Università degli Studi di Padova · Dipartimento di Fisica e Astronomia

---

Scuola di dottorato di ricerca in Fisica
Ciclo XXVII

B flavour tagging with leptons and measurement of the CP violation phase ϕ_s in the $B^0_s \to J/\psi\phi$ decay at the CMS experiment

Direttore della Scuola:
prof. **Andrea Vitturi**

Supervisore:
prof. **Franco Simonetto**

Co-supervisore:
prof. **Martino Margoni**

Dottorando: **Jacopo Pazzini**

Anno Accademico 2014-2015

Abstract

In questa tesi è presentato lo sviluppo e l'ottimizzazione di un algoritmo per la determinazione del sapore di produzione di mesoni neutri B^0 e B^0_s, che sfrutta muoni ed elettroni prodotti nel decadimento semileptonico dell'altro adrone-b prodotto nel processo $pp \to \mathrm{b\bar{b}X}$ a LHC. La carica dei leptoni è utilizzata per identificare il sapore del mesone-b neutro. Tre campioni di eventi simulati contenenti i decadimenti $\mathrm{B}^0_\mathrm{s} \to \mathrm{J}/\psi\,\phi$, $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ e $\mathrm{B}^0 \to \mathrm{J}/\psi\,\mathrm{K}^*$ sono sfruttati per lo sviluppo e il controllo delle prestazioni dell'algoritmo. Due reti neurali sono definite separatamente per muoni ed elettroni ed utilizzate per parametrizzare la probabilità di errata attribuzione del sapore ω dell'algoritmo, utilizzando rispettivamente 24 000 e 20 400 eventi simulati contenenti decadimenti $\mathrm{B}^0_\mathrm{s} \to \mathrm{J}/\psi\,\phi$.
Le prestazioni dell'algoritmo sono inoltre misurate e calibrate su un campione di eventi $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ raccolti dall'esperimento CMS nel corso del 2012, pari a 20 fb^{-1}. La carica del kaone determina univocamente il sapore del B^+ al momento della produzione e consente la misura diretta della probabilità di errata attribuzione del sapore. Il potere di identificazione dell'algoritmo $\mathcal{P}_{tag} = \varepsilon_{tag}(1-2\omega)^2$, è misurato essere pari a 0.833 ± 0.024 (stat.) $\pm$ 0.012 (syst.) % per i muoni e 0.483 ± 0.020 (stat.) $\pm$ 0.003 (syst.) % per gli elettroni. Il potere di identificazione risultante dalla combinazione dei due algoritmi è 1.307 ± 0.031 (stat.) $\pm$ 0.007 (syst.) %.
L'algoritmo combinato di identificazione del sapore è utilizzato nella misura dei parametri ϕ_s e $\Delta\Gamma_\mathrm{s}$ di violazione della simmetria CP, sensibili a potenziali processi di nuova fisica non inclusi nella descrizione del modello standard. Utilizzando il campione di dati raccolti dall'esperimento CMS nel corso del 2012, è eseguita una analisi angolare dipendente dal tempo dello stato finale $\mu^+\mu^-\mathrm{K}^+\mathrm{K}^-$ del decadimento $\mathrm{B}^0_\mathrm{s} \to \mathrm{J}/\psi\,\phi$, includendo nell'analisi l'algoritmo di identificazione del sapore. Un totale di 49 000 decadimenti ricostruiti del mesone B^0_s sono utilizzati per estrarre il valore sperimentale della fase debole ϕ_s e la differenza di vita media $\Delta\Gamma_\mathrm{s}$:

$$\phi_\mathrm{s} = -0.075 \pm 0.097\ \text{(stat.)} \pm 0.031\ \text{(syst.) rad}$$
$$\Delta\Gamma_\mathrm{s} = 0.095 \pm 0.013\ \text{(stat.)} \pm 0.007\ \text{(syst.) ps}^{-1}$$

Abstract

This thesis presents the development and optimization of an algorithm used to determine the flavour at production time of neutral B^0 and B^0_s mesons. The flavour tagging algorithm developed in this thesis exploits muons and electrons produced in the semileptonic decay of the additional b-hadron produced in $pp \to \mathrm{b\bar{b}X}$ collisions at the LHC. The charge of the lepton is used to infer the flavour of the neutral b-meson. Three simulated samples of $\mathrm{B}^0_\mathrm{s} \to \mathrm{J}/\psi\,\phi$, $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ and $\mathrm{B}^0 \to \mathrm{J}/\psi\,\mathrm{K}^*$ decays are exploited to develop and test the algorithm. Two independent neural networks are defined for muons and electrons, trained on 24 000 and 20 400 simulated $\mathrm{B}^0_\mathrm{s} \to \mathrm{J}/\psi\,\phi$ events respectively, to parametrize the probability of wrong flavour tag ω of the algorithm. The tagging performances are further measured and calibrated on a sample of self-tagging $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ decays collected by the CMS experiment during 2012, corresponding to 20 fb^{-1}. The charge of the kaon univocally determines the flavour of the B^+ at production time and allows the direct measure of the mis-identification probability. A tagging power $\mathcal{P}_{tag} = \varepsilon_{tag}(1-2\omega)^2$ of 0.833 ± 0.024 (stat.) $\pm$ 0.012 (syst.) % is measured using muons and 0.483 ± 0.020 (stat.) $\pm$ 0.003 (syst.) % using electrons. Combining the two algorithms results in the overall tagging power of 1.307 ± 0.031 (stat.) $\pm$ 0.007 (syst.) %.

The combined lepton flavour tagging algorithm is used in the measurement of the charge-parity (CP) violation parameters ϕ_s and $\Delta\Gamma_\mathrm{s}$, sensitive to potential new physics processes not included in the standard model description. A time-dependent and flavour-tagged full angular analysis of the $\mu^+\mu^-\mathrm{K}^+\mathrm{K}^-$ final state of the $\mathrm{B}^0_\mathrm{s} \to \mathrm{J}/\psi\,\phi$ decay is performed based on the 2012 CMS dataset. A total of 49 000 reconstructed B^0_s decays are used to extract the weak phase ϕ_s and decay with difference $\Delta\Gamma_\mathrm{s}$ values:

$$\phi_\mathrm{s} = -0.075 \pm 0.097\,(\text{stat.}) \pm 0.031\,(\text{syst.})\ \text{rad}$$
$$\Delta\Gamma_\mathrm{s} = 0.095 \pm 0.013\,(\text{stat.}) \pm 0.007\,(\text{syst.})\ \text{ps}^{-1}$$

Contents

Introduction

The standard model of particle physics provides the most accurate and consistent description of the subatomic nature of the universe developed so far as it has withstood a multitude of experimental tests over the years. Within the standard model the discrete charge-parity (CP) symmetry happens to be spontaneously violated due to the presence of a complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) matrix of the so-called flavour sector of the theory. The precise determination of CP violating effects and the measurement of the CKM matrix parameters is required to check the standard model consistency. Potential discrepancies from the predicted values could lead to infer the existence of new physics processes. One of the main fields for the search of indirect evidences of new physics is in fact represented by loop mediated processes, where new particles could appear as virtual contributions in box and penguin diagrams leading to sizeable effects on various flavour-related observables also potentially introducing new measurable sources of CP violation. Proceeding with the exchange of virtual particles, indirect searches of new physics can furthermore probe higher energy scales than those currently accessible with the direct search for new particles. The parameter ϕ_{s} governs the CP violation in the $\mathrm{B}^0_{\mathrm{s}}$ system and is one of the most sensitive physics observables to potential effects of new physics in the flavour sector. Given the small magnitude of the expected standard model value and the accurate theoretical predictions currently available, the precise experimental determination of this parameter in $\mathrm{c\bar{c}s}$ transitions such as the $\mathrm{B}^0_{\mathrm{s}} \to \mathrm{J}/\psi\,\phi$ decay make possible to infer stringent limits on the size of new physics effects. The phase ϕ_{s} arise from the quantum interference between the direct $\mathrm{B}^0_{\mathrm{s}} \to \mathrm{J}/\psi\,\phi$ decay and the analogous process where the $\mathrm{B}^0_{\mathrm{s}}$-$\overline{\mathrm{B}^0_{\mathrm{s}}}$ flavour oscillation occurred before the decay. Being the $\mathrm{J}/\psi\phi$ decay channel a non-definite final state, a time-dependent angular analysis of the $\mu^+\mu^-\mathrm{K}^+\mathrm{K}^-$ decay products is needed in order to disentangle the CP-even and CP-odd components. The experimental sensitivity on ϕ_{s} can be enhanced by the knowledge of the flavour of the b-meson at production time, since the final state of the decay is accessible to both $\mathrm{B}^0_{\mathrm{s}}$ and $\overline{\mathrm{B}^0_{\mathrm{s}}}$. This information can be obtained with flavour tagging algorithms, which infer the flavour of neutral b-mesons exploiting the properties of the particles reconstructed within the hadronization cone of the $\mathrm{B}^0_{\mathrm{s}}$ (same side algorithms) or by studying the decay products of the other b-hadron produced in the event from the $\mathrm{b\bar{b}}$ pair production mechanisms (opposite side algorithms). The topic of this thesis is the development of an opposite side flavour tagging algorithm for the CMS experiment at the LHC collider and the flavour tagged angular analysis of the $\mathrm{B}^0_{\mathrm{s}} \to \mathrm{J}/\psi\,\phi$ decay to measure the CP-violating weak phase ϕ_{s}.

This thesis is structured as follows. In Chapter 1 a brief introduction on the standard model is given. The flavour sector of the theory is presented focusing on the violation of the CP symmetry and the features related to the $\mathrm{b} \to \mathrm{c\bar{c}s}$ transition involved in $\mathrm{B}^0_{\mathrm{s}} \to \mathrm{J}/\psi\,\phi$

decay. Chapter 2 is dedicated to the description of the experimental apparatus, containing a summary of the features of the LHC hadron collider, and a more in-depth characterization of the CMS experiment in terms of the sub-detectors, event reconstruction and particle identification algorithms. The identification of the flavour of neutral b-mesons is first introduced in Chapter 3, where the basic principles of the various tagging techniques are reviewed, and the tagging strategy developed in CMS is illustrated. Chapter 3 also describe the reconstruction and selection of the data and simulated events used in this thesis. The development, characterization and calibration of the CMS opposite side lepton tagger are discussed in detail in Chapter 4. The features of the single-particle taggers are treated independently for electrons and muons first, and the results of the lepton-tagger combination are then summarized. Finally, the flavour tagged angular analysis of the $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ decay is described in Chapter 5, where the resulting weak CP violating phase ϕ_s and decay width difference $\Delta\Gamma_\mathrm{s}$ are reported and compared to the latest standard model predictions.

Chapter 1

Theory

1.1 The Standard Model of Particle Physics

The standard model of particle physics (SM) [1–3] is the most successful theoretical description of the elementary particles and their interactions. It is a renormalizable Quantum Field Theory (QFT) based on the fundamental principles of the special relativity and the quantum mechanics, and includes the description of the electroweak interaction (EW) [4, 5] and of the strong interaction (QCD) [6, 7]. The description of the SM relies on the invariance of its Lagrangian ($\mathcal{L}$) to the gauge transformation symmetry group G_{SM}, which determines the interactions between the fundamental particles.

$$G_{\text{SM}} = \text{SU}(3)_{\text{C}} \otimes \text{SU}(2)_{\text{L}} \otimes \text{U}(1)_{\text{Y}} \tag{1.1}$$

The $\text{SU}(2)_{\text{L}} \otimes \text{U}(1)_{\text{Y}}$ invariance group represents the theoretical unification of the electromagnetic and the weak forces in the electroweak interaction, while the $\text{SU}(3)_{\text{C}}$ group refers to the strong interaction. The particles arising from the excitations of the gauge fields are spin-1 vector bosons. The electroweak interactions are mediated by the massless photon γ and by the massive gauge bosons $W^{\pm}$ and Z^0, while the strong force is mediated by eight massless gluons g. The particles included in the SM description are three spin-1/2 fermion generations and a single spin-0 scalar, respectively represented under the G_{SM} group as Weyl spinors and as a complex doublet. The three fermionic families can further be divided into two fundamental field types: quarks and leptons. There are in fact six different quarks that can be divided in two groups, the up type quarks (u, c, t) and the down type quarks (d, s, b). The six leptons can be divided into charged (e, μ, τ) and the corresponding neutral spinors (ν_e , ν_μ , ν_τ). The representations of the fermionic fields can be described as follows:

$$Q_{\text{L}i}\,(3,2)_{+\frac{1}{6}}\,, \quad U_{\text{R}i}\,(3,1)_{+\frac{2}{3}}\,, \quad D_{\text{R}i}\,(3,1)_{-\frac{1}{3}}\,, \quad L_{\text{L}i}\,(1,2)_{-\frac{1}{2}}\,, \quad E_{\text{R}i}\,(1,1)_{-1} \tag{1.2}$$

where, for instance, the left-handed quarks representations Q_{L} are triplets of $\text{SU}(3)_{\text{C}}$ doublets of $\text{SU}(2)_{\text{L}}$ and carry the hypercharge $\text{Y} = +1/6$. The sub-index $i = 1, 2, 3$ represents the flavour (or generation) of the spinor fields. The spontaneous ElectroWeak Symmetry Breaking (EWSB) [8–12] mechanism provides mass terms in the Lagrangian to otherwise massless particles preserving the gauge invariance, and predicts the existence of an additional spin-0 (scalar) boson, included into the SM Lagrangian with the representation:

$$\phi\,(1,2)_{+\frac{1}{2}} \tag{1.3}$$

The SM Lagrangian is the most general renormalizable Lagrangian that is consistent to the gauge symmetry 1.1 and the particle content 1.2 and 1.3 hereby described. The SM has 18 free parameters to be determined experimentally. As any locally Lorentz invariant quantum field theory, the SM conserves CPT. In the SM, the discrete symmetries C and P are maximally violated, since both change the chirality of the fermion fields and the SM having different gauge representations for left-handed and right-handed representations of the fermionic fields. The SM also violates the CP symmetry. The violation of CP symmetry arises in the SM from the presence in the Lagrangian of complex coefficients in the couplings between fermions and the scalar field, as will be discussed in detail within this chapter.

1.1.1 The Strong Interaction

Quantum ChromoDynamics (QCD) is the gauge field theory which describes the strong interaction between the quark fermionic spinor fields and the QCD vector bosons, the gluons. This interaction is associated to the non-abelian (and unbroken) $SU(3)_C$ colour symmetry group. Due to the $SU(3)$ group structure, an octet of massless spin-one gauge bosons arise, and due to the non-abelian character of the QCD interaction, gluons can interact with each other, transforming under the adjoint representation of the colour group. The quarks carry the colour charge (denoted r,g, and b) and are in the fundamental representation of the group. From the product of quark fields representations two colourless QCD bound states can be formed, mesons and baryons, collectively denoted as hadrons. Mesons are bosons composed by a quark and an anti-quark, whereas baryons are instead fermions composed of three quarks. The QCD Lagrangian can be described by the following equation:

$$L^{\mathrm{QCD}} = \bar{\psi} i \gamma^{\mu} D_{\mu} \psi - \frac{1}{4} G^{a\mu\nu} G^{a}_{\mu\nu} \tag{1.4}$$

where ψ is a generic massless quark spinor field, and $G^{a}_{\mu\nu}$ is the field tensor of the gluonic fields G^{a}_{μ}. The covariant derivative of QCD is described using the λ_a Gell-Mann matrices, as:

$$D_{\mu} = \partial_{\mu} + \frac{i}{2} g_s \lambda_a G^{a}_{\mu} \tag{1.5}$$

where the QCD coupling constant g_s, can also be rewritten to as $\alpha_s \equiv g_s^2/4\pi$. The gluon tensor field $G^{a}_{\mu\nu}$ is defined as:

$$G^{a}_{\mu\nu} = \partial_{\mu} G^{a}_{\nu} - \partial_{\nu} G^{a}_{\mu} - g_s f_{abc} G^{b}_{\mu} G^{c}_{\nu} \tag{1.6}$$

The last term of Equation 1.6 is responsible for the self-interaction of the gluon fields. f_{abc} are the structure constants of the $SU(3)$ group. Due to the renormalization process of the theory, the QCD coupling constant α_s is subject to the so-called running, i.e. it varies depending on the transferred momentum Q. In the one-loop approximation, for a given transferred momentum Q^2, the value of α_s is described by:

$$\alpha_s(Q^2) = \frac{1}{\beta_0 \ln{(Q^2/\Lambda^2)}}, \quad \beta_0 = \frac{33 - 2n_f}{12\pi} \tag{1.7}$$

This formula holds true in an effective theory in the approximation of n_f quark flavours with masses $m_q \ll Q^2$ and the remaining quarks are to be considered decoupled from the strong interaction. The dimensional parameter Λ in Eq. 1.7 fixes the scale at which the coupling constant becomes large and the description of the QCD physics processes can not thus rely on perturbative evaluation. Thanks to the peculiar features of the running of α_s, in the high energy regime the QCD is well described in terms of weakly interacting quarks and gluons, the so-called asymptotic freedom, and a fully perturbative treatment of the strong interaction is possible. The strength of the interaction between coloured particles increases with the decreasing of the Q^2. Thus with the increasing of the distance among the interacting quarks a complex dynamic arise which cannot be described perturbatively, resulting in the so-called quark confinement.

1.1.2 The Electroweak Interaction

The gauge theory of the electroweak interaction is based on the invariance under $\mathrm{U}(1)_\mathrm{Y} \otimes \mathrm{SU}(2)_\mathrm{L}$ gauge transformations. The L subscript refers to the fact that the constituent of the electroweak interaction are left-handed weak isospin doublets ($T = 1/2$), while the right-handed components transforms as singlets. Both leptons and quarks interact under the electroweak symmetry, as:

$$\begin{pmatrix} \nu_L \\ \ell_L \end{pmatrix}, \quad \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad \ell_R, u_R, d_R \tag{1.8}$$

where ν and ℓ are the generic neutrino and lepton fields, while u and d are the generic up- and down-type quarks. It should be noted that the neutrino is the only elementary particle included in the SM with the left-handed component only. The projection of the weak isospin T_3 and the hypercharge Y are connected via the Gell Mann-Nishijima relation with the electric charge Q:

$$Y = 2\,(Q - T_3) \tag{1.9}$$

The boson fields involved in the SU(2) gauge symmetry are the three W^a_μ, related to the Pauli matrices τ^a and to the gauge coupling constant g, while the B_μ boson field arise from the U(1) symmetry, with coupling g'. The electroweak Lagrangian is written as:

$$L^{\mathrm{EW}} = \bar{\psi} i \gamma^\mu D_\mu \psi - \frac{1}{4} W^{a\mu\nu} W^a_{\mu\nu} - B_{\mu\nu} B^{\mu\nu} \tag{1.10}$$

where the covariant derivative is acting differently for the left doublets (L) and the right singlets (R) of the generic spinor field ψ:

$$D^L_\mu = \partial_\mu + \frac{i}{2} g \tau^a W^a_\mu + \frac{i\,'}{g} Y B_\mu \tag{1.11}$$

$$D^R_\mu = \partial_\mu + \frac{i\,'}{g} Y B_\mu \tag{1.12}$$

The gauge tensors are defined by:

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g \epsilon_{abc} W^b_\mu W^c_\nu \tag{1.13}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.14}$$

The W^a_μ and B_μ do not however represent the observable states of the vector bosons responsible for the weak (charged and neutral) currents and the electromagnetic interaction. The physical fields A_μ (corresponding to the photon), Z_μ , W^+_μ and W^-_μ (corresponding to the Z^0 and $W^\pm$), are in fact described within the Glashow-Salam-Weinberg electroweak theory as linear combination of the W_μ^a and B_μ fields:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp W_\mu^2\right) \tag{1.15}$$

$$Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W_\mu^3 \tag{1.16}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \tag{1.17}$$

where θ_W is the Weinberg electroweak mixing angle. The coupling constants g and g' are linked to θ_W and to the electromagnetic coupling constant (the electron's charge e) through the relation:

$$\tan\theta_W = \frac{g'}{g}, \quad e = g\sin\theta_W = g'\cos\theta_W \tag{1.18}$$

The electroweak interaction described so far is subject to the following issues:

- Local $\mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{U}(1)_\mathrm{Y}$ gauge invariance forbids massive gauge bosons
- Local $\mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{U}(1)_\mathrm{Y}$ gauge invariance forbids massive fermions
- Unitarity violation in processes like WW-scattering (i.e. non-renormalizable theory)

which are solved in the SM by the introduction of a scalar field with a peculiar mechanism of spontaneous breaking of the electroweak gauge symmetry.

1.1.3 The Electroweak Symmetry Breaking

An electroweak complex scalar field with four degrees of freedom is added to the electroweak Lagrangian. The field is described as a $\mathrm{SU}(2)_\mathrm{L}$ doublet:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \varphi_1 + i\varphi_2 \\ \varphi_3 + i\varphi_4 \end{pmatrix} \tag{1.19}$$

A generic Lagrangian describing the evolution of the scalar field ϕ can be written as follows:

$$L^\mathrm{S} = D_\mu\phi D^\mu\phi^\dagger + V(\phi) \tag{1.20}$$

where the covariant derivative under the $\mathrm{SU}(2)_\mathrm{L}$ group transformation is the one described in Eq. 1.11, and the potential term $V(\phi)$ is described as follows:

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda\left(\phi^\dagger\phi\right)^2 \tag{1.21}$$

The EW symmetry breaking mechanics requires the quadratic term $\mu^2 < 0$ and a positive quartic coupling $\lambda > 0$. For this conditions a degenerate set of minima in this potential is produced, corresponding to:

$$\phi^\dagger\phi = \frac{\mu^2}{2\lambda} \tag{1.22}$$

From Equation 1.21 a vacuum expectation value $v/\sqrt{2} = \mu/\lambda$ can be obtained, where $v \approx 246$ GeV. Although the full Lagrangian is still symmetric under the EW gauge transformation group, the vacuum is not. By choosing a vacuum description $\varphi_1 = \varphi_2 = \varphi_4 =$

0 and $\varphi_3 = v$ the $\mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{U}(1)_\mathrm{Y}$ symmetry breaks to a $\mathrm{U}(1)_\mathrm{em}$, the gauge symmetry of QED (quantum electrodynamic) [13], and any fluctuation of the ϕ field can be expressed as:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{1.23}$$

The field $h(x)$ is now the only residual physical degrees of freedom of ϕ in the particle spectrum. The degrees of freedom related to the broken generators of the EW symmetry are absorbed by the $W^\pm$ and Z^0 bosons, which in turn acquire their mass via the coupling terms between the h field and the gauge bosons, which are proportional to the vacuum expectation value v. The scalar boson itself (h) acquire mass. It is worth to be noted that as the scalar boson does not carry electrical charge (i.e., is invariant under $\mathrm{U}(1)_\mathrm{em}$) , it does not couple with the photon, which remains massless. The boson masses in the EW sector are given at tree level (i.e. to lowest order in perturbation theory) by:

$$m_{W^\pm} = \frac{1}{2} v g \tag{1.24}$$

$$m_{Z^0} = \frac{1}{2} v \sqrt{g^2 + g'^2} \tag{1.25}$$

$$m_\gamma = 0 \tag{1.26}$$

$$m_h = v\sqrt{2\lambda} \tag{1.27}$$

The mass of the fermions are obtained introducing Yukawa interaction terms which couple left-handed fermionic doublets ψ_L, to the corresponding right-handed singlets ψ_R and to the scalar field ϕ in the vertex $-\lambda_f \bar{\psi}_L \phi \psi_R$, where λ_f represents the Yukawa coupling constant of a generic fermion f. The Yukawa couplings are free parameters of the SM. As previously discussed, within the SM description the neutral leptonic fields (the neutrinos) are only described by left-handed components. Missing the right-handed singlet, neutrinos are therefore massless, while the charged leptons are massive:

$$m_\ell = \frac{v\lambda_\ell}{\sqrt{2}} \tag{1.28}$$

$$m_\nu = 0 \tag{1.29}$$

In the quark sector, the Yukawa couplings of the three quark doublets are mixed so that the quark eigenstates of the electroweak interaction are not proper states of the flavour basis. The flavour interaction eigenstates are instead linear combinations of the mass eigenstates. The convention normally chosen is such that the up-type mass and weak eigenstates are aligned, whilst the down-type mass eigenstates are rotated to form the weak eigenstates. A unitary mixing matrix is therefore introduced in the theory, corresponding to the base change from the interaction eigenstates to the mass eigenstates one.

1.2 Flavour Physics and CP Violation

Within the SM the only source of flavour-changing interactions arises from the quark Yukawa sector and it is originated from a rotation of the quarks flavour basis with respect to the weak-interaction basis by the 3 × 3 complex CKM matrix. As a result of the fact that the CKM matrix is not diagonal, the weak interactions mediated by $W^\pm$ bosons couples quarks of different generations via flavour-changing charged-currents, FCCC; flavour transitions mediated by neutral currents (flavour-changing neutral-currents, FCNC) are highly suppressed in the SM. The latter cannot in fact occur at tree-level within the SM, but they require the intermediate exchange of a quark and a W boson via loop transitions.

1.2.1 The CKM Matrix

The matrix responsible for the flavour-changing interactions is the Cabibbo-Kobayashi-Maskawa (or CKM) matrix [14, 15]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.30}$$

where the down quark vector on the left side correspond to the electroweak interaction eigenstates (q') whilst the vector to the right side correspond to the mass eigenstates (q). The elements of the CKM matrix V_{ij} represent the FCCC couplings between up-type (i) and down-type (j) quarks. Thanks to the EW symmetries of the SM and due to the unitarity constraint, the independent parameters of the matrix can be reduced down to 4 real parameters, $n(n-1)/2 = 3$ rotation angles and $(n-1)(n-2)/2 = 1$ phase. Among the infinite possible representations of the CKM matrix, the so called Standard parametrization is realized by using the three angles $\theta_{ij} = (\theta_{12}, \theta_{23}, \theta_{13})$ and a CP-violating phase δ.

$$\begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{1.31}$$

where $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$. It is immediately clear in this representation that although CP-violation is only present in the CKM matrix due to the single phase δ, the latter appears in multiple elements of the matrix with different coefficients, i.e. CPV is expected in all the three sectors for d, s and b-quarks. The current knowledge of the CKM matrix elements moduli, as obtained from [16], is the following:

$$|V_{CKM}| = \begin{pmatrix} 0.974235^{+0.000080}_{-0.000158} & 0.22551^{+0.00068}_{-0.00034} & 0.00357^{+0.00016}_{-0.00015} \\ 0.22537^{+0.00068}_{-0.00035} & 0.973395^{+0.000095}_{-0.000176} & 0.04136^{+0.00071}_{-0.00128} \\ 0.00855^{+0.00018}_{-0.00030} & 0.04062^{+0.00070}_{-0.00125} & 0.999138^{+0.000052}_{-0.000030} \end{pmatrix} \tag{1.32}$$

An alternative way to represent the CKM parameters is based on the observed hierarchy among the parameters. It is in fact possible to expand all the CKM terms in powers of $\lambda = |V_{us}|$, the sine of the Cabibbo angle ($\lambda \approx 0.23$). The expansion up to and including terms $\mathcal{O}\left(\lambda^5\right)$ is given by the Wolfenstein parametrization [17, 18]:

$$\begin{pmatrix} 1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + \frac{A^2\lambda^5}{2}\left[1 - 2\left(\rho + i\eta\right)\right] & 1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8}\left(1 + 4A^2\right) & A\lambda^2 \\ A\lambda^3\left[1 - \left(1 - \frac{\lambda^2}{2}\right)\left(\rho + i\eta\right)\right] & -A\lambda^2 - \frac{A\lambda^4}{2}\left[1 - 2\left(\rho + i\eta\right)\right] & 1 - \frac{A^2\lambda^4}{2} \end{pmatrix} \tag{1.33}$$

where A, ρ and η are the remaining real parameters, all of order unity ($A \approx 0.81$, $\rho \approx 0.15$, $\eta \approx 0.34$).

The unitarity of the CKM matrix leads to a set of 9 equations, 6 requiring the sum of three complex quantities to vanish

$$\sum_{k \in \{u,c,t\}} V_{ki}V^*_{kj} = \delta_{ij} \quad (i, j \in \{d, s, b\}) \tag{1.34}$$

these can be represented as triangles in the complex plane: the lengths of the sides of the triangles are the moduli of CKM matrix element products, while the angles are defined from the relative phases. Remarkably, the areas of all triangles are the same and can be calculated as half of the Jarlskog invariant (J) [19], which is a phase-convention independent measure of CP violation, defined by $J = c_{12}c_{13}^2c_{23}s_{12}s_{23}\sin\delta$. For the Jarlskog invariant to be non-zero it is required that there must be a global non-zero phase δ within the CKM and vice versa, as confirmed from the latest experimental result $J = 2.97^{+0.18}_{-0.20}$ [18]. Two unitarity triangles of particular interest are the ones shown in Figure 1.1, which are constructed from the relations:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.35}$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \tag{1.36}$$

These two conditions are usually referred to as the B^0 and B_s^0 unitarity triangles, where B^0 and B_s^0 are mesons composed by a $\overline{\mathrm{b}}$ anti-quark and a lighter down-type d and s quark respectively: B^0 ($\overline{\mathrm{b}}$-d), B_s^0 ($\overline{\mathrm{b}}$-s). The unitarity triangles are usually represented with respect to the best known term of the sum, in such a way that one of the normalised sides is the unity vector of one axis. For instance, in the case of the B^0 unitarity triangle, Equation 1.35 can be rewritten as:

$$R_t e^{-i\beta} + R_u e^{+i\gamma} = 1 \tag{1.37}$$

where R_t,R_u and β, γ are the non-trivial sizes and angles of the normalised unitarity triangle.

$$R_u \equiv \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| \simeq \sqrt{\rho^2 + \eta^2} \tag{1.38}$$

$$R_t \equiv \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| \simeq \sqrt{(1-\rho)^2 + \eta^2} \tag{1.39}$$

$$\beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \simeq \arg\left(\frac{1}{1-\rho-i\eta}\right) \tag{1.40}$$

$$\gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \simeq \arg\left(\rho + i\eta\right) \tag{1.41}$$

while the remaining angle α is defined as

$$\alpha \equiv \pi - \beta - \gamma \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \simeq \arg\left(-\frac{1-\rho-i\eta}{\rho+i\eta}\right) \tag{1.42}$$

The unitarity triangle can be therefore summarized by one complex number:

$$\bar{\rho} + i\bar{\eta} = R_u e^{i\gamma} \tag{1.43}$$

where $(\bar{\rho}, \bar{\eta})$ represent the coordinates of the non trivial apex (the others being $(0,0)$ and $(1,0)$) of the unitarity triangle in the complex plane.

Processes dominated by loop contributions in the SM are sensitive to new physics (NP). An important goal of flavour physics is to overconstrain the CKM elements, by performing and comparing many measurements in the $(\bar{\rho}, \bar{\eta})$ plane under the assumption that the flavour-changing processes are solely described by the SM. The global CKM fit provides so far a successful test of the validity of the SM. Given the excellent consistency of the global fit, flavour violation and CP violation, flavour-changing processes seem in fact to be dominated by the CKM mechanism of the SM. Potential source of NP contributions in flavour-changing processes are therefore highly suppressed with respect to SM contributions. A greater chance for the potential discovering of NP effects (not described in the SM) might therefore reside in the study of loop-mediated FCNC transitions, such as the ones that mediates neutral-mesons oscillations described in the next section.

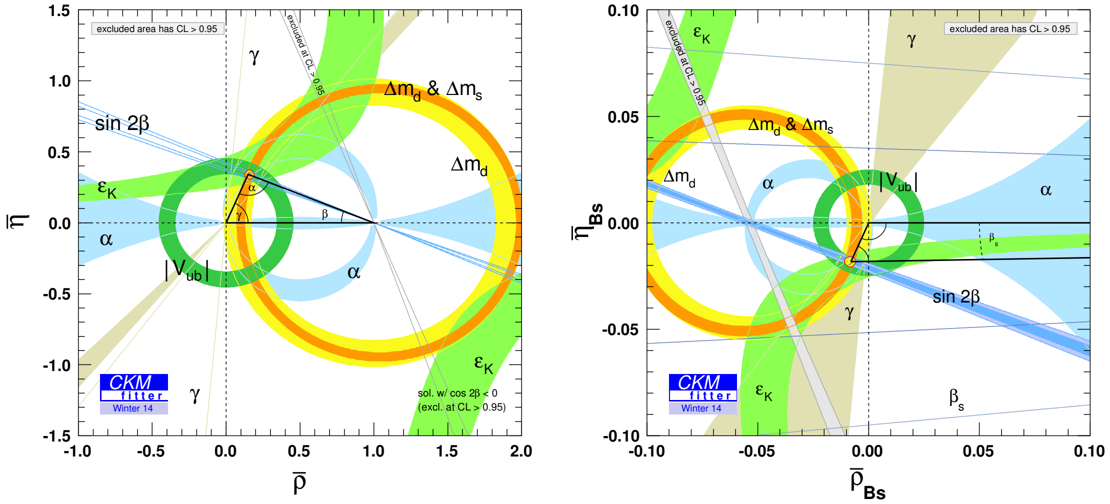


Figure 1.1: Individual constraints and the global fit result on the $(\bar{\rho} - \bar{\eta})$ plane for the B^0 (left) and B_s^0 (right) unitarity triangles. Shaded areas have 95% confidence level (CL). The allowed $(\bar{\rho}, \bar{\eta})$ apex region of coordinates at 95% CL is shown in yellow with a red contour. A good overall agreement is observed between the individual constraints.

1.2.2 Neutral B Meson Mixing

Neutral mesons, with the exception of the pion, are subject to the mixing, i.e., time-dependent oscillations from particle to antiparticle through weak FCNC transitions that change the meson flavour by two units, $\Delta F = 2$. The oscillation process arises directly from the disparity between the flavour eigenstates and the mass eigenstates of the quarks which leads to a mixing between the quark generations produced by the CKM mechanism. The phenomenology of CP violation is different in K^0, D^0, B^0 and B_s^0 decays primarily because each of these systems is governed by a different relation between decay rates, oscillations, and lifetimes. However, the overall underlying mechanism of CP violation is identical for all pseudoscalar mesons and it will therefore be discussed with reference to a generic neutral B-meson: B_q, where $q = s, d$ [20].

Before the time when the meson decays the system is described as a superposition of the two meson states B_q and $\overline{B_q}$, and its time evolution is governed by the Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix} |B_q(t)\rangle \\ |\overline{B_q}(t)\rangle \end{pmatrix} = \boldsymbol{H}\begin{pmatrix} |B_q(t)\rangle \\ |\overline{B_q}(t)\rangle \end{pmatrix} \equiv \left(\boldsymbol{M} - \frac{i}{2}\boldsymbol{\Gamma}\right)\begin{pmatrix} |B_q(t)\rangle \\ |\overline{B_q}(t)\rangle \end{pmatrix} \tag{1.44}$$

where $\boldsymbol{M}$ and $\boldsymbol{\Gamma}$ are 2×2 Hermitian matrices known as the mass and decay width matrices. The two matrices are associated with transitions via off-shell (dispersive) and on-shell (absorptive) intermediate states respectively. The CPT invariance of the SM requires the diagonal terms of $\boldsymbol{H}$ to be equal $H_{11} = H_{22}$. Diagonal elements of $\boldsymbol{M}$ and $\boldsymbol{\Gamma}$ are associated with the flavour-conserving transitions $\mathrm{B_q} \rightleftarrows \mathrm{B_q}$ ($\overline{\mathrm{B_q}} \rightleftarrows \overline{\mathrm{B_q}}$), while the off-diagonal terms M_{12} and Γ_{12} arise in the SM from the flavour-changing transitions $\mathrm{B_q} \rightleftarrows \overline{\mathrm{B_q}}$ ($\overline{\mathrm{B_q}} \rightleftarrows \mathrm{B_q}$) via box diagrams, as represented in Figure 1.2.

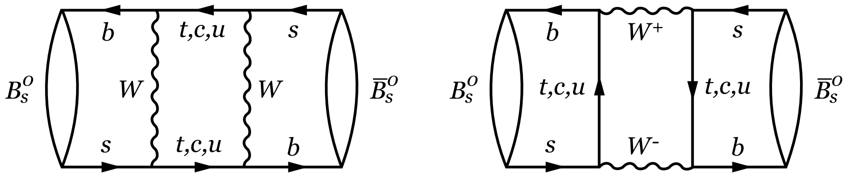


Figure 1.2: Dominant box diagrams for the $\mathrm{B_q} \to \overline{\mathrm{B_q}}$ transitions in the case of $q = s$.

The diagonal elements M_{11} and M_{22} describe in principle the masses of the $\mathrm{B_q}$ and $\overline{\mathrm{B_q}}$ states, and are therefore equal due to the CPT invariance. Similarly, also the lifetime of the two states must be the same, thus $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$. However, since $\boldsymbol{H}$ is not-diagonal, the meson states $\mathrm{B_q}$ and $\overline{\mathrm{B_q}}$ are not mass eigenstates and have not defined mass and widths. It is possible to denote the eigenstates of $\boldsymbol{H}$ as the heavy (H) and light (L) mass states, respectively $\mathrm{B_H}$ and $\mathrm{B_L}$, with masses $m_H > m_L$. Such eigenstates can be described by the superposition of the $\mathrm{B_q}$ and $\overline{\mathrm{B_q}}$ states:

$$|\mathrm{B_{L,H}}\rangle = p\,|\mathrm{B_q}\rangle \pm q\,|\overline{\mathrm{B_q}}\rangle \tag{1.45}$$

where the complex coefficients p and q obey the normalization condition $|p|^2 + |q|^2 = 1$. The time evolution of the mass eigenstates system is thus governed by the equation

$$i\frac{d}{dt}\begin{pmatrix} |\mathrm{B_L}(t)\rangle \\ |\mathrm{B_H}(t)\rangle \end{pmatrix} = \begin{pmatrix} M_L - \frac{i}{2}\Gamma_L & 0 \\ 0 & M_H - \frac{i}{2}\Gamma_H \end{pmatrix}\begin{pmatrix} |\mathrm{B_L}(t)\rangle \\ |\mathrm{B_H}(t)\rangle \end{pmatrix} \tag{1.46}$$

where the 2×2 matrix can be obtained by diagonalizing the $\boldsymbol{H}$ matrix from Eq. 1.44 in the new $(|\mathrm{B_L}\rangle, |\mathrm{B_H}\rangle)$ base. This is achieved by:

$$\boldsymbol{Q}^{-1}\boldsymbol{H}\boldsymbol{Q} = \begin{pmatrix} M_L - \frac{i}{2}\Gamma_L & 0 \\ 0 & M_H - \frac{i}{2}\Gamma_H \end{pmatrix} \tag{1.47}$$

with

$$\boldsymbol{Q} = \begin{pmatrix} p & q \\ p & -q \end{pmatrix} \tag{1.48}$$

It is useful to describe the average mass and width (m,Γ) and the mass and width differences (Δm, $\Delta\Gamma$) of the two mass eigenstates, which are defined as follows:

$$m = \frac{m_L + m_H}{2} = M_{11} = M_{22} \tag{1.49}$$

$$\Gamma = \frac{\Gamma_L + \Gamma_H}{2} = \Gamma_{11} = \Gamma_{22} \tag{1.50}$$

$$\Delta m = m_H - m_L \tag{1.51}$$

$$\Delta\Gamma = \Gamma_L - \Gamma_H \tag{1.52}$$

where Δm is positive by definition, whereas the sign of $\Delta\Gamma$ is to be determined experimentally. The sign of $\Delta\Gamma_s$ has been recently determined to be positive by the LHCb Collaboration [21]. The mass and width differences between B_L and B_H are related to the off-diagonal terms M_{12}, Γ_{12} and to the relative phase between the two: $\phi = \arg(-M_{12}/\Gamma_{12})$. It is possible to express the latter as the combination of the two complex phases of the M_{12} and Γ_{12} terms as:

$$\phi = \phi_M - \phi_\Gamma = \arg(M_{12}) - \arg(-\Gamma_{12}) \tag{1.53}$$

A simplified solution for Δm and $\Delta\Gamma$ can be derived when the $\Delta m \gg \Delta\Gamma$ and $M_{12} \gg \Gamma_{12}$ conditions hold true. In this case:

$$\Delta m \simeq 2\,|M_{12}| \tag{1.54}$$

$$\Delta\Gamma \simeq 2\,|\Gamma_{12}|\cos\phi \tag{1.55}$$

where the solutions are approximated at the order $|\Gamma_{12}/M_{12}|^2$, being $|\Gamma_{12}/M_{12}| \sim 10^{-3}$. Solving the eigenvalue equation 1.44 in terms of the Δm and $\Delta\Gamma$ parameters thus yields the two conditions:

$$(\Delta m)^2 - \frac{1}{4}(\Delta\Gamma)^2 = 4\left(|M_{12}|^2 - \frac{1}{4}\,|\Gamma_{12}|^2\right) \tag{1.56}$$

$$\Delta m \Delta\Gamma = 4\Re\left(M_{12}\Gamma_{12}^*\right) \tag{1.57}$$

The q/p ratio can now be extracted from Eq. 1.47 and 1.48 to be:

$$\frac{q}{p} = -\frac{\Delta m - \frac{i}{2}\Delta\Gamma}{2\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)} = -\frac{2\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}{\Delta m - \frac{i}{2}\Delta\Gamma} \tag{1.58}$$

An approximate expression for q/p can be found using once more the term $|\Gamma_{12}/M_{12}|$ as an expansion parameter

$$\frac{q}{p} \simeq -\frac{M_{12}^*}{M_{12}}\left[1 - \frac{1}{2}\Im\left(\frac{\Gamma_{12}}{M_{12}}\right)\right] \tag{1.59}$$

That is, the phase of q/p is essentially given by the phase induced through M_{12}, ϕ_M. In the SM $B_q - \overline{B_q}$ mixing is dominated by the box diagram of Figure 1.2 with internal top quark exchange. The leading contribution to q/p is hence due to the ratio $-(V_{tb}^* V_{tq})/(V_{tb}V_{tq}^*)$, with the usual definition $q = d, s$.

The time evolution of the flavour eigenstates can be expressed as

$$i\frac{d}{dt}\begin{pmatrix} |B_q(t)\rangle \\ |\overline{B_q}(t)\rangle \end{pmatrix} = \begin{pmatrix} g_+(t) & \frac{q}{p}g_-(t) \\ \frac{p}{q}g_-(t) & g_+(t) \end{pmatrix}\begin{pmatrix} |B_q(0)\rangle \\ |\overline{B_q}(0)\rangle \end{pmatrix} \tag{1.60}$$

where the time dependent functions $g_+(t)$ and $g_-(t)$ can be expressed in terms of the light and heavy time-dependent mass eigenvalues:

$$\begin{aligned} g_+(t) &= \frac{1}{2}\left(e^{-im_H t - \frac{\Gamma_H}{2}t} + e^{-im_L t - \frac{\Gamma_L}{2}t}\right) = \\ &= e^{-imt}e^{-\frac{\Gamma}{2}t}\left[+\cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta M t}{2} - i\sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta M t}{2}\right] \end{aligned} \tag{1.61}$$

$$\begin{aligned} g_-(t) &= \frac{1}{2}\left(e^{-im_H t - \frac{\Gamma_H}{2}t} - e^{-im_L t - \frac{\Gamma_L}{2}t}\right) = \\ &= e^{-imt}e^{-\frac{\Gamma}{2}t}\left[-\sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta M t}{2} + i\cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta M t}{2}\right] \end{aligned} \tag{1.62}$$

1.2.3 Decay of Neutral B Mesons

Four decay amplitudes can be defined for a $\mathrm{B_q}$ or $\overline{\mathrm{B_q}}$ meson that decays directly into a generic multi-particle final state f or into the charge-conjugated state $\bar{f}$:

$$\begin{aligned} A_f &\equiv \langle f|\, H \,|\mathrm{B_q}\rangle & \qquad A_{\bar{f}} &\equiv \langle \bar{f}|\, H \,|\mathrm{B_q}\rangle \\ \bar{A}_f &\equiv \langle f|\, H \,|\overline{\mathrm{B_q}}\rangle & \qquad \bar{A}_{\bar{f}} &\equiv \langle \bar{f}|\, H \,|\overline{\mathrm{B_q}}\rangle \end{aligned} \tag{1.63}$$

where H is the Hamiltonian governing weak interactions. All states may be redefined by an arbitrary phase transformation like $|f\rangle \to e^{i\gamma_f}\,|f\rangle$. Such transformations change the mixing parameters and the transition amplitudes. However, this has no physical effects due to the SM invariance under global phase transformations. It is obviously possible to identify a trivial set of rephasing-invariant quantities: the magnitudes of the transition amplitudes and the magnitude of q/p parameter

$$\left|\frac{q}{p}\right|, |A_f|, |\bar{A}_f| \tag{1.64}$$

Besides these magnitudes, other quantities which are invariant under the arbitrary phase redefinitions arise from the interference between the parameters describing the mixing and the parameters describing the transitions, and are:

$$\lambda_f \equiv \frac{q}{p}\frac{\bar{A}_f}{A_f}, \quad \lambda_{\bar{f}} \equiv \frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_{\bar{f}}} \tag{1.65}$$

1.2.4 CP violation in B decays

Three types of CP violation (CPV) processes can be classified in a model-independent way:

- CP violation in decay (often referred to as direct CPV), which occurs in both charged and neutral decays, when the amplitude for a decay and its CP conjugated process have different magnitudes
- CP violation (purely) in mixing, which occurs when the two neutral mass eigenstates are not CP eigenstates
- CP violation in the interference between decays of mixed and unmixed mesons, which occurs in decays into final states which are common to $\mathrm{B_q}$ and $\overline{\mathrm{B_q}}$

CP violation is included in the SM via complex coupling constants, but can only occur when (at least) two amplitudes contribute to the same final state. A relative phase difference must also exist between the two decay amplitudes. This can originate from the so-called strong phase δ, which does not change sign under CP conjugation, or from the weak phase ϕ, arising from the electroweak sector of the SM, which does change sign under CP conjugation.

1.2.4.1 CPV in Decay

Direct CP violation takes place when the decay rate of a B hadron to a final state f is different from the rate of the CP-conjugated process $\overline{\mathrm{B}} \to \bar{f}$. This CPV process is not limited to neutral B mesons, instead it can also occurs for all B hadrons, including charged meson and baryons. This implies that the amplitudes of the two processes $|A_f|$ and $|\bar{A}_{\bar{f}}|$ differ:

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| \neq 1 \tag{1.66}$$

It is possible to split the contribution to the decay amplitudes A in three parts: its magnitude $|A|$, its weak-phase term $e^{i\phi}$, and its strong phase term $e^{i\delta}$. Then, if several amplitudes contribute to the same $\mathrm{B} \to f$ ($\overline{\mathrm{B}} \to \bar{f}$) process, the amplitude A_f and the CP conjugate amplitude $\bar{A}_{\bar{f}}$ can be written, neglecting an arbitrary phase, as

$$A_f = \sum_i A_i e^{i(\phi_i + \delta_i)}, \quad \bar{A}_{\bar{f}} = \sum_i e^{2i\zeta_f} A_i e^{i(\phi_i - \delta_i)} \tag{1.67}$$

where if f is a CP-eigenstate then $e^{2i\zeta_f}$ represents its CP-eigenvalue ($e^{2i\zeta_f} = \pm 1$). CP violation in decay occurs if the decay amplitude A_f contains two interfering amplitudes with different phases. The largest direct CP violation occurs when the two amplitudes are of equal magnitude and have a weak and strong phase difference of $\pi/2$.

$$|A|^2 - |\bar{A}|^2 = -2 \sum_{i,j} A_i A_j \sin\left(\phi_i - \phi_j\right) \sin\left(\delta_i - \delta_j\right) \tag{1.68}$$

1.2.4.2 CPV in Mixing

CP violation in mixing implies that the probability of a neutral B meson to oscillate to the corresponding $\overline{\mathrm{B}}$ meson is different from the probability of a $\overline{\mathrm{B}}$ meson to oscillate to a B meson. This condition can be described by the inequality:

$$\left|\left\langle \mathrm{B_q}(0) \mid \overline{\mathrm{B_q}}(t) \right\rangle\right|^2 \neq \left|\left\langle \overline{\mathrm{B_q}}(0) \mid \mathrm{B_q}(t) \right\rangle\right|^2 \tag{1.69}$$

which is related to the q/p term defined in Eq. 1.59 through the time-evolution relation described in Eq. 1.60. The previous Eq. 1.69 can therefore be rewritten in the following form

$$\left|\frac{p}{q}\right|^2 |g_-(t)|^2 \neq \left|\frac{q}{p}\right|^2 |g_-(t)|^2 \tag{1.70}$$

from which it follows that CP is thus violated in mixing if

$$\left|\frac{q}{p}\right| \neq 1 \tag{1.71}$$

For CP to be conserved the mass eigenstates must therefore correspond to the CP eigenstates, thus making the relative phase between the terms M_{12} and Γ_{12} vanish.

1.2.4.3 CPV in Interference

CP violation can also occur as the result of the interference between the decay of mixed and unmixed neutral B mesons in the same final state, which is accessible to both B and $\overline{\mathrm{B}}$. An important feature of this CPV mechanism is that there can still be CP violation even thought it is neither observed in mixing nor in decay. The phase convention independent quantity of interest in this case is λ_f, already defined in Eq. 1.65. CP violation takes place when λ_f either carries a non-zero phase or has a non-unitary modulus. In the non-trivial case where CP is conserved both in decay and in mixing ($|q/p| = 1$ and $|\bar{A}_{\bar{f}}/A_f| = 1$), the relative phase between q/p and $\bar{A}_{\bar{f}}/A_f$ vanishes, but it is still possible to have CP violation if the following condition holds true:

$$\Im\left(\lambda_f\right) \neq 1, \quad \left|\lambda_f\right| = 1 \tag{1.72}$$

1.2.5 CP violating phase ϕ_s in $B^0_s \to J/\psi\,\phi$ decay

The $B^0_s \to J/\psi\,\phi$ decay is characterized by a vector-vector final state accessible to both the B^0_s and $\overline{B^0_s}$ pseudoscalar mesons. The CP violation can therefore occur in the $B^0_s \to J/\psi\,\phi$ system as the result of the interference between the direct and mixing-mediated contribution. Let's consider the quark transitions involved in the B^0_s decay and common for the two meson flavours B^0_s and $\overline{B^0_s}$. It is possible to expand the total amplitude in terms of all the quark amplitudes contributions a and the CKM matrix elements involved:

$$A\left(b \to c\bar{c}s\right) = V_{cb}V^*_{cs}a^T_{c\bar{c}s} + V_{ub}V^*_{us}a^T_{u\bar{u}s} + V_{ub}V^*_{us}a^P_{\bar{u}us} + V_{cb}V^*_{cs}a^P_{\bar{c}cs} + V_{tb}V^*_{ts}a^P_{\bar{t}ts} \tag{1.73}$$

where the apex T stands for tree-level amplitudes whilst P indicates loop-mediated transitions often referred to as "penguin" contributions. The two processes are shown in Figure 1.3.

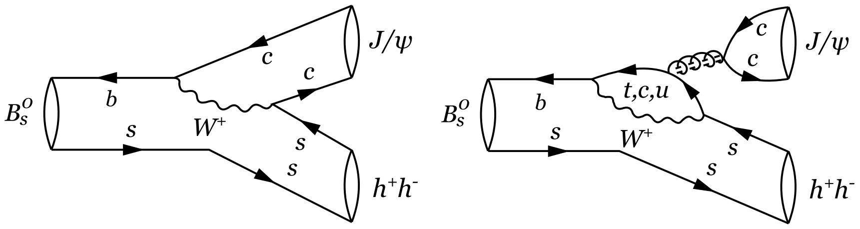


Figure 1.3: Feynman diagrams of the tree-level $B^0_s \to J/\psi\,\phi$ transitions (left) and of the loop-mediated, or penguin, transitions (right).

The penguin processes are mediated by the exchange of a up-type quark, and are therefore dominated (as in the case of the mixing diagrams) by the top contribution within the loop, with $V_{tb}V^*_{ts} \sim \mathcal{O}\left(\lambda^2\right)$. This process is however loop-suppressed with respect to the tree-level transitions [22]. The tree-level $b \to u\bar{u}s$ transition is doubly Cabibbo-suppressed, $V_{ub}V^*_{us} \sim \mathcal{O}\left(\lambda^4\right)$. The total decay amplitude is therefore dominated by the Cabibbo-suppressed $b \to c\bar{c}s$ tree-level process, for which $V_{cb}V^*_{cs} \sim \mathcal{O}\left(\lambda^2\right)$. Ignoring the penguin and λ^4 suppressed terms, the decay amplitudes can be described by the following:

$$A_{c\bar{c}s} = V^*_{cb}V_{cs}, \quad \bar{A}_{c\bar{c}s} = V_{cb}V^*_{cs} \tag{1.74}$$

Under the same assumption, since only one amplitude contribute to the decay, there is no direct CP violation. Thus, we set $\left|\bar{A}_{c\bar{c}s}\right| = \left|A_{c\bar{c}s}\right|$ and we have that the amplitude ratio

$$\frac{\bar{A}_{c\bar{c}s}}{A_{c\bar{c}s}} = \frac{V_{cb}V^*_{cs}}{V^*_{cb}V_{cs}} \tag{1.75}$$

is thus only given by the relative phase between the two amplitudes. Setting the direct weak phase $\phi_D = \arg(V^*_{\mathrm{cb}}V_{\mathrm{cs}})$, the total amplitude ratio is modulated by the argument

$$\arg\left(\frac{\bar{A}_{c\bar{c}s}}{A_{c\bar{c}s}}\right) = 2\phi_D \tag{1.76}$$

As previously discussed, the final J/$\psi\phi$ state can be reached by mixing-mediated processes, dominated by the $V_{\mathrm{tb}}V^*_{\mathrm{ts}}$ top-quark inner loop term. The q/p ratio for the B^0_s decay can thus be described as follows:

$$\frac{q}{p} = \frac{V_{\mathrm{tb}}V^*_{\mathrm{ts}}}{V^*_{\mathrm{tb}}V_{\mathrm{ts}}} \tag{1.77}$$

where as derived in Equation 1.59, the phase of q/p is essentially given by the phase induced through the off-diagonal mass term M_{12}, ϕ_M, hence giving

$$\arg\left(\frac{q}{p}\right) = \phi_M \tag{1.78}$$

Neglecting CP violation in mixing, thus assuming $|q/p| = 1$ as from [23], a source of CP violation through interference between decay and mixing could arise if the conditions of Equation 1.72 hold true. Combining the two phase-convention dependent expressions obtained for $\bar{A}_{c\bar{c}s}/A_{c\bar{c}s}$ and q/p, it is now possible to access the phase-convention independent parameter

$$\lambda_{c\bar{c}s} = \frac{q}{p}\frac{\bar{A}_{c\bar{c}s}}{A_{c\bar{c}s}} = \tag{1.79}$$

$$= \left|\frac{q}{p}\frac{\bar{A}_{c\bar{c}s}}{A_{c\bar{c}s}}\right| e^{+i\phi_M}e^{-i2\phi_D} \tag{1.80}$$

The weak phase ϕ_s is thus given by

$$\phi_\mathrm{s} = \arg\left(\lambda_{c\bar{c}s}\right) = \phi_M - 2\phi_D \tag{1.81}$$

It is moreover possible to relate the phase ϕ_s to the B^0_s unitarity triangle

$$\phi_\mathrm{s} = \arg\left(\frac{V_{\mathrm{tb}}V^*_{\mathrm{ts}}}{V^*_{\mathrm{tb}}V_{\mathrm{ts}}}\frac{V_{\mathrm{cb}}V^*_{\mathrm{cs}}}{V^*_{\mathrm{cb}}V_{\mathrm{cs}}}\right) = -2\beta_\mathrm{s} \tag{1.82}$$

In the SM the values of ϕ_s and β_s are expected to be very small. From a global fit to the experimentally determined CKM matrix elements, the SM value of β_s is evaluated to be $\sin(2\beta^{\mathrm{SM}}_\mathrm{s}) = 0.0363^{+0.0014}_{-0.0012}$ rad [16], resulting in the precise SM prediction $\phi^{\mathrm{SM}}_\mathrm{s} = -0.03634^{+0.00136}_{-0.00120}$ rad [18]. For this reason these quantities are of great experimental interest. A value significantly different from the SM expectations would in fact represent a smoking-gun signal of NP (NP). If NP, beyond the description of the SM, is present in the B^0_s mixing sector, it will contribute to both ϕ_s and β_s, which could be in fact rewritten, given the usual phase conventions, as combinations of SM and NP contributions:

$$\phi_\mathrm{s} = \phi^{\mathrm{SM}}_\mathrm{s} + \phi^{\mathrm{NP}}_\mathrm{s} \tag{1.83}$$

$$2\beta_\mathrm{s} = 2\beta^{\mathrm{SM}}_\mathrm{s} - \phi^{\mathrm{NP}}_\mathrm{s} \tag{1.84}$$

The precise measurement of the weak phase ϕ_s therefore represents a way to infer the presence of NP indirectly by detecting deviations from expectations precisely calculated in the SM, complementary to the direct searches carried out at the high energies available at the LHC.

Chapter 2

The CMS Experiment at LHC

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a two-ring superconducting proton-proton accelerator and collider installed in the existing 26.7 km tunnel that was constructed for the CERN LEP $e^+ \; e^-$ [24] machine. The LHC accelerator is designed to provide a centre of mass energy of 14 TeV and to reach the instantaneous luminosity of 10^{34} cm^{-2}s^{-1} [25]. A number of 1 232 superconducting niobium-titanium dipoles operate at 1.9 K generating a magnetic field up to 8.4 T to guide the particle beams into curvilinear trajectories together with 386 quadrupoles, 360 sextupoles and 336 octupoles used for beam focusing and control. The accelerating field is provided by 8 radio-frequency cavities per beam operating at a temperature of 4.5 K, which are able to generate a 5 MV/m field at the frequency of 400 MHz. Protons are accelerated up to 450 GeV prior to the injection into the LHC ring by the CERN injection chain, as shown in Figure 2.1. At its design luminosity, the LHC can be filled with 2 808 proton bunches, with a time bunch spacing of 25 ns.

The accelerator produced the first proton-proton collisions in November 2009. Starting from 2010 the LHC operated at the energy of 3.5 TeV per beam with the filling of 1 320 bunches at 50 ns of spacing (75 ns during 2010), reaching a maximum instantaneous luminosity of 4.0×10^{33} cm^{2}s^1 and delivering an integrated luminosity of about 6 fb^{-1} by the end of 2011. In 2012 the centre of mass energy was increased to 8 TeV and was filled with 1 380 bunches at 50 ns, with a maximum instantaneous luminosity of 7.7×10^{33} cm^{-2}s^{-1}, for a delivered integrated luminosity of about 23 fb^{-1} by the end of the first data taking period (Run I) of LHC, as shown in Figure 2.2. An important feature of the LHC collisions is the number of proton-proton collisions per bunch crossing, the so-called pileup, which is one of the main experimental challenges for the detectors. Given the high delivered instantaneous luminosity and the total cross section of the inelastic pp scattering of about 70 mb, an average number of 21 collisions per bunch crossing have been produced in 2012, as shown in Figure 2.3.

After the first scheduled long shut-down, the LHC will restart operations in 2015 (Run II) with the centre-of-mass energy increased up to 13 TeV and 25 ns bunch spacing, for a foreseen instantaneous luminosity of 10^{34} cm^{-2}s^{-1}. By the end of Run II a total integrated luminosity of about $100 - 300$ fb^{-1} is expected to be delivered by the LHC to the experiments.

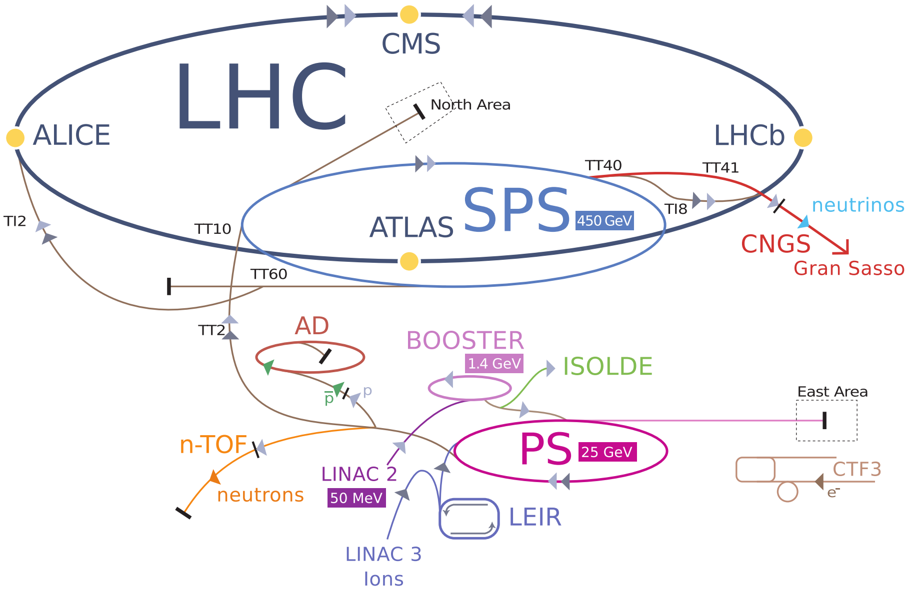


Figure 2.1: The CERN accelerator complex. The proton beam is accelerated by the chain LINAC2(50 MeV)$\mapsto$ BOOSTER(1.4 GeV)$\mapsto$ PS(25 GeV)$\mapsto$ SPS(450 GeV) prior to the injection in the LHC.

Four main experiments are installed at LHC. ATLAS (A Toroidal LHC ApparatuS) [26] and CMS (Compact Muon Solenoid) [27] are general-purpose detectors designed to investigate a wide range of physics, including the direct search for the scalar boson of the SM and particles that could arise from physics beyond the SM. ALICE (A Large Ion Collider Experiment) [28] is a heavy-ion detector, designed to study the physics of strongly interacting matter at extreme energy densities, where a phase of matter called quark-gluon plasma forms in heavy ions collisions. The Large Hadron Collider beauty (LHCb) [29] experiment is an asymmetric detector especially designed for the investigating of the properties of b-hadrons.

2.2 The CMS Experiment

The Compact Muon Solenoid experiment is one of the two general-purpose experiments at the LHC. The experiment is designed to investigate a wide physics program, including the electroweak symmetry breaking sector of the SM and searches for NP processes at the new energy and luminosity frontiers that the LHC is opening. The main features of the CMS detector are the superconducting solenoid, which allows a compact design with a strong magnetic field, a high-quality tracking system, an high resolution and high granularity electromagnetic calorimeter, an hermetic hadronic calorimeter and a redundant muon system. The CMS apparatus features a cylindrical symmetry around the beam axis, with an overall length of roughly 22 m for a 14 m diameter.

The right-handed coordinate system adopted by CMS has the origin in the nominal interaction point. The x-axis points radially towards the centre of the LHC and the z-axis runs tangent to the beam axis in the direction of the Jura mountains. This defines the direction of the y-axis to point vertically upward. The azimuthal angle $\phi \in [-\pi, \pi]$ is measured in the x-y plane (transverse plane) starting from the x-axis. The polar angle θ is

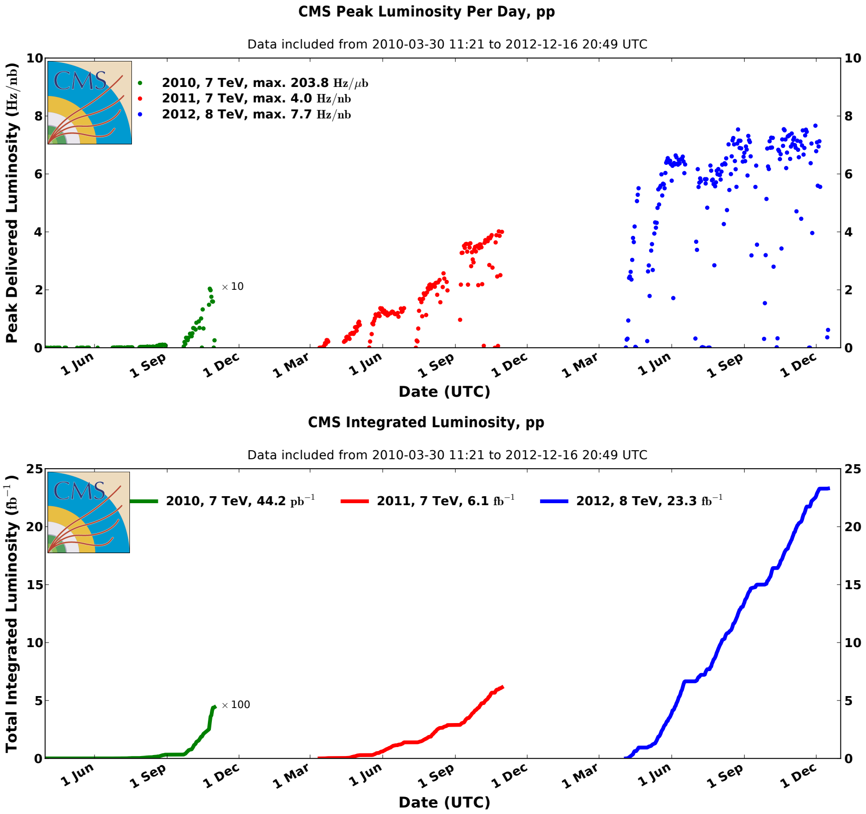


Figure 2.2: Instantaneous (top) and cumulative (bottom) luminosity delivered to the CMS experiment for $p-p$ collisions during the 2010 (green), 2011 (red) and 2012 (blue) data-taking periods.

measured with respect to the z-axis. The projection in the r-z plane, where $r = \sqrt{x^2+y^2}$, defines a longitudinal plane. The pseudorapidity η is defined as $\eta \equiv -\ln\left(\tan\left(\theta/2\right)\right)$. It is also useful to introduce the quantity usually employed to express the angular distance between two particles (1,2), the $\Delta\mathrm{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, where $\Delta\eta = \eta_1 - \eta_2$, $\Delta\phi = \phi_1 - \phi_2$.

The CMS detector, illustrated in Figure 2.4, is longitudinally segmented into three regions, one barrel and two endcaps. The barrel defines the central region of the experiment, approximately covering the $|\eta| \lesssim 1.5$ range, while the endcaps cover the so-called forward regions, with approximately $1.5 \lesssim |\eta| \lesssim 3$. Subdetectors are layered at increasing values of the radius r in the barrel, whereas they are arranged along the z coordinate in the endcaps. Both the barrel and the endcap are equipped with vertexing and tracking detectors, electromagnetic and hadronic calorimeters, and muon detectors. A preshower sampling calorimeter is hosted in front of the ECAL endcap. Additional coverage in the outermost forward regions, $3.0 < |\eta| < 5.0$, is provided by two so-called forward calorimeters. The detectors accommodated in the endcaps are subject to a higher flux of particles than in the barrel, thus requiring more radiation-hardness than for the former.

The main distinguishing feature of the CMS experiment is the 13 m long superconducting solenoid, which operates at a temperature of about 4 K providing an uniform magnetic field of 3.8 T. This corresponds to an amount of energy stored in the magnet of about 2.5 GJ. The magnetic field in the central region of the detector has a direction parallel to

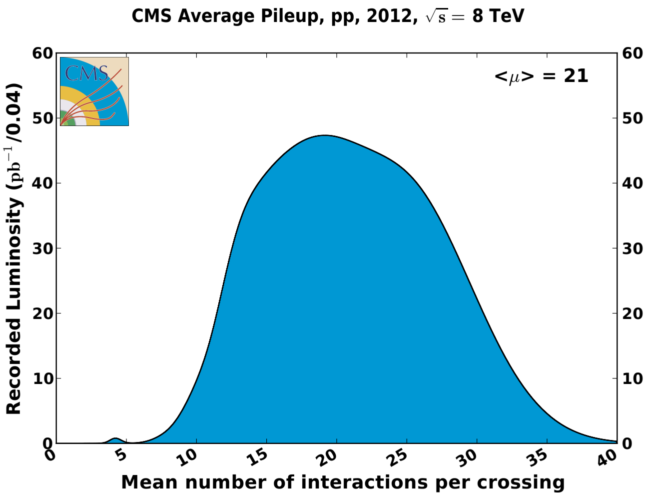


Figure 2.3: Distribution of the recorded number of collisions per bunch crossing (pile-up) measured during the 2012 pp data taking.

the beams, while the return field is guided by an external iron yoke, and is intense enough to saturate, to a value of about 2 T, the 1.5 m thick iron slabs. The magnet contains, from inside out, the tracker and the electromagnetic and hadronic calorimeters. Outside the magnet coil, the iron return joke of the magnet hosts the muon spectrometer, used for reconstruction of muon tracks.

In the following sections a brief description of the main features of the CMS detector is provided; a detailed description of the CMS detector can be found in Ref. [27].

2.2.1 Tracking System

The core of CMS is a Silicon Tracking System of 2.5 m diameter and 5.8 m length. It is designed to provide a precise and efficient measurement of the trajectories of charged particles emerging from LHC collisions and to allow the precise reconstruction of secondary vertices. The tracker is designed to cope with the high track multiplicity of about 1 MHz/mm^2 that are expected at 4 cm from the beamline at the design LHC luminosity and energy. The CMS Tracking System is composed of silicon pixel in the innermost region and silicon microstrips detectors in the outer region, as shown in Fig. 2.5. The full tracker coverage extends up to $|\eta| \approx 2.5$ with a surface of active silicon of about 210 m^2.

2.2.1.1 Pixel Detector

The Pixel Detector consists of matrices of for 1 440 modules of 100×150 μm^2 silicon pixels arranged in three barrel layers and two endcap disks for each side. It provides 2-dimensional points with spatial resolution of about 10 μm in the short pixel edge (r-ϕ in the barrel) and 20-45 μm in the long edge (z in the barrel) depending on the incidence angle. With a high hit efficiency well above 99%, the pixel detector typically measures three hits for particles within $|\eta| < 2.2$ and two hits for $2.2 < |\eta| < 2.5$. The full layout of the pixel detector consists of a barrel region (BPIX), with three barrels at radii of 4.4, 7.3 and 10.2 cm, complemented by two disks on each side (FPIX), at 34.5 and 46.5 cm from the nominal interaction point. This layout provides about 66 million pixels covering a

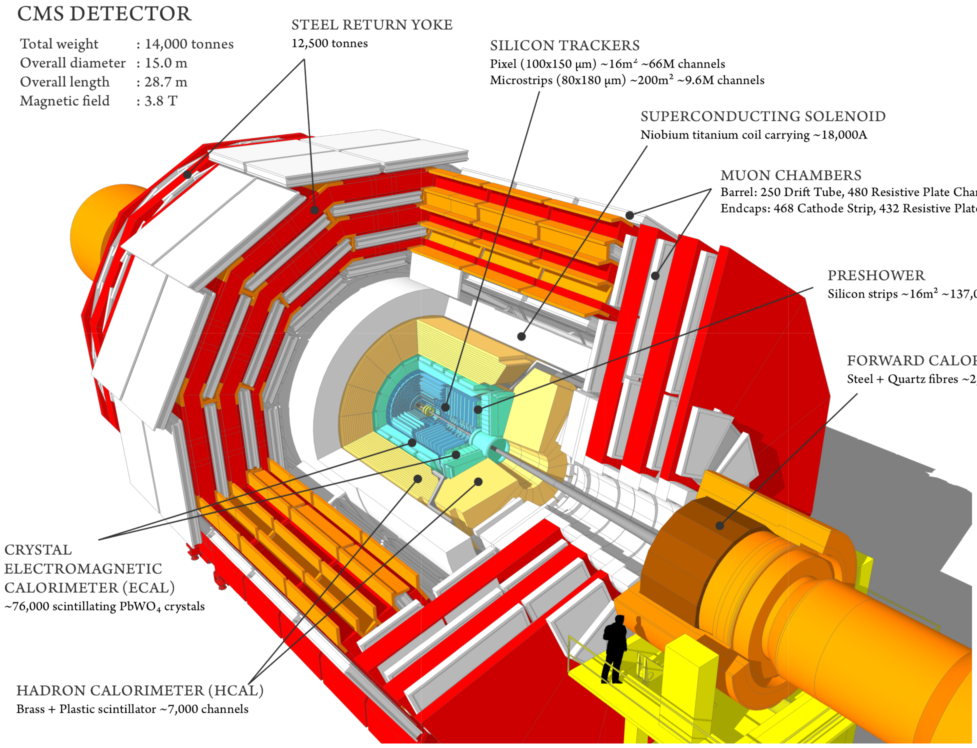


Figure 2.4: Schematic view of the CMS detector.

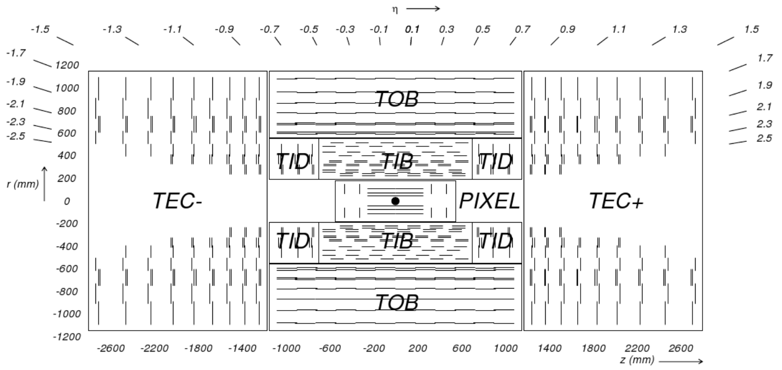


Figure 2.5: View of the CMS tracker in the rz-plane. Each line in the strip tracker represents a silicon strip detector, whereas lines in the pixel detector represent the structures on which the detectors are mounted in the barrel and endcaps, respectively.

total area of about 1 m^2. Detectors in FPIX disks are tilted by 20 degrees in a turbine-like geometry to induce charge sharing and achieve a spatial resolution of about 20 μm.

2.2.1.2 Strip Detector

To reduce the number of readout channels and hence the amount of material, power consumption and cost, silicon strip detectors are employed at larger radii of 20 to 116 cm, where the reduced track density permits a larger size of the active silicon cells. The strip tracker consists of 15 148 strip modules with about 10 million readout channels and is further subdivided into different parts. The Tracker Inner Barrel and Discs (TIB/TID) are composed of four barrel layers with strips parallel to the beam line and three endcap discs at each side with radial strips. The strips have a pitch of 80 to 141 μm, a length of 10 cm and a sensor thickness of 320 μm resulting in an occupancy of 2 to 3% and an r-ϕ resolution of 16 to 27 μm in the barrel. The Tracker Outer Barrel (TOB) with six layers and the Tracker Endcaps (TEC) with nine discs at each side use strips with 97 to 184 μm pitch, 25 cm length and 320 to 500 μm sensor thickness. This leads to an occupancy at the percent level and an r-ϕ resolution of 25 to 41 μm in the barrel. Some of the inner layers of all strip tracker subdivisions have an additional module mounted back-to-back on the first one under a stereo angle of 100 mrad, thereby providing also a position measurement in the direction along the strips (z in the barrel, r in the endcaps). The resulting z resolution is 230 μm in TIB and 530 μm in TOB. The design and performance of the tracker provide a p_{T} resolution at the percent level for tracks up to 100 GeV, a transverse impact parameter resolution of better than 30 μm for central tracks with $p_{\mathrm{T}} > 5$ GeV, and a primary vertex resolution of better than 30 μm for vertices with at least 30 tracks [30].

2.2.1.3 Track Reconstruction

Track reconstruction at CMS [31] is performed in the challenging environment of large track and hit densities and of a large amount of detector material leading to substantial multiple scattering, energy loss and interactions. The standard track-reconstruction algorithm capable of coping with these conditions is the combinatorial track finder (CFT) based on a Kalman Filter (KF) [32]. The first step consists of finding track seed in the inner tracker layers, which are identified either as hit triplets or as doublets with additional beam-spot constraint. Track candidates are best seeded from hits in the pixel detector because of the low occupancy, high efficiency and unambiguous two-dimensional position information. From this, a starting trajectory is determined and extrapolated to the next layer of the detector, which is scanned for compatible hits around the predicted point. In case a hit is found, it is added to the trajectory, which is updated by using the KF fit. Compatible hits are assigned to the track on the basis of the χ^2 between the predicted and measured positions. This procedure is repeated as long as compatible hits are found, maximally up to the last layer. At each stage the Kalman filter updates the track parameters with the new hits. In order to take into account possible inefficiencies, one further candidate is created without including any hit information. The tracks are assigned a quality based on the χ^2 and the number of missing hits and only the best quality tracks are kept for further propagation. Ambiguities between tracks are resolved during and after the track finding procedure. In case two tracks share more than 50% of their hits, the lower quality track is discarded. For each trajectory the finding stage results in an estimate of the track parameters. However, since the full information is only available at the last hit and the constraints applied during trajectory building can bias the estimate of the track parameters, all valid tracks are refitted with a standard Kalman filter and a second filter (smoother) is applied, running from the outermost hits towards the beam line.

2.2.1.4 Vertex reconstruction

The reconstruction of the primary vertices produced by the pp collisions is a two-step procedure. Reconstructed tracks are first clustered into sets that appear to come from the same interaction, their z coordinates at the beam closest approach point are evaluated, retaining only tracks with impact parameter with respect to the nominal beam-spot less than 3 cm. The main challenge for the clustering step in the presence of high pile-up is to avoid merging tracks from separate collisions into a single cluster, while preserving a good efficiency of assigning tracks to the proper clusters and a low rate of fake clusters not corresponding to a pp interaction. The algorithm used in CMS to perform clustering is the deterministic annealing [33]. With this algorithm it is possible to resolve vertices with separations down to about 1 mm, appropriate for the high multiplicity of interactions per bunch crossing experienced in CMS, as the longitudinal RMS spread of the luminous region is about 6 cm. The constituents of each cluster are then fitted with adaptive vertex fitter algorithm [34] in order to obtain a precise measurement of the vertex position. This algorithm addresses the issue of secondary and fake tracks in the cluster by iteratively downweighting the tracks which are not compatible with the common vertex being fitted.

2.2.2 Muon Spectrometer

Muons are typically considered as minimum-ionising particles (except at very high energies), and thus can be efficiently identified by dedicated muon detectors placed outside the calorimeters, which also provide an additional momentum measurement. The CMS muon spectrometer is located in the steel return yoke of the solenoid, covering the pseudorapidity region $|\eta| < 2.4$. The muon system is designed for three major functions: robust and fast identification of muons, good resolution of momentum measurement, and fast and reliable triggering. The muon spectrometer is composed of three types of gaseous detectors located inside the empty volumes of the iron yoke, and arranged in barrel and end-cap sections, as shown in Figure 2.6. The reason for using different detector technologies primarily lies in the different particle rates and occupancies, both higher in the endcaps, and in the intensity of the stray magnetic field, which is lower in the barrel.

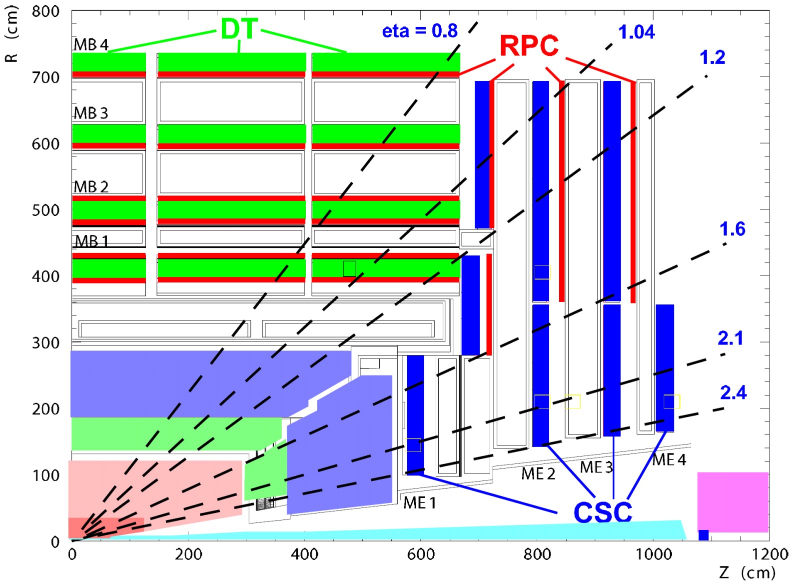


Figure 2.6: A longitudinal view of the muon system indicating the location of the three detector types contributing to the muon spectrometer.

2.2.2.1 Drift tubes

In the barrel region the muon rate is low, the background (mainly neutron-induced) is small and the magnetic field is uniform and contained in the iron yoke. For these reasons standard drift chambers with rectangular cells are used. The barrel drift tubes (DT) cover the $|\eta| < 1.2$ region and are organized in four stations composed of two or three stacked superlayers housed among the yoke layers. A honeycomb spacer is inserted between the superlayers measuring the ϕ coordinate to increase the lever-arm between the two. Cells in consecutive layers are shifted by half of their width in order to avoid un-instrumented regions and to eliminate any left-right signal ambiguity within a single cell. The muon position in each DT is reconstructed by measuring the drift time of the ionization electrons and by converting it into a distance from the wire. In all stations the tubes in the innermost and outermost superlayers have wires parallel to the z axis to provide accurate measurements of the ϕ coordinate. The intermediate superlayer, present in all but the outermost station, has wires along the ϕ coordinate and provides a measurement of the z coordinate. Each station can measure the muon position with 100 μm resolution in the r-ϕ direction and up to 150 μm in z, and its direction with $\sim$ 1 mrad accuracy [35, 36].

2.2.2.2 Cathode strip chambers

In the forward region the rate of muons and the amount of background is higher than what experienced in the barrel. The intensity of the magnetic field is also large and non uniform. Cathode strip chambers (CSC) are thus used in this region because of their fast response time, fine segmentation and radiation tolerance. Each end-cap is equipped with four stations of CSC, covering the $0.9 < |\eta| < 2.4$ pseudorapidity range, defining an overlap with the DT in the region $0.9 < |\eta| < 1.2$. The cathode strips are oriented radially and provide precise measurement in the bending plane; the anode wires run approximately perpendicular to the strips and are read out to measure the pseudorapidity and the beam-crossing time of a muon. Particle traverses the cathodes and produces an ionization trail through the chamber: the electrons are accelerated to the wire where an avalanche occurs, inducing a charge on the cathode. The coordinate orthogonal to the wire is measured by fitting the charge distribution on the cathode strips, while the longitudinal one is given by the coordinates of the wire that has been hit.

2.2.2.3 Resistive plate chambers

Resistive Plate Chambers (RPC) are used as trigger-dedicated muon detectors mainly added to help in measuring the correct beam-crossing time. These are gaseous detector operated in the avalanche mode, which can provide independent and fast trigger with high segmentation and sharp p_{T} threshold over a large portion of the pseudorapidity range. They provide good timing information, but a less precise position determination than the other systems. Thus, they are mainly used to improve the bunch-crossing assignment and for trigger purposes. Both the DT in the barrel and the CSC in the endcaps are complemented by RPC chambers up to $|\eta| < 1.6$.

2.2.2.4 Muon reconstruction

In the standard CMS reconstruction, tracks are first reconstructed independently in the inner tracker (tracker track) and in the muon system (standalone-muon track). The first step for muon identification at CMS is the reconstruction using the muon system alone. The algorithm used is hierarchical: segments are first reconstructed in the individual stations, and then combined to produce tracks. In the barrel the magnetic field is almost completely confined in the steel return joke of the magnet, thus the muon tracks are well approximated by straight lines within each DT station. Segments are first reconstructed independently in the transverse and longitudinal planes, through pattern matching and linear fitting of the hits; such 2-dimensional segments provide a measurement of one coordinate and one angle. The two views are then combined to produce the so-called 4-d segments, which carry information about two position coordinates and two angles. In the endcaps the CSC hits are reconstructed by combining the information from anode wires and cathode strips to obtain a 2-d point in the transverse plane. Segments are then formed from hits in all the six layers of each station. In order to reconstruct full muon tracks, the segments are extrapolated from one chamber to the next taking into account the magnetic field, the energy loss and the multiple scattering in the return yoke; if a compatible segment is found the two measurements are combined and the procedure is repeated for the next stations. The position of the beam-spot is combined with the angle at the entrance of the muon system allowing the p_{T} measurement exploiting the large bending power of the magnetic field inside the solenoid. Two muon reconstruction approaches are used:

- Global Muon reconstruction (outside-in approach). For each standalone-muon track, a matching tracker track is found by comparing parameters of the two tracks propagated onto a common surface. A global-muon track is fitted combining hits from the tracker track and standalone-muon track, using the Kalman-filter technique. At large transverse momenta ($p_{\mathrm{T}} \gtrsim 200$ GeV) the global-muon fit can improve the momentum resolution compared to the tracker-only fit.
- Tracker Muon reconstruction (inside-out approach). In this approach, all tracker tracks with $p_{\mathrm{T}} > 0.5$ GeV and total momentum $p > 2.5$ GeV are considered as possible muon candidates and are extrapolated to the muon system taking into account the magnetic field, the average expected energy losses, and multiple Coulomb scattering in the detector material. If at least one muon segment matches the extrapolated track, the corresponding tracker track qualifies as a Tracker Muon.

Tracker Muon reconstruction is more efficient than the Global Muon reconstruction at low momenta, $p \lesssim 5$ GeV, because it only requires a single muon segment in the muon system, whereas Global Muon reconstruction is designed to have high efficiency for muons penetrating through more than one muon station and typically requires segments in at least two muon stations. Muons reconstructed only as standalone-muon tracks have worse momentum resolution and higher admixture of cosmic-ray muons than the Global and Tracker Muons and are usually not exploited in physics analyses.

2.2.2.5 Muon identification

Physics analyses can set the desired balance between identification efficiency and purity by applying a selection based on various muon identification variables. Several standard algorithms for muon identification are available in CMS [37]:

- Soft Muon selection. This selection requires the candidate to be a Tracker

Muon, with the additional requirement that a muon segment is matched in both x and y coordinates with the extrapolated tracker track, such that the pull for local x and y is less than 3. Segments that form a better match with a different tracker track are not considered. These additional requirements are optimized for low $p_{\mathrm{T}} < 10$ GeV muons, and is often used in flavour physics analyses.

- Tight Muon selection. For this selection, the candidate must be reconstructed as a Global Muon with the χ^2/ndof of the global-muon track fit less than 10 and at least one muon chamber hit included in the global-muon track fit. In addition, its corresponding tracker track is required to be matched to muon segments in at least two muon stations, use more than 10 inner-tracker hits (including at least one pixel hit), and have a transverse impact parameter with respect to the primary vertex lower than 2 cm.
- Particle-Flow Muon selection. The details of the particle flow algorithm are described in Section 2.2.6. For muons, the particle-flow approach applies particular selection criteria to the muon candidates reconstructed with the Global and Tracker Muon algorithms described above. Depending on the environment of the muon (e.g. whether it is isolated or not) the selection criteria are adjusted making use of information from other subdetectors (like for instance the energy deposition in the calorimeters). In general, the selection is optimized in order to identify muons within jets with high efficiency, while maintaining a low rate for the misidentification of charged hadrons.
- Loose Muon selection. This selection requires the candidate to be identified by the Particle Flow event reconstruction and to be reconstructed as a Tracker or Global Muon, thus avoiding muons which are only reconstructed as Standalone Muons. The Loose identification criteria are often used alternatively to the Soft Muon ones, in the case specific muon quality cuts are going to be evaluated for the analysis.

Custom identification criteria can be studied and applied in the case of specific analyses, for which the muon identification is crucial, like in the case of the measurement of the $\mathrm{B}^0_{\mathrm{s}} \to \mu^+ \mu^-$ branching ratio recently performed by CMS [38–40].

2.2.3 Electromagnetic Calorimeter

The CMS electromagnetic calorimeter, ECAL, is a homogeneous calorimeter composed of nearly 76 000 lead tungstate ($PbWO_4$) crystals intended to measure the energy and position of (mainly) electrons and photons, from the analysis of the electromagnetic showers in the material. The ECAL surrounds the tracker and it is located within the CMS solenoid covering a region of $|\eta| < 3$, as can be seen in Figure 2.7. In the barrel region (EB) the calorimeter is located at a distance of 1.3 m from the beam pipe and $|\eta| < 1.479$. It is equipped with 61 200 crystals with a high granularity of 0.0174×0.0174 in η-ϕ corresponding to 22×22 mm^2 at the front face. These are slightly tilted by 3 degrees with respect to the direction towards the interaction point to avoid particles passing through the cracks of the crystals. In the endcaps (EE) are located 7 324 crystals per side covering the remaining $1.479 < |\eta| < 3$ region. The crystals have a front-face cross section of the is 28.6×28.6 mm^2 and are also slightly tilted. In the $1.653 < |\eta| < 2.6$ region a lead/silicon-strip preshower detector is also installed to improve the discrimination between a neutral pions decaying into two photons and single photons. Lead tungstate is characterized by high density, 8.28 g/cm^3, and short radiation length, 0.89 cm, allowing a very good longitudinal shower containment with a detector depth of about 26 radiation lengths in the barrel, and about $3 + 25$ radiation lengths in the preshower + endcaps. The small Molière radius, 2.2 cm, allows the calorimeter to be kept compact with fine granularity, while scintillation and optical properties of $PbWO_4$ make it fast ($\sim$ 80% of the light yield is delivered before a new collision occurs) and radiation tolerant. The readout of the ECAL crystals is performed by avalanche photodiodes (APD) in the barrel, and vacuum phototriodes (VPT) in the endcaps.

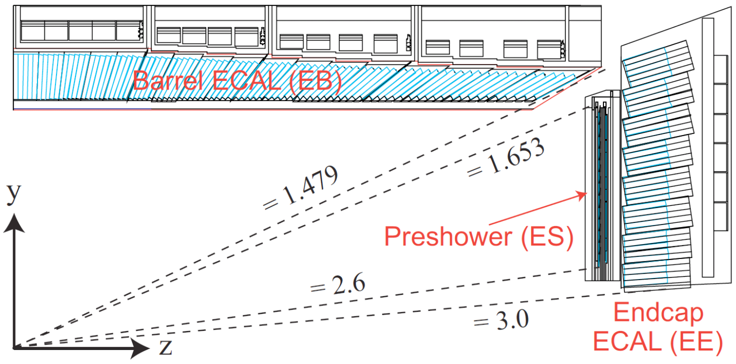


Figure 2.7: Layout of the CMS electromagnetic calorimeter presenting the arrangement of crystal modules for barrel, endcaps and the preshower.

2.2.3.1 Electrons and Photons Reconstruction

Electrons and photons are reconstructed in the CMS electromagnetic calorimeter by clustering channels into single showers, which are then merged into more extended superclusters. These are intended to collect all the informations associated to electrons or photons: the energy and all its associated radiation from bremsstrahlung and conversion tracks. The energy deposits are narrow along the η direction but tend to be large in the ϕ direction, especially in the barrel and at low momenta. This is due the strong magnetic field acting on the ECAL and the non-negligible amount of detector material upstream to the calorimeters. The reconstruction of electron tracks, both for primary electrons and for photon conversions, is performed by searching for track seeds compatible with the ECAL superclusters. Large search windows are used in the pattern recognition stage, since electrons can undergo substantial energy losses and large angle scatterings in the material. The final track fit is performed using the Gaussian sum filter (GSF) algorithm [41] which properly accounts for these effects. This algorithm allows also to measure the energy loss due to radiation processes by comparing the local curvatures of the trajectory at the two endpoints. The reconstruction algorithm based on ECAL superclusters is complemented by a particle-flow based one where electromagnetic energy deposits are searched for along tangents to the electron candidate tracks drawn in correspondence to the tracker layers, where most of the detector material is located. The combination of the two algorithms significantly improves the efficiency for electrons that have low momentum or are not isolated.

2.2.3.2 Electrons and Photons Identification

The requirements for the identification of electrons [42] and photons [43] are of extreme importance to separate the signal from the reducible backgrounds. Among the most discriminating variables are the ratio between the energy detected in the electromagnetic and in the hadronic calorimeters, and the transverse size of the shower in the η coordinates, the latter being insensitive to radiative losses. In the case of electrons, additional discrimination can be added by investigating the ratio between the energy measured in the tracker and in the ECAL, and the distance between the barycentre of the supercluster and the impact point of the GSF track extrapolated to the ECAL surface. Further improvements for electrons can be achieved by using the ratio between estimates of the electron momentum at the production point and at the ECAL. The initial momentum is estimated using the local curvature at innermost point of the GSF track; the final momentum is estimated either using the local curvature at the other endpoint of the GSF track, or using the most energetic cluster within the supercluster. Identification of electrons from photon conversions in the tracker is also crucial to reject $\pi^0 \to \gamma\gamma$ decays followed by $\gamma \to e^+e^-$, for which the previously mentioned variables provide scarce rejection power. Two main features can be used in order to identify electrons from photon conversions. Tracks from conversion are not supposed to have hits in the innermost tracker layers, while tracks from prompt electrons should provide hits whenever their trajectory crosses the sensitive area of an active tracker module. Conversions in which both the electrons from $\gamma \to e^+e^-$ are reconstructed in the inner tracker can be identified by the presence of two oppositely charged tracks with collinear momenta at the conversion vertex. The impact parameter of the electron tracks with respect to the primary vertex can also be used to discriminate against electrons from the decays of heavy flavour hadrons and from photon conversions if needed. As in the case of muons, specific approaches can be used in the various analyses to select electrons and photons on the basis of these variables, ranging from simple selection criteria applied to each variable individually to the use of multivariate algorithms able to combine all or some of these variables into a single discriminator.

2.2.4 Hadron Calorimeter

The hadron calorimeter (HCAL) is a sampling calorimeter that complements ECAL to form a complete calorimetry system. The HCAL is located outside the ECAL and within the coil of the solenoid, as shown in Figure 2.8. The design of the calorimeter is thus driven by the space restrictions, covering a $|\eta| < 1.3$ region in the barrel (HB) and the $1.3 < |\eta| < 3$ region in the endcaps (HE). The intense magnetic field acting within the solenoid also required the use of non-ferromagnetic materials. Both barrel and endcaps are in fact mainly composed of non-magnetic brass (the first and last layers made of stainless-steel) with interaction length of $\Lambda_I = 16.4$ cm, interspersed with plastic scintillator tiles. The calorimeter is segmented and arranged in towers of 0.087×0.087 in η-ϕ for $|\eta| < 1.6$ and 0.17×0.17 elsewhere, each tower matching a set of 5×5 ECAL crystals. The scintillation light coming from each tile of one tower is channelled by wavelength-shifting optic fibres, added together with the others, and finally translated into an analogical signal by hybrid photodiodes (HPD). Whereas the HE has a sufficient depth of about 10 interaction length, the HB is restricted by available space to only about 5.8 Λ_I for $\eta = 0$. An additional contribution of about 1 interaction length must be added by the contribution of the ECAL crystals. The HCAL is complemented in the barrel by the so-called outer calorimeter (HO), located outside the solenoid, which extends the depth to at least 11.8 Λ_I, taking the solenoid itself into account. An additional calorimeter is used in order to extend the angular coverage in the high pseudorapidity region $3 < |\eta| < 5.2$. The forward calorimeter (HF) is placed at $z = \pm 11.2$ m from the interaction point. Due to the high flux of particles in the forward regions, radiation-hard materials are chosen. Thus, Cherenkov-light-emitting quartz fibres are used as active material, together with steel as absorber material. The HF is segmented in towers of 0.175×0.175 in η-ϕ. To compensate the missing ECAL in this region, quartz fibres of different lengths are used to identify early-showering electrons and photons.

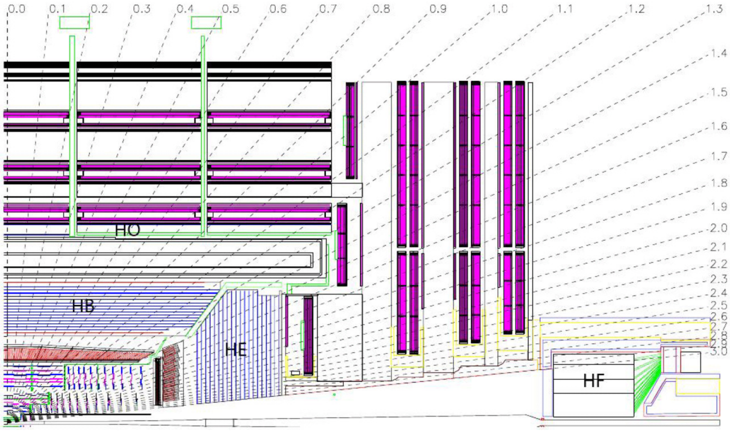


Figure 2.8: Longitudinal view of one quarter of the detector in the $r\eta$-plane, showing the positions of the HCAL parts: hadron barrel (HB), hadron outer (HO), hadron endcap (HE) and hadron forward (HF).

2.2.5 Trigger and Data Acquisition

The running conditions at LHC impose severe constraints on the design of trigger systems mainly due to the enormous data rate and the harsh radiation environment. As discussed previously in this Section, an average of about 20 superimposed inelastic non-diffractive interactions per bunch crossing have been observed during the 2012 pp data taking period. Given the high segmentation of the CMS detector, approaching 100 million channels, this corresponds to an enormous volume of data at the detector front-ends. The selection of signal events must thus start online. After zero suppression the data size per bunch crossing is $\sim$ 1MB resulting in a raw data rate of about 40 TB per second. Technical difficulties in handling, storing and processing such extremely large amounts of data impose a reduction factor on the rate of events that can be written to permanent storage. Furthermore the large number of minimum bias events per crossing combined with small cross-sections of possible discovery signatures or rare physics processes requires a sophisticated online event selection system. These figures are in fact many orders of magnitude larger than the archival storage capability of $\mathcal{O}\left(10^2\right)$ Hz at data rates of $\mathcal{O}\left(10^2\right)$ MB/s. The required rejection power of $\mathcal{O}\left(10^5\right)$ is too large to be achieved in a single processing step, if a high efficiency is to be maintained for rare physics phenomena. For this reason, the full selection task is split into two steps, as illustrated in Figure 2.9. The first step (Level-1 Trigger) is designed to reduce the rate of events accepted for further processing to less than 100 kHz. The second step (High-Level Trigger, or HLT) is designed to reduce this maximum Level-1 accept rate of 100 kHz to a final output rate of about

100 Hz.

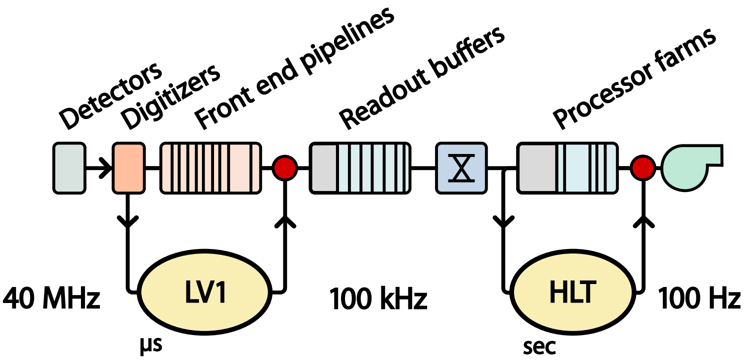


Figure 2.9: Data rates at each stage between the detectors and the data storage centre.

2.2.5.1 Level 1 Trigger

The Level-1 trigger is a high bandwidth, fixed latency system based on custom electronics: FPGAs (Field Programmable Gate Arrays) and ASICs (Application Specific Integrated Circuits). It is designed to to take a fast accept/reject decision every bunch crossing, on the basis of a rough reconstruction of the event. As the bandwidth requirement of the tracker is too large to allow read out of every bunch-crossing, the Level-1 trigger uses only rough information from coarse segmentation of calorimetric and muon data; the remaining detector data is kept in pipelines and passed to the Higher Level Trigger for further processing if the event is accepted by the Level-1. The trigger is designed to take a decision every 25 ns by processing data in a synchronous pipelined architecture, where every processing element takes less than 25 ns to complete its operation. During the Level-1 processing phase, the full detector information is stored in pipelined buffers for a constrained maximum time period of 3.2 μs (128 bunch crosses at the nominal 25 ns bunch spacing). The Level-1 trigger system can be further subdivided into three majors subsystems: the muon trigger, the calorimeter trigger, and the global trigger, as shown in Figure 2.10. The first two systems process informations coming from muon spectrometer and calorimeters respectively, and do not have to perform the task of rejecting/accepting events by themselves. Instead they identify and perform sorting on various types of trigger objects (such as electron/photon, jets and muons) and then they forward the four best “candidates” of each kind of trigger object to the Global Trigger where the final decision is taken. A multi level architecture is used. Trigger primitives generators (also referred as Local Triggers) perform a local reconstruction using a small fraction of the interested subdetector, to identify energy deposits in calorimeters or precise hit patterns in the muon chambers, producing basic objects, called trigger primitives. Regional triggers identify higher level objects matching the informations coming from the different primitives, and deliver physical quantities, such as p_T. The global calorimeter and muon triggers combine the information coming from different regional triggers, and perform the final ranking on the basis of physical informations as well as quality bits assigned from the previous components of the trigger chain. The global trigger finally selects the events on the basis of programmable trigger conditions. These can be requirements on the presence of different objects whose energy or momenta corresponds to predefined thresholds defined in advance, or topological constraints and correlations between the trigger objects themselves.

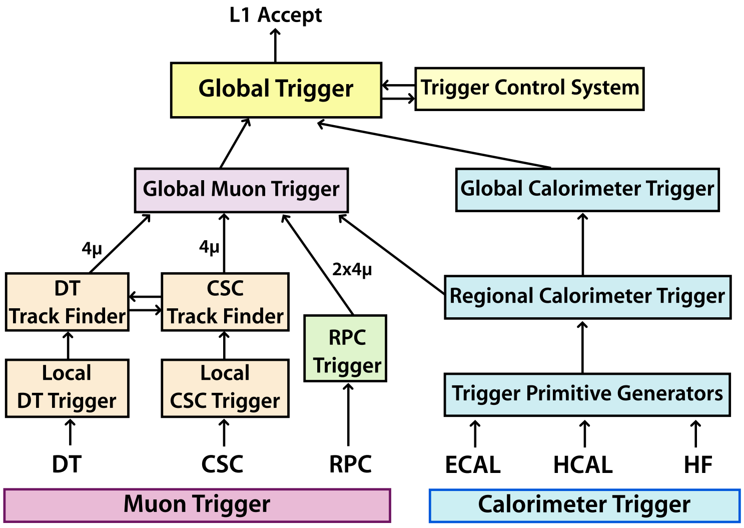


Figure 2.10: L1 Trigger decision flow of CMS before data is being transfered to the data acquisition (DAQ) system.

2.2.5.2 High Level Trigger and DAQ

The CMS high level trigger has the task to further reduce the (up to) 100 kHz event rate delivered from the Level-1 to the DAQ system to match the 100 Hz limit required from the storage system. A rejection factor of about $\mathcal{O}\left(10^3\right)$ is thus to be achieved. To fulfil this requirement, the HLT performs an analysis whose complexity is similar to the one used during off-line reconstruction, and it has to access the full information content available in the event. This computational complexity, together with the request to achieve maximum flexibility, brought to the decision to perform HLT reconstruction on a farm of commercial processors running the same software framework used for offline analysis. The HLT is only a part of the whole CMS DAQ system, whose architecture is schematically shown in Fig. 2.11. Data coming from the detector is initially stored inside 40 MHz pipeline buffers by the front end system (FES). Upon arrival of an acceptance signal from the Level-1, data is moved by the front end drivers (FEDs) to the front end readout links (FRLs) which are able to store informations coming from up to two FEDs. A total number of 626 FEDs and 458 FRLs is used to manage data coming from the different subdetectors and their trigger subsystems. Event fragments coming from different FRLs are then sent to the event builder system in charge to build up the full event. The latter is implemented in two different stages, referred as FED-builder and RU-builder. The first one is burden to transport data from underground to the CMS surface buildings, where the rest of the farm is located, and perform a first assembly phase. The FED-builder output consists of 72 super-fragments that are then stored in large readout units (RU) buffers waiting for the final building phase to be performed by the RU-builders. This two level architecture allows to deploy the system in up to 8 separated DAQ slices, each of them nearly independent from the others and able to handle a 12.5 kHz event rate. After the assembly phase is completed, the event is sent to the event filter where HLT algorithms, together with some data quality monitoring operations, are performed. Here the filter unit resource broker (FU-RB) takes care of sending the data coming from the event builder to the filter unit event processor (FU-EP) in charge to perform the HLT processing. As already mentioned, the FU-EP uses the CMS reconstruction software framework for HLT and data quality monitoring purposes. Filtered data are then separated into several online streams, whose content depend on trigger configuration, and are sent to a local storage system before being migrated to the CERN mass storage infrastructure.

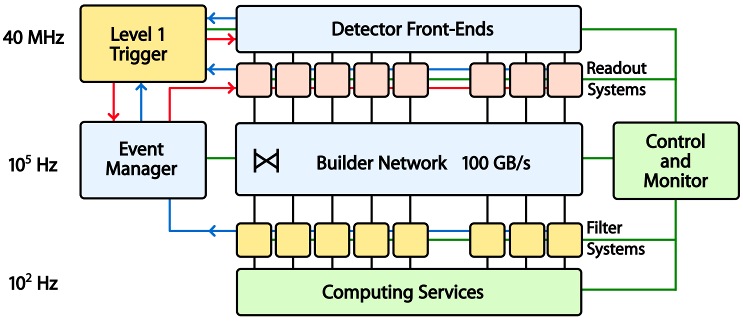


Figure 2.11: Architecture of the CMS data acquisition (DAQ) system.

2.2.6 Particle Flow Event Reconstruction

The particle-flow event reconstruction aims at reconstructing and identifying all stable particles in the event, i.e. electrons, muons, photons, charged hadrons and neutral hadrons, by mean of an optimized combination of informations from all CMS subdetectors. The algorithm is described in detail in References [44–46] where information on its commissioning with early data are also provided. The CMS PF algorithm relies on a efficient and pure track reconstruction, on a clustering algorithm able to disentangle overlapping showers, and on an efficient link procedure to connect together the deposits of each particle in the sub-detectors. Tracks are extrapolated through the calorimeters: if they fall within the boundaries of one or several clusters, the clusters are associated to the track. The set of track and cluster(s) constitute a charged hadron and are not considered anymore in the rest of the algorithm. The muons are identified beforehand so that their track does not give rise to a charged hadron. The electrons are more difficult to deal with. Indeed, due to the frequent Bremsstrahlung photon emission, a specific track reconstruction (the GSF already discussed in Section 2.2.3.1) is needed as well as a dedicated treatment to properly attach the photon clusters to the electron and avoid energy double counting. Once all the tracks are treated, the remaining clusters result in photons in case of the electromagnetic calorimeter (ECAL), and in neutral hadrons in the case a hadron calorimeter (HCAL) cluster is matched to an ECAL cluster. Once all the deposits of a particle are associated, its nature can be assessed, and the information of the subdetectors combined to determine optimally its four-momentum. In case the calibrated calorimeter energy of the clusters, which is simply a linear combination of the ECAL and HCAL energy deposits, associated to a track is found to be in excess with respect to the track momentum at more than one sigma, the excess is attributed to an overlapping neutral particle (photon or hadron), carrying an energy corresponding to the difference of the two measurements. The resulting list of particles, namely charged hadrons, photons, neutral hadrons, electrons and muons, is then used to reconstruct the jets, the missing transverse energy, to reconstruct and identify the τ from their decays products and to measure the isolation of the particles.

The association used in the linking stage is purely geometrical: tracks are linked to calorimetric clusters if their trajectory intersects one of the calorimetric cells of the cluster; and likewise clusters in the ECAL preshower, ECAL and HCAL are linked if the cluster position measured in the finer granularity subdetector lies within the envelop of the cluster in the coarser granularity subdetector. In order to account for uncertainties from multiple scattering in the track extrapolation and on the estimated position of the shower maximum in the calorimeters, a geometrical tolerance of the size of one calorimeter cell is included when defining links; this tolerance can also account for gaps and cracks in the calorimeters. By design, the linking algorithm is simple and robust, as it does not rely on the precise knowledge of the position resolution in each subdetector. Specialized algorithms are used for linking tracks to recover bremsstrahlung clusters in the case of electrons by considering tangents to electron trajectories at the crossing points with the tracker layers. Blocks of one or more linked objects are then processed to identify and reconstruct particle candidates. Isolated electrons and muons are selected first, and reconstructed using the dedicated algorithms developed for them; similarly, non-isolated tracks which satisfy tight muon identification criteria are immediately identified as muons. Charged hadrons are identified as tracks in the inner tracker, normally linked to calorimetric deposits if the particle p_{T} is sufficient for the trajectory to reach the calorimeters. If the momentum measurements from the track and calorimeter are compatible, after accounting for non-linearities and zero suppression effects, the best energy determination is obtained as a combination of the two. If the track momentum significantly exceeds the measured calorimetric energy, the particle is identified as muon if it satisfies very loose muon identification criteria; otherwise, tight track quality requirements are applied to to reject mis-reconstructed tracks. If instead an excess of calorimetric energy deposition is found with respect to the momentum of the associated track, e.g. in the case of collimated hadronic jets, the residual energy is identified as a photon or a neutral hadron. Additional photons and neutral hadrons are also identified from calorimetric deposits not linked to any track.

Chapter 3

Introduction to the Tagging of the Neutral B Flavour

3.1 Introduction and Definitions

Aim of flavour tagging algorithms is to determine whether, at production time, a physical state is produced as a particle or as an antiparticle. While this issue is trivial for self-tagging charged particles like D^+, B^+, or τ^+ leptons[1], it is far more complicated for neutral states such as D^0,B^0, and B^0_s which undergo flavour mixing and can decay to non-tagging final states (CP-eigenstates or other flavour neutral combinations of detectable particles). In this case, a more complex strategy must be deployed.

Flavour tagging algorithms are developed to infer the flavour of neutral B mesons at production time by associating to each B a tag decision (d) usually defined as $d = +1$ for B and $d = -1$ for $\overline{B}$. If the tagger is not able to provide a decision, the latter is usually set to $d = 0$. It is common to refer to the events corresponding to the $d = \pm 1$ tagger response as the *tagged* events, the complementary sample being the *un-tagged* events.

Tagging algorithms are characterized by their effective tagging efficiency, or tagging power, described by the following relation:

$$\mathcal{P}_{tag} = \varepsilon_{tag}\mathcal{D}^2 = \varepsilon_{tag}(1 - 2\omega)^2 \tag{3.1}$$

where ε_{tag} is defined as the tag efficiency, ω is the mistag (or wrong-tag) fraction, and $\mathcal{D}$ as the dilution.

The tagging efficiency of a tagger represents the fraction of events to which the algorithm is able to assign a tag decision, i.e. the fraction of tagged events. Given U the number of events for which the tagging algorithm is not able to determine the initial flavour, and defined R and W the number of events for which the decision taken by the algorithm is right and wrong respectively, the tagging efficiency can be represented by the following ratio:

$$\varepsilon_{tag} = \frac{R + W}{R + W + U} \tag{3.2}$$

The mistag is the fraction of tagged events with an incorrectly assigned tag decision, and

---

[1] Charge-conjugate states are thoroughly implied, unless the opposite is explicitly stated

it can be represented as the probability of having assigned a wrong tag decision to the event. Similarly to Equation 3.2, the mistag can be calculated as follows:

$$\omega = \frac{W}{R+W} \tag{3.3}$$

The dilution term $\mathcal{D}$ is simply related to the mistag, and defined as $\mathcal{D} = 1 - 2\omega$.

The tagging power $\mathcal{P}_{tag}$ is a variable representing the effective statistical reduction of a sample due to the imperfect tagging. It is widely used as the figure of merit in the optimization and evaluation of the performances of tagging algorithms. In other terms the tagging power is equivalent to the fraction of events that could be tagged in the hypothetical limit of a perfect tagger. It is in fact trivial to determine that in the case of null mistag $\omega = 0$, the tagging power definition of Equation 3.1 reduces to $\mathcal{P}_{tag} = \varepsilon_{tag}$. For any given event sample, the overall tagging power can be obtained by measuring the tagging efficiency ε_{tag} and the total mistag fraction ω.

The tagging power can be improved by exploiting multiple mutually exclusive categories, such as non-overlapping subsamples of events selected according to a variable able to differentiate between events with different average levels of mistag. For each category j, a fraction of tagged events ε_j and a mean dilution $\mathcal{D}_j$ can then be defined. The total tagging efficiency is clearly defined as $\varepsilon = \sum_j \varepsilon_j$, while the average value of the dilution is described as $\mathcal{D} = \sum_j \varepsilon_j \mathcal{D}_j / \varepsilon$. Using a categorization of events with different values of the dilution, one finds that the tagging power can be evaluated as follows:

$$\mathcal{P}_{tag} = \sum_j \varepsilon_j \mathcal{D}_j^2 = \varepsilon \mathcal{D}^2 + \sum_j \varepsilon_j \left(\mathcal{D}_j - \mathcal{D}\right)^2 \tag{3.4}$$

which is larger than the one obtained considering the entire sample as a single category $\varepsilon_{tag}\mathcal{D}^2$. The tagging power is the same for the two methods only in the case all the categories have the same value of dilution ($\mathcal{D}_j = \mathcal{D}_i,\ \forall i, j$).

As a natural extension of the described categorization procedure, when several alternative taggers are available the total tagging power can be obtained as the sum of the individual contributions, as long as the taggers do not overlap, i.e. if each event is tagged at most by one algorithm. In the case of multiple algorithms providing the tag for the same events, the results from the various taggers must be combined into a single tag decision, correctly taking into account the correlation between the tagging decisions and mistag of the different algorithms.

In CP-violation measurements the systematic uncertainty on the result is strongly correlated to the systematic uncertainty on the tagging power of the flavour tagging algorithm, which in turn depends on the knowledge of the fraction ω of wrong tags. The mistag fraction can be measured from data control samples, applying the tag to events where a final state with a well defined flavour content is identified (for instance, fully reconstructed B^+ final states). This allows a sizeable reduction of the systematic uncertainty.

3.2 Tagging Strategies

Two main families of tagging algorithms can be defined depending on the tagging strategy used in order to infer the flavour of the initial state for a B meson: Opposite Side (OS) and Same Side (SS) taggers. With reference to Figure 3.1 the reconstructed B meson in the event (e.g. a B_s^0) is regarded as the *signal* meson, hence defining the so-called same side of the event. Any secondary B hadron in the event is thus defined as the opposite side B meson.

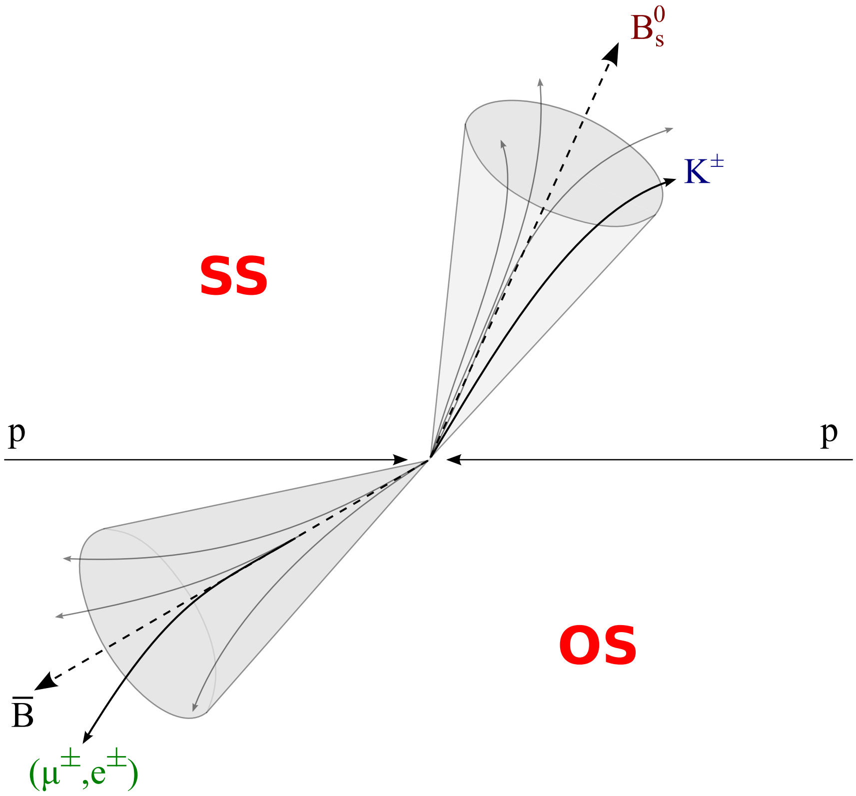


Figure 3.1: In this schematic representation of the tagging topologies, the B_s^0 (in red) is the fully reconstructed b-hadron decay. A Same Side tagger (SS) can be used to infer the flavour of the B_s^0 at production time by studying the properties of the nearby particles, e.g., the charge of a kaon (in blue) in the hadronization cone. Opposite Side taggers (OS) exploit the decay products of the other b-hadron ($\overline{B}$ in the picture) produced in the event, among which the charge and kinematic properties of the $\overline{B}$ semileptonic (μ,e in green) decays.

3.2.1 Opposite Side Taggers

Opposite side (OS) taggers exploit the features of flavour production of b-quarks at particle colliders. B hadrons arise from the hadronization of b quarks, which are produced in pairs of opposite flavour ($b\overline{b}$) at the LHC, as described in Figure 3.2.

Single b quark can also be produced at the LHC from topologies such as the s-channel single-t production, as shown in Figure 3.3. Being however heavily suppressed with respect to the $b\overline{b}$ flavour production, the single-b contributions are negligible.

When a B meson is reconstructed in a CP eigenstate, its flavour at production time can be inferred from the flavour of the *other* b. In practice, to ensure large-enough efficiency, opposite side taggers do not attempt a full reconstruction of the other B-hadron (OS-

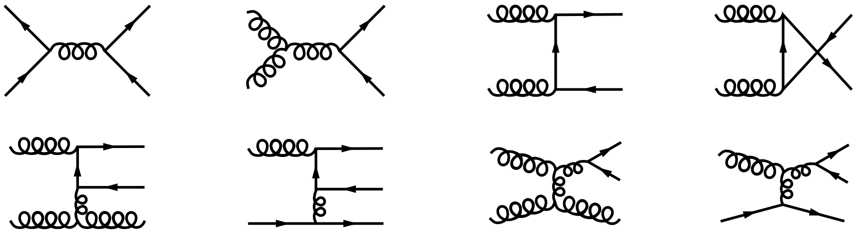


Figure 3.2: Leading (top row) and next-to-leading (bottom row) order Feynman diagrams for quark pair production.

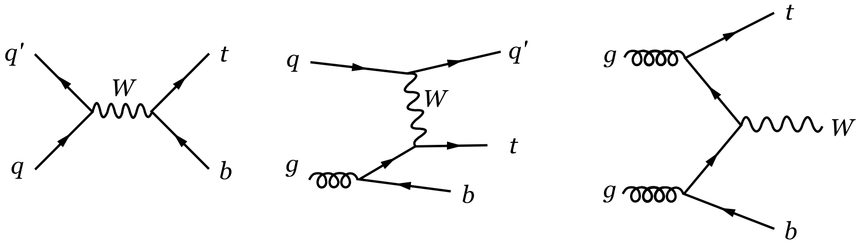


Figure 3.3: Feynman diagrams for single top quark production: (a) leading order s-channel; (b) next-to-leading t-channel, and (c) next-to-leading W-associated production.

B), but exploit inclusive tags, where only few particles from the OS-B decay are in fact analysed.

The performances of this kind of algorithms are subjected to two main sources of dilution: those affecting the tagging efficiency and those affecting the mistag. The first is related to the ability of the tagger to select and identify particles produced in the inclusively reconstructed decay of the opposite side B hadron. For instance, the products of the OS-B decay could fall outside the acceptance of the detector, determining a reduction of the tagging efficiency. The incorrect association of particles not involved in the decay of the opposite side B hadron spoils the performances of the tagger by inducing a (generally) random correlation between the OS-B flavour and the particle's properties. A second source of degradation of the OS tagging performances is related to the processes occurred between the production and decay time of the OS-B hadron, such as the flavour oscillation for neutral B mesons or the selection of particles coming from further cascade decays of the B hadron which could result in an inverted tagger response. Several independent OS tagging algorithms can be defined by exploiting different products of the OS-B hadrons: semileptonic and kaon decays, secondary vertices, and the hadronization jet.

Lepton taggers exploit the leptons (electrons or muons) produced in semileptonic $\mathrm{B} \to \ell X$, while kaon taggers acts similarly by investigating the properties of the kaon produced in the decay process. For both the taggers the tag decision is usually provided by the charge of the reconstructed particle (ℓ,K): a negative-charged lepton is mainly produced from a b-hadron, and therefore flags the other side as a $\overline{\mathrm{b}}$-hadron; the same relation holds true for the kaon tagger. The mistag of this tagging algorithm is affected by processes such as the cascade decay $\mathrm{b} \to \mathrm{c\,X} \to \ell\,\mathrm{X}'$ for the lepton tagger and the mis-identification of the kaon track for the kaon tagger, introducing dilutions of the tag informations. The tagging efficiency is limited to the branching fraction of the semileptonic or kaon decay, and is further reduced by the selection and identification requirements of the ℓ, K particles. The development of an opposite side lepton tagger at the CMS detector will be discussed in more detail in Section 4.

Jet tagging algorithms can be used in order to infer the properties of the OS b-quark at production time by the overall properties of the jet of particles produced in the hadronization of the b-quark and in the decay of the produced B hadron. Informations about the flavour of the OS b-quark could in fact be extracted on a statistical basis combining conveniently in a weighted sum the charges of all the associated particles.

Variables involving the charges of the tracks produced within the opposite side b-jets are usually defined to infer the flavour of the reconstructed B hadron at production. The efficiency of an OS jet tagger is generally very high compared to the lepton tagger, depending to the jet selection requirements. The tagging power is however quite limited due to the high amount of mistag to which the jet tagger is subjected mainly due to the poor correlation between the tagging variables and the actual flavour of the OS-B.

The above limitations are overcome by the use of vertex taggers, which rely on the reconstruction of secondary vertices (SV) produced by the decay of the OS-B hadrons. The features of the particles associated to the SV (such as the transverse momenta and the charges) can be combined in order to extract the information of the B hadron charge at decay time. This tagger is generally affected by low tagging efficiencies due to the reconstruction and selection of the secondary vertices. The vertex tagger is however characterized by a lower mistag compared to the other OS algorithms; in fact the tag decision is mainly spoiled only by the flavour oscillation of neutral B mesons between production and decay time and by the wrong assignment of the tracks to the reconstructed secondary vertex and by the loss of low-momentum particles.

3.2.2 Same Side Taggers

Same side (SS) taggers exploit the flavour correlation between the B-hadron and additional nearby particles produced in the hadronization process. SS taggers are widely used to tag the flavour of neutral B^0_s mesons by the sign of the charged particles produced nearby. In the case of the formation of a B^0_s meson, the $\overline{b}$-parton is bounded to a s-quark produced by a s-$\overline{s}$ pair, leaving an isolated $\overline{s}$ which then hadronizes about 50% of the times bounding with an u-quark into a positively charged kaon K^+. The charge of the kaon can therefore be used to determine the flavour of the neutral B^0_s meson. Analogously, the charge of the same side pion could be on principle exploited to tag the flavour of the B^0 meson. However, due to the much higher number of pions produced in a typical LHC event with respect to the number of kaons, and given the experimental difficulties in the selection and identification of soft-pions arising from hadronization processes, the same side algorithm is applied only to tag B^0_s decay channels.

3.3 Flavour tagging in other experiments

Flavour tagging algorithms are (and have been) widely applied in many experiments involved in the study of the physics of heavy flavour. Taggers are mainly exploited to perform measurements of flavour oscillation and CP-violation (CPV) in the evolution of neutral B mesons. Opposite side taggers have been developed and exploited by the beauty-factories experiments BaBar and Belle [47] by the CDF [48] and DØ [49] collaboration at Tevatron, and more recently by LHCb [50] and ATLAS [51] at the LHC. Same side taggers, experimentally more challenging than the opposite-side counterpart because they require kaon identification, have been developed and exploited by CDF [48] and LHCb [52] and have been used in addition to the opposite side taggers. In flavour-physics dedicated experiments located at B factories, e^+e^- colliders operating at $\sqrt{s} = m_{\Upsilon(4S)}$, tagging powers as high as 30% can be obtained, whereas hadron colliders experiments are able to reach $\mathcal{P}_{tag}$ a few percent at most.

The impressive tagging performances of beauty factory experiments such as Belle and BaBar could be achieved mainly due to the ability to access almost the entire phase-space of the B-$\overline{\text{B}}$ decays and the identification and reconstruction of tagging particles with very low momenta, together with the extremely low levels of background due to the absence of the fragmentation products. Due to the lack of fragmentation products SS tagging cannot be exploited at the beauty factories.

Table 3.1 reports the tagger performance obtained by different experiments at lepton and hadron colliders.

Table 3.1: Tagging power $\mathcal{P}_{tag}$ obtained in experiments at lepton and hadron colliders for the combined same side (SS) and opposite side (OS) flavour tagging algorithms.

Collider	Experiment	OS	SS
$e^+ \, e^-$	BaBar	33.1 ± 0.1	
	Belle	30.1 ± 0.4	
hadronic	CDF	1.8 ± 0.1	$3.5 \pm 0.9 (4.0 \pm 1.0)^*$
	DØ	2.48 ± 0.21	
	ATLAS	1.45 ± 0.05	
	LHCb	2.10 ± 0.25	2.42 ± 0.39

* estimated from simulations only in hadronic (semileptonic) decay samples [48]

3.4 Flavour Tagging in CMS

3.4.1 Strategy

Even if designed for high-p_{T} physics, due to the excellent track and vertex reconstruction and muon and electron identification, CMS can achieve competitive performances on most of the discussed flavour tagging algorithms. Lack of particle identification detectors such as Cherenkov or transition-radiation detectors inhibits however the possibility to develop tagging algorithms which require kaon identification. The possibility to identify kaons by exploiting the energy loss $\mathrm{d}E/\mathrm{d}x$ of the particles traversing the CMS silicon tracker is proved to extend only up to values of particle's momenta about $p \lesssim 1$ GeV [53], too low for the purpose of this studies.

The main topic of this thesis is the development of opposite side lepton taggers at the CMS experiment using electrons and muons. The strategy used in order to develop the algorithms is explained below.

Samples of simulated data containing a $\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\,\phi$, $\mathrm{B^+} \rightarrow \mathrm{J}/\psi\,\mathrm{K^+}$, and $\mathrm{B^0} \rightarrow \mathrm{J}/\psi\,\mathrm{K^*}$ decay are used as reference channels for the development and the optimization of the tagging algorithm. The generation of the simulated samples is described in Section 3.4.2. The tagging algorithms are defined and optimized using the $\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\,\phi$ simulated sample, exploiting the true Monte Carlo information (MC truth) to define the flavour of the reconstructed $\mathrm{B_s^0}$ at production time. As previously discussed in this section, the tagging power $\mathcal{P}_{tag}$ is used as the figure of merit for the optimization. The tagging algorithm is then applied to the $\mathrm{B^+} \rightarrow \mathrm{J}/\psi\,\mathrm{K^+}$ and $\mathrm{B^0} \rightarrow \mathrm{J}/\psi\,\mathrm{K^*}$ simulated events and the performances (efficiency, mistag and tagging power) of the three MC samples are compared to verify the independence of the tagger to the flavour of the reconstructed B meson. The performances of opposite side tagging algorithms should in fact be independent from the specific flavour of the reconstructed (same side) B hadron, and potential discrepancies are corrected and addressed by a dedicated systematic uncertainty.

Finally, the algorithm developed by studying the simulations is applied to the self tagging $\mathrm{B^+} \rightarrow \mathrm{J}/\psi\,\mathrm{K^+}$ data sample. The background events not associated to the $\mathrm{B^+} \rightarrow \mathrm{J}/\psi\,\mathrm{K^+}$ decay are subtracted by a fit to the $\mathrm{B^+}$ invariant mass spectrum, described in Section 3.4.3. It is important to remark that the simulated samples are used in this analysis in order to define the algorithm and to evaluate some of the systematic uncertainties. The final tagger performances are in fact measured directly and only on $\mathrm{B^+} \rightarrow \mathrm{J}/\psi\,\mathrm{K^+}$ data, and will be applied on the different measured samples, like for instance in the case of the CP-violation measurement in the $\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\,\phi$ or $\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\,\mathrm{f_0}$ decay.

3.4.2 Data and Monte Carlo Simulations

The data employed for the evaluation of the tagging algorithm were collected by the CMS detector during the 2012 run of the LHC, corresponding to an integrated luminosity of about 19.8 fb^{-1}.

As previously described, flavour tagging algorithms are developed and validated in CMS by means of Monte Carlo (MC) samples. The simulated processes involve the reference samples of $\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\,(\mu^+\mu^-)\;\phi\,(\mathrm{K^+K^-})$, $\mathrm{B^+} \rightarrow \mathrm{J}/\psi\,(\mu^+\mu^-)\;\mathrm{K^+}$ and $\mathrm{B^0} \rightarrow \mathrm{J}/\psi(\mu^+\mu^-)$ $\phi(\mathrm{K^+K^-})$ decays, considered as test channels for the development of the tagging algorithms. The hard parton process of the pp collisions is generated with PYTHIA [54], which performs the evaluation of matrix elements at leading order perturbative QCD and the convolution with the proton parton density functions defined by CTEQ [55]. The parton shower algorithm is based on a leading-logarithmic approximation for QCD radiation and a string fragmentation model is applied. The simulation of the the decay of short lived particles, including b and c hadrons is managed by EVTGEN [56]. Final state photon radiation is included in EVTGEN through the PHOTOS [57, 58] package. The propagation of the generated particles through the detector, including the full magnetic field mapping, and their interaction with the detector material is performed with the GEANT [59] package, which also manages the decay of long-lived generated particles, such as kaons or pions decaying in the detector material far from the vertex of the partonic interaction.

A dedicated sample of 5658339 simulated events containing at least one $\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\,\phi$ decay has been produced by selecting at generator level pp collisions resulting in the production of a $\mathrm{b\overline{b}}$ pair, with one of the two b quarks hadronizing as a $\mathrm{B_s^0}$, and forced then to decay according to the following cascade process: $\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\,\phi$, followed by $\mathrm{J}/\psi \rightarrow \mu^+\mu^-$, $\phi\,(1020) \rightarrow \mathrm{K^+K^-}$. The other b in the event is allowed to hadronize and decay without any constraint. A sample of 4898805 $\mathrm{B^+} \rightarrow \mathrm{J}/\psi\,\mathrm{K^+}$ events and a sample of 6093292 $\mathrm{B^0} \rightarrow \mathrm{J}/\psi\,\mathrm{K^*}$ events have been produced as well, where the J/ψ is always forced to decay into a $\mu^+\mu^-$ pair and the $\mathrm{K^*}(892)$ is naturally allowed to decay into the chain $\mathrm{K^*}(892) \rightarrow \mathrm{K^+}\pi^-$.

Table 3.2: Data and Monte Carlo samples used. Data were recorded into four periods of 2012 LHC run (named A,B,C,D). The corresponding integrated luminosity of the recorded data samples, and the total number of generated events for the MC samples are shown.

Data	Luminosity [fb^{-1}]
Run2012-A	0.90
Run2012-B	4.43
Run2012-C	7.15
Run2012-D	7.32
Run2012 (total)	19.8

Monte Carlo	Generated events
$\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\,\phi$	5658339
$\mathrm{B^+} \rightarrow \mathrm{J}/\psi\,\mathrm{K^+}$	4898805
$\mathrm{B^0} \rightarrow \mathrm{J}/\psi\,\mathrm{K^*}$	6093292

3.4.2.1 Trigger and Event Selection

Data are required to pass one of the displaced-J/ψ trigger paths. These require the online reconstruction of at least two muons with opposite charge, each with $p_\mathrm{T} > 4$ GeV, $|\eta| < 2.1$ and $\cos\theta > 0.9$, where θ is the angle between the muon momentum and the dimuon flight direction. The two opposite charge tracks are required to have a distance of closest approach of less than 0.5 cm. Candidates J/ψ are required to have a mass within the range $2.9 - 3.3$ GeV and a minimum transverse momentum of 6.9 GeV. The trigger requirements are summarized in Table 3.3.

Table 3.3: Online selection cuts of the displaced J/ψ trigger paths.

Variable	Online selection cut
$p_\mathrm{T}\ (\mu)$	> 4GeV
$|\eta|\ (\mu)$	< 2.1
$\cos(\theta)$	> 0.9
$L_{xy}/\sigma(L_{xy})$	> 3
M(J/ψ)	2.9-3.3 GeV
p_T (J/ψ)	> 6.9 GeV
distance of closest approach	< 0.5 cm

The offline event selection requires the muon tracks to be fitted to a common vertex and their momenta to be recomputed with the vertex constraint. Events are rejected if the fit probability is smaller than 15%. The dimuon invariant mass is required to be less than 200 MeV far ($2.9 - 3.3$GeV) from the J/ψ mass [18]. The J/ψ candidate must have $p_\mathrm{T} > 7$ GeV and a decay length significance $\mathrm{L_{xy}}/\sigma(\mathrm{L_{xy}}) > 3$, where $\mathrm{L_{xy}}$ is the distance between the primary vertex of the event and the dimuon vertex in the transverse plane, and $\sigma(\mathrm{L_{xy}})$ is the measurement uncertainty.

Candidates $\phi\,(1020)$ mesons are reconstructed from pairs of oppositely charged tracks with $p_\mathrm{T} > 0.7$ GeV. Tracks associated to the selected J/ψ muons are removed from the list of available tracks. The invariant mass of the tracks, evaluated under the assumption of the kaon mass for both, is required to be within 20 MeV/c^2 of the PDG $\phi\,(1020)$ mass value.

The $\mathrm{B^0_s}$ candidates are finally formed by combining a reconstructed J/ψ and a reconstructed $\phi\,(1020)$ candidates. The two muons and the two kaons are subjected to a combined vertex and kinematic fit, where the dimuon invariant mass is constrained to the nominal J/ψ mass. The $\mathrm{B^0_s}$ candidates with an invariant mass between 5.2 and 5.6 GeV and having a χ^2 vertex fit probability greater than 2% have been selected.

Candidates $\mathrm{B^+}$ are formed by combining a reconstructed J/ψ and a track with $p_\mathrm{T} > 2$ GeV under the kaon mass assumption. The two muons and the track are fitted to a common vertex and subjected to kinematic fit, where the dimuon invariant mass is constrained to the nominal J/ψ mass. The $\mathrm{B^+}$ candidates with an invariant mass between 5.0 and 5.55 GeV and having a χ^2 vertex fit probability $> 2\%$ have been selected.

$\mathrm{B^0}$ candidates are formed by combining a reconstructed J/ψ and two opposite charge tracks with $p_\mathrm{T} > 0.7$ GeV (tracks associated to the selected J/ψ muons are removed from the list of available tracks). The two muons and the tracks are fitted to a common vertex and subjected to kinematic fit, where the dimuon invariant mass is constrained to the nominal J/ψ mass. Two $\mathrm{B^0}$ mass hypotheses are evaluated, under the $\mathrm{K^+}\pi^-$ and $\pi^-\mathrm{K^+}$ mass assumptions for the two tracks. The $\mathrm{B^0}$ candidates with at least one of the two invariant mass hypotheses between 5.0 and 5.5 GeV and having a χ^2 vertex fit probability $> 2\%$ have been selected. The $\mathrm{B^0}$ mass ambiguity is resolved in the MC by using the true $\mathrm{B^0}$ flavour information.

In the case of multiple B candidates passing the reconstruction cuts in the event, the one with the highest vertex fit probability is retained.

The offline selection criteria discussed are summarized in Table 3.4.

Table 3.4: Selection cuts used in the $B \to \mathrm{J}/\psi X$ meson decay reconstruction.

Variable	$\mathrm{B^0_s} \to \mathrm{J}/\psi\,\phi$	$\mathrm{B^0} \to \mathrm{J}/\psi\,\mathrm{K}^*$	$\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$
vertex probability	> 0.02	> 0.02	> 0.02
$\mathrm{M(J/\psi)}$	3.08-3.11 GeV	3.08-3.11 GeV	3.08-3.11 GeV
$\mathrm{M}(\phi\,(1020))$	1.01-1.03 GeV	-	-
M(B)	5.2-5.6 GeV	5.00-5.50 GeV	5.00-5.55 GeV
p_T (J/ψ)	> 7 GeV	> 7 GeV	> 7 GeV
p_T (K)	> 0.7 GeV	-	> 2 GeV
p_T (tracks)	-	> 0.7 GeV	-

The number of events passing the $\mathrm{B} \to \mathrm{J}/\psi X$ selection cuts are respectively 608843 for the $\mathrm{B^0_s} \to \mathrm{J}/\psi\,\phi$ MC sample, 521855 for the $\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$ MC sample, and 469172 for the $\mathrm{B^0} \to \mathrm{J}/\psi\,\mathrm{K}^*$ MC sample. The number of events passing the $\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$ selection are

$707 \pm 2 \times 10^3$, extracted by the fit to the B^+ invariant mass spectrum, described in detail in Section 3.4.3.

3.4.2.2 Simulation Bias

The flavour composition of the unconstrained side in the simulated events differs from that which would be obtained in a generic $\mathrm{b\overline{b}}$ production, where the frequencies for a b-hadron B^+, B^0, B^0_s, or a B baryon (mostly Λ_b) are, respectively [18]: $f_u = f_d = 40.2 \pm 0.7\%$, $f_s = 10.5 \pm 0.6\%$, $f_\mathrm{baryon} = 9.2 \pm 1.5\%$. It is worth to be remarked that the opposite side tagger performances depend on the opposite-side b-hadron composition. For instance, let's assume that the opposite-side b-hadron is purely given by B^0_s mesons: the rapid flavour oscillations ($\chi_\mathrm{s} \simeq 0.5$) would highly dilute the charge-flavour correlation between the tag (B^0_s decay products) and the originally produced b-quark, resulting in a mistag $\omega \sim 0.5$ (hence $\mathcal{P}_{tag} \sim 0$). The measured flavour composition on MC and the predicted fractions are listed in Table 3.5.

Table 3.5: Relative B hadron fractions measured in the $\mathrm{B}^0_\mathrm{s} \to \mathrm{J}/\psi\,\phi$, $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ and $\mathrm{B}^0 \to \mathrm{J}/\psi\,\mathrm{K}^*$ simulated samples, before the correction of the bias. The reference values for the unbiased fractions quoted by the Particle Data Group [18] are shown in the last column of the table.

	$\mathrm{B}^0_\mathrm{s} \to \mathrm{J}/\psi\,\phi$	$\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$	$\mathrm{B}^0 \to \mathrm{J}/\psi\,\mathrm{K}^*$	Ref. values [18]
f'_d	42.23 ± 0.07 %	49.09 ± 0.08 %	25.26 ± 0.07 %	40.2 ± 0.7 %
f'_u	41.92 ± 0.07 %	25.25 ± 0.07 %	49.29 ± 0.08 %	40.2 ± 0.7 %
f'_s	6.83 ± 0.03 %	15.07 ± 0.06 %	14.91 ± 0.06 %	10.5 ± 0.6 %
f'_c	0.030 ± 0.002 %	0.027 ± 0.003 %	0.029 ± 0.003 %	-
f'_Λ	9.00 ± 0.04 %	10.57 ± 0.05 %	10.51 ± 0.05 %	9.2 ± 1.5 %

The origin of the bias can be understood as follows. According to the previous definition, given f_s the fraction of b-quarks hadronizing into a B^0_s meson in the $\mathrm{B}^0_\mathrm{s} \to \mathrm{J}/\psi\,\phi$ enriched sample, the fraction of events with at least one B^0_s selected is described by the following:

$$2f_s(1-f_s) + f_s^2 \tag{3.5}$$

where $2f_s(1-f_s)$ represents the fraction of events with exactly one produced B^0_s, and f_s^2 is the fraction of events with two B^0_s mesons.

The fraction f'_s of events with two selected B^0_s meson with respect to the number of events where at least one B^0_s is found in the unconstrained side is given by the ratio:

$$f'_s = \frac{f_s^2}{2f_s(1-f_s) + f_s^2} = f_s \frac{1}{2-f_s} < f_s \tag{3.6}$$

This result does not however take into account the effect of the selection efficiency, which acts differently in the data and in the simulation, and modifies Eq. 3.6 as:

$$f'_s = \frac{f_s^2\left(2\varepsilon_s\left(1-\varepsilon_s\right)+\varepsilon_s^2\right)}{2f_s(1-f_s) + f_s^2\left(2\varepsilon_s\left(1-\varepsilon_s\right)+\varepsilon_s^2\right)} = \frac{f_s\left(2-\varepsilon_s\right)}{2-\varepsilon_s f_s} \tag{3.7}$$

In the $B_s^0 \rightarrow J/\psi\,\phi$ simulated sample at least a B_s^0 meson is always produced within the detector's acceptance for each event; therefore, a large selection efficiency is obtained, which can be safely approximated to $\varepsilon_s = 1$, leading to a measurable fraction of B_s^0 mesons f_s' of:

$$f_s' = \frac{f_s}{2 - f_s} \tag{3.8}$$

Given the dedicated filtered production of the $B_s^0 \rightarrow J/\psi\,\phi$ simulated sample, the high selection efficiency in MC induces a bias which needs to be fixed in order to correctly account for the opposite side B flavour fractions.

On data on the other hand, the factorization of the acceptance, trigger, reconstruction and selection efficiencies for a $B_s^0 \rightarrow J/\psi\,\phi$ decay determines the overall efficiency to be very limited ($\varepsilon_s \ll 1$), and under the approximation $\varepsilon_s \simeq 0$, the measured f_s' fraction is given by:

$$f_s' = f_s \tag{3.9}$$

where it is easy to note that the fraction of B_s^0 mesons produced in the opposite side results therefore unbiased and no correction is needed.

For each sample of events from dedicated MC productions, a correction factor is hence computed as a multiplicative weight applied for each event. Only events with exactly two generated b partons are retained for this study, and will be considered in the following sections.

In the case of the $B_s^0 \rightarrow J/\psi\,\phi$ sample, the weight w_s is obtained by solving the following equation:

$$f_s = \frac{w_s f_s'}{w_s f_s' + f_u' + f_d' + f_c' + f_\Lambda'} \tag{3.10}$$

where f_x' are the biased B_x hadron fractions measured on MC, and f_s is the unbiased fraction of B_s^0 mesons. It clearly holds true the relation $f_s' + f_u' + f_d' + f_c' + f_\Lambda' = 1$, being the sum of all the possible b-hadron produced fractions.

The previous equation yields, taking also into account the $B_s^0 \rightarrow J/\psi\,\phi$ selection efficiency ε_s, and referring to Eq. 3.7:

$$f_s = w_s f_s \frac{2 - \varepsilon_s}{w_s f_s (2 - \varepsilon_s) + (2 - 2 f_s)} \tag{3.11}$$

from the previous it is trivial to obtain the resulting equation for w_s:

$$w_s = 2 \frac{1 - f_s}{(2 - \varepsilon_s)(1 - f_s)} = \frac{2}{2 - \varepsilon_s} \tag{3.12}$$

If the B_s^0 selection efficiency ε_s is large (close to unit, as in the case of filtered MC samples) then $w_s \approx 2$. If however ε_s is small ($\ll 1$, as in data) than $w_s \approx 1$. Similar relations apply to the $B^+ \rightarrow J/\psi\,K^+$ and $B^0 \rightarrow J/\psi\,K^*$ samples.

All the studied MC samples are corrected for a weight factor $w = 2$

3.4.3 Fit to B^+ Data

In order to extract the number of signal events in data, the B^+ invariant mass is fitted with an unbinned extended maximum likelihood procedure to subtract the non resonant background. The signal component is modelled by a double Gaussian with common mean, centred to the PDG reference value of the B^+ meson mass m_{B^+} [18], while the background PDF is defined as the convolution of a complementary error function and a first order Chebyshev polynomial:

$$\begin{aligned} f(m_{B^+}) = f_{\text{sig}} \left[f_{\text{gaus}} \cdot G_1\left(\mu, \sigma_1, m_{B^+}\right) + \left(1 - f_{\text{gaus}}\right) \cdot G_2\left(\mu, \sigma_2, m_{B^+}\right) \right] + \\ \left(1 - f_{\text{sig}}\right) \left[f_{\text{pol}} \cdot P\left(m_{B^+}, A\right) + \left(1 - f_{\text{pol}}\right) \cdot \text{Erf}\left(\left(m_{B^+} + 5.15\right)/\text{div} + 1\right) \right] \end{aligned} \tag{3.13}$$

where:

- $G_1\left(\mu, \sigma_1, m_{B^+}\right)$ and $G_2\left(\mu, \sigma_2, m_{B^+}\right)$ are the PDFs of the two Gaussian functions with common mean (μ) modelling the B^+ mass peak
- f_{gaus} is the relative fraction between G_1 and G_2
- $P\left(m_{B^+}, A\right)$ is the first order Chebyshev polynomial
- $\text{Erf}\left(\left(m_{B^+} + 5.15\right)/\text{div} + 1\right)$ is the error function, modulated by a constant factor ($\text{div} = 3.24 \times 10^{-2}$)
- f_{pol} is the relative fraction between the Chebyshev and Error function PDFs
- f_{sig} is the total fraction of signal events

The resulting fit, applied to the events passing the $B^+ \rightarrow J/\psi\, K^+$ selection in data yield a number of $(707 \pm 2) \times 10^3$ signal events is measuring, where the reported error is statistical only. Figure 3.4 show a comparison between the B^+ invariant mass distribution in data and the fit result.

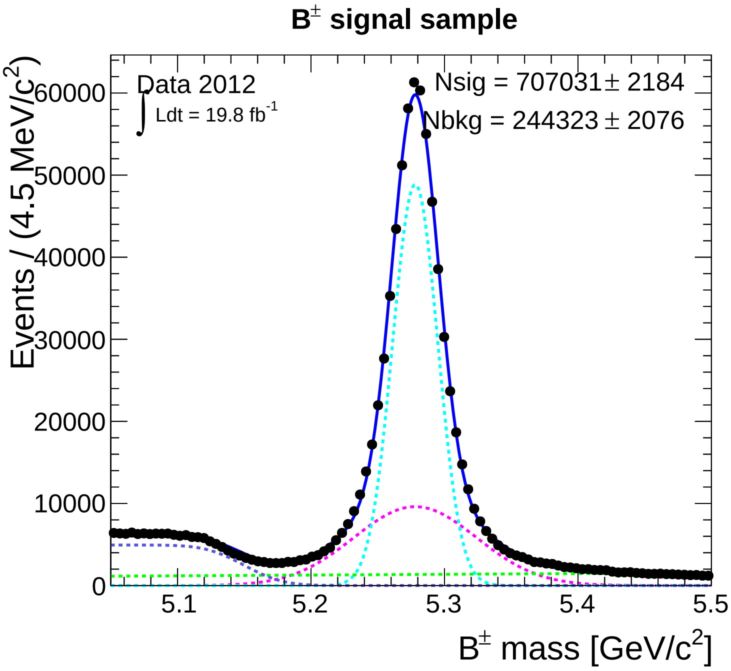


Figure 3.4: The invariant mass of the B^+ candidate (black markers) for the 2012 dataset, corresponding to 19.8 fb^{-1}. The overall fit is represented with the blue solid line, while the PDFs of the single fit components are shown with dashed lines: in light blue and in magenta the two Gaussian functions describing the signal B^+ peak, the first order Chebyshev polynomial in green and the error function in violet. The total yields of both signal (N_{sig}) and background (N_{bkg}) events are reported in figure.

Chapter 4

Opposite Side Lepton Tagging at CMS

4.1 Opposite Side Lepton Tagger Strategy

In an unbiased $\mathrm{b\overline{b}}$ sample, about 20% of the B hadrons decay via weak interaction to inclusive semileptonic $\mathrm{b} \rightarrow \ell\,\mathrm{X}$ final states, where ℓ is from now on considered to be either an electron or a muon. The charge of such leptons is a good estimator of the flavour of the related B mesons at decay time. Given a generic B hadron decay $\mathrm{b} \rightarrow \mathrm{c\,W^-}$, the charged lepton is in fact mainly originated from the further $\mathrm{W^-} \rightarrow \ell^- \nu$ decay, where it carries the same charge as the W boson exchanged in the process. If the tag-side B hadron does not undergo flavour oscillation the charge of the lepton provides a good estimator of the flavour of the opposite side B hadron at production time, and can thus be used to provide the tag decision of a flavour tagging algorithm.

Therefore, the charge of the lepton identifies the flavour of the reconstructed B hadron according to the following relation:

- OS-ℓ^+ tags an OS-B hadron at production time, thus a reco-$\mathrm{\overline{B}}$ at production time
- OS-ℓ^- tags an OS-$\mathrm{\overline{B}}$ hadron at production time, thus a reco-B at production time

The tagging efficiency of a lepton tagger algorithm is in practice reduced to at some percent level due to combination of the $\mathrm{B} \rightarrow \ell$ decay branching ratio and the trigger, reconstruction, identification and selection requirements applied.

Several processes contribute to the dilution of the tagging information, the following being the most relevant:

- The sequential $\mathrm{b} \rightarrow \mathrm{c\,W^-}$ decay, where the following semileptonic $\mathrm{c} \rightarrow \mathrm{W^+\,X} \rightarrow \ell^+ \nu\,\mathrm{X}$ decay produces a lepton with charge opposite to that expected from the direct decay.
- The neutral meson OS-B ($\mathrm{B^0}$,$\mathrm{B^0_s}$) mixing, $\mathrm{b} \rightarrow \mathrm{\overline{b}}$; also in this case the relation between the lepton charge and the b flavour is inverted.
- The flavour production via gluon splitting can result in the production of four heavy quarks ($\mathrm{b\overline{b}\,c\overline{c}}$, or $\mathrm{b\overline{b}\,b\overline{b}}$), where the lepton can be picked from an uncor-

related heavy hadron. In this case there is no relation between the charge of the lepton selected and the flavour of the same side B-hadron.

- Hadrons or photons wrongly identified as leptons due to $\gamma \to e^+e^-$ conversion, kaons or pions decaying into muons within the detector volume, hadrons absorbed in the ECAL, or sailing through the HCAL up to the muon chambers, carry weak or none charge correlation to the B-hadron flavour.

To reduce the dilution from background processes, several variables are employed. First a loose selection mostly based on the kinematic properties of the lepton is applied, then the variables are combined in a multivariate discriminator, trained on the simulated $\mathrm{B^0_s}$ sample, in order to effectively exploit the discriminating power and the correlations of all the variables involved. This variable is then used to enhance the separation between leptons carrying the correct and wrong tag decision, allowing to define regions of the leptons' phase space with various levels of mistag.

All the steps for the definition and optimization of the lepton tagger are performed independently for muons and electrons, and will be discussed in the following sections.

A detailed study of the relation between the origin of the lepton and its effect on the tagging is not performed. Instead, using MC-truth information, leptons are separated into three categories: those whose charge tags the correct flavour (labelled CC, correct charge); those carrying the wrong, i.e. opposite, charge-flavour relation (WC); and those not identified as produced from a B hadron decay, therefore contributing with a random tag information (RC). The complete Monte Carlo information is not available for leptons whose decays is managed by the GEANT package, like in the case of long lived hadron $\mathrm{b} \to \mathrm{K}(\pi)\,\mathrm{X} \to \mu\,\mathrm{X}'$ decays. These leptons are therefore assigned to the random background (RC) class.

4.2 Single Particle (e,μ) Taggers

4.2.1 Preselection

Events containing a reconstructed B meson ($\mathrm{B^0_s}$, $\mathrm{B^+}$ or $\mathrm{B^0}$ according to the simulated or data sample) are further selected requiring an additional lepton. The leptons are required to be reconstructed by the particle flow algorithm [44, 45] both in case of muons and electrons. All the tracks associated to the reconstructed B decay are explicitly excluded from the list of the candidates leptons, in order to avoid the selection of muons coming from the reconstructed J/ψ decay, or pions and kaons misidentified as muons or electrons in the case of the $\phi\,(1020)$, K^* or $\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$ decay tracks.

The tag muons are preselected according to the standard CMS muon identification prescriptions [60] using the Loose Muon criteria, which require the muon to be identified by the tracker informations only (tracker muon) or by the combination of the tracker and the muon spectrometers subdetectors (global muon), as described in Section 2.2.2.4. The PFmuon preselected with the highest transverse momentum is then retained as the tag muon.

The tag electrons are reconstructed by the particle flow algorithm with transverse momentum greater than 2 GeV. A minimum identification criterion is applied by default, requiring the electron to have the multivariate discriminator $PFmva > -0.2$ This identification variable is based on the electron features, exploiting the track and energy deposit characteristics and matching quality to specifically discriminate between electrons and pions in jets. The highest transverse momentum PFelectron of the event passing this preselection criteria is then chosen as the tag electron.

In the simulation, a reconstructed (reco) lepton is associated to the referring generated particle (gen) by asking a geometrical match, ΔR, between the track reconstructed by the inner tracker and the direction of the generated particle, and by comparing the transverse momenta, $p_{\mathrm{T}}^{\mathrm{reco}}$ and $p_{\mathrm{T}}^{\mathrm{gen}}$. The matching values have been optimized by the CMS collaboration for the best matching with the MC simulation: $\Delta\mathrm{R} < 0.5(0.12)$ and $|p_{\mathrm{T}}^{\mathrm{reco}} - p_{\mathrm{T}}^{\mathrm{gen}}|/p_{\mathrm{T}}^{\mathrm{gen}} < 0.5(0.3)$ for electrons (muons).

Using MC-truth informations, simulated leptons are assigned to one of three non-overlapping classes, depending on the relation of the muon (electron) charge with respect to the flavour of the associated B meson:

- Lepton from OS-B hadron decay, with correct ℓ charge - reconstructed B flavour relation (CC)
- Lepton from OS-B hadron decay, with wrong (opposite) ℓ charge - reconstructed B flavour relation (WC)
- Lepton not associated to OS-B hadron decays, with random ℓ charge - reconstructed B flavour relation (RC)

With the lepton preselection applied, about 20% of the events with a fully reconstructed B meson decay contain also an OS-muon and about 13% an OS-electron. However, due to the extremely loose selection and identification requirements applied at this stage, most of the preselected leptons (about 80%) do not originate from a B decay, as shown in Table 4.1.

In terms of tagging performances, despite having high tagging efficiency, both of the lepton's samples are therefore mostly composed by particles contributing with an almost null value of dilution, having a mistag close to $\omega \sim 50\%$. A tighter set of selection cuts is therefore investigated for both muons and electrons in order to reduce the contribution of the tag-uncorrelated background allowing to better exploit the tagging features of the leptons.

4.2.2 Selection

Two dedicated selections are optimized in order to reduce the RC backgrounds for tag muons and electrons. Independent cuts on a number of lepton variables have been applied, spanning two independent grids of about 10^6 alternative configurations for the two tag leptons. For each combination of cuts the tagging efficiency and the mistag fraction have been evaluated on the simulated $\mathrm{B}_s^0 \to \mathrm{J}/\psi\,\phi$ sample. Eventually, the best cut configuration is chosen to be the one which maximizes the tagging power; as discussed in the previous sections in fact, in order to obtain the best possible tagging performances, a trade-off between the efficiency and the mistag fraction is to be pursued. A selection mostly effective in the background reduction could in fact be applied; however the re-

Table 4.1: Lepton provenance evaluated for the three MC samples, after the preselection is applied. Leptons arising from a B hadron decay are separated into two categories: ℓ with correct charge correlation with respect to the B flavour (CC), and ℓ contributing to the tag dilution with opposite charge-flavour relation (WC). Leptons not originated from B decays have random charge-flavour correlation (RC).

	$B^0_s \to J/\psi\,\phi$ MC	$B^+ \to J/\psi\,K^+$ MC	$B^0 \to J/\psi\,K^*$ MC
N_{ev} reco-B	608 843	521 855	469 172
N_{ev} reco-B + OS-μ	121 029 (19.9%)	107 577 (20.6%)	93 694 (20.0%)
$B \to \mu$(CC)	15.1%	14.8%	15.4%
$B \to \mu$(WC)	5.4%	5.2%	5.4%
μ not from B (RC)	79.5%	80.0%	79.2%
N_{ev} reco-B + OS-e	77 529 (12.7%)	71 960 (13.8%)	60 984 (13.0%)
$B \to e$(CC)	14.6%	13.6%	14.6%
$B \to e$(WC)	4.9%	4.7%	5.0%
e not from B (RC)	80.5%	81.7%	80.4%

duction in tagging efficiency would not be compensated by the increase in the dilution, therefore leading to an overall reduction of the tagging power.

The following variables are used to select the candidate muons: muon transverse momentum p_T, and pseudorapidity η, muon's inner track 3-dimensional impact parameter d_{xyz}, the separation from the fully reconstructed B meson ΔR (B), and the isolation criterion evaluated with the particle flow algorithm using a ΔR cone of 0.4. The optimal working point obtained after this optimization procedure is found to be the one corresponding to the selections:

- muon transverse momentum $p_T > 2.2$ GeV/c
- muon impact parameter $d_{xyz} < 0.1$ cm
- muon separation from the reconstructed B meson $\Delta R(B) > 0.3$

When this selection is applied 22 981, 20 252 and 18 620 events are left respectively in the $B^0_s \to J/\psi\,\phi$, $B^+ \to J/\psi\,K^+$ and $B^0 \to J/\psi\,K^*$ simulated samples, while the number of events in data, extracted from the fit to the B^+ invariant mass shown in Figure 4.1, is $(324 \pm 2) \times 10^2$.

The tagging efficiency is then reduced from 20% obtained at the preselection step to about 4%. However, according to the simulation, the fraction of muons associated to the decay of the opposite side B hadron is increased to about 80%, the vast majority of them carrying the correct tag decision. The resulting fractions obtained from the true Monte Carlo informations are summarized in Table 4.2.

Similarly, a specific selection is evaluated for the electrons, spanning an independent grid of $\mathcal{O}\left(10^6\right)$ cut configurations, where the variables involved are analogous to the ones used for the muon selection: the electron transverse momentum p_T, and pseudorapidity η, electron's track 3-dimensional impact parameter d_{xyz}, the separation from the fully reconstructed B meson ΔR (B), and the isolation criterion evaluated with the particle

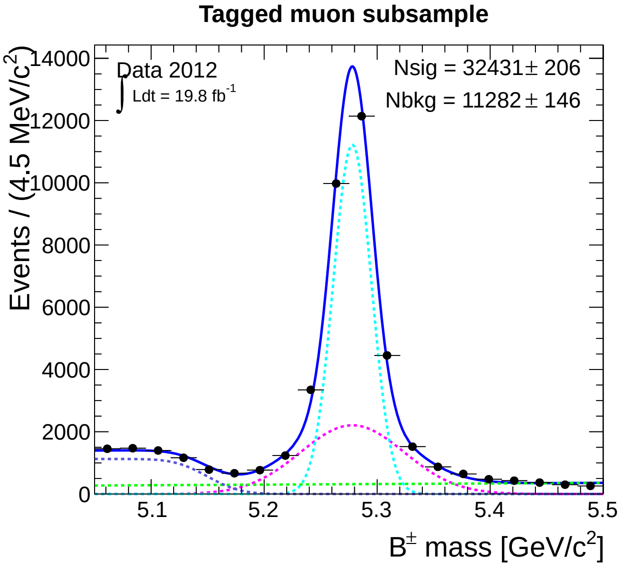


Figure 4.1: Fit to the B^+ invariant mass for the reconstructed $B^+ \to J/\psi K^+$ events containing a selected OS-muon in 2012 data. The overall fit result is represented with the blue solid line, while the PDFs of the single fit components are shown with dashed lines: in light blue and in magenta the two Gaussian functions describing the signal B^+ peak, the first order Chebyshev polynomial in green and the error function in violet. The total yields of both signal (N_{sig}) and background (N_{bkg}) events are reported in figure.

Table 4.2: Composition of the OS-muon sample at the selection stage in the three simulated samples.

	$B^0_s \to J/\psi\,\phi$ MC	$B^+ \to J/\psi\,K^+$ MC	$B^0 \to J/\psi\,K^*$ MC
N_{ev} reco-B	608 843	521 855	469 172
N_{ev} reco-B + OS-μ sel	22 981 (3.8%)	20 252 (3.9%)	18 620 (4.0%)
$B \to \mu$(CC)	60.1%	60.8%	60.4%
$B \to \mu$(WC)	18.5%	18.2%	18.2%
μ not from B (RC)	21.4%	21.0%	21.4%

flow algorithm using a ΔR cone of 0.4. In addition to the previously listed variables, the electron identification variable PFmva is exploited in the selection grid. The final electron selection cuts, optimized for the best $\mathcal{P}_{tag}$ require:

- electron transverse momentum $p_{\mathrm{T}} > 2.0$ GeV/c
- electron impact parameter $\mathrm{d}_{\mathrm{xyz}} < 0.1$ cm
- electron separation from the reco-B meson $\Delta\mathrm{R(B)} > 0.2$
- electron PFmva(e-π) > 0.2

After the selection is applied 20 381, 17 789 and 16 297 events are left respectively in the $\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\,\phi$, $\mathrm{B^+} \rightarrow \mathrm{J}/\psi\,\mathrm{K^+}$ and $\mathrm{B^0} \rightarrow \mathrm{J}/\psi\,\mathrm{K^*}$ simulated samples, corresponding to a tagging efficiency of about 3.5%. The number of events in data, extracted from the fit $\mathrm{B^+}$ invariant mass shown in Figure 4.2, is $(272 \pm 2) \times 10^2$.

The electron selection is able to reduce the fraction of uncorrelated background of a factor 1/2, from 80% to about 40%, as shown in Table 4.3. The fraction of electrons associated to the opposite side B hadron decay after the selection is applied is thus lower (about 45%) with respect to the corresponding fraction of muon. This effect, however not investigated in detail in this analysis, it is justified by the more difficult reconstruction and identification conditions of low p_{T} electrons with respect to low p_{T} muons. Given the high amount of material in the CMS tracker the electron background mainly arise from the decays of pions produced in a typical LHC collision, while the muon background is limited due to the great distance of the muon chambers with respect to the tracking system, and only a smaller contamination of sail-through is possible. A similar effect has already been studied in fact by CMS in the development of an identification algorithm for jets coming from b-quark (called b-tagging) exploiting the presence of muons and electrons within the jets [61].

The tagging performances obtained at this stage of the analysis are compared in Tables 4.4 and 4.5 for muons and electrons respectively.

Efficiencies, mistag probabilities and tagging powers of the three simulated samples are compared and are found to be approximately consistent within their statistical uncertainties. At this stage, the $\mathrm{B^0} \rightarrow \mathrm{J}/\psi\,\mathrm{K^*}$ sample is characterized by a slightly higher tagging efficiency with respect to the other simulated samples, however compensated by a worse mistag fraction, leading to an overall agreement at the level of $\lesssim 2.5\,\sigma$ for the measured tagging power. It is in fact important to observe that the larger values of the tagging efficiencies ε_{tag} measured on data are compensated by worse mistag fractions ω, so that in the end the tagging powers $\mathcal{P}_{tag}$ are found to be consistent with those obtained in the simulations.

With the lepton selection applied a better control over the flavour-uncorrelated background is obtained, thus increasing the tagging performances with respect to the preselection step, previously summarized in Table 4.1. Sizeable improvements of the tagging performances can be obtained my means of a multivariate technique, as described in the following Sections.

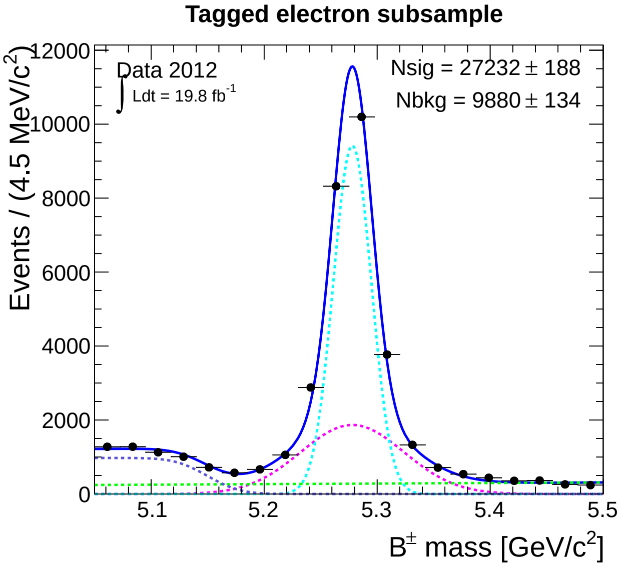


Figure 4.2: Fit to the B^+ invariant mass for the reconstructed $B^+ \rightarrow J/\psi K^+$ events containing a selected OS-electron in 2012 data. The overall fit result is represented with the blue solid line, while the PDFs of the single fit components are shown with dashed lines: in light blue and in magenta the two Gaussian functions describing the signal B^+ peak, the first order Chebyshev polynomial in green and the error function in violet. The total yields of both signal (N_{sig}) and background (N_{bkg}) events are reported in figure.

Table 4.3: Composition of the OS-electron sample at the selection stage in the three simulated samples.

	$B_s^0 \rightarrow J/\psi\,\phi$ MC	$B^+ \rightarrow J/\psi\,K^+$ MC	$B^0 \rightarrow J/\psi\,K^*$ MC
N_{ev} reco-B	608 843	521 855	469 172
N_{ev} reco-B + OS-e sel	20 381 (3.4%)	17 789 (3.4%)	16 297 (3.5%)
$B \rightarrow e$(CC)	43.6%	43.2%	43.2%
$B \rightarrow e$(WC)	13.6%	14.3%	14.1%
e not from B (RC)	42.8%	42.5%	42.7%

Table 4.4: OS-muon tagging performances evaluated on the $B_s^0 \to J/\psi\,\phi$, $B^+ \to J/\psi\,K^+$, and $B^0 \to J/\psi\,K^*$ simulations, and on the $B^+ \to J/\psi\,K^+$ channel of the 2012 data.

[%]	$B_s^0 \to J/\psi\,\phi$ MC	$B^+ \to J/\psi\,K^+$ MC	$B^0 \to J/\psi\,K^*$ MC	$B^+ \to J/\psi\,K^+$ data
ε_{tag}	3.77 ± 0.03	3.88 ± 0.03	3.97 ± 0.03	4.59 ± 0.03
ω	29.1 ± 0.3	28.3 ± 0.3	28.8 ± 0.3	31.0 ± 0.4
$\mathcal{P}_{tag}$	0.66 ± 0.02	0.73 ± 0.02	0.72 ± 0.02	0.67 ± 0.03

Table 4.5: OS-electron tagging performances evaluated on the $B_s^0 \to J/\psi\,\phi$, $B^+ \to J/\psi\,K^+$, and $B^0 \to J/\psi\,K^*$ simulations, and on the $B^+ \to J/\psi\,K^+$ channel of the 2012 data.

[%]	$B_s^0 \to J/\psi\,\phi$ MC	$B^+ \to J/\psi\,K^+$ MC	$B^0 \to J/\psi\,K^*$ MC	$B^+ \to J/\psi\,K^+$ data
ε_{tag}	3.35 ± 0.02	3.41 ± 0.03	3.47 ± 0.03	3.86 ± 0.03
ω	34.3 ± 0.3	34.8 ± 0.4	34.1 ± 0.4	35.5 ± 0.4
$\mathcal{P}_{tag}$	0.33 ± 0.01	0.32 ± 0.02	0.35 ± 0.02	0.33 ± 0.02

4.2.3 Multivariate Discriminator

To enhance the separation between leptons with correct tag information and those contributing to the mistag, several discriminating variables are combined using two independent neural networks, trained and optimized independently for muons and electrons. The multivariate tool is applied on the $B_s^0 \to J/\psi\,\phi$ simulated sample, the events used for the training being the ones for which a lepton (μ or e) is passing the selection described in the previous section.

4.2.3.1 Neural Network

A Neural Network (NN) multivariate technique can be defined as a simulated collection of interconnected neurons, with each neuron producing a certain response at a given set of input signals. The NN used in this study is the Multi Layer Perceptron, the recommended neural network of the TMVA package [62]. External signal applied to some input neurons will put the network into a defined state that can be measured from the response of output neurons. The neural network can therefore be viewed as a mapping from a space of input variables onto a one-dimensional space of an output variable. The behaviour of an artificial neural network is determined by the layout of the neurons, the weights of the inter-neuron connections, and by the response of the neurons to the input, described by the neuron response function. While in principle a network with n neurons can have n^2 directional connections, the complexity can be reduced by organising the neurons in layers and only allowing direct connections from a given layer to the following layer. The first layer of a multilayer perceptron is the input layer, the last one the output layer, and all others are hidden layers.

For a classification problem (e.g. in the case of signal-background separation) with n input variables, the input layer consists of n neurons that hold the input values, and one neuron in the output layer that holds the output variable, the neural network estimator output (from now on called MLP). The neuron response function maps the neuron input (from one or more neurons) onto the neuron output. This function can be generally separated into two components: the synapse function, which take care of the $\mathcal{R}^n \mapsto \mathcal{R}$

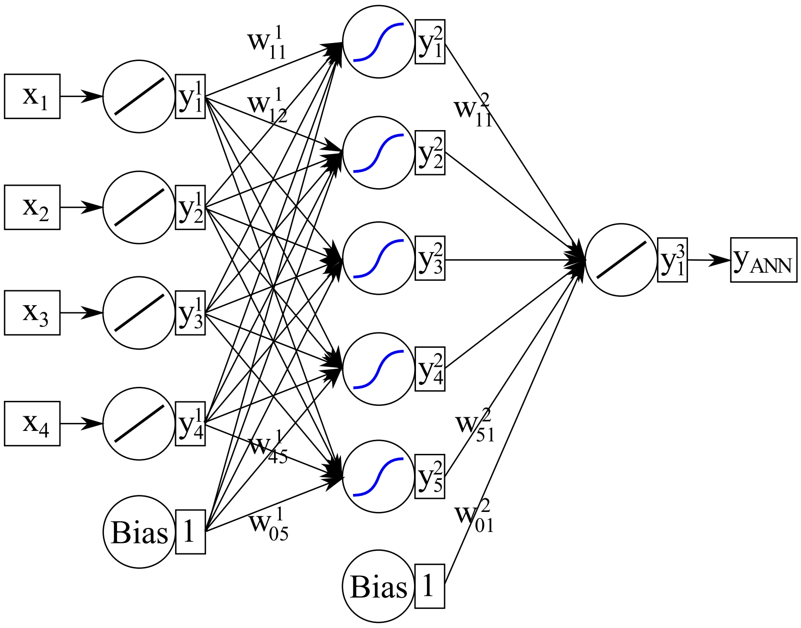


Figure 4.3: Sketch of a multilayer perceptron with four input variables ($x_1, \ldots, x_4$) and one hidden layer with five neurons. The neural network maps to an output variable y_{ANN}.

mapping starting from the values and weights of each input neuron; and the neuron activation function, which translates the result of the synapse function into the actual response of the neuron my means of $\mathcal{R} \mapsto \mathcal{R}$ functions (e.g. linear or sigmoid-like functions).

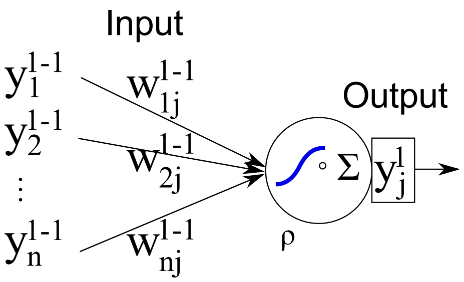


Figure 4.4: Structure of a neural network neuron. The inputs coming from the previous layer of neurons are combined according to the neuron response function (ρ).

The number of hidden layers in a network and the number of neurons in these layers determines the structure of the NN: they are fully configurable and need to be optimized for the best performances. Among the advantages of the neural networks lie their ability to represent both linear and non-linear relationships and their ability to learn these relationships directly from the data being modelled. The Multi Layer Perceptron learns in fact using back-propagation algorithms belonging to the family of supervised learning methods, where the desired output for every input event is known. In a first step, called the forward pass, the predicted outputs corresponding to the given inputs are evaluated thanks to the structure of the network. For each training event the neural network output is thus computed and compared to the desired output (e.g. 0 for background and 1 for signal). The difference between the two values can be described as functions (error functions) of the weights of the network's neuron. The second phase involves a backward pass through the network during which the partial derivatives of the error functions are minimized with respect to the various parameters and propagated back through the network. The network weights can then be adapted using a gradient-based optimisation algorithm. The whole process is iterated until the weights converge.

Several independent training tests were run in order to define the combination of input variables and network configuration parameters with the best signal vs background separation. The final configuration of the NN input variables is in fact chosen in order to maximize the integral of the area under the Receiver Operating Characteristic (ROC) curve. The ROC curve is generally presented as the background rejection versus signal efficiency for the test sample used in the classification procedure with the given NN output.

4.2.3.2 Variables Included in the NN

Several variables are involved in the process of optimization of the neural network inputs. Some of the variables are exploited for their ability to separate leptons arising from the decay of a B- vs $\overline{\mathrm{B}}$-hadron. Other quantities are mostly able to provide separation between leptons produced by the decay of B-hadrons from those arising from other sources, such as light hadron decays or misidentified particles, which are a source of flavour-uncorrelated (RC) background. Thanks to the ability of the multilayer perceptron to take into account the correlations among the input variables, the combination of all these information within the neural network is able to increase the discriminating power between leptons carrying the correct tag decision (CC) and those contributing to the mistag (WC).

Among the list of variables studied, the most relevant for the definition of the multivariate discriminators are listed in this section.

- p_{T}: transverse momentum
- η: pseudorapidity

These variables, related to the kinematics of the lepton, do not represent an effective source for the flavour identification, but are mostly helpful in the identification of RC background. Misidentified leptons (like, for instance kaons or pions) are in fact more easily produced in the forward region (for high values of pseudorapidity) and are characterized by lower values of transverse momenta compared to $\mathrm{b} \to \ell\,\mathrm{X}$ leptons.

Several versions of the lepton impact parameter (and their significances) are studied and compared during the optimization of the NN:

- $\mathrm{d_{xy}}$: 2D impact parameter with respect to the primary vertex of the event (PV)
- $\mathrm{S_{xy}}$: 2D impact parameter significance ($\mathrm{d_{xy}}/\sigma(\mathrm{d_{xy}})$)
- $\mathrm{d_{xyz}}$: 3D impact parameter with respect to the PV
- $\mathrm{S_{xyz}}$: 3D impact parameter significance ($\mathrm{d_{xyz}}/\sigma(\mathrm{d_{xyz}})$)

The impact parameter (also IP, for short) of a lepton is defined as the distance of closest approach of its inner track with respect to the primary vertex of the event in 3-dimensions

d_{xyz}, or its 2-dimensional projection in the transverse plane $r - \phi$ of the detector d_{xy}. Electrons and muons produced in the decay of a B hadron are subjected to the relatively long B lifetime, and are thus more likely to have a high value of impact parameter with respect to leptons in the decay of charmed hadrons. On the other hand, kaons and pions decaying in flight far from the primary vertex of the event can result into leptons with very large values of impact parameter, up to a several mm. Moreover, opposite side neutral B^0 meson are subject to flavour mixing, which results in a time-dependent probability of the B^0 flavour to oscillate between the production and decay time. Muons and electrons coming from a B^0 decay with small impact parameter are thus more likely to carry the correct flavour tag decision than leptons with high values of IP. For this reasons the variables related to the impact parameter of the leptons are largely important both in the classification of the uncorrelated backgrounds and in the flavour identification.

- ΔR: angular distance between the lepton and the reconstructed B meson

This variable acts as a veto-information of the particles involved in the same side hadronization and decay. As it was already discussed in Section 3.2.2, the properties of the same side of the event can be exploited in order to define the flavour of the reconstructed decay like in the case of the B^0_s. It is worth to be noted in fact that from the hadronization of a B^0_s meson a kaon is produced, its charge defining the flavour of the B^0_s at production time: K^+ tags the same side B^0_s and vice versa. In the case a K^+ produced from the hadronization of the same side B^0_s meson is misidentified as the lepton tagging the opposite side B hadron, the charge-flavour relation will result into a wrong OS-tag decision. The vast majority of such events is however already removed by the selection requirements, which include a ΔR cut both for electrons and muons.

An electron identification variable:

- PFmva: electron identification criterion.

This variable is computed combining in a multivariate discriminator several electron variables (such as the Bremsstrahlung deposit) and is optimized to discriminate electrons from background sources, mostly pions.

In addition to these variables, which are directly connected to the lepton and the reconstructed B meson, other variables sensitive to the features of the rest of the event are considered for this study.

- PFIsolation: isolation of the lepton from the surrounding detector's activity; computed from charged particles belonging to the same PV as the lepton, neutral hadrons, and photons, all contained within a $\Delta R < 0.4$ cone around the muon direction:

$$\text{PFIsolation} = \frac{\sum \mathrm{E_T}\,(\text{chHad from PV}) + \sum \mathrm{E_T}\,(\text{neutHad}) + \sum \mathrm{E_T}\,(\text{photons})}{p_{\mathrm{T}}^{\ell}}. \tag{4.1}$$

Muons and electrons produced in the decay of light hadrons are more likely to be surrounded by the particles contained in the hadronization of the light-quark jet than leptons originated from the decay of heavy-quarks. Therefore leptons arising from the $b \to \ell X$ decays are characterized by lower values of PFIsolation with respect to $b \to c X \to \ell X'$ decays and leptons produced from background sources.

A variable widely used in various analyses involving the discrimination of direct ($b \to \ell X$) and sequential ($b \to c X \to \ell X'$) heavy flavour decays is also investigated:

- $p_{\mathrm{T}}^{\mathrm{rel}}$: momentum relative to the axis of the jet associated to the lepton

This variable is defined as $p_{\mathrm{T}}^{\mathrm{rel}} = p_{\ell} sin\theta$, where $cos\theta = \frac{\vec{p}_{\ell} \cdot (\vec{p}_{jet} - \vec{p}_{\ell})}{p_{\ell} |(\vec{p}_{jet} - \vec{p}_{\ell})|}$. This quantity can be defined only in the case the lepton is associated with a well-identified jet. Jets are selected according to the Loose CMS particle flow jet-identification prescriptions[1] [44]. In order to reduce the residual pollution of jets due to the hadronization and fragmentation of the reconstructed (same-side) B meson, a minimum angular distance of $\Delta R > 0.5$ between the jet and the reconstructed B hadron flight direction is also required. The $p_{\mathrm{T}}^{\mathrm{rel}}$ is equal to the projection of the muon momentum in the direction perpendicular to the direction of the jet momentum. It reflects the momentum of the muon in the B-hadron rest frame, and is widely used to separate $b \to \ell X$ from sequential decays and other background processes. Due to the larger mass of the b-quark, muons from the direct decays populate higher values of $p_{\mathrm{T}}^{\mathrm{rel}}$, while leptons from lighter quarks are mostly characterized by lower values of $p_{\mathrm{T}}^{\mathrm{rel}}$. To increase the discriminating power, the momentum of the lepton is subtracted from the jet momentum.

In a b-hadron production, the algebraic sum of the charge of all the tracks involved in the hadronization and decay provides information about the flavour of the b-quark flavour at production time. A variable directly relate to this quantity is exploited in the neural network:

- The *lepton charge cone* (Q_{ℓ}), an estimator of the charge associated to the OS-b parton, described by the relation:

$$Q_{\ell}\left(\Delta R, k, \ell - \mathrm{in/out}\right) = \frac{\sum_{i}^{N_{\mathrm{tracks}}} q^{i} \cdot \left(p_{\mathrm{T}}^{i}\right)^{k}}{\sum_{i}^{N_{\mathrm{tracks}}} \left(p_{\mathrm{T}}^{i}\right)^{k}}, \tag{4.2}$$

where the sum runs over all the charged particle flow candidates that can be reconstructed within a ΔR cone from the lepton, and q^{i} is the charge of the i-th track in the sum. The tracks are selected with the minimal purity requirements: $p_{\mathrm{T}}^{i} > 500$ MeV, $|\eta|^{i} < 2.5$ and at least 5 tracker hits; . Tracks associated to the reconstructed B decay (same side) are explicitly excluded from the sum. If no track is found satisfying the listed requirements, the variable can not be defined and it is therefore not employed in the neural network. The Q_{ℓ} variable definition depends on three parameters: the width of the angular ΔR cone; the relevance of low p_{T} tracks in the Q_{ℓ} definition, that is enhanced or reduced by adjusting the k exponent; and the possibility to include (ℓ-in) or exclude (ℓ-out) the lepton inner track in the sum. Several definitions have been considered and have been applied during the optimization of the tagger.

Only a fraction of the selected muons are included in a reconstructed jet passing the Loose-jet identification criteria, and therefore can be used to define the $p_{\mathrm{T}}^{\mathrm{rel}}$ variable. Similarly, the fraction of selected leptons with a defined value of the Q_{ℓ} variable depends on the parameters ΔR and with the inclusion (or exclusion) of the inner track of the lepton

---

[1] Neutral Hadron Fraction < 0.99, Neutral EM Fraction < 0.99, Number of Constituents > 1, and, in addition if the jet is reconstructed in the $|\eta| < 2.4$ region, Charged Hadron Fraction > 0, Charged Multiplicity > 0, Charged EM Fraction < 0.99

in the sum. In principle, by including the lepton in the sum, the Q_ℓ (ℓ-in) variable could be expected to be defined for all the selected leptons and for any given size of the cone ΔR. However, the minimum number of hits required for a track to be included in the Q_ℓ definition happens to be not always fulfilled by the selected muon or electron, for which the identification requirements have been discussed in the previous section.

Tables 4.6, 4.7 and 4.8 report the number of events (N_{ev}) containing a reconstructed muon associated to a jet and with four definitions of Q_μ. The relative fraction of such events with respect to the total number of events containing the fully reconstructed $B \to J/\psi X$ decays are reported, as well as the relative efficiencies with respect to the previously discussed muon's selection cuts. Similarly, Tables 4.9, 4.10 and 4.11 report the analogous quantities for the electrons selected in the $B_s^0 \to J/\psi\,\phi$, $B^+ \to J/\psi\,K^+$, and $B^0 \to J/\psi\,K^*$ MC samples.

4.2.3.3 Training of the NN

Two independent training procedures are applied for electrons and muons separately. The neural network is trained with a classification procedure requiring to separate leptons correctly tagging the flavour of the reconstructed-B meson by their charge (signal for the NN), from leptons carrying the wrong tag information (background). Leptons from uncorrelated sources (RC in the Monte Carlo classification) are included in the training of the NN according to the relation between their charge and with the flavour of the reco-B meson. The NN is trained using the events of the $B_s^0 \to J/\psi\,\phi$ Monte Carlo sample containing an OS-leptons, for a total amount of about 24 000 events for the muon sample and 20 400 for the electron tagger. Half of each sample (randomly divided) is exploited to train the neural network, and the remaining events are then used to test the performance and the stability of the multilayer perceptron.

As previously stated, the choice of variables to be included into the multilayer perceptron as well as the list of the main parameters involved in the neural network configuration is part of the optimization process. The final network configuration is obtained by performing several independent training and tests in order to define the combination of input variables with the best signal vs background separation, described by the highest value of the integral of the ROC curve. The optimization proceeds independently for muons and electrons in order to optimise the two network architectures in a completely independent fashion.

The choice of the input variables proceed as follows: a kernel of inputs composed by the two kinetic variables, the transverse momentum and pseudorapidity, is considered as the starting point of the optimisation procedure. In a first step, each of the variables discussed in Section 4.2.3 is added to the NN and a new training is performed. Promising variables for which an increase in the value of the ROC integral is observed are selected for a second step of optimization, while variables not contributing to the signal-background separation are discarded. Several definitions of the charge cone variables have been studied independently, spanning the three parameters ruling their definitions. The most discriminating Q_ℓ definition is different for muons and electrons and will be discussed separately in the following paragraphs; however the inclusion of the lepton track within the charge cone ($Q_\ell(\ell - in)$) proved to perform as a better right vs wrong-tag discriminator that the one with the lepton removed from the cone. After a list of a number

Table 4.6: Number of selected events (and relative efficiencies) for which is possible to define the muon $p_{\mathrm{T}}^{\mathrm{rel}}$ and Q_{μ} variables, in the $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ MC.

Selection requirements	$\mathrm{N_{events}}$	$\varepsilon_{\text{reco-B}}$	$\varepsilon_{\text{OS-}\mu}$
Reconstructed $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$	608 843		
OS muon selection	22 981	3.77%	
Jet	20 482	3.36%	89.13%
Q_{μ} ($\Delta\mathrm{R} = 0.3, \mu$-in)	22 971	3.77%	99.96%
Q_{μ} ($\Delta\mathrm{R} = 0.5, \mu$-in)	22 981	3.77%	100%
Q_{μ} ($\Delta\mathrm{R} = 0.3, \mu$-out)	20 864	3.43%	90.97%
Q_{μ} ($\Delta\mathrm{R} = 0.5, \mu$-out)	22 774	3.74%	99.10%

Table 4.7: Number of selected events (and relative efficiencies) for which is possible to define the muon $p_{\mathrm{T}}^{\mathrm{rel}}$ and Q_{μ} variables, in the $\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$ MC.

Selection requirements	$\mathrm{N_{events}}$	$\varepsilon_{\text{reco-B}}$	$\varepsilon_{\text{OS-}\mu}$
Reconstructed $\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$	521 855		
OS muon selection	20 252	3.88%	
Jet	17 842	3.42%	88.10%
Q_{μ} ($\Delta\mathrm{R} = 0.3, \mu$-in)	20 237	3.88%	99.93%
Q_{μ} ($\Delta\mathrm{R} = 0.5, \mu$-in)	20 250	3.88%	99.99%
Q_{μ} ($\Delta\mathrm{R} = 0.3, \mu$-out)	18 509	3.55%	91.39%
Q_{μ} ($\Delta\mathrm{R} = 0.5, \mu$-out)	20 090	3.85%	99.20%

Table 4.8: Number of selected events (and relative efficiencies) for which is possible to define the muon $p_{\mathrm{T}}^{\mathrm{rel}}$ and Q_{μ} variables, in the $\mathrm{B^0} \to \mathrm{J}/\psi\,\mathrm{K^*}$ MC.

Selection requirements	$\mathrm{N_{events}}$	$\varepsilon_{\text{reco-B}}$	$\varepsilon_{\text{OS-}\mu}$
Reconstructed $\mathrm{B^0} \to \mathrm{J}/\psi\,\mathrm{K^*}$	469 172		
OS muon selection	18 620	3.97%	
Jet	16 648	3.55%	89.41%
Q_{μ} ($\Delta\mathrm{R} = 0.3, \mu$-in)	18 612	3.97%	99.96%
Q_{μ} ($\Delta\mathrm{R} = 0.5, \mu$-in)	18 619	3.97%	99.99%
Q_{μ} ($\Delta\mathrm{R} = 0.3, \mu$-out)	16 981	3.62%	91.20%
Q_{μ} ($\Delta\mathrm{R} = 0.5, \mu$-out)	18 434	3.93%	99.00%

Table 4.9: Number of selected events (and relative efficiencies) for which is possible to define the electron $p_{\mathrm{T}}^{\mathrm{rel}}$ and Q_e variables, in the $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ MC.

Selection requirements	$\mathrm{N_{events}}$	$\varepsilon_{\text{reco-B}}$	$\varepsilon_{\text{OS-}e}$
Reconstructed $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$	608 843		
OS electron selection	20 381	3.35%	
Jet	836	0.14%	4.10%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.3, e\text{-in})$	20 376	3.35%	99.98%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.5, e\text{-in})$	20 380	3.35%	100%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.3, e\text{-out})$	18 907	3.11%	92.77%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.5, e\text{-out})$	20 228	3.32%	99.25%

Table 4.10: Number of selected events (and relative efficiencies) for which is possible to define the electron $p_{\mathrm{T}}^{\mathrm{rel}}$ and Q_e variables, in the $\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$ MC.

Selection requirements	$\mathrm{N_{events}}$	$\varepsilon_{\text{reco-B}}$	$\varepsilon_{\text{OS-}e}$
Reconstructed $\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$	521 855		
OS electron selection	17 789	3.41%	
Jet	783	0.15%	4.40%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.3, e\text{-in})$	17 782	3.41%	99.96%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.5, e\text{-in})$	17 785	3.41%	99.98%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.3, e\text{-out})$	16 499	3.16%	92.75%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.5, e\text{-out})$	17 664	3.38%	99.30%

Table 4.11: Number of selected events (and relative efficiencies) for which is possible to define the electron $p_{\mathrm{T}}^{\mathrm{rel}}$ and Q_e variables, in the $\mathrm{B^0} \to \mathrm{J}/\psi\,\mathrm{K^*}$ MC.

Selection requirements	$\mathrm{N_{events}}$	$\varepsilon_{\text{reco-B}}$	$\varepsilon_{\text{OS-}e}$
Reconstructed $\mathrm{B^0} \to \mathrm{J}/\psi\,\mathrm{K^*}$	469 172		
OS electron selection	16 297	3.47%	
Jet	655	0.14%	4.02%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.3, e\text{-in})$	16 293	3.47%	99.98%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.5, e\text{-in})$	16 297	3.47%	100%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.3, e\text{-out})$	15 126	3.22%	92.81%
$\mathrm{Q}_e\,(\Delta\mathrm{R} = 0.5, e\text{-out})$	16 184	3.45%	99.31%

of meaningful variables is selected (p_{T}, η, IP, PFmva, ΔR, isolation, Q_{ℓ}-in, Q_{ℓ}-out, $p_{\mathrm{T}}^{\mathrm{rel}}$), a second step of optimization is performed. Only one of the four available definitions of the impact parameter variables (d_{xy}, S_{xy}, $\mathrm{d}_{\mathrm{xyz}}$, $\mathrm{S}_{\mathrm{xyz}}$) is kept, since the combination of multiple definitions proved not to increase the separation of the two training categories, while contributing to the steady increase of the correlation among the input variables. During the second step of optimization, all the interesting variables spanned all the possible combinations, and the configuration corresponding to the best value of the ROC integral is chosen to be the final set of input variables for the NN.

An analogous procedure was used in order to obtain the best multilayer perceptron configuration parameters, such as the hidden layer structure, the type of neuron activation function, etc.

Muon NN

The optimal choice of input variables for the muons multilayer perceptron is found to be the following:

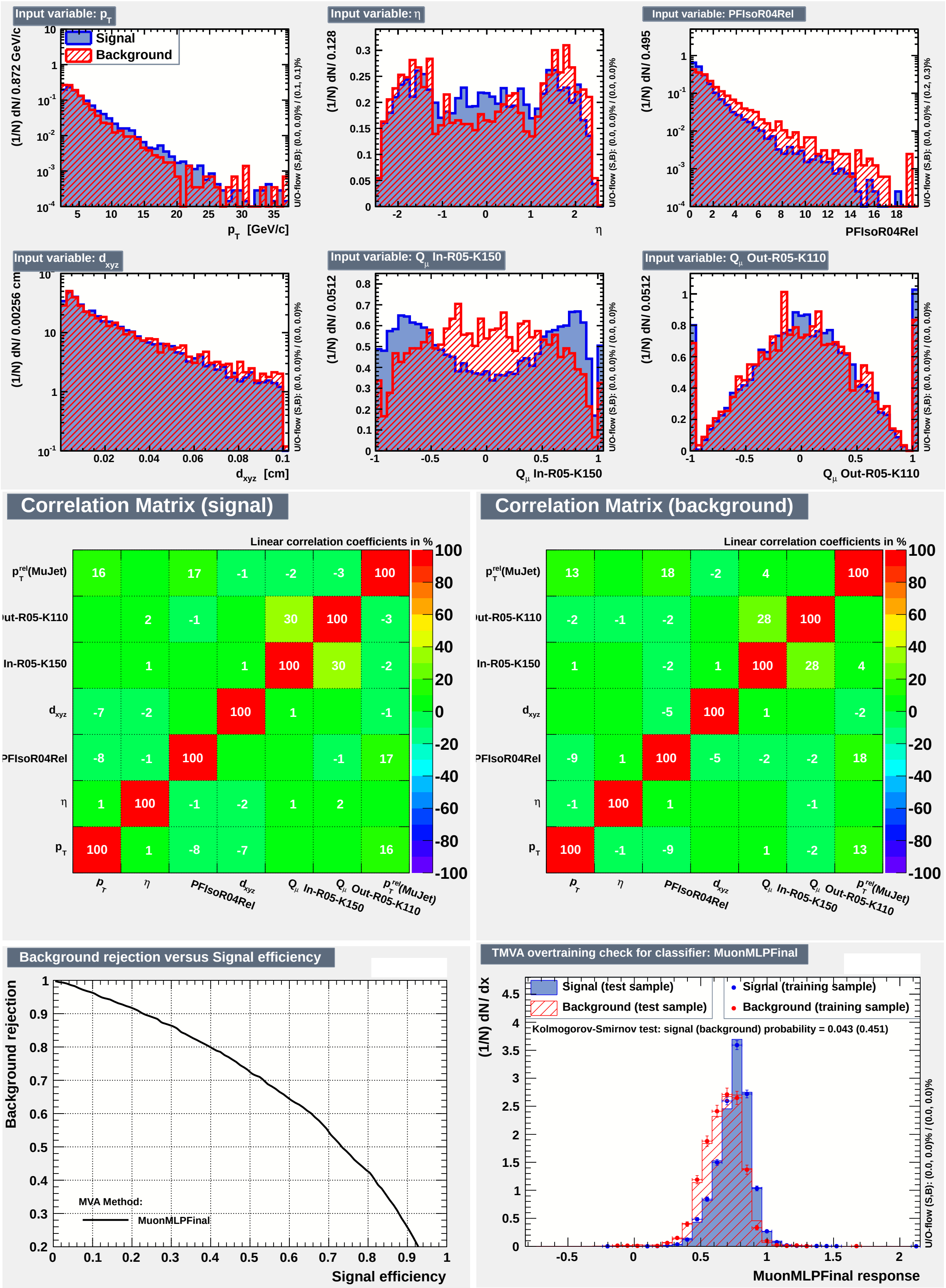
- muon p_{T}
- muon η
- muon PFIsolation
- muon $\mathrm{d}_{\mathrm{xyz}}$
- muon Q_{μ} (ΔR =0.5,k=1.50,μ-in)
- muon Q_{μ} (ΔR =0.5,k=1.10,μ-out)
- muon $p_{\mathrm{T}}^{\mathrm{rel}}$

The most discriminating definition of the Q_{μ} variable is found to be the one corresponding to a cone opening of $\Delta \mathrm{R} < 0.5$, a k-exponent $= 1.5$ and the muon track included in the sum. The inclusion of the second definition Q_{μ} (ΔR =0.5,k=1.10,μ-out), despite the low discriminating power provided by the single variable, allowed to further increase the overall separating power still maintaining a the correlation small ($\lesssim 30\%$). It is worth to note that since the $p_{\mathrm{T}}^{\mathrm{rel}}$ variable is defined only in the case of events where the muon is reconstructed as a constituent of a jet (i.e. $\approx 90\%$ of the events after muon selection), two categories are defined in the training of the MLP: in case of events where muons are not associated to a jet the $p_{\mathrm{T}}^{\mathrm{rel}}$ variable is not employed in the NN; otherwise the entire list of variables is used.

The final integral of the ROC curve is calculated by the TMVA to be 0.667. The distributions of the variables involved in the NN training and the comparison of these variables in the MC and data samples are collected in the Appendix A.1.

In Figure 4.5 the input variables and the standard TMVA control plots are shown. The correlation among the input variables is less than 30%, and no indication of overtraining can be observed.

The final optimized choice of parameters for the muon multilayer perceptron is preselected in Table 4.12.



Correlation Matrix (signal)

	p_T	η	PFIsoR04Rel	d_{xyz}	Q_μ In-R05-K150	Q_μ Out-R05-K110	p_T^{rel}(MuJet)
p_T^{rel}(MuJet)	16		17	-1	-2	-3	100
ut-R05-K110		2	-1		30	100	-3
In-R05-K150		1		1	100	30	-2
d_{xyz}	-7	-2		100	1		-1
PFIsoR04Rel	-8	-1	100			-1	17
η	1	100	-1	-2	1	2	
p_T	100	1	-8	-7			16

Correlation Matrix (background)

	p_T	η	PFIsoR04Rel	d_{xyz}	Q_μ In-R05-K150	Q_μ Out-R05-K110	p_T^{rel}(MuJet)
p_T^{rel}(MuJet)	13		18	-2	4		100
ut-R05-K110	-2	-1	-2		28	100	
In-R05-K150	1		-2	1	100	28	4
d_{xyz}			-5	100	1		-2
PFIsoR04Rel	-9	1	100	-5	-2	-2	18
η	-1	100	1			-1	
p_T	100	-1	-9		1	-2	13

Figure 4.5: TMVA control plots for the OS-muon NN showing: (first row) the input variables used in the training of the Multi Layer Perceptron, for signal (blue) and background (red); (second row) the correlation among the input variables for signal (left) and background (right) independently; (third row) the normalized output of the MLP$_\mu$ response for signal (blue) and background (red) in the training sample (markers) and in the test sample (histograms), and the Receiver Operating Characteristic (ROC) curve.

Table 4.12: Main parameters used in the training of the OS-muon tagger multilayer perceptron.

Parameter	Value
Number of cycles	750
Hidden Layers	N + 8
Neuron Type	tanh
Neuron Input Type	sum
Learning Rate	2×10^{-2}
Decay Rate	1×10^{-2}
Test Rate	5

Electron NN

The final choice of input variables for the electron NN is the following:

- p_{T}
- η
- PFIsolation
- $\mathrm{d}_{\mathrm{xyz}}$
- PFmva
- Q_e (ΔR =0.3,k=1.75,e-in)

Only one definition of the Q_e variable is used, corresponding to a cone opening of $\Delta\mathrm{R} < 0.3$, a k-exponent $= 1.75$ and the electron track included in the sum. Due to the very low jet association efficiency, the $p_{\mathrm{T}}^{\mathrm{rel}}$ variable is not exploited in the neural network.

The final integral of the ROC curve is found to be 0.629. Plots representing the distributions of all the variables involved in the NN training and the comparison of these variables in the MC and data samples are collected in the Appendix A.2.

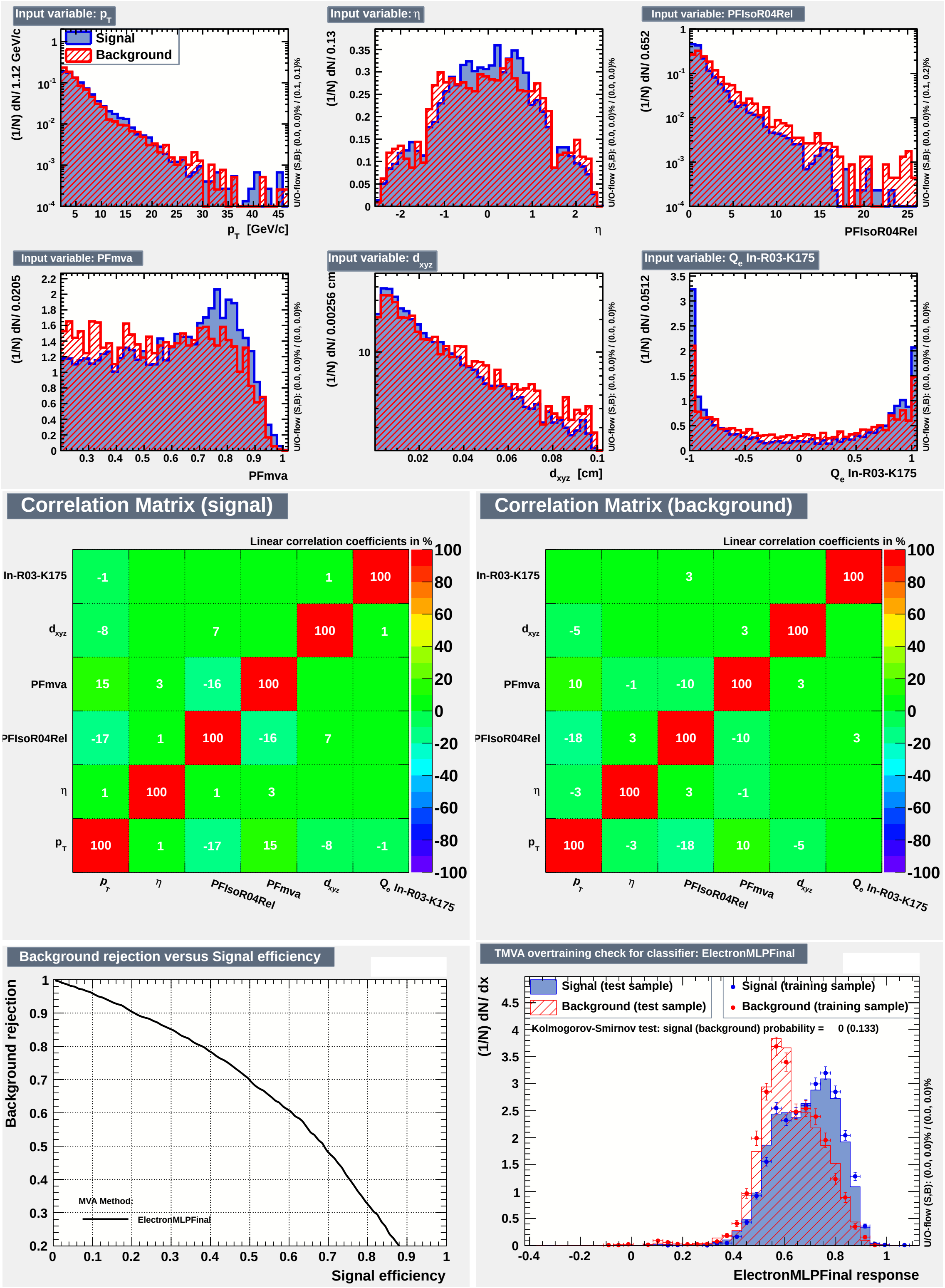
In Figure 4.6 the input variables and the standard TMVA control plots are shown. The correlation among the input variables is less than 20%, and no indication of overtraining can be observed.

The full list of parameter optimized for the electron NN is shown in Table 4.13.

4.3 Lepton Tagger Performances

The multilayer perceptron discriminator (MLP) can be exploited to develop a tagging algorithm using three different (and independent) approaches:

1. defining a correct-tag enriched region via a single cut on the discriminator
2. defining a number of non overlapping categories of the neural network discriminator; the tagging performances can then be evaluated independently for each category and then combined according to Equation 3.4



Correlation Matrix (signal)

	p_T	η	PFIsoR04Rel	PFmva	d_{xyz}	Q_e In-R03-K175
In-R03-K175	-1				1	100
d_{xyz}	-8		7		100	1
PFmva	15	3	-16	100		
PFIsoR04Rel	-17	1	100	-16	7	
η	1	100	1	3		
p_T	100	1	-17	15	-8	-1

Correlation Matrix (background)

	p_T	η	PFIsoR04Rel	PFmva	d_{xyz}	Q_e In-R03-K175
In-R03-K175			3			100
d_{xyz}	-5			3	100	
PFmva	10	-1	-10	100	3	
PFIsoR04Rel	-18	3	100	-10		3
η	-3	100	3	-1		
p_T	100	-3	-18	10	-5	

Figure 4.6: TMVA control plots for the OS-electron NN showing: (first row) the input variables used in the training of the Multi Layer Perceptron, for signal (blue) and background (red); (second row) the correlation among the input variables for signal (left) and background (right) independently; (third row) the normalized output of the MLP$_e$ response for signal (blue) and background (red) in the training sample (markers) and in the test sample (histograms) , and the Receiver Operating Characteristic (ROC) curve.

Table 4.13: Main parameters used in the training of the OS-electron tagger multilayer perceptron.

Parameter	Value
Number of Cycles	600
Hidden Layers	N + 5
Neuron Type	tanh
Neuron Input Type	sum
Learning Rate	2×10^{-2}
Decay Rate	1×10^{-2}
Test Rate	5

3. obtaining a per-event mistag probability evaluating ω as function of the MLP response with an analytic parametrization.

The results obtained for each of the three approaches will be discussed in the following paragraphs.

4.3.1 Single Cut

The most simple approach to the tagging is to exploit the MLP discriminator in order to remove from the sample the tagging leptons which are classified by the neural network to be predominantly contributing to the mistag. The cut on the two MLP discriminators is optimized independently for muons and electrons on the $\mathrm{B_s^0} \rightarrow \mathrm{J}/\psi\,\phi$ MC for the best tagging power. The tagging efficiency is hence reduced by the MLP cut.

The resulting optimal threshold for the muon discriminator is found to be $\mathrm{MLP}_\mu > 0.64$, which corresponds to a reduction of about 25% of events with respect to the opposite side muon selection. Incidentally, the independently optimized electron cut results to be $\mathrm{MLP}_e > 0.64$. By applying this cut the electron tagging efficiency is reduced by 43% with respect to the selection step, described in Section 4.2.2.

The tagging performances measured using the described approach are presented in Tables 4.14 and 4.15. Despite the lower tagging efficiency, the single-cut approach dramatically improves the average mistag value in the selected samples. The tagging power evaluated applying this approach on the 2012 data on the $\mathrm{B^+} \rightarrow \mathrm{J}/\psi\,\mathrm{K^+}$ self-tagging channel is in fact increased from 0.67%(0.35%), obtained in Section 4.2.2 after the lepton selection, to the values of about 0.78% (0.43%) for muons (electrons). The agreement among the MC samples is found to be excellent within the statistical uncertainties.

4.3.2 Categorization

As already discussed in Chapter 3, the tagging performances can be further improved without reducing the number of available events by segmenting the samples in a number of non overlapping categories. To show the benefits arising from this approach, the muon and electron samples are both divided in four categories. These categories are defined on the basis of the value of the MLP discriminators, in order to have a roughly equivalent number of tagged events in each class. For every category the tagging efficiency and

Table 4.14: OS-muon tagging performances evaluated on the $B_s^0 \to J/\psi\,\phi$, $B^+ \to J/\psi\,K^+$ and $B^0 \to J/\psi\,K^*$ simulated samples, and on the $B^+ \to J/\psi\,K^+$ channel of the 2012 data. The optimized cut on the MLP_μ discriminator $MLP_\mu > 0.64$ is applied.

[%]	$B_s^0 \to J/\psi\,\phi$ MC	$B^+ \to J/\psi\,K^+$ MC	$B^0 \to J/\psi\,K^*$ MC	$B^+ \to J/\psi\,K^+$ data
MLP_μ	> 0.64			
ε_{tag}	2.81 ± 0.02	2.86 ± 0.02	2.92 ± 0.03	3.34 ± 0.02
ω	23.1 ± 0.3	23.5 ± 0.4	23.5 ± 0.4	25.9 ± 0.4
$\mathcal{P}_{tag}$	0.82 ± 0.02	0.80 ± 0.02	0.82 ± 0.02	0.78 ± 0.03

Table 4.15: OS-electron tagging performances evaluated on the $B_s^0 \to J/\psi\,\phi$, $B^+ \to J/\psi\,K^+$ and $B^0 \to J/\psi\,K^*$ simulated samples, and on the $B^+ \to J/\psi\,K^+$ channel of the 2012 data. The optimized cut on the MLP_e discriminator $MLP_e > 0.64$ is applied.

[%]	$B_s^0 \to J/\psi\,\phi$ MC	$B^+ \to J/\psi\,K^+$ MC	$B^0 \to J/\psi\,K^*$ MC	$B^+ \to J/\psi\,K^+$ data
MLP_e	> 0.64			
ε_{tag}	1.90 ± 0.02	1.91 ± 0.02	1.96 ± 0.02	2.11 ± 0.02
ω	25.5 ± 0.4	26.4 ± 0.5	25.5 ± 0.5	27.5 ± 0.5
$\mathcal{P}_{tag}$	0.45 ± 0.02	0.42 ± 0.02	0.47 ± 0.02	0.43 ± 0.02

the wrong tag fraction are evaluated independently. The overall tagging power is then obtained as the sum of the tagging power obtained in each MLP category, as discussed in Chapter 3.

Muon Tagger

The tagging power evaluated on the $B^+ \to J/\psi\,K^+$ channel of the 2012 data, presented in Table 4.16, is increased to 0.83%, with a relative improvement of more than 6% with respect to the simple cut-based approach. The performances measured on the $B_s^0 \to J/\psi\,\phi$, $B^+ \to J/\psi\,K^+$ and $B^0 \to J/\psi\,K^*$ simulated samples are reported in Tables 4.17, 4.18, 4.19 and are found to be in good agreement.

Electron Tagger

The tagging power evaluated on the $B^+ \to J/\psi\,K^+$ channel of the 2012 data, presented in Table 4.20, is increased to 0.47%, with a relative improvement of more than 9% with respect to the simple cut-based approach. The performances measured on the $B_s^0 \to J/\psi\,\phi$, $B^+ \to J/\psi\,K^+$ and $B^0 \to J/\psi\,K^*$ simulated samples are reported in Tables 4.21, 4.22, 4.23 and are found to be in good agreement.

4.3.3 Per-Event Mistag

Extending the categorization approach, an unbinned analytic parametrization of the mistag fraction can be exploited in order to evaluate the wrong tag probability event per event. Each event can be in fact considered as an independent "category" for which a specific value of mistag should be defined. The mistag is parametrized as a function of the output of the multilayer perceptron. As already described in the previous sections, the NN

Table 4.16: OS-muon tagging performances evaluated on the reconstructed $B^+ \to J/\psi\, K^+$ events 2012 data. Four categories of MLP_μ discriminator are used.

[%]	$B^+ \to J/\psi\, K^+$ data			
MLP_μ	$[-0.05, 0.6405)$	$[0.6405, 0.7405)$	$[0.7405, 0.815)$	$[0.815, 1.35)$
ε_{tag}	1.219 ± 0.014	1.177 ± 0.014	1.13 ± 0.013	1.027 ± 0.013
ω	44.9 ± 0.9	33.3 ± 0.8	24.3 ± 0.7	19.3 ± 0.6
$\mathcal{P}_{tag}$	0.013 ± 0.005	0.132 ± 0.013	0.298 ± 0.017	0.387 ± 0.018
$\varepsilon_{tag}^{\text{tot}}$	4.55 ± 0.03			
$\mathcal{P}_{tag}^{\text{tot}}$	0.83 ± 0.03			

Table 4.17: OS-muon tagging performances evaluated on the $B_s^0 \to J/\psi\, \phi$ simulated sample. Four categories of MLP_μ discriminator are used.

[%]	$B_s^0 \to J/\psi\, \phi$ MC			
MLP_μ	$[-0.05, 0.6405)$	$[0.6405, 0.7405)$	$[0.7405, 0.815)$	$[0.815, 1.35)$
ε_{tag}	0.932 ± 0.012	0.938 ± 0.013	0.936 ± 0.013	0.929 ± 0.012
ω	47.2 ± 0.7	30.7 ± 0.6	22.8 ± 0.6	15.5 ± 0.5
$\mathcal{P}_{tag}$	0.003 ± 0.001	0.139 ± 0.009	0.277 ± 0.012	0.442 ± 0.014
$\varepsilon_{tag}^{\text{tot}}$	3.74 ± 0.03			
$\mathcal{P}_{tag}^{\text{tot}}$	0.86 ± 0.02			

Table 4.18: OS-muon tagging performances evaluated on the $B^+ \to J/\psi\, K^+$ simulated sample. Four categories of MLP_μ discriminator are used.

[%]	$B^+ \to J/\psi\, K^+$ MC			
MLP_μ	$[-0.05, 0.6405)$	$[0.6405, 0.7405)$	$[0.7405, 0.815)$	$[0.815, 1.35)$
ε_{tag}	0.991 ± 0.014	0.944 ± 0.014	0.952 ± 0.014	0.951 ± 0.014
ω	42.4 ± 0.7	30.7 ± 0.7	22.9 ± 0.6	16.9 ± 0.6
$\mathcal{P}_{tag}$	0.023 ± 0.004	0.141 ± 0.010	0.28 ± 0.013	0.418 ± 0.015
$\varepsilon_{tag}^{\text{tot}}$	3.84 ± 0.03			
$\mathcal{P}_{tag}^{\text{tot}}$	0.86 ± 0.02			

Table 4.19: OS-muon tagging performances evaluated on the $B^0 \to J/\psi\, K^*$ simulated sample. Four categories of MLP_μ discriminator are used.

[%]	$B^0 \to J/\psi\, K^*$ MC			
MLP_μ	$[-0.05, 0.6405)$	$[0.6405, 0.7405)$	$[0.7405, 0.815)$	$[0.815, 1.35)$
ε_{tag}	1.004 ± 0.015	0.936 ± 0.014	0.987 ± 0.015	0.994 ± 0.015
ω	44.2 ± 0.7	31.1 ± 0.7	22.0 ± 0.6	17.9 ± 0.6
$\mathcal{P}_{tag}$	0.013 ± 0.003	0.133 ± 0.01	0.309 ± 0.015	0.41 ± 0.016
$\varepsilon_{tag}^{\text{tot}}$	3.92 ± 0.02			
$\mathcal{P}_{tag}^{\text{tot}}$	0.87 ± 0.02			

Table 4.20: OS-electron tagging performances evaluated on the reconstructed $B^+ \to J/\psi K^+$ events 2012 data. Four categories of MLP_e discriminator are used.

[%]	$B^+ \to J/\psi K^+$ data			
MLP_e	$[0, 0.573)$	$[0.573, 0.6675)$	$[0.6675, 0.761)$	$[0.761, 1.1)$
ε_{tag}	1.003 ± 0.013	1.018 ± 0.013	0.978 ± 0.012	0.871 ± 0.012
ω	47.7 ± 1.1	40.4 ± 0.9	30.2 ± 0.8	21.7 ± 0.7
$\mathcal{P}_{tag}$	0.002 ± 0.002	0.038 ± 0.008	0.154 ± 0.014	0.279 ± 0.016
$\varepsilon_{tag}^{\text{tot}}$	3.87 ± 0.03			
$\mathcal{P}_{tag}^{\text{tot}}$	0.47 ± 0.02			

Table 4.21: OS-electron tagging performances evaluated on the $B_s^0 \to J/\psi \phi$ simulated sample. Four categories of MLP_e discriminator are used.

[%]	$B_s^0 \to J/\psi \phi$ MC			
MLP_e	$[0, 0.573)$	$[0.573, 0.6675)$	$[0.6675, 0.761)$	$[0.761, 1.1)$
ε_{tag}	0.838 ± 0.012	0.839 ± 0.012	0.835 ± 0.012	0.832 ± 0.012
ω	47.7 ± 0.7	40.2 ± 0.7	29.4 ± 0.7	19.6 ± 0.6
$\mathcal{P}_{tag}$	0.002 ± 0.001	0.032 ± 0.005	0.141 ± 0.009	0.307 ± 0.012
$\varepsilon_{tag}^{\text{tot}}$	3.34 ± 0.02			
$\mathcal{P}_{tag}^{\text{tot}}$	0.48 ± 0.02			

Table 4.22: OS-electron tagging performances evaluated on the $B^+ \to J/\psi K^+$ simulated sample. Four categories of MLP_e discriminator are used.

[%]	$B^+ \to J/\psi K^+$ MC			
MLP_e	$[0, 0.573)$	$[0.573, 0.6675)$	$[0.6675, 0.761)$	$[0.761, 1.1)$
ε_{tag}	0.876 ± 0.013	0.871 ± 0.013	0.831 ± 0.013	0.827 ± 0.013
ω	48.7 ± 0.8	39.9 ± 0.7	30.6 ± 0.7	18.8 ± 0.6
$\mathcal{P}_{tag}$	0.001 ± 0.001	0.036 ± 0.005	0.125 ± 0.009	0.322 ± 0.014
$\varepsilon_{tag}^{\text{tot}}$	3.41 ± 0.03			
$\mathcal{P}_{tag}^{\text{tot}}$	0.48 ± 0.02			

Table 4.23: OS-electron tagging performances evaluated on the $B^0 \to J/\psi K^*$ simulated sample. Four categories of MLP_e discriminator are used.

[%]	$B^0 \to J/\psi K^*$ MC			
MLP_e	$[0, 0.573)$	$[0.573, 0.6675)$	$[0.6675, 0.761)$	$[0.761, 1.1)$
ε_{tag}	0.856 ± 0.014	0.892 ± 0.014	0.879 ± 0.014	0.842 ± 0.014
ω	47.7 ± 0.8	40.3 ± 0.8	28.5 ± 0.7	19.7 ± 0.7
$\mathcal{P}_{tag}$	0.002 ± 0.001	0.033 ± 0.005	0.163 ± 0.011	0.309 ± 0.014
$\varepsilon_{tag}^{\text{tot}}$	3.47 ± 0.03			
$\mathcal{P}_{tag}^{\text{tot}}$	0.51 ± 0.02			

discriminator provides a separation from leptons carrying the right and reversed tag decisions, and can therefore be used in order to define regions with low levels of ω as well as regions characterized by high levels of mistag. Given the strategy pursued for the training of the neural networks, leptons with low values of the MLP output are more likely to contribute to the mistag than to the correct tag attribution, i.e. to have a mistag probability greater than 50%. Such leptons can effectively be used for the purpose of B flavour tagging. In fact, if a lepton's mistag probability exceed the 50% value, for instance $\omega = 0.55$, the tag decision of the lepton (d) should be simply reversed ($d' = -d$). The corresponding mistag must then be evaluated as $\omega' = 1 - \omega$, now corresponding in the previous example to a value of $\omega' = 0.45$. It is also worth to be noted that thanks to the formula defining the tagging power (see Equation 3.1), the two conditions are completely equivalent and would determine the same value of $\mathcal{P}_{tag}$.

In order to determine the function to be used to parametrize the per event mistag, the average wrong tag fraction is evaluated (separately on data and MC) in non-overlapping intervals of the MLP discriminator. The number of MLP bins involved in this study and the resulting boundaries and chosen in order to have comparable tagging efficiency for each division and to provide a sufficient number of events for the $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ fit on data.

Muon Tagger

Twenty MLP_μ bins of variable width are used to evaluate the average values of mistag and efficiency per category in the range $-0.05 < \mathrm{MLP}_\mu < 1.35$. Figure 4.7 shows the wrong tag fraction evaluated in twenty categories of MLP_μ for the three MC samples and for the $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ events reconstructed in 2012 data. From the comparison of the $\mathrm{B}^0_\mathrm{s} \to \mathrm{J}/\psi\,\phi$, $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$, and $\mathrm{B}^0 \to \mathrm{J}/\psi\,\mathrm{K}^*$ MC it can be observed a certain level of disagreement among the various simulated samples. The disagreement in the binned values of mistag ω^i has however a negligible effect on the final result. The tagging efficiencies ε^i_{tag} obtained for any given bin of MLP_μ are in fact slightly different among the three simulated samples: the discrepancies in ω^i and ε^i_{tag} compensate leading to a good bin-by-bin agreement among the tagging power values $\mathcal{P}^i_{tag}$ for the different MC samples, as shown in Figure 4.8. A systematic uncertainty dedicated to the tagger dependence on the specific flavour of the reconstructed B meson will however be considered and discussed in the following sections.

The mistag ω is interpolated with an analytic function. Several functions are tested, all requiring the mistag fraction to be positively defined in the whole MLP_μ range, and the one for which the mean of the residual distribution is found to be closer to zero is then chosen as the reference function. The family of functions that best represents the data is found to be a sigmoid parametrized with a complementary error function (Erf) of the form:

$$\omega\left(\mathrm{MLP}_\mu\right) = p_0 + p_1 \cdot \left[1 - \mathrm{Erf}\left(p_2 + p_3 \cdot \mathrm{MLP}_\mu\right)\right] \tag{4.3}$$

where MLP_μ is the neural network output, and the p_i terms are the four parameter of the function to be obtained by the fit of the analytic function to the binned data. The results

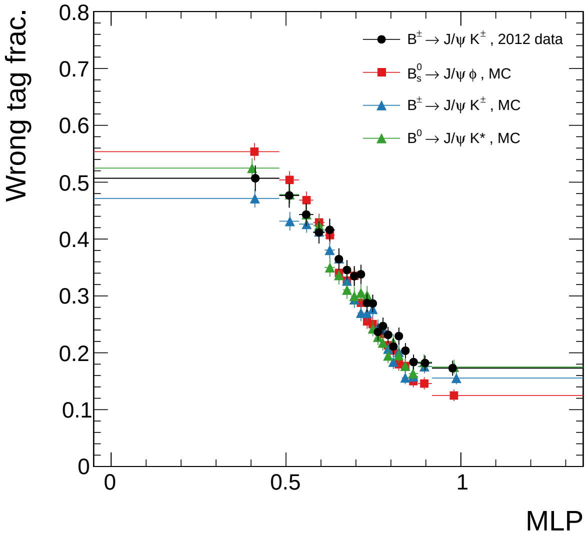


Figure 4.7: OS-muon wrong tag fraction evaluated in bins of the MLP_μ discriminator for the $B_s^0 \to J/\psi\,\phi$ (red), $B^+ \to J/\psi\,K^+$ (blue), and $B^0 \to J/\psi\,K^*$ (green) MC simulations, and in 2012 $B^+ \to J/\psi\,K^+$ data (black).

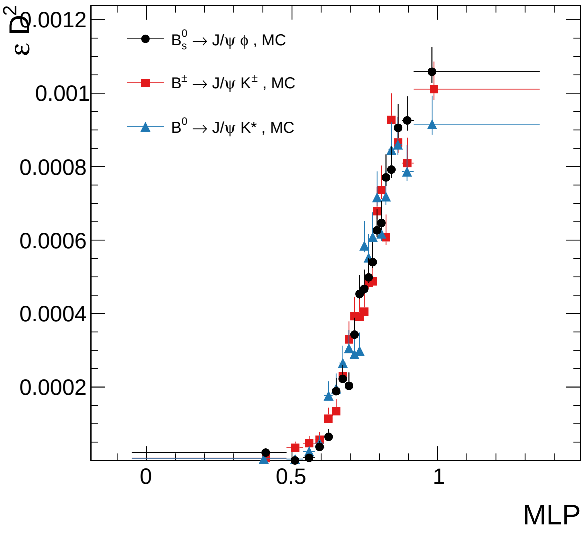


Figure 4.8: OS-muon tagging power evaluated in bins of the MLP_μ discriminator in $B_s^0 \to J/\psi\,\phi$ (black), $B^+ \to J/\psi\,K^+$ (red), and $B^0 \to J/\psi\,K^*$ (blue) MC simulations.

of the fit procedure are shown Figure 4.9 separately for the various MC and for the data sample.

The parameters of the error function evaluated by the fit to the binned ω distribution in the three MC samples and in 2012 $B^+ \to J/\psi\,K^+$ data are summarized in Table 4.24.

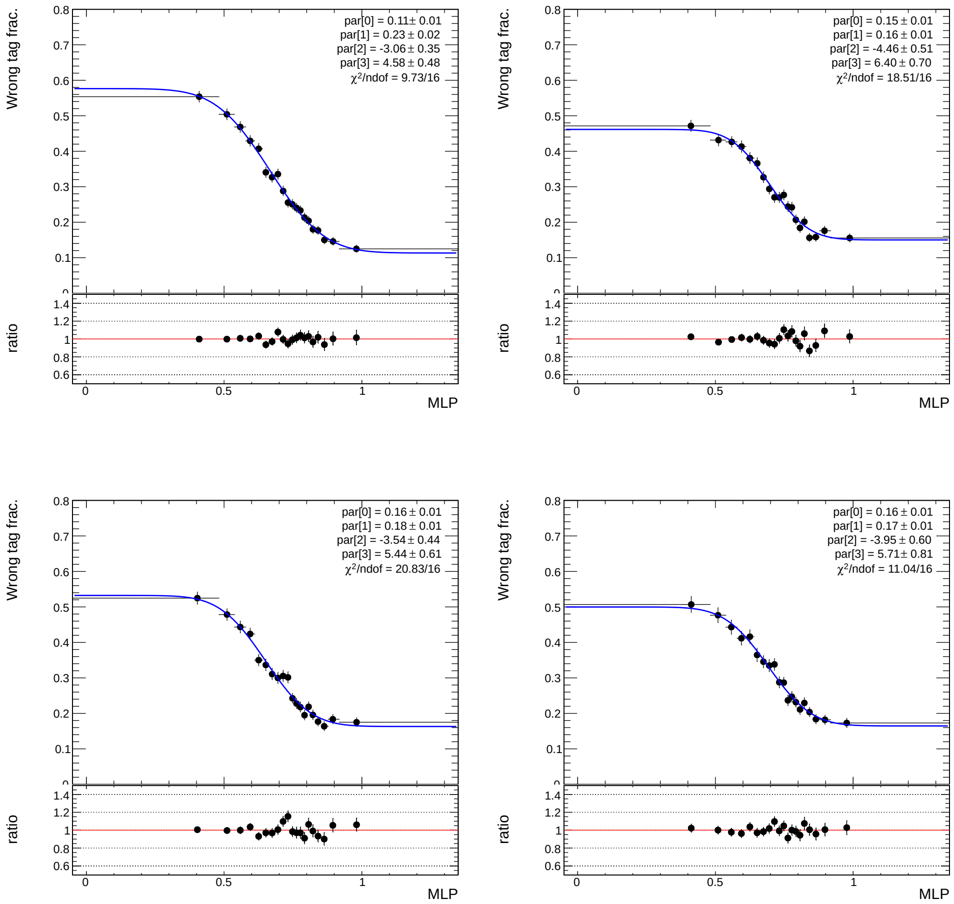


Figure 4.9: OS-muon wrong tag fraction evaluated as a function of the MLP_μ discriminator for the $B_s^0 \to J/\psi\,\phi$ MC (top-left), $B^+ \to J/\psi\,K^+$ MC (top-right), $B^0 \to J/\psi\,K^*$ MC (bottom-left), and 2012 $B^+ \to J/\psi\,K^+$ data (bottom-right). A function of the form $p_0 + p_1 \cdot \left[1 - \mathrm{Erf}\left(p_2 + p_3 \cdot \mathrm{MLP}_\mu\right)\right]$ is used for the fit.

Electron Tagger

An analogous procedure is carried out for electrons. The neural network discriminator is used to define twenty categories in the range $0.00 < \mathrm{MLP}_e < 1.10$.

Figure 4.10 compares the wrong tag fraction evaluated in twenty bins of the MLP_e discriminator for the three MC samples and for the $B^+ \to J/\psi\,K^+$ events reconstructed in 2012 data. An excellent agreement can be observed among the mistag measured for the $B_s^0 \to J/\psi\,\phi$, $B^+ \to J/\psi\,K^+$, and $B^0 \to J/\psi\,K^*$ simulated samples. The comparison of the tagging power for the different MC samples is shown in Figure 4.11.

Also in this case, data are best represented by the sigmoid defined by Eq. 4.3, but a different set of parameters (p_i) is obtained. Their values are reported in Table 4.25, and the parametrization curves are shown Figure 4.12 separately for the three simulated samples and for the 2012 data.

Table 4.24: Parameters resulting from the fit to the OS-muon ω distribution in the $\mathrm{B}_s^0 \rightarrow \mathrm{J}/\psi\,\phi$ MC, $\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$ MC, $\mathrm{B}^0 \rightarrow \mathrm{J}/\psi\,\mathrm{K}^*$ MC, and in 2012 $\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$ data. A function of the form $p_0 + p_1 \cdot \left[1 - \mathrm{Erf}\left(p_2 + p_3 \cdot \mathrm{MLP}_\mu\right)\right]$ is used.

parameters	$\mathrm{B}_s^0 \rightarrow \mathrm{J}/\psi\,\phi$ MC	$\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$ MC	$\mathrm{B}^0 \rightarrow \mathrm{J}/\psi\,\mathrm{K}^*$ MC	$\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$ data
p_0	0.113 ± 0.013	0.150 ± 0.009	0.163 ± 0.010	0.165 ± 0.012
p_1	0.232 ± 0.016	0.156 ± 0.009	0.185 ± 0.013	0.168 ± 0.014
p_2	-3.1 ± 0.4	-4.5 ± 0.5	-3.5 ± 0.4	-4.0 ± 0.6
p_3	4.6 ± 0.5	6.4 ± 0.7	5.4 ± 0.6	5.7 ± 0.8

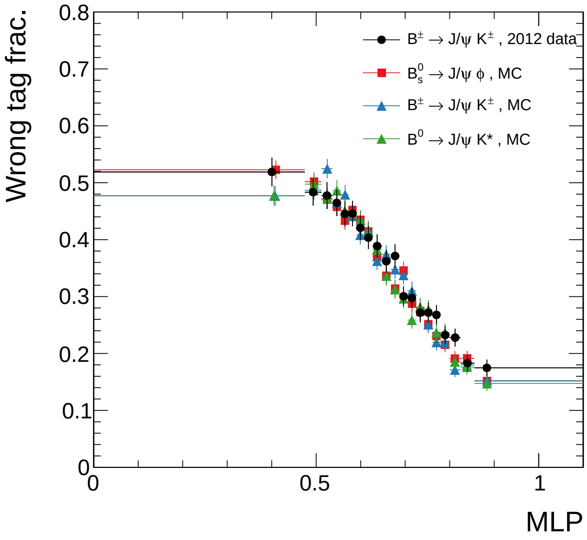


Figure 4.10: OS-electron wrong tag fraction evaluated in bins of the MLP_e discriminator for the $\mathrm{B}_s^0 \rightarrow \mathrm{J}/\psi\,\phi$ (red), $\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$ (blue), and $\mathrm{B}^0 \rightarrow \mathrm{J}/\psi\,\mathrm{K}^*$ (green) MC simulations, and in 2012 $\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$ data (black).

4.3.4 Single Particle Tagger Calibration

The predicted mistag evaluated from the parametrization function can be referred as the *calculated* mistag: $\omega^{\mathrm{calc}} = \omega(\mathrm{MLP})$. Using the self tagging $\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$ channel, for which the relation between the opposite lepton charge and the flavour of the reconstructed B meson is univocally defined by the charge of the kaon, the mistag can be directly extracted from data (*measured* mistag) ω^{meas} as a function of the calculated value ω^{calc}.

In absence of biases affecting the definition of the mistag function, a linear dependence between the measured and the calculated ω should be observed. The calibration function is defined as:

$$\omega^{\mathrm{meas}}(\omega^{\mathrm{calc}}) = p_0 + p_1 \times \left(\omega^{\mathrm{calc}} - \omega'\right) \tag{4.4}$$

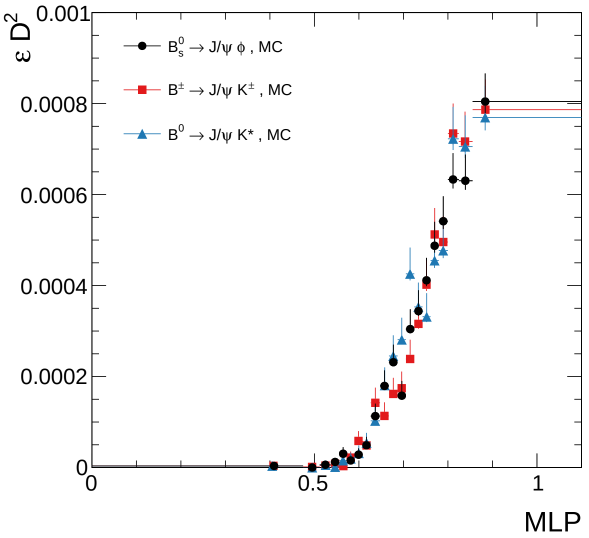


Figure 4.11: OS-electron tagging power evaluated in bins of the MLP$_e$ discriminator in $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ (black), $\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$ (red), and $\mathrm{B^0} \to \mathrm{J}/\psi\,\mathrm{K^*}$ (blue) MC simulations.

Table 4.25: Parameters resulting from the fit to the OS-electron ω distribution in the $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ MC, $\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$ MC, $\mathrm{B^0} \to \mathrm{J}/\psi\,\mathrm{K^*}$ MC, and in 2012 $\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$ data. A function of the form $p_0 + p_1 \cdot [1 - \mathrm{Erf}\,(p_2 + p_3 \cdot \mathrm{MLP}_e)]$ is used.

parameters	$\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ MC	$\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$ MC	$\mathrm{B^0} \to \mathrm{J}/\psi\,\mathrm{K^*}$ MC	$\mathrm{B^+} \to \mathrm{J}/\psi\,\mathrm{K^+}$ data
p_0	0.12 ± 0.03	0.13 ± 0.02	0.14 ± 0.02	0.14 ± 0.04
p_1	0.212 ± 0.018	0.184 ± 0.013	0.182 ± 0.016	0.20 ± 0.03
p_2	-3.4 ± 0.5	-4.4 ± 0.5	-4.1 ± 0.6	-3.3 ± 0.7
p_3	4.9 ± 0.8	6.3 ± 0.7	6.0 ± 0.8	4.8 ± 1.1

where the p_0 and p_1 calibration parameters are obtained by the fit, and $\omega' = 0.35$ is an arbitrary value, chosen to be approximately corresponding to the average of the calculated mistag range distribution.

The parametrization of Equation 4.4 is used instead of the usual first level polynomial parametrization ($p_0 + p_1 \times \omega^{\mathrm{calc}}$) in order to reduce the correlation between the parameters p_0 and p_1. Significant deviations from a linear dependence would indicate for instance the wrong choice of ω parametrization function. The the calibration curve is also exploited to evaluate the systematic uncertainties by comparing the mistag distributions of different samples.

Once the calibration curve is obtained, the mistag probability calculated according to the parametrization function should be corrected in order to fully reproduce the behaviour of the mistag measured on data. In order to extract p_0 and p_1, the measured mistag fraction is evaluated from fit to the $\mathrm{B^+}$ mass in twenty bins of ωcalc, and a fit to the resulting graph is performed.

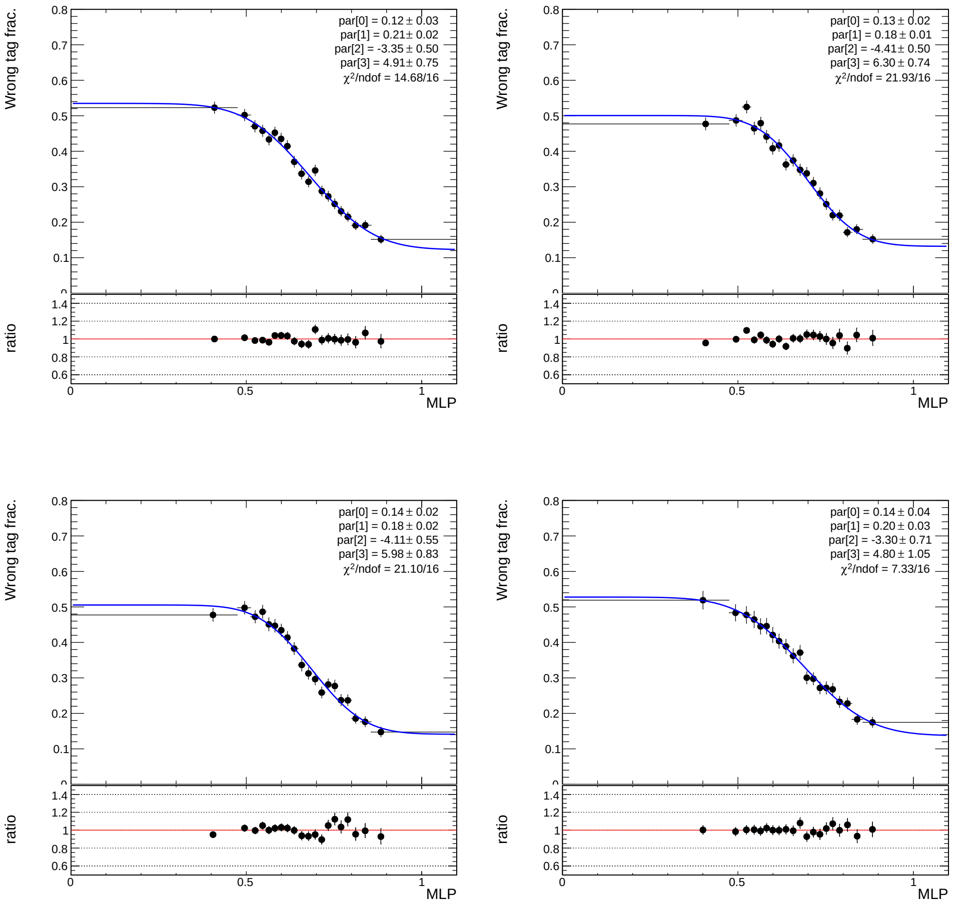


Figure 4.12: OS-electron wrong tag fraction evaluated as a function of the MLP_e discriminator for the $\mathrm{B}_s^0 \rightarrow \mathrm{J}/\psi\,\phi$ MC (top-left), $\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$ MC (top-right), $\mathrm{B}^0 \rightarrow \mathrm{J}/\psi\,\mathrm{K}^*$ MC (bottom-left), and 2012 $\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$ Data (bottom-right). A function of the form $p_0 + p_1 \cdot [1 - \mathrm{Erf}\,(p_2 + p_3 \cdot \mathrm{MLP}_e)]$ is used for the fit.

The calibration curves obtained for the muon and electron tagger are shown in Figure 4.13and the resulting calibration parameters are listed in Table 4.26. It can be observed that measured parameters are in an almost perfect agreement with the hypothesis for both of the single lepton taggers.

The defined mistag parametrization functions, corrected by the calibration curves just described, can be used to evaluate the tagging power and the wrong tag fraction for any given sample. For each tagged event of any given sample it is now possible to define the tag decision and the event-specific mistag probability. This allows the computation of the tagging power as the sum of the squared values of the per event dilutions:

$$\mathcal{P}_{tag} = \frac{1}{N_{tot}} \sum_{i}^{N} (1 - 2\omega_i)^2 \tag{4.5}$$

where we use the calibrated mistag value. The overall value of mistag associated to the

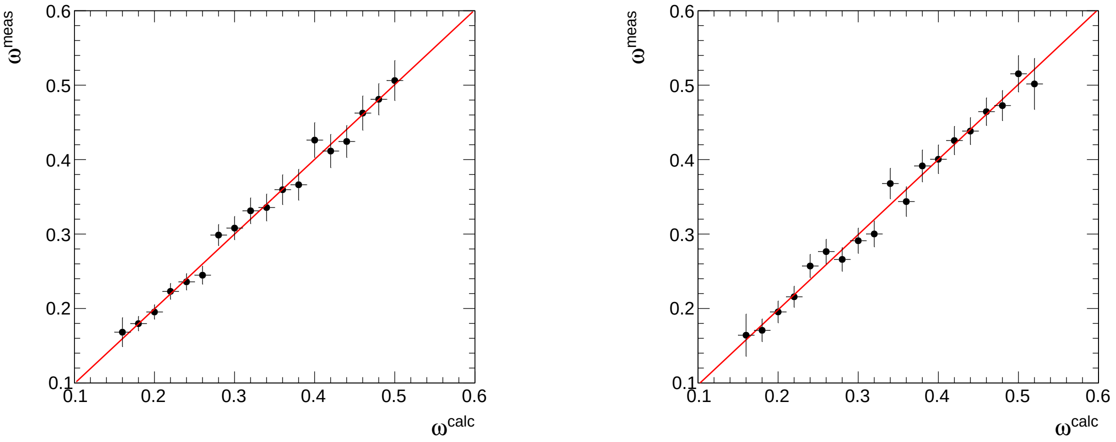


Figure 4.13: The measured wrong tag fraction extracted from fit to the B^+ mass in bins of the calculated mistag, obtained as function of the neural network discriminator $\omega = \omega(\text{MLP})$ for muons (left) and electrons (right).

Table 4.26: Calibration parameters of the OS-muon tagger. The uncertainties are statistical only.

	muon tagger	electron tagger
p_0	0.350 ± 0.004	0.350 ± 0.004
p_1	1.00 ± 0.04	1.01 ± 0.04
$\rho\,(p_0, p_1)$	0.642	0.317
$p_0 - p_1 \times \omega'$	-0.001 ± 0.011	-0.004 ± 0.013

tagging algorithm can thus be obtained from the total tagging power $\mathcal{P}_{tag}$ and the tagging efficiency ε_{tag}, the latter trivially defined as the fraction of tagged events in the sample, as:

$$\omega = \frac{1}{2}\left(1 - \sqrt{\frac{\mathcal{P}_{tag}}{\varepsilon_{tag}}}\right) \tag{4.6}$$

The calibration parameters evaluated on the B^+ data can then be used in order to estimate the statistical uncertainty on the values of ω and $\mathcal{P}_{tag}$, by propagating the uncertainties on p_0 and p_1 and their correlation, as described in detail within the Appendix B. The tagging power and mistag measured on the $B^+ \to J/\psi K^+$ data sample are summarized in Table 4.27, where the errors reported are only statistical. The statistical uncertainty is estimated in about 2% for the muon tagger and 4% for electrons.

Although it cannot be observed a clear improvement of the tagging power with respect to the binned approach, the parametrization of the per-event mistag is a reliable technique that does not rely on the estimation of a average value for each chosen interval of the NN discriminator, and it also results in the simplification of the flavour-tagged fit procedure to extract the CP violation parameters ϕ_s.

Table 4.27: Performances of the OS-muon tagger, using the per-event mistag curves. The uncertainties are statistical only.

	$B^+ \to J/\psi K^+$ data	
[%]	muon tagger	electron tagger
ε_{tag}^{tot}	4.56 ± 0.02	3.92 ± 0.02
ω^{tot}	28.6 ± 0.3	32.5 ± 0.4
$\mathcal{P}_{tag}^{tot}$	0.83 ± 0.02	0.48 ± 0.02

4.3.5 Consistency Checks and Systematic Uncertainties

Potential systematic uncertainties affecting the performances of the two taggers are studied by repeating the calibration procedure on the $B^+ \to J/\psi K^+$ channel on data for different conditions. The differences between the fitted calibration parameters and the reference ones for muons and electrons, summarized in Table 4.26, are propagated as the systematic uncertainties to the tagging power and mistag.

The various sources of systematic uncertainties are studied in detail the following paragraphs.

- choice of the parametrization function
- B^+ invariant mass fit model
- B^+ kinematics
- data taking period
- flavour of the reconstructed B meson

Choice of the Mistag Parametrization Function

As described above, in order to obtain the mistag function, a parametrization of the form $p_0 + p_1 \cdot [1 - \mathrm{Erf}\,(p_2 + p_3 \cdot \mathrm{MLP})]$ is chosen by selecting the functional form which minimizes the distribution of the fit residuals.

Several other functions have been tested, and similar compatibility to the binned data could be observed using other sigmoid function to describe the mistag parametrization. The residual dependence of the arbitrary choice of the selected sigmoid function can be tested comparing the resulting mistag functions.

Two fit functions are used for the purpose, defined similarly to the standard parametrization, as follows:

$$\omega\,(\mathrm{MLP}) = p_0 + p_1 \cdot [1 - \arctan\,(p_2 + p_3 \cdot \mathrm{MLP})] \tag{4.7}$$

$$\omega\,(\mathrm{MLP}) = p_0 + p_1 \cdot [1 - \tanh\,(p_2 + p_3 \cdot \mathrm{MLP})] \tag{4.8}$$

The fit obtained by using respectively the arctangent (arctan) and the hyperbolic tangent (tanh) parametrizations instead of the error function (Erf) formula can be found in the Appendix in Figs. A.12,A.13 for muons, and Figs. A.25 and A.26 for electrons.

The three tested sigmoid functions are compared and found to be in excellent agreement, as it can be observed by the comparison plot of Figure 4.14 and 4.15 respectively for the muon and electron tagger. For both the taggers the parametrization of the central region of the MLP discriminator range is consistently reproduced by all the investigated functions. Despite the excellent overall agreement, a localized disagreement between the various parametrizations can be observed in the extreme regions of the MLP distributions.

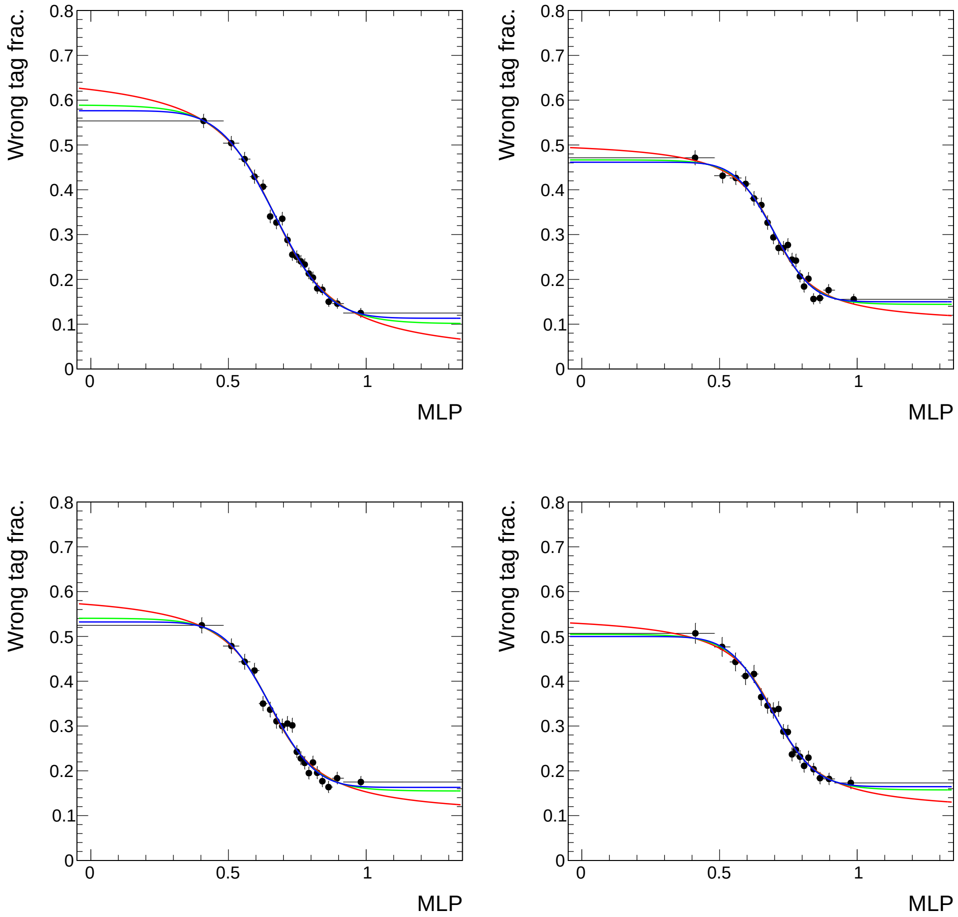


Figure 4.14: OS-muon wrong tag fraction evaluated as a function of the MLP$_{\mu}$ discriminator for the $B^0_s \rightarrow J/\psi\,\phi$ MC (top-left), $B^+ \rightarrow J/\psi\,K^+$ MC (top-right), $B^0 \rightarrow J/\psi\,K^*$ MC (bottom-left), and 2012 $B^+ \rightarrow J/\psi\,K^+$ data (bottom-right). Three resulting fit functions are overimposed in each figure: error function (blue), hyperbolic tangent (green), and arctangent (red).

To quantity the effect of this residual discrepancy on the function choice, the mistag and tagging power have been measured on data using the different parametrizations. The results have been compared with the legacy values. The maximum measured deviation induced on the tagging power is found to be of the order of 10^{-4}, way smaller than the statistical uncertainties (about 2% for the muon tagger and 4% for the electron tagger).

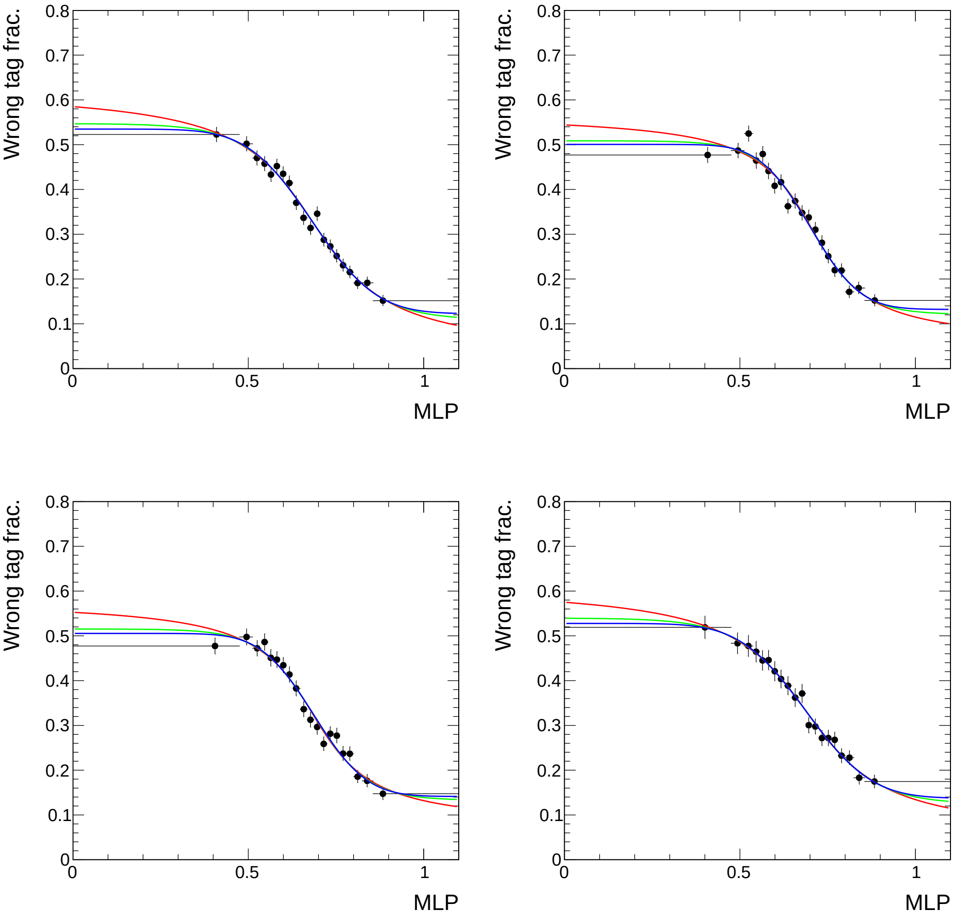


Figure 4.15: OS-electron wrong tag fraction evaluated as a function of the MLP$_e$ discriminator for the $B_s^0 \to J/\psi\,\phi$ MC (top-left), $B^+ \to J/\psi\,K^+$ MC (top-right), $B^0 \to J/\psi\,K^*$ MC (bottom-left), and 2012 $B^+ \to J/\psi\,K^+$ Data (bottom-right). Three resulting fit functions are overimposed in each figure: error function (blue), hyperbolic tangent (green), and arctangent (red).

B^+ Fit Model Dependence

The mistag is evaluated on data by extracting the number of signal events for each bin of the neural network discriminator through a fit the B^+ invariant mass distributions with the fit model described in Section 3.4.2.1. In order to test the potential dependence of the mistag function induced by the choice of the B^+ fit model, the signal and background PDFs are modified independently and the $B^+ \to J/\psi\,K^+$ data is fitted using the different PDFs hypothesis. The signal pdf is modified using a triple Gaussian with common mean, instead of a double Gaussian; the background pdf is modified using an error function convoluted with an exponential function, instead of the standard convolution of an error function and a first order Chebyshev polynomial.

For each MLP category used in order to define the mistag parametrization, the B^+ mass is then fitted using:

- standard signal + modified background
- modified signal + standard background

in order to extract the total, tagged, and wrongly-tagged number of events for each bin.

Figure 4.16 shows, for illustrative purpose only, the comparison of the fit to the B^+ mass using the three different PDFs, for the tagged events in a single bin of the muon MLP discriminator, corresponding to the interval $0.6395 < MLP_\mu < 0.6630$.

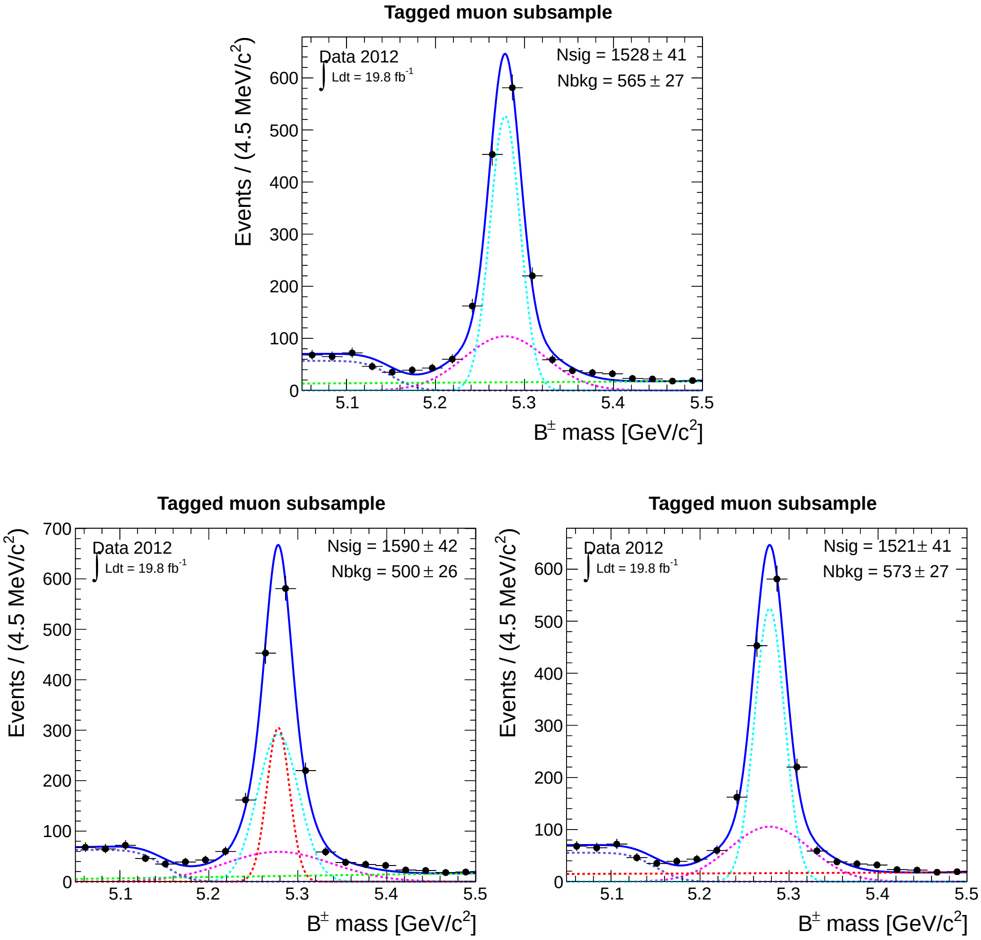


Figure 4.16: Fit to the B^+ mass performed for the tagged OS-muons in the $0.6395 < MLP < 0.6630$ range using the three different fit models: (top) standard fit model; (bottom left) standard background and modified signal model; (bottom right) standard signal and modified background model. In the top plot the overall fit is represented with the blue solid line, while the PDFs of the single fit components are shown with dashed lines: in light blue and in magenta the two Gaussian functions describing the signal B^+ peak, the first order Chebyshev polynomial in green and the error function in violet. In the bottom left plot, the signal pdf is described by the triple Gaussian in light blue, magenta and red. The background pdf of the bottom right plot is described by an error function in violet, and an exponential in red.

The mistag fractions evaluated using the two modified PDFs can be compared with the value measured fitting the B^+ invariant mass using the standard pdf (Eq. 3.13) parametrization for each of the MLP bins of the two taggers. Furthermore, the mistag distributions are fitted independently and the two mistag functions are extracted my means of the usual mistag parametrizations.

Although the number of signal and background events obtained by the B^+ mass fit differs reflecting the use of the three different PDFs, the values of the wrong tag fractions (as well as the tagging efficiencies) are found to be in excellent agreement, as shown in Figure 4.17 and 4.18.

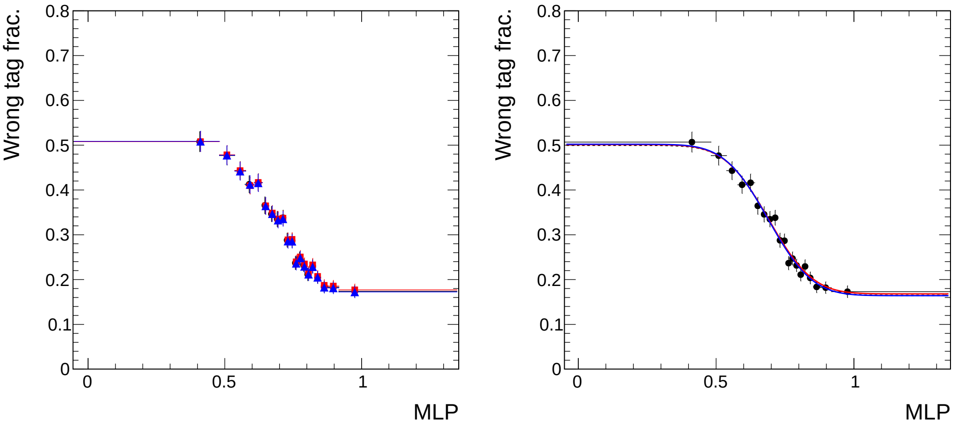


Figure 4.17: Bin-by-bin (left) and function (right) comparison of the OS-muon wrong tag fractions obtained by fitting the B^+ mass distributions using the three fit models: standard (black), modified signal + standard background PDFs (red), and modified background + standard signal PDFs (blue).

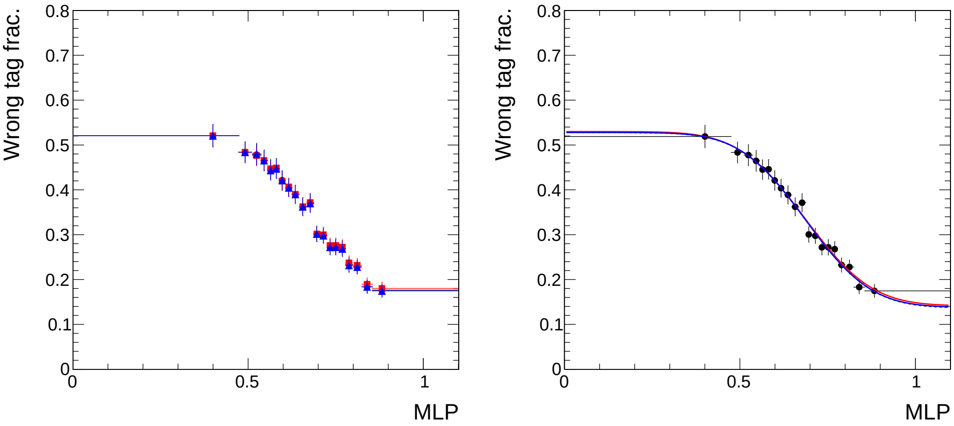


Figure 4.18: Bin-by-bin (left) and function (right) comparison of the OS-electron wrong tag fractions obtained by fitting the B^+ mass distributions using the three fit models: standard (black), modified signal + standard background PDFs (red), and modified background + standard signal PDFs (blue).

The deviations of the tagging performances obtained using the three different fit models are found to be completely negligible.

Data Taking Period

As described in Section 3.4.2, the 2012 data are recorded in four periods, named Run2012A, B, C, D, with variable values of maximum instantaneous luminosity delivered by the LHC and thus with different pile-up conditions. In order to verify the consistency of the tagging performances with the data taking periods, the full data sample is divided in three subsamples:

- Run2012 A + B
- Run2012 C
- Run2012 D

where the data of period A have been merged with the Run2012B. Due to the small amount of integrated luminosity recorded in the Run2012A period (see Table 3.2), the independent study of this period would not have been in fact statistically significant. A calibration function is obtained for each data taking period by measuring the mistag in the B^+ sub-samples and by comparing it with the corresponding value calculated using the ω parametrization function defined using the entire Run2012 dataset.

The calibration parameters obtained for the three sub-samples are then combined in a 2-dimensional fit, where the values of the parameters p_0, p_1 and the correlations are used. Since no dependence on the data taking period is expected, if a severe discrepancy is observed, the effect will be addressed by a correction procedure. Minor differences are instead going to be accounted with the estimation of dedicated systematic uncertainties.

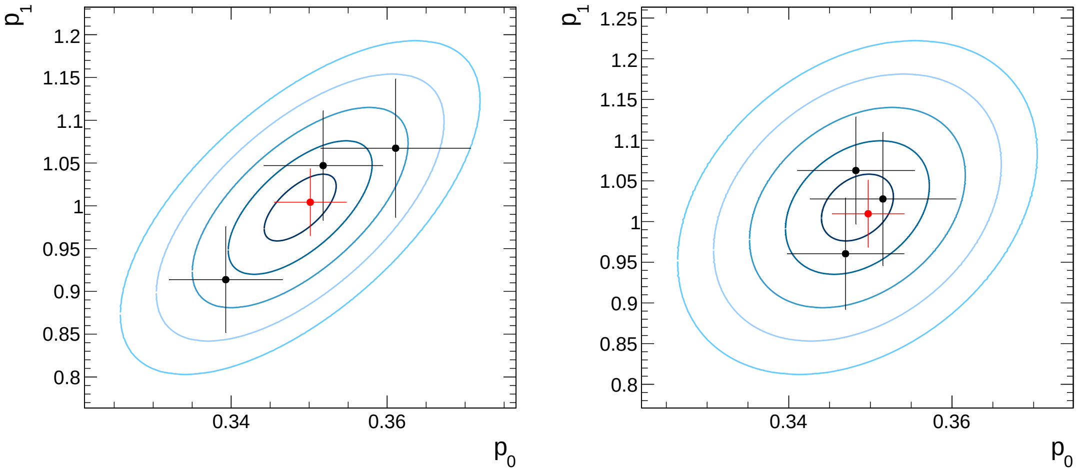


Figure 4.19: Results of the 2-dimensional combination of the OS-muon (left) and OS-electron (right) calibration parameters evaluated for the three data-taking periods. The three measured values (black) are shown along with the 1,2,3,4, and 5 σ levels from the result of combination. The parameters obtained by the standard calibration are shown in red.

The results of the 2-dimensional combinations for the muon and electron tagger are reported in Figures 4.19, while the calibration curves are shown in Fig. 4.20 and 4.21 for muons and electrons respectively.

The results of the various calibration procedures are summarized, together with the results of the 2-dimensional combination, in Table 4.28 for the muon tagger and in Table 4.29 for the electron tagger.

The various measurements are found to be in good agreement with the reference value of Table 4.26 for both the taggers, no correction is therefore applied.

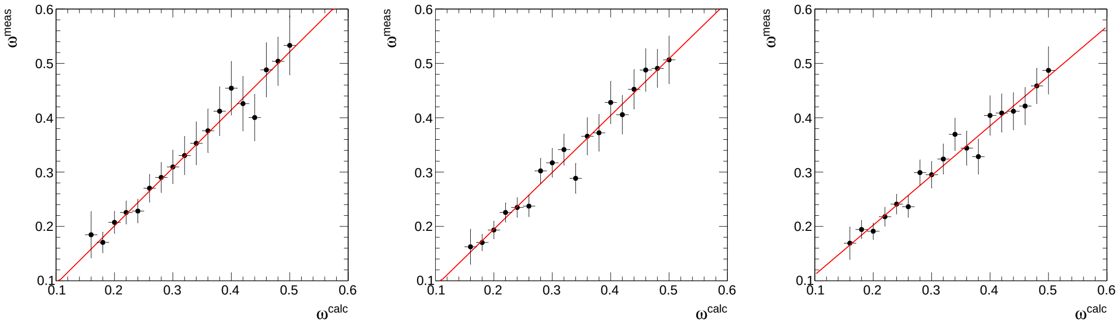


Figure 4.20: OS-muon calibration curves evaluated for the three data taking periods: Run2012 A+B (left); Run2012 C (center); Run2012 D (right).

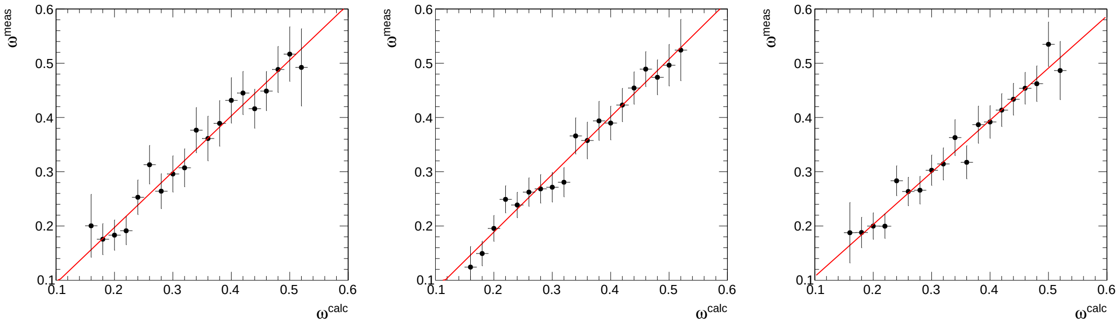


Figure 4.21: OS-electron calibration curves evaluated for the three data taking periods: Run2012 A+B (left); Run2012 C (center); Run2012 D (right).

Table 4.28: Calibration parameters of the OS-muon tagger for the $B^+ \to J/\psi K^+$ data sample divided for the three data taking period. The results from the combined 2-dimensional fit is also shown.

	Run2012 A+B	Run2012 C	Run2012 D	2-dimensional fit
p_0	0.361 ± 0.010	0.352 ± 0.008	0.339 ± 0.007	0.349 ± 0.005
p_1	1.07 ± 0.08	1.05 ± 0.06	0.91 ± 0.06	1.00 ± 0.04
$\sigma(p_0, p_1)$	0.659	0.651	0.620	0.641
$p_0 - p_1 \times \omega'$	-0.01 ± 0.02	-0.01 ± 0.02	-0.02 ± 0.02	

B^+ Kinematic

The full data sample is divided in subsamples according to the transverse momentum of the reconstructed B^+, in order to check for any dependence induced into the tagger by the

Table 4.29: Calibration parameters of the OS-electron tagger for the $B^+ \to J/\psi K^+$ data sample divided for the three data taking period. The results from the combined 2-dimensional fit is also shown.

	Run2012 A+B	Run2012 C	Run2012 D	2-dimensional fit
p_0	0.352 ± 0.009	0.348 ± 0.007	0.347 ± 0.007	0.348 ± 0.004
p_1	1.03 ± 0.08	1.06 ± 0.07	0.96 ± 0.07	1.02 ± 0.04
$\sigma(p_0, p_1)$	0.375	0.346	0.273	0.323
$p_0 - p_1 \times \omega'$	-0.01 ± 0.03	-0.02 ± 0.02	0.01 ± 0.02	

kinematic regime of the reconstructed B hadron. Three divisions are defined according to the B^+ transverse momentum:

- Low p_T: < 18 GeV/c
- Mid p_T: $18 - 26$ GeV/c
- High p_T: > 26 GeV/c

A calibration function is obtained independently for each period and for each tagger, as shown in Figs 4.22 and 4.23, respectively for the muon and electron taggers.

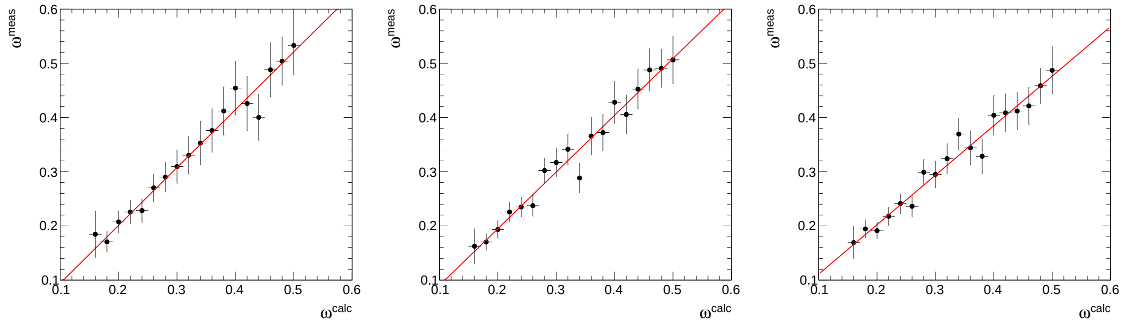


Figure 4.22: OS-muon calibration curves evaluated for the three reco-B^+ p_T ranges: Low (left); Mid (center); High (right).

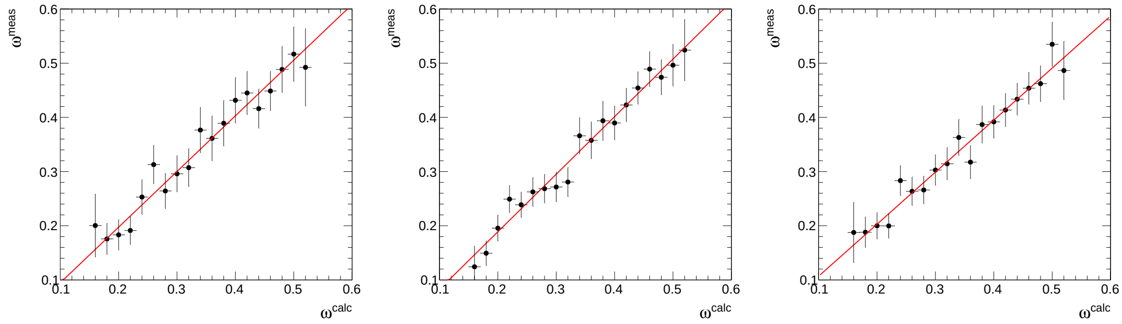


Figure 4.23: OS-electron calibration curves evaluated for the three reco-B^+ p_T ranges: Low (left); Mid (center); High (right).

The measured calibration parameters are combined within a 2-dimensional fit, where the values of the parameters p_0, p_1 and the correlation are used to infer the level of compatibility among the kinematic regions.

The calibration parameters are summarized in Table 4.30 (muons) and 4.31 (electrons) and the results of the 2-dimensional combination are presented in Figure 4.24. The resulting parameters are in agreement with the reference value of Table 4.26.

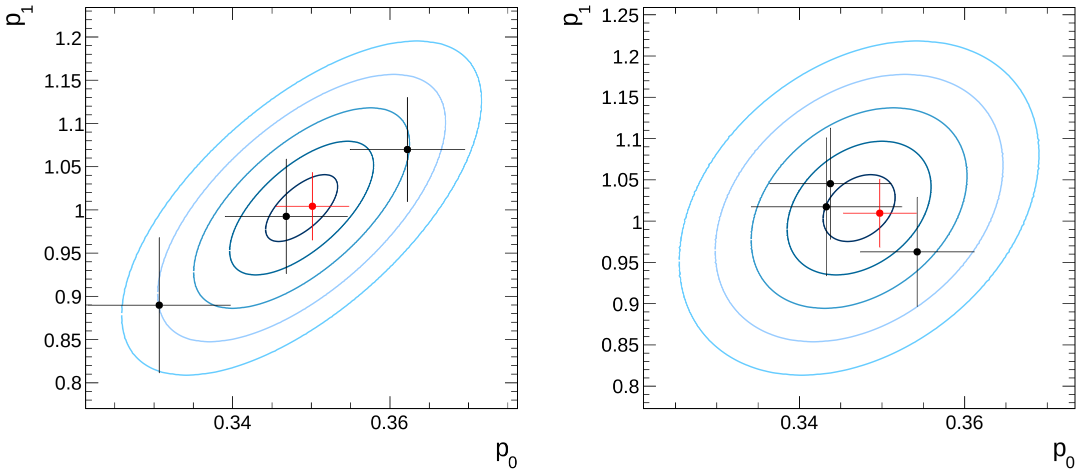


Figure 4.24: Results of the 2-dimensional combination of the OS-muon (left) and OS-electron (right) calibration parameters evaluated for the three ranges of transverse momentum of the reconstructed B meson. The three measured values (black) are shown along with the 1,2,3,4, and 5 σ levels from the result of combination. The parameters obtained by the standard calibration are shown in red.

Table 4.30: Calibration parameters of the OS-muon tagger for the $B^+ \to J/\psi K^+$ data sample divided for the three reco-B^+ p_T ranges. The results from the combined 2-dimensional fit is also shown.

	Low p_T	Mid p_T	High p_T	2-dimensional fit
p_0	0.331 ± 0.009	0.347 ± 0.008	0.362 ± 0.007	0.349 ± 0.005
p_1	0.89 ± 0.08	0.99 ± 0.07	1.07 ± 0.06	1.00 ± 0.04
$\sigma(p_0, p_1)$	0.682	0.665	0.586	0.641
$p_0 - p_1 \times \omega'$	0.02 ± 0.02	0.00 ± 0.02	-0.01 ± 0.02	

Same Side B Flavour

The mistag parametrization and the calibration functions extracted on data using the $B^+ \to J/\psi K^+$ self-tagging sample are designed to be applied to any required B sample, like for instance the B_s^0.

It was already stressed that on principle the behaviour of the opposite side lepton taggers rely solely on the properties of the selected leptons and therefore to the OS-B features. The complete independence of the two sides (same and opposite) of the event is

Table 4.31: Calibration parameters of the OS-electron tagger for the $B^+ \to J/\psi K^+$ data sample divided for the three reco-B^+ p_T ranges. The results from the combined 2-dimensional fit is also shown.

	Low p_T	Mid p_T	High p_T	2-dimensional fit
p_0	0.343 ± 0.009	0.344 ± 0.007	0.354 ± 0.007	0.347 ± 0.004
p_1	1.02 ± 0.08	1.04 ± 0.07	0.96 ± 0.07	1.02 ± 0.04
$\sigma(p_0, p_1)$	0.486	0.380	0.129	0.320
$p_0 - p_1 \times \omega'$	-0.01 ± 0.03	-0.02 ± 0.02	0.02 ± 0.02	

assumed. The tagger performances should thus not be affected by the specific flavour of the reconstructed (same side) B meson. The residual dependence of the taggers performances on the flavour of the reco-B meson is addressed with a dedicated systematic uncertainty.

The mistag functions evaluated using the 2012 data $B^+ \to J/\psi K^+$ decays, and those evaluated in the $B_s^0 \to J/\psi \phi$, $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^*$ MC samples are compared in Figures 4.25, where the data points are also shown for comparison.

It can be easily noted that in the case of the electron tagger the mistag parametrization functions evaluated for the different reconstructed B flavours are in excellent agreement.

The muon tagger does instead show relatively large mistag discrepancies among the three MC samples. In the central range of the neural network discriminator, approximately $0.5 < \text{MLP}_\mu < 0.9$, all the functions show a good agreement. It should be noted that the number of muons in this region contribute to more than the 85% of the total muon tagger efficiency. Larger level of tension, already introduced in Section 4.3.3, can be observed in the tails of the MLP_μ distribution, mostly localized in the low region, $\text{MLP}_\mu < 0.5$.

The $\text{MLP}_\mu < 0.5$ region of the muon tagger is therefore investigated in detail in order to control and possibly correct the source of the disagreement among the mistag function for the different flavours of the same side B hadron. As described in Section 3.4.2.2, the composition of the generated MC samples is corrected by a reweighting factor in order to reproduce the relative amounts of same side and opposite side B hadrons as in an unbiased sample. Sizeable discrepancies in the relative amounts of the B meson flavours among the various MC samples could result in a bias of the value of the mistag; for instance, if a MC sample reveals an higher amount of OS-B_s^0 mesons, the mistag will increase due to the high B_s^0 mixing probability.

The flavour composition of the opposite side B hadrons is therefore measured for the three MC samples specifically in the low MLP region. The relative amounts of B mesons are in good agreement for the three samples, as can be observed in Table 4.32, thus indicating that no major correction to the MC samples is missing for the correct reproduction of the unbiased B-flavour abundances. The residual small level of disagreement that could be observed in Table 4.32, always within 1.5σ, is however not enough to account for the measured difference of mistag.

As already discussed in Section 3.2.2, the charge of the kaon produced in the same side

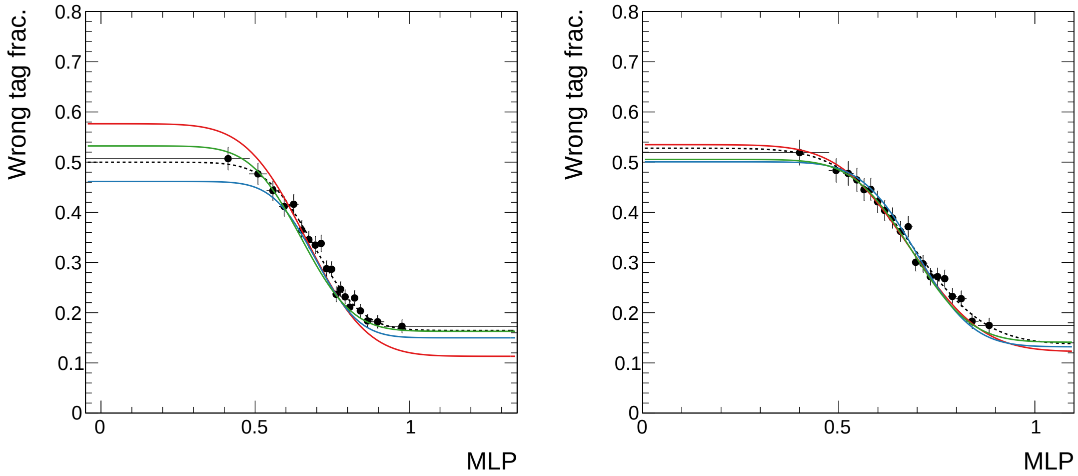


Figure 4.25: Comparison of the OS-muon (left) and OS-electron (right) wrong tag function evaluated with respect to different flavours of the reco-B mesons. The functions obtained studying the $B_s^0 \to J/\psi\, \phi$ (red), $B^+ \to J/\psi\, K^+$ (blue) and $B^0 \to J/\psi\, K^*$ (green) MC samples are compared to the function obtained with the 2012 $B^+ \to J/\psi\, K^+$ data (dashed black) and the data points.

Table 4.32: Fraction of opposite side B hadron flavours for the three MC samples in the events characterised by OS-muon $\mathrm{MLP}_\mu < 0.5$.

	Opposite Side B flavour			
[%]	B^+	B_s^0	B^0	Λ_b
$B^+ \to J/\psi\, K^+$ MC	42.2 ± 2.1	12.2 ± 2.6	39.8 ± 2.1	5.0 ± 2.7
$B_s^0 \to J/\psi\, \phi$ MC	39.1 ± 2.0	12.3 ± 2.4	40.5 ± 2.0	6.8 ± 2.5
$B^0 \to J/\psi\, K^*$ MC	40.7 ± 2.2	9.9 ± 2.7	42.9 ± 2.2	5.4 ± 0.7

hadronization process of a B_s^0 meson carries the correlation to the initial flavour of the same side b-parton. There is the possibility that selected opposite tag leptons are in fact mis-reconstructed kaons arising from the same side of the event. As charged kaons are more likely to be produced in the fragmentation of a b-quark to a B_s^0 meson, this might introduce a flavour-dependent bias in the tagging performance. A preliminary loose selection cut on the angular distance between the SS-B and the lepton is applied at selection level ($\Delta R > 0.3$ for muons and > 0.2 for electrons), as discussed in Section 4.2.2. This selection requirements act as a veto for particles involved in the same side decay and hadronization process. In order to check for the potential same side kaon pollution a set of tighter ΔR vetoes is applied to the B_s^0 MC sample in order to further remove possible SS-K → OS-μ inclusions. A set of increasing ΔR cuts ($\Delta R > 0.3, 0.5, 0.7, 1.0$) is then applied on top of the muon selection, and the average mistag is evaluated for the already defined twenty categories exploited in the definition of the standard mistag function. Parametrization curves are then produced for all the ΔR conditions by the independent fit of the different binned graphs, and the results are compared. As it can be observed in Figure 4.26, all the functions are completely in agreement within the entire spectrum of the multilayer perceptron output MLP_μ, thus suggesting that tag-side kaon pollution is not the source of the observed discrepancy.

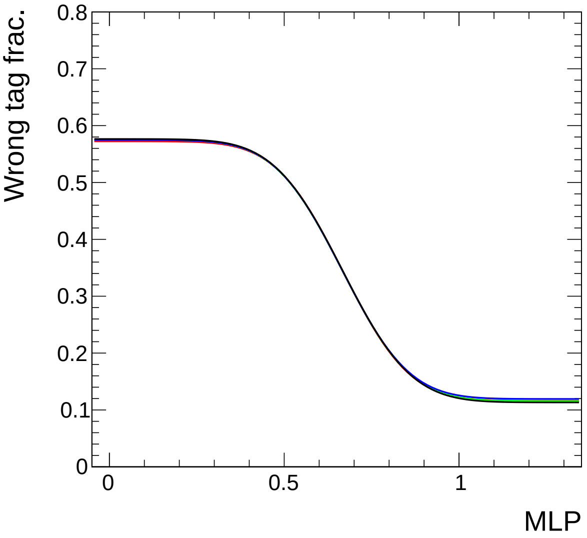


Figure 4.26: OS-muon mistag parametrization curves obtained in the $B_s^0 \to J/\psi\,\phi$ simulated sample using different values of ΔR vetoes. The standard $\Delta R > 0.3$ condition is presented in black, $\Delta R > 0.5$ in red, $\Delta R > 0.7$ in green, and $\Delta R > 1.0$ in blue.

The estimation of the flavour systematic is performed using the mistag function obtained on the $B^+ \to J/\psi\,K^+$ MC sample, and by comparing it with the mistag measured in the $B_s^0 \to J/\psi\,\phi$ MC and $B^0 \to J/\psi\,K^*$ MC samples my means of the true MC informations. A set of calibration curves can therefore be obtained by the comparison of the three MC mistag functions, as reported in Figures 4.27 and 4.28, where the mistag evaluated on the $B^+ \to J/\psi\,K^+$ MC is now exploited as the calculated ω^{calc} mistag. The calibration parameters obtained with this procedure are summarized in Tables 4.33 and 4.34, where it can be observed that the main discrepancy is due to the difference of the mistag in the $B_s^0 \to J/\psi\,\phi$ and $B^+ \to J/\psi\,K^+$ channels for the muon tagger, as it was expected (see Figure 4.7). While the parameter extracted from the $B^+ \to J/\psi\,K^+$ Monte Carlo vs $B^+ \to J/\psi\,K^+$ Monte Carlo calibration curve are directly related to the statistical finiteness of the B^+ MC sample, the calibration of the $B_s^0 \to J/\psi\,\phi$ vs $B^+ \to J/\psi\,K^+$ and the $B^0 \to J/\psi\,K^*$ vs $B^+ \to J/\psi\,K^+$ MC samples are directly related to the discrepancies of the mistag among the different flavours.

The new sets of calibration parameters related to the flavour dependence can thus be propagated for the estimation of the mistag and tagging power on data to evaluate the systematic uncertainty. As a conservative estimation the calibration parameters obtained from the $B_s^0 \to J/\psi\,\phi$ MC vs $B^+ \to J/\psi\,K^+$ MC fit are used. The systematic uncertainties are quantified in $\delta\omega = 0.15\%$ and $\delta\mathcal{P}_{tag} = 0.012\%$ for the muon tagger and $\delta\omega = 0.05\%$ and $\delta\mathcal{P}_{tag} = 0.003\%$ for the electron tagger, as measured from the $B^+ \to J/\psi\,K^+$ data sample.

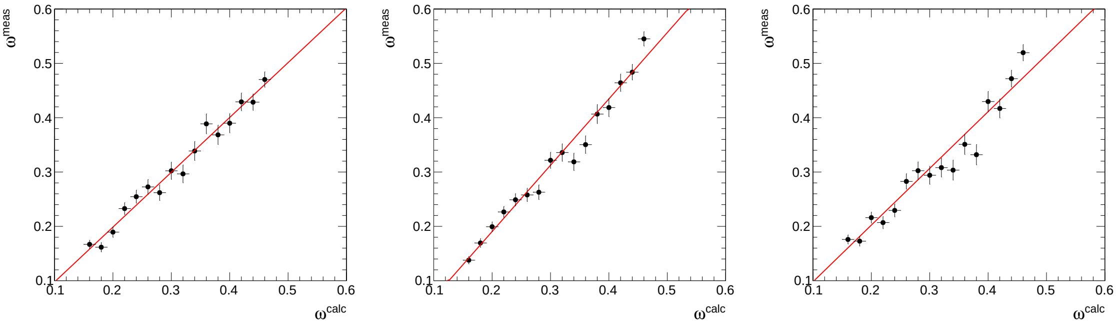


Figure 4.27: OS-muon calibration curves evaluated for the three MC samples using as the calculated mistag ω^{calc} the values obtained using the parametrization extracted from the $B^+ \to J/\psi K^+$ simulation, and as measured mistag ω^{meas} the values obtained from the $B^+ \to J/\psi K^+$ (left), $B_s^0 \to J/\psi \phi$ (center), and $B^0 \to J/\psi K^*$ (right) MC.

Table 4.33: OS-muon calibration parameters evaluated for the three MC samples using as the calculated mistag ω^{calc} the values obtained using the parametrization extracted from the $B^+ \to J/\psi K^+$ simulation, and as measured mistag ω^{meas} the values obtained from the $B^+ \to J/\psi K^+$, $B_s^0 \to J/\psi \phi$, and $B^0 \to J/\psi K^*$ MC.

	$B^+ \to J/\psi K^+$ MC	$B_s^0 \to J/\psi \phi$ MC	$B^0 \to J/\psi K^*$ MC
p_0	0.350 ± 0.004	0.374 ± 0.004	0.359 ± 0.004
p_1	1.01 ± 0.03	1.22 ± 0.03	1.04 ± 0.04
$\sigma(p_0, p_1)$	0.666	0.683	0.666
$p_0 - p_1 \times \omega'$	-0.003 ± 0.009	-0.054 ± 0.009	-0.006 ± 0.010

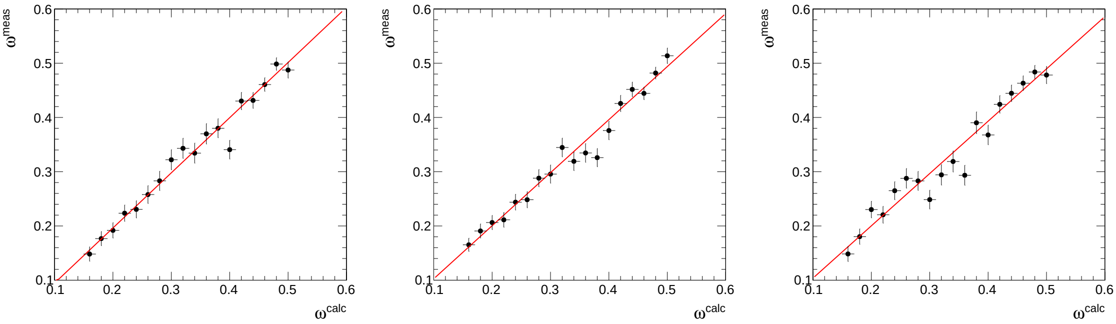


Figure 4.28: OS-electron calibration curves evaluated for the three MC samples using as the calculated mistag ω^{calc} the values obtained using the parametrization extracted from the $B^+ \to J/\psi K^+$ simulation, and as measured mistag ω^{meas} the values obtained from the $B^+ \to J/\psi K^+$ (left), $B_s^0 \to J/\psi \phi$ (center), and $B^0 \to J/\psi K^*$ (right) MC.

4.3.6 Single Particle Tagger Results

The tagging performances are evaluated for each of the two opposite side lepton taggers. The performances are measured on the $B^+ \to J/\psi K^+$ 2012 data using the per-event parametrization of the mistag as function of the Neural Network discriminator. The results are reported on Table 4.35, where the uncertainties are divided into their statistical

Table 4.34: OS-electron calibration parameters evaluated for the three MC samples using as the calculated mistag ω^{calc} the values obtained using the parametrization extracted from the $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ simulation, and as measured mistag ω^{meas} the values obtained from the $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$, $\mathrm{B}^0_s \to \mathrm{J}/\psi\,\phi$, and $\mathrm{B}^0 \to \mathrm{J}/\psi\,\mathrm{K}^*$ MC.

	$\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ MC	$\mathrm{B}^0_s \to \mathrm{J}/\psi\,\phi$ MC	$\mathrm{B}^0 \to \mathrm{J}/\psi\,\mathrm{K}^*$ MC
p_0	0.349 ± 0.004	0.347 ± 0.003	0.345 ± 0.004
p_1	1.02 ± 0.03	0.98 ± 0.03	0.96 ± 0.03
$\sigma(p_0, p_1)$	0.148	0.158	0.149
$p_0 - p_1 \times \omega'$	-0.007 ± 0.011	0.005 ± 0.010	0.007 ± 0.012

and systematic components. The systematic uncertainty for both the taggers is dominated by the same side flavour dependence, the other investigated sources being negligible in comparison. The tagging power estimated on $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ 2012 data for the muon tagger is 0.83 ± 0.02 (stat.) $\pm\ 0.012$ (syst.) %, and the final value of 0.48 ± 0.02 (stat.) $\pm$ 0.003 (syst.) % is obtained for the electron tagger, where the uncertainties are dominated by the statistical contribution.

Table 4.35: Performances of the OS-muon and OS-electron flavour tagging algorithms. The performances are measured on the 2012 $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ data decay channel, corresponding to 19.8 fb^{-1}.

[%]	muon tagger	electron tagger
ε_{tag}	4.56 ± 0.02 (stat.)	3.92 ± 0.02 (stat.)
ω	28.64 ± 0.32 (stat.) $\pm\ 0.15$ (syst.)	32.45 ± 0.36 (stat.) $\pm\ 0.05$ (syst.)
$\mathcal{P}_{tag}$	0.833 ± 0.024 (stat.) $\pm\ 0.012$ (syst.)	0.483 ± 0.020 (stat.) $\pm\ 0.003$ (syst.)

4.4 Lepton Tagger Combination

The muon and electron opposite side taggers can be combined to evaluate the overall opposite side lepton (OS-ℓ) tagger performances. From the total number of $(707 \pm 2) \times 10^3$ reconstructed $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ decays on data, a number of $(32.4 \pm 0.2) \times 10^3$ are tagged by the muon algorithm, and $(27.2 \pm 0.2) \times 10^3$ are tagged by the electron tagger. The overlap of the two taggers is minimal: the number of events containing both a selected OS-μ and a OS-e is in fact $1\,185 \pm 37$, corresponding to a fraction of about 3.5% of the events with both muon and electron tagger responses.

Due to the small overlap, the two tagger responses and mistags are not combined to a relatively complex algorithm, such as a multivariate tool. For those events where both the lepton taggers neural network responses are available, the tagger with the highest absolute value of dilution $|\mathcal{D}| = |1 - 2\omega|$ is in fact chosen, and its tag decision and mistag are used to assign the flavour of the reconstructed B hadron. This choice maximizes the tagging power for any given sample. However, due to the very limited overlap of the two taggers, the difference with other simple strategies, like for instance using the tagger with the lowest value of mistag, has been measured to be negligible.

The combined lepton tagger can therefore be used as a single-particle tagger, each event being characterised by a single tag decision and a single value of mistag. The calibration of the combined opposite side ℓ tagger, as well as the estimation of the systematic uncertainties, can thus be performed as on a single tagger, independently from the results previously obtained for the muon and electron algorithms.

The parameters resulting from the calibration on the $B^+ \to J/\psi K^+$ data are presented in Table 4.36 where the uncertainties are statistical only. The results agree with the expectations of a linear fit with $p_0 = \omega'$ and $p_1 = 1$ within the statistical uncertainties, as shown in Figure 4.29.

Table 4.36: OS-lepton tagger calibration parameters. The uncertainties are statistical only.

p_0	0.348 ± 0.003
p_1	1.01 ± 0.03
ρ	0.510
$p_0 - p_1 \times \omega'$	-0.004 ± 0.009

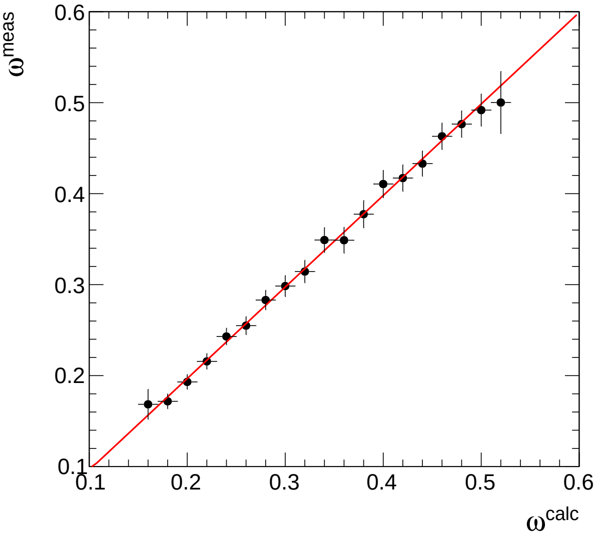


Figure 4.29: The measured OS-lepton tagger wrong tag fraction extracted from fit to the B^+ mass in bins of the calculated mistag, obtained as function of the neural network discriminator $\omega = \omega(\mathrm{MLP})$.

The fraction of events tagged by the lepton algorithm is measured to be $8.31 \pm 0.03\%$, where the uncertainty is only statistical. The final resulting performances are reported on Table 4.37, where the uncertainties are divided into their statistical and systematic components. The total tagging power obtained by the algorithm is 1.307 ± 0.031 (stat.) $\pm$ 0.007 (syst.) %. The main investigated source of systematic uncertainty is due to the residual dependence on the flavour of the reconstructed (same side) B meson. The total amount of systematic uncertainty is however limited, as the measured tagging performances are dominated by the statistical errors.

Table 4.37: Performances of the OS-lepton flavour tagging algorithm. The performances are measured on the $B^+ \to J/\psi K^+$ decay channel on 2012 data, corresponding to $19.8\ \text{fb}^{-1}$.

[%]	OS-ℓ tagger
ε_{tag}	8.31 ± 0.03 (stat.)
ω	30.2 ± 0.2 (stat.) ± 0.05 (syst.)
$\mathcal{P}_{tag}$	1.307 ± 0.031 (stat.) ± 0.007 (syst.)

Chapter 5

Flavour Tagged CP Violation Measurement with $B_s^0 \to J/\psi\,\phi$ Decays

As extensively discussed in Chapter 1, before decaying into the final state the B_s^0 mesons might be subject to the flavour oscillation between the two flavour eigenstates B_s^0 and $\overline{B_s^0}$. The eigenstates related to the time evolution of the B_s^0 system are the mass eigenstates B_H and B_L, characterized by sizeable mass and decay width difference. The decay width difference $\Delta\Gamma_s$ is predicted to be non-zero in the SM, and the theoretical prediction, assuming no NP in B_s^0 mixing, is $\Delta\Gamma_s = 0.087 \pm 0.021\ \mathrm{ps}^{-1}$ [63]. The CP-violating phase ϕ_s arise from the interference between direct B_s^0 meson decays into a $c\bar{c}s$ CP eigenstate and decays mediated by mixing to the same final state. A value of $\phi_s = 0.0363^{+0.0016}_{-0.0015}$ is predicted by the SM from a global fit to the experimental data [64]. Since the value of ϕ_s is small and precisely predicted, any deviation of the measured value would be particularly interesting as a possible hint of physics beyond the SM. The weak phase ϕ_s has been measured at the Tevatron experiments [65–68] and at the LHC by LHCb and ATLAS experiments [51, 69–72], using $B_s^0 \to J/\psi\,\phi$, $B_s^0 \to J/\psi\,f_0$, and $B_s^0 \to \phi\,\phi$ decays. Final states consisting of two vector mesons, as in the case of the $B_s^0 \to J/\psi\,\phi$ decay, are not CP eigenstates and a time-dependent angular analysis is needed to statistically disentangle the CP-even from the CP-odd contributions. A crucial component of the analysis, the flavour tagging is used in order to determine whether a given candidate was produced as a B_s^0 or $\overline{B_s^0}$, improving the sensitivity of the measurement of the weak CP violating phase ϕ_s.

5.1 Phenomenology of the $B_s^0 \to J/\psi\,\phi$ Decay

The $B_s^0 \to J/\psi\,\phi$ decay can either proceed through a *P*-wave transition, resulting is a CP-odd final state, or through *S*-, *D*-waves which result in a CP-even eigenstate. The $B_s^0 \to J/\psi\,\phi$ decay amplitude can be decomposed in three independent amplitudes, which can be expressed in terms of three equivalent bases: the basis of partial waves [73] (*S*, *P*, *D*), the transversity [74] basis (A_0, $A_\parallel$, $A_\perp$), and the helicity [75] basis (H_0, H_+, H_-). Whatever basis is chosen, the individual amplitudes and phases are inferred by analysing the angular distributions of the decay products of the two vector mesons. Partial waves are directly related to the states of relative angular momentum $L = 0, 1, 2$, while the helicity basis can be used to express the amplitude in terms of the helicity states $\lambda =$

$0, \pm 1$. The transversity basis employed here is related to the alignment of the projections of the J/ψ and ϕ spins in the plane orthogonal to the particles momenta, and it is the most convenient for identifying the CP parity of the final state.

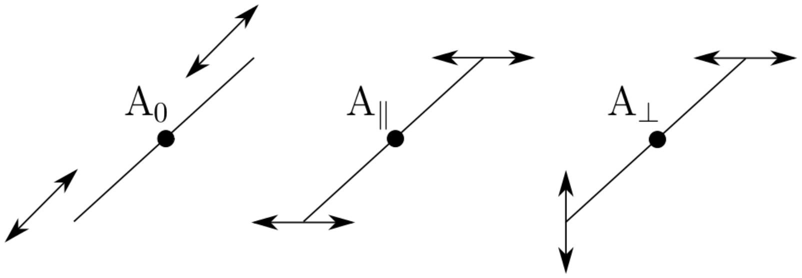


Figure 5.1: Schematic representation of the transversity amplitudes, as dependent to the relative alignment of the spin-1 mesons. A_0: longitudinal polarisation; $A_\parallel$: parallel transverse polarisation; $A_\perp$: perpendicular transverse polarisation.

As described in Figure 5.1, A_0 is the longitudinal polarization and corresponds to a CP-even state, while $A_\parallel$ and $A_\perp$ are transverse polarizations with particle spins parallel and perpendicular to each other, respectively. The $A_\perp$ amplitude is related to a CP-odd state, while $A_\parallel$ is CP-even. A fourth decay amplitude A_S is related to the non-resonant decay $B_s^0 \to J/\psi\, K, K$, where the two kaons are not produced by ϕ decay: this is related to a CP-odd configuration. Each amplitude has an associated complex strong phase, identified as δ_0, $\delta_\parallel$, $\delta_\perp$ and δ_S similarly to the related amplitudes. Since only the relative phases are important, it is used the convention where δ_0 is set to zero. Furthermore, the amplitudes are related by the normalization condition $|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2 + |A_S|^2 = 1$.

The amplitudes are expressed as functions of three angles $(\theta_T, \psi_T, \varphi_T)$ in the transversity basis [76], see Fig. 5.2. Here, θ and φ are the polar and azimuthal angle of the μ^+ in the rest frame of the J/ψ, where the x-axis is defined by the direction of the B_s^0 and the xy-plane by the decay plane of $\phi\,(1020) \to K^+K^-$. The angle ψ is the angle computed in the ϕ rest frame between the direction of the K^+ and that of the B_s^0.

The differential decay rate of the $B_s^0 \to J/\psi\,\phi$ decay in terms of proper decay time and angular variables is given, for a state originally produced as a B_s^0, by the sum of ten contributions where the angular and time-dependent components are factorised [77]:

$$\frac{d^4\Gamma\left(B_s^0\right)}{d\Theta dt} = f\left(\Theta, \alpha, t\right) \propto \sum_{i=1}^{10} O_i\left(\alpha, t\right) \cdot g_i\left(\Theta\right), \tag{5.1}$$

where O_i are the time-dependent functions, g_i are the angular functions, α is the set of physics parameters to be extracted from the fit to data, $\Theta = (\theta_T, \psi_T, \varphi_T)$ is the short notation representing the angles, and t represents the proper decay time of the B_s^0 meson. The angular functions are described in Table 5.1.

The time dependent O_i functions, described in Equation 5.2, explicitly contain the depen-

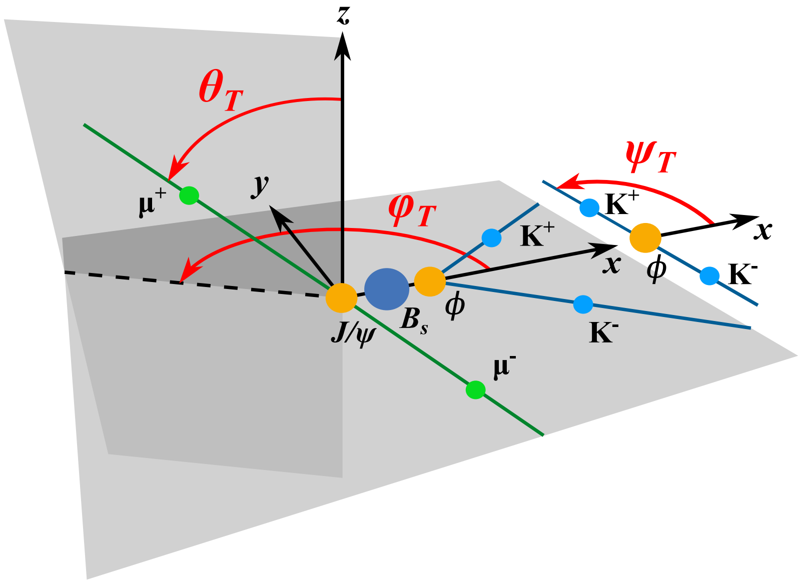


Figure 5.2: Definition of the three transversity angles (θ_T, ψ_T, φ_T) used for the description of the decay topology.

Table 5.1: Angular components of the signal model.

i	$g_i\,(\theta_T, \psi_T, \varphi_T)$
1	$2\cos^2\psi_T\left(1-\sin^2\theta_T\cos^2\varphi_T\right)$
2	$\sin^2\psi_T\left(1-\sin^2\theta_T\sin^2\varphi_T\right)$
3	$\sin^2\psi_T\sin^2\theta_T$
4	$-\sin^2\psi_T\sin 2\theta_T\sin\varphi_T$
5	$\frac{1}{\sqrt{2}}\sin 2\psi_T\sin^2\theta_T\sin 2\varphi_T$
6	$\frac{1}{\sqrt{2}}\sin 2\psi_T\sin 2\theta_T\sin\varphi_T$
7	$\frac{2}{3}\left(1-\sin^2\theta_T\cos^2\varphi_T\right)$
8	$\frac{1}{3}\sqrt{6}\sin\psi_T\sin^2\theta_T\sin 2\varphi_T$
9	$\frac{1}{3}\sqrt{6}\sin\psi_T\sin 2\theta_T\cos\varphi_T$
10	$\frac{4}{3}\sqrt{3}\cos\psi_T\left(1-\sin^2\theta_T\cos^2\varphi_T\right)$

dency on the $\Delta\Gamma_s$ and Δm_s parameters of the B_s^0 mixing:

$$\begin{aligned} O_i\,(\alpha, t) = N_i e^{-t/\tau}\Big[& a_i\cosh\left(\frac{1}{2}\Delta\Gamma_s t\right) + b_i\sinh\left(\frac{1}{2}\Delta\Gamma_s t\right) \\ & + c_i\cos\left(\Delta m_s t\right) + d_i\sin\left(\Delta m_s t\right)\Big] \end{aligned} \tag{5.2}$$

The N_i, a_i, b_i, c_i, d_i terms included within the O_i terms are described using the equations in Table 5.2, where it is clear the further dependency of decay rate on the sizes of the amplitudes and on the strong phases.

Finally the terms C, S and D are defined, using the same sign convention as LHCb [70],

Table 5.2: Time-dependent terms of the signal model.

i	N_i	a_i	b_i	c_i	d_i
1	$\lvert A_0(0)\rvert^2$	1	D	C	$-S$
2	$\lvert A_\parallel(0)\rvert^2$	1	D	C	$-S$
3	$\lvert A_\perp(0)\rvert^2$	1	$-D$	C	S
4	$\lvert A_\parallel(0)\rvert\lvert A_\perp(0)\rvert$	$C\sin(\delta_\perp-\delta_\parallel)$	$S\cos(\delta_\perp-\delta_\parallel)$	$\sin(\delta_\perp-\delta_\parallel)$	$D\cos(\delta_\perp-\delta_\parallel)$
5	$\lvert A_0(0)\rvert\lvert A_\parallel(0)\rvert$	$\cos(\delta_\parallel-\delta_0)$	$D\cos(\delta_\parallel-\delta_0)$	$C\cos(\delta_\parallel-\delta_0)$	$-S\cos(\delta_\parallel-\delta_0)$
6	$\lvert A_0(0)\rvert\lvert A_\perp(0)\rvert$	$C\sin(\delta_\perp-\delta_0)$	$S\cos(\delta_\perp-\delta_0)$	$\sin(\delta_\perp-\delta_0)$	$D\cos(\delta_\perp-\delta_0)$
7	$\lvert A_S(0)\rvert^2$	1	$-D$	C	S
8	$\lvert A_S(0)\rvert\lvert A_\parallel(0)\rvert$	$C\cos(\delta_\parallel-\delta_S)$	$S\sin(\delta_\parallel-\delta_S)$	$\cos(\delta_\parallel-\delta_S)$	$D\sin(\delta_\parallel-\delta_S)$
9	$\lvert A_S(0)\rvert\lvert A_\perp(0)\rvert$	$\sin(\delta_\perp-\delta_S)$	$-D\sin(\delta_\perp-\delta_S)$	$C\sin(\delta_\perp-\delta_S)$	$S\sin(\delta_\perp-\delta_S)$
10	$\lvert A_S(0)\rvert\lvert A_0(0)\rvert$	$C\cos(\delta_0-\delta_S)$	$S\sin(\delta_0-\delta_S)$	$\cos(\delta_0-\delta_S)$	$D\sin(\delta_0-\delta_S)$

as follows:

$$C = \frac{1-|\lambda|^2}{1+|\lambda|^2}, \qquad S = -\frac{2|\lambda|\sin\phi_s}{1+|\lambda|^2}, \qquad D = -\frac{2|\lambda|\cos\phi_s}{1+|\lambda|^2}, \tag{5.3}$$

Eq. 5.1 represents the model of a particle originally produced as B_s^0, the model of the $\overline{B_s^0}$ flavour case is obtained by reversing the sign of the c_i and d_i terms.

The complex parameters

$$\lambda_i = \frac{q}{p}\frac{A_i}{\overline{A_i}} \tag{5.4}$$

relate the decay amplitudes A_i ($\overline{A_i}$) of a B_s^0 ($\overline{B_s^0}$) meson to decay to the final state i, and the B_s^0-$\overline{B_s^0}$ mixing parameters p and q [20]. Assuming polarisation-independent CP-violation effects, λ_i can be simplified as $\lambda_i = \eta_i\lambda$, where η_i is the CP eigenvalue. No direct CP violation is assumed for this analysis, and therefore $|\lambda|$ is set to one, consistently with the results in Ref. [69]. A source of systematic uncertainty is associated to this simplification.

5.2 Event Selection and Simulated Samples

The data used in this analysis have been collected by the CMS experiment during the 2012 run of the LHC, corresponding to an overall integrated luminosity of 20.0 ± 0.1 fb^{-1} of pp collisions at the centre of mass energy of $\sqrt{s} = 8$ TeV.

The same trigger optimised for the detection of b-hadrons decaying to J/ψ, previously described in Section 3.4.2.1, is used to collect the data. It is useful to remind here that the trigger requirements involve a cut on the decay length significance (defined in the transverse detector plane $r-\phi$) of the dimuon candidate $L_{xy}/\sigma(L_{xy}) > 3$. Offline selection criteria, similar to the ones discussed in Section 3.4.2.1, are applied to select the B_s^0 candidates starting with requiring the individual muon candidates to lie within a kinematic acceptance region of $p_T > 4$ GeV and $|\eta| < 2.1$. Two oppositely charged muon candidates are paired and required to originate from a common vertex, which is determined by a Kalman vertex fit [32]. Dimuon candidates are kept if their invariant mass lies within 150 MeV of the world-average J/ψ mass [18]. Candidate $\phi(1020)$ mesons are reconstructed from pairs of oppositely charged tracks with $p_T > 0.7$ GeV, after removing the muon candidate tracks forming the J/ψ. Each selected track is assumed to be a kaon and the invariant mass of a track pair is required to be within 10 MeV of the world average $\phi(1020)$-meson mass [18].

The B_s^0 candidates are formed by combining a J/ψ with a $\phi(1020)$ candidate. The two muons and the two kaons are fitted with a combined vertex and kinematic fit, with a constraint of the dimuon invariant mass to be the nominal J/ψ mass [18]. A B_s^0 candidate is retained if the J/ψ-$\phi(1020)$ pair has an invariant mass between 5.20 and 5.65 GeV and the χ^2 vertex fit probability is larger than 2%. In each selected event at least one reconstructed primary vertex is required. In case of multiple primary vertices, the one which minimises the angle between the flight direction and the momentum of the B_s^0 is selected. The average number of primary vertices is approximately 16. For events with more than one B_s^0 candidate, the candidate with the highest vertex fit probability is selected.

Simulated samples, produced and selected according to the prescriptions of Section 3.4.2, are validated through comparison with the data and are then used to determine the signal reconstruction efficiencies, and to estimate the peaking backgrounds in the signal mass window.

The main background for the $\mathrm{B}_s^0 \to \mathrm{J}/\psi\,\phi$ decays originates from non-prompt J/ψ arising from the decay of b-hadrons, such as B^0, B^+, Λ_b, and B_c. The B_c cross section is expected to be very small and therefore B_c decays are not considered. The Λ_b contribution to the selected events is also found to be very small and the mass distribution in the selected mass range is observed to be flat. The mass distribution of the signal region is shown in Fig. 5.3, and the proper decay time distribution and its uncertainty in Fig. 5.4.

5.3 Efficiencies

Efficiency corrections due to the detector acceptance, the trigger selection, and the selection criteria applied in the data analysis are taken into account in the modelling of the proper decay time and angular observables.

The efficiency corrections $\epsilon(\Theta)$ are calculated as a function of the angular observables Θ using a fully simulated sample of $\mathrm{B}_s^0 \to \mathrm{J}/\psi\,(\mu^+\mu^-)\;\phi\,(\mathrm{K}^+\mathrm{K}^-)$ decays and comparing it to the original generator-level sample without any preselection. The samples used to determine the efficiency correction are simulated with $\Delta\Gamma$ set to zero, in order to avoid correlations between the proper decay time and the angular variables. Efficiency corrections are computed in the three dimensional space, and parametrised with a three-dimensional analytic function accounting for the cross terms between the angular observables.

At the analysis stage a sharp selection, $ct > 200\ \mu$m, is further applied, and the proper decay time efficiency is parametrised by the sum of a linear function and a sigmoid. It is verified that the contribution of the lifetime efficiency correction is very small and its residual effect is taken into account as a systematic uncertainty on the measurement.

5.4 Effect of Flavour Tagging

The decay model described so far within this chapter assumes that the flavour of the B_s^0 meson at production time is known. In reality the flavour is assigned only in tagged

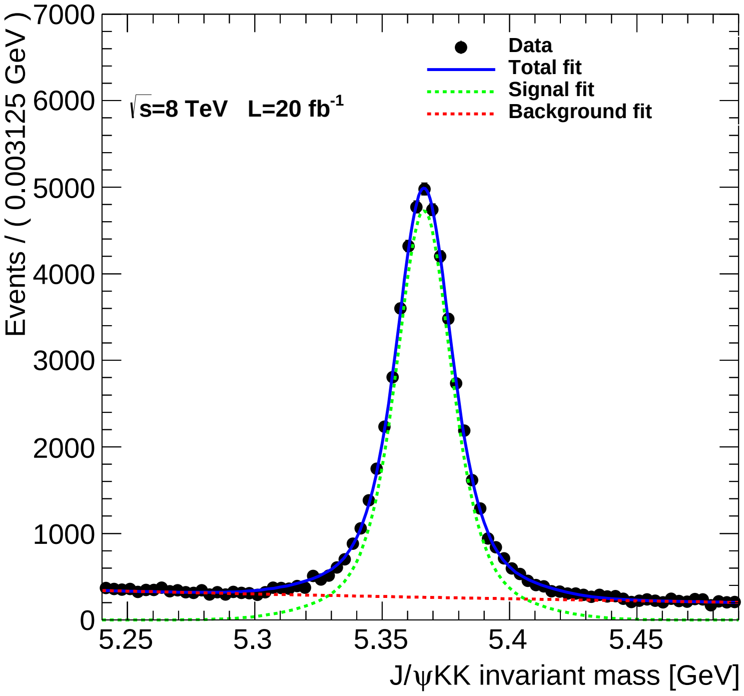


Figure 5.3: The $J/\psi K^+K^-$ invariant mass distribution of the B_s^0 candidates. Data is represented with solid markers, the blue line is the projection of the full likelihood, and the dashed green and red lines represent the fitted signal and background respectively.

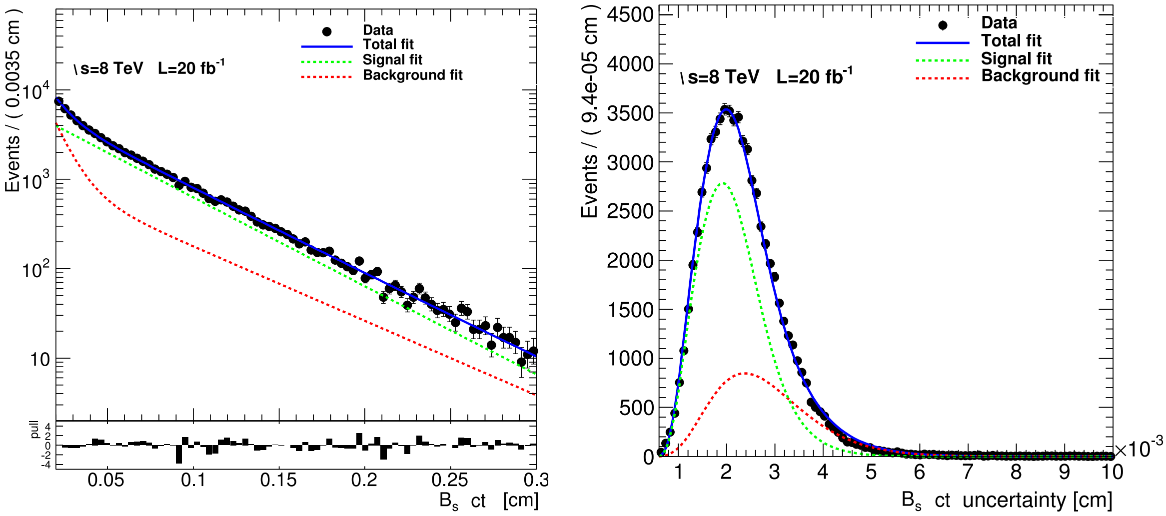


Figure 5.4: The proper decay time distribution (left) and uncertainty (right) of the B_s^0 candidates. Data is represented with solid markers, the blue line is the projection of the full likelihood, and the dashed green and red lines represent the fitted signal and background respectively. For the proper decay time distribution the pull between the data and the fitted function is displayed in the histogram below.

events using the algorithm described in the previous chapter. Equation 5.2 is then modified as:

$$O_i(\alpha,t) = N_i e^{-t/\tau}\Big[a_i \cosh\left(\frac{1}{2}\Delta\Gamma_s t\right) + b_i \sinh\left(\frac{1}{2}\Delta\Gamma_s t\right) \\ + c_i \xi (1-2\omega)\cos(\Delta m_s t) + d_i \xi (1-2\omega)\sin(\Delta m_s t)\Big] \tag{5.5}$$

to include the tag decision ξ and the tagging algorithm mistag ω. The tag decision takes the value +1 in the case of an identified B_s^0 at production time, and -1 in the case of a $\overline{B_s^0}$. For those events for which the flavour tagging algorithm is not able to provide a tag decision same model is used with $\omega = 0.5$. In this analysis the mistag probability ω is included by using the per event estimated from the tagging algorithms on the $B^+ \to J/\psi K^+$ data, as discussed in the previous Chapter 4. The statistical and systematic uncertainties associated to the flavour tagging algorithms are propagated to the measure of ϕ_s, as will be discussed in Section 5.6.

5.5 Maximum Likelihood Fit

An unbinned maximum likelihood fit to the data is performed by including the information on the invariant mass (m), the three decay angles (Θ) of the reconstructed B_s^0 candidates, the proper decay time (t) and its uncertainty (σ_t). The fit is applied to the sample of 70 000 events (49 000 signal candidates and 21 000 background events), selected in the mass range $[5.24 - 5.49]$ GeV and proper decay time range $ct = [200 - 3\,000]$ μm. From this multi-dimensional fit, the set of physics parameters of interest (α in Eq. 5.1) $\Delta\Gamma_s$, ϕ_s, the B_s^0 mean lifetime $c\tau$, $|A_\perp|^2$, $|A_0|^2$, $|A_S|^2$, and the strong phases $\delta_\parallel$, $\delta_\perp$ and $\delta_{S\perp} = \delta_S - \delta_\perp$ are determined.

The likelihood function is composed of probability density functions (PDFs) describing the signal and background components. The likelihood fit algorithm is implemented using the RooFit package from the ROOT framework [78]. The signal and background PDFs are formed as the product of PDFs, which model the invariant mass distribution and the time dependent decay rates of the reconstructed candidates. In addition, the signal PDF also includes the three-dimensional efficiency function. The event likelihood function $\mathcal{L}$ can be represented as

$$\mathcal{L} = L_{\text{sig}} + L_{\text{bkg}} \tag{5.6}$$

$$L_{\text{sig}} = N_{\text{sig}} \cdot \left(\tilde{f}(\Theta,\alpha,t) \otimes G(t,\sigma_t) \cdot \epsilon(\Theta)\right) \cdot \\ \cdot P_{\text{sig}}\left(m_{B_s^0}\right) \cdot P_{\text{sig}}(\sigma_t) \cdot P_{\text{sig}}(\xi) \tag{5.7}$$

$$L_{\text{bkg}} = N_{\text{bkg}} \cdot P_{\text{bkg}}(\cos\theta_T, \varphi_T) \cdot P_{\text{bkg}}(\cos\psi_T) \cdot \\ \cdot P_{\text{bkg}}(t) \cdot P_{\text{bkg}}\left(m_{B_s^0}\right) \cdot P_{\text{bkg}}(\sigma_t) \cdot P_{\text{bkg}}(\xi) \tag{5.8}$$

where L_{sig} is the PDF that describes the $B_s^0 \to J/\psi\,\phi$ signal model and L_{bkg} describes the background contributions. The PDF $\tilde{f}(\Theta,\alpha,t)$ is the differential decay rate function including the tagging information as described in Eq. 5.5. In the model $\tilde{f}\ \delta_0$ is set to zero, and the difference of phases $\delta_S - \delta_\perp$ is fitted with a unique variable $\delta_{S\perp}$ to reduce the correlation among the fitted parameters. Here $\epsilon(\Theta)$ is the angular efficiency function and G is a Gaussian resolution function, which makes use of the event-by-event proper decay time uncertainty σ_t scaled by a factor κ , which is a function of t. The kappa factor is a scale factor introduced to correct the proper decay time uncertainty in order to resemble the actual resolution. It is measured in simulated samples assuming that the kappa factor is the same as in data, for this assumption a systematic uncertainty is evaluated. All the parameters of the PDFs are left free to float in the final fit, unless explicitly stated otherwise. The value of $\Delta\Gamma_\mathrm{s}$ is constrained to be positive as recently measured by the LHCb collaboration [21].

The signal mass PDF $P_\mathrm{sig}(m_{\mathrm{B_s^0}})$ is given by the sum of three Gaussian functions with a common mean; the two smaller widths, the mean and the fraction of the Gaussians are fixed to the values obtained in a one-dimensional mass fit to the full dataset. The proper decay time uncertainty signal PDF $P_\mathrm{sig}(\sigma_t)$ is a sum of two Gamma functions; all the parameters are fixed to the values obtained fitting a sample of background-subtracted events. The background mass distribution $P_\mathrm{bkg}(m_{\mathrm{B_s^0}})$ is described by an exponential function. The background proper decay time component $P_\mathrm{bkg}(t)$ is described by the sum of two exponential functions. The angular part of the background PDFs $P_\mathrm{bkg}(\cos\theta_T, \varphi_T)$ and $P_\mathrm{bkg}(\cos\psi_T)$ are described analytically by a series of Legendre polynomials for $\cos\theta_T$ and $\cos\psi_T$ and sinusoidal functions used for the angle φ_T. For the $\cos\theta_T$ and φ_T variables a two-dimensional PDF is used to take into account the correlation among the variables. The proper decay time uncertainty background PDF $P_\mathrm{bkg}(\sigma_t)$ is represented by a single Gamma function; all the parameters are fixed to the values obtained fitting the mass peak sideband events. Finally, $P_\mathrm{sig}(\xi)$ and $P_\mathrm{bkg}(\xi)$ are the tag decision ξ PDFs which have been obtained from the data sample.

5.6 Results and Systematic Uncertainties

The fit on the 2012 CMS data is performed constraining the value of Δm_s to the current world average value $(17.69 \pm 0.08) \times 10^{12}\ \hbar/\mathrm{s}$ [18] by adding to the likelihood a Gaussian distribution centred on the world average and with the uncertainty as the width. The fit results are given in Table 5.3, where the uncertainties are statistical only. The observable distributions and the fit projections are shown in Figs. 5.3, 5.4, and 5.5. The 68%, 90% and 95% Confidence Level (C.L.) likelihood contours of the fit for ϕ_s and $\Delta\Gamma_\mathrm{s}$ are shown in Fig. 5.6.

Several sources of systematic uncertainties in the physics parameters are considered by testing the various assumptions made in the fit model and those associated with the fit procedure.

The systematic uncertainty associated with the hypothesis of a flat proper decay time efficiency is evaluated by fitting again the data with the a proper decay time efficiency which takes into account a small contribution of the time significance cut at small t and a first order polynomial variations at high t.

The uncertainties associated with the variables of three-dimensional angular efficiency

Table 5.3: Results of the fit to the 2012 data. Only the statistical uncertainties are reported in the table.

Parameter	Fit result
$\|A_0\|^2$	0.510 ± 0.005
$\|A_S\|^2$	0.012 ± 0.008
$\|A_\perp\|^2$	0.243 ± 0.008
$\delta_\parallel$ [rad]	3.48 ± 0.08
$\delta_{S\perp}$ [rad]	0.37 ± 0.17
$\delta_\perp$ [rad]	2.98 ± 0.36
$c\tau$ [μm]	447 ± 3
$\Delta\Gamma_s$ [ps^{-1}]	0.095 ± 0.013
ϕ_s [rad]	-0.075 ± 0.097

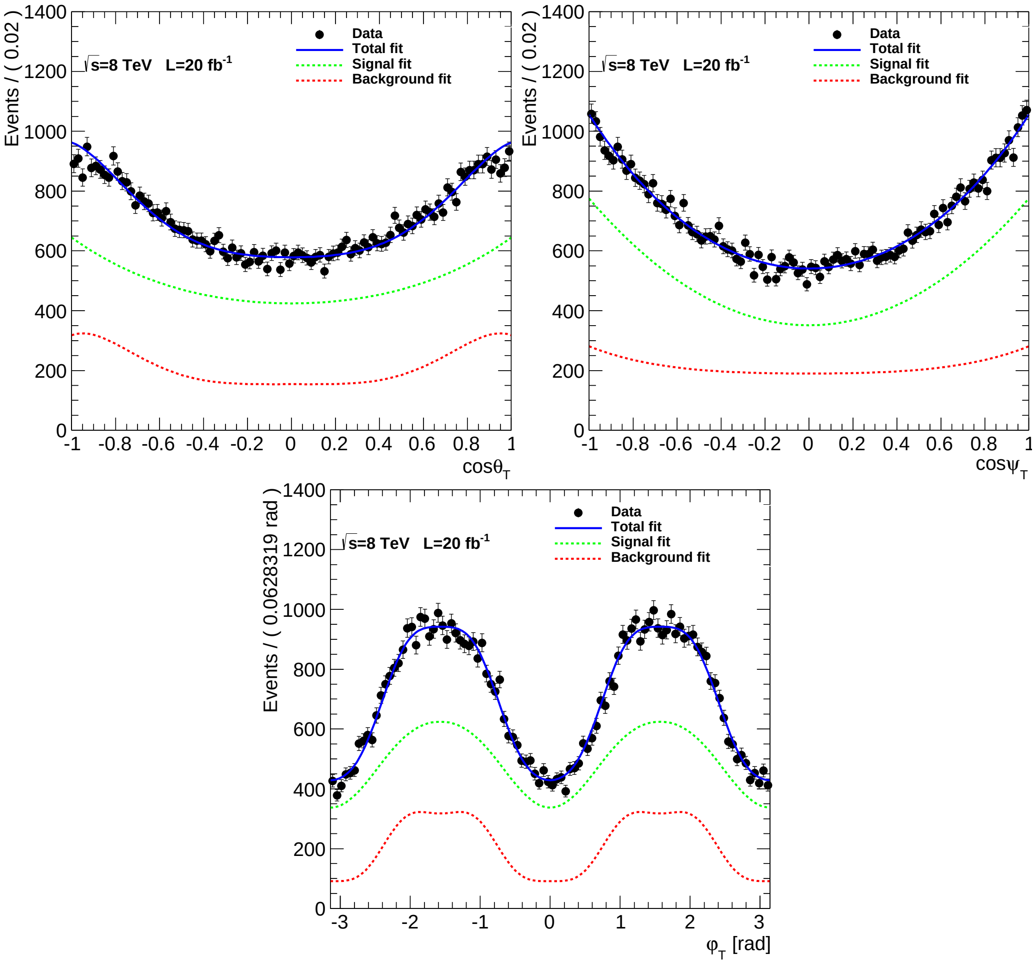


Figure 5.5: Projections of the data (solid markers) along the three angular variables ($\cos\theta_T$, $\cos\psi_T$, φ_T), compared to the fit result. The blue line is the projection of the full likelihood, while the dashed green and red lines represent the fitted signal and background respectively.

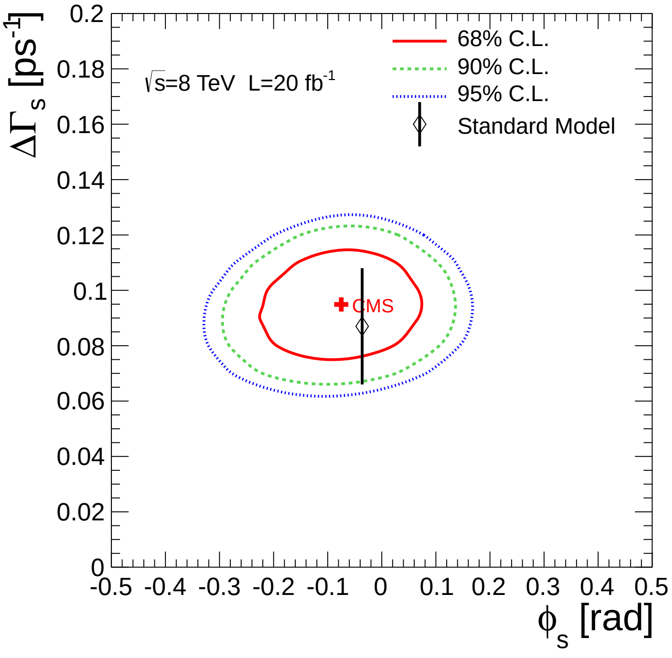


Figure 5.6: The 68%, 90% and 95% C.L. contours in the $\Delta\Gamma_s$ versus ϕ_s plane, together with the SM fit prediction. Uncertainties are statistical only.

function $\cos\theta_T$, φ_T, and $\cos\psi_T$ are propagated to the fit results by varying the corresponding parameters within their statistical uncertainties and accounting for the covariances among the parameters. The maximum variation in the physics parameters is taken as the systematic uncertainty. The systematic uncertainty due to a small discrepancy in the kaon p_T spectrum between the data and the simulations is evaluated by re-weighting the simulated kaon pt spectrum to agree with the data.

The biases, which could be intrinsic to the fit model itself, are also taken into account. The nominal model function is tested by using simulated pseudo-experiments and using the bias of the pull as a systematic uncertainty if it exceeds the statistical uncertainty of the bias.

The uncertainty in the proper decay time resolution associated with the proper decay time uncertainty scale factor κ is propagated to the results. In order to evaluate the systematic uncertainties arising from the scale factor, hundreds of test samples are produced by varying the $\kappa(t)$ factors within their statistical error in a Gaussian fashion. The difference with respect to the nominal fit is investigated, and one standard deviation of the obtained distribution is taken as the systematic uncertainty. Since the $\kappa(t)$ factors are obtained from simulation, the associated systematic uncertainty is assessed by using a sample of prompt J/ψ decays obtained with an unbiased trigger and comparing them to similarly processed simulated data. In this way the proper decay time resolution for $t \approx 0$ is obtained. The difference of the resolutions between data and simulation are used to vary the $\kappa(t)$ factor. The resulting variations of the physics parameters are added to the systematics uncertainties.

Although the likelihood function makes use of a per-event mistag parameter, it does not contain a PDF model for the mistag distribution. Therefore a systematic uncertainty is estimated by generating simulated pseudo-experiments with different independent mistag distributions for signal and background and fitting them with the nominal fit.

The systematic uncertainty due to flavour tagging is assessed by propagating independently the statistical and systematic uncertainties of the combined lepton tagger per-event mistag to the final fit. As previously described in Section 4.3.5 the mistag statistical uncertainty is obtained by evaluating the mistag calibration curve using the $B^+ \to J/\psi K^+$ data sample, and the systematic uncertainty (largely dominated by the same side B flavour dependency) is estimated by the calibration curve assessed comparing the mistag in the simulated B_s^0 and B^+ samples. The systematic contributions on the fit results is evaluated by measuring the average biases induced on the each fit parameters by generating simulated pseudo-experiments according to the statistical and systematic mistag uncertainties. The total tagging systematic effect is finally evaluated as the square root of sum of squares of statistical and systematic mistag uncertainty contributions for each physics parameter extracted by the fit. An additional uncertainty is assigned to the possible bias in the mistag fractions coming from the probability distribution function used to extract wrongly tagged and tagged signal events from the background.

The various hypotheses that have been assumed when building the likelihood function are tested by generating simulated pseudo-experiments with different hypotheses in the generated samples and fitting the samples with the nominal likelihood function. The obtained pull histograms of the physics variables are fitted with Gaussian functions, and the bias of the pull is used as a systematic uncertainty if the difference with respect to the model bias exceeds the statistical uncertainty. The model of the $J/\psi KK$ invariant mass distribution the background model is changed to a Chebyshev function PDF from the nominal exponential PDF. The proper decay time background PDF is changed to a triple-exponential instead of the double-exponential of the nominal fit. The angular background PDF is generated by using the background simulation angular shapes instead of the fitted ones. The effect of not including the angular resolution is also tested. The contribution to the systematic uncertainty from the background tagging asymmetry is negligible.

Finally the hypothesis that $|\lambda| = 1$ is tested by leaving that parameter free in the fit. The value extracted by the fit agrees with the hypothesis $|\lambda| = 1$ within one standard deviation. The differences found in the fit results with respect to the nominal fit are used as systematic uncertainties.

The systematic uncertainties are summarised in Table 5.4. The uncertainties of the ϕ_s and $\Delta\Gamma_s$ results are dominated by statistical uncertainties.

The measured values for the weak phase and the decay width difference between the B_s^0 mass eigenstates are:

$$\phi_s = -0.075 \pm 0.097 \text{ (stat.)} \pm 0.031 \text{ (syst.) rad} \tag{5.9}$$

$$\Delta\Gamma_s = 0.095 \pm 0.013 \text{ (stat.)} \pm 0.007 \text{ (syst.) ps}^{-1} \tag{5.10}$$

The value of ϕ_s is in agreement with the previous measurements and with the SM fit prediction. The value of $\Delta\Gamma_s$ is confirmed to be non-zero. The uncertainties of the ϕ_s and $\Delta\Gamma_s$ results are mainly statistical.

Table 5.4: Summary of the uncertainties. If no value is reported, then the systematic uncertainty is negligible with respect to the statistical and other systematic uncertainties. The total systematic uncertainty is the square root of sum of squares of the listed systematic uncertainties.

Source of uncertainty	$\lvert A_0\rvert^2$	$\lvert A_S\rvert^2$	$\lvert A_\perp\rvert^2$	$\Delta\Gamma_s$ [ps^{-1}]	$\delta_\parallel$ [rad]	$\delta_{S\perp}$ [rad]	$\delta_\perp$ [rad]	ϕ_s [rad]	$c\tau$ [μm]
Statistical uncertainty	0.0058	0.016	0.0077	0.0138	0.092	0.24	0.36	0.109	3.0
Proper time efficiency	0.0015	-	0.0023	0.0057	-	-	-	0.002	1.0
Angular efficiency	0.0060	0.008	0.0104	0.0021	0.674	0.14	0.66	0.016	0.8
Model bias	0.0008	-	-	0.0012	0.025	0.03	-	0.015	0.4
Proper time resolution	0.0009	-	0.0008	0.0021	0.004	-	0.02	0.006	2.9
Background mistag modelling	0.0006	-	-	0.0018	0.008	0.01	-	0.004	0.1
Flavour tagging	-	-	-	0.0003	0.006	0.02	-	0.003	-
PDF modelling assumptions	0.0016	0.002	0.0021	0.0003	0.010	0.03	0.04	0.006	0.2
$\lvert\lambda\rvert$ as a free parameter	0.0001	0.005	0.0001	0.0003	0.002	0.01	0.03	0.015	-
Kaon p_T re-weighting	0.0094	0.020	0.0041	0.0015	0.085	0.11	0.02	0.014	1.1
Total systematics	0.0114	0.022	0.0116	0.0069	0.680	0.18	0.66	0.031	3.4

Conclusions

This thesis presented the development and optimization of an algorithm for the identification of the production flavour of neutral b-mesons at the CMS experiment. The algorithm infers the b-meson flavour by the charge of the muons and electrons arising from the semileptonic decay of the other b-hadron produced in the event by the $\mathrm{b\overline{b}}$ pair production mechanism. A neural network was trained on simulated events to separate leptons contributing to the right and wrong decision of the algorithm. The probability of wrong flavour attribution (mistag) was parametrised according to the output of the neural network in order to provide a per-event value of the predicted mistag probability. The mistag was measured on data using the self-tagging decay of charged meson $\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$, and the performances of the algorithm was evaluated in $\mathcal{P}_{tag} = \varepsilon_{tag}(1-2\omega)^2$ of 0.833 ± 0.024 (stat.) $\pm\, 0.012$ (syst.) % for muons and 0.483 ± 0.020 (stat.) $\pm\, 0.003$ (syst.) % for electrons. The two algorithms were combined resulting in the overall flavour identification power of 1.307 ± 0.031 (stat.) $\pm\, 0.007$ (syst.) %. The systematic uncertainty of the flavour tagging algorithm was found to be dominated by the residual dependence on the flavour of the reconstructed b-meson, and was evaluated using simulated events of $\mathrm{B^0_s} \rightarrow \mathrm{J}/\psi\,\phi$, $\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\,\mathrm{K}^+$ and $\mathrm{B}^0 \rightarrow \mathrm{J}/\psi\,\mathrm{K}^*$ decays.

The flavour tagging algorithm was exploited to measure the weak CP-violating phase ϕ_s and the decay width difference $\Delta\Gamma_\mathrm{s}$ in $\mathrm{B^0_s} \rightarrow \mathrm{J}/\psi\,\phi$ decays. A complete time-dependent angular analysis was performed on 49 000 events reconstructed $\mathrm{B^0_s}$ using the data collected during 2012 by the CMS experiment, corresponding to 20 fb^{-1}. The measured parameters are:

$$\begin{aligned} \phi_\mathrm{s} &= -0.075 \pm 0.097\ \text{(stat.)} \pm 0.031\ \text{(syst.) rad} \\ \Delta\Gamma_\mathrm{s} &= 0.095 \pm 0.013\ \text{(stat.)} \pm 0.007\ \text{(syst.) ps}^{-1} \end{aligned}$$

The value of $\Delta\Gamma_\mathrm{s}$ was confirmed to be non-zero, while ϕ_s was found to be in agreement with the most recent standard model predictions. The uncertainties of ϕ_s and $\Delta\Gamma_\mathrm{s}$ resulted dominated by the statistical contributions.

References

[1] S. L. Glashow, "Partial-symmetries of weak interactions," *Nucl. Phys.*, vol. 22, no. 4, pp. 579–588, 1961.

[2] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967.

[3] A. Salam and W. J. C., "Electromagnetic and weak interactions," *Phys. Lett.*, vol. 13, no. 2, pp. 168–171, 1964.

[4] J. Goldstone, A. Salam, and S. Weinberg, "Broken symmetries," *Phys. Rev.*, vol. 127, pp. 965–970, Aug 1962.

[5] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak interactions with lepton-hadron symmetry," *Phys. Rev. D*, vol. 2, pp. 1285–1292, Oct 1970.

[6] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, "Advantages of the colour octet gluon picture," *Phys. Lett. B*, vol. 47, no. 4, pp. 365–368, 1973.

[7] M. Gell-Mann, "Symmetries of baryons and mesons," *Phys. Rev.*, vol. 125, pp. 1067–1084, Feb 1962.

[8] P. W. Anderson, "Plasmons, gauge invariance, and mass," *Phys. Rev.*, vol. 130, pp. 439–442, Apr 1963.

[9] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.

[10] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.

[11] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global conservation laws and massless particles," *Phys. Rev. Lett.*, vol. 13, pp. 585–587, Nov 1964.

[12] T. W. B. Kibble, "Symmetry breaking in non-abelian gauge theories," *Phys. Rev.*, vol. 155, pp. 1554–1561, Mar 1967.

[13] R. P. Feynman, "Mathematical formulation of the quantum theory of electromagnetic interaction," *Phys. Rev.*, vol. 80, pp. 440–457, Nov 1950.

[14] N. Cabibbo, "Unitary symmetry and leptonic decays," *Phys. Rev. Lett.*, vol. 10, pp. 531–533, Jun 1963.

[15] M. Kobayashi and T. Maskawa, "CP violation in the renormalizable theory of weak interaction," *Progr. Theor. Exp. Phys.*, vol. 49, no. 2, pp. 652–657, 1973.

[16] J. Charles *et al.*, "Cp violation and the CKM matrix: Assessing the impact of the asymmetric B factories," *Eur. Phys. J.*, vol. C41, pp. 1–131, 2005. Updated results and plots available at http://ckmfitter.in2p3.fr.

[17] L. Wolfenstein, "Parametrization of the kobayashi-maskawa matrix," *Phys. Rev. Lett.*, vol. 51, pp. 1945–1947, Nov 1983.

[18] K. A. Olive *et al.*, "Review of particle physics," *Chin. Phys.*, vol. C38, p. 090001, 2014.

[19] C. Jarlskog, "Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP nonconservation," *Phys. Rev. Lett.*, vol. 55, pp. 1039–1042, Sep 1985.

[20] G. C. Branco, L. Lavoura, and J. P. Silva, *"CP Violation"*. International series of monographs on physics, Clarendon Press, 1999.

[21] R. Aaij *et al.*, "Determination of the sign of the decay width difference in the B_s^0 system," *Phys. Rev. Lett.*, vol. 108, p. 241801, 2012.

[22] I. Dunietz, R. Fleischer, and U. Nierste, "In pursuit of new physics with B_s^0 decays," *Phys. Rev. D*, vol. 63, p. 114015, May 2001.

[23] R. Aaij *et al.*, "Measurement of the semileptonic CP asymmetry in B_s^0-$\overline{B_s^0}$ mixing," 2014.

[24] S. Myers, *The LEP Collider, from design to approval and commissioning*. John Adams' Lecture, Geneva: CERN, 1991.

[25] E. Lyndon and B. Philip, "LHC machine," *JINST*, vol. 3, no. 08, p. S08001, 2008.

[26] G. Aad *et al.*, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST*, vol. 3, p. S08003, 2008.

[27] S. Chatrchyan *et al.*, "The CMS experiment at the CERN LHC," *JINST*, vol. 3, p. S08004, 2008.

[28] K. Aamodt *et al.*, "The ALICE experiment at the CERN LHC," *JINST*, vol. 3, p. S08002, 2008.

[29] A. A. J. Alves *et al.*, "The LHCb Detector at the LHC," *JINST*, vol. 3, p. S08005, 2008.

[30] S. Chatrchyan *et al.*, "Description and performance of track and primary-vertex reconstruction with the CMS tracker," *JINST*, vol. 9, no. 10, p. P10009, 2014.

[31] S. Chatrchyan *et al.*, "Description and performance of track and primary-vertex reconstruction with the CMS tracker," *JINST*, vol. 9, no. 10, p. 10009, 2014.

[32] R. Frühwirth, "Application of kalman filtering to track and vertex fitting," *Nucl. Instr. Meth. Phys. Res. Section A*, vol. 262, no. 2-3, pp. 444–450, 1987.

[33] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems," *Proc. IEEE*, vol. 86, pp. 2210–2239, Nov 1998.

[34] W. Waltenberger, R. Frühwirth, and P. Vanlaer, “Adaptive vertex fitting,” *J. Phys. G*, vol. 34, no. 12, p. 343, 2007.

[35] S. Chatrchyan *et al.*, “The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC,” *JINST*, vol. 8, no. 11, p. 11002, 2013.

[36] S. Chatrchyan *et al.*, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV,” *JINST*, vol. 7, no. 10, p. 10002, 2012.

[37] S. Chatrchyan *et al.*, “Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV,” 2010.

[38] S. Chatrchyan *et al.*, “Measurement of the $\mathrm{B}^0_\mathrm{s} \to \mu^+ \mu^-$ branching fraction and search for $\mathrm{B}^0 \to \mu^+ \mu^-$ with the CMS experiment,” *Phys. Rev. Lett.*, vol. 111, p. 101804, Sep 2013.

[39] V. Khachatryan *et al.*, “Observation of the rare $\mathrm{B}^0_\mathrm{s} \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data,” 2014. Submitted to Nature.

[40] J. Pazzini *et al.*, “Observation of $\mathrm{B}^0_\mathrm{s} \to \mu^+\mu^-$ and search for $\mathrm{B}^0 \to \mu^+\mu^-$ with the 2011 and 2012 data,” CMS Note 2013/009.

[41] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, “Reconstruction of electrons with the gaussian-sum filter in the CMS tracker at the LHC,” *J. Phys. G*, vol. 31, no. 9, p. 9, 2005.

[42] “Electron reconstruction and identification at $\sqrt(s) = 7$ TeV,” Tech. Rep. CMS-PAS-EGM-10-004, CERN, Geneva.

[43] “Photon reconstruction and identification at $\sqrt(s) = 7$ TeV,” Tech. Rep. CMS-PAS-EGM-10-005, CERN, Geneva.

[44] C. Collaboration, “Particle-flow event reconstruction in CMS and performance for jets, taus, and $\mathrm{E}_\mathrm{T}^\mathrm{miss}$,” Tech. Rep. CMS-PAS-PFT-09-001, CERN, Geneva.

[45] C. Collaboration, “Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector,” Tech. Rep. CMS-PAS-PFT-10-001, 2010.

[46] G. Petrucciani, *“Observation of a new state in the search for the Higgs boson at CMS”*. 2013.

[47] A. J. Bevan *et al.*, “The physics of the b factories,” 2014.

[48] A. Abulencia *et al.*, “Observation of B^0_s-$\overline{\mathrm{B}^0_\mathrm{s}}$ oscillations,” *Phys. Rev. Lett.*, vol. 97, p. 242003, 2006.

[49] V. M. Abazov *et al.*, “First direct two-sided bound on the B^0_s oscillation frequency,” *Phys. Rev. Lett.*, vol. 97, p. 021802, 2006.

[50] R. Aaij *et al.*, “Opposite-side flavour tagging of B mesons at the LHCb experiment,” *Eur. Phys. J.*, vol. C72, p. 2022, 2012.

[51] G. Aad *et al.*, "Flavour tagged time dependent angular analysis of the $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ decay and extraction of $\Delta\Gamma_\mathrm{s}$ and the weak phase ϕ_s in ATLAS," *Phys. Rev.*, vol. D90, p. 052007, 2014.

[52] G. A. Krocker, "Development and calibration of a same side kaon tagging algorithm and measurement of the $\mathrm{B_s^0}$-$\overline{\mathrm{B_s^0}}$ oscillation frequency Δm_s at the LHCb experiment," 2013.

[53] V. Khachatryan *et al.*, "CMS tracking performance results from early LHC operation," *Eur. Phys. J.*, vol. C70, no. 4, pp. 1165–1192, 2010.

[54] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 physics and manual," *JHEP*, vol. 0605, p. 026, 2006.

[55] J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP*, vol. 0207, p. 012, 2002.

[56] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instr. Meth. A*, vol. 462, no. 1-2, pp. 152–155, 2001.

[57] E. Barberio, B. van Eijk, and Z. Was, "PHOTOS - a universal monte carlo for QED radiative corrections in decays," *Comput. Phys. Commun.*, vol. 66, no. 1, pp. 115–128, 1991.

[58] E. Barberio and Z. Was, "PHOTOS - a universal monte carlo for QED radiative corrections: version 2.0," *Comput. Phys. Commun.*, vol. 79, no. 2, pp. 291–308, 1994.

[59] S. Agostinelli *et al.*, "Geant4 - a simulation toolkit," *Nucl. Instr. Meth. A*, vol. 506, no. 3, pp. 250–303, 2003.

[60] S. Chatrchyan *et al.*, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV," *JINST*, vol. 7, p. 10002, 2012.

[61] S. De Visscher and A. Zucchetta, "Tagging b-jets with electrons and muons in 8 TeV pp collisions at CMS," CMS Note 2013/142.

[62] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, "TMVA: Toolkit for multivariate data analysis," *PoS*, vol. ACAT, p. 040, 2007.

[63] A. Lenz and U. Nierste, "Theoretical update of $\mathrm{B_s^0}$-$\overline{\mathrm{B_s^0}}$ mixing," *JHEP*, vol. 2007, no. 06, p. 072, 2007.

[64] J. Charles *et al.*, "Predictions of selected flavor observables within the standard model," *Phys. Rev. D*, vol. 84, p. 033005, Aug 2011.

[65] T. Aaltonen *et al.*, "First Flavor-Tagged Determination of Bounds on Mixing-Induced CP Violation in $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ Decays," *Phys. Rev. Lett.*, vol. 100, p. 161802, 2008.

[66] V. M. Abazov *et al.*, "Measurement of $\mathrm{B_s^0}$ mixing parameters from the flavor-tagged decay $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$," *Phys. Rev. Lett.*, vol. 101, p. 241801, 2008.

[67] T. Aaltonen *et al.*, "Measurement of the CP Violating Phase $\beta_\mathrm{s}^{\mathrm{J}/\psi\phi}$ in $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ Decays with the CDF II Detector," *Phys. Rev.*, vol. D85, p. 072002, 2012.

[68] V. M. Abazov *et al.*, “Measurement of the CP Violating phase $\phi_s^{J/\psi\phi}$ using the flavor-tagged decay $B_s^0 \to J/\psi\,\phi$ in 8 fb^{-1} of $p\bar{p}$ collisions,” *Phys. Rev.*, vol. D85, p. 032006, 2012.

[69] R. Aaij *et al.*, “Measurement of the CP Violating phase ϕ_s in $B_s^0 \to J/\psi\,\pi^+,\pi^-$ decays,” 2014. Submitted to Phys. Lett. B.

[70] R. Aaij *et al.*, “Measurement of CP Violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi\,K^+,K^-$ and $B_s^0 \to J/\psi\,\pi^+,\pi^-$ decays,” *Phys. Rev.*, vol. D87, p. 112010, 2013.

[71] R. Aaij *et al.*, “Measurement of the CP Violating phase ϕ_s in $B_s^0 \to J/\psi\,\pi^+,\pi^-$ decays,” *Phys. Lett.*, vol. B713, pp. 378–386, 2012.

[72] G. Aad *et al.*, “Time-dependent angular analysis of the decay $B_s^0 \to J/\psi\,\phi$ and extraction of $\Delta\Gamma_s$ and the CP Violating weak phase ϕ_s by ATLAS,” *JHEP*, vol. 1212, p. 072, 2012.

[73] M. Prim *et al.*, “Angular analysis of $B^0 \to J/\psi\,K^*$ decays and search for CP Violation at Belle,” *Phys. Rev.*, vol. D88, no. 7, p. 072004, 2013.

[74] I. Dunietz, H. Quinn, A. Snyder, W. Toki, and H. J. Lipkin, “How to extract CP violating asymmetries from angular correlations,” *Phys. Rev. D*, vol. 43, pp. 2193–2208, Apr 1991.

[75] T. Aaltonen *et al.*, “Measurement of polarization and search for CP violation in $B_s^0 \to \phi\,\phi$ decays,” *Phys. Rev. Lett.*, vol. 107, p. 261802, Dec 2011.

[76] A. S. Dighe, I. Dunietz, and R. Fleischer, “Extracting CKM phases and B_s^0-$\overline{B_s^0}$ mixing parameters from angular distributions of non-leptonic b decays,” *Eur. Phys. J. C*, vol. 6, p. 647, 1999.

[77] A. S. Dighe, I. Dunietz, H. J. Lipkin, and J. L. Rosner, “Angular distributions and lifetime differences in $B_s^0 \to J/\psi\,\phi$ decays,” *Phys. Lett.*, vol. B369, pp. 144–150, 1996.

[78] I. Antcheva *et al.*, “ROOT - a C++ framework for petabyte data storage, statistical analysis and visualization,” *Comput. Phys. Commun.*, vol. 180, no. 12, pp. 2499–2512, 2009.

Appendix A

Additional Plots and Tables

A.1 OS-Muon tagger

Figures A.1 and A.2 show the distribution of the OS-muon variables for the simulated samples $B_s^0 \to J/\psi\,\phi$ and $B^+ \to J/\psi\,K^+$ after the OS-muon selection. The comparison between the two MC samples is shown in Figure A.4.

In Figures A.5 and A.7 the muons variables are divided into the three categories defined in MC: muons from B with correct charge-flavour relation (CC), muons from B with opposite charge-flavour relation (WC), background muons not originated from a B decay (RC).

The data distributions are shown in Figure A.8. For each variable's bin a fit to the B^+ invariant mass is performed in order to obtain the signal yield of $B^+ \to J/\psi\,K^+$ events. The comparison among the OS-muon distributions obtained in 2012 data and in the $B^+ \to J/\psi\,K^+$ Monte Carlo samples is shown in Figure A.9.

Figures A.10 and A.11 show the fit to B^+ invariant mass for the reconstructed $B^+ \to J/\psi\,K^+$ events, performed for each of the 20 bins of the neural network discriminator (MLP) for OS-muons.

The bin-by bin tagging performances evaluated using twenty neural network discriminator categories are presented in Table A.1, and the combined results are summarized in Table A.2, for the $B_s^0 \to J/\psi\,\phi$ MC. Analogous tables are presented for the $B^+ \to J/\psi\,K^+$ MC in Tables A.3,A.4; for the $B^0 \to J/\psi\,K^*$ MC in Tables A.5,A.6; and for the $B^+ \to J/\psi\,K^+$ 2012 data in Tables A.7,A.8.

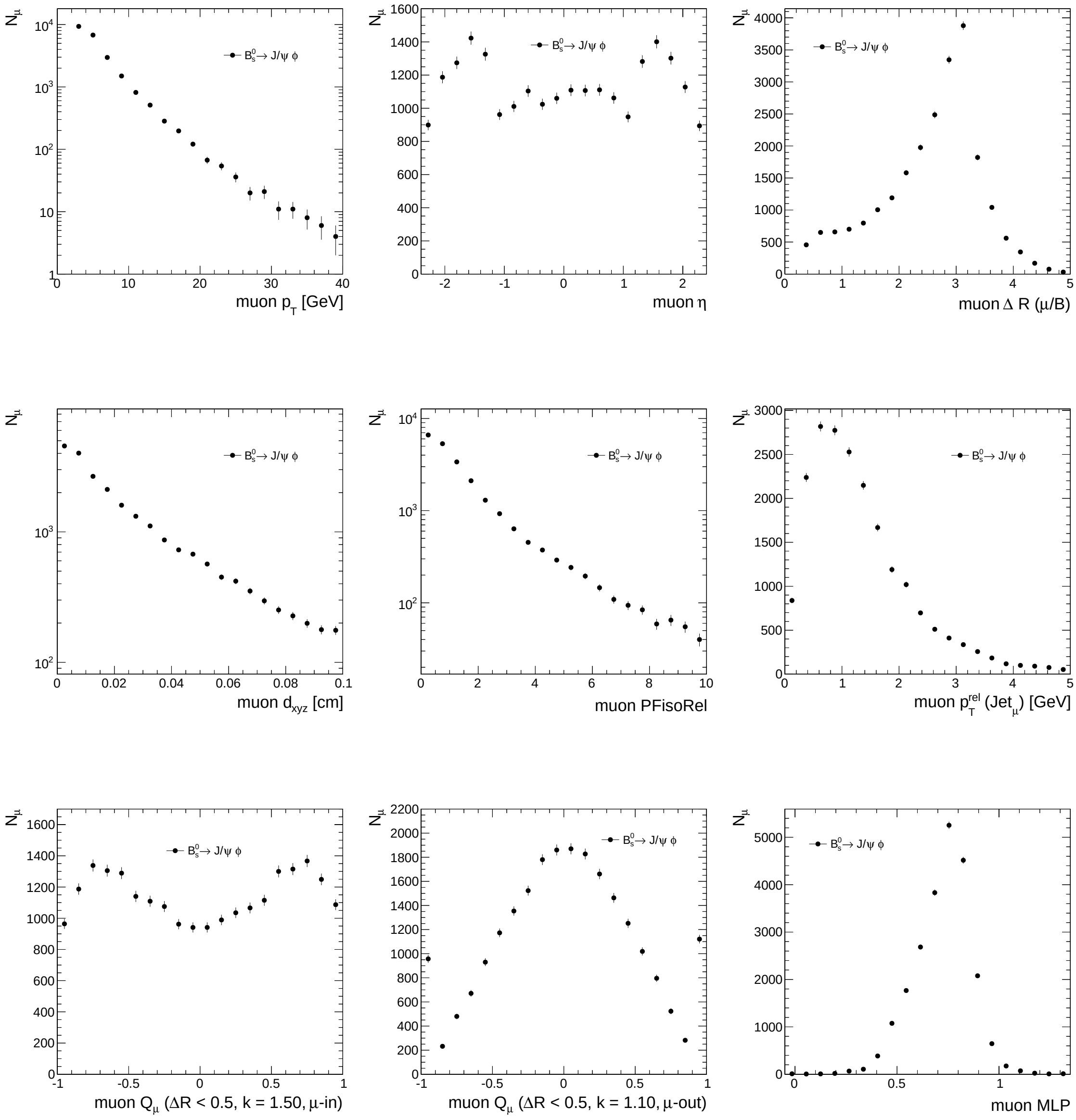


Figure A.1: OS-muon variables in the $B_s^0 \to J/\psi\,\phi$ simulation, after the muon selection. From left to right, top to bottom: muon transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, p_T^{rel}, Q_μ ($\Delta R < 0.5$, $k = 1.50$, μ-in), Q_μ ($\Delta R < 0.5$, $k = 1.10$, μ-out), Neural Network output (MLP).

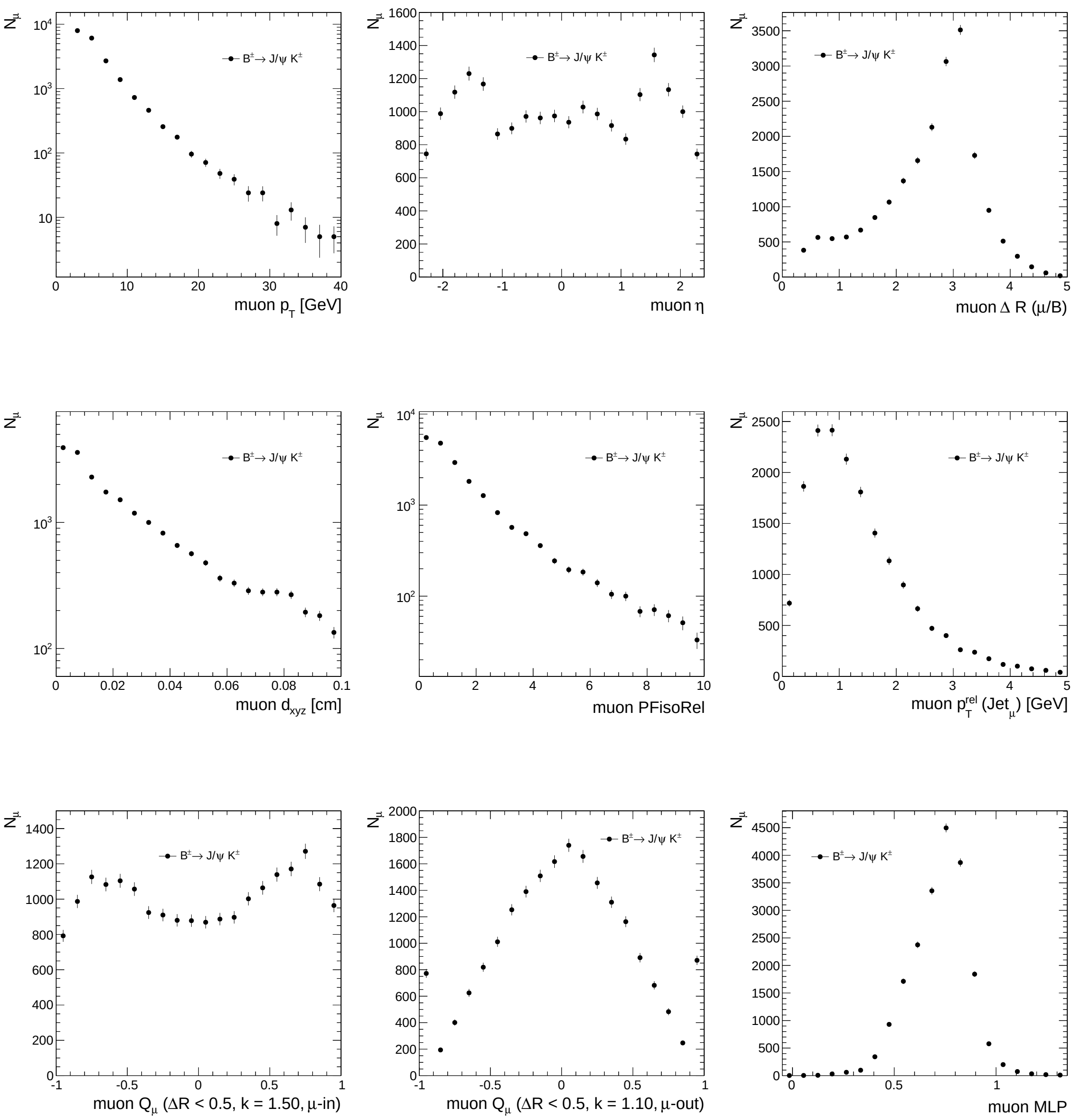


Figure A.2: OS-muon variables in the $B^+ \rightarrow J/\psi\, K^+$ simulation, after the muon selection. From left to right, top to bottom: muon transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, p_T^{rel}, Q_μ ($\Delta R < 0.5$, $k = 1.50$, μ-in), Q_μ ($\Delta R < 0.5$, $k = 1.10$, μ-out), Neural Network output (MLP).

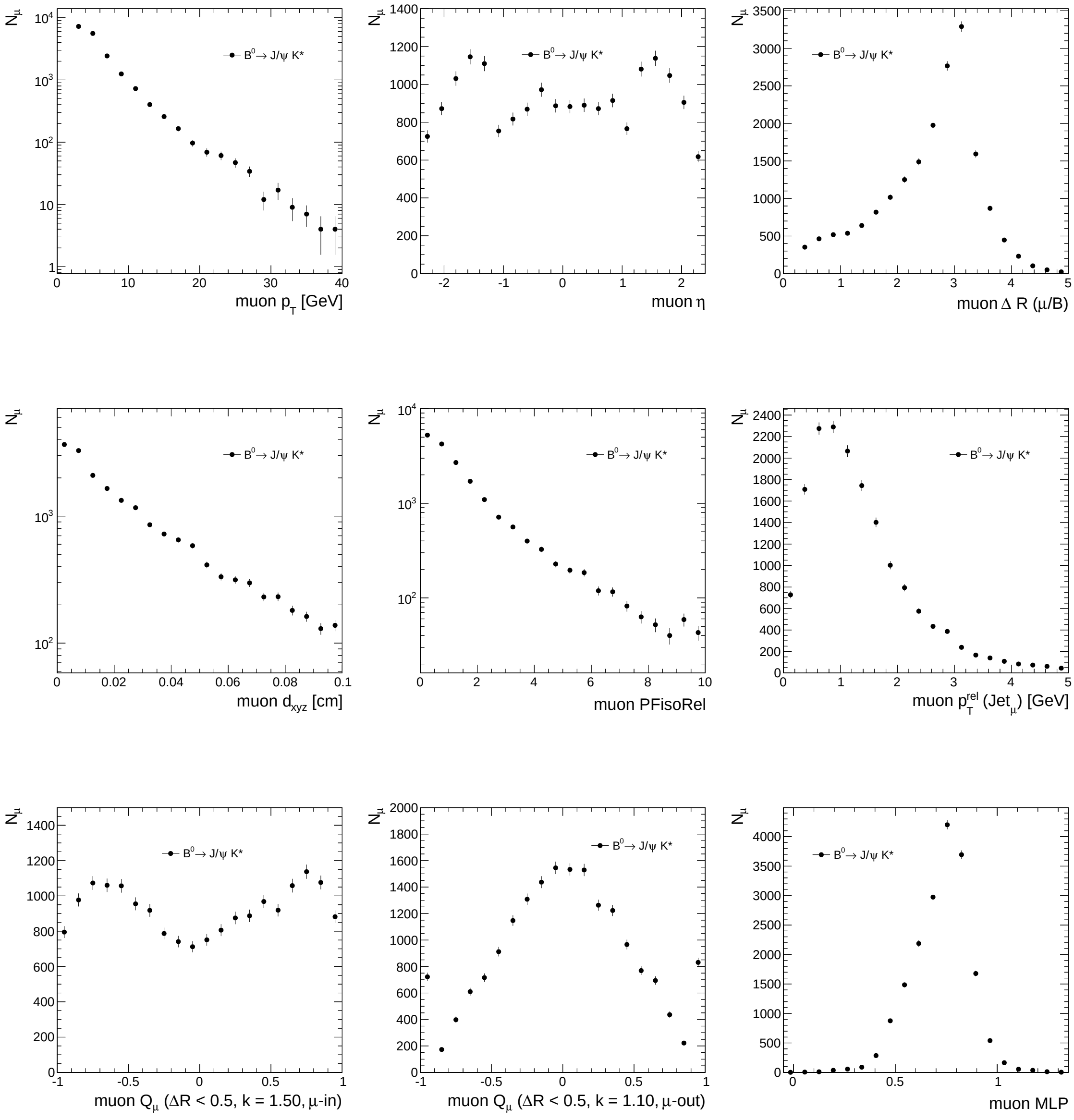


Figure A.3: OS-muon variables in the $B^0 \to J/\psi\, K^*$ simulation, after the muon selection. From left to right, top to bottom: muon transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, p_T^{rel}, Q_μ ($\Delta R < 0.5$, $k = 1.50$, μ-in), Q_μ ($\Delta R < 0.5$, $k = 1.10$, μ-out), Neural Network output (MLP).

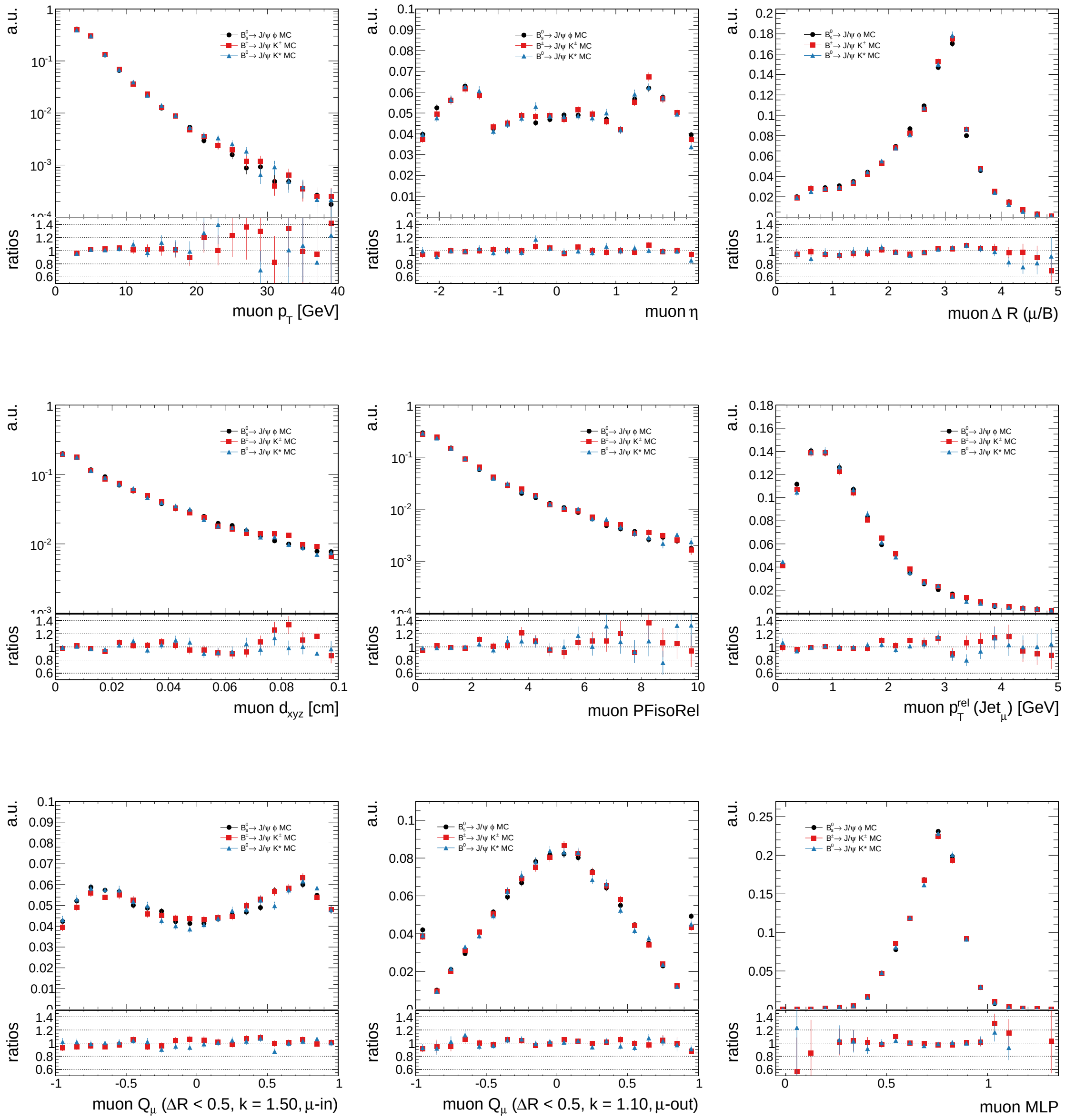


Figure A.4: Comparison of $B_s^0 \to J/\psi\,\phi$ (black), $B^+ \to J/\psi\,K^+$ (red) and $B^0 \to J/\psi\,K^*$ (blue) simulations after the muon selection is applied. The distributions are normalized to the same integral, and the bin-by-bin ratio of the $B^+ \to J/\psi\,K^+$ ($B^0 \to J/\psi\,K^*$) over the $B_s^0 \to J/\psi\,\phi$ is also shown. From left to right, top to bottom: muon transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, p_T^{rel}, Q_μ ($\Delta R < 0.5$, $k = 1.50$, μ-in), Q_μ ($\Delta R < 0.5$, $k = 1.10$, μ-out), Neural Network output (MLP).

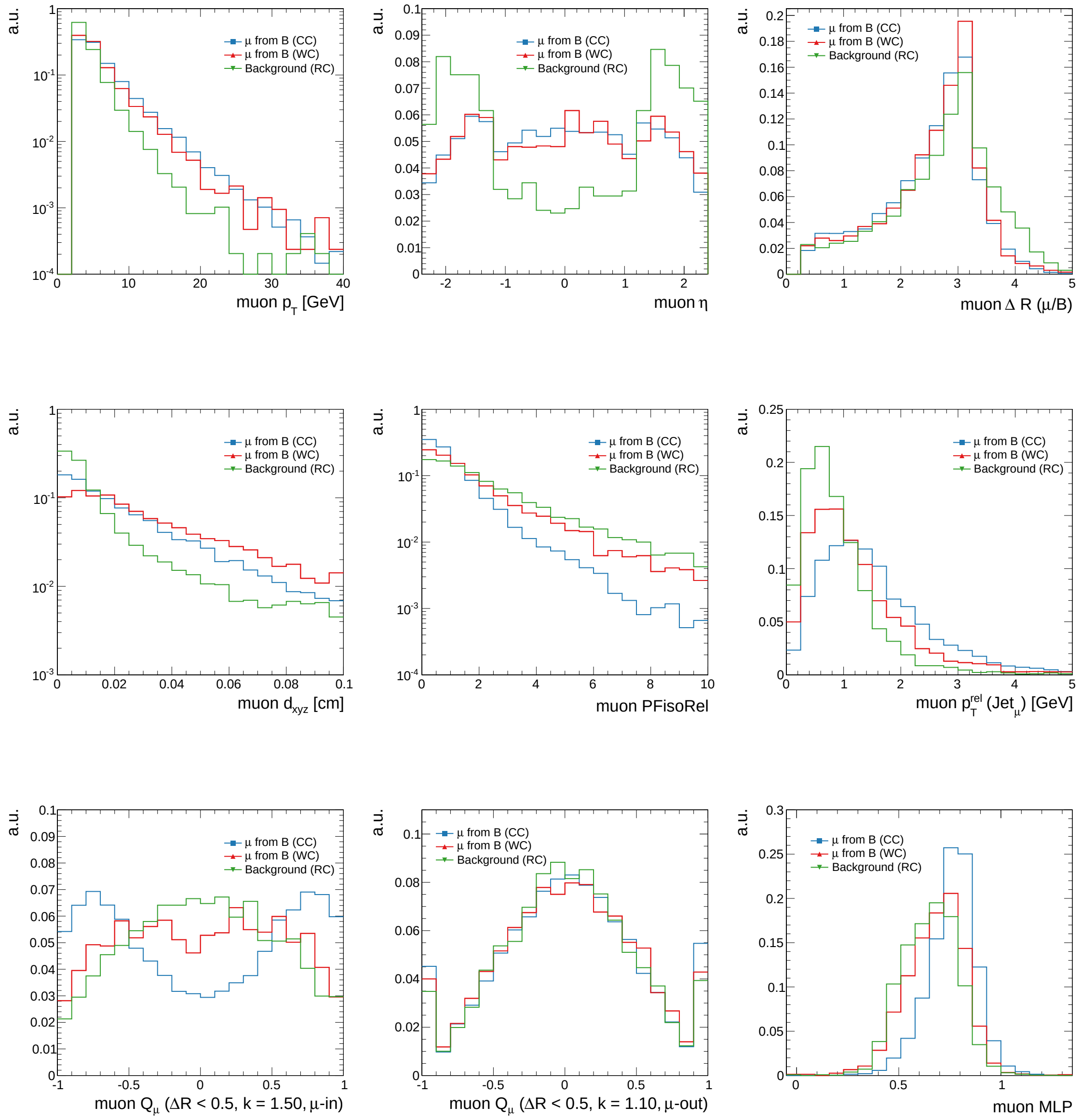


Figure A.5: OS-muon variables in the $B_s^0 \rightarrow J/\psi\,\phi$ simulation, divided in three categories: green, muons coming from B with correct charge-flavour relation (CC); blue, muons from B with opposite charge-flavour relation (WC); red, background muons with random charge-flavour relation (RC). Histograms are normalized to the same area. OS-muon selection is applied. From left to right, top to bottom: muon transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, p_T^{rel}, Q_μ ($\Delta R < 0.5$, $k = 1.50$, μ-in), Q_μ ($\Delta R < 0.5$, $k = 1.10$, μ-out), Neural Network output (MLP).

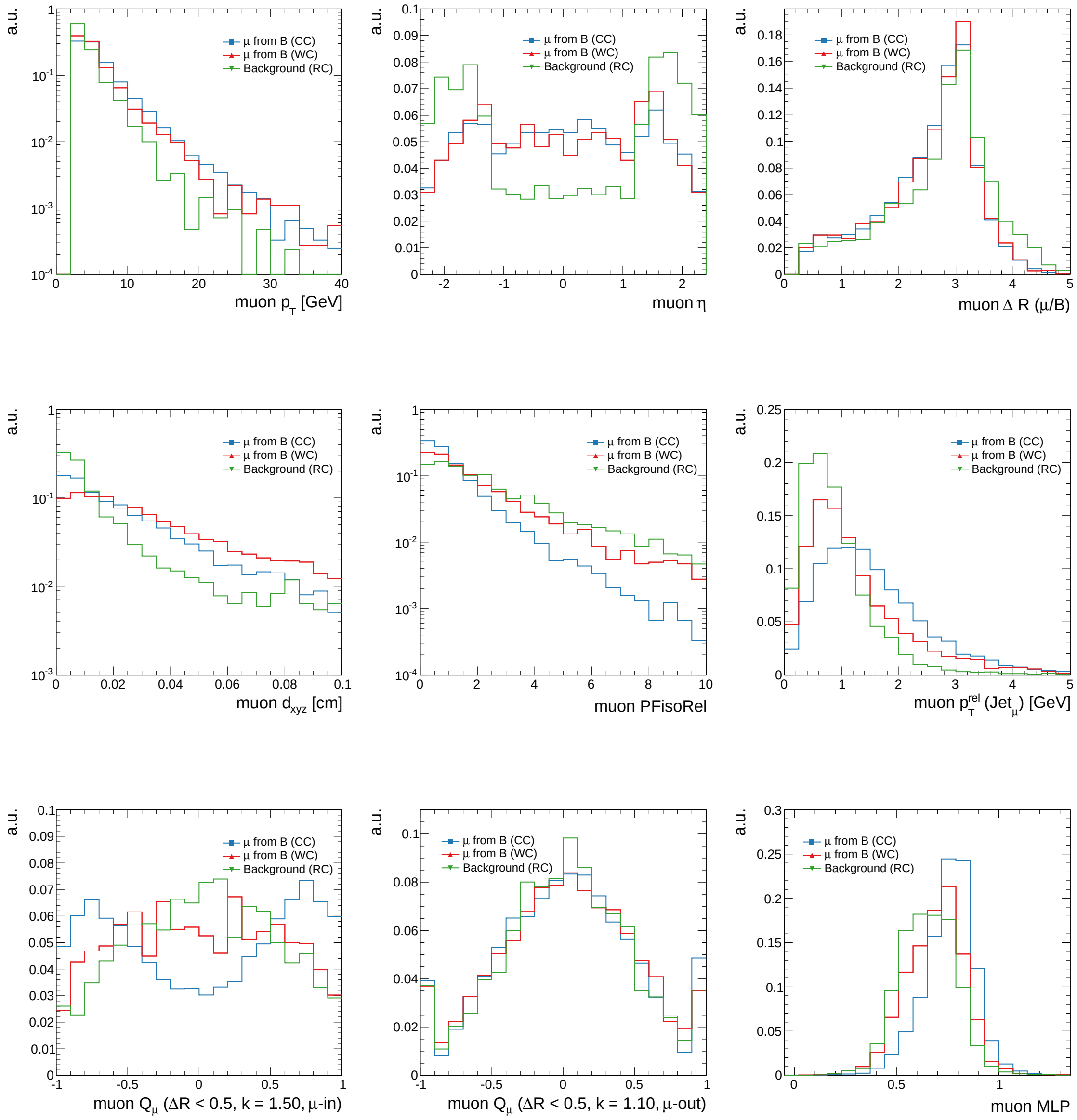


Figure A.6: OS-muon variables in the $B^+ \to J/\psi K^+$ simulation, divided in three categories: green, muons coming from B with correct charge-flavour relation (CC); blue, muons from B with opposite charge-flavour relation (WC); red, background muons with random charge-flavour relation (RC). Histograms are normalized to the same area. OS-muon selection is applied. From left to right, top to bottom: muon transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, p_T^{rel}, Q_μ ($\Delta R < 0.5$, $k = 1.50$, μ-in), Q_μ ($\Delta R < 0.5$, $k = 1.10$, μ-out), Neural Network output (MLP).

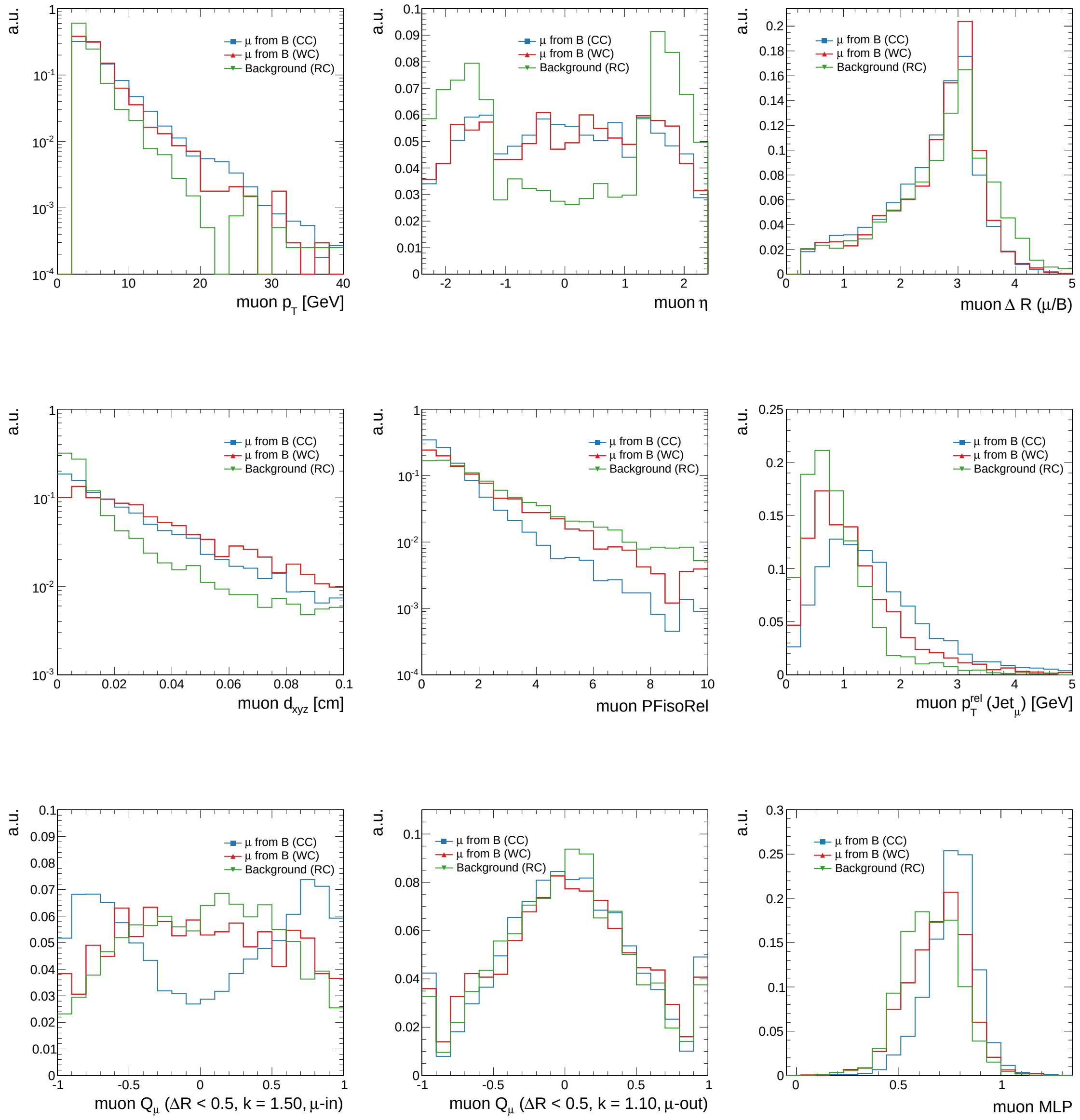


Figure A.7: OS-muon variables in the $B^0 \to J/\psi\, K^*$ simulation, divided in three categories: green, muons coming from B with correct charge-flavour relation (CC); blue, muons from B with opposite charge-flavour relation (WC); red, background muons with random charge-flavour relation (RC). Histograms are normalized to the same area. OS-muon selection is applied. From left to right, top to bottom: muon transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, $p_T^{\rm rel}$, Q_μ ($\Delta R < 0.5$, $k = 1.50$, μ-in), Q_μ ($\Delta R < 0.5$, $k = 1.10$, μ-out), Neural Network output (MLP).

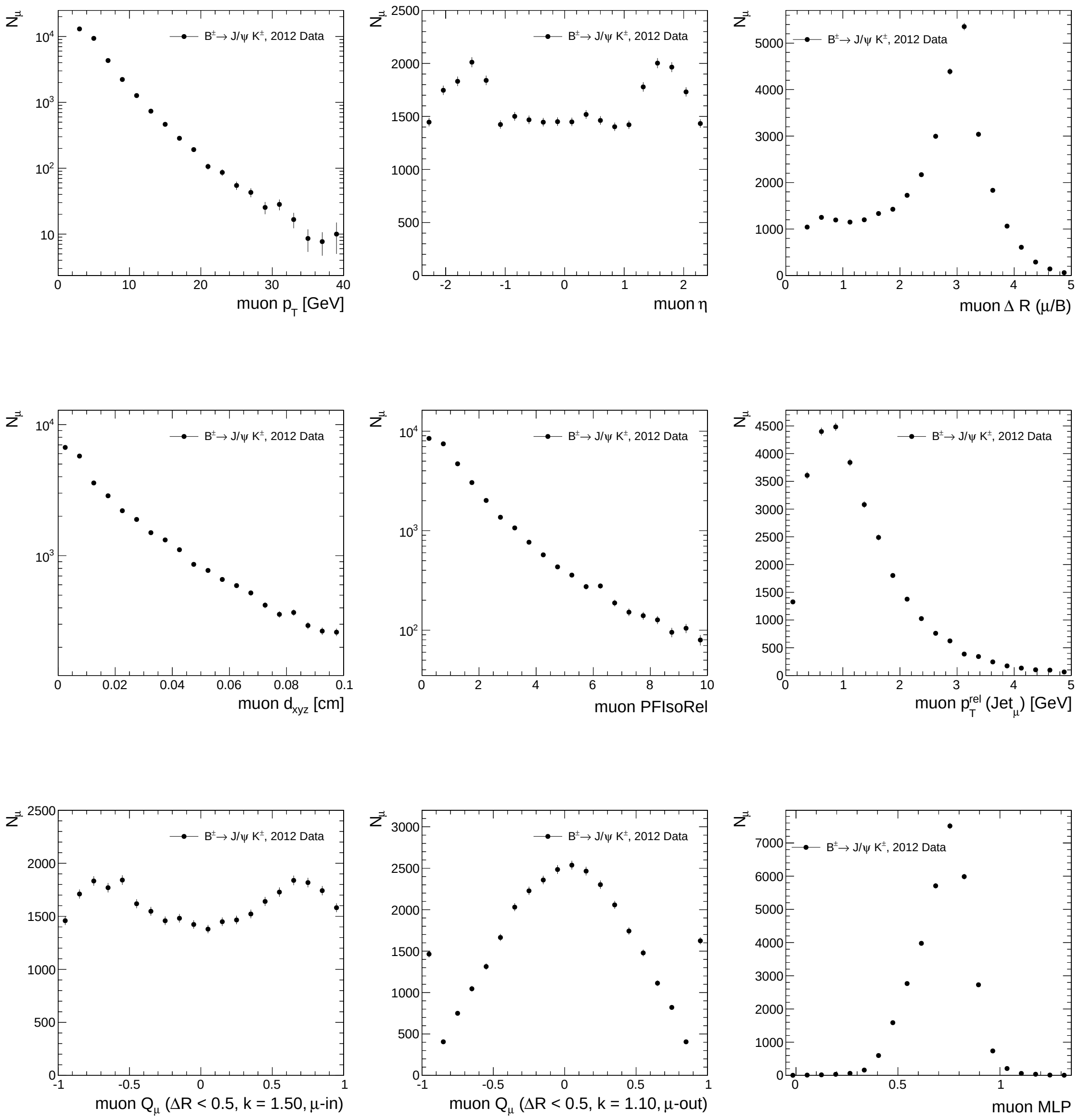


Figure A.8: OS-muon variables in the 2012 data sample, corresponding to an integrated luminosity of 19.7 fb^{-1}, after the muon selection. From left to right, top to bottom: muon transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, p_T^{rel}, Q_μ ($\Delta R < 0.5$, $k = 1.50$, μ-in), Q_μ ($\Delta R < 0.5$, $k = 1.10$, μ-out), Neural Network output (MLP).

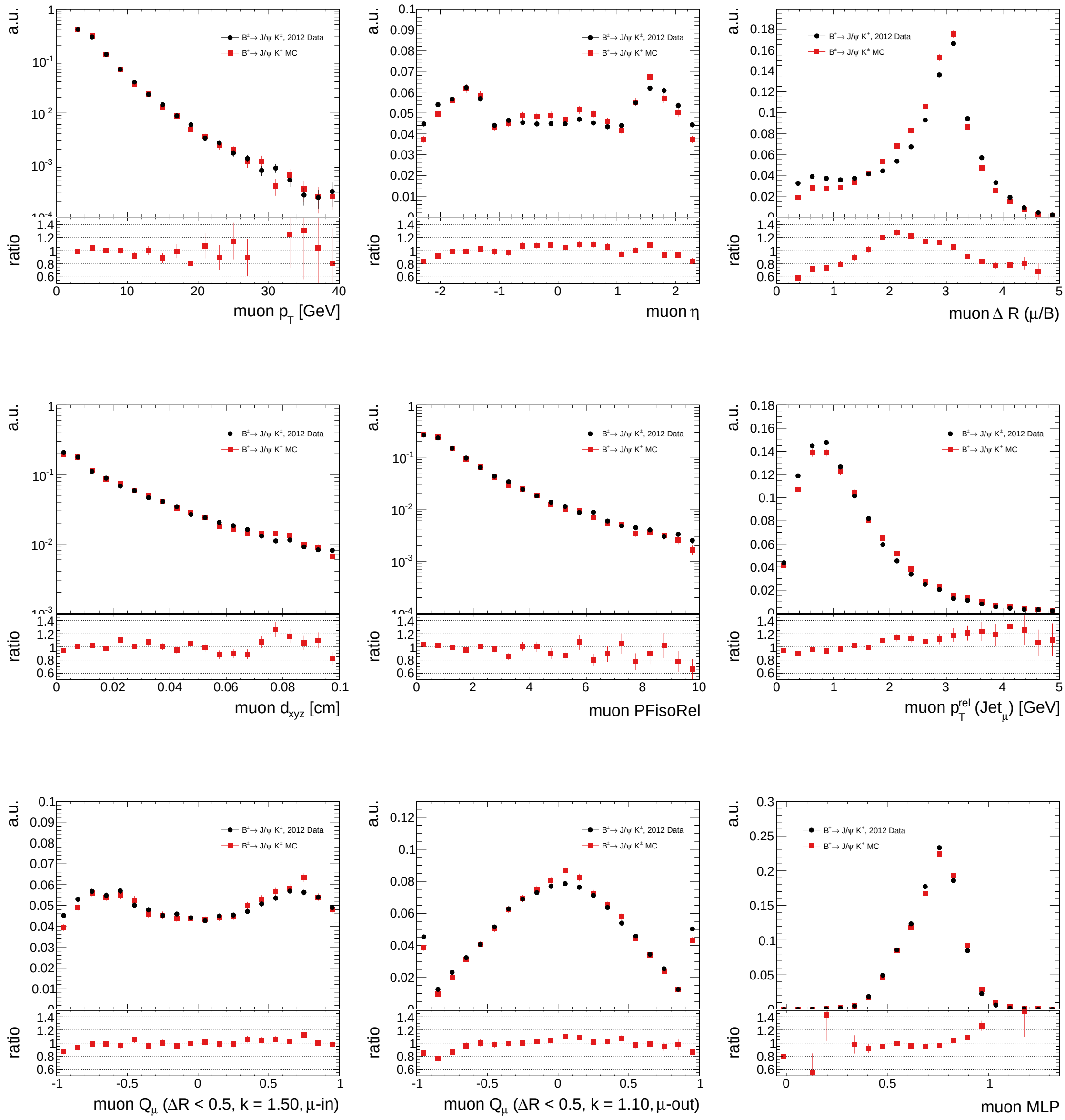


Figure A.9: Comparison of $B^+ \to J/\psi\, K^+$ 2012 data (black) and MC (red), after the muon selection. The MC and data distributions are normalized to the same area, and the bin-by-bin ratio of the MC over Data is also shown. From left to right, top to bottom: muon transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, p_T^{rel}, Q_μ ($\Delta R < 0.5$, $k = 1.50$, μ-in), Q_μ ($\Delta R < 0.5$, $k = 1.10$, μ-out), Neural Network output (MLP).

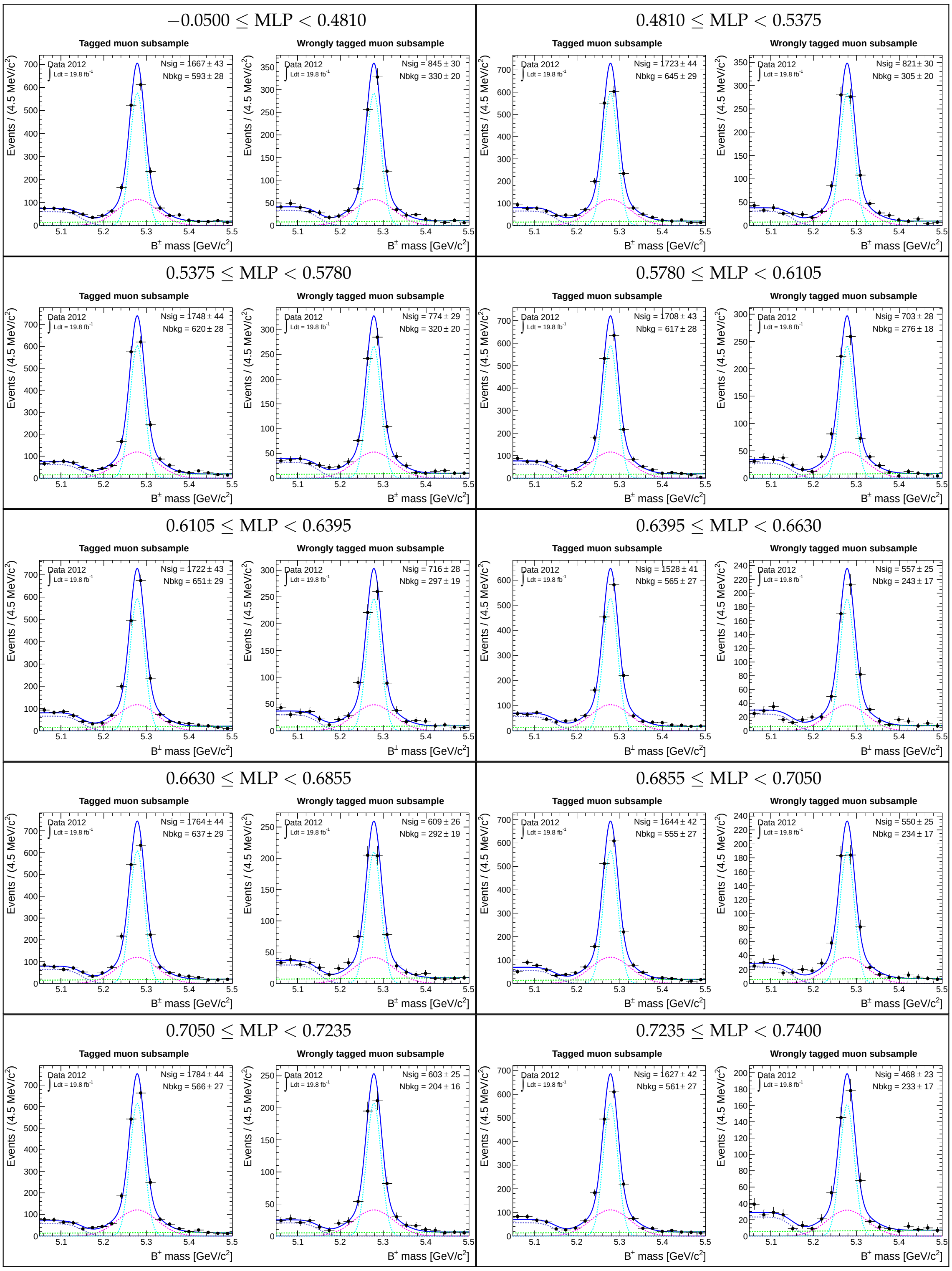


Figure A.10: Fit to B^+ invariant mass for the reconstructed $B^+ \rightarrow J/\psi\, K^+$ events containing a selected OS-muon in ranges of MLP.

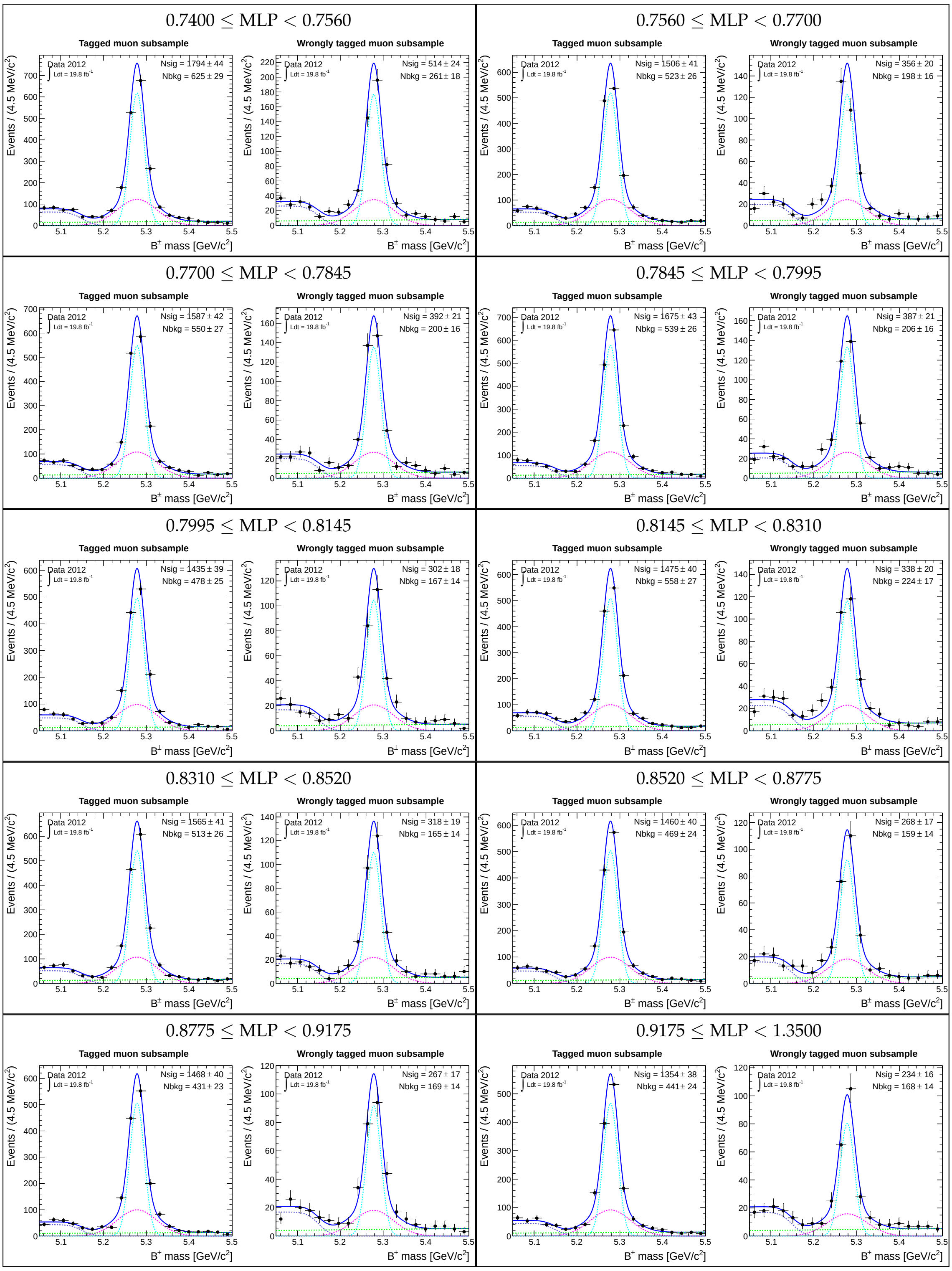


Figure A.11: Fit to B^+ invariant mass for the reconstructed $B^+ \to J/\psi\, K^+$ events containing a selected OS-muon in ranges of MLP.

Table A.1: Muon tagging performances evaluated on the reconstructed $B_s^0 \to J/\psi\,\phi$ simulated events. Twenty categories of MLP discriminator are used.

MLP	$[-0.05, 0.481)$	$[0.481, 0.5375)$	$[0.5375, 0.578)$	$[0.578, 0.6105)$	$[0.6105, 0.6395)$
ε_{tag}	0.185 ± 0.006	0.184 ± 0.006	0.185 ± 0.006	0.186 ± 0.006	0.186 ± 0.006
ω	55.368 ± 1.52	50.401 ± 1.535	46.844 ± 1.534	42.92 ± 1.522	40.688 ± 1.511
$\mathcal{P}_{tag}$	0.002 ± 0.001	0 ± 0	0.001 ± 0.001	0.004 ± 0.002	0.006 ± 0.002

MLP	$[0.6395, 0.663)$	$[0.663, 0.6855)$	$[0.6855, 0.705)$	$[0.705, 0.7235)$	$[0.7235, 0.74)$
ε_{tag}	0.185 ± 0.006	0.185 ± 0.006	0.188 ± 0.006	0.191 ± 0.006	0.189 ± 0.006
ω	34.043 ± 1.469	32.684 ± 1.455	33.538 ± 1.455	28.793 ± 1.391	25.521 ± 1.349
$\mathcal{P}_{tag}$	0.019 ± 0.004	0.022 ± 0.004	0.02 ± 0.004	0.034 ± 0.005	0.045 ± 0.005

MLP	$[0.74, 0.756)$	$[0.756, 0.77)$	$[0.77, 0.7845)$	$[0.7845, 0.7995)$	$[0.7995, 0.8145)$
ε_{tag}	0.188 ± 0.006	0.185 ± 0.006	0.19 ± 0.006	0.19 ± 0.006	0.184 ± 0.006
ω	25.044 ± 1.348	24.021 ± 1.342	23.31 ± 1.311	21.311 ± 1.271	20.392 ± 1.274
$\mathcal{P}_{tag}$	0.047 ± 0.005	0.05 ± 0.005	0.054 ± 0.006	0.063 ± 0.006	0.065 ± 0.006

MLP	$[0.8145, 0.831)$	$[0.831, 0.852)$	$[0.852, 0.8775)$	$[0.8775, 0.9175)$	$[0.9175, 1.35)$
ε_{tag}	0.188 ± 0.006	0.19 ± 0.006	0.185 ± 0.006	0.185 ± 0.006	0.188 ± 0.006
ω	17.991 ± 1.208	17.678 ± 1.195	15.009 ± 1.141	14.591 ± 1.13	12.489 ± 1.055
$\mathcal{P}_{tag}$	0.077 ± 0.006	0.079 ± 0.006	0.091 ± 0.007	0.093 ± 0.007	0.106 ± 0.007

Table A.2: Combined muon tagging performances evaluated on the reconstructed $B_s^0 \to J/\psi\,\phi$ simulated events using twenty categories of MLP discriminator.

	$B_s^0 \to J/\psi\,\phi$ MC
ε_{tag}^{tot}	$3.74 \pm 0.03\%$
$\mathcal{P}_{tag}^{tot}$	$0.88 \pm 0.02\%$

Table A.3: Muon tagging performances evaluated on the reconstructed $B^+ \to J/\psi K^+$ simulated events. Twenty categories of MLP discriminator are used.

MLP	$[-0.05, 0.481)$	$[0.481, 0.5375)$	$[0.5375, 0.578)$	$[0.578, 0.6105)$	$[0.6105, 0.6395)$
ε_{tag}	0.198 ± 0.006	0.184 ± 0.006	0.217 ± 0.007	0.187 ± 0.006	0.2 ± 0.006
ω	47.144 ± 1.604	43.139 ± 1.655	42.63 ± 1.519	41.309 ± 1.634	38.063 ± 1.562
$\mathcal{P}_{tag}$	0.001 ± 0.001	0.003 ± 0.002	0.005 ± 0.002	0.006 ± 0.002	0.011 ± 0.003

MLP	$[0.6395, 0.663)$	$[0.663, 0.6855)$	$[0.6855, 0.705)$	$[0.705, 0.7235)$	$[0.7235, 0.74)$
ε_{tag}	0.186 ± 0.006	0.191 ± 0.006	0.194 ± 0.006	0.186 ± 0.006	0.185 ± 0.006
ω	36.588 ± 1.609	32.663 ± 1.554	29.377 ± 1.502	27.01 ± 1.501	26.971 ± 1.505
$\mathcal{P}_{tag}$	0.013 ± 0.003	0.023 ± 0.004	0.033 ± 0.005	0.039 ± 0.005	0.039 ± 0.005

MLP	$[0.74, 0.756)$	$[0.756, 0.77)$	$[0.77, 0.7845)$	$[0.7845, 0.7995)$	$[0.7995, 0.8145)$
ε_{tag}	0.204 ± 0.006	0.185 ± 0.006	0.183 ± 0.006	0.197 ± 0.006	0.184 ± 0.006
ω	27.684 ± 1.441	24.403 ± 1.463	24.188 ± 1.465	20.68 ± 1.34	18.399 ± 1.335
$\mathcal{P}_{tag}$	0.041 ± 0.005	0.048 ± 0.006	0.049 ± 0.006	0.068 ± 0.007	0.074 ± 0.007

MLP	$[0.8145, 0.831)$	$[0.831, 0.852)$	$[0.852, 0.8775)$	$[0.8775, 0.9175)$	$[0.9175, 1.35)$
ε_{tag}	0.17 ± 0.006	0.196 ± 0.006	0.185 ± 0.006	0.193 ± 0.006	0.213 ± 0.007
ω	20.135 ± 1.436	15.625 ± 1.218	15.822 ± 1.262	17.594 ± 1.283	15.558 ± 1.164
$\mathcal{P}_{tag}$	0.061 ± 0.006	0.093 ± 0.007	0.087 ± 0.007	0.081 ± 0.007	0.101 ± 0.008

Table A.4: Combined muon tagging performances evaluated on the reconstructed $B^+ \to J/\psi K^+$ simulated events using twenty categories of MLP discriminator.

	$B^+ \to J/\psi K^+$ MC
ε_{tag}^{tot}	$3.84 \pm 0.03\%$
$\mathcal{P}_{tag}^{tot}$	$0.88 \pm 0.02\%$

Table A.5: Muon tagging performances evaluated on the reconstructed $B^0 \to J/\psi\, K^*$ simulated events. Twenty categories of MLP discriminator are used.

MLP	$[-0.05, 0.481)$	$[0.481, 0.5375)$	$[0.5375, 0.578)$	$[0.578, 0.6105)$	$[0.6105, 0.6395)$
ε_{tag}	0.189 ± 0.007	0.213 ± 0.007	0.194 ± 0.007	0.204 ± 0.007	0.196 ± 0.007
ω	52.477 ± 1.728	47.848 ± 1.632	44.334 ± 1.708	42.38 ± 1.656	35 ± 1.643
$\mathcal{P}_{tag}$	0 ± 0.001	0 ± 0.001	0.002 ± 0.002	0.005 ± 0.002	0.018 ± 0.004

MLP	$[0.6395, 0.663)$	$[0.663, 0.6855)$	$[0.6855, 0.705)$	$[0.705, 0.7235)$	$[0.7235, 0.74)$
ε_{tag}	0.183 ± 0.006	0.185 ± 0.006	0.191 ± 0.007	0.191 ± 0.007	0.189 ± 0.007
ω	33.644 ± 1.689	31.07 ± 1.649	29.978 ± 1.611	30.546 ± 1.615	30.146 ± 1.618
$\mathcal{P}_{tag}$	0.02 ± 0.004	0.027 ± 0.005	0.031 ± 0.005	0.029 ± 0.005	0.03 ± 0.005

MLP	$[0.74, 0.756)$	$[0.756, 0.77)$	$[0.77, 0.7845)$	$[0.7845, 0.7995)$	$[0.7995, 0.8145)$
ε_{tag}	0.221 ± 0.007	0.186 ± 0.007	0.191 ± 0.007	0.192 ± 0.007	0.195 ± 0.007
ω	24.251 ± 1.406	22.769 ± 1.507	21.763 ± 1.467	19.491 ± 1.408	21.858 ± 1.452
$\mathcal{P}_{tag}$	0.059 ± 0.007	0.055 ± 0.006	0.061 ± 0.007	0.072 ± 0.007	0.062 ± 0.007

MLP	$[0.8145, 0.831)$	$[0.831, 0.852)$	$[0.852, 0.8775)$	$[0.8775, 0.9175)$	$[0.9175, 1.35)$
ε_{tag}	0.194 ± 0.007	0.202 ± 0.007	0.19 ± 0.007	0.196 ± 0.007	0.217 ± 0.007
ω	19.56 ± 1.404	17.684 ± 1.326	16.368 ± 1.334	18.35 ± 1.365	17.502 ± 1.274
$\mathcal{P}_{tag}$	0.072 ± 0.007	0.085 ± 0.007	0.086 ± 0.007	0.079 ± 0.007	0.092 ± 0.008

Table A.6: Combined muon tagging performances evaluated on the reconstructed $B^0 \to J/\psi\, K^*$ simulated events using twenty categories of MLP discriminator.

	$B^0 \to J/\psi\, K^*$ MC
ε_{tag}^{tot}	$3.92 \pm 0.03\%$
$\mathcal{P}_{tag}^{tot}$	$0.89 \pm 0.03\%$

Table A.7: Muon tagging performances evaluated on the reconstructed $B^+ \to J/\psi K^+$ events in 2012 data. Twenty categories of MLP discriminator are used.

MLP	$[-0.05, 0.481)$	$[0.481, 0.5375)$	$[0.5375, 0.578)$	$[0.578, 0.6105)$	$[0.6105, 0.6395)$
ε_{tag}	0.236 ± 0.006	0.244 ± 0.006	0.247 ± 0.006	0.242 ± 0.006	0.244 ± 0.006
ω	50.691 ± 2.258	47.66 ± 2.134	44.307 ± 2.023	41.162 ± 1.947	41.611 ± 1.958
$\mathcal{P}_{tag}$	0 ± 0	0.001 ± 0.001	0.003 ± 0.002	0.008 ± 0.003	0.007 ± 0.003

MLP	$[0.6395, 0.663)$	$[0.663, 0.6855)$	$[0.6855, 0.705)$	$[0.705, 0.7235)$	$[0.7235, 0.74)$
ε_{tag}	0.216 ± 0.006	0.25 ± 0.006	0.233 ± 0.006	0.252 ± 0.006	0.23 ± 0.006
ω	36.46 ± 1.914	34.556 ± 1.726	33.475 ± 1.748	33.803 ± 1.676	28.765 ± 1.598
$\mathcal{P}_{tag}$	0.016 ± 0.005	0.024 ± 0.006	0.025 ± 0.006	0.026 ± 0.006	0.042 ± 0.007

MLP	$[0.74, 0.756)$	$[0.756, 0.77)$	$[0.77, 0.7845)$	$[0.7845, 0.7995)$	$[0.7995, 0.8145)$
ε_{tag}	0.254 ± 0.006	0.213 ± 0.006	0.225 ± 0.006	0.237 ± 0.006	0.203 ± 0.006
ω	28.669 ± 1.533	23.67 ± 1.497	24.699 ± 1.482	23.153 ± 1.398	21.103 ± 1.432
$\mathcal{P}_{tag}$	0.046 ± 0.007	0.059 ± 0.008	0.058 ± 0.008	0.068 ± 0.008	0.068 ± 0.008

MLP	$[0.8145, 0.831)$	$[0.831, 0.852)$	$[0.852, 0.8775)$	$[0.8775, 0.9175)$	$[0.9175, 1.35)$
ε_{tag}	0.209 ± 0.006	0.222 ± 0.006	0.207 ± 0.006	0.208 ± 0.006	0.192 ± 0.005
ω	22.933 ± 1.5	20.368 ± 1.333	18.365 ± 1.308	18.194 ± 1.307	17.298 ± 1.32
$\mathcal{P}_{tag}$	0.061 ± 0.008	0.078 ± 0.008	0.083 ± 0.008	0.084 ± 0.008	0.082 ± 0.008

Table A.8: Combined muon tagging performances evaluated on the reconstructed $B^+ \to J/\psi K^+$ events in 2012 data using twenty categories of MLP discriminator.

	$B^+ \to J/\psi K^+$ data
ε_{tag}^{tot}	$4.56 \pm 0.03\%$
$\mathcal{P}_{tag}^{tot}$	$0.84 \pm 0.03\%$

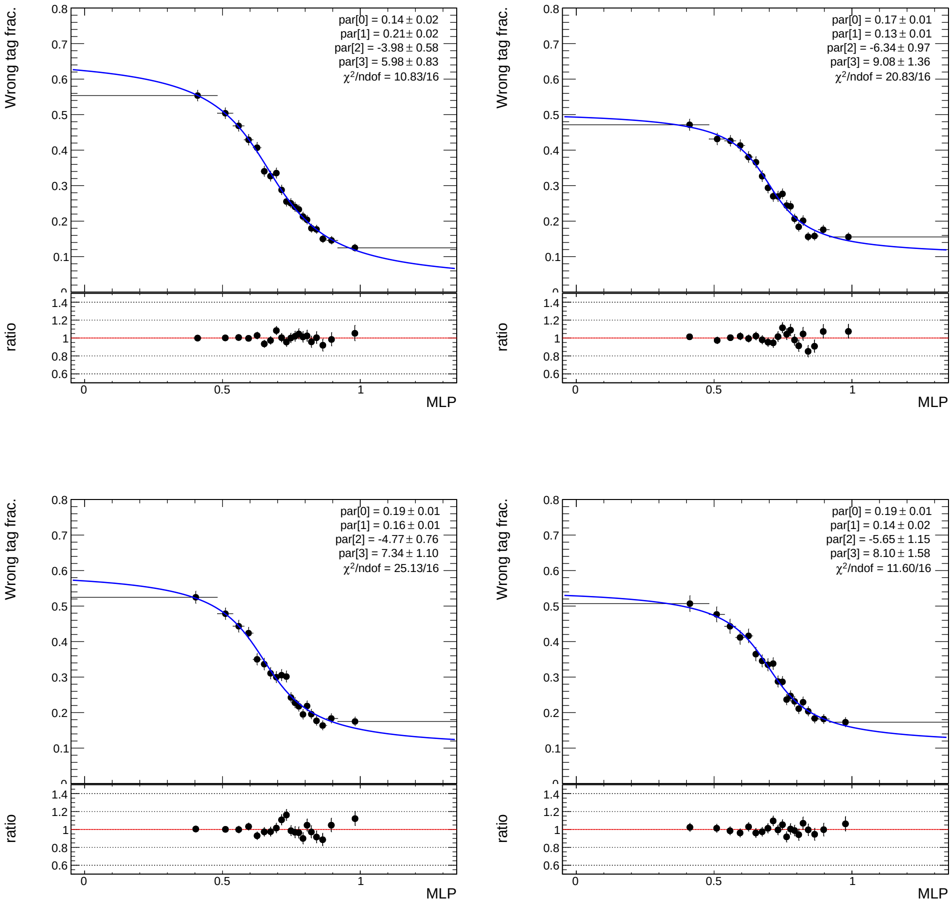


Figure A.12: OS-muon wrong tag fraction evaluated as a function of the MLP discriminator for the $B_s^0 \to J/\psi\,\phi$ MC (top-left), $B^+ \to J/\psi\,K^+$ MC (top-right), $B^0 \to J/\psi\,K^*$ MC (bottom-left), and 2012 $B^+ \to J/\psi\,K^+$ data (bottom-right). A function of the form $p_0 + p_1 \cdot [1 - \arctan(p_2 + p_3 \cdot \mathrm{MLP})]$ is used for the fit.

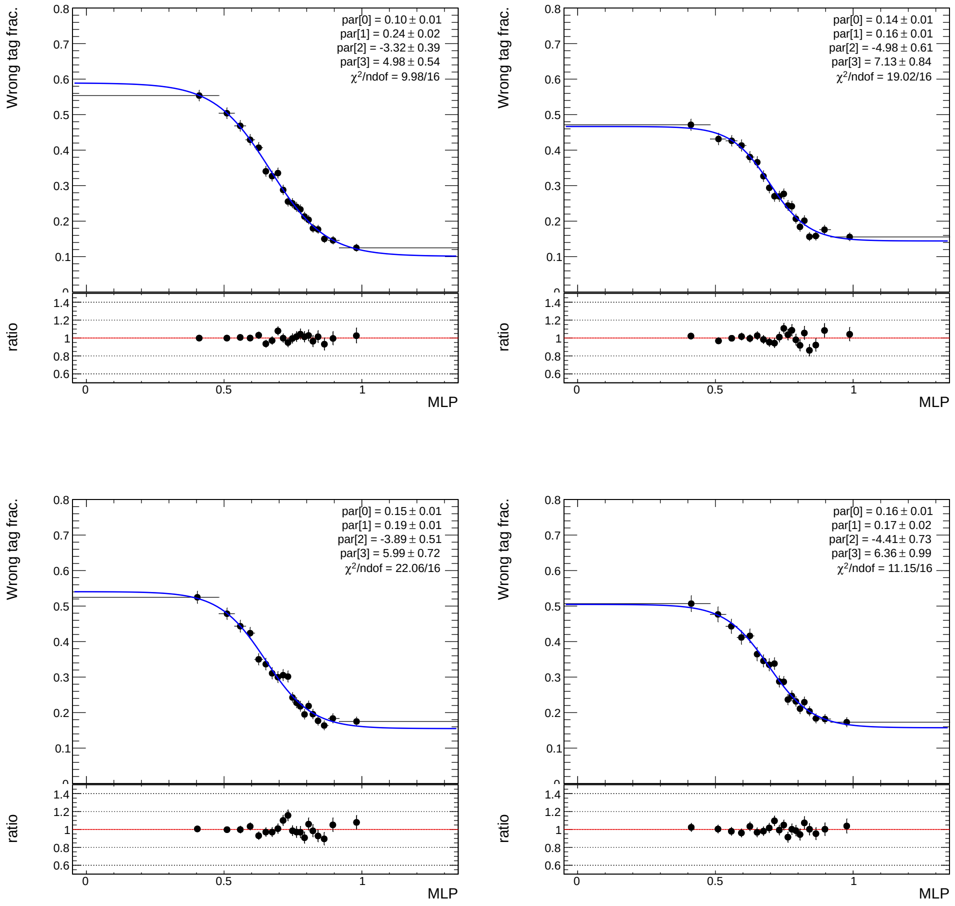


Figure A.13: OS-muon wrong tag fraction evaluated as a function of the MLP discriminator for the $B_s^0 \to J/\psi\,\phi$ MC (top-left), $B^+ \to J/\psi\,K^+$ MC (top-right), $B^0 \to J/\psi\,K^*$ MC (bottom-left), and 2012 $B^+ \to J/\psi\,K^+$ data (bottom-right). A function of the form $p_0 + p_1 \cdot [1 - \tanh(p_2 + p_3 \cdot \mathrm{MLP})]$ is used for the fit.

A.2 OS-Electron tagger

Figures A.14 and A.16 show the distribution of the OS-electron variables for the simulated samples $B_s^0 \to J/\psi\,\phi$ and $B^+ \to J/\psi\,K^+$ after the OS-electron selection. The comparison between the two MC samples is shown in Figure A.17.

In Figures A.18 and A.20 the electrons variables are divided into the three categories defined in MC: electrons from B with correct charge-flavour relation (CC), electrons from B with opposite charge-flavour relation (WC), background electrons not originated from a B decay (RC).

The data distributions are shown in Figure A.21. For each variable's bin a fit to the B^+ invariant mass is performed in order to obtain the signal yield of $B^+ \to J/\psi\,K^+$ events. The comparison among the OS-electron distributions obtained in 2012 data and in the $B^+ \to J/\psi\,K^+$ Monte Carlo samples is shown in Figure A.22.

Figures A.23 and A.24 show the fit to B^+ invariant mass for the reconstructed $B^+ \to J/\psi\,K^+$ events, performed for each of the 20 bins of the neural network discriminator (MLP) for OS-electrons.

The bin-by bin tagging performances evaluated using twenty neural network discriminator categories are presented in Table A.9, and the combined results are summarized in Table A.10, for the $B_s^0 \to J/\psi\,\phi$ MC. Analogous tables are presented for the $B^+ \to J/\psi\,K^+$ MC in Tables A.11,A.12; for the $B^0 \to J/\psi\,K^*$ MC in Tables A.13,A.14; and for the $B^+ \to J/\psi\,K^+$ 2012 data in Tables A.15,A.16.

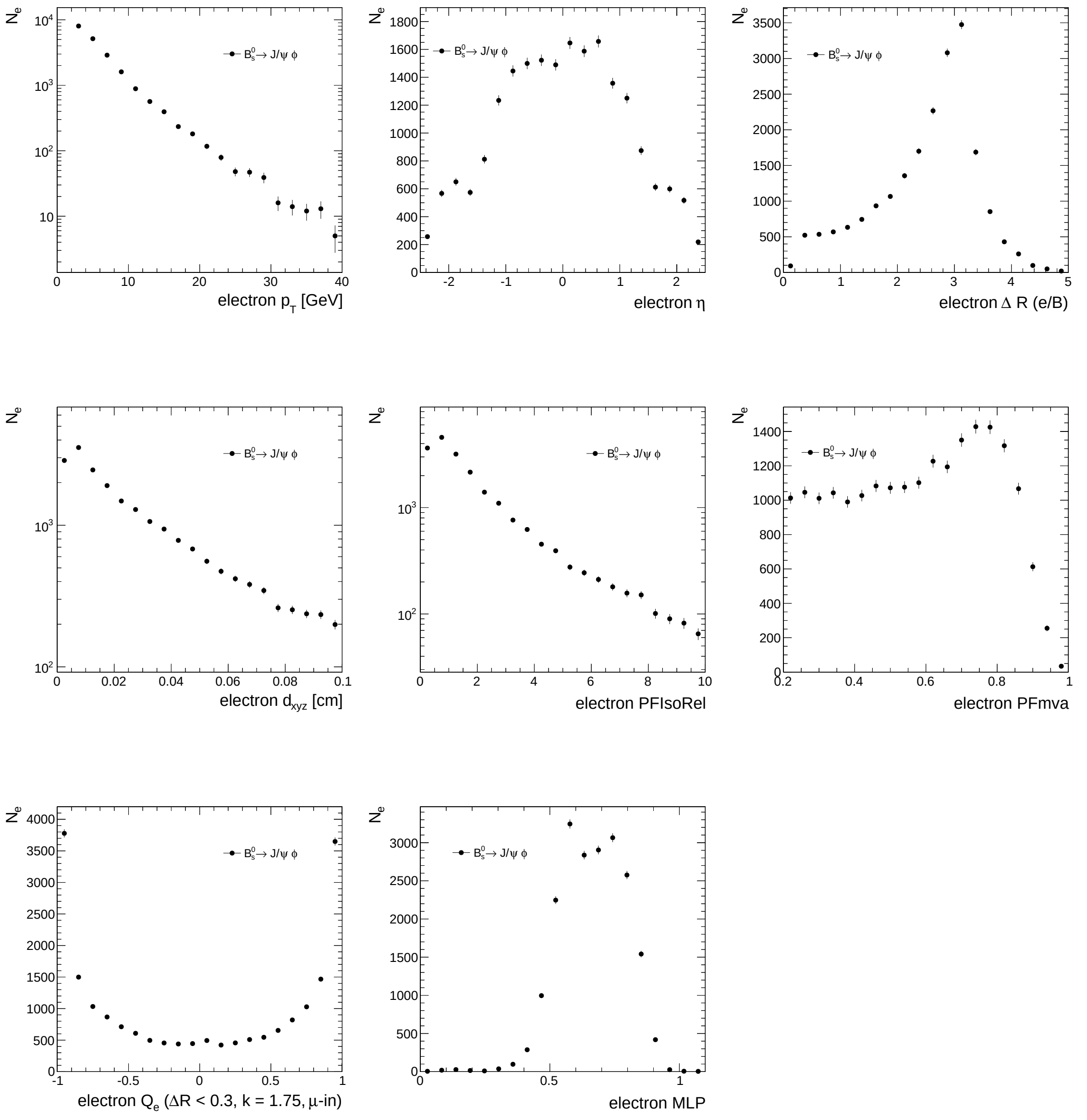


Figure A.14: OS-electron variables in the $B_s^0 \rightarrow J/\psi\,\phi$ simulation, after the electron selection. From left to right, top to bottom: electron transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, PFmva, Q_e ($\Delta R < 0.3$, $k = 1.75$, e-in), Neural Network output (MLP).

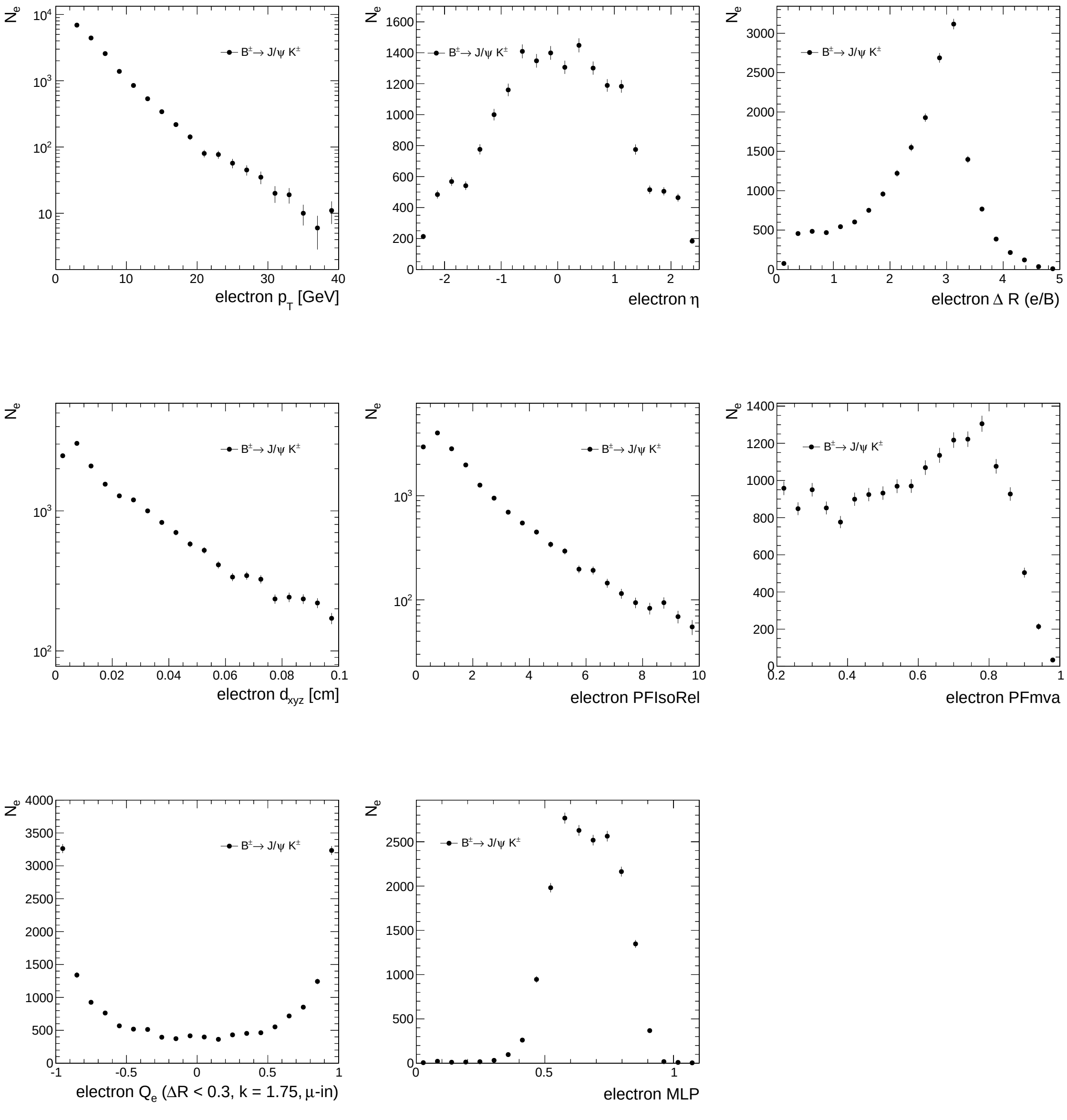


Figure A.15: OS-electron variables in the $B^+ \to J/\psi K^+$ simulation, after the electron selection. From left to right, top to bottom: electron transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, PFmva, Q_e ($\Delta R < 0.3$, $k = 1.75$, e-in), Neural Network output (MLP).

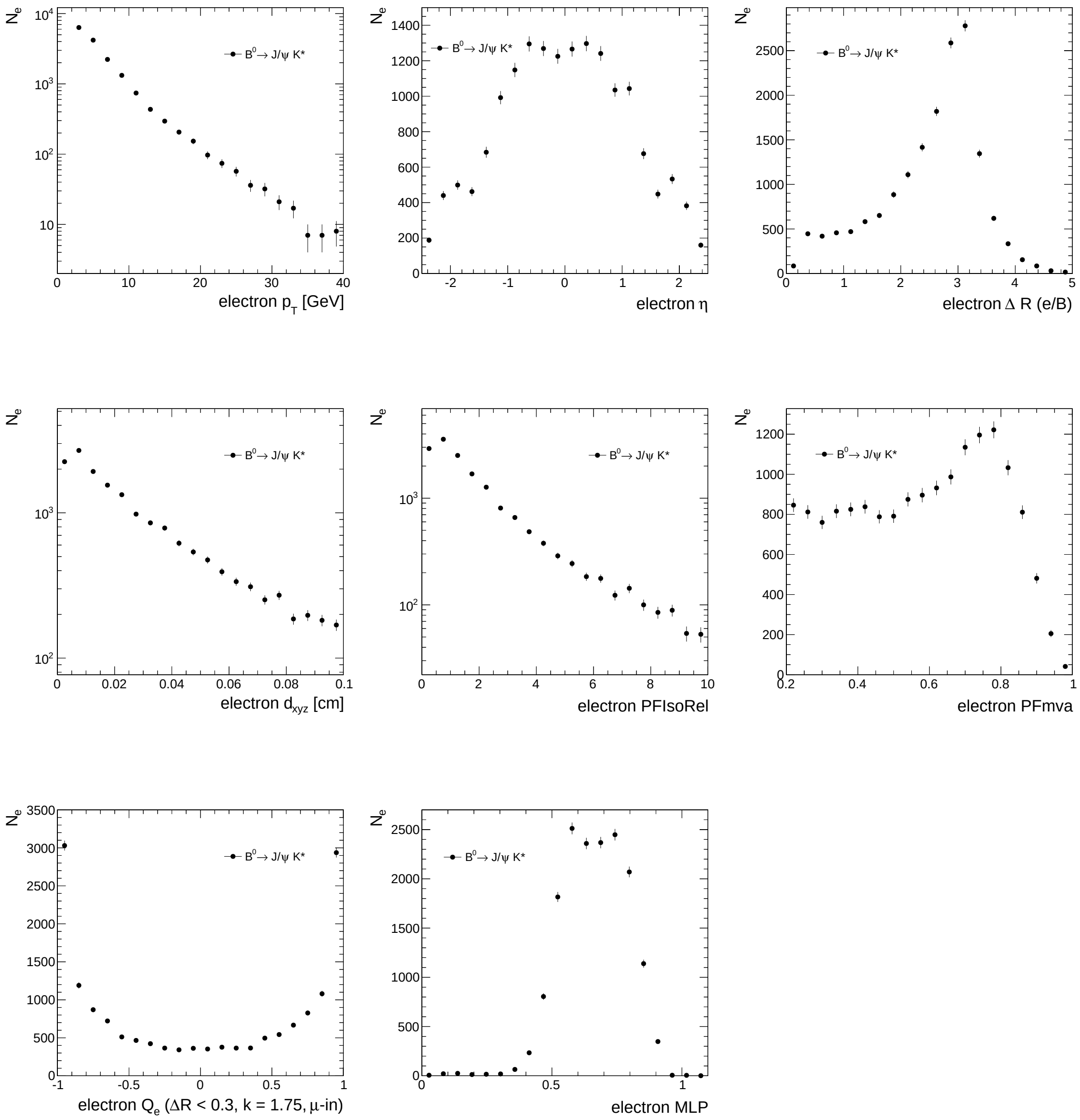


Figure A.16: OS-electron variables in the $B^0 \to J/\psi\, K^*$ simulation, after the electron selection. From left to right, top to bottom: electron transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, PFmva, Q_e ($\Delta R < 0.3$, $k = 1.75$, e-in), Neural Network output (MLP).

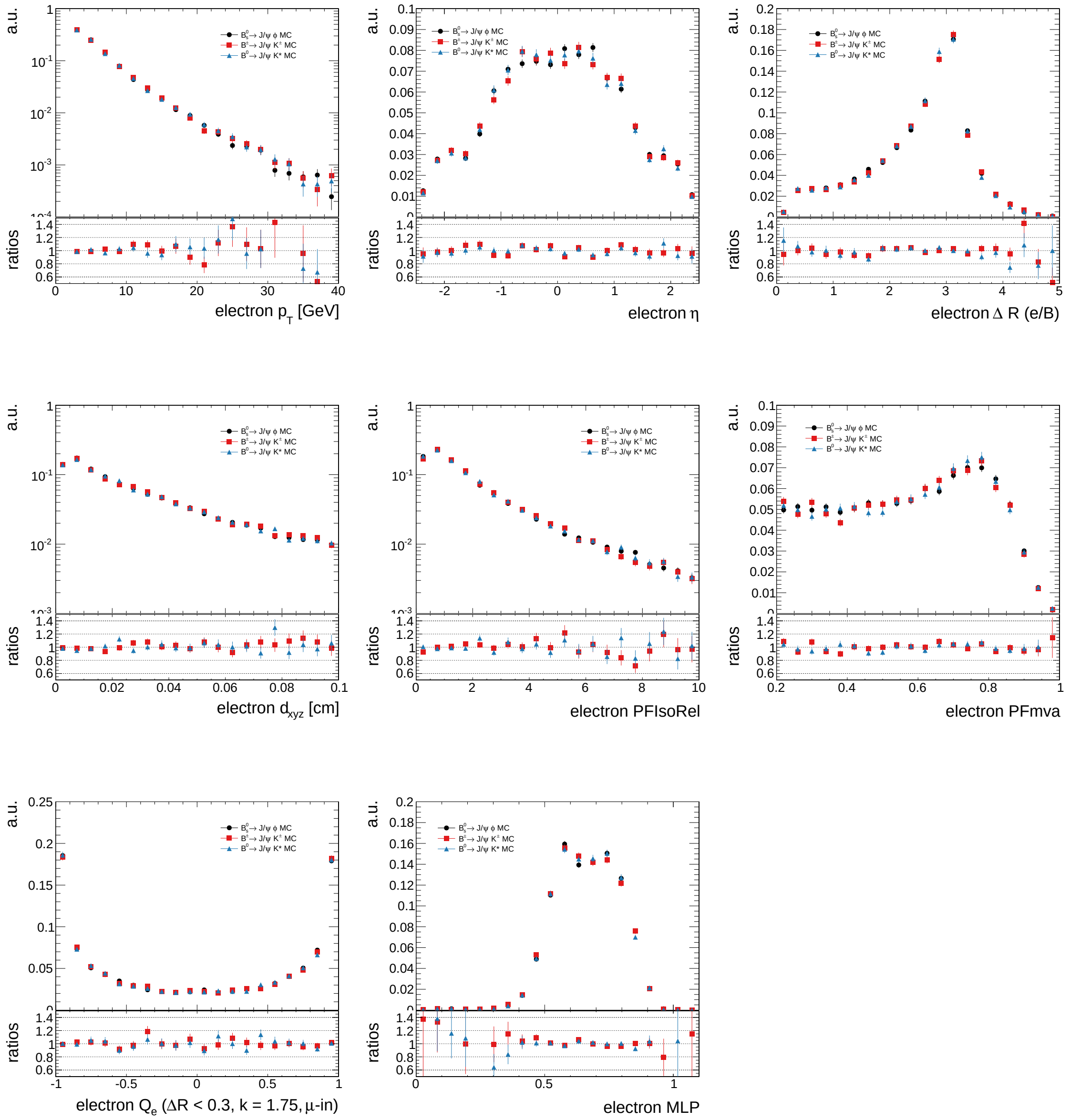


Figure A.17: Comparison of $B_s^0 \to J/\psi\,\phi$ (black), $B^+ \to J/\psi\,K^+$ (red) and $B^0 \to J/\psi\,K^*$ (blue) simulations after the electron selection is applied. The distributions are normalized to the same integral, and the bin-by-bin ratio of the $B^+ \to J/\psi\,K^+$ ($B^0 \to J/\psi\,K^*$) over the $B_s^0 \to J/\psi\,\phi$ is also shown. From left to right, top to bottom: electron transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, PFmva, Q_e ($\Delta R < 0.3$, $k = 1.75$, e-in), Neural Network output (MLP).

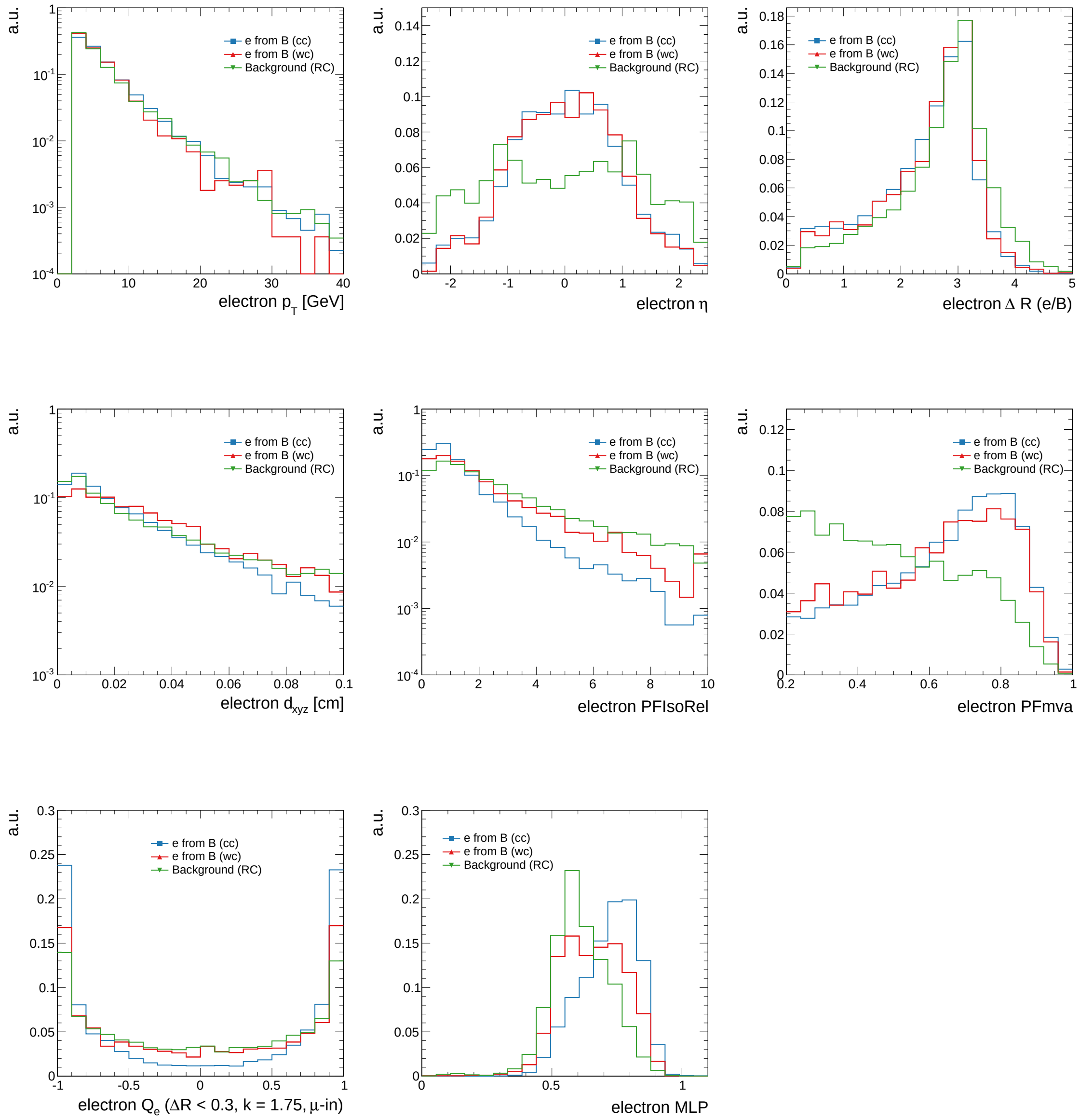


Figure A.18: OS-electron variables in the $B_s^0 \to J/\psi\,\phi$ simulation, divided in three categories: green, electrons coming from B with correct charge-flavour relation (CC); blue, electrons from B with opposite charge-flavour relation (WC); red, background electrons with random charge-flavour relation (RC). Histograms are normalized to the same area. OS-electron selection is applied. From left to right, top to bottom: electron transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, PFmva, Q_e ($\Delta R < 0.3$, $k = 1.75$, e-in), Neural Network output (MLP).

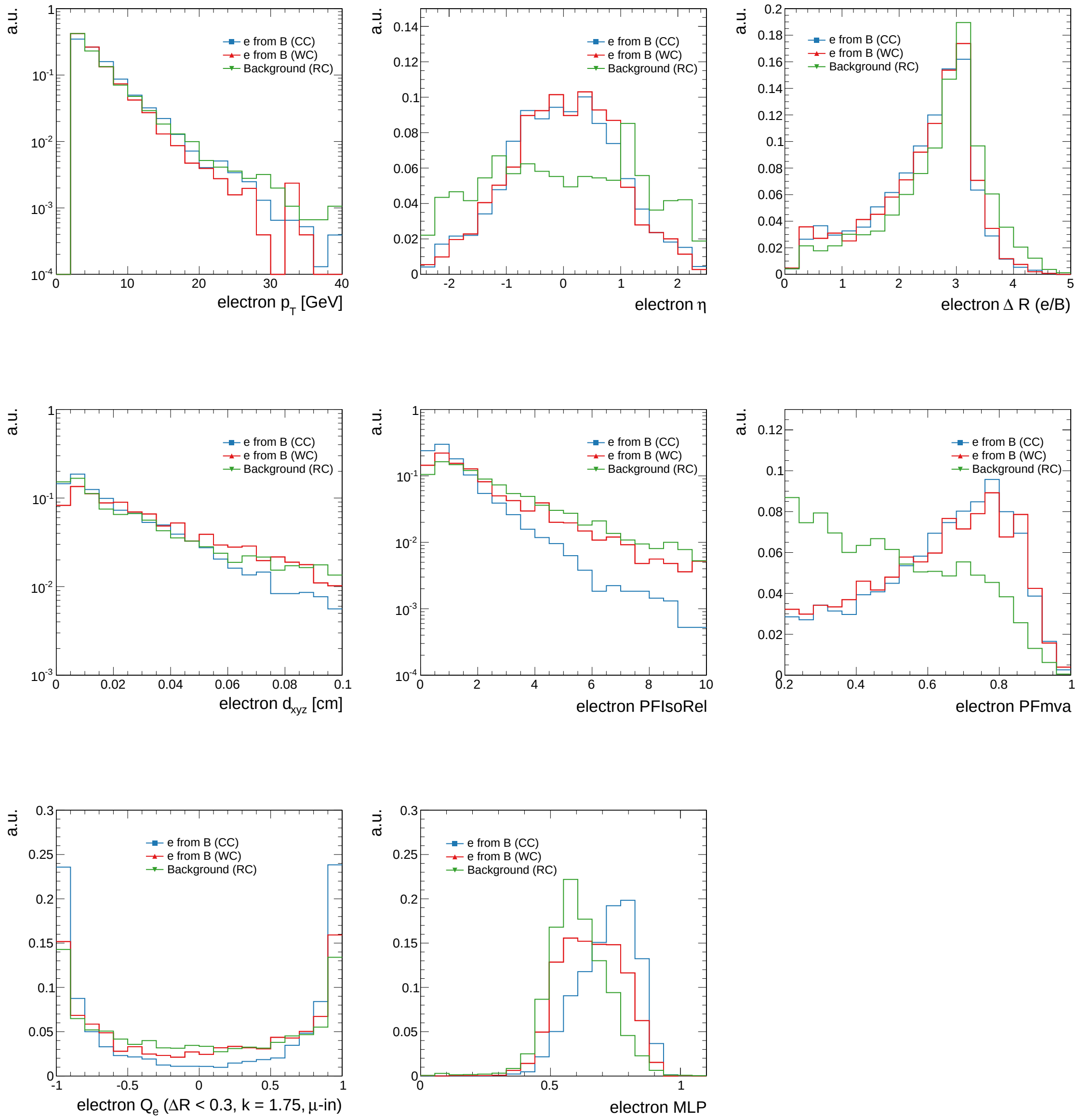


Figure A.19: OS-electron variables in the $B^+ \to J/\psi K^+$ simulation, divided in three categories: green, electrons coming from B with correct charge-flavour relation (CC); blue, electrons from B with opposite charge-flavour relation (WC); red, background electrons with random charge-flavour relation (RC). Histograms are normalized to the same area. OS-electron selection is applied. From left to right, top to bottom: electron transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, PFmva, Q_e ($\Delta R < 0.3$, $k = 1.75$, e-in), Neural Network output (MLP).

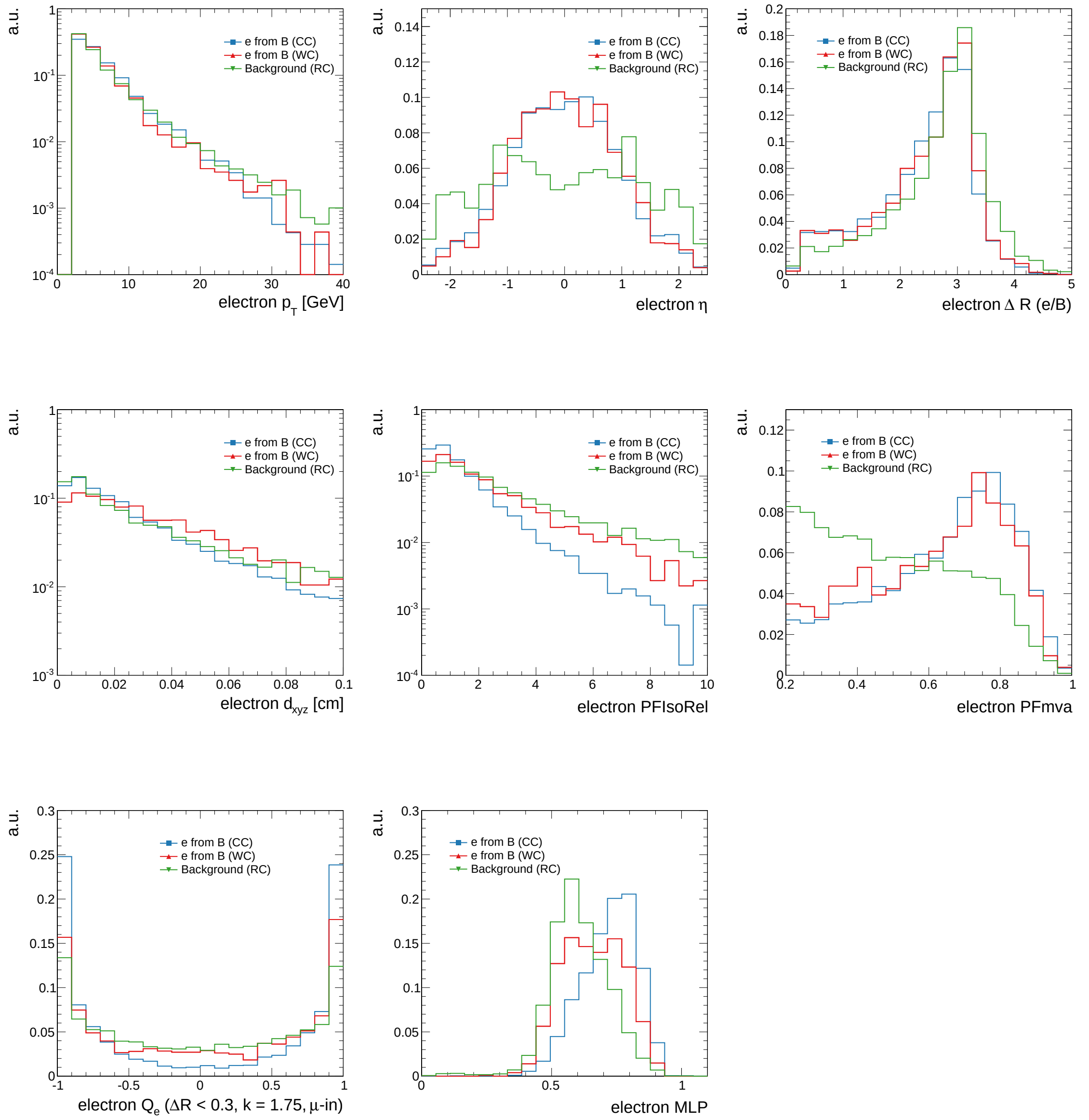


Figure A.20: OS-electron variables in the $B^0 \to J/\psi K^*$ simulation, divided in three categories: green, electrons coming from B with correct charge-flavour relation (CC); blue, electrons from B with opposite charge-flavour relation (WC); red, background electrons with random charge-flavour relation (RC). Histograms are normalized to the same area. OS-electron selection is applied. From left to right, top to bottom: electron transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, PFmva, Q_e ($\Delta R < 0.3$, $k = 1.75$, e-in), Neural Network output (MLP).

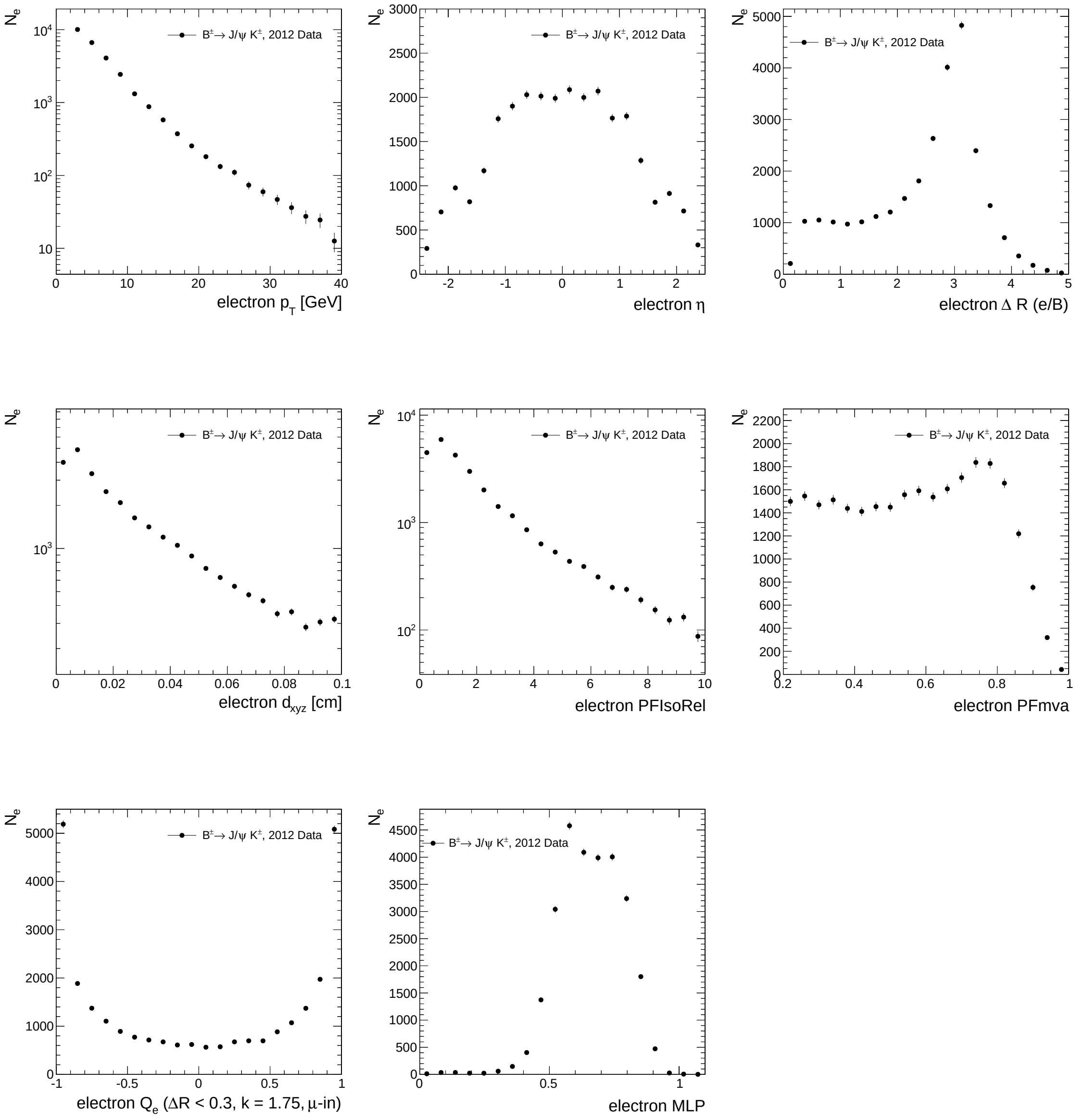


Figure A.21: OS-electron variables in the 2012 data sample, corresponding to an integrated luminosity of $19.7\ \text{fb}^{-1}$, after the electron selection. From left to right, top to bottom: electron transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, PFmva, Q_e ($\Delta R < 0.3$, $k = 1.75$, e-in), Neural Network output (MLP).

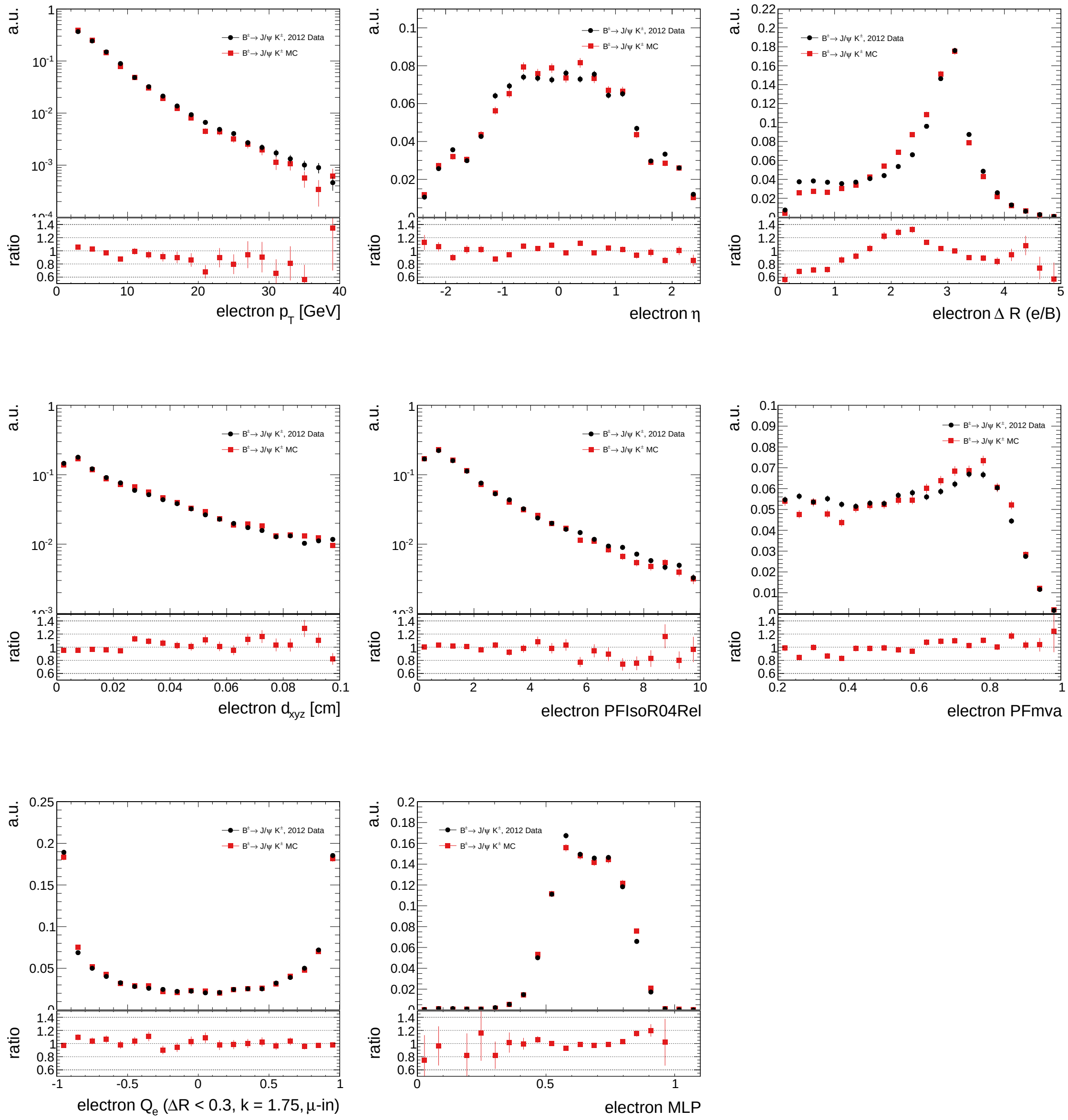


Figure A.22: Comparison of $B^+ \to J/\psi\, K^+$ 2012 data (black) and MC (red), after the electron selection. The MC and data distributions are normalized to the same area, and the bin-by-bin ratio of the MC over Data is also shown. From left to right, top to bottom: electron transverse momentum, pseudorapidity, ΔR with respect to the reco-B, 3D impact parameter, PFIsolation, PFmva, Q_e ($\Delta R < 0.3$, $k = 1.75$, e-in), Neural Network output (MLP).

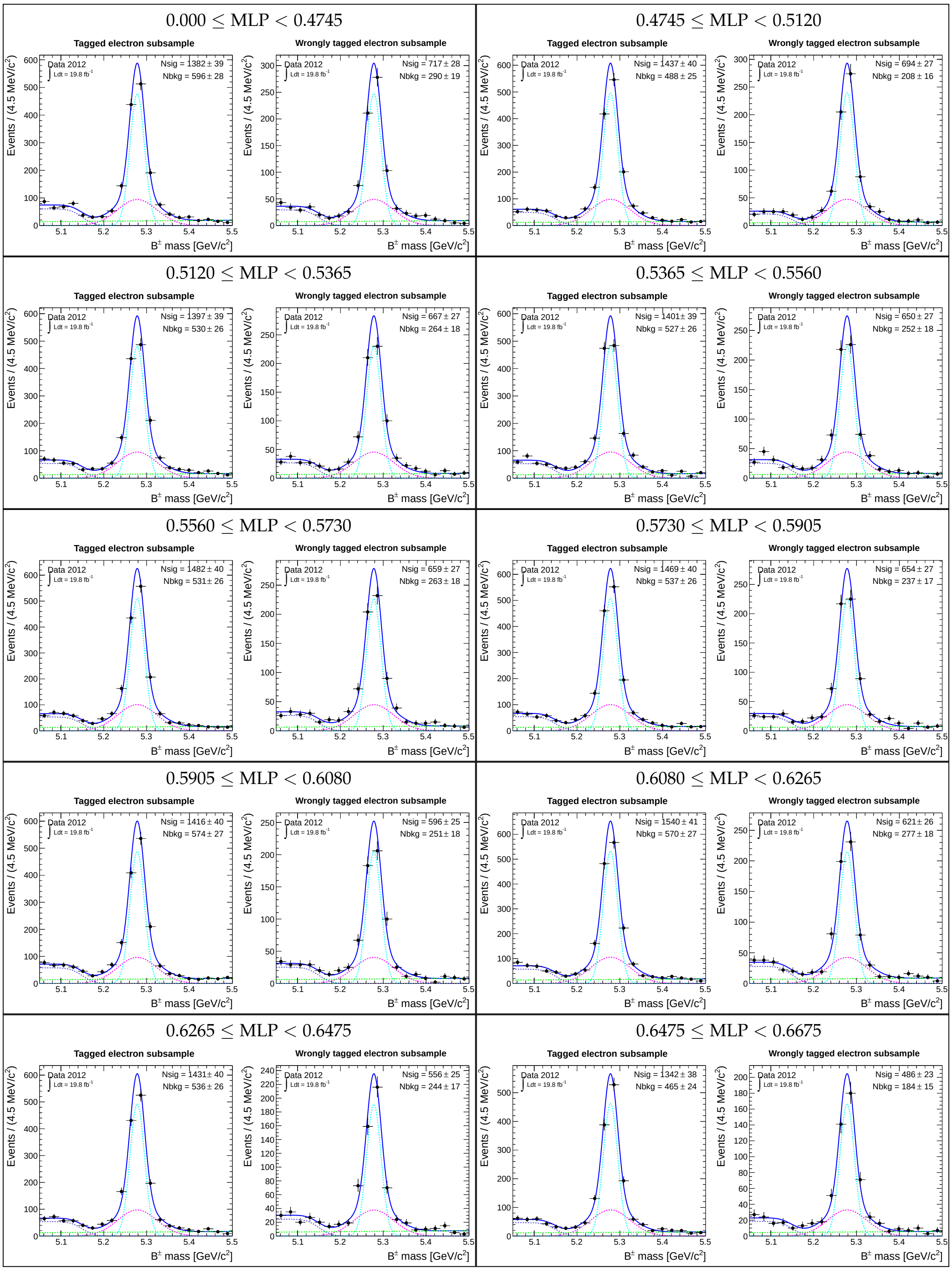


Figure A.23: Fit to B^+ invariant mass for the reconstructed $B^+ \to J/\psi\, K^+$ events containing a selected OS-electron in ranges of MLP.

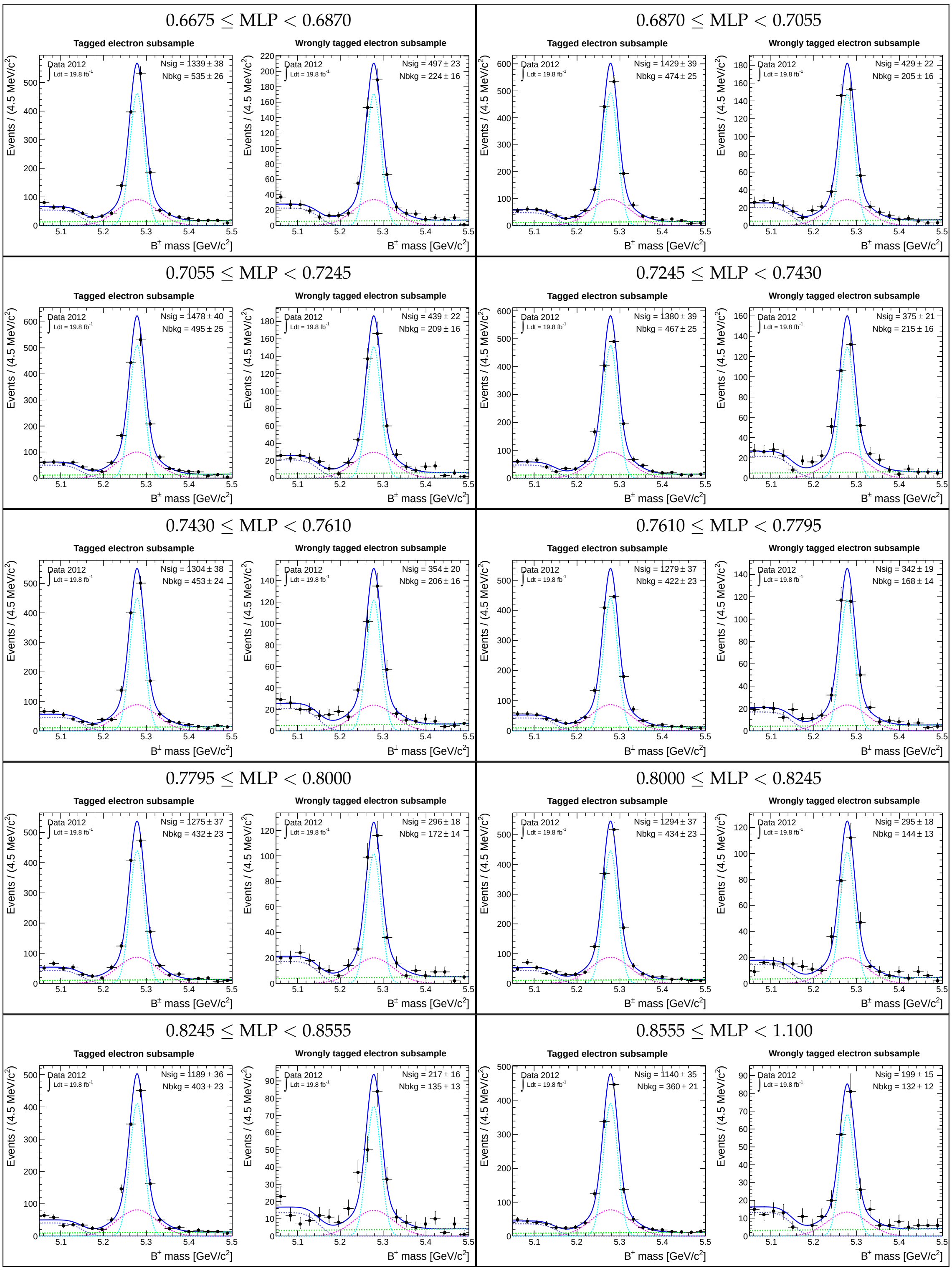


Figure A.24: Fit to B^+ invariant mass for the reconstructed $B^+ \to J/\psi\, K^+$ events containing a selected OS-electron in ranges of MLP.

Table A.9: Electron tagging performances evaluated on the reconstructed $B_s^0 \to J/\psi\,\phi$ simulated events. Twenty categories of MLP discriminator are used.

MLP	$[0, 0.4745)$	$[0.4745, 0.512)$	$[0.512, 0.5365)$	$[0.5365, 0.556)$	$[0.556, 0.573)$
ε_{tag}	0.166 ± 0.005	0.166 ± 0.005	0.168 ± 0.005	0.168 ± 0.005	0.17 ± 0.005
ω	52.277 ± 1.618	50.198 ± 1.623	47.024 ± 1.61	45.756 ± 1.608	43.327 ± 1.595
$\mathcal{P}_{tag}$	0 ± 0	0 ± 0	0.001 ± 0.001	0.001 ± 0.001	0.003 ± 0.001

MLP	$[0.573, 0.5905)$	$[0.5905, 0.608)$	$[0.608, 0.6265)$	$[0.6265, 0.6475)$	$[0.6475, 0.6675)$
ε_{tag}	0.17 ± 0.005	0.167 ± 0.005	0.167 ± 0.005	0.168 ± 0.005	0.168 ± 0.005
ω	45.217 ± 1.599	43.491 ± 1.612	41.437 ± 1.603	37.023 ± 1.573	33.659 ± 1.543
$\mathcal{P}_{tag}$	0.002 ± 0.001	0.003 ± 0.001	0.005 ± 0.002	0.011 ± 0.003	0.018 ± 0.003

MLP	$[0.6675, 0.687)$	$[0.687, 0.7055)$	$[0.7055, 0.7245)$	$[0.7245, 0.743)$	$[0.743, 0.761)$
ε_{tag}	0.167 ± 0.005	0.166 ± 0.005	0.169 ± 0.005	0.167 ± 0.005	0.167 ± 0.005
ω	31.403 ± 1.521	34.585 ± 1.56	28.752 ± 1.483	27.291 ± 1.47	25.148 ± 1.437
$\mathcal{P}_{tag}$	0.023 ± 0.004	0.016 ± 0.003	0.03 ± 0.004	0.034 ± 0.005	0.041 ± 0.005

MLP	$[0.761, 0.7795)$	$[0.7795, 0.8)$	$[0.8, 0.8245)$	$[0.8245, 0.8555)$	$[0.8555, 1.1)$
ε_{tag}	0.168 ± 0.005	0.167 ± 0.005	0.166 ± 0.005	0.166 ± 0.005	0.166 ± 0.005
ω	23.069 ± 1.393	21.534 ± 1.367	19.109 ± 1.318	19.147 ± 1.32	15.164 ± 1.214
$\mathcal{P}_{tag}$	0.049 ± 0.005	0.054 ± 0.005	0.063 ± 0.006	0.063 ± 0.006	0.08 ± 0.006

Table A.10: Combined electron tagging performances evaluated on the reconstructed $B_s^0 \to J/\psi\,\phi$ simulated events using twenty categories of MLP discriminator.

	$B_s^0 \to J/\psi\,\phi$ MC
ε_{tag}^{tot}	$3.35 \pm 0.02\%$
$\mathcal{P}_{tag}^{tot}$	$0.50 \pm 0.02\%$

Table A.11: Electron tagging performances evaluated on the reconstructed $B^+ \rightarrow J/\psi K^+$ simulated events. Twenty categories of MLP discriminator are used.

MLP	[0, 0.4745)	[0.4745, 0.512)	[0.512, 0.5365)	[0.5365, 0.556)	[0.556, 0.573)
ε_{tag}	0.174 ± 0.006	0.189 ± 0.006	0.174 ± 0.006	0.165 ± 0.006	0.175 ± 0.006
ω	47.692 ± 1.712	48.679 ± 1.645	52.481 ± 1.71	46.449 ± 1.763	47.917 ± 1.71
$\mathcal{P}_{tag}$	0 ± 0.001	0 ± 0	0 ± 0.001	0.001 ± 0.001	0 ± 0

MLP	[0.573, 0.5905)	[0.5905, 0.608)	[0.608, 0.6265)	[0.6265, 0.6475)	[0.6475, 0.6675)
ε_{tag}	0.158 ± 0.006	0.173 ± 0.006	0.173 ± 0.006	0.188 ± 0.006	0.179 ± 0.006
ω	44.121 ± 1.796	40.819 ± 1.7	41.62 ± 1.706	36.253 ± 1.599	37.406 ± 1.651
$\mathcal{P}_{tag}$	0.002 ± 0.001	0.006 ± 0.002	0.005 ± 0.002	0.014 ± 0.003	0.011 ± 0.003

MLP	[0.6675, 0.687)	[0.687, 0.7055)	[0.7055, 0.7245)	[0.7245, 0.743)	[0.743, 0.761)
ε_{tag}	0.174 ± 0.006	0.165 ± 0.006	0.166 ± 0.006	0.164 ± 0.006	0.162 ± 0.006
ω	34.763 ± 1.651	33.759 ± 1.687	31.019 ± 1.653	28.089 ± 1.618	25.089 ± 1.58
$\mathcal{P}_{tag}$	0.016 ± 0.004	0.017 ± 0.004	0.024 ± 0.004	0.032 ± 0.005	0.04 ± 0.005

MLP	[0.761, 0.7795)	[0.7795, 0.8)	[0.8, 0.8245)	[0.8245, 0.8555)	[0.8555, 1.1)
ε_{tag}	0.163 ± 0.006	0.157 ± 0.006	0.17 ± 0.006	0.175 ± 0.006	0.162 ± 0.006
ω	21.974 ± 1.512	21.883 ± 1.542	17.136 ± 1.36	17.982 ± 1.363	15.212 ± 1.335
$\mathcal{P}_{tag}$	0.051 ± 0.006	0.05 ± 0.006	0.073 ± 0.007	0.072 ± 0.007	0.079 ± 0.007

Table A.12: Combined electron tagging performances evaluated on the reconstructed $B^+ \rightarrow J/\psi K^+$ simulated events using twenty categories of MLP discriminator.

	$B^+ \rightarrow J/\psi K^+$ MC
ε_{tag}^{tot}	$3.41 \pm 0.03\%$
$\mathcal{P}_{tag}^{tot}$	$0.49 \pm 0.02\%$

Table A.13: Electron tagging performances evaluated on the reconstructed $B^0 \to J/\psi\, K^*$ simulated events. Twenty categories of MLP discriminator are used.

MLP	$[0, 0.4745)$	$[0.4745, 0.512)$	$[0.512, 0.5365)$	$[0.5365, 0.556)$	$[0.556, 0.573)$
ε_{tag}	0.174 ± 0.006	0.174 ± 0.006	0.184 ± 0.006	0.163 ± 0.006	0.161 ± 0.006
ω	47.736 ± 1.81	49.756 ± 1.808	47.222 ± 1.759	48.624 ± 1.876	45.093 ± 1.884
$\mathcal{P}_{tag}$	0 ± 0.001	0 ± 0	0.001 ± 0.001	0 ± 0	0.002 ± 0.001

MLP	$[0.573, 0.5905)$	$[0.5905, 0.608)$	$[0.608, 0.6265)$	$[0.6265, 0.6475)$	$[0.6475, 0.6675)$
ε_{tag}	0.173 ± 0.006	0.184 ± 0.006	0.181 ± 0.006	0.186 ± 0.007	0.169 ± 0.006
ω	44.691 ± 1.814	43.438 ± 1.754	41.363 ± 1.756	38.259 ± 1.715	33.628 ± 1.763
$\mathcal{P}_{tag}$	0.002 ± 0.001	0.003 ± 0.002	0.005 ± 0.002	0.01 ± 0.003	0.018 ± 0.004

MLP	$[0.6675, 0.687)$	$[0.687, 0.7055)$	$[0.7055, 0.7245)$	$[0.7245, 0.743)$	$[0.743, 0.761)$
ε_{tag}	0.175 ± 0.006	0.17 ± 0.006	0.183 ± 0.006	0.185 ± 0.006	0.167 ± 0.006
ω	31.258 ± 1.704	29.648 ± 1.707	25.874 ± 1.582	28.143 ± 1.61	27.714 ± 1.692
$\mathcal{P}_{tag}$	0.025 ± 0.005	0.028 ± 0.005	0.043 ± 0.006	0.035 ± 0.005	0.033 ± 0.005

MLP	$[0.761, 0.7795)$	$[0.7795, 0.8)$	$[0.8, 0.8245)$	$[0.8245, 0.8555)$	$[0.8555, 1.1)$
ε_{tag}	0.165 ± 0.006	0.172 ± 0.006	0.183 ± 0.006	0.168 ± 0.006	0.155 ± 0.006
ω	23.705 ± 1.63	23.668 ± 1.591	18.553 ± 1.424	17.617 ± 1.462	14.738 ± 1.435
$\mathcal{P}_{tag}$	0.046 ± 0.006	0.048 ± 0.006	0.072 ± 0.007	0.071 ± 0.007	0.077 ± 0.007

Table A.14: Combined electron tagging performances evaluated on the reconstructed $B^0 \to J/\psi\, K^*$ simulated events using twenty categories of MLP discriminator.

	$B^0 \to J/\psi\, K^*$ MC
$\varepsilon_{tag}^{\mathrm{tot}}$	3.47 ± 0.03
$\mathcal{P}_{tag}^{\mathrm{tot}}$	0.52 ± 0.02

Table A.15: Electron tagging performances evaluated on the reconstructed $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ events in 2012 data. Twenty categories of MLP discriminator are used.

MLP	$[0, 0.4745)$	$[0.4745, 0.512)$	$[0.512, 0.5365)$	$[0.5365, 0.556)$	$[0.556, 0.573)$
ε_{tag}	0.196 ± 0.006	0.203 ± 0.006	0.198 ± 0.006	0.198 ± 0.006	0.21 ± 0.006
ω	51.895 ± 2.528	48.343 ± 2.342	47.747 ± 2.379	46.459 ± 2.331	44.511 ± 2.207
$\mathcal{P}_{tag}$	0 ± 0.001	0 ± 0.001	0 ± 0.001	0.001 ± 0.001	0.003 ± 0.002

MLP	$[0.573, 0.5905)$	$[0.5905, 0.608)$	$[0.608, 0.6265)$	$[0.6265, 0.6475)$	$[0.6475, 0.6675)$
ε_{tag}	0.208 ± 0.006	0.2 ± 0.006	0.218 ± 0.006	0.203 ± 0.006	0.19 ± 0.005
ω	44.581 ± 2.213	42.092 ± 2.178	40.361 ± 2.027	38.874 ± 2.058	36.21 ± 2.028
$\mathcal{P}_{tag}$	0.002 ± 0.002	0.005 ± 0.003	0.008 ± 0.004	0.01 ± 0.004	0.014 ± 0.004

MLP	$[0.6675, 0.687)$	$[0.687, 0.7055)$	$[0.7055, 0.7245)$	$[0.7245, 0.743)$	$[0.743, 0.761)$
ε_{tag}	0.189 ± 0.005	0.202 ± 0.006	0.209 ± 0.006	0.195 ± 0.006	0.184 ± 0.005
ω	37.129 ± 2.065	30.051 ± 1.756	29.752 ± 1.712	27.173 ± 1.703	27.181 ± 1.745
$\mathcal{P}_{tag}$	0.013 ± 0.004	0.032 ± 0.006	0.034 ± 0.006	0.041 ± 0.007	0.038 ± 0.006

MLP	$[0.761, 0.7795)$	$[0.7795, 0.8)$	$[0.8, 0.8245)$	$[0.8245, 0.8555)$	$[0.8555, 1.1)$
ε_{tag}	0.181 ± 0.005	0.18 ± 0.005	0.183 ± 0.005	0.168 ± 0.005	0.161 ± 0.005
ω	26.78 ± 1.735	23.247 ± 1.598	22.795 ± 1.572	18.308 ± 1.454	17.474 ± 1.45
$\mathcal{P}_{tag}$	0.039 ± 0.006	0.052 ± 0.007	0.054 ± 0.007	0.068 ± 0.007	0.068 ± 0.007

Table A.16: Combined electron tagging performances evaluated on the reconstructed $\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ events in 2012 data using twenty categories of MLP discriminator.

	$\mathrm{B}^+ \to \mathrm{J}/\psi\,\mathrm{K}^+$ data
$\varepsilon_{tag}^{\mathrm{tot}}$	$3.88 \pm 0.03\%$
$\mathcal{P}_{tag}^{\mathrm{tot}}$	$0.48 \pm 0.02\%$

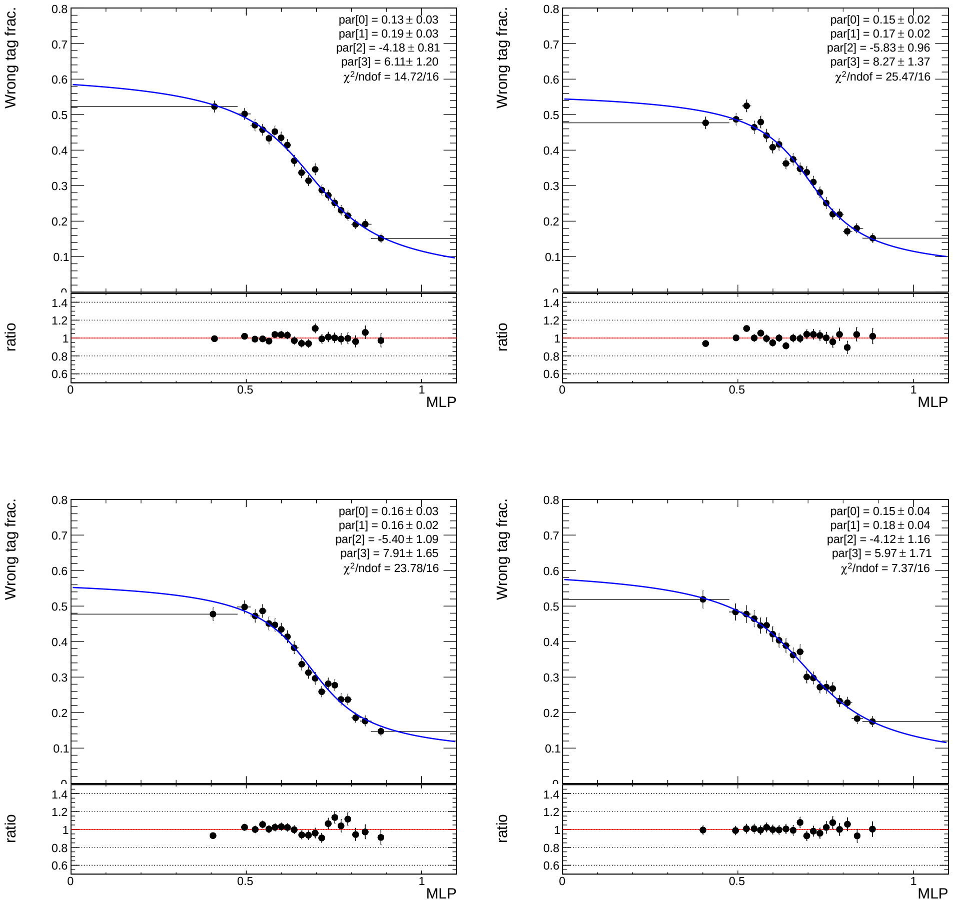


Figure A.25: OS-electron wrong tag fraction evaluated as a function of the MLP discriminator for the $B_s^0 \to J/\psi\,\phi$ MC (top-left), $B^+ \to J/\psi\,K^+$ MC (top-right), $B^0 \to J/\psi\,K^*$ MC (bottom-left), and 2012 $B^+ \to J/\psi\,K^+$ Data (bottom-right). A function of the form $p_0 + p_1 \cdot [1 - \arctan(p_2 + p_3 \cdot \mathrm{MLP})]$ is used for the fit.

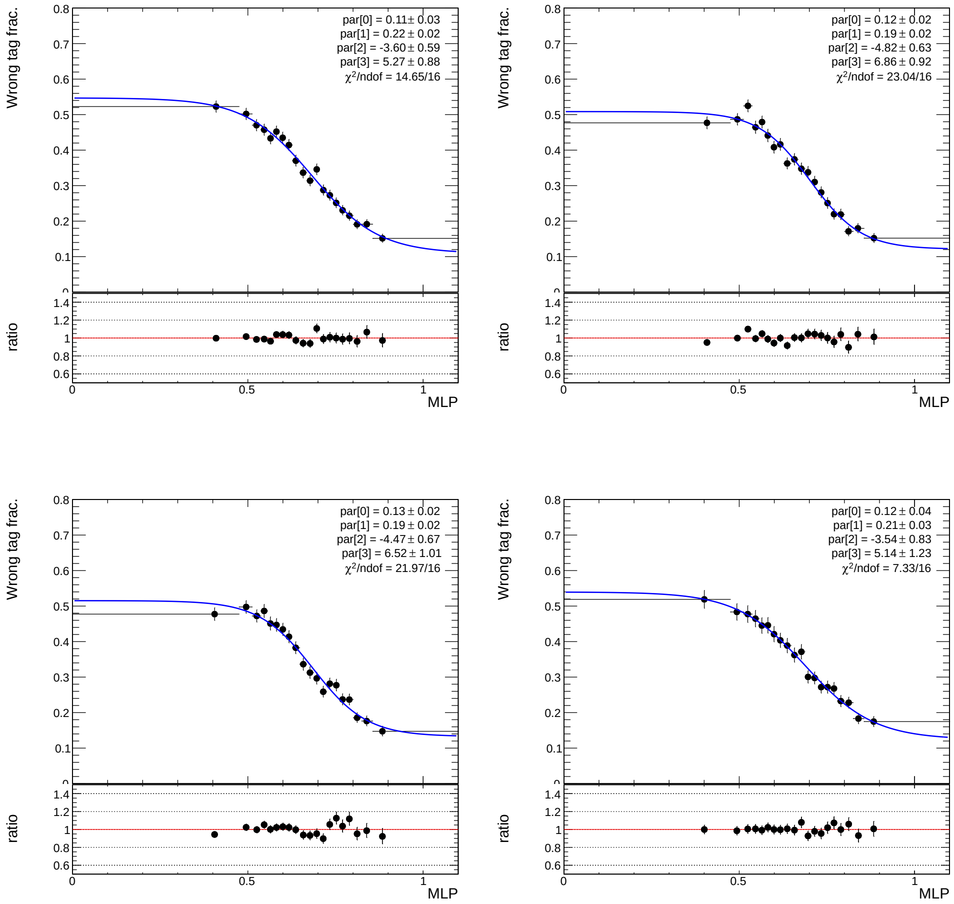


Figure A.26: OS-electron wrong tag fraction evaluated as a function of the MLP discriminator for the $B_s^0 \to J/\psi\,\phi$ MC (top-left), $B^+ \to J/\psi\,K^+$ MC (top-right), $B^0 \to J/\psi\,K^*$ MC (bottom-left), and 2012 $B^+ \to J/\psi\,K^+$ Data (bottom-right). A function of the form $p_0 + p_1 \cdot [1 - \tanh(p_2 + p_3 \cdot \mathrm{MLP})]$ is used for the fit.

Appendix B

Notes

B.1 Propagation of the uncertainties on $\mathcal{P}_{tag}$ and ω

The calibration curve is parametrized according to Equation B.1

$$\omega^{meas} = p_0 + p_1 \left(\omega^{calc} - \omega' \right) \tag{B.1}$$

where ω' is a constant value fixed to approximate the average value of the mistag distribution

$$\omega' = 0.350 \tag{B.2}$$

In order to fully exploit the mistag informations, the same calibration procedure described above can be applied for each event, in order to obtain a per event calibrated mistag

$$\omega_i^{meas} = p_0 + p_1 \left(\omega_i^{calc} - \omega_i' \right) \tag{B.3}$$

The overall tagging power can therefore be obtained for any given sample of N_{tot} events by summing the dilution terms

$$\begin{aligned}
\mathcal{P}_{tag} &= \frac{1}{N_{tot}} \sum_i^N \mathcal{D}_i^2 = & \text{(B.4)} \\
&= \frac{1}{N_{tot}} \sum_i^N (1 - 2\omega_i^{meas})^2 = & \text{(B.5)} \\
&= \frac{1}{N_{tot}} \sum_i^N \left[1 - 2 \left(p_0 + p_1 \left(\omega_i^{calc} - \omega' \right) \right) \right]^2 & \text{(B.6)}
\end{aligned}$$

The overall mistag fraction is obtained by the tagging power and the tagging efficiency

$$\omega = \frac{1}{2} \left(1 - \sqrt{\frac{\mathcal{P}_{tag}}{\varepsilon_{tag}}} \right) \tag{B.7}$$

which can be re-written in the following form

$$\omega = \frac{1}{2}\left(1 - \sqrt{\frac{\sum_i^N \left(1 - 2\omega_i^{meas}\right)^2}{N_{tag}}}\right) \tag{B.8}$$

where N_{tag} is the number of tagged events.

The tagging efficiency can be easily evaluated as fraction of tagged events within the investigated sample.

$$\varepsilon_{tag} = \frac{N_{tag}}{N_{tot}} \tag{B.9}$$

It it clear that ε_{tag} does not depend on the calibration parameters, and can be defined for any events sample.

The uncertainties on the calibration parameters (δp_0, δp_1), and the correlation between them (ρ), are then propagated in order to obtain the correct estimation of the uncertainties on $\mathcal{P}_{tag}$.

$$\delta\mathcal{P}_{tag} = \sqrt{\left(\frac{\partial\mathcal{P}_{tag}}{\partial p_0}\delta p_0\right)^2 + \left(\frac{\partial\mathcal{P}_{tag}}{\partial p_1}\delta p_1\right)^2 + 2\rho\frac{\partial\mathcal{P}_{tag}}{\partial p_0}\frac{\partial\mathcal{P}_{tag}}{\partial p_1}\delta p_0\delta p_1} \tag{B.10}$$

where

$$\frac{\partial\mathcal{P}_{tag}}{\partial p_0} = -\frac{4}{N_{tot}}\sum_i^N \left(1 - 2\omega_i^{meas}\right) \tag{B.11}$$

$$\frac{\partial\mathcal{P}_{tag}}{\partial p_1} = -\frac{4}{N_{tot}}\sum_i^N \left(1 - 2\omega_i^{meas}\right)\left(\omega_i^{calc} - \omega'\right) \tag{B.12}$$

The propagation of the uncertainties on the mistag follows it's analogous to what previously done for $\delta\mathcal{P}_{tag}$:

$$\delta\omega = \sqrt{\left(\frac{\partial\omega}{\partial p_0}\delta p_0\right)^2 + \left(\frac{\partial\omega}{\partial p_1}\delta p_1\right)^2 + 2\rho\frac{\partial\omega}{\partial p_0}\frac{\partial\omega}{\partial p_1}\delta p_0\delta p_1} \tag{B.13}$$

where

$$\frac{\partial\omega}{\partial p_0} = \frac{1}{\sqrt{N_{tag}}}\frac{\sum_i^N \left(1 - 2\omega_i^{meas}\right)}{\sqrt{\sum_j^N \left(1 - 2\omega_j^{meas}\right)^2}} \tag{B.14}$$

$$\frac{\partial\omega}{\partial p_1} = \frac{1}{\sqrt{N_{tag}}}\frac{\sum_i^N \left(1 - 2\omega_i^{meas}\right)\left(\omega_i^{calc} - \omega'\right)}{\sqrt{\sum_j^N \left(1 - 2\omega_j^{meas}\right)^2}} \tag{B.15}$$