CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD ZACATENCO
DEPARTAMENTO DE FÍSICA

"Mediciones de precisión de producción y decaimientos raros de hadrones b en el experimento CMS"

Tesis que presenta

Gabriel Artemio Ayala Sánchez

para obtener el Grado de

Doctor en Ciencias

en la Especialidad de

Física

Directores de tesis: Dr. Iván Heredia de la Cruz
Dr. Jhovanny Andrés Mejía Guisao

Ciudad de México Agosto, 2023

CENTER FOR RESEARCH AND ADVANCED STUDIES OF THE
NATIONAL POLYTECHNIC INSTITUTE

PHYSICS DEPARTMENT

"Precision measurements of b hadron production and rare decays in the CMS experiment"

by

Gabriel Artemio Ayala Sánchez

In order to obtain the

Doctor of Science

degree, speciality in

Physics

Advisors: Ph. D. Iván Heredia de la Cruz
Ph. D. Jhovanny Andrés Mejía Guisao

Mexico City August, 2023

Asesor:	Dr. Ivan Heredia de la Cruz	Cinvestav, Departamento de Física
Asesor:	Dr. Jhovanny Mejia Guisao	Universidad de Antioquia
Sinodal:	Dr. Heriberto Castilla Valdez	Cinvestav, Departamento de Física
Sinodal:	Dr. Alonso Contreras Astorga	Cinvestav, Departamento de Física
Sinodal:	Dr. Mateo Ramírez García	Universidad iberoamericana
Sinodal:	Dra. Estela Alejandra Garcés García	Instituto de Fisica, UNAM

Contents

Resumen

El descubrimiento en 2015 del decaimiento raro $B_s \to \mu^+\mu^-$ impuso estrictas constricciones a teorías de nueva física, tales como supersimetría. Esta sola observación ejemplifica como las mediciones de precisión de hadrones B tienen gran potencial para ya sea descartar modelos de nueva física o encontrar desviaciones al Modelo Estándar.

Este trabajo reporta la medición de la dependencia cinemática a las razones de fracciones de producción, f_s/f_u y f_d/f_u, en colisiones protón-protón en el detector CMS. Interesantemente, la razón f_s/f_u es constante para alto momento transverso, consistente con los resultados de LEP, pero incrementa para valores más pequeños, tal como lo observó LHCb. La observación de esta transición es el resultado principal de esta tesis. No se encontró dependencia en rapidity. Por otro lado, la medición de f_d/f_u representa la primera prueba de invariancia de isospín en producción de mesones B en colisiones protón-protón. Es independiente de momento transverso y rapidity. Estos resultados proveerán mejoras a próximas mediciones a las fracciones de desintegración a los decaimientos $B_s \to \mu^+\mu^-$ y $B^0 \to \mu^+\mu^-$, pues f_s/f_u y f_d/f_u son fuentes relevantes de incertidumbre sistemática.

Adicionalmente, un segundo análisis se enfoca en el decaimiento $B^0 \to K^0_S\mu^+\mu^-$. Exploramos la posibilidad de medir las fracciones de desintegración y parámetros de decaimiento angular en el experimento CMS asi como evaluar la competitividad de los resultados esperados con otros experimentos. El análisis usa una muestra de datos llamada "BParked", grabada con un novedoso trigger dinámico, basado en un muon desplazado, el cual no restringe el vértice de decaimiento desplazado del mesón K^0_S . Los resultados proveen suficiente evidencia de la factibilidad de realizar tales mediciones con los datos actuales (Run 2) y una mejora con los próximos datos (Run 3).

Abstract

The discovery in 2015 of the rare decay $B_s \to \mu^+\mu^-$ imposed stringent constraints on new physics theories, such as Supersymmetry. This single observation exemplifies how precision measurements of B hadrons have great potential to either discard new physics models or find deviations from the Standard Model.

This work reports the measurement of the kinematic dependence of the B hadron production fraction ratios, f_s/f_u and f_d/f_u, in proton-proton collisions with the CMS detector. Interestingly, the f_s/f_u ratio is constant at large transverse momenta, consistent with LEP results, but increases at smaller values, as observed by the LHCb experiment. The observation of this transition is the main result of this thesis. No dependence was found in rapidity. On the other hand, the measurement of f_d/f_u represents the first test of isospin invariance in B meson production in proton-proton collisions. It shows to be independent of transverse momentum and rapidity. These results will improve future measurements of the branching ratios of the decays $B_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$, as f_s/f_u and f_d/f_u are relevant sources of systematic uncertainty.

Additionally, a second analysis focuses on the $B^0 \to K^0_S\mu^+\mu^-$ decay. We explore the possibility to measure the branching fraction and decay angular parameters in the CMS experiment and assess the competitiveness of the expected results with other experiments. The analysis uses the so-called "BParked" data sample, recorded with a novel, dynamic, displaced muon trigger, which does not restrict the significantly-displaced decay-vertex of the K^0_S meson. Results provide sufficient evidence of the feasibility to realize such measurements with current (Run 2) data and provide a significant improvement with incoming (Run 3) data.

Agradecimientos

Al CONACYT por el apoyo económico proporcionado para realzar mis estudios de maestría y doctorado.

Al CINVESTAV por permitirme realizar mis estudios en sus instalaciones, y con la enseñanza de sus investigadores.

A mi amada Jareth León, que me ha acompañado en todo momento durante todos mis estudios de posgrado.

A mis sinodales el Dr. Ivan Heredia y el Dr. Jhovanny Mejia, gracias por todo el apoyo brindado.

A mis hermanos cuyo apoyo incondicional me ha permitido seguir cumpliendo mis objetivos.

A mis padres, por su apoyo y cariño.

A mis compañeros de CMS, Horacio Crotte, Rogelio Reyes, Iraq Rabadán, Daniel Pérez (el burgués), Y Osvaldo Miguel. Siempre me brindaron sus consejos y amenas discusiones de física.

A todos los miembros del equipo de CMS del CINVESTAV liderado por el Dr. Heriberto Castilla.

A mis buenos amigos y compañeros del CINVESTAV, Marco, José, Gonzalo, Marcela, Salvio.

A Eléctrica ARSA, Vero y todos sus colaboradores.

A Rosemary, Flor, Mariana y todas las secretarias del departamento de física, por su amabilidad y paciencia.

A mi gatita griselin, que llego a mi vida cuando más la necesitaba.

Chapter 1

Introduction

1.1 Standard Model of Elementary Particles

The Standard Model (SM) of particle physics is the best theory achieved so far to describe and classify the observed elementary particles and three of the four known interactions between them: electromagnetic, weak, and strong. In the SM, the first of these interactions can be described by the quantum electrodynamics theory (QED), which unifies the weak (EW) and the electromagnetic (EM) interactions together. The strong interaction is described through the theory of Quantum Chromodynamics (QCD). These theories are built under the paradigm of Quantum Field Theory (QFT), which allows the combination of quantum mechanics and special relativity (general relativity is not included), where particles appear as excitations of the quantum fields. Meaning its fundamental objects are quantum fields, which are defined at all points in space-time, these fields are:

- The fermion fields, which account for matter particles.
- The electroweak boson fields.
- The gluon fields.
- The Higgs field.

The mathematical consequence of these being quantum, rather than classical fields is that they are operator-valued. In particular, their values generally do not commute, as operators act upon a quantum state.

The fourth known interaction, which is not considered in the SM, is the gravitational force, the carrier of which would be the graviton. While it is true that the theory of quantum gravity has not been established yet, it is usually assumed that it can be treated as a QFT, at low energies, and in the context of an effective field theory (EFT), shows no actual impact at the energies considered in this work (neither in any current observation).

The elementary particles considered in the SM are illustrated in figure 1.1. These particles are usually divided by their spin, fermions have half-integer spin, and bosons have integer spin. Fermions include quarks that interact with all forces, and leptons that do not interact with the strong force. In the case of neutrinos, they are also blind to the EM interaction. The fermions are organized in a 3-generation (3-fold) structure, in this context, a generation or family is a division where the particles differ by their mass and flavor quantum number, but the electrical and strong interactions are identical. There is an important mass hierarchy between generations, and each generation contains two types of lepton and two types of quarks. The leptons are classified as one with electric charge -1 (electron-like) with neutral charge 0 (neutrino), and the two quarks in each generation are classified with charge $-1/3$ (down-type) and with charge $+2/3$ (up-type).

The spin 1 bosons are the carriers of the three different forces: gluons (g) carry the strong force, photons (γ) carry the electromagnetic force, and the weak force is carried by the Z and $W^{\pm}$ bosons. In the case of the Higgs boson (spin 0), it is modeled as the excitation of the Higgs doublet field, whose non-vanishing vacuum expectation value is responsible for the electroweak symmetry breaking.

The SM has a simple and elegant structure: it is a chiral gauge theory. It reveals a rich phenomenology that can account for strong and electroweak interactions, confinement and spontaneous symmetry breaking, hadronic and leptonic flavour physics, etc. [79]. The chiral gauge structure is built upon the unitary product group:

$$G_{\mathrm{SM}}^{\mathrm{gauge}} = SU(3)_C \times SU(2)_L \times U(1)_Y, \tag{1.1}$$

where each gauge group corresponds to a different gauge symmetry, i.e. invariance under local transformations (the subscripts C, W, and Y denote color, weak isospin, and hypercharge, respectively). For the strong interaction, the QCD sector defines interactions between quarks and gluons, under the $SU(3)_C$ symmetry. The weak interaction respect $SU(2)_L$ symmetry and the hypercharge interaction a $U(1)_Y$ symmetry.From a theoretical standpoint, the SM has four additional global symmetries, usually de-

noted as accidental symmetries, which are continuous $U(1)$ global symmetries. By Noether's theorem, each symmetry has an associated conservation law; the conservation of baryon number, electron number, muon number, and tau number. Each quark is assigned a baryon number of $1/3$, and $-1/3$ for antiquarks. This simply implies that the number of quarks minus the number of antiquarks is a constant (proved correct within experimental uncertainties). Each lepton is associated with a quantity named lepton family number. In a version of the SM (where neutrinos are massless), it predicts that each of these three numbers should be conserved separately in a manner similar to baryon. Experimentally, neutrino oscillations have been observed, demonstrating that individual electron, muon, and tau numbers are not conserved. For a complete description of the structure of the SM Lagrangian see Ref. [82]. For the purposes of this work, we will continue describing some features of the SM relevant to the physics behind the experimental measurements presented.

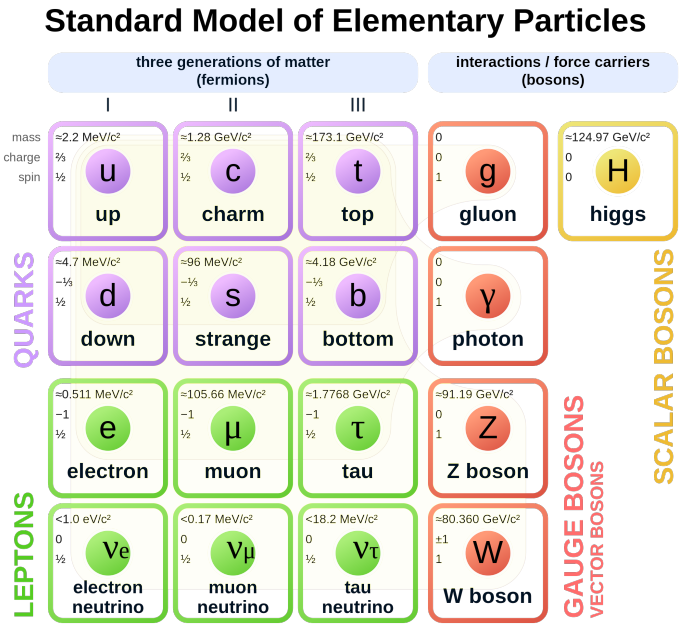

Figure 1.1: Standard Model schematic figure. Taking the column ordering from left to right, the 1st, 2nd, and 3rd generations of fermions are portraited, then the next column is populated by the gauge bosons, and lastly the Higgs boson is shown. Image from Ref. [40].

1.1.1 Virtual Particles

It has been stated that the fundamental particle interactions are mediated by force carriers (gauge bosons). For example, in a nuclear beta decay, a neutron decays into a proton, an electron, and an electron anti-neutrino (Figure 1.2). In this case, the weak interaction is mediated by a virtual W^- boson, but the mass of the W^- boson is 80.401(38) GeV, while the mass difference between the neutron and the proton is about 1.3 MeV. In this context, the word virtual means that the process violates energy and momentum conservation laws for a very short period of time, and it is also said that the W^- boson is "off-shell".

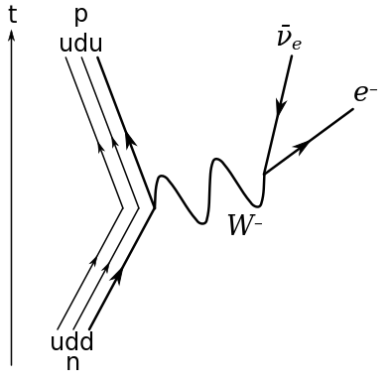

Figure 1.2: Feynman diagram of the beta decay

1.1.2 GIM Mechanism: $K_L^0 \to \mu^+\mu^-$ decay

The **GIM** mechanism (or Glashow–Iliopoulos–Maiani mechanism) was the solution to a problem arising in the simplest weak interaction theory with one charged vector boson coupled to the Cabibbo currents (weak hadronic interaction introduced by Nicola Cabibbo in 1963 [31], see Sec. 1.1.3). In strangeness-changing neutral-current processes, such as $K_L^0 \to \mu^+\mu^-$ and $K^0 - \bar{K^0}$ mixing (a special case of FCNC explained in Sec. 1.1.6), these are not allowed to happen at tree level, so they are generated at one loop with amplitudes of order $\backsim G \sin\theta_c \cos\theta_c (G\Lambda^2)$, where G is the Fermi constant, Λ is an ultraviolet cutoff, and $G\Lambda^2$ is the first (dimensionless) term in a perturbative expansion which could be continued to take higher order diagrams into account. According to the experimental results of that time, one had to require a small value of the Λ (around 2-3 GeV), compared to the expected value of $\Lambda = G^{-1/2} \backsim 300$GeV, so

Glashow–Iliopoulos–Maiani wrote in their work: "it appears necessary to depart from the original phenomenological model of weak interactions" [49].

To solve this problem, it was proposed the existence of a fourth quark (at that time only u, d, and s have been discovered), the charm quark, the idea of the GIM mechanism is that $K_L^0 \to \mu^+\mu^-$ only occurs via loops, one involving the u quark and the other the c quark (Figure 1.3), so amplitudes for s → d, with u or c on the same fermion line, would cancel exactly for $m_c = m_u$, providing a mean to suppress strangeness-changing neutral-current processes. Therefore the rare-observation (non-observation at the time) of this decay could be explained by adding a new particle to the theory, the c quark, which was eventually discovered in 1974 independently by teams at the Stanford Linear Accelerator Center (SLAC) [23] and at the Brookhaven National Laboratory (BNL) [22] [1].

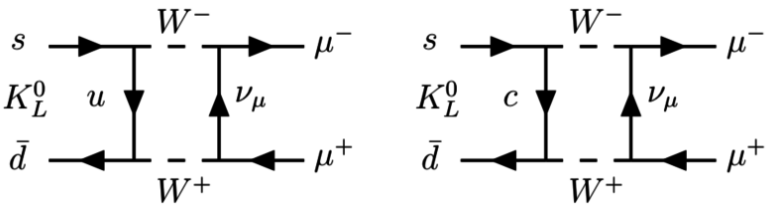

Figure 1.3: Feynman diagrams of loops contributing to the decay $K_L^0 \to \mu^+\mu^-$.

This is an example of an observation of New Physics mediated by a new virtual particle. There is a great interest in the $B_s \to \mu^+\mu^-$ decay for the same reasons exposed here.

1.1.3 CKM Matrix

As mentioned before in Sec. 1.1, the fermions are organized in a 3-fold family structure, an EW structure for all fermions, and just a color structure for quarks. This 3-generation description was achieved in a joint effort starting with Cabbibo [31], followed up by Glashow, Iliopoulos, and Maiani in 1970 [49], and a few years later by Kobayashi and Maskawa (introduced a three-generation quark mixing) in 1973 [56]. In strong interactions (modeled by QCD), quarks with strong eigenstates can change color by emission or absorption of a gluon; in weak interactions, quarks with weak eigenstates can change flavour by emission or absorption of an intermediate vector

[1] Individual quarks can not be observed due to confinement. These teams found a $c\bar{c}$ state, and they called it J and ψ respectively. It's now called J/ψ.

bosons $W^{\pm}$. Color transformations do not change Flavour. Flavours come in three generations and the generations can be mixed by flavour transformations but not by color transformations. The mixing between generations is a consequence of the mixing between mass eigenstates, and flavour eigenstates. The mass eigenstates u, c, t, d, s, b are the base for the strong interactions, while the weak eigenstates u', c', t', d', s', b' constitute the base of weak interactions. The change of basis matrix is the Cabibbo-Kobayashi-Maskawa matrix:

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.2}$$

It is chosen (as a convention) so that the up-type flavors are the same for both interactions and the down-type are transformed with the matrix . The CKM matrix describes the probability of a transition from one flavour j quark to another flavour i quark. These transitions are proportional to $|V_{ij}|^2$, schematized in Fig. 1.4.

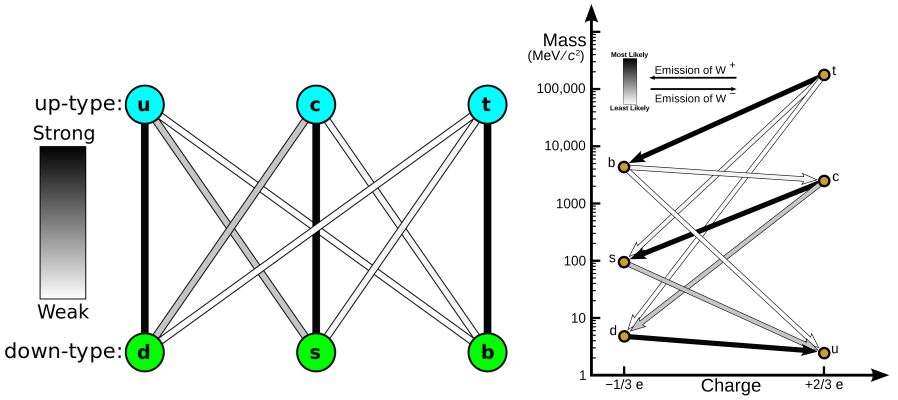

Figure 1.4: Schematic figures representing the allowed transitions between quarks. On the left, a representation of the six quarks in the usual structure (as in Fig. 1.1) [39]. On the right, a diagram of the decay routes allowed at tree level in a charged weak interaction, and a representation of their likelihood. The intensity of the lines is proportional to the measured squared values of the CKM entries [41]

The SM says nothing about the elements of the CKM-matrix except for the unitarity requirement , corresponding to assuming transformations to stay within exactly three generations. This assumption reduces the number of independent parameters to three angles and one phase (associated with CP-violation). Although experimental

values are aligned with a 3-generations unitarity matrix, it is still not a closed chapter in the SM.

1.1.4 Z mediated weak interactions

All fermions couple to the $Z-$boson, but its important to emphasize the following points:

- $Z-$boson exchanges transfer momentum, spin, and energy, which leave the interacting particle's quantum numbers unaffected (flavour, baryon number, lepton number, etc.).

- The couplings of the $Z-$boson are diagonal and as a result of this there are no Z-mediated flavour changing neutral currents (Sec. 1.1.6): The reason is evident, as mentioned in section 1.1.3, quark mixing occurs due to a misalignment in the flavour basis and the mass basis. The hypercharge and the EM current cannot fail to be flavor diagonal, the interaction states and mass eigenstates are identical, and there is no possibility for misalignment. Diagonality has been tested with Ref. [66]:

$$\begin{aligned} BR(Z \to e^+e^-) &= (3.3632 \pm 0.0042)\%, \\ BR(Z \to \mu^+\mu^-) &= (3.3662 \pm 0.0066)\%, \\ BR(Z \to \tau^+\tau^-) &= (3.3658 \pm 0.0023)\%, \end{aligned} \tag{1.3}$$

and

$$\begin{aligned} BR(Z \to e^+\mu^-) &< 7.5 \times 10^{-7}, \\ BR(Z \to e^+\tau^-) &< 5.0 \times 10^{-6}, \\ BR(Z \to \mu^+\tau^-) &< 6.5 \times 10^{-6}. \end{aligned} \tag{1.4}$$

- The $Z-$boson couplings to different fermion generations are universal. The universality is a result of the chiral representation $SU(2)_L \times U(1)_Y$. Recent tests of lepton flavour universality from Z-decays are [66]:

$$\begin{aligned} \Gamma(\mu^+\mu^-)/\Gamma(e^+e^-) &= 1.0009 \pm 0.0028, \\ \Gamma(\tau^+\tau^-)/\Gamma(e^+e^-) &= 1.0019 \pm 0.0032. \end{aligned} \tag{1.5}$$

1.1.5 Flavour changing charged currents (FCCC) in the SM

It has just been discussed, that any flavor-changing process will be mediated by a $W^{\pm}$ boson. The tree-level currents are called Flavour Changing Charged Currents (FCCC), as a charged W boson mediates them. In the quark sector, these transitions will manifest between up-type and down-type quarks, whereas in the lepton sector, between leptons and their corresponding neutrinos. The quark transitions are possible due to the misalignment of the electroweak basis and the mass basis discussed in the CKM Sec. 1.1.3.

The three types of FCCC decays are:

- *fully leptonic decays* ($\mu \to e\bar{\nu}_e\nu_\mu$): natural processes, used to test Lepton Flavour Universality (LFU).

- *semileptonic decays* ($B \to Dl\bar{\nu}_l$): used to test LFU and are in general good probes to the SM, the CKM matrix and its consistency. *non-leptonic decays* ($B \to D\pi$). These are no longer probes for LFU, since there are no leptons involved, in these processes, non-perturbative QCD regime is involved and they require a more complex treatment.

1.1.6 Flavour Changing Neutral Currents (FCNC)

Flavour-changing neutral currents (whereby a quark changes its flavour without altering its electric charge) are in some ways more attractive. These have been categorized as radiative ($B \to K^*\gamma$), semileptonic ($B \to Kl^+l^-$), and non-leptonic (divided in $\Delta F = 2$, like $B - \bar{B}$ mixing, and $\Delta F = 1$, like $B \to K\pi$).

FCNCs are loop suppressed in the SM (they do not appear at tree level, as in Sec. 1.1.5). The semileptonic decays occur through a penguin diagram or box diagram. Compared to FCCCs, FCNC are rare [2], so there is an additional complexity when making measurements. However, in the case of charged leptons in the final states without neutrinos, they are experimentally appealing since there is no missing energy and are relatively clean, implying good detection efficiency. As a final remark, it is important to say that these decays experience further suppression due to the CKM unitarity and the GIM mechanism (see Ref. [49], Sec. 1.1.2).

[2] The suppression when comparing $B \to Dl\nu$ to $B \to Kll$ is in the order of 10^{-5}.

1.1.7 *b*-hadron Decays

Having discussed new physics can be found indirectly trough **FCNC** decays, the focus will change to the study of rare decays of hadrons containing a b quark. $B_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays (purely leptonic) are the simplest FCNC process of B mesons, they are theoretically well understood. Small QED corrections have been computed in works as Ref. [27], and non-perturbative QCD effects are parametrized by the B meson decay constants, which are known from lattice QCD.

In these channels, it is possible to observe small contributions from virtual new particles that are too heavy to be produced at colliders and may lead to measurable deviations from the expected properties in the SM.

Electroweak penguin diagrams as in Fig. 1.5 are of particular interest. Here quark-level transitions cannot be measured directly as the quarks form immediately hadrons. These processes are experimentally favored but, come with larger theoretical uncertainties.

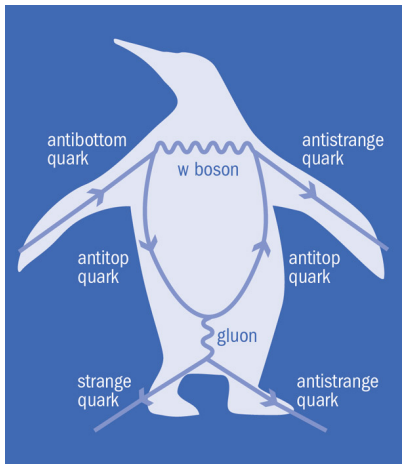

Figure 1.5: Original penguin diagram. (Named by John Elis in Ref. [46]), describing the $b \to s$ decay, where a gluon produces an $s\bar{s}$ pair (Image: Symmetry/Fermilab).

The rare decays theoretical description is challenging due to the different energy scales involved, which are represented in Fig. 1.6. These energy scales are more or less:

$$\mathcal{O}(\Lambda_{QCD}) \sim 0.2\,\text{GeV}, \quad \mathcal{O}(m_b) \sim 4\,\text{GeV}, \quad \mathcal{O}(\mu_{\text{EW}}) \sim 80\,\text{GeV}. \tag{1.6}$$

In some treatments, people add an extra high energy scale for new physics (NP)

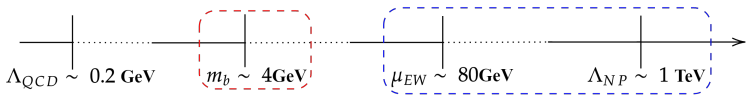

Figure 1.6: Energy scales present in flavour physics. The blue dashed area will be "integrated-out" (EFT sub-section).

(Λ_{NP}), and it is assumed that this scale will be higher than $\mu_{\rm EW}$, about 1 TeV which is reasonable since direct searches at the LHC have not been successful. Having this in mind, the energy hierarchy will be:

$$\Lambda_{\rm QDC} \ll m_b \ll \mu_{\rm EW} \ll \Lambda_{\rm NP}. \tag{1.7}$$

One of the main reasons why it is important to classify this phenomenon in terms of its energy is due to the confinement in QCD, which can not be computed perturbatively at low energies, and also to avoid large logarithms of the different scales in the use of RGE (Renormalization Group Theory) perturbation theory. In this case, effective field theory is a convenient methodology because it takes advantage of the energy hierarchy.

Effective Field Theories

Effective field theories (EFTs) are a useful tool in the computation of QFTs in a multi-scale energy range [67]. The main framework was developed by Wilson and Zimmermann in Ref. [80]. Consider a QFT with a high energy scale M (it could be the mass of a heavy particle or large momentum transfer), and there is an interest to describe its behavior at lower energy, say E, such that $E \ll M$. It is possible to define a "cut-off" [3] energy Λ at or slightly below M, such that the fields can be divided into high energy modes and low energy modes such that $E \ll \Lambda < M$ (it is possible to add several energy scales, but the treatment is to take them one by one)and where ϕ_L is for the Fourier modes with low frequency $\omega < \Lambda$ and ϕ_H for high-frequency modes with $\omega > \Lambda$.

$$\phi = \phi_L + \phi_H \tag{1.8}$$

By construction, the low-energy physics will be contained in terms of ϕ_L fields, so all computations can be derived from these fields (scattering amplitudes, cross sections,

[3] Ref. [64] describes the "cut-off" energy as a "threshold of ignorance", where it is pretended to know nothing about the theory for scales above μ.

decay rates, etc). The high-energy modes do not propagate on long distances, they only appear as virtual particles (Sec. 1.1.1) and they are "integrated out" or "removed" from the calculations (calculate the path integral over these modes, with the "Wilsonian effective action). The formal procedure can be found in Refs. [64] or [50].". Then again, by construction, the action which is integrated with the high-frequency modes depends on the choice of Λ, and this theory will be non-local on scales $\Delta x^\mu \sim \frac{1}{\Lambda}$, because the fluctuations of the high energy modes have been removed from the theory. To tackle this problem, the resulting non-local theory can be expanded in powers of $1/\Lambda$, in terms of local operators composed of light fields (this is called Wilsonian operator product expansion or OPE). The expansion is possible because it has been established that $E \ll \Lambda$. The resulting mathematical object is called the "effective Lagrangian". It is an infinite sum over local operators $\mathcal{O}_i$, in powers of $1/M$,

$$\mathcal{L}_{\text{eff}} = \sum_i \frac{1}{M^{d_i-4}} \mathcal{C}_i(\Lambda)\mathcal{O}_i(\Lambda). \tag{1.9}$$

Here, each operator $\mathcal{O}_i$ is matched with a coefficient $\mathcal{C}_i$, whose value comes from the short-distance dynamics that have been integrated. $\mathcal{C}_i$ are the Wilson coefficients and their value depends logarithmically on Λ. The power of M^{d_i-4} is determined by the mass-dimension of the operator, $[\mathcal{O}_i] = d_i$. At low energy, the contribution of each operator $\mathcal{O}_i$ to an observable is expected to behave as:

$$\left(\frac{E}{M}\right)^{d_i-4} = \begin{cases} \gg 1 \text{ if } d_i < 4, \\ \mathcal{O}(1) \text{ if } d_i = 4, \\ \ll 1 \text{ if } d_i > 4. \end{cases} \tag{1.10}$$

So the importance of the operator at low energies is determined by its dimension and the expansion on E/M terms in Eq. 1.9 can be truncated at a given operator dimension.

Now that it has been discussed the general ideas behind the effective field theories, it is time to wrap up these concepts and discuss how they are applied to b-hadron rare decays. FCNC decays can be described with an EFT as an operator product expansion. The generic effective Hamiltonian of a $b \to s$ transition is:

$$\mathcal{H}_{eff} = \frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^* \frac{e^2}{16\pi^2} \sum_i [\mathcal{C}_i(\Lambda)\mathcal{O}_i(\Lambda) + \mathcal{C}_i'(\Lambda)\mathcal{O}_i'(\Lambda)], \tag{1.11}$$

where G_F is the Fermi constant and V_{ij} CKM elements. As mentioned before, the

Wilson coefficients, $\mathcal{C}_i$, contain perturbative short-distance effects, and the nonperturbative QCD and long-distance effects are encapsulated in their respective operators $\mathcal{O}_i$. The primed operators, with their respective primed Wilson coefficients, involve right-handed quark currents and describe genuine new physics effects. Any other new physics effects will either modify a Wilson coefficient or be a new operator not considered in the SM. The sensibility of a certain Wilson coefficient will depend on the decay channel (final states) and the energy scale or q^2 region (see Fig. 1.7). From now on q^2 denotes the invariant mass of the lepton pair in the final state. Table 1.1 shows the operators that can be measured through certain decays.

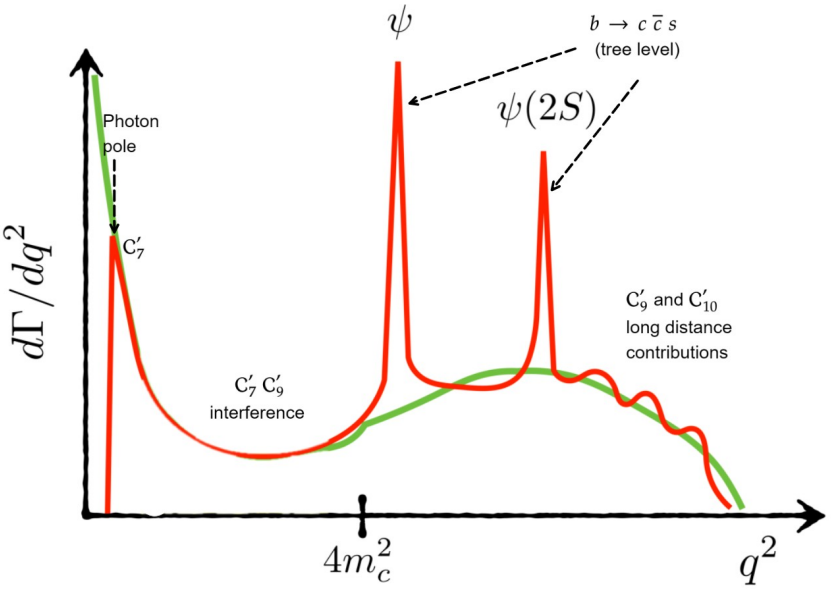

Figure 1.7: Schematic picture of the differential branching fraction as a function of the invariant dilepton mass squared (q^2) in a $b \to s l^+ l^-$ decay. It shows how different Wilson coefficients intervene in different kinematic regions. The Green line is the SM contribution from operators $\mathcal{O}_{7,9,10}$, and the red line shows the effect of including the lepton masses and the $c\bar{c}$ resonances into account. The figure ([45]) is taken from the Scholarpedia publication Ref. [58].

1.1.8 $B^0 \to \mu^+\mu^-$ and $B_s \to \mu^+\mu^-$ decays

The $B^0 \to \mu^+\mu^-$ and $B_s \to \mu^+\mu^-$ are very rare, so any small deviation from SM expectations can be a signal of New Physics (NP). As it has been mentioned, they can

Operator	$B_{d,s} \to X + \mu^+\mu^-$	$B_{d,s} \to \mu^+\mu^-$	$B_{d,s} \to X + \gamma$	Type
$\mathcal{O}_7$	X		X	electromagnetic penguin
$\mathcal{O}_9$	X			vector semileptonic
$\mathcal{O}_{10}$	X	X		axial-vector semileptonic
$\mathcal{O}_S$		X		scalar
$\mathcal{O}_P$		X		pseudoscalar

Table 1.1: Operators that contribute to $b \to s\mu^+\mu^-$ decays.

only occur through loop diagrams and, additionally, they are helicity suppressed. The SM value expected for the Branching Ratio (BR) is accurately known.

These particular decays are very sensitive to many NP models with scalar and pseudoscalar operators, which may change the decay rate. Additionally, the ratio between the $B^0 \to \mu^+\mu^-$ and $B_s \to \mu^+\mu^-$ BR also provides discrimination among NP theories. As can be seen in Fig. 1.8, these measurements have been crucial to dismiss almost all the minimal SUSY models.

An important input for the $B_s \to \mu^+\mu^-$ (and $B^0 \to \mu^+\mu^-$) is the ratio of the B_S to B^+ production fractions, f_s/f_u (and f_d/f_u). Will be discussed in detail in chapter 3.

1.1.9 Semileptonic $b \to s(d)l^+l^-$ transitions

Finally, in the scope of the theoretical framework that has been discussed so far, it is important to talk about semileptonic $b \to s(d)l^+l^-$ decays (first column in Table 1.1). These provide a rich spectrum of observables, sensitive to NP contributions. These are divided into three different kinds of measurements:

- Differential branching fraction measurement: These are performed in bins of q^2, the principal limitation comes from the theoretical uncertainties from hadronic form factors and non-factorizable effects (see Fig. 1.7). The latest LHCb experimental results on the decays $B^+ \to K^+\mu^+\mu^-$, $B \to K^{(*)}\mu^+\mu^-$, $B^0 \to K^0\mu^+\mu^-$, $B^+ \to K^+\mu^+\mu^-$ and $B^+ \to K^{*+}\mu^+\mu^-$ [6], $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ [5], and $B_s^0 \to \phi\mu^+\mu^-$ [4] show a tendency to lie bellow the SM prediction, in the entire q^2 region. Nonetheless, even the most discrepant results (below 6GeV^2 in the $B^0 \to K^0\mu^+\mu^-$) are within 3σ from the SM prediction, $B^0 \to K_S^0\,\mu^-\mu^+$ is explored with the CMS detector in chapter 4.

- Angular analyses: These can measure several observables, which gives complementary information on the previous differential BR measurements. Additionally it is possible to build observables that are free from form factors uncer-

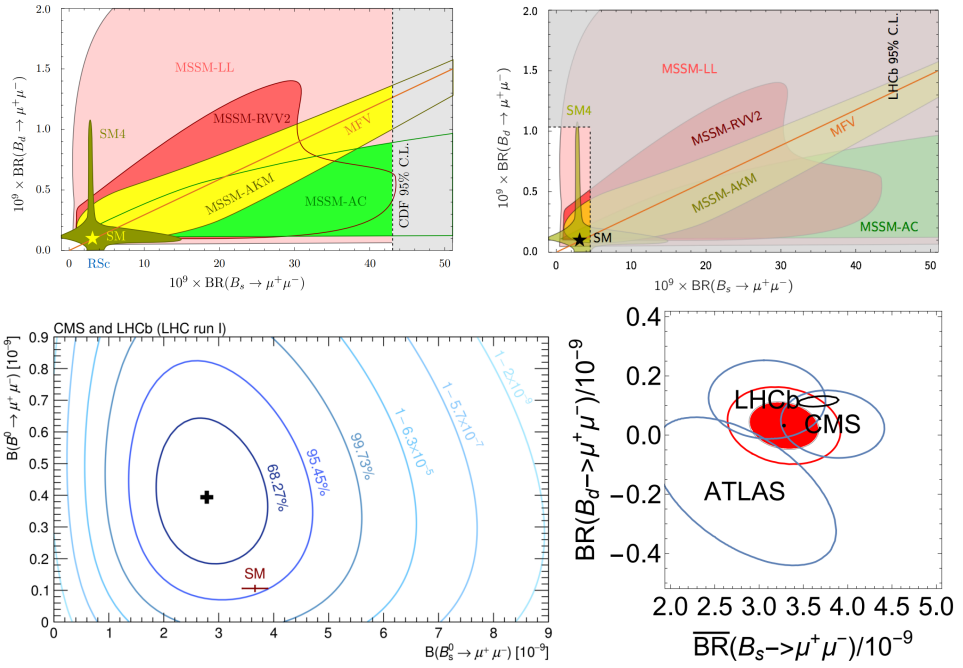

Figure 1.8: Top left plot, made by D.M Staub Ref. [75], shows the correlation between the branching ratios of $B_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$, comparing with MFV (Minimal Flavour Violation [43]), SM4 (SM with 4 generations of quarks and leptons [29]), and MSSM (Minimal Super Symmetric Model). The gray area was excluded by CDF results in 2012 Ref. [62]. The top right image shows the same plot, compared with the updated measurements made by LHCb in 2012 [12]. Bottom left image shows the same measurements using the full Run 1 data, namely a joint publication by LHCb and CMS [10]. The bottom right shows the LHCb, ATLAS, and CMS results up to 2022, recently presented in Moriond 2023 [18].

tainties at leading order. In chapter 4, a preliminary angular analysis in the $B^0 \to K^0\mu^+\mu^-$ channel is performed with the CMS detector, where the F_H parameter (which is related to a possible scalar and pseudoscalar contribution) is measured.

- Lepton Flavour Universality Violation (FLUV): BR of semileptonic decays are expected to be the same (universal) except for lepton mass effects (this topic is related with Sec. 1.1.4, third kind of tests in Eq. 1.5, but they are not the same!)

1.1.10 Global fits

All measurements mentioned so far, are in reasonable agreement with the SM if they are considered separately. However, an additional way to extract information from experimental searches is to perform a combined analysis, which takes into account all the observables present in the $b \to s(d)ll$, $b \to ll$, and $b \to h\gamma$ decays, for an estimate of around 90 parameters (depending the models considered). These kind of analyses are called global fits. They provide a powerful tool to minimize data-prediction deviations by allowing to vary the Wilson coefficients and allowing NP contributions (as a fit to certain parameters would do). Results from global fits are constantly changing with every new meausrement. Nowadays (2023) tensions are in the $\mathcal{C}_{9,10}$ parameters and $R_{K^{(*)}}$ ratios (see Ref. [63]).

Chapter 2

The Large Hadron Collider and the CMS detector

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) consists of a two-ring-superconducting-hadron accelerator and collider. It is the largest and most powerful particle accelerator until now [19]. It accelerates particles to nearly the velocity of light in clockwise and anticlockwise directions. It is installed in the existing 26.7 km tunnel constructed for the LEP machine. The tunnel has eight straight sections and eight arches lying at a max depth of 170 meters below the surface.

The machine comprises 1232 dipole magnets, with radiofrequency (r.f) cavities providing a kick that results in an increase in the proton energy of 0.5 MeV per turn. The luminosity is given by

$$\mathcal{L} = \frac{\gamma f k_B N_p^2}{4\pi\epsilon_n \beta^*} F, \tag{2.1}$$

where γ is the Lorentz factor, f is the revolution frequency, k_B is the number of proton bunches, N_p is the number of protons per bunch, ϵ_n is the normalized transverse emittance (with a design value of 3.75 μm), β^* is the betatron function at the interaction point, and F is the reduction factions due to the crossing angle. The nominal energy of each proton beam is 7 TeV. The bunches are formed in a 26 GeV Proton Synchrotron (PS) with the correct 25 ns spacing. After that, the beam is accelerated to 450 GeV in the super Proton Synchrotron (SPS) and transferred to the LHC. This

operation is repeated 12 times for each beam.

2.1.1 Coordinate system

The coordinate system adopted by CMS (other detectors around LHC) has the origin centered at the nominal collision point inside the experiment. The y-axis is pointing vertically upward, and the x-axis is pointing radially inward toward the center of the LHC. Thus, the z-axis points along the beam; in the case of CMS, directed toward the Jura mountains (west). The azimuthal angle ϕ is measured from the x-axis in the $x-y$ plane. The polar angle θ is measured from the z-axis. Pseudorapidity is defined as $\eta \equiv -\ln(\tan(\theta/2))$. Therefore, the momentum and energy measured transverse to the beam direction, denoted by p_T and E_T, respectively, are computed from the x and y components. The imbalance of energy measured in the transverse plane is denoted by E_T^{miss}.

2.2 CMS Detector

The Compact Muon Solenoid (CMS) is at one of four collision points (point 5) of the LHC. It is a 14,000-tonne detector that owns its name to the following:

- It is 15 meters high and 21 meters long, which is quite **compact** for all the detector material it contains.
- It is designed to detect **muons** very accurately.
- It is a powerful **solenoid** capable of producing a magnetic field of around 4 Tesla.

The general layout of the detector is shown in figure 2.1.

In the design phase of the CMS and ATLAS detectors in the early 1990s, the detection of the standard model Higgs boson was used as a benchmark to test the performance of the proposed designs. It was a particularly appropriate benchmark since there is a wide range of decay modes depending on the mass of the Higgs boson. The lower mass limit detected by LEP was 114.4 GeV. The branching fractions of the Higgs bosons are dominated by hadronic channels, which were difficult to use as discovery channels due to large QCD backgrounds and the relatively poor mass resolution that is obtainable with jets. Hence, the search was conducted using final states that contained isolated leptons and photons, despite the smaller branching ratios. Apart from

the Higgs boson, there were other important areas that needed to be tested, which is why CMS is called a multipurpose detector (it was not designed with a single objective in mind) [26], in particular:

- Search for supersymmetric particles: In the 1990s it was expected that the decays of supersymmetric particles, such as s-quarks or gluinos (supersymmetric partners of quarks and gluons), would involve cascades that, if R-parity is conserved [1], must contain the lightest supersymmetric particle (LSP). The latter is expected to interact very weakly, thus leading to significant E_T^{miss} in the final state. The rest of the cascade results would be an abundance of leptons and jets (especially b/τ-jets). In gauge-mediated supersymmetry breaking (GMSB) schemes, with the LSP decaying into a photon [2] an increased number of hard isolated photons were expected.

- Search for new massive vector bosons: In case of discovery of an object like a Z' boson, to distinguish between different models it is important to measure the natural width and the forward-backward asymmetry. This requires sufficiently good momentum resolution at high p_T ($\sim$ $\Delta p_T/p_T < 0.1$ in the range of $p_T \approx 1$ TeV) to determine the sign of the leptons and a pseudorapidity coverage up to $\eta = 2.4$. Both are needed to determine high momenta final states of massive objects, such as $Z' \to l^+l^-$ (l can be either μ or e).

- SM precision studies: The LHC opened the possibility to study QCD, electroweak, and flavor physics. Precision studies (such as the ones presented in this thesis) can give indications for physics beyond the SM, providing complementary information with respect to direct searches. The top quark will be produced at the LHC; so, to test its spin and couplings, good identification of b-jets is needed. Also, searches for FCNC and lepton flavor violation through $\tau \to 3\mu$ or $\tau \to \mu\gamma$, measurements of $B_s^0 \to \mu\mu$, and measurements of triple/quartic-gauge couplings are aimed as a search for new physics. A joint effort with the TOTEM experiment (TOTal cross-section, Elastic scattering, and diffraction dissociation Measurement at the LHC) and CMS was established to cover a full range of diffractive physics as well.

[1] R-symmetry is the largest subgroup of the automorphisms group of the supersymmetry algebra which commutes with the Lorentz group [81].

[2] Christopher Kolda writes: *In gauge-mediated models the gravitino is light, roughly 1 eV to 1 GeV, making it the LSP. It is still too weak to be directly produced in experiments, but as the LSP, all other SUSY particles must eventually decay into it. The phenomenology of gauge-mediation, which is otherwise so much like that of supergravity, has a new component, the search for decays into gravitinos.* in [57].

- Extra dimensions: The existence of extra dimensions can lead to a characteristic energy scale of quantum gravity, M_D, which is the analog of the Plank mass in a D-dimensional theory, possibly just beyond the electroweak scale. In terms of experimental signatures, 3 regimes were established as distinguishable [68]: i) Cis-Planckian, where $E \ll M_D$, leading to signals involving the emission of gravitinos that escape into extra dimensions ($pp \rightarrow \text{jet} + \text{graviton}$, i.e. E_T^{miss}) ii) Planckian, $E \approx M_D$, leading to model-dependent signatures, some of which would be visible as Z-like resonances with large separations in mass ($\sim$ TeV). iii) Trans-Planckian, $E \gg M_D$, leading to mini black hole production with decays involving equal production of fundamental particles such as leptons, photons, neutrinos, W, Z, jets, etc. The resulting production and kinematic distributions could allow the determination of the Hawking temperature, the mass of the black holes, or the number of extra dimensions.

- Heavy-ion physics: The collision energy of 5.5 TeV (for heavy-ion) allows the study for the studies of jet quenching at high $\mathrm{p_T}$, the identification of fully formed jets, and hard probes like Υ and Z^0. These studies in heavy-ion collisions require large acceptance, high-resolution calorimeters, and a flexible trigger.

These were the main goals when CMS was designed, so the detector requirements to meet the goals of the LHC physics program are summarized as follows:

- Good muon identification and momentum resolution are required over a wide range of momenta in the region $|\eta| < 2.5$, good dimuon mass resolution ($\approx 1\%$ at 100 GeV), with the ability to determine without doubt the charge of the muons with $p < 1$ TeV.

- Good charged particle momentum resolution and efficiency in the inner tracker, efficient triggering, and offline tagging of $\tau/b-$jets, needing pixel detectors close to the collision point.

- Good electromagnetic energy resolution: good diphoton/dielectron mass resolution (at least 1% at 100 GeV), wide pseudorapidity coverage ($|\eta| < 2.5$), a good measurement of the direction of photons (correct localization of the primary vertex), π^0 rejection and efficient photon and lepton isolation at a high luminosity environment.

- Good E_T^{miss} and dijet mass resolution: requires hadron calorimeters with a large hermetic geometric coverage ($|\eta| < 5$) and with fine lateral segmentation ($\Delta\eta \times \Delta\phi < 0.1 \times 0.1$).

The design of CMS can be consulted in detail in Ref. [26]. The main distinguishable features of CMS (that will be described briefly below) are a high-field solenoid, a full silicon-based inner tracking system, and a fully active scintillating crystals-based electromagnetic calorimeter.

2.2.1 Solenoid

A solenoid is needed to bend the charged particle's trajectories emerging from the high-energy collision in the LHC. The more momentum a particle has, the less its path is curved by the magnetic field, so tracing its path gives a measure of momentum [78]. The CMS strong magnet, combined with high-precision measurements in the tracker and muon detectors, allows for accurate measurement of the momentum of high-energy particles. The CMS magnet is a solenoid that produces a uniform magnetic field. It is made of stabilized NbTi superconducting wire coils, cooled down to -268.5 ºC. The inner coil radius is big enough so that the inner tracker and the calorimeter detectors (ECAL and HCAL) can fit inside (Fig. 2.1), whilst the muon detectors are interleaved with a 12-sided iron structure that surrounds the magnet and contains and guides the field. Made up of three layers, this "return yoke" reaches out 14 meters in diameter and also acts as a filter, allowing through only muons and weakly interacting particles such as neutrinos. The enormous magnet also provides the detector's structural support and must be very strong to withstand the forces of its magnetic field.

2.2.2 Tracker system

Measuring the momentum of particles is crucial to building up a picture of what happened at the heart of the collision. So, as mentioned before, to calculate the momentum, CMS tracks the charged particle through a magnetic field. The tracker needs to record particle paths accurately, yet be lightweight so as to disturb the particle as little as possible.

CMS obtains position measurements (hits) so accurately that tracks can be reliably reconstructed using just a few measurement points. Each measurement is accurate to 10 μm [38]. The innermost part of the tracking system is a silicon pixel detector. It provides 3-dimensional space points in the region closest to the interaction point that allows for high-precision, charged particle tracking and vertex reconstruction [36]. The pixel detector is located in a particularly harsh radiation environment characterized by a high track density. The original pixel detector consisted of three barrel layers at radii of 44, 73, and 102 mm and two endcap disks on each end at distances of 345 and

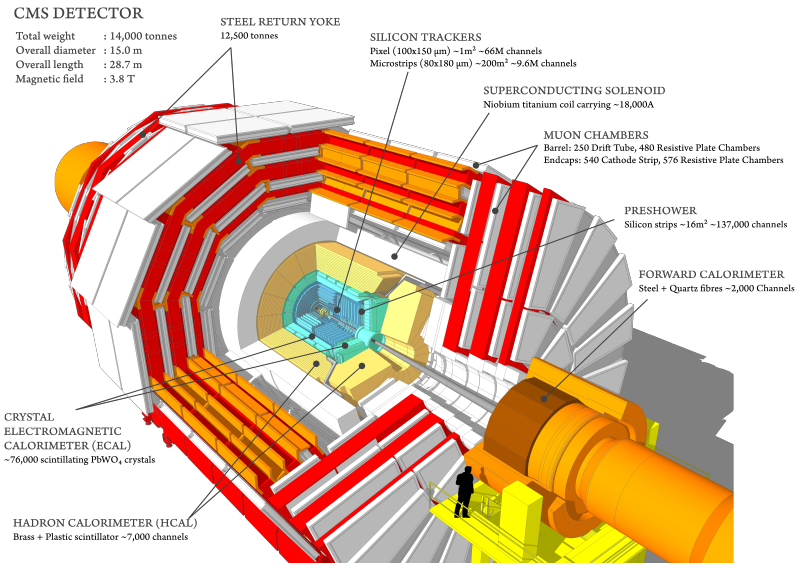

Figure 2.1: CMS detector diagram

465 mm from the interaction point. It was designed for a maximum instantaneous luminosity of $1 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ and a maximum average pileup (number of inelastic interactions per bunch crossing) of 25 in LHC operation with 25 ns bunch spacing. After the upgrade done to the accelerator during the first long shutdown (LS1, 2013-2014), these parameters have been exceeded and the luminosity and pileup have more than doubled compared to the original design values. In order to keep a robust and efficient tracking system under these conditions, the pixel detector has been changed by a new system, referred to as CMS phase-1 pixel detector [38]. The installation of the CMS phase-1 pixel detector took place during the extended year-end technical stop in 2016/2017.

2.2.3 ECAL

The CMS Electromagnetic calorimeter (ECAL) [76] was designed to measure the energies of the collision emerging particles. Of particular interest are electrons and photons because of their use in discovering the Higgs boson and other new phenomena.

Measuring energies with the necessary precision in the very strict conditions of the LHC - a high magnetic field, high levels of radiation, and only 25 nanoseconds between collisions - requires dedicated detector materials. The ECAL it is built as a compact, hermetic, fine-grain, homogeneous calorimeter made of 75848 lead tungstate

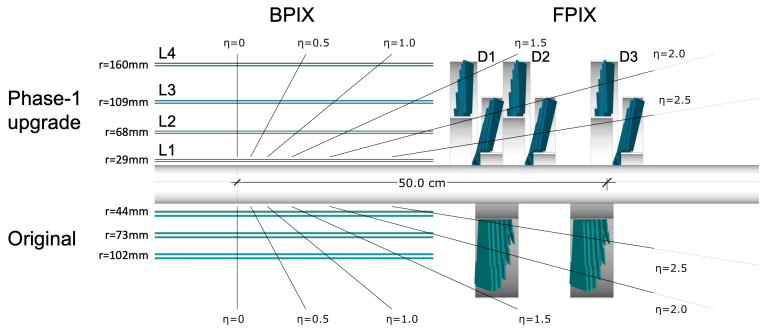

Figure 2.2: Layout of the CMS phase-1 pixel detector compared to the original layout in longitudinal view. The layout upgraded pixel detector is optimized to have four-hit coverage over the pseudorapidity range $|\eta| < 2.52$, improved pattern recognition and track reconstruction, and added redundancy to cope with hit losses. The total silicon area of the CMS Phase-1 pixel detector is 1.9 m^2, while the total silicon area of the original pixel detector was 1.1 m^2. The CMS Phase-1 pixel detector is built from 1856 segmented silicon sensor modules, where 1184 modules are used in the barrel pixel detector (BPIX) and 672 modules are used for the forward disks (FPIX). Each module consists of a sensor with 160 × 416 pixels connected to 16 readout chips (ROCs). In total there are 124 million readout channels.

(PbWO4) scintillating crystals, primarily made of metal heavier than stainless steel (density = 8.28g/cm^3). It is highly transparent and "scintillates" when electrons and photons pass through it. It also has a small radiation length ($X_0 = 0.89$ cm), and a Molière radius of $rm = 2.19\ cm$ (smaller Molière radius means better shower position resolution). This means it produces light in proportion to the impinging particle's energy. These high-density crystals produce light in fast, short, well-defined photon burst that allows for a fast, precise, and compact detector. The crystals are arranged in a quasi-projective geometry and distributed in a central barrel section (EB) covering the pseudorapidity range $|\eta| < 1.48$, and two end caps (EE) extending the coverage up to $|\eta| < 3.0$. For extra spatial precision, the ECAL also contains a pre-shower detector that sits in front of the endcaps. These allow CMS to distinguish between single high-energy photons (often signs of exciting physics) and the less interesting close pairs of low-energy photons. A diagram of the ECAL is shown in Fig. 2.3.

2.2.4 HCAL

The CMS hadron calorimetry system (HCAL) has four major sections [77]: the HCAL Barrel (HB), HCAL Endncap (HE), HCAL Outer (HO), and HCAL Froward (HF). The locations of each of these calorimeters are shown in Fig. 2.4. The HB covers the

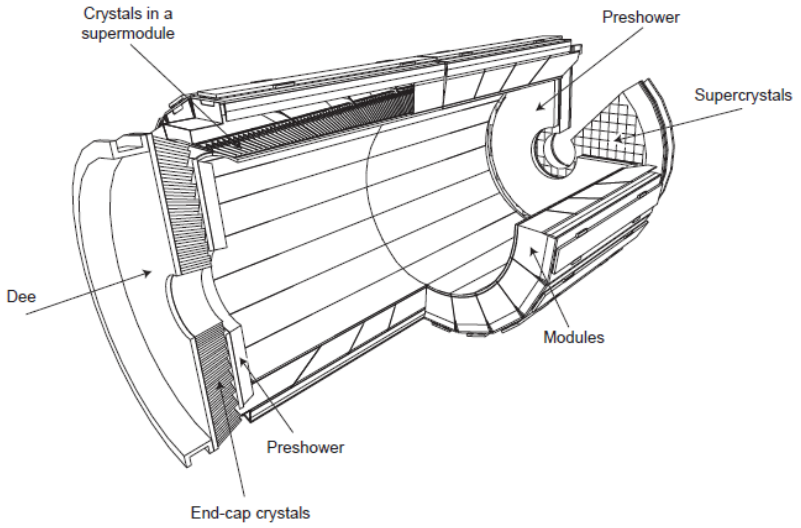

Figure 2.3: CMS ECAL diagram [76].

region $|\eta| > 1.3$ and consists of 36 azimuthal wedges assembled into two half barrels (HB+ and HB-). HE covers the range $1.3 < |\eta| < 3.0$. The HF extends the coverage to $3.0 < |\eta| < 5.0$, after that only additional scintillators from HO are located outside of the solenoid and act as "tail catchers", effectively increasing the thickness of the calorimeter in the central region.

The HB and HE calorimeters are sampling calorimeters that use alternating plates of brass as the absorber and plastic scintillator as the active material. Wavelength from the light coming from the plastic scintillator is shifted and captured in optical fibers for transport to the photo-transducers and front-end electronics.

The HF subdetector is a Cherenkov calorimeter base on a 165 cm steel absorber and quartz fibers which run longitudinalñy through the absorber and collect Cherenkov radiation. There are two types of fibers within the HF: long fibers which span the length of the HF, and short fibers as short as 22 cm. The differences between signals received from the long and short fibers are used to distinguish between electrons/photons and hadrons. PMTs (photomultiplier tubes) are connected to these fibers so the energy can be reconstructed through digitalized charge measurements. The quartz fibers are inserted into the HF with a spacing of 5 mm and the fibers associated with a particular $\eta \times \phi$ region are bundled.

Finally, the HO accomplishes its tail-catcher objective (for hadronic showers helps for muon identification) using the same active material and WLS fiber as the HB and

HE calorimeters, but uses the steel return yoke and magnet material of CMS as the absorber.

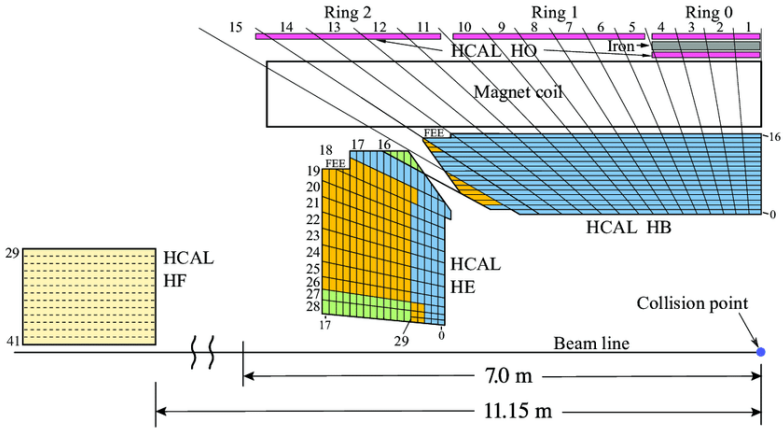

Figure 2.4: CMS HCAL r-Z schematic drawing of a quarter of the CMS detector locating the HB, HE, HO and HF calorimeters in CMS [77]

2.2.5 Muon System

Detecting muons is one of the most important tasks of CMS because muons are expected to be produced in the decay of a number of potential new particles and yield the clearest signatures for important decays. Because muons can penetrate several meters of material losing little energy (unlike most particles), they are not stopped by any of the CMS calorimeters, and therefore the detectors are placed in the outer part of the experiment where they are the only particles expected to produce a clear signal. Muons are measured by fitting a curve to the hits registered in the four muon stations. As mentioned before, the path is measured by tracking its position through the multiple layers of each station. This information is combined with the measurements taken with the CMS silicon tracker. The muon system covers a range of $|\eta| < 2.4$, and it is composed of gaseous detector systems interposed among layers of steel (return yoke), allowing a flying muon to be detected at multiple points such that the particle path can be measured (Figure 2.5). The CMS muon system was designed with physics considerations that lead to the following list of guidelines for the muon trigger design:

- Flexibility: In order to access all the interesting physics channels and to tune the rate to a level acceptable for the Event Filter, the p_T threshold must be adjustable, so a range between 2-100 GeV is covered.

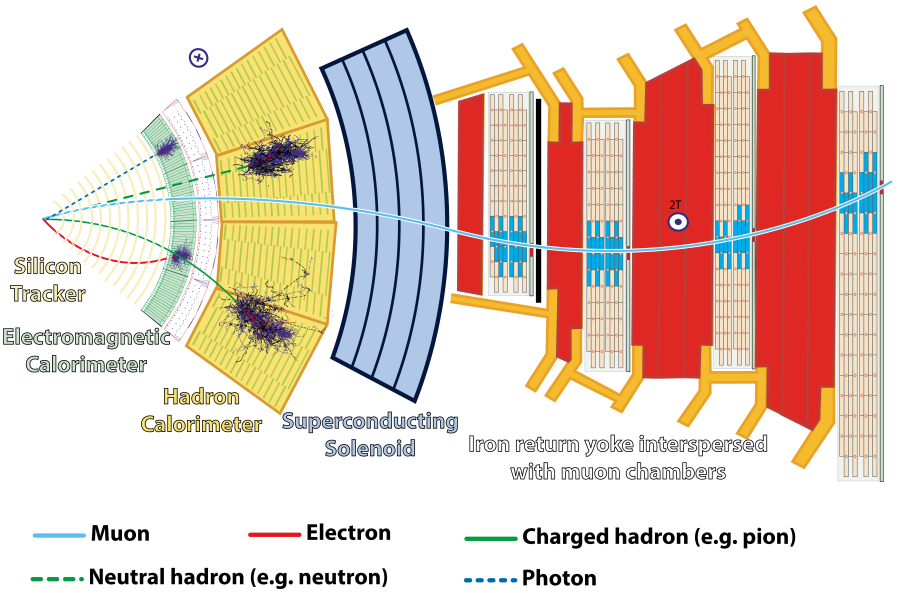

Figure 2.5: Different particles in the plane perpendicular to the LHC beams. The muon leaves a curved trajectory in four layers of muon detectors [25].

- Time resolution: The first Trigger Level (L1) must be able to assign an event to the proper bunch crossing, thus the time resolution must be smaller than the bunch crossing interval, i.e. 25 ns.

- Speed: The trigger decision must be available in about 3 μs after the collision. This includes the propagation time of the signals from the detector to the control room (∼120 m) and back.

- High acceptance: Searches for rare events (such as the ones studied in this thesis) require an optimal acceptance close to 100 %. Therefore, the muon stations are arranged in such a way that most of the tracks cross 4 triggering plates, but every track crosses at least 3 triggering planes.

- Redundancy: The trigger system has to deal with all the possible inefficiencies, noise, pileup, and background from muon radiation. Thus it has to have substantial redundancy. In CMS this is ensured by having two complementary subsystems (which will be described in more detail below), one based on fast dedicated trigger detectors, RPCs, and the other using precise multilayer muon chambers, DTs and CSCs. For important low p_T (3-5 GeV) muon physics, such

as B-physics and heavy ion collisions, a double system of RPCs is included in the two innermost muon stations on the barrel region.

The muon system uses three different technologies to detect and measure the muons: drift tubes (DT) in the barrel region, cathode strip chambers (CSC) in the endcap region, and resistive plate chambers (RPC) in both the barrel and endcap regions [60]. A muon trigger in the barrel region is generated using a mean-timer to identify patterns. In the endcap, the trigger is generated from the cathode readout patterns and the wire timing. For both, barrel and endcap, the RPCs provide an additional trigger signal which has a different sensitivity to backgrounds. All the muon chambers are aligned roughly perpendicular to the muon trajectories and distributed to provide hermetic coverage over the $0.0 < |\eta| < 2.4$ range. The barrel DTs cover roughly $0.0 < |\eta| < 1.3$, while the endcap CSCs cover $0.9 < |\eta| < 2.4$. The RPCs cover the region $0.0 < |\eta| < 2.1$.

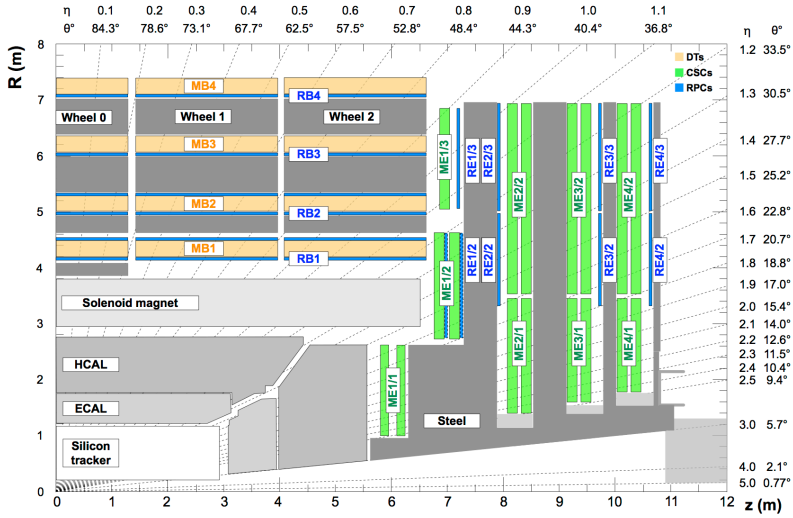

Figure 2.6: R-z cross-section of a quadrant of the CMS muon system, showing DT (yellow), CSC (green), and RPC chambers (light blue). The interaction point is at the lower left corner

Chapter 3

Measurement of $b-$hadron production fraction ratios f_s/f_u and f_d/f_u

3.1 Introduction

Precise knowledge of the hadron production fractions is essential for measuring b-hadron decay rates at the LHC. It has been extensively discussed in the Introduction Chap. 1 why b-hadron decays are so relevant to get precise tests to the SM and glimpses of new physics. Given the current knowledge of the $b\bar{b}$ cross-section at the LHC, any direct measurement of $\mathrm{B}_s \to \mu^+\mu^-$ and $\mathrm{B}^0 \to \mu^+\mu^-$ would have a large uncertainty. The usual fix to this problem is to calculate the signal branching fraction by normalizing to another decay channel of the B meson, whose branching fraction is well known and whose final state has a clear signal. The $B^+ \to J/\psi K^+$ is the best candidate for this task, so the signal branching fraction can be extracted as:

$$\mathcal{B}(\mathrm{B}_s \to \mu^+\mu^-) = \frac{N_{\mathrm{B}_s\to\mu^+\mu^-}}{N_{\mathrm{B}^+}} \frac{f_u}{f_s} \frac{\epsilon_{\mathrm{B}^+}}{\epsilon_{\mathrm{B}_s\to\mu^+\mu^-}} \times \mathcal{B}(B^+ \to J/\psi K^+) \times \mathcal{B}(J/\psi \to \mu^+\mu^-), \tag{3.1}$$

and

$$\mathcal{B}(\mathrm{B}^0 \to \mu^+\mu^-) = \frac{N^{\mathrm{B}_s \to \mu^+\mu^-}}{N^{\mathrm{B}^0}} \frac{f_u}{f_d} \frac{\epsilon^{\mathrm{B}^0}}{\epsilon^{\mathrm{B}^0 \to \mu^+\mu^-}} \times \mathcal{B}(\mathrm{B}^0 \to J/\psi\, \mathrm{K}^{*0}) \times \mathcal{B}(J/\psi \to \mu^+\mu^-), \tag{3.2}$$

where N_x is the number of fitted candidates for the decay x and ϵ_x is the corresponding full selection efficiency.

In particular, one of the main systematic uncertainties affecting the measurement of the branching fraction is the uncertainty in the relative production yields of the $\mathrm{B_s}^0$, B^0, and B^+ mesons, which is directly related to the ratio of the respective hadron fractions, $f_\mathrm{s}/f_\mathrm{u}$ and $f_\mathrm{d}/f_\mathrm{u}$. Refs. [35, 1, 71, 7, 8]:

The relative abundances of b-hadrons (B^+, B^0, $\mathrm{B_s}^0$, and Λ_b^0) have been measured at LEP in electron-positron collisions [16, 30, 15, 14] and at the Tevatron in proton-antiproton collisions [13], the two sets of results being consistent with each other [20] and with theoretical expectations [61].

More recently, the LHCb Collaboration, using large event samples of proton-proton (pp) collisions at $\sqrt{s} = 7$, 8, and 13 TeV, was able to study how the b-hadron abundances depend on their transverse momentum (p_T). In particular, the b-baryon fraction (f_baryon, derived from the Λ_b^0 yields) decreases by around a factor of 2, with respect to the B^+ and B^0 fractions, from 5 to 25 GeV, a clear observation of a p_T dependence of a b-hadron fraction [3].

Considering that $f_u + f_d + f_s + f_\mathrm{baryon} = 1$, neglecting the very small contributions of b-hadrons made of several heavy quarks, it is reasonable to expect that also the B meson fractions are p_T dependent. Indeed, LHCb has also reported a significant p_T dependence of the ratio between the $\mathrm{B_s}^0$ and B^+ fragmentation fractions [9, 11], in the 2.0–4.5 rapidity (y) range and at relatively low p_T . On the other hand, the ATLAS Collaboration reported a measurement consistent with no p_T dependence, although the study was performed in different regions of p_T and (y) [2].

Our analysis aims at establishing if and how the relative $\mathrm{B_s}^0$ and B^+ production rates change with p_T in a kinematic region relevant for the ATLAS and CMS experiments at the CERN LHC, $p_\mathrm{T} > 12$ GeV and $|y| < 2.4$, approximately complementary to that of the LHCb detector. Additionally, it is the first test of the isospin invariance in B meson production at hadron colliders, measuring the $f_\mathrm{d}/f_\mathrm{u}$ ratio in the same kinematic regime. The measurement uses a sample of pp collisions collected by the CMS experiment in 2018 at a center-of-mass energy of 13 TeV and corresponding to an integrated luminosity of 61.6 fb^{-1} [37, 70].

Charge-conjugate states are implicitly included, and K^{*0} and ϕ represent the

$K^{*0}(892)$ and $\phi(1020)$, respectively. The B^+ and ${B_s}^0$ mesons are reconstructed using the $B^+ \to J/\psi\, K^+$ and ${B_s}^0 \to J/\psi\, \phi(1020)$ decay channels, with the J/ψ and ϕ mesons decaying as $J/\psi \to \mu^+\mu^-$ and $\phi(1020) \to K^+K^-$. The respective event yields, N_{B^+} and $N_{B_s{}^0}$, are measured with corresponding detection efficiencies ϵ_{B^+} and $\epsilon_{B_s{}^0}$. The ratio of the efficiency-corrected meson yields, $\mathcal{R}_s = (N_{B_s{}^0}/\epsilon_{B_s{}^0}) \,/\, (N_{B^+}/\epsilon_{B^+})$, is directly proportional to the f_s/f_u ratio,

$$\mathcal{R}_s = \frac{f_s}{f_u} \frac{\mathcal{B}({B_s}^0 \to J/\psi\, \phi(1020))\mathcal{B}(\phi(1020) \to K^+K^-)}{\mathcal{B}(B^+ \to J/\psi\, K^+)} = \frac{N_{B_s{}^0}}{N_{B^+}} \frac{\epsilon_{B^+}}{\epsilon_{B_s{}^0}}, \tag{3.3}$$

where $\mathcal{B}({B_s}^0 \to J/\psi\, \phi(1020))$, $\mathcal{B}(\phi(1020) \to K^+K^-)$, and $\mathcal{B}(B^+ \to J/\psi\, K^+)$ are the branching fractions of the indicated decay channels; the $\mathcal{B}(J/\psi \to \mu^+\mu^-)$ factor cancels in the ratio. Given that the available measurements of the ${B_s}^0 \to J/\psi\, \phi(1020)$ branching fraction and of f_s are correlated, we report measurements of $\mathcal{R}_s$ rather than of f_s/f_u .

Similarly the $\mathcal{R}_d$ ratio is defined as,

$$\mathcal{R}_d = \frac{f_d}{f_u} \frac{\mathcal{B}(B^0 \to J/\psi\, K^{*0}) \times \mathcal{B}(K^{*0} \to \pi^+K^-)}{\mathcal{B}(B^+ \to J/\psi\, K^+)} = \frac{N_{B^0}}{B_{B^+}} \frac{\epsilon_{B^+}}{\epsilon_{B^0}} \tag{3.4}$$

3.2 Branching fractions

The final states involving this measurement are well-known and studied. Their world-average branching fractions will be used in this analysis, summarized in table 3.1. For the case of the $K^{*0} \to \pi^+K^-$ channel, the branching fraction is estimated from the P.D.G. [51], ($\mathcal{B}(K^{*0} \to \pi^+K^-) = 99.754 \pm 0.021\%$), which takes into account all possible final decay products, including the one where both the kaon and the pion are neutral, which CMS can not detect. Assuming isospin symmetry, a value around 2/3 of the referenced value is used as our estimation for the charged decays. Note that the quoted K^{*0} branching value it is not exactly 66.67% since the additional $K^{*0} \to K^0\gamma$ decay ($\mathcal{B}(K^{*0} \to K^0\gamma) = (2.46 \pm 0.21) \times 10^{-3}$) also contributes to the total decay rate.

3.3 Data and Montecarlo samples

3.3.1 Trigger description

This measurement was made on "Run 2" pp collisions at $\sqrt{s} = 13$ TeV, collected during 2018, using the `HLT_DoubleMu4_JpsiTrk_Displaced` trigger. This high-level trigger

Variable	Value
$\mathcal{B}(B_s^0 \to J/\psi\ \phi(1020))$	$(1.08 \pm 0.08) \cdot 10^{-3}$
$\mathcal{B}(B^+ \to J/\psi\ K^+)$	$(1.026 \pm 0.031) \cdot 10^{-3}$
$\mathcal{B}(B^0 \to J/\psi\ K^{*0})$	$(1.27 \pm 0.05) \cdot 10^{-3}$
$\mathcal{B}(\phi(1020) \to K^+K^-)$	$(48.9 \pm 0.5)\%$
$\mathcal{B}(K^{*0} \to K^+\ \pi^-)$	$(66.503 \pm 0.014)\%$

Table 3.1: Relevant branching fractions. Values taken from Ref. [51]

(HLT) requires two opposite charged muons within $|\eta| < 2.5$ and transverse momentum $p_{\mathrm{T}} > 4.0$ GeV, a distance of closest approach (between the two muons) smaller than 0.5 cm, dimuon vertex fit χ^2 probability larger than 10%, dimuon invariant mass in the range $2.9 - 3.3$ GeV, and transverse distance between the dimuon vertex and the beam axis larger than three times its uncertainty. Additionally, the dimuon p_{T} must be aligned with the transverse vector, $\cos\theta > 0.9$, where $\cos\theta = \vec{L}_{xy} \cdot \vec{p}_{\mathrm{T}}/(L_{xy}\ p_{\mathrm{T}})$, where $\vec{L}_{xy}$ is the transverse decay displacement vector of the dimuon computed with respect to the beam spot Fig. 3.1 (note that this angle is a different angle θ, with respect to the one used in the CMS coordinate system). Finally, there must exist a third track in the event compatible with being produced at the dimuon vertex. The HLT path requires this track with $p_{\mathrm{T}} > 1.2$ GeV, an impact parameter significance $d_{xy}/\sigma_{d_{xy}}$ greater than 2, and a three-object vertex fit χ^2 probability larger than 10%.

During the 2018 data taking two main versions of the trigger were used. Table 3.2 expose the details.

Year	HLT path	L_{int} (fb^{-1})	
		Recorded	Certified
2018	`HLT_DoubleMu4_JpsiTrk_Displaced_v14`	9.4	9.1
	`HLT_DoubleMu4_JpsiTrk_Displaced_v15`	54.7	52.5
	Total luminosity		61.6

Table 3.2: HLT Trigger paths. The respective recorded and integrated luminosities (L_{int}) for certified data, as computed with the `brilcalc` tool [28]. The list of certified runs and lumisections were taken from the JSON file `Cert_314472-325175_PromptReco_Collisions18_JSON_MuonPhys`.

3.3.2 Generation Software

For the efficiency studies and analysis selection, samples were generated with `PYTHIA` 8 [73] for the production and hadronization processes. After that, `EVTGEN` [59] was used to

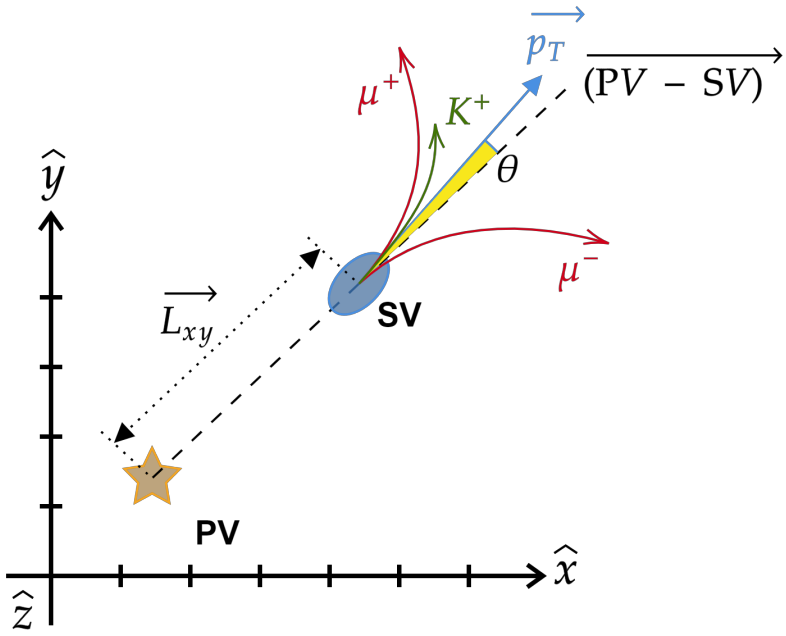

Figure 3.1: Schematic picture of how the pointing angle, θ is calculated in the $B^+ \to J/\psi\, K^+$ channel, the idea holds for all the other channels involving two different vertices. It is basically the angle between the PV-SV line and the $b-$hadron p_T.

simulate the decay of the b hadrons. The QED final state radiation is modeled with `PHOTOS` [44]. The generated events were processed through a simulation of the CMS detector, using the `GEANT4` package [17], with the exact same trigger and reconstruction algorithms used in the recorded data. All the simulated events include several proton-proton interactions in the same or nearby beam crossings (called pileup), with a distribution matching the one observed in the data. These Monte-Carlo samples are public and can be found on the CMS Data Acquisition System (DAS) as:

- `BsToJpsiPhi_BMuonFilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen` (5 M generated events).
- `BuToJpsiK_BMuonFilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen` (10 M generated events).
- `BdToJpsiKstar_BMuonFilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen` (5 M generated events).

To measure the generation prefilter efficiency (which is a filter used to remove

events that will not be reconstructed due to the CMS acceptance), two samples were used: one without any requirements, and a second one with the following generation cuts:

- $|\eta(\mu)| < 2.5$;
- $|\eta(K)| < 2.5$;
- $p_{\mathrm{T}}(\mu) > 3.8$; GeV
- $p_{\mathrm{T}}(K) > 0.5$; GeV

The acceptance (the efficiency for this specific generation prefilter) was calculated as the ratio of these two MC samples in bins of η and p_{T}.

3.4 Reconstruction and Selection

The b-hadron candidates were reconstructed with two tracks (only one in the case of B^+) and a pair of muons in all the studied channels. The dimuon pair must fulfill a set of requirements that will be discussed in detail in this section, but they basically aim for a J/ψ resonance.

3.4.1 J/ψ Selection

The J/ψ quarkonium states were reconstructed taking advantage of the standard CMS muon reconstruction. The approved Muon-POG (POG: Physics Object Group) `Soft Muon` selection [69] uses the following requirements. The track identified by the tracker is matched with at least one segment in any muon station, in both x and y coordinates (`TMOneStationTight`), and arbitrated[1]. The track needs to qualify as `high-purity` [32] and must include more than five hits in the tracker layers, at least one of them being in a pixel layer. The impact parameter of the single muons with respect to the beam-spot must be smaller than 0.3 cm in the transverse plane and smaller than 20 cm along the beam axis. The single muons also need to have $|\eta| < 2.4$ and, a minimum transverse momentum of $p_{\mathrm{T}} > 4$ GeV. In the offline reconstruction, it is required that both opposite charged muons are compatible with the muons that triggered the detector readout, a process called trigger matching.

[1] Tracker muon arbitration is the process where the ambiguity of sharing segments is resolved by picking the best matching based on position and pull requirements.

Additionally, a *Kinematic Vertex Fit* (see appendix B) is performed to improve the mass resolution of the dimuon resonance, after this it is required that the fit converges and gives a probability greater than 1%. The now fitted invariant mass of the dimuon candidate must be within the $[2.9, 3.3]$ GeV window (from now on, the invariant mass of the dimuon or J/ψ candidate will be the one obtained from the *Kinematic Vertex Fit*).

3.4.2 b-hadron reconstruction

From the sample of collected J/ψ events, candidate b-hadrons are reconstructed by pairing the J/ψ candidate (dimuon pair) with extra tracks, depending on the decay mode. To simplify things, hereafter these last tracks are denoted as K_1 for $B^+ \to J/\psi\, K^+$ decay mode, and K_2 and K_3 for ${B_s}^0 \to J/\psi\, \phi(1020)$ decay mode. The charge of the K_2 and K_3 tracks are required to be opposite. As in the case of the muon, a trigger matching process is applied for K_1 and K_2 (or K_3), to ensure they fired the trigger. For K_1 and K_2, an impact parameter significance dxy/σ_{dxy} greater than 2 is required. Only tracks that pass the `high-purity` [32] selection are used.

The charged hadron tracks, K_1, K_2, and K_3 are required to have at least 1-pixel hit, at least 5-tracker hits (strips and pixels together), and must have $|\eta| < 2.4$. The p_T of the charged hadron tracks K_1, K_2, and K_3 are required to be greater than 1.2 GeV.

The b-hadron candidate is fitted to a common vertex with the appropriate masses assigned to the charged tracks and the dimuon invariant mass constrained to the world average J/ψ mass [51](using a *Kinnematic Vertex Fit*, see appendix B). In the case of the ${B_s}^0 \to J/\psi\, \phi(1020)$ decay mode, the intermediate candidate state $\phi(1020)$ is selected if $|M(K^+K^-) - m^{\mathrm{PDG}}_{\phi(1020)}| < 10$ MeV. That corresponds to 2 times the experimental resolution or natural width around the nominal $\phi(1020)$ mass.

Primary collision vertices (PVs) are fitted from the reconstructed tracks. As expected, there are several PVs per bunch crossing. The best PV, is chosen based on the pointing angle between the line formed by the b-hadron vertex (or secondary vertex SV) and the PV (see Fig. 3.1), and the line formed by the b-hadron momentum, so the PV that gives the best pointing angle (closest to zero) it is selected as the b-candidate primary vertex. The decay length of the b-hadron, denoted by l, is computed as the distance between the best PV and the b-hadron reconstructed vertex (assumed to be, respectively, the b-meson production and decay vertices). The chosen PV is refitted removing all tracks from the b-candidate if they were present. Only b-hadron candi-

dates, with p_{T} in the range of $12 < p_{\mathrm{T}} < 70$ GeV and a rapidity between $|y| < 2.4$, a decay length larger than five times its uncertainty, and a vertex χ^2 probability greater than 10% are selected.

There are a few (<1%) events in which more than one b-hadron candidate is found. In such cases, only the candidate with the highest χ^2 vertex fit probability is retained.

3.5 Signal Extraction

The $\mathrm{B_s}^0$, B^+ and B^0 yields are extracted using unbinned maximum likelihood estimations to the reconstructed $\mathrm{B_s}^0 \to J/\psi\,\phi(1020)$, $\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+$ and $\mathrm{B}^0 \to J/\psi\,\mathrm{K}^{*0}$ invariant mass distributions, respectively. The signal's nominal fit function is composed of the sum of two Gaussian functions with a single mean and different widths. The mean and widths are open parameters in the fit. The two Gaussian functions have different individual widths, contributing to the total yield with some fractions so that the b-mass resolution is computed as the (weighted) sum in quadrature of the signal components.

For the $\mathrm{B_s}^0 \to J/\psi\,\phi(1020)$ channel, the nominal fit model for the combinatorial background is an exponential distribution.

In the case of the $\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+$ decay, it is necessary to introduce a couple of additional background contributions. Partially reconstructed $\mathrm{B}^+ \to J/\psi K^+ X$ decays that contribute to the invariant mass distribution, are modeled with an error function with floating parameters in the fit. In addition, a misidentification of a pion as a kaon may occur, and a small contribution from $\mathrm{B}^+ \to J/\psi\pi$ decays is present, with a shape fixed from simulation studies and a normalization fixed by the $\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+$ yield, scaled by the ratio of the corresponding branching fractions [51].

For the $\mathrm{B_s}^0 \to J/\psi\,\phi(1020)$ measurement other possible sources of peaking background as $\Lambda_b \to J/\psi Kp$ or $B^0 \to J/\psi K\pi$ were considered. No evidence of a sizable peaking background from $\Lambda_b \to J/\psi Kp$ was found. However, the $B^0 \to J/\psi K\pi$ contribution was modeled taking as reference a previous B physics analysis [72], using a Johnson function [52] whose parameters are fixed from MC.

In the case of the B^0 channel, which has the same topology as the $\mathrm{B_s}^0$ channel (J/ψ plus two tracks), the main difference is the K^{*0} meson. In the case of the $\mathrm{B}^0 \to J/\psi\,\mathrm{K}^{*0}$ decay mode, the intermediate candidate state K^{*0} is selected if $|M(\pi^+\mathrm{K}^-) - m^{PDG}_{\mathrm{K}^{*0}}| <$ 50 MeV.

Additionally, tge $\mathrm{K}^{*0} \to \pi^+\mathrm{K}^-$ decay may be reconstructed with swapped hadron mass assignments: either the first track is a K and the second is a π, or the other

way around. The combination whose invariant mass lies closest to the K^{*0} nominal mass is selected. Using MC samples with event selection applied as in data, for each of the p_T and rapidity bins, the fraction of candidates where this criteria results in the incorrect assignment is determined. These candidates are called swapped. The swapped component shape and amount are fixed from simulation studies for each of the p_T and rapidity bins. The resulting swapped yield is included in the total B^0 signal yield.

The $B^0 \to J/\psi\, K^{*0}$ distribution has a couple of additional background contributions: the Cabibbo Suppressed $B_s^0 \to J/\psi K^{*0}$ component (small peak at 5.367 GeV) is described by a double Gaussian function with widths parameters equal to those of the B^0 peak, scaled by the ratio of effective resolutions found in MC: $\sigma_{effMC}^{B_s^0}/\sigma_{effMC}^{B^0}$. The B_s^0 and B^0 signals share the fraction of the first Gaussian parameter.

Another source of peaking background is $B^0 \to J/\psi K\pi$, where $K\pi$ is not resonant. This contribution is modeled with a sum of a double Crystal Ball and Gaussian functions, whose parameters were fixed from MC. All background sources were considered for each bin of transverse momentum and rapidity. This background component is included in the fit model with the yield fraction defined relative to the signal contribution and Gaussian constrained to the value determined on the total phase space fit.

The dataset is explored both with respect to transverse momentum and rapidity. The binning was chosen to have enough data for a robust maximum likelihood fit estimation and to have similar uncertainties (same statistical significance). The fitted distributions and the results of each fit are shown in the next subsections.

3.5.1 Distributions on Transverse Momentum Bins

b-hadron yields are extracted in bins of p_T, defined by the boundaries of 12, 13, 14, 15, 16, 18, 20, 23, 26, 29, 34, 45, and 70 GeV.

In figure 3.2, the $B_s^0 \to J/\psi\,\phi(1020)$ mass distribution in each p_T bin is displayed. Similarly, figure 3.3 shows $B^+ \to J/\psi\, K^+$. The quality of the fits can be quantified through the χ^2 per degrees of freedom ratio available in the plots. Figures 3.2 and 3.4 show the corresponding plots for the $B_s^0 \to J/\psi\,\phi(1020)$ and $B^0 \to J/\psi\, K^{*0}$ channels.

3.5.2 Distributions on Rapidity Bins

Rapidity boundaries are chosen: 0.0, 0.25, 0.50, 0.75, 1.00, 1.30, 1.60, and 2.40. The distributions for the $B_s^0 \to J/\psi\,\phi(1020)$ candidates are shown in figure 3.5. The

3.5. Signal Extraction

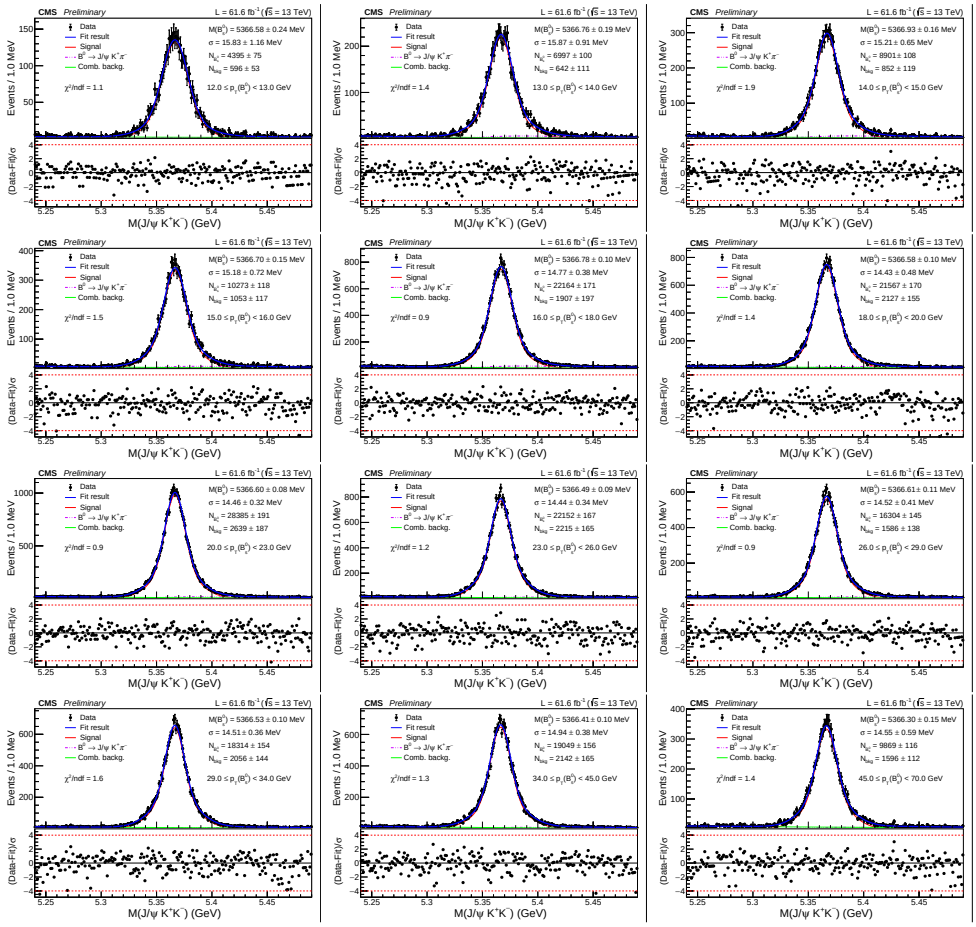

Figure 3.2: Nominal fits performed to the B_s^0 candidates in transverse momentum bins. The signal component is fitted to the sum of two Gaussian functions with a single mean and different widths. The combinatorial background is modeled with an exponential distribution. The $B^0 \to J/\psi K\pi$ contribution was modelled with a Johnson function fixed from MC.

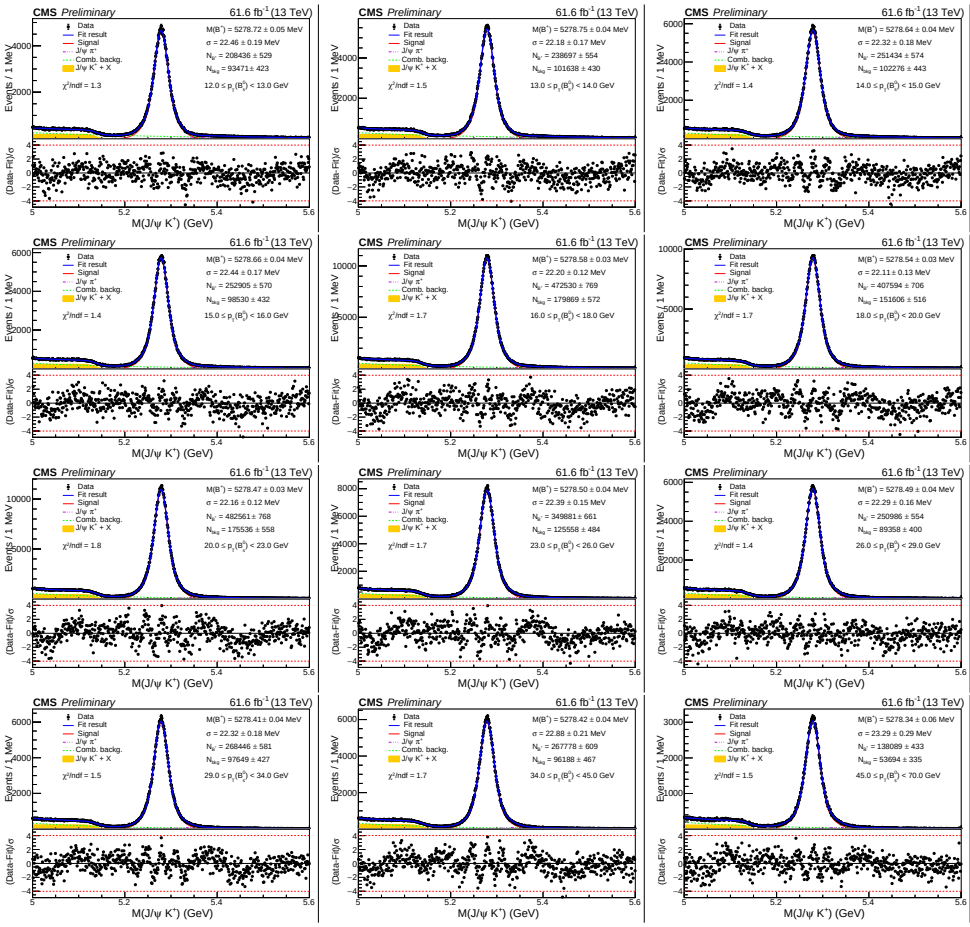

Figure 3.3: Nominal fits performed to the B^+ candidates in transverse momentum bins. The signal is modeled with the sum of two Gaussian functions with a single mean and different widths. The combinatorial background component is fit to an exponential distribution. An error function models the contribution of partially reconstructed decays. The $B^+ \to J/\psi\pi^+$ component is fixed from simulation studies.

3.5. Signal Extraction

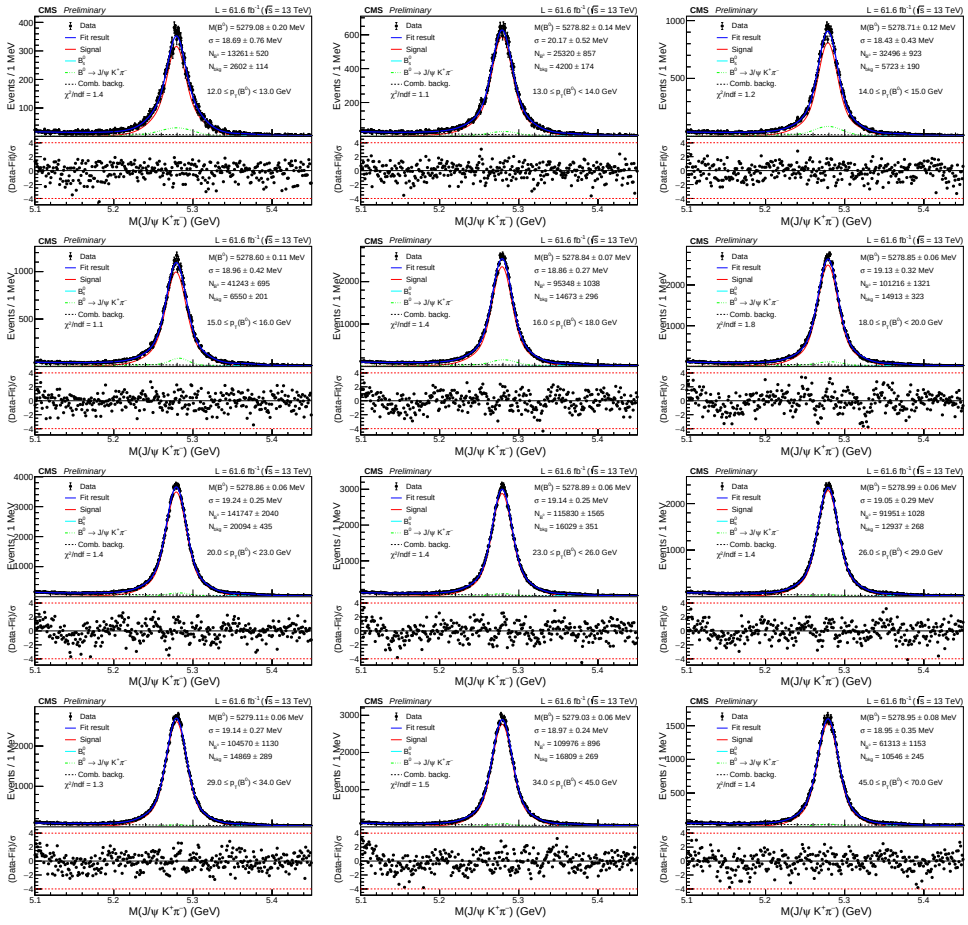

Figure 3.4: Nominal fits performed to the B^0 candidates in transverse momentum bins. The signal component is fitted to the sum of two Gaussian functions with a single mean and different widths. The combinatorial background is modeled with an exponential distribution. The $B_s{}^0$ and swapped contributions were modeled as described in the text.

corresponding distributions for $B^+ \to J/\psi\, K^+$ channel are shown in figure 3.6, and in figure 3.7 for $B^0 \to J/\psi\, K^{*0}$.

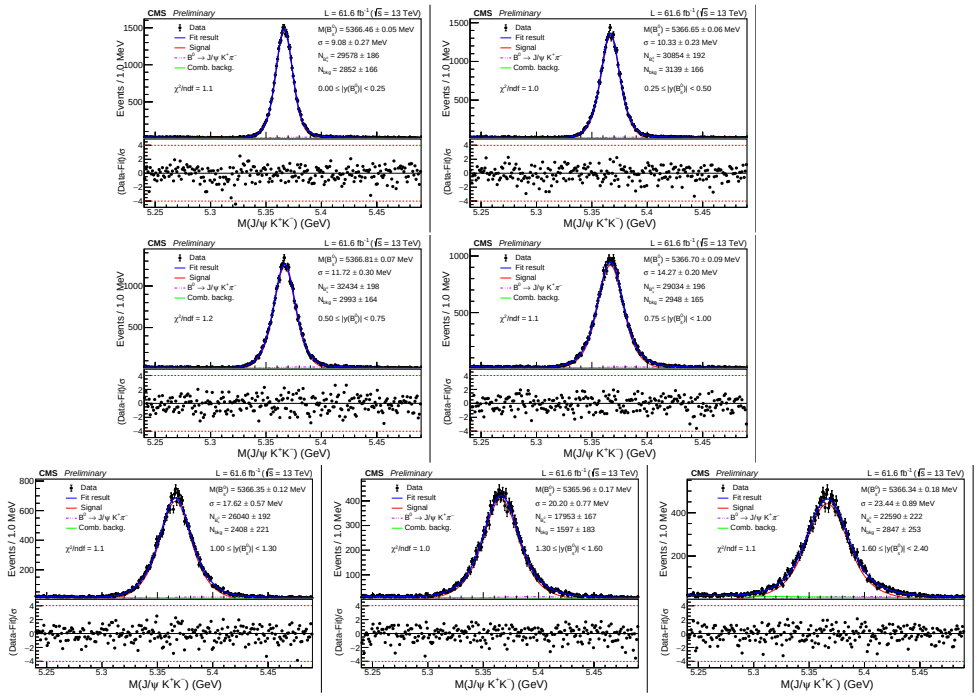

Figure 3.5: Nominal fits performed to the $B_s{}^0$ candidates in rapidity bins. The signal component is fitted to the sum of two Gaussian functions with a single mean and different widths. The combinatorial background is modeled with an exponential distribution. The $B^0 \to J/\psi K\pi$ contribution was modelled with a Johnson function fixed from MC.

3.6 Efficiency Determination

Efficiency is the number of reconstructed B events in MC after the complete selection, divided by the number of generated B decays in the fiducial region of the analysis specified by the B kinematic window($12 < p_{\mathrm{T}}(B) < 70$ GeV and $|y(B)| < 2.4$). The efficiency includes both the MC acceptance and offline selection. This strategy requires the use of two MC samples per channel:

- A sample with no generation prefilters.

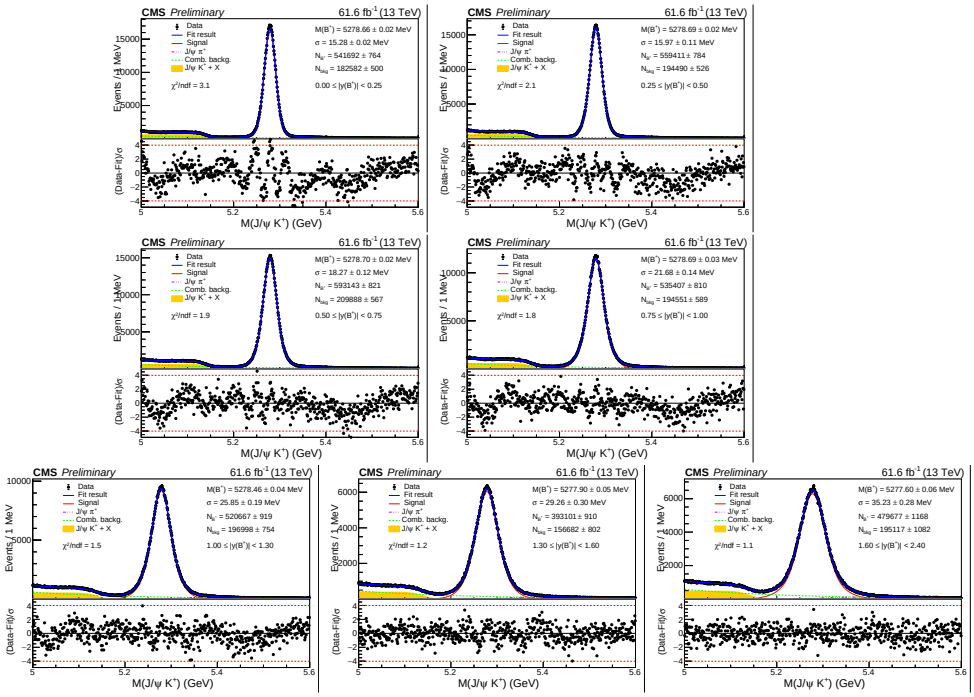

Figure 3.6: Nominal fits performed to the B^+ candidates in rapidity bins. The signal is modeled with the sum of two Gaussian functions with a single mean and different widths. The combinatorial background component is fitted with an exponential distribution. An error function models the contribution of partially reconstructed decays. The $B^+ \to J/\psi\pi$ component is fixed from simulation studies.

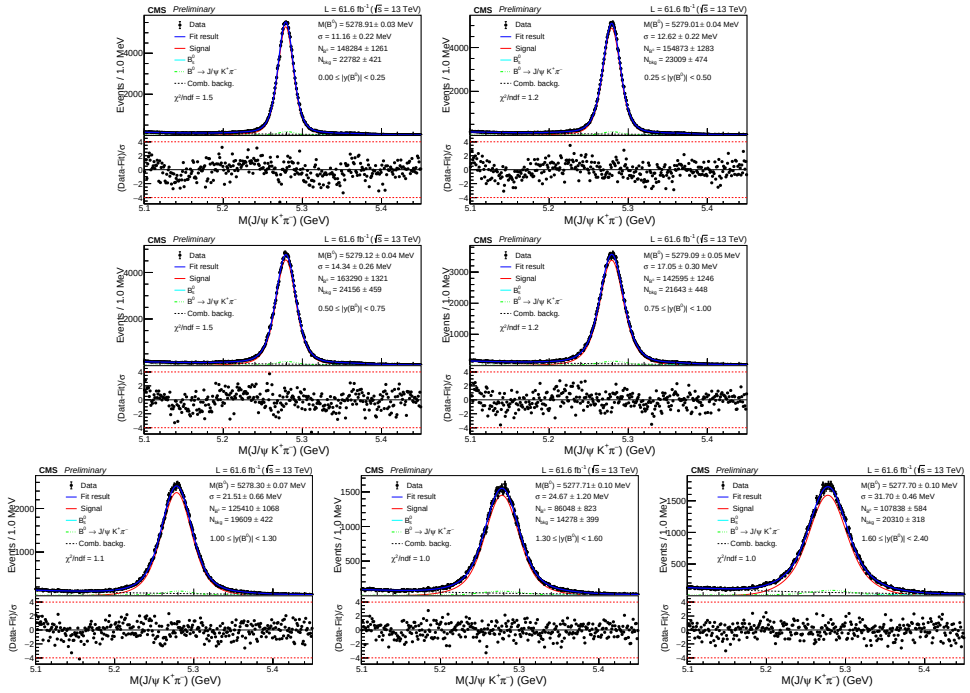

Figure 3.7: Nominal fits performed to the B^0 candidates in rapidity bins. The signal component is fitted to the sum of two Gaussian functions with a single mean and different widths. The combinatorial background is modeled with an exponential distribution. The $B_s{}^0$ and swapped contributions were modeled as described in the text.

- A second sample with generator filters: $|\eta(\mu)| < 2.5$, $|\eta(K)| < 2.5$, $p_{\mathrm{T}}(\mu) > 3.8$ GeV and $p_{\mathrm{T}}(K) > 0.5$ GeV (which is the fiducial region of the analysis).

The efficiency (ϵ) is split into two terms: the pre-filter efficiency (ϵ_1) and the reconstruction efficiency (ϵ_2). The multiplication of this pair of efficiencies ($\epsilon = \epsilon_1 \times \epsilon_2$) is regarded as the total efficiency and it is computed bin per bin (both in p_{T} and rapidity ($|y|$) matching the analysis regions).

The number of events passing the generation filter cuts divided by the number of events generated is regarded as the acceptance or pre-filter efficiency ϵ_1. The efficiency per p_{T} bin i is defined as:

$$\epsilon_1^i = \frac{N(\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+; |y^{\mathrm{gen}}_{\mathrm{B}^+}| < 2.4, c < p^{\mathrm{gen}}_{\mathrm{T}}(\mathrm{B}^+) < d, \text{filter cuts})}{N(\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+; |y^{gen}_{\mathrm{B}^+}| < 2.4, c < p^{\mathrm{gen}}_{\mathrm{T}}(\mathrm{B}^+) < d)}. \tag{3.5}$$

To calculate the efficiency as a function of $|y^B|$, a similar procedure is followed. The efficiency for rapidity bin j, in $c < |y_{\mathrm{B}^+}| < d$, is defined as:

$$\epsilon_1^j = \frac{N(\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+; 12.0 < p^{\mathrm{gen}}_{\mathrm{T}}(\mathrm{B}^+) < 70.0, c < |y^{\mathrm{gen}}_{\mathrm{B}^+}| < d, \text{filter cuts})}{N(\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+; 12.0 < p^{\mathrm{gen}}_{\mathrm{T}}(\mathrm{B}^+) < 70.0, c < |y^{\mathrm{gen}}_{\mathrm{B}^+}| < d)}. \tag{3.6}$$

The efficiency shown in these formulas is for the $\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+$ channel. The same idea is applied to ${\mathrm{B_s}}^0 \to J/\psi\,\phi(1020)$ and $\mathrm{B}^0 \to J/\psi\,\mathrm{K}^{*0}$. The corresponding results are shown in figure 3.8.

The second component of the total efficiency (ϵ_2) is the reconstruction efficiency, estimated from MC samples that are generated with pre-filter. It is defined as the ratio of the number of reconstructed events after the complete selection with respect to the number of generated decays in the appropriate binning. For p_{T} bin i, $c < p_{\mathrm{T}}(\mathrm{B}^+) < d$, and rapidity bin j, $c < |y_{\mathrm{B}^+}| < d$,

$$\epsilon_2^i = \frac{N(\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+, |y^{\mathrm{reco}}_{\mathrm{B}^+}| < 2.4, c < p^{\mathrm{reco}}_{\mathrm{T}}(\mathrm{B}^+) < d, \text{filter cuts, full selection})}{N(\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+, |y^{\mathrm{gen}}_{\mathrm{B}^+}| < 2.4, c < p^{\mathrm{gen}}_{\mathrm{T}}(\mathrm{B}^+) < d, \text{filter cuts})}, \tag{3.7}$$

$$\epsilon_2^j = \frac{N(\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+, 12.0 < p^{\mathrm{reco}}_{\mathrm{T}}(\mathrm{B}^+) < 70.0, c < |y^{\mathrm{reco}}_{\mathrm{B}^+}| < d, \text{filter cuts, full selection})}{N(\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+, 12.0 < p^{\mathrm{gen}}_{\mathrm{T}}(\mathrm{B}^+) < 70.0, c < |y^{gen}_{\mathrm{B}^+}| < d, \text{filter cuts})}. \tag{3.8}$$

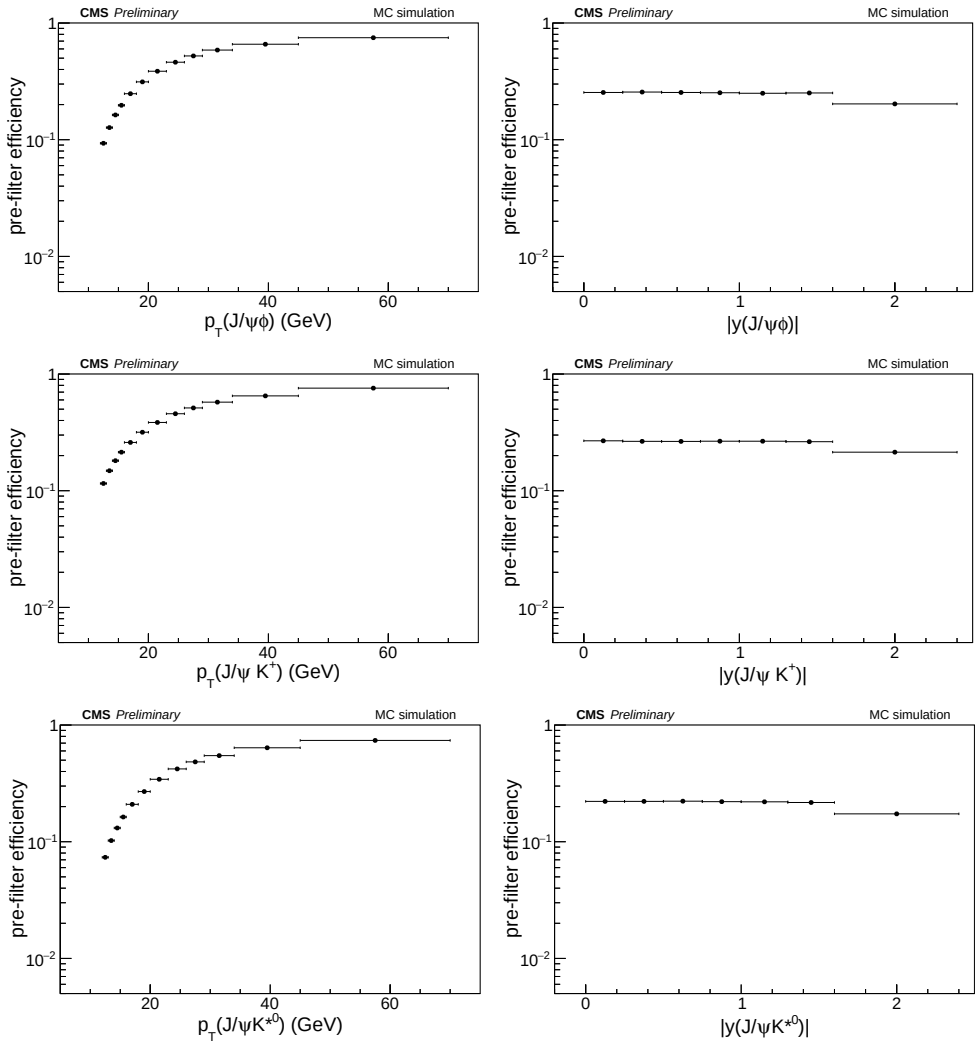

Figure 3.8: Efficiencies at generator level (ϵ_1) are shown for the $B_s{}^0$ (top), B^+ (middle) and B^0 (bottom) decays in p_T (left) and rapidity (right) bins.

These efficiencies are shown in figure 3.9 for all channels.

The total efficiency, computed as the product of pre-filter and reconstruction efficiency, is shown in figure 3.10. Numerical values are shown the tables 3.3– 3.8.

Table 3.3: Numerical values obtained for the prefilter efficiency, the reconstruction efficiency, and the total efficiency in transverse momentum bins for B_s^0 candidates.

$p_T(B_s^0)$ (GeV)	Pre-Filter	Reco	Total
12 – 13	0.09321 ± 0.00058	0.02826 ± 0.00032	0.00263 ± 0.00003
13 – 14	0.12703 ± 0.00076	0.04422 ± 0.00038	0.00562 ± 0.00006
14 – 15	0.16305 ± 0.00095	0.05976 ± 0.00044	0.00974 ± 0.00009
15 – 16	0.19729 ± 0.00115	0.07350 ± 0.00049	0.01450 ± 0.00013
16 – 18	0.24793 ± 0.00104	0.09037 ± 0.00041	0.02240 ± 0.00014
18 – 20	0.31315 ± 0.00139	0.10939 ± 0.00049	0.03425 ± 0.00022
20 – 23	0.38627 ± 0.00154	0.12488 ± 0.00049	0.04824 ± 0.00027
23 – 26	0.46154 ± 0.00208	0.14111 ± 0.00062	0.06513 ± 0.00041
26 – 29	0.52248 ± 0.00268	0.15473 ± 0.00078	0.08085 ± 0.00058
29 – 34	0.58665 ± 0.00277	0.16488 ± 0.00080	0.09673 ± 0.00065
34 – 45	0.65828 ± 0.00295	0.17897 ± 0.00085	0.11781 ± 0.00077
45 – 70	0.74855 ± 0.00417	0.18217 ± 0.00123	0.13636 ± 0.00119

Table 3.4: Numerical values obtained for the prefilter efficiency, the reconstruction efficiency and the total efficiency in rapidity bins for B_s^0 candidates.

$\lvert y(B_s^0)\rvert$	Pre-Filter	Reco	Total
0.00 – 0.25	0.25407 ± 0.00113	0.11834 ± 0.00048	0.03007 ± 0.00018
0.25 – 0.50	0.25619 ± 0.00113	0.12324 ± 0.00049	0.03157 ± 0.00019
0.50 – 0.75	0.25407 ± 0.00114	0.13452 ± 0.00052	0.03418 ± 0.00020
0.75 – 1.00	0.25266 ± 0.00116	0.12810 ± 0.00051	0.03237 ± 0.00020
1.00 – 1.30	0.25010 ± 0.00107	0.11167 ± 0.00045	0.02793 ± 0.00016
1.30 – 1.60	0.25162 ± 0.00110	0.09284 ± 0.00043	0.02336 ± 0.00015
1.60 – 2.40	0.20275 ± 0.00067	0.06471 ± 0.00027	0.01312 ± 0.00007

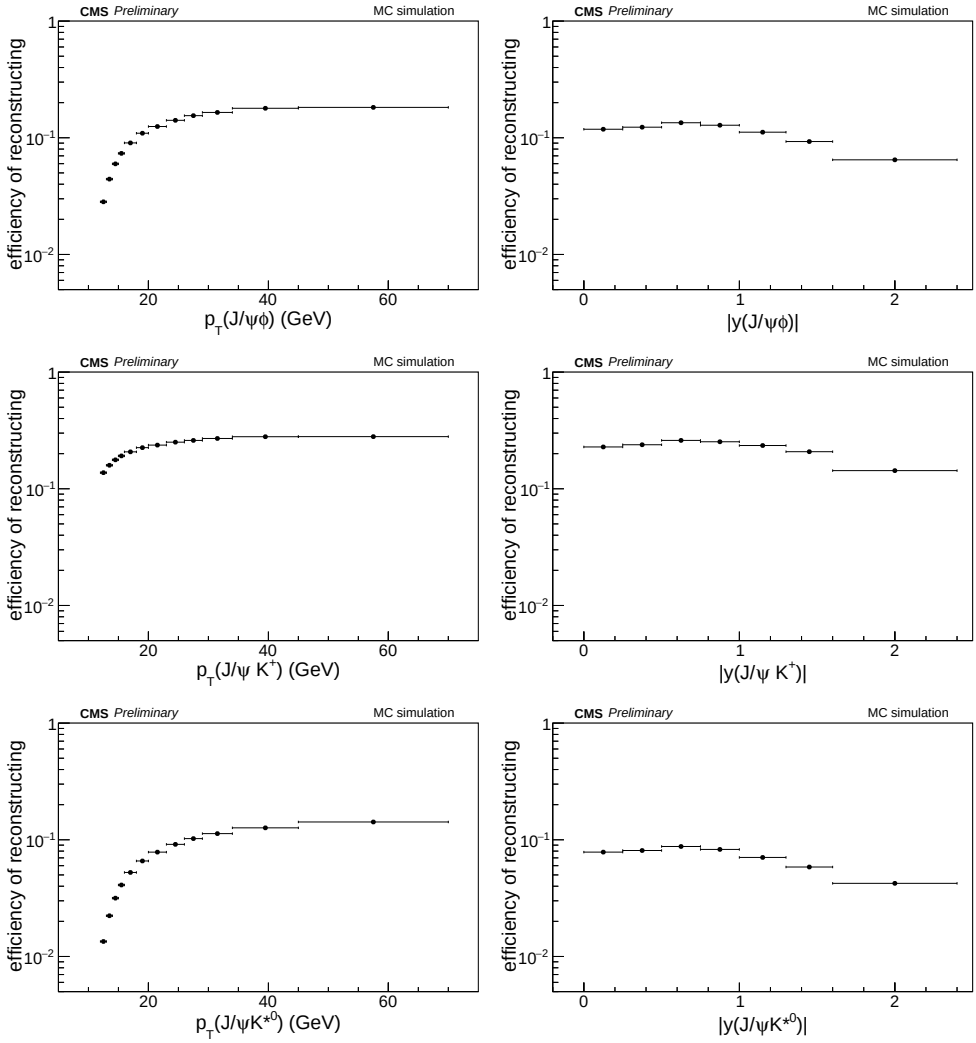

Figure 3.9: Reconstruction efficiencies (ϵ_2) are shown for the $B_s{}^0$ (top), B^+ (middle) and B^0 (bottom) decays in p_T (left) and rapidity (right) bins.

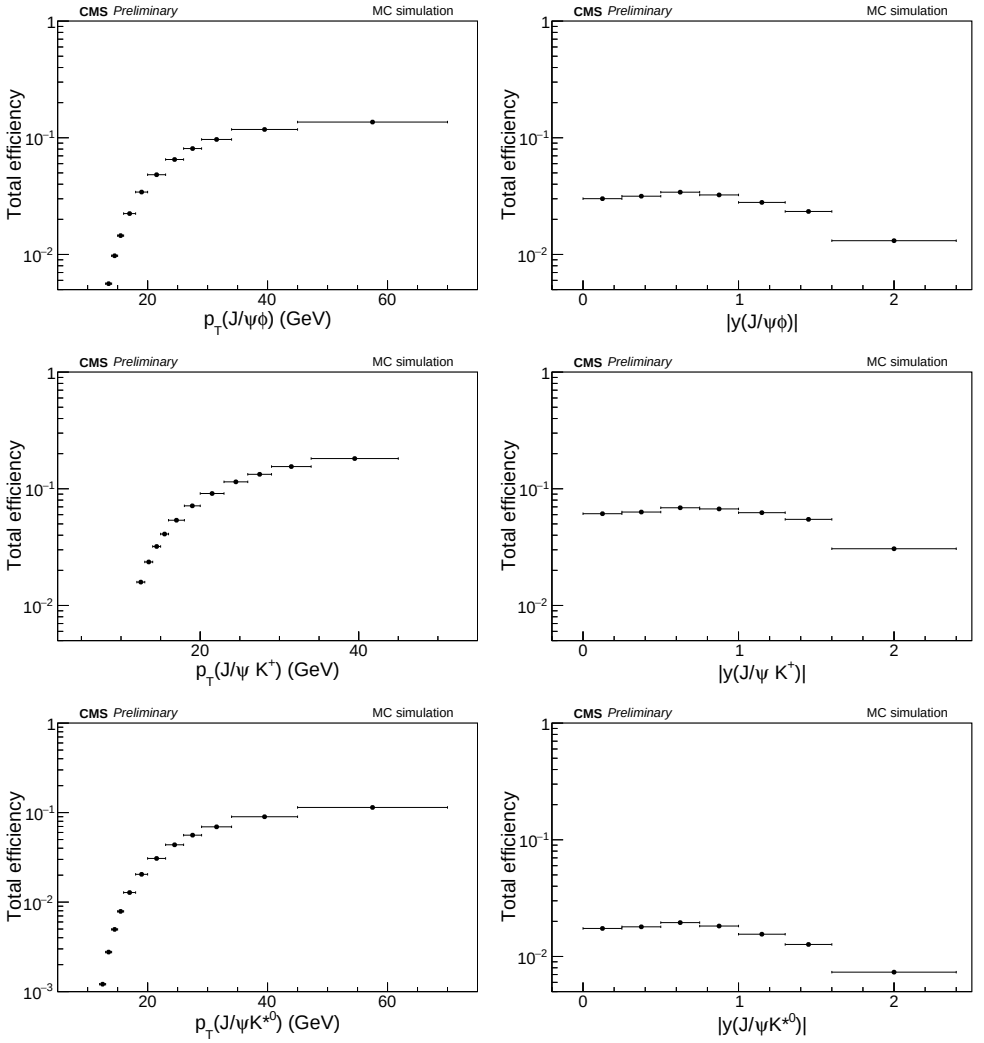

Figure 3.10: Total efficiency distributions for the $B_s{}^0$ (top), B^+ (middle) and B^0 (bottom) decays in p_T (left) and rapidity (right) bins.

Table 3.5: Numerical values obtained for the prefilter efficiency, the reconstruction efficiency, and the total efficiency in as transverse momentum bins for B^+ candidates.

$p_T(B^+)$ (GeV)	Pre-Filter	Reco	Total
12 – 13	0.11538 ± 0.00062	0.13742 ± 0.00069	0.01586 ± 0.00012
13 – 14	0.14836 ± 0.00078	0.15906 ± 0.00074	0.02360 ± 0.00016
14 – 15	0.18091 ± 0.00095	0.17706 ± 0.00078	0.03203 ± 0.00022
15 – 16	0.21400 ± 0.00114	0.19135 ± 0.00084	0.04095 ± 0.00028
16 – 18	0.25936 ± 0.00102	0.20730 ± 0.00066	0.05377 ± 0.00027
18 – 20	0.31720 ± 0.00135	0.22517 ± 0.00076	0.07143 ± 0.00039
20 – 23	0.38471 ± 0.00148	0.23680 ± 0.00074	0.09110 ± 0.00045
23 – 26	0.45694 ± 0.00200	0.25090 ± 0.00092	0.11465 ± 0.00065
26 – 29	0.51235 ± 0.00260	0.25970 ± 0.00112	0.13306 ± 0.00089
29 – 34	0.57434 ± 0.00269	0.26991 ± 0.00113	0.15502 ± 0.00097
34 – 45	0.64990 ± 0.00288	0.27942 ± 0.00118	0.18159 ± 0.00111
45 – 70	0.75724 ± 0.00396	0.27989 ± 0.00169	0.21194 ± 0.00169

Table 3.6: Numerical values obtained for the prefilter efficiency, the reconstruction efficiency, and the total efficiency in rapidity bins for B^+ candidates.

$\lvert y(B^+)\rvert$	Pre-Filter	Reco	Total
0.00 – 0.25	0.26781 ± 0.00110	0.22831 ± 0.00072	0.06114 ± 0.00032
0.25 – 0.50	0.26502 ± 0.00111	0.23844 ± 0.00073	0.06319 ± 0.00033
0.50 – 0.75	0.26464 ± 0.00111	0.25978 ± 0.00076	0.06875 ± 0.00035
0.75 – 1.00	0.26573 ± 0.00113	0.25290 ± 0.00077	0.06720 ± 0.00035
1.00 – 1.30	0.26568 ± 0.00105	0.23487 ± 0.00069	0.06240 ± 0.00031
1.30 – 1.60	0.26315 ± 0.00108	0.20786 ± 0.00069	0.05470 ± 0.00029
1.60 – 2.40	0.21401 ± 0.00066	0.14310 ± 0.00043	0.03063 ± 0.00013

Table 3.7: Numerical values obtained for the prefilter efficiency, the reconstruction efficiency, and the total efficiency in transverse momentum bins for B^0 candidates.

$p_T(B^0)$ (GeV)	Pre-Filter	Reco	Total
12 – 13	0.07352 ± 0.00051	0.01645 ± 0.00018	0.00121 ± 0.00002
13 – 14	0.10231 ± 0.00067	0.02702 ± 0.00022	0.00276 ± 0.00003
14 – 15	0.13116 ± 0.00084	0.03777 ± 0.00026	0.00495 ± 0.00005
15 – 16	0.16295 ± 0.00104	0.04834 ± 0.00030	0.00788 ± 0.00007
16 – 18	0.20916 ± 0.00096	0.06113 ± 0.00025	0.01279 ± 0.00008
18 – 20	0.26918 ± 0.00130	0.07591 ± 0.00030	0.02043 ± 0.00013
20 – 23	0.34299 ± 0.00146	0.08966 ± 0.00030	0.03075 ± 0.00017
23 – 26	0.42082 ± 0.00200	0.10382 ± 0.00038	0.04369 ± 0.00026
26 – 29	0.48345 ± 0.00260	0.11577 ± 0.00048	0.05597 ± 0.00038
29 – 34	0.54698 ± 0.00272	0.12677 ± 0.00049	0.06934 ± 0.00044
34 – 45	0.63859 ± 0.00289	0.14078 ± 0.00053	0.08990 ± 0.00053
45 – 70	0.73845 ± 0.00405	0.15465 ± 0.00078	0.11420 ± 0.00085

Table 3.8: Numerical values obtained for the prefilter efficiency, the reconstruction efficiency, and the total efficiency in rapidity bins for B^0 candidates.

$\lvert y(B^0)\rvert$	Pre-Filter	Reco	Total
0.00 – 0.25	0.22171 ± 0.00105	0.08941 ± 0.00030	0.01982 ± 0.00012
0.25 – 0.50	0.22174 ± 0.00105	0.09249 ± 0.00031	0.02051 ± 0.00012
0.50 – 0.75	0.22263 ± 0.00106	0.09990 ± 0.00032	0.02224 ± 0.00013
0.75 – 1.00	0.22051 ± 0.00107	0.09446 ± 0.00032	0.02083 ± 0.00012
1.00 – 1.30	0.21982 ± 0.00099	0.08064 ± 0.00028	0.01773 ± 0.00010
1.30 – 1.60	0.21693 ± 0.00102	0.06655 ± 0.00026	0.01444 ± 0.00009
1.60 – 2.40	0.17333 ± 0.00062	0.04835 ± 0.00017	0.00838 ± 0.00004

3.7 Fragmentation Fractions Ratio

Following the ideas stated in section 3.1, the observable of interest is presented: the ratio of the fragmentation fraction $f_{\mathrm{s}}/f_{\mathrm{u}}$ times the branching fractions Eq. 3.3 ($\mathcal{R}_S$) and similarly for $f_{\mathrm{d}}/f_{\mathrm{u}}$ ($\mathcal{R}_d$). So the raw yields obtained and the efficiencies computed are shown in the tables 3.9 and 3.10 for $\mathcal{R}_S$ and in tables 3.11 and 3.12 for $\mathcal{R}_d$ Eq. 3.4.

3.8 Systematic Uncertainties

The results showed in section 3.7 only include statistical uncertainties. Systematic uncertainties related to muon reconstruction & identification, and trigger efficiency, are considered to cancel out in the measured ratios. On the other hand, systematic uncertainties affecting $\mathcal{R}$ are: signal and background models, the MC size, tracking, and MC reweight. These contributions are summarized in tables 3.13 and 3.14 for $\mathcal{R}_S$ in p_{T} and rapidity. And in tablers 3.15 and 3.16 for $\mathcal{R}_d$ for p_{T} and rapidity respectively.

The total systematic uncertainty is the sum in quadratures of the individual systematic uncertainties. The details of each contribution are discussed below.

3.8.1 Signal and Background Models

The uncertainties associated with the fitting models are evaluated by varying the probability density functions used in the maximum likelihood fit for signal and background components separately. For signal instead of using a double Gaussian distribution, a non-standardized Studentś t-distribution is used. The systematic uncertainty is estimated as the difference between the yields obtained by varying the model.

Similarly, the combinatorial background model is changed to a Chebyshev polynomial. Signal yields differences were quoted as systematic uncertainties due to the combinatorial background model.

Finally, the fitting procedure is tested using randomly generated event samples (toy MC), with a sample size matching the data size. This reflects the nominal likelihood probability distribution functions and fitted parameters. No significant fit biases are found in the central values and uncertainties.

3.8.2 Monte Carlo Size

Given the finite nature of the MC samples used, a systematic uncertainty is associated with the statistical uncertainty of the efficiencies. Given the adequate size of the MC samples produced for this analysis, this systematical uncertainty is sub-dominant.

3.8.3 Tracking

In the determination of the ratio $\mathcal{R}$, the number of tracks involved in the ratio is not the same in the numerator and denominator. The extra track in the decays of the B_s^0 and B^0 mesons may induce a systematic uncertainty, due to the track efficiency determination. To take into account this effect, the results produced by the tracking POG (CMS AN-2017/166 and CMS DP-2018/050) are used it is assumed conservatively to be valid for the 2018 data-taking period. Therefore, the systematic uncertainty for one track is 2.3%.

3.8.4 Monte Carlo Reweighting

When making comparisons of different kinematic variables between the signal in data and MC, a few discrepancies were observed. To match the MC to data reweighting procedures were performed for the b-hadron and the tracks (p_{T} and rapidity). In addition, other reweighting procedures related to the ϕ-helicity angle have been tested without significant changes to the measurement.

Considering the b hadron rapidity variable as an example, the reweighting procedure consists of computing weights as the ratio of background subtracted data and MC in each rapidity bin:

$$w_{Bin} = \frac{N_{\mathrm{data}}(\mathrm{Bin})}{N_{\mathrm{MC}}(\mathrm{Bin})}. \tag{3.9}$$

These weights applied to the MC will make a perfect match between the data and the reweighted MC in that specific variable. Then the efficiency and $\mathcal{R}$ are recalculated. The difference between the results with and without MC reweighting is considered as systematic uncertainty associated with the MC reweighting.

Table 3.9: Raw yields and efficiencies with their respective statistical uncertainties. The value of $\mathcal{R}_S$ computed directly from these results are also shown. For $\mathcal{R}_S$, just the yield error propagation is present.

p_T (GeV)	B_s^0 Yield	B^+ Yield	B^+ Efficiency	B_s^0 Efficiency	$\mathcal{R}_S$
12 – 13	4395 ± 75	208436 ± 529	0.01586 ± 0.00012	0.00263 ± 0.00003	0.12695 ± 0.00219
13 – 14	6997 ± 100	238697 ± 554	0.02360 ± 0.00016	0.00562 ± 0.00006	0.12312 ± 0.00177
14 – 15	8901 ± 108	251434 ± 574	0.03203 ± 0.00022	0.00974 ± 0.00009	0.11637 ± 0.00144
15 – 16	10273 ± 118	252905 ± 570	0.04095 ± 0.00028	0.01450 ± 0.00013	0.11470 ± 0.00134
16 – 18	22164 ± 171	472530 ± 769	0.05377 ± 0.00027	0.02240 ± 0.00014	0.11256 ± 0.00089
18 – 20	21567 ± 170	407594 ± 706	0.07143 ± 0.00039	0.03425 ± 0.00022	0.11033 ± 0.00089
20 – 23	28385 ± 191	482561 ± 768	0.09110 ± 0.00045	0.04824 ± 0.00027	0.11109 ± 0.00077
23 – 26	22152 ± 167	349881 ± 661	0.11465 ± 0.00065	0.06513 ± 0.00041	0.11144 ± 0.00087
26 – 29	16304 ± 145	250986 ± 554	0.13306 ± 0.00089	0.08085 ± 0.00058	0.10691 ± 0.00098
29 – 34	18314 ± 154	268446 ± 581	0.15502 ± 0.00097	0.09673 ± 0.00065	0.10934 ± 0.00095
34 – 45	19048 ± 156	267778 ± 609	0.18159 ± 0.00111	0.11781 ± 0.00077	0.10965 ± 0.00093
45 – 70	9869 ± 116	138089 ± 433	0.21194 ± 0.00169	0.13636 ± 0.00119	0.11109 ± 0.00135

Table 3.10: Raw yields and efficiencies. $\mathcal{R}_S$ is also computed from the definition in section 1. For $\mathcal{R}_S$, just the yield error propagation is present.

$\|y\|$	B_s^0 Yield	B^+ Yield	B^+ Efficiency	B_s^0 Efficiency	$\mathcal{R}_S$
0.00 – 0.25	29578 ± 186	541692 ± 764	0.06114 ± 0.00032	0.03007 ± 0.00018	0.11105 ± 0.00071
0.25 – 0.50	30854 ± 192	559411 ± 784	0.06319 ± 0.00033	0.03157 ± 0.00019	0.11039 ± 0.00070
0.50 – 0.75	32434 ± 198	593143 ± 821	0.06875 ± 0.00035	0.03418 ± 0.00020	0.10999 ± 0.00069
0.75 – 1.00	29034 ± 196	535407 ± 810	0.06720 ± 0.00035	0.03237 ± 0.00020	0.11260 ± 0.00078
1.00 – 1.30	26040 ± 192	520667 ± 919	0.06240 ± 0.00031	0.02793 ± 0.00016	0.11174 ± 0.00085
1.30 – 1.60	17953 ± 167	393101 ± 910	0.05470 ± 0.00029	0.02336 ± 0.00015	0.10694 ± 0.00102
1.60 – 2.40	22590 ± 222	479677 ± 1168	0.03063 ± 0.00013	0.01312 ± 0.00007	0.10994 ± 0.00111

Table 3.11: Raw yields and efficiencies with their respective statistical uncertainties. The value of $\mathcal{R}_d$ computed directly from these results is also shown. For $\mathcal{R}_d$, just the yield error propagation is present.

p_T (GeV)	B^0 Yield	B^+ Yield	B^+ Efficiency	B^0 Efficiency	$\mathcal{R}_d$
12 – 13	13261 ± 520	208436 ± 529	0.01586 ± 0.00012	0.00121 ± 0.00002	0.82210 ± 0.03228
13 – 14	25320 ± 857	238697 ± 554	0.02360 ± 0.00016	0.00276 ± 0.00003	0.90562 ± 0.03073
14 – 15	32496 ± 923	251434 ± 574	0.03203 ± 0.00022	0.00495 ± 0.00005	0.83264 ± 0.02372
15 – 16	41243 ± 695	252905 ± 570	0.04095 ± 0.00028	0.00788 ± 0.00007	0.83779 ± 0.01423
16 – 18	95348 ± 1038	472530 ± 769	0.05377 ± 0.00027	0.01279 ± 0.00008	0.83733 ± 0.00922
18 – 20	101216 ± 1321	407594 ± 706	0.07143 ± 0.00039	0.02043 ± 0.00013	0.85811 ± 0.01130
20 – 23	141747 ± 2040	482561 ± 768	0.09110 ± 0.00045	0.03075 ± 0.00017	0.86955 ± 0.01259
23 – 26	115830 ± 1565	349881 ± 661	0.11465 ± 0.00065	0.04369 ± 0.00026	0.86398 ± 0.01179
26 – 29	91951 ± 1028	250986 ± 554	0.13306 ± 0.00089	0.05597 ± 0.00038	0.87443 ± 0.00997
29 – 34	104570 ± 1130	268446 ± 581	0.15502 ± 0.00097	0.06934 ± 0.00044	0.86285 ± 0.00951
34 – 45	109976 ± 896	267778 ± 609	0.18159 ± 0.00111	0.08990 ± 0.00053	0.82683 ± 0.00699
45 – 70	61313 ± 1153	138089 ± 433	0.21194 ± 0.00169	0.11420 ± 0.00085	0.81379 ± 0.01555

Table 3.12: Raw yields and efficiencies. $\mathcal{R}_d$ is computed from the definition in section 1. For $\mathcal{R}_d$, just the yield error propagation is present.

$\lvert y\rvert$	B^0 Yield	B^+ Yield	B^+ Efficiency	B^0 Efficiency	$\mathcal{R}_d$
0.00 – 0.25	148284 ± 1261	541692 ± 764	0.06114 ± 0.00032	0.01982 ± 0.00012	0.84440 ± 0.00728
0.25 – 0.50	155504 ± 1450	559411 ± 784	0.06319 ± 0.00033	0.02051 ± 0.00012	0.85652 ± 0.00808
0.50 – 0.75	163290 ± 1322	593143 ± 821	0.06875 ± 0.00035	0.02224 ± 0.00013	0.85098 ± 0.00699
0.75 – 1.00	142595 ± 1246	535407 ± 810	0.06720 ± 0.00035	0.02083 ± 0.00012	0.85926 ± 0.00762
1.00 – 1.30	125410 ± 1068	520667 ± 919	0.06240 ± 0.00031	0.01773 ± 0.00010	0.84790 ± 0.00737
1.30 – 1.60	86048 ± 823	393101 ± 910	0.05470 ± 0.00029	0.01444 ± 0.00009	0.82934 ± 0.00816
1.60 – 2.40	107838 ± 584	479677 ± 1168	0.03063 ± 0.00013	0.00838 ± 0.00004	0.82149 ± 0.00487

Table 3.13: Yield variations and systematic uncertainties on $\mathcal{R}_S$ vs p_T estimated from alternative B_s^0 and B^+ fitting strategies described in the text. The total systematic uncertainty is the sum in quadrature of the individual uncertainties. Statistical uncertainty is shown too.

p_T (GeV)	B_s^0 signal	B_s^0 bkg	B^+ signal	B^+ bkg	MC size	MC reweight	Total Systematic uncertainty (%)	Statistical uncertainty (%)
12 – 13	1.3	0.6	0.5	1.0	1.5	1.9	3.8	2.1
13 – 14	0.5	2.0	0.8	0.8	1.3	0.9	3.6	1.6
14 – 15	0.9	1.6	0.4	0.6	1.2	1.6	3.6	1.3
15 – 16	1.0	1.7	0.6	0.5	1.1	1.6	3.7	1.2
16 – 18	0.5	1.4	0.6	0.2	0.8	2.0	3.6	0.8
18 – 20	1.1	1.6	0.4	0.3	0.8	2.0	3.7	0.8
20 – 23	0.7	1.4	0.7	1.1	0.7	2.1	3.8	0.7
23 – 26	0.9	1.1	0.1	0.8	0.9	1.9	3.5	0.8
26 – 29	0.6	1.2	1.1	1.9	1.0	1.6	4.0	0.9
29 – 34	0.6	0.9	0.2	0.5	0.9	1.8	3.3	0.9
34 – 45	0.4	1.2	0.1	1.3	0.9	1.9	3.6	0.9
45 – 70	1.0	1.1	1.0	0.0	1.2	1.4	3.5	1.3

Table 3.14: Yields and systematic uncertainties on $\mathcal{R}_S$ vs $|y|$ estimated from alternative B_s^0 and B^+ fitting strategies described in the text. The total systematic uncertainty is the sum in quadrature of the individual uncertainties. Statistical uncertainty is show too.

$\|y\|$	B_s^0 signal	B_s^0 bkg	B^+ signal	B^+ bkg	MC size	MC reweight	Total Systematic uncertainty (%)	Statistical uncertainty (%)
0.00 – 0.25	0.3	0.1	0.5	0.1	0.8	1.6	2.9	0.6
0.25 – 0.50	0.6	0.1	0.4	0.3	0.8	1.7	3.0	0.6
0.50 – 0.75	0.5	0.4	0.5	0.2	0.8	1.5	3.0	0.6
0.75 – 1.00	0.7	0.5	1.1	1.8	0.8	3.1	4.5	0.7
1.00 – 1.30	1.2	0.7	1.4	2.0	0.8	3.2	4.9	0.8
1.30 – 1.60	1.8	1.2	0.9	1.7	0.8	3.1	4.9	1.0
1.60 – 2.40	2.1	1.0	0.7	1.6	0.7	3.3	4.9	1.0

3.9 Results

The measured $\mathcal{R}_s$ values, displayed in figure 3.11 (right), do not show any signs of a rapidity dependence. Nevertheless, on the p_{T} variable, it shows a clear dependence at low-p_{T} , followed by a flat high-p_{T} trend. Averaging the $p_{\mathrm{T}} > 18$ GeV measurements gives $\mathcal{R}_{\mathrm{s}} = 0.1102 \pm 0.0027$, where the uncertainty includes all contributions added in quadrature (see table 3.13). The low-p_{T} dependence is compatible with the LHCb measurements (for $2 < y < 4.5$) Refs. [9, 11], also shown in figure 3.11, while the asymptotically flat dependence observed at high p_{T} is in agreement with the LEP results in Refs. [16, 30, 15, 14].

The measured $\mathcal{R}_d$ ratio can be used to probe isospin invariance in B meson production provided that, in the conversion from $\mathcal{R}_d$ to $f_{\mathrm{d}}/f_{\mathrm{u}}$ in Eq. 3.4 $\mathcal{B}(\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+)/\mathcal{B}(\mathrm{B}^0 \to J/\psi\,\mathrm{K}^{*0})$, is evaluated without the isospin invariance assumption.

For $\mathcal{B}(\mathrm{B}^0 \to J/\psi\,\mathrm{K}^{*0})$, the world-average value [66], which is dominated by at the $\Upsilon(4S)$ resonance that asume isospin invariance: $\mathcal{R}^{\pm,0} = \mathcal{B}(\Upsilon(4S) \to \mathrm{B}^+\mathrm{B}^-)/\mathcal{B}(\Upsilon(4S) \to \mathrm{B}^0\bar{\mathrm{B}}^0) = 1$. Whereas for $\mathcal{B}(\mathrm{B}^+ \to J/\psi\,\mathrm{K}^+)$, it is used its most precise measurement [34], after correcting for their assumption, $\mathcal{R}^{\pm,0} = 1.058 \pm 0.024$ [21], to make it compatible with the branching fractions that use $\mathcal{R}^{\pm,0} = 1$. The ratio of branching fractions in Eq. 3.4 is then divided by the most recent $\mathcal{R}^{\pm,0}$ value (1.059 ± 0.027) [20] to remove the isospin conservation assumption (as in Ref. [53]). The obtained $f_{\mathrm{d}}/f_{\mathrm{u}}$ ratios are plotted versus p_{T} and $|y|$ in Fig. 3.12, with no dependence on either variable observed. The average value of 0.998 ± 0.063, with the uncertainty including all contributions, is compatible with unity within the 6% precision of the measurement, consistent with isospin invariance in B meson production at hadron colliders.

In summary, while no $\mathcal{R}_s$ dependence on the B meson rapidity is seen, a strong variation is observed in the $12 < p_{\mathrm{T}} < 18$ GeV range, followed by a flat trend for

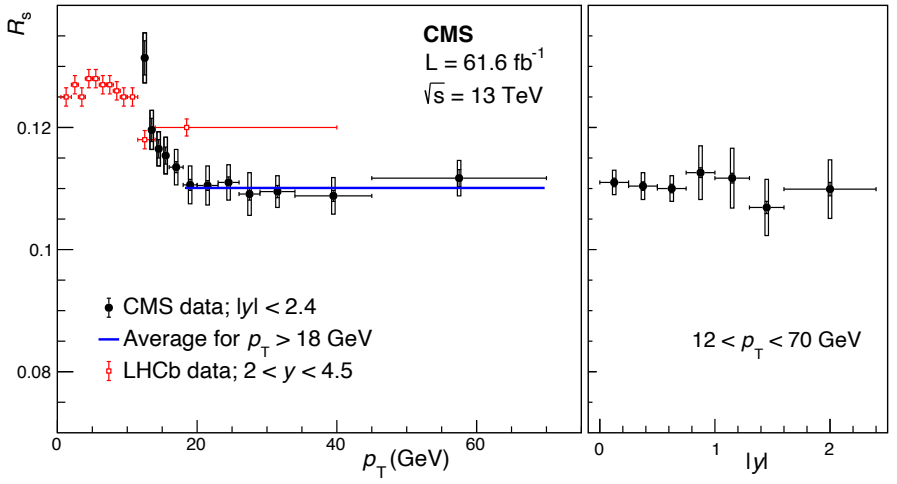

Figure 3.11: Efficiency-corrected yield ratio $\mathcal{R}_s$, as a function of p_{T} (left) and $|y|$ (right). The vertical bars (boxes) represent the statistical (bin-to-bin systematic) uncertainties, while the horizontal bars give the bin widths. The global uncertainty (of 2.3%) is not graphically represented. The blue line represents the average for $p_{\mathrm{T}} > 18$ GeV. For comparison, the LHCb measurement [3] is also shown.

Table 3.15: Yield variations and systematic uncertainties on $\mathcal{R}_d$ vs p_{T} estimated from alternative B_s^0 and B^0 fitting strategies described in the text. The total systematic uncertainty is the sum in quadrature of the individual uncertainties. Statistical uncertainty is shown too.

p_{T} (GeV)	B_s^0 signal	B_s^0 bkg	B^0 signal	B^0 bkg	MC size	MC reweight	Total Systematic uncertainty (%)	Statistical uncertainty (%)
12 – 13	1.1	2.8	1.0	2.7	1.5	1.2	4.6	1.6
13 – 14	0.3	1.4	0.0	2.4	1.3	1.1	3.3	1.3
14 – 15	1.1	0.3	0.2	2.0	1.2	1.6	3.1	1.2
15 – 16	0.3	1.6	0.2	2.0	1.2	2.1	3.5	1.2
16 – 18	0.8	1.8	0.2	1.7	0.8	1.3	3.0	0.8
18 – 20	0.9	3.1	0.3	1.5	0.9	1.8	4.1	0.8
20 – 23	0.2	1.5	0.2	1.3	0.8	1.4	2.5	0.7
23 – 26	0.8	0.9	0.2	1.2	0.9	2.6	3.2	0.8
26 – 29	0.0	1.3	0.3	1.2	1.0	2.6	3.3	1.0
29 – 34	0.0	1.6	0.2	1.2	0.9	2.8	3.6	0.9
34 – 45	0.3	1.7	0.2	1.2	0.9	2.7	3.6	0.9
45 – 70	0.3	1.3	0.3	1.8	1.1	3.2	4.1	1.3

Table 3.16: Yield variations and systematic uncertainties on $\mathcal{R}_d$ vs p_{T} estimated from alternative $\mathrm{B_s^0}$ and $\mathrm{B^0}$ fitting strategies described in the text. The total systematic uncertainty is the sum in quadrature of the individual uncertainties. Statistical uncertainty is shown too.

$\lvert y\rvert$	$\mathrm{B_s^0}$ signal	$\mathrm{B_s^0}$ bkg	$\mathrm{B^0}$ signal	$\mathrm{B^0}$ bkg	MC size	MC reweight	Total Systematic uncertainty (%)	Statistical uncertainty (%)
0.00 – 0.25	0.4	0.3	0.3	1.1	0.8	4.9	5.1	0.6
0.25 – 0.50	0.7	0.3	0.2	1.1	0.8	5.0	5.2	0.6
0.50 – 0.75	0.7	0.6	0.2	1.2	0.8	4.8	5.0	0.6
0.75 – 1.00	0.9	0.6	0.3	1.3	0.8	5.3	5.6	0.7
1.00 – 1.30	1.9	1.0	0.3	1.3	0.8	6.4	6.9	0.7
1.30 – 1.60	2.0	1.2	0.4	1.4	0.9	4.9	5.7	0.9
1.60 – 2.40	2.4	1.1	1.2	1.6	0.7	1.4	3.7	0.9

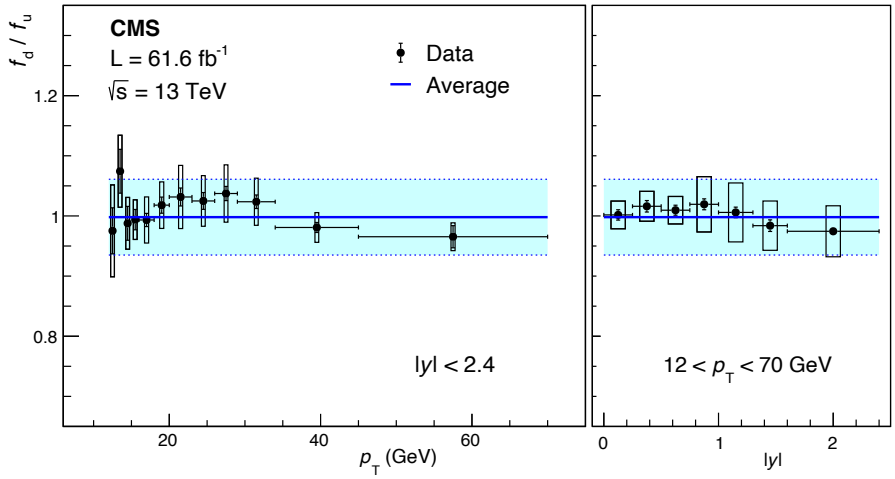

Figure 3.12: The ratio of the $\mathrm{B^0}$ to $\mathrm{B^+}$ hadron production fractions $f_{\mathrm{d}}/f_{\mathrm{u}}$, as a function of p_{T} (left) and $|y|$ (right). The vertical bars (boxes) represent the statistical (bin-to-bin systematic) uncertainties, while the horizontal bars give the bin widths. The horizontal blue line and band represent the average value and uncertainty. The global uncertainty (of 5.7%) is included in the blue bands but not in the individual data points.

higher p_{T} values. The $f_{\mathrm{d}}/f_{\mathrm{u}}$ ratio, measured for the first time in proton-proton collisions using the $\mathrm{B}^0 \to J/\psi\,\mathrm{K}^{*0}$ decay channel, is found to be compatible with unity and independent of rapidity and p_{T} . This is the first direct measurement of isospin invariance in B meson production at hadron colliders.

Chapter 4

$B^0 \to K^0_S \mu^- \mu^+$ Branching Fraction and Angular Analysis

4.1 Introduction

The $B^0 \to K^0_S \mu^- \mu^+$ decay is interesting in the search for new physics. The angular distribution and the differential branching ratio can measure the F_H (contribution from scalar, pseudoscalar, and tensor amplitudes) and A_I (isospin asymmetry) respectively. This channel was previously not explored in CMS because of the poor performance of the regular triggers. This analysis uses the BParked data set, which was recorded during the 2018 data taking. This dataset was recorded with a displaced muon trigger which opens the possibility to observe this decay and measure these observables.

The $B^0 \to K^0_S \mu^- \mu^+$ decay is a rare, flavor-changing neutral-current process that is mediated by electroweak box and penguin amplitudes in the SM. The loop-order suppression of the SM amplitudes increases the sensitivity to new virtual particles that can influence the decay amplitude at a similar level to the SM contribution (this has been explored in-depth in Chapter 1).

It is also possible to look for the contribution of new particles through deviations from the angular distribution of the final-state particles predicted by the SM:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_l} = \frac{3}{4}(1 - F_H)(1 - \cos^2\theta_l) + \frac{1}{2}F_H + A_{FB}\cos\theta_l, \tag{4.1}$$

where θ_l is the angle between the direction of the muon and the kaon. This differen-

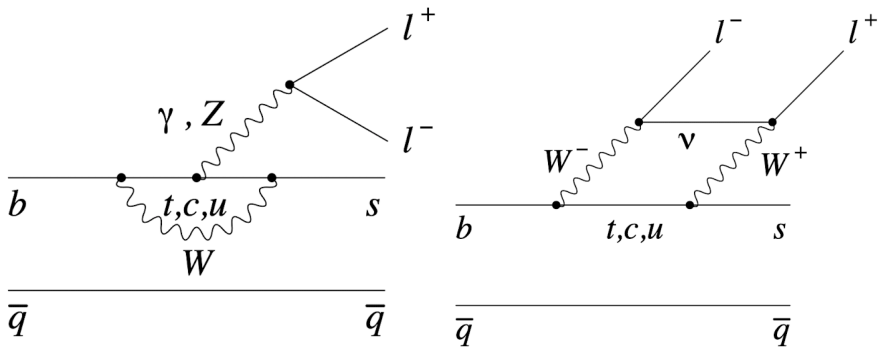

Figure 4.1: Electroweak diagrams for the $B^0 \to K^0_S \mu^- \mu^+$ decay.

tial decay rate depends on two parameters: the forward-backward asymmetry of the dimuon system, A_{FB}; and F_H the fractional contribution of (pseudo)scalar and tensor amplitudes to the decay width in the approximation that muons are massless. The decay width, A_{FB} and F_H, all depend on q^2.

Since the B^0 and $\bar{B}^0$ mesons can decay to the same $K^0\mu^+\mu^-$ final state, it is not possible to determine the flavor of the B meson from the decay products. Without tagging the flavor of the neutral B meson at production time, it is not possible to unambiguously chose the correct muon to determine $\cos\theta_l$. However if no attempt is made to measure B^0 and $\bar{B}^0$ decay rates separately, we can combine both contributions and measure the effective $F_H = \frac{F_H^{B^0} + F_H^{\bar{B}^0}}{2}$ and $A_{FB} = A_{FB}^{B^0} - A_{FB}^{\bar{B}^0}$. This way, $\cos\theta_l$ can be defined with respect to the μ^+. In this case, any visible A_{FB} would indicate that there is either a difference in the number of B^0 and $\bar{B}^0$ mesons produced (i.e. CP violation in the decay) or that the A_{FB} of the B^0 and $\bar{B}^0$ decay differ. Any residual asymmetry can be canceled by performing the analysis in terms of $|\cos\theta_l|$:

$$\frac{1}{\Gamma}\frac{d\Gamma}{d|\cos\theta_l|} = \frac{3}{2}(1 - F_H)(1 - \cos^2\theta_l) + F_H, \tag{4.2}$$

where the constraint $0 \leq F_H < 3$ is needed for this expression to remain positive at all values of $|\cos\theta_l|$.

On the other hand, $B^0 \to K^0\mu^+\mu^-$ decays are highly sensitive to contributions from vector or axial-vector-like particles predicted in extensions of the SM. However, despite recent progress in lattice calculations, theoretical predictions of the decay rates suffer from relatively large uncertainties in the $B^0 \to K^0$ form factor calculations. To maximize sensitivity, observables can be constructed from ratios or asymmetries where

the leading form factor uncertainties cancel. The CP-averaged isospin asymmetry (A_I) is such observable, and it's defined in equation 4.3.

$$\begin{aligned} A_I &= \frac{\Gamma\left(B^0 \to K^0\ \mu^+\mu^-\right) - \Gamma\left(B^+ \to K^+\ \mu^+\mu^-\right)}{\Gamma\left(B^0 \to K^0\ \mu^+\mu^-\right) + \Gamma\left(B^+ \to K^+\ \mu^+\mu^-\right)} \\ &= \frac{\mathcal{B}\ \left(B^0 \to K^0 \mu^+\mu^-\right) - (\tau_0/\tau_+) \cdot \mathcal{B}\ \left(B^+ \to K^+ \mu^+\mu^-\right)}{\mathcal{B}\ \left(B^0 \to K^0 \mu^+\mu^-\right) + (\tau_0/\tau_+) \cdot \mathcal{B}\ \left(B^+ \to K^+ \mu^+\mu^-\right)} \end{aligned} \tag{4.3}$$

where $\Gamma(f)$ and $\mathcal{B}(f)$ are the partial width and branching fraction of the $B \to f$ decay and τ_0/τ_+ is the ratio of the lifetimes of the B^0 and B^+ mesons.

To calculate the A_I per bin of q^2 it is needed for the differential branching fractions of $B^0 \to K^0_S \mu^- \mu^+$ and $B^+ \to K^+ \mu^- \mu^+$. The charged channel is not explored in the scope of this work, but the individual branching fraction (preliminary results) of $B^0 \to K^0_S \mu^- \mu^+$ are reported in Sec. 4.6.

4.2 Data and Monte Carlo samples

The data used in this analysis have been collected by the CMS detector during pp (proton-proton) collisions at the LHC in 2018. Specifically, the data were recorded in a special high rate ($\sim$ 3 kHz) stream which, due to resource management, did not undergo the prompt reconstruction step and is known as `ParkingBPH` or `BParked dataset` [24] (see appendix: A).

The BParked dataset was recorded with a special set of trigger paths firing on events in which a displaced muon is present, targeting the production of B mesons decaying to a muon in the final state. The HLT paths present in ParkingBPH are named `HLT_MuX_IPY`, where `X` and `Y` represent respectively the threshold in transverse momentum (p_{T}) in GeV (requested at L1 and HLT) and impact parameter significance (IP, HLT only) applied on the muon at trigger level. The triggers p_{T} threshold varies from 7 GeV to 12 GeV and the IP from 4 to 6. In order to meet the CMS data acquisition system requirements, these paths, which are always present in the trigger menu, are artificially and independently switched off in portions of the LHC fill imposing a pre-scale factor of 0. As a result, each HLT path has a different integrated luminosity: the set of active triggers evolves allowing non-zero prescale for looser paths as the instantaneous luminosity falls. Each path is further divided into $n = 5$ (6) parts for eras C and D (A and B, the switch from 5 to 6 happened halfway through era B) by the creation of n copies of the same trigger path (appending `_partZ` at the end of the name, `Z`=1, ... n). Each copy is prescaled a factor n in order to

retain the totality of the triggered events.

The usual `DoubleMuon+Track` HLT used in the usual b-physics analysis, has a low reconstruction efficiency for this particular channel because the track requirement removes all the events where a daughter particle "fly" away from the b-hadron secondary vertex (K^0, in this case). The BParked dataset solves this problem, by removing the track requirement, opening the possibility to make measurements on rare channels such as $B^0 \to K^0_S \mu^- \mu^+$, or $\Lambda^0_b \to \Lambda \mu^+ \mu^-$. Using the BParked data set adds an extra challenge to the analysis, due to the difficulty to calculate adequately the trigger efficiency (it is a dynamic trigger that turns on and off during data taking), that discussion is carried out in the appendix A.

In this analysis, the selection requires a trigger muon (match between the HLT path and the muon object), but there is no selection for a specific HLT path, so the whole `ParkingBPH` dataset is used. The input format for data is the centrally produced MINIAOD which is then further analyzed using the global tag `102X_dataRun2_v11` and the certified luminosity mask [1].

Datasets and their corresponding luminosities are listed in table 4.1.

Table 4.1: Dataset name, run ranges, and the corresponding luminosity for the B Parked dataset.

Data set	Delivered Luminosity
/ParkingBPH[1-5]/Run2018D-05May2019promptD-v1/MINIAOD	26.50 fb^{-1}
/ParkingBPH[1-5]/Run2018C-05May2019-v1/MINIAOD	5.51 fb^{-1}
/ParkingBPH[1-6]/Run2018B-05May2019-v2/MINIAOD	4.93 fb^{-1}
/ParkingBPH[1-6]/Run2018A-05May2019-v1/MINIAOD	4.64 fb^{-1}
Sum	41.58 fb^{-1}

For efficiency studies and selection optimization, simulated samples are generated with `PYTHIA` 8 [73] for the production and hadronization processes. After that, `EVTGEN` [59] was used to simulate the decay of the b hadrons. The QED final state radiation is modeled with `PHOTOS` [44]. The generated events were processed through a simulation of the CMS detector, using the `GEANT4` package [17], with the exact same trigger and reconstruction algorithms used in the recorded data. All the simulated events include several proton-proton interactions in the same or nearby beam crossings (called pileup), with a distribution matching the one observed in the data.

[1] https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/Collisions18/13TeV/ReRecoCert_314472-325175_13TeV_17SeptEarlyReReco2018ABC_PromptEraD_Collisions18_JSON.txt

These MC samples are public and can be found on the CMS Data Aggregation System (DAS). Due to the dynamic nature of the trigger menu in data, the MC has kinematic differences with respect to the data. This is discussed in appendix A.

Control (resonant channel $B^0 \to K^0_S J/\psi(\mu^-\mu^+)$) and analysis (non-resonant channel $B^0 \to K^0_S \mu^-\mu^+$) MC samples were used to determine the total (absolute) efficiency as well as the angular efficiency as a function of $\cos\theta_l$. The samples can be found in DAS as:

- `BdToK0sMuMu_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`

  `RunIIAutumn18MiniAOD-PUPoissonAve20_BParking_102X_upgrade2018_realistic_v15-v`

- `BdToK0sJPsi_JPsiToMuMu_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`

  `RunIIAutumn18MiniAOD-PUPoissonAve20_BParking_102X_upgrade2018_realistic_v15-v`

4.3 Reconstruction and Selection

The process $B^0 \to K^0_S \mu^-\mu^+$, and the $B^0 \to K^0_S J/\psi(\mu^-\mu^+)$ control channel, were reconstructed in the final charged states $J/\psi \to \mu^+\mu^-$ and $K^0_s \to \pi^+\pi^-$.

Each event is required to pass any of the HLT paths from the BParked menu.

4.3.1 Dimuon selection

The $\mu^+\mu^-$ offline selection cuts can be summarized as follows:

- Leading muon $p_T > 7.0$ GeV;
- Trailing muon $p_T > 3.0$ GeV;
- Both muon impact parameter significance with respect to the chosen primary vertex $\frac{IP}{\sigma_{IP}} > 3.0$;
- Both muon $|\eta| < 2.4$;
- Both muon pass the `soft muon` ID criteria.

Both muons are used to perform a `KinematicVertexFit` (Appendix B) without any mass constraint. The invariant mass of the dimuon pair must have a vertex probability greater than 1.0 %.

4.3.2 K^0_S (Displaced Vertex) Selection

The K^0_S is extracted from the `SlimmedKs0Vertices` container produced by the CMS software framework (CMSSW). It is reconstructed by joining two opposite-charged tracks passing the following conditions:

- $p_T > 0.55$ GeV;
- `numberOfValidHits` > 2;
- $\frac{IP}{\sigma_{IP}} > 2.0$;
- `HighPurity` ID criteria;

Then, a `KinemmaticVertexFit` (appendix B) is done including a mass constraint for the $\pi^+\pi^-$ candidate to the K^0_S mass hypothesis. For an extra background sample (used to train a Boosted Decision Tree (BDT) with higher statistics), data is extracted emulating this process with the `lostTracks` and `packedPFCandidates` containers in CMSSW.

B^0 Reconstruction

The B^0 hadron is reconstructed using again a `KinematicVertexFit` which uses the dimuon and K^0_S candidates. The B^0 candidate must have a secondary vertex (SV) probability $> 2\%$. Once a B^0 candidate is reconstructed, there is a selection of the primary vertex (PV) by the best pointing angle α_{B^0}, the angle α_{B^0} between the line that joins the PV and SV and the B^0 momentum. This quantity is also discussed in section 3.4, and shown in Fig. 3.1 . Also, we calculate the pointing angle of the K^0_S with respect to the SV, $\alpha_{K^0_S}$. We require:

- $\cos\alpha_B > 0.8$;
- $\cos\alpha_{K_s} > 0.8$.

Resonances and Anti-Radiation Veto

The discrimination between the signal ($B^0 \to K^0_S\,\mu^-\mu^+$) and the two resonant channels ($B^0 \to K^0_S\,J/\psi\,(\mu^-\mu^+)$ and $B^0 \to K^0_S\psi'(\mu^-\mu^+)$), is done by excluding mass windows in the dimuon invariant mass, $q = m(\mu^+\mu^-)$. The mass window was estimated using a mass fit to the J/ψ resonance in a significant subsample (10M for J/ψ and 500K for ψ'). The J/ψ and the ψ' resonances fits are shown in figure 4.2.

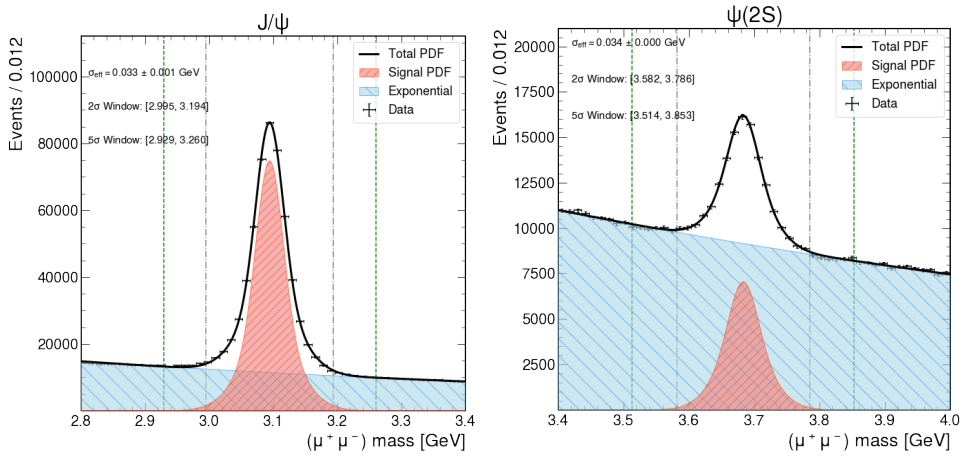

Figure 4.2: Resonant fits, using a double Gaussian plus exponential model in both cases. The mass windows are estimated using $\sigma_{eff} = \sqrt[2]{f\sigma_1^2 + (1-f)\sigma_2^2}$ where f, is the fraction of the first Gaussian.

The J/ψ resonance is vetoed with a $5\sigma_{eff}$ window and the $\psi(2S)$ with a $3\sigma_{eff}$ window. The antiradiation veto is taken from the studies made on other CMS studies (AN-21-020), which suppress resonant decays where the muons lose energy through radiation, as shown in Fig. 4.3. Vetoes are summarized in table 4.2.

Table 4.2: Resonances vetoes, and anti-radiation veto.

Resonance	Mass veto [GeV]	Anti-radiation veto (only if q is bellow the resonance mass)[GeV]
J/ψ	[2.92, 3.26]	$\lvert(m(\mathrm{B}^0)_{\mathrm{PDG}} - m(\mathrm{K}^0_S\mu^-\mu^+)) - (m(J/\psi)_{\mathrm{PDG}} - m(\mu^-\mu^+))\rvert \leq 0.130$
$\psi(2S)$	[3.5841, 3.7881]	$\lvert(m(\mathrm{B}^0)_{\mathrm{PDG}} - m(\mathrm{K}^0_S\mu^+\mu^-)) - (m(\psi(2S))_{\mathrm{PDG}} - m(\mu^-\mu^+))\rvert \leq 0.085$

4.3.3 XGBoost Classifier

The final selection is based on an XGBoost classifier (BDT) [33].The nonresonant MC is used as the signal and mass sidebands in data samples as background to train our classifiers[2]. The variables used as features to train the classifier are:

- cosalfaB2D: cosine of the angle between the PV and the SV calculated in the transversal plane $x - y$;
- cosalfaKs2D: cosine of the angle between the SV and the K^0_S vertex, calculated in the transversal plane $x - y$;

[2]These samples were fed with K^0_S candidates reconstructed with tracks from `lostTracks` and `packedPFCandidates` CMSSW track containers.

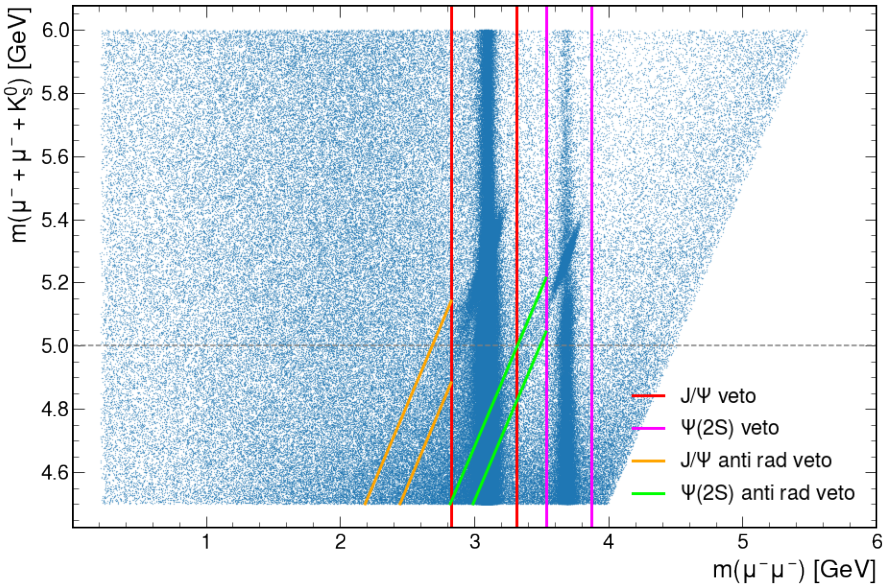

Figure 4.3: Dimuon invariant mass region in data.

- ksBdl: K^0_S proper decay length (PDL = ct, where t is the proper decay time, see below) calculated from the SV to the K^0_S vertex;
- sigLxyB: Significance of the SV, defined as the distance from the PV to the SV divided by the SV uncertainties (obtained in the kinematic vertex fit, appendix B);
- Bdl: B^0 PDL, calculated from the PV to the SV;
- priVtxCL: PV confidence level;
- lm_pt: Leading muon p_T;
- tm_pt: Trailing muon p_T;
- B_pT: B^0 candidate p_T;
- PVTriggDz: Smallest distance measured on the longitudinal coordinate (z) from the PV to a trigger muon;
- muon_dca: Distance of closest approach between the two muons;
- Bprob: SV probability.

- Jprob: dimuon pair vertex probability.
- Ks0prob: K^0_S candidate vertex probability.

The proper decay length is defined as $l_{xy}\frac{M}{X_{p_T}}$, where l_{xy} is the distance between the production and decay vertices of the particle, projected on the transverse plane, X_{p_T} is the p_T of the particle, and M is its mass. The correlation between variables is shown in Fig. 4.4.

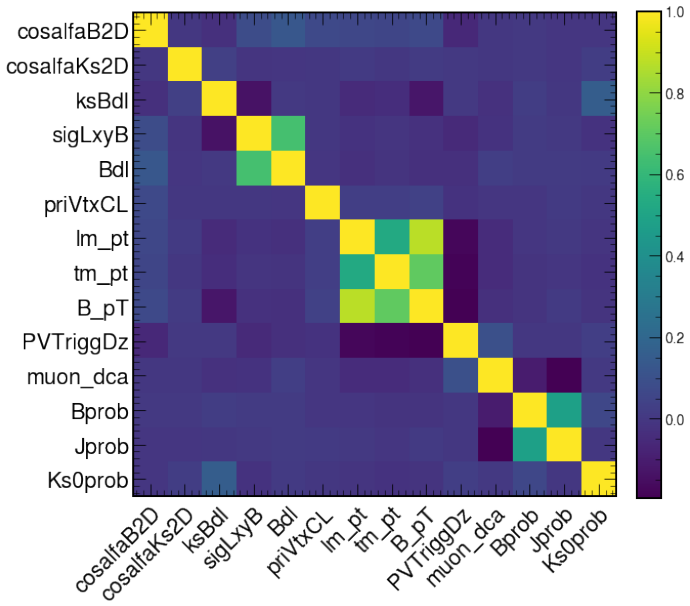

Figure 4.4: Correlation matrix of the variables used for the XGB training. These are calculated on MC true signal.

The XGB classification worked with one signal sample taken from MC ($B^0 \to K^0_S \mu^- \mu^+$) and two background samples: the first one made of the left mass-sideband which is a sample taken from data that had passed the preselection cuts and the B^0 candidate has an invariant mass less than 5.15 GeV; and similarly the right-sideband, from B^0 candidates with an invariant mass larger than 5.40 GeV, above the B^0 resonance.

Two models were trained, each with one sideband sample as background and the true matched $B^0 \to K^0_S \mu^- \mu^+$ MC as signal. The idea behind using two samples of background is based on the fact that each background sample is composed of events from different phenomena: the left sideband is composed of majorly partially recon-

structed B hadrons, and the right sideband is largely composed of combinatorial background.

4.3.4 Final Selection

To choose the working points for the XGB variables, an iterative method to find the optimal cut is used. The figure of merit (FOM) $FOM = \frac{S}{\sqrt{S+B}}$ is chosen as a metric to evaluate the best working points. The following steps describe how this process flows:

- Evaluate the FOM without any cut as a control point.
- Evaluate the FOM in a series of sequential increasing cuts on the left sideband XGB variable.
- Select the best cut (largest FOM value) on the left sideband XGB model value and fix that working point, Wp_1.
- Evaluate the FOM in a series of sequential increasing cuts on the right sideband XGB variable (all the evaluations are done with fixed in Wp_1).
- Select the best cut on the right sideband XGB model, Wp_2.
- Now fix Wp_2 and evaluate again Wp_1.
- Repeat the search for Wp_2 fixing Wp_1.

This iteration is done over 3 cycles to find the best cuts on the XGB variables. A graphical scheme of these iteration is shown in Fig. 4.6.

The best cut values for the left sideband model were found to be 0.39 and 0.92 for the right sideband model. The final selection is the trigger selection, which requires at least one muon fires a BParked trigger.

4.4 Angular Analysis Modeling

The 2-dimensional probability density function (P.D.F.) used in the angular analysis is

$$\text{P.D.F.}(m, \theta_l, F_H) = Y_S[S_m(m)S_A(\theta_l, F_H)\epsilon_A(\theta_l)] + Y_B[B_m(m)B_A(\theta_l)], \tag{4.4}$$

where:

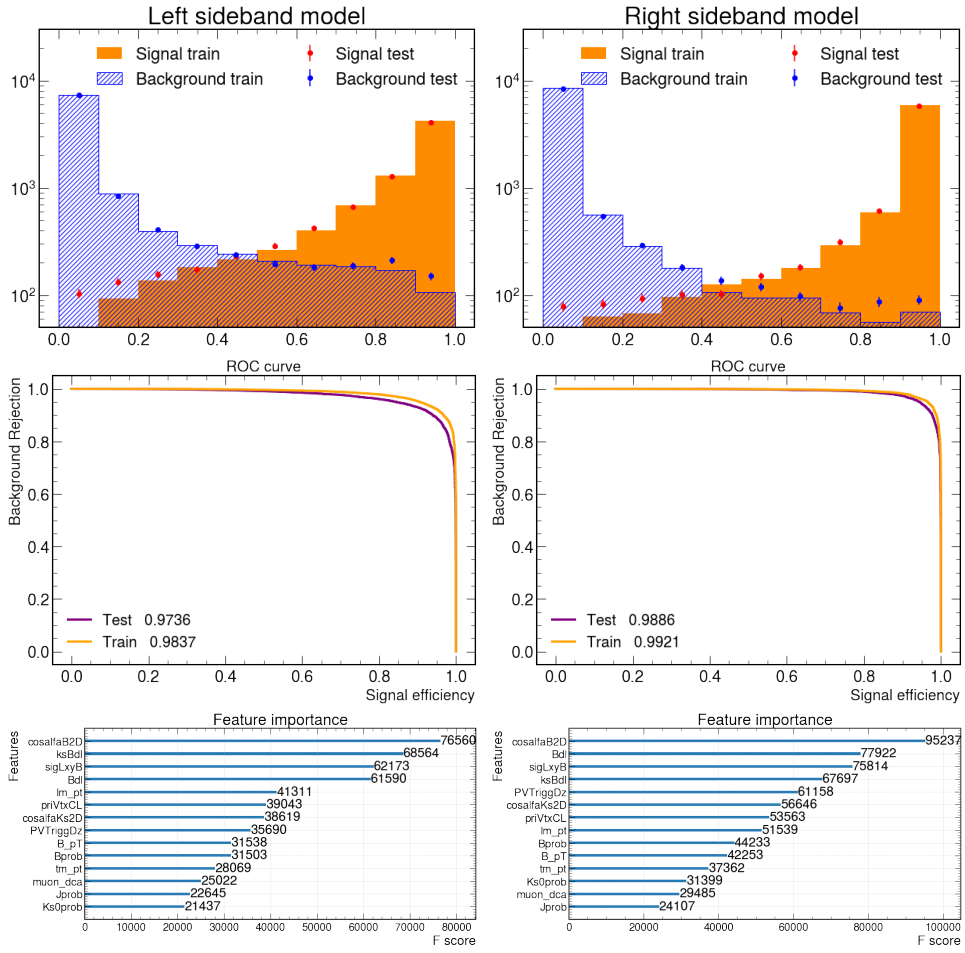

Figure 4.5: XGB performance. The left column are the left mass sideband model results, and the right column are the right mass sideband model results. The first row are the signal/background scores, for the training and testing samples. The second row is the comparison of the ROC curve on the training and testing samples. The last row is the feature (variable) importance for the model.

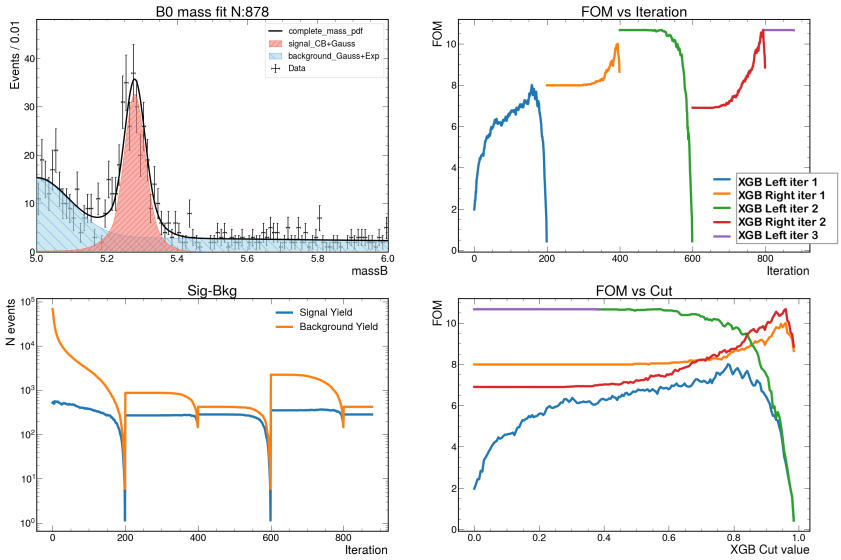

Figure 4.6: FOM iteration algorithm results.

- Y_S is the signal yield;
- S_m is the mass signal model (Gaussian + Crystal Ball);
- m is the invariant mass of $\mu^- + \mu^+ + \mathrm{K}^0_S$;
- S_A is the angular signal model Eq. 4.2;
- θ_l is the angle formed between the K^0_S and μ^+;
- F_H is the parameter described in section 4.1;
- ϵ_A is the angular efficiency;
- Y_B is the background yield;
- B_m is the mass background model (Gaussian for partially reconstructed B hadrons + exponential for combinatorial background);
- B_A is the angular background model (2 Bernstein polynomials one for each mass sideband).

The signal S_m shape is extracted from MC, so the parameters for the fit in data are Gaussian-constrained to the parameters extracted from MC. The angular efficiency is

also taken from the corrected MC (through the accept-reject method and scale factors, see appendix A). The MC sample is generated with a flat angular shape (phase space). The reconstruction distorts the flat shape and, the new distribution is representative of the angular efficiency Fig. 4.7 (ϵ_A). As shown in Eq. 4.4, the acceptance will be multiplied by the signal shape for the final fit.

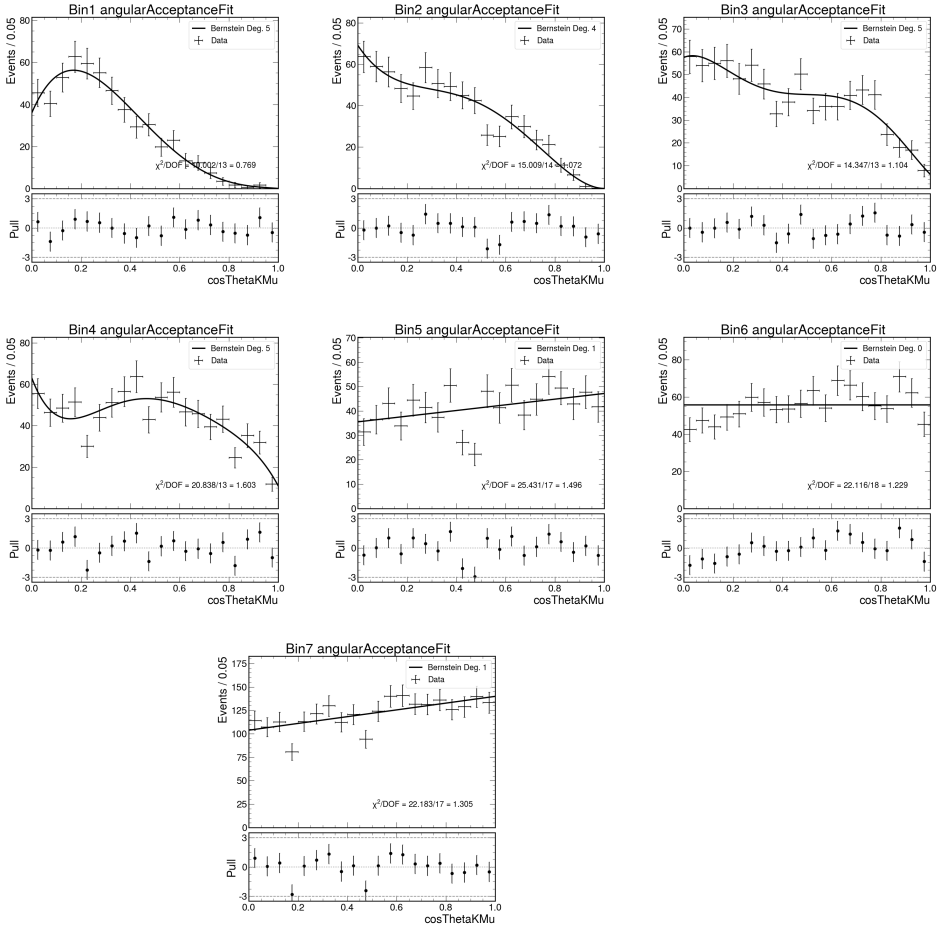

Figure 4.7: Angular efficiency plots. Obtained from MC signal after all selection requirements and corrections in bins of q^2.

The angular background is modeled as the sum of two Bernstein polynomials, one (P_1) taken from left mass sideband of B^0 candidates (a mass range of [5.0, 5.15] GeV, which is largely composed of partially reconstructed B hadrons), and the other

(P_2) taken from the right mass sideband (mass range of [5.40, 6.0] GeV, composed by combinatorial background): $f * p_1 + (1 - f) * p_2$, where the fraction f is shared for both, angular and mass background models. Each mass sideband is fitted to an n-dimensional Bernstein polynomial, and the degree of the polynomial is chosen by fitting different degrees and finding the best (χ^2/n_{dof}).

To measure on real data the fitting procedure is as follows:

- perform a fit to the mass observable and obtain the mass model parameters Fig. 4.8;
- extract the left and right mass sidebands events and plot these background regions on the angular variable.
- obtain the best angular models (χ^2/n_{dof}) that fit the mass sidebands backgrounds, figures 4.9 and 4.10.
- Fit the data to the complete model 2-D and estimate, through the sequence Minos-HESSE, the F_H value (Fig. 4.11).

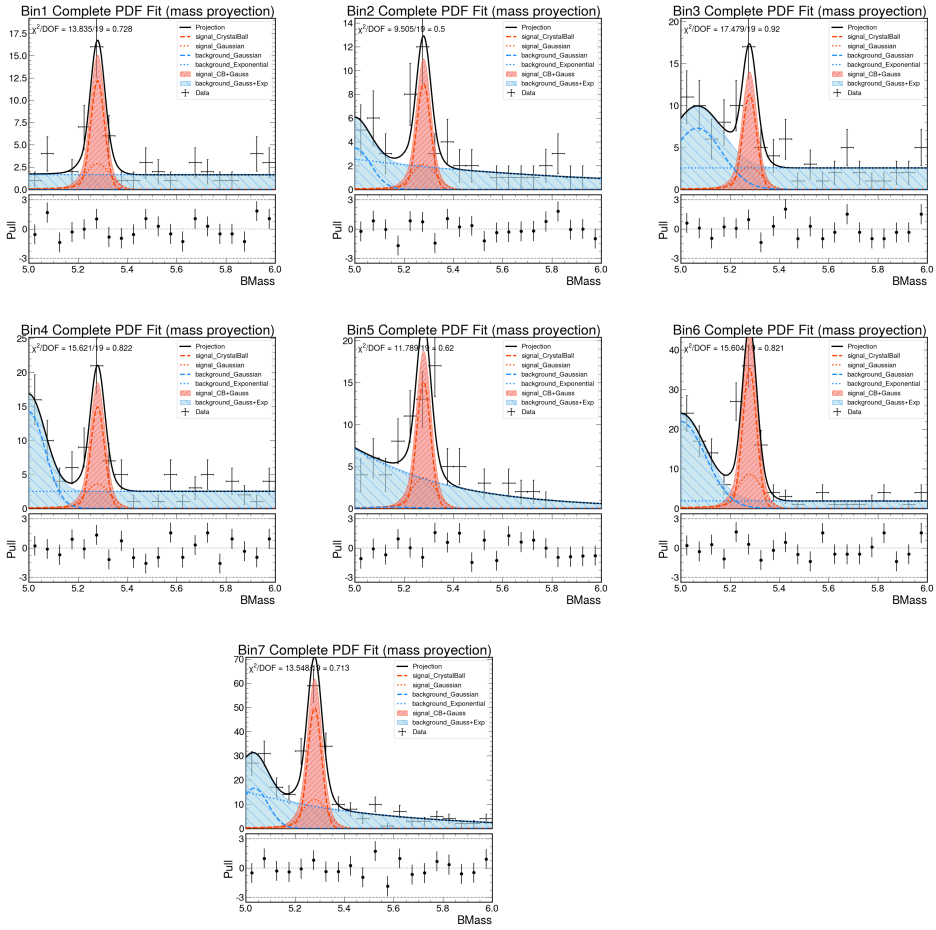

Figure 4.8: Mass projections after a 2-D maximum likelihood fit $(m, \cos\theta_l)$ in bins of q^2.

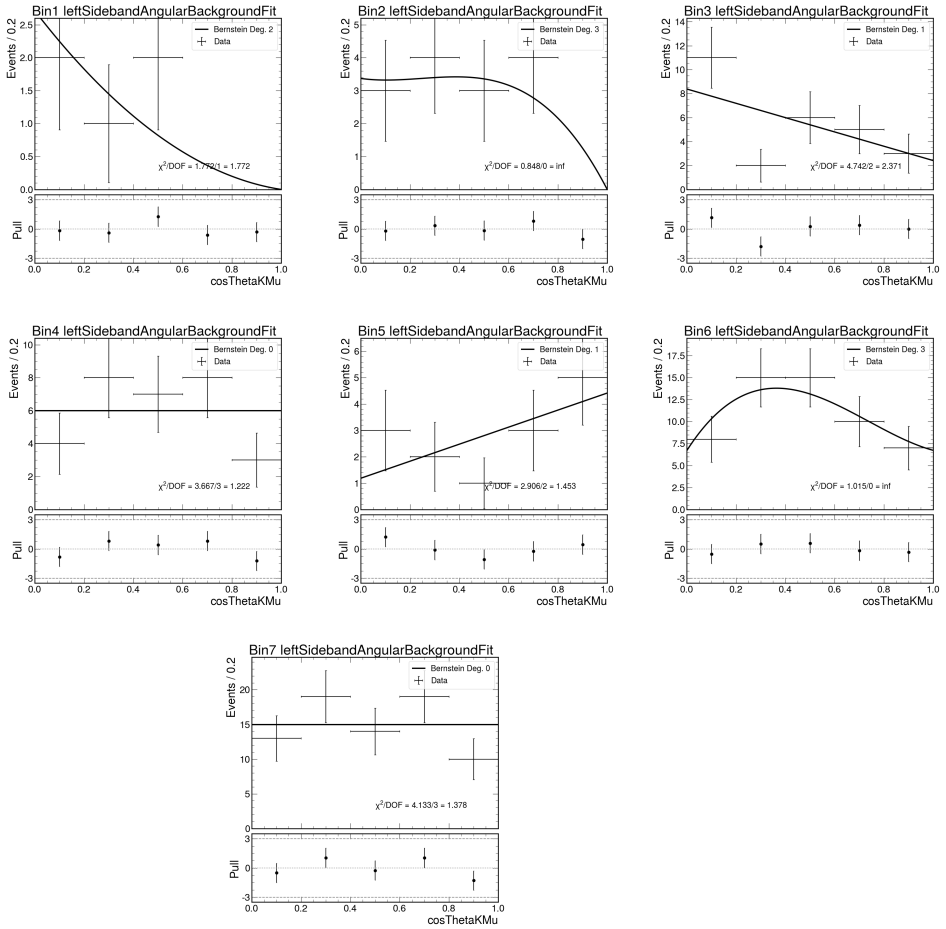

Figure 4.9: Left mass sideband fits in the angular variable $\cos\theta_l$ in bins of q^2.

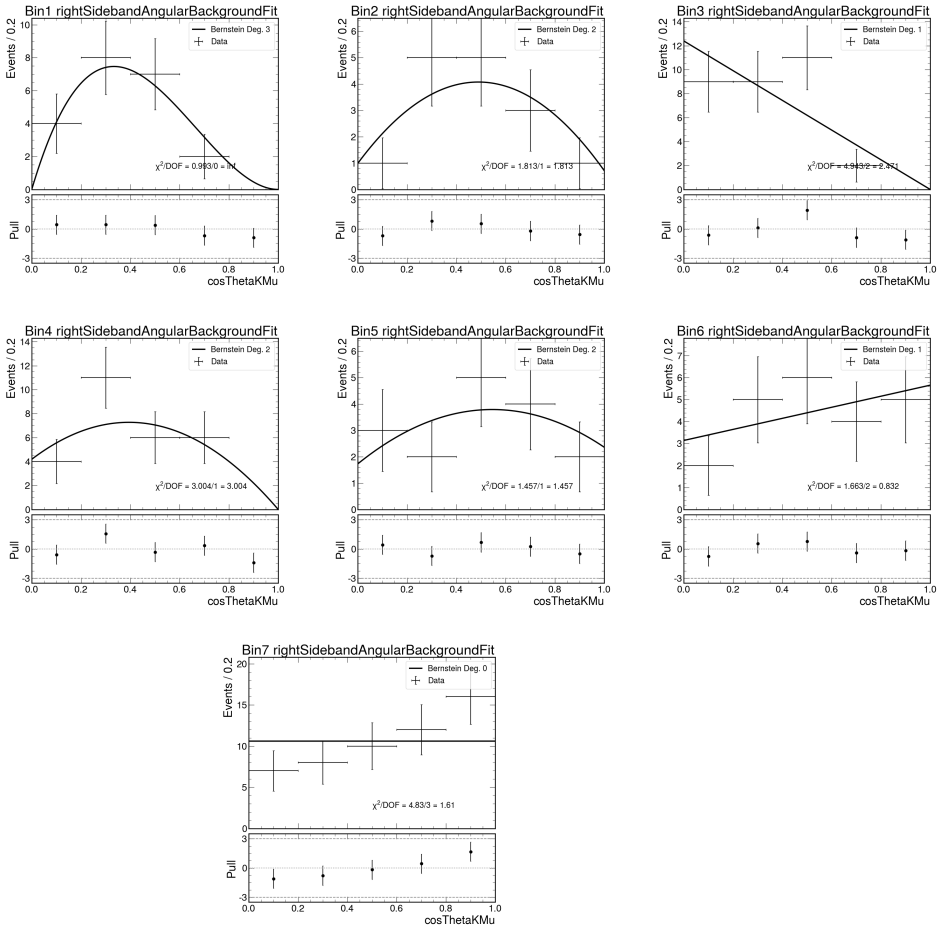

Figure 4.10: Right mass sideband fits in the angular variable $\cos\theta_l$ in bins of q^2.

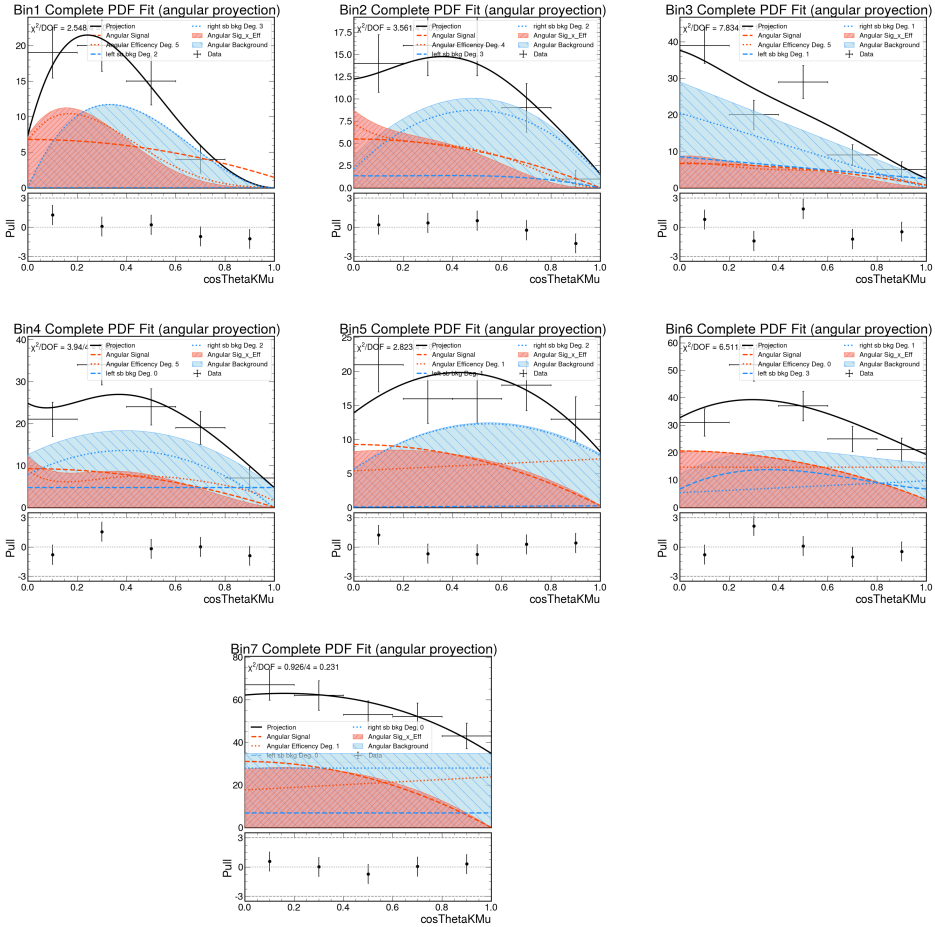

Figure 4.11: Angular fit projection after a 2-D maximum likelihood fit $(m, \cos\theta_l)$ in bins of q^2.

4.5 Confidence Interval Estimation

4.5.1 Neyman Classical Confidence Interval Construction

Classical confidence intervals are the traditional way in which scientists report uncertainties of the results of their experiments. These confidence intervals can be computed through approximations such as the likelihood ratio or with the classical construction proposed by Neyman [65], which can be oversimplified as "*Choose an ordering rule, build a confidence belt, invert the confidence belt and then get a confidence interval*".

These confidence intervals require a method called "confidence belts", build for every measurement of a quantity and one unknown parameter. Fig. 4.12 shows this construction, of the interest parameter μ vs. the measured quantity x (observable). For each value of μ, one can evaluate $P(x|\mu)$ along the horizontal line through μ. The interval $[x_1, x_2]$ is a sub set of this line, it is selected such that

$$P(x \in [x_1, x_2]|\mu) = \int_{x_1}^{x_2} P(x|\mu)\, dx = \alpha. \tag{4.5}$$

These intervals are drawn as horizontal line segments, there is one line for every value of μ. The interval $[x_1, x_2]$ is called the "acceptance region" or "acceptance interval", for μ. To "choose" uniquely a certain acceptance region, there are common choices like [3]:

$$P(x < x_1|\mu) = 1 - \alpha, \tag{4.6}$$

which leads to "upper confidence limits"(satisfy $P(\mu > \mu_2) = 1 - \alpha$); and

$$P(x < x_1|\mu) = P(x > x_2|\mu) = (1 - \alpha)/2, \tag{4.7}$$

which leads to "central confidence intervals" (this ones satisfy $P(\mu < \mu_1) = P(\mu > \mu_2) = (1-\alpha)/2$). To complete the construction, horizontal acceptance intervals should be drawn for every value of μ. Once there there is an experiment to measure x, with result value x_0, then a vertical line in x_0 is drawn and the confidence interval $[\mu_1, \mu_2]$ will be given by the union of values of μ for which the corresponding horizontal interval is intercepted by the vertical line.

The confidence belt has the property, that as long as the interval $[x_1, x_2]$ fulfills equation 4.5, each vertical intersection gives an interval of coverage α (See Figure 4.12, a vertical intersection refers to a line in the μ direction). But, as pointed out by many

[3] There is total freedom to make this choice, these are common choices when there is no influence from the data [48]

authors, equation 4.5 does not determine a unique pair $[x_1, x_2]$; quite the contrary, there is a big family of possible pairs that satisfy Eq. 4.5. That is where the ordering rule is essential since it gives a unique solution for x_1 and x_2. The classical intervals are obtained by including the integration intervals dx in order of their probability. This yields the additional requirement that

$$P(x_1|\mu) = P(x_2|\mu). \tag{4.8}$$

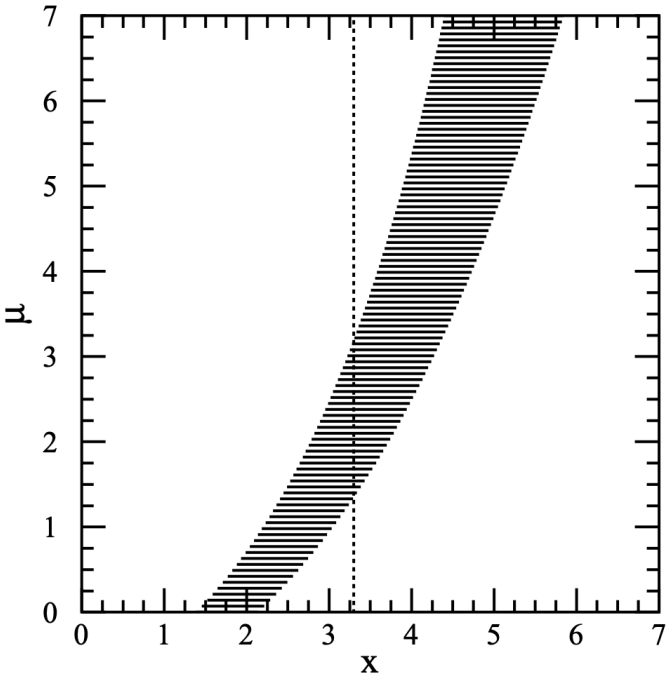

Figure 4.12: Figure made by Feldman and Cousins in Ref. [47]. It illustrates a generic confidence belt construction (all the horizontal lines produce an "area"). For any value of μ it is possible to draw a horizontal acceptance interval $[x_1, x_2]$ such that $P(x \in [x_1, x_2]|\mu) = \alpha$ (Eq. 4.5). After that, an experiment is performed to measure x, say the value x_0 is measured, then a vertical line (dashed line) is drawn at x_0. The confidence interval $[\mu_1, \mu_2]$ is formally built as the union of all the values of μ for which the corresponding acceptance is intercepted by the vertical line.

However, this rule to get confidence intervals is problematic for certain distributions because it could lead to empty sets for confidence intervals (when the confidence belt is built, certain "measured values" does not intersect with the obtained confidence belt,

leading to empty confidence intervals, see Figure 1 of Ref. [47]) or unphysical confidence intervals (negative regions for a mass measurement). For these cases, a "flip-flop" policy, is exemplified in Ref. [47] as "*If my measured value of a physically positive quantity is negative, I will pretend that I measured zero when quoting a confidence interval*". This is also problematic because it leads to incorrect confidence intervals that do not match the quoted coverage.

4.5.2 Feldman Cousins Method

Feldman and Cousins [47] developed a method for computing confidence intervals with well-defined frequentist coverage. It solves the aforementioned issues with a classical construction, offering a transition between upper (or lower) limits and two-sided confidence intervals. The method defines an alternative ordering principle, such that dx is included in order of decreasing likelihood ratio

$$R(x, \mu) = \frac{\mathcal{L}(\mu|x)}{\mathcal{L}(\mu_{\text{best}}|x)}, \tag{4.9}$$

were μ_{best} is the value that maximizes the likelihood $\mathcal{L}(\mu|x)$ and is in the allowed (physical) region (notice, that in Eq. 4.9, we interchanged $P(x|\mu)$ for $\mathcal{L}(\mu|x)$ [4]). Using the likelihood ratio ordering rule is possible to find the interval $[x_1, x_2]$ for a given value of μ such that

$$R(x_1) = R(x_2), \tag{4.10}$$

and at the same time equation 4.5 remains true. Feldman and Cousins made sure to clarify that in the case of discrete probability distributions such as the Poisson distribution, these requirements can not be met. [47].

4.5.3 Karbach Algoritmic Apporach

The original publication by Feldman and Cousins [47] provides a recipe to perform a numerical computation based on toy Monte-Carlo to solve equations 4.5 to 4.10 and get a confidence region. Later on, Karbach published [55] the Feldman-Cousins method

[4] This change is very subtle but essential, given that the first represents a probability distribution and the latter the likelihood. Although the likelihood is numerically identical to the probability density function $P(x|\mu)$ it can be both interpreted as a function of x or μ; if it is interpreted as a function of x is the probability density function; if it is interpreted as a function of the parameter of interest μ then it is called likelihood. It may be clear if it is generalized for a vector $x = (x^1, ..., x^n)$ that represents a set of different values, i.e. more than a single experimental result. The likelihood is built as the product of probability distributions $\mathcal{L}(\mu|x) = \prod_{i=1}^{n} P(x^i|\mu)$, in which the likelihood is used to find the "best" μ for a given data $x = (x^1, ..., x^n)$.

in an algorithmic way (the computed confidence interval for this angular analysis is based on Karbach's work, but the branching fraction analysis is not). The algorithm used in this work was written by Antonio Cota for his master's thesis Ref. [42].

Karbach's algorithm [55] consists of the construction of the confidence belts using Neymann's ordering rule and the Feldman-Cousins ordering rule, to build $1 - CL$ curves (Confidence Level) using the likelihood ratio. Feldman and Cousins introduced a new ordering rule given by the likelihood ratio in Eq. 4.9, i.e. for a given P.D.F. $f(x|\mu)$ with a parameter μ, the likelihood ratio used as ordering rule

$$R = \frac{\mathcal{L}(\mu|x)}{\mathcal{L}(\mu_{\text{best}}|x)} > k_\alpha \tag{4.11}$$

where k_α determines the acceptance interval C_α such that the integral in Eq. 4.5 over the interval C_α is fulfilled. The acceptance interval is given by all the values of x such that $C_\alpha = \{x|R > k_\alpha\}$, $R(x_{min}) = R(x_{max})$. This likelihood ratio will be used as a statistical test. Its distribution will be a comparison between real data and MC obtaining the p-value or $1 - CL$ value. $1 - CL$ plots display confidence regions for a specific value of x. This value is often the measured value. So to interpret let's say a 68% CL interval $[\mu_1, \mu_2]$, a horizontal line at $1 - CL = 1 - 0.68 = 0.32$ is drawn, and the interval $[\mu_1, \mu_2]$ is given by the intersection of this line and the generated curve.

The innovation of this method consists of treating the data sets as an ensemble of data, including nuisance parameters inside the likelihood function.

Taking the negative natural logarithm of equation 4.11 (this reduces the computational load and the minimization-maximization point is the same), this new quantity will be labeled $\Delta\chi^2_{toys}$

$$-2\ln R \equiv \Delta\chi^2_{toys} = -2\ln\mathcal{L}(\mu, x) - 2\ln\mathcal{L}(\mu_{best}, x). \tag{4.12}$$

This equation can be extended to consider nuisance parameters and, use an ensemble of data instead of a single measurement

$$R = \frac{\mathcal{L}(\mu, \theta_{fix\,best}|\vec{x})}{\mathcal{L}(\mu_{best}, \theta_{best}|\vec{x})}, \tag{4.13}$$

$$\Delta\chi^2_{toys} = -2\ln\mathcal{L}(\mu, \theta_{fix\,best}|\vec{x}) - 2\ln\mathcal{L}(\mu_{best}, \theta best|\vec{x}), \tag{4.14}$$

where $\theta_{fix\,best}$ is the value of the nuisance parameter θ that maximizes the likelihood when μ is fixed, whereas θ_{best} and μ_{best} are the values that maximize the likelihood given by the estimator when both parameters are allowed to fluctuate during the

maximization process.

Returning to the physical problem in this angular analysis. It is recalled that this is a measurement of the parameter F_H. Looking at equation 4.2, it is easily seen that the only physical region is $0 \leq F_H \leq 3$. The Feldman-Cousins algorithm is used to get confidence intervals that could achieve these goals:

- Physical boundaries: When the parameter F_H has its value restricted to some physical region ($0 \leq F_H \leq 3$), the classical confidence interval can be outside this region, but it is desirable that the confidence interval could be contained inside the physical region.
- Agreement with the punctual estimation: the confidence interval must contain the estimated value obtained in the fit.
- Validation with coverage: the basic idea is that, if the confidence intervals are correct, when toy MC experiments are carried out and the F_H parameter is calculated, 68.27% of the values will be within the confidence interval built at the 68.27% C.L.

Using Cota's implementation [42] of the Karbach algorithm [55] is possible to build curves $1 - C.L.$ to obtain confidence intervals $1 - \alpha \in [0, 1]$ in a graph associated with the F_H measurement. The algorithm used can be described as follows:

- Consider the value F_{H_0} the true value (this value is assumed, and used to calculate the $1 - CL$ value for a given F_H). Generate N random samples of size n using the distribution probability corresponding to the true value F_{H_0}.
- Compute $\Delta\chi^2_{toy}$ for each of the N toy MC samples.
- Compute $\Delta\chi^2_{data}(\vec{x}, F_{H_0})$ from the observed data $\vec{x}_{obs}$.
- The value $1 - CL$ is computed as the fraction of toy MC ($\Delta\chi^2_{toy}$) bigger than the true value (measured value):

$$1 - CL = \frac{N[\Delta\chi^2_{data}(F_{H_0}, \vec{x}_{obs}) < \Delta\chi^2_{toy}(F_{H_0}, \vec{x})]}{N}. \tag{4.15}$$

The physical restriction on the F_H parameter is included in every single minimization. As mentioned in the first step, the F_{H_0} value gradually changes in every iteration of the algorithm, trying to scan the whole region of interest, creating a curve of $1 - CL$ values (y-axis) for every F_{H_0} (x-axis), figures 4.13 - 4.15.

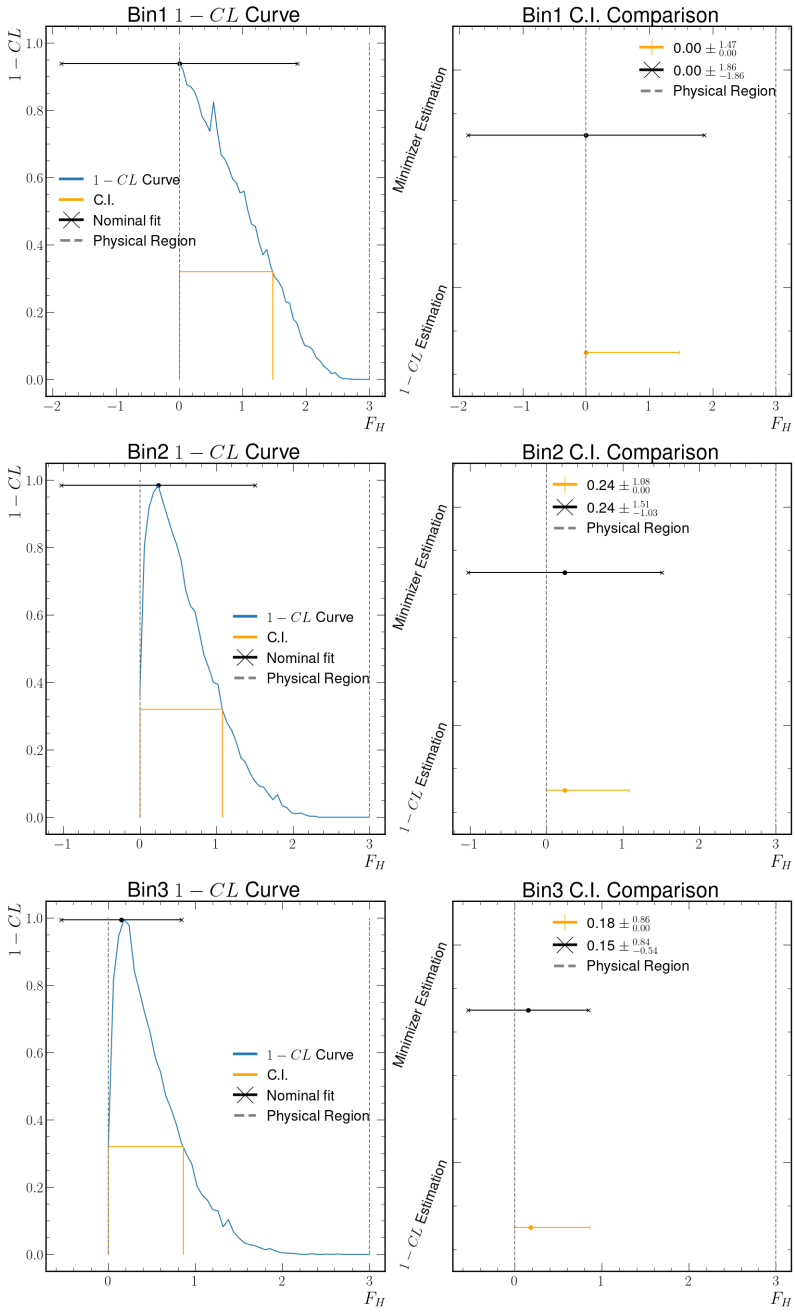

Figure 4.13: Confidence intervals calculated using Karbach-Cota method taking the results from the angular fits in bins of q^2. Only bins 1,2 and 3 are shown

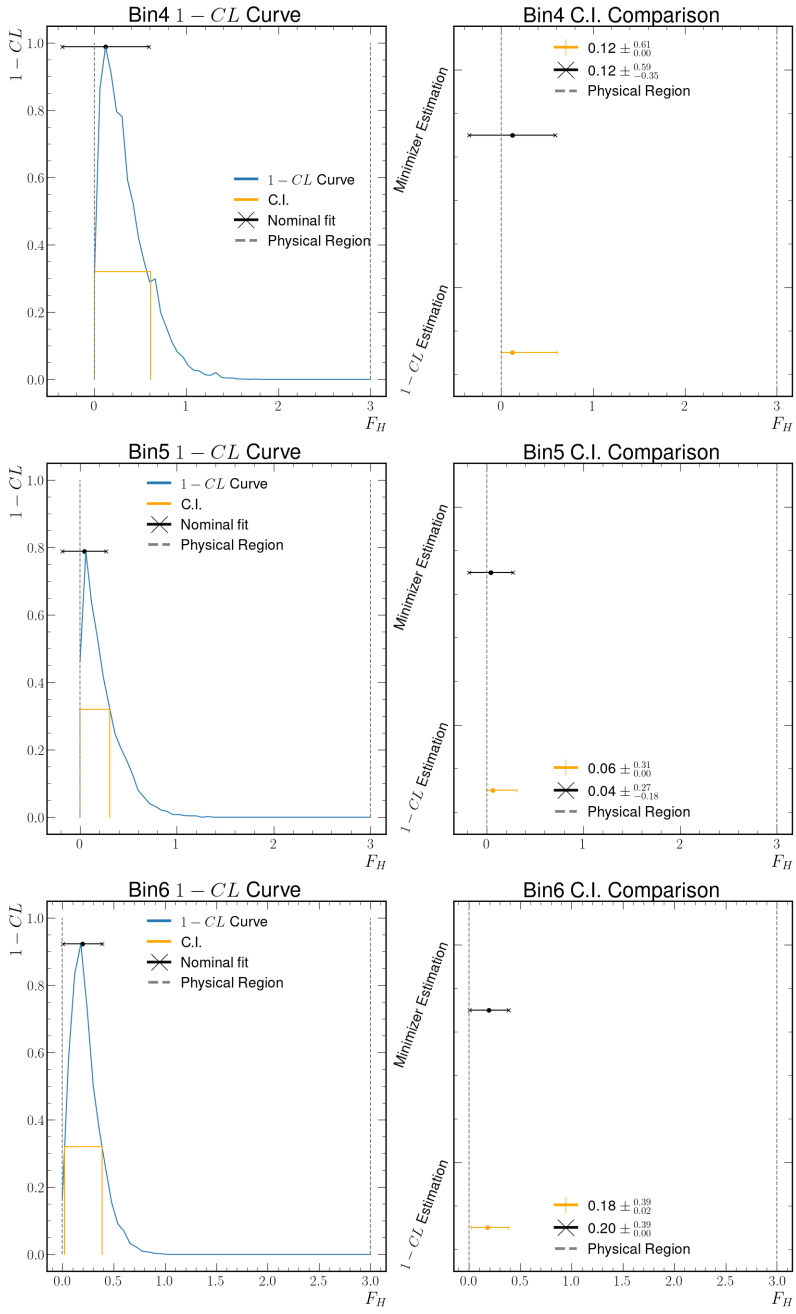

Figure 4.14: Confidence intervals calculated using Karbach-Cota method taking the results from the angular fits in bins of q^2. Only bins 4,5 and 6 are shown

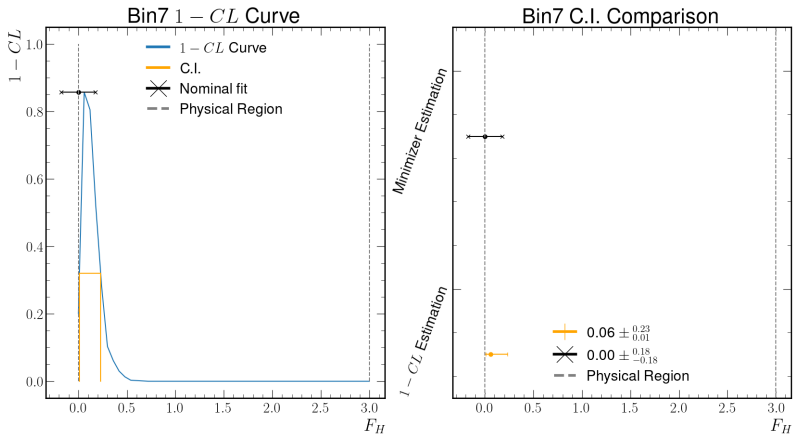

Figure 4.15: Confidence intervals calculated using Karbach-Cota method taking the results from the angular fits in bins of q^2. Only bin 7 is shown

After obtaining the $1 - C.L.$ curves, we have reliable confidence intervals[5]. It is possible to compare the statistical uncertainties with the current results Fig. 4.16, and the SM prediction (SM predictions obtained with `Flavio2.5.5` [74]).

[5] See Ref. [42], to see the coverage cross-check done to this method

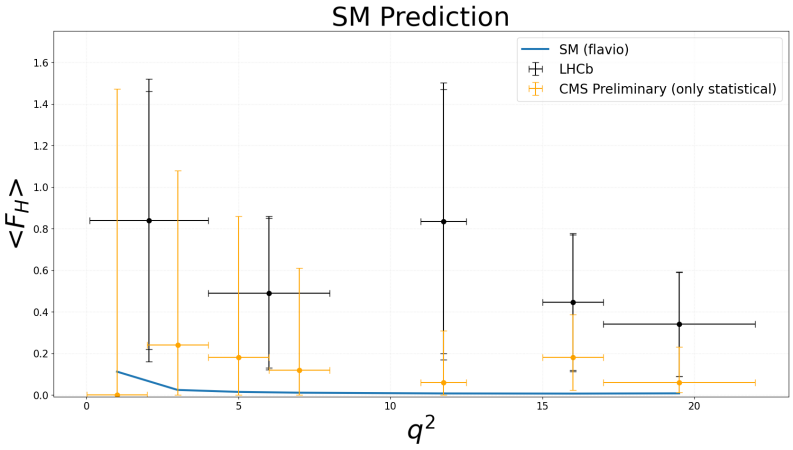

Figure 4.16: F_H parameter (preliminary measurement only statistical uncertainties) compared with the SM in bins of q^2.

4.6 Branching Ratio Analysis

The signal mode is normalized with respect to the corresponding resonant channel ($B^0 \to K^0_S J/\psi\,(\mu^-\mu^+)$). This channel has similar kinematic properties and the same final state particles, so the normalization results in an almost complete cancellation of systematic uncertainties when measuring the ratio of branching fractions. The differential branching fraction averaged over a q^2 bin of width $q^2_{\text{max}} - q^2_{\text{min}}$ is given by

$$\frac{d\mathcal{B}}{dq^2} = \frac{N(B^0 \to K^0_S \mu^-\mu^+)}{N(B^0 \to K^0_S J/\psi\,(\mu^-\mu^+))} \cdot \frac{\epsilon(B^0 \to K^0_S J/\psi\,(\mu^-\mu^+))}{\epsilon(B^0 \to K^0_S \mu^-\mu^+)} \cdot \frac{\mathcal{B}(B^0 \to J/\psi K^0_S)\mathcal{B}(J/\psi \to \mu^+\mu^-)}{(q^2_{\text{max}} - q^2_{\text{min}})}, \quad (4.16)$$

where $N(B^0 \to K^0_S \mu^-\mu^+)$ is the number of signal B candidates in the respective bin, $N(B^0 \to K^0_S J/\psi\,(\mu^-\mu^+))$ is the number of resonant candidates, $\epsilon(B^0 \to K^0_S J/\psi\,(\mu^-\mu^+))/\epsilon(B^0 \to K^0_S \mu^-\mu^+)$ is the relative efficiency between the resonant channel and the signal in the q^2 bin, and finally $\mathcal{B}$ are the branching fractions of the resonant B^0 decay and J/ψ into muons.

The P.D.F. function used to model the B^0 candidate invariant mass is used in both the resonant and nonresonant channels, which is the same mass model used in the angular analysis (section 4.4). The mass projection for Eq. 4.4 would be

$$\text{P.D.F.}(m) = Y_S[S_m(m)] + Y_B[B_m(m)], \quad (4.17)$$

where the signal mass shape S_m would be composed by the sum of a Gaussian and a Crystal Ball functions, while the background B_m mass model is made of a Gaussian (partially reconstructed B hadrons) and an exponential function (combinatorial background), the Y_S and Y_B are the corresponding yields to signal and background, Y_S will be extracted from all the fits to obtain the number of signal candidates. The mass fits per bin of q^2 can be seen in figure 4.8, and the resonant fit in figure 4.17.

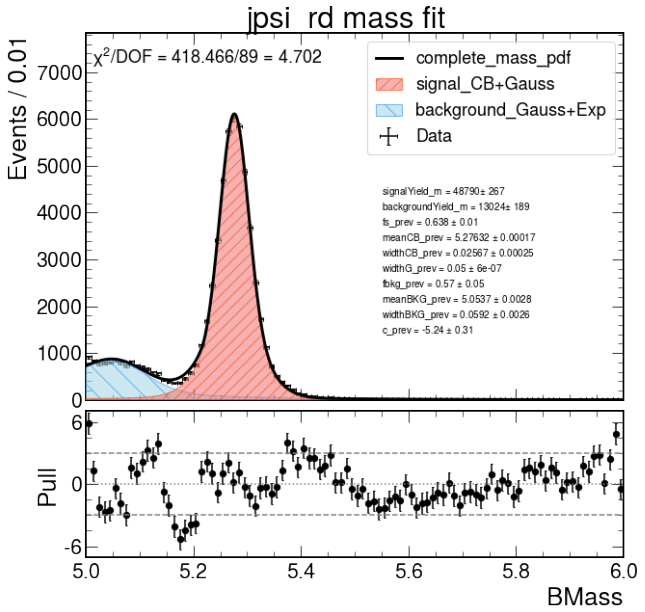

Figure 4.17: $B^0 \to K^0_S\, J/\psi\, (\mu^-\mu^+)$ invariant mass fit.

These fits do not fall into the same problem as the F_H measurement, because the yield fit will give us a positive number and the errors will be contained within the physical region (positive yield), so there is no need for a Feldman-Cousins construction of the confidence intervals. Due to a lack of time, only statistical uncertainties were explored, still, it is expected that most of the systematic uncertainties cancel out, and the statistics will be the main source of uncertainties. The results are shown in Fig. 4.18.

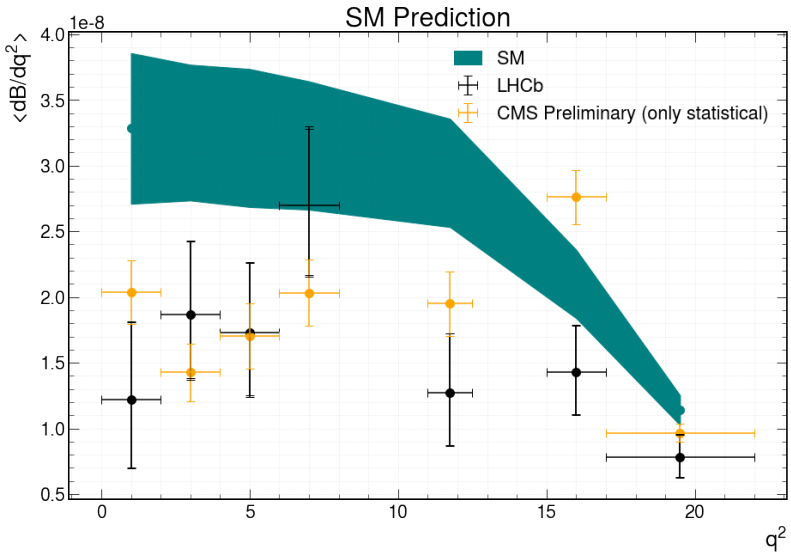

Figure 4.18: $B^0 \to K^0_s \mu^+ \mu^-$ differential branching ratio (preliminary results only statistical uncertainties) compared with the SM and the reported results from LHCb [5].

Chapter 5

Conclusions and Perspectives

5.0.1 Measurement of the dependence of the hadron production fraction ratio f_s/f_u and f_d/f_u

We have successfully measured the kinematic (p_{T} and rapidity) dependence of the ratio between the $\mathrm{B_s}^0$ and B^+ hadron production fractions: $f_{\mathrm{s}}/f_{\mathrm{u}}$. This analysis use data from the CMS experiment from 2018 with an integrated luminosity of 61.6 fb^{-1}. The $f_{\mathrm{s}}/f_{\mathrm{u}}$ ratio shows to be dependent on the B hadron p_{T} and consistent with being asymptotically constant at large p_{T} . In rapidity, no dependence was observed. In the case of the $f_{\mathrm{d}}/f_{\mathrm{u}}$ ratio, this is the first measurement in proton-proton collisions and is found to be within one standard deviation of the unity, independent of p_{T} and rapidity.

Tables 5.1 and 5.2 summarize the $\mathcal{R}_{\mathrm{s}}$ results for bins of p_{T} and $|y|$, respectively, also shown in Fig. 3.11. Tables 5.3 and 5.4 report $f_{\mathrm{d}}/f_{\mathrm{u}}$ results in bins of p_{T} and $|y|$, respectively, also shown in Fig. 3.12. Besides the central values, the tables include the statistical and systematic uncertainties for each bin. Not included in the tables an additional systematic uncertainty of 2.3% associated with the track reconstruction efficiency that applies to all results. Neither are included for the $f_{\mathrm{d}}/f_{\mathrm{u}}$ tables, the systematic uncertainties of 4.6% and 2.5% associated to uncertainties in the branching fractions from Eq. 3.4 and the $\mathcal{R}^{\pm,0}$ correction factor, respectively.

These results will help to improve the experimental uncertainties of the $\mathrm{B}_s \to \mu^+\mu^-$ and $\mathrm{B}^0 \to \mu^+\mu^-$ measurements. Achieving this is a significant step toward matching the experimental and theoretical uncertainties, and hence getting a more clear picture in the search for NP and SM deviations.

Table 5.1: Measured $\mathcal{R}_s$ values as a function of p_T , with statistical (σ_{stat}) and bin-to-bin systematic (σ_{sys}) uncertainties, in percent. Not included in the table is an additional systematic uncertainty of 2.3% that is common to all bins and is associated with the track reconstruction efficiency.

p_T (GeV)	$\mathcal{R}_s$	σ_{stat} (%)	σ_{sys} (%)
12–13	0.1314	2.1	3.1
13–14	0.1196	1.6	2.7
14–15	0.1165	1.3	2.4
15–16	0.1154	1.2	2.6
16–18	0.1135	0.8	2.6
18–20	0.1106	0.8	2.8
20–23	0.1105	0.7	2.9
23–26	0.1110	0.8	2.6
26–29	0.1091	0.9	3.2
29–34	0.1095	0.9	2.3
34–45	0.1088	0.9	2.8
45–70	0.1117	1.3	2.6

Table 5.2: Measured $\mathcal{R}_s$ values as a function of $|y|$, with statistical (σ_{stat}) and bin-to-bin systematic (σ_{sys}) uncertainties, in percent. Not included in the table is an additional systematic uncertainty of 2.3% that is common to all bins and is associated with the track reconstruction efficiency.

$\|y\|$	$\mathcal{R}_s$	σ_{stat} (%)	σ_{sys} (%)
0.00–0.25	0.1110	0.6	1.8
0.25–0.50	0.1104	0.6	2.0
0.50–0.75	0.1100	0.6	1.9
0.75–1.00	0.1126	0.7	3.9
1.00–1.30	0.1117	0.8	4.3
1.30–1.60	0.1069	1.0	4.3
1.60–2.40	0.1099	1.0	4.4

Table 5.3: Measured f_d/f_u values as a function of p_T , with the statistical (σ_{stat}) and bin-to-bin systematic (σ_{sys}) uncertainties, in percent. Not included in the table is an additional systematic uncertainty of 5.7% that is common to all bins and is the sum in quadrature of uncertainties associated with the track reconstruction efficiency, branching fractions from Eq. 3.4, and the $\mathcal{R}^{\pm,0}$ correction factor.

p_T (GeV)	f_d/f_u	σ_{stat} (%)	σ_{sys} (%)
12–13	0.975	3.9	7.8
13–14	1.074	3.4	5.6
14–15	0.988	2.8	4.4
15–16	0.994	1.7	3.3
16–18	0.993	1.1	3.9
18–20	1.018	1.3	3.8
20–23	1.032	1.4	5.1
23–26	1.025	1.4	4.1
26–29	1.037	1.1	4.6
29–34	1.024	1.1	3.8
34–45	0.981	0.8	2.5
45–70	0.965	1.9	2.4

Table 5.4: Measured f_d/f_u values as a function of $|y|$, with the statistical (σ_{stat}) and bin-to-bin systematic (σ_{sys}) uncertainties, in percent. Not included in the table is an additional systematic uncertainty of 5.7% that is common to all bins and is the sum in quadrature of uncertainties associated with the track reconstruction efficiency, branching fractions from Eq. 3.4, and the $\mathcal{R}^{\pm,0}$ correction factor.

$\|y\|$	f_d/f_u	σ_{stat} (%)	σ_{sys} (%)
0.00–0.25	1.002	0.9	2.3
0.25–0.50	1.016	0.9	2.4
0.50–0.75	1.010	0.8	2.3
0.75–1.00	1.019	0.9	4.5
1.00–1.30	1.006	0.9	4.9
1.30–1.60	0.984	1.0	4.2
1.60–2.40	0.975	0.6	4.3

5.0.2 $B^0 \to K^0_S \mu^- \mu^+$ Branching Fraction and Angular Analysis

The B-Parked data set showed positive results extracting signal from a double displaced vertex like in the $B^0 \to K^0_S \mu^- \mu^+$ decay. This opens to study other channels such as $\Lambda^0_b \to \Lambda \mu^+ \mu^-$, that were previously not explored in CMS due to the poor performance of the standard *Two muons + track* triggers for B decays with Λ or K^0_S in the final state. Differential branching ratio and F_H measurements in bins of q^2 summarized in figures 4.16, and 4.18 show promising results that are competitive with LHCb. Due to the lack of time, systematic uncertainties were not explored yet. Finally, it is expected that statistical uncertainties will be significantly reduced with the expected luminosity in Run-III of about 110fb^{-1} [54].

Appendix A

Parking BPH Trigger

A.1 Trigger description

The `ParkingBPH` dataset was designed to record 10^{10} unbiased B hadron events in 2018 (the last year of Run-II). The main idea was to use the capabilities of the Data Acquisition (DAQ) system of CMS, which exceeds the computing capacity, and write on tape more data than normally, "park"-it, and reconstruct them later during the LHC downtime (when the computing farms are available).

Having this in mind, the trigger strategy was optimized to maximize the number of B hadrons recorded on tape, given a fixed bandwidth. As the luminosity changed, different `Single Muon` $|\eta|$-restricted L1-seeds (Table A.1) were turned on, keeping the L1 rate constant and increasing the HLT toward the end of each proton-proton fill in the LHC.

Table A.1: L1 seeds for the `ParkingBPH` dataset.

L1 seed name	p_T Threshold	η restrictions
L1_SingleMu7er1p5	7.0 GeV : 0xf	MU-ETA_1p5 [-1.506, 1.506] : [0x176, 0x8a]
L1_SingleMu8er1p5	8.0 GeV : 0x11	MU-ETA_1p5 [-1.506, 1.506] : [0x176, 0x8a]
L1_SingleMu9er1p5	9.0 GeV : 0x13	MU-ETA_1p5 [-1.506, 1.506] : [0x176, 0x8a]
L1_SingleMu10er1p5	10.0 GeV : 0x15	MU-ETA_1p5 [-1.506, 1.506] : [0x176, 0x8a]
L1_SingleMu12er1p5	12.0 GeV : 0x19	MU-ETA_1p5 [-1.506, 1.506] : [0x176, 0x8a]
L1_SingleMu14er1p5	14.0 GeV : 0x1d	MU-ETA_1p5 [-1.506, 1.506] : [0x176, 0x8a]
L1_SingleMu18er1p5	18.0 GeV : 0x25	MU-ETA_1p5 [-1.506, 1.506] : [0x176, 0x8a]
L1_SingleMu22	22.0 GeV : 0x2d	
L1_SingleMu25	25.0 GeV : 0x33	

The HLT paths were named `HLT_MuX_IPY`, where `X` is the pT [GeV] and `Y` is the IP (Impact Parameter) significance thresholds applied to the trigger muon. These paths were always present in the trigger menu and turned on / turned off by portions of the LHC fill by imposing a prescale factor of 0. The active triggers evolved, turning on the more `loose` triggers as the instantaneous luminosity decayed. Every part was divided into $n = 5(6)$ streams, creating n copies/parts of the same trigger path. Each copy was prescaled by a factor of n and sent to a different stream, and at the end, it retained the totality of the triggered events.

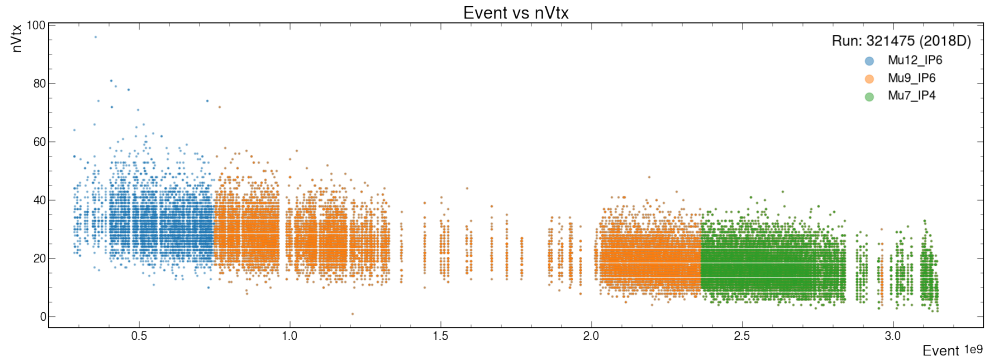

Figure A.1: Number of reconstructed primary vertices vs the event number as an example of the turn on / turn off of the `ParkingBPH` HLT trigger paths during the 2018 data taking.

A.1.1 Tag And Probe method

To calculate the trigger scale factor[1] is necessary to have a measurement of the trigger efficiencies in Data and MC. The Tag and Probe (T&P) method was used following the general guidelines in the official CMS documentation: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/TagAndProbe`. The idea is to use the J/ψ resonance yield as a counting method to calculate the trigger efficiency as the ratio of J/ψ yields, when the Tag and Probe requirements are fulfilled in the numerator, and when the tag muon passes the tag requirements in the denominator:

$$\epsilon = \frac{\text{Signal yield (Tag \& Probe)}}{\text{Signal yield (Tag)}}. \tag{A.1}$$

The whole `BParked` data set (the same as in the analysis) is used to calculate the

[1] The raw efficiency in of the triggers expected to be different in the real data and in MC, the scale factor is the trigger correction applied to MC in order to match the trigger efficiency in real data.

data efficiency. For the MC efficiencies, it was used a large combination of MC samples (labeled as MC_mix in the plots), trying to reduce the systematic uncertainties due to the MC limited size. These MC samples can be found in the CMS DAS (Data Aggregation System) as:

- `JpsiToMuMu_JpsiPt8_TuneCP5_13TeV-pythia8/`
  `RunIIAutumn18MiniAOD-102X_upgrade2018*`
- `BuToKStarJPsi_ToMuMu_Mufilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`
  `RunIIAutumn18MiniAOD-PUPoissonAve20_BParking_102X_upgrade2018*`
- `BuToKStarJPsi_ToMuMu_Mufilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`
  `RunIIAutumn18MiniAOD-PUPoissonAve20_BParking_102X_upgrade2018*`
- `BuToK10Jpsi_ToMuMu_probefilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`
  `RunIIAutumn18MiniAOD-PUPoissonAve20_BParking_102X_upgrade2018*`
- `BsToPhiJpsi_ToKKMuMu_probefilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`
  `RunIIAutumn18MiniAOD-PUPoissonAve20_BParking_102X_upgrade2018*`
- `BsToJpsiPhi_BMuonFilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`
  `RunIIAutumn18MiniAOD-102X_upgrade2018*`
- `BdToKstarJpsi_ToKPiMuMu_probefilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`
  `RunIIAutumn18MiniAOD-PUPoissonAve20_BParking_102X_upgrade2018*`
- `BdToK0sJPsi_ToMuMu_Mufilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`
  `RunIIAutumn18MiniAOD-PUPoissonAve20_BParking_102X_upgrade2018*`
- `BdToK0sJPsi_JPsiToMuMu_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`
  `RunIIAutumn18MiniAOD-PUPoissonAve20_BParking_102X_upgrade2018*`
- `BdToJpsiKstar_BMuonFilter_SoftQCDnonD_TuneCP5_13TeV-pythia8-evtgen/`
  `RunIIAutumn18MiniAOD-N1_102X_upgrade2018*`

A.1. Trigger description

The dimuon pair is reconstructed making a kinematic vertex fit with a pair of opposite charge muons that have valid inner tracks. Because muons are expected to have different efficiencies in different kinematic regions, the events are separated by bins of p_{T} , η, and IP significance where the probe muon is located.

To model the J/ψ resonance from the dimuon pair, a double Gaussian and an exponential are used for signal, and background modeling, respectively.

The J/ψ resonance used in the T&P technique fulfilled these requirements:

- Selected trigger has a positive prescale;
- Kinematic dimuon vertex probability greater than 5%;
- Dimuon invariant mass range is 0.25 GeV around the J/ψ resonance peak;
- Both muons have $|\eta| < 2.5$;
- Both muons has passed the `soft muon` ID.

The tag requirements are:

- muon has fired any `HLT_MuX_IPY` HLT path;
- angular distance with the other muon $\Delta R > 0.10$.

The probe selection is defined as:

- muon has fired a specific `HLT_MuX_IPY` HLT path.

Only `HLT_Mu7_IP4_part*_v*`, `HLT_Mu8_IP3_part*_v*`, `HLT_Mu8_IP5_part*_v*`, `HLT_Mu9_IP5_part*_v*`, `HLT_Mu9_IP6_part*_v*`, `HLT_Mu12_IP6_part*_v*` have exclusive events recorded (more details in a moment) and are present in the MC trigger menu. So these are the only trigger paths that are used for the correction. Figures A.2 to A.8 show the efficiencies corresponding to these HLT triggers.

The efficiencies are estimated with the root class `TEfficiency`, with a 68% Clopper-Pearson confidence interval.

After obtaining the efficiencies, a scale factor is calculated as the ratio of the Data efficiency over the MC efficiency (SF $= \epsilon_{RD}/\epsilon_{MC}$). The uncertainty is propagated from the Clopper-Pearson confidence intervals from both data and MC efficiencies.

Very similar efficiencies are expected once the trigger has reached the plateau (a scale factor near to 1 at large p_{T}), and have larger discrepancies between data and MC below the thresholds and, hence, scale factors with values far from 1. The thresholds are different for every HLT path.

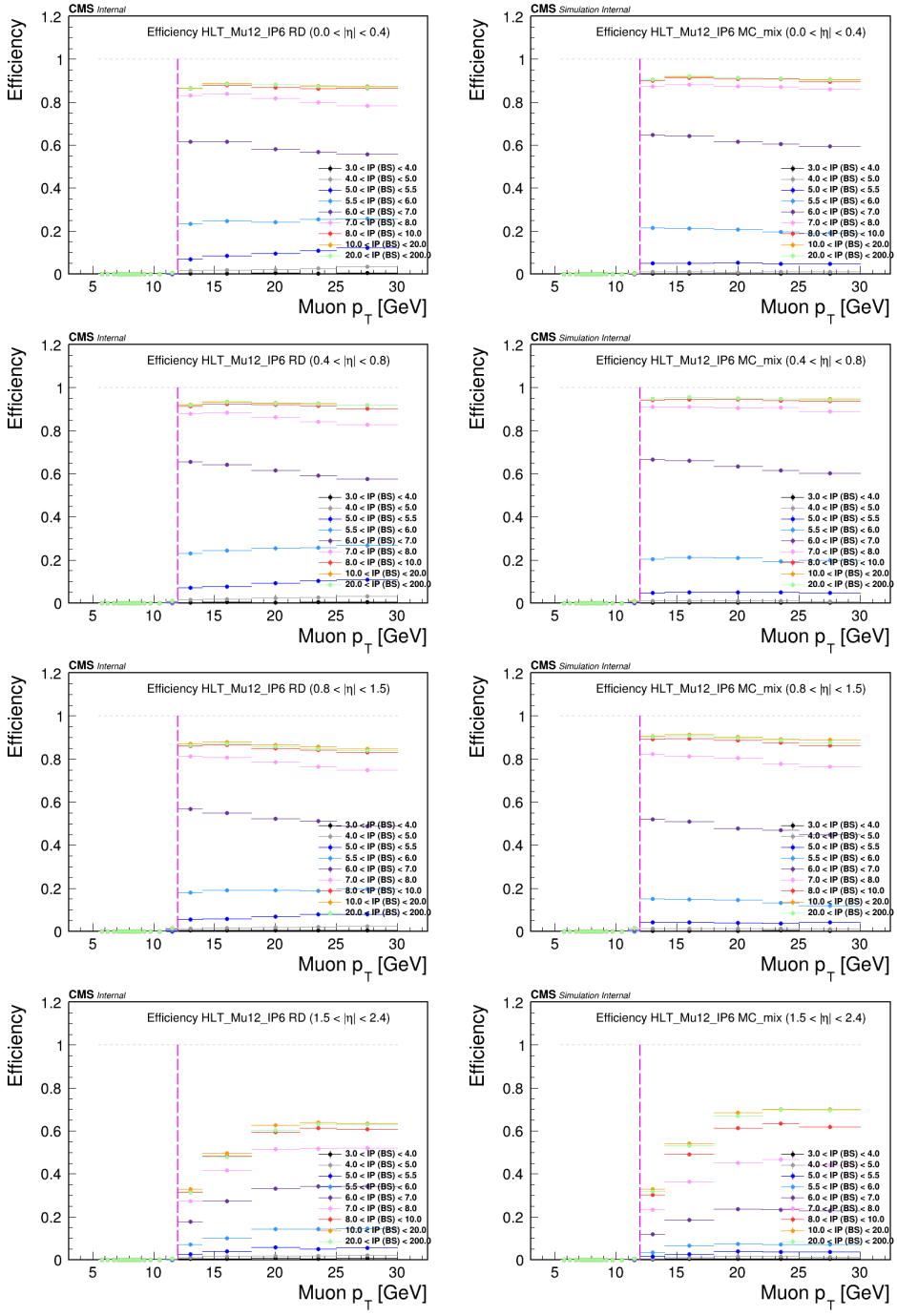

Figure A.2: HLT_Mu12_IP6 trigger efficiencies in 4 regions of $|\eta|$. The left side is data and the right column is MC.

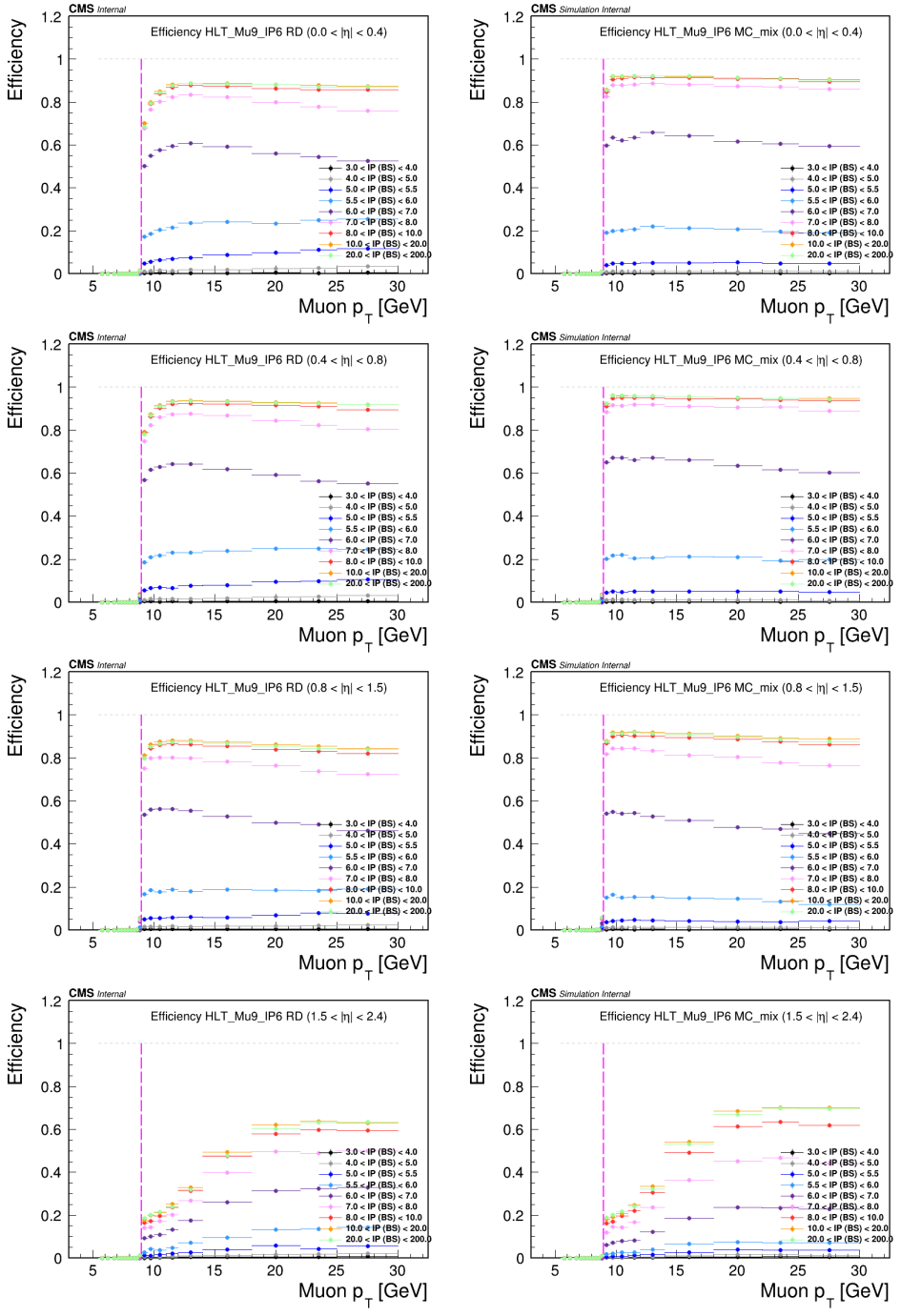

Figure A.3: HLT_Mu9_IP6 trigger efficiencies in 4 regions of $|\eta|$. The left side is data and the right column is MC

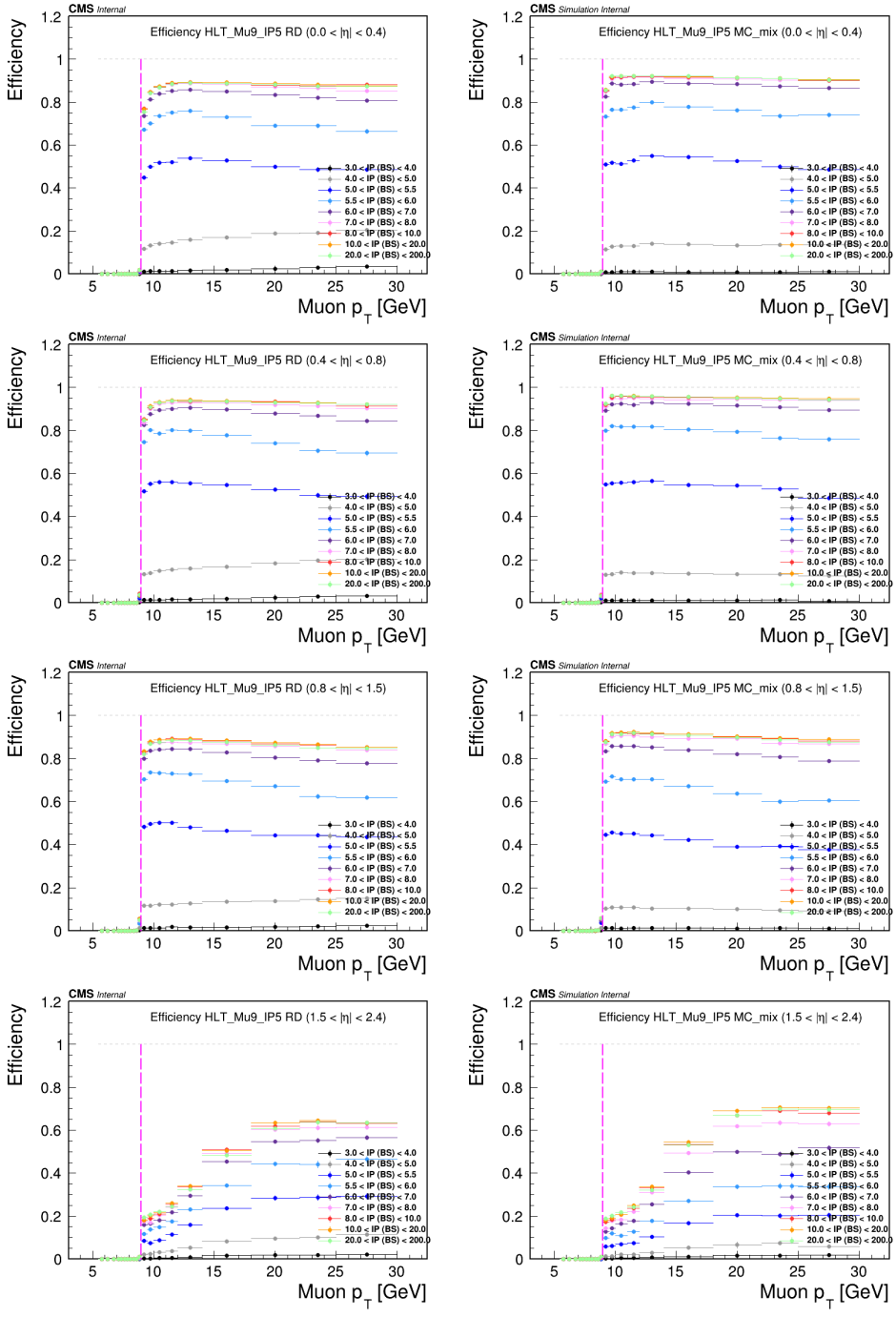

Figure A.4: HLT_Mu9_IP5 trigger efficiencies in 4 regions of $|\eta|$. The left side is data and the right column is MC

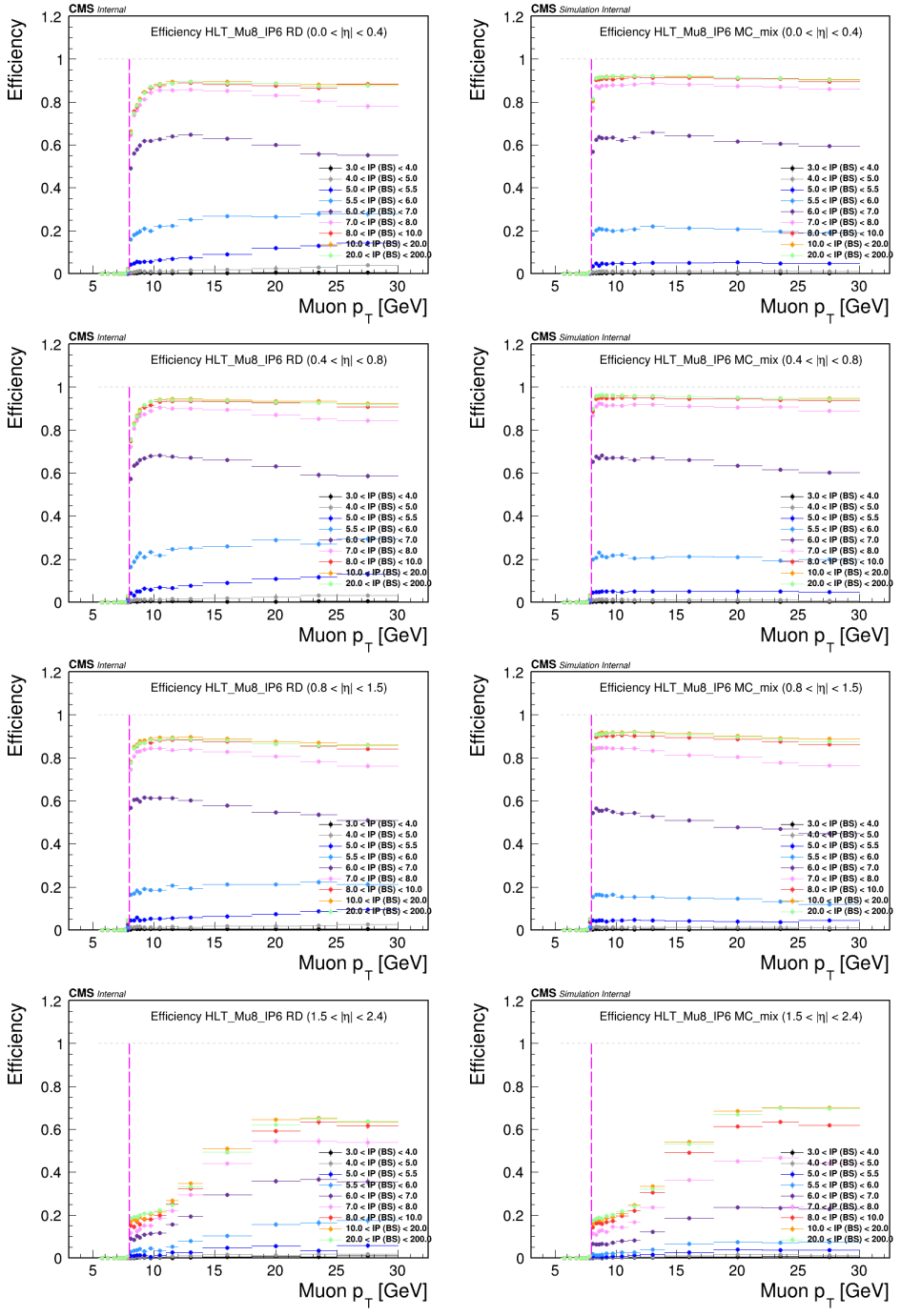

Figure A.5: HLT_Mu8_IP6 trigger efficiencies in 4 regions of $|\eta|$. The left side is data and the right column is MC

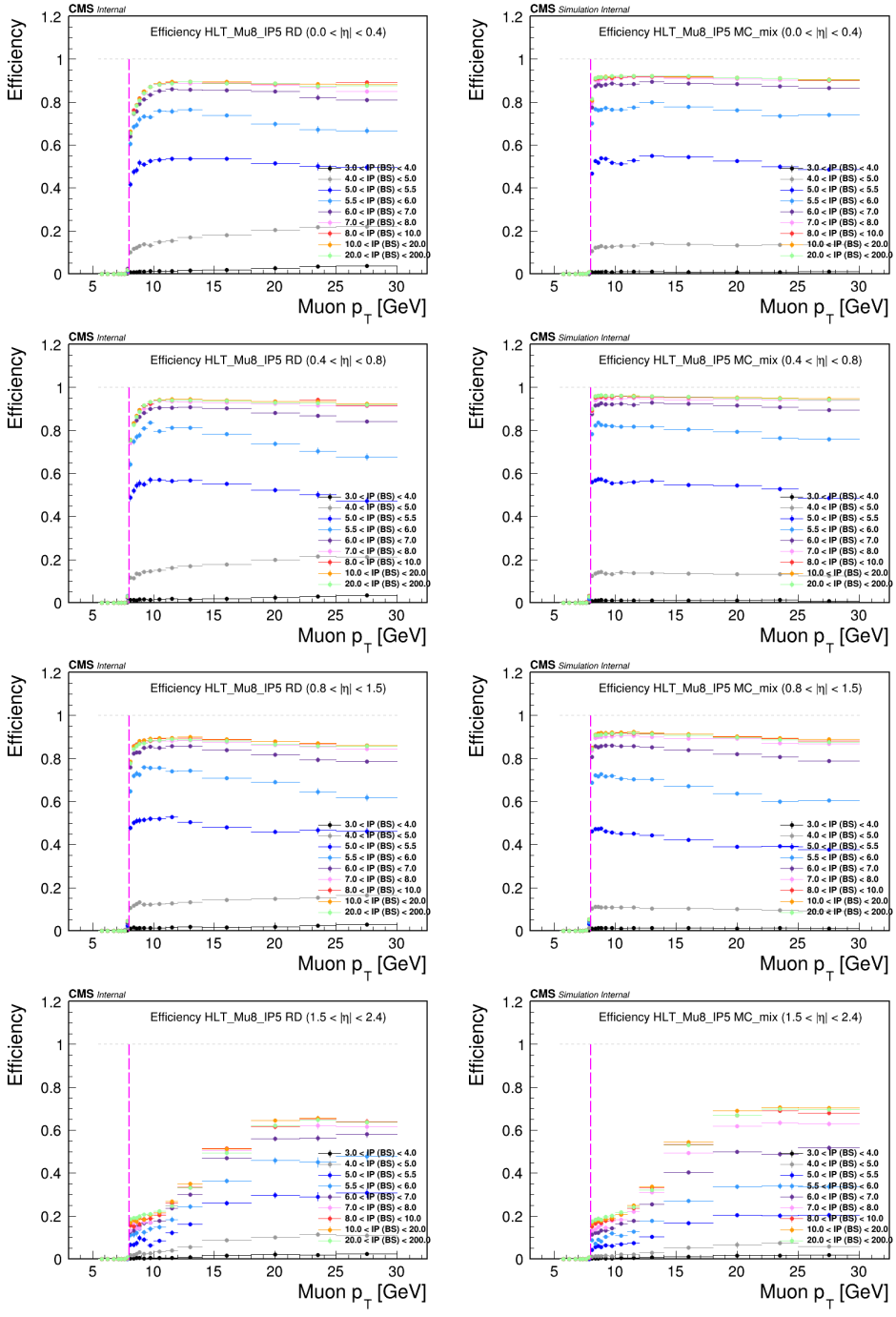

Figure A.6: HLT_Mu8_IP5 trigger efficiencies in 4 regions of $|\eta|$. The left side is data and the right column is MC.

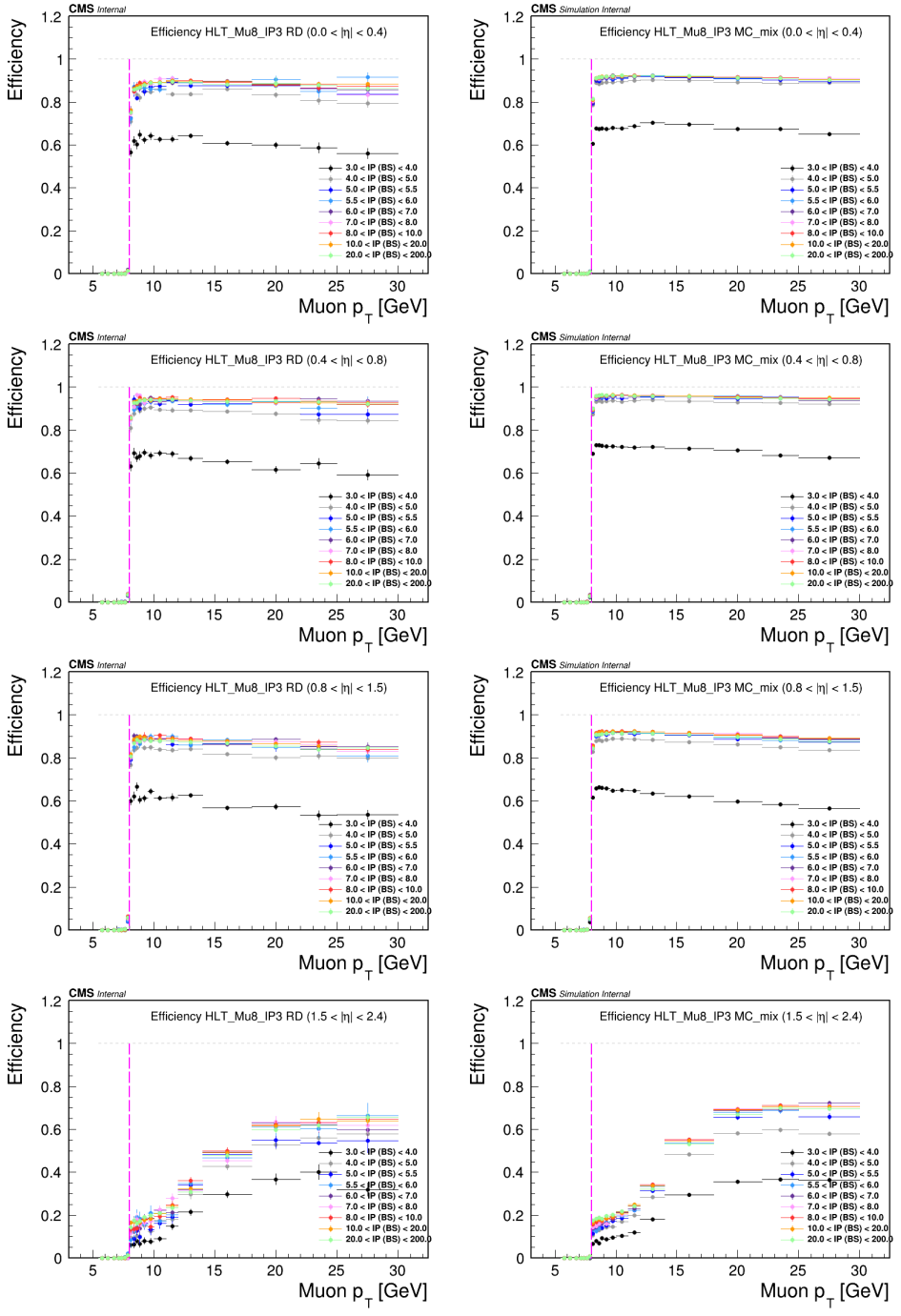

Figure A.7: HLT_Mu8_IP3 trigger efficiencies in 4 regions of $|\eta|$. The left side is data and the right column is MC

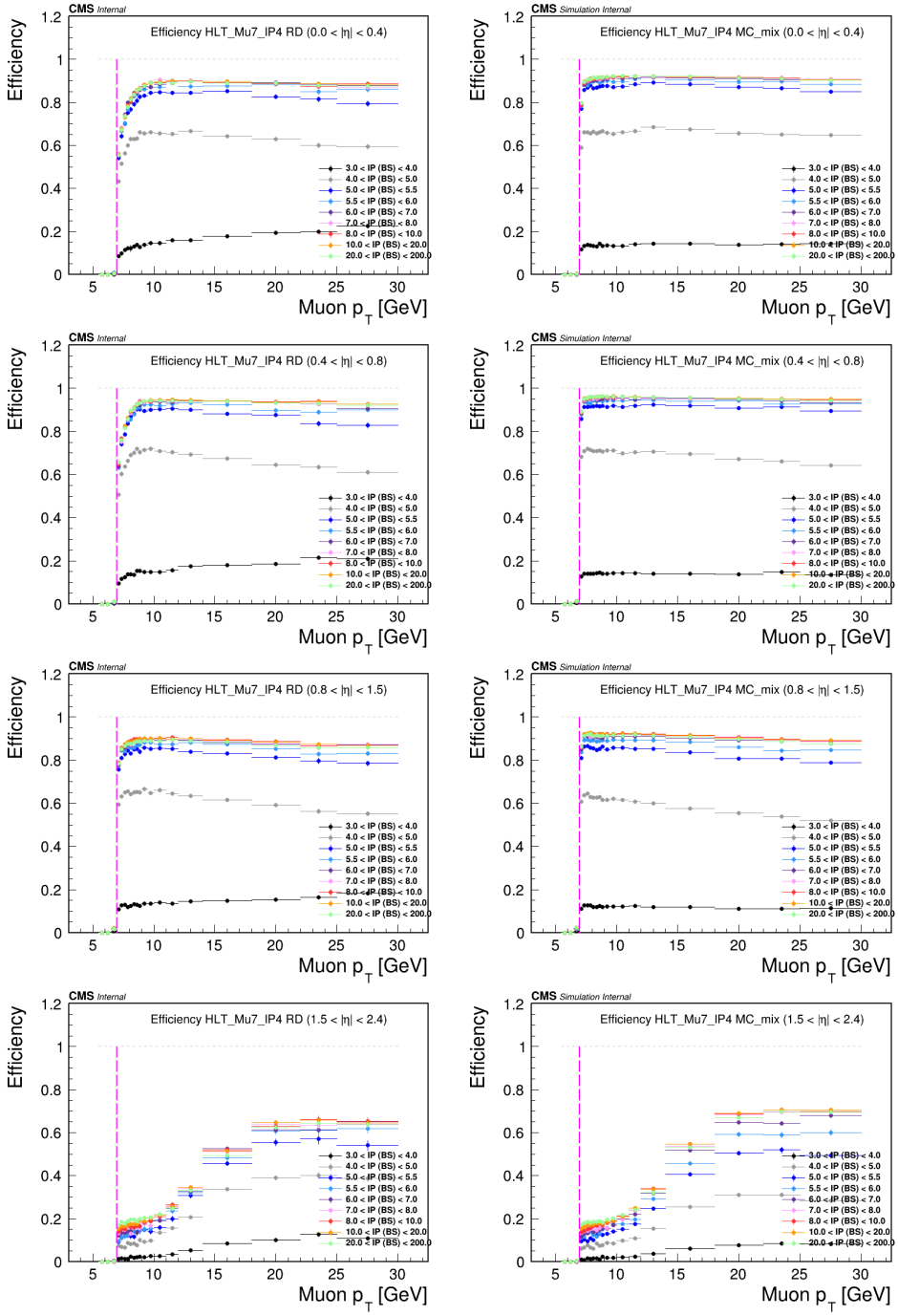

Figure A.8: HLT_Mu7_IP4 trigger efficiencies in 4 regions of $|\eta|$. The left side is data and the right column is MC

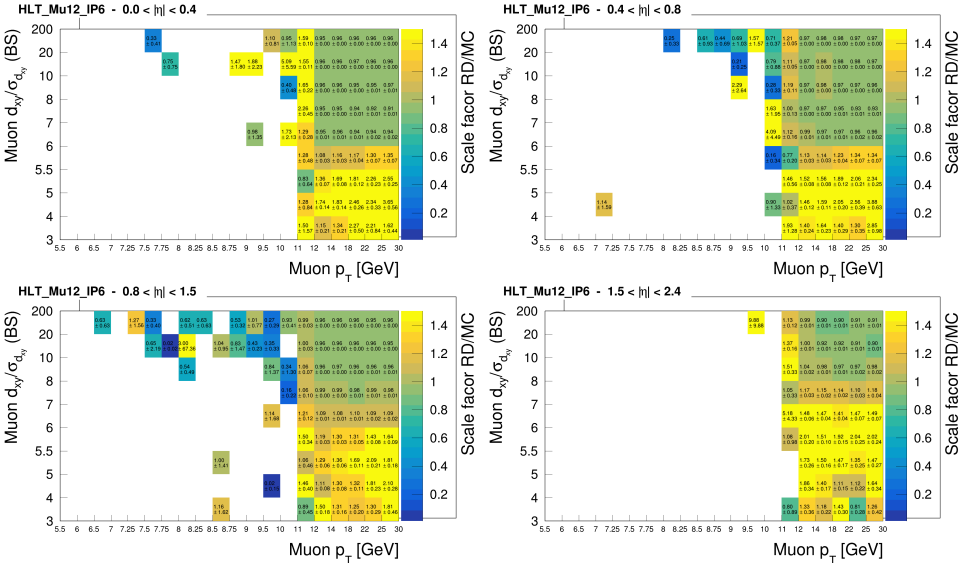

Figure A.9: HLT_Mu12_IP6 trigger scale factors in bins of muon $\frac{d_{xy}}{\sigma_{d_{xy}}}$ and p_{T}.

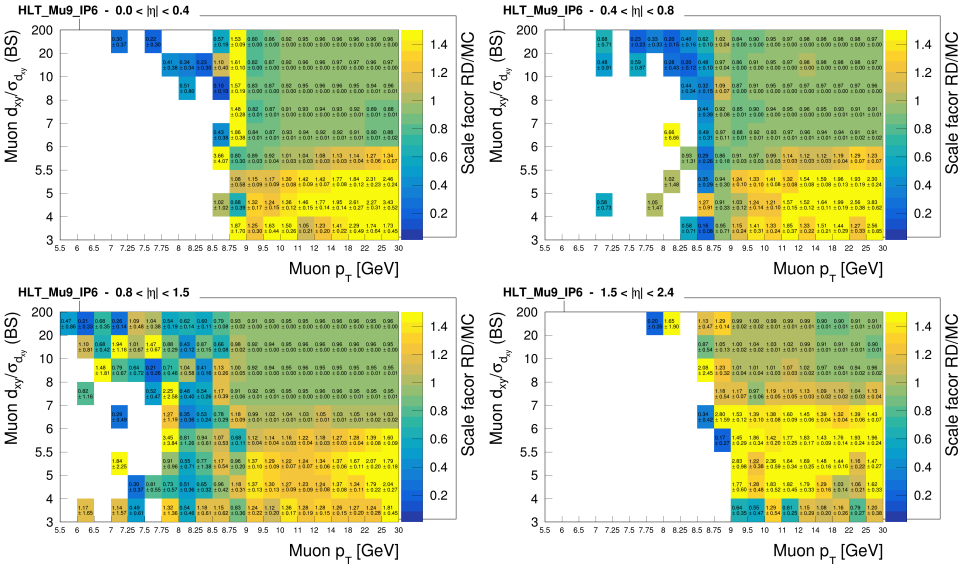

Figure A.10: HLT_Mu9_IP6 trigger scale factors in bins of muon $\frac{d_{xy}}{\sigma_{d_{xy}}}$ and p_{T}.

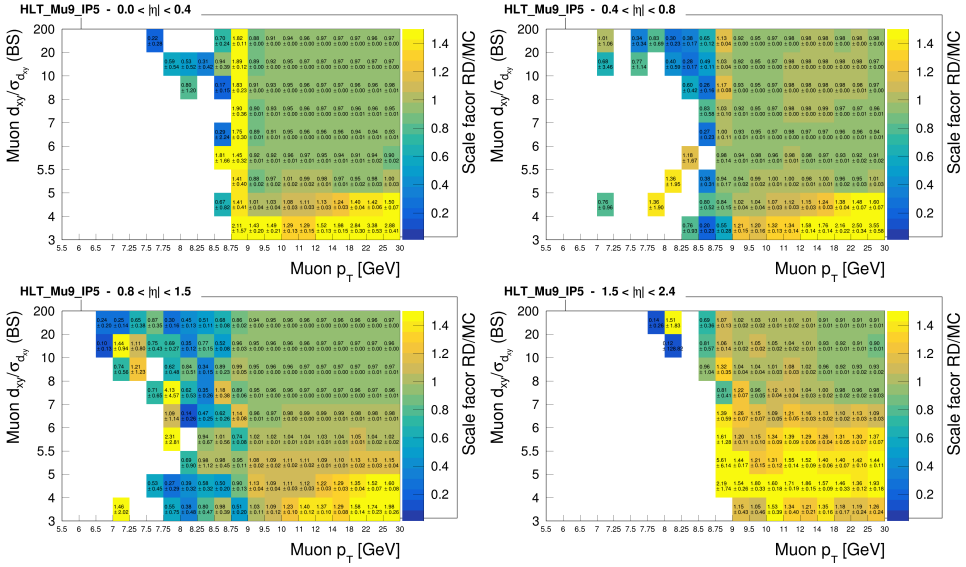

Figure A.11: HLT_Mu9_IP5 trigger scale factors in bins of muon $\frac{d_{xy}}{\sigma_{d_{xy}}}$ and p_{T}.

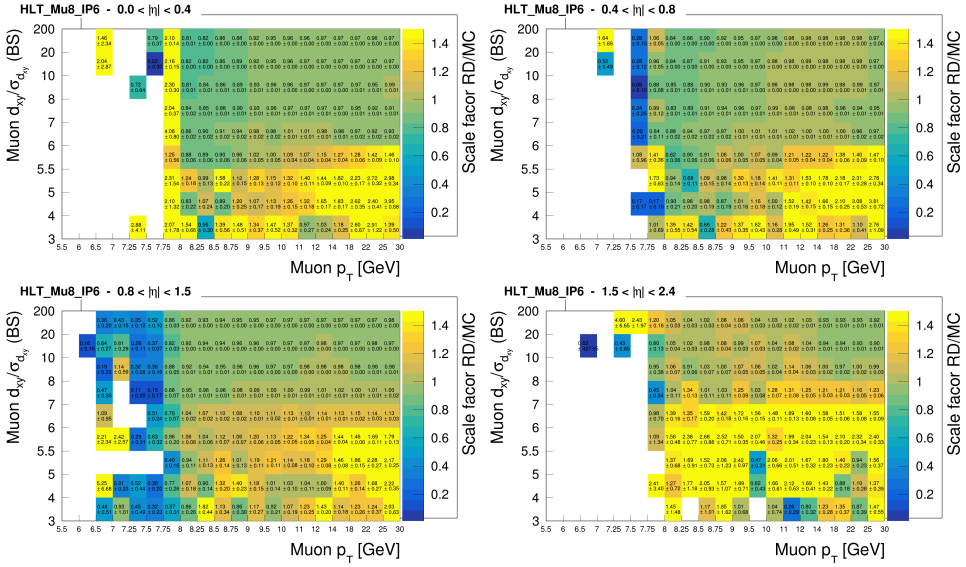

Figure A.12: HLT_Mu8_IP6 trigger scale factors in bins of muon $\frac{d_{xy}}{\sigma_{d_{xy}}}$ and p_{T}.

A.1.2 Accept-Reject method

Another observed discrepancy between data and MC is the trigger composition profile, which represents the distribution of fired HLT triggers present in the dataset. The discrepancy is expected since more inclusive triggers are always active in MC, but not in the real data, due to the dynamic nature of the HLT triggers discussed in Sec. A.1. For instance, in MC the trigger `HLT_Mu7_IP4_*` will be fired any time `HLT_Mu12_IP6_*` is fired, so the more inclusive triggers are also the more dominant (more frequently fired). In data, it has been discussed that the trigger menu changes with an artificial zero prescale.

A correction to this effect is done through an accept-reject (A/R) method applied to the MC. The idea was proposed by the $\tau \to 3\mu$ CMS analysis group[2], and it consists in accepting events in MC accordingly to the collected luminosity per trigger. So the first step is to calculate the *exclusive* [3] luminosity recorded by each trigger. Results[4] are reported in the table A.2.

Table A.2: HLT paths, run eras, and the if corresponding integrated luminosity for the B Parked dataset. The last column reports the fraction of events recorded by each trigger.

HLT path	Run A [fb^{-1}]	Run B [fb^{-1}]	Run C [fb^{-1}]	Run D [fb^{-1}]	Total [fb^{-1}]	fraction
`HLT_Mu7_IP4`	0.000	0.365	0.150	6.418	6.933	0.167
`HLT_Mu8_IP3`	0.442	0.353	0.000	0.000	0.795	0.019
`HLT_Mu8_IP5`	0.000	0.000	0.000	3.328	3.328	0.080
`HLT_Mu8_IP6`	0.000	0.000	0.000	0.000	0.000	0.000
`HLT_Mu9_IP5`	0.000	1.002	2.400	7.079	10.482	0.253
`HLT_Mu9_IP6`	4.200	2.603	1.140	3.949	11.891	0.287
`HLT_Mu12_IP6`	0.000	0.610	1.677	5.736	8.023	0.194
Total	4.642	4.932	5.368	26.511	41.453	1.000

The next step is to generate a random number between 0 and 1 for every MC event, $X_{[0,1]}$. Then accept the event if $X_{[0.1]}$ is within a specific range (related to the fraction of events recorded exclusively by each trigger in data) and fired a specific HLT path (see following rules), otherwise reject it:

- `HLT_Mu7_IP4` fired and $0.0 < X_{[0,1]} < 0.167$

[2] https://indico.cern.ch/event/895802/contributions/3777180/attachments/2007850/3353822/Tau3mu_HF_2018.pdf, slide 23

[3] In this context we are calling *exclusively* to the integrated luminosity collected by the most inclusive trigger, i.e: if in a given lumi section, several triggers were fired, we assign the recorded luminosity to the lowest pT HLT.

[4] There is a small difference with the luminosity reported in table 4.1, it is because the `HLT_Mu8p5_IP3p5` trigger path is not being taken into account because is not present in the MC trigger menu.

- `HLT_Mu8_IP3` fired and $0.167 \leq X_{[0,1]} < 0.186$, $(0.167 + 0.019)$
- `HLT_Mu8_IP5` fired and $0.186 \leq X_{[0,1]} < 0.266$, $(0.186 + 0.080)$
- `HLT_Mu9_IP5` fired and $0.266 \leq X_{[0,1]} < 0.519$, $(0.266 + 0.253)$
- `HLT_Mu9_IP6` fired and $0.519 \leq X_{[0,1]} < 0.806$, $(0.519 + 0.287)$
- `HLT_Mu12_IP6` fired and $0.806 \leq X_{[0,1]} < 1.00$, $(0.806 + 0.194)$

After this accept-reject procedure is done, the specific trigger scale factor is applied (if the event passed the `HLT_MuX_IPY` A/R, the `HLT_MuX_IPY` SF is used) using this formula:

	μ_1 triggered	μ_1 not triggered
μ_2 triggered	$\epsilon_1 \times \epsilon_2$	$(1-\epsilon_1) \times \epsilon_2$
μ_2 not triggered	$\epsilon_1 \times (1-\epsilon_2)$	$(1-\epsilon_1) \times (1-\epsilon_2)$

Table A.3: Rules to apply the SF to each event. The case where no trigger is fired is never used because of selection requirements.

After the MC sample has been corrected by the A/R and each passing MC event has been assigned a scale factor, it is compared with data (background subtraction is obtained using the s-weights technique). Figures A.13 and A.14 show the comparison of the p_{T} spectrum of the B hadron and the pair of muons with the other variables that are used in the HLT pat h(η and IP significance).

In order to use the efficiency as shown in the table A.3, it is only considered the case when at least one muon is a trigger muon, considering first the case of two trigger muons

$$\begin{aligned} \mathrm{SF}' &= \frac{[\epsilon(\mu_1(p_{\mathrm{T}}), \mu_1(|\eta|), \mu_1(IP)) \cdot \epsilon(\mu_2(p_{\mathrm{T}}), \mu_2(|\eta|), \mu_2(IP))]_{\mathrm{data}}}{[\epsilon(\mu_1(p_{\mathrm{T}}), \mu_1(|\eta|), \mu_1(IP)) \cdot \epsilon(\mu_2(p_{\mathrm{T}}), \mu_2(|\eta|), \mu_2(IP))]_{\mathrm{MC}}} \\ &= \mathrm{SF}(\mu_1(p_{\mathrm{T}}), \mu_1(|\eta|), \mu_1(IP)) \cdot \mathrm{SF}(\mu_2(p_{\mathrm{T}}), \mu_2(|\eta|), \mu_2(IP)), \end{aligned} \tag{A.2}$$

in the case that μ_1 fired a trigger, but μ_2 did not, then the formula would be

$$\begin{aligned} \mathrm{SF}' &= \frac{[\epsilon(\mu_1(p_{\mathrm{T}}), \mu_1(|\eta|), \mu_1(IP)) \cdot (1 - \epsilon(\mu_2(p_{\mathrm{T}}), \mu_2(|\eta|), \mu_2(IP)))]_{\mathrm{data}}}{[\epsilon(\mu_1(p_{\mathrm{T}}), \mu_1(|\eta|), \mu_1(IP)) \cdot (1 - \epsilon(\mu_2(p_{\mathrm{T}}), \mu_2(|\eta|), \mu_2(IP)))]_{\mathrm{MC}}} \\ &= \mathrm{SF}(\mu_1(p_{\mathrm{T}}), \mu_1(|\eta|), \mu_1(IP)) \cdot (\mathrm{SF} - 1)(\mu_2(p_{\mathrm{T}}), \mu_2(|\eta|), \mu_2(IP)), \end{aligned} \tag{A.3}$$

where $\mathrm{SF} - 1$ it is just the ratio of $1 - \epsilon$, ϵ being the efficency. So two quantities are calculated, the ratio of efficiencies and the ratio of $1 - \epsilon$ to apply the MC correction.

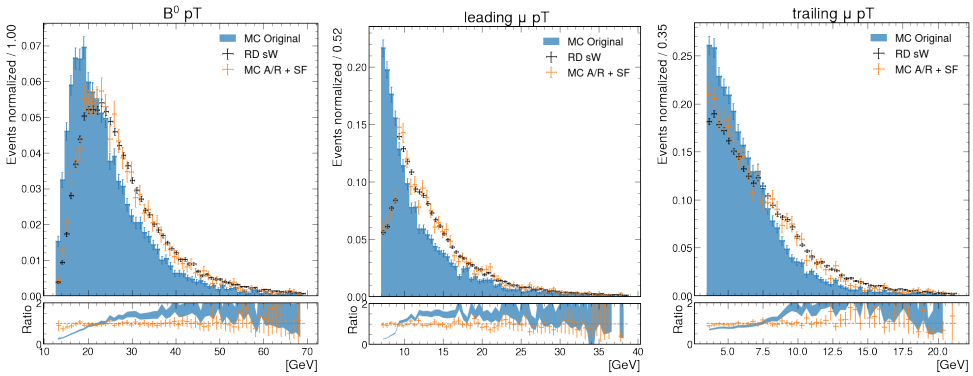

Figure A.13: Comparison between the original MC, the data (background subtracted), and the MC after all the corrections. Shown are the B^0 and muons p_T .

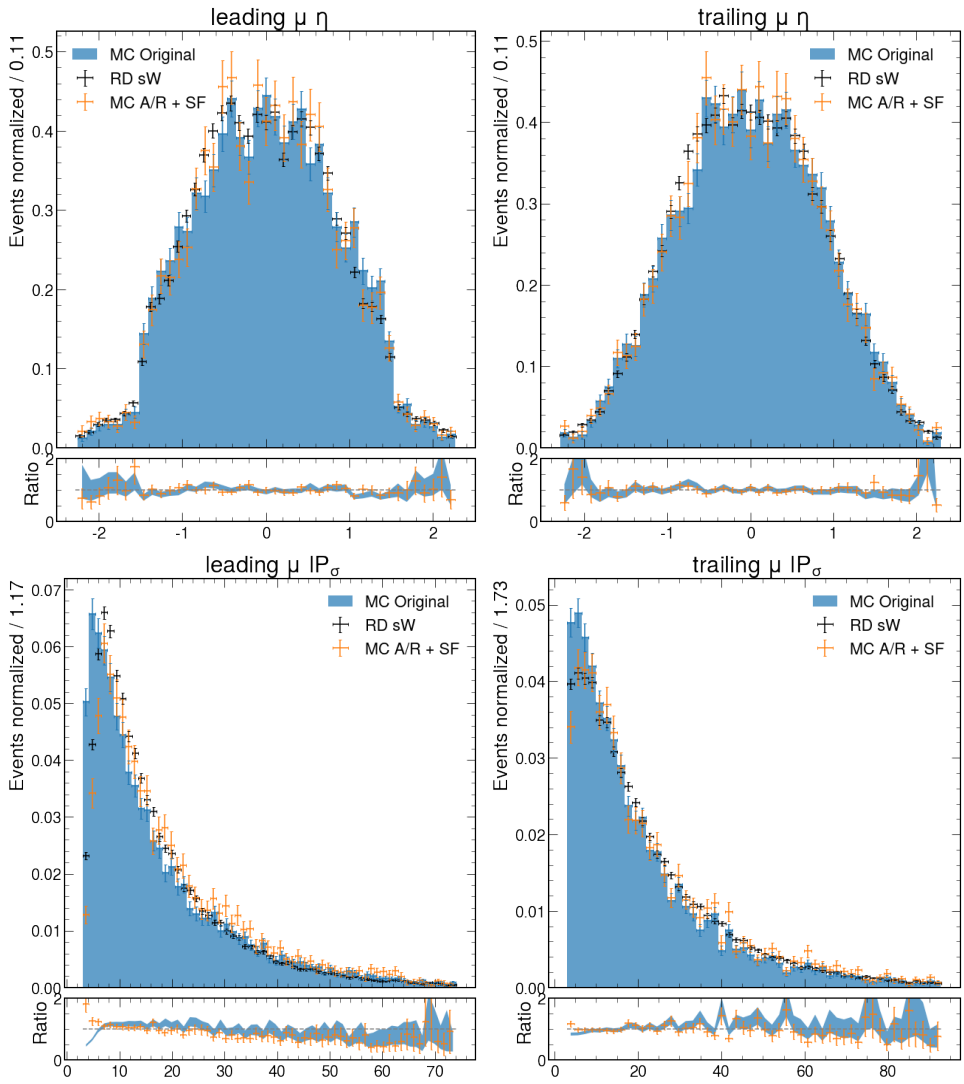

Figure A.14: Comparison between the original MC, the data (background subtracted), and the MC after all the corrections. Shown are the muons η and IP significance (IP_σ).

Appendix B

Kinematic Vertex Fitting

The kinematic vertex fit of a reconstructed particle decay is a mathematical procedure done to improve the experimental resolution of the measurements. It achieves this goal by applying the known physics governing the decay. This algorithm has been programmed and introduced as a package in the CMSSW reconstruction program (see the CMS internal note 2004/020 from K. Prokofiev and Th. Speer).

The example that is always used to explain the Kinematic Vertex Fit (KVF), is the four tracks decay from $D^0 \to K^-\pi^+\pi^+\pi^-$. All tracks must come from a common point that is used to improve the 4-momentum and position resolution of all the daughter particles, and doing so improves the mass resolution of the D^0. The physical information is combined via *constraints*. Each constraint is written in the form of an equation expressing the physical condition that the process must satisfy. Vertexing is one kind of constraint, but it is also possible to require that the invariant mass of a certain known resonance is fixed, known as *mass constraint*.

Constraints are generally implemented through Least Mean Squared minimization (LMS) (this is the most widely used technique for a constrained fit), but sequential methods such as the Kalman filter optimization algorithm and global methods (global minimization with Lagrange multipliers, global minimization with penalty functions and others) are also used in high energy physics. In CMS, the vertex reconstruction is done using LMS with Lagrange multipliers and Kalman filter techniques. Constrained refit on tracks and vertices (such as the Kinematic Vertex Fitting) is performed via LMS minimization with Lagrange multipliers.

Obtaining the solution to the minimization with constraints is only one part of the full process. It is also important to know about the errors and correlations of the

new parameters. It is expected that the constraints help to reduce the errors from the original measurement. This is solved using a correlation matrix to obtain the new errors, from the refitting process.

Bibliography

[1] Morad Aaboud et al. "Study of the rare decays of B_s^0 and B^0 mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector". In: *JHEP* 04 (2019), p. 098. DOI: 10.1007/JHEP04(2019)098. arXiv: 1812.03017 [hep-ex].

[2] Georges Aad et al. "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". In: *Phys. Rev. Lett.* 115 (2015), p. 262001. DOI: 10.1103/PhysRevLett.115.262001. arXiv: 1507.08925 [hep-ex].

[3] R. Aaij et al. "Measurement of b-hadron fractions in 13 TeV pp collisions". In: *Phys. Rev. D* 100 (2019), p. 031102. DOI: 10.1103/PhysRevD.100.031102. arXiv: 1902.06794 [hep-ex].

[4] Roel Aaij et al. "Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$". In: *Journal of High Energy Physics* 2015.9 (), p. 179. DOI: 10.1007/jhep09(2015)179. arXiv: 1506.08777 [hep-ex].

[5] Roel Aaij et al. "Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays". In: *Journal of High Energy Physics* 2015.6 (), p. 115. DOI: 10.1007/JHEP06(2015)115. arXiv: 1503.07138v3 [hep-ex].

[6] Roel Aaij et al. "Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ decays". In: *Journal of High Energy Physics* 2014.6 (), p. 133. DOI: 10.1007/jhep06(2014)133. arXiv: 1403.8044 [hep-ex].

[7] Roel Aaij et al. "Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays". In: *Phys. Rev. Lett.* 118 (2017), p. 191801. DOI: 10.1103/PhysRevLett.118.191801. arXiv: 1703.05747 [hep-ex].

[8] Roel Aaij et al. "Measurement of the $B_s^0 \to \mu^+\mu^-$ decay properties and search for the $B^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-\gamma$ decays". In: *Phys. Rev. D* 105 (2022), p. 012010. DOI: 10.1103/PhysRevD.105.012010. arXiv: 2108.09283 [hep-ex].

[9] Roel Aaij et al. "Measurement of f_s/f_u variation with proton-proton collision energy and B-meson kinematics". In: *Phys. Rev. Lett.* 124 (2020), p. 122002. DOI: 10.1103/PhysRevLett.124.122002. arXiv: 1910.09934 [hep-ex].

[10] Roel Aaij et al. "Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data". In: *Nature* 522 (2015), pp. 68–72. DOI: 10.1038/nature14474. arXiv: 1411.4413 [hep-ex].

[11] Roel Aaij et al. "Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B_s^0 decay branching fractions". In: *Phys. Rev. D* 104 (2021), p. 032005. DOI: 10.1103/PhysRevD.104.032005. arXiv: 2103.06810 [hep-ex].

[12] Roel Aaij et al. "Strong Constraints on the Rare Decays $B_s^0 \to \mu^+\mu-$ and $B^0 \to \mu^+\mu^-$". In: *Phys. Rev. Lett.* 108 (2012), p. 231801. DOI: 10.1103/PhysRevLett.108.231801. arXiv: 1203.4493 [hep-ex].

[13] T. Aaltonen et al. "Measurement of ratios of fragmentation fractions for bottom hadrons in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV". In: *Phys. Rev. D* 77 (2008), p. 072003. DOI: 10.1103/PhysRevD.77.072003. arXiv: 0801.4375 [hep-ex].

[14] J. Abdallah et al. "A measurement of the branching fractions of the b-quark into charged and neutral c-hadrons". In: *Phys. Lett. B* 576 (2003), p. 29. DOI: 10.1016/j.physletb.2003.09.070. arXiv: hep-ex/0311005.

[15] M. Acciarri et al. "Measurements of the $b\bar{b}$ production cross-section and forward backward asymmetry at center-of-mass energies above the Z pole at LEP". In: *Phys. Lett. B* 485 (2000), p. 71. DOI: 10.1016/S0370-2693(00)00676-6. arXiv: hep-ex/0005023.

[16] P. D. Acton et al. "Evidence for the existence of the strange b-flavoured meson B_s^0 in Z^0 decays". In: *Phys. Lett. B* 295 (1992), p. 357. DOI: 10.1016/0370-2693(92)91578-W.

[17] S. Agostinelli et al. "GEANTfour -a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8.

[18] Ben Allanach and Joe Davighi. "The Rumble in the Meson: a leptoquark versus a Z' to fit $b \to s\mu^+\mu^-$ anomalies including 2022 LHCb $R_{K^{(*)}}$ measurements. The Rumble in the Meson: a leptoquark versus a Z' to fit $b \to s\mu^+\mu^-$ anomalies". In: *JHEP* 2304 (2023). 31 pages, 6 figures. v3 is updated with a new fit including December 2022 LHCb reanalysis of $R_{K^{(*)}}$ measurements. Plots and fits significantly changed but the main conclusion is left unchanged. v4 has minor changes. v4 has a typo fixed in the caption of Fig 6, p. 033. DOI: 10.1007/JHEP04(2023)033. arXiv: 2211.11766. URL: http://cds.cern.ch/record/2841928.

[19] Wolfgang Altmannshofer et al. "Symmetries and asymmetries of $B \to K^*\mu^+\mu^-$ decays in the Standard Model and beyond". In: *Journal of High Energy Physics* 2009.01 (Jan. 2009), pp. 019–019. DOI: https://doi.org/10.1088/1126-6708/2009/01/019. URL: https://doi.org/10.10882F1126-67082F20092F012F019.

[20] Yasmine Sara Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021". In: (2022). DOI: 10.48550/ARXIV.2206.07501. arXiv: 2206.07501 [hep-ex].

[21] Yasmine Sara Amhis et al. "Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018". In: *Eur. Phys. J. C* 81.3 (2021), p. 226. DOI: 10.1140/epjc/s10052-020-8156-7. arXiv: 1909.12524 [hep-ex].

[22] J. J. Aubert et al. "Experimental Observation of a Heavy Particle J". In: *Phys. Rev. Lett.* 33 (23 Dec. 1974), pp. 1404–1406. DOI: 10.1103/PhysRevLett.33.1404. URL: https://link.aps.org/doi/10.1103/PhysRevLett.33.1404.

[23] J. -E. Augustin et al. "Discovery of a Narrow Resonance in e^+e^- Annihilation". In: *Phys. Rev. Lett.* 33 (23 Dec. 1974), pp. 1406–1408. DOI: 10.1103/PhysRevLett.33.1406. URL: https://link.aps.org/doi/10.1103/PhysRevLett.33.1406.

[24] Robert Bainbridge. "Recording and reconstructing 10 billion unbiased b hadron decays in CMS". In: *EPJ Web of Conferences* 245 (2020). Ed. by C. Doglioni et al., p. 01025. DOI: 10.1051/epjconf/202024501025. URL: https://doi.org/10.1051/epjconf/202024501025.

[25] David Barney. "CMS Detector Slice". CMS Collection. 2016. URL: https://cds.cern.ch/record/2120661.

[26] G. L. Bayatian et al. "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". In: (2006).

[27] Martin Beneke, Christoph Bobeth, and Robert Szafron. "Power-enhanced leading-logarithmic QED corrections to $B_q \to \mu^+\mu^-$". In: *JHEP* 2019 (2019), p. 232. DOI: 10.1007/jhep10(2019)23. arXiv: 1908.07011. URL: https://arxiv.org/abs/1908.07011.

[28] "Brilcalc documentation". CMS internal documentation available at http://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html. 2015.

[29] Andrzej J. Buras et al. "Patterns of flavour violation in the presence of a fourth generation of quarks and leptons". In: *Journal of High Energy Physics* 2010.9 (2010), p. 106. DOI: 10.1007/jhep09(2010)106.

[30] D. Buskulic et al. "Measurement of the B_s^0 lifetime and production rate with $D_s^- l^+$ combinations in Z decays". In: *Phys. Lett. B* 361 (1995), p. 221. DOI: 10.1016/0370-2693(95)01173-N.

[31] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[32] Serguei Chatrchyan et al. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *JINST* 9.10 (2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009. arXiv: 1405.6569 [physics.ins-det].

[33] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. San Francisco, California, USA: ACM, 2016, pp. 785–794. ISBN: 978-1-4503-4232-2. DOI: 10.1145/2939672.2939785. URL: http://doi.acm.org/10.1145/2939672.2939785.

[34] S. Choudhury et al. In: *JHEP* 03 (2021), p. 105. DOI: 10.1007/JHEP03(2021)105. arXiv: 1908.01848 [hep-ex].

[35] CMS and LHCb Collaborations. "Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data". In: *Nature* 522 (), p. 68. DOI: 10.1038/nature14474. arXiv: 1411.4413 [hep-ex].

[36] The Tracker Group of the CMS Collaboration. *The CMS Phase-1 Pixel Detector Upgrade*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2745805.

[37] *CMS luminosity measurement for the 2018 data-taking period at* $\sqrt{s} = 13$ *TeV*. CMS Physics Analysis Summary CMS-PAS-LUM-18-002. 2019. URL: https://cds.cern.ch/record/2676164.

[38] The CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *Journal of Instrumentation* 9.10 (Oct. 2014), P10009–P10009. DOI: 10.1088/1748-0221/9/10/p10009. URL: https://doi.org/10.10882F1748-02212F92F102Fp10009.

[39] Wikimedia Commons. *Quark weak interactions*. File: Quark_weak_interactions.svg. 2008. URL: https://commons.wikimedia.org/wiki/File:Quark_weak_interactions.svg.

[40] Wikimedia Commons. *Standard Model of Elementary Particles*. File: Standard_Model_of_Ele 2019. URL: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[41] Wikimedia Commons. *Weak Decay (flipped)*. File: Weak_Decay_(flipped).svg. 2013. URL: https://commons.wikimedia.org/wiki/File:Weak_Decay_(flipped).svg.

[42] Antonio Cota Rodriguez. "Aplicación de la Técnica de Feldman-Cousins al Decaimiento $B^0 \to K^0_S \mu^+ \mu^-$". PhD thesis. 2021. URL: https://repositorio.cinvestav.mx/handle/cinvestav/3671.

[43] G. D'Ambrosio et al. "Minimal flavour violation: an effective field theory approach". In: *Nuclear Physics B* 645.1-2 (2002), pp. 155–187. DOI: 10.1016/s0550-3213(02)00836-2.

[44] N. Davidson, T. Przedzinski, and Z. Was. "PHOTOS interface in C++: Technical and Physics Documentation". In: *Comput. Phys. Commun.* 199 (2016), p. 86. DOI: 10.1016/j.cpc.2015.09.013. arXiv: 1011.0937 [hep-ph].

[45] Zdenek Dolezal. *Haisch-Yellow.png*. File: Haisch-Yellow.png. 2016. URL: http://www.scholarpedia.org/w/images/4/4f/Haisch-Yellow.png.

[46] J. Ellis et al. "The phenomenology of the next left-handed quarks". In: *Nuclear Physics B* 131.2 (1977), pp. 285–307. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(77)90374-1. URL: https://www.sciencedirect.com/science/article/pii/0550321377903741.

[47] Gary J Feldman and Robert D Cousins. "Unified approach to the classical statistical analysis of small signals". In: *Phys. Rev. D Part. Fields* 57.7 (Apr. 1998), pp. 3873–3889. DOI: 10.1103/PhysRevD.57.3873. URL: https://link.aps.org/doi/10.1103/PhysRevD.57.3873.

[48] Gary J. Feldman and Robert D. Cousins. "Unified approach to the classical statistical analysis of small signals". In: *Physical Review D* 57.7 (Apr. 1998), pp. 3873–3889. DOI: 10.1103/physrevd.57.3873. URL: https://doi.org/10.48550/arXiv.physics/9711021.

[49] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Phys. Rev. D* 2 (7 Oct. 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285. URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1285.

[50] Benjamin Grinstein, Roxanne Springer, and Mark B. Wise. "Effective hamiltonian for weak radiative B-meson decay". In: *Physics Letters B* 202.1 (1988), pp. 138–144. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(88)90868-4. URL: https://www.sciencedirect.com/science/article/pii/0370269388908684.

[51] Particle Data Group, M. Tanabashi, et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[52] N. L. Johnson. "Systems of Frequency Curves Generated by Methods of Translation". In: *Biometrika* 36.1/2 (1949), pp. 149–176. ISSN: 00063444. URL: http://www.jstor.org/stable/2332539.

[53] Martin Jung. "Branching ratio measurements and isospin violation in B-meson decays". In: *Phys. Lett. B* 753 (2016), p. 187. DOI: 10.1016/j.physletb.2015.12.024. arXiv: 1510.03423 [hep-ph].

[54] N Karastathis et al. "LHC Run 3 Configuration Working Group Report". In: (2019), pp. 273–284. URL: http://cds.cern.ch/record/2750302.

[55] Till Moritz Karbach. *Feldman-Cousins Confidence Levels - Toy MC Method*. 2011. DOI: https://doi.org/10.48550/arXiv.1109.0714. arXiv: 1109.0714 [physics.data-an].

[56] Makoto Kobayashi and Toshihide Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. URL: https://doi.org/10.1143/PTP.49.652.

[57] Christopher Kolda. "Gauge-mediated supersymmetry breaking: Introduction, review and update". In: *Nuclear Physics B - Proceedings Supplements* 62 (1998), pp. 266–275. DOI: 10.1016/s0920-5632(97)00667-1. arXiv: hep-ph/9707450. URL: https://arxiv.org/abs/hep-ph/9707450.

[58] Patrick Koppenburg, Zdenek Dolezal, and Maria Smizanska. "Rare decays of b hadrons". In: *Scholarpedia* 11.6 (2016), p. 32643. DOI: 10.4249/scholarpedia.32643. URL: https://doi.org/10.4249%5C%2Fscholarpedia.32643.

[59] D.J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth. A* 462 (2001). Ed. by S. Erhan, P. Schlein, and Y. Rozen, p. 152. DOI: 10.1016/S0168-9002(01)00089-4.

[60] J. G. Layter. *The CMS muon project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/343814.

[61] B. Mele and P. Nason. "The fragmentation function for heavy quarks in QCD". In: *Nucl. Phys. B* 361 (1991). [Erratum: DOI:10.1016/0550-3213(91)90597-Q], p. 626. DOI: 10.1016/0550-3213(91)90597-Q.

[62] Mario Greco, ed. *Les Rencontres de Physique de la Vallée d'Aoste - La Thuile 2012*. 2012. DOI: 10.1393/ncc/i2012-11366-6. URL: https://www.sif.it/riviste/sif/ncc/econtents/2012/035/06.

[63] Mario Fernandez Navarro and Stephen F. King. "B-anomalies in a twin Pati-Salam theory of flavour including the 2022 LHCb $R_{K^{(*)}}$ analysis". In: *Journal of High Energy Physics* 2023.2 (), p. 188. DOI: 10.1007/jhep02(2023)188. arXiv: 2209.00276 [hep-ex].

[64] Matthias Neubert. "Effective field theory and heavy quark physiscs". In: *Physics In D ≥ 4 Tasi 2004*. NordiCHI. Cornell University: World Scientific, 2006, pp. 1–36.

[65] J. Neyman. "20. Outline of a theory of statistical estimation based on the classical theory of probability. Phil. Trans. R. S. of London, Ser. A, No. 767, 236, 333-80 (1937)". In: *A Selection of Early Statistical Papers of J. Neyman*. Berkeley: University of California Press, 1967, pp. 250–290. ISBN: 9780520327016. DOI: doi:10.1525/9780520327016-022. URL: https://doi.org/10.1525/9780520327016-022.

[66] Particle Data Group, R. L. Workman, et al. "Review of particle physics". In: *Prog. Theor. Exp. Phys.* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[67] Joseph Polchinski. *Effective Field Theory and the Fermi Surface*. 1999. arXiv: hep-th/9210046 [hep-th].

[68] Riccardo Rattazzi. *Transplanckian collisions at future accelerators*. 2002. DOI: 10.48550/ARXIV.HEP-PH/0205265. URL: https://arxiv.org/abs/hep-ph/0205265.

[69] A. M. Sirunyan et al. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$ =13 TeV". In: *JINST* 13 (2018), P06015. DOI: 10.1088/1748-0221/13/06/P06015. arXiv: 1804.04528 [physics.ins-det].

[70] A. M. Sirunyan et al. "Precision luminosity measurement in proton-proton collisions at $\sqrt{s}$ = 13 TeV in 2015 and 2016 at CMS". In: *Eur. Phys. J. C* 81 (2021), p. 800. DOI: 10.1140/epjc/s10052-021-09538-2. arXiv: 2104.01927 [hep-ex].

[71] Albert M Sirunyan et al. "Measurement of properties of $B_s^0 \to \mu^+\mu^-$ decays and search for $B^0 \to \mu^+\mu^-$ with the CMS experiment". In: *JHEP* 04 (2020), p. 188. DOI: 10.1007/JHEP04(2020)188. arXiv: 1910.12127 [hep-ex].

[72] Albert M Sirunyan et al. "Measurement of the CP-violating phase ϕ_s in the $B_s^0 \to J/\psi\phi$ channel in proton-proton collisions at $\sqrt{s}$=13 TeV". In: *submitted to Phys. Lett. B.* (2020). arXiv: 1902.00571 [hep-ex].

[73] Torbjörn Sjöstrand et al. "An Introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[74] David M. Straub. *flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond*. 2018. arXiv: 1810.08132 [hep-ph].

[75] David M. Straub. *New physics correlations in rare decays*. https://arxiv.org/abs/1012.3893. Accessed: 2023-04-20. 2011. arXiv: 1012.3893 [hep-ph].

[76] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/349375.

[77] *The CMS hadron calorimeter project: Technical Design Report*. Technical design report. CMS. The following files are from http://uscms.fnal.gov/pub/hcal$_t$*drandmaynotbethever*. Geneva: CERN, 1997. URL: http://cds.cern.ch/record/357153.

[78] *The CMS magnet project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. DOI: 10.17181/CERN.6ZU0.V4T9. URL: https://cds.cern.ch/record/331056.

[79] S. Weinberg. "The making of the Standard Model". In: *The European Physical Journal C* 34.1 (May 2004), pp. 5–13. DOI: 10.1140/epjc/s2004-01761-1. URL: https://doi.org/10.11402Fepjc2Fs2004-01761-1.

[80] Kenneth G. Wilson and Wolfhart Zimmermann. "Operator product expansions and composite field operators in the general framework of Quantum Field theory". In: *Communications in Mathematical Physics* 24.2 (1972), pp. 87–106. DOI: 10.1007/bf01878448.

[81] Bernard de Wit. *Supergravity*. 2002. DOI: 10.48550/ARXIV.HEP-TH/0212245. URL: https://arxiv.org/abs/hep-th/0212245.

[82] Julia Woithe, Gerfried J Wiener, and Frederik F Van der Veken. "Let's have a coffee with the Standard Model of particle physics!" In: *Physics Education* 52.3 (Mar. 2017), p. 034001. DOI: 10.1088/1361-6552/aa5b25. URL: https://dx.doi.org/10.1088/1361-6552/aa5b25.