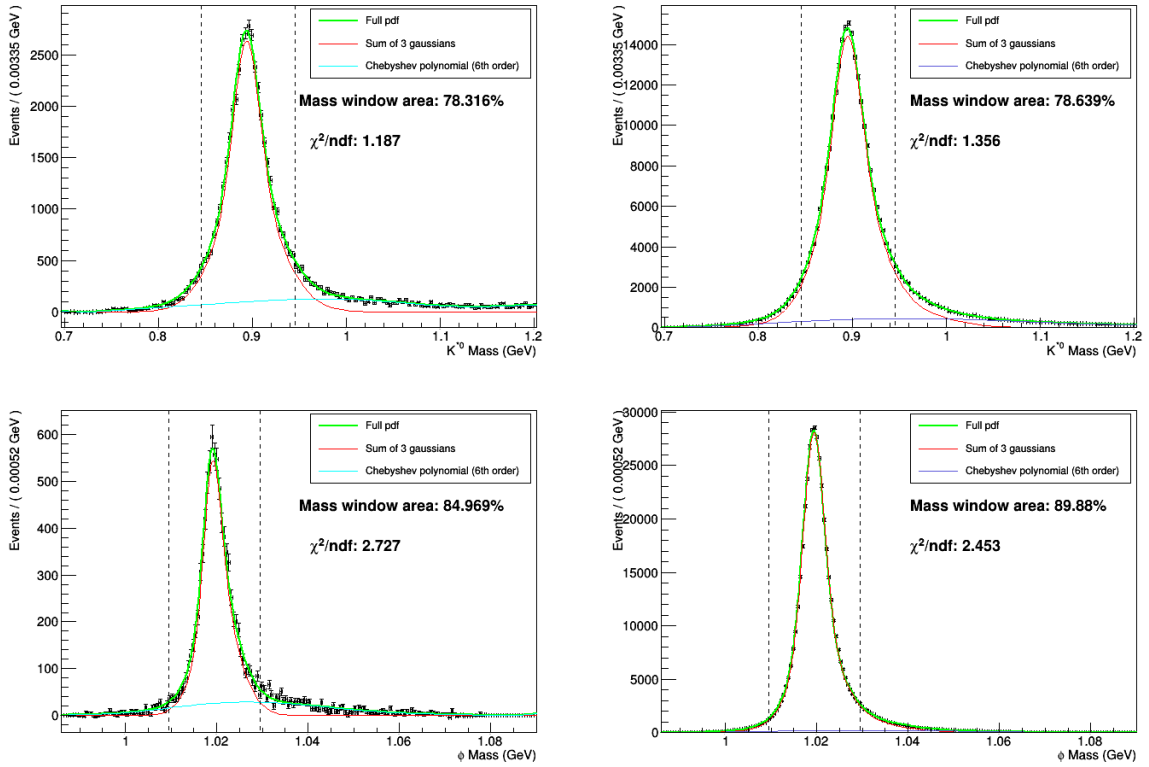


Figure 7.3: Double track invariant mass spectra: $K^{*0} \rightarrow K^+\pi^-$ channel (top) and $\phi \rightarrow K^+K^-$ channel (bottom); the left column refers to data and the right column to MC. The data distributions are obtained after applying the SSM technique. The spectra are fitted with a maximum likelihood fit consisting of a sum of three Gaussians plus a Chebyshev polynomial to describe some asymmetries.

7.4 Tag and Probe

Muons are copiously produced at the LHC. The high rates involved require the use of triggers (described in Section 2.4). These only select the muons which might contribute to 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 (Chapter 6) 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 single muon efficiencies¹ by exploiting a very clever concept. It works as follows: tight selection criteria, including single muon triggers, are applied to a group of reconstructed muons. Their misidentification is then very unlikely. These particles are almost certainly real muons: the “tag” muons. They 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. A trigger matching is imposed to the tag muons, not to the probe muons, and as a consequence the probe muons will not be biased with respect to the trigger. The matching between the tag and the probe muons 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 thus 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. We give an example: if one measures the single muon efficiency of the L3 part of the trigger (see Chapter 2), the passing probes will be the ones which get selected by the L3 trigger, while the failing probes will be all the muons with loose selection criteria that did not pass this specific selection stage. 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 (see Fig. 7.4). The efficiency is directly extracted from the signal, according to:

$$\varepsilon_{\text{cut}} = \frac{N_{\text{signal passing probes}}}{N_{\text{signal passing probes}} + N_{\text{signal failing probes}}}, \quad (7.7)$$

where ‘N’ refers to the yields, and ‘cut’ to the selection applied to the probe muons.

The data-driven efficiencies and their resemblance to MC single muon efficiencies were measured by the CMS Collaboration applying the same muon selection to the probes as the muon selection applied

¹ It is very often used to measure single electron efficiencies as well.

in this work. The total single muon efficiency was split into four terms, according to:

$$\varepsilon_{\mu} = \varepsilon_{\text{Track}} \cdot \varepsilon_{\text{SoftMuonID}} \cdot \varepsilon_{\text{L1L2}} \cdot \varepsilon_{\text{L3}} ,$$

where the offline muon tracking efficiency in the silicon tracker, $\varepsilon_{\text{Track}}$, was extensively studied by the CMS Collaboration in the past. Its value was measured to be $(99 \pm 1)\%$, and it is independent of the p_T or η of the muons. The other three quantities have the following meaning:

- $\varepsilon_{\text{SoftMuonID}}$: offline SoftMuonID cuts efficiency (see Section 3.2) with respect to a silicon tracker muon²;
- $\varepsilon_{\text{L1L2}}$: combined L1 and L2 muon trigger efficiency with respect to an offline reconstructed muon which passed the SoftMuonID and the silicon tracker cuts;
- ε_{L3} : L3 muon trigger efficiency with respect to an offline reconstructed muon which passed all previous cuts.

We note that each individual term refers to an efficiency calculated given a set of cuts that were applied before. For example, $\varepsilon_{\text{L1L2}}$ is the L1 and L2 efficiency given the previous SoftMuonID and silicon tracking cuts. The L1, L2 and L3 triggers were already described in Chapter 2. The efficiencies were measured in terms of the $p_T(\mu)$ and $\eta(\mu)$ bins of the single muons. The used binning was the following:

- $p_T(\mu) = 2, 2.5, 3, 3.5, 4, 4.75, 5.5, 7.5, 10, 15, 20, 40 \text{ GeV}$;
- $|\eta(\mu)| = 0, 0.2, 0.3, 0.9, 1.2, 1.6, (2.1, 2.4)$;

where the parenthesis indicate that for those bins only the $\varepsilon_{\text{SoftMuonID}}$ was calculated. The other efficiencies, $\varepsilon_{\text{L1L2}}$ and ε_{L3} , were measured up to $|\eta(\mu)| = 1.6$. The passing and failing distributions were fitted to the J/ψ peak (see Fig. 7.4).

We take advantage of $\mathbb{W}(p_T(\mu), \eta(\mu))$ binned weights obtained in the past by the CMS Collaboration by comparing single muon data-driven with single muon MC-driven efficiencies, which are given by:

$$\mathbb{W}(p_T(\mu), \eta(\mu)) = \frac{\text{Data}(p_T(\mu), \eta(\mu))}{\text{MC}(p_T(\mu), \eta(\mu))} . \quad (7.8)$$

Since the nominal efficiencies are obtained using a purely MC approach, the Tag and Probe procedure is used to determine the systematic uncertainties which arise from the differences between data and MC. The weights in Eq. 7.8 quantify the difference. Three different set of weights are available, corresponding to the $\varepsilon_{\text{SoftMuonID}}$, $\varepsilon_{\text{L1L2}}$ and ε_{L3} efficiencies. We show an example of those weights in Fig. 7.5. In this work we calculate the systematic uncertainty of the SoftMuonID, L1 and L2, and L3 single muon efficiencies separately 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 $J/\psi \rightarrow \mu^+ \mu^-$ decay) is obtained. Looking at the available binned weights distributions (Eq. 7.8) the corresponding weights are extracted. For each B

²Technically, these probes are not necessarily muons; they are tracks which passed the invariant mass range requirement (relative to the tag muon), but no information coming from the muon chambers is used.

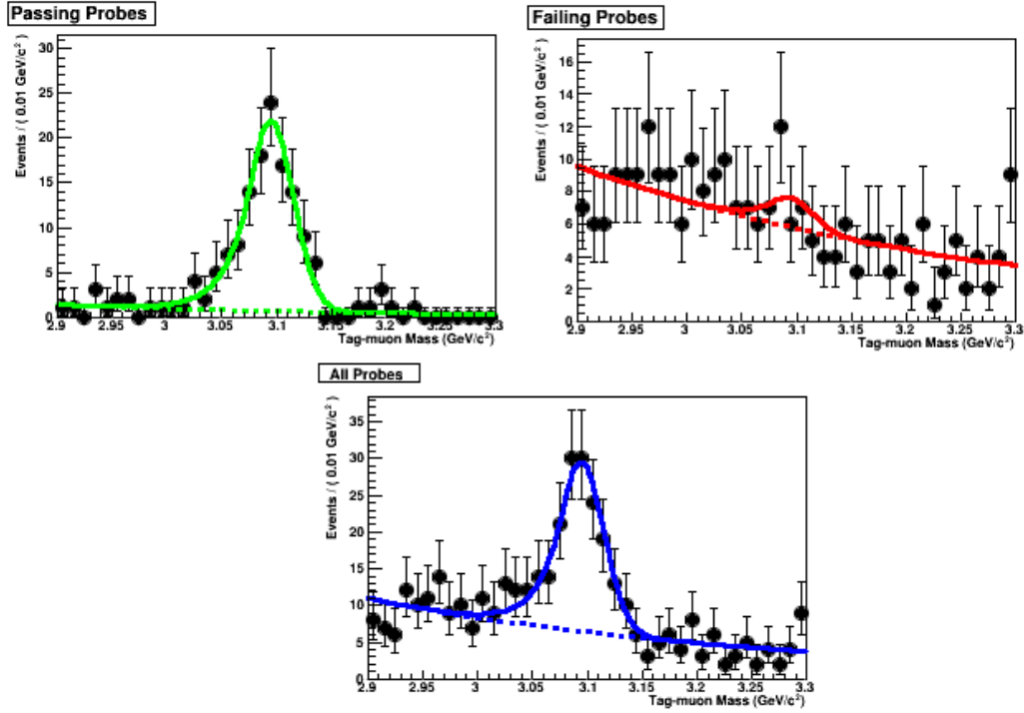


Figure 7.4: Example of Tag and Probe study performed using early 2015 CMS data. The signal and background were described by different fit functions, and were simultaneously fitted to the J/ψ resonance peak. The upper-left plot shows the passing probes distribution, while the upper-right plots shows the failing probes distribution. The bottom plot depicts the distributions of the two groups combined. These plots were created by Leonardo Cristella.

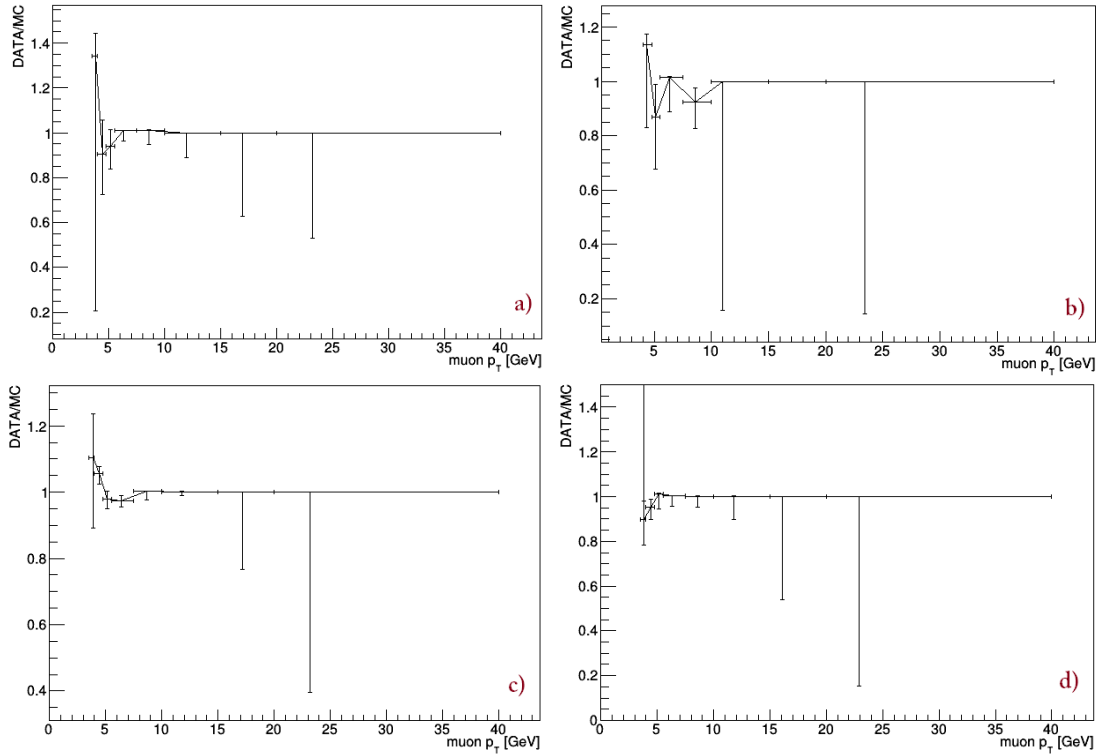


Figure 7.5: Example of a subset of Tag and Probe weights which are used in this work. Here the weights refer to L3 efficiencies. The bins of rapidity shown are, from top left to bottom right: $0 < |\eta(\mu)| < 0.2$ (a), $0.2 < |\eta(\mu)| < 0.3$ (b), $0.3 < |\eta(\mu)| < 0.9$ (c) and $0.9 < |\eta(\mu)| < 1.2$ (d). The points were connected by a line. The weights were obtained by Leonardo Cristella.

meson, the two weights are multiplied³. This procedure is repeated for all reconstructed B mesons in the three channels of interest. When all the pairs of muons have their corresponding weights assigned, the systematic uncertainty is calculated per channel and per bin of p_T and y of the B meson using the following expression:

$$\text{Systematic uncertainty} = \left| 1 - \frac{\sum_{\text{B mesons}} (\mathbb{W}_{\mu_1} \cdot \mathbb{W}_{\mu_2})}{\mathcal{N}} \right| \quad (7.9)$$

where $\mathcal{N}$ refers to the number of reconstructed B mesons in a certain $(p_T(B), y(B))$ bin. The $\mathbb{W}$ dependence on the p_T and η of the muons was omitted to simplify the notation. We draw the reader's attention to a detail which can be at first confusing: while the weights $\mathbb{W}$ are split into $(p_T(\mu), \eta(\mu))$ bins, as shown in Eq. 7.8, the obtained systematic uncertainties are calculated per $p_T(B)$ and $y(B)$ bins. The last step is necessary since the fragmentation fraction ratio measurements are performed in $p_T(B)$ and $y(B)$ bins. The results can be seen in Tables 7.4, 7.5 and 7.6 as a function of $p_T(B)$ and in Tables 7.7, 7.8 and 7.9 as a function of $y(B)$. The contributions from the three sources were quadratically added, and the final results can be seen Table 7.10 for the three channels of interest.

What about some potential events which had muons with $p_T > 40$ GeV or $\eta > 1.6$ (or 2.1 for the ϵ_{MuonID})? No weights were calculated for these events, and as such we imposed $\mathbb{W} = 1$. We do not expect to have many muons with those extreme kinematic values. Still, it was necessary to check whether they could significantly change the obtained systematic uncertainty. For the B^+ , B^0 and B_s^0 events, respectively, only 5.6%, 3.9% and 3.8% of the muons were in this situation. We believe that these numbers are low enough so that no significant change to the final systematic occurs.

Table 7.4: Soft Muon ID trigger systematic uncertainties in $p_T(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.

$p_T(B)$ (GeV)	10-16	16-20	20-25	25-30	30-35	35-42	42-50	50-60	60-70	70-90	90-120
B^+	0.517	0.292	0.443	1.352	3.080	5.813	9.860	13.438	14.488	13.719	11.190
B^0	0.465	0.229	0.244	0.909	2.377	4.788	8.722	12.604	13.034	12.880	—
B_s^0	0.535	0.234	0.214	0.772	2.276	5.060	9.297	12.073	13.257	12.128	—

Table 7.5: L1.L2 trigger systematic uncertainties in $p_T(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.

$p_T(B)$ (GeV)	10-16	16-20	20-25	25-30	30-35	35-42	42-50	50-60	60-70	70-90	90-120
B^+	0.852	1.687	2.582	3.189	3.255	2.864	1.889	0.642	0.135	0.547	0.679
B^0	0.529	1.388	2.348	3.098	3.368	3.127	2.265	0.904	0.128	0.590	—
B_s^0	0.624	1.486	2.495	3.252	3.437	3.027	1.930	0.615	0.366	0.921	—

³Possible correlations between the two muons induced by the detector or by the trigger system are here ignored. This correlation cannot be studied using the Tag and Probe technique, since the latter deals with single muon triggers which do not store any information regarding the second muon. In this work this detail is not very significant, since it would refer to an uncertainty of the uncertainty we are calculating (second-order effect).

Table 7.6: L3 trigger systematic uncertainties in $p_T(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.

$p_T(B)$ (GeV)	10-16	16-20	20-25	25-30	30-35	35-42	42-50	50-60	60-70	70-90	90-120
B^+	1.170	0.875	0.592	0.445	0.365	0.282	0.215	0.188	0.141	0.118	0.107
B^0	1.308	1.122	0.765	0.563	0.432	0.347	0.266	0.211	0.198	0.140	—
B_s^0	1.377	1.099	0.716	0.515	0.413	0.317	0.246	0.210	0.171	0.131	—

Table 7.7: Soft Muon ID trigger systematic uncertainties in $y(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.

$y(B)$	0 - 0.25	0.25 - 0.5	0.5 - 0.75	0.75 - 1	1 - 1.25	1.25 - 1.5	1.5 - 1.75	1.75 - 2	2 - 2.25
B^+	0.198	0.144	0.146	0.708	1.554	1.458	6.029	9.947	6.603
B^0	0.140	0.102	0.052	0.925	2.041	1.791	7.242	12.337	7.721
B_s^0	0.163	0.106	0.119	0.821	1.770	1.611	6.723	11.112	7.112

Table 7.8: L1.L2 trigger systematic uncertainties in $y(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.

$y(B)$	0 - 0.25	0.25 - 0.5	0.5 - 0.75	0.75 - 1	1 - 1.25	1.25 - 1.5	1.5 - 1.75	1.75 - 2	2 - 2.25
B^+	3.014	2.620	1.686	1.661	1.624	1.604	0.867	0.129	0.003
B^0	3.135	2.646	1.672	1.718	1.767	1.702	0.921	0.110	0.002
B_s^0	3.085	2.594	1.675	1.678	1.699	1.606	0.870	0.105	0.003

Table 7.9: L3 trigger systematic uncertainties in $y(B)$ bins measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.

$y(B)$	0 - 0.25	0.25 - 0.5	0.5 - 0.75	0.75 - 1	1 - 1.25	1.25 - 1.5	1.5 - 1.75	1.75 - 2	2 - 2.25
B^+	1.875	1.953	1.141	0.498	0.127	0.186	0.195	0.047	0.001
B^0	1.818	1.900	1.063	0.492	0.121	0.192	0.196	0.048	0.001
B_s^0	1.911	1.937	1.122	0.490	0.100	0.195	0.192	0.052	0.001

Table 7.10: Global systematic uncertainties measured using weights calculated with the Tag and Probe technique. The uncertainties were obtained using Eq. 7.8 for the three channels of interest. All values are expressed in %.

	Soft Muon ID	L1.L2	L3	Total
B^+	1.481	1.811	0.744	2.455
B^0	1.478	1.964	0.793	2.583
B_s^0	1.303	1.920	0.825	2.463

A fundamental aspect of the fragmentation fraction ratio was still not considered in this section. Looking at Eq. 1.23, we see a ratio of efficiencies. Both of them are calculated using muons which are kinematically similar. We can therefore argue that the systematic uncertainty coming from the Tag and Probe method will cancel for the case of the f_s/f_d measurement. Comparing the values in Table 7.10 for the various channels, we indeed see that they are quite similar. By looking at the bin-dependent tables we also see that they follow the same trend. Furthermore, we obtain other systematic contributions which have a stronger impact than this one (see Section 7.7). As such, we decided not to include the systematic uncertainty retrieved with the Tag and Probe method in the final result.

By looking at Eq. 1.5, we see that this work already contains all the necessary ingredients to perform a cross section measurement for the three channels of interest. In addition to this, no CMS cross section measurement for the B^0 and B_s^0 channels at 13 TeV has been done so far (see [53] for the B^+ cross section measurement). Since the cross section does not include a ratio in its definition, the Tag and Probe method here considered becomes extremely important for the determination of the cross section's systematic uncertainties, which will be equal to the values shown in Table 7.10. The work here performed will be therefore used in a future cross section measurement.

Finally, we discuss an interesting detail. In Section 7.2 we describe the systematic uncertainty calculated from differences between data and MC. Are we not doing the same in this section? Not exactly. If the overall rate of muons that we observe gets very small, assuming that the simulation does not change, the Tag and Probe systematic uncertainty will increase a lot. For the case of the reweighting, since the MC distribution is normalized to the data distribution, this effect would be undetectable. At the same time, the reweighting considers variables that have nothing to do with the muons. Double-counting still happens, but that should not be a problem: as long as the Tag and Probe is not the dominant systematic source (usually it is not, and that is the case for this work as well), there is no issue in considering a systematic which is larger than the real one.

7.5 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 varied to check whether the obtained yield remains the same or not⁴. The difference between the nominal and the alternative yields of every variation is divided by the nominal yield in order to retrieve a relative uncertainty.

All used (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 $B^+ \rightarrow J/\psi K^+$ decay channel includes more components, and so more systematic sources are considered: both the $B^+ \rightarrow J/\psi \pi^+$ and the $B^+ \rightarrow J/\psi h X$ decays are accordingly modified, respectively, by removing the former and fitting the latter with a Gaussian rather than with an error function. We show some examples in

⁴Contrary to all the previous systematic uncertainties mentioned, this one required a computational framework that was already in place. The author is responsible for running it and tuning the parameters whenever needed.

Figure 7.6. For the $B^0 \rightarrow J/\psi K^{*0}$ channel an additional systematic is as well considered, due to the existence of the previously discussed “K π swap” component. The constraint on the fraction of swapped events in the total fit is released to assess the fit dependence on this particular component. The obtained differences in the signal yields across all the bins are found to be negligible. No additional component is considered for the B^0 channel. Regarding the B_s^0 channel, the fit is the simplest among the three and does not contain any additional PDF. Moreover, the mass window bounds are increased both on the right and on the left side of the mass spectrum by a factor of 10%, to analyse possible variations in the background parameters obtained when modifying the mass-window fitting range. An example of such mass window extension is shown in Fig. 7.6.

The systematic uncertainties mentioned in this section represent in general the dominant systematic source for the final fragmentation fraction ratios measurement. Their values can be seen in Tables 7.13, 7.14 and 7.15. Their impact is specially notorious in the bins containing lower yields, which is the case for the $1.75 < |y(B)| < 2.00$, $2.00 < |y(B)| < 2.25$ and $70 \text{ GeV} < p_T < 90 \text{ GeV}$. We suggest that the chosen alternative fitting functions do not describe the data well enough. This leads to a conservative determination of the systematic uncertainties. In addition, both the variation in the background fitting function and the change in the B meson mass window range act on the background PDF. These two sources are then highly correlated, but they are conservatively added in quadrature for the final fragmentation fraction ratio uncertainty.

7.6 Other systematic sources

The systematic uncertainty of the integrated luminosity of the datasets, mentioned in Section 3, does not affect the measurement of the fragmentation fraction ratio at first order, since it affects the denominator and numerator of the ratios in the same way. As a consequence it is not included in the final result.

The reconstruction of the tracks also represents a systematic source. A previous study of charged hadron track reconstruction was carried out at CMS. The ratio of the tracking efficiency in data and simulation is determined by exploring two and four body decays of the D^0 meson. A 2.8% and 2.3% systematic uncertainty per track is obtained for, respectively, 2015 and 2016 CMS data. For the f_s/f_d ratio, involving final states with the same number of charged hadron tracks, this systematic source is neglected. For the f_d/f_u and f_s/f_u ratios, where the involved final states differ by one track, the systematic uncertainty assigned for this source is equal to the systematic of one track (see Table 7.12).

The used MC simulations do not have infinite yields. As a consequence, the statistical uncertainty in the measurement of the MC-based efficiencies contributes with a systematic. Those statistical uncertainties are taken as systematic uncertainties in the fragmentation fraction ratio measurements. Given the adequate size of the MC samples produced for this analysis, this systematic source is sub-dominant.

Finally, the used branching fractions involve uncertainties that are obtained from Ref. [5], and represent a significant uncertainty in the final measurement. These uncertainties can be found in Table 7.11.

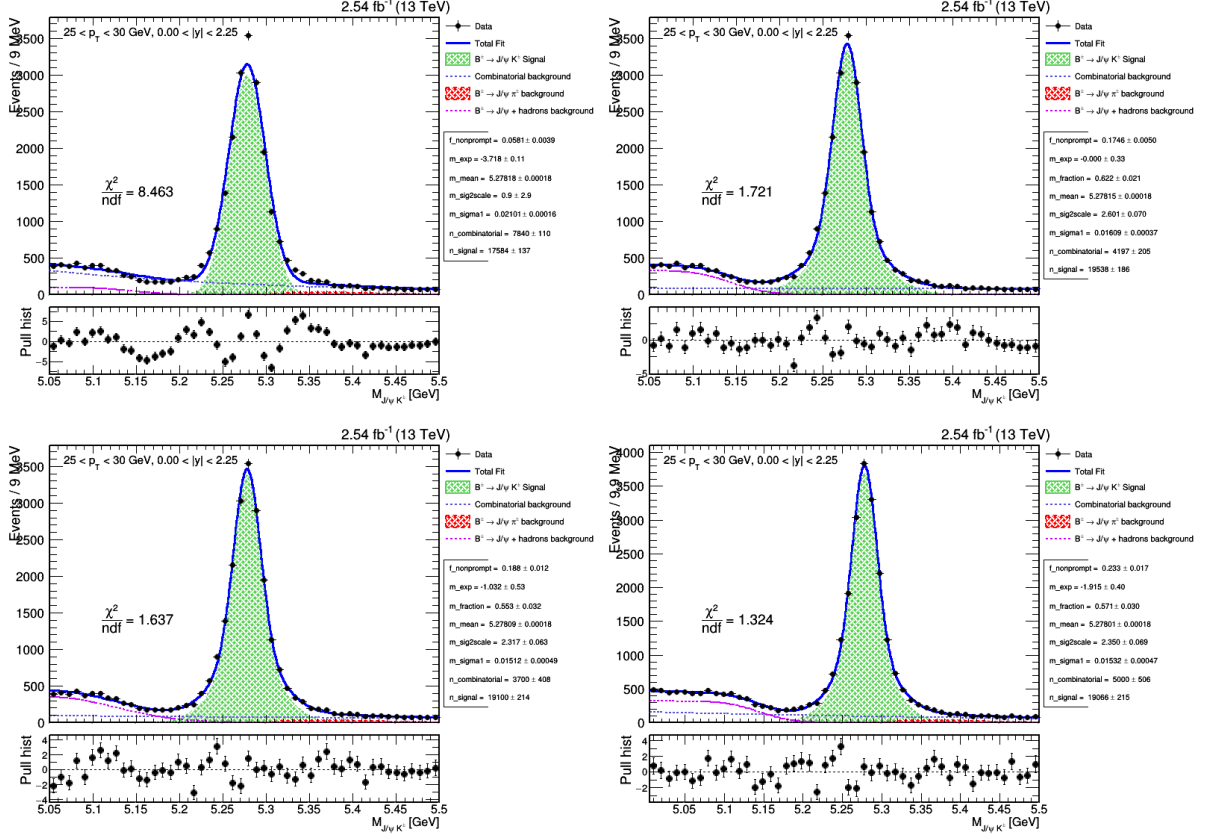


Figure 7.6: Examples of modifications done to the nominal B^+ PDF in order to obtain systematic uncertainties for the measurement of the yields. In the top left plot the signal was fitted with a Gaussian. In the top right plot no $J/\psi \pi^+$ component was included. In the bottom left plot the $J/\psi h X$ component was fitted with a Gaussian. In the bottom right plot the mass window was increased to the left by 10%. The three plots contain data from the $25 \leq p_T(B) \leq 30$ GeV bin.

Table 7.11: Branching fraction systematics for single channels and respective ratios.

Code	Channels	Syst. uncertainty	Ratios	Syst. uncertainty
A	$\mathcal{B}(B^+ \rightarrow J/\psi K^+)$	3.021%	A / B	4.938%
B	$\mathcal{B}(B^0 \rightarrow J/\psi K^{*0}) \mathcal{B}(K^{*0} \rightarrow K^+ \pi^-)$	3.906%	A / C	8.064%
C	$\mathcal{B}(B_s^0 \rightarrow J/\psi \phi) \mathcal{B}(\phi \rightarrow K^+ K^-)$	7.477%	B / C	8.436%

7.7 Summary

Many potentially significant systematic sources are studied, and a quantitative estimate for their contribution to the final result is obtained. The global systematics are depicted in Table 7.12 and the bin-dependent ones can be found in Tables 7.13, 7.14 and 7.15. The uncertainty due to the branching fractions is considered as a separate uncertainty source, since it does not depend on the measurements performed in this work (see Table 7.11). The variation of the fitting functions represent the most significant systematic source. As a future improvement, it would be possible to change the chosen alternative PDFs for them to better describe the data. Considering different alternative PDFs depending on the bin yields would also give more control on the systematics of the low yield bins. In certain bins the MC reweighting also provides non-negligible systematic uncertainties, which mostly arise due to the bad

description of the L_{xy} (and L_{xy}/σ_{xy}) variable in the B^0 channel. The low yield bins are particularly sensitive to the reweighting procedure. Would this systematic source become dominant, and alternatives would have to be discussed. Some possibilities include considering weights which take into account many variables at once (multidimensional weights), or directly correcting the MC distributions to match the data (as done by ATLAS in Ref. [34]).

Table 7.12: Global relative systematic uncertainties for the three channels of interest and for the three measured fragmentation fraction ratios. The values are expressed in %.

Channel	Tracking	Resolution	Pull	Total	Ratio	Tracking	Resolution	Pull	Total
B^+	2.8	–	0.518	2.85	f_d/f_u	2.8	0.04	0.603	2.87
B^0	5.6	0.04	0.309	5.99	f_s/f_u	2.8	5.51	1.090	6.03
B_s^0	5.6	5.51	0.959	6.92	f_s/f_d	0	5.51	1.008	5.60

Table 7.13: Bin-dependent relative systematic uncertainties for the f_s/f_d ratio. The values shown are expressed in %.

	Signal shape	Comb. bkg shape	Mass fit window	MC statistics	MC reweight	Overall systematic
Full bin	7.6	3.7	3.7	0.3	2.5	9.5

pT [GeV]	Signal shape	Comb. bkg shape	Mass fit window	MC statistics	MC reweight	Overall systematic
10 to 16	12.2	4.8	7.6	0.7	0.5	15.2
16 to 20	10.9	3.9	5.1	0.7	1.2	12.7
20 to 25	8.9	4.8	4.8	0.7	0.4	11.2
25 to 30	4.8	4.9	1.9	1.0	2.2	7.6
30 to 35	6.0	1.8	4.3	1.2	3.3	8.3
35 to 42	1.6	7.9	2.3	1.4	2.7	8.9
42 to 50	14.6	3.0	4.8	1.9	2.0	15.9
50 to 60	4.5	15.6	7.8	2.5	1.0	18.2
60 to 70	6.2	3.6	6.0	3.8	0.6	10.1
70 to 90	5.6	1.0	8.2	4.4	0.2	10.9

$ y $	Signal shape	Comb. back. shape	Mass fit window	MC statistics	MC reweight	Overall Systematic
0.00 to 0.25	8.7	5.6	7.1	0.9	0.2	12.6
0.25 to 0.50	6.5	3.1	4.3	0.9	0.4	8.4
0.50 to 0.75	10.1	7.3	7.4	0.9	1.3	14.6
0.75 to 1.00	6.2	4.9	4.0	0.9	0.0	8.9
1.00 to 1.25	6.1	6.3	2.8	0.9	2.9	9.7
1.25 to 1.50	4.6	5.4	1.9	1.0	0.7	7.4
1.50 to 1.75	0.9	13.0	4.9	1.3	0.5	14.0
1.75 to 2.00	16.9	27.9	13.5	1.9	0.9	35.4
2.00 to 2.25	18.6	16.3	23.5	2.8	2.0	34.3

Table 7.14: Bin-dependent relative systematic uncertainties for the f_s/f_u ratio. The values shown are expressed in %.

	Signal shape	Comb. back. shape	Mass fit window	MC statistics	Jpsi pi	Jpsi X	MC reweight	Overall Systematic
Full bin	9.9	3.8	3.0	0.2	0.8	0.7	3.1	11.5

pT [GeV]	Signal shape	Comb. bkg shape	Mass fit window	MC statistics	Jpsi pi	Jpsi X	MC reweight	Overall systematic
10 to 16	11.3	5.6	5.9	0.4	2.3	2.6	0.4	14.3
16 to 20	12.0	3.7	3.5	0.5	0.9	1.0	0.1	13.1
20 to 25	10.1	4.4	3.3	0.5	0.4	0.7	0.2	11.5
25 to 30	8.0	3.3	2.0	0.7	3.4	1.1	3.6	10.3
30 to 35	9.3	1.0	4.3	0.9	2.7	0.4	5.2	11.9
35 to 42	4.8	3.0	3.0	1.0	4.1	2.8	5.2	9.6
42 to 50	14.4	3.4	6.8	1.3	7.8	4.1	4.7	19.2
50 to 60	5.4	14.9	6.5	1.8	4.0	3.3	4.2	18.5
60 to 70	9.1	3.9	6.6	2.7	7.9	1.1	4.2	15.2
70 to 90	6.7	20.1	15.5	3.2	20.8	13.8	3.3	36.5

y	Signal shape	Comb. bkg shape	Mass fit window	MC statistics	Jpsi pi	Jpsi X	MC reweight	Overall systematic
0.00 to 0.25	10.0	6.5	5.4	0.6	2.2	2.6	7.3	15.4
0.25 to 0.50	8.9	3.2	3.6	0.6	1.9	2.2	4.2	11.4
0.50 to 0.75	9.8	6.8	6.2	0.6	0.0	0.5	1.1	13.5
0.75 to 1.00	6.8	3.9	2.8	0.6	2.1	0.3	0.8	8.7
1.00 to 1.25	7.7	5.7	2.8	0.6	1.5	0.4	0.1	10.1
1.25 to 1.50	7.6	2.3	2.0	0.7	5.5	0.5	2.0	10.1
1.50 to 1.75	5.3	10.6	4.8	0.8	9.1	0.8	3.7	16.2
1.75 to 2.00	9.9	20.3	7.4	1.1	2.9	4.6	8.0	25.7
2.00 to 2.25	20.3	13.5	25.2	1.6	15.1	5.2	26.5	46.8

Table 7.15: Bin-dependent relative systematic uncertainties for the f_d/f_u ratio. The values shown are expressed in %.

	Signal shape	Comb. back. shape	Mass fit window	MC statistics	Jpsi pi	Jpsi X	MC reweight	Overall systematic
Full bin	9.5	1.0	2.6	0.3	0.8	0.7	0.6	9.9

$p_T(B)$ [GeV]	Signal shape	Comb. back. shape	Mass fit window	MC statistics	Jpsi pi	Jpsi X	MC reweight	Overall systematic
10 to 16	13.4	2.9	6.4	0.6	2.3	2.6	0.1	15.5
16 to 20	11.8	2.1	4.2	0.6	0.9	1.0	1.1	12.9
20 to 25	9.3	2.7	3.8	0.6	0.4	0.7	0.1	10.4
25 to 30	7.5	3.9	1.2	0.8	3.4	1.1	1.5	9.4
30 to 35	7.5	1.6	0.6	1.0	2.7	0.4	1.9	8.4
35 to 42	5.0	8.4	2.7	1.1	4.1	2.8	2.5	11.6
42 to 50	4.8	4.3	5.4	1.5	7.8	4.1	2.8	12.6
50 to 60	6.9	7.8	6.2	2.0	4.0	3.3	3.3	13.7
60 to 70	7.4	2.3	4.5	3.0	7.9	1.1	3.6	12.9
70 to 90	4.6	20.1	13.6	3.4	20.8	13.8	3.5	35.4

$y(B)$	Signal shape	Comb. back. shape	Mass fit window	MC statistics	Jpsi pi	Jpsi X	MC reweight	Overall Systematic
0.00 to 0.25	10.3	3.4	6.1	0.7	2.2	2.6	7.1	14.8
0.25 to 0.50	8.7	3.3	4.5	0.7	1.9	2.2	4.6	11.8
0.50 to 0.75	8.3	2.9	4.4	0.7	0.0	0.5	2.5	10.2
0.75 to 1.00	6.7	2.9	2.8	0.7	2.1	0.3	0.8	8.2
1.00 to 1.25	5.8	2.7	0.5	0.7	1.5	0.4	2.9	7.3
1.25 to 1.50	7.9	4.9	0.7	0.8	5.5	0.5	2.7	11.2
1.50 to 1.75	5.4	7.5	2.3	1.1	9.1	0.8	4.2	13.9
1.75 to 2.00	19.1	20.0	13.6	1.5	2.9	4.6	8.9	32.6
2.00 to 2.25	8.2	16.6	10.1	2.3	15.1	5.2	25.0	36.5

Chapter 8

Results and Discussion

The fragmentation fraction ratios are determined according to Eqs. 1.21, 1.22 and 1.23. The final results combine all ingredients discussed throughout this work: yields, efficiencies and systematic uncertainties. The branching fractions specified in Table 1.3 represent an external input. The obtained integrated results are measured to be:

$$\begin{aligned}\frac{f_d}{f_u} &= 1.164 \pm 0.008 \text{ (stat.)} \pm 0.120 \text{ (syst.)} \pm 0.057 \text{ (BF)}, \\ \frac{f_s}{f_u} &= 0.216 \pm 0.003 \text{ (stat.)} \pm 0.029 \text{ (syst.)} \pm 0.017 \text{ (BF)}, \\ \frac{f_s}{f_d} &= 0.186 \pm 0.002 \text{ (stat.)} \pm 0.021 \text{ (syst.)} \pm 0.016 \text{ (BF)}.\end{aligned}\tag{8.1}$$

The branching fraction uncertainties are separately marked as ‘BF’ and represent a sizeably contribution to the uncertainty of the measurement. 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 uncertainty sources shown in Eq. 8.1. The most precise result is obtained with the latter, as expected.

8.1 Kinematic dependence

In order to quantitatively verify possible kinematic dependencies of the fragmentation fraction ratios, the differential results, as a function of $p_T(B)$ and $y(B)$, including all uncertainties but the global sources (*i.e.*, those that are common across all bins), are fitted with two alternative models: constant and linear. These results are displayed in Fig. 8.1. The central points are chosen according to the $p_T(B)$ and $y(B)$ distributions in each bin. The global uncertainty appears in the top left corner. The pink boxes show the statistical uncertainty only, while the full vertical error bars denote the quadratic sum of the statistical and systematic uncertainties. By using the linear model, one verifies that the slope parameters obtained for the f_s/f_u and f_s/f_d 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. This conclusion can only be draw in the acceptance region of the measurement: $10 < p_T(B) < 90$ GeV and $|y(B)| < 2.25$.

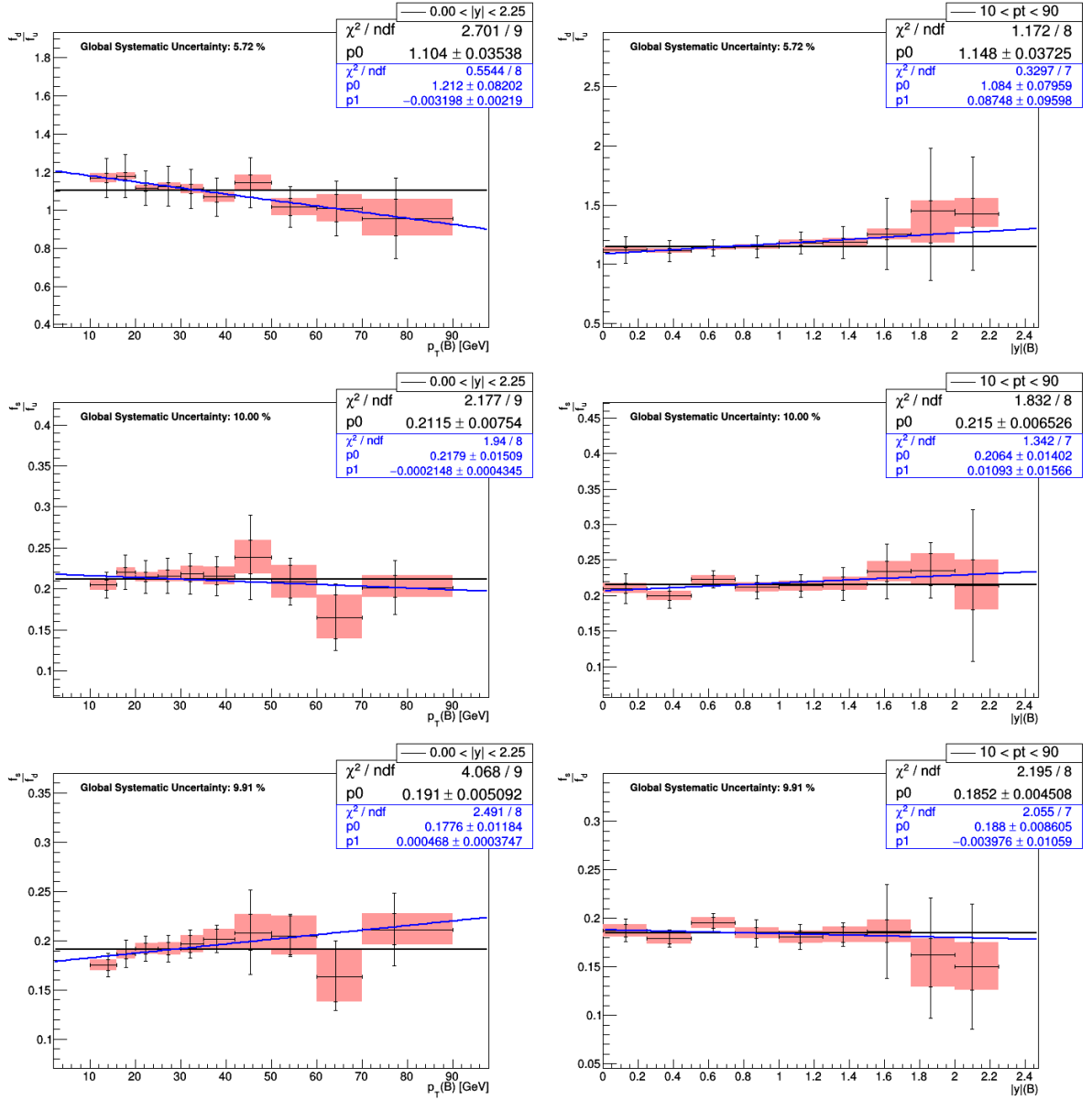


Figure 8.1: Fragmentation fraction ratios f_d/f_u (top), f_s/f_u (middle), f_s/f_d (bottom) as a function of $p_T(B)$ (left) and $|y(B)|$ (right). The pink boxes denote statistical uncertainties only. The vertical error bars show the quadratic sum of statistical and systematic uncertainties. The global systematic uncertainties are not displayed as vertical error bars but rather appear indicated in black in the top left corner. The fit results for the constant and linear fits are displayed in the top right corner and repeated in Table 8.1.

Table 8.1: Constant and linear fit parameters, for the three ratios under study, for $p_T(B)$ and $y(B)$ bins.

	Bins	Constant fit	Linear fit
f_d/f_u	$p_T(B)$	1.104 ± 0.035	$(1.212 \pm 0.082) - (0.003 \pm 0.002) p_T(B)$
	$y(B)$	1.148 ± 0.037	$(1.084 \pm 0.080) + (0.087 \pm 0.096) y(B)$
f_s/f_u	$p_T(B)$	0.212 ± 0.008	$(0.2179 \pm 0.0151) - (0.0002 \pm 0.0004) p_T(B)$
	$y(B)$	0.215 ± 0.007	$(0.206 \pm 0.014) + (0.011 \pm 0.016) y(B)$
f_s/f_d	$p_T(B)$	0.191 ± 0.005	$(0.1776 \pm 0.0118) + (0.0005 \pm 0.0004) p_T(B)$
	$y(B)$	0.185 ± 0.005	$(0.188 \pm 0.009) - (0.004 \pm 0.011) y(B)$

8.2 Comparison with other results

Both ATLAS and LHCb have recently measured the f_s/f_d ratio [34, 35], such that a comparison is possible. We must nevertheless note that the LHCb experiment covers a different pseudorapidity region than CMS and ATLAS, and that it does not have access to high p_T mesons. A comparison of the kinematic dependence must be therefore done with care. On the other hand, ATLAS covers roughly the same kinematic region of CMS, such that both results are expected to be similar. Still, the measurement performed in this work is the only one which used a 13 TeV centre of mass collision energy; the two before mentioned results were obtained with a centre of mass collision energy of 7 TeV.

Starting from the LHCb results, the measured global ratio turns out to be somewhat higher than the one obtained in this work: $f_s/f_d = 0.256 \pm 0.020$ [35], within a 2.70σ distance. The previous value was obtained by averaging three different values: one of them was measured using inclusive semileptonic B decays involving D mesons of the type $B \rightarrow D l \nu_l X$, where 'l', ' ν ' and 'X' denote, respectively, a lepton, its respective neutrino and anything else [54], while the other two used hadronic decays of the type $B \rightarrow D h$, where 'h' represents a hadron ($h = K, \pi$) [35, 55]. Regarding the differential measurements, the LHCb Collaboration reported a $p_T(B)$ dependence with a 3σ significance, while no $\eta(B)$ dependence was observed (see Figure 8.2). The measurement performed by the ATLAS Collaboration used the

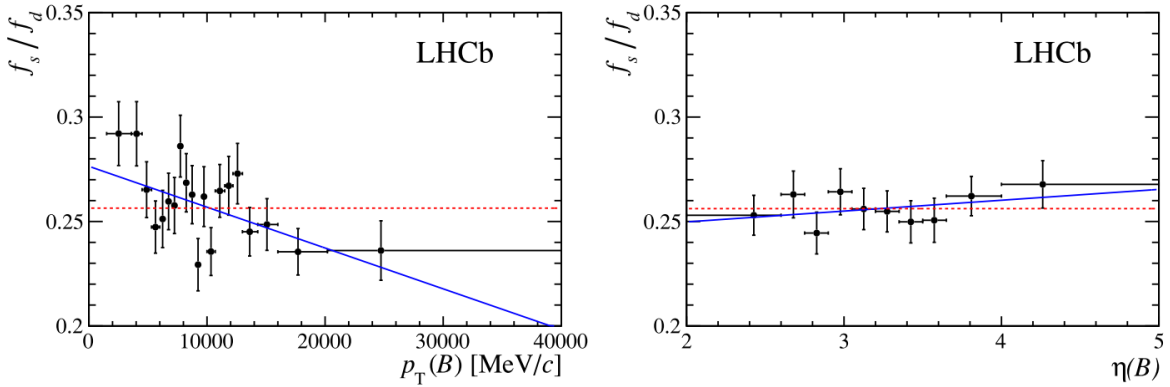


Figure 8.2: Ratio of fragmentation fractions f_s/f_d as functions of $p_T(B)$ (left) and $\eta(B)$ (right). The errors on the data points are the statistical and uncorrelated systematic uncertainties added in quadrature. The solid line is the result of a linear fit, and the dashed line corresponds to the fit for the no-dependence hypothesis. The average value of $p_T(B)$ or $\eta(B)$ is determined for each bin and used as the centre of the bin. The horizontal error bars indicate the bin size. Note that the scale is zero suppressed [35].

decay channels used in this work. Its global result was the following: $f_s/f_d = 0.240 \pm 0.004$ (stat.) ± 0.010 (syst.) ± 0.017 (theor.), where the theoretical uncertainty arises since the $\mathcal{B}(B_s^0 \rightarrow J/\psi \phi)/\mathcal{B}(B^0 \rightarrow J/\psi K^{*0})$ ratio was inserted using a perturbative QCD prediction. The result lies at a 2.09σ distance from the result here obtained. The kinematic dependence was also studied, but in this case no $p_T(B)$ or $\eta(B)$ dependence was found (see Figure 1.10). The latter conclusion agrees with the result here obtained, and confers robustness to the idea that no dependence exist in the $10 < p_T(B) < 90$ GeV range.

8.3 Future Prospects

The LHC will increase its centre of mass collision energy to 14 TeV and its integrated luminosity to 300 fb⁻¹ by 2023 (see Fig. 8.3). Starting in 2025, following the third maintenance long shutdown, the High Luminosity LHC is expected to deliver 250 fb⁻¹ per year during 10 years of operation, at the same centre of mass collision energy. On the one hand, rare decays will be seen much often. On the other hand, the CMS detector must be refurbished in order to deal with both the larger radiation exposure and high pile-up (the number of primary vertices per bunch crossing), which will be around 140. The identification of B meson vertexes will be indeed more difficult, since the chance to get pile-up vertexes close to the vertexes where the B mesons under study are produced is going to increase.

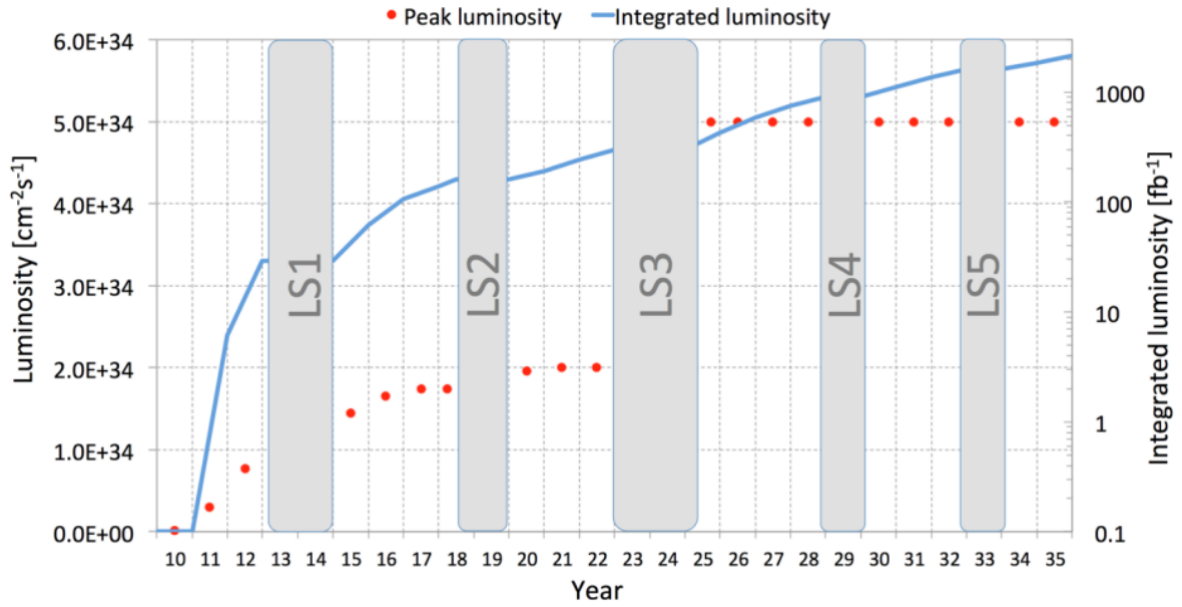


Figure 8.3: Projected LHC performance until 2035, showing preliminary dates for long shutdowns and projected luminosities. The integrated luminosity is displayed according to a logarithmic scale [56].

The tracker will be upgraded in order to sustain an integrated luminosity of up to 3000 fb⁻¹ and to enhance its functionality. Its read-out channel density will increase. The tracker's ability to distinguish two close-by muons or tracks¹ will also be improved. Furthermore, the pixel tracker acceptance will be extended up to about $\eta = 4$, and its passive volume will be reduced to a minimum to reduce multiple scattering, nuclear interactions, *bremssstrahlung* and photon conversions. Finally, the tracking system will be compatible with the new trigger system envisioned by the CMS Collaboration; the L1 trigger will use tracking information at hardware level. All these upgrades translate into the replacement of the whole tracker.

The muon system will handle the high radiation and the trigger changes as well. Candidate tracks with p_T of at least 2 GeV in the inner silicon tracker will be matched to muons at L1. This kind of matching is currently performed at the HLT only. Three general upgrades will be put into action: the existing muon detectors will be refurbished to grant better performance and extended longevity, additional muon

¹For the tracker a muon or a different track is exactly the same thing. Muons are later identified in the muon chambers (see Chapter 2).

detectors will be placed in the CMS forward region ($1.6 < |\eta| < 2.4$) to increase the redundancy currently granted by the RPC as well as to augment the trigger and reconstruction capabilities, and the muon coverage will be extended up to $|\eta| = 3$ to keep up with the tracker acceptance extension. This latter goal will be fulfilled thanks to the addition of a small but precise muon detector².

The inclusion of tracking information at hardware level is not only beneficial but also crucial. The predicted 140 pile-up would require a L1 trigger acceptance rate fifteen times larger than the current one. If one considers a potentially higher pile-up of 200, the L1 would require almost 4000 kHz, which is technically infeasible. By adding a tracker trigger to the system, the acceptance rate drops dramatically, even considering extra uncertainties: a 750 kHz rate is expected to maintain with some margin an acceptance of 200 pile-up vertexes. The trigger latency will increase from $4\text{ }\mu\text{s}$ up to $12.5\text{ }\mu\text{s}$ for the new L1 tracker trigger to have more time to reconstruct tracks and to match them with information coming from the muon systems and calorimeters. These upgrades will allow a better identification of charged leptons and the improvement of the determination of the p_T of the muons, as well as the determination of isolation of leptons and photons, jets and leptons vertexes and transverse missing energy carried by L1 tracks. In addition, the already existing parts of the trigger will be upgraded as well. With these new tools it will be possible for the Global Trigger to increase its selectivity. In turn, the HLT will need a processing power 24 times larger than before for a pile-up of 140, and it will output data at a rate of up to 7.5 kHz, while maintaining a rejection factor of 100 to L1 events [56].

The performance of the analysis of the $B^0 \rightarrow \mu^+ \mu^-$ decays is expected to improve, which could lead to the observation of the $B^0 \rightarrow \mu^+ \mu^-$ decay channel with an excess of 5σ . A prediction for the

²Interestingly, the addition of this detector will reduce the background where missing energy calculations are wrong due to high p_T muons which have $|\eta| > 2.4$. The missing transverse energy variable is not used in the work here reported, but it is often used to infer the existence of non detectable particles: this is done in SUSY searches, for example.

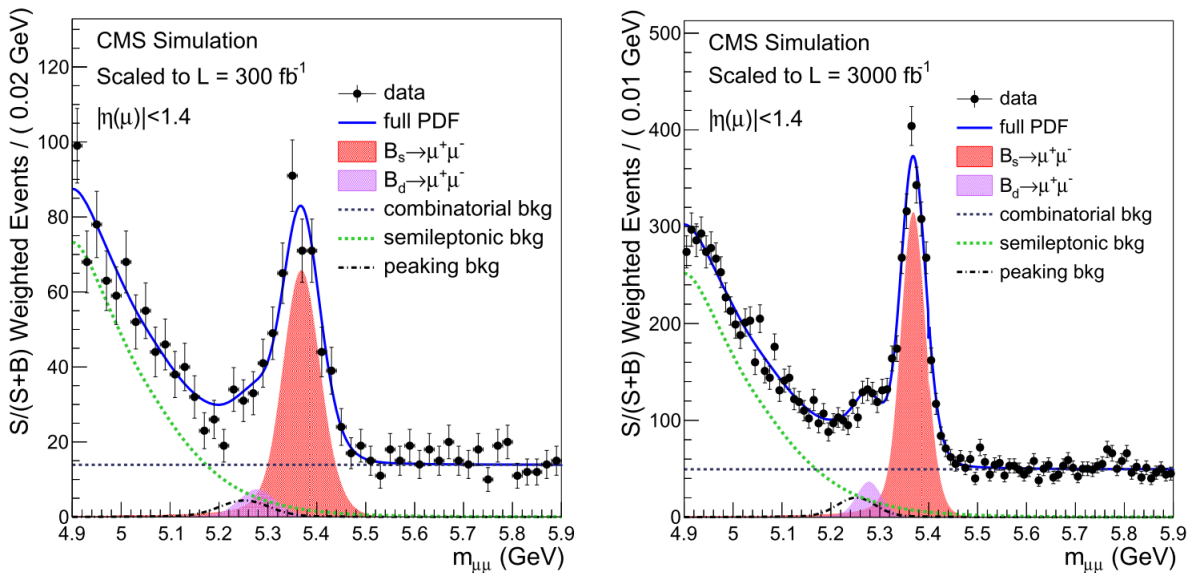


Figure 8.4: CMS simulations for the expected invariant mass distributions before (left) and during (right) the High Luminosity LHC running. The plot on the left considers an integrated luminosity of 300 fb^{-1} , while the plot on the right reaches 3000 fb^{-1} . Both simulations only take the barrel region ($|\eta| < 1.4$) into account. The $B^0 \rightarrow \mu^+ \mu^-$ peak will probably be disentangled from the $B_s^0 \rightarrow \mu^+ \mu^-$ peak, making its experimental observation possible [56].

B_s^0 observed invariant mass distributions is shown in Fig. 8.4, for two different scenarios. In the first one, the simulation refers to the conditions prior to long shutdown 3, with an integrated luminosity of 300 fb^{-1} , while in the second one the integrated luminosity reaches 3000 fb^{-1} , in order to reproduce the expected conditions at the High Luminosity LHC. The latter scenario takes into account the planned detector's upgrades and also conservatively assumes a loss of efficiency of 35% for the signal and 30% for the background with respect to Run-I (2010-2012). A striking difference can be seen: the B^0 peak can only be observed in the second scenario, with an expected sensitivity of 6.8σ thanks to the improvement in the mass resolution from 42 to 28 MeV for both the B_s^0 and B^0 peaks [56]. The observed distribution for the 300 fb^{-1} scenario is quite similar to the one shown in Fig. 1.6. The LHC and CMS upgrades will therefore surely test the SM in a more stringent way.

Bibliography

- [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. doi: 10.1016/j.physletb.2012.08.021.
- [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. doi: 10.1016/j.physletb.2012.08.020.
- [3] I. J. R. Aitchison and A. J. G. Hey. *Gauge Theories in Particle Physics: A Practical Introduction*, volume 1. CRC Press, 2003. ISBN 9781466512993.
- [4] F. Halzen and A. D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, 1984.
- [5] C. Patrignani *et al.* (Particle Data Group), Chin. Phys. C, 40, 100001 (2016).
- [6] D. Griffiths. *Introduction to Elementary Particles*. Wiley-VCH Verlag GmbH & Co. KGaA, 2nd, revised edition, 2008.
- [7] C. Giunti and C. W. Kim. *Fundamentals of Neutrino Physics and Astrophysics*. Oxford University Press, 2007. ISBN 9780198508717. doi: 10.1093/acprof:oso/9780198508717.001.0001.
- [8] E. Roulet. Beyond the standard model. Lectures given at the CERN-CLAF School, Itacuruca, Brasil, May 2001, last revised 2007. arXiv:hep-ph/0112348v2.
- [9] N. Leonardo. *Matter antimatter fluctuations, search, discovery and analysis of B_s flavor oscillations*. LAP, 2011. ISBN 978-3843376938. URL http://home.fnal.gov/~leonardo/oscillations_book.htm.
- [10] P. Koppenburg, Z. Dolezal, and M. Smizanska. Rare decays of b hadrons. *Scholarpedia*, 11(6): 32643, 2016. ISSN 1941-6016. doi: 10.4249/scholarpedia.32643. URL <http://arxiv.org/abs/1606.00999>.
- [11] G. C. Branco, L. Lavoura, and J. P. Silva. *CP Violation*. Oxford University Press, 1999. ISBN 9780198503996.
- [12] T. Feldmann. Short and long-distance QCD effects in B meson decays, 2003. Invited talk given at Hadrons 03, Aschaffenburg, Germany.

- [13] I. J. R. Aitchison and A. J. G. Hey. *Gauge Theories in Particle Physics: A Practical Introduction*, volume 2. CRC Press, 2003. ISBN 9781466513075.
- [14] T. E. Browder and K. Honscheid. B mesons. *Progress in Particle and Nuclear Physics*, 35(C): 81–219, 1995. ISSN 01466410. doi: 10.1016/0146-6410(95)00042-H.
- [15] J. Lewis and R. van Kooten. Review of Physics results from the Tevatron: Heavy Flavour Physics. *Int. J. Mod. Phys.*, A30(06):1541003, 2015. doi: 10.1142/S0217751X15410031.
- [16] The CMS Collaboration. Inclusive b-hadron production cross section with muons in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 03:090, 2011. doi: 10.1007/JHEP03(2011)090.
- [17] The CMS Collaboration. Measurement of the Strange B Meson Production Cross Section with $J/\psi \phi$ Decays in pp Collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev.*, D84:052008, 2011. doi: 10.1103/PhysRevD.84.052008.
- [18] J. P. Silva. Phenomenological aspects of CP violation. In *Central European School in Particle Physics Prague, Czech Republic*, September 14-24, 2004.
- [19] K. De Bruyn, R. Fleischer, R. Kneijens, P. Koppenburg, M. Merk, and N. Tuning. Branching Ratio Measurements of B_s Decays. *Phys. Rev.*, D86:014027, 2012. doi: 10.1103/PhysRevD.86.014027.
- [20] The CMS Collaboration. Measurement of the CP-violating weak phase ϕ_s and the decay width difference $\Delta\Gamma_s$ using the $B_s^0 \rightarrow J/\psi\phi(1020)$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV. *Phys. Lett.*, B757:97–120, 2016. doi: 10.1016/j.physletb.2016.03.046.
- [21] The CMS and LHCb Collaborations. Observation of the rare $B_s^0 \rightarrow \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data. *Nature*, 522:68–72, 2015. doi: 10.1038/nature14474.
- [22] The LHCb Collaboration. Measurement of the effective $B_s^0 \rightarrow K^+K^-$ lifetime. *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics*, 707(3-4):349–356, 2012. ISSN 03702693. doi: 10.1016/j.physletb.2011.12.058.
- [23] The LHCb Collaboration. Measurement of the $B_s^0 \rightarrow \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \rightarrow \mu^+\mu^-$ decays. *Phys. Rev. Lett.*, 118:191801, May 2017. doi: {10.1103/PhysRevLett.118.191801}. URL <https://link.aps.org/doi/10.1103/PhysRevLett.118.191801>.
- [24] The LHCb Collaboration. Test of Lepton Universality Using $B \rightarrow K^+l^+l^-$ Decays. *Physical Review Letters*, 113(15):151601, oct 2014. ISSN 0031-9007. doi: 10.1103/PhysRevLett.113.151601. URL <https://link.aps.org/doi/10.1103/PhysRevLett.113.151601>.
- [25] C. Bobeth, M. Gorbahn, T. Hermann, M. Misiak, E. Stamou, and M. Steinhauser. $B_{s,d} \rightarrow l^+l^-$ in the standard model with reduced theoretical uncertainty. *Physical Review Letters*, 112(10):3–7, 2014. ISSN 10797114. doi: 10.1103/PhysRevLett.112.101801.

- [26] M. Beneke, C. Bobeth, and R. Szafron. Enhanced Electromagnetic Corrections to the Rare Decay. *Physical Review Letters*, 120(1):011801, Jan 2018. ISSN 0031-9007. doi: 10.1103/PhysRevLett.120.011801. URL <http://arxiv.org/abs/1708.09152><https://link.aps.org/doi/10.1103/PhysRevLett.120.011801><http://dx.doi.org/10.1103/PhysRevLett.120.011801>.
- [27] M. Hussain, M. Usman, M. A. Paracha, and M. J. Aslam. Constraints on two-Higgs-doublet model parameters in light of rare B decays. *Physical Review D*, 95(7):075009, apr 2017. ISSN 2470-0010. doi: 10.1103/PhysRevD.95.075009. URL <http://link.aps.org/doi/10.1103/PhysRevD.95.075009><http://arxiv.org/abs/1703.10845><http://dx.doi.org/10.1103/PhysRevD.95.075009>.
- [28] R. Gauld. Understanding forward B hadron production. 2017.
- [29] R. Fleischer. Probing new physics with $B_s^0 \rightarrow \mu^+ \mu^-$: Status and perspectives. *Int. J. Mod. Phys.*, A29:1444004, 2014. doi: 10.1142/S0217751X14440047.
- [30] The ATLAS Collaboration. Study of the rare decays of B_s^0 and B^0 into muon pairs from data collected during the LHC Run 1 with the ATLAS detector. *European Physical Journal C*, 76(9), 2016. ISSN 14346052. doi: 10.1140/epjc/s10052-016-4338-8.
- [31] D. M. Straub. Overview of constraints on new physics in rare B decays. page 6, 2012. URL <http://arxiv.org/abs/1205.6094>.
- [32] The LHCb Collaboration. Observation of $B_c \rightarrow D^0 K^+$ Decays. *Physical Review Letters*, 118(11):111803, mar 2017. ISSN 0031-9007. doi: 10.1103/PhysRevLett.118.111803. URL <http://arxiv.org/abs/1701.01856><http://dx.doi.org/10.1103/PhysRevLett.118.111803><https://link.aps.org/doi/10.1103/PhysRevLett.118.111803>.
- [33] 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. ISSN 1434-6044. doi: 10.1140/epjc/s10052-017-5058-4. URL <http://link.springer.com/10.1140/epjc/s10052-017-5058-4>.
- [34] 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. doi: 10.1103/PhysRevLett.115.262001.
- [35] 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. ISSN 10298479. doi: 10.1007/JHEP04(2013)001.
- [36] The CMS Collaboration. Measurement of the $B_s^0 \rightarrow \mu^+ \mu^-$ Branching Fraction and Search for $B^0 \rightarrow \mu^+ \mu^-$ with the CMS Experiment. *Phys.Rev.Lett.*, 111(10):101804, 2013. doi: 10.1103/PhysRevLett.111.101804.
- [37] The CMS Collaboration. CMS Luminosity - Public Results, . URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>. Accessed on April 6th 2018.

- [38] CERN. LHC Machine Outreach. URL <http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/>. Accessed on April 6th 2018.
- [39] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3: S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.
- [40] A. De Angelis and M. J. M. Pimenta. *Introduction to Particle and Astroparticle Physics*. Undergraduate Lecture Notes in Physics. Springer Milan, Milano, 2015. ISBN 978-88-470-2687-2. doi: 10.1007/978-88-470-2688-9. URL <http://link.springer.com/10.1007/978-88-470-2688-9>.
- [41] The CMS Collaboration. The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC. *Journal of Instrumentation*, 8(11):P11002–P11002, Nov 2013. ISSN 1748-0221. doi: 10.1088/1748-0221/8/11/P11002. URL <http://arxiv.org/abs/1306.6905><http://dx.doi.org/10.1088/1748-0221/8/11/P11002><http://stacks.iop.org/1748-0221/8/i=11/a=P11002?key=crossref.104bb466b78b67be3cffb0850293beb4>.
- [42] The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation*, 9(10), 2014. ISSN 17480221. doi: 10.1088/1748-0221/9/10/P10009.
- [43] The CMS Collaboration. *The CMS Tracker Technical Design Report*. April 1998.
- [44] The CMS Collaboration. *The Muon Project Technical Design Report*. December 1997.
- [45] The CMS Collaboration. The CMS Trigger System. *Journal of Instrumentation*, 12, 2016. doi: 10.1088/1748-0221/12/01/P01020.
- [46] The CMS Collaboration. WorkbookDataflowTrigger - From Collision to Primary Datasets, . URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkbookDataflowTrigger>. Accessed on April 6th 2018.
- [47] The CMS Collaboration. WorkbookCollisionsDataAnalysis - Analysis of the collision data, . URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkbookCollisionsDataAnalysis>. Accessed on April 6th 2018.
- [48] The CMS Collaboration. WorkbookDataFormats - Data Formats and Data Tiers, . URL <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkbookDataFormats>. Accessed on April 7th 2018.
- [49] M. Stoye. *Calibration and Alignment of the CMS Silicon Tracking Detector*. PhD thesis, 2007. URL <http://bib-pubdb1.desy.de/record/288878>.
- [50] M. C. Morais. *Notas de Apoio da Disciplina de Probabilidades e Estatística*. Milano, February 2012.
- [51] T. J. Adye. *A study of J/ψ production at the LEP $e^+ e^-$ collider and the implementation of the DELPHI slow controls system*. PhD thesis, 1998. URL <http://weblib.cern.ch/abstract?CERN-THESIS-2006-049>.

- [52] R. Barlow. Extended maximum likelihood. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 297(3):496–506, dec 1990. ISSN 01689002. doi: 10.1016/0168-9002(90)91334-8. URL <http://linkinghub.elsevier.com/retrieve/pii/0168900290913348>.
- [53] The CMS Collaboration. Measurement of differential inclusive B^+ hadron cross section in pp collisions at $\sqrt{s} = 13$ TeV. *Physics Letters B*, 771:435–456, August 2017. doi: 10.1016/j.physletb.2017.05.074.
- [54] The LHCb Collaboration. Measurement of b hadron production fractions in 7 TeV pp collisions. *Physics Letters B*, 718(3):902–909, nov 2011. ISSN 03702693. doi: 10.1103/PhysRevD.85.032008. URL <http://linkinghub.elsevier.com/retrieve/pii/S0370269312012038><http://arxiv.org/abs/1111.2357><http://dx.doi.org/10.1103/PhysRevD.85.032008>.
- [55] The LHC Collaboration. Determination of f_s/f_d for 7 TeV pp collisions and a measurement of the branching fraction of the decay $B_d \rightarrow D^- K^+$. *Nuclear Physics*, (3):7, 2011. ISSN 00319007. doi: 10.1103/PhysRevLett.107.211801. URL <http://arxiv.org/abs/1106.4435>.
- [56] The CMS Collaboration. Technical Proposal for the Phase-II Upgrade of the CMS Detector. Technical Report CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02, Geneva, Jun 2015. URL <https://cds.cern.ch/record/2020886>.

Appendix A

Data and MC Comparison

A.1 B^+ channel

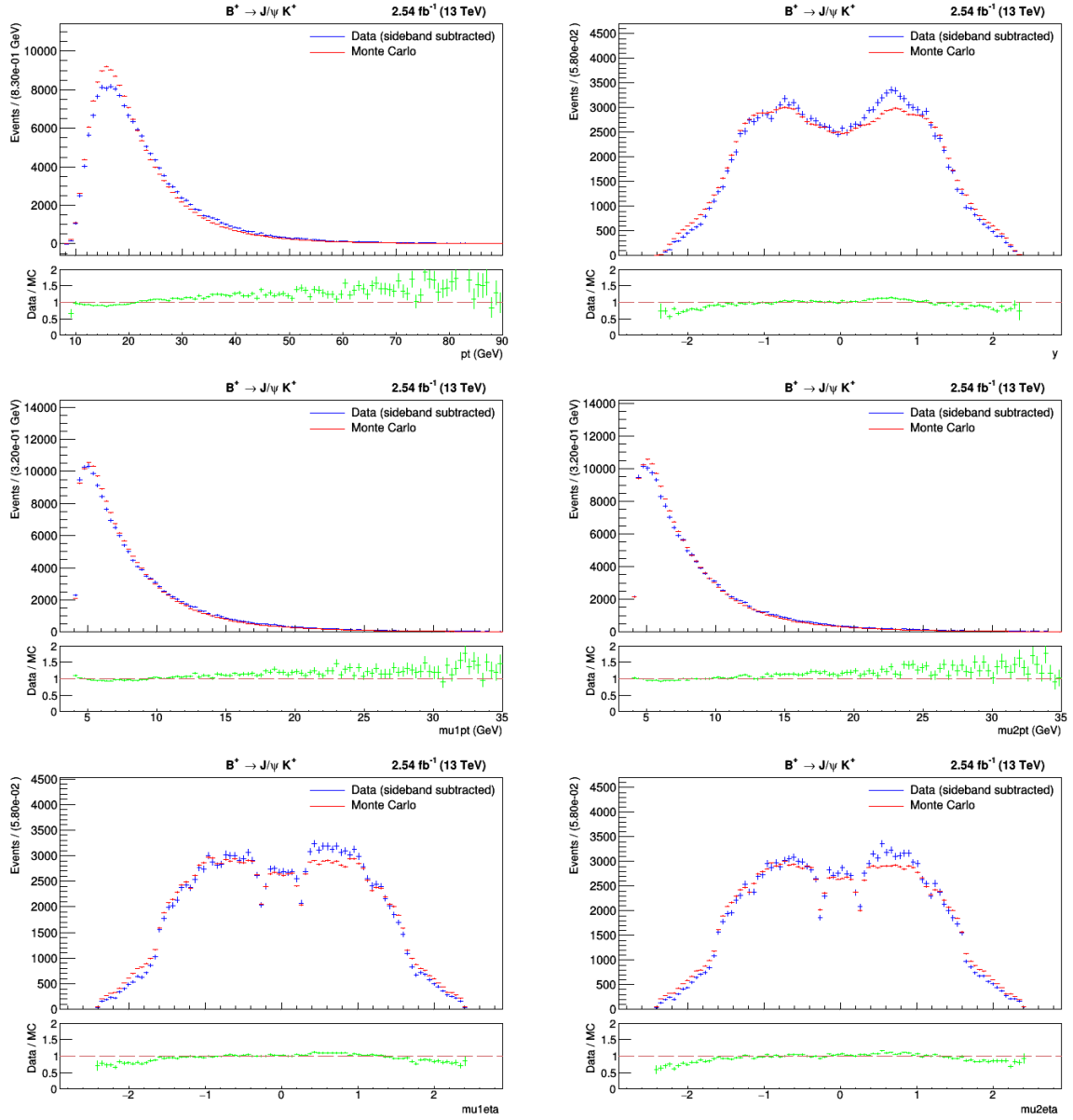


Figure A.1: Data / MC comparison for the $B^+ \rightarrow J/\psi K^+$ channel. The distributions correspond, from top left to bottom right, to the following variables: $p_T(B)$, $y(B)$, $p_T(\mu_1)$, $p_T(\mu_2)$, $\eta(\mu_1)$ and $\eta(\mu_2)$.

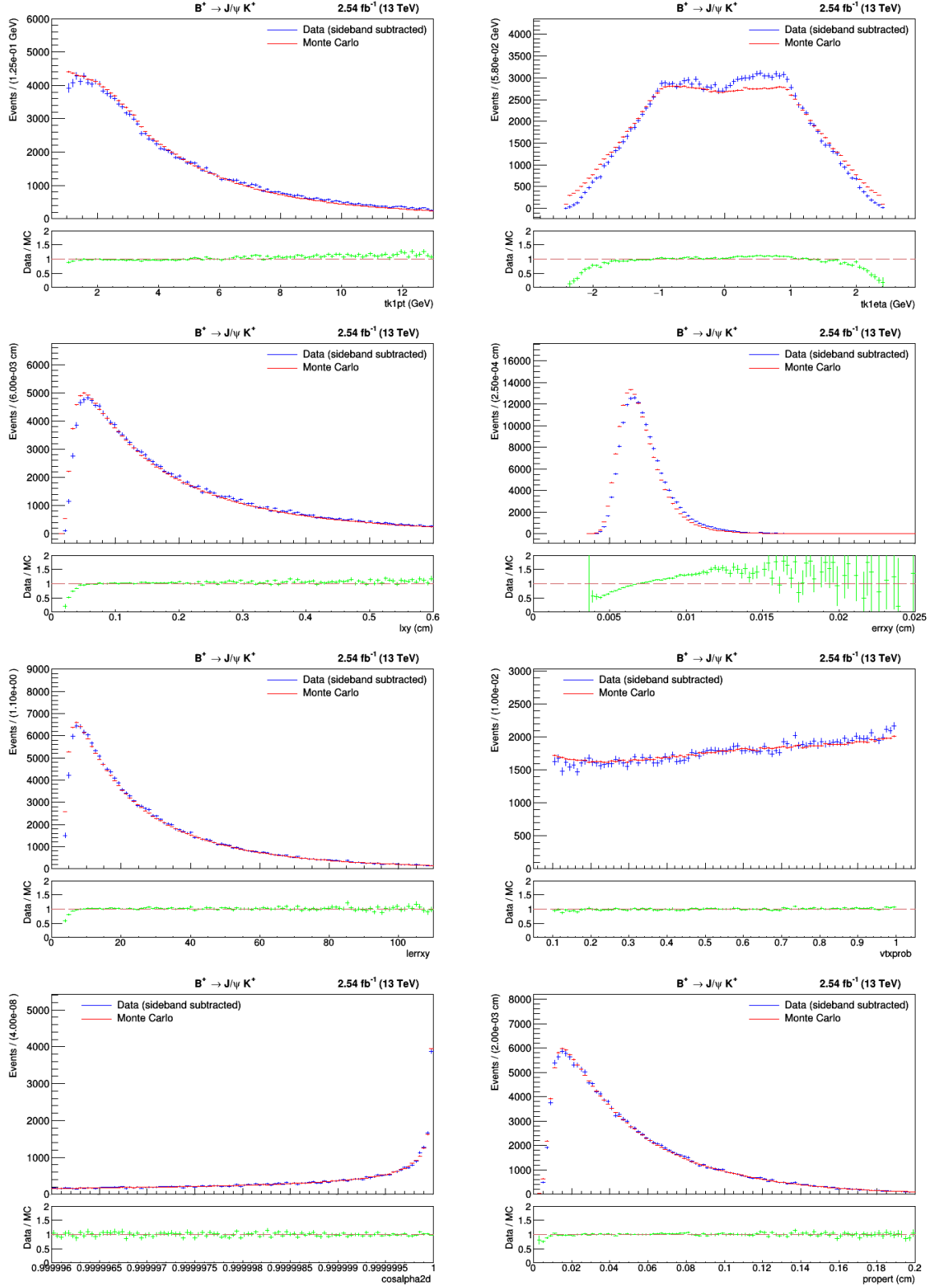


Figure A.2: Data / MC comparison for the $B^+ \rightarrow J/\psi K^+$ channel. The distributions correspond, from top left to bottom right, to the following variables: p_T (track $_1$), η (track $_1$), L_{xy} , σ_{xy} , L_{xy}/σ_{xy} , the vertex probability, $\cos(\alpha_{2D})$ and the proper time.

A.2 B^0 channel

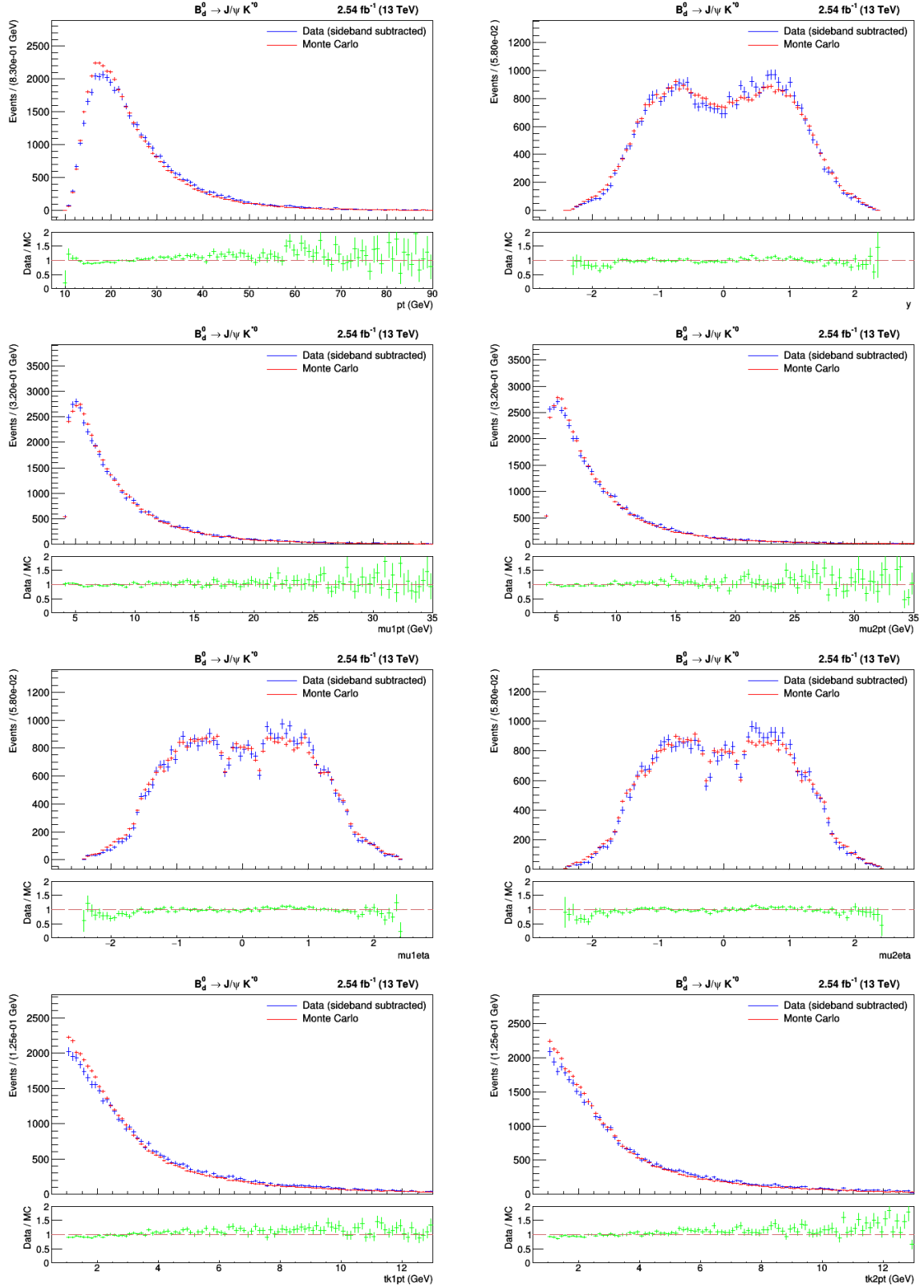


Figure A.3: Data / MC comparison for the $B^0 \rightarrow J/\psi K^{*0}$ channel. The distributions correspond, from top left to bottom right, to the following variables: $p_T(B)$, $y(B)$, $p_T(\mu_1)$, $p_T(\mu_2)$, $\eta(\mu_1)$, $\eta(\mu_2)$, $p_T(\text{track}_1)$ and $p_T(\text{track}_2)$.

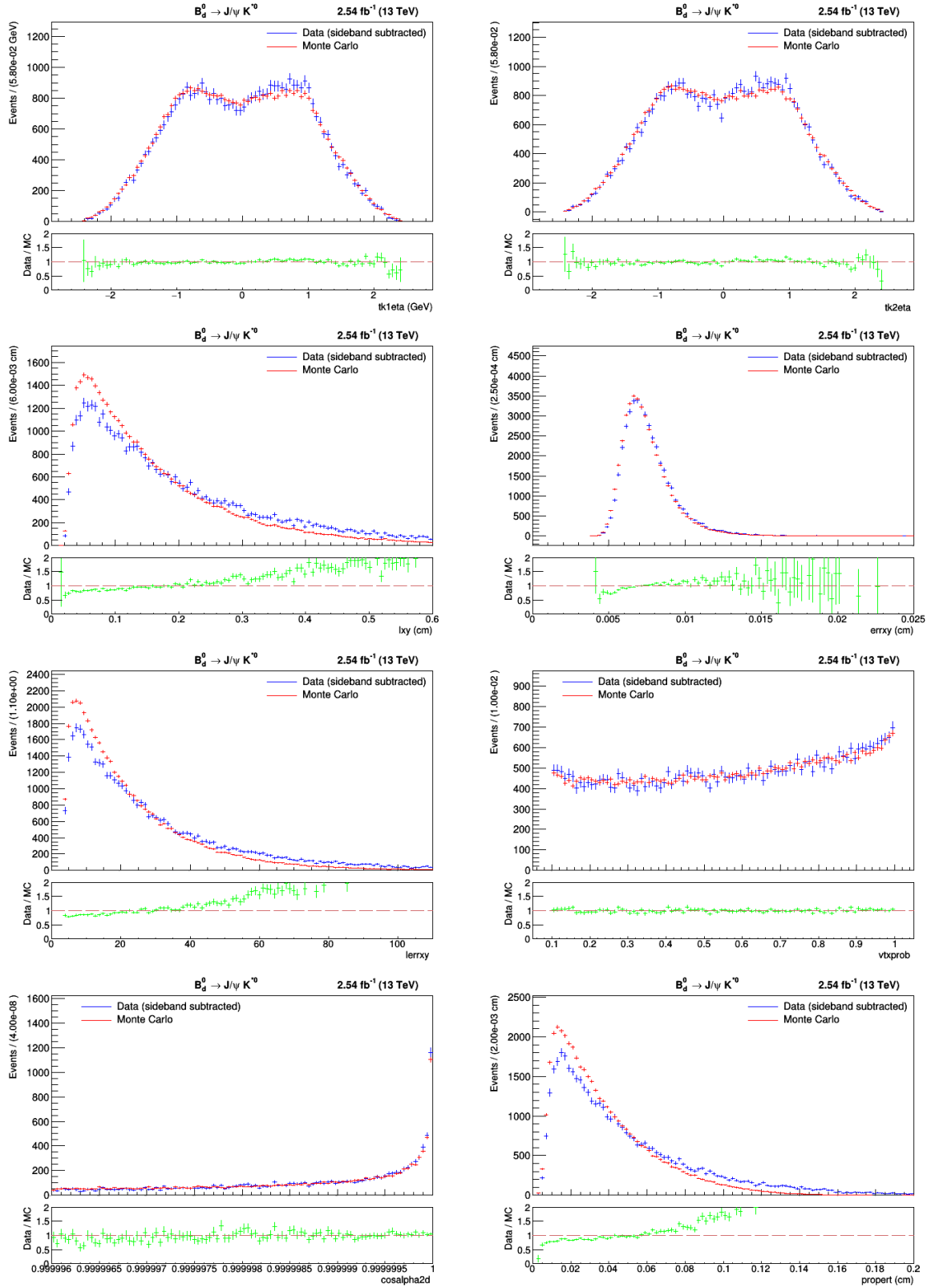


Figure A.4: Data / MC comparison for the $B^0 \rightarrow J/\psi K^0$ channel. The distributions correspond, from top left to bottom right, to the following variables: η (track₁), η (track₂), L_{xy} , σ_{xy} , L_{xy}/σ_{xy} , the vertex probability, $\cos(\alpha_{2D})$ and the proper time.

A.3 B_s^0 channel

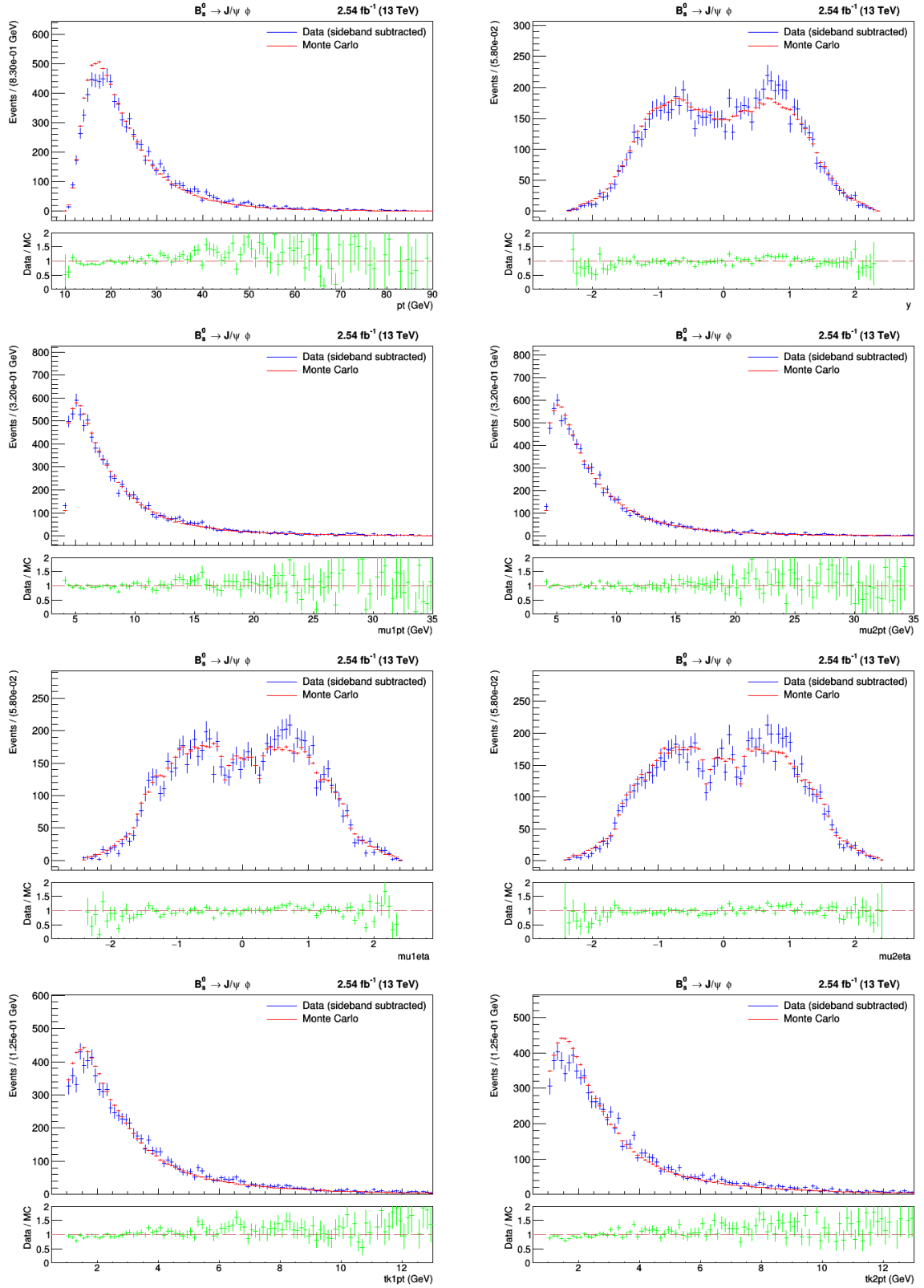


Figure A.5: Data / MC comparison for the $B_s^0 \rightarrow J/\psi \phi$ channel. The distributions correspond, from top left to bottom right, to the following variables: $p_T(B)$, $y(B)$, $p_T(\mu_1)$, $p_T(\mu_2)$, $\eta(\mu_1)$, $\eta(\mu_2)$, $p_T(\text{track}_1)$ and $p_T(\text{track}_2)$.

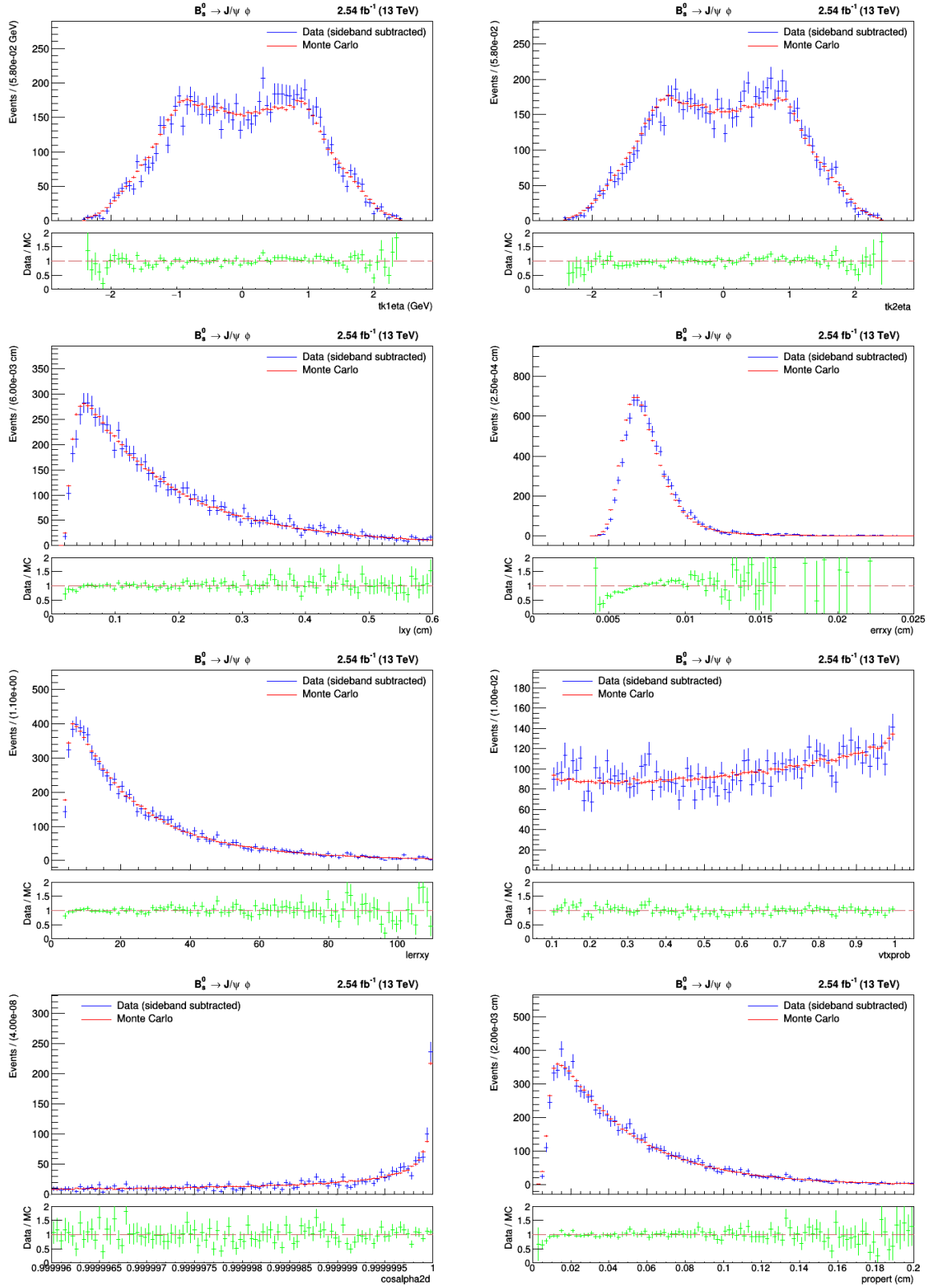


Figure A.6: Data / MC comparison for the $B_s^0 \rightarrow J/\psi \phi$ channel. The distributions correspond, from top left to bottom right, to the following variables: η (track₁), η (track₂), L_{xy} , σ_{xy} , L_{xy}/σ_{xy} , the vertex probability, $\cos(\alpha_{2D})$ and the proper time.

Appendix B

MC Reweighting

B.1 $B^0 \rightarrow J/\psi K^{*0}$ channel

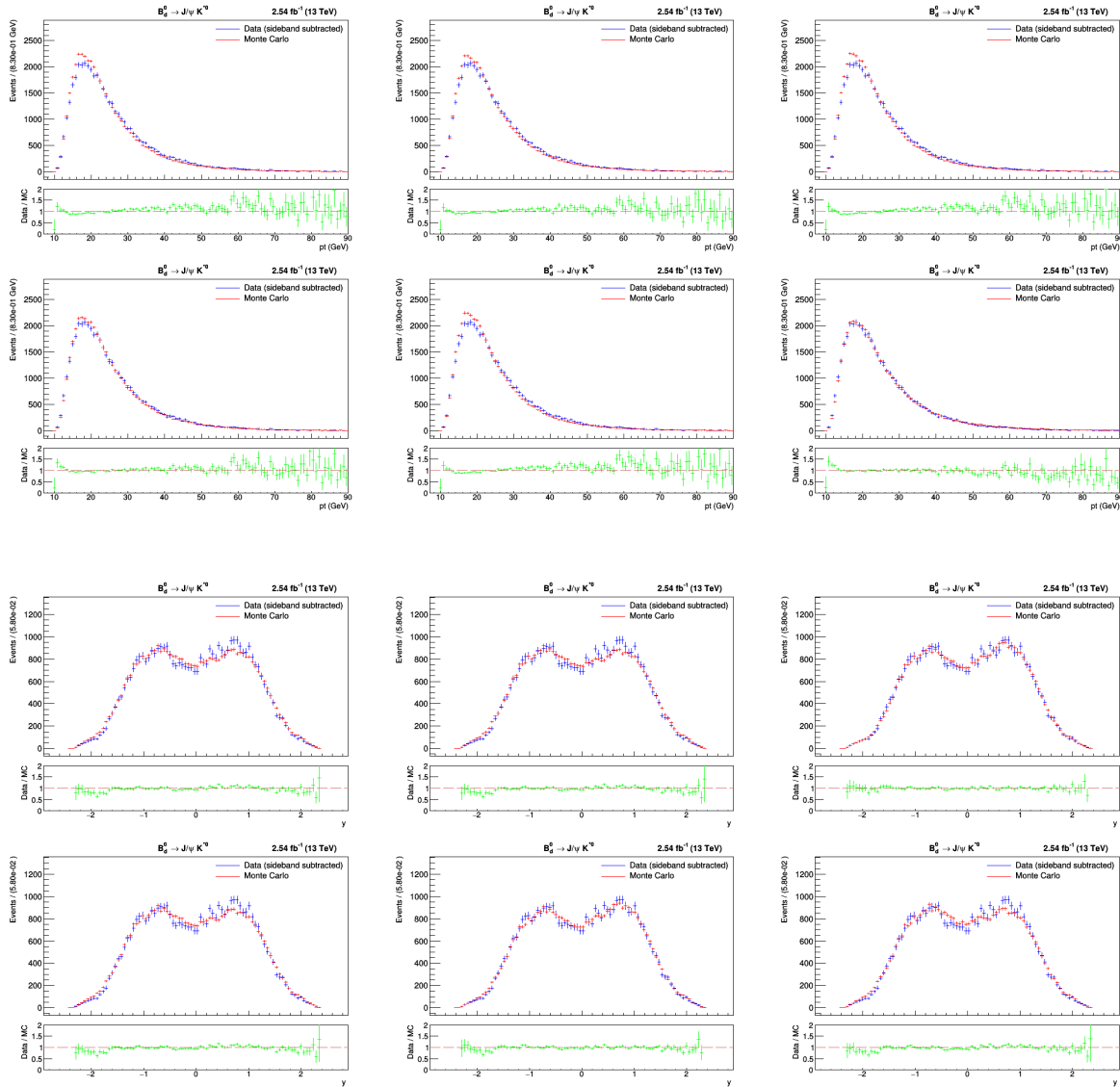


Figure B.1: MC reweighting applied to the $p_T(B)$ (top group) and $y(B)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.

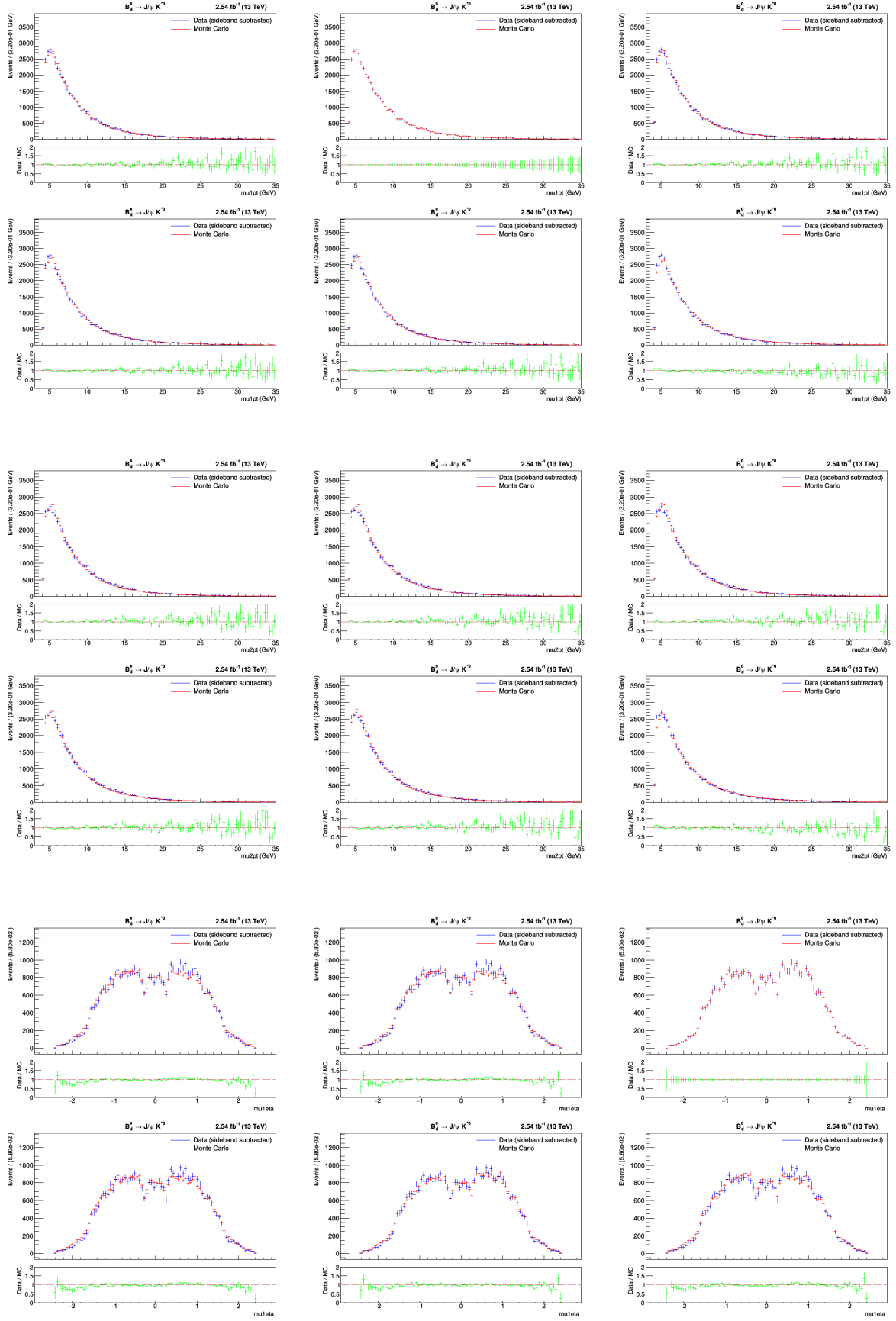


Figure B.2: MC reweighting applied to the $p_T(\mu_1)$ (top group), $p_T(\mu_2)$ (middle group) and $\eta(\mu_1)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.

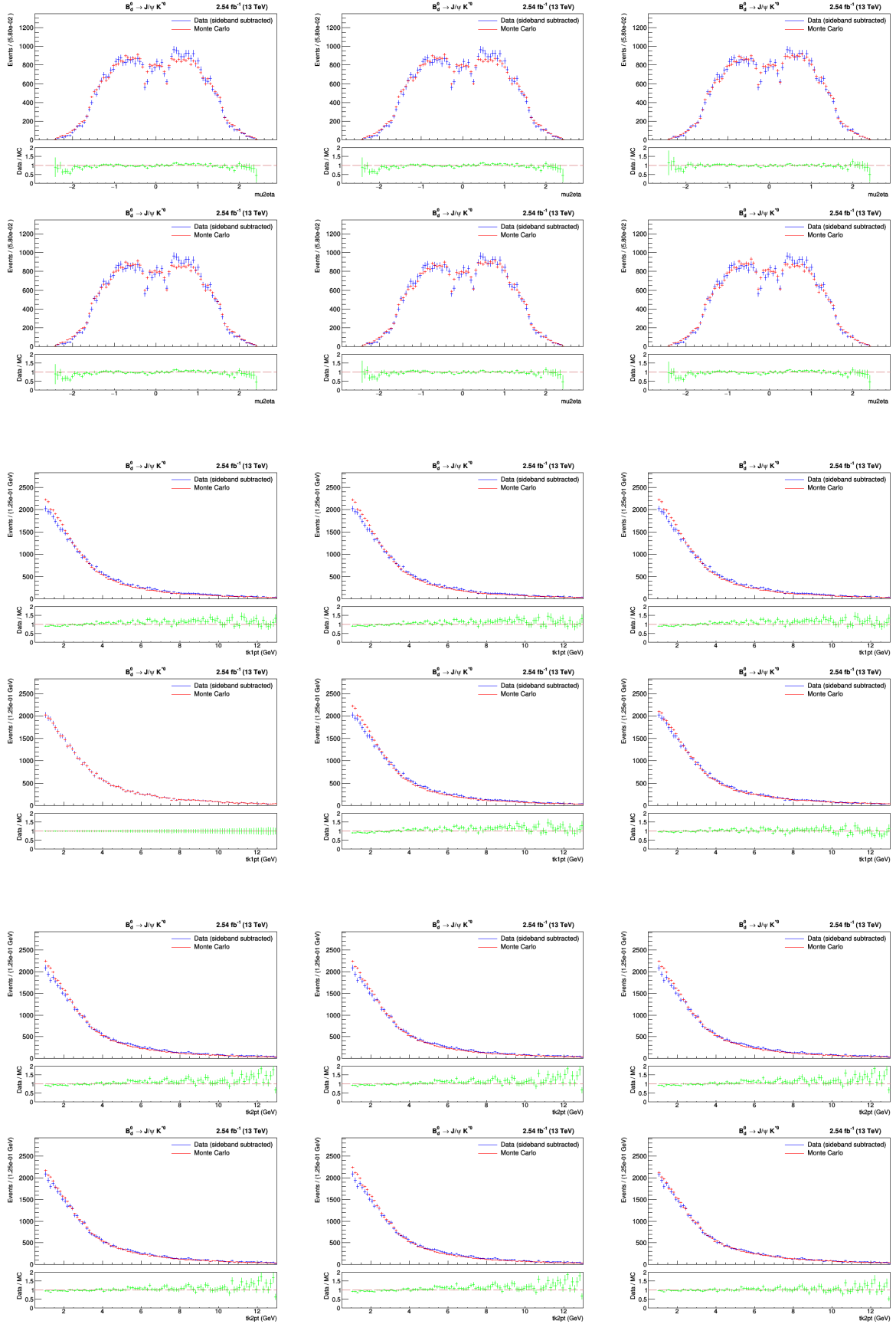


Figure B.3: MC reweighting applied to the $p_T(\mu_2)$ (top group), $p_T(\text{track}_1)$ (middle group) and $p_T(\text{track}_2)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.

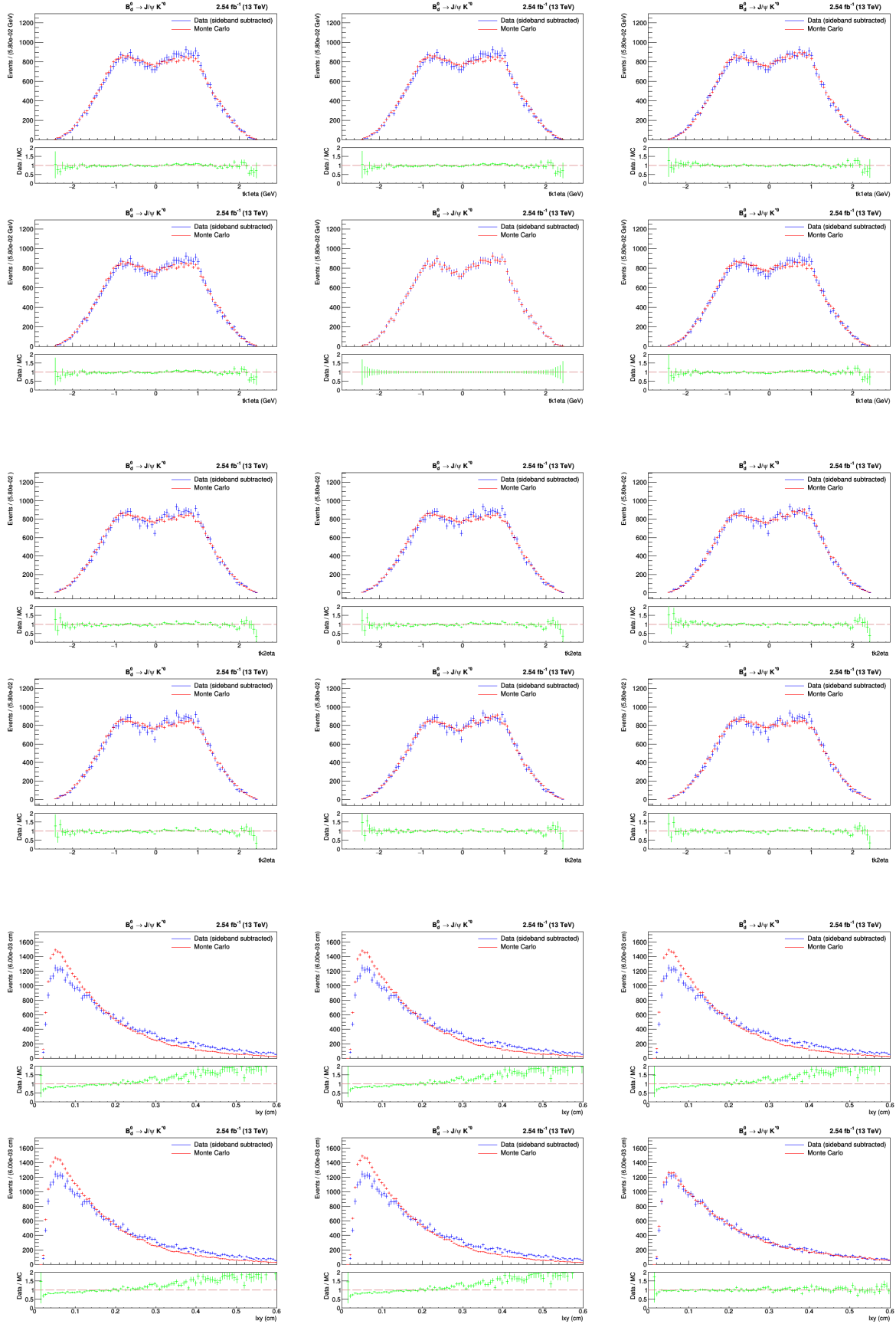


Figure B.4: MC reweighting applied to the $p_T(\mu_2)$ (top group), $p_T(\text{track}_1)$ (middle group) and $p_T(\text{track}_2)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.

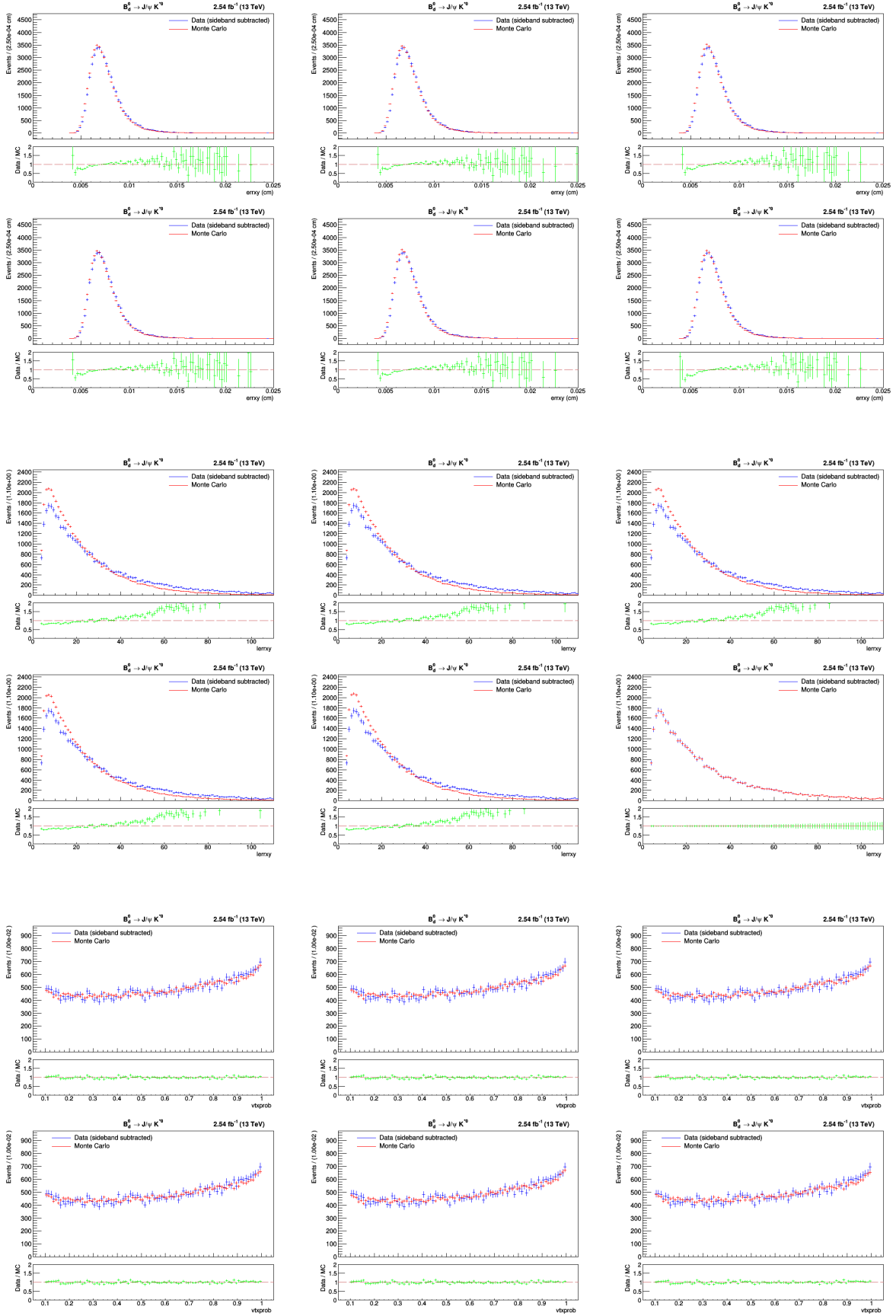


Figure B.5: MC reweighting applied to the σ_{xy} (top group), L_{xy}/σ_{xy} (middle group) and vertex probability (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.

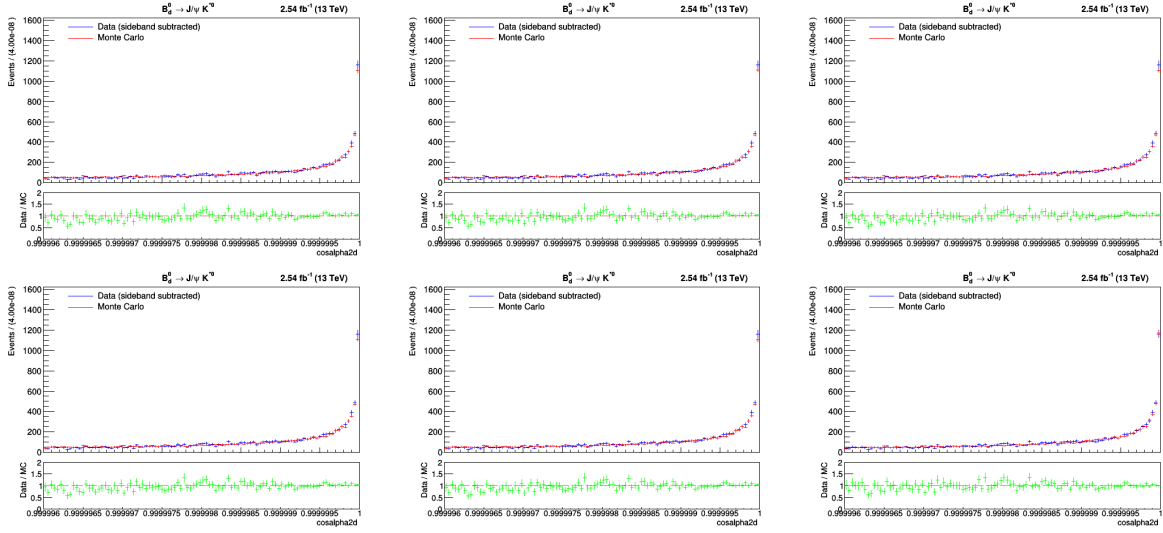


Figure B.6: MC reweighting applied to the $\cos(\alpha_{2D})$ variable. The top left plot shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.

B.2 $B_s^0 \rightarrow J/\psi \phi$ channel

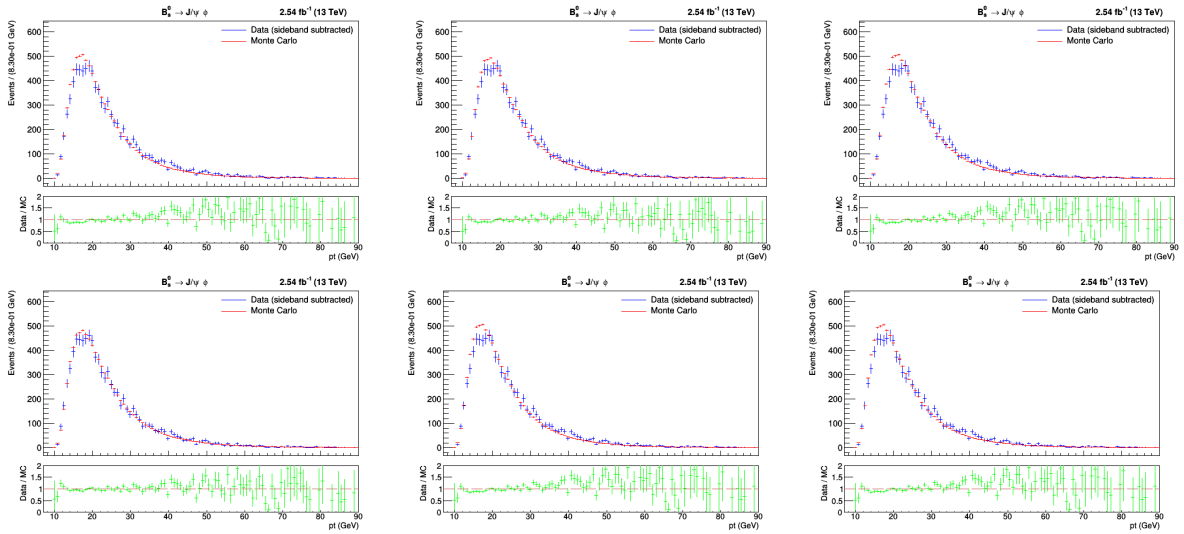


Figure B.7: MC reweighting applied to the $p_T(B)$ variable. The top left plot shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.

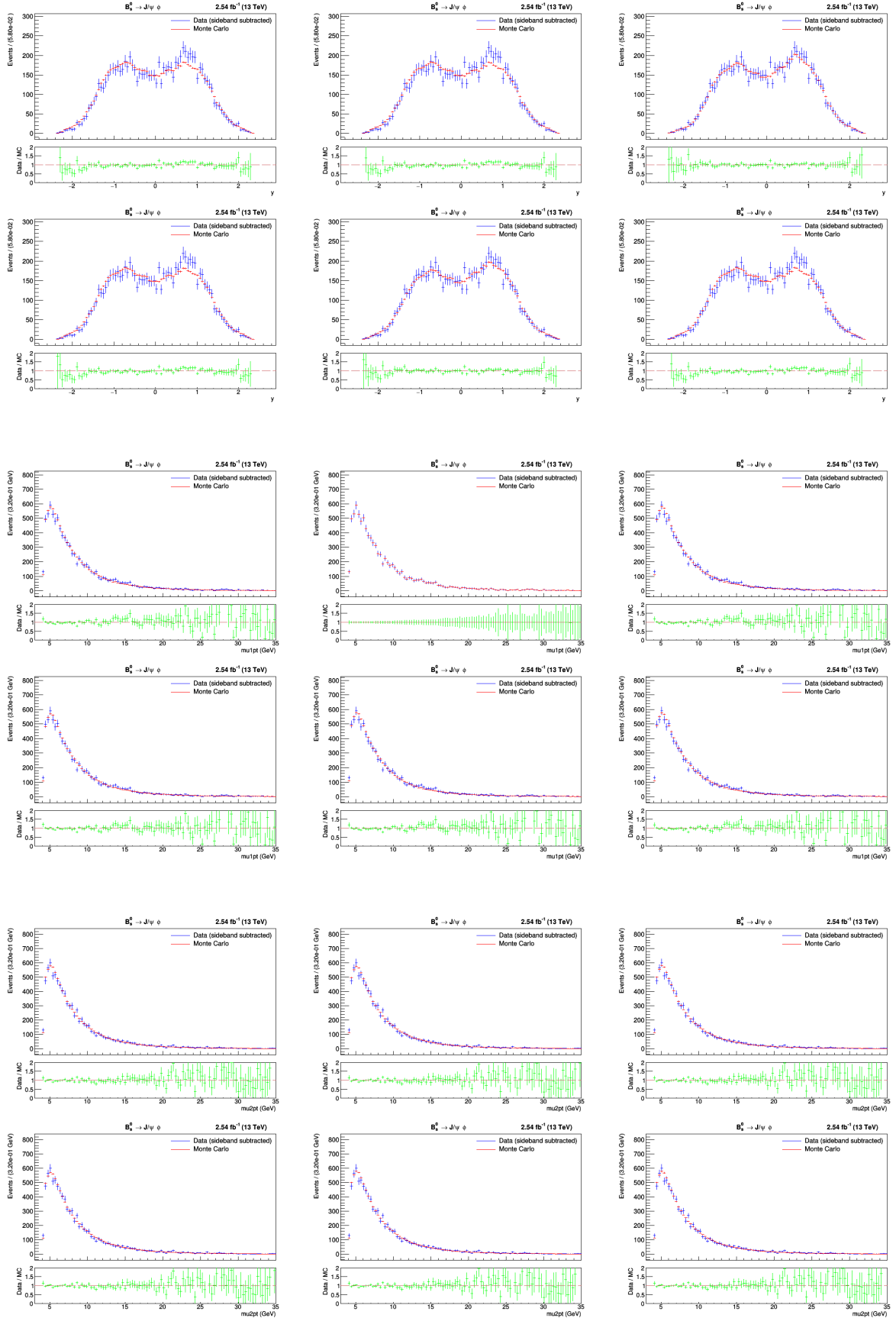


Figure B.8: MC reweighting applied to the γ (B) (top group), $p_T(\mu_1)$ (middle group) and $p_T(\mu_2)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.

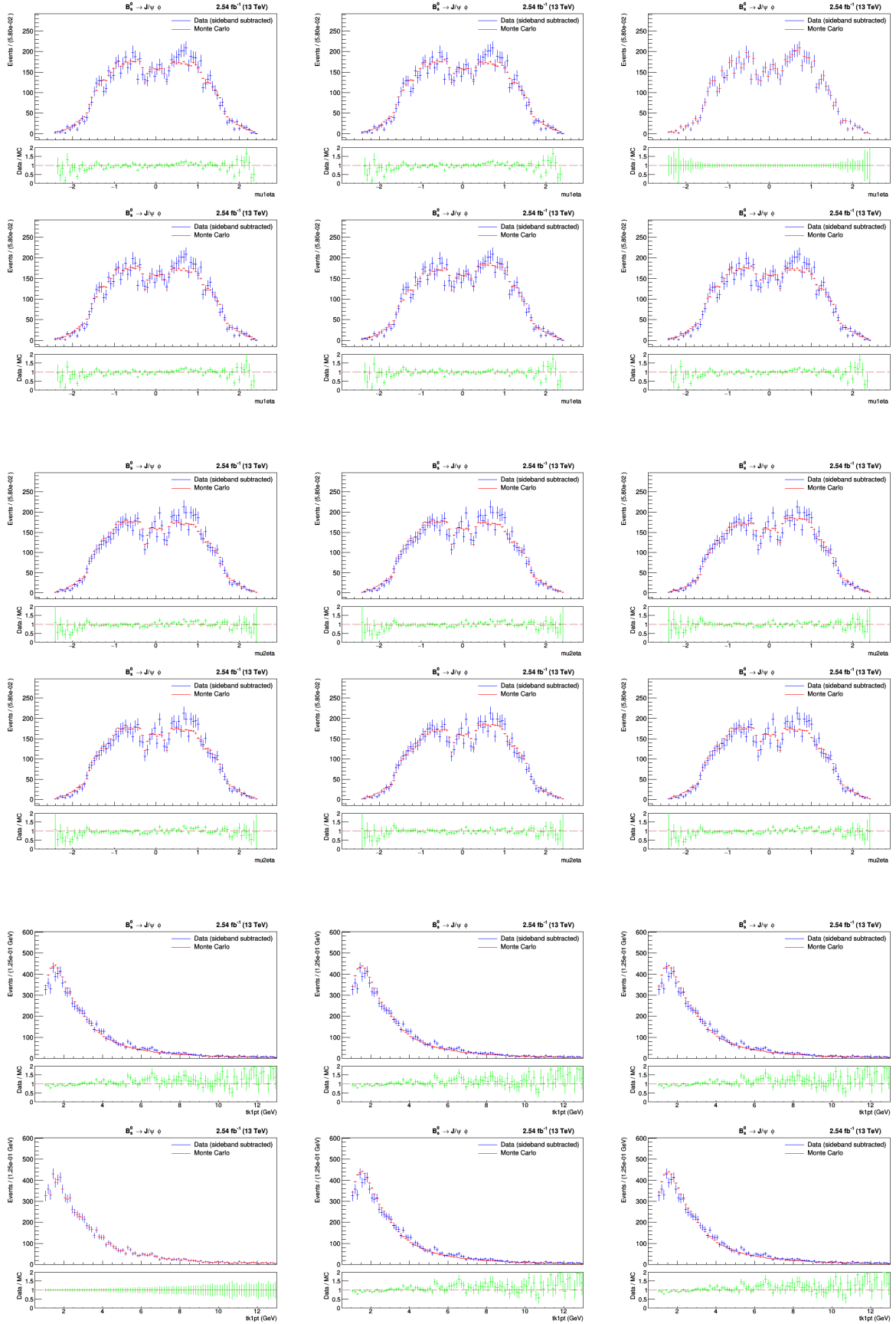


Figure B.9: MC reweighting applied to the $\eta(\mu_1)$ (top group), $\eta(\mu_1)$ (middle group) and $\eta(\text{track}_1)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.

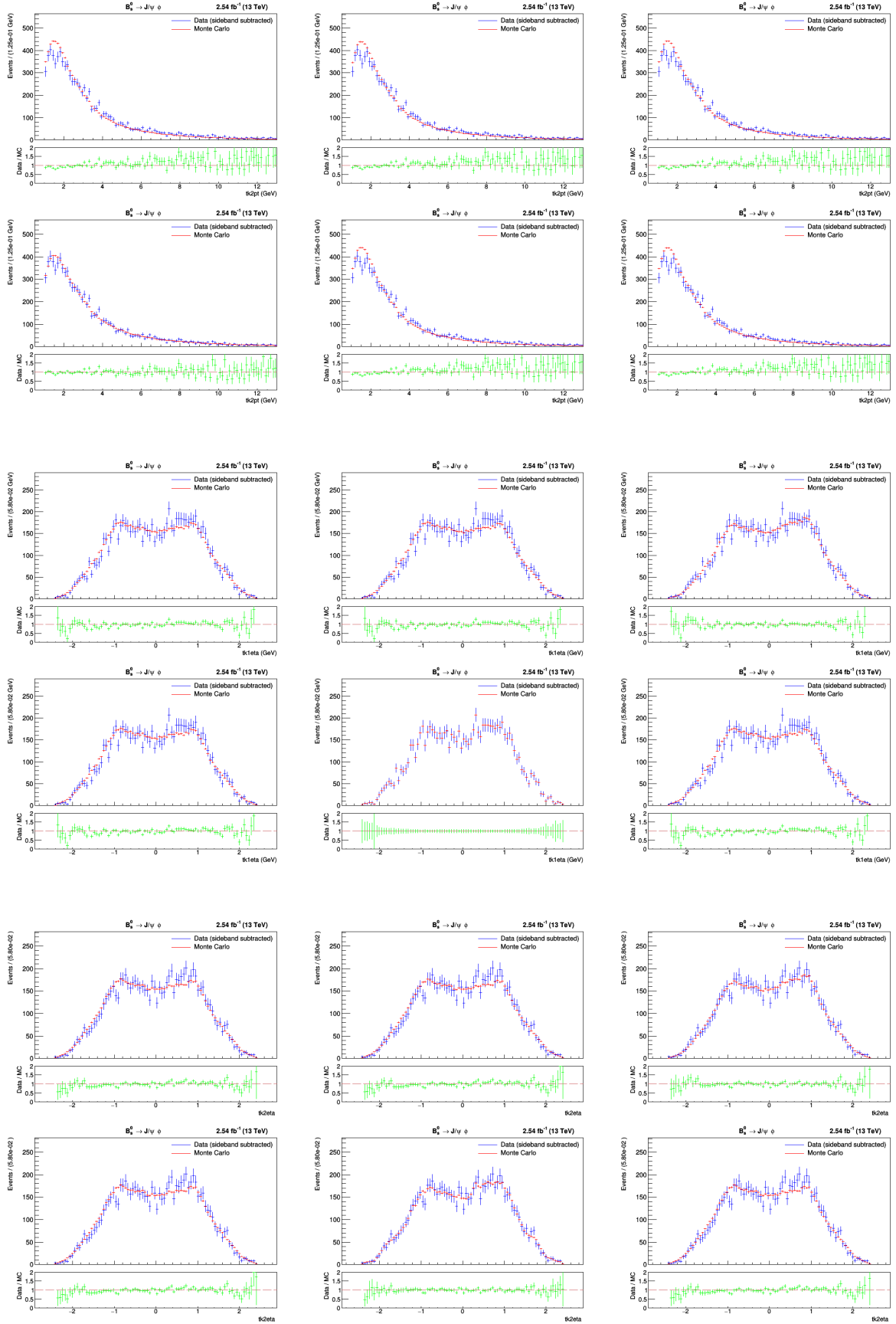


Figure B.10: MC reweighting applied to the $p_T(\text{track}_2)$ (top group), $\eta(\text{track}_1)$ (middle group) and $\eta(\text{track}_2)$ (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.

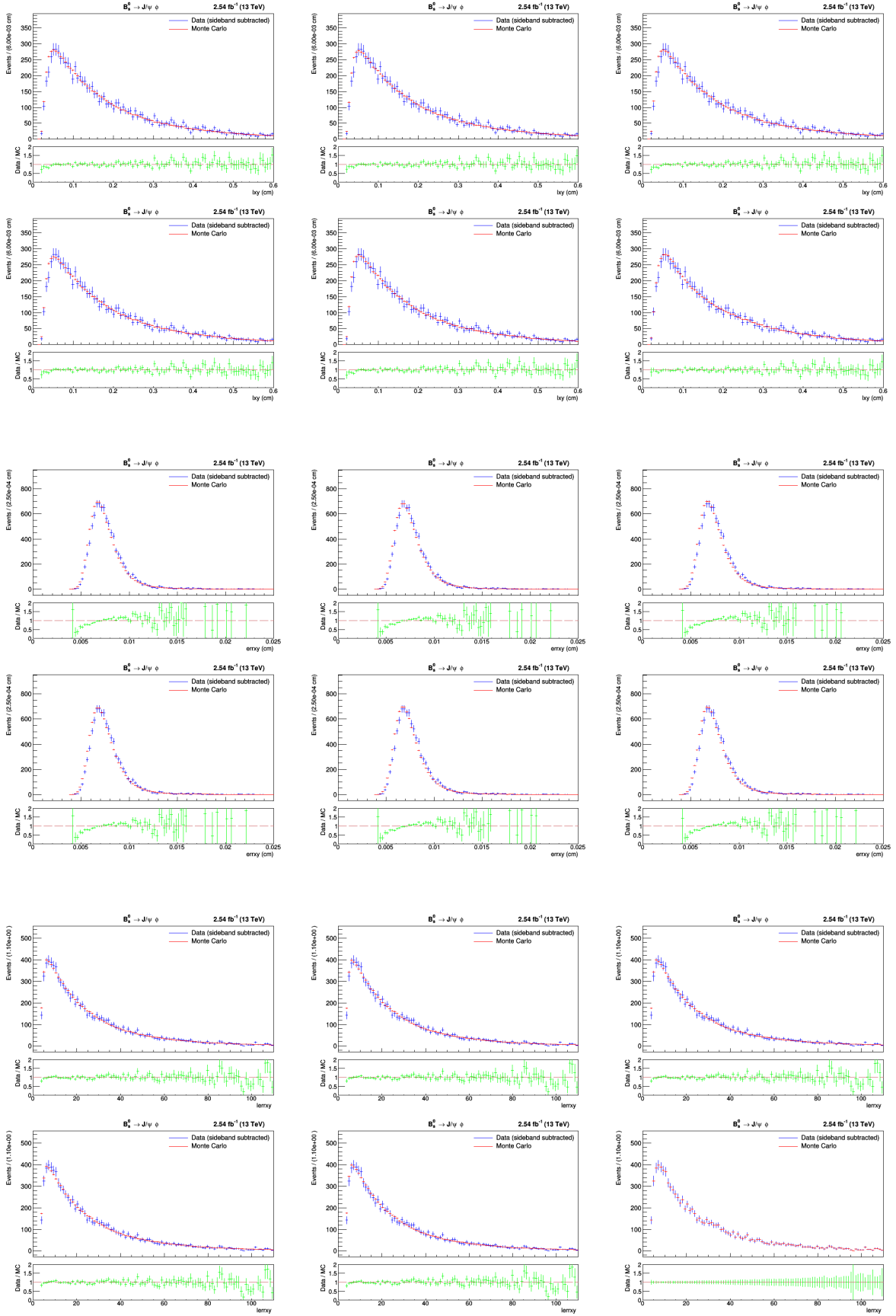
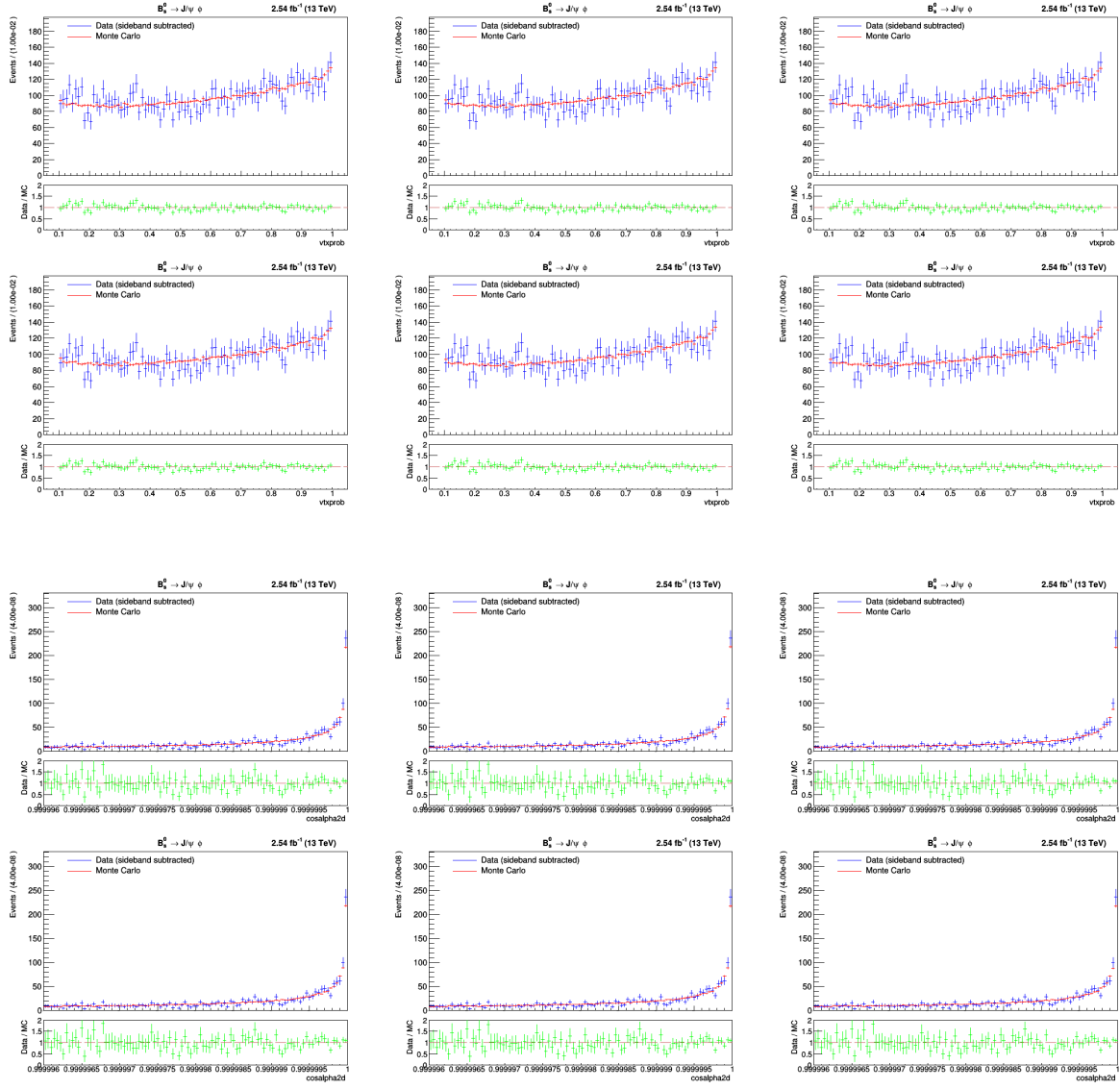


Figure B.11: MC reweighting applied to the L_{xy} (top group), σ_{xy} (middle group) and L_{xy}/σ_{xy} (bottom group) variables. The top left plot of each group shows the original data/MC comparison. The other plots show the reweighting using as reweighting variables, from top middle to bottom right, $p_T(\mu_1)$, $\eta(\mu_1)$, $p_T(\text{track}_1)$, $\eta(\text{track}_1)$ and L_{xy} significance.



Appendix C

Invariant mass fits

C.1 B^+ channel

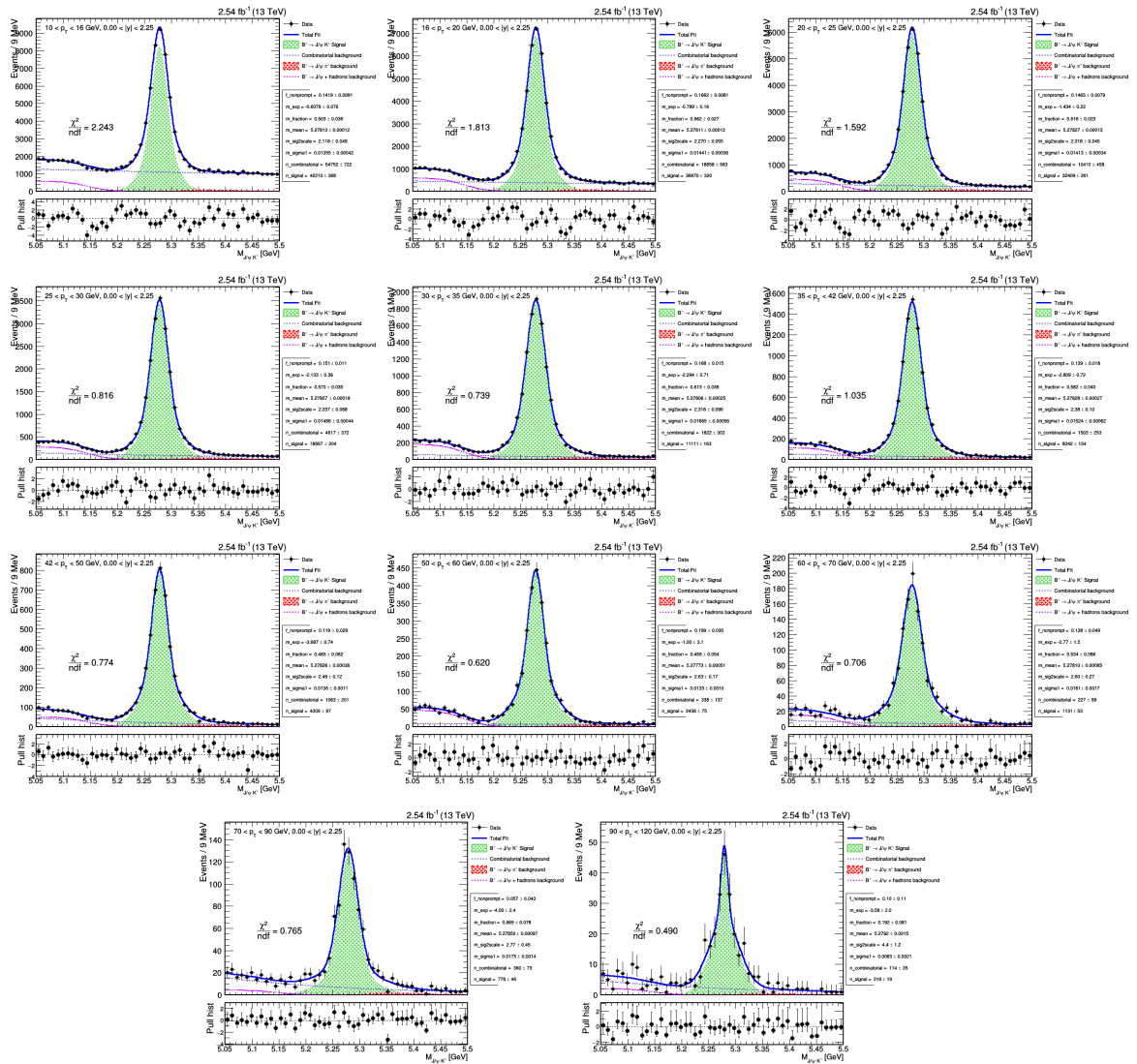


Figure C.1: Unbinned extended maximum likelihood fit for the B^+ channel, using bins of p_T . The same kind of fits were performed for γ (B) bins.

C.2 B^0 channel

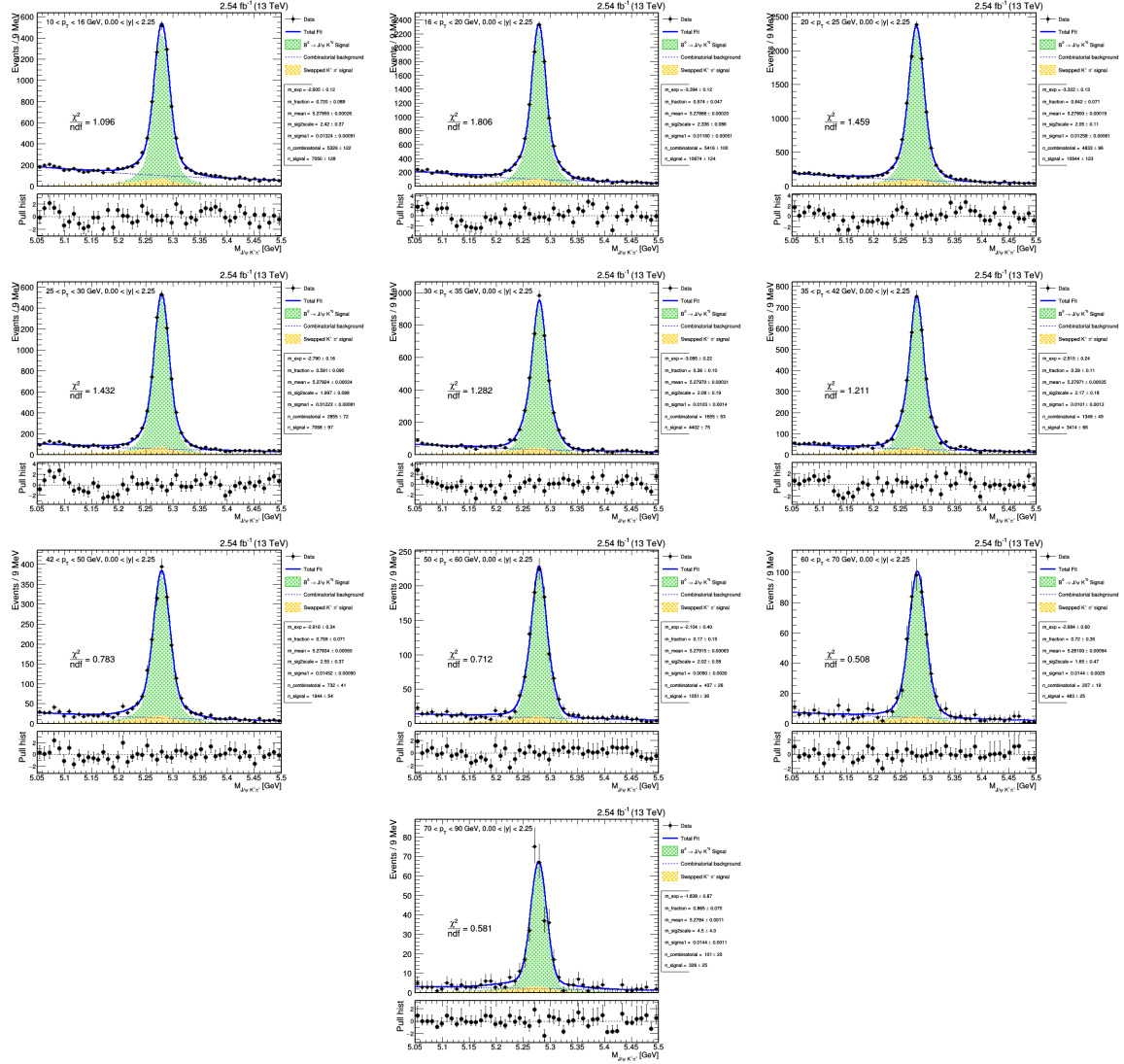


Figure C.2: Unbinned extended maximum likelihood fits for the B^0 channel, using bins of p_T . The same kind of fits were performed for γ (B) bins.

C.3 B_S^0 channel

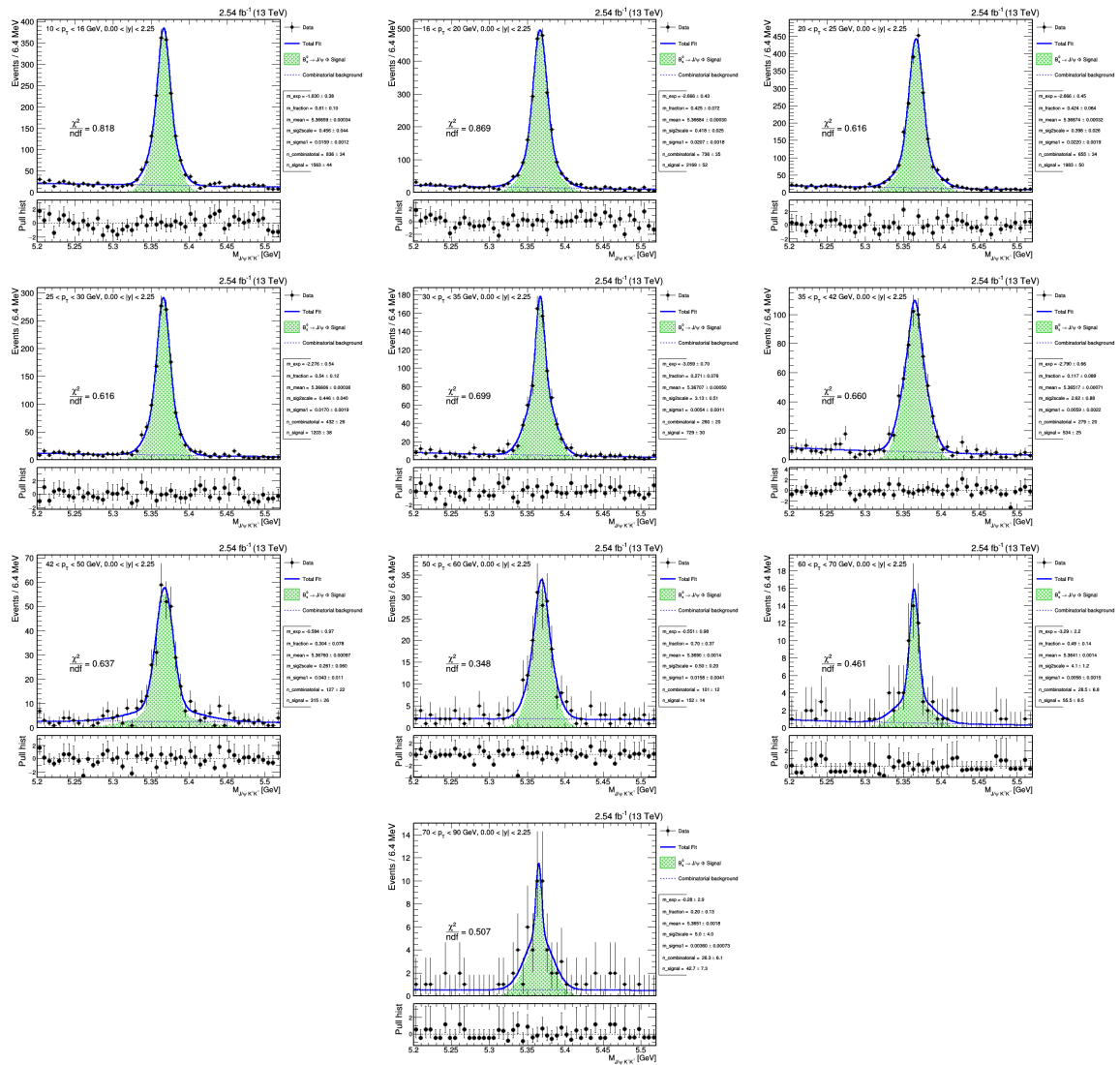


Figure C.3: Unbinned extended maximum likelihood fits for the B_s^0 channel, using bins of p_T . The same kind of fits were performed for $\gamma(B)$ bins.