Department of Physics Giuseppe Occhialini

PhD program in Physics and Astronomy                    Cycle XXXIV

Curriculum in Subnuclear Physics

# Test of lepton flavour universality with the simultaneous measurement of $R(D^{+})$ and $R(D^{*+})$ with $\tau^{-} \rightarrow \mu^{-}\,\nu_{\tau}\,\overline{\nu}_{\mu}$ decays at the LHCb experiment

Surname: Meloni          Name: Simone

Registration number: 763674

Tutor: Prof. Marta Calvi

Co-tutor: Dr. Julián García Pardiñas

Coordinator: Prof. Marta Calvi

**ACADEMIC YEAR 2020/2021**

*"The important thing is not to stop questioning; curiosity has its own reason for existing. One cannot help but be in awe when contemplating the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of the mystery every day. The important thing is not to stop questioning; never lose a holy curiosity."*

- Albert Einstein

*"Non temete i momenti difficili, il meglio viene da lì."*

- Rita Levi Montalcini

*"When you have eliminated the impossible, whatever remains, however improbable, must be the truth."*

- *Sherlock Holmes* - Sir Arthur Conan Doyle

## Abstract

In the Standard Model of particle physics, the coupling of the electroweak gauge bosons to the leptons is independent of the lepton flavour. This property, known as Lepton Flavour Universality, is an accidental symmetry of the Standard Model, which can be tested in semileptonic $b$-meson decays.

The variables used to test the Lepton Flavour Universality hypothesis are ratios of branching fractions between decays with the $\tau$ lepton and the ones with the $\mu$ lepton in the final state:

$$\mathcal{R}(H_c) = \frac{\mathcal{B}(B \to H_c \tau \nu)}{\mathcal{B}(B \to H_c \mu \nu)},$$

with $H_c$ a charmed meson produced in the decay. Any sign of deviation with respect to the Standard Model predictions in these variables could be a clear sign of New Physics effects.

A tension at the level of $3\sigma$ [1] with respect to the Standard Model predictions has been observed in the combination of the measurements of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ performed by the Belle, BaBar and LHCb collaborations. At the time of writing of this thesis, no measurement of the $\mathcal{R}(D)$ parameter has been performed by any hadron collider experiment.

This thesis reports a simultaneous measurement of the $\mathcal{R}(D^+)$ and $\mathcal{R}(D^{*+})$ parameters performed using $\overline{B}^0 \to D^{(*)+} \ell^- \overline{\nu}_\ell$ decays. This measurement exploits leptonic decays of the $\tau$ lepton, $\tau^- \to \mu^- \overline{\nu}_\mu \nu_\tau$, using a data sample corresponding to an integrated luminosity of $2.0 \text{ fb}^{-1}$ collected in proton-proton collisions at a centre-of-mass energy of $13 \text{ TeV}$ at the LHCb experiment during the 2015 and 2016 data taking years. All the steps of the analysis have been performed and the main systematic uncertainties have been studied. The values of the measured parameters remain still blinded and the analysis is in internal review within the LHCb collaboration. The expected uncertainties on the parameters of interest are given by

$$
\begin{aligned}
\mathcal{R}(D^+) &= xxx \pm 0.033 (\text{stat.}) \pm 0.037 (\text{syst.}), \\
\mathcal{R}(D^{*+}) &= xxx \pm 0.040 (\text{stat.}) \pm 0.070 (\text{syst.}).
\end{aligned}
$$

## Sommario

Nel Modello Standard delle interazioni fondamentali l'accoppiamento dei bosoni di gauge ai leptoni è indipendente dal flavour leptonico. Questa proprietà, conosciuta come Universalit di Flavour leptonico, è una simmetria accidentale del modello che pu essere testata in decadimenti semileptonici di mesoni contententi quark di tipo $b$.

Le variabili usate per testare l'Universalità di Flavour leptonico sono rapporti di ratei di decadimento tra transizioni con un leptone $\tau$ e transizioni con un leptone $\mu$ nello stato finale:

$$\mathcal{R}(H_c) = \frac{\mathcal{B}(B \to H_c \tau \nu)}{\mathcal{B}(B \to H_c \mu \nu)},$$

con $H_c$ che rappresenta un mesone contenente un quark $c$ prodotto nel decadimento. L'osservazione di qualsiasi segno di deviazione in queste variabili rispetto alle previsioni del Modello Standard potrebbe essere un chiaro segno dell'effetto di effetti di nuova fisica.

Combinando le misure dei parameteri $\mathcal{R}(D)$ e $\mathcal{R}(D^*)$ effettuate dalle collaborazioni Belle, BaBar e LHCb, è stata osservata una tensione rispetto alle previsioni del Modello Standard a livello di circa $3\sigma$. Ad oggi nessuna misura del parametro $\mathcal{R}(D)$ è mai stata effettuata a collider adronici.

Questa tesi riporta una misura simultanea dei parametetri $\mathcal{R}(D^+)$ e $\mathcal{R}(D^{*+})$ effettuata tramite l'analisi di decadimenti $\overline{B}^0 \to D^{(*)+}\ell^-\overline{\nu}_\ell$. Questa misura utilizza il decadimento leptonico del $\tau$, $\tau^- \to \mu^-\overline{\nu}_\mu\nu_\tau$, sfruttando un campione di $2.0\,\mathrm{fb}^{-1}$ di dati raccolto in collisioni protone-protone, ad un'energia nel centro di massa di $13\,\mathrm{TeV}$ dall'esperimento LHCb durante la presa dati degli anni 2015 e 2016. Tutti i passi dell'analisi sono stati effettuati e le principali incertezze sistematiche sono state valutate. Il valore dei parametri è ancora *blinded* e l'analisi è in revisione interna presso la collaborazione LHCb. L'incertezza attesa associata ai parametri di interesse è:

$$\begin{aligned}
\mathcal{R}(D^+) &= xxx \pm 0.033(\text{stat.}) \pm 0.037(\text{syst.}), \\
\mathcal{R}(D^{*+}) &= xxx \pm 0.040(\text{stat.}) \pm 0.070(\text{syst.}).
\end{aligned}$$

# Contents

# I  Foreword

Due to the huge amount of work required by the analysis that will be presented in this thesis, the results shown are the outcome of a fruitful collaboration of many individuals working on the LHCb experiment. The work performed by the Author would not have been possible without the help, support, guidance and advice of the rest of the analysis team, in the people of Marta Calvi, Julián García Pardiñas, Lucia Grillo, Patrick Owen and Nicola Serra.

The work presented in the following chapters is the result of independent efforts of the Author or work with guidance, advice and collaboration with the analysis team, except for what reported in the following list, which amounts to work entirely attributed to others:

- Optimization and definition of the selections reported in Sections 8.3.3, 8.3.4, 8.3.6;

- Emulation of the PID, neutral isolation and L0 response, reported in Sections 10.1, 10.2, 10.3;

- Data driven estimation of the $\mu$-MisID and $D\mu$-Combinatorial backgrounds reported in Sections 12.1 and 12.3;

- Data/MC corrections reported in Sections 13.1.

The LHCb experiment is a collaboration of $\approx 1000$ scientists working on different aspects of the detector and its data taking. The analysis presented in this thesis would not have been possible without the excellent performance of the LHCb detector, the data collected and the tools and infrastructure provided by the collaboration. The Author is grateful for the work of any individual that contributed to this endeavour.

October 2021,
Simone Meloni

_____

# 1 Theoretical introduction

## 1.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics [2–4], introduced in 1961 by Glashow, Weinberg and Salam, is the quantum theory of fields that best describes three of the four fundamental forces in Nature: the electromagnetic force, the weak force and the strong force.

Its lagrangian is constructed imposing the gauge invariance principle under the $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ group. The $SU(3)_C$ group, where C stands for colour, describes the strong interactions (Quantum Chromo Dynamics, QCD). The $SU(2)_L \otimes U(1)_Y$ group, where L stands for Left and Y for hypercharge, describes the electroweak (EW) interactions. The gravitational force is still not accounted for in the SM.

The SM describes the interaction between all elementary particles in nature, that are classified as either spin-$\frac{1}{2}$ fermions or spin-1 bosons, which constitute the force mediators. The only spin-0 particle introduced in the SM, called Higgs-Boson, is included in the theory in order to provide masses to all the particles in a gauge invariant way, through the mechanism of spontaneous symmetry breaking.

### 1.1.1 Matter content

The fermions constitute the matter content in the SM and are divided into leptons, particles that are affected only by the EW interactions, and quarks, particles that are affected by the EW and QCD interactions. Both fermions and quarks are divided into three generations, also called families. Particles in different families have the same quantum numbers, but are characterized by different masses.

The leptons are also divided into charged leptons, $(e, \mu, \tau)$ with unitary electric charge, and the corresponding neutral leptons $(\nu_e, \nu_\mu, \nu_\tau)$, called neutrinos, electrically neutral.

Six quarks exist: up $(u)$, down $(d)$, charm $(c)$, strange $(s)$, top $(t)$ and bottom $(b)$. The positively charged quarks $(u, c, t)$ carry $\frac{2}{3}$ of the fundamental electric charge, while the negatively charged quarks $(d, s, b)$ carry $-\frac{1}{3}$ of the fundamental electric charge. Both leptons and quarks appear as right- and left-handed spinors in the theory.

Left-handed quarks and leptons are introduced as doublets under $SU(2)_L$ transformations. Right-handed fermions are instead included as singlets under $SU(2)_L$ transformations.

The quarks $SU(2)_L$ doublets are composed by a $T_3 = \frac{1}{2}$ component (with $T_3 = \frac{\sigma_3}{2}$ being the third generator of the $SU(2)_L$ group, with $\sigma_i$, $i = 1, 2, 3$ Pauli matrices), represented by the up-type quark $(u, c, t)$, and a $T_3 = -\frac{1}{2}$ part, represented by the corresponding down-type quark $(d, s, b)$. In the same fashion, the lepton $SU(2)_L$ doublets are made up by the left-handed electron $(e_L)$, muon $(\mu_L)$ and tau $(\tau_L)$, that constitute the $T_3 = -\frac{1}{2}$ component, paired with their corresponding $T_3 = \frac{1}{2}$ neutrinos $(\nu_e, \nu_\mu, \nu_\tau)$.

The charges of the Standard Model particles satisfy the Klein-Nishijima condition

$$Q = Y + T_3, \tag{1}$$

where $Q$ is the electric charge and $Y$ is the hypercharge associated with the $U(1)_Y$ group. Following this relation, all the right-handed quarks $(u_R, d_R, c_R, s_R, t_R, b_R)$ and right-handed charged fermion fields $(e_R, \mu_R, \tau_R)$ are charged with respect to the $U(1)_Y$ group, whereas the

right-handed neutrinos, neutral with respect to the electric charge and singlets with respect to $SU(2)_L$, are neutral also with respect to the $U(1)_Y$ group.

Being completely neutral to any interaction, right handed neutrinos are not included in the SM and therefore no Yukawa coupling or Majorana mass term is present for them. For this reason neutrinos do not acquire a mass from the spontaneous symmetry breaking mechanism. This is an obvious shortcoming of the SM, since an indirect indication that neutrinos instead do have a mass comes from the observation of the neutrino flavour oscillation mechanism. The matter content of the model is summarised in Table 1.

| symbol | field | $T$ | $T_3$ | $Y$ | $Q$ |
|--------|-------|-----|-------|-----|-----|
| $L_L$ | $\binom{\nu_{eL}}{e_L}\binom{\nu_{\mu L}}{\mu_L}\binom{\nu_{\tau L}}{\tau_L}$ | $\frac{1}{2}$ | $\binom{+\frac{1}{2}}{-\frac{1}{2}}$ | $-\frac{1}{2}$ | $\binom{0}{-1}$ |
| $E_R$ | $e_R\ \mu_R\ \tau_R$ | $0$ | $0$ | $-1$ | $0$ |
| $Q_L$ | $\binom{u_L}{d_L}\binom{c_L}{s_L}\binom{t_L}{b_L}$ | $\frac{1}{2}$ | $\binom{+\frac{1}{2}}{-\frac{1}{2}}$ | $\frac{1}{6}$ | $\binom{\frac{2}{3}}{-\frac{1}{3}}$ |
| $U_R$ | $u_R\ c_R\ t_R$ | $0$ | $0$ | $\frac{2}{3}$ | $\frac{2}{3}$ |
| $D_R$ | $d_R\ s_R\ b_R$ | $0$ | $0$ | $-\frac{1}{3}$ | $-\frac{1}{3}$ |

Table 1: Matter content of the Standard Model of Particle Physics.

### 1.1.2 Interactions

The interactions between particles in the SM are introduced in the theory by imposing a local gauge invariance of the lagrangian under transformations of the group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. This requirement uniquely determines the interactions between the particles, that are interpreted as the exchange of spin-1 mediators called gauge bosons. The number of the force-carrying bosons is determined by the number of generators of the gauge group. Their mass is reported in Table 2.

| interaction | bosons | mass |
|-------------|--------|------|
| Electromagnetic | $\gamma$ | $0$ |
| Weak | $\binom{W^\pm}{Z^0}$ | $\binom{80.4\ \text{GeV}}{91.2\ \text{GeV}}$ |
| Strong | $g_1 \dots g_8$ | $0$ |

Table 2: Gauge bosons of in the Standard Model along with their (measured) masses.

The interaction terms in the lagrangian are generated by substituting the covariant derivative associated to the gauge group in the most general Lorentz invariant lagrangian that describes the free fields defined in Table 1.

**The electroweak interactions** The SM is able to provide a unification framework for two of the known fundamental interactions: the weak and the electromagnetic interaction.

The gauge fields associated to the generators of the weak isospin $SU(2)_L$ group and the one associated to the weak hypercharge group $U(1)_Y$ are denoted by $W_i^\mu (i = 1, 2, 3)$ and $B^\mu$, respectively. The electroweak gauge covariant derivative is

$$D_\mu = \left( \partial_\mu - igW_\mu^i T^i - i\frac{1}{2}g'YB_\mu \right), \ (i = 1, 2, 3), \tag{2}$$

where $g$ and $g'$ are the coupling constants of the $SU(2)_L$ and $U(1)_Y$ groups, respectively and $T^i$ are the generators of the $SU(2)_L$ group. In the unification process, the physical mediators are expressed as linear combinations of the gauge fields:

$$
\begin{aligned}
W_\mu^\pm &= \frac{1}{\sqrt{2}} \left( W_\mu^1 \mp iW_\mu^2 \right), \\
A_\mu &= -\sin\theta_W W_\mu^3 + \cos\theta_W B_\mu, \\
Z_\mu &= \cos\theta_W W_\mu^3 + \sin\theta_W B_\mu,
\end{aligned} \tag{3}
$$

where $\theta_W$, called Weinberg angle, is defined by the following relation

$$\theta_W = \arctan\left( \frac{g'}{g} \right). \tag{4}$$

The mediator of the electromagnetic force is the photon ($\gamma$), associated to the field $A^\mu$, and the mediators of the weak force are the neutral and charged bosons $Z^0$ and $W^\pm$.

The electroweak lagrangian can be written, for a single generation of fermions and quarks, as

$$\mathcal{L}_{EW} = i\overline{L}_L \slashed{D} L_L + i\overline{\ell}_R \slashed{D} \ell_R + \tag{5}$$

$$+ i\overline{Q}_L \slashed{D} Q_L + i\overline{u}_R \slashed{D} u_R + i\overline{d}_R \slashed{D} d_R. \tag{6}$$

This can be split into a kinematic part

$$\mathcal{L}_{kin} = i\overline{L}_L \slashed{\partial} L_L + i\overline{\ell}_R \slashed{\partial} \ell_R + \tag{7}$$

$$+ i\overline{Q}_L \slashed{\partial} Q_L + i\overline{u}_R \slashed{\partial} u_R + i\overline{d}_R \slashed{\partial} d_R, \tag{8}$$

a charged current part

$$\mathcal{L}_{CC} = \frac{g}{\sqrt{2}} W_\mu^+ \overline{\nu}_L \gamma^\mu \ell_L + \frac{g}{\sqrt{2}} W_\mu^- \overline{\ell}_L \gamma^\mu \nu_L + \tag{9}$$

$$+ \frac{g}{\sqrt{2}} W_\mu^+ \overline{u}_L \gamma^\mu d_L + \frac{g}{\sqrt{2}} W_\mu^- \overline{d}_L \gamma^\mu u_L, \tag{10}$$

and a neutral current part

$$\mathcal{L}_{NC} = e\overline{\Psi}\gamma_\mu \mathcal{Q}\Psi A_\mu + \overline{\Psi}\gamma^\mu \mathcal{Q}_Z \Psi Z_\mu, \tag{11}$$

$$\tag{12}$$

where $\Psi$ is a column vector containing all the left and right-handed fields and $\mathcal{Q}$ and $\mathcal{Q}_Z$ are diagonal matrices containing, respectively, their electric charge and their coupling to the $Z$ boson, which reads

$$(\mathcal{Q}_Z)_{ii} = \frac{e}{\cos\theta_W \sin\theta_W}(T_3^i - Q_i \sin^2\theta_W). \tag{13}$$

**The strong interactions**   The theory of strong interactions, called Quantum Chromo Dynamics (QCD), is defined by the $SU(3)_C$ gauge group. The force is mediated by 8 massless gauge bosons, called gluons and associated to the fields $A_\mu^a (a = 1, \ldots, 8)$. The QCD gauge covariant derivative is:

$$D_\mu = \left(\partial\mu - i g_S A_\mu^a t^a\right), \ (a = 1, \ldots, 8), \tag{14}$$

where $t^a$ are the generators of the $SU(3)_C$ group and $g_S$ is the strong interaction coupling constant.

For a single generation of quarks, the QCD lagrangian reads

$$\mathcal{L}_{QCD} = i\overline{Q}_L \slashed{D} Q_L + i\overline{u}_R \slashed{D} u_R + i\overline{d}_R \slashed{D} d_R. \tag{15}$$

### 1.1.3   The Higgs mechanism and the flavour structure of the Standard Model

Starting from a free massless Dirac lagrangian with all the fermions listed in Table 1, and substituting the covariant derivatives of Eq. 2 and Eq. 14 results in the generation of the interaction terms for the electroweak and strong interactions, respectively, but not for mass terms for any particle in the model. Insertion by hand of such terms would violate the gauge symmetry of the lagrangian. In fact, a fermion mass term, that takes the form of

$$m\overline{\psi}\psi = m\left(\overline{\psi}_L \psi_R + \overline{\psi}_R \psi_L\right), \tag{16}$$

where $\psi$ is the fermion field, would violate the $SU(2)_L$ gauge invariance whereas a vectorial mass term of the form

$$m^2 A^\mu A_\mu, \tag{17}$$

where $A^\mu$ is the vector field, would violate the $U(1)_Y$ gauge invariance.

Therefore, an alternative way of dinamically generate the mass terms is needed. This is achieved in the SM through the Spontaneous Symmetry Breaking (SSB) mechanism, also commonly called the Higgs mechanism [5–8].

A new field $\phi$, called the Higgs field, is introduced, along with a potential that depends on it. In order to provide mass terms for the electroweak gauge bosons in a gauge invariant way, the $SU(2)_L \otimes U(1)_Y$ invariance is spontaneously broken, meaning that the ground state of the lagrangian is not invariant under local gauge transformations, while the full lagrangian remains fully invariant.

The Higgs field is represented as a doublet of complex scalar fields:

$$\phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}, \tag{18}$$

4

transforming as a doublet ($T = 1/2$) under the $SU(2)_L$ transformations and with a hypercharge $Y(\phi) = 1$. The terms relative to the dynamics of this field in the SM lagrangian are given by:

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi^\dagger \phi) \tag{19}$$

where $V(\phi^\dagger \phi)$ is a scalar potential that can depend only on the modulo of the $\phi$ field. In order to spontaneously break the gauge symmetry, the vacuum of this potential should be non-invariant with respect to the gauge symmetry, but to preserve gauge invariance the potential itself should be invariant under gauge transformations. For these reasons, the potential should exhibit a set of vacuum states that is gauge invariant as a whole, but each of them having a non-vanishing vacuum expectation value:

$$\langle 0 | \phi | 0 \rangle \neq 0. \tag{20}$$

The Higgs potential inserted in the SM has the following expression:

$$V(\phi^\dagger \phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \tag{21}$$

with $\mu^2$ and $\lambda$ two real valued parameters. As it can be seen from Fig. 1, the degenerate ground states lie on a circumference centered in $\phi \equiv 0$ and of radius $|\phi| = \dfrac{\mu}{\sqrt{2\lambda}} = \dfrac{v}{\sqrt{2}}$, where $v = \dfrac{\mu}{\lambda}$ is called the vacuum expectation value of the Higgs field.
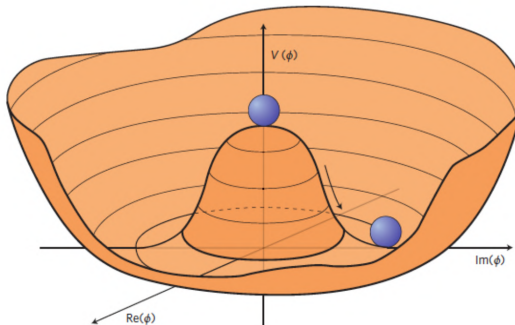


Figure 1: Sketch of the Higgs potential as a function of the real and imaginary part of the Higgs field.

When the system chooses one of the many possible vacuum states, the vacuum will not be symmetric anymore under local gauge transformations and the symmetry is said to be spontaneously broken. Expanding the field $\phi$ around the minimum, one can parameterize the fields $\phi$ in this way:

$$\phi = \frac{1}{\sqrt{2}} \exp\left[\frac{i\sigma_i \theta^i(x)}{v}\right] \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{22}$$

This shows the freedom in choosing one of the many minima of the potential: all minima in the set are linked to each other by an $SU(2)_L$ gauge transformation, $U(x) = \exp\left[\frac{i\sigma_i \theta^i(x)}{v}\right]$, and the fields $\theta_i(x)$ that appear in this transformation are called Goldstone bosons. These

fields correspond to unphysical degrees of freedom which can be rotated away with a gauge transformation $U^{-1}(x)$.

$$\phi(x) \rightarrow \phi'(x) = U^{-1}(x)\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{23}$$

The field $H(x)$ is called the Higgs field, and this gauge choice, also called unitary gauge, corresponds to having chosen a vacuum state with the vacuum expectation value of

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}. \tag{24}$$

In this choice, the Goldstone bosons have been reabsorbed and their degrees of freedom will end up giving mass to the $W^{\pm}$ and $Z^0$ vector bosons. It can be shown that this vacuum is explicitly invariant under the electric charge operator. Therefore the electric charge symmetry remains unbroken and neither photons nor gluons acquire any mass. By expanding the potential using this vacuum choice one gets the following:

$$V(\phi^{\dagger}\phi) = -\mu^2\phi^{\dagger}\phi + \lambda(\phi^{\dagger}\phi)^2 = \frac{1}{2}(2\lambda\nu^2)H^2(x) + \lambda\nu H^3(x) + \frac{\lambda}{4}H^4(x) - \frac{\lambda}{4}\nu^4 \tag{25}$$

The Higgs field acquires a mass of $m_H = \sqrt{2\lambda v^2}$, and cubic and quartic self-coupling interactions appear in the lagrangian. Additionally a constant term, irrelevant in the SM, appeared.

**Generation of the vector boson masses**   Masses for the vector bosons are generated after spontaneous symmetry breaking by the kinematic terms of the SM lagrangian

$$\mathcal{L}_{\text{Higgs,kin}} = (D_{\mu}\phi)^{\dagger}(D^{\mu}\phi) \tag{26}$$

$$= \frac{1}{2}\partial^{\mu}H\partial_{\mu}H + \left[\left(\frac{gv}{2}\right)W^{\mu+}W^-_{\mu} + \frac{1}{2}\frac{(g^2 + g'^2)v^2}{4}Z^{\mu}Z_{\mu}\right]\left(1 + \frac{H}{v}\right)^2. \tag{27}$$

This contains a kinematic term for the Higgs field, two mass terms for the $W^{\pm}$ and $Z^0$ bosons, which have acquired masses

$$m_W^2 = \frac{g^2 v^2}{4} \tag{28}$$

$$m_Z^2 = \frac{(g^2 + g'^2)v^2}{4} = \frac{m_W^2}{\cos^2\theta_W}, \tag{29}$$

and terms for the interaction of $Z$ and $W$ bosons with the Higgs boson.

## 1.2   Generation of the fermion masses

Along with the inclusion of the kinematic and potential terms of the field $\phi$ inside the SM lagrangian, all the possible Lorentz invariant, gauge invariant and renormalizable interaction terms of the field $\phi$ with fermions are included. These take the name of Yukawa interactions

$$\mathcal{L}_{\text{Yukawa}} = -\,\Gamma_d^{ij}\overline{Q'}_L^i \phi d_R^{\prime j} - \Gamma_d^{ij*}\overline{d'}_R^i \phi^\dagger Q_L^{\prime j} + \tag{30}$$

$$-\Gamma_u^{ij}\overline{Q'}_L^i \phi_c u_R^{\prime j} - \Gamma_u^{ij*}\overline{u'}_R^i \phi_c^\dagger Q_L^{\prime j} + \tag{31}$$

$$-\Gamma_\ell^{ij}\overline{L'}_L^i \phi \ell_R^{\prime j} - \Gamma_\ell^{ij*}\overline{\ell'}_R^i \phi^\dagger L_L', \tag{32}$$

where $\phi_c = i\sigma_2 \phi^* = \frac{1}{\sqrt{2}}\binom{v+H(x)}{0}$ is the charge conjugate of the field $\phi$. The matrices $\Gamma$ are $3 \times 3$ complex matrices, with indices that span the generations space. They are not diagonal, and therefore no mass term can be identified in this lagrangian. For this reason the fields have been denoted with a $'$ sign, to testify the fact that these are generic linear combinations of the mass eigenstates.

After spontaneous symmetry breaking, in the unitary gauge the Yukawa interactions can be written as

$$\mathcal{L}_{\text{Yukawa}} = -\Gamma_d^{ij}\frac{(v+H(x))}{\sqrt{2}}\overline{d'}_L^i d_R^{\prime i} - \Gamma_u^{ij}\frac{(v+H(x))}{\sqrt{2}}\overline{u'}_L^i u_R^{\prime i} - \Gamma_\ell^{ij}\frac{(v+H(x))}{\sqrt{2}}\overline{\ell'}_L^i \ell_R^{\prime i} + h.c. \tag{33}$$

$$= -\left[ M_d^{ij}\overline{d'}_L^i d_R^{\prime j} + M_u^{ij}\overline{u'}_L^i u_R^{\prime j} + M_\ell^{ij}\overline{\ell'}_L^i \ell_R^{\prime j} \right]\left(1 + \frac{H}{v}\right), \tag{34}$$

with $M^{ij} = \Gamma^{ij}\frac{v}{\sqrt{2}}$. It is known that for any generic complex square matrix $C$, there exists two unitary matrices $U, V$ such that

$$D = U^\dagger C V \tag{35}$$

is diagonal with real positive entries. Using this fact it can be seen that after spontaneous symmetry breaking, mass terms for the fermion fields are generated by the Yukawa interactions. This can be achieved by defining, for each fermion type $f = u, d, \ell$, two unitary matrices $U_L^i$ and $U_R^f$ such that $(U_L^f)^\dagger M_f U_R^f$ is diagonal with positive entries.

By rotating the fermionic fields in this way:

$$f'_{Li} = (U_L^f)_{ij} f_{Lj} \tag{36}$$

$$f'_{Ri} = (U_R^f)_{ij} f_{Rj} \tag{37}$$

the Yukawa lagrangian takes the following form

$$\mathcal{L}_{\text{Yukawa}} = -\sum_{f',i,j} \overline{f'}_L^i M_f^{ij} f_R^{\prime j} \left(1 + \frac{H}{v}\right) + h.c. \tag{38}$$

$$= -\sum_{f,i,j} \overline{f}_L^i \left[(U_L^f)\dagger M_f U_R^f\right]_{ij} f_R^j \left(1 + \frac{H}{v}\right) + h.c. \tag{39}$$

$$= -\sum_f m_f (\overline{f}_L f_R + \overline{f}_R f_L)\left(1 + \frac{H}{v}\right). \tag{40}$$

One can recognize in this lagrangian both mass terms and interaction terms between fermions and the Higgs boson, that will be proportional to the mass of the fermion.

The transformation performed on the fermion fields preserves unvaried all the kinematic and dynamic terms of the lagrangian, apart from the charged current lagrangian, $\mathcal{L}_{CC}$.

For the lepton sector

$$\mathcal{L}_{CC}^{\text{leptons}} = \frac{g}{\sqrt{2}} W_\mu^+ \overline{\nu}_L \gamma^\mu \ell'_L + h.c.. \tag{41}$$

The redefinition of the field $\ell'_L$ can be reabsorbed in the redefinition of the $\nu_L$ field: in the SM, as a consequence of the absence of $\nu_R$ fields, the neutrinos remain massless and the charged current interactions remain diagonal on the mass basis. Therefore any interaction in the SM will conserve the lepton flavour number, defined as $N_f = \#(f) + \#(\nu_f) - \#(\overline{f}) - \#(\overline{\nu}_f)$.

For the quark sector

$$\mathcal{L}_{CC}^{\text{quarks}} = \frac{g}{\sqrt{2}} W_\mu^+ \overline{u'}_L^i \gamma^\mu d'^i_L + h.c. \tag{42}$$

$$= \frac{g}{\sqrt{2}} W_\mu^+ \overline{u}_L^i \gamma^\mu d_L^J \left[ (U_L^u)^\dagger U_L^d \right]_{ij}. \tag{43}$$

The charged current interactions are not diagonal in the mass eigenstates basis. The non-diagonal complex matrix that arises in the interactions is called the Cabibbo-Kobayashi-Maskawa matrix

$$V_{CKM} = (U_L^u)^\dagger U_L^d = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{td} & V_{tb} \end{pmatrix} \tag{44}$$

For $N$ flavours, it depends on $(N-1)^2$ parameters, out of which $(N-1)(N-2)/2$ are complex phases. In the SM, where $N = 3$, this matrix will have one complex phase which is the only source of CP violation in the Standard Model.

## 1.3 The unsolved problems in the Standard Model

The SM of Particle Physics has been able to describe most observed physical phenomena in terms of a relatively small set of parameters and it is, at this date, the most tested theory available to describe the fundamental interactions of particles. It has been tested with outstanding accuracy in many experiments and so far no sizeable deviation from its predictions has been found and the theory still holds up to the threshold of the TeV scale. However, the SM is thought not to be a completely exhaustive theory because some issues still remain open and cannot be solved in its framework:

- **Hierarchy and unification:** the coupling constants of the different forces have values that span over many orders of magnitude and, although they run with the energy scale, no convergence point between them can be found.

- **Naturalness:** the finiteness of the mass of the Higgs boson can be achieved only by a massive and unnatural *fine-tuning* of the radiative corrections.

- **Dark Matter:** there is no good particle candidate for Dark Matter, that has been shown, from astrophysical evidences, to constitute almost 30% of the matter in the universe.

- **Baryogenesis:** the level of CP violation predicted in the SM is insufficient to explain the huge matter/anti-matter asymmetry observed in the Universe.

The hierarchy and the naturalness problems are intimately related; in fact the SM can be considered an effective field theory valid from the tested energy scales up to an unknown energy scale $\Lambda$. Effects coming from unobserved degrees of freedom that have higher mass than the energy scale $\Lambda$ can contribute in an indirect way to the physics at lower energy, and therefore the low-energy observables (masses, couplings...) are expected to be functions of the parameters of a more fundamental theory valid at a scale $Q > \Lambda$.

# 2 Lepton Flavour universality

With the discovery of the Higgs boson, the list of particles that make up the matter content of the Standard Model of Particle physics has been entirely discovered. Bounds on New Physics contributions have been put by both precision tests of the Standard Model performed and even more stringent constraints have been put thanks to direct New Physics searches performed at colliders.

With the lack of a direct evidence of BSM particles, a renewed interest in indirect searches of New Physics, through the precise measurement of Standard Model observables, has been seen in the High Energy Physics community. In the $B$ physics sector this has been especially true, thanks to the hints of anomalies with respect to the Standard Model prediction in quantities related to the *Lepton Flavour Universality* (LFU) hypothesis in $B$ meson decays that have been observed in the latest years.

In this chapter the LFU hypothesis and its origin in the Standard Model will be briefly discussed. Then, the tests that have been performed at the LEP collider in this context will be outlined, and lastly $b \rightarrow c\ell\nu$ decays as a test bed for the LFU hypothesis as a precision test of the Standard Model will be discussed.

## 2.1 Lepton Flavour universality in the Standard Model

One of the key features of the Standard Model is the presence of three lepton families and three quark families, which are inserted in the Model without any distinction between them except the mass of their consistuents and the value of their Yukawa couplings. The three fermion field families have the same gauge charge assignment, which leads to universality.

After Electroweak Symmetry Breaking, the Yukawa terms get diagonalised and mass terms for all the particles get generated. As it has been described in the previous chapter, this rotation has the effect of generating interactions between the $W^{\pm}$ gauge bosons and the quark mass eigenstates which are not diagonal anymore. The interactions between the quarks and the $W^{\pm}$ bosons in the Standard Model are therefore not diagonal in the three quark families and weighted through the CKM matrix.

This redefinition of the interactions, however, does not take place in the neutral interactions sector, thanks to the unitarity of the rotation matrices used to diagonalise the quarks and leptons mass terms. For this reason the interaction of $Z$ bosons with the leptons and quarks turns out to be universal amongst the three families. For the same reason transitions of the type $Z \rightarrow \ell\ell'$, with $\ell \neq \ell'$ are prohibited in the Standard Model. This in literature is read as the absence of Flavour Changing Neutral Currents (FCNC) at tree level.

The universality of the couplings between the gauge bosons and the particles takes place also in the charged electroweak lepton sector, thanks to the absence of right-handed neutrino fields. In fact this enables a redundancy in the definition of the Yukawa rotation matrix, that can be reabsorbed in the definition of the fields. For this reason, also the interaction between the $W^{\pm}$ bosons and the leptons is diagonal and universal with respect to the three lepton families.

The absence of transitions of the type $W^{-} \rightarrow \ell\overline{\nu}_{\ell'}$ and $Z \rightarrow \ell\ell'$, with $\ell \neq \ell'$, in the Standard Model is called *Lepton Flavour Number Conservation*. The equality of the gauge bosons couplings to the three lepton families in the Standard Model is called *Lepton Flavour Universality*. Both properties in the Standard Model arise as consequence of an accidental symmetry, a by product of the way the matter content is organized in the model; any discrepancy with respect to the

predictions of these two hypotheses could be a clear sign of the presence of physics Beyond the Standard Model.

This thesis will present a test of Lepton Flavour Universality with $b \to c\ell\nu_\ell$ transitions.

## 2.2 Tests of Lepton Flavour Universality in low energy physics and $W$ and $Z$ decays

Precision tests of lepton flavour universality have been carried out in both the charged current and neutral current sectors by many experiments, exploiting both low and high energy processes and providing stringent tests of the Lepton Flavour universality hypothesis.

Amongst the most precise low energy physics tests of $e - \mu$ universality in the charged sector is the comparison between the decay rates of the $K$ and $\pi$ mesons to leptons [9–13], $(K^-/\pi^-) \to \mu^- \overline{\nu}_\mu$ and $(K^-/\pi^-) \to e^- \overline{\nu}_e$, and semileptonic decays of the $K$ meson, $K \to \pi\mu\nu$ and $K \to \pi e\nu$ [14], which provide results compatible with the universality hypothesis at the order of $\approx 0.2\%$.

Very stringent constraint then came from the decays of the tau lepton and the precise measurement of its lifetime and mass. Thanks to the precision to which the decays of the $\tau$ can be predicted in the Standard Model and thanks to the high variety of open final states, consequence of its high mass, this particle has always been an ideal test bench for Lepton Flavour Universality.

By comparing the decay rates of $\tau^- \to \mu^- \nu_\tau \overline{\nu}_\mu$ and $\tau^- \to e^- \nu_\tau \overline{\nu}_e$, it has been possible to further reduce the constraint on $e - \mu$ universality in charged current decays. Very stringent constraints on $\tau - \mu$ universality in the charged current sector have also been obtained by comparing the decay of the $\tau$ and $\mu$ lepton to electrons. Furthermore the $\tau - \mu$ universality has been tested comparing the decay rates of $\tau \to \mu\nu\nu$ and the lifetime of the $\mu$, since its branching ratio is saturated by decays to electrons, $\mu \to e\nu\nu$ [15].

The most stringent constraints on lepton flavour universality of the $W^\pm$ bosons has been achieved at LEP2, when the energy in the electron beams was enough to produce real $W$ bosons pairs through $e^+e^- \to W^+W^-$. The comparison between the $W \to \ell\nu_\ell$ branching ratios provides the most stringent constraints on lepton universality in the charged sector, with a precision of the order of 0.1% [16]. The only notable discrepancy was in the ratio between the $W \to \tau\nu$ and $W \to \mu\nu$ rates, which has been found to be discrepant with the lepton flavour universality hypothesis at a level of $2.7\sigma$. This discrepancy has then been resolved by a recent measurement of the $\mathcal{B}(W \to \tau\nu)/\mathcal{B}(W \to \mu\nu)$ performed by the ATLAS collaboration [17].

At LEP, lepton flavour universality has also been precisely tested in the neutral sector, by precisely measuring the partial widths for the $Z \to \ell\overline{\ell}$ process in the various lepton final states, which have been found to be equal to one another at the 0.2% level [18].

Lepton Flavour universality of the $Z$ bosons has also been tested by the precision measurements of the couplings of the $Z$ boson to the leptons, extracted by a global fit to the Forward Backward asymmetries measured by LEP and the Left Right asymmetries measured by experiments at the SLC, which operated with polarized $e^+e^-$ beams. The result, as reported in Fig. 2, has been found to be compatible with the SM expectation of Lepton Flavour universality [18].
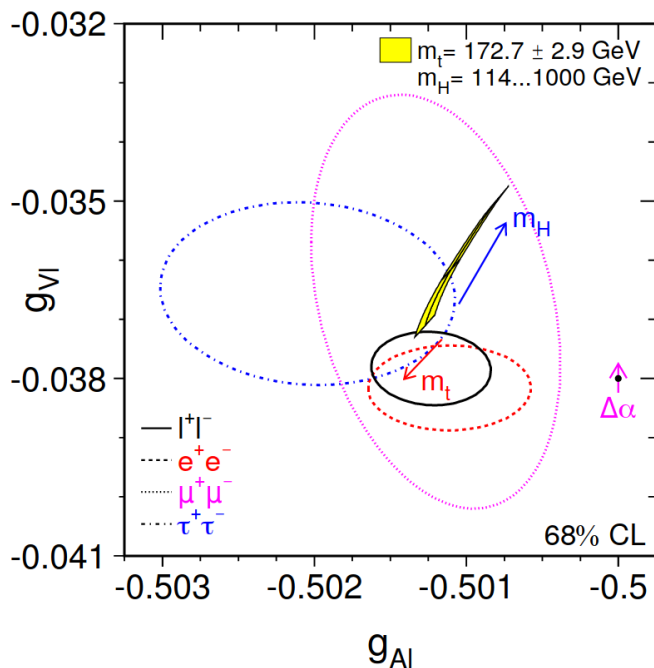
Figure 2: Combined LEP and SLD measurements of the vector and axial couplings of leptons to the $Z$ boson. The shaded region shows the SM prediction. [18]

## 2.3  Tests of lepton flavour universality in $b$ meson decays

Lepton flavour universality in both the charged and neutral current interactions seemed to be well established up until recently, when a new interest in this topic has been sparked by some hints of violation of LFU in measurements, performed at collider experiments, exploiting $B$-mesons semileptonic decays. The experiments that have performed these measurements are LHCb, and the Belle [19] and BaBar [20] experiments, working at two different $e^+e^-$ colliders operating in the same period, called KEKB and SLAC respectively. Both accelerators are $e^+e^-$ circular asymmetric colliders, operating at $\Upsilon(4S)$ resonance, that decays into coherent $b\bar{b}$ pairs.

Lepton Flavour universality has been tested in both the neutral currents sector, through measurements performed in $b \to s\ell\ell$ transitions, and in the charged currents sector, through measurements performed in $b \to c\ell\nu$ transitions.

### 2.3.1  $b \to s\ell\ell$ transitions

Decays of $B$ hadrons mediated by a $b \to s\ell\ell$ transition provide an ideal laboratory for testing the Lepton Flavour Universality of neutral currents. Since the Flavour Changing Neutral Currents are prohibited at tree level in the Standard Model, these decays must proceed through a loop level diagram, and being suppressed, they offer an increased sensitivity to NP contributions.

The LFU observables being measured in this sector are ratios of the type:

$$R_{H_s} = \frac{\int_{q_{\min}^2}^{q_{\max}^2} \frac{d\Gamma(H_b \to H_s \mu^+ \mu^-)}{dq^2} dq^2}{\int_{q_{\min}^2}^{q_{\max}^2} \frac{d\Gamma(H_b \to H_s e^+ e^-)}{dq^2} dq^2} \tag{45}$$

where $H_b$ is a hadron containing a $b$ quark decaying into a hadron containing an $s$ quark ($H_s$) and $q^2$ is the invariant mass of the $\ell^+\ell^-$ pair, integrated from a minimum, $q^2_{\min}$, and a maximum, $q^2_{\max}$. In this ratio the hadronic uncertainties on the form factors of the decay largely cancel out and the radiative corrections are controlled at a very high precision [21]. These ratios are expected to be very close to unity, given the smallness of the muon and the electron mass [22].

The Belle and BaBar collaborations have performed measurements of the $R_{K^{(*)}}$ ratios, using data samples of about 433 and 605 fb$^{-1}$ of $e^+e^-$ collisions each [23,24]. The analyses have been performed using both neutral and charged $B$ mesons, and the hadronic system used can be a $K^0_S$, $K^\pm$, $K^{*0}$ or a $K^{*\pm}$ meson. At $e^+e^-$ collisions, the center of mass energy is precisely known from the energy of the beams. This is used to identify candidates consistent with a $B \to K^{(*)}\ell^+\ell^-$ decay and to reduce the background contamination. The ratios have been measured by Belle in one region of $q^2$ and two regions of $q^2$ by the BaBar collaboration, excluding regions in which the decay amplitude is dominated by charmonium resonances.

All the measurements have been found to be consistent with the Standard Model expectation, with a precision of $20-50\%$.

The LHCb experiment has also measured the $R_{K^{(*)}}$ ratios for the first time at a hadron collider [25,26]. The harsh hadronic environment introduces significant challenges with respect to analyses performed at an $e^+e^-$ machine. The higher energy of the $B$ meson, and therefore of its daughters, translates in a higher bremsstrahlung rate from the electrons. This is partly coof aboutrrected with a bremsstrahlung recovery procedure which improves the electron momentum reconstruction. Nonetheless, the resolution on the reconstructed $B$ invariant mass is broader than the one obtained at Belle or BaBar, and this poses additional challenges in controlling the backgrounds. Systematic uncertainties due to the different experimental efficiencies in the reconstruction of muons and electrons are reduced by measuring a double ratio, using a resonant charmonium $B \to K^{(*)}(J/\psi \to \ell^+\ell^-)$ mode as a reference:

$$R_{K^*} = \frac{\frac{\mathcal{B}(B \to K^{(*)}\mu^+\mu^-)}{\mathcal{B}(B \to K^{(*)}(J/\psi \to \mu^+\mu^-))}}{\frac{\mathcal{B}(B \to K^{(*)}e^+e^-)}{\mathcal{B}(B \to K^{(*)}(J/\psi \to e^+e^-))}}. \tag{46}$$

The measurements are performed using both the neutral and charged $B$ meson modes, but only using reconstructed charged $K$ and $\pi$ mesons. $R_{K^*}$ has been measured in two $q^2$ bins, whereas the $R_K$ in only one $q^2$ bin. Even with the experimental challenges of a harsh hadron environment, with a dataset of 3 fb$^{-1}$ of data collected during RunI, at a center of mass energy of 7 and 8 TeV, the LHCb experiment achieved a precision in both observables of about half the one observed at Belle and BaBar. Both measurement have been found discrepant with respect to the Standard model, at about $2.6\sigma$ and $2.4-2.5\sigma$ for $R_K$ and $R_{K^*}$, respectively. The analysis of $R_K$ has been recently updated with the full RunI+RunII statistics, corresponding to a total data sample of 9 fb$^{-1}$ of data, collected at a center of mass energy of 7,8 and 13 TeV [27]. This resulted in the first evidence of a violation of Lepton Flavour Universality, with the value of $R_K$ being measured smaller than the Standard Model and discrepant with its expectation value by $3.1\sigma$.

LHCb has also studied Lepton Flavour Universality in $b \to s\ell\ell$ transitions of the $\Lambda_b$ meson, $\Lambda_b \to pK\ell\ell$ decays [28], obtaining a consistent result with respect to the analyses performed with $B$ mesons. Table 3 reports an overview of all the measurements performed in this sector by the various experiments.

| Experiment | $H_s$ | $q^2$ range GeV$^2/c^4$ | Value | Ref. |
|---|---|---|---|---|
| Belle (2009) | $K$ | 0.0 - kin. endpoint | $1.03 \pm 0.19 \pm 0.06$ | [24] |
| Belle (2009) | $K^*$ | 0.0 - kin. endpoint | $0.83 \pm 0.17 \pm 0.08$ | [24] |
| BaBar (2012) | $K$ | $0.10 - 8.12$ | $0.74^{+0.40}_{-0.31} \pm 0.06$ | [23] |
| BaBar (2012) | $K$ | $> 10.11$ | $0.74^{+0.40}_{-0.31} \pm 0.12$ | [23] |
| BaBar (2012) | $K^*$ | $0.10 - 8.12$ | $1.06^{+0.48}_{-0.33} \pm 0.08$ | [23] |
| BaBar (2012) | $K^*$ | $> 10.11$ | $1.18^{+0.55}_{-0.37} \pm 0.11$ | [23] |
| LHCb (2017) | $K^*$ | $0.045 - 1.1$ | $0.66^{+0.11}_{-0.03} \pm 0.05$ | [25] |
| LHCb (2017) | $K^*$ | $1.1 - 6.0$ | $0.69^{+0.11}_{-0.07} \pm 0.05$ | [25] |
| LHCb (2020) | $K$ | $1.0 - 6.0$ | $0.745^{+0.090}_{0.074} \pm 0.036$ | [27] |
| LHCb (2019) | $pK$ | $0.1 - 6.0$ | $0.86^{+0.14}_{0.11} \pm 0.05$ | [28] |

Table 3: Summary of the $R_{K^{(*)}}$ measurements performed at the $B$-factories and by the LHCb experiment. The first uncertainty is statistical and the second is systematic.

### 2.3.2 $b \to c\ell\nu$ transitions

The charged current decays of $b$-hadrons are an ideal laboratory to study the LFU hypothesis. They proceed at tree level in the Standard Model, and therefore they are precisely described theoretically and the visible branching ratios are very high, providing big data samples to be analysed.

Semileptonic $B \to D^{(*)}\mu\nu$ and $B \to D^{(*)}e\nu$ have been since long used by $B$-factory experiments to test the Lepton Flavour Universality hypothesis in the first two lepton families, without finding any discrepancy with respect to the Standard Model expectations. Table 4 reports a list of analyses testing lepton flavour universality through the ratio of branching fractions of decays to the first two generations. For this reason these decays are often assumed to be NP free and are used to perform measurements of $|V_{cb}|$.

| Experiment (year) | $H_c$ type | Ref. |
|---|---|---|
| CLEO (2002) | $D^{*\pm}$ and $D^{*0}$ | [29] |
| BaBar (2008) | $D^{*\pm}$ | [30] |
| BaBar (2009) | $D^0$ and $D^{*0}$ | [31] |
| Belle (2010) | $D^{*\pm}$ | [32] |
| Belle (2016) | $D^0$ and $D^+$ | [33] |
| Belle (2018) | $D^{*\pm}$ | [34] |

Table 4: Tests of LFU in $b \to c\ell\nu$ transitions using the first two generations of leptons.

Nonetheless, due to the large mass of the $\tau$ lepton, semitauonic decays of the type $H_b \to H_c \tau \nu$ could exhibit larger sensitivity to NP effects. The variables used to test lepton flavour universality with $b \to c\ell\nu$ decays are therefore ratios of branching ratios of the type:

$$\mathcal{R}(H_c) = \frac{\mathcal{B}(H_b \to H_c \tau^- \overline{\nu}_\tau)}{\mathcal{B}(H_b \to H_c \ell^- \overline{\nu}_\ell)}, \tag{47}$$

with $\ell$ representing a light lepton. In this ratio, a good fraction of experimental uncertainties due to the difference in selection and reconstruction efficiency of the two leptons, and the value of $|V_{cb}|$ cancel out, providing clean variables that can be known theoretically with a precision of $\approx 1\%$.

Measurements of these ratios can be performed with various decays of the $\tau$ lepton. The various decay modes of the $\tau$ lepton used in the measurements of $\mathcal{R}(H_c)$ have been reported in Table 5.

| Decay | $\mathcal{B}[\%]$ |
|---|---|
| $\tau^- \to \mu^- \overline{\nu}_\mu \nu_\tau$ | $17.39 \pm 0.04$ |
| $\tau^- \to e^- \overline{\nu}_e \nu_\tau$ | $17.82 \pm 0.04$ |
| $\tau^- \to \pi^- \pi^0 \nu_\tau$ | $25.49 \pm 0.09$ |
| $\tau^- \to \pi^- \nu_\tau$ | $10.82 \pm 0.05$ |
| $\tau^- \to \pi^- \pi^+ \pi^- \nu_\tau$ | $9.02 \pm 0.05$ |
| $\tau^- \to \pi^- \pi^+ \pi^- \pi^0 \nu_\tau$ | $4.49 \pm 0.05$ |

Table 5: Branching fractions of $\tau$ decays that have been used to perform measurements in $b \to c\ell\nu$ decays. The 3-prong hadronic modes do not include the $K^0$ contribution.

Each analysis using different decay modes has various advantages and disadvantages:

- Analyses using leptonic $\tau^- \to \mu^- \nu_\tau \overline{\nu}_\mu$ and $\tau^- \to e^- \nu_\tau \overline{\nu}_e$ aim at reconstructing one charmed meson and a light lepton. The final state will contain both signal $H_b \to H_c \tau \nu$ decays and normalization $H_b \to H_c \ell \nu$ decays, allowing an extraction of $\mathcal{R}(H_c)$ from a single fit to the data sample. Being the reconstruction of the two modes identical, with only kinematic differences, the systematic uncertainties due to the different efficiencies largely cancel out, but the analysis has to face a large background coming from partially reconstructed semileptonic decays of the $H_b$ hadrons.

- Hadronic $\tau$ final states contain only one neutrino in the final state with respect to the leptonic case. Furthermore the relatively long lifetime of the $\tau$ lepton enables the reconstruction of the $\tau$ decay vertex, if the final state includes more than one charged final state pion. For these reasons, these decays are usually more constrained than the previous ones, allowing for a better resolution on the $B$ momentum and a better partially reconstructed background rejection. The signal and normalization decays are not reconstructed in the same final state. For this reason the measurements are usually normalized to a fully hadronic decay with a similar topology to the signal and then external constraints are used to extract the

value of $\mathcal{R}(H_c)$. This in turn translates in higher systematic uncertainties due to external constraints with respect to the analyses exploiting leptonic $\tau$ decays.

$B-$factories exploit the fact that the $\Upsilon(4S)$ decays exclusively to $B^0\overline{B}^0$ and $B^+B^-$ pairs, by reconstructing one of the two $B$ hadrons using hadronic or semileptonic decays. Althoughthe decay of a semileptonic tag of the event guarantees a big data sample, thanks to the big branching fraction of these decays, an hadronic tag enables a good reconstruction of the momentum of the $B$-signal candidate momentum, obtaining a good resolution on the invariant mass of the undetected particles in the signal decay, $m_{\mathrm{miss}}^2 = (p_{e^+e^-} - p_{\mathrm{tag}} - p_{H_c} - p_\ell)^2$, where $p_{e^+e^-}, p_{\mathrm{tag}}, p_{D^*}$ and $p_\ell$ are the four momenta of the colliding electrons, the $B$-tag candidate, the $H_c$ meson and the lepton, and on the energy of the lepton in the $B$ rest frame, $E_\ell^*$. The $m_{\mathrm{miss}}^2$ variable is especially good at separating the signal from the normalization channel in analyses with leptonic decays of the $\tau$, peaking at 0 in the presence of only one neutrino missing in the final state and having a broader distribution at higher values for decays when additional final state particles go unreconstructed.

The first simultaneous measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ was performed by the BaBar collaboration, using a leptonic decay of the $\tau$ and a hadronic tag [35]. This measurement has been performed by fitting the dataset in $m_{miss}^2$ and $E_\ell^*$. The analysis was performed using decays of both charged and neutral $B$ mesons assuming isospin conservation. The collaboration reported an excess of $2.0\sigma$ and $2.7\sigma$ in $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, which sparked a lot of interest in the community. Since then, several measurements have been performed also by other experiments.

The Belle collaboration has reported two simultaneous measurements of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ [36, 37], using hadronic and semileptonic tag of the events, both in the leptonic final state of the $\tau$, and two independent measurements of $\mathcal{R}(D^*)$, one with a semileptonic tag and a leptonic $\tau$ final state [38], and one with a hadronic tag and a hadronic decay of the $\tau$ [39].

Also LHCb has actively contributed to this endeavour by measuring $\mathcal{R}(D^*)$ with both a leptonic $\tau$ final state and a hadronic $\tau$ final state. At a $pp$ collider, the $b\bar{b}$ pairs are produced with a broad energy spectrum, and therefore the momentum of the decaying $B$ particle cannot be analytically reconstructed. However, the good vertexing resolution of the detector, united with the long decay time of the $B$ meson can be used at LHCb to approximate this quantity, as it will be explained later on in the text.

In the measurement of $\mathcal{R}(D^*)$ with the leptonic decay of the $\tau$ [40], the parameter of interest has been extracted with a fit to three variables evaluated in this approximated rest frame: the muon energy $E_\mu^*$, the missing mass squared $m_{\mathrm{miss}}^2$ and the momentum transfer to the lepton system, $q^2 = (p_B - p_{D^*})^2$. A Multivariate isolation tool has been used in order to suppress partially reconstructed backgrounds.

For the $\mathcal{R}(D^*)$ measurement with a hadronic decay of the $\tau$, a three prong $\tau^- \to \pi^-\pi^+\pi^-(\pi^0)$ decay has been used [41]. Thanks to the flight distance of the $\tau$ lepton and the excellent vertex reconstruction of the detector, the LHCb experiment has been able to reconstruct the $\tau$ decay vertex. Events in which the significance of the $\tau$ flight distance does not exceed a given threshold have been rejected, which enabled a big suppression of the $B \to D^* 3\pi X$ background. To minimize the experimental systematic uncertainties the signal is normalized to a fully hadronic, $B \to D^* 3\pi$ decay, and external inputs have been used to extract $\mathcal{R}(D^*)$. The extraction of the signal yield has been performed by fitting $q^2$, the $\tau$ decay time and the output of a BDT separating the signal from the $B \to D^* D(X)$ background.

Thanks to the large $B_c^+$ production at the LHC, the LHCb experiment has also measured

$\mathcal{R}(J/\psi)$ [42], using leptonic $\tau$ decays and a $J/\psi$ reconstructed in a $\mu^+\mu^-$ final state. A strategy for the reconstruction of the $B$ momentum similar to the one of the $\mathcal{R}(D^*)$ measurement with leptonic $\tau$ decays has been used. The parameter of interest has been extracted in a four dimensional fit to $m^2_{\text{miss}}$, the $B_c$ decay time, $q^2$ and $E^*_\mu$. The measured value of $\mathcal{R}(J/\psi)$ has shown an excess with respect to the Standard Model predictions of around $2\sigma$.

Table 6 reports an overview of all the measurements performed in this sector by the various experiments.

By combining the available measurements of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, the average shows an excess with respect to the Standard Model predictions of about $3.1\sigma$ [1]. An overview of these measurements and their combination is reported in Fig. 3. It can be noticed that the parameter $\mathcal{R}(D)$, using ground states $D^+$ or $D^0$ mesons has never been measured in $pp$ collisions. No vertical constraint in the vertical direction of the $\mathcal{R}(D^*)$ vs $\mathcal{R}(D)$ plane is therefore present from hadron colliders. This is an essential piece of information that is missing and could be able to resolve or confirm the long-standing discrepancy being present in these channels. A simultaneous measurement of $\mathcal{R}(D^+)$ and $\mathcal{R}(D^{*+})$ is the subject of this thesis.
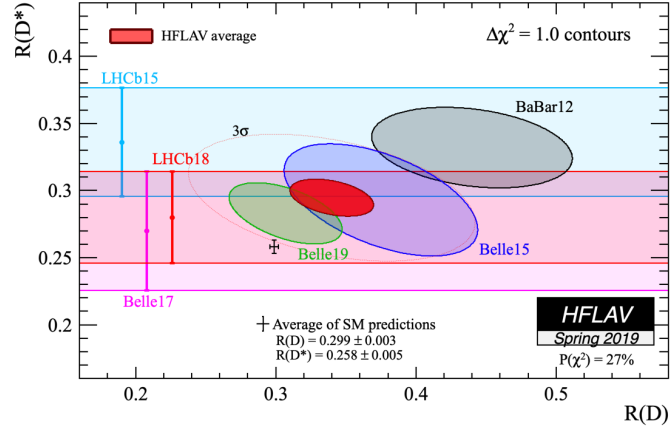


Figure 3: Experimental results on $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ and comparison with the SM prediction [1].

| Experiment (year) | Observable | $\tau$ decay type | Tag type | Ref. |
|---|---|---|---|---|
| BaBar (2012) | $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ | leptonic | hadronic | [35] |
| Belle (2015) | $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ | leptonic | hadronic | [36] |
| Belle (2019) | $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ | leptonic | semileptonic | [37] |
| Belle (2016) | $\mathcal{R}(D^*)$ | leptonic | semileptonic | [38] |
| Belle (2016) | $\mathcal{R}(D^*)$ | hadronic | hadronic | [39] |
| LHCb (2015) | $\mathcal{R}(D^*)$ | leptonic | - | [40] |
| LHCb (2017) | $\mathcal{R}(D^*)$ | hadronic | - | [41] |
| LHCb (2018) | $\mathcal{R}(J/\psi)$ | leptonic | - | [42] |

Table 6: Summary of the Lepton Flavour Universality measurements performed in the $b \to c\ell\nu$ sector.

# 3 Phenomenology of semileptonic decays in the Standard Model

Semileptonic decays of the type $H_b \to H_c \ell \nu$ such as the one under study in this thesis, even though difficult to study experimentally due to the presence of neutrinos in the final stats, are theoretically clean thanks to the fact that leptonic system, being colour singlet, is not affected by QCD effects, as illustrated in Fig. 4, and therefore the leptonic current involved in the decay can be studied independently of the hadronic current.

In this chapter the phenomenology of these decays, with particular focus on the $B \to D^{(*)}\ell\nu$ decays, will be briefly discussed. The general structure of the decay will be studied. Then the concept of Form Factor will be introduced and some parameterization used in this thesis will be introduced.
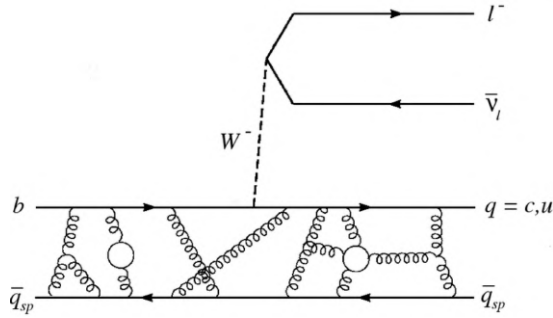


Figure 4: Semileptonic decay of a meson containing a $b$ quark and a "spectator" $u$ or $d$ quark, denoted as $q_{sp}$, [43].

## 3.1 Differential decay rate

$H_b \to H_c \ell \nu$ decays proceed via a $b \to c\ell\nu$ transition at the parton level, which in the SM is mediated by the exchange of a $W^{\pm}$ boson. The SM operator mediating this transition is therefore given by

$$\mathcal{O}_{\text{SM}} \propto G_F V_{cb} \left( \bar{c}\gamma^{\mu} \frac{1-\gamma_5}{2} b \right) \left( \bar{\ell}\gamma_{\mu} \frac{1-\gamma_5}{2} \nu_{\ell} \right). \tag{48}$$

Being the lepton system not affected by the strong interactions, the matrix element can be factorized in a hadronic part and a leptonic part. Using the notation of [44], in fact, the SM matrix element for this transition can be written in the following way:

$$M_{\lambda_X}^{\lambda_\ell}(q^2, x) = \frac{G_F}{\sqrt{2}} V_{cb} \frac{M_W^2}{M_W^2 - q^2} \sum_{\lambda_W} \eta_{\lambda_W} L_{\lambda_W}^{\lambda_\ell} H_{\lambda_W}^{\lambda_X}, \tag{49}$$

where $\lambda_W$ is the virtual $W$ helicity ($\lambda_W = \pm, 0, s$), $\lambda_X$ is the helicity of the $H_c$ system, $\lambda_\ell = \pm\frac{1}{2}$ is the spin of the lepton and $\eta_{\lambda_W}$ is a metric factor defined as $\eta_{\pm}, 0 = 1$ and $\eta_s = (q^2 - M_W^2)/M_W^2$.

$$L_{\lambda_W}^{\lambda_\tau} = \epsilon_\mu(\lambda_W)\langle \ell\bar{\nu}_\ell | \bar{\ell}\gamma^\mu(1 - \gamma^5)\nu_\ell | 0 \rangle, \tag{50}$$

$$H_{\lambda_W}^{\lambda_M} = \epsilon_\mu^*(\lambda_W)\langle D^{(*)} | \bar{c}\gamma^\mu(1 - \gamma^5)b | \overline{B} \rangle, \tag{51}$$

19

where $\epsilon_\mu(\lambda_W)$ is the polarization vector of the virtual W boson.

The leptonic matrix elements can be evaluated using the Dirac algebra of the Electroweak theory and using an explicit choice for the helicity vectors of the $W$ mediator. In the rest frame of the $H_b$ mother, with the $z$ axis aligned along the trajectory of the $H_c$ daughter, a suitable basis for the $W$ helicities is

$$\varepsilon_\mu(\pm) = \frac{1}{\sqrt{2}}(0, \pm 1, -i, 0) \tag{52}$$

$$\varepsilon_\mu(0) = \frac{1}{\sqrt{q^2}}(|\vec{p}_{H_c}|, 0, 0 - q_0) \tag{53}$$

$$\varepsilon_\mu(s) = \frac{1}{\sqrt{q^2}}(q_0, 0, 0, -|\vec{p}_{H_c}|), \tag{54}$$

where $q^2 = q_\mu q^\mu$ is the momentum transferred to the leptonic system, $q_\mu = p_{H_b} - p_{H_c}$. They satisfy the following normalization and completeness relation:

$$\varepsilon_\mu^*(m)\varepsilon^\mu(m') = g_{mm'}, \tag{55}$$

$$\sum_{m,m'} \varepsilon_\mu(m)\varepsilon_\nu^*(m')g_{mm'} = g_{\mu\nu}. \tag{56}$$

With this choice for the helicity basis, the leptonic amplitudes are given by [44]:

$$L_0^- = -2\sqrt{q^2}\sqrt{(1 - m_\ell^2/q^2)}\sin\theta_\ell, \tag{57}$$

$$L_0^+ = 2m_\ell\sqrt{(1 - m_\ell^2/q^2)}\cos\theta_\ell, \tag{58}$$

$$L_s^- = 0, \tag{59}$$

$$L_s^+ = -2m_\ell\sqrt{(1 - m_\ell^2/q^2)}, \tag{60}$$

where $\theta_\ell$ is the helicity angle of $\ell$ in the $W$ rest frame. Notice that all leptonic amplitudes other than $L_0^-$ disappear for massless leptons.

In $B \to D^{(*)}\ell\nu$ decays, the initial state has spin-0, and therefore the final state must be spin-0 due to angular momentum conservation. For a spin-0 daughter $H_c$, whose only helicity state is 0, only $H_0^0$ and $H_s^0$ are non-zero. For a spin-1 daughter, with additional $\pm 1$ helicity states $H_+^+$ and $H_-^-$ are also non-zero. A spin-2 daughter has $\pm 2$ helicity states, but these are not accessible from any $W$ helicity state. So, without any ambiguity, the non-zero hadronic tansor helicity components can be denoted as $H_\pm, H_0$ and $H_s$, where the last two are non-zero only for scalar daughter particles.

Therefore, the differential decay rate can be entirely expressed in terms of $\theta_\ell$ and the momentum transfer $q^2$:

$$\frac{d\Gamma}{dq^2 d\cos\theta} = \frac{\eta_{ew}^2 G_F^2 |V_{cb}|^2 |\vec{p}_{H_c}|}{256\pi^3 M^2}\left(1 - \frac{m_\ell^2}{q^2}\right)^2 \tag{61}$$

$$\times [(1 - \cos\theta)^2 H_+^2 + (1 + cos\theta)^2 H_-^2 + 2\sin\theta H_0^2 + \tag{62}$$

$$+ \frac{m_\ell^2}{q^2}(\sin^2\theta(H_+^2 + H_-^2) + 2(H_s - H_0\cos\theta)^2)], \tag{63}$$

20

where $|\vec{p}_{H_c}|$ is the momentum of the daughter $H_c$ in the parent $B$ rest frame. By applying a simple energy momentum conservation, this quantity can be written as a function of $q^2$:

$$|\vec{p}_{H_c}| = \sqrt{\left(\frac{m_{H_b}^2 + m_{H_c}^2 - q^2}{2m_{H_b}}\right)^2 - m_{H_c}^2} = m_{H_c}\sqrt{w^2 - 1} \tag{64}$$

where $w$ is the scalar product of the $4-$velocities of the $B$ and $H_c$ particles, a quantity often used as expansion parameter in Heavy Quark Effective Field Theories:

$$w = v_B \cdot v_{H_c} = \frac{m_B^2 + m_{H_c}^2 - q^2}{2m_B m_{H_c}}. \tag{65}$$

The factor $\eta_{ew} = 1 + \alpha/\pi \ln M_z/m_B \approx 1.0066$ [45] takes into account the short distance QED corrections due to the running of the four-Fermi operator from the weak to the $B$ scale and represents the leading electroweak correction.

Integrating the doubly differential decay rate over $\cos\theta$ gives the differential decay rate as a function of $q^2$:

$$\frac{d\Gamma}{dq^2} = \frac{\eta_{ew}^2 G_F^2 |V_{cb}|^2 |\vec{p}_{H_c}| q^2}{96\pi^3 M^2} \left(1 - \frac{m_l^2}{q^2}\right)^2 \left[\left(1 + \frac{m_l^2}{2q^2}\right)[H_+^2 + H_-^2 + H_0^2] + \frac{3m_\ell^2}{2q^2}H_s^2\right] \tag{66}$$

## 3.2  Form Factors

The hadronic matrix elements are difficult to be evaluated due to the intrinsic non perturbative nature of the transition.

Thanks to Lorentz invariance, these matrix elements can be written in terms of combinations of the Lorentz covariant quantities in the decay. The possible combinations are governed by the Lorentz nature of the matrix element. The Lorentz covariant quantities must be written as combinations of Lorentz vectors available in the process, namely the momenta of the particles $(p_B^\mu, p_{H_c}^\mu)$ and the spin vector of the $H_c$ state $(\varepsilon^\mu)$. These quantities will enter the expression of the matrix element multiplied by functions that depend only on invariants of the system, like $q^2$ or the analogous $w$. These functions are usually called Form Factors, and are the quantities which are measured experimentally.

Various theoretical frameworks can be employed to be able to evaluate the Form Factors or to understand their properties. Two of the most famous parametrizations, that arise from different theoretical assumptions imposed on the matrix elements, will be described later on.

### 3.2.1  $B \to D\ell\nu$ form factors

Given that the $D$ is a scalar state, no pseudovector is available for describing the hadronic matrix element. Therefore, the only operators that will contribute to this process will be scalar, vectors and tensors, out of which only the vector operator is implemented in SM. The Form Factors are defined as follows:

$$\langle D|\bar{c}b|\overline{B}\rangle = f_S(q^2), \tag{67}$$

$$\langle D|\bar{c}\gamma_\mu b|\overline{B}\rangle = f_+(q^2)(p_B + p_D)\mu + f_-(q^2)q^\mu, \tag{68}$$

$$= f_+(q^2)(p_B + p_D)^\mu + \left[f_0(q^2) - f_+(q^2)\right]\frac{m_B^2 - m_D^2}{q^2}q^\mu, \tag{69}$$

$$\langle D|\bar{c}\sigma^{\mu\nu}b|\overline{B}\rangle = if_T(q^2)\left[(p_B + p_D)^\mu q^\nu - (p_B + p_D)^\nu q^\mu\right], \tag{70}$$

where $f_0(q^2) = f_+(q^2) + \frac{q^2}{m_B^2 - m_D^2}f_-(q^2)$, from which follows that $f_0(0) = f_+(0)$. This Form Factor basis, $\{f_+, f_0, f_S\}$, is denoted as the *spectroscopic basis*.

Another common definition for the $B \to D\ell\nu$ form factors, usually employed in parametrizations using Heavy Quark Symmetry, is the following:

$$\langle D|\bar{c}b|\overline{B}\rangle = \sqrt{m_B m_D}h_S(w+1), \tag{71}$$

$$\langle D|\bar{c}\gamma_\mu b|\overline{B}\rangle = \sqrt{m_B m_D}\left[h_+(v + v')_\mu + h_-(v - v')_\mu\right], \tag{72}$$

$$\langle D|\bar{c}\sigma^{\mu\nu}b|\overline{B}\rangle = i\sqrt{m_B m_D}\left[h_T(v'_\mu v_\nu - v_\mu v'_\nu)\right]. \tag{73}$$

This Form Factor basis, $\{h_+, h_-, h_S\}$ is called the *Heavy Quark Symmetry (HQS) basis*.

### 3.2.2 $B \to D^*\ell\nu$ form factors

Given that the $D^*$ is a spin-1 particle, its polarization pseudovector $\varepsilon_\mu^*$ is available for writing the expression of the matrix elements. Therefore, also the pseudo-scalar and pseudo-vector operators will contribute to the process. Given that all the matrix elements should depend on $\varepsilon_\mu^*$, the scalar operator will not contribute to the process. The only matrix elements implemented in the SM are the vector and the pseudovector.

The Form Factors are defined as follows:

$$\langle D^*|\bar{c}\gamma^5 b|B\rangle = a_0\ \varepsilon^* \cdot p_B \tag{74}$$

$$\langle D^*|\bar{c}\gamma^\mu b|B\rangle = -ig\ \varepsilon^{\mu\nu\rho\sigma}\varepsilon_\nu^*(p_B + p_{D^*})_\rho q_\sigma \tag{75}$$

$$\langle D^*|\bar{c}\gamma^\mu\gamma^5 b|B\rangle = \varepsilon^{*\mu}f + a_+\varepsilon^* \cdot p_B(p_B + p_{D^*})^\mu + a_-\varepsilon^* \cdot p_B q^\mu \tag{76}$$

$$\langle D^*|\bar{c}\sigma^{\mu\nu}b|B\rangle = -a_{T_+}\varepsilon^{\mu\nu\rho\sigma}\varepsilon_\rho^*(p_B + p_{D^*})_\sigma - a_{T_-}\varepsilon^{\mu\nu\rho\sigma}\varepsilon_\rho^* q_\sigma + \tag{77}$$

$$- a_{T_0}\varepsilon^* \cdot_B \varepsilon^{\mu\nu\rho\sigma}(p_B + p_{D^*})_\rho q_\sigma. \tag{78}$$

This basis, $\{a_0, g, f, a_+, a_-, a_{T_+}, a_{T_-}, a_{T_0}\}$ is called the *spectroscopic basis*. Sometimes in the literature the differential decay rate is reported as a function of two form factors, denoted as $\mathcal{F}_1$ and $\mathcal{F}_2$, which are defined as

$$\mathcal{F}_1 = \frac{1}{m_{D^*}}\left[2p_{D^*}^2 q^2 a_+ - \frac{1}{2}(q^2 - m_B^2 + m_{H_c}^2)f\right], \tag{79}$$

$$\mathcal{F}_2 = \frac{1}{m_{D^*}}\left[f + (m_B^2 - m_{D^*}^2)_+ + q^2 a_-\right]. \tag{80}$$

The *HQS basis*, $\{h_P, h_V, h_{A_1}, h_{A_2}, h_{A_3}, h_{T_1}, h_{T_2}, h_{T_3}\}$, in this case, is defined in this way:

22

$$\langle D^*|\bar{c}\gamma^5 b|\overline{B}\rangle = -\sqrt{m_B m_{D^*}} h_P(\varepsilon \cdot v), \tag{81}$$

$$\langle D^*|\bar{c}\gamma^\mu b|\overline{B}\rangle = i\sqrt{m_B m_{D^*}} h_V \varepsilon^{\mu\nu\alpha\beta} \varepsilon_\nu^* v_\alpha' v_\beta, \tag{82}$$

$$\langle D^*|\bar{c}\gamma^\mu \gamma^5 b|\overline{B}\rangle = \sqrt{m_B m_{D^*}} \left[ h_{A1}(w+1)\varepsilon^{*\mu} - h_{A_2}(\varepsilon \cdot v) - h_{A_3}(\varepsilon \cdot v') \right], \tag{83}$$

$$\langle D^*|\bar{c}\sigma^{\mu\nu} b|\overline{B}\rangle = -\sqrt{m_B m_{D^*}} \varepsilon^{\mu\nu\alpha\beta} \left[ h_{T_1}\varepsilon_\alpha^*(v+v')_\beta + h_{T_2}\varepsilon_\alpha^*(v-v')_\beta + h_{T_3}(\varepsilon^* \cdot v)v_\alpha v_\beta' \right] \tag{84}$$

When in the literature the HQS basis is employed, usually the decay rate is expressed in terms of $h_{A_1}$ and the three form factor ratios

$$R_1(w) = \frac{h_V}{h_{A_1}}, \tag{85}$$

$$R_2(w) = \frac{h_{A_3} + r^* h_{A_2}}{h_{A_1}}, \tag{86}$$

$$R_0(w) = \frac{(w+1)h_{A_1} - (w-r^*)h_{A_3} - (1-wr^*)h_{A_2}}{(1+r^*)h_{A_1}}, \tag{87}$$

where $r^* = m_{D^*}/m_B$ and $R_0$ enters the decay amplitude always accompanied by a $m_\ell$ suppression factor.

### 3.2.3  BGL parametrization

In order to study the properties of the form factors, one can impose a different variety of theoretical constraints. The BGL parametrization [45] tries to be as model independent as possible and arises from imposing a very small number of constraints which come only from:

- Crossing symmetry;

- Unitarity;

- Analytical properties of the matrix elements, thought as functions of the complex plane;

- Dispersion relations of the two point functions of the theory;

Crossing symmetry is related to the observation that the matrix element for a $H_b \to H_c$ process, $\langle H_c|J|H_b\rangle$, mediated by a current $J = \bar{c}\Gamma b$, may be described by the analytical continuation of the matrix element of another related process which describes another phenomenon in which some of the initial state particles are swapped with final state particles, or vice versa. Semileptonic decays of the type $H_b \to H_c \ell \nu$ have a $q^2$ range that goes from $m_\ell^2$ to $t_- = (m_{H_b} - m_{H_c})^2$. Crossing symmetry, united with analytical properties of the matrix elements, says that the analytical continuation of the Form Factors for this process beyond the physical region, $q^2 > t_-$ also describe the Form Factors for the $H_b \overline{H}_c$ pair production in the region in which $q^2 > (m_{H_b} + m_{H_c})^2 = t_+$.

One of the consequences of the unitarity of the theory and of analiticity of the matrix elements is the presence of dispersion relations for two-point functions. These relations are important for Form Factors parametrizations because they lead to a tower of bounds on Form Factors that can be used to constraint the Form Factors.

The existence of dispersion relations can be shown in full generality for two point functions in the momentum space of the type

$$\Gamma(p^2) = \int d^4x e^{ip\cdot x}\langle 0|T\{J(x)J^\dagger(0)\}|0\rangle, \tag{88}$$

where $T$ stands for the time ordered product. This quantity satisfies the following dispersion relation

$$\Gamma(p^2) = i\int_0^\infty \frac{\rho(q^2)}{q^2 - p^2 + i\epsilon}dq^2, \tag{89}$$

where $\rho(q^2) = \frac{1}{\pi}\text{Im}(\Gamma(q^2))$ is a spectral function containing information about the physical content of the theory. In a general interacting theory $\rho(q^2)$ is a function with delta functions around bound states and a continuous spectrum above threshold for the production of multiparticle states, as can be seen in Fig. 5.
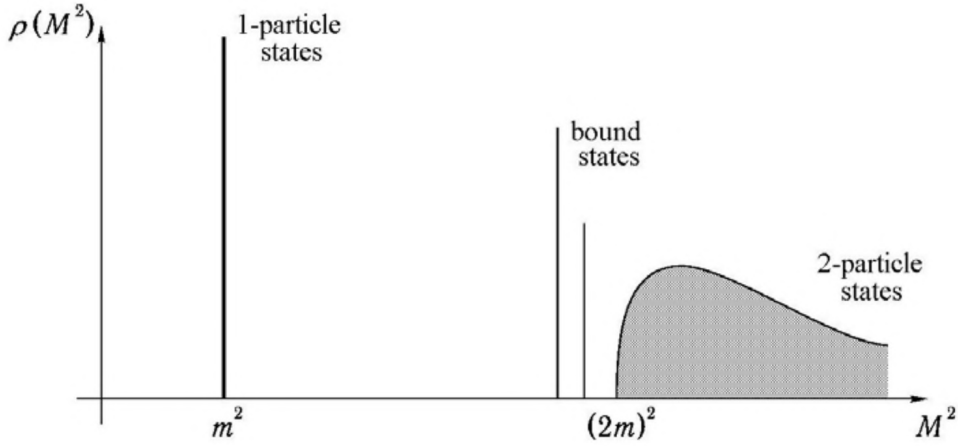


Figure 5: The spectral function for a typical interacting field theory, [46].

Therefore the two point functions will have poles around bound states of the theory and cut branches on the real line, above the threshold for production of multiparticle states. Of interest for $H_b \to H_c$ transition form factors is the two point function:

$$\Gamma_J^{\mu\nu}(q^2) = i\int d^4x\ e^{iqx}\langle 0|T\{J^\mu J^{\dagger\nu}(0)\}|0\rangle = (q^\mu q^\nu - q^2 g^{\mu\nu})\Gamma_J^T(q^2) + g^{\mu\nu}\Gamma_J^L(q^2), \tag{90}$$

where $J$ can be a V, A, or V-A current.

This correlation function and its derivatives satisfy dispersion relations and, in particular, the one used in the BGL expression reads

$$\chi_J^{T,L} = \frac{\partial \Gamma_J^{T,L}}{\partial q^2} = \frac{1}{\pi^2}\int_0^\infty \frac{\text{Im}\Gamma_J^{T,L}(t)}{(t - q^2)^2}. \tag{91}$$

The left-hand side of the equation is evaluated at one loop in perturbative QCD in the region of $q^2 > t_+$ and then, using crossing symmetry, continued into the semileptonic $q^2 < t_-$ region.

The integral on the right-hand side can be written as the sum over phase-space weighted matrix elements for the production of on shell hadronic states

$$\mathrm{Im}\Gamma_J^{T,L}(q^2) = \sum_X \int d\mathcal{PS}_x (2\pi)^4 \delta^4(p - p_X) |\langle 0|J(0)|X\rangle|^2. \tag{92}$$

In this case, the states $X$ represents a complete set of $H_b\overline{H}_c$ hadron pairs, with the appropriate spin parity, but the positivity of each of the summands enables to write a lower bound on the sum by using only a single term with $X = H_b\overline{H}_c$. In this way one can derive lower bounds on the form factors describing the pair production matrix element $\langle 0|J|H_b\overline{H}_c\rangle \sim F_J(q^2)$. Using crossing symmetry, this form factor can be thought as the analytical continuation, in the pair production region, of the form factor describing the semileptonic transition mediated by the matrix element $\langle H_c|J|H_b\rangle$.

The lower bound can, therefore, be expressed in the form

$$\int_{|z|=1} \frac{dz}{2\pi i} \sum_i |P_i^J(z)\phi_i^J F_i^J(z)|^2 \leq 1, \tag{93}$$

where $z$ is a conformal mapping, defined as

$$z(q^2, q_0^2) = \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - q_0^2}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - q_0^2}}. \tag{94}$$

This transformation maps the $|q^2| < t_+$ ($|q^2| > t_-$) to the interior (exterior) of the unit circle $|z| = 1$, centered around $q^2 = q_0^2$. A common choice is $q_0^2 = t_-$, for which $z(w = 0) = 0$. This allows the semileptonic form factors to be written as analytical functions in the interior of the unit $|z| = 1$ circle, apart from sub-threshold poles, corresponding to $B_c^{(*)}$ resonances, located on the real axis, and the pair production branch cut, being mapped onto the unit circle.

$\phi_i^J$ are complex functions, called *outer functions*, that do not have any singularity and encode the $q^2$ dependency of the matrix element and the $1/\sqrt{\pi\chi_J}$ prefactor. The $P_i^J$ functions, called *inner* or *Blaschke functions* have zeroes at the positions of the poles and unit modulus along the physical cut placed at $|z| = 1$. In this way one can cancel the poles, for which the residual is not known, and obtain an analytical function in the inner $|z| \leq 1$ disc. The explicit form of the Blaschke functions is

$$B_j(z) = \prod_n \frac{z - z_n^{(j)}}{1 - zz_n^j} \tag{95}$$

where the index $n$ runs over all the $B_c$ bound states that couple to a given form factor and

$$z_n^j = \frac{\sqrt{(m_{B^{(*)}} + m_{D^{(*)}})^2 - M_n^2} - 2a\sqrt{m_{B^{(*)}} m_{D^{(*)}}}}{\sqrt{(m_{B^{(*)}} + m_{D^{(*)}})^2 - M_n^2} + 2a\sqrt{m_{B^{(*)}} m_{D^{(*)}}}} \tag{96}$$

is the image of the pole mass $M_n$ in the $z$ plane. The functions $\chi_J$ are, instead, known from perturbative QCD calculations.

25

Due to the bounds above, each term in the sum must be analytic, and therefore the functions $f_i(z) = P_i^J(z)\phi_i^J(z)F_i^J(z)$ can be expanded in Taylor series, which gives the BGL expression for the Form Factors

$$F_i(z) = \frac{1}{P_i(z)\phi_i(z)}\sum_{n=0}^{\infty}a_n^i z^i. \tag{97}$$

The bounds translate also on bounds on sums of the coefficients:

$$\sum_{n=0}^{\infty}|a_n|^2 \leq 1. \tag{98}$$

For $B \to D\ell\nu$, the SM form factors have the following BGL parametrization

$$f_+(z) = \frac{1}{P_{f_+}(z)\phi_{f_+}(z)}\sum_{n=0}^{\infty}a_{+n}z^n, \tag{99}$$

$$f_0(z) = \frac{1}{P_{f_0}(z)\phi_{f_0}(z)}\sum_{n=0}^{\infty}a_{0n}z^n. \tag{100}$$

For $B \to D^*\ell\nu$, the SM form factors have the following BGL parametrization

$$g(z) = \frac{1}{P_g(z)\phi_g(z)}\sum_{n=0}^{\infty}a_n z^n, \tag{101}$$

$$f(z) = \frac{1}{P_f(z)\phi_f(z)}\sum_{n=0}^{\infty}b_n z^n, \tag{102}$$

$$\mathcal{F}_1(z) = \frac{1}{P_{\mathcal{F}_1}(z)\phi_{\mathcal{F}_1}(z)}\sum_{n=0}^{\infty}c_n z^n, \tag{103}$$

$$\mathcal{F}_2(z) = \frac{1}{P_{\mathcal{F}_2}(z)\phi_{\mathcal{F}_2}(z)}\sum_{n=0}^{\infty}d_n z^n. \tag{104}$$

### 3.2.4 CLN parametrization

The parametrization described in the previous section is a very general one, with a minimal set of assumptions on the hadron modelling. Additional theoretical input might be combined with the dispersive bound approach in order to generate SM predictions for lepton flavour universality observables.

One input may come from Heavy Quark Effective Theory (HQET). A heavy hadron is a bound state containing a heavy quark, $b, c$. In this theory the heavy quark, thought as the only source of gluons in the system, is dressed by a light quark wave function, called the *brown muck*, that describes the degrees of freedom of the light quarks. HQET can be thought as an effective theory for the interactions of the brown muck with the heavy quark, which are studied in a perturbative way with expansion parameter $1/m_Q$, with $m_Q$ the mass of the heavy quark. In this

context the key result is that any $B \to D^{(*)}$ matrix element is predicted to be proportional to a universal, spin and flavour independent, scalar function, called the *leading Isgur-Wise function*

$$\langle D^{(*)}|\bar{c}\Gamma b|\overline{B}\rangle \propto \sqrt{m_{D^{(*)}} m_B}\xi(w) + \mathcal{O}\left(\frac{\Lambda_{QCD}}{2m_c}, \frac{\Lambda_{QCD}}{2m_b}\right). \tag{105}$$

Higher order terms in $\frac{\Lambda_{QCD}}{2m_{c,b}}$ can be incorporated in terms of universal *subleading Isgur-Wise functions*. In the CLN parametrization [47] the $B \to D$ from factor called $V_1$ is expanded up to a cubic order around the zero recoil point $w \approx 1$ as

$$V_1(w) = V_1(1)\left[1 - \rho_1^2(w-1) + c_1(w-1)^2 + d_1(w-1)^3 + \dots\right], \tag{106}$$

and then dispersion relation bounds are applied to this expression. Approximate relations between the $\rho_1^2$ and the $c_1$ and $d_1$ parameters are found by saturating these bounds at the $1\sigma$ uncertainty of $\chi_J$ QCD correlators. This form factor is then used to get the functional shape of all the other form factors, by using HQET symmetry. The subleading terms are included by incorporating subleading IsgurWise functions, expanded at the second order in $(w-1)$

$$R_i(w) = R_i(1) + R_i'(1)(w-1) + R_i''(1)(w-1) + \dots. \tag{107}$$

In the HQET expansion, the first and second derivative of $R$ are correlated with the intercept $R_i(1)$ at each order.

For the $B \to D$ transition the form factors in the CLN parametrization take the following form:

$$V_1(w) = V_1(1)[1 - 8\rho^2 z + (51\rho^2 - 10)z^2 - (252\rho^2 - 84)z^3] \tag{108}$$

$$S_1(w) = [1 + \Delta(w)]V_1(w) \tag{109}$$

$$\Delta(w) = -0.019 + 0.041(w-1) - 0.015(w-1)^2. \tag{110}$$

where $\Delta(w)$ denotes the QCD $1/m_Q$ corrections. If written in terms of $f_+(w)$ and $f_0(w)$, one obtains the following expressions:

$$f_+(w) = f_+(1)[1 - 8\rho^2 z + (51\rho^2 - 10)z^2 - (252\rho^2 - 84)z^3] \tag{111}$$

$$f_0(w) = \frac{2r(w+1)}{(r+1)^2}[1 + \Delta(w)]f_+(w). \tag{112}$$

For the $B \to D^*$ transition they take the following form [48], in terms of the universal form factor $h_{A1}$:

$$A_1(q^2) = R_{D^*}\frac{w+1}{2}h_{A1} \tag{113}$$

$$A_0(q^2) = \frac{R_0(w)}{R_{D^*}}h_{A_1(w)} \tag{114}$$

$$A_2(q^2) = \frac{R_1(w)}{R_{D^*}}h_{A_1(w)} \tag{115}$$

$$V(q^2) = \frac{R_1(w)}{R_{D^*}}h_{A_1(w)} \tag{116}$$

where $R_{D^*} = \frac{2\sqrt{m_B m_{D^*}}}{(m_B + m_D^*)}$, and:

$$h_{A_1}(w) = h_{A_1}(1)[1 - 8\rho^2 z + (53\rho^2 - 15)z^2 - (231\rho^2 - 91)z^3], \tag{117}$$

$$R_1(w) = R_1(1) - 0.12(w - 1) + 0.05(w - 1)^2, \tag{118}$$

$$R_2(w) = R_2(1) + 0.11(w - 1) - 0.06(w - 1)^2, \tag{119}$$

$$R_0(w) = R_0(1) - 0.11(w - 1) + 0.01(w - 1)^2 \tag{120}$$

# 4  Standard Model prediction of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$

As it can be read in the previous section, the decay rates for the semileptonic decays involving $\tau$ or light leptons depend on terms that are proportional to $m_\tau$, and therefore the SM predictions for $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ depend on ratios of form factors. For this reason, the ratios are not expected to be equal to 1, and the precision to which they can be known depends on the precision to which the form factors are known.

The predicted value of $\mathcal{R}(D)$ is $0.299 \pm 0.003$, which relies on two independent calculations of the form factors by the MILC and HPQCD collaborations and the input from the Belle and BaBar collaborations providing extra constraint on the $f_+$ form factor.

For the $\mathcal{R}(D^*)$, the strong decay of the $D^*$ meson makes the theoretical evaluation of the form factors more challenging. The average of the predictions leads to a value of $\mathcal{R}(D^*) = 0.258 \pm 0.005$. This is an arithmetic combination of various predictions using different assumptions on the $B \to D^*$ Form Factors. The first prediction [49] used a CLN parametrization of the Form Factors. Since 2017, the BGL parametrization has been used in conjunction with input unfolded data from Belle and corrections of order $1/m_c^2$. The values of the various predictions for $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ are reported in Table 7.

| Observable | SM prediction | Ref. |
|:---:|:---:|:---:|
| $\mathcal{R}(D)$ | $0.299 \pm 0.011$ | [50] |
| $\mathcal{R}(D)$ | $0.300 \pm 0.008$ | [51] |
| $\mathcal{R}(D)$ | $0.299 \pm 0.003$ | [45] |
| $\mathcal{R}(D)$ | $0.299 \pm 0.004$ | [52] |
| $\mathcal{R}(D)$ | $0.299 \pm 0.003$ | [53] |
| $\mathcal{R}(D^*)$ | $0.252 \pm 0.003$ | [49] |
| $\mathcal{R}(D^*)$ | $0.260 \pm 0.008$ | [54] |
| $\mathcal{R}(D^*)$ | $0.257 \pm 0.005$ | [52] |
| $\mathcal{R}(D^*)$ | $0.257 \pm 0.003$ | [53] |

Table 7: SM predictions for the $\mathcal{R}(D^{(*)})$ ratios.

# 5   Development of RooHammerModel

Analyses of semileptonic $b \to c\ell\nu$ decays, like the one presented in this thesis, heavily rely on detailed MC simulations to model the detector responses and acceptance. Distributions of events contributing to the final state under study are then constructed out of these simulations. These *templates* are then employed in fits in order to interpret the data and to extract measurements of observables like $\mathcal{R}(D^{(*,+)})$ or $|V_{cb}|$.

The shape of these templates and the acceptance of the events in general may depend on the underlying theoretical model employed in the generation of the events. When trying to reinterpret the data with a different model, this has to be taken this into account properly. This is usually done by analyzing the data, after it has been manipulated in order to remove the acceptance and reconstruction effects. This procedure is called *unfolding*, and it is a difficult experimental procedure which is also usually accompanied by additional systematic uncertainties.

This chapter starts advocating for a more suitable procedure, especially in the context of analyses of $b \to c\ell\nu$ decays. In the second part, it introduces a tool specifically developed to perform forward folding in analyses of $b \to c\ell\nu$ data, called *Hammer* [55]. Finally it will describe the development and validation of an interface to Hammer, called *RooHammerModel*, that can be used to perform forward folding in analyses implemented with *RooFit*.

This work has resulted in the publication of an article on preprint [56].

## 5.1   Fitting for Form Factors and New Physics parameters in $b \to c\ell\nu$ analyses

As said in the previous chapter, the shape of the templates used to interpret $b \to c\ell\nu$ data, as well as the acceptance of the events, may depend on the theoretical model employed during the generation of MC events. The theoretical model for the hadronic $B \to D^{(*)}$ transition is encoded in the choice of the parametrization for the Form Factors which, as described in the previous section, encode the long distance physics phenomena occurring in the non perturbative QCD part of the transition. The Form Factors depend on parameters which are usually measured from data: the impact that the choice of the value of the Form Factor parameters has one the measured values is important, as it will be clear in the next chapters of this thesis.

The situation becomes even more complex when a New Physics analysis of the data has to be performed. In this case the number of parameters of interest to be measured increases, including also parameters encoding the physics structure of NP phenomena that may contribute to the process under study.

In the SM, after integrating out the degrees of freedom of the heavy $W$ boson, $b \to c\ell\nu$ transitions are mediated by a V-A 4-Fermi operator of this type

$$\mathcal{O}_{SM} = 2\sqrt{2} G_F V_{cb} \left( \bar{c} \gamma^\mu P_L b \right) \left( \bar{\ell} \gamma_\mu P_L \nu_\ell \right), \tag{121}$$

with $P_L = \frac{1-\gamma_5}{2}$ the left helicity projector. If New Physics contributes to the process, new 4-Fermi field operators may arise. The most general interaction operator mediating $b \to c\ell\nu$ transitions can be written as heavy

$$\mathcal{O}_{4-\text{Fermi}} = 2\sqrt{2} G_F V_{cb} \sum_{XY} c_{XY} \left( \bar{c} \; \Gamma_X \; b \right) \left( \bar{\ell} \; \Gamma_Y \; \nu_\ell \right), \tag{122}$$

where $\{\Gamma_X\}$ is a complete set of Dirac currents, encoding the structure of the currents contributing to the process and $c_{XY}$ are coefficients arising when the heavy degrees of freedom

are integrated out of the theory. These parameters, called *Wilson Coefficients*, are the parameters of interest when performing NP analyses.

These Wilson Coefficients can be written as $c_{XY} = S_{XY}, V_{XY}, T_{XY}$, with $S, V, T$ denoting the Lorentz Structure (Scalar, Vector, Tensor respectively) and $X, Y = L, R$ denoting the chirality of the quarks and leptons fields, respectively. The complete set of operators, with the associated Wilson Coefficients, is summarized in Table 8.

| Current | Wilson Coefficient, $c_{XY}$ | Operator |
|---------|------------------------------|----------|
| SM | $1$ | $[\bar{c}\gamma^\mu P_L b]\,[\bar{\ell}\gamma_\mu P_L \nu]$ |
| Vector | $V_{LL}$ | $[\bar{c}\gamma^\mu P_L b]\,[\bar{\ell}\gamma_\mu P_L \nu]$ |
|  | $V_{LR}$ | $[\bar{c}\gamma^\mu P_L b]\,[\bar{\ell}\gamma_\mu P_R \nu]$ |
|  | $V_{RL}$ | $[\bar{c}\gamma^\mu P_R b]\,[\bar{\ell}\gamma_\mu P_L \nu]$ |
|  | $V_{RR}$ | $[\bar{c}\gamma^\mu P_R b]\,[\bar{\ell}\gamma_\mu P_R \nu]$ |
| Scalar | $S_{LL}$ | $[\bar{c} P_L b]\,[\bar{\ell} P_L \nu]$ |
|  | $S_{LR}$ | $[\bar{c} P_L b]\,[\bar{\ell} P_R \nu]$ |
|  | $S_{RL}$ | $[\bar{c} P_R b]\,[\bar{\ell} P_L \nu]$ |
|  | $S_{RR}$ | $[\bar{c} P_R b]\,[\bar{\ell} P_R \nu]$ |
| Tensor | $T_{LL}$ | $[\bar{c}\sigma^{\mu\nu} P_L b]\,[\bar{\ell}\sigma_{\mu\nu} P_L \nu]$ |
|  | $T_{RR}$ | $[\bar{c}\sigma^{\mu\nu} P_R b]\,[\bar{\ell}\sigma_{\mu\nu} P_R \nu]$ |

Table 8: Operator basis for $b \to c\ell\nu$, with the associated Wilson Coefficients.

The mixed tensorial terms, with operators corresponding, to $[\bar{c}\sigma^{\mu\nu} P_L b]\,[\bar{\ell}\sigma_{\mu\nu} P_R \nu]$ and $[\bar{c}\sigma^{\mu\nu} P_R b]\,[\bar{\ell}\sigma_{\mu\nu} P_L \nu]$, always vanish thanks to the equality

$$\sigma^{\mu\nu}\gamma_5 = \frac{i}{2}\varepsilon^{\mu\nu\rho\sigma}\sigma_{\rho\sigma}. \tag{123}$$

## 5.2 Forward folding, rather than unfolding

Analyses of semileptonic data require large MC samples, due to the high branching fractions for decays like $B \to D^{(*)}\ell\nu$. If the analysis is performed with an underlying theoretical model and one would like to reinterpret the data with another underlying model, ideally the MC would have to be reproduced in the new configuration. This is also true when the systematic uncertainty due to the choice of *e.g.* the Form Factor parameters is not negligible. To assess the impact of this choice one would have to generate MC samples for a bunch of Form Factor parameters configurations and repeat the analysis, taking as a systematic uncertainty the envelope of the results obtained. The situation becomes even more complicated if the analysis is aiming at measuring New Physics parameters, in which case the number of parameters that have to be measured from the data increases dramatically, as explained in the previous section.

Another complication that arises in analyses of $b \to c\ell\nu$ data is due to the fact that the unfolding procedure becomes difficult especially due to the high level of backgrounds to be expected. Furthermore, in this case, the underlying process involved in the decays of many background processes is the same $b \to c\ell\nu$ as the signal. If NP is to be expected in the signal, in principle it could even contribute to the shape of the background templates, and therefore, before unfolding, a subtraction made by assuming the backgrounds to have a Standard Model like shape, could ideally become inconsistent.

For these reasons, another procedure, alternative to unfolding, could play a better role in analyses like the one presented in this thesis. This procedure is usually referred to as *forward folding*. Instead of unfolding the response of the detector from the measured, background subtracted data, the physics modelling is folded into the smearing and acceptance of the detector. The detector response is taken from a single MC dataset, generated from an underlying theory model. The events are then reweighted according to the ratio of the decay rate of the target model and the one of the generation model. In this way the shape and normalization of the templates will depend on the parameters defining the theory model under inspection, with them being either Form Factor parameters, Wilson Coefficients, or both.

This has many advantages. In this way a likelihood can in principle be constructed also as a function of the theory parameters of interest, that can then be either measured directly on data or floated as nuisance parameters to include systematic uncertainties in the fit. Furthermore, the same procedure could be applied to all samples that are expected to be affected by $b \to c\ell\nu$ NP effects, eliminating the inconsistency of model dependent subtractions. Finally the procedure can be performed with a single MC sample produced, removing the need to produce many MC samples in a grid of the parameters of interest.

## 5.3  The Hammer reweighting tool

*Hammer* [55] is a software tool specifically developed to reweight large MC samples of $b \to c\ell\nu$ events to any desired NP model or to any description of the hadronic matrix elements. Hammer makes use of efficient amplitude-level calculations strategies and it is designed to interface with the main existing HEP experimental analysis frameworks. The only required input are the event-level true kinematic quantities of the particles involved in the decay. The reweight can be applied both at a per-event basis on or a histogram basis, and it is fast enough to be performed at each minimisation step of a fit procedure, so that it is the ideal tool to implement forward folding, as described in the previous paragraph.

In order to reweight a MC event, indexed by $I$, from the theory used at generation level, denoted as 'old' to a different one, denoted as 'new', the weight to be assigned is the truth-level ratio of the decay rates

$$r_I = \frac{d\Gamma_I^{\text{new}}/d\mathcal{PS}}{d\Gamma_I^{\text{old}}/d\mathcal{PS}}. \tag{124}$$

Hammer makes use of the fact that, in general, a matrix element can be written in the form

$$\mathcal{M}^{\{s\}}(\{q\}) = \sum_{\alpha,i} c_\alpha F_i(\{q\}) \mathcal{A}_{\alpha i}^{\{s\}}(\{q\}), \tag{125}$$

where $\{s\}$ is a set of external quantum numbers (like spin, parity), $\{q\}$ the set of four-moment, $c_\alpha$ are Wilson coefficients and $F_i$ are Form Factors. The object $\mathcal{A}_{\alpha,i}$ is a NP- and FF- generalized

amplitude tensor. The decay rate can then be written as

$$
\frac{d\Gamma^{\{s\}}}{d\mathcal{PS}} = \sum_{\alpha,i,\beta,j} c_\alpha c_\beta^\dagger F_i(\{q\}) F_j^\dagger(\{q\}) \mathcal{A}_{\alpha i}^{\{s\}}(\{q\}) \mathcal{A}_{\beta j}^{\dagger\{s\}}(\{q\}) = \sum_{\alpha,i,\beta,j} c_\alpha c_\beta^\dagger F_i(\{q\}) F_j^\dagger(\{q\}) \mathcal{W}_{\alpha i \beta j}
$$
(126)

where the term $\mathcal{W}_{\alpha i \beta j}$ is a weight tensor evaluated from the generalized amplitude tensor. The object $\sum_{ij} F_i F_j \mathcal{W}_{ij}$ is independent of the Wilson Coefficients and therefore, once evaluated for a specific event it can be reused. To generate an event weight it can be contracted with a choice of Wilson Coefficients. This can also be done for all the events falling inside a single bin, denoted as $\Omega$, because the Wilson Coefficients can be factored out of the phase space integral

$$
\Gamma_\Omega^{\{s\}} = \sum_{\alpha\beta} c_\alpha c_\beta^\dagger \int_\Omega d\mathcal{PS} \sum_{ij} F_i F_j^\dagger(\{q\}) \mathcal{W}_{\alpha i \beta j}(\{(q)\}).
$$
(127)

In this way efficient reweighting for a different choice of Wilson coefficients can be made, but in order to generate a weight for a different choice of Form Factor parameters, the entire Form Factor $F_i$ has to be evaluated from scratch, because they depend on the kinematics of an event and they cannot be factored out of the phase-space integral. This incovenience is surpassed by Hammer employing a linearization of the Form Factors. A Form Factor parametrization with a set of parameters $\{\mu\}$ can be linearized around a point in the parameters space $\{\mu^0\}$ so that

$$
F_i(\{q\};\{\mu\}) = F_i(\{q\},\{\mu^0\}) + \sum_a F'_{i,a}(\{q\},\{\mu^0\})\Delta\mu_a,
$$
(128)

with $\Delta\mu_a = \mu_a - \mu_a^0$, here $a$ is one or more variational indices.

This linearization in general brings some level of approximation, but is exact for the BGL parametrization.

Defining $\xi_a = (1, \mu_a - \mu_a^0)$ and $\Phi_{i,a+1} = (F_i, F'_{i,a})$, the previous equation becomes

$$
F_i(\{q\};\{\mu\}) = \sum_a \xi_a \Phi_{i,a},
$$
(129)

and the differential rate can be therefore written in this way

$$
\frac{d\Gamma^{\{s\}}}{d\mathcal{PS}} = \sum_{\alpha,a\beta,b} c_\alpha c_\beta^\dagger \xi_a \xi_b^\dagger \mathcal{U}_{\alpha a \beta b}^{\{s\}}
$$
(130)

$$
\mathcal{U}_{\alpha a \beta b}^{\{s\}} = \sum_{ij} \Phi_{i,a} \Phi_{j,b}^\dagger(\{q\}) \mathcal{W}_{\alpha i \beta j}^{\{s\}}.
$$
(131)

With this linearization, the parameters $\xi_a$ are now independent on the phase space and can be factored out of the space integral which now acts only on a NP- and FF- generalized tensor

$$
\Gamma_\Omega^{\{s\}} = \sum_{\alpha,\beta,a,b} c_\alpha c_\beta^\dagger \xi_a \xi_b^\dagger \int_\Omega d\mathcal{PS}\, \mathcal{U}_{a\alpha b\beta}^{\{s\}},
$$
(132)

and the rates can be evaluated by contracting this generalized tensor with the new choice of $c_\alpha$ and $\xi_a$.

## 5.4 RooHammerModel: integrating hammer into RooFit analyses

The Hammer software tool has been used in this thesis in order to propagate the systematic uncertainties due to the choice of the Form Factor parameters to the parameters of interest. As it will be described later on, the fit performed is a binned maximum likelihood fit, whose likelihood is constructed with the *RooFit* [57] analysis framework.

To perform forward folding of Form Factor parameters, the templates used to construct the model used in the fit have to have a shape which depends on the Form Factor parameters. A new class, called *RooHammerModel* [56], has been developed to serve as an interface between the RooFit analysis package and the Hammer software tool. The class represents a binned pdf that depends on a set of Wilson Coefficients and Form Factor parameters. It can be configured to take in input the set of parameters the pdf has to depend on and the files in which the Hammer related objects are persisted (created in the preprocessing phase of the MC samples).

This pdf class can be used in any RooFit application, and in particular to perform binned template maximum likelihood fits in RooFit.

Internally the class takes care of communicating with the minimization program and to reweight the template used in the model only when one or more parameters relevant to Hammer have been changed during the likelihood maximization. If no parameters have been changed in the current minimization step, no reweight is performed and a cached version of the histogram is used instead. The RooHammerModel pdf is normalized after each reweighting step. In general, if the number of MC events in a given bin is zero but the number of observed events is not null, the definition of binned maximum likelihood might suffer from numerical inconsistencies. In order to avoid this issue, all the empty bins in a RooHammerModel template are corrected with a small value, after each reweighting step. The full details on how to use this tool in a real analysis and the more technical aspects of the implementation can be read in [56].

Before its deployment in the measurement, this tool has been tested by means of fits to pseudo data that make use of the same technologies used in the fit to data. The physics scenario used is a simplified one with respect to the one expected in real data. Datasets in which only $\overline{B}^0 \rightarrow D^{*+}\mu^-\overline{\nu}_\mu$ and $\overline{B}^0 \rightarrow D^{*+}\tau^-\overline{\nu}_\tau$ contribute, with $\tau^- \rightarrow \mu^-\overline{\nu}_\mu\nu_\tau$ and $D^{*+} \rightarrow D^+\pi^0$, have been generated. It is assumed that both decays have been reconstructed in the same $D^+\mu^-$ final state, with the $\pi^0$ going unreconstructed.

For each pseudo-experiment three samples have been generated, representing the pseudo-data and the $\overline{B}^0 \rightarrow D^{*+}\mu^-\overline{\nu}_\mu$ and $\overline{B}^0 \rightarrow D^{*+}\tau^-\overline{\nu}_\tau$ MC samples used to generate the RooHammer-Model functions. They have been all generated using RapidSim [58] with a LHCb specific configuration and EvtGen [59]. The pseudo-data and the $\overline{B}^0 \rightarrow D^{*+}\tau^-\overline{\nu}_\tau$ MC samples have been generated using a CLN parametrization of the Form Factors, whereas the $\overline{B}^0 \rightarrow D^{*+}\mu^-\overline{\nu}_\mu$ MC samples have been generated with a pure Phase Space model and then are reweighted to a CLN model using Hammer before fitting. For each pseudo-experiment the number of generated events in each sample is fluctuated according to a Poisson distribution.

The pseudo-data is fitted with a model constructed out of two RooHammerModel pdfs, one for each sample contributing. The expected number of events in a bin centered at $x$ is parametrized in this way

$$f(x,\theta) = N_{\text{total}} \left( \frac{1}{1 + \mathcal{R}(D^*)} h_\mu(x,\theta) + \frac{\mathcal{R}(D^*)}{1 + \mathcal{R}(D^*)} h_\tau(x,\theta) \right) \qquad (133)$$

where $N_{\text{total}}$ represents the total sum of events, $\mathcal{R}(D^*)$ the fraction of observed events in the

$\tau$ sample with respect to the ones observed in the $\mu$ sample and $h_\mu(x,\theta)$ and $h_\tau(x,\theta)$ are the RooHammerModel pdfs. The parameters being measured by the fit are $N_{total}$, $\mathcal{R}(D^*)$ and the parameters of the CLN parametrization, encoded in the parameters $\theta$ in the equation above. The CLN parameters that are left floating in the fit are $\rho^2$, $R_1$ and $R_2$ (to be read as their value at zero recoil, $\rho^2(w=1)$, $R_1(w=1)$, $R_2(w=1)$). As a technical note, the parameter being measured by the fit are differences with respect to the central value set during the preprocessing step of the samples with Hammer, namely $\Delta\rho^2$, $\Delta R_1$, $\Delta R_2$. In each pseudo-experiment the pseudo-dataset has been generated with $N_{\text{total}} = 10^6$, $\mathcal{R}(D^*) = 0.3$ and $\Delta\rho^2 = \Delta R_1 = \Delta R_2 = 0$ (with central values of $\rho^2 = 1.207$, $R_1 = 1.401$, $R_2 = 0.854$).

Three variables are used in the fit: the muon energy in the $\overline{B}^0$ rest frame, the invariant mass of the lepton-neutrino system, $q^2 = (p(\overline{B}^0) - p(D^+))^2$ and the squared missing mass, $m_{miss}^2 = (p(\overline{B}^0) - p(D^+) - p(\mu^-))^2$, where $p(P)$ denotes the 4-momentum of the particle $P$. No experimental resolution effects have been taken into account in these studies, but it has to be noticed that the usage of the 4-momentum of the $D^+$ instead of the one of the $D^*$ introduces some level of smearing into the distribution of these variables.

A set of 1000 pseudo-experiments have been generated and fitted, and for each of them the pull of the fit variables is evaluated as

$$\theta_{\text{pull}} = \frac{\hat{\theta} - \theta_{\text{true}}}{\hat{\sigma}_{\hat{\theta}}}, \tag{134}$$

where $\hat{\theta}$ and $\hat{\sigma}_{\hat{\theta}}$ are the fit estimates of the parameter $\theta$ and of its uncertainty. The distribution of the pulls obtained in this way are fitted with a Gaussian pdf, and the result is shown in Fig. 6.

All the distributions are found to be compatible with a standard unit gaussian. This serves as a validation of the RooHammerModel as a tool to perform forward folding of Form Factor parameters. Part of the events have also been generated using EvtGen as a generator, configured to use the CLN parametrization. The good coverage of the fit results is also an indication of the compatibility between the EvtGen and Hammer implementation of the CLN parametrization. Finally, since the CLN parametrization is not linear in the form factor parameters, this test serves also as a validation of the good approximation of the linearized CLN model as a function of the Form Factor parameters, within their statistical uncertainties.
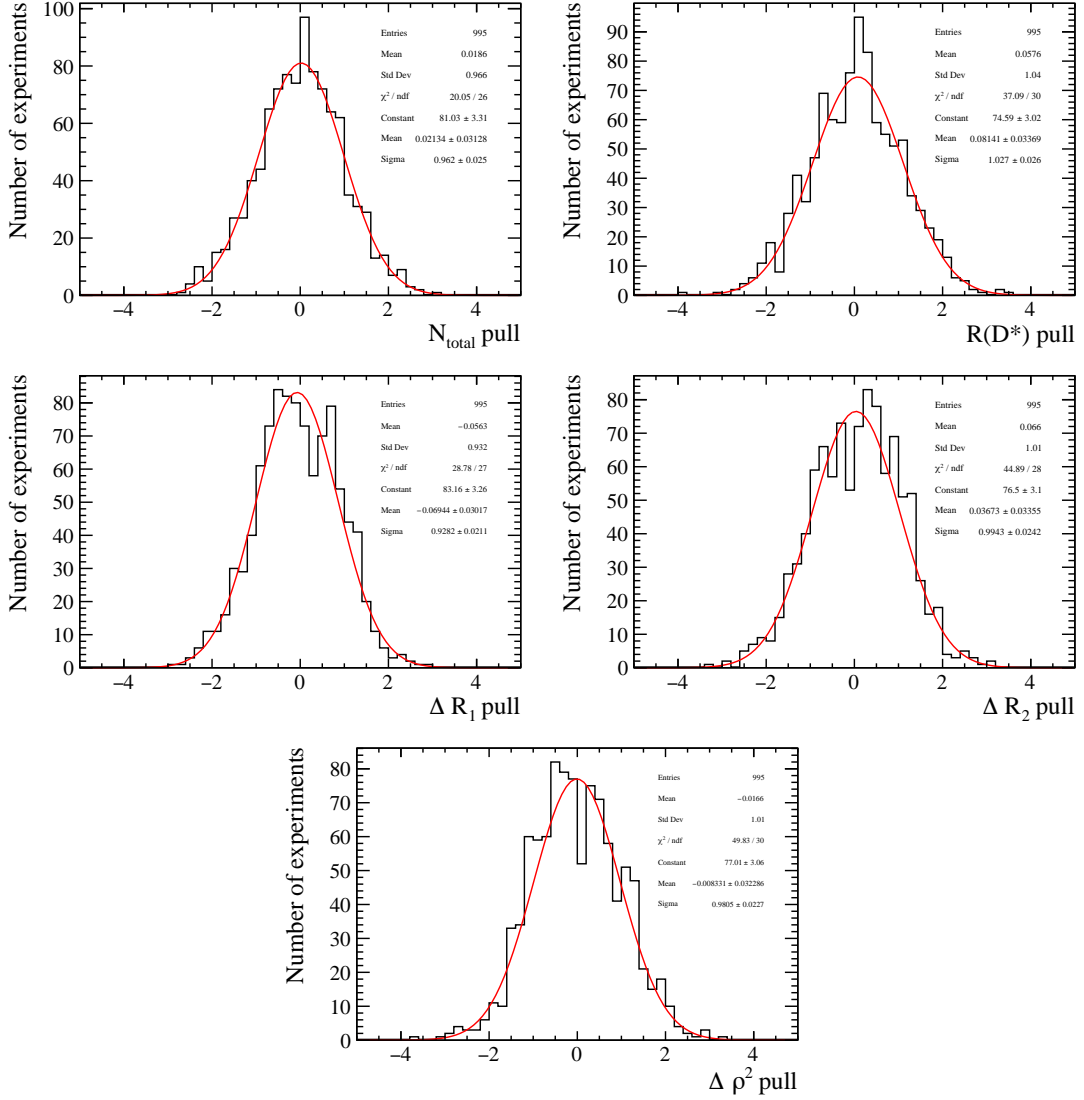
Figure 6: Pull distributions for the fit parameters. The result of a guassian fit to each distribution is superimposed as a red curve.

## 5.5 Looking ahead: a NP analysis strategy using forward folding

The Hammer tool enables a fast reweighting of the matrix elements that can be exploited to implement forward folding in analyses of $b \to c\ell\nu$ data. This can be done to measure both the form factors parameters and New Physics Wilson coefficients directly from data.

This would be very important especially in the case in which the discrepancies in the $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ observables were established. In this case an extraction of the Wilson Coefficients directly from data would be very important, since in [55] it has been shown that performing an analysis assuming the SM could lead to biases both in the $\mathcal{R}(D^{(*)})$ and in the Wilson Coefficient extracted subsequently when interpreting the results of the measurements.

Up to now the measurements of $\mathcal{R}(D^{(*)})$ proceeded with extracting the parameters of interest

fitting the data using the shape of the single contributions extracted from SM decays. When moving to an analysis employing forward folding, it would be very important to understand in which decays, with which variables and in which type of analysis one would be sensitive to a given set of Wilson Coefficients. In this section, a very preliminary study in this direction, using pseudo data of $B \to D^{(*)}\ell\nu$ decays drawn from the LHCb simulation will be shown. This is still very preliminary and no solid quantitative conclusion should be drawn from the results of it. The aim is to have an idea of what type of NP analysis could be performed with the data available and with the decay under study.

The idea is to get a grasp of what parameters the analysis of $B \to D^{(*)+}\tau\nu$ decays such as the ones studied in this thesis is sensitive to. Pseudo datasets have been generated using the number of events expected in the $\mathcal{R}(D^+)$ analysis presented in this thesis, after selections, in a dataset collected during the 2015 and 2016 datasets.

The datasets generated contain only the signal and normalization modes:

- $\overline{B}^0 \to (D^+ \to K^-\pi^+\pi^+)\mu^-\overline{\nu}_\mu$;

- $\overline{B}^0 \to (D^+ \to K^-\pi^+\pi^+)(\tau^- \to \mu^-\nu_\tau\overline{\nu}_\mu)\overline{\nu}_\tau$;

- $\overline{B}^0 \to (D^{*+} \to (D^+ \to K^-\pi^+\pi^+)\pi^0/\gamma)\mu^-\overline{\nu}_\mu$;

- $\overline{B}^0 \to (D^{*+} \to (D^+ \to K^-\pi^+\pi^+)(\tau^- \to \mu^-\nu_\tau\overline{\nu}_\mu)\overline{\nu}_\tau$.

generated assuming Standard Model holds and assuming a value of $\mathcal{R}(D^+)$ and $\mathcal{R}(D^*)$ equal to the current world average. In order to generate the pseudodatasets, the expected number of events is evaluated for each bin and then smeared according to a Poisson distribution, using the BLPR [53] form factor parametrization.

The pseudodata is then fitted using RooHammerModel templates for all decay modes, with NP Wilson Coefficients floating for the two tauonic modes, whereas no NP contribution is assumed to contribute to the muonic modes. In this preliminary analysis, NP is assumed to be left-handed in the lepton current and therefore the $V_{LR}, V_{RR}, S_{LR}, S_{RR}$ and $T_{RR}$ parameters have been fixed in the fit to zero.

The sensitivity to a given set of Wilson Coefficients depends on the decay under study. The $B \to D\ell\nu$ transition is a pseudo-scalar to pseudo-scalar transition, and only the scalar and vector matrix elements are non-zero in the decay rate. For this reason, a data sample of exclusive $B \to D\ell\nu$ transitions would only offer sensitivity to the scalar Wilson coefficient $(S_L^+ = \frac{S_{LL}+S_{RL}}{2})$, as opposed to the pseudo-scalar Wilson Coefficient $(S_L^- = \frac{S_{LL}-S_{RL}}{2})$. In $B \to D^*\ell\nu$ transitions, instead, being the $D^*$ a pseudo-vector, also the pseudo-scalar matrix element is non-zero. Therefore, having both $B \to D\ell\nu$ and $B \to D^*\ell\nu$ samples in the same data-set would in principle give experimental access to $S_{LL}, S_{RL}$ Wilson Coefficients together.

In a first version of the fit, the parameters $S_{LL}$ and $S_{RL}$ were left floating simultaneously, but a high correlation between these parameters has been found which would lead to large instabilities in the fit. For this reason the fit is performed in a rotated basis, floating the parameters $S_L^-$ and $S_L^+$ instead.

With these caveats, a shape analysis of this kind of datasets would give some sensitivity to the $S_{LL}, S_{RL}, V_{RL}$ and $T_{LL}$ Wilson Coefficients. The $V_{LL}$ Wilson Coefficient, instead, would play a different role. In fact this coefficient is accompanied by a SM-like current structure. A value of this coefficient different from zero would be reflected only in a change of the rate of

the $B \to D^{(*)}\tau\nu$ decays with respect to the SM predictions, and no sensitivity could be reached through a shape-only analysis.

Nevertheless, a very good opportunity is given by analyses of $b \to c\ell\nu$ analyses using $\tau \to \mu\nu\nu$ decays. In these analyses, the final state of the signal and normalization modes is the same and reconstructed in the same data sample. This enables to be sensitive also to the $V_{LL}$ Wilson Coefficient, when assuming no New Physics in the normalization sample and measuring also the rate of the signal modes with respect to the normalization mode as a function of the NP Wilson Coefficients. This requires a slight shift in the analysis strategy with respect to the previous ones. In $b \to c\ell\nu$ with $\tau \to \mu\nu\nu$ analyses published so far, the measured parameters were the yield of the normalization sample and the ratio of the signal and normalization yields. In order to get some information on Wilson Coefficients also from the rate of the signal sample, a reference normalization must be assumed. This can be done for example fixing the expected yield of the signal sample with respect to the normalization sample to the expected value in the SM, and to fit for the ratio of the signal rate as a function of the Wilson Coefficients.

This is exactly what has been done in this study. The expected number of events in each bin used to fit the pseudo dataset is expressed in this way:

$$N_i \quad = \frac{N(D^+\mu\nu)}{\int h_{D^+\mu\nu}(x)dx} h_{D^+\mu\nu}(x_i) + \frac{N(D^+\mu\nu)}{\int h_{D^+\tau\nu}(x,\theta=0)dx} h_{D^+\tau\nu}(x_i,\theta) \times \mathcal{R}_{SM}(D^+) + \qquad (135)$$

$$+ \frac{N(D^*\mu\nu)}{\int h_{D^*\mu\nu}(x)dx} h_{D^*\mu\nu}(x_i) + \frac{N(D^*\mu\nu)}{\int h_{D^*\tau\nu}(x,\theta=0)dx} h_{D^*\tau\nu}(x_i,\theta) \times \mathcal{R}_{SM}(D^*) \qquad (136)$$

where $N(j)$ is the total yield for mode $j$, $x_i$ is the value of the fit variables in the center of bin $i$, $h_j$ is the unnormalized RooHammerModel template for the mode $j$ and $\theta$ is the set of NP Wilson Coefficients. The parameters $\mathcal{R}_{SM}(D^{(*)})$ are the LFU parameters assuming the SM and are fixed in the fit. The change in the yield of the signal model assuming a different value of the Wilson Coefficients is encoded in the change of the normalization of the $h_{D^*\tau\nu}(x_i,\theta)$ and $h_{D^+\tau\nu}(x_i,\theta)$ RooHammerModel templates. This is possible thanks to the fact that the weights evaluated in Hammer correspond to the ratio of the differential rates in the NP model with respect to the SM assumption. Furthermore, with this approach, a change in the value of the acceptance given a change of the NP Wilson Coefficients would be already taken into account. Therefore, using this approach, information about $V_{LL}$ could be extracted directly from data as well.

The fit is performed on the set of variables used for the analysis that will be presented in the next chapters of this thesis. The Wilson Coefficients are also assumed to be complex and their imaginary part is floated in the fit, apart from $V_{LL}$, which is assumed to be real. In this way, given that no specific $CP$ violating variable has been used in the fit and therefore no absolute $CP$ violating phase could be extracted from the fit, the value of the phase of the Wilson Coefficients has to be interpreted as the one relative to phase of the $V_{LL}$ Wilson Coefficient, which will remain unknown.

The result of the fit to one of the pseudodatasets is shown in Fig. 7 and in Fig. 8 in various $q^2$ bins. The generation and fit is performed around 500 times, and the scatter plots of the result of the fit in the space of the real and imaginary part of the Wilson Coefficients is reported in Fig. 9.

It can be noticed that some information on all the Wilson Coefficients floated in the fit can be extracted exploiting this analysis strategy, although on some parameters the uncertainty is large already in this configuration in which no backgrounds are present and no systematic uncertainty

is taken into account. The larger uncertainty has been observed in the imaginary part of the Wilson Coefficients, as expected.

The purpose of this study was to propose a starting point for a measurement of the Wilson Coefficients using forward folding in a $b \to c\ell\nu$ analysis, and to study the parametrization needed to extract the highest information on NP parameters, using $\tau \to \mu\nu\nu$ analyses. The results obtained have shown that some level of sensitivity could be reached in all the NP parameters left floating in the fit, with the analysis strategy described in the text. No quantitative conclusion can be extracted from this study, as additional and more systematic checks have to be performed on this analysis. The next steps for this kind of study could be, for example, to include backgrounds and systematic uncertainties to get a grasp of the expected dilution on the sensitivity on the Wilson Coefficients, and to study new fit variables, especially the ones related the angular distribution of the decays, to boost the experimental sensitivity to NP coefficients.
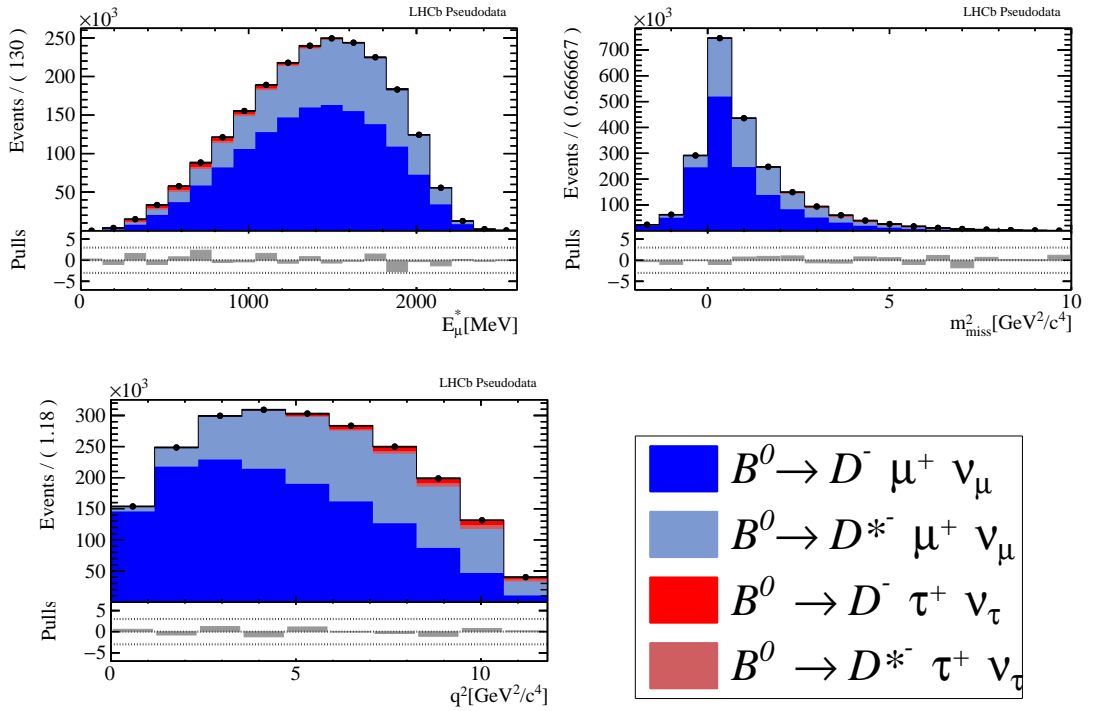


Figure 7: Result of the fit to a single pseudodataset.
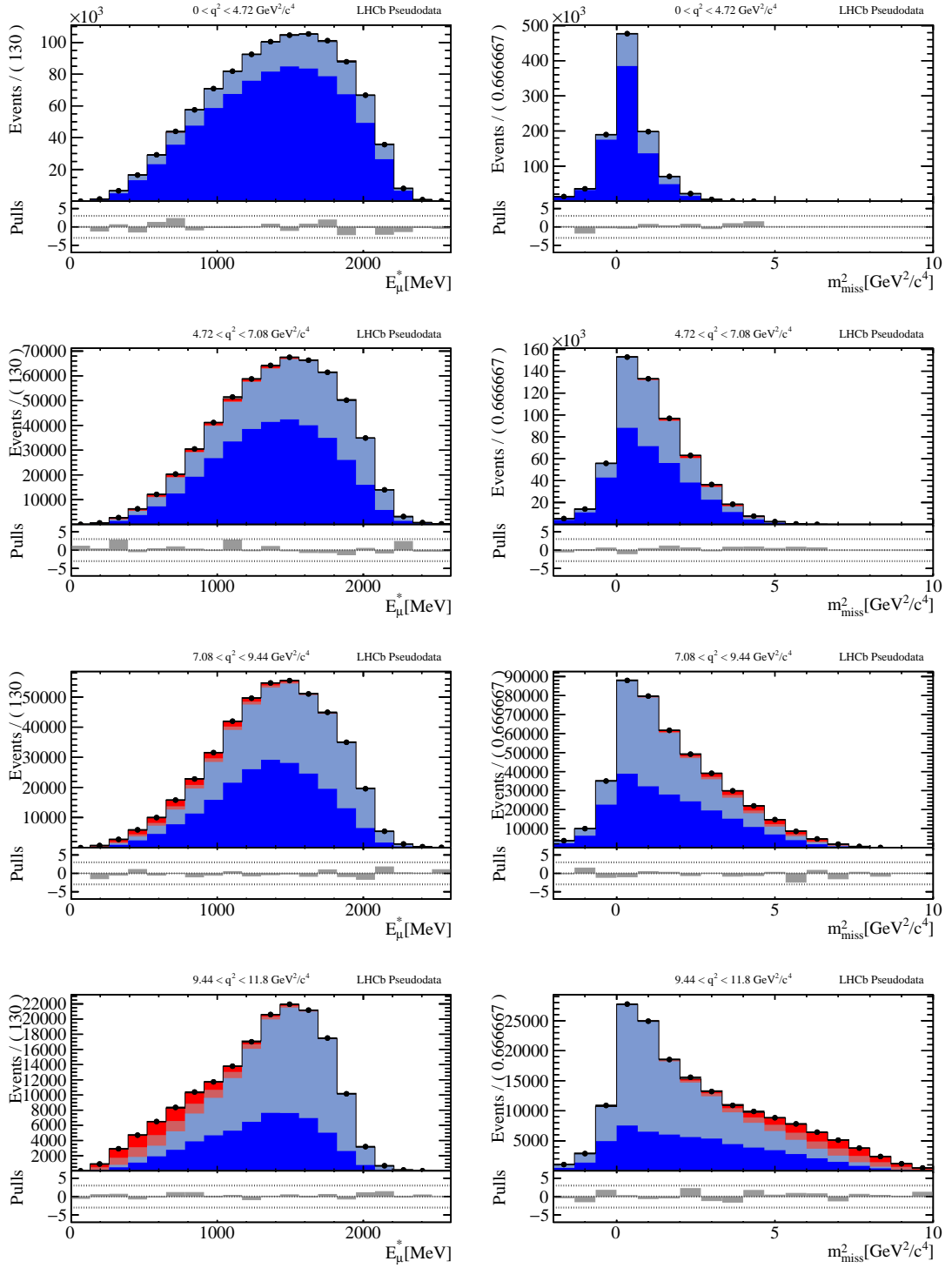
Figure 8: Result of the fit to a single pseudodataset, projected in four different $q^2$ regions.
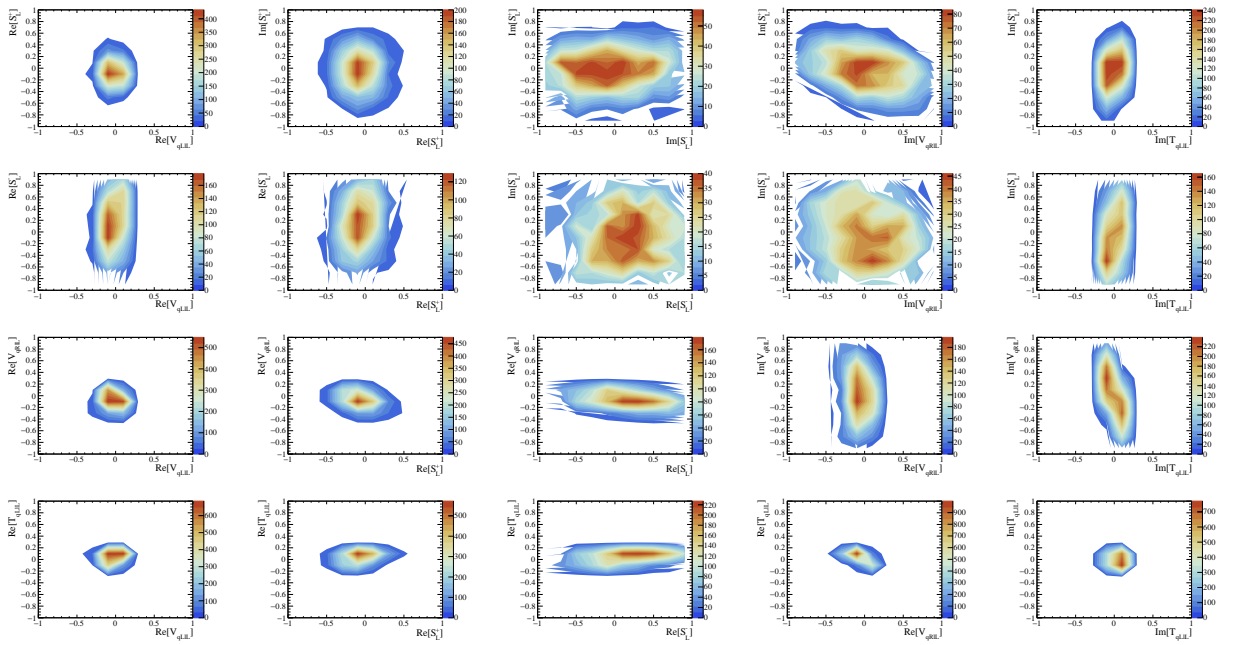
Figure 9: Scatter plot of the result of the fit to the pseudodatasets, in the space of the (real and imaginary part of the) New Physics Wilson Coefficients.

# 6 LHC and the LHCb experiment

LHCb is one of the four main experiments located at the Large Hadron Collider (LHC) at CERN, the European organization for Nuclear Research. It is designed to perform precision measurements for heavy-quark decays. This chapter will give an overview of the experimental setup in which the measurement takes place. The first sections will give a brief overview of the LHC and of the production mechanism of $b$-quarks and $B$ mesons in $pp$ collisions. Then the LHCb detector is described in all the subsystems needed to perform this analysis. Finally the trigger system and the event reconstruction, needed to select and save the high level objects used by this measurement, are discussed.

## 6.1 The Large Hadron Collider

The LHC [60] at CERN is the world's largest particle accelerator. It is a storage ring, consisting of two parallel pipes accelerating protons and heavy ions in opposite directions and colliding them in four interaction points. The accelerator is located 100 m underground, in the Geneva area, and it has been built inside the 27 km tunnel that was constructed for the Large Electron Positron (LEP) collider. A schematic view of the accelerator and the experiments installed on the interaction points is reported in Fig. 10.
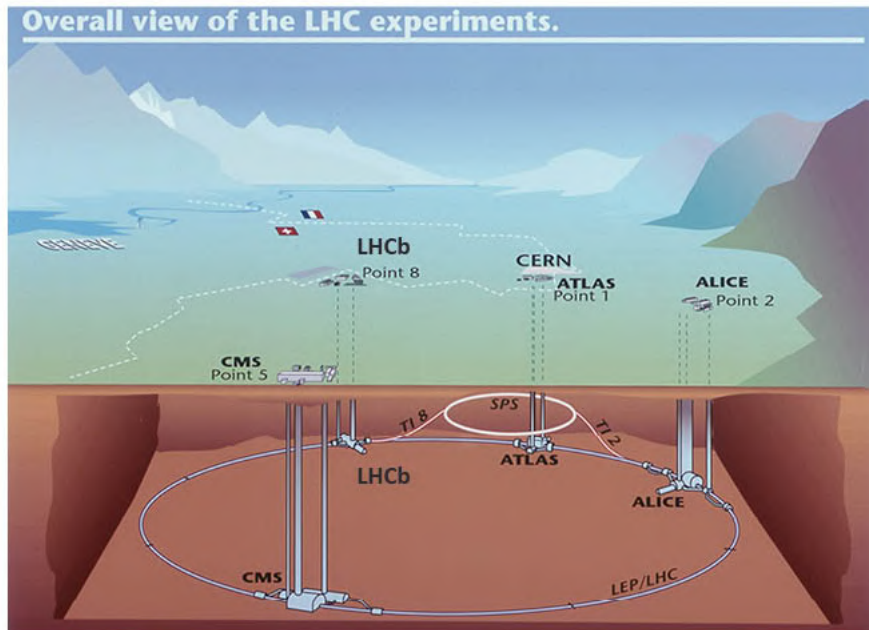


Figure 10: A schematic representation of the LHC collider.

The machine was built to reach a center-of-mass energy of $\sqrt{s} = 14$ TeV and an instantaneous luminosity of $\mathcal{L} = 10^{34} \text{cm}^{-2}\text{s}^{-1}$ in $pp$ and heavy ions collisions. In the nominal regime the LHC

stores 2808 proton bunches per ring, each of them made of around $1.1 \times 10^{11}$ protons, and the bunch crossing frequency is 40 MHz.

The protons are extracted from ionized hydrogen atoms and are then accelerated in various steps, by various preaccelerators: the linear accelerator Linac2, the Proton Synchroton Booster (PBS), the Proton Synchroton (PS) and the Super Proton Synchroton (SPS). In order to keep the protons on a closed orbit and to reduce the transverse beam size, a series of 1232 dipole and 392 quadrupole magnets are used. Radio Frequency (RF) cavities are employed to accelerate the proton beams, after their injection into the LHC.
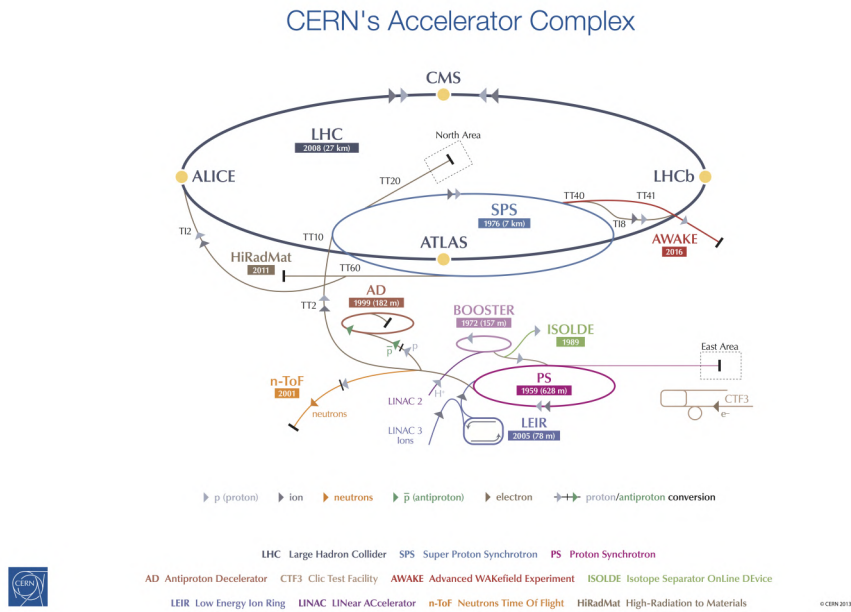


Figure 11: A schematic representation of the complex accelerators system at CERN.

The magnetic field generated by the dipoles ranges up to 8.34 T, which can be reached only through a very high intensity current. For this reason the magnets are made of a superconducting material, NbTi, cooled at a temperature of 1.9 K by a liquid helium cryogenic system.

Right before the interaction points, a special magnet configuration, called low-$\beta$ triplet, is used to reduce at its minimum the $\beta$ function of the beam, in order to have maximum luminosity during bunch crossing.

### 6.1.1  LHC experiments

The LHC is host to many experiments, covering a wide range of physics reach. A brief description is reported in the following. Many other experiments, not listed here, are hosted at CERN, using beams from previous stages of the accelerator system or looking for specific signatures after the beam dumps.

- **ALICE**: A Large Ion Collider Experiment [61], is an experiment dedicated to the study of heavy ions collisions. The detector covers the $4\pi$ and is designed to operate at a very high density of tracks. Its main purpose is the study of a particular state of matter, called

Quark-Gluon Plasma, characterised by temperatures and densities comparable to the ones of the early stages of the universe.

- **ATLAS**: A Toroidal LHC Apparatus [62], is one of the two general purpose detectors. It is a cylindrical detector, constructed with the aim of detecting the decays of the Higgs boson and of beyond the Standard Model particles, at the TeV scale. The name refers to the geometry of the external magnetic field, generated by three sets of toroidal magnets, which, along with the inner solenoidal magnetic field, is used to measure the momentum of the muons in the event with high precision.

- **CMS**: Compact Muon Solenoid [63], is a general purpose detector with the same geometry of ATLAS. It is optimised to reconstruct high energy objects with high precision, with the same purpose as ATLAS to detect the Higgs boson and the discovery of new particles at the TeV scale. The main characteristic of the experiment is the high density of the electromagnetic calorimeter, which enables to have both hadronic and electromagnetic calorimeters inside the solenoidal magnet. An external iron yoke enables to have a return magnetic field outside the calorimeters, in order to increase the accuracy in the momentum measurement of the muons in the event.

- **LHCb**: Large Hadron Collider beauty experiment, it is the experiment dedicated to the study of heavy-flavour physics. It will be extensively discussed in 6.3

- **LHCf**: Large hadron Collider forward [64], it is a special purpose detector, dedicated to the study of cosmic rays. It uses the particles generated at a very small angle, from the diffractive interactions taking place between protons during the bunch crossing, to simulate the showering in cosmic rays. It consists of two detectors placed at the two sides of the ATLAS experiment. Each detector is places 140 m from the interaction point, in order to let the particles diffracted at very small angles with respect to the beam axis, to exit from the beam pipe.

- **TOTEM**: Total Elastic and diffractive cross-section measurement [65], is a special purpose experiment whose aim is very similar to the one of LHCf. It is located near the CMS experiment and it is made of two telescopes that detect the particles produced in the interaction point at a very small angle with respect to the beam line.

- **MoEDAL**: Monopole and Exotics Detector at the LHC [66], is dedicated to the direct search of the magnetic monopole and highly ionizing pseudo-stable massive particles. It is a passive detector consisting of 400 tiles which are installed on the walls of the same cavern hosting the LHCb experiment. The clear signature of monopoles being created in the interaction point would be an aligned set of holes through the layers of the material constituting the tiles.

### 6.1.2  LHC performance

### 6.2   $b$-quarks production in $pp$ collisions

In the $pp$ collision heavy quarks and in particular $b$-quarks can be produced as a result of strong interactions between partons. Since the strong interactions conserve the flavour, $b$-quarks should be produced in pairs. The main production mechanisms, at leading order, are *gluon-gluon fusion*,

$gg \to b\bar{b}$ and *quark-quark fusion*, $qq \to b\bar{b}$, whose Feynman diagrams are reported in Fig. 12. The integrated cross-section for $b\bar{b}$ production in $pp$ collisions is 527 μb at a center of mass energy of 14 TeV [67].

After being produced the two $b$-quarks undergo a process called *hadronization*: the quarks start emitting secondary particles of smaller energy and the process stops only if bound colorless states are produced, called hadrons. The most abundant $b$-hadrons produced in this process are $B$ mesons, but higher mass states such as $B_c$ mesons and $\Lambda_b$ baryons can be produced at the LHC.

Due to the small production threshold for $b\bar{b}$ pair production, compared to the center of mass energy in the parton-parton interaction at LHC, the two quarks are usually produced at small angles with respect to the beam axis. Furthermore the direction of the two $b$-hadrons is highly correlated, being both produced either in the forward or backward region, as shown in Fig. 13.
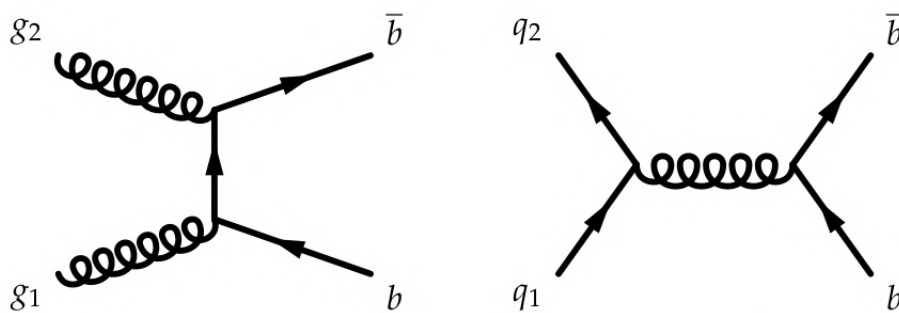


Figure 12: Feynman diagrams for the *gluon-gluon fusion* and *quark-quark fusion* processes.
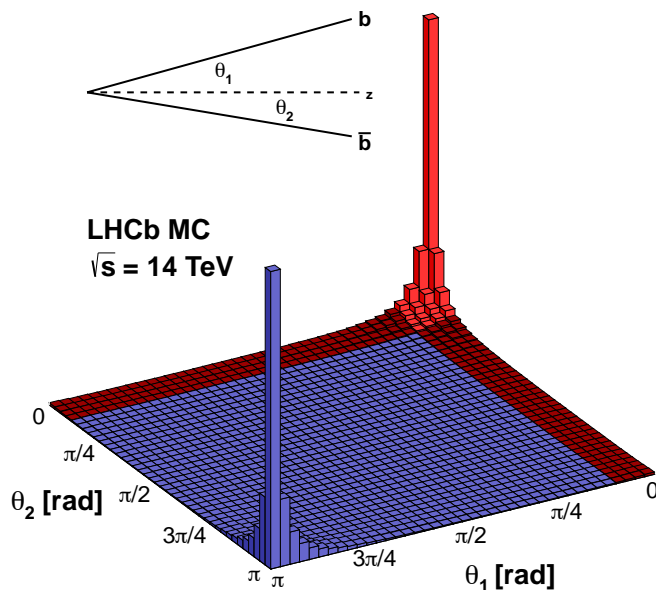
Figure 13: Angular distribution of the $b\bar{b}$ production at LHC. The red part of the distribution corresponds to quark pairs produced inside the LHCb acceptance

## 6.3   The LHCb experiment

The LHCb experiment [68–83] is housed in the same cavern where the DELPHI [84] experiment at LEP was located. It detects particles coming out of the interactions between protons at the interaction point 8 (IP8).

With a luminosity of $2 \times 10^{32}$cm$^{-2}$s$^{-1}$ for LHCb, $10^{12}$ $b\bar{b}$ pairs will be produced in $10^{7}s$, corresponding to the canonical one year of data taking. The LHCb detector must be able to exploit this large number of $b$ hadrons for high quality physics analysis and in order to fulfil its complex and varied physics program.

- An excellent identification of electrons and muons, as well as charged hadrons (pions, kaons and protons) is needed to precisely study and identify the wide variety of $B$ meson decays.

- The momentum of the charged particles must be reconstructed with high precision to be able to have a good invariant mass resolution in order to both distinguish fully reconstructed decays from combinatorics and to distinguish the signal from mesons with masses close to each other, *e.g.* $B$ and $B_{\mathrm{s}}^{0}$ mesons.

- A good vertexing system is needed to identify and separate the $pp$ interaction and $B$ hadron decay vertices. This is crucial to both efficiently identify heavy hadrons, characterized by a long lifetime, precisely measure their lifetime, and to study time dependent $CP$ violation effects and fastly oscillating systems, *e.g.* $B_{\mathrm{s}}^{0}$-$\bar{B}_{\mathrm{s}}^{0}$

- The pseudorapidity typical of heavy hadrons is characterized by an elevated particle density and therefore by a high detector occupancy. This requires a very efficient trigger that copes with the harsh hadronic environment, efficiently selects the signal and rejects the background, to also reduce the data sample to a suitable size.

In order to fulfill these requirements, the main components of the LHCb detector are organized in three sections: a *tracking and vertexing system*, a *particle identification system* and a *trigger system*. Each of these systems is composed of different subdetectors or subsystems. After a description of the geometry of the detector, the subdetectors employed during the Run1 and Run2 data taking will be described in detail in the following sections.

### 6.3.1 Geometry of the detector

Due to the geometry of the $b$-hadrons production, described in the previous sections, the LHCb detector is a single arm, forward spectrometer with an acceptance very close to the beam axis, from 10 mrad to 300 (250) mrad in the horizontal (vertical) plane. The difference in the angular acceptance in the horizontal and vertical plane is due to the presence of a dipolar magnet at the center of the experiment, whose magnetic field is aligned with the vertical axis: when charged particles coming out of the interaction point traverse this magnetic field, they are bent in the direction of the horizontal plane.

The pseudorapidity [1] range for particles falling in the LHCb acceptance is restricted from 1.8 to 4.9, which is enough to capture 41% of $b\bar{b}$ pairs produced in the collisions.

The layout of the detector is shown in Fig. 14. The right-handed orthogonal coordinate system is chosen such that the $z$ axis corresponds to the beam pipe axis, the $y$ axis corresponds to vertical upwards direction and $x$ is the horizontal direction, pointing towards the center of the LHC ring.

Another commonly used coordinate system exploits the $(r, \phi, \eta)$ variables, where r is the distance from the beam axis in the transverse ($x$-$y$) plane, $\phi$ is the azimuthal angle in the transverse plane and $\eta$ is the pseudorapidity,

---

[1]The pseudorapidity is defined as

$$\eta = \ln\left(\tan\frac{\theta}{2}\right) \tag{137}$$

where $\theta$ is the angle between the particle flight direction and the beam axis. It is related to the rapidity, defined as

$$\eta = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{138}$$

where E and p are the energy and momentum of the particle and the $z$ axis has been aligned with the beam direction. The pseudorapidity is the limit of the rapidity at high energies.
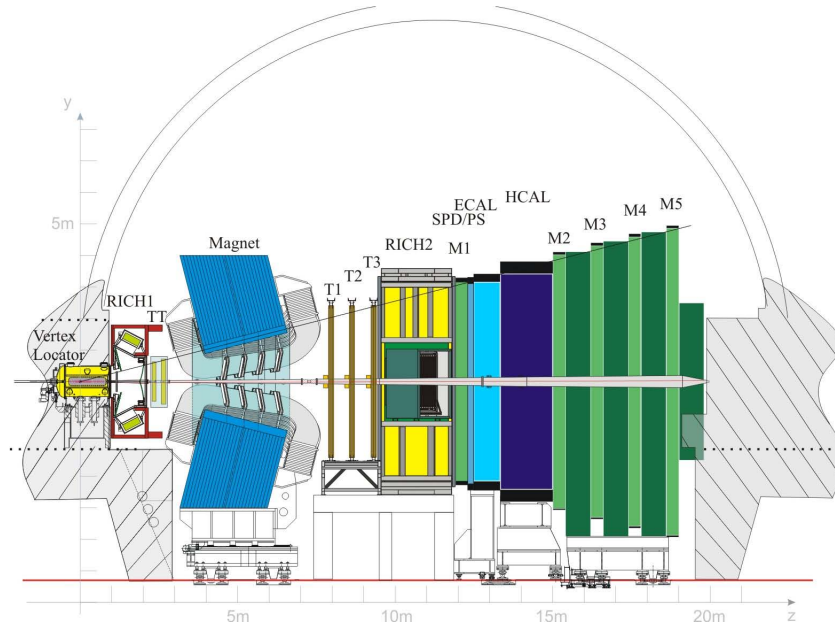
Figure 14: A section of the LHCb detector.

### 6.3.2 The track reconstruction system

The track reconstruction system [82] is designed to determine the trajectory of charged particles (tracks) traversing the detector material. It consists of the Vertex Locator (VELO), the Trigger Tracker (TT) and three tracking stations (T1-T3). Everything is complemented with a dipolar magnet, which bends the charged particles trajectory, enabling the measurement of their momentum. All the subdetectors need to have a high spacial resolution and a low material budget and is constructed into two halves, separated by the beam pipe, that can be opened when intervention is needed.

**The Magnet**   A dipole magnet [68] is used to bend the path of charged particles in order to measure their momentum. It is a warm magnet constructed using two saddle-shaped coils in a window-frame yoke placed mirror-symmetrically to each other, as shown in Fig. 15.

It is located after the TT detector, just before the first tracking station (T1), about 5 m from the interaction point.

The design of the magnet, with an integrated magnetic field of 4 Tm for tracks of 10 m length, has to accomodate the contrasting needs for a field level inside the RICHs envelope less than 2 mT and a field as high as possible in the regions between the Vertex Locator and the Tracking Stations.

The main component of the magnetic field is aligned with the $y$ axis and a scan along the $z$ direction is shown in Fig. 16. The particles are therefore bent in the $x$-$y$ plane and the polarity can be inverted to reduce systematic errors due to detector asymmetries that can limit the precision of $CP$ asymmetry measurements.

The bending of the magnet can be approximated as a single kick at the center of the magnet as shown in Fig. 17. The momentum of the particle is inversely proportional to the difference of the track slope as measured in the tracking subdetectors placed before and after the magnet
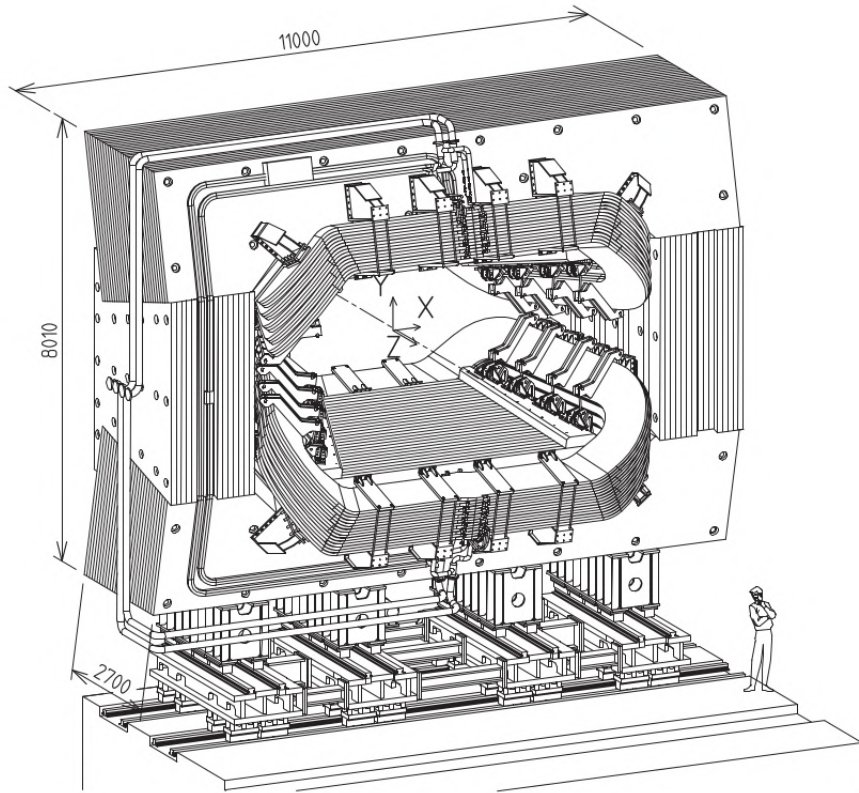
Figure 15: Perspective view of the dipole magnet. The interaction point is located upstream.
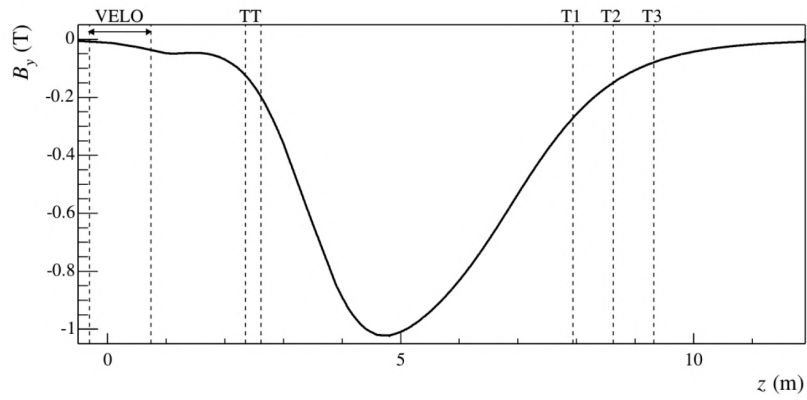


Figure 16: Scan of the magnetic field along the beam axis. The integrated value of the field is 4 Tm.

(*momentum kick* method).
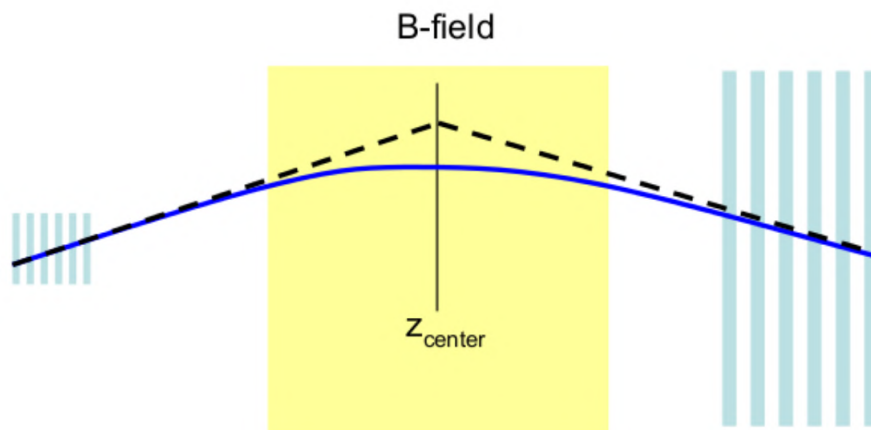
Figure 17: Schematic representation of the *momentum kick method*. The bending of the track can be represented as a single kick at the center of the dipole magnet.

**The Vertex Locator** Vertex reconstruction is a fundamental requirement for the LHCb experiment. Displaced secondary vertices are a distinctive feature of heavy hadron decays that show a typically long lifetime. The Vertex Locator (VELO) [72, 80] is the subdetector closest to the beam pipe and provides precise measurements of track coordinates in the proximity of the interaction point. These are used to reconstruct production and decay vertices of beauty- and charm-hadrons, to provide accurate measurement of their lifetimes, and to measure the impact parameter of particles with respect to production and decay vertices.

The VELO consists of a series of silicon modules arranged along the beam direction as shown in Fig. 18. Each module is composed by two half moon shaped sensors, providing the measurement of the $r$ and $\phi$ coordinates respectively. In order to avoid high levels of radiation damage during beam injection into the LHC, the sensors are retractable: the radial distance of the sensors from the beam in the fully closed position is around 8 mm. Two of the modules are placed upstream of the nominal interaction point and are used as pile-up (PU) veto sensors at the first stage (Level0, L0) of the trigger.

The detectors are mounted in a vessel that mantains vacuum around the sensors and is separated from the machine vacuum by a thin corrugated aluminium foil.

The sensors are made by 300 μm thick silicon microstrips and their conceptual layout is shown in Fig. 19. For the $r$-sensors the strips are concentric semicircles with their centre at the nominal LHC beam position. The minimum pitch at the innermost radius is 38 μm, increasing linearly to 101.6 μm at the outer radius of 41.9 μm. The $\phi$-sensor is designed to read out the orthogonal coordinate to the $r$-sensor. It is subdivided into two regions, inner and outer. The outer region starts at a radius of 17.25 μm and its pitch is set to be roughly half (39.3 μm) that of the inner region (78.3 μm), which ends at the same radius. The pitch at the innermost radius is 38 μm.

The design of the strips in the $\phi$-sensor is complicated by the introduction of a skew to improve pattern recognition. At 8 mm from the beam the inner strips have an angle of approximately 20° with respect to the radial direction, whereas the outer strips make an angle of approximately 10 deg at 17 mm. The direction of the skew of the outer part is reversed to the one of the inner part and the modules are placed so that adjacent $\phi$-sensors have the opposite skew with respect to each other.
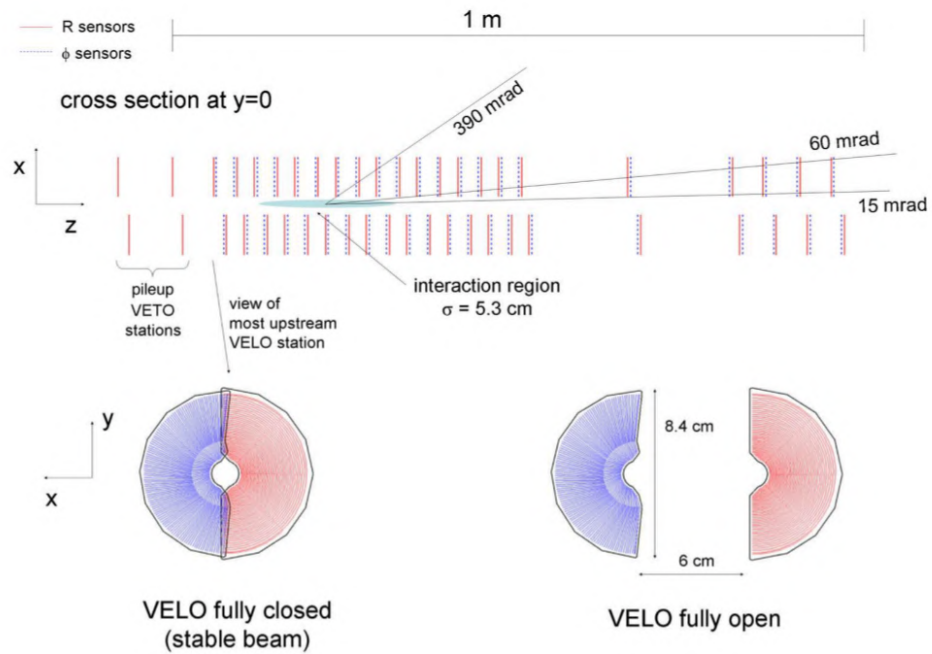
Figure 18: Cross-section in the $(x,z)$ plane of the VELO sensors with the detector in the fully closed position. The two pile-up veto stations are located upstream of the VELO sensors.

All tracks inside the LHCb acceptance pass through at least three modules.

The VELO reaches the best spatial resolution of about 4 μm on the PV, which represents the best vertex detector resolution achieved at the LHC.
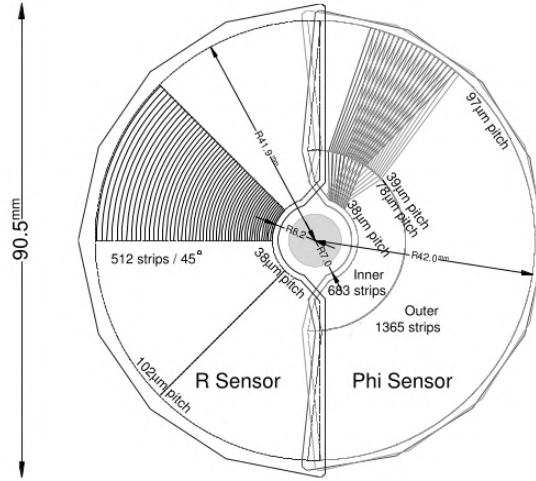
Figure 19: Illustration of the $r\phi$ geometry of the VELO sensors. In the $\phi$-sensors, the strips on two adjacent modules are indicated to highlight the stereo angle.

**The Tracker Turicensis**  The Tracker Turicensis (TT), also known as Trigger Tracker, is a silicon microstrip detector placed before the dipole magnets. It is placed approximately 2.4 m from the interaction point and it is made of four rectangular layers 150 cm wide and 120 cm tall. The strips are vertically oriented, in order to achieve a better spatial resolution in the horizontal plane, which is the bending plane for the dipole magnet. Each strip has a pitch of 183 μm providing a single hit resolution of around 50 μm. The layers are organized in a particular configuration named $(x - u - v - x)$ geometry, in which the strips of the two inner layers are tilted by $+5°$ ($u$) and $-5°$ ($v$). This configuration enables to reconstruct the hits in three dimensions through a stereo view. The geometry of the four layers is shown in Fig. 20. Each sensor is 9.46 cm wide and 9.44 cm long. The sensors are grouped in readout sectors containing one or two sensors, in the innermost part of the detector, closest to the beamline, and three or four sensors in the outer region.

**The tracking stations**  Te tracking stations, T1-T3, are placed downstream the dipole magnet, before RICH2. They are characterized by two detector technologies:

- *The Inner Tracker:* [75] the innermost part of the detector, closest to the beam pipe, uses silicon microstrip sensors.

- *The Outer Tracker:* [73] the outer part of the detector, uses straw-tube drift chambers.

The Outer Tracker (OT) consists of three stations, each of them consisting of four modules in a $(x - u - v - x)$ geometry. Each module contains two staggered layers of drift-tubes, as shown in Fig 21, with inner diameters of 4.9 mm. The tubes are wound from two layers of foil material. An inner carbon-doped Kapton (Kapton XC) wire acts as a cathode for the collection of the positive ions. The outer layer, made of polymide-aluminium laminate, provides shielding

(a) First x layer          (b) u layer

(c) v layer          (d) Second x layer
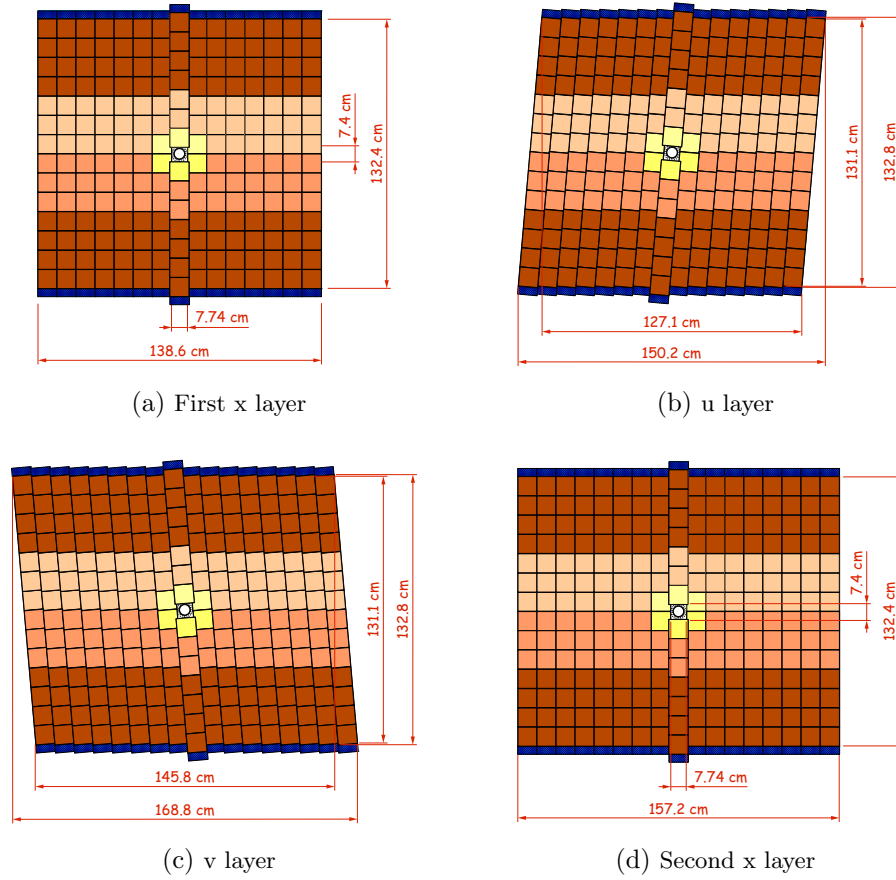
Figure 20: Layout of the TT detection layers. Different colours represent different readout sectors while the blue edge indicates the readout electronics.

and together with the anode wire forms a transmission line for the effective transport of the high-frequency signals. As a counting gas, a mixture of Argon (70%) and $CO_2$ (30%) is chosen in order to guarantee a fast drift time (below 50 ns). The pitch between each tube is 5.25 and the spatial resolution is about 200 $\mu$m.

The Inner Tracker (IT) consists of three stations, each one containing three layers arranged in a $(x - u - v - z)$ geometry. The sensors are read out in sectors containing one or two sensors, as shown in Fig 22. Each sensor is 7.6 cm wide and 11 cm tall, with a thickness of 320 $\mu$m or 410 $\mu$m. The pitch of the microstrips is about 198 $\mu$m, which allows for a single hit resolution comparable to the one of the TT.
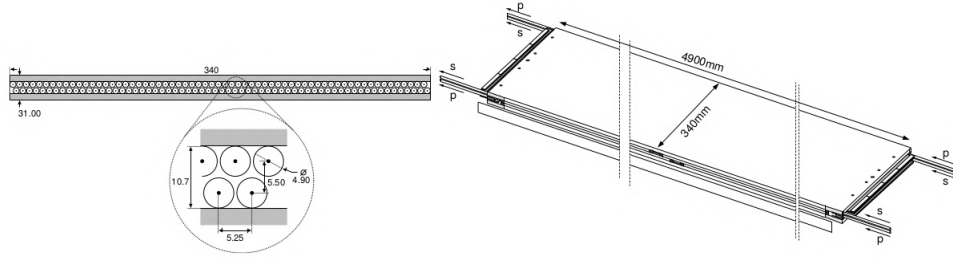
Figure 21: Layout of the OT sub-detector. On the left the configuration of the layers is shown. On the right, the cross-section of a single module, showing the straw tubes structure, is shown.
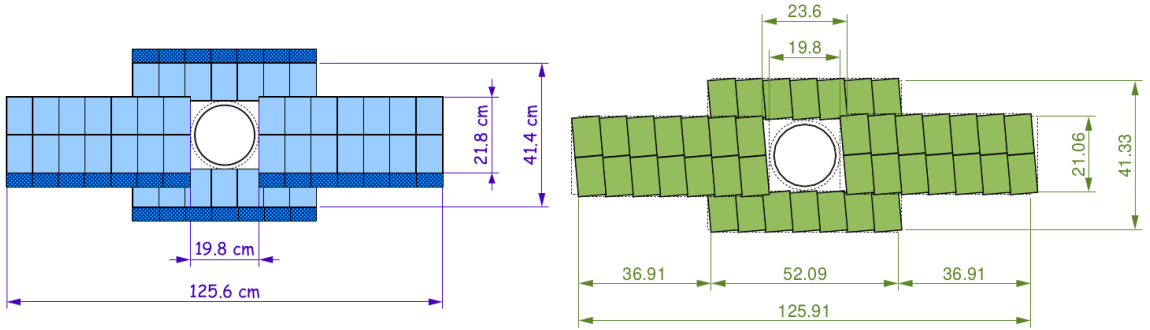


Figure 22: Layout of the $x$ (left) and $u$ layer of the Inner Tracker. On the left the sensors are represented in light blue, while the readout electronics are shown in dark blue.

### 6.3.3 The Particle Identification System

Particle identification (PID)is a fundamental requirement for the LHCb experiment. The ability to distinguish between pions and kaons is essential for the study of *CP*-violation in selected *B* hadron decays. The system [81] consists of: two Ring Imaging Cherenkov (RICH) detectors, to distinguish between charged tracks; two calorimeters, (Electromagnetic, ECAL, and Hadronic, HCAL) that complement the RICH detectors in the recognition of electrons and charged hadrons, and enable the recognition of photons and neutral hadrons; and the Muon Stations (M1-M5), that perform an efficient muon recognition.

**The Ring Imaging Cherenkov detectors**   Rich Imaging Cherenkov (RICH) detectors [70] exploit the Cherenkov effect: when a charged particle travels in a medium with a speed that exceeds the speed of light $c' = c/n$, where $n$ is the refraction index of the medium. In this case photons are emitted from the medium in a cone aligned with the direction of the track and whose opening angle is correlated with the speed of the particle and the refraction index by the following expression:

$$\cos\theta_C = \frac{1}{n\beta} = \frac{1}{n \cdot v/c}. \tag{139}$$

By combining the information on the momentum of the particle, measured by the tracking system, it is possible to estimate the mass of the particle, as shown by the following expression:

$$\cos\theta_C = \frac{1}{n}\sqrt{\left(\frac{m}{p}\right)^2 + 1} \tag{140}$$

When having particles with a wide range of momenta, it is crucial to employ different radiating material. As it is shown in Fig 23, in fact, the Cherenkov angle saturates at the value of $\theta_C = \arccos(1/n)$: having different radiating materials is therefore fundamental to cover different momentum ranges.
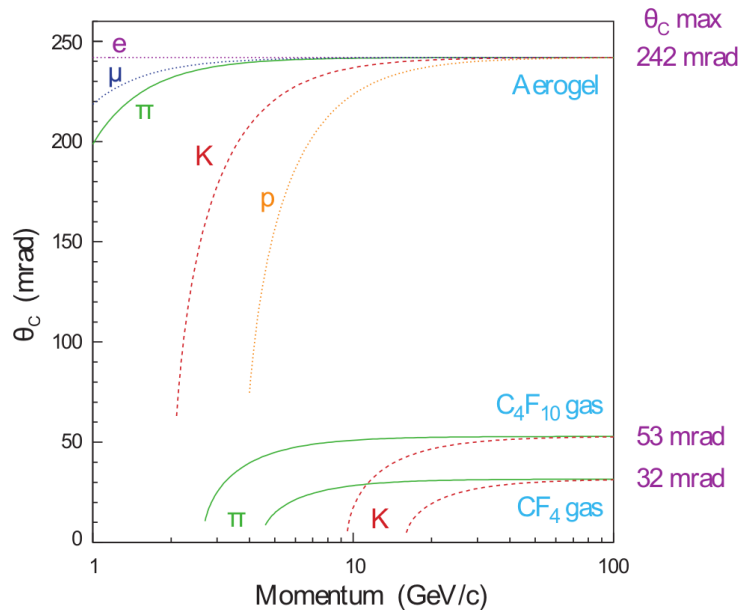


Figure 23: Cherenkov angle as a function of the particle momentum for the various RICH radiators.

The momentum spectrum of particles produced at large polar angles is softer while at small polar angles the is harder; therefore, two Ring Imaging Cherenkov (RICH) detectors have been placed at LHCb, on to cover the full momentum range, one upstream and one downstream the magnet. The upstream detector, RICH1, covers the low momentum range (from 1 to 60 GeV/$c$) and has a wide acceptance to cover the full LHCb acceptance. It is located after the VELO and before the magnet, to detect the low momentum particles before they are swept out of the LHCb acceptance by the magnetic field. During RunI, two radiators were employed: an aerogel tile ($n = 1.03$) and a gas of $C_4F_{10}$ ($n = 1.0014$). During RunII the aerogel was removed. The downstream detector, RICH2, covers the high momentum range from $15\,\mathrm{GeV}/c$ up to and beyond $100\,\mathrm{GeV}/c$ using a $CF_4$ ($n = 1.005$) radiator. It is located downstream of the magnet and has a limited angular acceptance, from $\pm15$ mrad to $\pm120$ mrad (horizontal plane) and $\pm100$ mrad (vertical plane). A sketch view of the two RICH detectors is shown in Fig. 24.

In both the sub-detectors the focusing of the Cherenkov light is accomplished using a combination of spherical and flat mirrors to reflect the image out of the spectrometer acceptance. In the RICH1 the mirrors are made of carbon fiber, a light material in order to reduce the

(a) Side view of RICH1        (b) Top view of RICH2
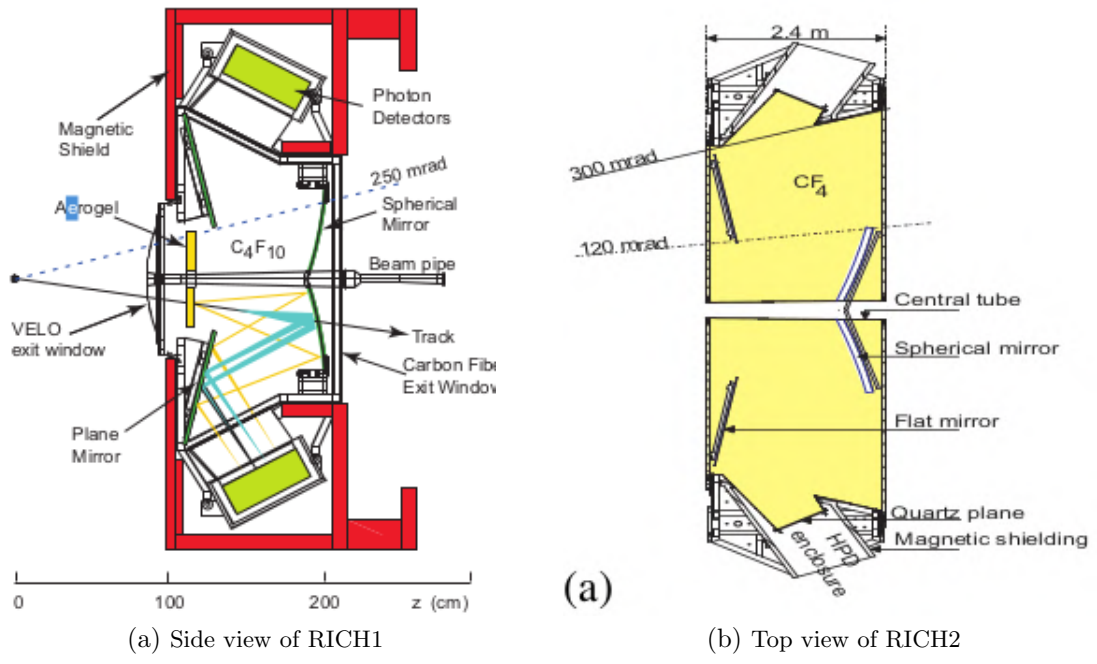
Figure 24: Schematic view of the RICH detectors.

multiple scattering of the particles before the measurement of their momentum. In the RICH2 the mirrors are made of hexagonal shaped glass tiles. To measure the spatial position of the emitted Cherenkov light, photons are detected by arrays of Hybrid Photon Detectors (HPDs). The HPD is a photon detector in which a photoelectron, released from the conversion in a photocatode of an incident photon, is accelerated by an applied high voltage on to a Silicon pixels array with a pitch area of $500\,\mu\mathrm{m} \times 500\,\mu\mathrm{m}$, as sketched in Fig. 25.
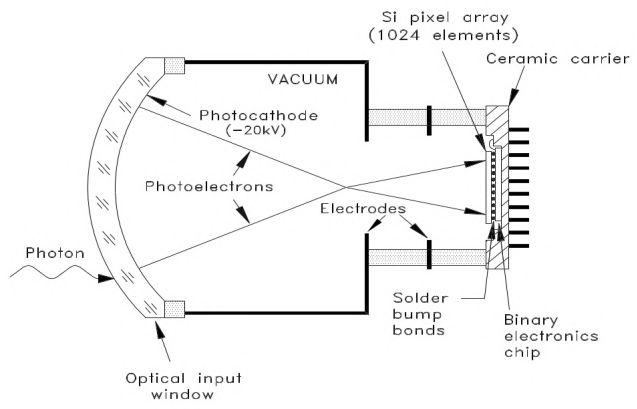


Figure 25: Schematic of a single HPD sensor.

**The Calorimeters**  The calorimeter system [69] provides measurement of the energy of electrons, photons and hadrons and provides their identification as well as the measurement of their position. The reconstruction with good accuracy of $\pi^0$ that decay into two photons is essential for flavour tagging and for the study of $B$-meson decays. Fast measurements made by the calorimeters are also fundamental for taking decisions at the first stage of the trigger (L0) to select interesting events to be further analyzed.

The LHCb calorimeter system consists of several layers. In order of distance from the interaction point:

- the Scintillating Pad Detector (SPD)

- the Pre-Shower Detector (PS)

- the Electromagnetic Calorimeter (ECAL)

- the Hadron Calorimeter (HCAL)

Since the hit density varies by two orders of magnitude over the calorimeter surface, the calorimeters are segmented in the $x - y$ plane such that the channel density is higher towards the beam pipe. A lateral segmentation into three different sections has been chosen for the ECAL, as shown in Fig. 26, and projectively for the SPD/PS. Given the dimensions of the hadronic showers, the HCAL is segmented into two zones with larger cell sizes.
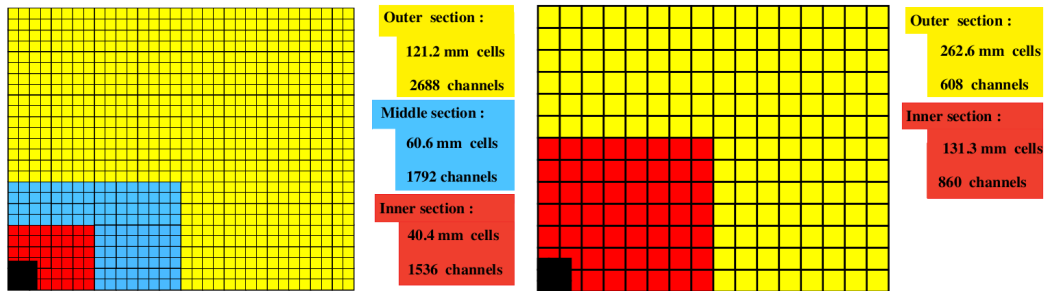


Figure 26: Lateral segmentation of the SPD/PS and ECAL (left) and the HCAL (right).

All calorimeters follow the same principle: the scintillation light emitted by the passage of charged particles in the active part of the detector is transmitted to a Photon Multiplier (PMT) by wavelenght-shifting (WLS) fibers. The single fibers for the SPD/PS are read out using multianode photomultipliers tubes (MAPMT), while the fiber bunches in the ECAL and HCAL modules are read out by individual phototubes.

The SPD and the PS are placed upstream the ECAL and are divided by a 12 mm thick lead absorber. The detector elements are 15 mm thick scintillation pads, of about 4 cm × 4 cm in the inner region and 12 cm × 12 cm in the outer region. The scintillation light is collected by a helicoidal WLS fiber and is sent to photomultipliers that are located above or below the detector. The SPD is mainly used at the trigger level (L0) in association with the ECAL to separate of electrons from photons. This is possible due to the fact that electrons, being charged, produce light in the SPD, while photons don't. Similarly the PS detector is used to separate photons

from neutral pions both at the trigger and offline level. In depth the two detectors correspond to a 2.5 $X_0$ thickness.

The energy and position of the electromagnetic showers initiated in the SPD/PS is measured by ECAL. The shashlik technology has been chosen in order to ensure a good energy resolution and a variable transverse granularity while keeping a fast response time and a good radiation hardness. The detector is constructed alternating 2 mm lead sheets with 4 mm thick scintillator plates. In depth the 66 lead converters correspond to 25 $X_0$ thickness, with a Moliere radius of 3.5 cm. The cells' size varies from $4 \times 4$ cm in the inner part of the detector, to $6 \times 6$ cm and $12 \times 12$ cm in the middle an outer parts. The overall detector dimensions are $7.76 \times 6.30$ m covering an acceptance of 25 mrad $< \theta <$ 300 mrad in the horizontal plane and 25 mrad $< \theta <$ 250 mrad in the vertical plane. Light is collected by WLS fibers and detected by PMT's. The energy resolution of ECAL is given by

$$\frac{\sigma_E}{E} = \frac{8.5 - 9.5\%}{\sqrt{E}} \oplus 0.8\%.$$

The parametrization $\sigma_E/E = a/\sqrt{E} \oplus b \oplus c/E$ (E in GeV) is used, where a, b and c stand for the stochastic, constant and noise terms respectively.

The HCAL is positioned downstream the ECAL and measures the energy and the position of hadronic showers. HCAL is made of a total of 6 layers, alternating 16 mm iron absorber layers with 4 mm thick scintillation tiles, readout by WLS fibers and PMTs. the scintillator and iron plates being parallel to the beam. In the lateral direction the tiles are interspersed with 1 cm of iron (corresponding to 1 $X_0$), whereas in the longitudinal direction the length of tiles and iron spacers correspond to the hadron interaction length $\lambda_I$ in steel. The segmentation is similar to the one of ECAL, with modules of size 13cm $\times$ 13cm and 26cm $\times$ 26cm in the inner and outer region, respectively. The total length of 6 layers corresponds to 5.6 $\lambda_I$. The energy resolution of the HCAL modules is given by

$$\frac{\sigma_E}{E} = \frac{69\%}{\sqrt{E}} \oplus 9\%.$$

**Muon stations**  Muons are the only particles (except neutrinos) able to pass through the calorimetrs without losing all their energy. They play a major role in many aspects of the LHCb physics program: they are present in the final states of many $CP$-sensitive $B$ decays and they are very important for $CP$ asymmetry and oscillations, since muons from semileptonic $b$-decays provide a tag of the initial state flavour of the accompanying neutral $B$-meson. In addition, the study of rare $B$ decays such as $B_s^0 \to \mu^+\mu^-$, may reveal new physics beyond the Standard Model, and muons play a major role in semileptonic decays such as the one at study in this thesis. Therefore, muon triggering and offline muon identification are fundamental requirements of the LHCb experiment.

The muon system [71] provides fast information for the high-$p_T$ muon trigger at the earliest level (L0) and muon identification for the high-level trigger (HLT) and offline analysis. It is composed by five stations (M1-M5) of rectangular shape, as shown in Fig. 27. Station M1 is placed in front of the calorimeters and it is used to improve the $p_T$ measurement in the trigger. Stations M2 to M5 are placed downstream the calorimeters and are interleaved with iron absorbers 80 cm thick to select penetrating muons. The total absorber thickness, including the calorimeters, is approximately 20 interaction lengths and the minimum momentum of a muon to

cross the five stations is approximately 6 GeV/$c$. The geometry of the five stations is projective, with an acceptance coverage of 20 (16) mrad and 306 (258) mrad in the bending (non-bending) plane respectively.
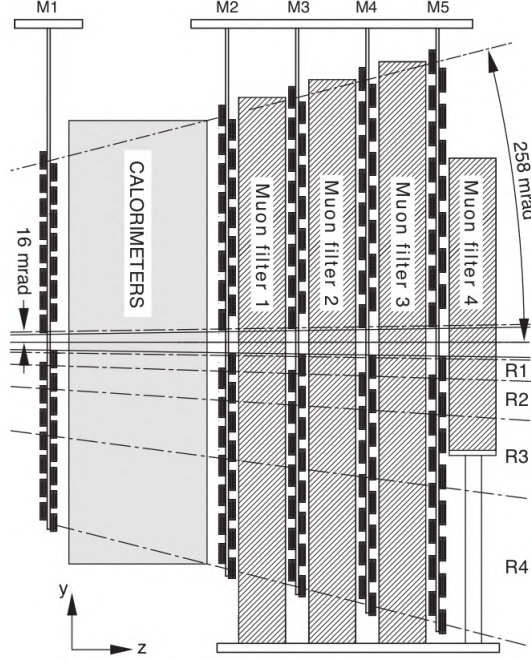


Figure 27: Side view of the muon system.

The layout of the muon stations is shown in Fig. 28. Each muon station is divided into four regions, R1 to R4 with increasing distance from the beam axis with a segmentation scale in the ratio 1:2:4:8. With this geometry the particle flux is expected to be roughly the same over the four regions of a given section.

Multi-Wire proportional chambers (MWPC) are used for all regions except the inner region of station M1 where the expected particle rate exceeds safety limits for ageing. In this region triple-GEM detectors are used.

In total, the muon system comprises of 1368 MWPC. They are made by wire planes of 2 mm spacing, symmetrically placed in a 5 mm gas gap containing a gas mixture of $Ar/CO_2/CF_4$. An exploded schematic is shown in Fig. 29. Each chamber is made of four equal gas gaps superimposed and rigidly stacked together.

In the innermost region R1 of the station M1 each of the 12 chambers consists of two triple-GEM detectors superimposed and forming two sensitive layers. The triple-GEM detector consists of three gas electron multiplier (GEM) foils sandwiched between anode and cathode planes. A cross-section of the detector, showing the different elements and their physical dimensions, is shown in Fig. 30. The GEM foils, 50 $\mu$m thick Kapton with two-sided Cu cladding of 5 $\mu$m, are punctuated with bi-conical holes of 70 $\mu$m in diameter. The hole-hole pitch is 140 $\mu$m and the gas gaps are filled with a $Ar/CO_2/CF_4$ gas mixture.
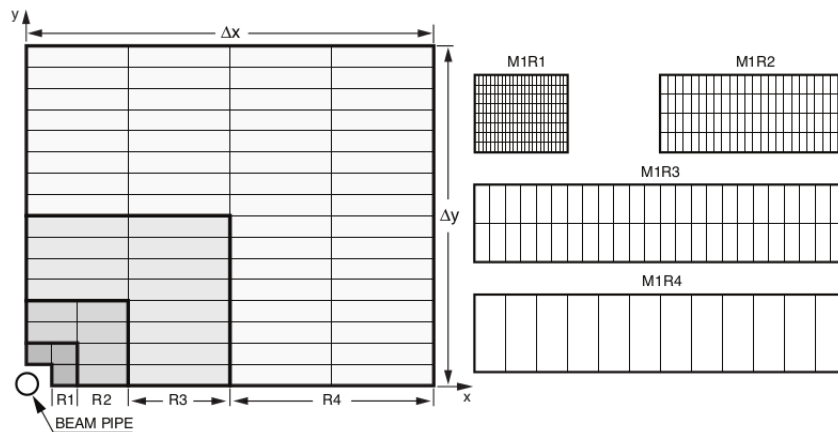
Figure 28: Left: front view of a quadrant of a muon station. Each rectangle represents one chamber. Right: division into logical pads of four chambers belonging to station M1.
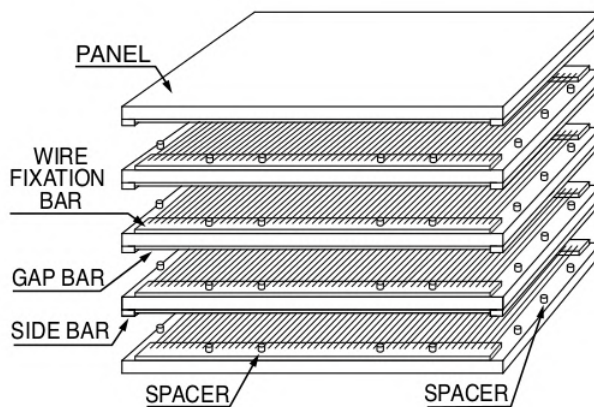


Figure 29: Exploded schematic view of a chamber showing the various elements of a Multi Wire Proportional Chamber.
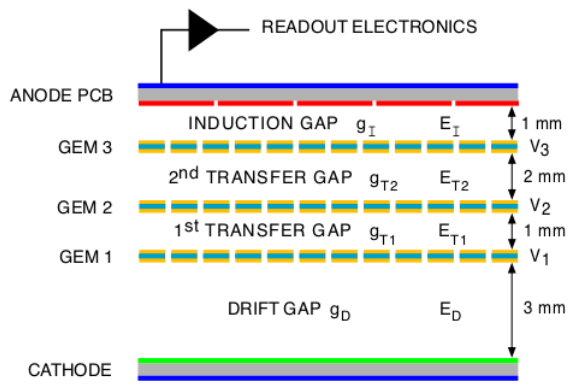
Figure 30: Exploded schematic view of a chamber showing the various elements of a Gas Electron Multiplier chamber.

### 6.3.4 Trigger and reconstruction at the LHCb experiment

The LHC collider is designed to work at a bunch crossing frequency of 40 MHz. In order to avoid filling the available disk size while still keeping as many interesting events as possible for further analysis. A fast and reliable trigger system [77, 83] is needed to quickly and efficiently select and store interesting events. The LHCb trigger system has the goal to select the events at a rate of about 5 kHz during RunI and of about 12.5 kHz during RunII, that correspond to an output throughput of 0.6 GB/s to disk.

It is divided into three levels: the first one (L0), operating at the bunch crossing frequency, is implemented in hardware, while the second and the third level ones (HLT1 and HLT2) implemented as software applications. In RunII LHCb has introduced a novel real-time detector alignment and calibration, working on events selected by the second and third level trigger. Data collected at the start of the fill are processed in a few minutes and used to update the alignment, while the calibration is performed at each run. A sketch of the trigger system is shown in Fig. 31.
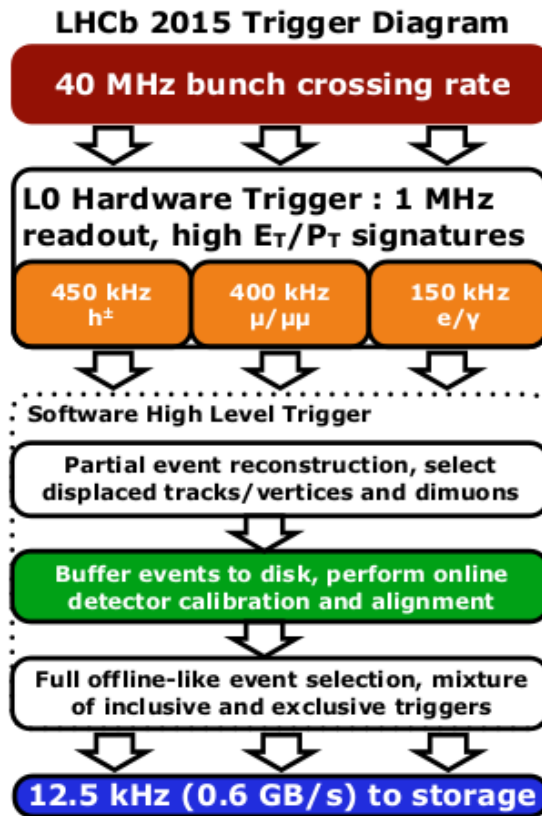


Figure 31: Trigger layout in RunII illustrating the central place of the detector alignment and calibration.

**L0 Trigger**  The Level 0 (L0) trigger is implemented in hardware and has to give a fast response, synchronous to the LHC frequency. Due to the stringent timing requirements, the L0 trigger has to use a limited amount of information coming from the detector. It reduces the rate to 1 MHz,

using only calorimeter and muon chambers information and a few information from the VELO. The trigger selections are divided into three categories, each using a different set of subdetectors.

The LOPileUp trigger uses information coming from the pile-up detector in the VELO. This system estimates the number of primary vertices in each bunch crossing and offers the possibility to veto events with multiple interactions. The pile-up system consists of two planes, (A and B), made up of two overlapping sensors perpendicular to the beam-line and upstream of the VELO. In both planes the radii of track hits, $r_a$ and $r_b$ are measured. Under the assumption that a track is coming from the beam line, from two hits coming from the same track it is possible to estimate the position of the origin vertex along the $z$-axis using the following formula:

$$z_v = \frac{k(z_a - z_b)}{k - 1}, \tag{141}$$

where $k = r_b/r_a$ and $z_a$ and $z_b$ are the detector positions along the beam line. The LOPileUp system combines all the hits in the same octant of both planes and evaluates the $z_v$ position of all the possible combinations according to Eq. 141. As shown in Fig. 32, the $z_v$ values are put into a histogram onto which a peak search is performed. All the hits contributing to the highest peak in the histogram are then masked, after which a second peak is searched for. The height of the second peak is a measure of the number of tracks coming from a second vertex and can be used to distinguish events with multiple interactions. If an event exceeds a threshold in this multiplicity measure, it is vetoed.
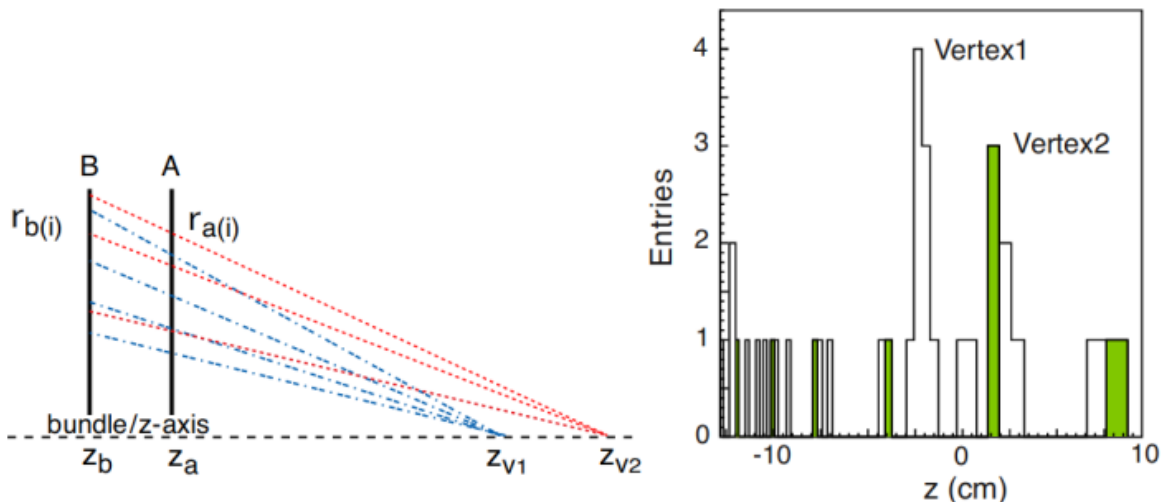


Figure 32: The basic principle of detecting vertices in an event. The hits of plane A and B are combined in a coincidence matrix. All coincidences are projected onto a $z_v-$histogram. After the first vertex finding iteration, the hits corresponding to the highest bins are masked, resulting in the green shaded histogram.

The other L0 triggers exploit the high-$p_T$ signature of $B$ decays that decay in high transverse momentum ($p_T$) and high transverse energy ($E_T$) particles, due to their high mass.

The energies deposited in the SPD, PS ECAL and HCAL are used in the LOCalorimeter systems to trigger the selection of events. The LOCalorimeter triggers reconstruct the hadron, electron and photon with the highest transverse energy ($E_T$) by summing up the energy deposits

in $2 \times 2$ clusters of the calorimeters. The transverse energy is evaluated assuming the particles come from the nominal proton-proton interaction region and are evaluated as

$$E_T = \sum_{i=1}^{4} E_i \sin \theta_i, \qquad (142)$$

where $E_i$ is the energy deposited in cell $i$ and $\theta_i$ is the angle between the $z-$axis and the line from the cell center to the average proton-proton interaction point. The deposits are associated to a photon if they are only present in ECAL, to an electron if associated hits in the SPD are also found. Lastly they are associated to a hadron if both ECAL and HCAL hits (and SPD for charged hadrons) are present. The event is then triggered by the L0Hadron, L0Photon or L0Electron algorithms if the relative particle's $E_T$ exceeds a certain threshold. As some subdetectors also read out hits associated to other bunch crossings (spillover), event with a large $E_T$ ($> 24$ GeV) are also rejected in most of the L0 lines.

The L0Muon trigger tries to reconstruct the two muon tracks with the highest $p_T$ for each of the 4 muon quadrants, using hits coming from all the M1-M5 stations. A dedicated reconstruction algorithm reconstructs the tracks assuming that they form a straight line in the muon stations, they come from the nominal proton-proton interaction region and they get a single kink from the magnetic field. The transverse momentum of the track is measured with a momentum resolution of about 20%, using the position of the track in the first two stations. The L0Muon trigger returns a positive decision if the $p_T$ of the highest track exceeds a given threshold, the L0DiMuon trigger when the product of the $p_T$ of the two tracks does.

The information from all the algorithms is then collected by the L0 Decision Unit and are logically OR-ed to obtain the L0 decision. An event that passes either one of the trigger lines is said to pass the L0Global trigger decision.

The thresholds for the different trigger lines for each year of RunII data taking used in this thesis are reported in Table 9.

| L0 trigger | $E_T/p_T$ threshold | | SPD threshold |
|---|---|---|---|
| | 2015 | 2016 | |
| Hadron | $> 3.6$ GeV | $> 3.7$ GeV | $< 450$ |
| Photon | $> 2.7$ GeV | $> 2.78$ GeV | $< 450$ |
| Electron | $> 2.7$ GeV | $> 2.4$ GeV | $< 450$ |
| Muon | $> 2.8$ GeV | $> 1.8$ GeV | $< 450$ |
| Muon high $p_T$ | $> 6.0$ GeV | $> 6.0$ GeV | - |
| Dimuon | $> 1.69$ GeV$^2$ | $> 2.25$ GeV$^2$ | $< 900$ |

Table 9: The L0 thresholds for the different trigger lines used to take the majority of the data for each indicated year.

**High Level Trigger**   The High Level Trigger (HLT) analyses the events that are selected by the L0 trigger, implemented as two levels, written as C++ applications, which run on every CPU of the Event Filter Farm. The HLT is based on the same software framework used throughout the LHCb offline reconstruction software. Given the available resources in the EFF, the available time per event is much smaller with respect to the one available in the offline processing.

The first level (HLT1), performs an inclusive selection of events performing a partial reconstruction, reducing the 1 MHz of the L0 output to a rate of events of 150 kHz, which are further processed by the HLT2. This trigger level exploits the features of a heavy hadron decay, which features displaced vertices, by selecting tracks or combinations of tracks with high $p_T$ and impact parameter with respect to the primary vertex. Events selected by an HLT1 trigger line are buffered to the local hard-disk of each node. This is done for two purposes: events can be processed during the inter-fill periods, and the detector can be calibrated and aligned run-by-run before the HLT2 stage.

Once the detector is aligned and calibrated, events are passed to the second level of HLT (HLT2), where a full event reconstruction is performed and the event rate is further reduced to 12.5 kHz. Different trigger selections (or lines) are available in HLT, each one specifically developed to select a particular decay or decay topology.

In RunII, the increased EFF and improvements of the software allowed to run in the HLT a reconstruction with almost the same quality as the one that can be reached during the offline reconstruction.

**TIS, TOS, TOB definition**   During offline analysis, events can be classified in categories, according to the portion of the signal under study being responsible for having triggered the event [85].

- Triggered On Signal (TOS): events for which the presence of the signal is sufficient to generate a positive trigger decision.

- Triggered Independent of Signal (TIS): events for which the rest of the event is sufficient to generate a positive trigger decision, where the rest of the event is defined through an operational procedure consisting in removing the signal and all detector hits belonging to it.

- Triggered On Both (TOB): these are events that are neither TIS nor TOS; neither the presence of the signal alone nor the rest of the event alone are sufficient to generate a positive trigger decision, but rather both are necessary.

Note that a single event can be simultaneously TIS and TOS if the presence of the signal alone as well as the rest of the event alone are sufficient to generate a positive trigger decision. TOB events, on the other hand, can be neither TIS nor TOS.

In order to classify an event in one of the categories, a comparison between the online object that has triggered the event and the offline reconstructed objects that define the signal need to be performed. Every single sub-detector in LHCb has a LHCbID which is unique across the whole detector. Physics objects, such as tracks, can be defined as the sets of LHCbIDs of the subdetectors that have been used to construct them (e.g. subdetector parts around the hits used to reconstruct a track or the calorimeter cells used to construct a L0 cluster). When a trigger decision is made, the LHCbID of the subdetector parts that have been used to construct the

trigger object are saved in the raw event. This allows offline objects reconstructed later to be compared with the objects used in the trigger. For decisions taken at the HLT level, the set of LHCbIDs of the detector components for hits associated to the tracks that fired the trigger are saved. For decisions taken at the L0 level, all the charged signal tracks are extrapolated to the calorimeters and the LHCbIDs of the intersected calo elements are saved to the reconstructed object. For ECAL charged objects, photons or merged $\pi^0$s also the LHCbIDs of the ECAL cluster(s) associated to the particle is (are) saved.

An event is considered to be TOS with respect to a trigger selection if the LHCbIDs of object that caused the trigger selection overlap for more than 70% with the LHCbIDs of all the objects defining the offline signal candidate. An event is considered to be TIS with respect to a trigger selection if the event has been triggered, but the LHCbIDs of the object that caused the trigger selection overlap for less than 1% with the LHCbIDs of all the objects defining the offline signal candidate.

**Event reconstruction**   The reconstruction of the charged particle trajectories (tracks), the vertex and particle identifications are the fundamental steps needed in order to have available the high level object on to which physics analysis can be performed.

The tracks are reconstructed from the combination of hits in the tracking sub-detectors, and the vertices are reconstructed as points from which collection of tracks are originating. Different types of tracks can be defined, based on the detector crossed by the track, as shown in Fig. 33.
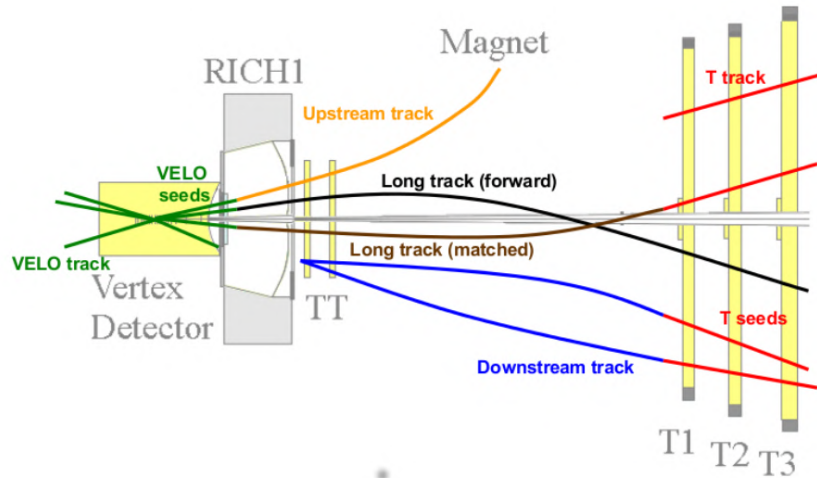


Figure 33: Illustration of the different types of tracks reconstructed in LHCb.

- **VELO tracks:** defined by hits only in the VELO. They are used as the seed for the long and upstream tracks reconstruction. If they cannot be extended out of the velo, they are used to determine the primary vertices.

- **T tracks:** reconstructed with hits in the tracking stations and used as input to the long and downstream tracks reconstruction.

- **Long tracks:** they have hits in all the tracking sub-detectors, and therefore have the best momentum resolution. They are the tracks being used to perform physics analysis.

- **Upstream tracks:** they have hits only in the VELO and the TT stations. They belong to particles with low momentum which are bent out of the LHCb acceptance before reaching the tracking stations.

- **Downstream tracks:** they have hits in the TT and tracking stations. They are usually tracks belonging to decays of long-lived particles, like $\Lambda^0$ or $K_S$, that decay outside the VELO acceptance.

The long tracks reconstruction is performed in different steps [76]. The first step consists into a pattern recognition of the VELO hits, constructing the VELO tracks. The VELO track constitute the seed for the Long tracks. A VELO track should have at least three hits in both the $R$ and $\phi$ sensors.

Vertices are reconstructed from a minimum of five intersecting VELO tracks and the ones within a radius of 300 μm of the mean position of the proton-proton interaction envelope are considered to be primary vertices.

The VELO tracks that have a significant $IP_{\chi^2}$² with respect to all PVs or that are matched to muon hits by the previous step are then passed to two complementary algorithms: the *forward tracking* and the *track matching*. In the *forward tracking* algorithm, hits found around the VELO seed direction are added to the VELO track. In the *track matching* algorithm, track segment in the T station, reconstructed with hits in the T station, are combined with VELO seeds. The track candidates found by each algorithm are then combined and duplicates are removed and hits in the TT consistent with them are added.

Downstream tracks are constructed by extrapolating T tracks through the magnetic field and searching for corresponding hits in the TT. Upstream tracks are found by extrapolating VELO tracks similarly to what is done in the downstream tracking. For these algorithms, at least three hits have to be present in the TT.

Finally a Kalman filter [86] is run on each track candidate, taking into account ionization and scattering effects and to precisely measure its momentum and charge. Clone tracks are then removed, by inspecting tracks which share many hits and retaining the one with most hits. After having evaluated the quality of a track, in the measure of its $\chi^2/ndof$ in the Kalman filter, fake tracks not associated to any particle (named ghosts) are removed by means of a selection on a dedicated neural network [87].

Different level of precision in the reconstruction can be achieved at the different stages of the HLT trigger, with respect to the offline reconstruction, which has the best resolution due to smaller time constraints.

The first stage of the HLT performs a partial event reconstruction. The full 3D reconstruction of the VELO tracks is fast enough to be run online. The only difference from the offline VELO reconstruction algorithm is that unused hits, which are further analyzed offline to increase the reconstruction efficiency for tracks pointing far away from the beam-line, are discarded during the online reconstruction.

In order to reduce the processing time, the forward track search at this stage is performed by using upstream tracks as a seed, with respect to the reconstruction performed at later stages. In RunII, upstream tracks are subject to a transverse momentum cut of 0.4 GeV/$c$ prior to the forward tracking, which is furthermore tightened to 0.5 GeV/$c$ after the forward tracking.

---

²The $IP_{\chi^2}$ of a track or a composite particle with respect to a vertex is the difference in the Vertex $\chi^2$ when fitted with or without that track or object. It is a measure of the pointing, or of the Impact parameter, of the object with respect to the vertex.

The HLT1 output event rate is low enough to allow the forward tracking of all VELO tracks in HLT2. Remaining VELO tracks which were not extended into Long Tracks, are propagated to the T stations, this time without any minimum transverse momentum requirement and without requiring clusters in the TT stations.

**Particle Identification**   Another feature of paramount importance for the LHCb experiment is particle identification (PID), which is the association of a mass hypothesis to charged particles. This is possible thanks to the information coming from the two calorimeters, the two RICH detectors and the muon stations. The charged track direction is extrapolated in the RICH, the calorimeter systems and in the muon chambers. Combining information coming from the Cherenkov angle of the associated ring in the RICH systems, the energy deposits in the calorimeters and associated hits in the muon chambers, a probability for a specific particle hypothesis $(x)$, defined by a likelihood $L_x$ can be associated to a track. Since the value of the likelihood spans over many orders of magnitude, its logarithm is used instead. Since pions are the most abundant particle specie at LHCb, the likelihood for a specific particle hypothesis is evaluated against the $\pi$ hypothesis, using differences of log-likelihoods:

$$DLL_x = \log L_x - \log L_\pi. \tag{143}$$

A second set of PID variables has been developed to improve upon the log likelihood approach. In this case both the log likelihood values from each subdetector, along with additional information, are convoluted into an artificial neural network which provides a single probability for each particle hypothesis, referred as $ProbNN_x$.

# 7    Analysis strategy

Various measurements of the parameters $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ have been performed at both $e^+e^-$ colliders, by the Belle and BaBar collaborations, and at the $pp$ LHC collider, by the LHCb experiment. As discussed in Sect. 2.3.2, a discrepancy between the Standard Model predictions and the average of these measurements has been observed, at a level of $3.1\sigma$.

The LHCb collaboration contributed to the average with two measurements of the $\mathcal{R}(D^*)$ parameter, and no measurement of $\mathcal{R}(D)$ has been performed at a hadron collider so far. These measurements would be able to further reduce the uncertainty on the world average, and are of paramount importance in order to confirm or reject the observed discrepancy.

The aim of this thesis is the measurement of the $\mathcal{R}(D^+)$ with $\overline{B}^0 \to D^+\ell^-\overline{\nu}_\ell$ decays. In this chapter the main strategy and the biggest challenges of the analysis will be described.

## 7.1    Measuring $\mathcal{R}(D^+)$ at LHCb

Thanks to the high $b\overline{b}$ production cross-section and the high branching fraction, of around 1.08%, $\overline{B}^0 \to D^+\tau^-\overline{\nu}_\tau$ decays are copiously present in $pp$ collisions recorded by the LHCb experiment. The measurement reported in this thesis exploits $\overline{B}^0 \to D^+\tau^-\overline{\nu}_\tau$ and $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ decays to measure $\mathcal{R}(D^+)$. The $D^+$ and $\tau$ candidates are reconstructed in their $D^+ \to K^-\pi^+\pi^+$ ($\mathcal{B} = 9.38 \pm 0.16\%$) and $\tau^- \to \mu^-\nu_\tau\overline{\nu}_\mu$ ($\mathcal{B} = 17.39 \pm 0.04\%$) decays.

A nearly irreducible background of $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays, in which neutral particles from the $D^{*+} \to D^+\pi^0/\gamma$ decays are not reconstructed, is expected in the measurement. This background could be reduced with the help of isolation selections on the signal candidates, but the low efficiency of the neutral reconstructions at the LHCb experiment makes this task very difficult. For this reason, it is not possible to reduce this contribution to a negligible level, and it must be taken into account in the measurement. Therefore in this analysis it will be treated on the same footing as signal decays, and the result will be a simultaneous measurement of the $\mathcal{R}(D^+)$ and $\mathcal{R}(D^{*+})$ parameters.

Reconstructed samples of $D^+\mu^-$ candidates contain both $\overline{B}^0 \to D^+\ell^-\overline{\nu}_\ell$ and $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays. No explicit reconstruction of the neutral $\pi^0$ and $\gamma$ accompanying the $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays has been performed. For this reason, the expected correlation between $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ is higher than the one observed in the previous simultaneous measurements of these parameters performed by $B-$factory experiments.

The signal and normalization samples are also present simultaneously in the same dataset and the two contributions are very similar. From a topological point of view the signal sample differs from the normalization sample only for the presence of two additional neutrinos in the final state, resulting in a higher missing mass signature.

In order not to spoil the kinematic difference between the $\mu$ and $\tau$ decays, the event is therefore triggered either on the hadronic part of the signal or on the rest of the event. This is feasible only thanks to the introduction of a fully hadronic trigger during the RunII data taking. The data sample used for this analysis is the one collected during the 2015+2016 RunII data taking, corresponding to an integrated luminosity of 2 fb$^{-1}$. The data collected during RunI is not used due to the lack of HLT2 selection lines dedicated to semitauonic decays analyses.

The reconstruction of the events is performed by looking for high$-p_T$ $D^+$ candidates as combinations of three hadronic tracks making a good vertex, displaced with respect to the

primary vertices of the event. To this vertex then a track, identified as a muon from the muon stations, is added to make the reconstructed $B$ candidate.

The signal and normalization yields is extracted from a binned maximum likelihood fit of the data sample, passing the selection that will be described in the next chapters. The variables used in the fit to data are described in the next section.

## 7.2   Approximated $B$-momentum with $\tau^- \to \mu^- \nu_\tau \overline{\nu}_\mu$ decays

One of the most discriminant variables being used by the $B-$factories is called $m^2_{\text{miss}}$, which is the invariant mass of the particles not being reconstructed in the decay, and the energy of the $\mu$ in the $B-$rest frame, $E^*_\mu$. These variables are especially important in the discrimination between the signal and the normalization samples. The $\overline{B}^0 \to D^+ \mu^- \overline{\nu}_\mu$ decays will peak around zero in $m^2_{\text{miss}}$, being a single neutrino the only missing particle in the decay, and at high $E^*_\mu$, being the muon produced promptly from the $B$ decay. $\overline{B}^0 \to D^+ \tau^- \overline{\nu}_\tau$ decays, instead will show a broader spectrum in $m^2_{\text{miss}}$, due to the presence of three neutrinos in the final state, and a softer $E^*_\mu$ spectrum, being the $\mu$ produced in the decay of a $\tau$ lepton.

The computation of these variables requires the knowledge of the momentum of the $B$ meson undergoing decay. This information is well constrained at $B$-factories, in which the collision of elementary $e^+ e^-$ particles enables the knowledge of the center of mass energy of the muon and the clean environment permits a full reconstruction of the missing energy.

At a hadron collider the $B$ rest frame cannot be fully reconstructed due to the fact that the energy of the colliding partons is not known. The reconstruction of the $B$ hadron momentum is therefore done in an approximated way, using information coming from the topology of the decay. This strategy is the same that has been used in the first $\mathcal{R}(D^*)$ measurement, with $\tau^- \to \mu^- \nu_\tau \overline{\nu}_\mu$ decays, performed by the LHCb experiment [40]. The approximated $B$ rest frame is reconstructed assuming that the proper velocity, $\gamma\beta$ of the "visible" part of the decay ($D\mu$) along the beam ($z$) axis is equal to the one of the $B$ meson. This is a good assumption due to the topology of the $b\overline{b}$ production and the forward acceptance of the LHCb detector. This leads to the following relationship:

$$p_z(B) = \frac{m_B}{m_{D\mu}} p_z(D\mu). \tag{144}$$

The other components of the $B$ momentum vector are then determined from the knowledge of its flight direction, inferred from the position of the $B$ decay vertex with respect to the associated PV. Denoting the angle between the unit vector of the $B$ flight direction and the $z-$axis as $\alpha$, the momentum of the $B$ meson can be approximated as

$$P^{\text{RECO}} = |p(B)| = \frac{m_B}{m_{D\mu}} p_z(D\mu) \sqrt{1 + \tan^2 \alpha}. \tag{145}$$

The resolution obtained on the $B$ momentum in this way is of the order of 25%, as can be noticed in Fig. 34.

Using this estimate for the $B$ momentum, three rest-frame variables are constructed:

- $m^2_{\text{miss}} = (p(B) - p(D) - p(\mu))^2$, the squared invariant mass of the unreconstructed system.

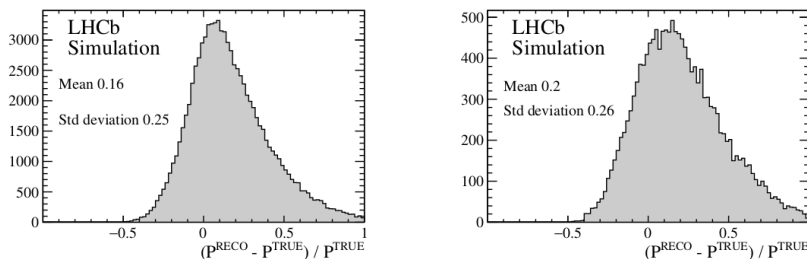- $E^*_\mu$, the energy of the muon in the $B$ rest frame.

Figure 34: *B* momentum resolution obtained with the rest frame approximation described in the text. The plot on the right shows the resolution obtained on $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ simulated events, and the plot on the right the one obtained in $\overline{B}^0 \to D^+\tau^-\overline{\nu}_\tau$

- $q^2 = (p(B) - p(D))^2$, the invariant mass transferred to the lepton system.

The resolution obtained on these variables is around $15 - 20\%$, which is sufficient in order to have a good separation power between the signal and normalization decays. The yields of the signal and normalization samples are then extracted from a three-dimensional maximum likelihood binned fit to these variables. The shape of the signal and normalization contribution as a function of the fit variables is reported in Fig. 35.
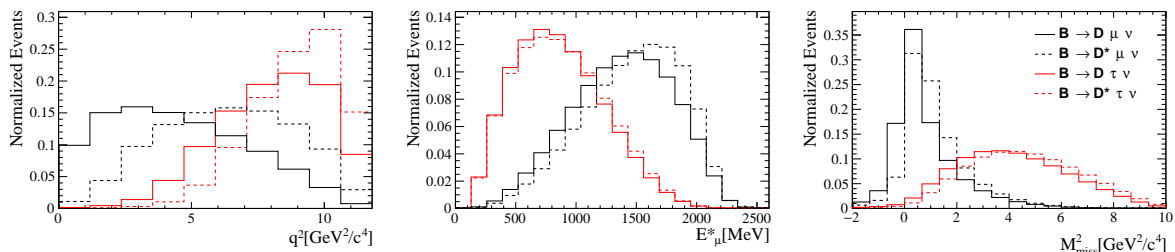


Figure 35: Distribution of the signal and normalization samples as a function of the three fit variables.

## 7.3 Expected background sources

The branching fraction for the $B \to D\ell\nu X$ decays is high, of the order of 10%. For this reason a high background from partially reconstructed decays is expected to contribute to the reconstructed final state.

Two major families of partially reconstructed backgrounds can be identified: *Feed-Down* and *Double-Charm*.

The Feed-Down background is composed by decays of the type $B \to D^{**}\mu\nu$, in which the $D^{**}$ denotes one of the excited states of the $D^{(*)}$ meson.

D mesons are bound states of a charm quark and a (light) antiquark. The ground $D$ and $D^*$ states have orbital quantum number $L = 0$ and a spin of $S = 0$ and $S = 1$ [88], respectively. Given that for a fermion-antifermion pair the parity is given by $(-1)^{l+1}$, the quantum numbers $J^P$ are the following

$$D = 0^- \tag{146}$$

$$D^* = 1^-$$ (147)

The decay of the ground states $S = 1$ are summarized in Table 10.

| Meson | Mass (MeV) | Decay modes |
|---|---|---|
| $D^*(2010)^\pm$ | $2010.26 \pm 0.05$ | $D^0\pi^+$ $(67.7 \pm 0.5\%)$ |
| | | $D^+\pi^0$ $(30.7 \pm 0.5\%)$ |
| | | $D^+\gamma$ $(1.6 \pm 0.4\%)$ |
| $D^*(2007)^0$ | $2006.85 \pm 0.05$ | $D^0\pi^0$ $(64.7 \pm 0.9\%)$ |
| | | $D^0\gamma$ $(35.3 \pm 0.9\%)$ |

Table 10: Properties of the $S = 1$ $c\bar{u}$ ground states [88].

The orbitally excited states can have a total angular momentum of $L = 0, 1, 2, 3$. and are referred in the spectroscopic notation with letters $S, P, D, F$. Other excitations are expected for the radial component, resulting in states with the same quantum numbers but higher masses. These are referred as $2S$ states.

The charm meson spectrum has been predicted in the 1980s [89] and it is reported in Fig. 36.
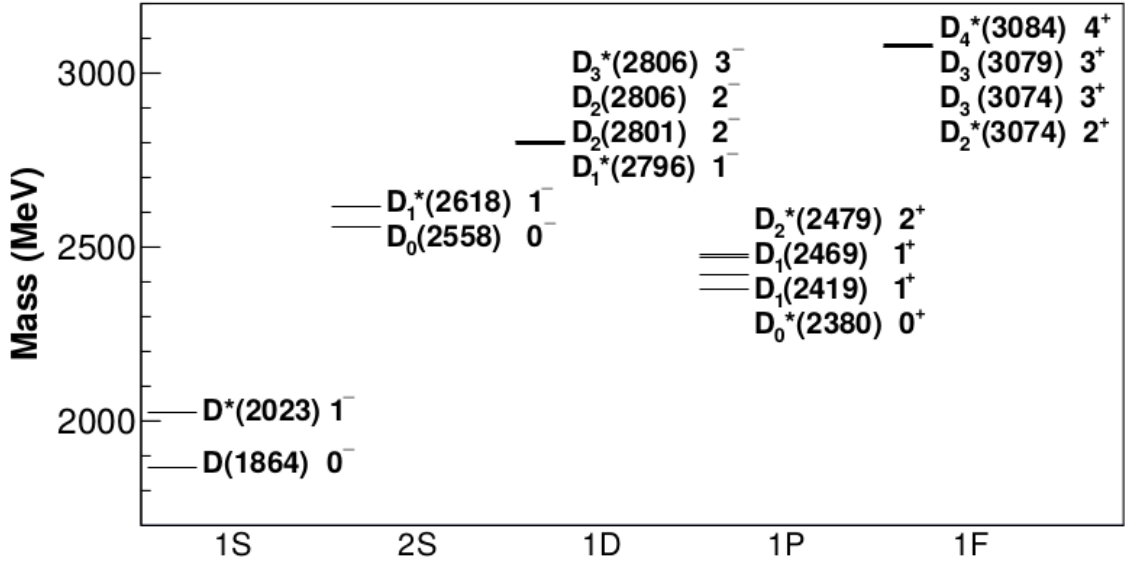


Figure 36: Modified Godfrey-Isgur mass prediction [89, 90]. The $c\bar{u}$ spectrum has been scaled such that the ground state coincides with the $D^0$ mass. The $2^-$ states have been inserted following the splitting structure of the 1P states.

The $L = 1$ mass states can have a total angular momentum given by $|L - S| < J < L + S$. The parity is always positive $((-1)^{l+1})$. The 1P states are therefore 4, named

$$D^{**} = \{D_0^*, D_1', D_1, D_2^*\}.$$ (148)

71

The $D^{**}$ states decay hadronically or electromagnetically to the signal $D^+$, that will be accompanied by a variable number of charged or neutral pions. The status of their knowledge is summarized in Table 11 [88]. It is important to notice that not all the predicted resonances have been observed. For both the neutral and charged charm states, the broad $J = 1$ states are not observed. For the neutral charged states, only the $J = 2$ states have been observed.

| Resonance | Mass (MeV) | Width (MeV) | Decay modes |
|---|---|---|---|
| $D_0^*(2300)^0$ | $2300 \pm 19$ | $274 \pm 40$ | $D^+\pi^-$ (seen) |
| $D_1(2420)^0$ | $2420.8 \pm 0.5$ | $31.7 \pm 2.5$ | $D^*(2010)^+\pi^-$ (seen) |
| | | | $D^0\pi^+\pi^-$ (seen) |
| | | | $D^+\pi^-$ (not seen) |
| | | | $D^{*0}\pi^+\pi^-$ (not seen) |
| $D_2^*(2460)^0$ | $2460.7 \pm 0.4$ | $47.5 \pm 1.1$ | $D^+\pi^-$ (seen) |
| | | | $D^*(2010)^+\pi^-$ (seen) |
| | | | $D^0\pi^+\pi^-$ (not seen) |
| | | | $D^{*0}\pi^+\pi^-$ (not seen) |
| $D_2^*(2465)^\pm$ | $2465.5 \pm 1.3$ | $46.7 \pm 1.2$ | $D^0\pi^+$ (seen) |
| | | | $D^{*0}\pi^+$ (seen) |
| | | | $D^+\pi^+\pi^-$ (not seen) |
| | | | $D^{*+}\pi^+\pi^-$ (not seen) |

Table 11: Properties of the $1P$ states [91].

From Tab. 10 one can infer that background event from decays of the type $B \to D^{**}\mu\nu$ decays, with the $D^{**}$ being one of the $1P$ states, has to come from either a direct decay of the $D^{**}$ to a ground $D^+$ decay, or it has to go through the charged $D^*(2010)^+$ state, since the neutral $D^*(2007)^0$ state does not decay into charged $D$ mesons. This reduces the expected Feed-Down contribution with respect to the analyses using neutral $D^{(*)}$ mesons. The Feed Down coming from 2S states is usually poorly known experimentally, and its shape is calibrated directly from data, as it will be described later on in the text.

The other family of partially reconstructed background takes the name of Double-Charm. This consists of events of the type $B \to (H_c \to \mu\nu X)D^+X'$, in which the signal $\mu$ comes from the decay of an additional charm meson produced in the $B$ decay. The decays in this background can be of two types: Two-Body $B \to (H_c \to \mu\nu X)D^+$ decays and Multi-Body $B \to (H_c \to \mu\nu X)D^+X'$ decays, with the latter being less known experimentally with respect to the former. For this reason the shape of the Multi-Body decays will be calibrated directly from data, as it will be described later on in the text. A source of baryonic Double-Charm background, from $\Lambda_b$ decays for example, can also be expected, albeit with a very small fraction.

The shape of these backgrounds is estimated with a MC simulation, as it will be explained

later in the text.

The rest of the expected backgrounds is instead estimated through a data driven method. Various sources of combinatorial-like background are to be expected. The first background is called $D$-combinatorial, which consists of random three tracks combinations that are being reconstructed as a $D^+$. This background is characterized by a broad distribution in the invariant $K\pi\pi$ mass spectrum, whereas true $D^+$ candidates are characterized by a peak at the nominal value of the $D^+$ mass.

The sample of tracks reconstructed and selected as the signal $\mu$ candidate is polluted also by electron, hadrons and fake tracks passing all the PID selections and being reconstructed as true muons. This background takes the name of $\mu$-MisID, and constitutes one of the most dangerous backgrounds for this analysis, with a very difficult modelling.

Even when a true $\mu$ and a true $D^+$ have been reconstructed, it is not granted the two come from the decay of the same ancestor. The last data driven estimated background is called $D\mu$-combinatorial, and is composed of random combination of true $\mu$ and true $D$ candidates.

These three backgrounds have been estimated in a fully data-driven way, as it will be described in Sect. 12. Some overlap between them is also expected, and special care must be taken in order to avoid double counting of any background source.

Given the amount of background events expected in this analysis, the calibration of the shape (and sometimes of the rate) of the background sources is of fundamental importance. Orthogonal control regions have been defined with the purpose of calibrating the shape of specific background sources in a data driven way. These control regions, that will be defined later on in the text, are fitted simultaneously with the signal region, in order to correctly take into account the background related systematic uncertainties and check the agreement between the data and the fit model. This also offers the advantage of having systematic uncertainties that will scale with the luminosity recorded by the experiment. Finally some validation regions have also been used to evaluate and to cross-check the shape and the rate of some specific backgrounds, as it will be later on described in the text.

## 7.4   Fast simulation

In the first $\mathcal{R}(D^*)$ measurement performed by LHCb, using leptonic decays of the $\tau$ [40], the low statistics of the MC sample used to estimate the background sources was one of the biggest sources of systematic uncertainty. This is due to the very big data sample to be analysed. In the 2015 + 2016 combined sample, the number of signal events to be analysed, after all the selections, is expected to be around $3 \times 10^6$ in the case of the analysis presented in this thesis. In order to reduce the systematics due to the statistics of the MC sample used to a negligible level, one would need at least 50-100 simulated events for each real data event used.

An optimistic estimation of the CPU time needed to simulate the event, propagate the final state particles through the material and simulate the response of the detector, would give a ballpark time of about 1 minute for each event. It is clear that in order to reduce to a negligible level this systematic uncertainty, the full simulation of the event would not be sustainable.

Roughly 85% of the time required for processing a single event in MC is taken during the simulation of the response of the RICH detectors and the response of the calorimeters system, in which the development of hadronic and electromagnetic showers is being simulated.

In this analysis a novel *fast-simulation* technique, taking the name of *Tracker-Only* simulation, has been used. This is a fast simulation technique in which the whole detector material is present

for the interaction of the final state particles, but the response of only the tracking subdetectors is simulated. This is achieved by setting the RICH, the Calorimeters and the Muon systems to passive material in Geant4, as illustrated in Fig. 37.
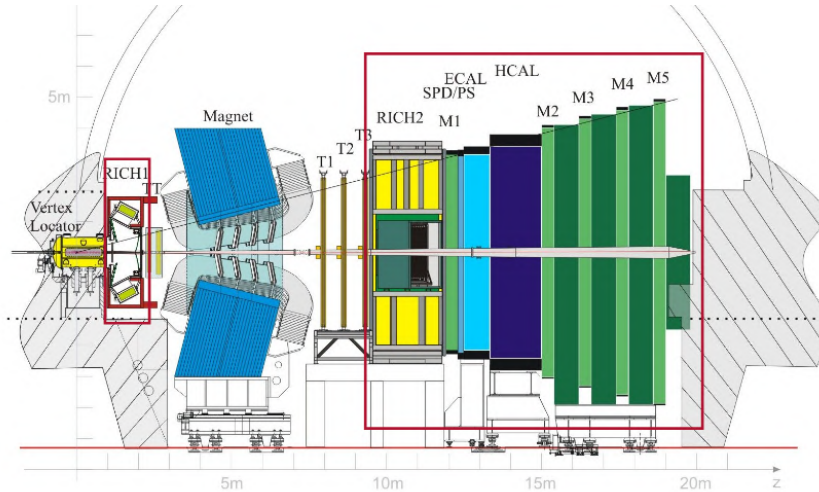


Figure 37: Schematic view of the LHCb detector showing the components set to passive material in tracker only simulation in red.

Tracker-Only simulation is 8 times faster and the event size required on disk is approximately 40% low than full simulation.

The price that has to be paid in order to exploit this advantage is not having the full information available for the event reconstruction. The relevant information used for the analysis presented in this thesis can, nonetheless, be successfully emulated offline using only information coming from the tracking system or reproducing it from calibration real data samples. This will be explained in detail in Sect. 10.

# 8 Datasets and selections

## 8.1 Real Data

The analysis presented in this thesis is based on $\approx 300$ pb$^{-1}$ of data recorded in 2015 and 1600 pb$^{-1}$ of data recorded in 2016 by the LHCb experiment in proton-proton collisions at a center of mass energy of $\sqrt{s} = 13$ TeV.

The data used is composed of events in which a $D^+\mu^-$ (and the corresponding charged conjugated final state), with the $D^+ \to K^-\pi^+\pi^+$ have been reconstructed and selected. This will be referred as the *nominal* data sample. A sample in which the $D$ and $\mu$ candidates have the same charge have also been constructed. This sample, as it will be shown later on, has been used in order to estimate the background coming from random combinations of $D$ and $\mu$ candidates ($D\mu$-combinatorial background), and it is constructed combining $D^+\mu^+$ and $D^-\mu^-$, where the single particles are required to pass the same selections as the nominal data sample. This will be called the *Wrong Sign (WS)* sample from now on.

Lastly, a sample in which the muon candidate is required to pass all the selections applied to the nominal sample, but it is required to have no hits in the muon chambers associated to its track, is constructed. This sample is enriched in events in which the $\mu$ candidate is actually a true hadron being misidentified as a muon. As it will be described later on, this sample has been used to estimate the $h \to \mu$ misidentification background ($\mu$-MisID). This sample is constructed in both the wrong-sign ($D^+h^+$, $D^-h^-$) and right-sign ($D^+h^-$, $D^-h^+$) combination. In the following these will be denoted as *wrong-* and *right-sign MisID* samples.

## 8.2 Monte Carlo samples

All the physical backgrounds, the signal and the normalization channels, have been simulated by means of a Monte Carlo simulation. MC events for both 2015 and 2016 data taking conditions and for both magnet polarities configurations (Magnet Up and Magnet Down), have been generated.

The *pp* collision and fragmentation of the partons produced in the hard scattering process have been generated with *Pythia 8* [67]. The decay of the mesons and hadrons produced in the fragmentation is handled by the *EvtGen* [59] package. Each process has been generated in a specific sample, by running the production and the EvtGen application with a specific configuration, detailed in a single file with a unique identifier called EventType. Finally, the interaction of the stable particles with the detector, and the detector response itself, are simulated with the *Geant4* [92] toolkit, configured with a detailed description of the LHCb detector. After the response of the detector has been simulated and the raw data has been saved, the simulated events are reconstructed with the same reconstruction sequence employed on events coming from real collisions.

The list of MC generated samples and their relative event type is reported in Table 12.

In order to reduce the systematic uncertainties due to the limited size of the MC samples used to obtain the templates, high statistics MC samples, in which only the response of the tracking system has been simulated, have been produced. These samples are referred as Tracker-Only MC (TO MC) samples. This fast simulation, omitting the simulation of the response of the RICH, Calorimeters and Muon stations, enables to reduce the generation time by a factor of $\approx 8$ and a save of $\approx 40\%$ of disk space. As a downside, it requires the emulation of a set of information not accessible due to the non-active subdetectors in the simulation steps. The emulation of those features will be explained in the following sections.

| Sample | Event type |
|---|---|
| $\overline{B}^0 \to D^+ \tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 11574060 |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 11574401 |
| $\overline{B}^0 \to D^+ \mu\overline{\nu}_\mu$ | 11574061 |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\overline{\nu}_\mu$ | 11574402 |
| $\overline{B}^0 \to D^{**+}(\to D^+X)\mu\overline{\nu}_\mu$ | 11574403 |
| $\overline{B}^0 \to D^{**+}(\to D^+X)\mu\overline{\nu}_\mu$, high mass | 11574070 |
| $B^\pm \to D^{**0}(\to D^+X)\mu\overline{\nu}_\mu$ | 12874050 |
| $\overline{B}^0 \to D^+ H_c(\to \mu\overline{\nu}_\mu X')X$ | 11995204 |
| $B^- \to D^+ H_c(\to \mu\overline{\nu}_\mu X')X$ | 12995604 |
| $\Lambda_b \to (\Lambda_c \to \mu\nu X)DX'$ | 15976000 |
| $B^0 \to D^\pm(D_s \to \tau\nu)X$ | 11995214 |
| $B^\pm \to D^\pm(D_s \to \tau\nu)X$ | 12995615 |

Table 12: List of Monte Carlo samples used.

A small sample of events, in the 2015 data taking conditions, has been generated with the full simulation of the LHCb detector. This was requested to set up the analysis and in order to validate the Tracker only MC production.

In order not to spend CPU time in the simulation of events that would not be reconstructed or that have signal particles falling outside the LHCb acceptance, the generated events are required to pass some loose selections at the generator level, before their propagation in the detector has been simulated. These selections, that will be referred as *generator-level* selections in the rest of the thesis, are listed in Table 13. They require that all the signal tracks fall inside the LHCb acceptance and have a minimum (transverse) momentum requirement. Furthermore, they require a minimum (transverse) momentum on the daughters of the $D$ candidate.

In order to reduce the size of the MC files saved after the simulation and not store unnecessary events that are not going to pass the selections anyway, only reconstructed events that pass a set of selections are saved. This set of selection tries to mimic as closely as possible the selections that will be later on called *Stripping* selections. This is going to be called *MC-filtering* selection in the rest of the thesis. The set of MC-filtering selections applied is reported in Table 14.

The number of reconstructed MC events, for each year, MC type and magnet polarity, passing the filtering selections is reported in Table 15.

Simulated events are also truth matched to avoid double counting of non-physical backgrounds. All the samples are required to pass a common truth matching selection, in which the reconstructed $B$ candidate is required to have a true ID of a $b$ meson resonance. Its daughters are required to have been correctly matched to MC truth objects, effectively removing ghost tracks and clone candidates, and to come from the same true parent, effectively removing combinatorial

| Variable | Requirement |
|---|---|
| $\theta(\pi)$ | $\in [0.01, 0.4]$ rad |
| $\theta(K)$ | $\in [0.01, 0.4]$ rad |
| $\theta(\mu)$ | $\in [0.01, 0.4]$ rad |
| $p_T(\pi)$ | $> 150$ MeV/c |
| $p_T(K)$ | $> 150$ MeV/c |
| $p(\mu)$ | $> 2500$ MeV/c |
| $p(K^+) + p(\pi^-) + p(\pi^-)$ | $> 15000$ MeV/c |
| $p_T(K^+) + p_T(\pi^-) + p_T(\pi^-)$ | $> 2300$ MeV/c |

Table 13: List of generator level selections.

background. The reconstructed $D$ and $\mu$ candidates are also required to have the correct true ID, effectively removing any $\mu$-MisID and any $D$-combinatorial backgrounds.

Subsequently, a specific truth matching is performed on the single MC samples. For the $B \to D\mu\nu$ sample, the mother of the $D$ and the mother of the $\mu$ candidates are required to be a $B^0$ particle, and the neutrino is required to have a true positive energy.

For the $B \to D^*\mu\nu$ sample, the mother of the $D$ candidate and the daughter of the $B$ candidate are required to have a true ID compatible with the one of a $D^*(2010)$ meson, and the grandmother of the $D$ and the mother of the $\mu$ the one of a true $B^0$ meson. As before, the neutrino is required to have a positive true energy. Lastly, the true ID of the charm daughter of the $D^*$ candidate is asked to be consistent with the one of a $\pi^0$, with a positive true energy.

In the $B \to D\tau\nu$ sample the mother of the $D$ candidate and the grandmother of the $\mu$ candidate are required to be a true $B^0$ and the mother of the $\mu$ is required to be a true $\tau$. The true energy of the $\tau$ and the one of all neutrinos in the event are required to be positive.

In the $B \to D^*\tau\nu$ sample the mother of the $D$ is required to be a true $D^*$ and the mother of the $\mu$ is required to be a true $\tau$. The grandmother of the $\mu$ and of the $D$ are required to be a $B$ and the true energy of all neutrinos in the event is required to be positive. As in the $B \to D^*\mu\nu$ case, the true ID of the daughters of the $D^*$ candidate are required to be compatible with the ones of a $D$ and a $\pi^0$.

For the Double Charm samples, the only requirement applied is that the first two daughters of the $B$ are compatible with a charmed hadron. For events involving a $D_s \to \tau\nu$ decay, the ID of the mother of the $\mu$ is required to be compatible with a $\tau$, and to be compatible with a charmed hadron in all the other cases.

Finally, for the $B \to D^{**}\mu\nu$ samples, the ID of the first daughter of the $B$ is required to be compatible with a 1P $D^{**}$ state.

The truth matching selections are summarised in Table 16.

| Particle | Variable | Requirement |
|---|---|---|
| $K, \pi$ | Track $p$ [GeV/$c$] | $> 2$ |
| | Track $p_T$ [MeV/$c$] | $> 300$ |
| | Track $IP_{\chi^2}^{PV}$ | $> 9$ |
| | Track GhostProb | $< 0.5$ |
| | MC match to true particle | $True$ |
| $\mu$ | Track $p$ [GeV/$c$] | $> 3$ |
| | Track GhostProb | $< 0.5$ |
| | Track $IP_{\chi^2}^{PV}$ | $> 16$ |
| | MC match to true particle | $True$ |
| $D$ | $\sum$ tracks $p_T$ [MeV/$c$] | $> 2500$ |
| | Mass interval[MeV/$c$] | 1770-1970 |
| | Vertex $\chi^2$/d.o.f. | $< 4$ |
| | $DIRA$ | $> 0.999$ |
| | Flight Distance $\chi^2_{BPV}$ | $> 25$ |
| $B$ | Vertex $\chi^2$/d.o.f. | $< 6$ |
| | $DIRA$ | $> 0.999$ |

Table 14: Summary of the MC-filtering selections applied. The $IP_{\chi^2}$ of a track with respect to a vertex is the change in the vertex fit $\chi^2$ of the vertex reconstructed with and without the track under consideration. The flight distance of a particle is given by the length of the separation between the production and its decay vertex. The flight distance $\chi^2$ is defined as the significance of this separation in units of $\chi^2$ and it is evaluated as $\vec{v}^T \cdot M \cdot \vec{v}$, where $M$ is the sum of the covariance matrices of the production vertex and decay vertex and $\vec{v}$ is the separation vector between the decay vertex and the production vertex. The $DIRA$ of a composite particle is defined as the cosine of the angle between its flight direction and its momentum vector.

| Sample | DecFile Number | Full MC | TO MC (2015) | TO MC (2016) |
|---|---|---|---|---|
| $\overline{B}^0 \to D^+ \mu \overline{\nu}_\mu$ | 11574061 | 254574<br>260659 | 2697651<br>2575258 | 14826442<br>14840102 |
| $\overline{B}^0 \to D^{*+}(\to D^+ \pi^0) \mu \overline{\nu}_\mu$ | 11574402 | 262662<br>253462 | 2515725<br>2684522 | 14905871<br>14913692 |
| $\overline{B}^0 \to D^+ \tau(\to \mu \overline{\nu}_\mu \nu_\tau) \overline{\nu}_\tau$ | 11574060 | 54544<br>52128 | 504758<br>504119 | 3054641<br>3044256 |
| $\overline{B}^0 \to D^{*+}(\to D^+ \pi^0) \tau(\to \mu \overline{\nu}_\mu \nu_\tau) \overline{\nu}_\tau$ | 11574401 | 52485<br>51284 | 522630<br>505642 | 3057081<br>3072491 |
| $B^- \to D^{**0}(\to D^+ X) \mu \overline{\nu}_\mu$ | 12874050 | 70011<br>73182 | 529407<br>500930 | 3024193<br>3029403 |
| $\overline{B}^0 \to D^{**+}(\to D^+ X) \mu \overline{\nu}_\mu$ | 11574403 | 69903<br>69506 | 515467<br>521194 | 3023983<br>3003310 |
| $\overline{B}^0 \to D^{**+}(\to D^+ X) \mu \overline{\nu}_\mu$, high mass | 11574070 | 33179<br>33740 | 593073<br>251550 | 3100683<br>3099873 |
| $\overline{B}^0 \to D^+ H_c(\to \mu \overline{\nu}_\mu X') X$ | 11995204 | 169986<br>173472 | 1715847<br>1714838 | 8988227<br>8989294 |
| $B^- \to D^+ H_c(\to \mu \overline{\nu}_\mu X') X$ | 12995604 | 79481<br>77967 | 764100<br>781436 | 3612289<br>3618327 |
| $\overline{B}^0 \to D^+ H_c(\to \tau \overline{\nu}_\tau X') X$ | 11995214 | | 129257<br>103398 | 646084<br>612330 |
| $B^- \to D^+ H_c(\to \tau \overline{\nu}_\tau X') X$ | 12995615 | | 102780<br>102197 | 681352<br>655164 |
| $\Lambda_b \to (\Lambda_c \to \mu \overline{\nu_\mu} X') D^+ X$ | 15976000 | | 53070<br>53872 | 656017<br>632889 |
| $\overline{B}^0 \to D^{**+}(\to D^+ X) \tau(\to \mu \overline{\nu}_\mu \nu_\tau) \overline{\nu}_\tau$ | 11874002 | | 65119<br>63074 | 300407<br>298796 |
| $B^- \to D^{**0}(\to D^+ X) \tau(\to \mu \overline{\nu}_\mu \nu_\tau) \overline{\nu}_\tau$ | 12874001 | | 64026<br>66564 | 299820<br>310589 |
| $\overline{B}_s^0 \to D_s^{**+}(\to D^+ X) \mu \overline{\nu}_\mu$ | 13874003 | | 93442<br>98911 | 505121<br>492409 |

Table 15: Number of events passing the filtering selections. For each MC production, the number of events for the Magnet Up polarity (first row) and Magnet Down polarity (second row) is reported.

79

| Sample | Variable | Requirement |
|---|---|---|
| All samples | BKGCAT($\boldsymbol{B}$) | $< 60$ |
| | $\|\text{ID}(\boldsymbol{B})\|$ | compatible with a $b$ meson |
| | $\|\text{ID}(\boldsymbol{D})\|$ | compatible with $D^\pm$ |
| | $\|\text{ID}(\boldsymbol{\mu})\|$ | compatible with $\mu^\pm$ |
| $\overline{B}^0 \to D^+\mu\overline{\nu}_\mu$ | $\|\text{MC MOTHER ID}(\boldsymbol{D})\|$ | compatible with $B^0$ |
| | $\|\text{MC MOTHER ID}(\boldsymbol{\mu})\|$ | compatible with $B^0$ |
| | $E(\nu)$ | $> 0$ MeV |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\overline{\nu}_\mu$ | $\|\text{MC MOTHER ID}(\boldsymbol{D})\|$ | compatible with $D^{*+}(2010)$ |
| | $\|\text{MC GD MOTHER ID}(\boldsymbol{D})\|$ | compatible with $B^0$ |
| | $\|\text{MC MOTHER ID}(\boldsymbol{\mu})\|$ | compatible with $B^0$ |
| | $E(\nu)$ | $> 0$ MeV |
| | $\|\text{ID}(D_0)\|$ | compatible with $D^{*+}(2010)$ |
| | $\|\text{ID(DAUGHTER 0 }(D_0))\|$ | compatible with $\pi^0$ |
| | $\|\text{ID(DAUGHTER 1 }(D_0))\|$ | compatible with $D^\pm$ |
| | $E(D_0)$ | $> 0$ MeV |
| | $E(\text{DAUGHTER 0}(D_0))$ | $> 0$ MeV |
| $\overline{B}^0 \to D^+\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | $\|\text{MC MOTHER ID}(\boldsymbol{D})\|$ | compatible with $B^0$ |
| | $\|\text{MC MOTHER ID}(\boldsymbol{\mu})\|$ | compatible with $\tau^\pm$ |
| | $\|\text{MC GD MOTHER ID}(\boldsymbol{\mu})\|$ | compatible with $B^0$ |
| | $E(\tau)$ | $> 0$ MeV |
| | $E(\nu_\mu)$ | $> 0$ MeV |
| | $E(\nu_\tau)$ | $> 0$ MeV |
| | $E(\overline{\nu}_\tau)$ | $> 0$ MeV |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | $\|\text{MC MOTHER ID}(\boldsymbol{D})\|$ | compatible with $D^{*+}(2010)$ |
| | $\|\text{MC GD MOTHER ID}(\boldsymbol{D})\|$ | compatible with $B^0$ |
| | $\|\text{MC MOTHER ID}(\boldsymbol{\mu})\|$ | compatible with $\tau^\pm$ |
| | $\|\text{MC GD MOTHER ID}(\boldsymbol{\mu})\|$ | compatible with $B^0$ |
| | $E(\tau)$ | $> 0$ MeV |
| | $E(\nu_\mu)$ | $> 0$ MeV |
| | $E(\nu_\tau)$ | $> 0$ MeV |
| | $E(\overline{\nu}_\tau)$ | $> 0$ MeV |
| | $\|\text{ID}(D_0)\|$ | compatible with $D^{*+}(2010)$ |
| | $\|\text{ID(DAUGHTER 0 }(D_0))\|$ | compatible with $\pi^0$ |
| | $\|\text{ID(DAUGHTER 1 }(D_0))\|$ | compatible with $D^\pm$ |
| | $E(D_0)$ | $> 0$ MeV |
| | $E(\text{DAUGHTER 0}(D_0))$ | $> 0$ MeV |
| $B \to D^+(H_c \to \mu\nu X)X$ | $\|\text{MC MOTHER ID}(\boldsymbol{\mu})\|$ | compatible with $D^0, D^\pm$ or $D_s$ |
| | $\|\text{ID}(D_0)\|$ | compatible with charmed particle |
| | $\|\text{ID}(D_1)\|$ | compatible with charmed particle |
| $B \to D^+(D_s \to \tau\nu X)X$ | $\|\text{MC MOTHER ID}(\boldsymbol{\mu})\|$ | compatible with $\tau$ |
| | $\|\text{ID}(D_0)\|$ | compatible with charmed particle |
| | $\|\text{ID}(D_1)\|$ | compatible with charmed particle |
| $\overline{B}^0 \to D^{**+}\mu\overline{\nu}$ | $\|\text{ID}(D_0)\|$ | compatible with $D^{**}$ $1P$ state |
| $B^- \to D^{**0}\mu\overline{\nu}$ | $\|\text{ID}(D_0)\|$ | compatible with $D^{**}$ $1P$ state |

Table 16: Summary of truth matching selections. The particles reported in bold text denote the reconstructed candidates. $D_0$ and $D_1$ refer to the first and second hadron encountered in the decay chain, when starting from the true $B$ hadron produced in the event and traversing the decay chain downwards.

## 8.3 Selection

The selections employed in this analysis aim at efficiently reconstruct and select $D\mu$ candidates with high (transverse) momentum and a good vertexing quality. Particle identification selections are also put on stable particle tracks, to reduce combinatorial and misidentification backgrounds. A pointing and displacement requirement on the $D$ candidate with respect to the associated primary vertex is also put in order to remove contamination coming from charm particles promptly produced at the primary vertex.

From the very first stages, the selection and reconstruction sequence are aimed to have a minimum impact on the muon kinematics, in order to preserve as much as possible the kinematic difference between $\mu$ directly produced at the $B$ decay with respect to those coming from a $\tau$ decay. The latter are expected to have softer kinematics, and any hard kinematic selection could spoil the separation power between the $B \to D^{(*)}\mu\nu$ and $B \to D^{(*)}\tau\nu$, of fundamental importance for this analysis. The distribution of the momentum of the muons in the signal and normalization samples is reported in Fig. 38.
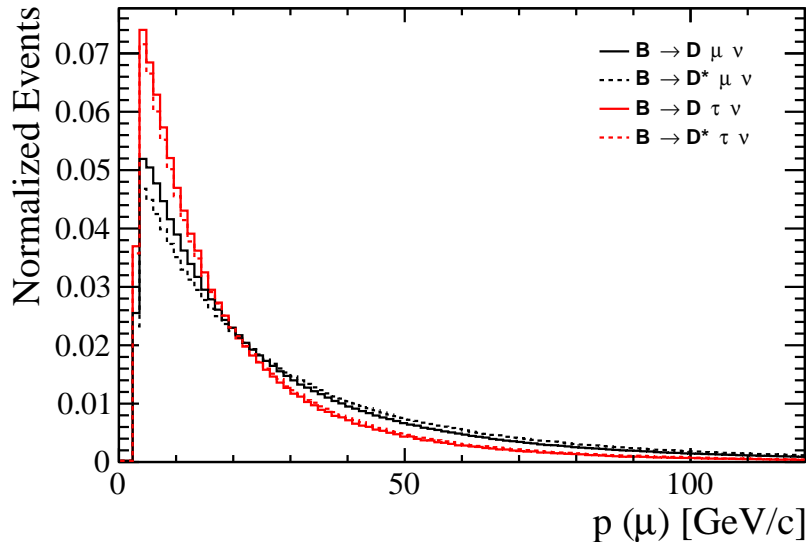


Figure 38: Momentum distribution of events from simulated signal and normalization samples.

During the online data taking, the events are selected and reconstructed using trigger lines which operate entirely on the hadronic part of the event.

During offline selection, a first set of loose selection is put. This step is called *Stripping*, and the selections employed are very close to the ones used at the second level of the software trigger. This selection is performed centrally, and the Stripping line used will be described in the next sections.

During offline selection, a set of Particle identification cuts and MVA selections, aimed at reducing the $D$-combinatorial, and optimized on the decay under study, have been put.

A selection to prevent Cross-Feeds coming from different charm hadrons, polluting the $D$ mass spectrum, is put on the events surviving the previous selections.

Finally, a multivariate Isolation selection is employed to reduce the fraction of events in which additional charged and neutral particles have been produced at the $B$ decay vertex but have not

been associated with the signal.

All the selections are detailed in the following sections.

### 8.3.1 Trigger

At the hardware level of the trigger system (L0), all the events are required to have triggered either the `L0Global` trigger (TIS) or the `L0Hadron` (TOS) lines. The L0Global line, requires that the requirements of at least one of the L0 trigger lines have been satisfied by the event. The requirements of the L0 trigger lines have been reported in Table 9.

No hardware level $\mu$ trigger requirement is imposed in order not to bias the shape of the discriminating kinematical variables used in the fit.

In order to protect the reconstruction sequences against saturated events, global cuts on variables correlated with the detector occupancy are applied during the data taking. Before the HLT1 reconstruction sequence, for example, the event is discarded if the number of hits in the tracking system exceeds some thresholds. The list of Global Event Cuts applied during data taking before the HLT1 reconstruction sequence is reported in Table 17.

The selection imposed at the first stage of the software trigger require the presence of one or two tracks, with high momentum, detached from the primary vertex. This selection is implemented with two separate trigger lines, called `Hlt1TrackMVA` and `Hlt1TwoTrackMVA`, respectively, which are required to be firing due to the hadronic part of the event (at least one of the two lines is required to be TOS on the $D$.) In order to reduce the CPU time of the reconstruction strategy, events with too high an occupancy in the VELO, OT, and IT are rejected. The maximum number of hits allowed is 6000 in the VELO (`nVeloClusters`), 3000 in the IT (`nITClusters`) and 15000 in the OT (`nOTClusters`).

The `Hlt1TrackMVA` line requires the presence of one track, with a good fit quality and high momentum. The requirement of being detached from the PV is imposed by means of an MVA. For speed reasons, the MVA is not fully run online, but its decision is approximated with an analytical formula, which acts on the $p_T$ and $IP_{\chi^2}$ of the track with respect to the PV. The requirements of this trigger line are reported in Table 18 and Table 19, for the two years of data taking.

The `Hlt1TwoTrackMVA` line requires the presence of a two-track combination, with a high momentum, good vertex quality and high impact parameter with respect to the primary vertex. After a first preselection on the tracks and two-tracks combinations, this is implemented with a MVA selection. The selections of this trigger line are reported in Table 20 and Table 20, for the two years of data taking.

| Global Event Cuts |
| :---: |
| $50 < $ `nVeloClusters` $ < 6000$ |
| $50 < $ `nITClusters` $ < 3000$ |
| $50 < $ `nOTClusters` $ < 15000$ |

Table 17: Global Event cuts applied before the HLT1 reconstruction during data taking

| Hlt1TrackMVa |
|---|
| Input tracks selections (after Track-Fit) |
| $p_T > 500$ MeV |
| $p > 3$ GeV |
| Track $\chi^2/\text{d.o.f} < 2.5$ |
| $\left(p_T > 25 \text{ GeV}/c \wedge IP_{\chi^2} > 7.4\right) \vee [(1 \text{ GeV}/c < p_T < 25 \text{ GeV}/c) \wedge$ $\log(IP_{\chi^2}) > \left(\dfrac{1}{p_T[\text{GeV}/c] - 1}\right)^2 + \left(\dfrac{1.1}{25 \text{ GeV}/c}\right)(25 \text{ GeV}/c - p_T) + \log(7.4)]$ |

Table 18: Requirements of the Hlt1TrackMVa trigger line in 2015

| Hlt1TrackMVa |
|---|
| Input tracks selections (after Track-Fit) |
| $p_T > 600$ MeV |
| $p > 5$ GeV |
| Track $\chi^2/\text{d.o.f.} < 2.5$ |
| $\left(p_T > 25 \text{ GeV}/c \wedge IP_{\chi^2} > 7.4\right) \vee [(1 \text{ GeV}/c < p_T < 25 \text{ GeV}/c) \wedge$ $\log(IP_{\chi^2}) > \left(\dfrac{1}{p_T[\text{GeV}/c] - 1}\right)^2 + \left(\dfrac{1.1}{25 \text{ GeV}/c}\right)(25 \text{ GeV}/c - p_T) + \log(7.4)]$ |

Table 19: Requirements of the Hlt1TrackMVa trigger line in 2016

At the second stage of the software trigger, the events are required to pass the selection defined in the Hlt2XcMuXForTauB2XcMu trigger line. The events are required to be TOS on the $D$ with respect to these selections. This line was specifically designed for this analysis and is aimed at minimizing the biases on the $\mu$ kinematic distributions, which may affect the distinction between the tauonic and the muonic modes during the fit. This trigger line requires the hadronic tracks to have high momentum, to form a good vertex, well separated from the PV, for the $D$ candidate. A loose PID selection is also imposed on them. No hard kinematic selection is imposed on the $\mu$, which is only required to have a high impact parameter with respect to the primary vertex and to form a good vertex with the $D$ candidate. The selections of this line are reported in Table 22, along with the Stripping and the filtering selections used on the MC samples.

A similar line, called Hlt2XcFakeMuForTauB2XcMu, is employed to define the MisID samples. In this line all the selections are identical to the previous one, except that the reconstructed

|  | Hlt1TwoTrackMVa |
|---|---|
| Requirement on the single tracks | $p_T > 0.5$ GeV/$c$ <br> $p > 5.0$ GeV/$c$ <br> $\chi^2/d.o.f. < 2.5$ |
| Requirements on the track pair before vertexing | $(p_1 + p_2)_T > 2$ GeV/$c$ <br> $DOCA(1,2) < 10$ |
| Requirements on the track pair combination | Vertex $\chi^2 < 10$ <br> $m_{corr} > 1$ GeV/$c^2$ <br> $2 < \eta < 5$ <br> $DIRA > 0$ |
| MVA requirements | MVA Output $> 0.95$ |
| MVA training variables | Vertex $\chi^2$ <br> Vertex distance $\chi^2$ <br> $p_{T,1} + p_{T,2}$ <br> Number of tracks with $IP_{\chi^2} < 16$. |

Table 20: Requirements of the `Hlt1TwoTrackMVa` trigger line in 2015

$\mu$ candidates are required to not have any hit in the $\mu$ stations associated to them. This is done requiring the $\mu$ to not pass the `isMuon` selection, while still falling in the muon chambers acceptance (`inMuon`). The former is a boolean flag which is set to true for tracks recognized as muons. The evaluation of `IsMuon` relies on the number of hits around the tracks extrapolated through the muon stations. The sizes of the hit search windows, named FoI (Field of Interest), are parametrised accounting for the particle momentum and the muon detector regions crossed. This flag is set by the Muon Identification algorithm, run both at the HLT1 and the offline reconstruction levels.

This line has been used construct the $\mu-$MisID sample, and it will be denoted as the Fake-$\mu$ (HLT2) trigger line. A prescale of 10% is applied to it.

### 8.3.2 Stripping selections

Stripping selections are used during the offline selection which is performed centrally at the LHCb experiment, to classify the events that have passed the online trigger selections. The line called `Strippingb2DpMuXB2DMuForTauMuLine` is used to perform this step. The line consist in a set of selections that closely match the requirements applied at the HLT2 trigger stage. An

| | Hlt1TwoTrackMVa |
|---|---|
| Requirement on the single tracks | $p_T > 0.6$ GeV/$c$ <br> $p > 5.0$ GeV/$c$ <br> $\chi^2/d.o.f. < 2.5$ <br> $GhostProb < 0.2$ |
| Requirements on the track pair before vertexing | $(p_1 + p_2)_T > 2$ GeV/$c$ <br> $DOCA(1,2) < 10$ |
| Requirements on the track pair combination | Vertex $\chi^2 < 10$ <br> $m_{corr} > 1$ GeV/$c^2$ <br> $2 < \eta < 5$ <br> $DIRA > 0$ |
| MVA requirements | MVA Output $> 0.95$ |
| MVA training variables | Vertex $\chi^2$ <br> Vertex distance $\chi^2$ <br> $p_{T,1} + p_{T,2}$ <br> Number of tracks with $IP_{\chi^2} < 16$. |

Table 21: Requirements of the `Hlt1TwoTrackMVa` trigger line in 2016

additional PID selection is applied at this stage on all the signal tracks, which are required to have a small probability to be fake (or ghost) tracks. The selections of this line are reported in Table 22, along with the HLT2 and the filtering selections used on the MC samples.

Analogously to the HLT line, the MisID background is constructed from candidates in output to the `Strippingb2DpMuXFakeB2DMuForTauMuLine` module, which differs from the previous one only for the requirement of the $\mu$ to fail the muon identification requirement.

| Particle | Variable | Hlt2 cuts | Stripping cuts | Filtering cuts |
|---|---|---|---|---|
| $K, \pi$ | $K\ PIDK$ | $> 2$ | $> 4$ | $-$ |
| | $\pi\ PIDK$ | $< 4$ | $< 2$ | $-$ |
| | Track $IP_{\chi^2}$ | $> 9$ | $> 9$ | $> 9$ |
| | Track $p_T$ [MeV/$c$] | $> 200$ | $> 300$ | $> 300$ |
| | Track $p$ [MeV/$c$] | $> 5000$ | $> 2000$ | $> 2000$ |
| | $\geq 1$ track $p_T$ [MeV/$c$] | $> 800$ | $-$ | $-$ |
| | $\sum$ track $p_T$ [MeV/$c$] | $> 2500$ | $> 2500$ | $> 2500$ |
| | Track GhostProb | $-$ | $< 0.5$ | $< 0.5$ |
| $D$ | $D$ mass interval [MeV/$c^2$] | $1830 - 1910$ | $1790 - 1950$ | $1770 - 1970$ |
| | $D\ p_T$ [MeV/$c$] | $> 2000$ | $-$ | $-$ |
| | $D$ child pair DOCA [mm] | $< 0.10$ | $-$ | $-$ |
| | $D\ \chi^2_{vertex}/ndf$ | $< 10$ | $< 4$ | $< 4$ |
| | $D$ DIRA | $> 0.999$ | $> 0.999$ | $> 0.999$ |
| | $D$ Flight Distance $\chi^2$ | $> 25$ | $> 25$ | $> 25$ |
| $\mu$ | $\mu\ IP_{\chi^2}$ | $> 16$ | $> 16$ | $> 16$ |
| | $\mu\ PID\mu$ | $-$ | $> -200$ | $-$ |
| | $\mu$ GhostProb | $-$ | $< 0.5$ | $< 0.5$ |
| | $\mu\ p$ [MeV/$c$] | $-$ | $> 3000$ | $> 3000$ |
| $D\mu$ | $D\mu\ \chi^2_{vertex}/ndf$ | $< 15$ | $< 6$ | $< 6$ |
| | $D\mu$ DIRA | $> 0.999$ | $> 0.999$ | $0.999$ |
| | $D\mu$ DOCA [mm] | $< 0.50$ | $-$ | $-$ |
| | $D\mu$ Flight Distance $\chi^2$ | $> 50$ | $-$ | $-$ |
| | $D\mu$ mass interval [MeV/$c^2$] | $< 10500$ | $0 - 10000$ | $-$ |
| | $D\mu$ mass interval [MeV/$c^2$] (Before vert.) | $< 11000$ | $< 10200$ | $-$ |

Table 22: Summary of Hlt2 and stripping selection requirements. The flight distance of a particle is given by the length of the separation between the production and its decay vertex. The flight distance $\chi^2$ is defined as the significance of this separation in units of $\chi^2$ and it is evaluated as $\vec{v}^T \cdot M \cdot \vec{v}$, where $M$ is the sum of the covariance matrices of the production vertex and decay vertex and $\vec{v}$ is the separation vector between the decay vertex and the production vertex. The $DIRA$ of a composite particle is defined as the cosine of the angle between its flight direction and its momentum vector. The $DOCA$ of two particles is the distance of closest approach between the flight direction vectors of the two particles.

### 8.3.3  Particle identification

On top of the particle identification selections applied at the Stripping level, additional selections have been placed during the offline data processing.

To suppress the contribution from the $\mu-$MisID background, the $\mu$ candidates are required to pass a $PID_\mu > 2$ selection and a $ProbNN_{\text{ghost}} < 0.2$.

In order to suppress the contribution from the $D$-combinatorial background, on top of the PID requirements based on DLL variables that are applied in the stripping, a tighter filtering is done basing on the ProbNN variables that identify pions and kaons. Specifically, the combination $ProbNNpik \equiv ProbNNpi \times (1 - ProbNNk)$ is used for the pions and the combination $ProbNNkpi \equiv ProbNNk \times (1 - ProbNNpi)$ is used for the kaons. Since the Monte Carlo is known not to properly reproduce the PID variable distributions, the requirements are optimised using the real data sample. For this purpose, the fraction of signal and non-$D^+$ background for a given PID requirement is determined through a fit to the $K\pi\pi$ invariant mass distribution. In order to model the background contribution, a linear polynomial with a free coefficient is used.

The pion and kaon PID variables are assumed not to be correlated with each other, and are thus optimised independently. Each variable is sampled along its range, applying the corresponding PID cut to data at each step and computing the associated signal significance, $S/\sqrt{S+B}$, where the signal $S$ and the background $B$ are determined from the three-body-mass fits. The maximum significance is achieved for $ProbNNpik > 0.14$ and $ProbNNkpi > 0.12$.

### 8.3.4  Decay kinematic and topology

The $D$-combinatorial background is further suppressed by applying a cut on the score of a BDT trained using $B \to D\tau\nu$ simulated events as signal and data mass sidebands $1830 < M(K\pi\pi) < 1840$ MeV/$c^2$ and $1900 < M(K\pi\pi) < 1910$ MeV/$c^2$ as background. The PID cuts on the PID variables described in the previous section are applied to the real data sample but not to the simulated one in order to avoid biases due to an incorrect description of the PID variables in the MC simulation. Half of the sample has been used in the training and half in the testing of the BDT performances. The variables used in the training of the BDT are:

- $\pi_1 \ \log p_T$: natural logarithm of the transverse momentum of the first pion

- $\pi_2 \ \log p_T$: natural logarithm of the transverse momentum of the second pion

- $K \ \log p_T$: natural logarithm of the transverse momentum of the kaon

- $\pi_1 \ IP^{PV}_{\chi^2}$: Significance of the impact parameter of the first pion track with respect to the associated primary vertex

- $\pi_2 \ IP^{PV}_{\chi^2}$: Significance of the impact parameter of the second pion track with respect to the associated primary vertex

- $K \ IP^{PV}_{\chi^2}$: Significance of the impact parameter of the kaon track with respect to the associated primary vertex

- $\pi_1$ Ghost Probability: Ghost Probability of the first pion track

- $\pi_2$ Ghost Probability: Ghost Probability of the first pion track
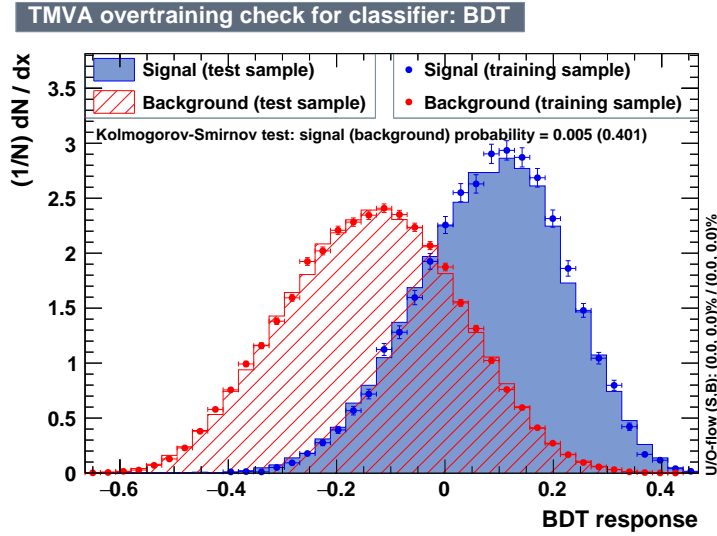
Figure 39: Output distributions and test of overtraining for the BDT used to suppress the non $D^+$ background.

- $K$ Ghost Probability: Ghost Probability of the first pion track

- $D$ $\chi^2_{\mathrm{VTX}}/d.o.f.$: $\chi^2$ of the fit to the $D$ vertex, normalized to its degrees of freedom

- $B$ $\log(d_{XY})$: Natural logarithm of he $B^0$ transverse flight distance

The output of the BDT, for training and testing samples superimposed, is reported in Fig. 39.

In order to find the optimal BDT cut, the signal significance defined as $S/\sqrt{S+B}$ is used as figure of merit, where S and B are the signal and background yields at each cut value, obtained from the fractions of simulated signal events and data sideband events that pass the corresponding BDT cut, together with the initial (pre-cut) yields for signal and background that are obtained through a $D^+$ mass fit, as described in the previous section. The significance is found to be maximal when the BDT output values are required to be larger than 0.23.

To suppress obvious $D$-combinatorial events, only events falling in between a $D^+$ mass window with boundaries of 1843 and 1897 MeV$/c^2$ are further analyzed. To reduce the $D\mu-$combinatorial background, instead, the invariant mass of the $D\mu$ system is required to be smaller than the physical mass of the $B$ meson ($m(D\mu) < 5.3\,\mathrm{GeV}\,/c^2$). This requirement has also been inverted during the analysis to define a control region of events with $m(D\mu) > 5.3\,\mathrm{GeV}\,/c^2$.

For events passing selections described up to this point, almost half of the combinatorial background has been found to be coming from random combinations of a $\mu$ track with a $D$ that has been produced in the primary vertex. It is fairly simple to remove this background by cutting on the significance of the impact parameter of the $D$ with respect to the associated primary vertex ($IP_{\chi^2}$[3]). This value is smaller for $D$ particles directly produced in the primary vertex with respect to the ones produced in a decay of a $B$ meson, as in the signal case. An excess

---

[3]The $IP_{\chi^2}$ of a track with respect to a vertex is the change in the vertex fit $\chi^2$ of the vertex reconstructed with and without the track under consideration.

attributed to combinatorial events with a prompt $D$ can be observed around $\log{(IP_{\chi^2})} \approx 0$. This contribution can be safely neglected by requiring the $D$ candidate to have a $\log{(IP_{\chi^2})} > 2.425$, and only combinatorial events with secondary $D$ candidates are expected after this selection.

### 8.3.5 Charged tracks Isolation

Charged tracks additional to the $D\mu$ candidates are suppressed by means of a MVA classifier developed for the $\mathcal{R}(D^*)$ measurement performed on RunI data [40].

Long, Upstream and VELO tracks present in the event are analyzed, and a BDT is evaluated using information coming from the kinematics of the track and quantities associated to the $B$ vertex before and after adding the new track to the vertex fit. The list of variables in input to the BDT is reported in Table 23. The BDT is evaluated on all the tracks and the BDT score of the most background like track is used as MVA output.

The BDT has been trained using associated tracks taken from the $D^{**} \to D^{(*)}$ decay in the simulated $B \to D^{**}\mu\nu$ events and unassociated tracks taken from $B \to D\mu\nu$ events, with the signal decay excluded.

The cut value on the MVA output is optimized for this analysis by comparing the MVA output distribution of a simulated background sample of $B^0 \to DDX$ events with the distribution of a simulated signal sample of $B \to D\mu\nu$ events. The distributions are shown in Fig 40. The ROC curve is shown on Fig. 41 and the working point is chosen where the signal efficiency starts to drop, corresponding to a threshold value of 0.32.
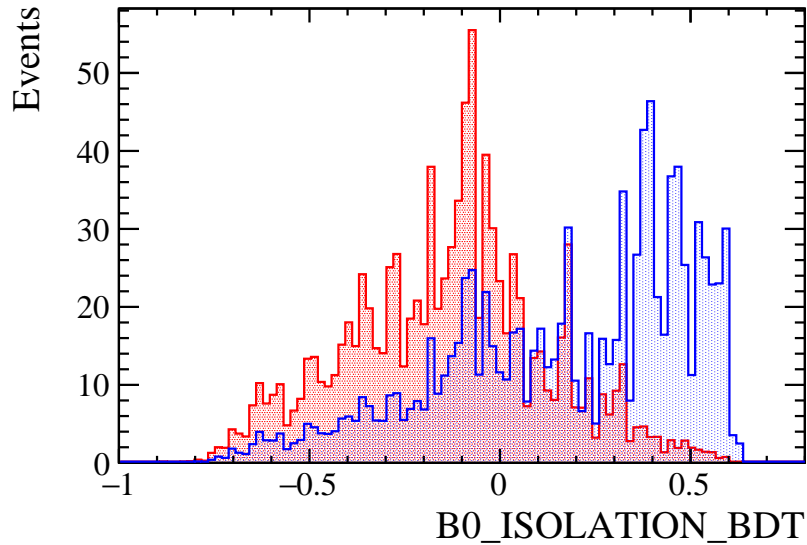


Figure 40: MVA output distribution for signal (red) and background (blue) simulated sample.

In addition to the MVA output, the BDT score along with additional information is stored for the three most background like tracks. This information can be further analyzed to define control samples to be used in the fit, as it will be described later on in the thesis.
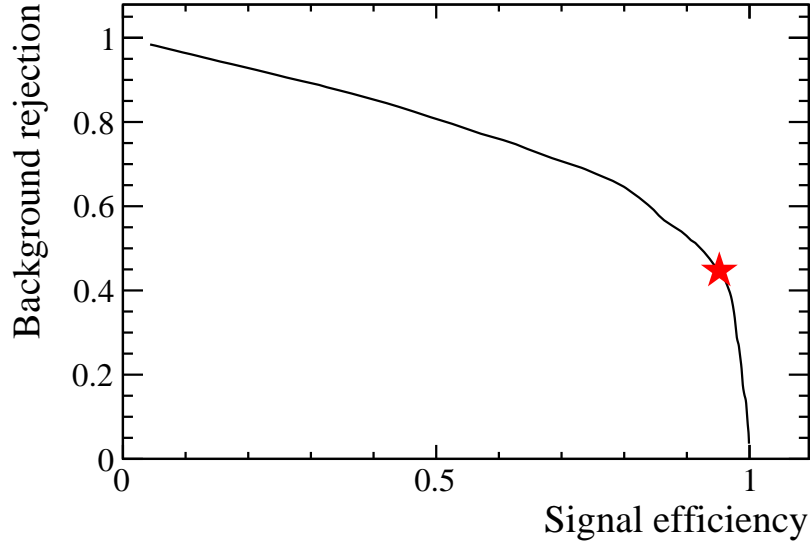
Figure 41: ROC curve for the considered samples. The optimal working point is indicated by the red star.

| variable | description |
|----------|-------------|
| $min(IP\chi^2_{PV})$ | Minimum of the $IP\chi^2$ of the track with respect to all the primary vertices. |
| $IP\chi^2_{SV}$ | $IP\chi^2$ of the track with respect to the $B$ vertex. |
| $p_T$ | Transverse momentum of the track. |
| $\cos(D\mu, \text{track})$ | Cosine of the angle between the visible momentum and the track direction. |
| $FD\chi^2$ | $\chi^2$ of the Flight distance of the refitted $B$ vertex with respect to the PV which has the minimum $IP\chi^2$ with respect to the track. |
| $\Delta(FD\chi^2)$ | The difference between the $FD\chi^2$ after and before including the track in the $B$ vertex fit. |

Table 23: Input variables to the Charged Isolation BDT.
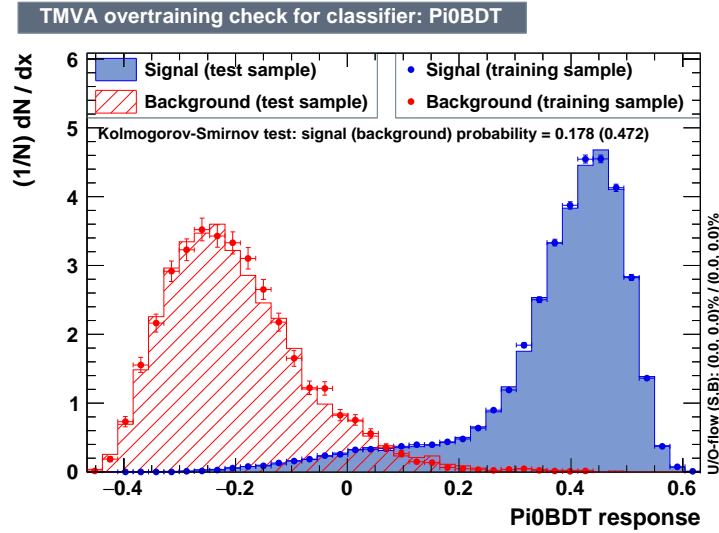
Figure 42: Output distributions and test of overtraining for the resolved $\pi^0$ BDT.

### 8.3.6 Neutral particles Isolation

Isolation against events with accompanying neutral pions and photons can also help reduce the background of partially reconstructed decays. Two Multivariate isolation tools are first evaluated on each event and then combined in a single neutral isolation response. The two tools are taking input from different objects. The first one is a BDT which is associated to each neutral pion in the event which have decayed into two photons that have been resolved and reconstructed as separate clusters in the ECAL (resolved $\pi^0$). The BDT is trained using resolved $\pi^0$s from $B \to D\mu\nu$ Full MC as false positives and truth matched resolved $\pi^0$s from $B \to (D^* \to D\pi^0)\mu\nu$ Full MC as true positives. Half of the datasets have been used during the training while the other half is kept for the validation of the BDT. The variables used to train the BDT, which will be called *Resolved Neutral Pions BDT* from now on, are:

- The invariant mass of the resolved $\pi^0$;

- The difference in invariant mass between the $D^+$ and the $D^+\pi^0$ candidate;

- The natural logarithm of the $pi^0$ transverse momentum;

- The natural logarithm of the $\pi^0$ longitudinal momentum;

- The confidence level of each of the two daughter photons.

The output of the Resolved $\pi^0$ BDT for the training and testing samples, is shown in Fig. 42. No overtraning of the BDT is observed.

The second isolation tool is a BDT that associates for each event the probability of having an additional neutral particle that has not been associated to the decay. The input to the BDT is taken from information from neutral cones constructed around the $D^+$ flight direction. The radius of the cone is defined as $\rho = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. Neutral cones with different radii,
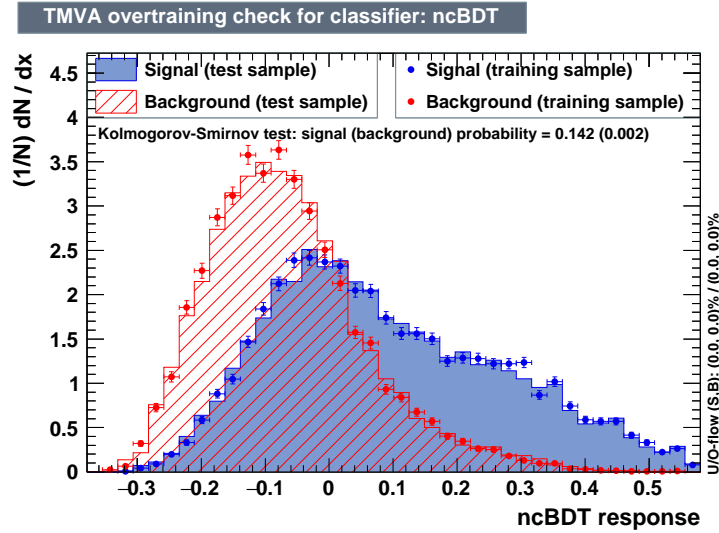
Figure 43: Output distributions and test of overtraining for the neutral cone BDT.

$\rho \in [0.1, 0.2, 0.3, 0.4, 0.5, 0.6]$ have been investigated. The energy of the neutral object within the cone is summed and a BDT for each cone radius has been trained using the following variables:

- The natural logarithm of the cone's longitudinal momentum;

- The photon multiplicity of the cone;

- The average photon confidence level in the cone;

- The difference in pseudorapidity between the 3-momentum of the $D^+$ and the one of the neutral cone;

- The difference in the azimuthal angle between the 3-momentum of the $D^{*+}$ and the one of the neutral cone.

The best cone radius has been chosen by comparing the integral of the ROC curves and has been found to be $\rho = 0.2$. The BDT, which will be called from now on *Neutral Cone BDT*, has been trained using $B \to D\mu\nu$ Full MC events as signal and $B \to (D^* \to D\pi^0)\mu\nu$ Full MC events as background. The output of the Neutral Cone BDT for the training and testing samples, is shown in Fig 43.

No overtraining of the BDT has been observed.

The two isolation tools are then combined into a single, per-event, *neutral isolation BDT*. This BDT takes as input:

- The number of resolved $\pi^0$ that have the resolved $\pi^0$ BDT output smaller than 0;

- The output of the neutral cone BDT.

|  | (a) | (b) |
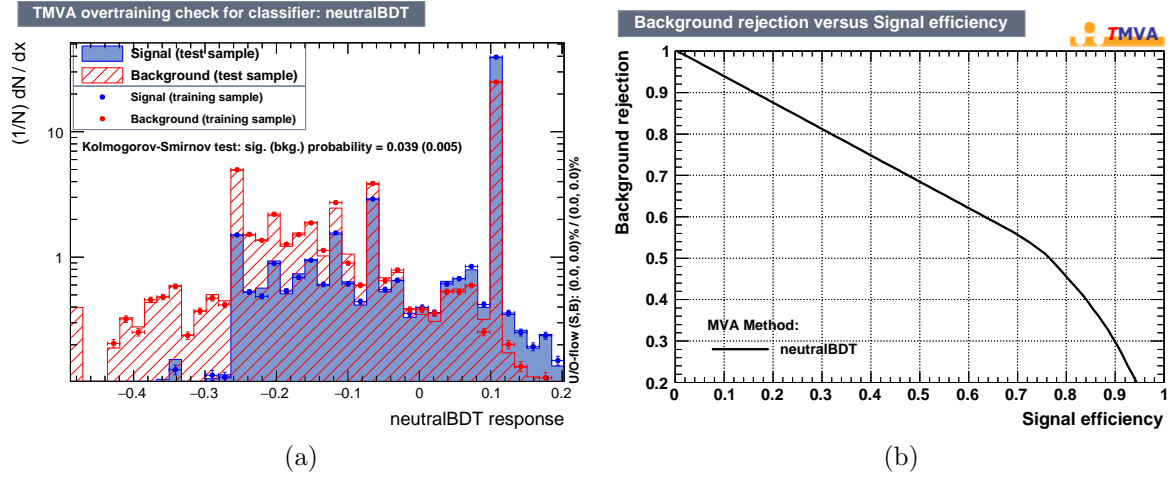
Figure 44: Performance of the final BDT used for neutral isolation. Figure (a) shows the output distributions, in logarithmic scale, and the test of overtraining. Figure (b) shows the associated ROC curve.

This BDT is trained with the same signal and background samples as before, but without any truth matching applied to them. This situation represents more accurately the one found in real data and the overall performances get worse due to the low neutral reconstruction efficiency of the LHCb detector.

The BDT output distributions for the training and testing samples are shown in Fig. 44a. No overtraining is observed. The corresponding ROC curve is shown in Fig. 44b. The working point is set at a signal efficiency level of 90%, which corresponds to a BDT output requirement of $> -0.16$ and a background rejection power of 30%.

### 8.3.7 Cross-Feed vetoes

The cuts explained in 8.3.3 have been optimized to suppress the background of $D$-combinatorial events. Another source of background could be due to other charm hadron decays in which one or multiple $D$-daughter tracks have been assigned a wrong mass hypothesis. This background would be peaking in the distribution of the mass of the $D^+ \to K^-\pi^+\pi^+$ candidate, and is called Cross-feed background.

In order to check if these backgrounds are present after all the selections described before, a check has been performed in which the mass hypothesis of one of tracks $D$ daughters tracks is changed, looking for peaks in the reconstructed mass distribution.

Given the charge of the tracks involved in the decay, no cross-feed from $D_s^+ \to K^+K^-\pi^+$ and from $\Lambda_c^+ \to pK^-\pi^+$ decays by the mis-identification of two tracks simultaneously can be expected. The probability of a simultaneous misidentification of three tracks is considered small enough to be neglected. This leaves out the cross-feed with the misidentification of one single track as the only possible source of this background.

Cross-feeds coming from $D_s^+ \to K^+K^-\pi^+$ decays can originate from the misidentification of one kaon as one of the two pions in the event. This can be looked for by changing the mass

hypothesis of one of the two pions in the event to the one of a $K$ and looking for a peak around the $D_s$ mass.

Another possible cross-feed background is from $\Lambda_c^+ \to pK^-\pi^+$ decays, originating from the misidentification of a proton as one of the two signal pions. This can be looked for by changing the mass hypothesis of one of the two pions in the event and looking for a peak around the $\Lambda_c$ mass.

This procedure is repeated for events passing all the selections described up to now for all the real data samples, defined in Section 8.1, in both 2015 and 2016 data taking conditions.

A peak corresponding to $\Lambda_c$ cross-feeds has been observed in the data, as can be seen in Figure 45. Both the 2015 and 2016 data show an excess around the nominal $\Lambda_c$ mass, when changing the mass hypothesis of one of the two pions into a proton.

By defining a window of $\pm 15$ MeV around it, it is possible to compare the distribution of the $ProbNNp$ of the corresponding pion for events falling inside or outside this region, as reported in Fig. 45



(a) 2015, RS data sample



(b) 2016, RS data sample
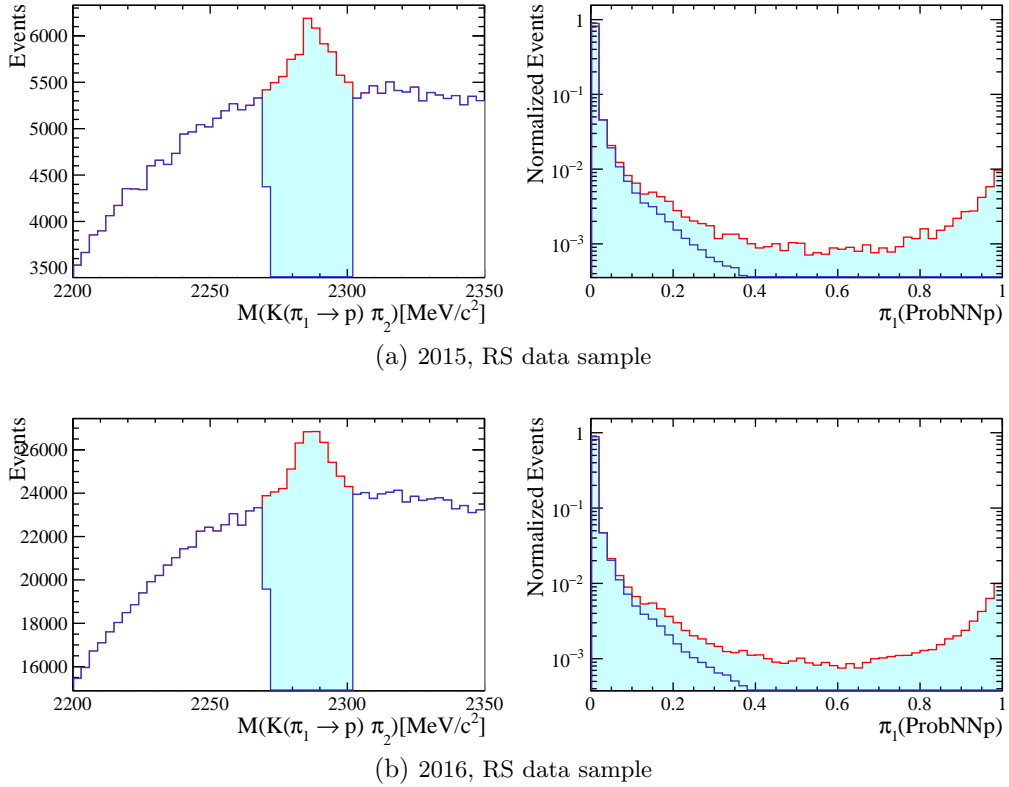
Figure 45: Left: distribution of the $K\pi\pi$ mass when changing the mass hypothesis of the first pion to a proton. Right: ProbNNp distribution for the first pion for events falling inside a $\pm 15$ MeV mass window around the $\Lambda_c$ mass (cyan) and for the ones falling outside the mass window (blue).

This distribution clearly shows that events falling within the $\Lambda_c$ mass window have a bigger tail of high $ProbNNp$ for the pions, when instead the distribution for the events falling outside the window is peaked towards low values.

To reject this background, a cut of $ProbNNp < 0.4$ is put on both signal pions, in all the

data samples. The $K\pi\pi$ mass distribution, when changing the mass hypothesis of either of the signal pions, is shown in Figures 46 and 47, for 2015 and 2016 datasets respectively. The dark blue histogram represents the events without any cross-feed veto, whereas the cyan histogram represents events that pass the veto selection.

No contribution from $D_s \to KK\pi$ decays has been instead observed and, therefore, no veto selections against it is applied. The $K\pi\pi$ mass distribution, when changing the mass hypothesis of one of the two signal pions to a kaon, is shown in Figures 48 and49, for 2015 and 2016 respectively. The color scheme is the same as before.
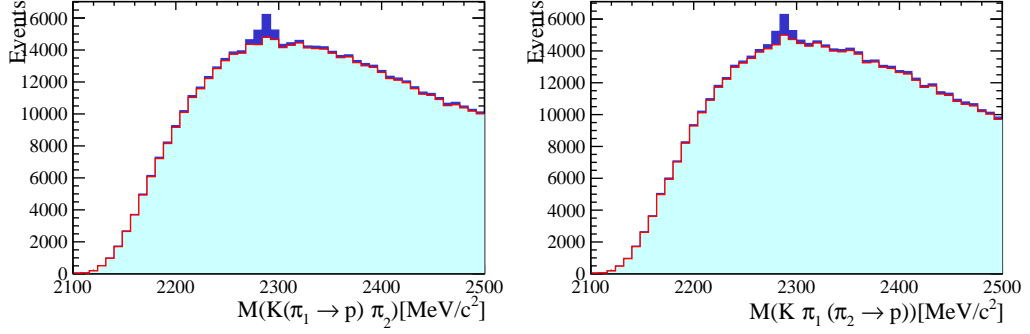
(a) RS data
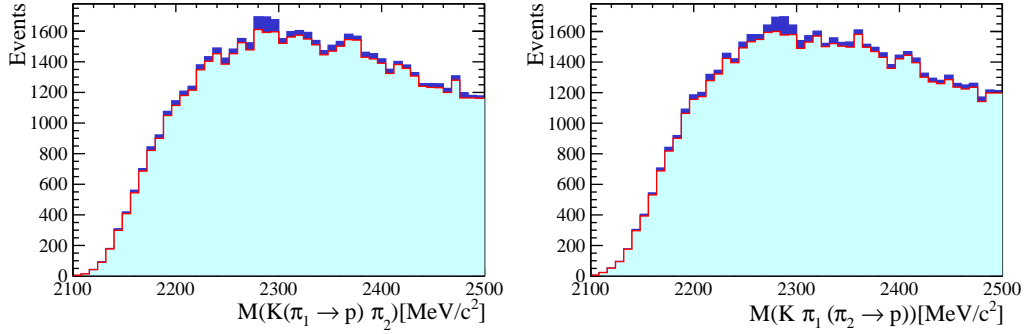


(b) WS data
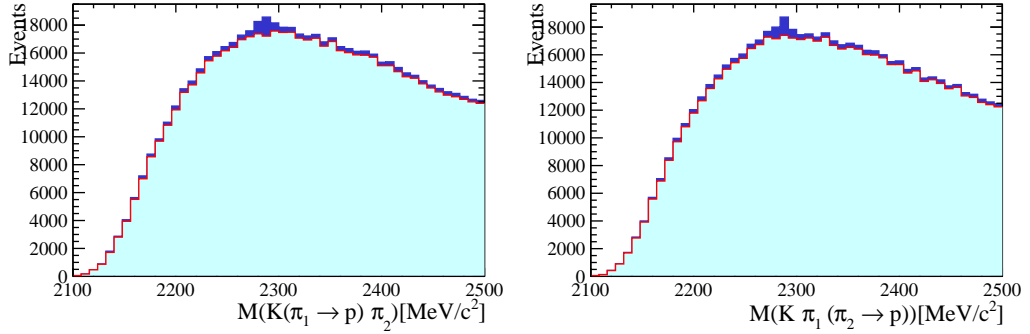


(c) MisID,RS data
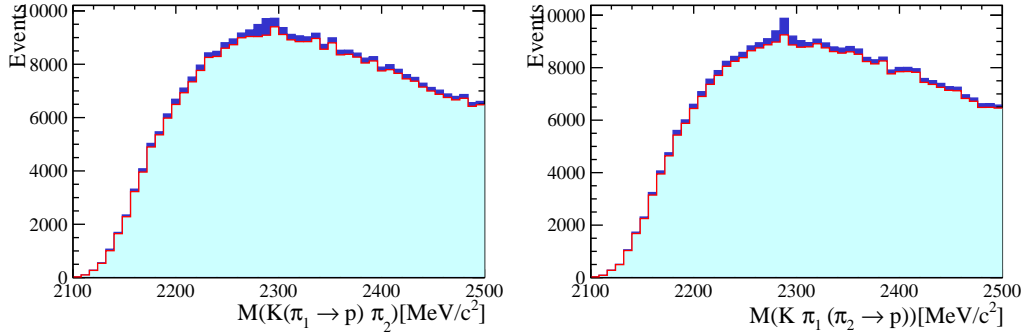


(d) MisID,WS data

Figure 46: Distribution of the $K\pi\pi$ mass when changing the mass hypothesis of one of the signal pions to a proton, 2015 data set.

96
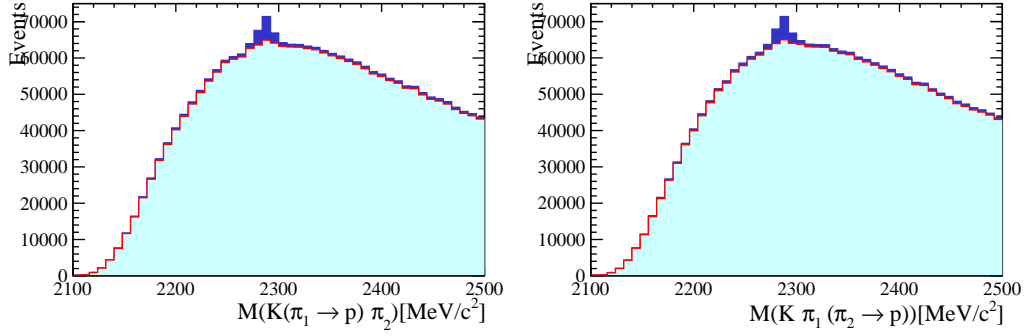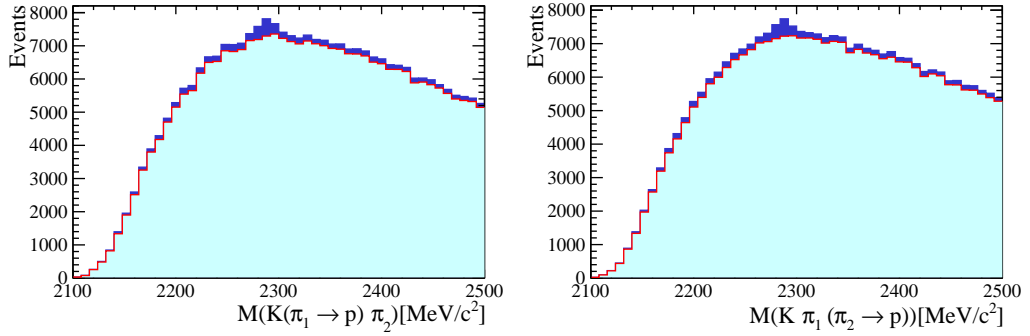
(a) RS data

(b) WS data

(c) MisID,RS data

(d) MisID,WS data

Figure 47: Distribution of the $K\pi\pi$ mass when changing the mass hypothesis of one of the signal pions to a proton, 2016 data set.
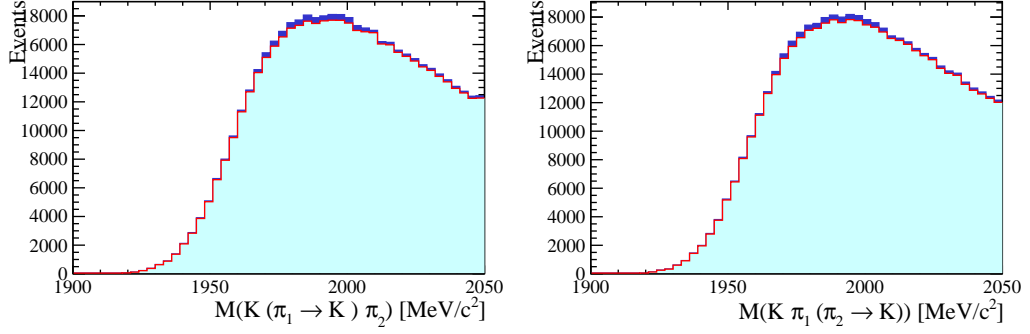
(a) RS data
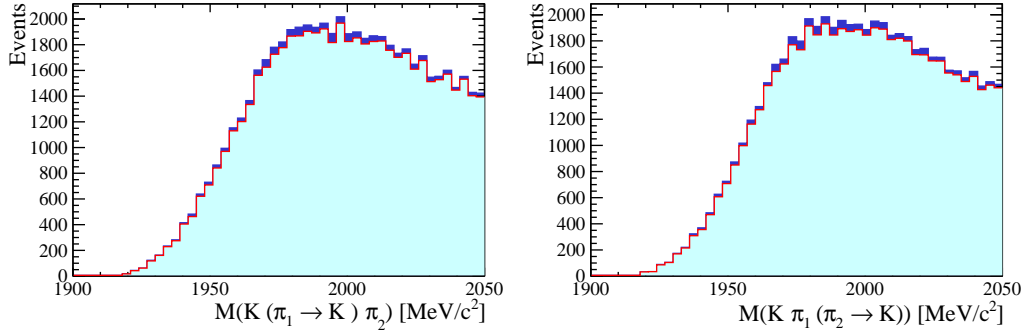


(b) WS data
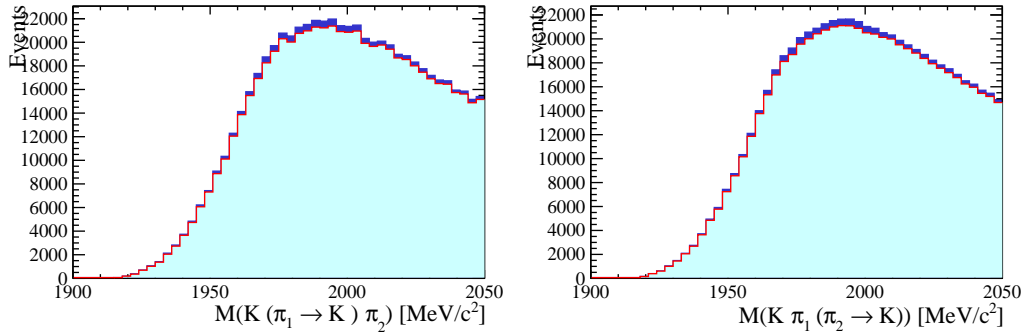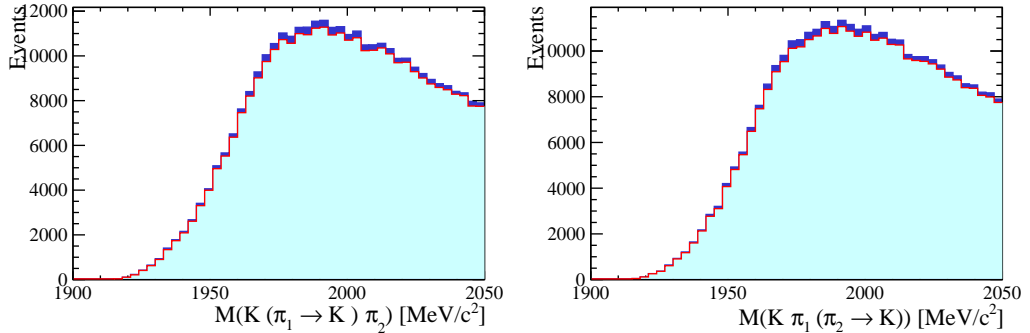


(c) MisID,RS data



(d) MisID,WS data

Figure 48: Distribution of the $K\pi\pi$ mass when changing the mass hypothesis of one of the signal pions to a kaon, 2015 data set.
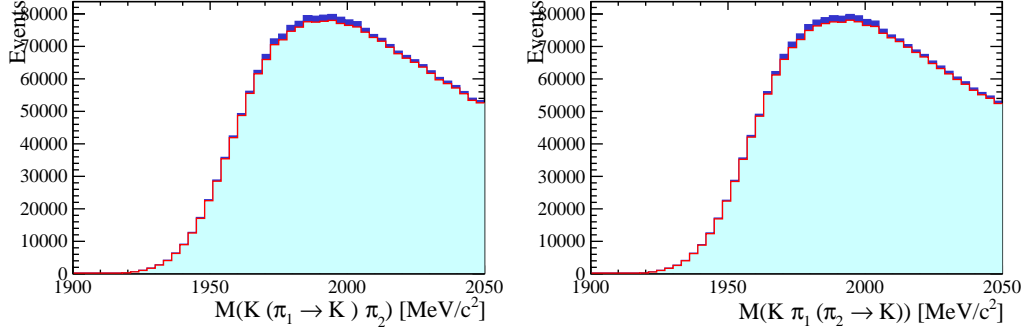
(a) RS data
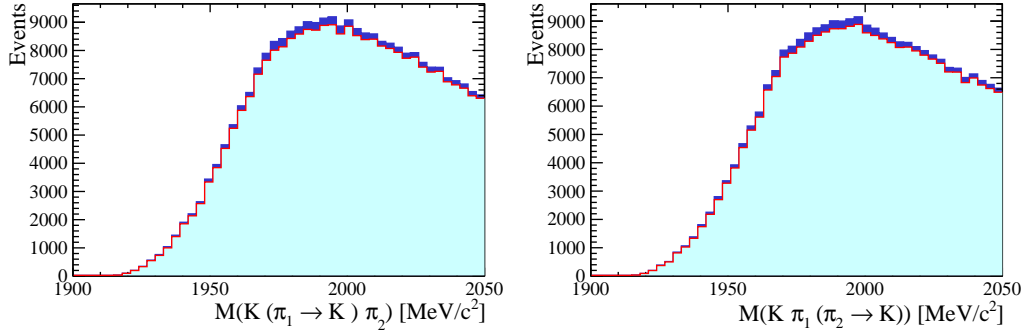


(b) WS data
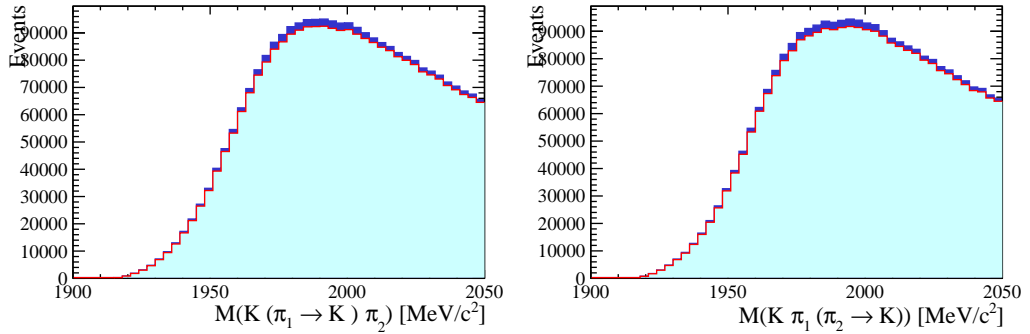


(c) MisID,RS data
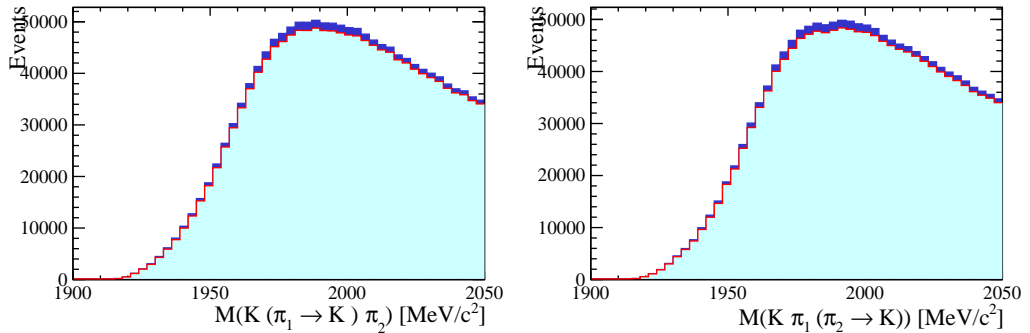


(d) MisID,WS data

Figure 49: Distribution of the $K\pi\pi$ mass when changing the mass hypothesis of one of the signal pions to a kaon, 2016 data set.

# 9 Control regions

The sample constructed from the events passing all the selections described in the previous section constitutes the one in which the signal $B \to D\tau\nu$ decays are expected to contribute the most. This region of the data is the one that will be used as main driver of the sensitivity on $\mathcal{R}(D^+)$. It will be referred to as the signal region or Isolated sample (ISO) from now on.

Apart from the signal region, additional control regions have been defined in order to calibrate and study the background events that pollute the measurement. The events in these control regions are required to pass all the nominal requirements except for some specific selections which have been inverted with respect to the signal region in order to target specific background contributions.

The analysis presented in this thesis heavily relies on the validation that can be performed thanks to the control regions: the underlying assumption is the one that a good modelling of the data in the control regions, which are enriched of specific background events, directly reflects in a good modelling of the same background in the signal region.

Physical backgrounds usually come accompanied by additional (charged or neutral) particles which have been generated in the $B$ decay vertex but have not been reconstructed. To select these events, the charged and neutral isolation selections is inverted in the control regions. Thanks to the information saved by the charged isolation tool, several control regions, with a different number of additional charged tracks can be defined in order to define orthogonal regions aimed at the study of specific background processes. Additional requirements on the charge or PID of the additional charged tracks that fail the charged isolation requirement can be applied.

When additional charged tracks are being selected to define the control regions, they are always required to be long tracks, to have a $ProbNNGhost < 0.2$ and to not have been identified as a muon (`isMuon`= 0) and to have a charged isolation BDT score higher than 0.35. They are also classified to be either *pion-like*, when $ProbNN\pi > ProbNNk$, or *kaon-like*, when $ProbNNk > ProbNN\pi$.

All the signal and control regions have been defined to be orthogonal to each other, in the sense that an event being selected in one of the control regions will not be selected by any of the others. In this way these regions can be simultaneously employed in the fit, without risk of double counting events.

Finally, two validation regions have also been defined. These regions have been defined only to check the agreement between the data and the model being used to fit it. These validation regions have not been used in the simultaneous fit to data, and their specific usage will be described later on in the thesis.

The selections employed in all the control and validation regions, along with their purpose, will be described in the following sections and are reported in Table 24.

## 9.1 One charged pion (1OS) control sample

This region is defined by inverting the charged isolation by requiring the presence of one anti-isolated track being compatible with the pion hypothesis. The charge of the track is required to be opposite to the one of the $D^+$. The decay vertex is required to be isolated with respect to all the other charged tracks in the event.

This region is enriched in contributions coming from 1P $D^{**}$ states, which are expected to decay mostly with the presence of one additional pion into a $D^+$. The requirement of the

| Region | Isolation selection | Anti-isolated track selection | $m(D\mu)$ selection |
|---|---|---|---|
| 1OS | `IsoBDTHighest` $> 0.35$ & `IsoBDTSecondHighest` $< 0.32$ | `isLong ==1` & `ProbNNghost < 0.2` & `isMuon ==0` & `ProbNN`$\pi$ `< ProbNNK` | $< 5.3~\mathrm{GeV}^2/c^4$ |
| 2OS | `IsoBDTHighest` $> 0.35$ & `IsoBDTSecondHighest` $> 0.35$ & `IsoBDTThirdHighest` $< 0.1$ | `isLong ==1` & `ProbNNghost < 0.2` & `isMuon ==0` & `ProbNN`$\pi$ `< ProbNNK` | $< 5.3~\mathrm{GeV}^2/c^4$ |
| DD | `IsoBDTHighest` $> 0.35$ or `IsoBDTSecondHighest` $> 0.35$ or `IsoBDTThirdHighest` $> 0.35$ | `isLong ==1` & `ProbNNghost < 0.2` & `isMuon ==0` & `ProbNNK > ProbNN`$\pi$ | $< 5.3~\mathrm{GeV}^2/c^4$ |
| $\pi^0$ | `NeutralBDT` $< 0.26$ | N/A | $< 5.3~\mathrm{GeV}^2/c^4$ |
| high $B^0$-mass | Same as ISO | N/A | $> 5.3~\mathrm{GeV}^2/c^4$ |

Table 24: Summary of control and validation regions definitions. `IsoBDTHighest`,`IsoBDTSecondHighest` and `IsoBDTThirdHighest` refer to the value of the Charged Isolation BDT for the most, the second to most and the third to most anti-isolated track, respectively.

relative charge of the pion with respect to signal $D^+$ candidate, selects the charge of all possible contributions in this region. This region is important in the analysis in order to control the modelling of the Feed-Down and the branching fractions that make up for this background. The fractions for the various $D^{**}$ 1P states take the majority of their sensitivity from this region, and are then converted in the other regions, taking efficiency ratios from MC and isospin relations into account, as it will be explained later on.

## 9.2 Two charged pions (2OS) control sample

The events falling inside this region are required to have exactly two charged tracks failing the charged isolation requirement. The two tracks are required to be compatible with the pion hypothesis and to have the opposite charge with respect to each other. In this way this region is enriched in decays of Feed-Down events in which the $D^{**}$ decays with multiple pions in the decay. This region has been designed specifically to target Feed-Down background events with $D^{**}$ states with a higher mass with respect to the 1P states ($B \to D_J^{**}\mu\nu$). The decays of these states are not precisely known in the literature and for this reason a phenomenological shape correction is applied in the fit, as it will be described later on. The nuisance parameters controlling this correction take their sensitivity from the data of this control region.

In order to suppress the big amount of MisID polluting this region, the third most anti-isolated track for the events falling in this region is required to pass a BDT tighter than the nominal one, with a charged isolation score $< 0.1$ and the mass difference $m(D^+\pi^+\pi^-) - m(D^+)$ is required to be smaller than $1400\,\mathrm{MeV}/c^2$.

## 9.3 Double Charm (DD) control sample

This region has been designed to be enriched in Double Charm decays ($B \to D^{(*,+)}(H_c \to \mu\nu X)(X')$), in which the $\mu$ is coming from an additional charmed hadron produced in the $B$

decay. The decays of the charmed mesons usually come accompanied by kaons, thanks to the high branching fraction for $V_{cs}$ induced transitions. The events falling in this region are required to have at least one of the three most anti-isolated charged tracks failing the charged isolation selection and being compatible with the kaon hypothesis.

As it will be described more in detail later on, the template for the Double Charm events is divided into two-body decays $(B \rightarrow D^{(*,+)}(H_c \rightarrow \mu \nu X))$ and multi-body decays $(B \rightarrow D^{(*,+)}(H_c \rightarrow \mu \nu X)X')$. The multi-body decays are known to be poorly measured and their shape is therefore varied in a phenomenological way. The shape parameter, the fractions of multi-body and two body decays and the fraction of Double Charm events from $B^0$ or $B^\pm$ are taking their highest sensitivity in the measurement from this region.

## 9.4   Neutral pion $(\pi^0)$ control sample

The selection in place targets at selecting $B \rightarrow D^+ \ell \nu$ events with high efficiency. Due to the irreducible presence of events of $B \rightarrow D^{*+} \ell \nu$, with the $D^{*+} \rightarrow D^+ \pi^0$, the $\mathcal{R}(D^*)$ parameter has to be measured simultaneously with $\mathcal{R}(D^+)$. In order to increase the sensitivity to $\mathcal{R}(D^*)$, the events selected by this region are required to fail the neutral isolation selection, therefore enhancing the presence of neutral pions accompanying the reconstructed signal.

The cut on the neutral isolation in this region is inverted and a tighter anti-isolation cut of $< -0.26$ with respect to the direct cut is employed to enhance the purity of the control region.

It has to be noticed that no explicit reconstruction of the $\pi^0$ has been performed in this region.

## 9.5   high-$B^0$ mass validation sample

This validation sample shares all the selections with the signal region, except the cut on the mass of the visible system, which is inverted with respect to the signal region $(m(D\mu) > 5.3 \, \text{GeV}^2/c^4)$. Thanks to this cut, no physical background has to be expected in this validation region: the only contributions expected in this region are the ones of the $D\mu$-combinatorial and $\mu-$MisID events.

This region is extensively used to estimated and validate the shape of the $D\mu$-combinatorial background, as it will be explained later on.

## 9.6   Normalization enriched validation sample

This validation region is defined as a subregion of the signal sample, and it is used in order to validate the modelling of the MC simulation. It is defined as a region in which the signal is not expected to contribute and the background contribution is expected to be small. This sample is defined by selecting the left tail of the distribution of the $m^2_{miss}$ variable, as defined in Section 7.2, where the normalization $(B \rightarrow D^{(*)}\mu\nu)$ events vastly dominate.

This validation region has been used to check the detector modelling of the MC simulation by looking at a cocktail of events obtained by projecting the fit result in this region, as it will be explained later in the thesis.

# 10 Emulation of missing features in the Tracker Only simulation

As explained in Sec. 7.4, in order to reduce the systematic uncertainty due to the statistics of the MC samples used to construct the templates used in the fit for the physical backgrounds, a fast simulation in which only the response of the tracking subdetectors has been simulated has been employed.

For this reason, many of the features of the event that are accessible in real data, are not available in Tracker Only MC. Some selections must then be emulated offline, relying only on the response of the tracking subdetectors.

In the following sections, the emulation of these selections is reported.

## 10.1 PID response

In tracker only simulation, the response of all detectors from which the PID information is extracted (RICH, muon stations, ECAL and HCAL) is not simulated. For this reason, the effect all the PID selections described in the previous section have to be emulated in Tracker Only MC.

Furthermore, the simulation is known to not fully well reproduce the PID related quantities especially the ones related to the response of the RICH gas at the passage of charged particles. In fact the RICH gas response is very sensitive to pressure and temperature variations. For these reasons, a calibration of the PID response coming from real data for both the Tracker Only and Full MC samples is needed.

The PID selections are calibrated through a data-driven method using real data calibration samples of $D^{*+} \rightarrow D^0(K^-\pi^+)\pi^+$ and $J/\psi \rightarrow \mu^+\mu^-$ samples in which the background has been removed by means of the sPlot technique [93]. Using these samples, the efficiency of the PID selections has been measured using a *tag-and-probe* method, as a function of some variables.

The variables chosen for the calibration follow from the fact that the ProbNN and $\Delta \log \mathcal{L}$ variables mostly depend on the momentum of the final-state particle, due to the correlation with the emission angle of the Cherenkov photons and the fact that a $\mu$ has a different probability of traversing all the Muon stations depending on its momentum, on the pseudorapidity of the particle, since the RICH detectors have been designed with different angular acceptances and optimised for different momentum regions, and on the detector occupancy.

The PID efficiencies are evaluated from the calibration samples in bins of $p, \eta$ and the number of tracks in the event (`nTracks`). This is done by applying the PID selections to the calibration samples and comparing the original number of events in a given bin with the number of events in that bin surviving the PID selections.

The binning scheme used for all particles except for the signal $\mu$ is reported in Table 25. A finer binning for the $\mu$ signal tracks has been used, as reported in Table 26.

| Variable | Bin edges |
|:---:|:---:|
| $|\vec{p}|$ | [3., 6., 9., 15., 30., 50., 100.] GeV/$c$ |
| $\eta$ | [1.5, 2.5, 3.5, 5.] |
| nTracks | [0, 225, 500] |

Table 25: Bin edges used for the PID efficiencies. `nTracks` refers to the number of tracks in the event.

| Variable | Bin edges |
|---|---|
| $\|\vec{p}\|$ | [3., 5., 6.5, 9.3, 10.2, 12., 13.2, 15.6, 19., 24.4, 29.8, 35.2, 46., 51.4, 56.8, 62.2, 67.6, 73., 78.4, 83.8, 89.2, 94.6, 100.] GeV/$c$ |
| $\eta$ | [1.5, 1.9735, 2.375, 2.8125, 3.25, 3.6875, 4.125, 4.5625, 5.] |
| nTracks | [0, 225, 500] |

Table 26: Bin edges used for the PID efficiency to reconstruct pions, kaons, protons and electrons as muons. `nTracks` refers to the number of tracks in the event.

A per-event weight, corresponding to the product of the PID efficiencies evaluated from the calibration samples for each track of the event is associated to each MC event.

A complication arises when evaluating the PID selection efficiency for the additional tracks in the event selected by inverting the charged isolation selection. In order for the control regions to be orthogonal to each other, each MC event not passing the charged isolation selection should be selected exclusively by one of the control regions or be discarded. Weighting the events would not be sufficient and MC events falling in multiple control regions, albeit with different weights, would spoil the orthogonality of the control regions, mining the assumptions needed for a simultaneous fit. For this reason, for the additional tracks selected by inverting the charged isolation selection, also the decision on the category to which the event should be assigned has to be emulated.

As described in Sect. 9, the tracks failing the charged isolation selection are required to fail the muon identification. This selection is assumed to be 100% efficient in removing true muons. Furthermore, they are required to have small probability of being ghost tracks; also this cut is assumed to be 100% efficient in removing true ghosts from the events. For this reason events in which the extra track failing the isolation requirement has a true ID of either a true muon or a true ghost are removed from the simulation samples.

For the remaining events, the following procedure is applied:

- for each track two efficiencies are computed: the efficiency for the *pion-like* and for the *kaon-like* selection, as described in Sect. 9;

- since the *pion-like* and *kaon-like* selections on the additional tracks are orthogonal to each other but do not cover the entire phase space, they are used to compute per event probabilities for each track to be selected as a pion, as a kaon or to fail the PID selections:

$$p_{\pi-\text{like}} = \varepsilon_{\pi-\text{like}}, \tag{149}$$

$$p_{K-\text{like}} = \varepsilon_{K-\text{like}}, \tag{150}$$

$$p_{\text{fail}} = 1 - \varepsilon_{\pi-\text{like}} - \varepsilon_{K-\text{like}}; \tag{151}$$

- finally a PID decision for each track is sampled from these probabilities.

This procedure has been followed for the three most anti-isolated tracks in the event with respect to the charged isolation tool. In this way all the Tracker Only MC events are assigned uniquely to one of the control regions used in the fit.

## 10.2 Neutral isolation

Due to the absence of the simulation of the ECAL and HCAL calorimeters, the neutral isolation cannot be directly evaluated in the Tracker Only MC. The efficiency of this selection has been observed to induce an (albeit small) acceptance effect on the fit variables. The effect of this acceptance has been emulated via a reweighting procedure.

The emulation is performed on variables constructed using tracking information of all the signal particles and the truth level four momentum of all true photons emitted in the decays. The variables used are:

- the angle ($\alpha$) between the $D^+$ flight direction and the beam line, defined in the laboratory frame;

- the angle ($\beta$) between the $D^+$ flight direction and the sum of the three momenta of the true photons emitted in the decay;

- the invariant mass of the emitted photons.

A *bin-by-bin* reweighting is trained on a combination of $\overline{B}^0 \rightarrow D^+ \mu \bar{\nu}_\mu$, $\overline{B}^0 \rightarrow D^{*+}(D^+\pi^0)\mu\bar{\nu}_\mu$ and $\overline{B}^0 \rightarrow DDX$ (double-charm cocktail) Full MC samples, before and after having applied the neutral isolation selection on them. The weights are then re-scaled by the average efficiency found in the combination of the training samples.

For the $\pi^0$ control region, the procedure followed is the same as the one described above, but with the neutral isolation cut inverted.

## 10.3 L0 emulation

The L0 trigger selection used in this analysis is a combination of the `L0Hadron` line, being TOS on the hadronic part of the event, and the `L0Global`, being TIS on the whole signal.

The effect of the two requirements are emulated using procedures that are going to be explained in this section, and then combined into a single correction to be applied to the Tracker Only MC samples used in the analysis.

The `L0Hadron` trigger, as explained in 6.3.4, is an algorithm that looks for $2 \times 2$ HCAL clusters and requires the $E_T$ of the sum of the clusters to exceed a given threshold. If the cluster is matched to a signal track extrapolated to the HCAL plane, then the event is said to be TOS with respect to the `L0Hadron` trigger selection.

The challenge for an emulation of the response of this trigger using only information coming from the tracking system lies in the fact that clusters due to energy deposits in the HCAL are broad and therefore the effect of multiple particles cannot be factorised. The emulation is therefore done in steps of complexity.

First, the cluster deposits charged final state hadrons interacting with the HCAL are emulated from simulated events in which single pions are produced. This is done by storing the difference between the $E_T$ as measured by the tracker (denoted as $E_T^{trk}$) and the $E_T$ as measured by the HCAL, denoted as $E_T^{HCAL}$, as a function of kinematics. Example distributions for two kinematic bins are shown in Fig. 50.

Often, no energy recorded by the HCAL is available, which corresponds to events falling into the right peak in the resolution distribution. The energy of the signal tracks in the Tracker Only simulation is smeared by random sampling from this distribution.
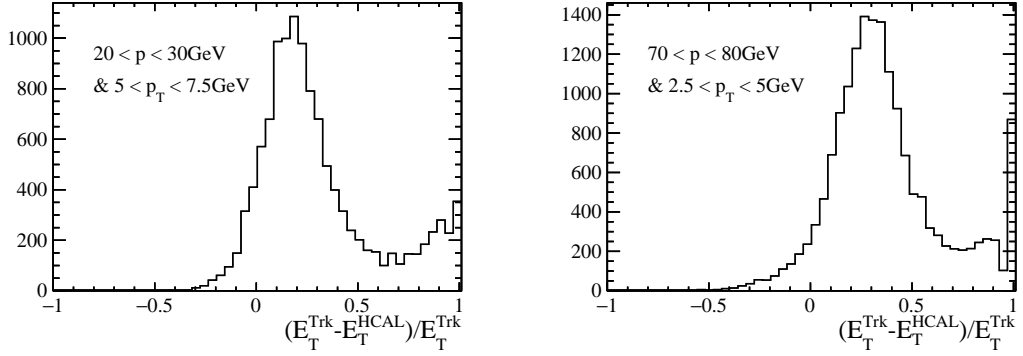
Figure 50: HCAL $E_T$ resolution for a couple of example kinematic bins.
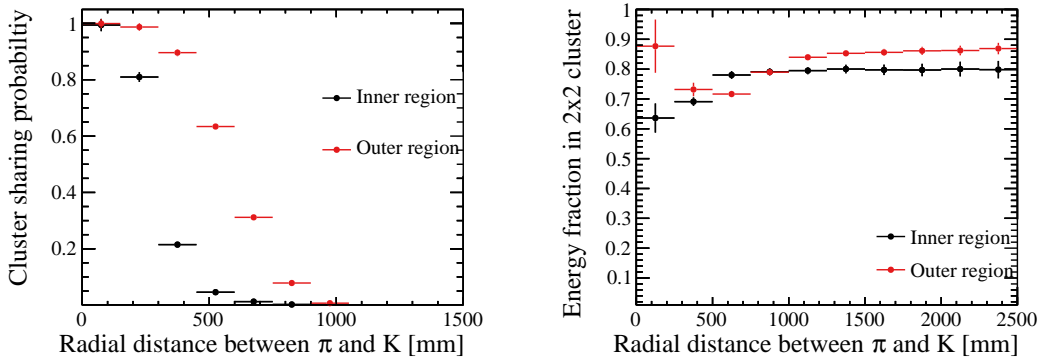


Figure 51: Probability to share an HCAL cluster (left) and fraction of energy contained in a 2x2 cluster (right) as a function of the radial distance between the kaon and pion on the HCAL surface.

When having multiple tracks in the event, the clusters belonging to the two particles end up overlapping in the HCAL, due to its coarse granularity and the shape of the hadronic shower. This can have the effect of enhancing the trigger efficiency, especially when neither of the clusters alone would have enough energy to trigger the selection by itself. The effect of this is studied using $D^0 \to K\pi$ simulated events. The probability of sharing a cluster is measured as a function of the separation of the impact point of the two tracks on the HCAL surface, as extrapolated from the tracking system. This is done separately for the inner and the outer region of the HCAL, due to the different granularity of the detector. As shown in Fig. 51, the fraction of merged clusters is higher in the outer region with respect to the one observed in the inner region.

The sum of the energies of the emulated clusters in the Tracker Only MC samples due to different signal tracks are summed up together taking into account the probability of two clusters to overlap as a function of their separation on the surface of the HCAL.

It has to be noticed that clusters that are the result of the overlap of multiple single particle deposits tend to be broader with respect to single particle clusters. This means that the fraction of energy contained in a $2 \times 2$ cell tends to be smaller in this case, as shown in Fig. 51. This is corrected for by evaluating the fraction of energy in $2 \times 2$ cells for merged clusters in the $D^0 \to K\pi$ sample and applying a proper correction to the Tracker Only MC samples.

Up to now, the emulated energy deposit takes into account only signal particles. In real
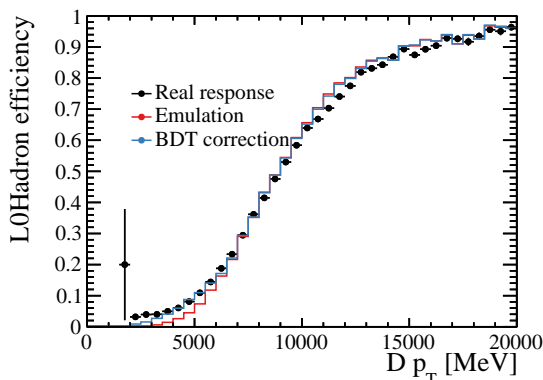
Figure 52: L0 hadron TOS efficiency as a function of the $D^+$ kinematics for the real and emulated response with and without the BDT correction.

conditions, the rest of the event could increase the trigger efficiency, as it constitutes a background energy deposit on most cells. To take this into account in the emulation, a regression BDT has been trained on the difference of the true and emulated trigger $E_T$. The features used are listed in Table 27. The emulated $E_T$ of the Tracker Only MC particles deposit is then corrected accordingly.

| Features |
| --- |
| $D^+ \ p_T$ |
| $D^+ \ p$ |
| $\pi_1^+$ tracker $E_T$ |
| $\pi_2^+$ tracker $E_T$ |
| $K^-$ tracker $E_T$ |
| $\pi_1^+ K^-$ radial distance |
| $\pi_2^+ K^-$ radial distance |
| $\pi_1^+ \pi_2^+$ radial distance |

Table 27: Features included in the regression BDT used for correcting for the underlying event in the emulation of L0Hadron TOS.

This correction has the effect of bringing the emulated trigger efficiency closer to the real efficiency as expected from the Full MC simulation especially at low momentum, where most of the signal tracks are expected to fall.

The comparison between the emulated L0Hadron TOS efficiency and the real efficiency as measured from Full MC simulation, as a function of the $D^+ \ p_T$ is shown in Fig. 52.

The L0Global TIS selection effect is emulated in a fully data-driven way. This is done using a *tag-and-probe* method, using a control sample of $B^+ \to J/\psi K^+$ real data events. The events

are first selected with a L0 selection that requires the presence of one or a combination of two $\mu$ with high $p_T$ in the event (`LOMuon` or `LODiMuon`). The events are required to be TOS on at least one of the two trigger lines. The efficiency of the `LOGlobal` TIS selection is then evaluated on events passing this selection, in bins of $B^{\pm}$ $p_T$ and $p_Z$. The efficiencies are then applied to the Tracker Only MC as efficiency weights.

One of the assumptions of this method is that the muonic TOS efficiency is independent of its TIS efficiency. This can be achieved if a fine enough binning has been chosen. Another assumption is that the TIS efficiency is independent of the decay under study, so that the efficiencies can be directly ported to the MC samples used in this analysis. This has been carefully checked with the help of the Full MC simulation.

Non-trivial correlations between the muonic TOS decision and the global TIS decision may be induced by global event selections and, in particular, by the selection on the SPD multiplicity. For the `LODiMuon` selection the number of hits in the SPD is required to not exceed 900, whereas for the rest of the L0 bandwidth is required not to exceed 450. To ensure this difference does not cause any problem, all the events are required to have a number of hits in the SPD detector never exceeding 450. Given that this variable is not available in the Tracker Only simulation, due to the lack of the simulation of the SPD detector response, this selection is emulated in a data driven way, by looking at the effect this cut has on a sample of $B \to J/\psi K$ events which are required only to fulfill the `LODiMuon` trigger line requirements.

Finally, the emulation weights for the `LOHadron` TOS and `LOGlobal` TIS selections are multiplied to get an overall weight for the L0 selection emulation.

## 10.4   HLT1 emulation

Also, the HLT1 trigger application is not run in Tracker Only simulation, due to the absence of the L0 decision. The HLT1 trigger selection employed in this analysis is a combination of two lines, called `HLT1TrackMVA` and `HLT1TwoTrackMVa`, required to have triggered due to signal tracks or combinations of them. Since the lines are required to be TOS and they act on quantities reconstructed solely by the tracking system, their efficiency can be evaluated by looping over all the signal tracks, evaluating the related quantities, and emulating the response of the selections offline.

As described in 8.3.1, the single track line `HLT1TrackMVA` requires the presence of a single track with a non-linear combination of $p_T$ and $IP_{\chi^2}$ exceeding a given threshold. The two track line `HLT1TwoTrackMVA` requires the presence of a two-tracks combination passing a requirement on a MatrixNet MVA that acts on the $p_T$ and $IP_{\chi^2}$ of the combination. For each event, all the relevant offline information from all the signal tracks and all the two particle combinations are saved and then the selections are applied.

Although the tracks given in input to the HLT1 application are of the long type, as the ones available after the offline reconstruction, the track reconstruction itself is not exactly the same between offline and online reconstruction, due to speed limitations at the HLT1 level. While during the offline reconstruction the Long tracks are reconstructed using the `Forward Tracking` algorithm starting from VELO tracks, in the HLT1 the algorithm is run on VeloTT tracks, which are constructed matching Velo Tracks with hits in the TT, to give a rough estimate (with 20% resolution) of the momentum of the track and discard the very low $p_T$ ones before the full tracking sequence.

The VeloTT tracking algorithm [94] requires the presence of at least three TT hits around

the Velo track direction, and in the algorithm run at HLT1 to promote VeloTT tracks to Long tracks the search window in the OT and IT is defined by the maximum deflection of charged particles with $p_T$ larger than 500 GeV/c. During HLT2 and offline reconstruction, instead, after repeating this steps, the Velo Tracks not matched with any TT segment are used again as input to the Forward tracking to construct Long Tracks without any $p_T$ requirement.

The emulation of the HLT1 selections is run on a sample of tracks reconstructed by the HLT2 and offline reconstruction sequence, which may contain both tracks that have been reconstructed by the online sequence and tracks which could or could not be reconstructed in the online sequence but have been reconstructed by the offline algorithms.

In order to have a sample as uniform as possible to a sample of online reconstructed tracks, selection on the tracks $p_T > 500\,\mathrm{MeV}/c$ and a requirement on the number of hits in the TT associated to the track (nTTHits > 3) has been applied, so that tracks which have been reconstructed by the offline sequence but were not reconstructible by the VeloTT algorithm are removed.

As reported in [94], the VeloTT-Forward reconstruction algorithm has an average per-track reconstruction efficiency which is lower than the Velo-Forward reconstruction chain. This effect is not captured by the emulation described above, but can be easily emulated by randomly prescaling each track with an inefficiency factor which encodes the difference in reconstruction efficiency between the online and offline reconstruction sequences. The inefficiency factor used is around 4.2%.

The HLT1 selections also require some Global Event selections to be fulfilled. These selections, reported in Table 17, are imposed in the emulation as well, being the related quantities available in Tracker Only simulation.

To validate the procedure, the emulation has been run on Full MC events, in which the true HLT1 selection efficiency can be evaluated and compared to the emulated efficiency. Both efficiencies have been evaluated on truth matched events passing the Stripping, HLT2 and L0 requirements reported in the previous chapters. Fig. 53 and 54 report the comparison between the real and emulated efficiency on 2015 $B \to D\mu\nu$ Full MC events, as a function of kinematic variables and as a function of fit variables. The ratio of the emulated and real efficiency is reported below each plot, along with a red error band constructed by varying the inefficiency factor by 1%. In Fig. 55 and 55 the real and emulated efficiencies, their difference and their ratio is reported in 2D planes of fit variables. It can be noticed that an excellent agreement between the real and the emulated efficiency has been reached in both the single and two track trigger lines selection emulations. This has been also carefully checked for all the other MC samples used in the analysis.
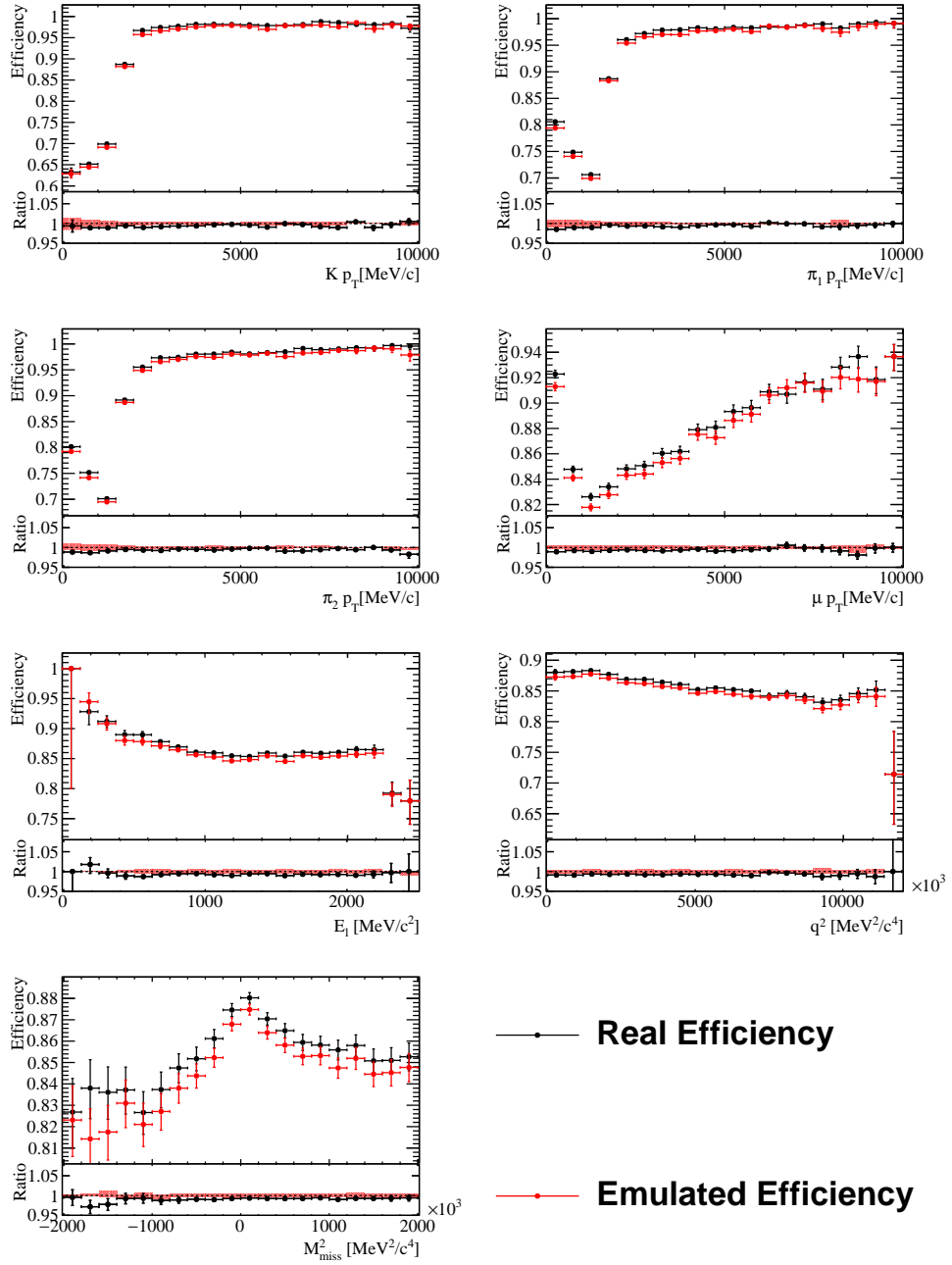
Figure 53: Efficiency of the `HLT1TrackMVA` TOS on the $D$ requirement as a function of the momenta of the signal tracks, after the Efficiency correction on the $B \to D\mu\nu$ 2015 Full MC sample
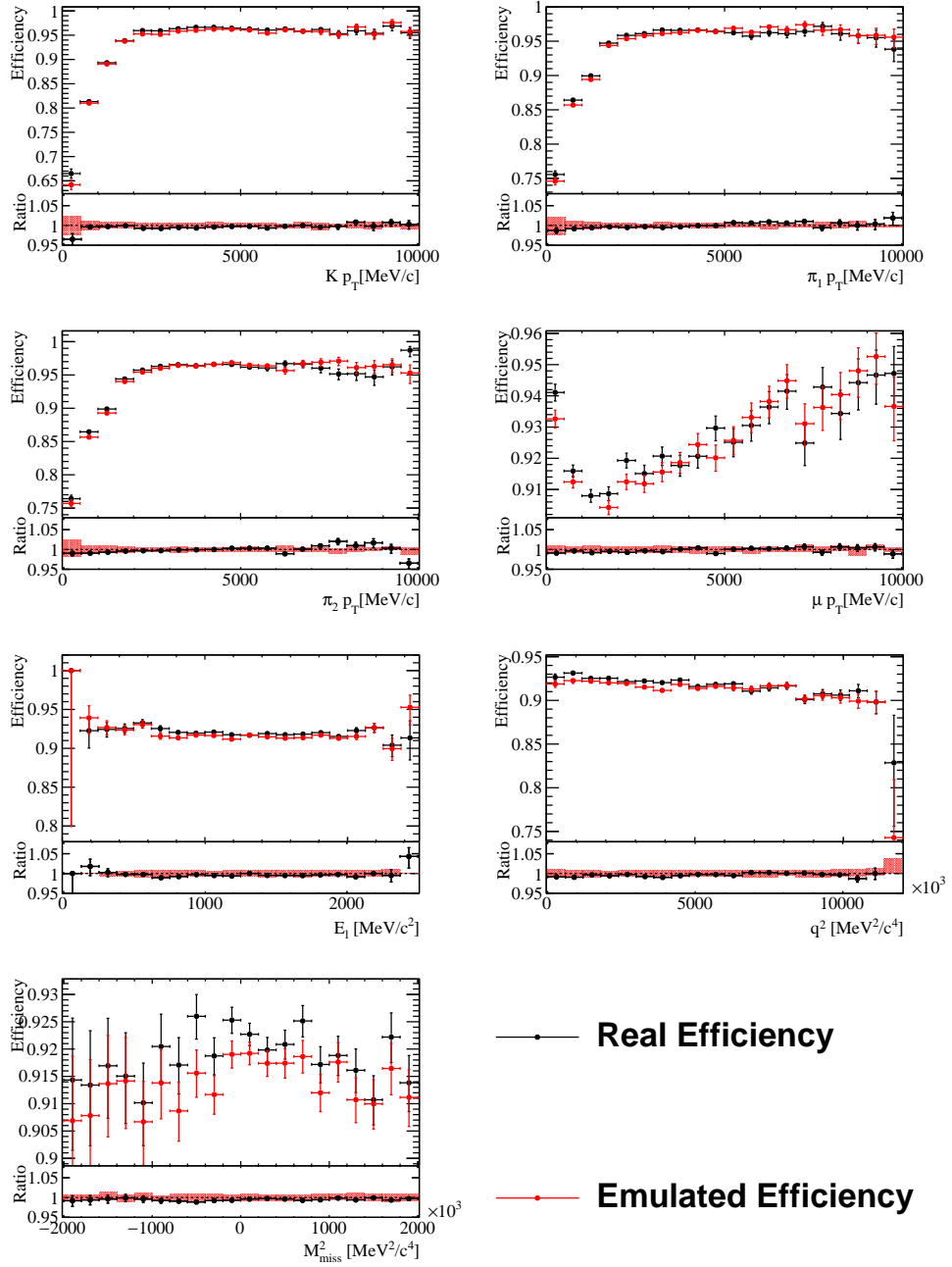
Figure 54: Efficiency of the `HLT1TwoTrackMVA` TOS on the $D$ requirement, after the Efficiency correction on the $B \to D\mu\nu$ 2015 Full MC sample
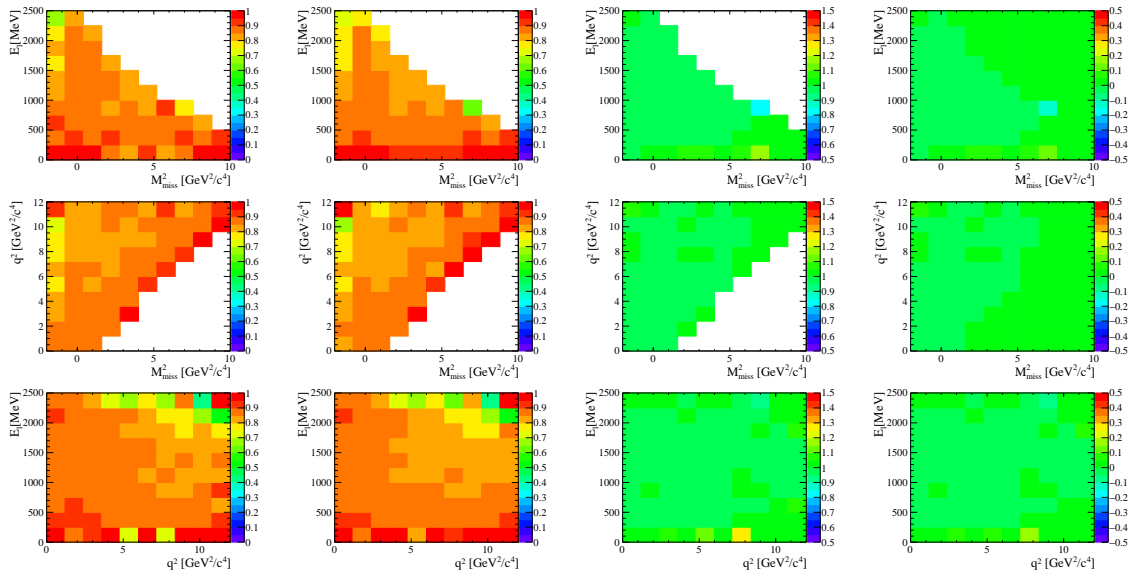
Figure 55: Left: Real Efficiency. Center Left: Emulated Efficiency. Center Right: Ratio of Emulated over Real Efficiency. Right: Difference between Emulated and Real Efficiency
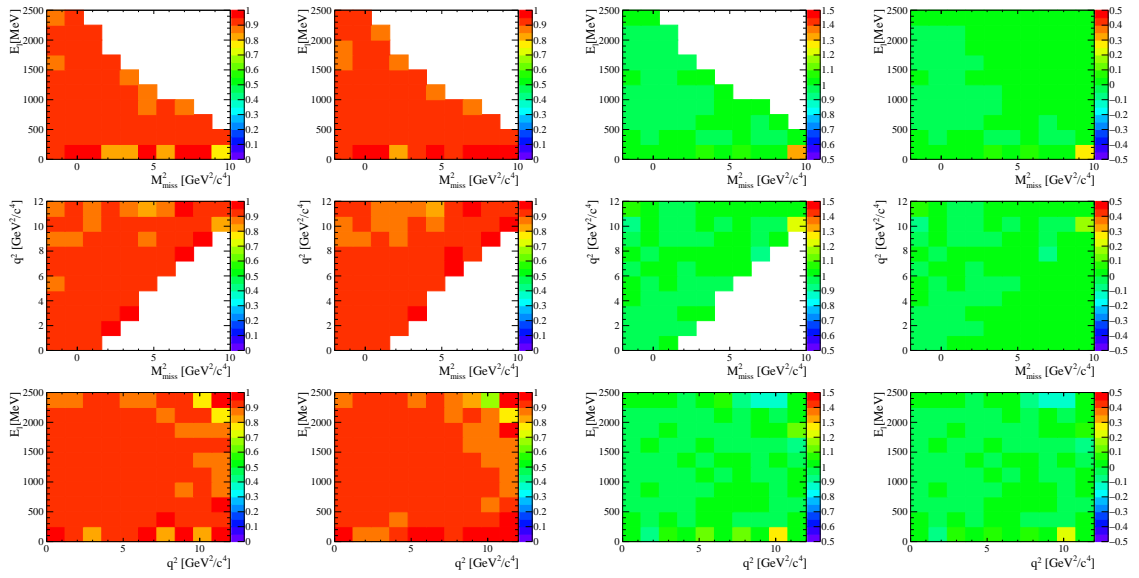


Figure 56: Left: Real Efficiency. Center Left: Emulated Efficiency. Center Right: Ratio of Emulated over Real Efficiency. Right: Difference between Emulated and Real Efficiency

# 11   Efficiencies and Efficiency ratios

The parameters extracted from the fit are ratios of observed yields between the signal and normalization modes, called $\mathcal{R}_{\text{raw}}(D^+)$ and $\mathcal{R}_{\text{raw}}(D^*)$. In order to convert these parameters into the physical result, two efficiency ratios are necessary. These ratios are:

$$\eta_{D^+} = \frac{\varepsilon(\overline{B}^0 \to D^+\tau^-\overline{\nu}_\tau)}{\varepsilon(\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu)}, \tag{152}$$

$$\eta_{D^{*+}} = \frac{\varepsilon(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)}{\varepsilon(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)}, \tag{153}$$

where $\varepsilon(X) = \frac{N^{\text{RECO}}(X)}{N^{\text{GEN}}(X)}$ denotes the efficiency of selecting events for mode $X$ in the signal region.

With these definitions the physical parameters of interest can be extracted using the observed number of events in this way:

$$\mathcal{R}(D^+) = \frac{1}{\eta_{D^+}} \frac{\mathcal{R}_{\text{raw}}(D^+)}{\mathcal{B}(\tau^- \to \mu^-\nu_\tau\overline{\nu}_\mu)} \tag{154}$$

$$\mathcal{R}(D^{*+}) = \frac{1}{\eta_{D^{*+}}} \frac{\mathcal{R}_{\text{raw}}(D^{*+})}{\mathcal{B}(\tau^- \to \mu^-\nu_\tau\overline{\nu}_\mu)} \tag{155}$$

The efficiencies for selecting the events in all the control and validation regions used in the analysis are also evaluated. The efficiencies evaluated using the Tracker Only simulation, after all the corrections to the MC simulation, that will be described later on throughout the text, have been applied.

The efficiency for a given sample is given by the number of reconstructed, selected and truth matched events, corrected with data/MC correction weights ($N_{RECO} = \sum_{i=RECO} w_i$), divided by the number of generated events, before any selection or generator level cut.

In order to obtain the number of generated events, the number of events passing the filtering selections are divided by the generator level efficiency and the filtering efficiency:

$$N_{GEN} = \frac{N_{FIL}^{\text{ACC}}}{\varepsilon_{GEN} \times \varepsilon_{FIL}} \tag{156}$$

The generator level efficiencies ($\varepsilon_{GEN}$) represent the ratio of events that pass the generator level selections over the number of total generated events. These values are provided by the simulation separately for the Magnet Up and Magnet Down conditions and are averaged between the two polarities. The values of the the generator level efficiencies are reported in Table 28.

The filtering efficiencies ($\varepsilon_{FIL}$) represent the number of events that have been reconstructed and pass the filtering selections. The values of the filtering efficiencies, $\varepsilon_{FIL}$, are reported in Table 29.

The total efficiencies and their ratios, for the signal and normalization events, and separately for 2015, 2016 and the combined sample are reported in Table 30. For comparison, the same quantities computed with Full MC instead of Tracker Only are reported in Table 31.

It can be seen that, while a small discrepancy between the absolute efficiencies evaluated using the Tracker Only MC samples and the Full MC samples can be observed the ratio of

efficiencies is compatible to a very high degree between the two simulation versions. A small discrepancy on the absolute efficiencies can anyhow be expected, since part of the emulations done on the Tracker Only MC samples are data driven, making the efficiencies estimations more in line with the expected ones in real data than the ones estimated with the Full MC.

The overall agreement of the efficiency ratios between the two simulation version therefore gives very high confidence to the chain of emulations used in the Tracker Only MC.

| Sample | DecFile Number | Full MC | TO MC (2015) | TO MC (2016) |
|---|---|---|---|---|
| $\overline{B}^0 \to D^+ \mu\overline{\nu}_\mu$ | 11574061 | $0.098 \pm 0.00025\%$ <br> $0.09801 \pm 0.000023\%$ | $0.09813 \pm 0.00016\%$ <br> $0.09787 \pm 0.00017\%$ | $0.09801 \pm 0.00013\%$ <br> $0.09791 \pm 0.00013\%$ |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\overline{\nu}_\mu$ | 11574402 | $0.08798 \pm 0.000233\%$ <br> $0.08772 \pm 0.00022\%$ | $0.08763 \pm 0.00015\%$ <br> $0.0878 \pm 0.00021\%$ | $0.08756 \pm 0.00012\%$ <br> $0.08789 \pm 0.00012\%$ |
| $\overline{B}^0 \to D^+\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 11574060 | $0.08023 \pm 0.00021\%$ <br> $0.08017 \pm 0.00021\%$ | $0.08006 \pm 0.00019\%$ <br> $0.08015 \pm 0.00021\%$ | $0.0798 \pm 0.00015\%$ <br> $0.0801 \pm 0.00015\%$ |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 11574401 | $0.07508 \pm 0.0002\%$ <br> $0.07482 \pm 0.00018\%$ | $0.07486 \pm 0.00011\%$ <br> $0.07463 \pm 0.00019\%$ | $0.07462 \pm 0.00014\%$ <br> $0.07456 \pm 0.00014\%$ |
| $B^- \to D^{**0}(\to D^+X)\mu\overline{\nu}_\mu$ | 12874050 | $0.08448 \pm 0.00022\%$ <br> $0.08451 \pm 0.00021\%$ | $0.08435 \pm 0.00017\%$ <br> $0.08444 \pm 0.00011\%$ | $0.08468 \pm 0.00016\%$ <br> $0.08456 \pm 0.00016\%$ |
| $\overline{B}^0 \to D^{**+}(\to D^+X)\mu\overline{\nu}_\mu$ | 11574403 | $0.08468 \pm 0.00022\%$ <br> $0.08425 \pm 0.00022\%$ | $0.08427 \pm 0.00012\%$ <br> $0.08444 \pm 0.00011\%$ | $0.08424 \pm 0.00016\%$ <br> $0.08429 \pm 0.00016\%$ |
| $\overline{B}^0 \to D^{**+}(\to D^+X)\mu\overline{\nu}_\mu$, high mass | 11574070 | $0.08093 \pm 0.00022\%$ <br> $0.08089 \pm 0.00022\%$ | $0.08067 \pm 0.00009\%$ <br> $0.08092 \pm 0.00012\%$ | $0.081 \pm 0.00015\%$ <br> $0.08055 \pm 0.00015\%$ |
| $\overline{B}^0 \to D^+H_c(\to \mu\overline{\nu}_\mu X')X$ | 11995204 | $0.03054 \pm 0.00008\%$ <br> $0.03058 \pm 0.00008\%$ | $0.03042 \pm 0.00005\%$ <br> $0.03048 \pm 0.00006\%$ | $0.03047 \pm 0.00004\%$ <br> $0.03054 \pm 0.00004\%$ |
| $B^- \to D^+H_c(\to \mu\overline{\nu}_\mu X')X$ | 12995604 | $0.10261 \pm 0.00024\%$ <br> $0.10211 \pm 0.00024\%$ | $0.10201 \pm 0.00014\%$ <br> $0.1023 \pm 0.00015\%$ | $0.10224 \pm 0.00014\%$ <br> $0.10237 \pm 0.00014\%$ |
| $\Lambda_b \to D^+\Lambda_c(\to \mu\overline{\nu}_\mu X')X$ | 15976000 | - <br> - | $0.05277 \pm 0.00013\%$ <br> $0.05297 \pm 0.00013\%$ | $0.05313 \pm 0.00009\%$ <br> $0.05307 \pm 0.00009\%$ |
| $\overline{B}^0 \to D^+H_c(\to \tau\overline{\nu}_\tau)X$ | 11995214 | - <br> - | $0.08127 \pm 0.00008\%$ <br> $0.08144 \pm 0.00009\%$ | $0.08123 \pm 0.00011\%$ <br> $0.08132 \pm 0.00011\%$ |
| $B^- \to D^+H_c(\to \tau\nu_\tau)X$ | 12995615 | - <br> - | $0.07553 \pm 0.00008\%$ <br> $0.07527 \pm 0.00008\%$ | $0.07565 \pm 0.00011\%$ <br> $0.07572 \pm 0.00011\%$ |
| $\overline{B}^0 \to D^{**+}(\to D^+X)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 11874002 | - <br> - | $0.0724 \pm 0.00013\%$ <br> $0.07236 \pm 0.00013\%$ | $0.07263 \pm 0.00008\%$ <br> $0.0726 \pm 0.00008\%$ |
| $B^- \to D^{**0}(\to D^+X)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 12874001 | - <br> - | $0.07297 \pm 0.00013\%$ <br> $0.0726 \pm 0.00013\%$ | $0.07297 \pm 0.00008\%$ <br> $0.07295 \pm 0.00008\%$ |
| $\overline{B}_s^0 \to D_s^{**+}(\to D^+X)\mu\overline{\nu}_\mu$ | 13874003 | - <br> - | $0.08215 \pm 0.00013\%$ <br> $0.08212 \pm 0.00013\%$ | $0.08185 \pm 0.0001\%$ <br> $0.0821 \pm 0.0001\%$ |

Table 28: Generator Level efficiencies. For each year and MC type, the first (second) line represents the value of the efficiency for the Magnet Up (Down) condition.

| Sample | DecFile Number | Full MC | TO MC (2015) | TO MC (2016) |
|---|---|---|---|---|
| $\overline{B}^0 \to D^+ \mu\overline{\nu}_\mu$ | 11574061 | 12.5496 ± 0.0233% | 12.6479 ± 0.00072% | 13.2782 ± 0.0032% |
| | | 12.6131 ± 0.0231% | 12.662 ± 0.0074% | 13.2989 ± 0.0032% |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\overline{\nu}_\mu$ | 11574402 | 12.9335 ± 0.0235% | 12.9114 ± 0.00076% | 13.5574 ± 0.0033% |
| | | 12.9449 ± 0.024% | 12.948 ± 0.0074% | 13.5851 ± 0.0033% |
| $\overline{B}^0 \to D^+\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 11574060 | 10.403 ± 0.0422% | 10.4696 ± 0.01399% | 10.9601 ± 0.0059% |
| | | 10.3749 ± 0.043% | 10.417 ± 0.01399% | 10.9608 ± 0.0059% |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 11574401 | 10.3585 ± 0.0428% | 10.4795 ± 0.0137% | 10.9851 ± 0.0059% |
| | | 10.3937 ± 0.0434% | 10.5021 ± 0.014% | 11.0032 ± 0.0059% |
| $B^- \to D^{**0}(\to D^+X)\mu\overline{\nu}_\mu$ | 12874050 | 13.0201 ± 0.0459% | 12.6705 ± 0.0163% | 13.3366 ± 0.0071% |
| | | 13.0785 ± 0.0451% | 12.6679 ± 0.0167% | 13.3642 ± 0.0071% |
| $\overline{B}^0 \to D^{**+}(\to D^+X)\mu\overline{\nu}_\mu$ | 11574403 | 12.6797 ± 0.0448% | 12.5021 ± 0.0163% | 13.1611 ± 0.0071% |
| | | 12.7825 ± 0.0453% | 12.5335 ± 0.0162% | 13.1853 ± 0.0071% |
| $\overline{B}^0 \to D^{**+}(\to D^+X)\mu\overline{\nu}_\mu$, high mass | 11574070 | 13.0986 ± 0.067% | 12.4144 ± 0.0151% | 13.1044 ± 0.0069% |
| | | 13.2694 ± 0.0673% | 12.4404 ± 0.02332% | 13.1191 ± 0.0069% |
| $\overline{B}^0 \to D^+H_c(\to \mu\overline{\nu}_\mu X')X$ | 11995204 | 9.4398 ± 0.02218% | 8.7517 ± 0.0064% | 9.2113 ± 0.0029% |
| | | 9.4186 ± 0.02215% | 8.7611 ± 0.0064% | 9.2269 ± 0.0029% |
| $B^- \to D^+H_c(\to \mu\overline{\nu}_\mu X')X$ | 12995604 | 1.6054 ± 0.00566% | 1.3053 ± 0.0015% | 1.3749 ± 0.0007% |
| | | 1.6013 ± 0.0057% | 1.3071 ± 0.0015% | 1.3764 ± 0.0007% |
| $\overline{B}^0 \to D^+H_c(\to \tau\overline{\nu}_\tau X')X$ | 11995214 | - | 7.1042 ± 0.019% | 7.384 ± 0.0088% |
| | | - | 7.1023 ± 0.0213% | 7.391 ± 0.0091% |
| $B^- \to D^+H_c(\to \tau\overline{\nu}_\tau X')X$ | 12995615 | - | 7.5628 ± 0.0227% | 7.8274 ± 0.0091% |
| | | - | 7.5514 ± 0.0227% | 7.8427 ± 0.0093% |
| $\Lambda_b \to (\Lambda_c \to \mu\overline{\nu}_\mu X')D^+X$ | 15976000 | - | 9.8268 ± 0.0405% | 10.1742 ± 0.0119% |
| | | - | 9.8158 ± 0.0402% | 10.1769 ± 0.0121% |
| $\overline{B}^0 \to D^{**+}(\to D^+X)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 11874002 | - | 10.57 ± 0.0392% | 10.9233 ± 0.0188% |
| | | - | 10.5192 ± 0.0396% | 10.9346 ± 0.0189% |
| $B^- \to D^{**0}(\to D^+X)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 12874001 | - | 10.7525 ± 0.0401% | 11.0636 ± 0.0191% |
| | | - | 10.7528 ± 0.0394% | 11.1019 ± 0.0188% |
| $\overline{B}^0_s \to D^{**+}_s(\to D^+X)\mu\overline{\nu}_\mu$ | 13874003 | - | 12.9949 ± 0.0397% | 13.3918 ± 0.0175% |
| | | - | 13.0335 ± 0.0386% | 13.4313 ± 0.0178% |

Table 29: Filtering Efficiencies. For each year and MC type, the first (second) line represents the value of the efficiency for the Magnet Up (Down) condition.

116

| Sample | 2015 | | 2016 | | comb. | |
|---|---|---|---|---|---|---|
| | $\varepsilon$ | ratio | $\varepsilon$ | ratio | $\varepsilon$ | ratio |
| $\overline{B}^0 \to D^+\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | $0.8513 \pm 0.0027‰$ | $0.5631 \pm 0.0019$ | $0.9061 \pm 0.0011‰$ | $0.5730 \pm 0.0008$ | $0.8970 \pm 0.0010‰$ | $0.5714 \pm 0.0007$ |
| $\overline{B}^0 \to D^+\mu\overline{\nu}_\mu$ | $1.5116 \pm 0.0019‰$ | | $1.5814 \pm 0.0008‰$ | | $1.5699 \pm 0.0008‰$ | |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | $0.5891 \pm 0.0021‰$ | $0.5851 \pm 0.0023$ | $0.6198 \pm 0.0009‰$ | $0.5823 \pm 0.0009$ | $0.6147 \pm 0.0008‰$ | $0.5828 \pm 0.0009$ |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\overline{\nu}_\mu$ | $1.0068 \pm 0.0015‰$ | | $1.0644 \pm 0.0007‰$ | | $1.0549 \pm 0.0006‰$ | |

Table 30: Total efficiencies and efficiency ratios ($\eta$) for the signal and normalization samples, Tracker Only MC samples.

| Sample | 2015 | | 2016 | | comb. | |
|---|---|---|---|---|---|---|
| | $\varepsilon$ | ratio | $\varepsilon$ | ratio | $\varepsilon$ | ratio |
| $\overline{B}^0 \to D^+\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | $0.761 \pm 0.008‰$ | $0.556 \pm 0.006$ | - | - | - | - |
| $\overline{B}^0 \to D^+\mu\overline{\nu}_\mu$ | $1.367 \pm 0.006‰$ | | - | | - | |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | $0.522 \pm 0.006‰$ | $0.583 \pm 0.008$ | - | - | - | - |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\overline{\nu}_\mu$ | $0.895 \pm 0.004‰$ | | - | | - | |

Table 31: Total efficiencies and efficiency ratios ($\eta$) for the signal and normalization samples, Full MC samples.

# 12 Backgrounds estimated from data

The non-physical backgrounds relevant for this analysis are three:

- $D$-combinatorial: random three-tracks combinations which are reconstructed as a $D^+$ candidate;

- $D\mu$-combinatorial: also referred to as *combinatorial*, made of real $D$ and $\mu$ candidates coming from uncorrelated decays in the event;

- $\mu$-MisID: events in which the $\mu$ candidate is actually a hadron, electron or ghost track being misidentified as a muon.

The physical backgrounds are either estimated in a fully data-driven way or are subtracted in a statistical way from the data samples. The procedure to do so is described in this chapter. The procedure is repeated independently for the two years of data taking and independently in each control region used in the fit.

## 12.1 $\mu$-misidentification

The $\mu$-MisID background is made up of events in which hadrons, electrons or ghosts tracks have passed the $\mu$ PID identification criteria. This background is estimated in a data driven way.

In order to estimate the shape of the $\mu$-MisID background, one needs to know the amount of true particles ($h$) being reconstructed as a muon $\hat{\mu}$. This depends on the MisID efficiency, $P(\hat{\mu}|h)$ and the true number of particles in the sample $N_h$. The number of $\mu$-MisID events is then given by

$$N(\hat{\mu}|h) = \sum_h P(\hat{\mu}|h)N_h. \tag{157}$$

The $h \to \hat{\mu}$ mis-identification rate, $P(\hat{\mu}|h)$ (that is, the probability of having selected a true hadron $h$ as a muon, $\hat{\mu}$), for the PID selection used in the analysis, are taken from calibration samples using a *tag-and-probe* technique. These efficiencies are evaluated for each data taking year (2015 and 2016) and for each *true* particle specie ($h \in \{e, p, \pi, K, g\}$). The calibration samples are taken from real data, except for the one used to evaluate the $g \to \hat{\mu}$ which is taken from Full MC simulation.

The number of true hadrons in the event, $N_h$, although is not known a priori. This is evaluated using an unfolding technique that makes use of $D^+ h^\pm$ candidates selected in the *misID sample*, as described in Sect. 8.1. This sample is first divided into 5 reconstructed particle specie categories, $\hat{h} \in \{\hat{e}, \hat{p}, \hat{\pi}, \hat{K}, \hat{g}\}$, using mutually exclusive PID selections. These selections are reported in Table 32.

In this sample, the number of hadrons reconstructed in each specie, $N_{\hat{h}}$ is known, and is linked to the number of true particles in each specie by the following relation

$$N_{\hat{h}} = \sum_h P(\hat{h}|h)N_h. \tag{158}$$

| Species | Selection |
|---------|-----------|
| $\hat{\pi}$ | $(ProbNN_{\pi} > 0.75 ~\&\&~ DLL_K < 0.0 ~\&\&~ DLL_p < 0.0)$ |
|  | $\&\&~ ProbNN_K * (1 - ProbNN_p) < 0.1$ |
|  | $\&\&~ ProbNN_p * (1 - ProbNN_K) < 0.1$ |
|  | $\&\&~ ProbNN_e < 0.8$ |
|  | $\&\&~ ProbNNghost < 0.2$ |
| $\hat{K}$ | $(ProbNN_{\pi} < 0.75 ~||~ DLL_K > 0.0 ~||~ DLL_p > 0.0)$ |
|  | $\&\&~ ProbNN_K * (1 - ProbNN_p) > 0.1$ |
|  | $\&\&~ ProbNN_p * (1 - ProbNN_K) < 0.1$ |
|  | $\&\&~ ProbNN_e < 0.8$ |
|  | $\&\&~ ProbNN_g < 0.2$ |
| $\hat{p}$ | $(ProbNN_{\pi} < 0.75 ~||~ DLL_K > 0.0 ~||~ DLL_p > 0.0)$ |
|  | $\&\&~ ProbNN_K * (1 - ProbNN_p) < 0.1$ |
|  | $\&\&~ ProbNN_p * (1 - ProbNN_K) > 0.1$ |
|  | $\&\&~ ProbNN_e < 0.8$ |
|  | $\&\&~ ProbNN_g < 0.2$ |
| $\hat{e}$ | $(ProbNN_{\pi} < 0.75 ~||~ DLL_K > 0.0 ~||~ DLL_p > 0.0)$ |
|  | $\&\&~ ProbNN_K * (1 - ProbNN_p) < 0.1$ |
|  | $\&\&~ ProbNN_p * (1 - ProbNN_K) < 0.1$ |
|  | $\&\&~ ProbNN_e > 0.8$ |
|  | $\&\&~ ProbNN_g < 0.2$ |
| $\hat{g}$ | not passing any of the above selections |

Table 32: Definition of reconstructed PID categories.

By inverting this relation, it would be possible to evaluate the number of true particles in each hadron specie, $N_h$. Once the number of true particles $N_h$ is known, the number of them being reconstructed as a $\hat{\mu}$ can be evaluated in each bin with the following formula:

$$N(\hat{\mu}) = \sum_{i \in \{\pi, K, p, e, g\}} N_i P(\hat{\mu}|i), \tag{159}$$

where $P(\hat{\mu}|i)$ is the probability of a true particle $i$ passing the reconstructed $\mu$ selections, measured from calibration samples. In order to measure $N_i$ from the inversion of Eq. 157, a

matrix inversion would be required. This procedure is known to be an ill-posed problem in statistics and could lead to high instabilities in the procedure. For this reason, another approach is followed.

The number $N_i$ is effectively measured by assigning a per event weight to the events, defined in the following way. The measured yield in each of the five reconstructed PID categories, $N_{\hat{h}}$ is fitted with templates constructed using cross-feed probabilities $P(\hat{h}|h)$ measured from calibration samples. As before, these calibration samples used for ghost tracks are taken from Full MC samples, whereas the rest is taken from real data. The templates are constructed for each data taking condition, and in different kinematical bins defined in the 3D space formed by the track $p$, track $\eta$ and $nTracks$ variables. The binning used in this step is described in Table 33.

| Variable | Edges |
|----------|-------|
| $|\vec{p}|$ | [3.,6., 9., 15., 30., 50., 100.] GeV |
| $\eta$ | [1.5, 2.5, 3.5, 5.] |
| nTracks | [0, 225, 500] |

Table 33: Bin edges used for PID efficiencies to model cross-feeds for misID template.

As a reference, an example of the templates constructed in a given kinematic bin is reported in Fig. 57.
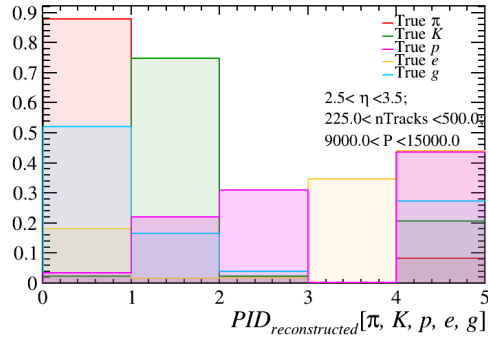


Figure 57: Distributions of true particle species in reconstructed PID categories.

By weighting each event with the s$\mathcal{W}$eight (s$\mathcal{W}_i$) relative to a given true particle specie $i \in \{\pi, K, p, e, g\}$ extracted from the previous fit, only true particles being reconstructed in either one of the reconstructed particles categories $\hat{j} \in \{\hat{\pi}, \hat{K}, \hat{p}, \hat{e}, \hat{g}\}$ are statistically selected. The weight is then divided by the probability of a true particle $i$, belonging to the considered kinematic bin, to be reconstructed in either one of the reconstructed particles categories, $\sum_{\hat{j} \in \{\hat{\pi}, \hat{K}, \hat{p}, \hat{e}, \hat{g}\}} P(\hat{j}|i)$ Finally the weight is multiplied by the inverse of the prescale of the MisID stripping line, to align the integrated luminosity of the MisID sample with the one of the data sample. By summing this per event weight over all the events would result in an estimation of the particles in the true

category $i$. This is inserted in Eq. 159 to obtain the per event weight assigned to all the events in the MisID sample:

$$w_{\mu-\text{MisID}} = \sum_{i\in\{\pi,K,p,e,g\}} \frac{s\mathcal{W}_i}{\sum_{\hat{j}\in\{\hat{\pi},\hat{K},\hat{p},\hat{e},\hat{g}\}} P(\hat{j}|i)} \cdot \frac{1}{0.10} \cdot P(\hat{\mu}|i). \tag{160}$$

The procedure is repeated separately for each data taking year (2015 and 2016) and separately for the RS MisID and WS MisID samples. The results of the fits in the MisID RS sample and MisID WS samples, for both 2015 and 2016 data taking conditions, are reported in Fig. 58,59,60 and 61. In both samples the $D$-combinatorial background is subtracted, with the procedure explained in Sect. 12.2. The events of the MisID RS sample are then used to estimate the shape (and expected yield) of the $\mu$-MisID background in each region used in the fit.

The events in the WS misID sample are instead subtracted from the WS sample used to estimate the $D\mu$-combinatorial background, as it will be explained in Sect. 12.3
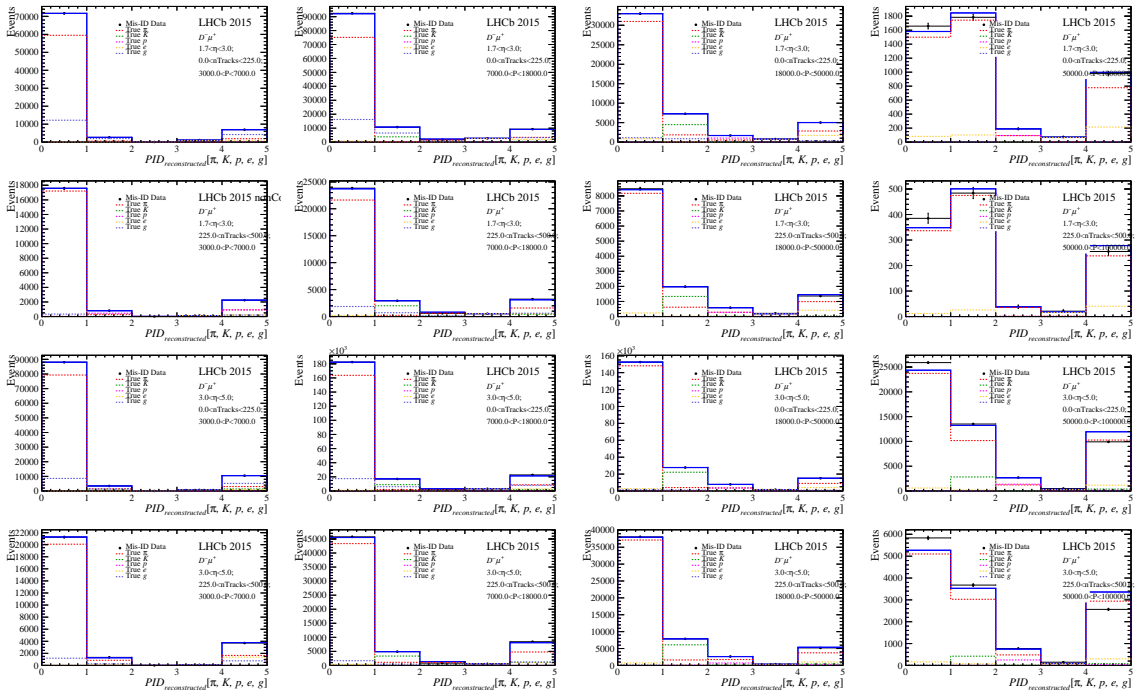
Figure 58: Fit to the reconstructed hadron categories histograms, MisID RS, 2015 sample.
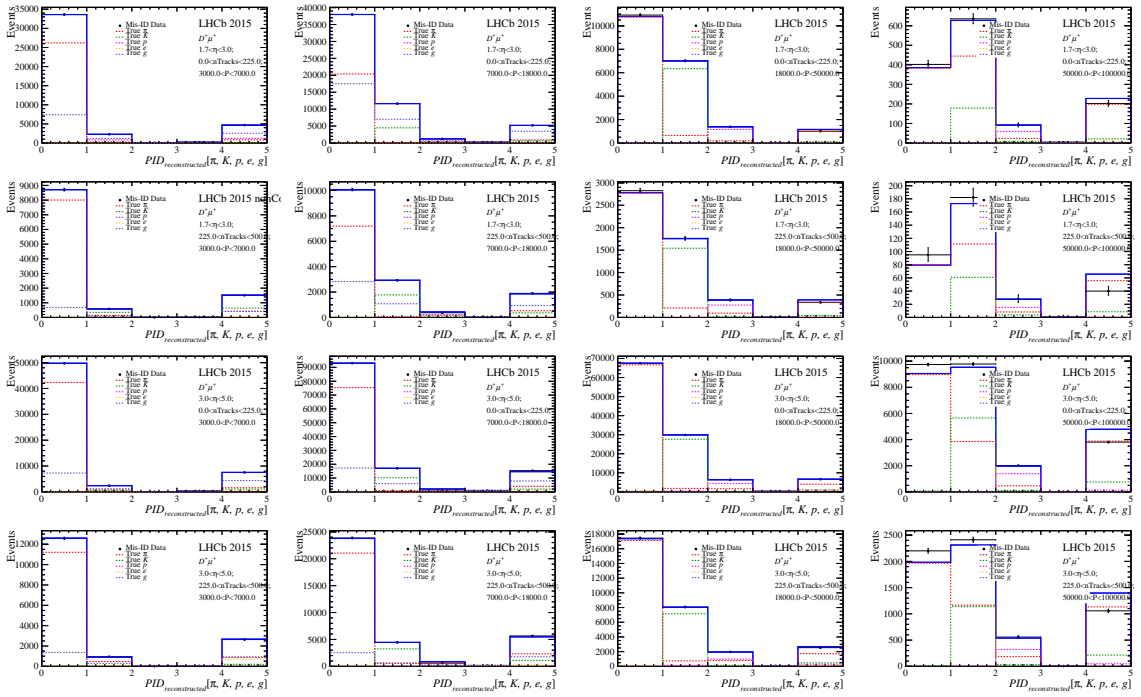


Figure 59: Fit to the reconstructed hadron categories histograms, MisID WS, 2015 sample.
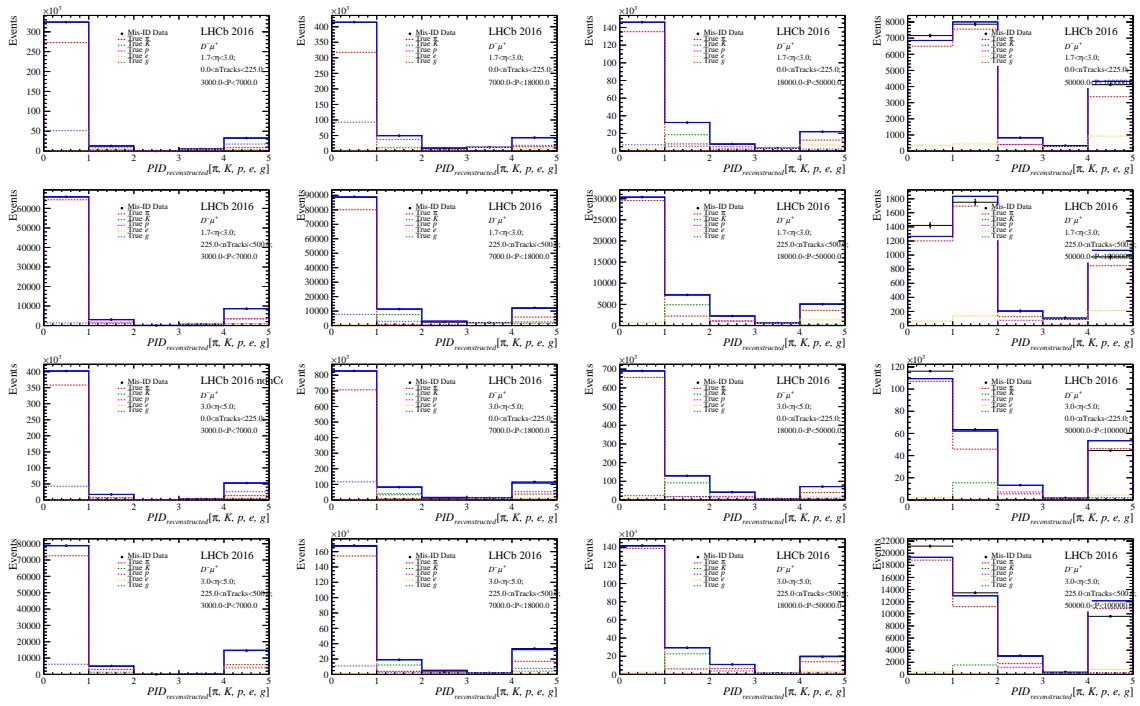
Figure 60: Fit to the reconstructed hadron categories histograms, MisID RS, 2016 sample.
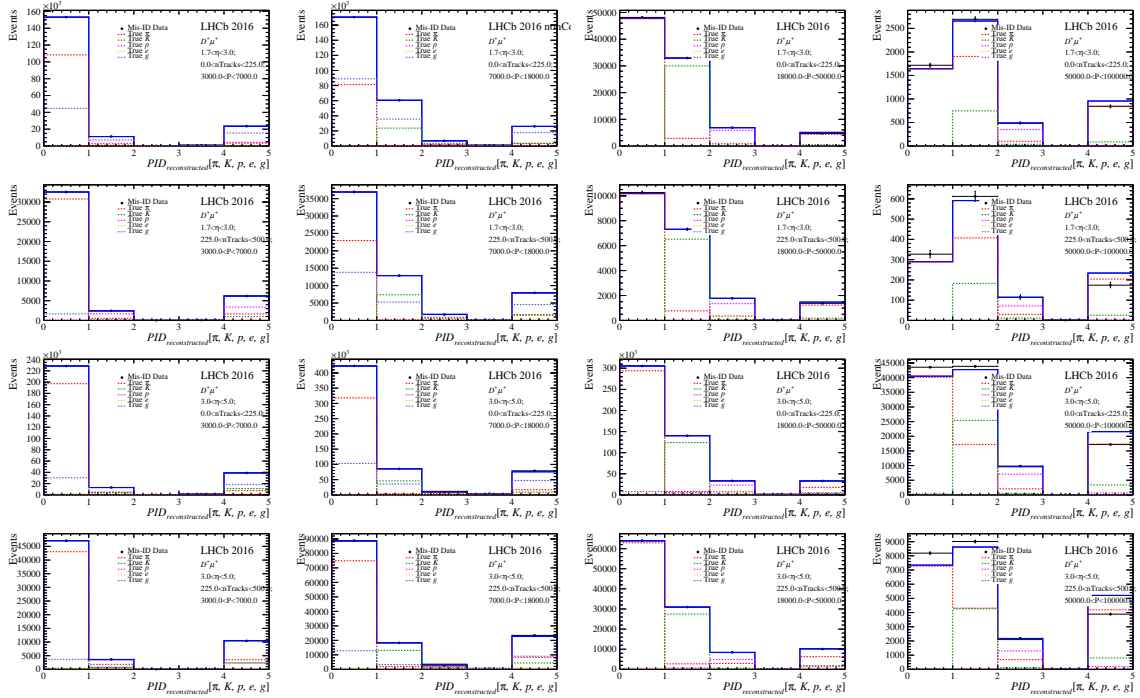


Figure 61: Fit to the reconstructed hadron categories histograms, MisID WS, 2016 sample.
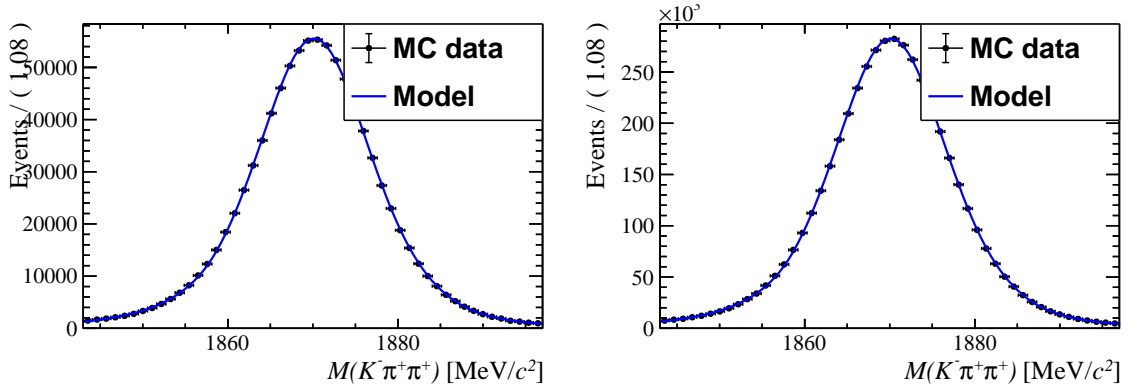
Figure 62: Result of the fits to the $M(K\pi\pi)$ distribution of (left) 2015 and (right) 2016 $\overline{B}^0 \to D^+\mu^-\overline{\nu}_\mu$ tracker-only Monte Carlo events, with all the previous selection requirements applied.

## 12.2  $D$-combinatorial

This background consists of combinations of random three tracks that have passed all the reconstruction and selection requirements and have been reconstructed as $D^+$ candidates. This background can be distinguished from the rest by their distribution on the invariant mass of the three tracks combination, $M(K^-\pi^+\pi^+)$, in which the events containing a true $D^+ \to K\pi\pi$ candidate show a peak around the mass of the $D^+$ meson and the $D$-combinatorial is instead flat.

This background is statistically subtracted by means of the sPlot technique [93], obtained by fitting the $M(K^-\pi^+\pi^+)$ distribution. The background has been modelled with an exponential function whose slope is left freely floating. The $D^+$ peak is modelled using an Ipatia function [95], defined as:

$$
I(m, \mu, \sigma, \lambda, \zeta, \beta, a_1, a_2, n_1, n_2) \propto
$$
$$
\begin{cases}
\frac{A}{(B+m-\mu)^{n_1}}, & \text{if } m - \mu < -a_1\sigma \\
\frac{C}{(D+m-\mu)^{n_2}}, & \text{if } m - \mu > a_2\sigma \\
((m-\mu)^2 + \delta^2)^{\frac{1}{2}\lambda - \frac{1}{4}} e^{\beta(m-\mu)} K_{\lambda - \frac{1}{2}}(\alpha\sqrt{(m-\mu)^2 + \delta^2}), & \text{otherwise,}
\end{cases}
\tag{161}
$$
,

where $K_\nu(z)$ is the modified Bessel function of the second kind, $\delta \equiv \sigma\sqrt{\frac{\zeta K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)}}$, $\alpha \equiv \frac{1}{\sigma}\sqrt{\frac{\zeta K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}}$, and $A$, $B$, $C$ and $D$ are obtained by imposing continuity and differentiability. The parameters $\zeta$ and $\beta$ are fixed to zero, the central mass and the width parameters are floated from the fit, whereas all the other parameters are fixed to the estimates found in fits to the 2015 and 2016 $B \to D\mu\nu$ tracker only MC events which pass all the selections described in the previous sections. The result of the fits to the MC samples is reported in Fig. 62.

The fit to the $M(K^-\pi^+\pi^+)$ distribution is performed separately for each signal, control and validation region, and independently for the 2015 and 2016 data taking conditions. The fit is performed on both RS and WS samples. For the MisID sample, the fit is performed on events which have been firstly weighted according to the MisID weights, described in the previous section. The result of the fits is reported in Fig. 63 and 64.
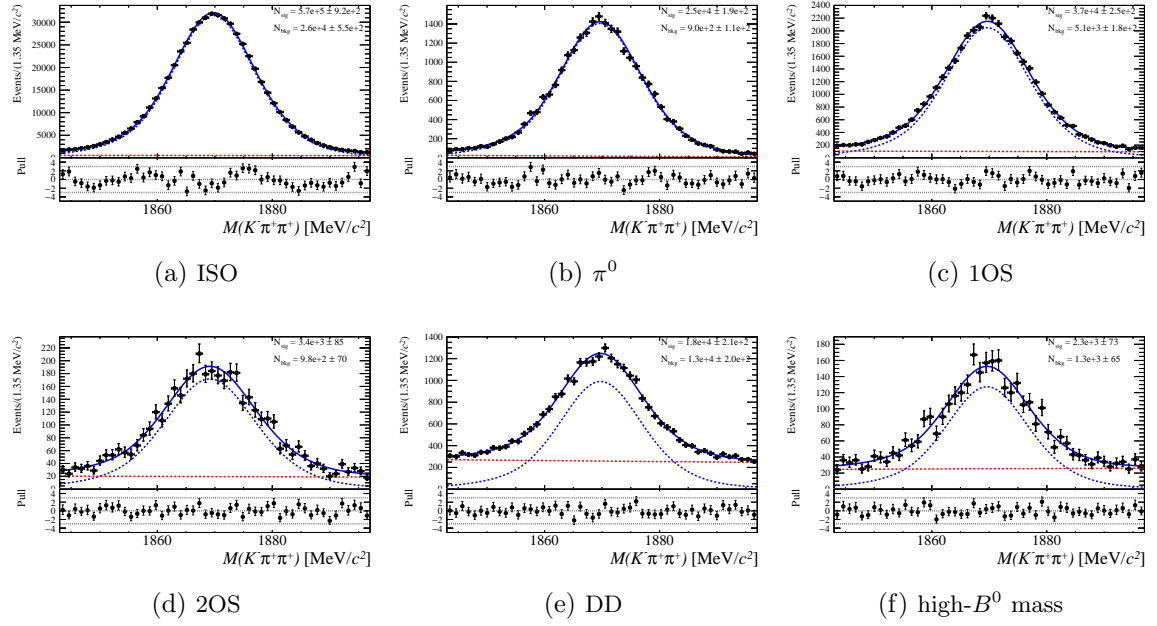
Figure 63: Fit to the reconstructed $D^+$ mass used to subtract the $D$-combinatorial background, 2015 data sample.
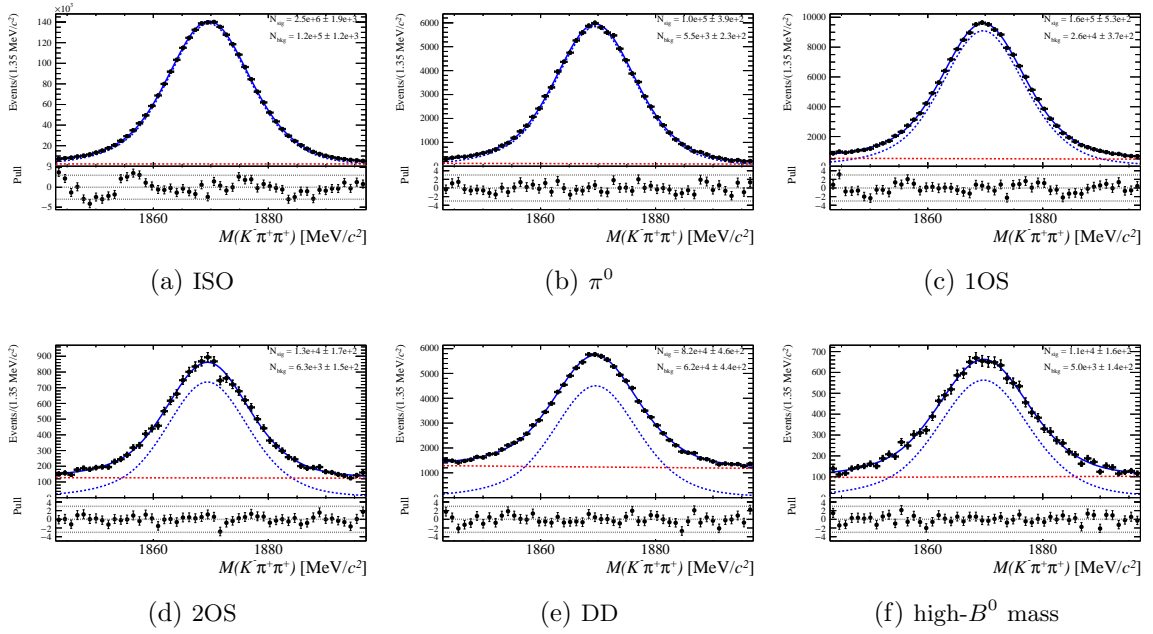


Figure 64: Fit to the reconstructed $D^+$ mass used to subtract the $D$-combinatorial background, 2016 data sample.

(a) ISO      (b) $\pi^0$      (c) 1OS

(d) 2OS      (e) DD      (f) high-$B^0$ mass
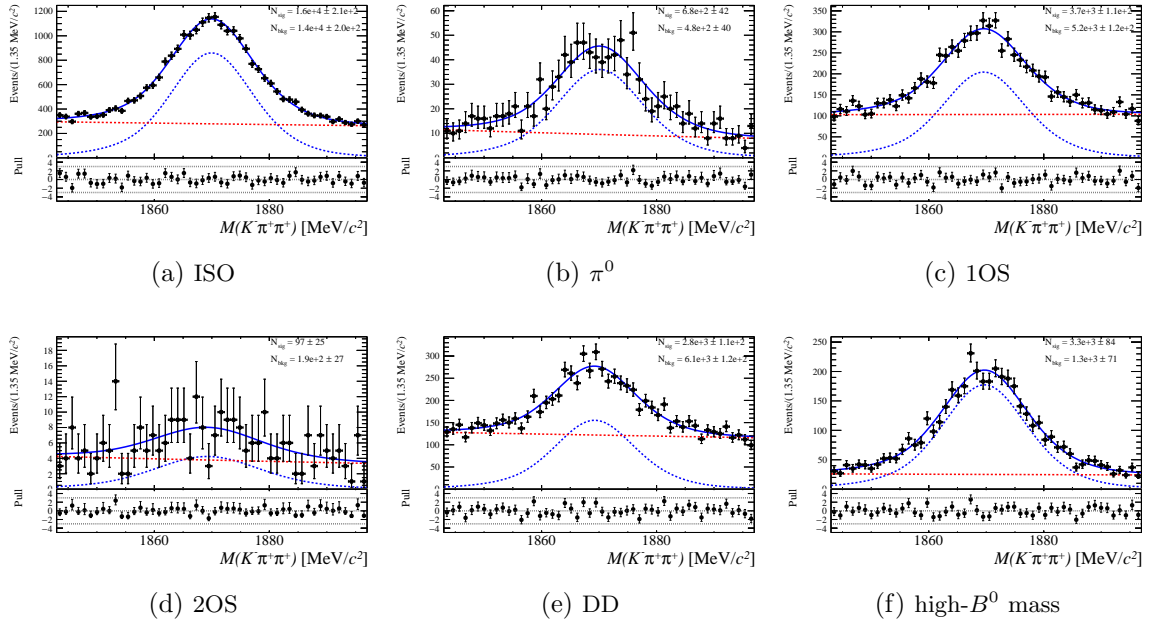
Figure 65: Fit to the reconstructed $D^+$ mass used to subtract the $D$-combinatorial background, 2015 WS data sample.



(a) ISO      (b) $\pi^0$      (c) 1OS

(d) 2OS      (e) DD      (f) high-$B^0$ mass
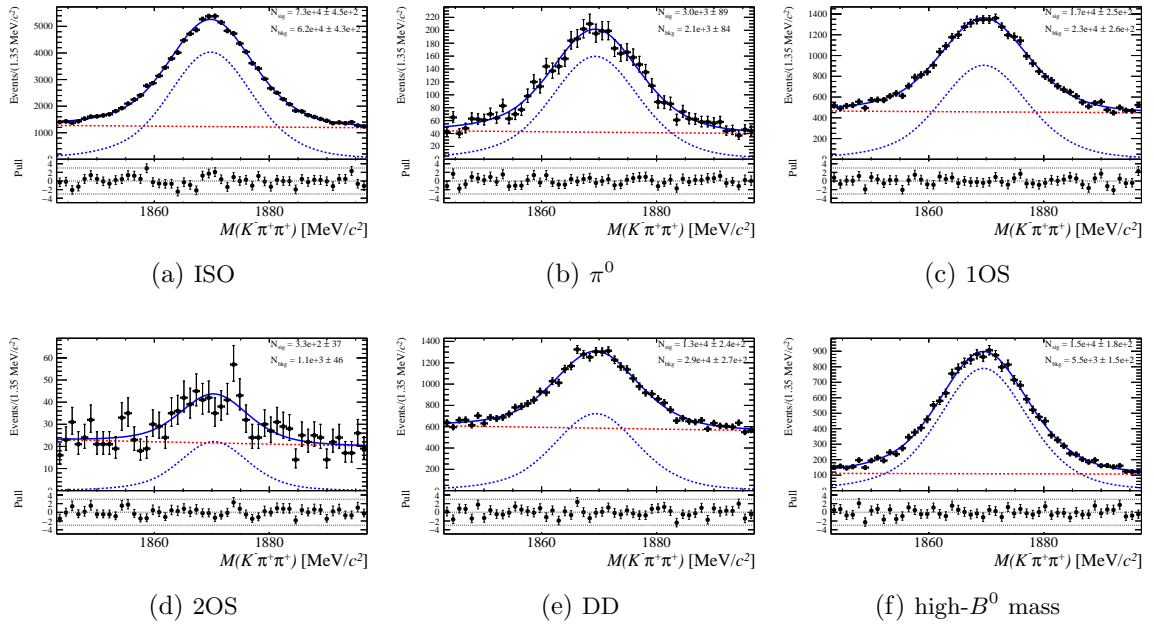
Figure 66: Fit to the reconstructed $D^+$ mass used to subtract the $D$-combinatorial background, 2016 WS data sample.

(a) ISO     (b) $\pi^0$     (c) 1OS

(d) 2OS     (e) DD     (f) high-$B^0$ mass
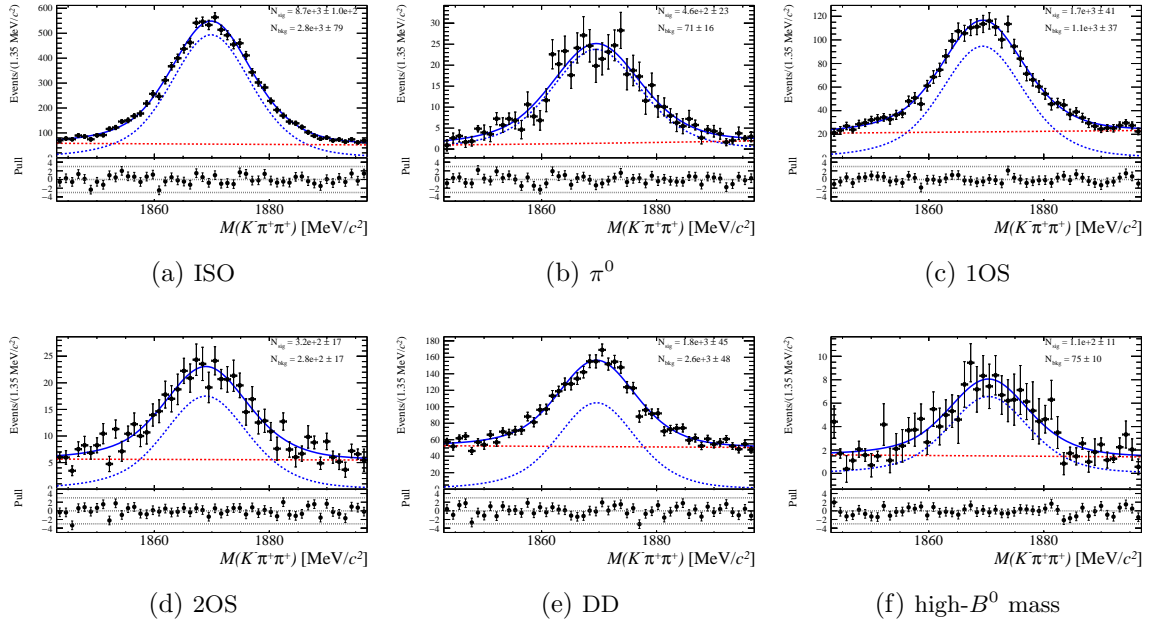
Figure 67: Fit to the reconstructed $D^+$ mass used to subtract the $D$-combinatorial background, 2015 MisID RS data sample.



(a) ISO     (b) $\pi^0$     (c) 1OS

(d) 2OS     (e) DD     (f) high-$B^0$ mass
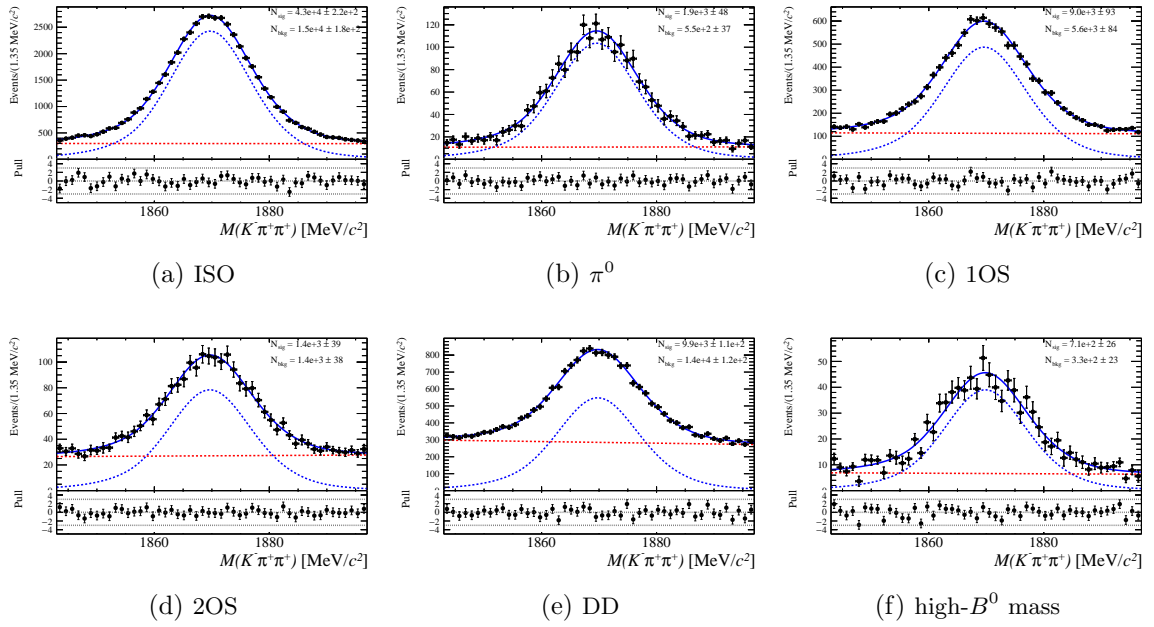
Figure 68: Fit to the reconstructed $D^+$ mass used to subtract the $D$-combinatorial background, 2016 MisID RS data sample.

127
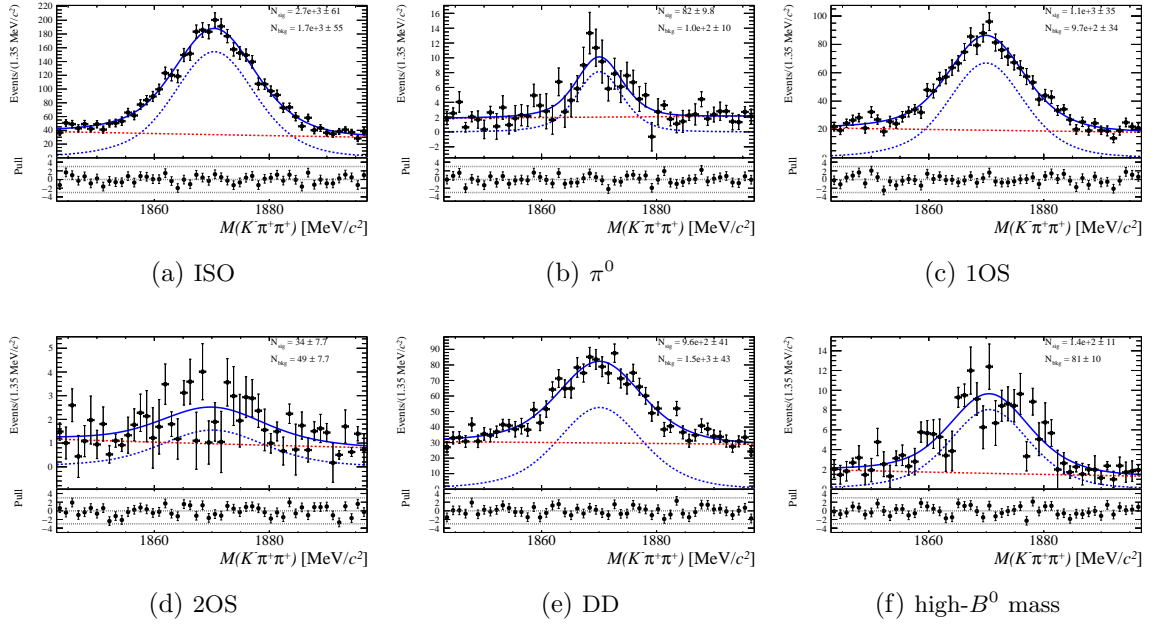
Figure 69: Fit to the reconstructed $D^+$ mass used to subtract the $D$-combinatorial background, 2015 MisID WS data sample.
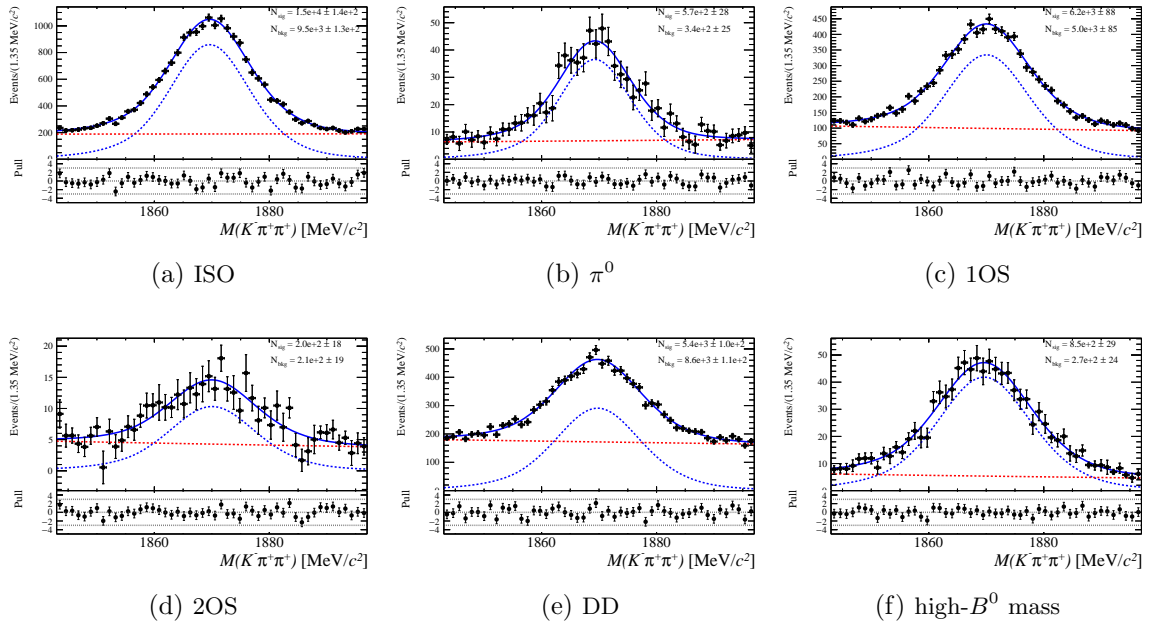


Figure 70: Fit to the reconstructed $D^+$ mass used to subtract the $D$-combinatorial background, 2016 MisID WS data sample.

## 12.3   $D\mu$-combinatorial

The $D\mu$-combinatorial background is made of events in which a true $D$ and a true $\mu$ coming from decays of uncorrelated particles in the event have been associated to the same signal vertex.

This background is modelled using the WS sample, after having subtracted the $D$-combinatorial, as described in Sect 12.2. After this subtraction, the WS sample is constituted by random combinations of a true $D^+$ candidate with tracks that have been identified as $\mu$. Combinations of a true $D^+$ candidate with hadrons tracks misidentified as a $\mu$ are subtracted, in order to avoid double counting of this background, that is already taken into account in the $\mu$-MisID sample. This is done by statistically subtracting the WS MisID sample, after having it weighted using the procedure detailed in Sect. 14.9.

The assumption for the evaluation of this background is that the shape of the WS sample combinatorial well describes the shape of the RS combinatorial. Even if this assumption is satisfied, the physical processes contributing to the WS sample may be different with respect to the ones contributing to the RS sample. This may manifest itself with a difference in the normalization of the WS combinatorial with respect to the one of the RS combinatorial. In order to correct for this effect, the RS combinatorial is obtained scaling the WS combinatorial sample by a weight depending on the $D\mu$ mass. This function is determined fitting the ratio of the yield of the RS and WS combinatorial as a function of the $D\mu$ mass, in the high-$B^0$ mass region ($m(D\mu) > 5.3$ GeV/$c^2$), in which only the $D\mu$-combinatorial and $\mu$-MisID are present. The ratio is fitted with an exponential function, as reported in Fig. 71, and is then extrapolated to the signal region.
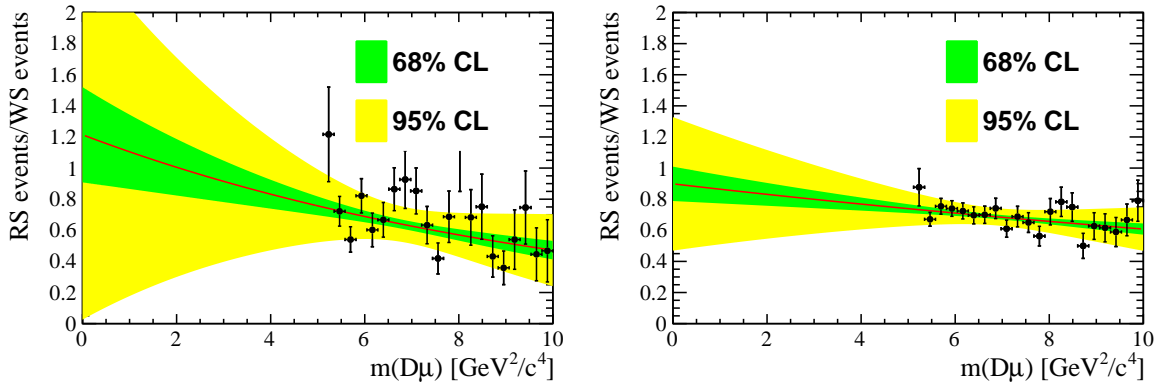


Figure 71: Fit to the RS/WS ratio of $D\mu$ invariant mass distributions for the (right) 2015 and (left) 2016 datasets.

## 12.4   $D\mu$-combinatorial validation

The high-$B^0$ mass region is expected to be populated mainly by combinatorial background, apart from a small contribution of $\mu$-MisID. This offers the opportunity to check if the WS sample is a good proxy for the RS combinatorial background, by checking the shape of the two data samples in this region, after having applied the corrections described in the previous section.

The normalization of both the $D\mu$-combinatorial and $\mu$-MisID templates is fixed to be the

same as the one expected for the two backgrounds in this region. The comparison is made for the variables used in the fit as well as for the $B^0$ mass, used for the correction. Additional variables, related to the kinematic of the decays and to the detector response are also checked. The comparison is made for the 2015+2016 combined sample and the result is reported in Figures 72,73,74,75 and 76.

As can be noticed in Fig. 72, the mass dependent correction reported in Fig. 71 completely accounts for the discrepancy in the $B^0$ mass distribution between the WS and RS samples, and the fit variables distributions are very well reproduced by the templates shapes and their expected normalization.
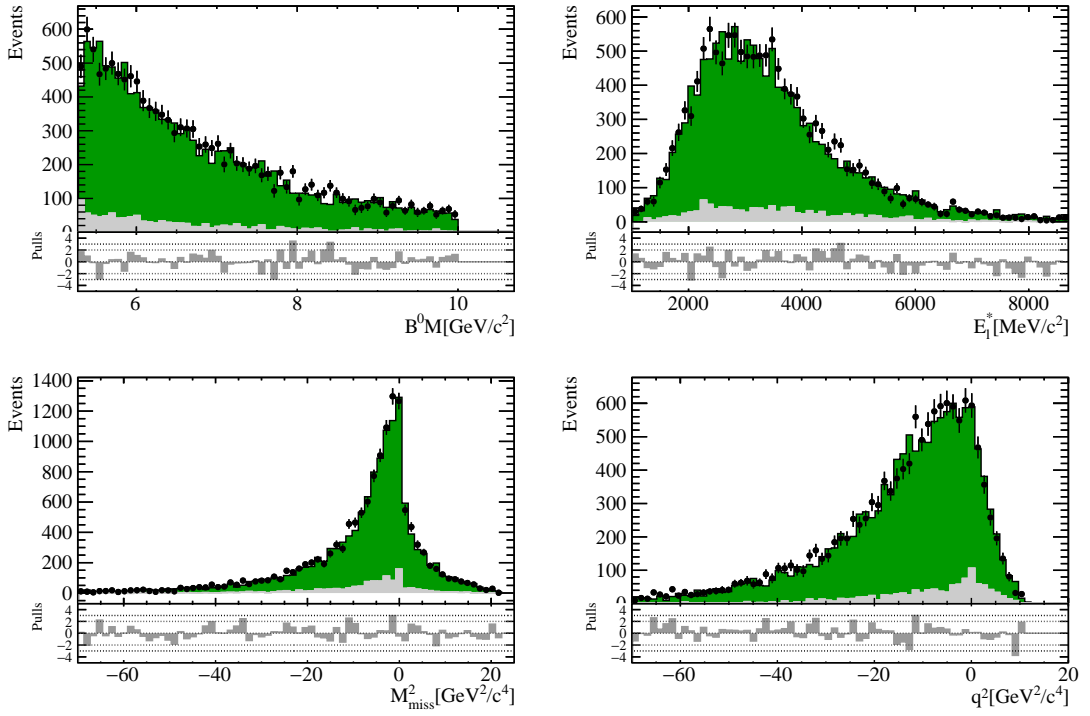


Figure 72: Comparison between the data, the $D\mu$-combinatorial (green) and $\mu$-MisID (gray) templates, in the high $B^0$ mass sideband region, 2015 + 2016 combined sample combined, for the fit variables and the visible $B^0$ mass.
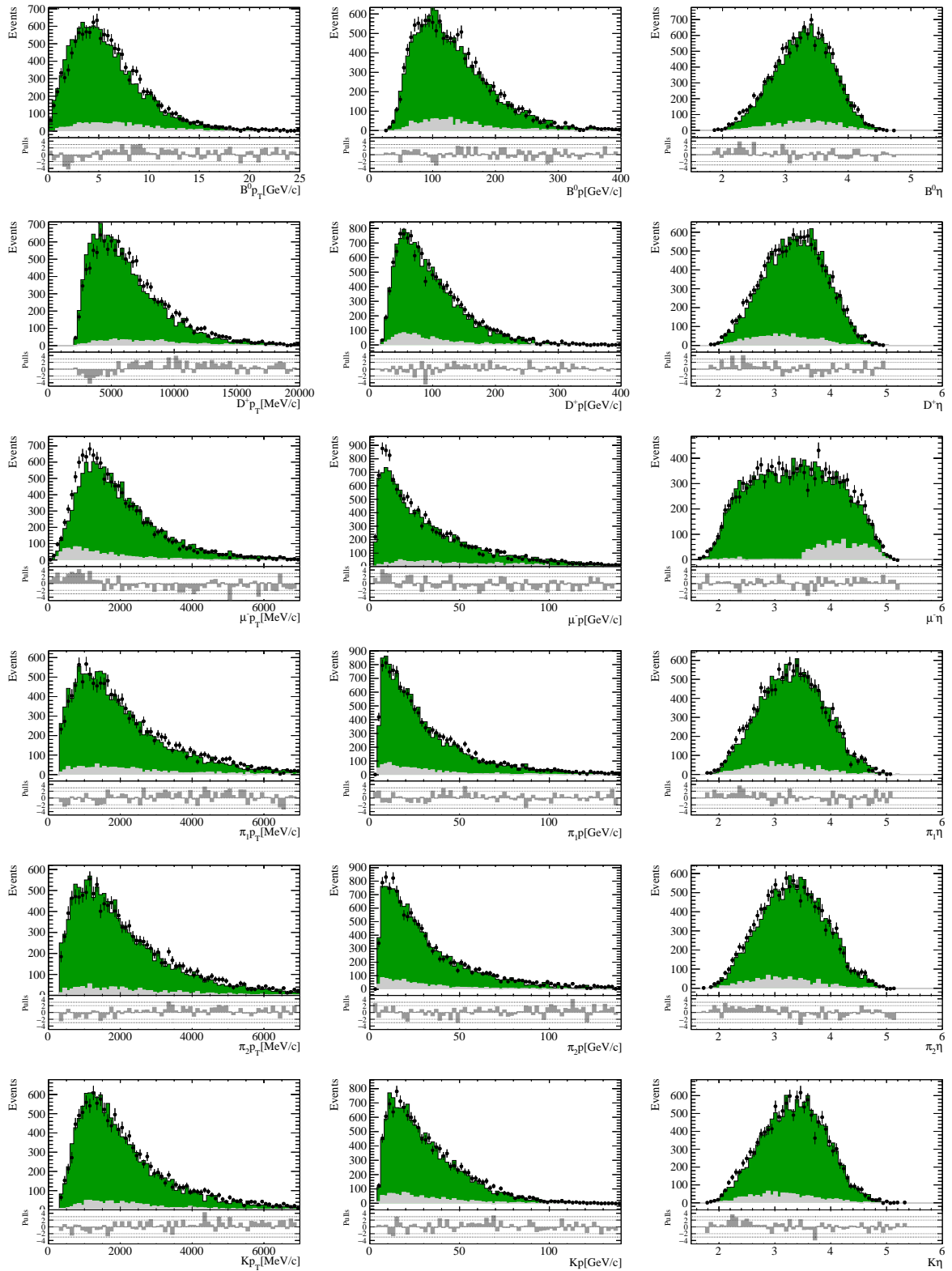
Figure 73: Comparison between the data, the $D\mu$-combinatorial (green) and $\mu$-MisID (gray) templates, in the high $B^0$ mass sideband region, 2015 + 2016 combined sample, for kinematic distributions of the signal candidates.

131
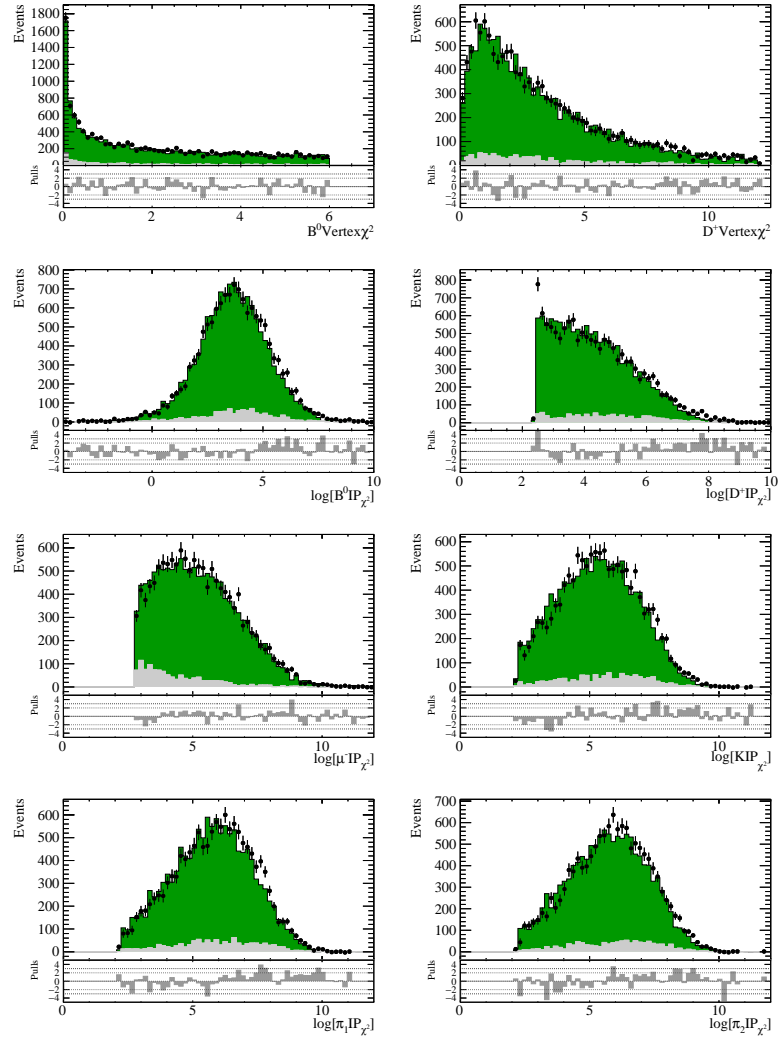
Figure 74: Comparison between the data, the $D\mu$-combinatorial (green) and $\mu$-MisID (gray) templates, in the high $B^0$ mass sideband region, 2015 + 2016 combined sample, for the $IP_{\chi^2}$ of the candidates and the Vertex $\chi^2$ of the $B^0$ and $D^+$.
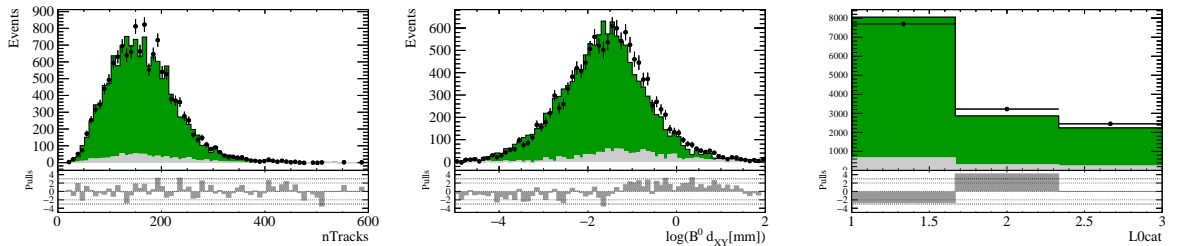


Figure 75: Comparison between the data, the $D\mu$-combinatorial (green) and $\mu$-MisID (gray) templates, in the high $B^0$ mass sideband region, 2015 + 2016 combined sample, for the distributions of nTracks, the transverse plane flight distance of the $B^0$ and the L0 category (1 = TIS & !TOS, 2 = TOS & ! TIS, 3 = TIS & TOS)
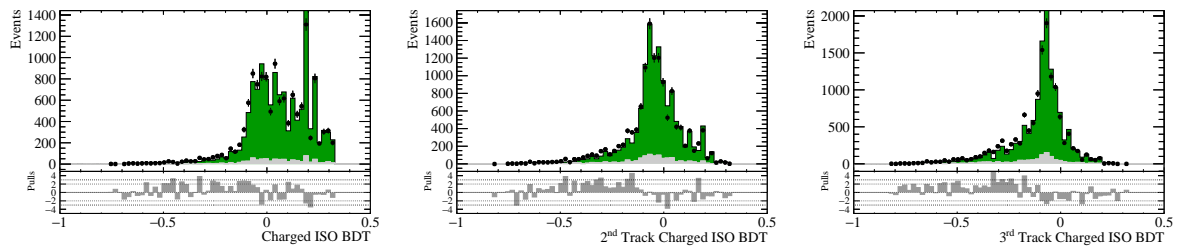
Figure 76: Comparison between the data, the $D\mu$-combinatorial (green) and $\mu$-MisID (gray) templates, in the high $B^0$ mass sideband region, 2015 + 2016 combined sample, for the distributions of the Charged Isolation BDT of the first, second and third most anti-isolated charged track.

# 13 Corrections on Monte Carlo simulation

## 13.1 $b-$production

The hadrons produced within the $pp$ collision and their kinematics is produced with Pythia. It is known that the kinematic of the $b$ hadrons and the tracks multiplicity are not well described by Pythia. In order to correct for this, a control channel of $B^+ \to J/\psi K^+$ is used to compare data and MC and to correct for data/MC differences.

The data and MC events are reconstructed and selected using the same selection. The trigger lines used require the presence of one or two muons combinations, with high $p_T$ and detached from the primary vertex. In order to better align the MC events to our datasets, the number of hits in the SPD is required not to exceed the value of 450. This cut has been placed directly on real data events and emulated on the MC events in the same way as it has been described in Section 10.3.

The MC events have been truth matched to the signal, whereas the combinatorial background is removed from real data events by means of sWeights [93] evaluated with a maximum likelihood fit to the $J/\psi K$ invariant mass distribution. The model used for the signal is an Ipatia function, whose coefficients have been calibrated from the MC calibration and fixed in the fit, apart from the mean and width which are left floating in the fit. The model for the combinatorial background is an exponential function whose slope is left floating in the fit. The results of the fits to the MC and data samples are shown in Fig. 77, separately for the 2015 and 2016 data taking conditions.

A *bin-by-bin* per event weight is evaluated by looking at the normalized distributions of data and background subtracted MC in bins of the following variables:

- $\log(B^+ p_T)$;

- Pseudorapidity ($\eta$) of the $B^+$;

- Number of charged tracks in the event ($nTracks$).

The weights are normalized such that they preserve the normalization of the MC sample with which they have been evaluated, and then are applied to all the MC used in the analysis. Given that the MC samples used in the analysis are made of semileptonic decays, in which the definition of the $B$ momentum and pseudorapidity is not evaluated in the same way as in the control channel due to the presence of the neutrino in the final state, the truth level quantities are instead used to assign the weights. The effect of the reweighting on the training variables and some kinematic variables for the daughter particles, on the control channel used to evaluate the weights, are shown in Fig 78.

The simulation version used to produce the $B \to J/\psi K$ control channel is different with respect to the one used to produce the MC used for the background estimation of the analysis. For part of the MC used, the Pythia version used has been found to be affected by a bug for which the broad resonances are not being decayed in the correct way, leading to a softer $nTracks$ spectrum. Given that this bug is not present in the control channel used to train the MC corrections, this bug would not be corrected for properly by using only the correction weights described above. Given that this bug is present in the simulation version that is used to produce the $B \to D\mu\nu$ 2016 sample but not in the one used to produce the $B \to D\mu\nu$ 2015 sample, $nTracks$ distribution of these two MC samples has been used to generate a *bin-by-bin* correction factor to be applied to all MC samples affected by the issue.
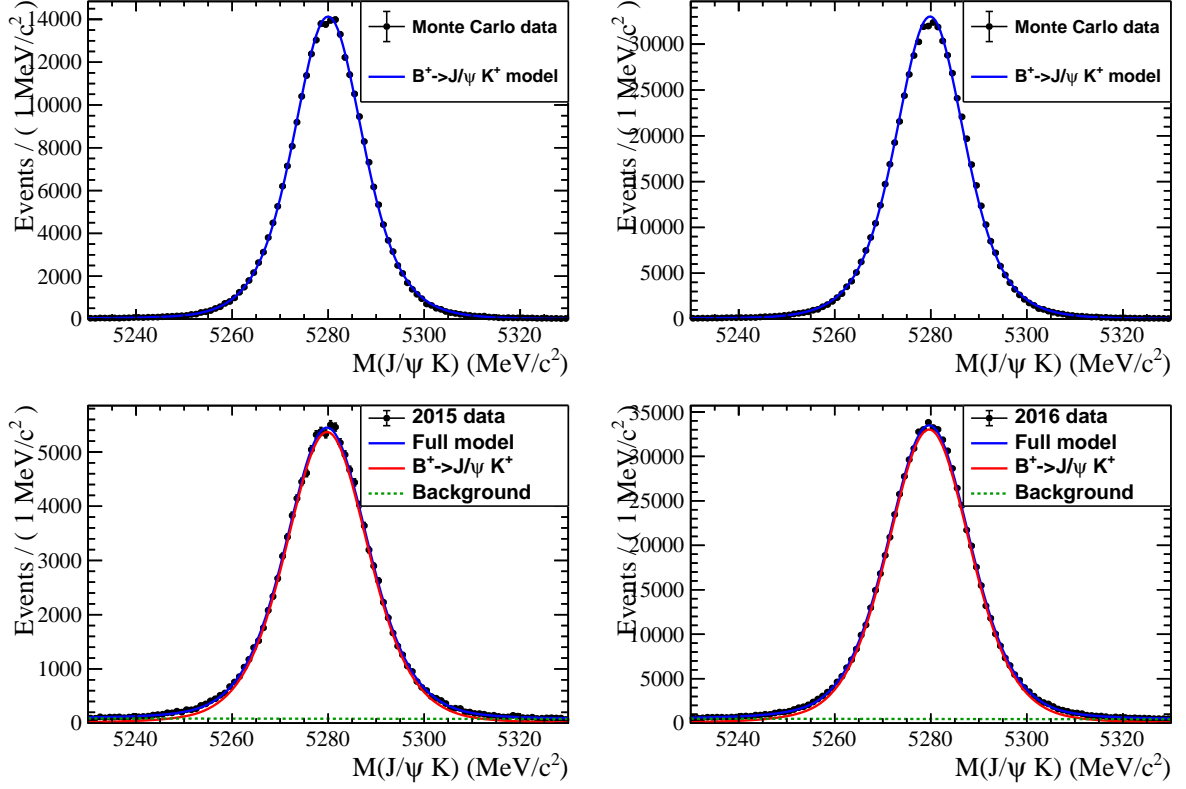
Figure 77: Mass fit to the (left) 2015 and (right) 2016 $B^+ \to J/\psi K^+$ (top) MC and (bottom) data samples.

The effect of this reweighting on $nTracks$ and other kinematical variables of $B \to D\mu\nu$ 2016 events is shown in Fig. 79.

The weights to correct the bug on the nTracks and the ones to correct data/MC discrepancies are then multiplied together in what will be called the *kinematical reweighting*, from now on.

The effect of the kinematical reweighting on the shape of the variables has been found to be very small. The effect that it has on the efficiencies and efficiency ratios is reported in Table 34. This is evaluated as the ratio of the value of the efficiencies and efficiency ratios, as evaluated using the weighted samples, and the ones evaluated using samples without the weights applied.
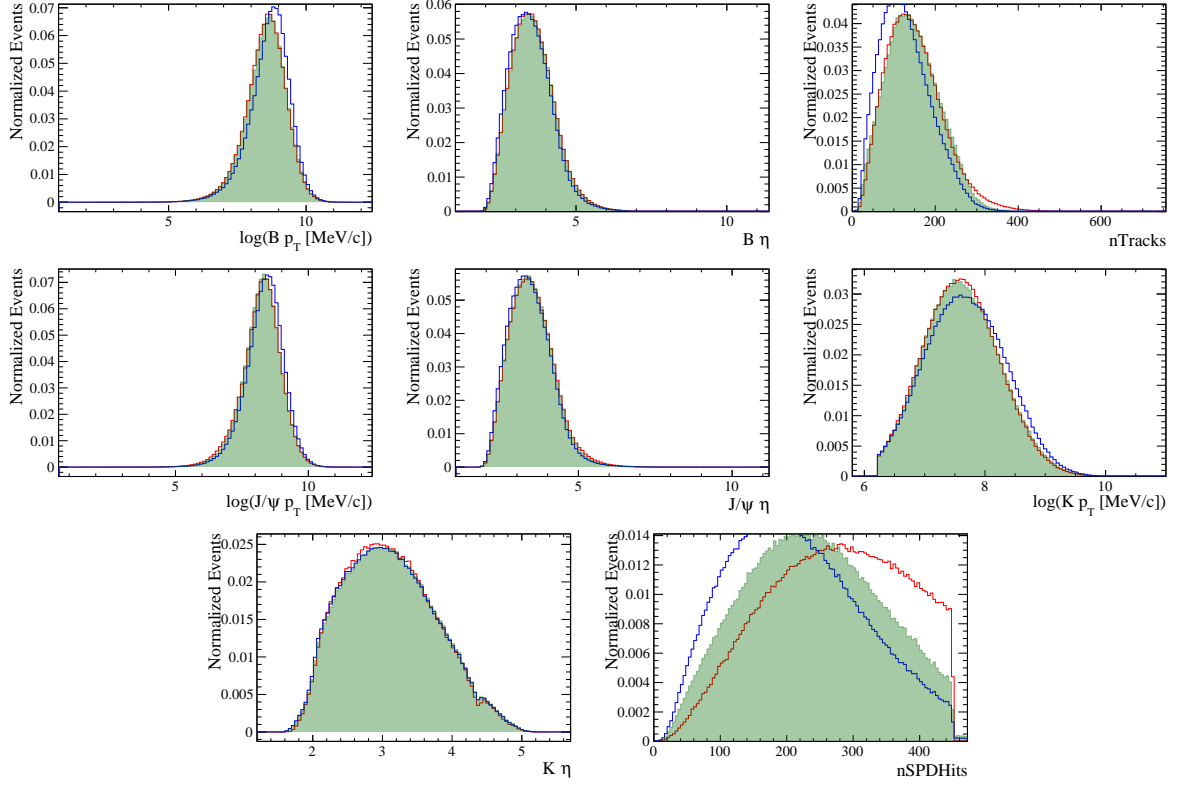
Figure 78: Effect of the reweight on several $B^+ \to J/\psi K^+$ 2016 MC distributions. The original MC distributions are shown in blue, the sWeighted 2016 data sample is shown in red, and the weighted MC distributions are shown in green.

| Sample | 2015 | | 2016 | |
|---|---|---|---|---|
| | $\varepsilon$ | ratio | $\varepsilon$ | ratio |
| $\bar{B}^0 \to D^+\tau(\to \mu\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$ | 0.881 | 0.981 | 0.837 | 0.978 |
| $\bar{B}^0 \to D^+\mu\bar{\nu}_\mu$ | 0.898 | | 0.856 | |
| $\bar{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$ | 0.878 | 0.982 | 0.835 | 0.978 |
| $\bar{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\bar{\nu}_\mu$ | 0.894 | | 0.854 | |

Table 34: Effect of the $B$-kinematic reweighting on the efficiencies and efficiency ratios.
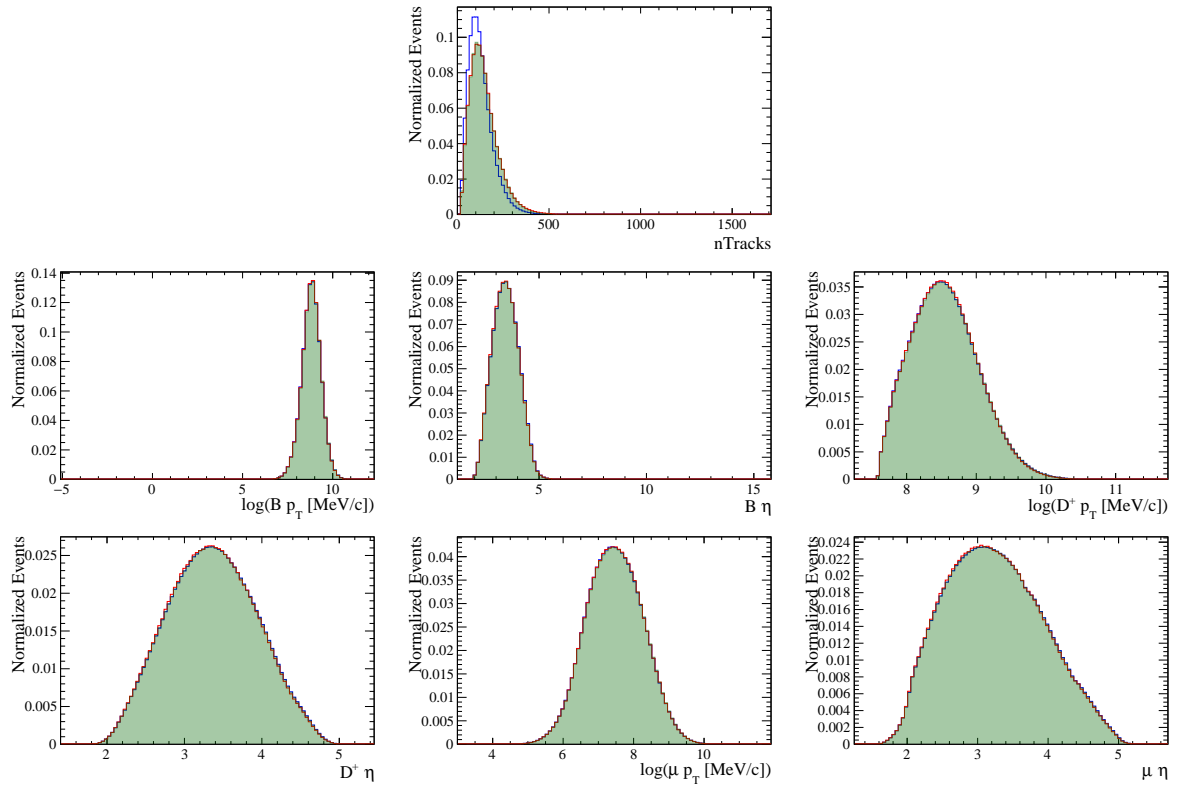
Figure 79: Effect of the reweight on nTracks on several $B \to D\mu\nu$ MC distributions. The original MC distributions in 2015 MC sample are shown in blue, the reference MC distributions in 2016 MC sample is shown in red, and the weighted 2015 MC distributions are shown in green.

## 13.2   Tracking efficiency

The track reconstruction efficiency measures the probability that the trajectory of a charged particle passing through the full tracking system is reconstructed.

The reconstruction sequence for charged particles is run in simulation in the same way as in data, after having simulated the interaction of the particles in the event with the material of the detector. Therefore, the knowledge of the track reconstruction efficiency for charged particles in the event relies on the correct simulation of the particles interactions with the material. To take possible differences between simulation and real data into account, the track reconstruction efficiency is corrected using a data-driven technique, using standard tools and calibration samples provided by the LHCb collaboration [96].

The track reconstruction efficiency is evaluated in data for $\mu$ tracks coming from a $J/\psi \to \mu^+\mu^-$ decays. This is an ideal channel to calibrate the tracking efficiency as it is very abundant and clean. The purity of the sample is further increased by triggering on events in which the $J/\psi$ candidate is likely to come from a $b$-hadron decay.

The efficiency is evaluated using the *tag-and-probe* method. One of the two tracks, called the *tag*, is fully reconstructed and well identified as a muon. The other, called *probe*, is only partially reconstructed, using the information coming from a subset of all the tracking detectors, used to get an estimate of the particle momentum. Using this information, the invariant mass of the $J/\psi$ candidate is reconstructed, which allows for discrimination against combinatorial background. The track reconstruction efficiency for Long tracks is then determined by matching the partially reconstructed probe to a Long track. The fraction of signal events with a matched probe provides a data driven determination of the track reconstruction efficiency. Depending on the tracking subdetector used to define the probe tracks, various methods can be defined.

In the *VELO method*, the tracking efficiency in the VELO stations is measured using downstream tracks as probes. A probe track is considered to be matched to a Long track if they share at least 50% of the hits in the T stations.

In the *T method*, the tracking efficiency in the T stations is evaluated using tracks that have segments in both the VELO and muon stations. The probe tracks used are the ones matched to VELO segments using a dedicated matching algorithm, starting from a muon segment. The probe is considered matched to a Long track if they stem from the same VELO seed and at least two hits in the muon stations are compatible with the extrapolation of the Long track.

In the *Long method* the tracking efficiency for long tracks is evaluated using tracks that have hits in both the muon and TT stations, because long tracks do not require the presence in neither of these two subdetectors. These probe tracks are reconstructed by matching the muon tracks to the TT hits by a dedicated algorithm. These TT-muon tracks are then considered matched to a Long track, if more than 70% of the hits in the muon stations are compatible with the extrapolation of the long track in the muon stations, and they share more than 60% of the hits in the TT stations, if present in the Long track.

The efficiencies are determined with the three methods, in bins of $p$ and $\eta$ of the tracks. The output of the various methods are then combined. The output of the *T* and *VELO methods* is multiplied, in each kinematic bin, given that they are independent of one another (*Combined method*). This gives an estimate of the Long tracking efficiency, which can be compared with the output of the *Long method*. Small systematic differences between the two can be expected, due to the difference hits and matching requirement of the methods. The *method*, for example measures the efficiency for tracks with hits in the TT station, which is also a requirement for the

probe tracks of the *VELO* method, but not of the *T method*. Furthermore, both the *T* and the *VELO methods* include the efficiency that, given a VELO or T track, the corresponding Long track is found. This matching efficiency is not present in the *Long* method, which uses as probe tracks constructed with hits not needed by the Long track reconstruction, but it is counted twice when combining the *VELO* and *T* methods. The output of the *Combined* and *Long* method are then combined using a weighted average.

The same procedure is then performed in MC after having reweighted the distribution of the number of hits in the SPD ($nSPDHits$), because the detector occupancy is known to be not well reproduced in simulation and very correlated with the track reconstruction efficiency. The ratio between the track reconstruction efficiency in data and MC, $\frac{\varepsilon_{data}}{\varepsilon_{MC}}$, is then measured in each kinematic bin.

At the time of the writing of this thesis, for the 2016 samples, only the *Long method* is available for 2016 data taking conditions. Therefore, for this portion of the dataset, the efficiency ratio is evaluated using only the *Long method*. The simulation versions used to produce the 2016 MC is either *Sim*09*c* or *Sim*09*d*, but only *Sim*09*b* MC calibration samples are provided. Given that this analysis aims at measuring a ratio of branching fractions, any discrepancy due to the simulation version is assumed to be cancelling in the ratio, since it would affect both signal and normalization modes.

For bins in which the statistical uncertainty exceeds 3%, the correction is set to 1 with a 5% statistical uncertainty. The value of the correction in the various kinematic bins is reported in Fig. 80, where it can be noticed that the ratio of efficiencies has been found to be always very close to unity. The trigger selection used to record the calibration samples requires a minimum transverse momentum of 100 MeV and a minimum momentum of 5 GeV for the probe tracks. All signal hadrons have tighter momentum requirements, but muon tracks used in our analysis are, instead, kept as less biassed as possible and no minimum momentum requirement is present. They are required to have a momentum higher than 3 GeV, and no requirement is placed on their transverse momentum. Furthermore, the correction tables provided are defined for a range between 1.9 and 4.9 in pseudorapidity: a small portion of the pseudorapidity spectrum of all the signal tracks in the event falls outside this range. For all the tracks falling outside the support of the correction tables, the correction associated to the closest bin within the boundaries, with an uncertainty of 5% is associated. The $p$ and $\eta$ distribution, for the signal and normalization modes, is reported in Fig. 81, where the bin boundaries used in the table are reported as the vertical red lines and the portion of the spectrum which falls outside the correction tables is highlighted in darker blue.

For each event, the track efficiency correction associated to each signal track of the event is read from the tables. The per-event tracking efficiency correction is then defined as the product of each contribution.

The effect of the tracking correction efficiency on the efficiencies and on the efficiency ratios is reported in Table 35. The effect is estimated by evaluating the efficiency and efficiency ratios before and after the tracking efficiency corrections are applied to the MC samples. The table reports the ratio between the value of the efficiencies after the correction with respect to the one before the correction.
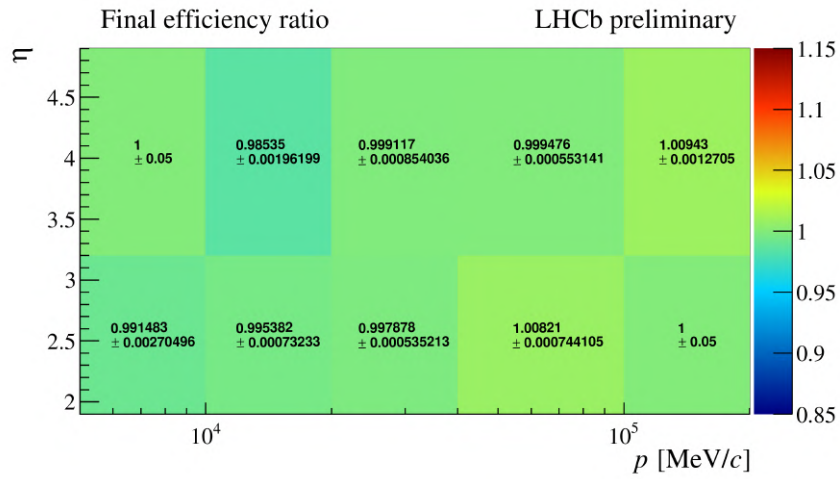
(a) 2015
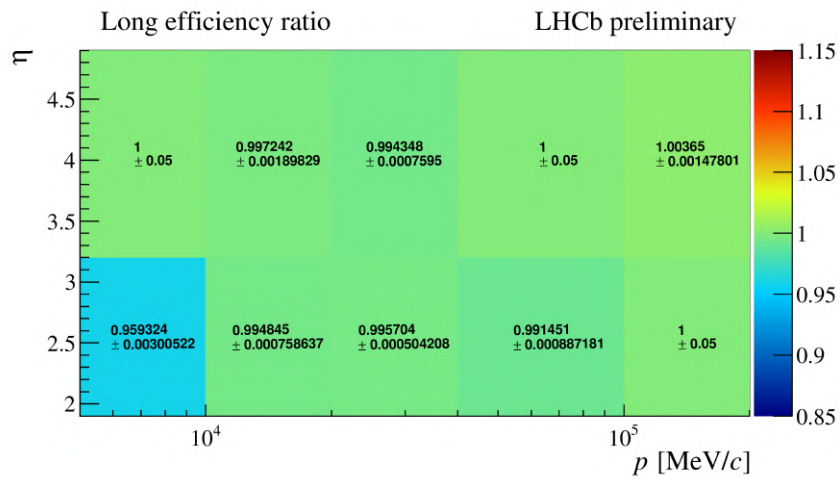


(b) 2016

Figure 80: Track reconstruction efficiency ratio tables $(\frac{\varepsilon_{data}}{\varepsilon_{MC}})$ used in the analysis for 2015 and 2016 datasets.

(b) $B \to D\mu\nu$



(d) $B \to D\tau\nu$
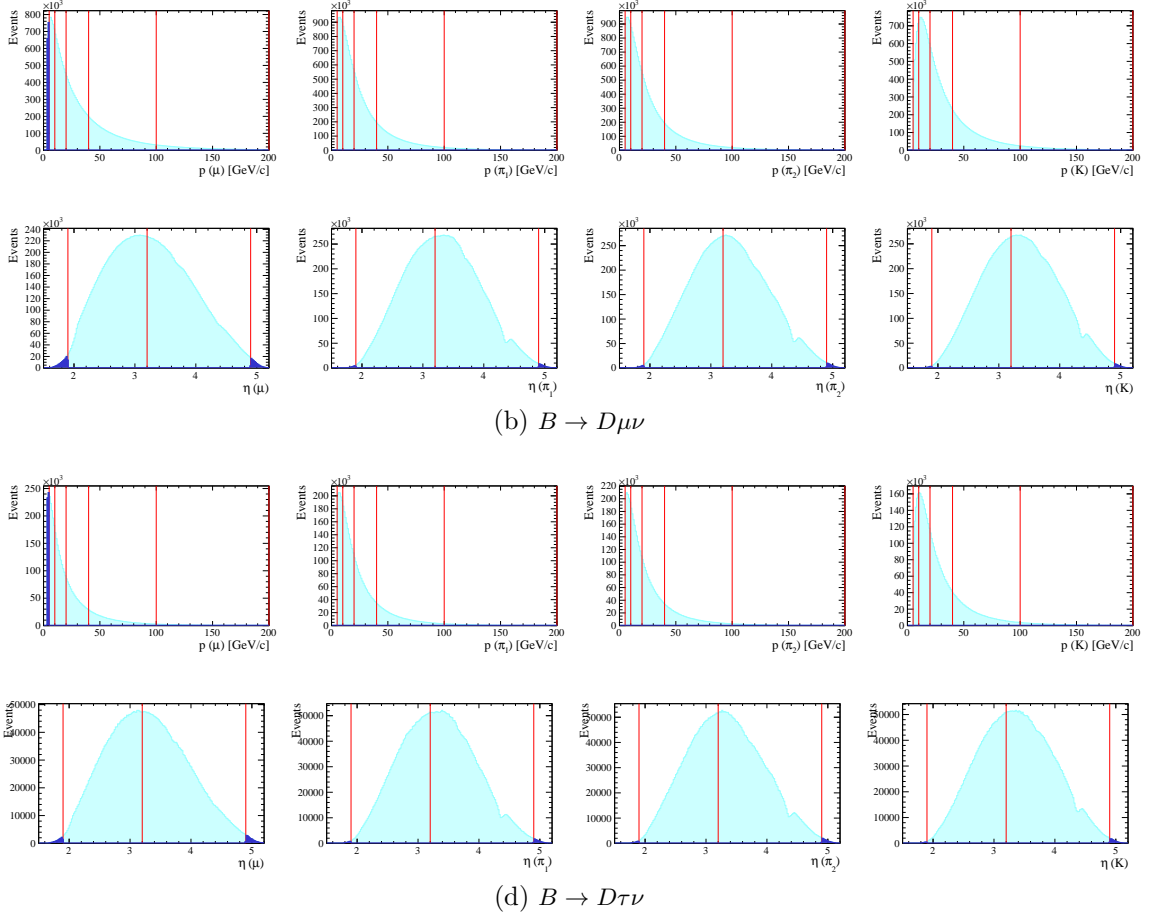
Figure 81: Kinematic distribution for signal and normalization 2016 MC samples. The red lines indicate the bin boundaries used in the correction tables. Events in darker blue indicate events that do not fall inside the corrections table support

| | 2015 | | 2016 | |
|---|---|---|---|---|
| Sample | $\varepsilon$ | ratio | $\varepsilon$ | ratio |
| $\overline{B}^0 \to D^+\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 0.984 | 0.997 | 0.961 | 0.997 |
| $\overline{B}^0 \to D^+\mu\overline{\nu}_\mu$ | 0.987 | | 0.964 | |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | 0.983 | 0.997 | 0.96 | 0.997 |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\overline{\nu}_\mu$ | 0.986 | | 0.963 | |

Table 35: Effect of the tracking correction efficiency on the efficiencies and on the efficiency ratios.

141

## 13.3    $\mathcal{B}(D_q \to \ell\nu X)$ in production

To estimate the Double Charm background, data-sets containing cocktails of $B \to D^{*,+}D_q(X')$ events, where $D_q = (D_s, D^0, D^\pm)$ are bound to decay into a $\mu$, have been generated with EvtGen. The various decays with the different charm states are produced in the MC samples with different fractions, which depend on the $\mathcal{B}(D_q \to \mu\nu X)$ branching fractions that are configured in EvtGen. When producing events with EvtGen, although, a subtlety plays a role in the definition of the fractions of the different charm states.

If the branching fractions of any particle declared in the configuration of EvtGen do not sum up to 1, the branching fractions are firstly renormalized so that they sum up to unity. After this normalization, the fraction of events from the different $D_q$ contributions do not take into account the inclusive $\mathcal{B}(D_q \to \mu\nu X)$ branching fractions anymore. This could in principle affect the shape of the templates if the shape of the various $D_q$ contributions was much different.

In order to correct for this, a weight equal to the inclusive $\mathcal{B}(D_q \to \mu\nu X)$ branching fraction has been assigned to each event. This weight is enough for the purpose of this analysis, since only the relative contributions matter in the fit, being the overall yield of the double-charm component measured in the fit in each control region independently.

For the $D^0$ and $D^+$ states, the inclusive branching fraction for their semimuonic decay is available in the PDG and directly used in the reweighting. These are reported in the Tables 36 and 37, along with all the exclusive branching fractions used in the cocktails, as well as those reported in the PDG, for comparison. Some semileptonic decays have only been measured in the electron final state: for them, the electron mode branching fraction is used and reported in the table, highlighted in blue.

For the $D_s$, no available measurement of the inclusive $\mathcal{B}(D_s \to \mu\nu X)$ is available. To evaluate it, the sum of the exclusive muonic decays branching fractions reported in Table 38 is used.

The result of this reweighting on the shape of the muonic double charm templates is reported in Fig. 82 and 83. No appreciable shape difference is seen between the templates before and after the reweighting. For this reason, no further systematic uncertainty is associated to this reweighting.

| decay | $\mathcal{B}$ in decfile (%) | $\mathcal{B}$ PDG (%) | $\mathcal{B}$ used (%) |
|---|---|---|---|
| $D^0 \to \mu^+ \nu_\mu X$ | - | $6.8 \pm 0.6$ | 6.8 |
| $D^0 \to K^- \mu^+ \nu_\mu$ | 3.33 | $3.41 \pm 0.04$ | |
| $D^0 \to K^{*-} \mu^+ \nu_\mu$ | 1.92 | $1.89 \pm 0.04$ | |
| $D^0 \to K^- \pi^0 \mu^+ \nu_\mu$ | 0.040 | $1.6^{+1.3}_{-0.5}$ | |
| $D^0 \to \bar{K}^0 \pi^- \mu^+ \nu_\mu$ | - | $1.44 \pm 0.04$ | |
| $D^0 \to \pi^- \mu^+ \nu_\mu$ | 0.238 | $0.267 \pm 0.012$ | |
| $D^0 \to \rho^- \mu^+ \nu_\mu$ | - | $0.150 \pm 0.012$ | |
| $D^0 \to \pi^- \pi^0 \mu^+ \nu_\mu$ | - | $0.145 \pm 0.007$ | |
| $D^0 \to (a^- \to \eta\pi^-)\mu^+ \nu_\mu$ | - | $0.133 \pm 0.034$ | |
| $D^0 \to K_1^{'-} \mu^+ \nu_\mu$ | 0.076 | $0.076 \pm 0.004$ | |
| $D^0 \to K_2^{*-} \mu^+ \nu_\mu$ | 0.11 | - | |

Table 36: Inclusive $D^0 \to \mu\nu X$ branching fractions used in the reweight and branching fractions for exclusive semi-leptonic decays found in the PDG, along with the ones used in the MC cocktail. The branching fractions reported in blue are the ones for the corresponding $D^0 \to e\nu X$ mode, as the one for the muonic decay is not present on the PDG.

| decay | $\mathcal{B}$ in decfile (%) | $\mathcal{B}$ PDG (%) | $\mathcal{B}$ used (%) |
|---|---|---|---|
| $D^+ \to \mu^+ \nu_\mu X$ | - | $17.6 \pm 3.2$ | 17.6 |
| $D^+ \to \bar{K}^0 \mu^+ \nu_\mu$ | 9.2 | $8.76 \pm 0.19$ | |
| $D^+ \to \bar{K}^{*0} \mu^+ \nu_\mu$ | 5.28 | $5.27 \pm 0.15$ | |
| $D^+ \to \pi^0 \mu^+ \nu_\mu$ | - | $0.350 \pm 0.015$ | |
| $D^+ \to \pi^+ \pi^- \mu^+ \nu_\mu$ | - | $0.245 \pm 0.010$ | |
| $D^+ \to \rho^0 \mu^+ \nu_\mu$ | - | $0.24 \pm 0.04$ | |
| $D^+ \to K^- \pi^+ \mu^+ \nu_\mu$ | 0.292 | $0.19 \pm 0.05$ | |
| $D^+ \to \omega \mu^+ \nu_\mu$ | - | $0.169 \pm 0.0011$ | |
| $D^+ \to \tau^+ \nu_\tau$ | - | $0.120 \pm 0.027$ | |
| $D^+ \to \eta \mu^+ \nu_\mu$ | - | $0.111 \pm 0.007$ | |
| $D^+ \to K_1^0 \mu^+ \nu_\mu$ | 0.277 | $0.106 \pm 0.015$(incl. $K_1^0$ decay) | |
| $D^+ \to \mu^+ \nu_\mu$ | - | $0.0374 \pm 0.0017$ | |
| $D^+ \to \eta' \mu^+ \nu_\mu$ | - | $0.020 \pm 0.004$ | |
| $D^+ \to (a \to \eta \pi^0) \mu^+ \nu_\mu$ | - | $0.017^{+0.008}_{-0.007}$ | |
| $D^+ \to K_2^{*0} \mu^+ \nu_\mu$ | 0.293 | - | |
| $D^+ \to \bar{K}^0 \pi^0 \mu^+ \nu_\mu$ | 0.120 | - | |

Table 37: Inclusive $D^+ \to \mu \nu X$ branching fractions used in the reweight and branching fractions for exclusive semi-leptonic decays found in the PDG, along with the ones used in the MC cocktail. The branching fractions reported in blue are the ones for the corresponding $D^+ \to e \nu X$ mode, as the one for the muonic decay is not present on the PDG.

| decay | $\mathcal{B}$ in decfile (%) | $\mathcal{B}$ PDG (%) | $\mathcal{B}$ used (%) |
|---|---|---|---|
| $D_s \to \mu^+\nu_\mu X$ | - | - | 7.829 |
| $D_s \to \phi\mu^+\nu_\mu$ | 2.59 | $3.03 \pm 0.24$ | 3.03 |
| $D_s \to \eta\mu^+\nu_\mu$ | 2.67 | $2.4 \pm 0.5$ | 2.4 |
| $D_s \to \eta'\mu^+\nu_\mu$ | 0.99 | $1.1 \pm 0.5$ | 1.1 |
| $D_s \to \mu^+\nu_\mu$ | 0.59 | $0.549 \pm 0.016$ | 0.549 |
| $D_s \to K^0\mu^+\nu_\mu$ | 0.37 | $0.34 \pm 0.04$ | 0.37 |
| $D_s \to K^{*0}\mu^+\nu_\mu$ | 0.18 | $0.215 \pm 0.028$ | 0.18 |
| $D_s \to f_0\mu^+\nu_\mu$ | 0.20 | $0.13 \pm 0.04$ | 0.20 |

Table 38: Exclusive $D_s \to \mu\nu X$ branching fractions used in the reweight and branching fractions for exclusive semi-leptonic decays found in the PDG, along with the ones used in the MC cocktail. The branching fractions reported in blue are the ones for the corresponding $D_s \to e\nu X$ mode, as the one for the muonic decay is not present on the PDG.
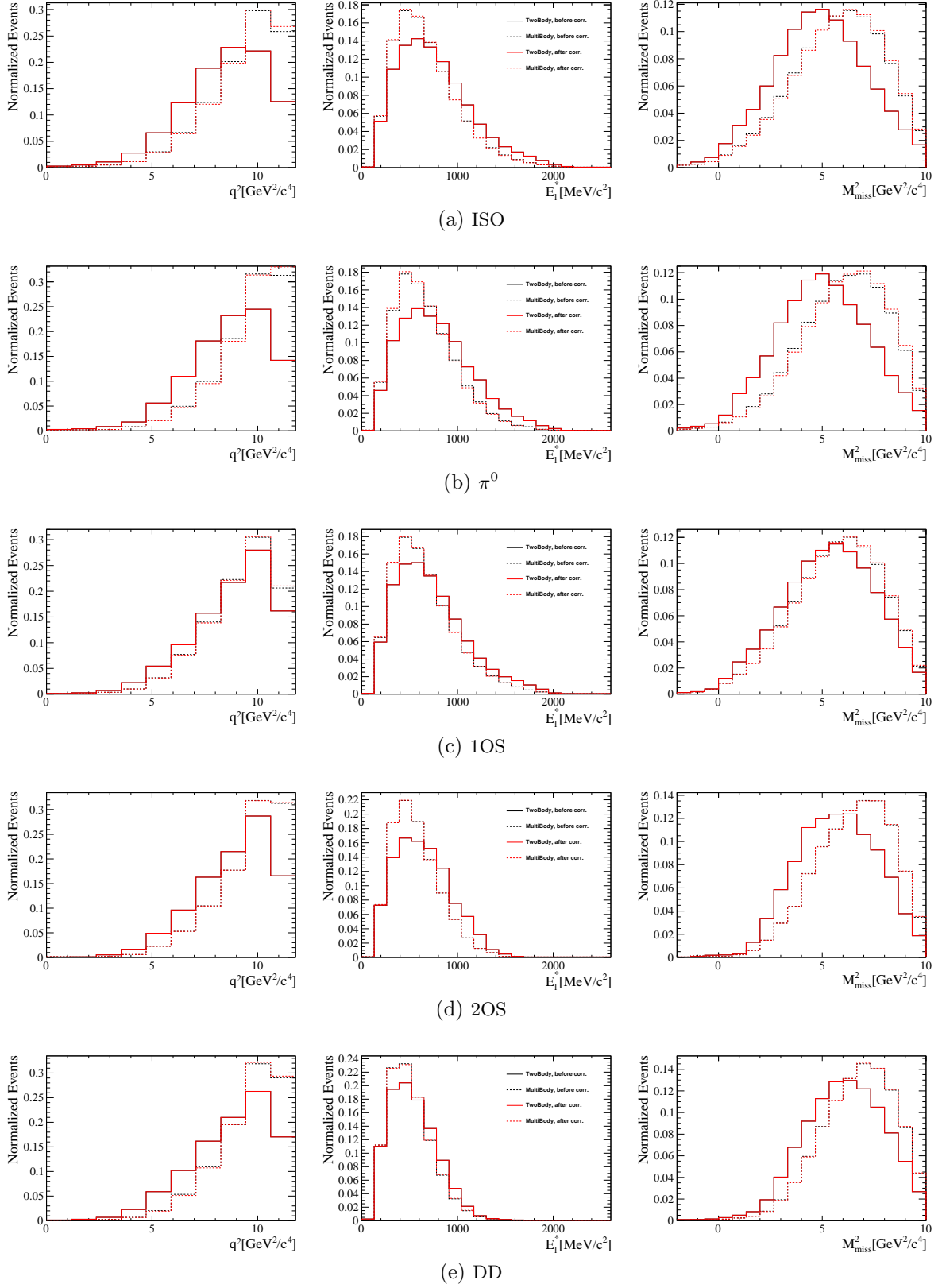
Figure 82: Effect of the corrections of the inclusive $\mathcal{B}(D_q \to \mu\nu X)$ on the $B^\pm \to D(X_c \to \mu\nu X)X'$ templates.

146
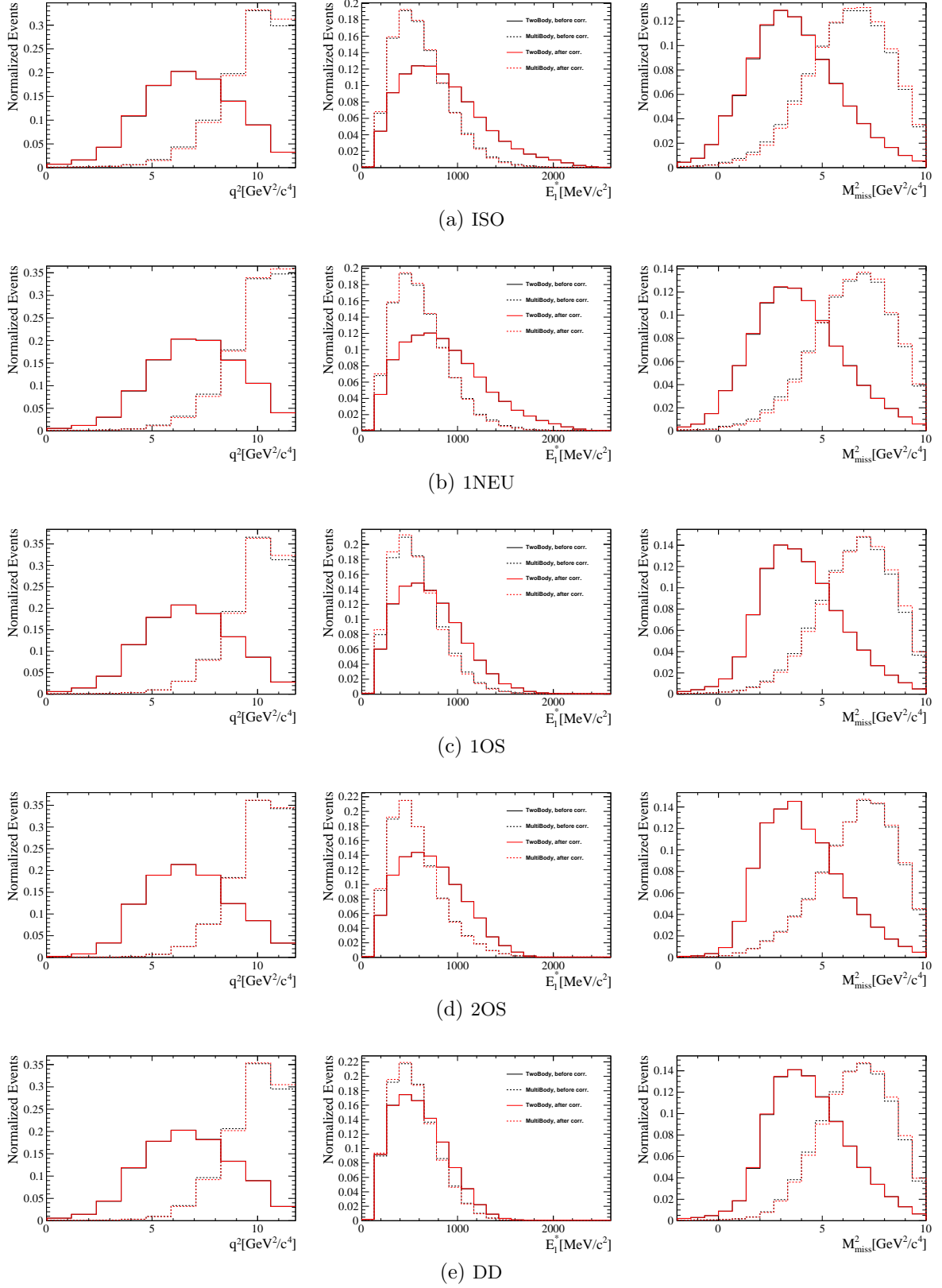
Figure 83: Effect of the corrections of the inclusive $\mathcal{B}(D_q \to \mu\nu X)$ on the $B^0 \to D(X_c \to \mu\nu X)X'$ templates.

## 13.4 Post-fit corrections

Many background sources contribute to the $D^+\mu$ final state, and many of them require a difficult modelling. Any comparison between the data and the MC used to estimate the shape of our templates must take this into account, since any discrepancy could either be due to a real data/MC difference, or to a mismodelling on the backgrounds if not to a bias in their expected yield.

In order to be able to decouple the two possible sources of mismatch, all the comparisons between data and simulation are performed in a region in which the dependency on the fit modelling, on the shape of the backgrounds and on the amount of signal events is reduced.

After a first iteration of the fit has been performed, the fit model is projected into a subregion of the ISO sample, which corresponds to the left tail of the $m^2_{\mathrm{miss}}$ spectrum ($m^2_{\mathrm{miss}} < -0.67\,\mathrm{GeV}^2/c^4$), where the normalization modes ($B \to D^{(*,+)}\mu\nu$) vastly dominate the events.

The number of events for each mode in this region is evaluated, and a cocktail of simulated events containing the right amount of each process is constructed. This cocktail is weighted using all the corrections described in the previous sections. The only corrections not applied to these events are the ones for the shape of the MultiBody Double Charm and Higher mass $D^{**}$ events, and the ones for the Form Factors of the normalization sample.

Given that the Combinatorial and MisID templates are derived in a fully data-driven way, their contribution is subtracted from the dataset, using their expected contribution in this region as derived from the fit result. The resulting dataset and the cocktail of background events constructed as described above are then normalized to the same area.

A `GBReweighter` instance is trained with the resulting dataset. The weights are normalized a posteriori to preserve the normalization of the cocktail sample in this region. The normalized weights are then applied to all the simulated events, in all the fit regions used for the fit.

Variables used during training are mostly coming from four sets: variables related to the modelling of the *rest of the event*, variables related to the Charged Isolation, variables related to the kinematics of the signal and variables related to the $D^+$ decay model.

For the *rest of the event*, a discrepancy is observed in the distribution of nTracks, even after the initial reweighting trained on $B \to J/\psi K$ events. Given that this variable is found to be very correlated with the trigger decision, the fraction of events triggered with each L0 trigger selection is inspected. Three mutually exclusive categories are defined: the first category, containing the fraction of events that pass the L0TIS but do not pass the L0TOS requirement; the second category, containing the fraction of events that pass the L0TOS but do not pass the L0TIS requirement; and finally the third category, containing the fraction of events passing both the L0TOS and L0TIS requirements. Since the simulation used is *tracker-only*, the trigger decision cannot be implemented on a per-event basis, but only using the trigger emulations described in the previous chapters. Each event is assigned randomly to a given category based on the per-event probability weights evaluated by the trigger emulations. In this way the events can be classified in the trigger categories on a statistical level, mantaining some degree of correlation between the trigger category and other variables.

By inspecting the L0 category, it can be seen that the fraction of events in the exclusive TIS category, before the reweighting, is overestimated with respect to the one in data. This might be one of the causes for the mismodelling of the nTracks distribution. It has also been seen that a single correction to the L0 category variable is not enough to correct for the nTracks distribution. Furthermore, it has been seen that this would also break the agreement between

the data and the model in the $D^+$ $p_T$ distribution. Since this variable is the relevant one for the L0TOS decision, this could be a signal of a mismodelling of the distributions within each L0 trigger category, especially given the fact that the L0TOS is fully emulated on simulated events, without any data correction. These discrepancies would not be corrected for by the kinematical reweight on $B \to J/\psi K$ events, since the L0 response was not taken into account. Therefore, all these variables are used in the reweighting to correct them simultaneously, and the $D^+$ $p_T$ is considered in order to correct for any discrepancy within each of the trigger categories related to the L0 modelling. The $D^+$ $p$ along with $p_T$ is also provided, in order to be able to catch any acceptance dependent effect.

Another discrepancy has been observed in the Charged Isolation output of the three most anti-isolated tracks. Also this variable has been observed to be correlated with the fraction of events triggered by the L0TIS decision. In order to be able to generalize the reweighting also to other samples in which the cut on the isolation output is the opposite to the one used in this region, instead of correcting the isolation directly, the input variables of the isolation BDT have been inspected. The input variables available offline are the Flight distance significance of the $D\mu$ vertex with respect to the associated primary vertex, the Track Type (Long, Velo, Downstream), and the angle between the track and the $D\mu$ candidate. The most significant discrepancies are observed in the two latter variables. All these variables are provided in input to the reweighting, and this has been found to be sufficient to be able to correct the discrepancies observed in the charged isolation output at an acceptable level.

Correcting for the Flight distance significance has been found to negatively affect the agreement between the data and the model in the $\mu$ impact parameter significance. Therefore, this variable is added to the list of reweighting variables, which is enough to correct for the correlated effect and improve the agreement in this variable.

Lastly, even though the model used to simulate the $D^+$ decay in EvtGen takes into account measurements of the $D^+ \to K\pi\pi$ Dalitz plot, some discrepancies have been observed in variables related to the $D^+$ decay modelling. This can be seen for example in the one dimensional distributions for the mass of pairs of daughters of the $D$ candidate. In order to correct for this, all the Dalitz variables of the decay ($m^2_{K\pi_1}$, $m^2_{K\pi_2}$ and $m^2_{\pi_1\pi_2}$) are used as input variables.

After the correction the agreement between the data and the model is found to be acceptable, with some residual differences being observed in the kinematics of the signal tracks, especially for events with low transverse momentum. This could be caused by tracking efficiencies and PID selection efficiencies at low momentum, for which the calibration samples are known to have small statistics.

The variables used in input for the training of the reweighter are reported in Table 39. The figures reported below show the comparison of the data and the model in this region, after reweighting. The model before this reweight is also overlaid, with a black dashed line. The pulls between the data and the model are shown below each plot, as a dark (light) grey histogram for the model after (before) the correction. The legend is reported in Fig. 84.

Fig. 91 also reports the Dalitz plots of the $D$ decay in the data and in the model. The pulls between the data and the model, before and after reweighting are also shown in the same Figure.

The effect of the data/model reweighting on the efficiencies and on the efficiency ratios is reported in Table 40. The effect is estimated by evaluating the efficiency and efficiency ratios before and after the corrections are applied to the simulation samples. The table reports the ratio between the value of the efficiencies after the correction with respect to the one before.

The effect of the data/model reweighting on the templates shapes is reported in Fig. 92

and 93.

| Objects | Variables |
|---|---|
| $D^+$, Most Anti-Isolated tracks | $p$, $p_T$ |
| $\pi_1$, $\pi_2$, $K$ | $m^2_{K\pi_1}$, $m^2_{K\pi_2}$, $m^2_{\pi_1\pi_2}$ |
| $(D^+\mu)$ | $\mathrm{FD}^{\mathrm{PV}}_{\chi^2}$ |
| $\mu$ | $\mathrm{IP}^{\mathrm{PV}}_{\chi^2}$ |
| Most Anti-Isolated tracks | $\cos(\theta_{D\mu})$, Track Type |
| Event | nTracks, L0 Category |

Table 39: List of variables used during the training of the data/model reweighter.



Figure 84: Legend

| | 2015 | | 2016 | |
|---|---|---|---|---|
| Sample | $\varepsilon$ | ratio | $\varepsilon$ | ratio |
| $\bar{B}^0 \to D^+\tau(\to \mu\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$ | 1.032 | | 1.005 | |
| | | 1.036 | | 1.036 |
| $\bar{B}^0 \to D^+\mu\bar{\nu}_\mu$ | 0.996 | | 0.97 | |
| $\bar{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$ | 1.045 | | 1.018 | |
| | | 1.034 | | 1.034 |
| $\bar{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\bar{\nu}_\mu$ | 1.011 | | 0.985 | |

Table 40: Effect of the data/model reweighting on the efficiencies and efficiency ratios.
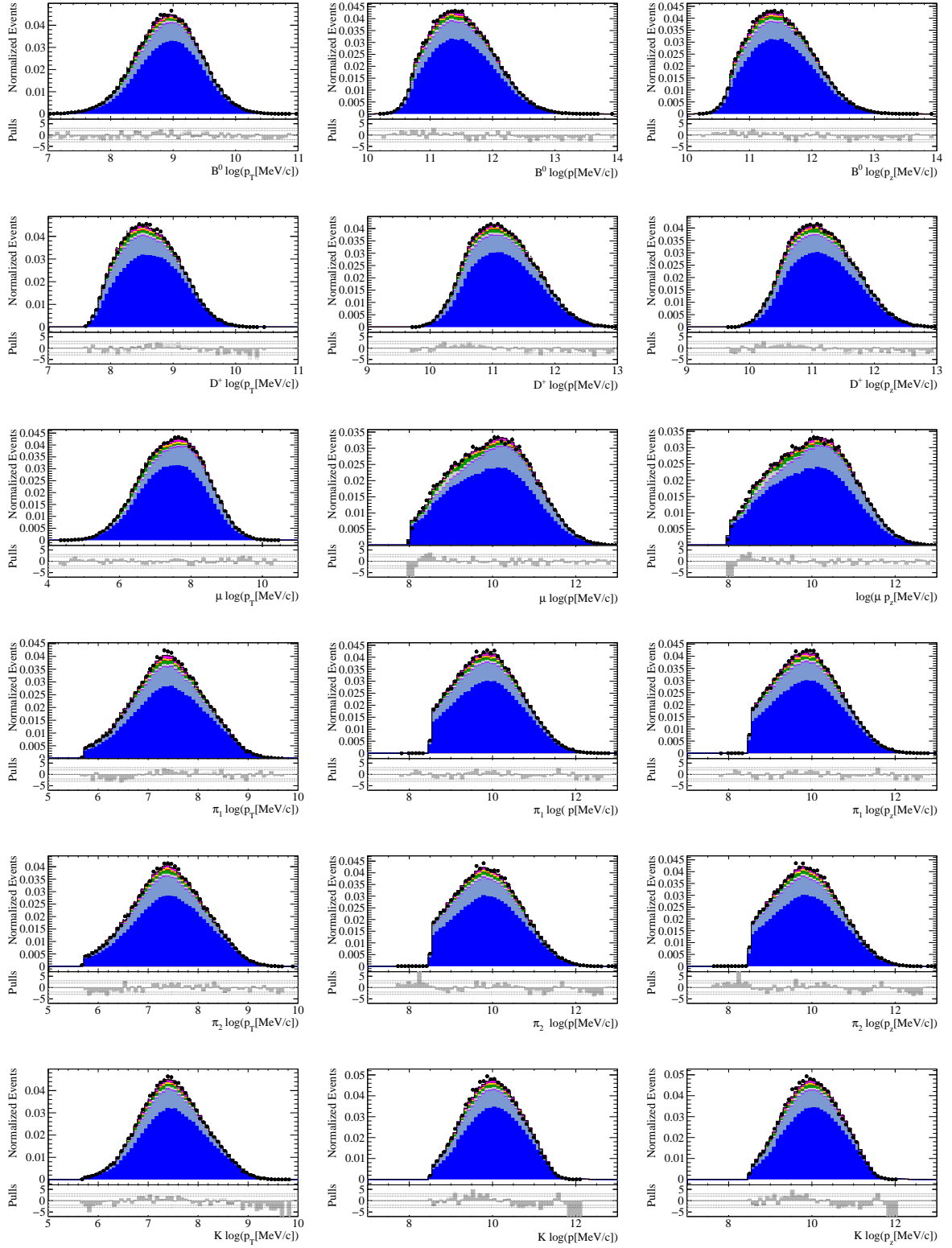
Figure 85: Comparison in the kinematic distributions of the normalization enriched region, between data and the fit model. The black dashed line represents the model before the correction. The pulls are reported for both the models, before (light grey) and after (dark grey) the correction.

152
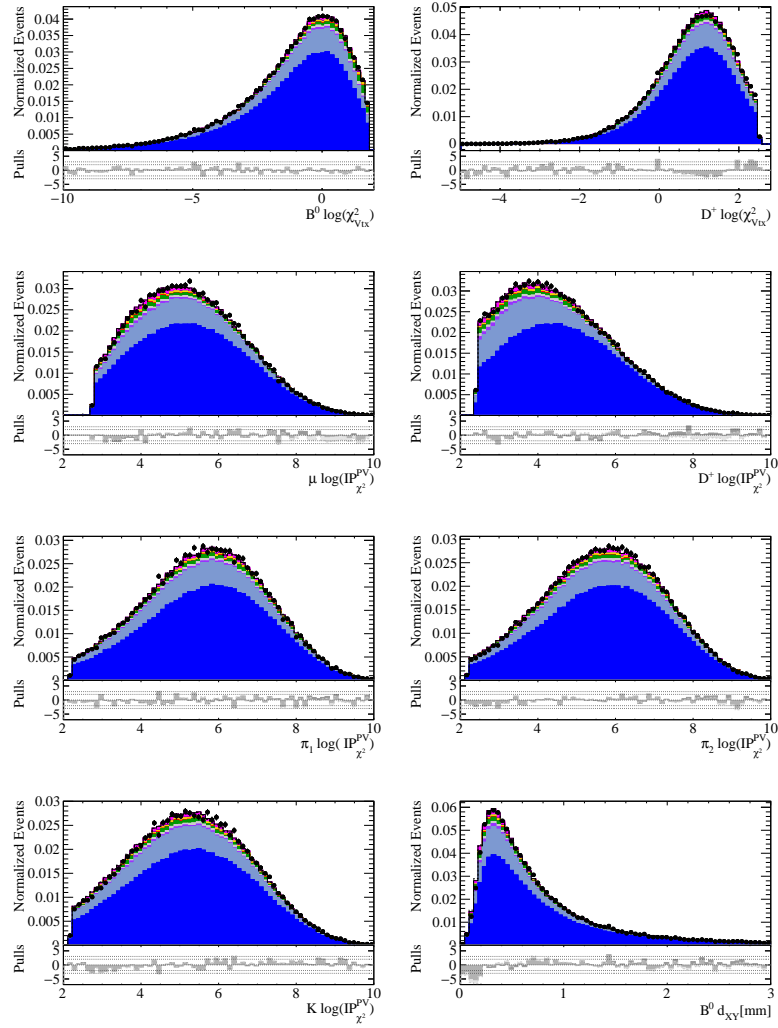
Figure 86: Comparison in the vertex $\chi^2$ of the $B^0$ and $D^+$, in the $IP_{\chi^2}$ of $D^+$ and signal tracks with respect to the primary vertex and in the transverse flight distance of the $B^0$, for the normalization enriched region, between data and the fit model. The black dashed line represents the model before the correction. The pulls are reported for both the models, before (light grey) and after (dark grey) the correction.
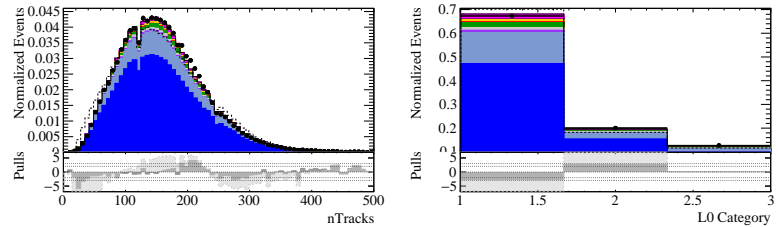


Figure 87: Comparison in nTracks and in the L0 category, for the normalization enriched region, between data and the fit model. The black dashed line represents the model before the correction. The pulls are reported for both the models, before (light grey) and after (dark grey) the correction.
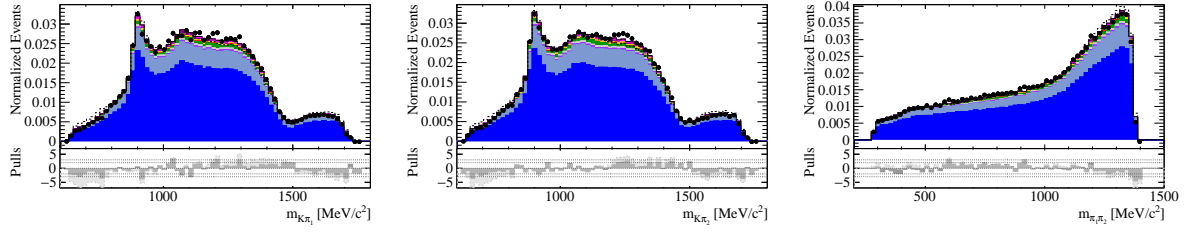
Figure 88: Comparison in variables related to the $D^+$ Dalitz, for the normalization enriched region, between data and the fit model. The black dashed line represents the model before the correction. The pulls are reported for both the models, before (light grey) and after (dark grey) the correction.
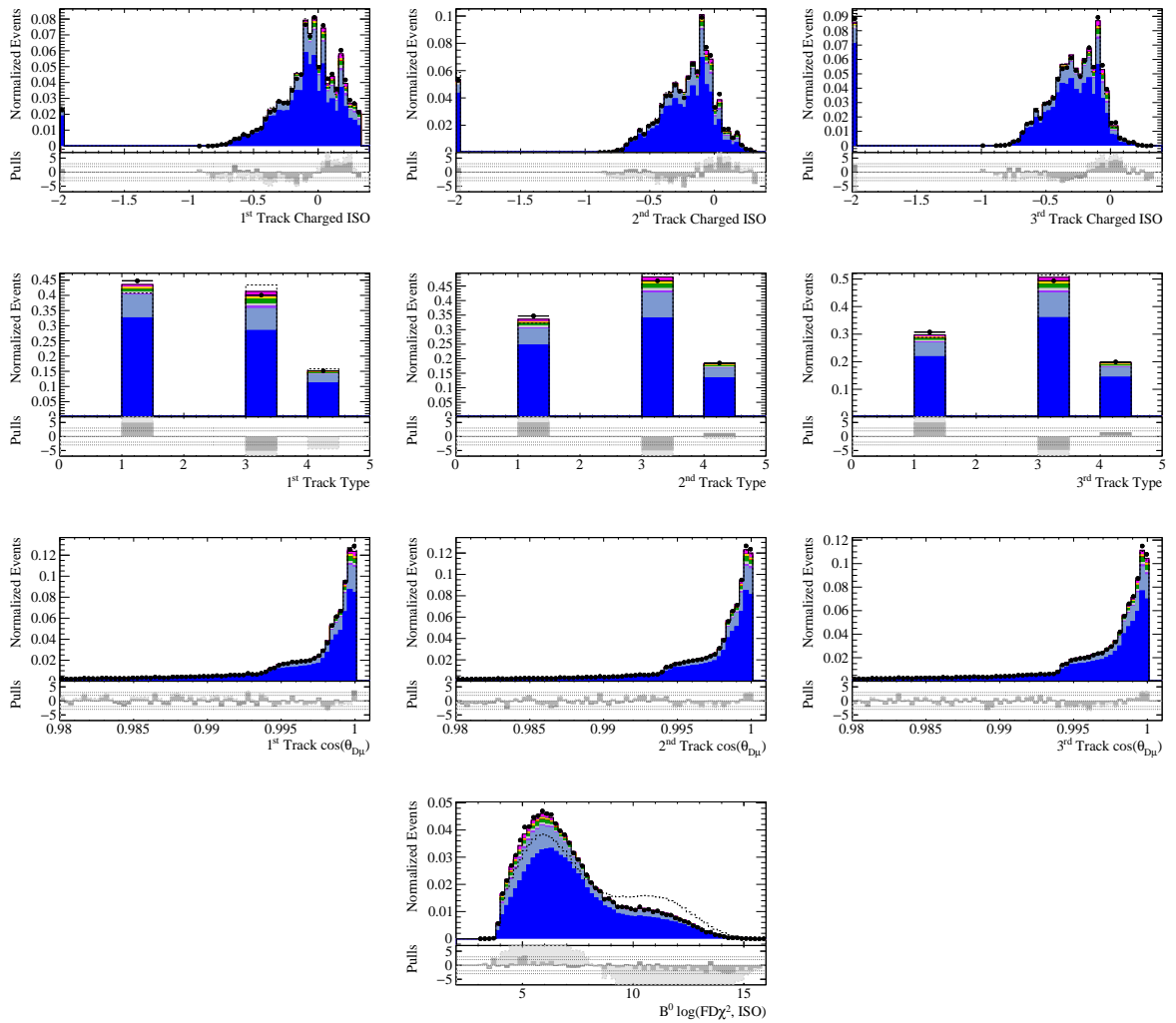


Figure 89: Comparison in the charged isolation output and some variables in input to the charged isolation BDT of the three most anti-isolated tracks in the event, for the normalization enriched region, between data and the fit model. The black dashed line represents the model before the correction. The pulls are reported for both the models, before (light grey) and after (dark grey) the correction.

154
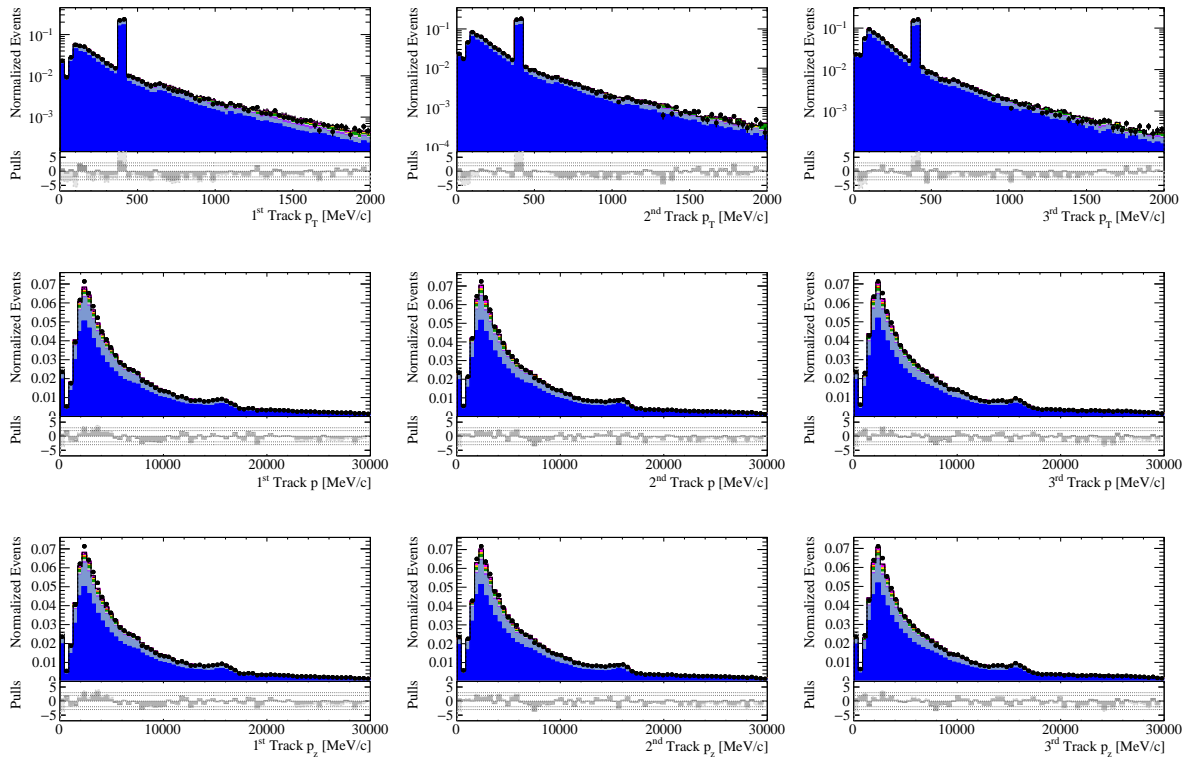
Figure 90: Comparison in the kinematics of the three most anti-isolated tracks in the event, for the normalization enriched region, between data and the fit model. The black dashed line represents the model before the correction. The pulls are reported for both the models, before (light grey) and after (dark grey) the correction.

(a) data



(b) model



(c) pulls before correction



(d) pulls after correction
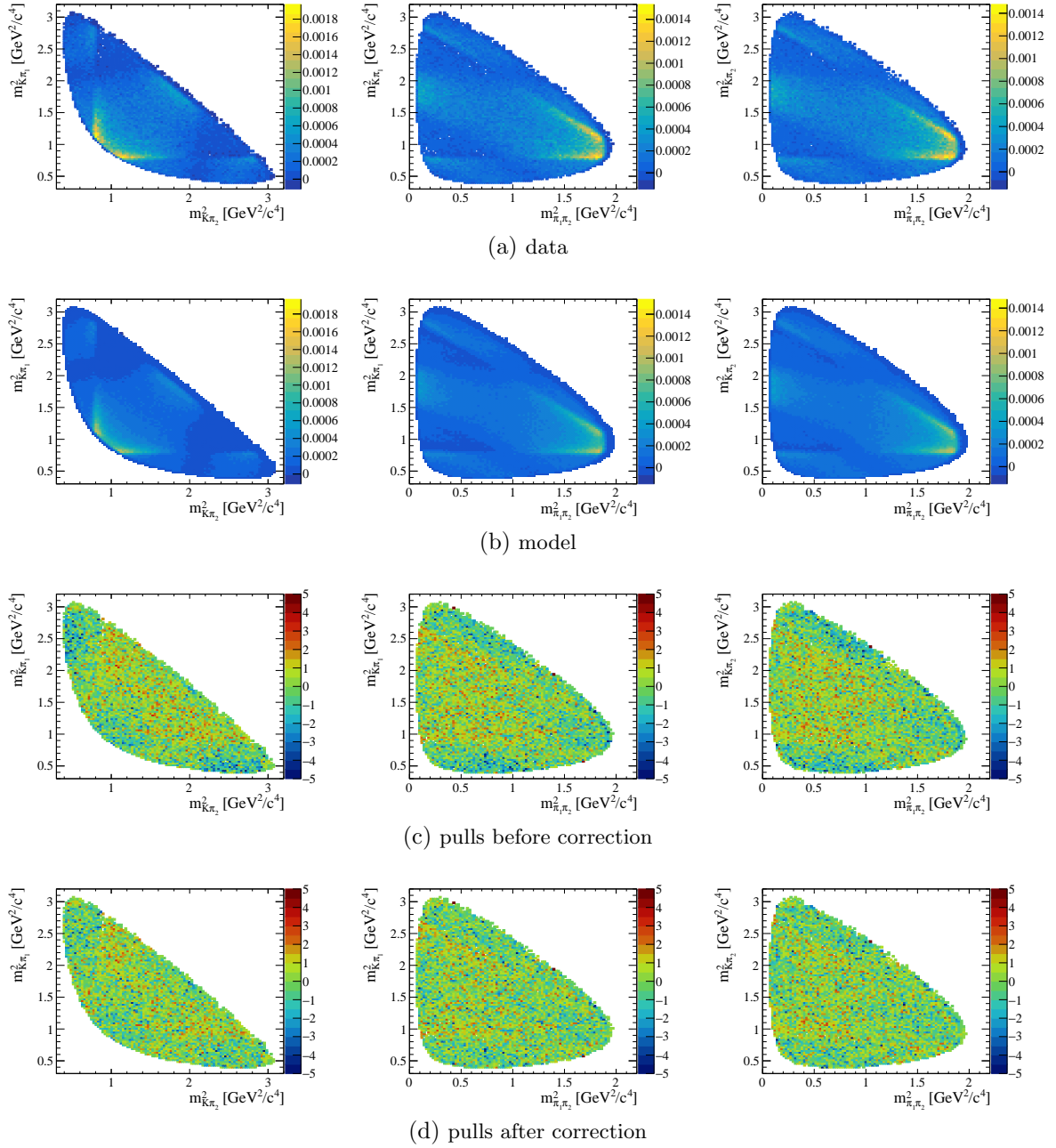
Figure 91: Dalitz plots of the $D^+$ decay, in data (a) and in the model (b). The pulls between the data and the model, before and after the corrections, are reported in (c) and (d), respectively.
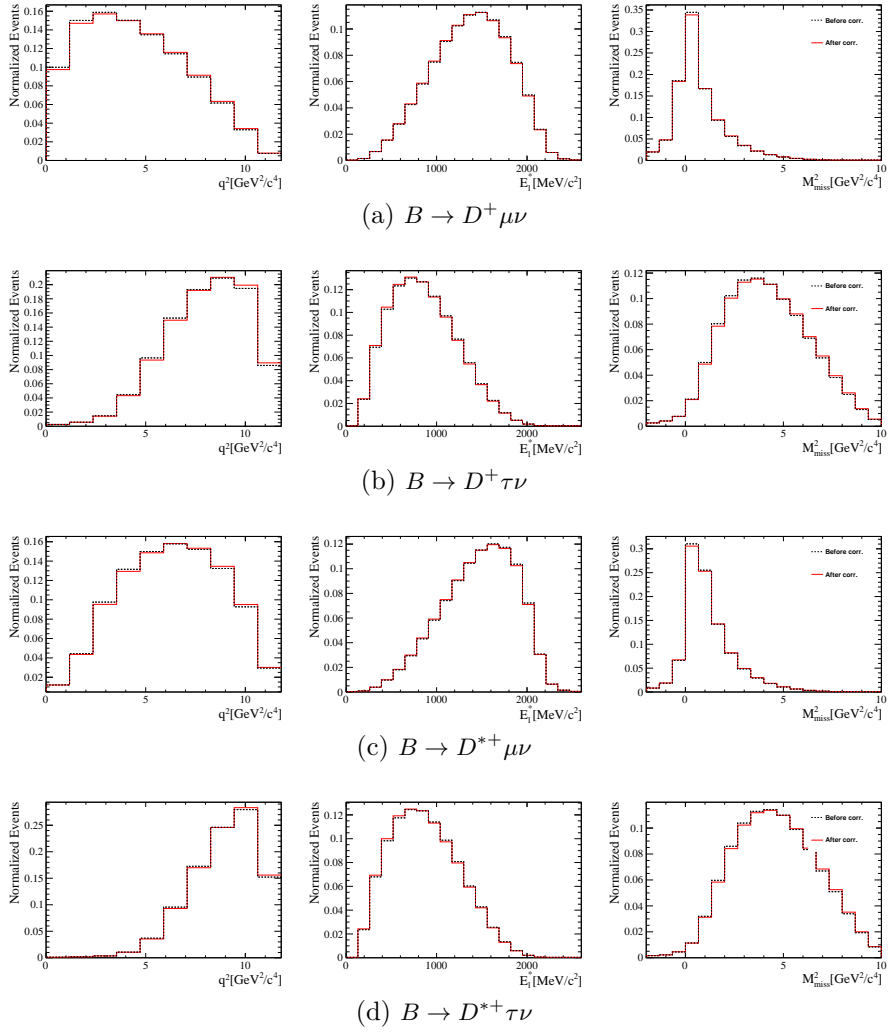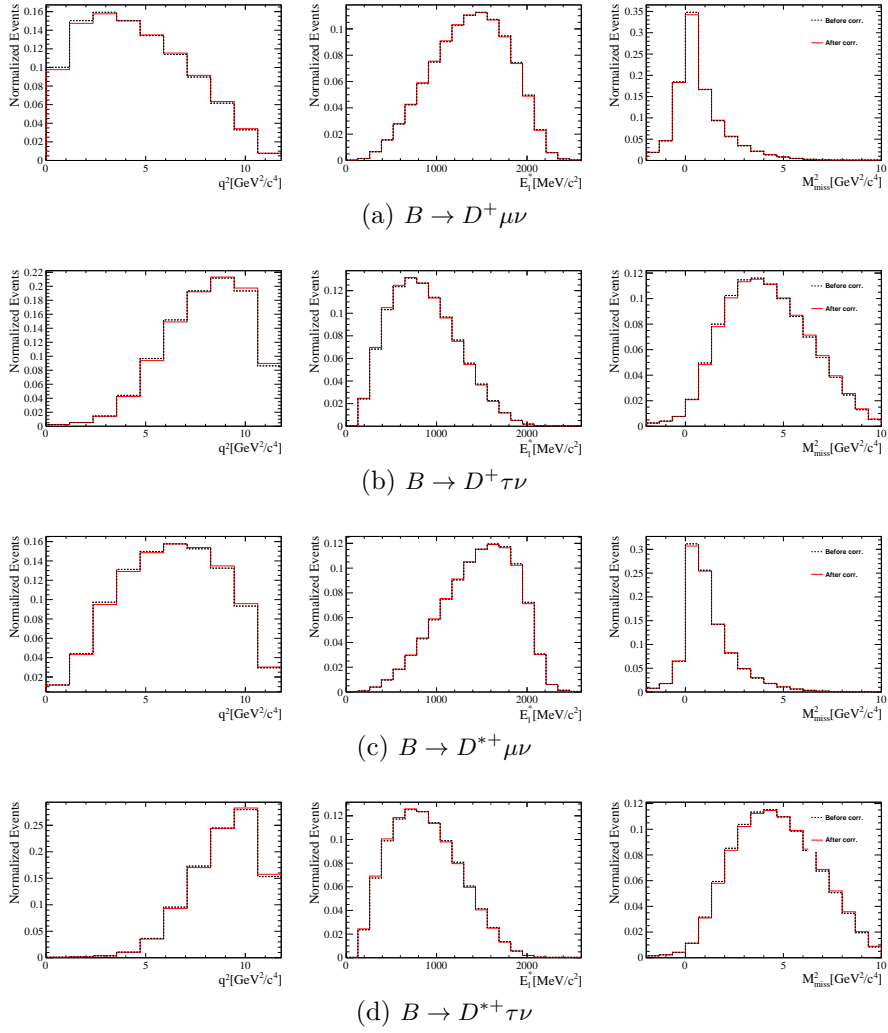
Figure 92: Effect of the data/model corrections on the two signal and normalization modes 2015 templates, ISO sample. The red histograms represent the templates after the correction, the black dashed histograms the ones before.

(a) $B \to D^+ \mu\nu$

(b) $B \to D^+ \tau\nu$

(c) $B \to D^{*+} \mu\nu$

(d) $B \to D^{*+} \tau\nu$

Figure 93: Effect of the data/model corrections on the two signal and normalization modes 2016 templates, ISO sample. The red histograms represent the templates after the correction, the black dashed histograms the ones before.

# 14 Fit model and nuisance parameters

In this chapter the fit model and the description of all the components used is reported, along with the description of all the nuisance parameters and the parameters of interest.

The fit is performed by means of the maximum likelihood method on binned data. The binning scheme used is the following:

- $q^2$: 10 equal-width bins, in the range $[0, 11.8]\,\mathrm{GeV}^2/c^4$;

- $E_\mu^*$: 20 equal-width bins, in the range $[0, 2.6]\,\mathrm{GeV}^2/c^4$;

- $m_{miss}^2$: 18 equal-width bins, in the range $[-2, 10]\,\mathrm{GeV}^2/c^4$.

The fit has been performed on the combined 2015+2016 sample. The data, MisID and Combinatorial samples have been added together without any correction, whereas the MC samples used to estimate the different background sources have been added together with a proportion that takes into account the integrated luminosity collected in the two years and the number of produced events in the two data taking conditions.

The fit has been performed simultaneously in the ISO, 1OS, $\pi^0$, 2OS and DD regions. The combined likelihood being minimized has been constructed with the help of the `HistFactory` framework.

The function being minimised is the sum of the following, constructed in each region

$$\log \mathcal{L} = \sum_i n_i \log \mu_i(\theta) - \mu_i(\theta), \tag{162}$$

where $i$ is an index running over the bins, $n_i$ is the number of observed events in the $i^{th}$ bin and $\mu_i(\theta)$ is the number of expected events in the $i^{th}$ bin. This number depends on $\theta$, a vector of nuisance parameters and parameters of interest. The number of events expected in each bin is given by the sum of expected events from each background or signal contribution, expressed as the bin content of a normalized histogram (template) and a yield parameter. The parametrization of each contribution will be detailed in the following sections. The number of observed events in each bin is given by the sum of the sWeights used to subtract the $D$-combinatorial background ($n_i = (\sum s\mathcal{W})$). Being the data weighted, the coverage properties of the maximum likelihood method are not expected to hold nicely, being the bin content distributed in a non Poisson way. For this reason the estimates of the parameters are known not have the correct coverage probability. This is corrected for by using the so-called SPD method [97], in which the likelihood is effectively modified to have the following form

$$\log \mathcal{L} = \sum_i n_i' \log \mu_i'(\theta) - \mu_i'(\theta), \tag{163}$$

where

$$n_i' = \quad n_i/s_i, \tag{164}$$

$$\mu_i' = \quad \mu_i/s_i, \tag{165}$$

$$s_i = \quad (\textstyle\sum s\mathcal{W}^2)_i/(\sum s\mathcal{W})_i. \tag{166}$$

A derivation of this correction is reported in Appendix A.

The templates are constructed using the Tracker Only MC, with all the correction weights and selections described previously in the text. For contributions in which the Form Factor modelling is either simply reweighted or it is being varied in the fit to include the related systematic uncertainty in the measurement, the template is evaluated using the RooHammerModel class described in Sect. 5.

External constraints on the nuisance parameters are inserted in the fit using multiplicative *p.d.f.* terms to the joint likelihood of the measurement.

## 14.1 Signal and normalization

Two signal and two normalization samples contribute to the measurement. These are made of $B \to D^+\ell\nu$ and $B \to (D^{*+} \to D^+\pi^0)\ell\nu$ decays, with $\ell = \mu, \tau$ for the normalization and signal modes, respectively.

The normalization mode is characterized by a narrow peak around $m^2_{miss} \approx 0$, due to the presence of only one neutrino in the final state, whereas the signal modes have a more continuous distribution in $m^2_{miss}$ and are characterized by a softer $E^*_\mu$ spectrum, being the muon coming from a decay of the $\tau$ with multiple neutrinos in the final state.

The shape of these contributions is taken from Tracker Only MC, produced with a CLN parametrization of the Form Factor parameters. The shape of the templates used in the fit is reported in Fig. 94.
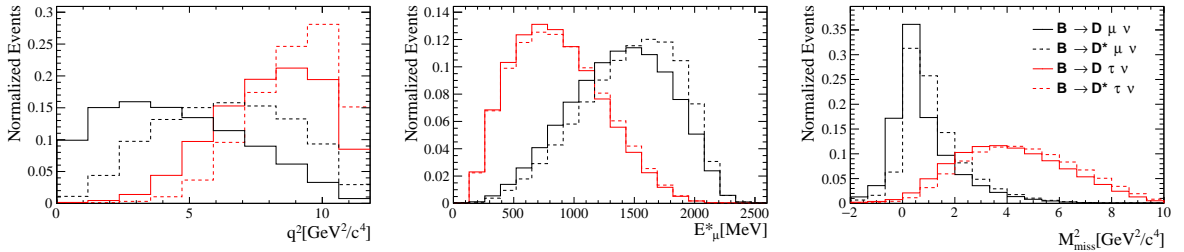


Figure 94: Projections of $B \to D^*\mu\nu$ template. The $B \to D^{(*)}\tau\nu$ template is overlaid with the same normalization in red.

They are inserted in the fit using RooHammerModel templates, and the Form Factor parameters are reweighted to the BGL parametrization, as reported in Sect.3.2.3. The $B \to D$ Form Factors are expanded up at the second order in BGL

$$f_+(z) = \frac{1}{P_{f_+}(z)\phi_{f_+}(z)} \sum_{n=0}^{2} a_{+n} z^n, \tag{167}$$

$$f_0(z) = \frac{1}{P_{f_0}(z)\phi_{f_0}(z)} \sum_{n=0}^{2} a_{0n} z^n. \tag{168}$$

and the $B \to D^*$ Form Factors are expanded at first order in BGL, except for $\mathcal{F}_1$ which is

expanded at third order:

$$g(z) = \frac{1}{P_g(z)\phi_g(z)} \sum_{n=0}^{1} a_n z^n, \tag{169}$$

$$f(z) = \frac{1}{P_f(z)\phi_f(z)} \sum_{n=0}^{1} b_n z^n, \tag{170}$$

$$\mathcal{F}_1(z) = \frac{1}{P_{\mathcal{F}_1}(z)\phi_{\mathcal{F}_1}(z)} \sum_{n=0}^{2} c_n z^n, \tag{171}$$

$$\mathcal{F}_2(z) = \frac{1}{P_{\mathcal{F}_2}(z)\phi_{\mathcal{F}_2}(z)} \sum_{n=0}^{1} d_n z^n. \tag{172}$$

The Form Factor parameters are varied in the fit, in order to include the related systematic uncertainty in the measurement. The $B \to D$ form factors are constrained in the fit, using the external measurement and their uncertainty as reported in [45], approximation $N = 2$. In this paper the parameters are measured by fitting the differential distribution of $B \to D\ell\nu$ experimental analyses from the Belle and BaBar collaborations and coming from FNAL/MILC and HPQCD lattice calculations. The $B \to D$ Form Factor parameters and their uncertainty is reported in Table 41.

| parameter | value | | error |
|:---:|:---:|:---:|:---:|
| $a_{+0}$ | 0.01566 | $\pm$ | 0.00011 |
| $a_{+1}$ | -0.0342 | $\pm$ | 0.0031 |
| $a_{+2}$ | -0.090 | $\pm$ | 0.022 |
| $a_{00}$ | 0.07935 | $\pm$ | 0.00058 |
| $a_{01}$ | -0.205 | $\pm$ | 0.014 |
| $a_{02}$ | -0.23 | $\pm$ | 0.10 |

Table 41: $B \to D$ BGL Form Factor parameters [45].

In this fit, the parameter $a_{00}$ is fixed to the other form factor parameters through the relation at maximum recoil

$$f_+(q^2 = 0) = f_0(q^2 = 0), \tag{173}$$

which gives the following formula using the measured Form Factor parameters

$$a_{00} = 4.99 \cdot a_{+0} + 0.32 \cdot a_{+1} + 0.021 \cdot a_{+2} - 0.065 \cdot a_{01} - 0.004 \cdot a_{02}. \tag{174}$$

The $B \to D^*$ Form Factor parameters are constrained using the external measurement and their uncertainty presented in [98], approximation $\text{BGL}_{222}$. In that paper, the parameters are

measured by fitting Belle unfolded data, fixing the parameter $c_0$ to the value of $b_0$ through the following formula

$$c_0 = \left( \frac{m_B - m_{D^*} \phi_{\mathcal{F}_1}(0)}{\phi_f(0)} \right) b_0, \tag{175}$$

which follows from the zero recoil relation between the $\mathcal{F}_1$ and $f$ Form Factors ($\mathcal{F}_1(z = 0) = (m_B - m_{D^*} f(z = 0))$). The fit in the paper have been performed in the massless lepton limit, in which the $\mathcal{P}_1$ form factor does not contribute, and therefore no value for the $d_i$ parameters has been provided. In our fit, we therefore make use of the central value of the $d_1$ and $d_2$ parameters and we constrain them with a 100% uncertainty, in an uncorrelated way with respect to all the other form factor parameters.

The $B \to D^*$ Form Factor parameters have been reported in Table 42.

| parameter | value | | error |
|:---:|:---:|:---:|:---:|
| $a_0$ | 0.000379 | $\pm$ | 0.000249 |
| $a_1$ | 0.026954 | $\pm$ | 0.009320 |
| $b_0$ | 0.000550 | $\pm$ | 0.000023 |
| $b_1$ | -0.002040 | $\pm$ | 0.001064 |
| $c_1$ | -0.000433 | $\pm$ | 0.000264 |
| $c_2$ | 0.005350 | $\pm$ | 0.004606 |
| $d_0$ | 0.002 | $\pm$ | 0.002 |
| $d_1$ | -0.013 | $\pm$ | 0.013 |

Table 42: $B \to D^*$ form factor parameters [98].

The values being actually fitted are differences with respect to the nominal Form Factor parameters value (denoted as $\Delta(p)$, where $p$ is the Form Factor parameter under study). The correlation matrices used in the fit are reported in Table 43 and Table 44.

| | $\Delta(a_{+0})$ | $\Delta(a_{+1})$ | $\Delta(a_{+2})$ | $\Delta(a_{01})$ | $\Delta(a_{02})$ |
|:---:|:---:|:---:|:---:|:---:|:---:|
| $\Delta(a_{+0})$ | 1 | 0.304 | -0.294 | 0.212 | 0.161 |
| $\Delta(a_{+1})$ | 0.304 | 1 | -0.422 | 0.747 | 0.190 |
| $\Delta(a_{+2})$ | -0.294 | -0.422 | 1 | -0.034 | 0.148 |
| $\Delta(a_{01})$ | 0.212 | 0.747 | -0.034 | 1 | -0.210 |
| $\Delta(a_{02})$ | 0.161 | 0.190 | 0.148 | -0.210 | 1 |

Table 43: Correlation matrix for the $B \to D$ BGL parameters reported in [45].

|            | $\Delta(a_0)$ | $\Delta(a_1)$ | $\Delta(b_0)$ | $\Delta(b_1)$ | $\Delta(c_1)$ | $\Delta(c_2)$ |
|------------|---------------|---------------|---------------|---------------|---------------|---------------|
| $\Delta(a_0)$ | 1        | -0.952 | -0.249 | 0.417  | 0.137  | -0.054 |
| $\Delta(a_1)$ | -0.952   | 1      | 0.383  | -0.543 | -0.268 | 0.165  |
| $\Delta(b_0)$ | -0.249   | 0.383  | 1      | -0.793 | -0.648 | 0.461  |
| $\Delta(b_1)$ | 0.417    | -0.543 | -0.793 | 1      | 0.542  | -0.333 |
| $\Delta(c_1)$ | 0.137    | -0.268 | -0.648 | 0.542  | 1      | -0.953 |
| $\Delta(c_2)$ | -0.054   | 0.165  | 0.461  | -0.333 | -0.953 | 1      |

Table 44: Correlation matrix for the $B \to D^*$ BGL parameters reported in [98].

When inserting multidimensional gaussian constraints defined by these correlation matrices, some instabilities have been observed in the fit. This can be due to the fact that these correlation matrices define very narrow and very correlated multidimensional constraint, leading to an unstable minimization in Minuit. To overcome this problem, the covariance matrices are firstly diagonalised. The parameters being floated in the fit are then linear combinations of the $\Delta$ parameters, corresponding to the eigenvectors of the covariance matrices. These parameters are denoted with the $\tilde{\Delta}$ throughout the rest of the thesis. In this basis, the multidimensional gaussian constraint reduces to a product of $N$ unidimensional gaussian constraints, with $N$ the number of Form Factor parameters being floated and variances given by the eigenvalues of the covariance matrices.

The yield parameters being fit are the yield of the $B \to D^+\mu\nu$ and of the $B \to D^*\mu\nu$ samples. The yield of the signal sample is related to the one of the normalization samples through the parameters $\mathcal{R}_{\mathrm{raw}}(D^+)$ and $\mathcal{R}_{raw}(D^*)$, that represent the ratio of yields between the signal and normalization sample. These parameters are essentially measured in the signal region. The yield is then converted to the yield in the control regions using Transfer Factors (TF), representing ratios of efficiencies between the control region and the signal region.

Since the amount of normalization events in the control regions is known to be small, the Transfer Factors are left floating in the fit to allow for differences in the isolation efficiency. No signal is expected in the 1OS, 2OS and DD region, and is therefore vetoed in there. To relate the measurements of the LFU ratios between the ISO and the $\pi^0$ regions in a way which does not rely much on the absolute neutral isolation efficiency, the double ratios of efficiencies

$$\frac{\varepsilon(B \to D^+\tau\nu, \pi^0)}{\varepsilon(B \to D^+\tau\nu, \mathrm{ISO})} \Big/ \frac{\varepsilon(B \to D^+\mu\nu, \pi^0)}{\varepsilon(B \to D^+\mu\nu, \mathrm{ISO})} \tag{176}$$

and

$$\frac{\varepsilon(B \to D^*\tau\nu, \pi^0)}{\varepsilon(B \to D^*\tau\nu, \mathrm{ISO})} \Big/ \frac{\varepsilon(B \to D^*\mu\nu, \pi^0)}{\varepsilon(B \to D^*\mu\nu, \mathrm{ISO})} \tag{177}$$

are set to be equal to their MC estimates. This amounts to expressing the TFs of the tauonic component in the $\pi^0$ sample as a linear function of the TFs of their muonic counterparts.

The model being used for the signal and normalization samples is reported in Table 45.

| Component | Shape | Normalization |
|---|---|---|
| $B^0 \to D^- \tau^+ \nu_\mu$ | MC + Hammer BGLVar | $N(D\mu) \times \mathrm{TF} \times \mathcal{R}_{\mathrm{raw}}(D^+)$ |
| $B^0 \to D^{*-} \tau^+ \nu_\mu$ | MC + Hammer BGLVar | $N(D^*\mu) \times \mathrm{TF} \times \mathcal{R}_{\mathrm{raw}}(D^*)$ |
| $B^0 \to D^- \mu^+ \nu_\mu$ | MC + Hammer BGLVar | $N(D\mu) \times \mathrm{TF}$ |
| $B^0 \to D^{*-} \mu^+ \nu_\mu$ | MC + Hammer BGLVar | $N(D^*\mu) \times \mathrm{TF}$ |

Table 45: Fit components and their normalizations. Parameters denoted in blue are shared between the fit regions.

## 14.2 Feed Down from 1P states, $B \to D^{**}\mu\nu$

Feed Down from the 1P states with prompt muons is included in eight separate templates, one for each state and each flavour of the $B$ mother.

The $D^{**}$ states decay primarily to $D^{(*)}\pi$, although the $D_1$ state has also been observed decaying into the $D\pi\pi$ final state. The shape of this contribution is similar to the normalization modes, with a peaking $m_{\mathrm{miss}}^2$ distribution around 0, although with a larger width and a slightly softer $E_\mu^*$ distribution, due to the presence of additional unrecostructed pions in the event. The templates projections in the signal region are reported in Fig. 95 and 96, for the $B^\pm$ and $B^0$ components respectively.

The templates for the $B^0$ modes are very similar to the templates of the $B^\pm$ modes, since the $D^{**}$ decays are very similar, the only difference being the charge of the final state particles. To avoid instabilities in the fit, the normalization of only one of the 8 templates (*i.e.* the one of the $B^\pm D_1\mu\nu$ mode) has been left freely floating in the fit in each region of the fit. The yields of all the other components are constrained to the one of this reference mode using the ratio of efficiencies and branching fractions, as summarised in Table 46.

| Component | Shape | Normalization |
|---|---|---|
| $B^\pm \to (D_1 \to D^- X)\mu^+\nu_\mu$ | MC + Hammer BLRVar | $N(D_1^0)$ |
| $B^\pm \to (D_0^* \to D^- X)\mu^+\nu_\mu$ | MC + Hammer BLRVar | $N(D_1^0) \times \varepsilon_{D_1^0}^{D_0^{*0}} \times \hat{\mathcal{B}}_{D_1^0}^{D_0^{*0}}$ |
| $B^\pm \to (D_1' \to D^- X)\mu^+\nu_\mu$ | MC +Hammer BLRVar | $N(D_1^0) \times \varepsilon_{D_1^0}^{D_1'^{0}} \times \hat{\mathcal{B}}_{D_1^0}^{D_1'^{0}}$ |
| $B^\pm \to (D_2^* \to D^- X)\mu^+\nu_\mu$ | MC + Hammer BLRVar | $N(D_1^0) \times \varepsilon_{D_1^0}^{D_2^{*0}} \times \hat{\mathcal{B}}_{D_1^0}^{D_2^{*0}}$ |
| $B^0 \to (D_1 \to D^- X)\mu^+\nu_\mu$ | MC + Hammer BLRVar | $N(D_1^0) \times \varepsilon_{D_1^0}^{D_1^{\pm}} \times \hat{\mathcal{B}}_{D_1^0}^{D_1^{\pm}}$ |
| $B^0 \to (D_0^* \to D^- X)\mu^+\nu_\mu$ | MC + Hammer BLRVar | $N(D_1^0) \times \varepsilon_{D_1^0}^{D_0^{*\pm}} \times \hat{\mathcal{B}}_{D_1^0}^{D_0^{*\pm}}$ |
| $B^0 \to (D_1' \to D^- X)\mu^+\nu_\mu$ | MC + Hammer BLRVar | $N(D_1^0) \times \varepsilon_{D_1^0}^{D_1'^{*\pm}} \times \hat{\mathcal{B}}_{D_1^0}^{D_1'^{*\pm}}$ |
| $B^0 \to (D_2^* \to D^- X)\mu^+\nu_\mu$ | MC + Hammer BLRVar | $N(D_1^0) \times \varepsilon_{D_1^0}^{D_2^{*\pm}} \times \hat{\mathcal{B}}_{D_1^0}^{D_2^{*\pm}}$ |

Table 46: Fit components and their normalizations. Parameter denoted in red are fixed in the fit. Parameters denoted in blue are shared between the fit regions. Parameters with a ˆ are constrained in the fit.

For each component X, the ratio of efficiencies measured in each category, $\varepsilon_{D_1^\pm}^X$, between component X and the reference sample, is defined as

$$\varepsilon_{D_1^\pm}^X = \frac{N_{\mathrm{MC}}^{\mathrm{RECO}}(X)}{N_{\mathrm{MC}}^{\mathrm{GEN}}(X)} \times \frac{N_{\mathrm{MC}}^{\mathrm{GEN}}(D_1^\pm)}{N_{\mathrm{MC}}^{\mathrm{RECO}}(D_1^\pm)}, \tag{178}$$

where $N_{\mathrm{MC}}^{\mathrm{RECO}}$ is the number of reconstructed and truth matched events selected in the control region under consideration and $N_{\mathrm{MC}}^{\mathrm{GEN}}$ is the number of generated events.

The number of generated events is evaluated by multiplying the number of events in input to the filtering selections for the $B^0$ and $B^\pm$ MC samples with the expected fractions of events generated in each state. This fraction is evaluated by looking at the branching fractions reported

in the EvtGen decfiles, and it is reported in Table 47. The efficiency ratio is fixed in the fit, and variations are considered for systematic uncertainties.

| Component | | fraction |
|---|---|---|
| $B^0 \to D^{**}\mu\nu$ | | |
| | $D_1^\pm$ | 33.3% |
| | $D_1'^\pm$ | 11.1% |
| | $D_0^\pm$ | 25.9% |
| | $D_2^\pm$ | 29.8% |
| $B^\pm \to D^{**}\mu\nu$ | | |
| | $D_1^0$ | 27.0% |
| | $D_1'^0$ | 9.0% |
| | $D_0^0$ | 34.1% |
| | $D_2^0$ | 29.9% |

Table 47: Fractions of generated events for each $D^{**}$ state, before any selection.

For each component X, the ratio of branching fractions $\mathcal{B}_{D_1^\pm}^X$, relative to the reference channel is given by

$$\mathcal{B}_{D_1^\pm}^X = \frac{\mathcal{B}(B \to (X \to D^\pm X')\mu\nu)}{\mathcal{B}(B \to (D_1^\pm \to D^\pm X')\mu\nu)}. \tag{179}$$

During the evaluation of the branching fractions, both the $B \to D^{**} \to D^* \to D^+$ and the $B \to D^{**} \to D^+$ decay chains are accounted for, using $\mathcal{B}(D^{*0} \to D^+X) = 0$ and $\mathcal{B}(D^{*+} \to D^+X) = 0.323 \pm 0.005$ [88]. The measured branching fractions at the time of the writing of this thesis are of the type $\mathcal{B}(B \to D^{**}\mu\nu) \times \mathcal{B}(D^{**} \to D^{(*)}\pi)$, with charged slow pions in the decay. To generalize to all the possible charge configurations for the final state, the following relation is used, coming from the Clebsch-Gordan coefficient of the two-body, isospin conserving decays:

$$\frac{\mathcal{B}(D^{**\pm} \to D^{(*)0}\pi^\pm)}{\mathcal{B}(D^{**\pm} \to D^{(*)}\pi)} = \frac{\mathcal{B}(D^{**0} \to D^{(*)\pm}\pi^\pm)}{\mathcal{B}(D^{**0} \to D^{(*)}\pi)} = \frac{2}{3}. \tag{180}$$

In [88] only branching fractions for decays to the $D^*\pi$ final state for the $D_1$ and $D_1'$ states, and to the $D\pi$ final state for the $D_0^*$ are reported. The $D_2^*$ is reported decaying into both final states. Therefore the following saturation relations are used in evaluating the branching fraction ratios:

$$\mathcal{B}(D_1 \to D^*\pi) = 1, \tag{181}$$

$$\mathcal{B}(D_1' \to D^*\pi) = 1, \tag{182}$$

$$\mathcal{B}(D_0^* \to D\pi) = 1, \tag{183}$$

$$\mathcal{B}(D_2^* \to D^*\pi) + \mathcal{B}(D_2^* \to D\pi) = 1. \tag{184}$$

As mentioned before, though, the $D_1$ state has also been observed decaying into the $D\pi\pi$ state, effectively contributing to the $B \to D^{**} \to D^+$ decay chain. In order to account for this, the error on the branching fraction for the decay involving $D_1$ states is increased by a factor of 10%.

With these inputs, the branching fraction for each decay and its uncertainty is evaluated and reported in Table 48.

| Mode | $\mathcal{B}_{\to D^* \to D^+}(10^{-3})$ | $\mathcal{B}_{\to D^+}(10^{-3})$ | $\sigma/\mathcal{B}$ |
|---|---|---|---|
| $\bar{B}^0 \to D_1^+\mu^-\nu$ | 0.45 | 0.0 | 11.1% |
| $\bar{B}^0 \to D_2^{*+}\mu^-\nu$ | 0.11 | 0.61 | 23.2% |
| $\bar{B}^0 \to D_1'^+\mu^-\nu$ | 0.5 | 0.0 | 29.1% |
| $\bar{B}^0 \to D_0^{*+}\mu^-\nu$ | 0.0 | 1.5 | 40.0% |
| $B^- \to D_1^0\mu^-\nu$ | 0.98 | 0.0 | 7.5% |
| $B^- \to D_2^{*0}\mu^-\nu$ | 0.33 | 1.53 | 9.6% |
| $B^- \to D_1'^0\mu^-\nu$ | 0.87 | 0.0 | 22.3% |
| $B^- \to D_0^{*0}\mu^-\nu$ | 0.0 | 2.5 | 20.0% |

Table 48: Branching ratios and their uncertainties used in determining the nominal $D^{**}$ branching fractions used in the fit.

The branching fraction ratios are then constrained in the fit using a multidimensional gaussian, taking also into account the correlation between them, defined in Table 49.

The events for these decays have been produced with Form Factors taken from the ISGW2 quark model [99], which employs Heavy Quark Symmetry to express all the Form Factors as a function of a universal Isgur Wise function. The templates are then reweighted in order to have Form Factors parametrized using the results obtained in [100]. The four $D^{**}$ states can be grouped in two heavy quark spin symmetry doublets: $D^{1/2^+} = \{D_0^*, D_1^*\}$ and $D^{3/2^+} = \{D_1, D_2^*\}$. The $B^0$ and the $B^\pm$ decays into charm hadrons of the same doublet share the same form factor parameters. In the fit three Form Factor parameters for the $D^{1/2^+}$ doublet, $\zeta(1)$, $\zeta'$ and $\hat{\zeta}(1)$, and four Form Factor parameters for the $D^{3/2^+}$ doublet, $\tau(1)$, $\tau'$, $\hat{\tau}_1$ and $\hat{\tau}_2$ are floated with a constraint coming from [100].

| | $\mathcal{B}_{D_1^0}^{D_2^+}$ | $\mathcal{B}_{D_1^0}^{D_1'^+}$ | $\mathcal{B}_{D_1^0}^{D_0^+}$ | $\mathcal{B}_{D_1^0}^{D_1^+}$ | $\mathcal{B}_{D_1^0}^{D_1'^0}$ | $\mathcal{B}_{D_1^0}^{D_2^0}$ | $\mathcal{B}_{D_1^0}^{D_0^0}$ |
|---|---|---|---|---|---|---|---|
| $\mathcal{B}_{D_1^0}^{D_2^+}$ | 1.0 | 0.06 | 0.05 | 0.15 | 0.08 | 0.16 | 0.09 |
| $\mathcal{B}_{D_1^0}^{D_1'^+}$ | 0.06 | 1.0 | 0.04 | 0.12 | 0.06 | 0.13 | 0.07 |
| $\mathcal{B}_{D_1^0}^{D_0^+}$ | 0.05 | 0.04 | 1.0 | 0.09 | 0.05 | 0.1 | 0.05 |
| $\mathcal{B}_{D_1^0}^{D_1^+}$ | 0.15 | 0.12 | 0.09 | 1.0 | 0.16 | 0.31 | 0.17 |
| $\mathcal{B}_{D_1^0}^{D_1'^0}$ | 0.08 | 0.06 | 0.05 | 0.16 | 1.0 | 0.16 | 0.09 |
| $\mathcal{B}_{D_1^0}^{D_2^0}$ | 0.16 | 0.13 | 0.1 | 0.31 | 0.16 | 1.0 | 0.18 |
| $\mathcal{B}_{D_1^0}^{D_0^0}$ | 0.09 | 0.07 | 0.05 | 0.17 | 0.09 | 0.18 | 1.0 |

Table 49: Correlation matrix for the $B \to D^{**}\mu\nu$ Branching fraction ratios used in the fit.
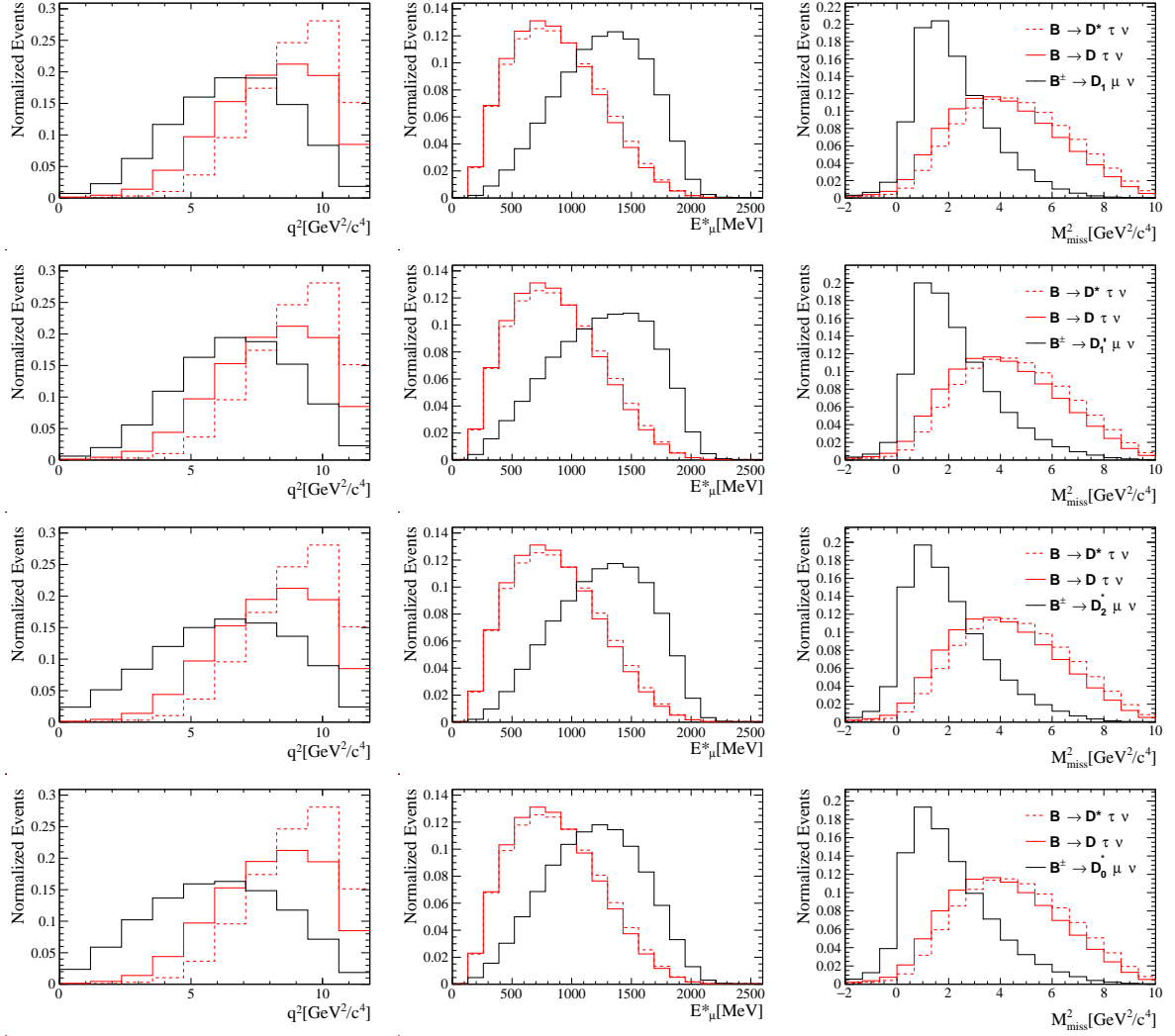
Figure 95: Projections of the $B^{\pm} \rightarrow D_1^0 \mu^+ \nu_{\mu}$ (first row), $B^{\pm} \rightarrow D_1^{'0} \mu^+ \nu_{\mu}$ (second row), $B^{\pm} \rightarrow D_2^{*0} \mu^+ \nu_{\mu}$ (third row) and $B^{\pm} \rightarrow D_0^{*0} \mu^+ \nu_{\mu}$ (fourth row) templates. $B \rightarrow D^{(*)} \tau \nu$ template is overlaid with the same normalization in red
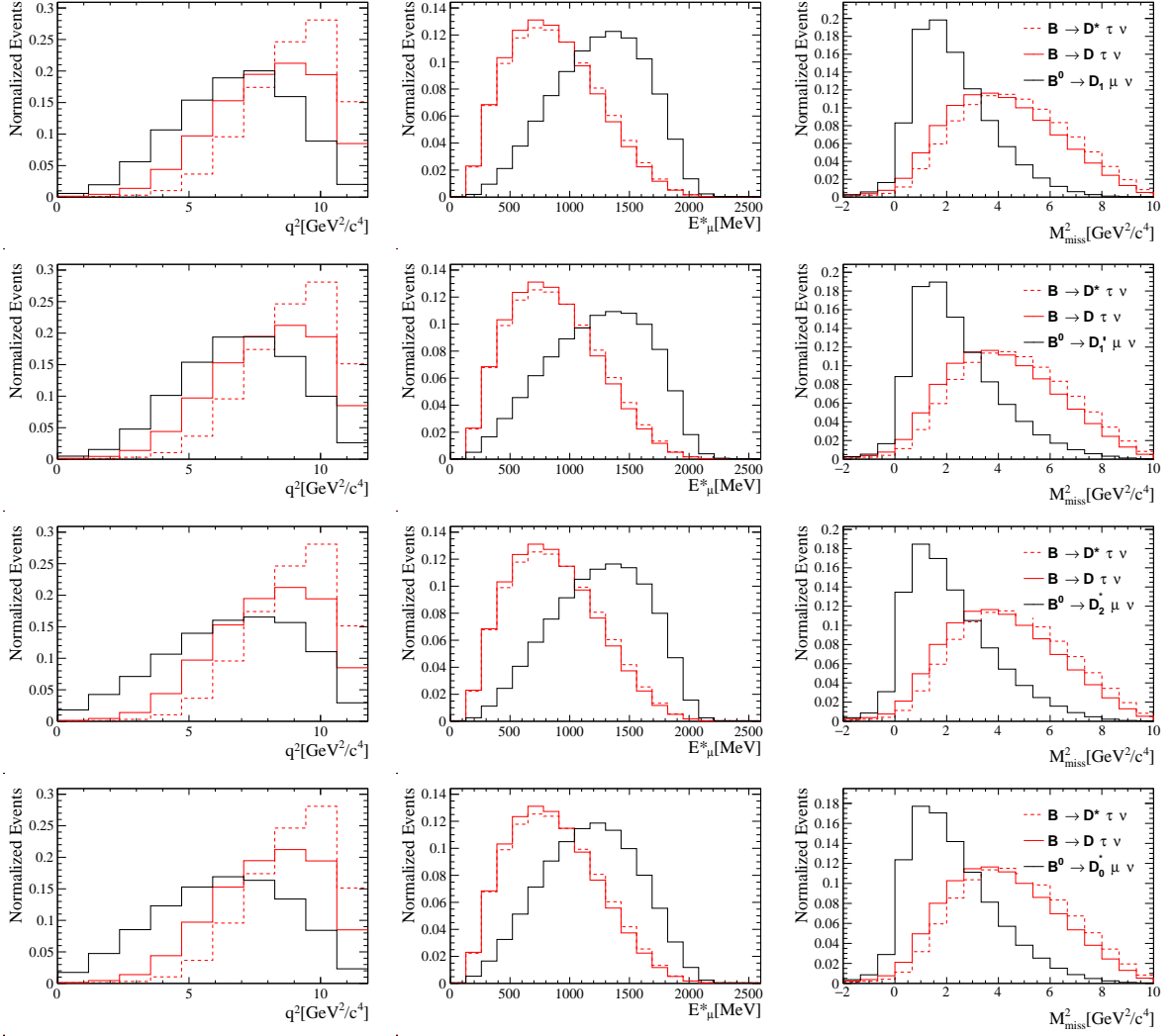
Figure 96: Projections of the $B^0 \to D_1^- \mu^+ \nu_\mu$ (first row), $B^0 \to D_1'^- \mu^+ \nu_\mu$ (second row), $B^0 \to D_2^{*-} \mu^+ \nu_\mu$ (third row) and $B^0 \to D_0^{*-} \mu^+ \nu_\mu$ (fourth row) templates. $B \to D^{(*)} \tau \nu$ template is overlaid with the same normalization in red

## 14.3   Feed Down from 1P states, $B \to D^{**}\tau\nu$

Events in which the same decays as the ones outlined in the previous section, but with a $\tau$ lepton in the final state instead of a $\mu$, may contribute as a background source when the decay $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$ takes place. This contribution is expected to be small and therefore some approximation in their parametrization is employed.

The four 1P states are combined, using the fractions expected from the MC simulation used. The projection of the templates along the fit variables for this background, when combining the four $D^{**}$ modes, is reported in Fig. 97.
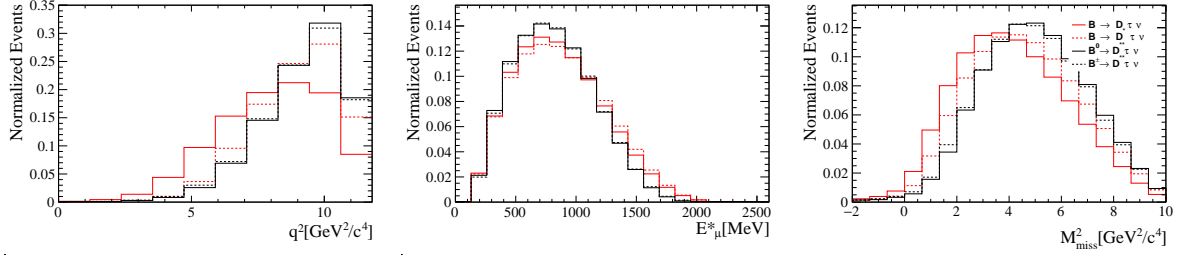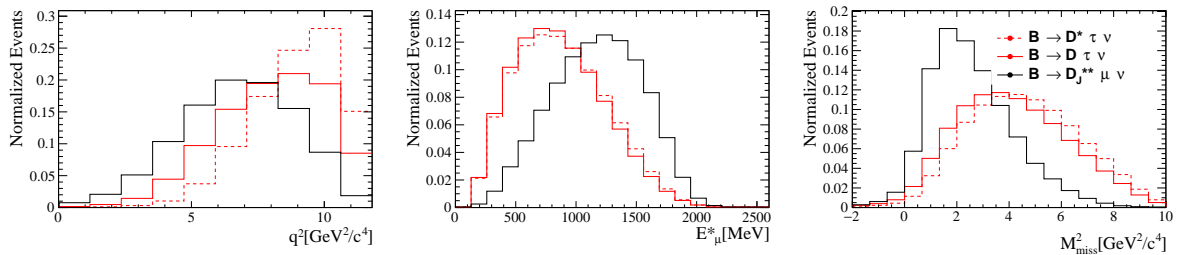


Figure 97: Projections of $B \to D^{**}\tau\nu$ templates. The $B \to D^{(*)}\tau\nu$ template is overlaid with the same normalization in red.

The yield of this component is constrained in the fit relative to the combination of the yields of the muonic counterpart, as shown in Table 50. The constraint takes into account the ratio of efficiencies, evaluated as

$$\varepsilon_{B^{0,\pm}D^{**}\mu}^{B^{0,\pm}D^{**}\tau} = \frac{N_{\text{MC}}^{\text{RECO}}(B^{0,\pm} \to D^{**}\tau^+\nu_\tau)}{N_{\text{MC}}^{\text{GEN}}(B^{0,\pm} \to D^{**}\tau^+\nu_\tau)} \times \frac{N_{\text{MC}}^{\text{GEN}}(B^{0,\pm} \to D^{**}\mu^+\nu_\mu)}{N_{\text{MC}}^{\text{RECO}}(B^{0,\pm} \to D^{**}\mu^+\nu_\mu)}, \qquad (185)$$

and the branching fraction ratio between the tauonic and the muonic mode. This is assumed to be common for the $B^0$ and $B^\pm$ modes, and it is expressed as follows:

$$\mathcal{B}_{D^{**}\mu}^{D^{**}\tau} = \frac{\mathcal{B}(B \to D^{**}\tau^+\nu_\tau)}{\mathcal{B}(B \to D^{**}\mu^+\nu_\mu)} \times \mathcal{B}(\tau(\to \mu\bar{\nu}_\mu\nu_\tau)) \equiv \mathcal{R}(D^{**}) \times \mathcal{B}(\tau(\to \mu\bar{\nu}_\mu\nu_\tau)), \qquad (186)$$

where the SM expectation for the $\mathcal{R}(D^{**})$ parameter, coming from [101], has been used. The value of $\mathcal{B}_{D^{**}\mu}^{D^{**}\tau}$ is left floating in the fit with a 50% variation to cover the knowledge of all the factors with a generous systematic uncertainty.

| Component | Shape | Normalization |
|---|---|---|
| $B^{\pm} \to (D^{**} \to D^- X)\mu^+ \nu_\mu$ | MC | $N(D_1^0) \times (1 + \varepsilon_{D_1^0}^{D_0^{*0}} \times \hat{\mathcal{B}}_{D_1^0}^{D_0^{*0}} + $ <br> $\varepsilon_{D_1^0}^{D_1'^{0}} \times \hat{\mathcal{B}}_{D_1^0}^{D_1'^{0}} + \varepsilon_{D_1^0}^{D_2^{*0}} \times \hat{\mathcal{B}}_{D_1^0}^{D_2^{*0}})$ <br> $\times \textcolor{red}{\varepsilon_{B^{\pm}D^{**}\mu}^{B^{\pm}D^{**}\tau}} \times \textcolor{blue}{\hat{\mathcal{B}}_{D^{**}\mu}^{D^{**}\tau}}$ |
| $B^0 \to (D^{**} \to D^- X)\mu^+ \nu_\mu$ | MC | $N(D_1^0) \times (\varepsilon_{D_1^0}^{D_1^{\pm}} \times \hat{\mathcal{B}}_{D_1^0}^{D_1^{\pm}} + \varepsilon_{D_1^0}^{D_0^{*\pm}} \times \hat{\mathcal{B}}_{D_1^0}^{D_0^{*\pm}} + $ <br> $\varepsilon_{D_1^0}^{D_1'^{*\pm}} \times \hat{\mathcal{B}}_{D_1^0}^{D_1'^{*\pm}} + \varepsilon_{D_1^0}^{D_2^{*\pm}} \times \hat{\mathcal{B}}_{D_1^0}^{D_2^{*\pm}})$ <br> $\times \textcolor{red}{\varepsilon_{B^0 D^{**}\mu}^{B^0 D^{**}\tau}} \times \textcolor{blue}{\hat{\mathcal{B}}_{D^{**}\mu}^{D^{**}\tau}}$ |

Table 50: Fit components and their normalizations. The parameters denoted in red are fixed in the fit. Parameters with a ˆ are constrained. Parameters denoted in blue are shared between the fit regions.

## 14.4 Feed Down from higher mass states

This contribution is constituted by semileptonic decays of the type $B \to D_J^{**} \mu\nu$, where $D_J$ denotes higher mass excitations of the $D$ meson. These decays have been modelled using a dedicated cocktail of events $B \to (D_J^{**} \to D\pi\pi X)\mu\nu$ events, decayed with a pure Phase-Space model. These decays are also known to be poorly measured. For this reason, their normalization is left floating in each region independently from the other Feed Down contributions and in order to insert a systematic uncertainty due to the knowledge of their shape, a phenomenological shape variation is implemented in the fit to calibrate the shape of the templates of this contribution directly in data.

The shape of the template for this contribution, compared to the one of the signal decays, is reported in Fig. 98.



Figure 98: Projections of $B \to D_J^{**} \mu\nu$ templates. The $B \to D^{(*)}\tau\nu$ template is overlaid with the same normalization in red. The templates are drawn from 2015 tracker only MC, before any shape reweight.

The shape variation is implemented as follows. A multiplicative deformation depending on the true $(p_\mu + p_\nu)^2$ in the decay is applied as a per-event weight:

$$ w(\alpha^{D_J}) = 1 + 2\alpha^{D_J} \left( \frac{(p_\mu + p_\nu)^2 - m_\mu^2}{8\,\text{GeV}^2/c^4} - \frac{1}{2} \right). \tag{187} $$

The $8\,\text{GeV}^2/c^4$ value has been chosen to be approximately equal to the maximum value of the numerator, so that the value in the parenthesis is approximately bound to stay within the range of $[-0.5, 0.5]$. The weights depend on a nuisance parameter $\alpha^{D_J}$, which is fitted from data. The weight values defined for the values of $\alpha^{D_J} = \pm 1$ are used to define alternative $\pm 1\sigma$ templates, between which the fit interpolates, choosing the best estimation for $\alpha^{D_J}$ from data, as explained in Appendix B. The parameter $\alpha^{D_J}$ is able to float freely without any constraint in the fit, changing the shape of the template in all the regions simultaneously. The alternative templates are normalized to the same area as the nominal template, so that the value of $\alpha^{D_J}$ only affects the shape of the templates.

In Fig. 99 the shape of the nominal template, along with the one of the $\pm 1\sigma$ alternative ones, in all the control region of the fit, is reported.
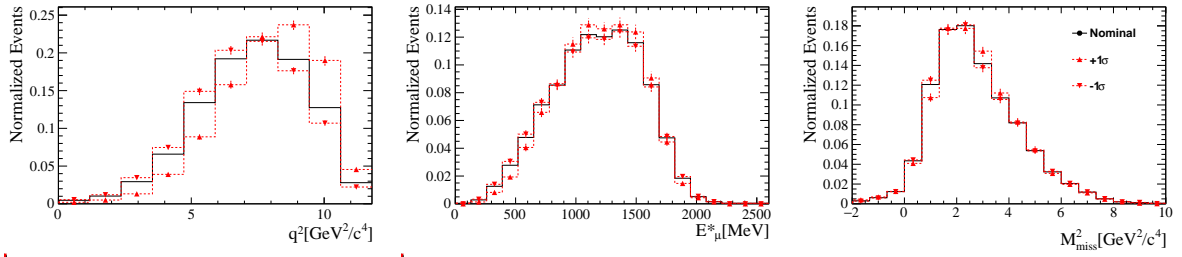
Table 51 summarises the model used in the fit.

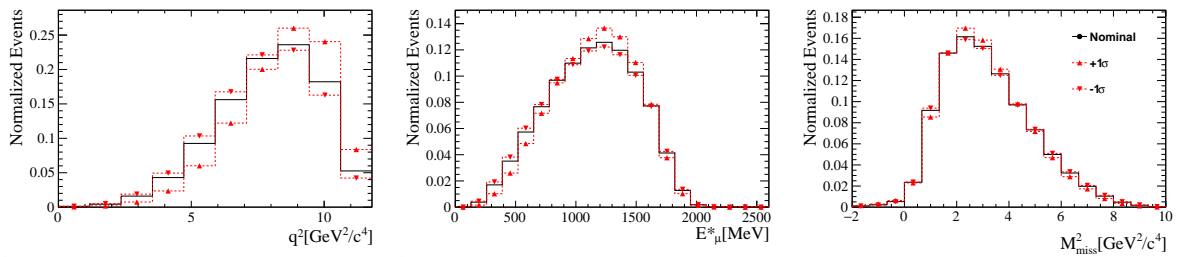| Component | Shape | Normalization |
|-----------|-------|---------------|
| $B^0 \to (D_J^{**} \to D^- X)\mu^+ \nu_\mu$ | MC + Shape Var. | $N(D_J^{**})$ |

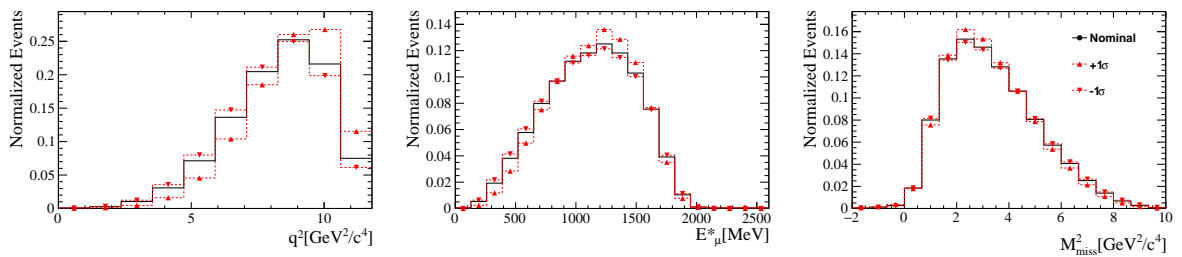Table 51: Fit components and their normalizations. Parameters denoted in blue are shared between the fit regions.
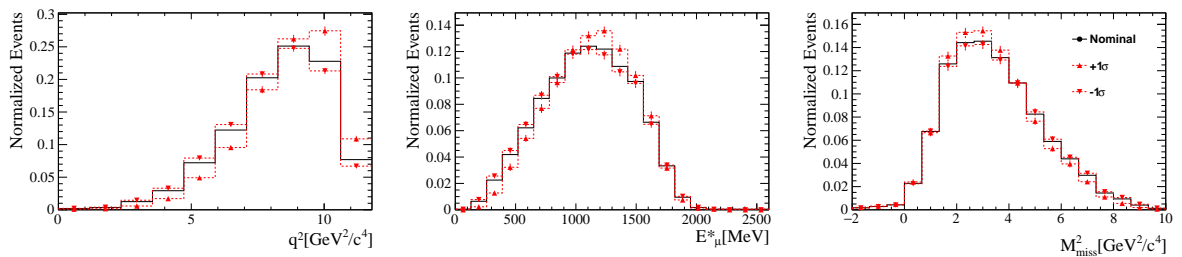
(a) ISO



(b) $\pi^0$
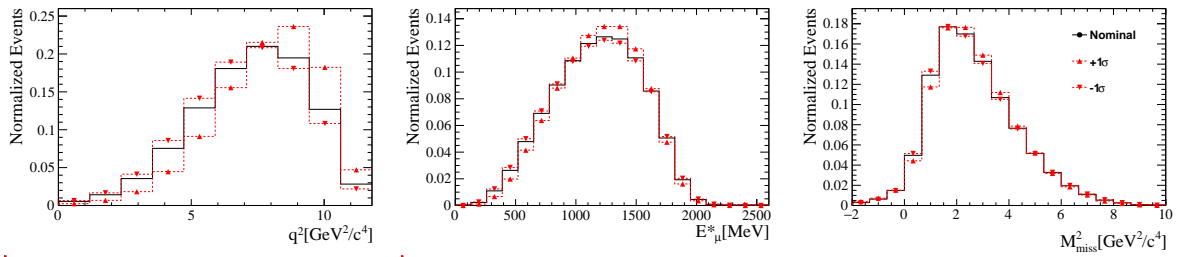


(c) 1OS



(d) 2OS



(e) DD

Figure 99: Projection of the nominal $B \to D_J^{**}\mu\nu$ template along with the two histograms used in the `HistFactory` shape variation.

175

## 14.5 Feed Down from $B_s \to D_s^{**} \mu\nu$

The Feed Down coming from $B_s$ decays is included in the fit by means of a single template independent from all the other Feed Down contributions. This component is expected to be very small. Therefore its normalization is controlled by the yield in the signal region, and the yield in all the other regions is taken from this one using ratio of efficiencies taken from MC.

The projection of the template for this component along the fit variables in the signal region is reported in Fig. 100.
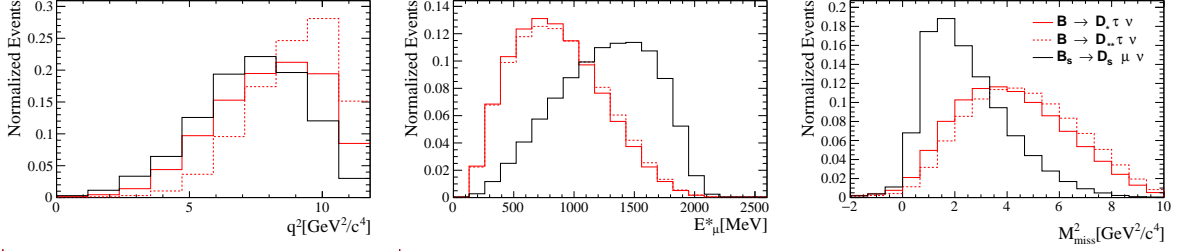


Figure 100: Projections of the $\bar{B}_s \to D_s^{**} \mu\nu$ template. The $B \to D^{(*)}\tau\nu$ template is overlaid with the same normalization in red.

## 14.6 Double Charm, $B \to D(H_c \to \mu\nu X')X$

This background takes into account $b \to c\bar{c}q$ transitions, in which one of the charm particles in the final state decays into a muon. This is taken into account in the fit by means of four templates in each control region, according to the charge of the $B$ mother ($B^0$ and $B^\pm$) and the type of decay (Two Body, $B \to D(H_c \to \mu\nu X)$ or Multi Body, $B \to D(H_c \to \mu\nu X)X'$)

The shape of the Two Body and the Multi Body templates as a function of the fit variables is shown in Fig. 101 and Fig. 102, respectively.
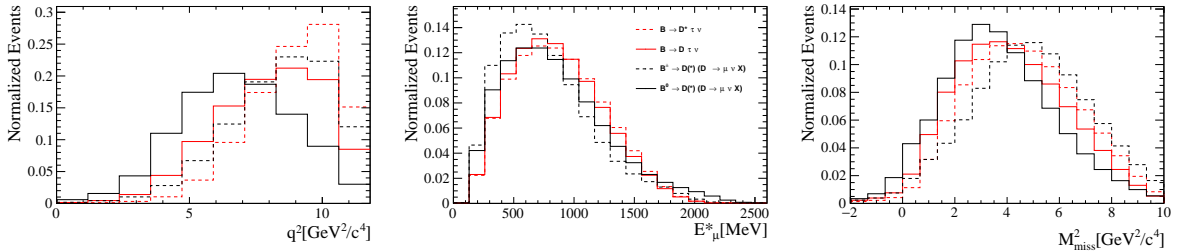


Figure 101: Projections of the $B \to D(X_c \to \mu\nu X)$ template. The $B \to D^{(*)}\tau\nu$ templates are overlaid with the same normalization in red. The templates are drawn from 2015 tracker only MC.

The Multi Body decays are known to be poorly measured. Therefore, in order not to rely much on the MC simulation for the shape of the templates, the shape of their template is interpolated in a data driven way, applying a phenomenological correction which is controlled by parameters that are directly measured from data.

The shape variation is implemented as follows: events with Multi Body decays of the type $B \to D_1 D_2 X$ are reweighted using two multiplicative deformations depending on the kinematics

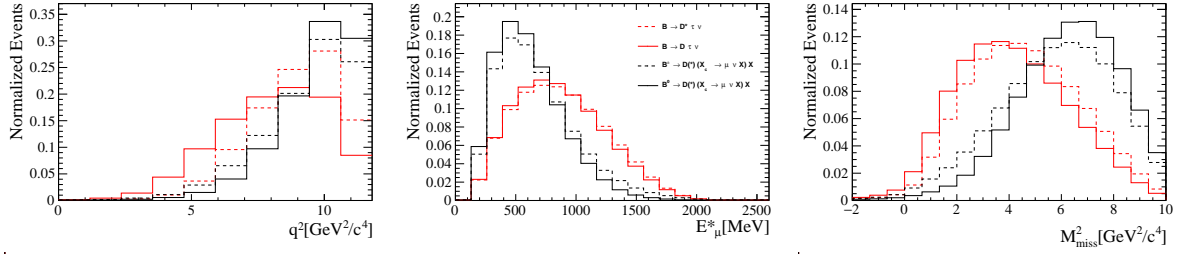Figure 102: Projections of the $B \to D(X_c \to \mu\nu X)X$ template. The $B \to D^{(*)}\tau\nu$ templates are overlaid with the same normalization in red. The templates are drawn from 2015 tracker only MC, before any shape reweight.

of the $D_1 D_2$ system:

$$w(\alpha_1) = 1 + 2\alpha_1 \left( \sqrt{\left( \frac{m_{D_1 D_2}^2 - (m_{D_1} + m_{D_2})^2}{(m_B - m_K)^2 - (m_{D_1} + m_{D_2})^2} \right)} - \frac{1}{2} \right), \tag{188}$$

$$w(\alpha_2) = (1 - \alpha_2) + 8\alpha_2 \left( \sqrt{\left( \frac{m_{D_1 D_2}^2 - (m_{D_1} + m_{D_2})^2}{(m_B - m_K)^2 - (m_{D_1} + m_{D_2})^2} \right)} - \frac{1}{2} \right)^2. \tag{189}$$

Two sets of alternative templates are then defined at the values of $\alpha_1, \alpha_2 = \pm 1$. The fit is then able to interpolate between the nominal and the alternative templates, effectively measuring the best estimate of the $\alpha_1$ and $\alpha_2$ parameters. These parameters are common to every region, and the highest sensitivity to them is coming from the DD region.

In Fig. 103 and Fig. 104, the shape of the nominal template, along with the one of the $\pm 1\sigma$ alternative ones, in all the control region of the fit, is reported.

The normalization of these templates is constructed in the fit as follows. The total yield for the $B \to D(H_c \to \mu\nu X')X$ decays for each region $r$, $N^r(DD)$, is floated independently. The fractions of the different components are floated, shared across all the regions. The yield for each component $i$ in a region $r$ is expressed as follows:

$$N_i^r = N^r(DD) \times \frac{f_i^{ISO} \times \varepsilon_i^{ISO \to r}}{\sum_j f_j^{ISO} \times \varepsilon_j^{ISO \to r}} = N^r(DD) \times \mathcal{F}_i \tag{190}$$

where $N^r(DD)$ is the total DD yield in region $r$, $f_{i(j)}^{ISO}$ denotes the fraction of component $i(j)$ in the ISO region, and $\varepsilon_{i(j)}^{ISO \to r}$ denotes the ratio of efficiencies for component $i(j)$ between region $r$ and the signal region.

The ratios of efficiencies are taken from simulation. Since the efficiency for the Multi Body component depends on the value of the $\alpha$ parameters, this is evaluated at three values of the $\alpha$ parameters ($\alpha = 0, \pm 1$) and the fit interpolates between the three points, effectively taking into account the change of the efficiencies due to the change in the shape of this component. Since the fractions in the signal sample should be normalized to 1, the fractions are parametrized in the following way:

$$f_{B_d,2body}^{ISO} = (1 - f_{B_u}) \times f_{DD}^{B_d}, \tag{191}$$

$$f_{B_d,multibody}^{ISO} = (1 - f_{B_u}) \times (1 - f_{DD}^{B_d}), \tag{192}$$

$$f_{B_u,2body}^{ISO} = f_{B_u} \times f_{DD}^{B_u}, \tag{193}$$

$$f_{B_u,multibody}^{ISO} = f_{B_u} \times (1 - f_{DD}^{B_u}), \tag{194}$$

where $f_{B_u}$ is the fraction of events expected to be coming from the decay of $B^{\pm}$, and $f_{DD}^{B_u}$ and $f_{DD}^{B_d}$ are the fractions of Two-body decays from the $B^0$ and $B^{\pm}$, respectively.

The model used for this contribution is summarised in Table 52.

| Component | Shape | Normalization |
|---|---|---|
| $B^0 \to D^-(X_c \to \mu^+ \nu_\mu X)$ | MC | $N(DD) \times \mathcal{F}_1(f_{B_u}, f_{DD}^{B_u}, f_{DD}^{B_d}, \alpha_{DD})$ |
| $B^0 \to D^-(X_c \to \mu^+ \nu_\mu X)X$ | MC + Shape Var. | $N(DD) \times \mathcal{F}_2(f_{B_u}, f_{DD}^{B_u}, f_{DD}^{B_d}, \alpha_{DD})$ |
| $B^+ \to D^-(X_c \to \mu^+ \nu_\mu X)$ | MC | $N(DD) \times \mathcal{F}_3(f_{B_u}, f_{DD}^{B_u}, f_{DD}^{B_d}, \alpha_{DD})$ |
| $B^+ \to D^-(X_c \to \mu^+ \nu_\mu X)X$ | MC + Shape Var. | $N(DD) \times \mathcal{F}_4(f_{B_u}, f_{DD}^{B_u}, f_{DD}^{B_d}, \alpha_{DD})$ |

Table 52: Fit components and their normalizations. The $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_3$ and $\mathcal{F}_4$ terms denote the functions that have been described in the main text, and $\alpha_{DD}$ collectively denotes the parameters that control the multi-body shape corrections.
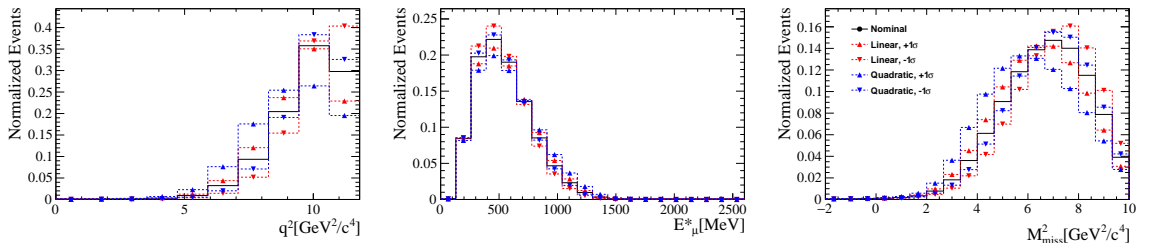
(a) ISO



(b) $\pi^0$



(c) 1OS



(d) 2OS
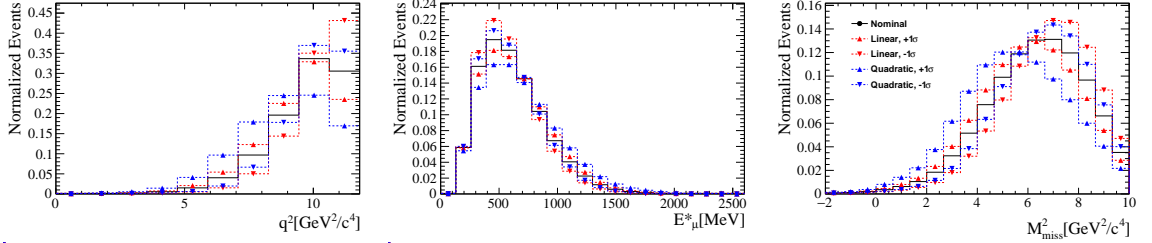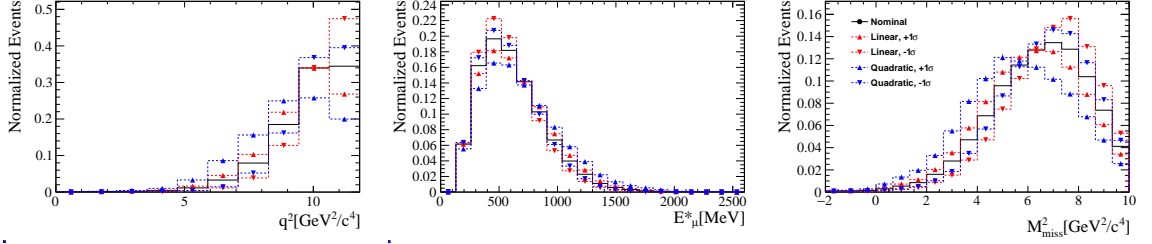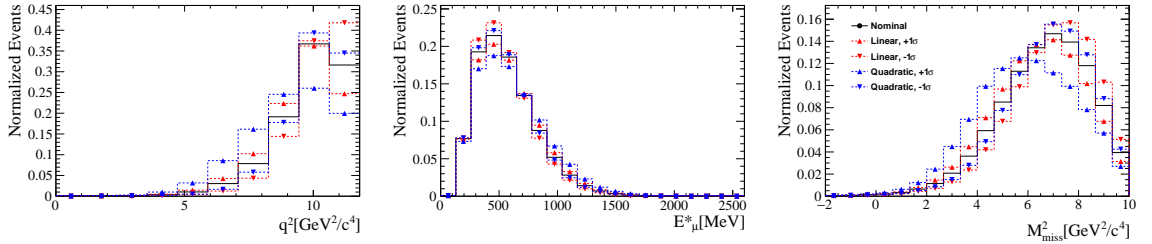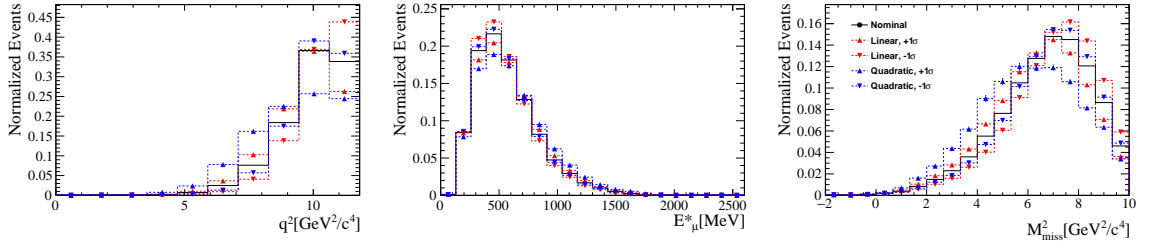


(e) DD

Figure 103: Projection of the nominal $B^0 \to DDX$ multibody template along with the histograms corresponding to the variation parameters $\alpha_i = \pm 1$. The term Linear (quadratic), $\pm 1\sigma$ in the legend corresponds to templates weighted with $w_1(\alpha_1 = \pm 1)$ $(w_2(\alpha_2 = \pm 1))$.
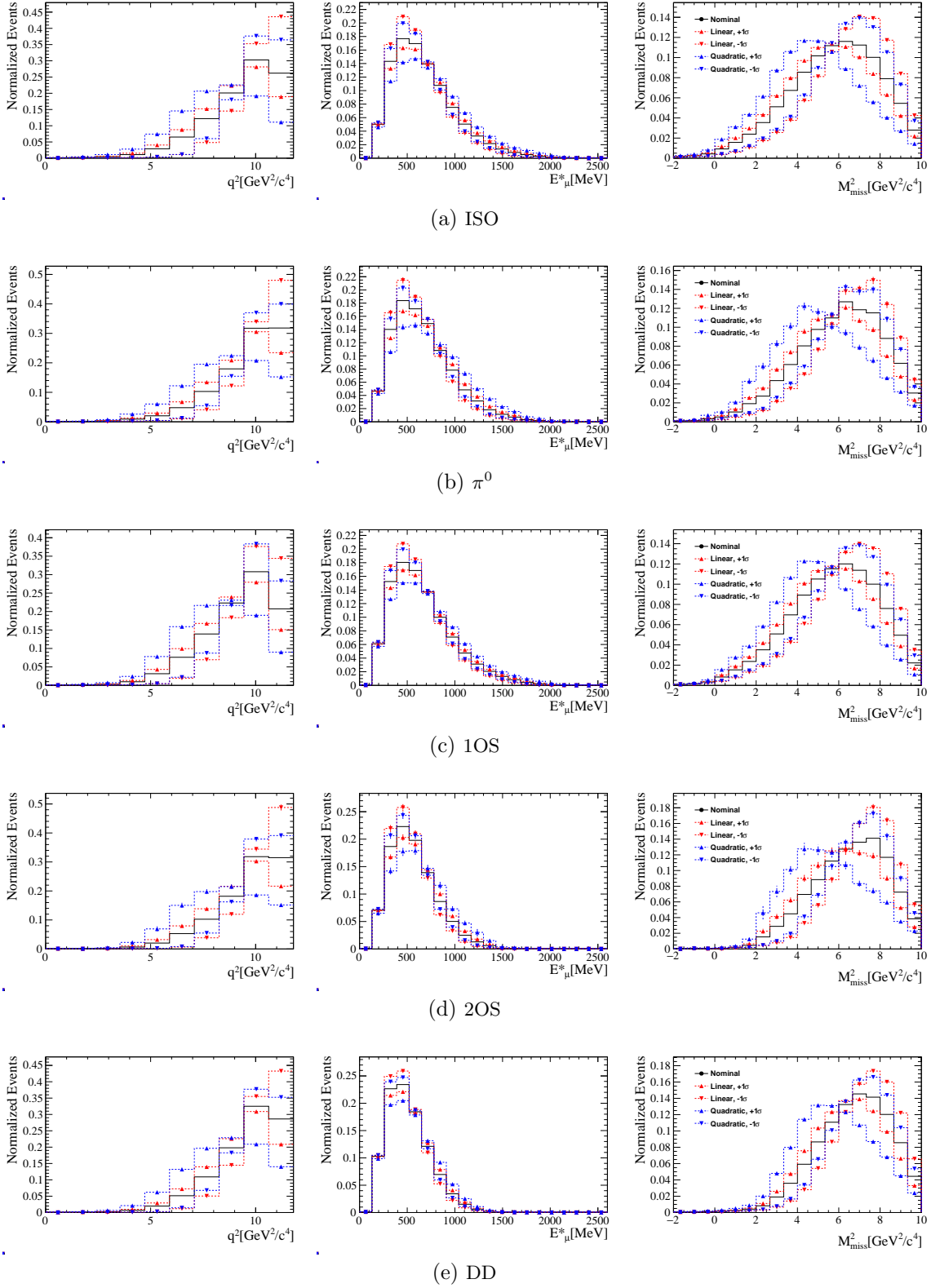
179

(a) ISO



(b) $\pi^0$



(c) 1OS



(d) 2OS



(e) DD

Figure 104: Projection of the nominal $B^{\pm} \rightarrow DDX$ multibody template along with the histograms corresponding to the variation parameters $\alpha_i = \pm 1$. The term Linear (quadratic), $\pm 1\sigma$ in the legend corresponds to templates weighted with $w_1(\alpha_1 = \pm 1)$ $(w_2(\alpha_2 = \pm 1))$.

## 14.7 Double Charm, $B \to D(D_s \to \tau\nu)X$

When the additional charm particle in the decay, coming from $b \to c\bar{c}q$ transitions is a $D_s$ meson, an additional contribution coming from events in which this meson decays into a $\tau$, followed by $\tau \to \mu\nu_\mu\nu_\tau$, can contribute to the background. This background source is simulated using a separate MC sample. Even though the yield of this contribution is expected to be very small, it could be dangerous due to the similarity of its shape with respect to the signal. The shape of the templates as a function of the fit variables is reported in Fig. 105 and Fig 106.
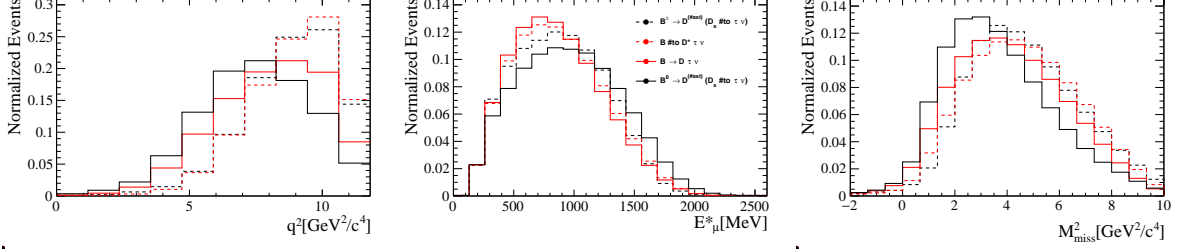


Figure 105: Projections of the $B^0 \to D^{(*)}(D_s \to \tau\nu)X$ and $B^\pm \to D^{(*)}(D_s \to \tau\nu)X$, Two Body decay templates. The $B \to D^{(*)}\tau\nu$ templates are overlaid with the same normalization in red. The templates are drawn from 2015 tracker only MC.
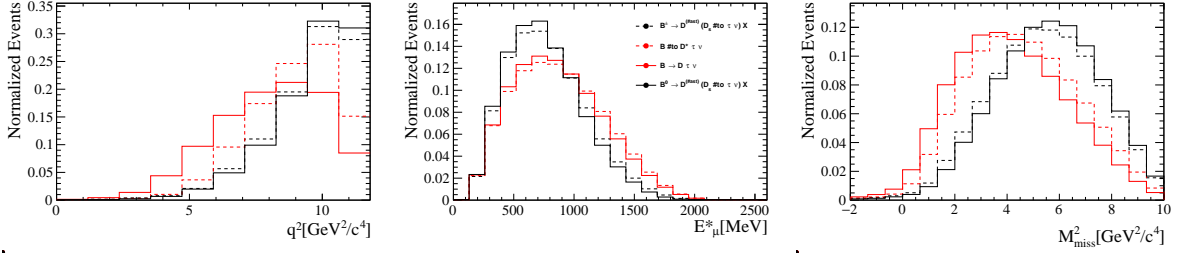


Figure 106: Projections of the $B^0 \to D^{(*)}(D_s \to \tau\nu)X$ and $B^\pm \to D^{(*)}(D_s \to \tau\nu)X$, Multi-Body decay templates. The $B \to D^{(*)}\tau\nu$ templates are overlaid with the same normalization in red. The templates are drawn from 2015 tracker only MC, before any shape reweight.

This contribution is included in the fit in the same way as the muonic counterpart, dividing it into four templates and with the shape parameters shared with the muonic component. The model for this contribution is reported in Table 53.

Each contribution is normalized to the respective muonic mode, using ratios of efficiencies and branching fractions.

The efficiency ratios are encoded in the fraction $f_\mu^\tau$, which has the following expression

$$f_\mu^\tau = f_\mu(D_s) \times \frac{N_{MC}^{RECO}(\tau)}{N_{MC}^{GEN}(\tau)} \times \frac{N_{MC}^{GEN}(\mu)}{N_{MC}^{RECO}(\mu)}, \tag{195}$$

where $f_\mu(D_s)$ is the fraction of $D_s$ events generated in each of the muonic double charm subsamples, $N_{MC}^{RECO}(\tau)$ is the number of reconstructed and truth matched events that enter each template and $N_{MC}^{GEN}(\tau)$ is the number of events generated with the given subdecay in the $D_s \to \tau\nu$ sample. $N_{MC}^{RECO}(\mu)$ and $N_{MC}^{GEN}(\mu)$ are the corresponding numbers in the muonic

double-charm sample. In order to evaluate $N_{MC}^{GEN}(\tau)$ and $N_{MC}^{GEN}(\tau)$, the number of events in input to the filtering selections for the $B^0$ and $B^\pm$ decfiles are divided by the generator level efficiencies and multiplied by the expected fractions of events generated in each sub-decay. These numbers, along with the various fractions and with $f_\mu(D_s)$, are reported in Table 54.

These numbers have been evaluated using the EvtGen configurations files used in the MC production.

The branching fraction ratio, $\mathcal{B}_\nu^\tau$ is defined as follows:

$$\mathcal{B}_\mu^\tau = \frac{\mathcal{B}(D_s \to \tau\nu) \times \mathcal{B}(\tau \to \mu\nu\nu)}{\mathcal{B}(D_s \to \mu\nu X)}. \tag{196}$$

The branching fractions used to evaluate this number are summarised in Table 55. Some of the $D_s$ decays have not been measured in the semi-muonic decay, but only in decays involving electrons. In these cases we take the value of the corresponding electron mode from the PDG. The branching fractions ratio $B_\mu^\tau$ is left floating in the fit with a 30% gaussian uncertainty in order to avoid too strong constraints on this background contribution.

| Component | Shape | Normalization |
|---|---|---|
| $B^0 \to D^-(D_s^+ \to \tau^+\nu_\tau)$ | MC | $N(DD) \times \mathcal{F}_1(f_{B_u}, f_{DD}^{B_u}, f_{DD}^{B_d}, \alpha_{DD}) \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$ |
| $B^0 \to D^-(D_s^+ \to \tau^+\nu_\tau)X$ | MC + Shape Var. | $N(DD) \times \mathcal{F}_2(f_{B_u}, f_{DD}^{B_u}, f_{DD}^{B_d}, \alpha_{DD}) \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$ |
| $B^+ \to D^-(D_s^+ \to \tau^+\nu_\tau)$ | MC | $N(DD) \times \mathcal{F}_3(f_{B_u}, f_{DD}^{B_u}, f_{DD}^{B_d}, \alpha_{DD}) \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$ |
| $B^+ \to D^-(D_s^+ \to \tau^+\nu_\tau)X$ | MC + Shape Var. | $N(DD) \times \mathcal{F}_4(f_{B_u}, f_{DD}^{B_u}, f_{DD}^{B_d}, \alpha_{DD}) \times f_{\tau/\mu} \times \hat{\mathcal{B}}_\mu^\tau$ |

Table 53: Fit components and their normalizations. Parameters denoted in red are fixed in the fit. Parameters with a ˆ are constrained. Parameters denoted in blue are shared between the fit regions.

| Component | | fraction |
|---|---|---|
| $B^0 \to D(H_c \to \mu\nu X')X$ | | |
| | Multi-body | 8.6% |
| | Two-body | 20.8% |
| $B^\pm \to D(H_c \to \mu\nu X')X$ | | |
| | Multi-body | 11.6% |
| | Two-body | 11.1% |
| $B^0 \to D(H_c \to \tau\nu)X$ | | |
| | Multi-body | 16.3% |
| | Two-body | 83.7% |
| $B^\pm \to D(H_c \to \tau\nu)X$ | | |
| | Multi-body | 33.8% |
| | Two-body | 66.2% |

Table 54: Fraction of events generated with a $D_s$ in the various Double Charm (sub-)samples.

| decay | $\mathcal{B}$ in decfile (%) | $\mathcal{B}$ PDG (%) | $\mathcal{B}$ used (%) |
|---|---|---|---|
| $D_s \to \phi\mu^+\nu_\mu$ | 2.59 | $3.03 \pm 0.24$ | 3.03 |
| $D_s \to \eta\mu^+\nu_\mu$ | 2.67 | $2.4 \pm 0.5$ | 2.4 |
| $D_s \to \eta'\mu^+\nu_\mu$ | 0.99 | $1.1 \pm 0.5$ | 1.1 |
| $D_s \to \mu^+\nu_\mu$ | 0.59 | $0.549 \pm 0.016$ | 0.549 |
| $D_s \to K^0\mu^+\nu_\mu$ | 0.37 | $0.34 \pm 0.04$ | 0.37 |
| $D_s \to K^{*0}\mu^+\nu_\mu$ | 0.18 | $0.215 \pm 0.028$ | 0.18 |
| $D_s \to f_0\mu^+\nu_\mu$ | 0.20 | $0.13 \pm 0.04$ | 0.20 |
| $D_s \to \tau^+\nu_\mu$ | 100 | $5.48 \pm 0.23$ | 5.48 |
| $\tau \to \mu\nu\nu$ | 100 | $17.39 \pm 0.04$ | 17.39 |

Table 55: Branching fraction values used in the evaluation of $\mathcal{B}_\mu^\tau$. The ones reported in blue are inferred from the electron modes.

## 14.8 Baryonic double charm

This background is made of decays of $\Lambda_b$ baryons to open charm particles, in which the $\mu$ is coming from the subsequent decay of one of them, usually from $\Lambda_c \to \Lambda\mu\nu$. This contribution is expected

to be small, and it is inserted in the fit through a single template normalized independently with respect to the rest of the double charm contributions. The normalization of this component is measured in the signal region, and the expected normalization in the control regions is then constrained in the fit using the ratios of efficiencies measured from MC.

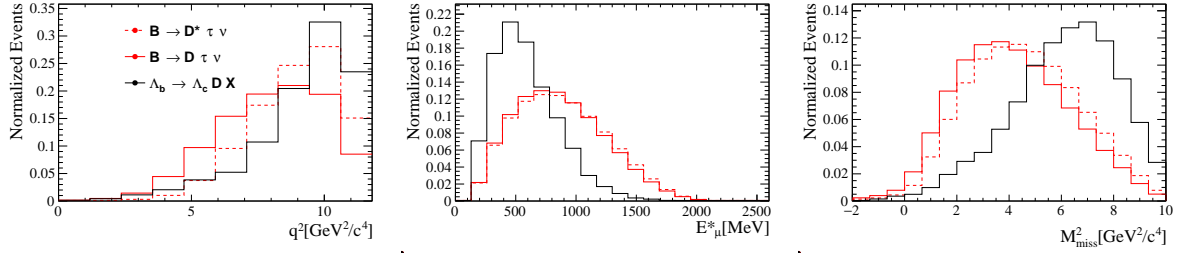The shape of the template of this contribution as a function of the fit variables is reported in Fig. 107.



Figure 107: Projections of the $\Lambda_b \to \Lambda_c DX$ template. The $B \to D^*\tau\nu$ template is overlaid with the same normalization in red. The templates are drawn from 2015 tracker only MC.

## 14.9  $\mu$-MisID

The shape of the $\mu$-MisID background is evaluated as reported in Sect. . The template is then normalized and the normalization is left floating freely in the fit, independently within each region. The template projections for this component in the signal region are shown in Fig. 108.
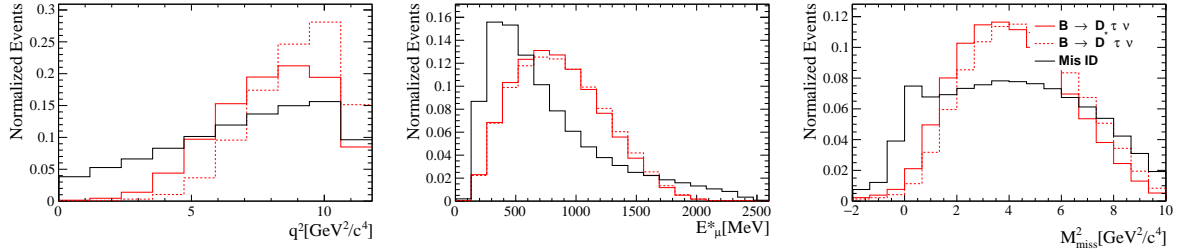


Figure 108: Projections of the MisID template. The $B \to D^{(*)}\tau\nu$ template is overlaid with the same normalization in red.

## 14.10  $D\mu$-combinatorial

The shape of the $D\mu$-combinatorial background is evaluated as reported in Sect. 12.3. The template is then normalized and the normalization is left floating freely in the fit, independently within each region. The template projections for this component in the signal region are shown in Fig. 109.
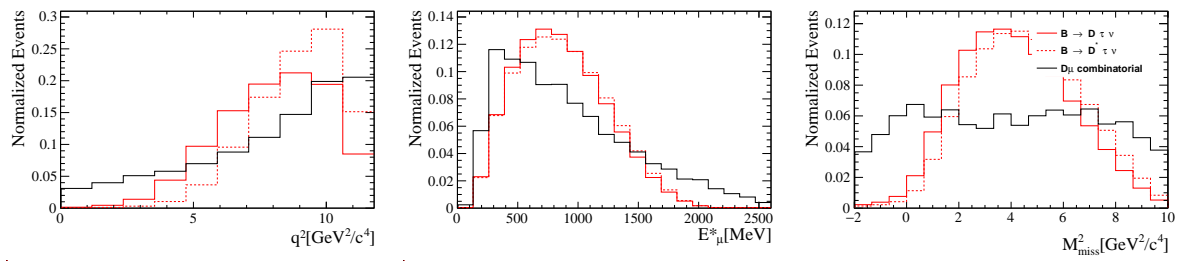
Figure 109: Projections of the $D\mu$ combinatorial template. The $B \to D^{(*)}\tau\nu$ template is overlaid with the same normalization in red.

# 15 Fit result

In this section, the results of the fit to the 2015+2016 data sample are reported. The fit is performed simultaneously to the ISO, $\pi^0$, 1OS, 2OS and $DD$ samples. The projections of the fit are reported in Figures 110, 111, 112,113 and 114. In Figures 110, 116, 117,118 and 119, the fit result is reported in 4 bins of $q^2$. In order to visually blind the value of $\mathcal{R}_{\mathrm{raw}}(D^+)$ and $\mathcal{R}_{\mathrm{raw}}(D^*)$, the two signal components are plotted with the same colour.

The list of nuisance parameters is reported in Table 57 and 58.

It can be noticed that some branching fraction ratios controlling the $B \to D^{**}\mu\nu$ background composition are not compatible with the constraint in input to the fit. This can be either due to a mismodelling of the single components in the MC used to produce the templates or to a lack of decays, especially non-resonant decays of the $D^{**}$ states. In the first case, an additional systematic uncertainty has been evaluated for this purpose, as it will be described in the next chapter. Any discrepancy coming from the second case should be absorbed from the varying shape and normalization of the $B \to D_J^{**}\mu\nu$ template. In any case, the goodness of fit in all the fitted regions, especially the control ones, is an assurance of the goodness of the overall description of the background processes in the variables used in the fit. Some other smaller discrepancies between the fitted parameters of interest and the constraint in input to the fit are being investigated at the time of writing of this thesis.

The uncertainty on the parameters of interest $\mathcal{R}_{\mathrm{raw}}(D^+)$ and $\mathcal{R}_{\mathrm{raw}}(D^*)$ are reported in Table 56. This uncertainty contains the statistical uncertainty and part of the systematic uncertainties that will be listed in the next section. The correlation between the two parameters has been measured to be equal to $-0.528$.

| Parameter | Fit value | correlation |
|-----------|-----------|-------------|
| $\mathcal{R}_{raw}(D^+)$ | $xxx \pm 0.0046$ | -0.528 |
| $\mathcal{R}_{raw}(D^*)$ | $xxx \pm 0.0057$ | |

Table 56: Errors on the parameters of interest determined by the fit.
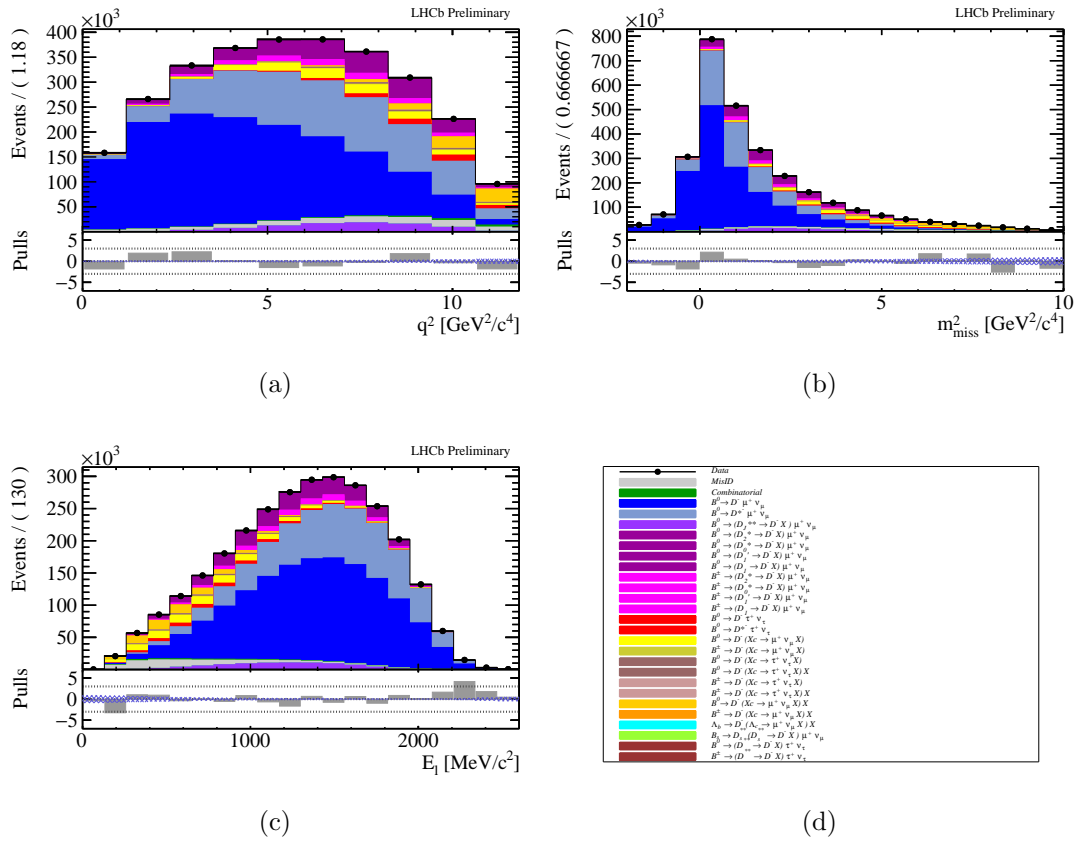
Figure 110: Projection of the nominal fit result in the ISO category, combined 2015 and 2016 dataset.
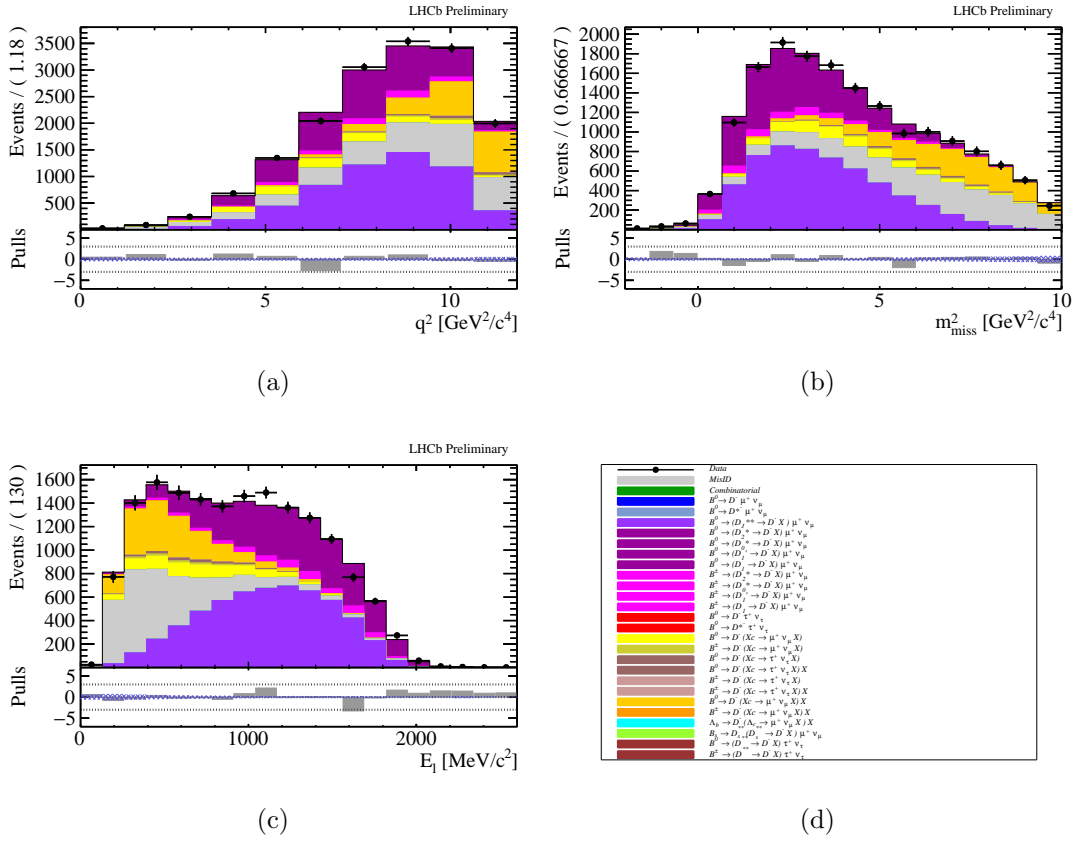
Figure 111: Projection of the nominal fit result in the $\pi^0$ category, combined 2015 and 2016 dataset.

Figure 112: Projection of the nominal fit result in the 1OS category, combined 2015 and 2016 dataset.

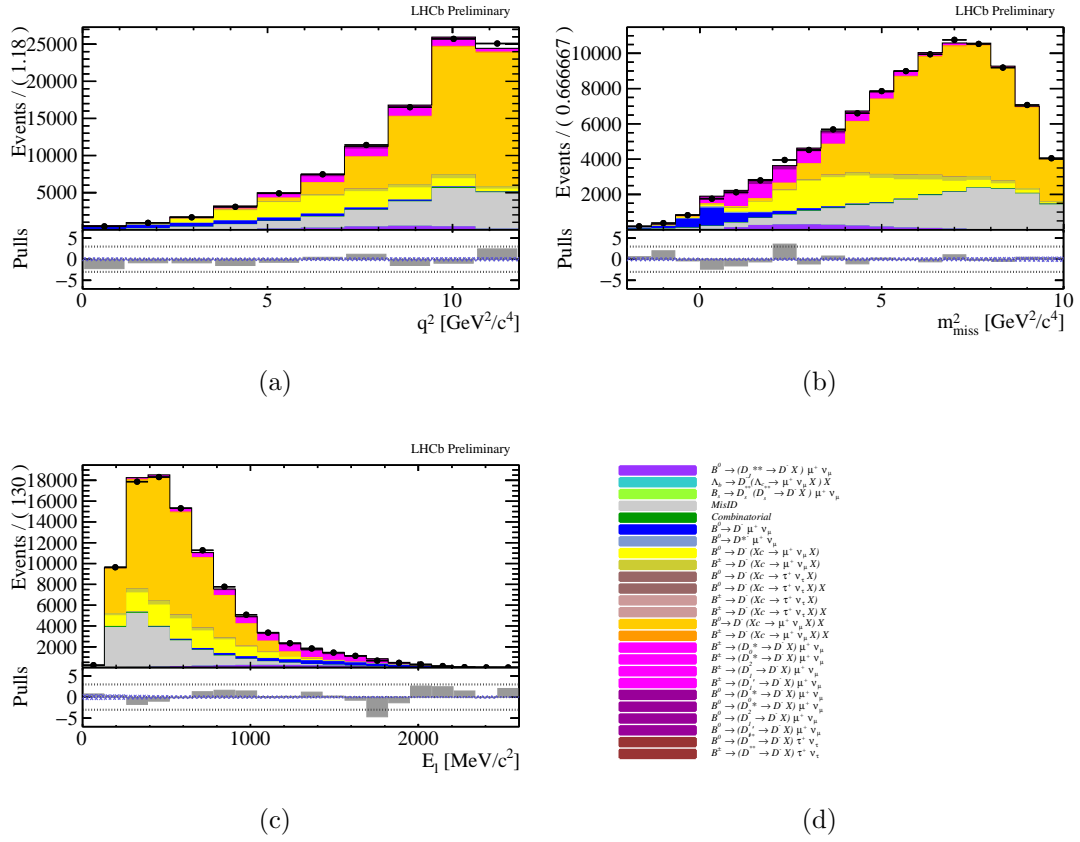Figure 113: Projection of the nominal fit result in the 2OS category, combined 2015 and 2016 dataset.
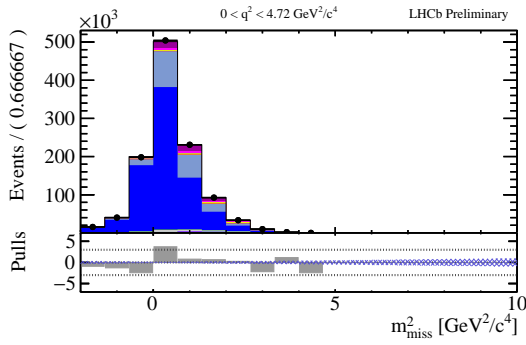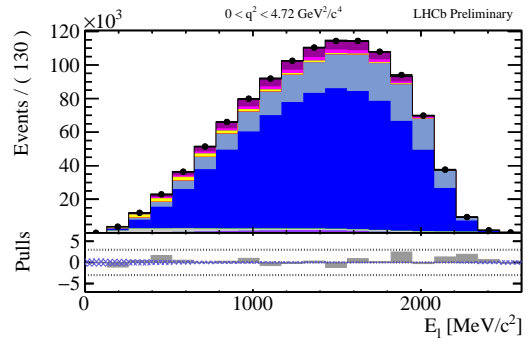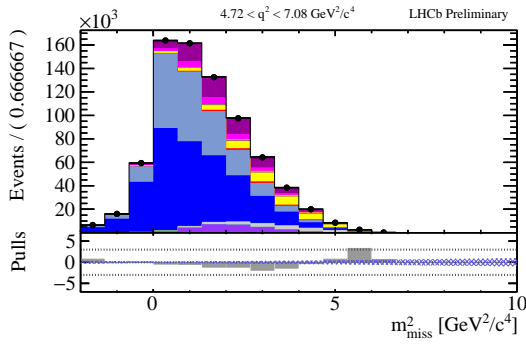
Figure 114: Projection of the nominal fit result in the DD category, combined 2015 and 2016 dataset.

(a)

(b)

(c)

(d)

(e)

(f)

(g)

192
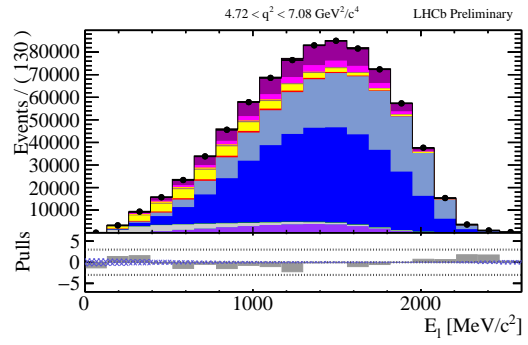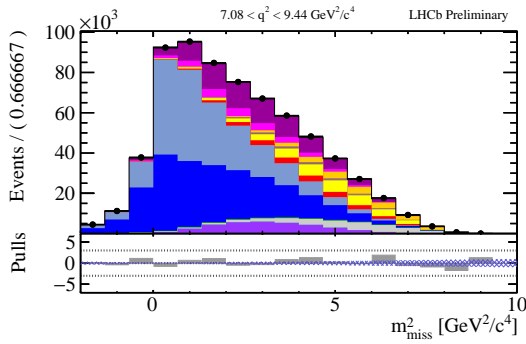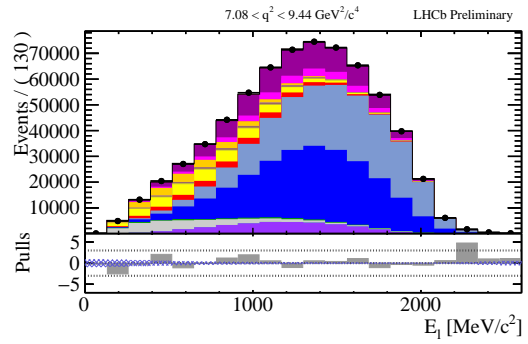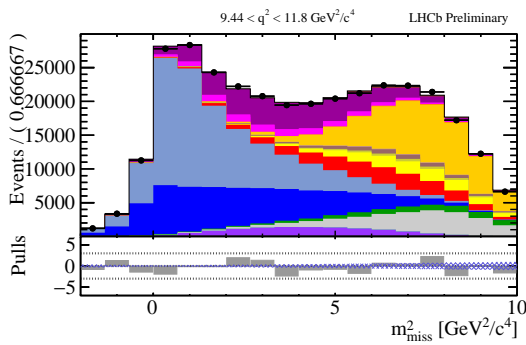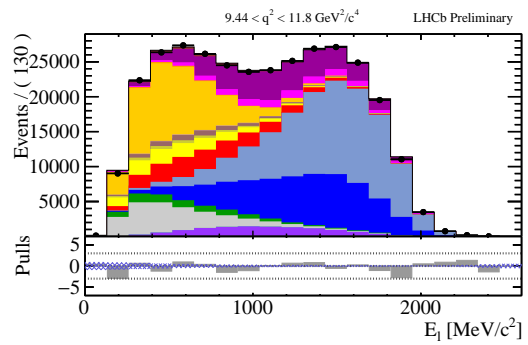
(h)

Figure 115: Separate projections in $q^2$ bins of the nominal fit result in the ISO category, combined 2015 and 2016 dataset.

(a)

(b)

(c)

(d)

(e)

(f)

193
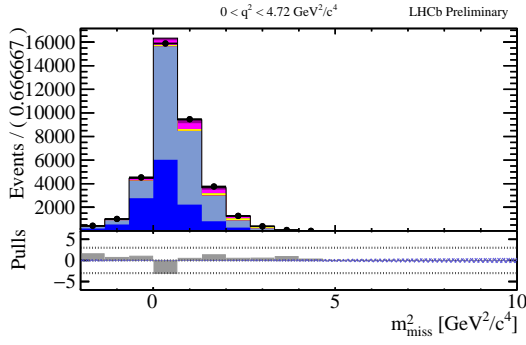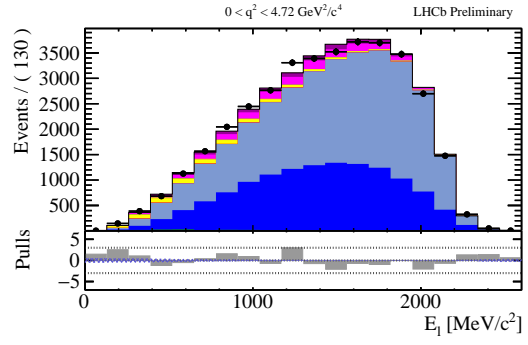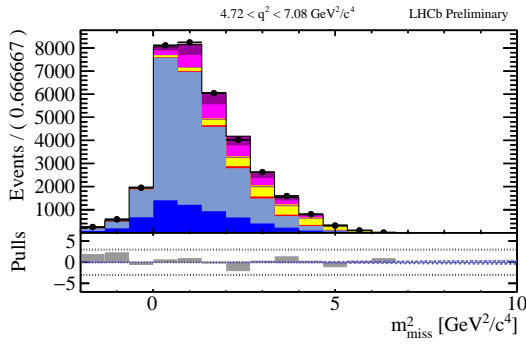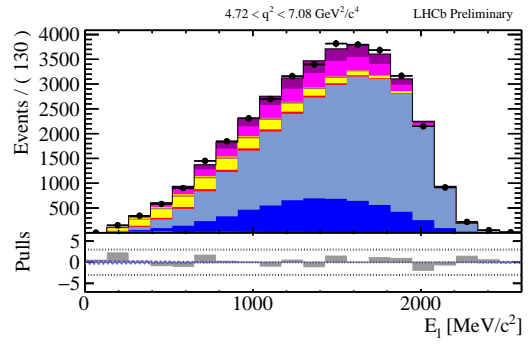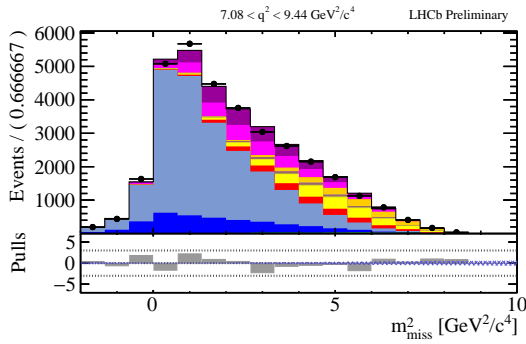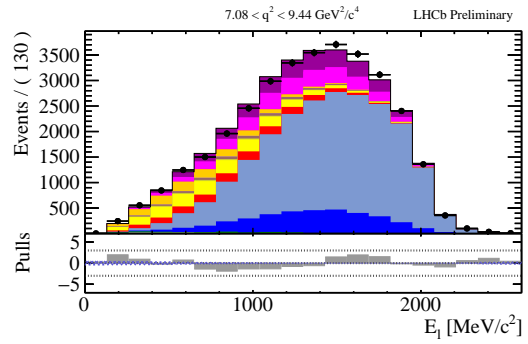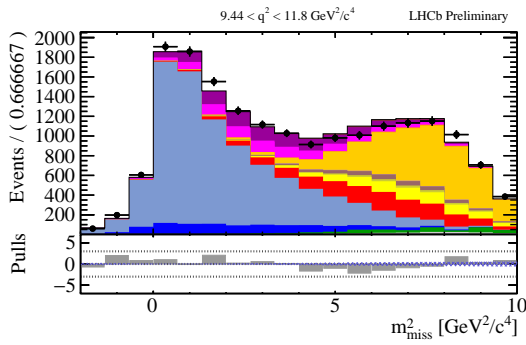
(g)
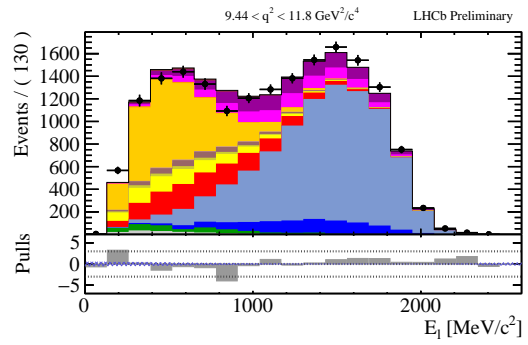
(h)

Figure 116: Separate projections in $q^2$ bins of the nominal fit result in the $\pi^0$ category, combined 2015 and 2016 dataset.
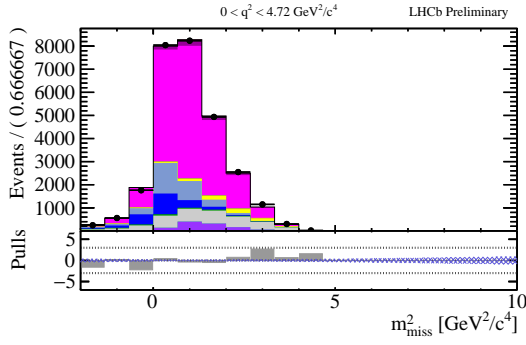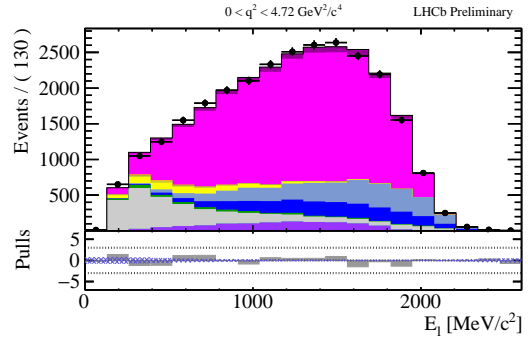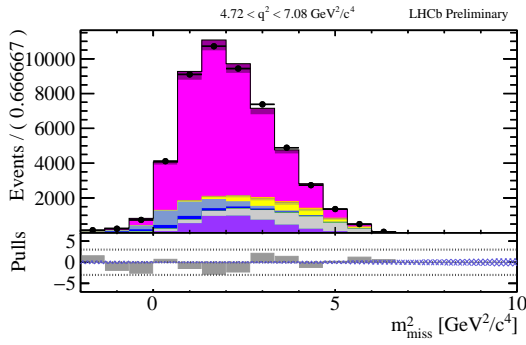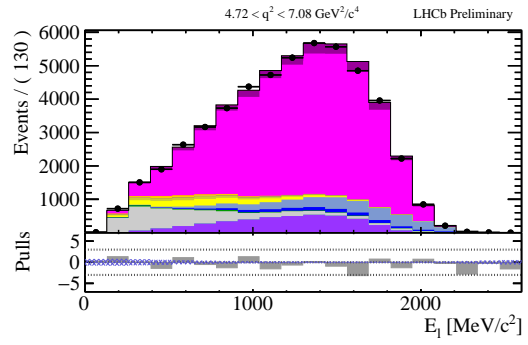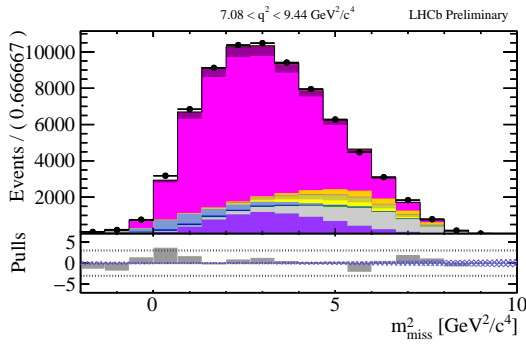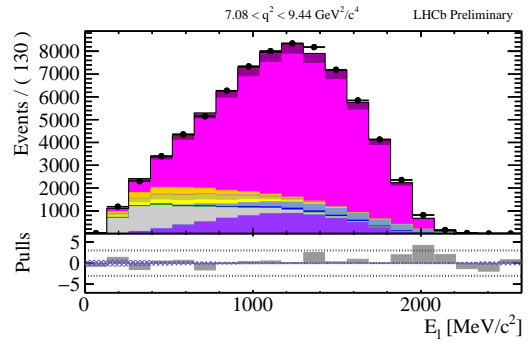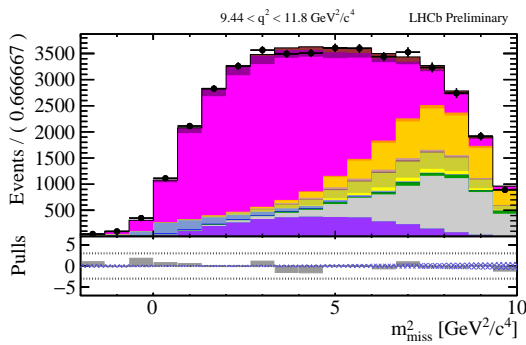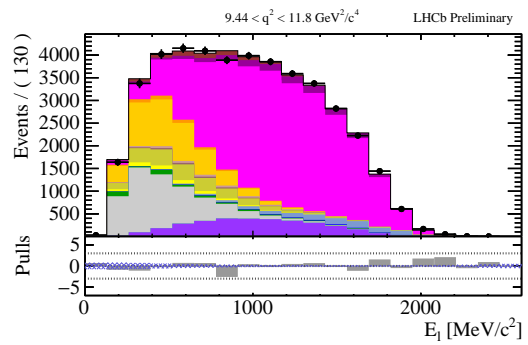
Figure 117: Separate projections in $q^2$ bins of the nominal fit result in the 1OS category, combined 2015 and 2016 dataset.

(a)

(b)

(c)

(d)

(e)

(f)

(g)

195

(h)
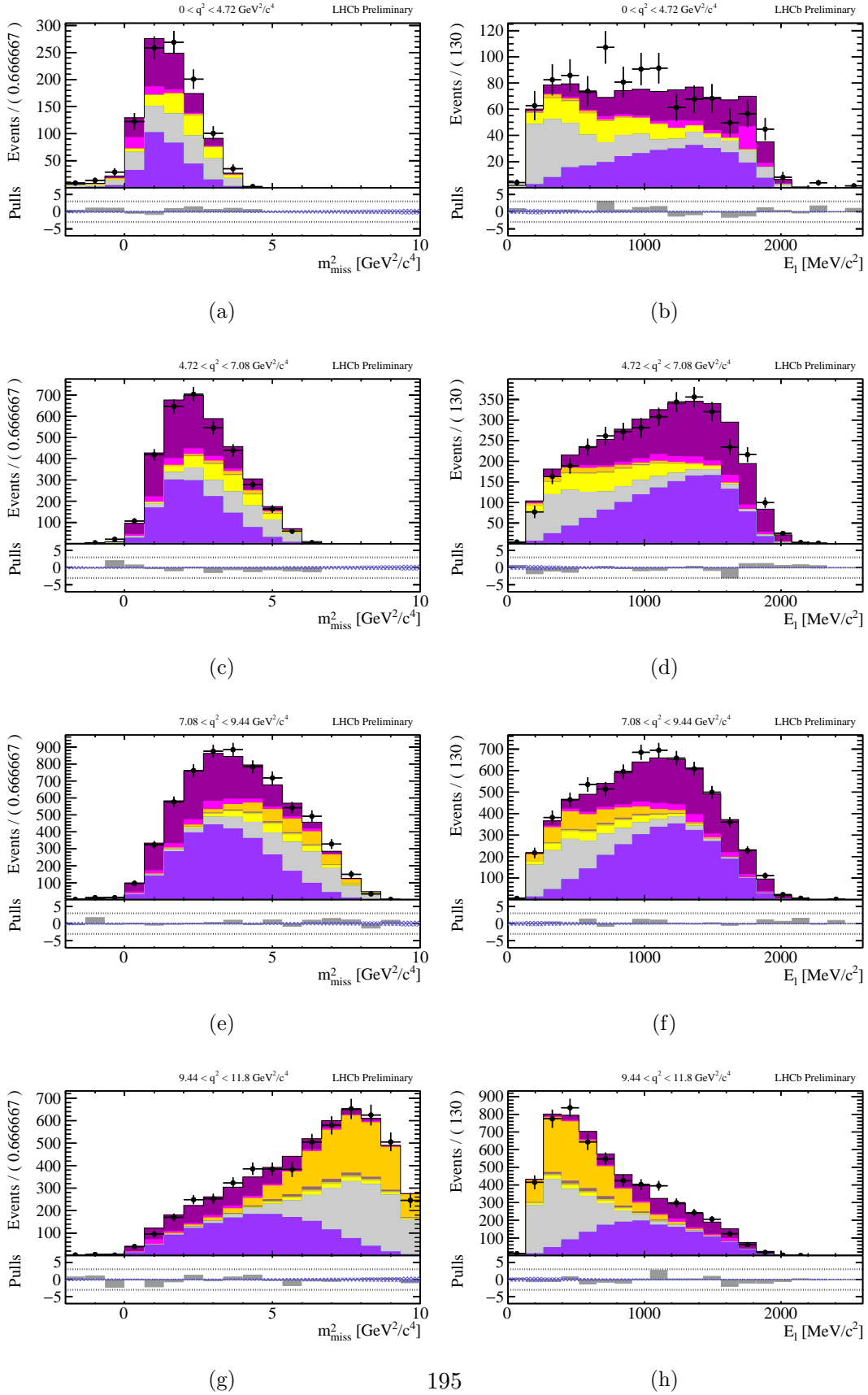
Figure 118: Separate projections in $q^2$ bins of the nominal fit result in the 2OS category, combined 2015 and 2016 dataset.

(a)　　　　　　　　(b)

(c)　　　　　　　　(d)

(e)　　　　　　　　(f)
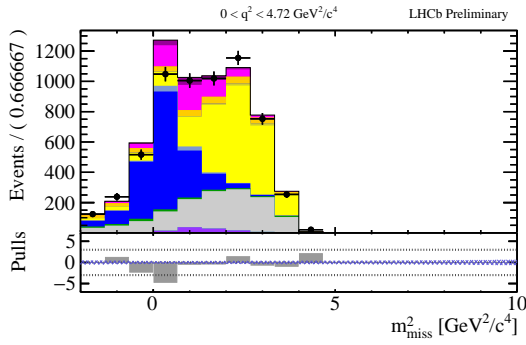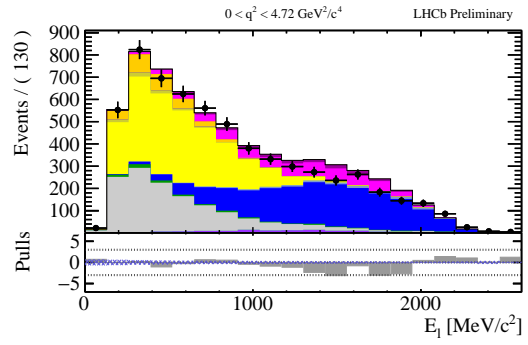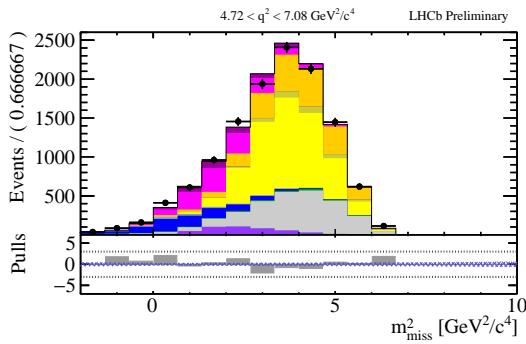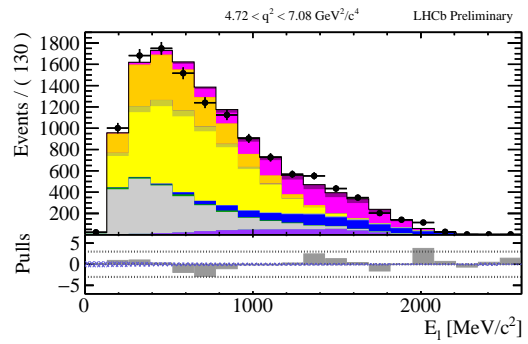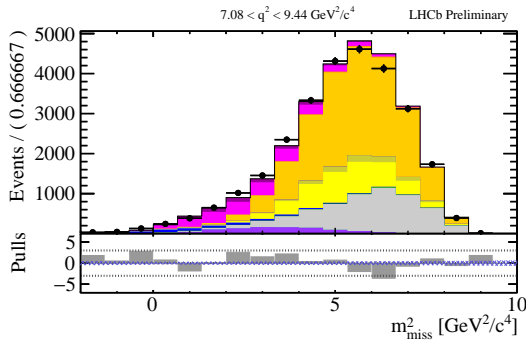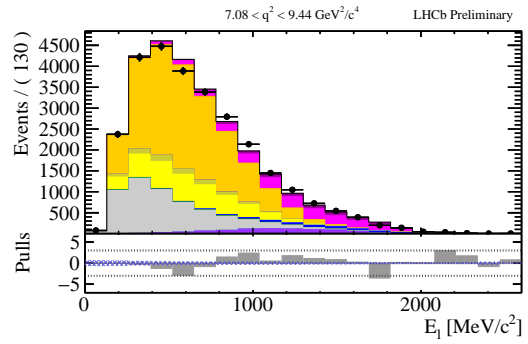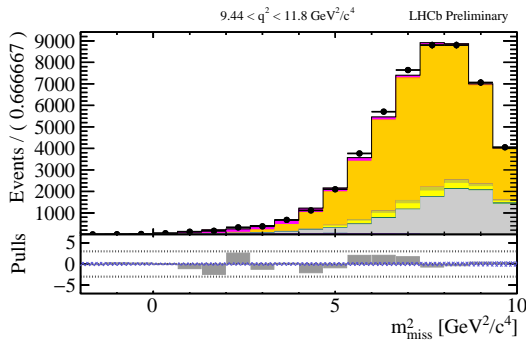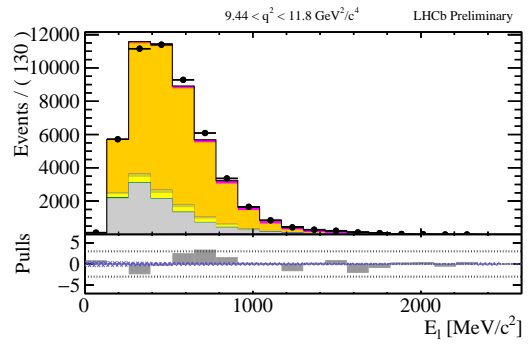
(g)　　　196　　　(h)

Figure 119: Separate projections in $q^2$ bins of the nominal fit result in the DD category, combined 2015 and 2016 dataset.

| Parameter | Fit value | External constraint |
|---|---|---|
| $\mathcal{B}(B^0 \to D_0 \mu\nu)/\mathcal{B}(B^\pm \to D_1 \mu\nu)$ | $2.32 \pm 0.449$ | $1.532 \pm 0.621$ |
| $\mathcal{B}(B^0 \to D_1 \mu\nu)/\mathcal{B}(B^\pm \to D_1 \mu\nu)$ | $2.12 \pm 0.0578$ | $0.462 \pm 0.055$ |
| $\mathcal{B}(B^0 \to D_1' \mu\nu)/\mathcal{B}(B^\pm \to D_1 \mu\nu)$ | $1.04 \pm 0.146$ | $0.511 \pm 0.152$ |
| $\mathcal{B}(B^0 \to D_2 \mu\nu)/\mathcal{B}(B^\pm \to D_1 \mu\nu)$ | $1.84 \pm 0.167$ | $0.730 \pm 0.176$ |
| $\mathcal{B}(B^\pm \to D_0 \mu\nu)/\mathcal{B}(B^\pm \to D_1 \mu\nu)$ | $3.58 \pm 0.399$ | $2.554 \pm 0.539$ |
| $\mathcal{B}(B^0 \to D_1' \mu\nu)/\mathcal{B}(B^\pm \to D_1 \mu\nu)$ | $2.04 \pm 0.184$ | $0.891 \pm 0.206$ |
| $\mathcal{B}(B^\pm \to D_2 \mu\nu)/\mathcal{B}(B^\pm \to D_1 \mu\nu)$ | $4.14 \pm 0.212$ | $1.896 \pm 0.221$ |
| $\mathcal{B}(D_s \to (\tau \to \mu\nu\nu)\nu)/\mathcal{B}(D_s \to \mu\nu)$ | $0.0846 \pm 0.0264$ | $0.142 \pm 0.043$ |
| $N(B_s \to D_s^{**}\mu\nu)$ | $0.0 \pm 2770$ | |
| $N(B \to D\mu\nu)$ | $1.43 \times 10^6 \pm 3.99 \times 10^3$ | |
| $N(B \to D^*\mu\nu)$ | $7.27 \times 10^5 \pm 6.78 \times 10^3$ | |
| $N(\Lambda_b \to \Lambda_c DX)$ | $0.000618 \pm 400$ | |
| $\mathcal{R}_{raw}(D^+)$ | $xxx \pm 0.00456$ | |
| $\mathcal{R}_{raw}(D^*)$ | $xxx \pm 0.00573$ | |
| $\alpha_1$ | $-0.792 \pm 0.0528$ | |
| $\alpha_2$ | $0.379 \pm 0.0331$ | |
| $\alpha^{DJ}$ | $-0.634 \pm 0.255$ | |
| $\tilde{\Delta}(a_0)$ | $3.57 \times 10^{-6} \pm 1.09 \times 10^{-5}$ | $0 \pm 1.165 \times 10^{-5}$ |
| $\tilde{\Delta}(a_{01})$ | $-0.0142 \pm 0.0214$ | $0 \pm 0.02179$ |
| $\tilde{\Delta}(a_{02})$ | $0.091 \pm 0.0962$ | $0 \pm 0.1001$ |
| $\tilde{\Delta}(a_1)$ | $0.000128 \pm 1.78 \times 10^{-5}$ | $0 \pm 4.931 \times 10^{-5}$ |
| $\tilde{\Delta}(a_{+0})$ | $2.72 \times 10^{-5} \pm 9.5 \times 10^{-5}$ | $0 \pm 9.579 \times 10^{-5}$ |
| $\tilde{\Delta}(a+1)$ | $-0.00101 \pm 0.000971$ | $0 \pm 0.001$ |
| $\tilde{\Delta}(a_{+2})$ | $0.0564 \pm 0.00827$ | $0 \pm 0.01391$ |
| $\tilde{\Delta}(b_0)$ | $2.32 \times 10^{-5} \pm 2.79 \times 10^{-5}$ | $0 \pm 6.789 \times 10^{-5}$ |
| $\tilde{\Delta}(b_1)$ | $0.000203 \pm 0.000175$ | $0 \pm 0.0008532$ |
| $\tilde{\Delta}(c_1)$ | $0.0279 \pm 0.00117$ | $0 \pm 0.009383$ |
| $\tilde{\Delta}(c_2)$ | $0.00107 \pm 0.00144$ | $0 \pm 0.004536$ |
| $\tilde{\Delta}(d_0)$ | $-0.0052 \pm 0.0018$ | $0 \pm 0.007$ |
| $\tilde{\Delta}(d_1)$ | $-0.0079 \pm 0.0314$ | $0 \pm 0.036$ |
| $\tilde{\Delta}(\tau(1))$ | $-0.0342 \pm 0.0215$ | $0 \pm 0.0214$ |
| $\tilde{\Delta}(\hat{\tau}_1)$ | $0.843 \pm 0.162$ | $0 \pm 0.210$ |
| $\tilde{\Delta}(\hat{\tau}_2)$ | $-5.34 \pm 0.569$ | $0 \pm 1.426$ |
| $\tilde{\Delta}(\tau')$ | $-0.0782 \pm 0.0978$ | $0 \pm 0.130$ |
| $\tilde{\Delta}(\zeta(1))$ | $-2.8 \pm 0.43$ | $0 \pm 1.426$ |
| $\tilde{\Delta}(\hat{\zeta}')$ | $0.166 \pm 0.0623$ | $0 \pm 0.053$ |
| $\tilde{\Delta}(\hat{\zeta}(1))$ | $0.0903 \pm 0.111$ | $0 \pm 0.239$ |
| $f_{B_u}$ | $0.0369 \pm 0.0169$ | |
| $f_{DD}^{B_d}$ | $0.572 \pm 0.00819$ | |
| $f_{DD}^{B_u}$ | $0.797 \pm 0.142$ | |
| $\mathcal{B}(B \to D^{**}\tau\nu) \cdot \mathcal{B}(\tau \to \mu\nu\nu)/\mathcal{B}(B \to D^{**}\mu\nu)$ | $0.0153 \pm 0.00663$ | $0.0148 \pm 0.0074$ |

Table 57: Values of the common parameters determined by the fit. Also shown are the external constraints. Note that the constraints on the form-factor parameters are defined along the diagonal directions of the external correlation matrices, as explained in Sec. 14.

| Parameter | Fit value |
|---|---|
| Parameters in the ISO region | |
| $N(B \to D_J^{**}\mu\nu)(\text{ISO})$ | $1.23 \times 10^{+5} \pm 1.07 \times 10^4$ |
| $N(B^{\pm} \to D_1^{**}\mu\nu)$ (ISO) | $1.11 \times 10^4 \pm 839$ |
| $N(DD)$ ISO | $1.74 \times 10^5 \pm 4.8 \times 10^3$ |
| $N(\text{MisID})$ (ISO) | $9.03 \times 10^4 \pm 2.73 \times 10^3$ |
| $N(\text{Combinatorial})$ (ISO) | $1.87 \times 10^4 \pm 1.77 \times 10^3$ |
| Parameters in the $1\pi^{\pm}$ region | |
| $N(B \to D_J^{**}\mu\nu)(1\pi^{\pm})$ | $3.4 \times 10^{+4} \pm 2.41 \times 10^3$ |
| $N(B^{\pm} \to D_1^{**}\mu\nu)$ $(1\pi^{\pm})$ | $8.13 \times 10^3 \pm 529$ |
| $N(DD)$ $(1\pi^{\pm})$ | $1.46 \times 10^4 \pm 893$ |
| $N(\text{MisID})$ $(1\pi^{\pm})$ | $1.84 \times 10^4 \pm 657$ |
| $N(\text{Combinatorial})(1\pi^{\pm})$ | $1.4 \times 10^3 \pm 263$ |
| $TF(B \to D\mu\nu)$ $(1\pi^{\pm})$ | $0.00208 \pm 0.000222$ |
| $TF(B \to D^*\mu\nu)$ | $0.0117 \pm 0.000875$ |
| Parameters in the $\pi^0$ region | |
| $N(B^{\pm} \to D_1^{**}\mu\nu)$ $(\pi^0)$ | $817 \pm 58.7$ |
| $N(DD)(\pi^0)$ | $9.88 \times 10^3 \pm 389$ |
| $N(\text{MisID})$ | $562 \pm 193$ |
| $N(\text{Combinatorial})$ | $805 \pm 221$ |
| $TF(B \to D\mu\nu)$ | $0.0157 \pm 0.000316$ |
| $TF(B \to D^*\mu\nu)$ | $0.0963 \pm 0.00119$ |
| Parameters in the $2\pi^{\pm}$ region | |
| $N(B \to D_J^{**}\mu\nu)$ $(2\pi^{\pm})$ | $6.08 \times 10^3 \pm 379$ |
| $N(B^{\pm} \to D_1^{**}\mu\nu)$ $(2\pi^{\pm})$ | $13 \pm 1.16$ |
| $N(DD)$ 2OS $(2\pi^{\pm})$ | $3.06 \times 10^3 \pm 303$ |
| $N(\text{MisID})$ $(2\pi^{\pm})$ | $3.32 \times 10^3 \pm 266$ |
| Parameters in the DD region | |
| $N(B \to D_J^{**}\mu\nu)$ (DD) | $3.6 \times 10^3 \pm 624$ |
| $N(B^{\pm} \to D_1^{**}\mu\nu)(\text{DD})$ | $241 \pm 31.5$ |
| $N(DD)$ DD | $6.55 \times 10^4 \pm 971$ |
| $N(\text{MisID})(\text{DD})$ | $1.95 \times 10^4 \pm 914$ |
| $N(\text{Combinatorial})(\text{DD})$ | $520 \pm 159$ |
| $TF(B \to D\mu\nu)(\text{DD})$ | $0.00202 \pm 0.000104$ |
| $TF(B \to D^*\mu\nu)(\text{DD})$ | $0.000498 \pm 0.000216$ |

Table 58: Values of the parameters specific to the different data regions.

# 16    Systematic uncertainties

In this chapter the systematic uncertainties associated with the measurement of $\mathcal{R}(D^+)$ and $\mathcal{R}(D^*)$ are reported.

The systematic uncertainties have been divided into Additive uncertainties, which affect the fitted values of $\mathcal{R}_{\mathrm{raw}}(D^+)$ and $\mathcal{R}_{\mathrm{raw}}(D^*)$, and Multiplicative uncertainties, which affect the value of the factors used to convert $\mathcal{R}_{\mathrm{raw}}(D^+)$ and $\mathcal{R}_{\mathrm{raw}}(D^*)$ into the parameters of interest. A summary of the systematic uncertainties has been reported in Table 59 and Table 60. The first additive systematic uncertainties reported in Table 59 have been already taken into account in the fit by means of appropriate nuisance parameters and therefore were already present in the uncertainty reported in Table 56. In these cases the systematic uncertainty is evaluated as the quadrature difference of the uncertainty on the parameters of interest when leaving the associated nuisance parameters floating and when fixing them. In all the other cases, the additive uncertainties have been evaluated using had-hoc checks, as it will be explained in the next sections.

It can be noticed that, since the analysis relies on a ratio of efficiencies between the signal and normalization modes, which share the same final state, the Multiplicative systematic uncertainties are found to be subdominant with respect to the Additive systematic uncertainties.

The way each systematic uncertainty has been evaluated will be described in the next sections.

| Source | $\mathcal{R}_{raw}(D^+)$ | $\mathcal{R}_{raw}(D^*)$ |
|---|---|---|
| Statistical uncertainties | 0.0034 | 0.0042 |
| Systematic uncertainties | 0.0037 | 0.0073 |
| *Evaluated with nuisance parameters* | | |
| $B \to D$ and $B \to D^*$ form factors | 0.0022 | 0.0021 |
| $B \to D^{**}$ form factors | 0.0011 | 0.0014 |
| External BF constraints | 0.0015 | 0.0023 |
| Shape parameters | 0.0012 | 0.0028 |
| *Evaluated with additional checks* | | |
| Mis-ID bkg shape | 0.0017 | 0.0042 |
| $D^{**}$ hadronic fractions | 0.0004 | 0.0026 |
| Simulation size | 0.00088 | 0.0030 |

Table 59: Summary of additive systematic uncertainties (absolute value) on the raw $\mathcal{R}(D^+)$ and $\mathcal{R}(D^*)$ parameters. The statistical uncertainties are obtained from a fit in which all the parameters not related to yields are fixed to their best fit point values.

| Source | Relative size (%) | |
| --- | --- | --- |
| | $\eta_{D^+}$ | $\eta_{D^{*+}}$ |
| Simulation size | 0.14 | 0.15 |
| Muon ID | 0.44 | 0.65 |
| Tracking efficiency correction | $< 0.1$ | $< 0.1$ |
| Neutral Isolation emulation | $< 0.1$ | 0.19 |

Table 60: Summary of multiplicative systematic uncertainties (relative to the central value of the efficiency ratios) on $\mathcal{R}(D^+)$ and $\mathcal{R}(D^*)$.

## 16.1 Form Factor parameters

The systematic uncertainty due to the knowledge of the Form Factor uncertainties is already inserted in the fit using parameters that are allowed to float, as described in Sect. 14. A correction factor to the efficiency ratios depending on the value of the form factor parameters, encoding the change of the efficiencies with the floating form factor parameters, is inserted by hand in the fit. For this reason, the systematic due to form factors is only an additive one. The uncertainty is reported as the change in the value of the fit uncertainties when floating and when fixed to their best fit point value.

## 16.2 External $\mathcal{B}$ constraints

External constraints on branching fractions, as taken from [88], are used in the fit. In particular, the following quantities are constrained using external measurements:

- The relative $B \rightarrow D^{**}\mu\nu$ branching fractions.

- The branching fractions of semitauonic $B \rightarrow D^{**}\tau\nu$ modes with respect to the semi-muonic decays.

- The branching fraction of the $D_s \rightarrow \tau\nu$ decay with respect to the $D_s \rightarrow X\mu\nu$ decays.

The uncertainty on the knowledge of these quantities is evaluated by comparing the fit uncertainties when floating these parameters or fixing them to their best fit value.

## 16.3 Shape parameters

The shape of the high-mass $D^{**}$ Feed Down background and of the Multi-Body Double charm backgrounds are poorly known, and therefore calibrated directly from the data itself, floating dedicated shape parameters. Furthermore, the fractions of Two and Multi-Body Double charm decays, as well the fraction of Double Charm decays from $B^{\pm}$ and $B^0$ decays are directly fit from data. The systematic uncertainty linked to these nuisance parameters is obtained by comparing the total uncertainty on $\mathcal{R}(D^+)_{\text{raw}}$ and $\mathcal{R}(D^*)_{\text{raw}}$ when floating and fixing them to the best fit value.

## 16.4  Simulation size

The uncertainty due to the finite simulation size affects the template shapes of the signal, normalization and all the backgrounds. The size of this effect on the values of $\mathcal{R}_{raw}(D^+)$ and $\mathcal{R}_{raw}(D^*)$ has been evaluated using a bootstrapping technique. The MC samples used to create the templates used for fitting the data have been bootstrapped on a per-event basis.

For the templates that are constructed using RooHammerModel, the bootstrapping is not a feasible option, since this would require the reprocessing of the entire data samples through Hammer many times, which would be too CPU costly. For this reason, a *block bootstrapping* method has been used in those cases. The MC samples have been preprocessed with Hammer in bunches, splitting the data in subfiles. For each template, 8 sub-samples have been defined and preprocessed in order to be joined later on, before the construction of the RooHammerModel template. During bootstrapping, for each component processed with Hammer, 8 random subfiles have been chosen, reinserting the extracted file in the list of available ones. In this way, each bootstrapping iteration, a random sample of events has been used to construct the RooHammerModel template. The number of possible combinations without repetition that could be constructed in this way is around 6000.

Random pseudo datasets have been generated around 500 times. Half of them is fitted with the nominal templates, half of them is fitted with templates constructed with the bootstrapping technique described above. The value of the $\mathcal{R}_{raw}(D^+)$ and $\mathcal{R}_{raw}(D^*)$ used in generation is the one expected if the World Average held. The additive systematic uncertainty is then evaluated as the quadrature difference between the standard deviation of the distribution of fitted values $\mathcal{R}_{raw}(D^+)$ and $\mathcal{R}_{raw}(D^*)$ between the iterations in which the MC templates have been bootstrapped and the cases in which they haven't. The result is reported in Fig. 120, in which the distributions are fitted with a gaussian p.d.f. and the fit result is reported, along with the 68% Confidence Interval on the fitted function, in blue for the toys with nominal MC templates and in red for toys with bootstrapped MC templates.
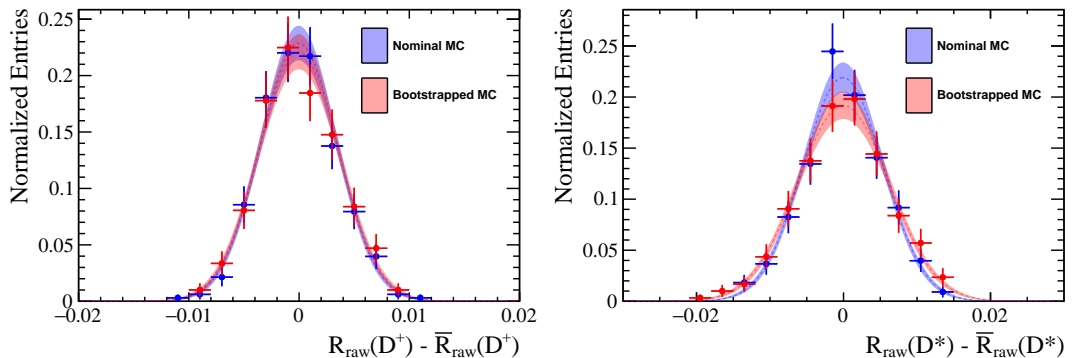


Figure 120: Comparison between the distribution of fitted parameters in toy studies, when the templates used are the nominal ones (blue) or the ones taken from a bootstrapped MC sample (red).

The uncertainty found in this way is around $8.8 \times 10^{-4}$ on $\mathcal{R}(D^+)$ and $3.0 \times 10^{-3}$ on $\mathcal{R}_{raw}(D^*)$, that would translate in a relative uncertainty of 2.5% and 9.6%, respectively, if the World Average held.

We also rely on MC simulation to determine the ratio of efficiencies needed to convert the

fitted yields into branching fraction ratios. The statistical uncertainty on the efficiency ratios, evaluated as a binomial uncertainty, is taken as the effect of their finite statistics on the efficiency ratios.

## 16.5 MisID background shape

The normalization of the $\mu$-MisID background is already floating in the fit, therefore only systematic uncertainties affecting the shape of the $\mu$-MisID template have to be taken into account. A systematic uncertainty coming from the method used to evaluate the PID efficiency, through calibration samples in which the background is subtracted with the sPlot technique, is evaluated. This is done by refitting the data with a template constructed with the same procedure as before, but evaluating the PID efficiencies used in the $\mu$-MisID background estimation using a *fit-and-count* procedure on the calibration samples. The efficiency of selecting pions and kaons as muons is extracted from a two-dimensional fit to the $m(D^*) - m(D^0)$ and $m(D^0)$ variables for each kinematic bin described in section 14.9. The difference in the fitted values of $\mathcal{R}_{raw}(D^+)$ and $\mathcal{R}_{raw}(D^*)$ obtained with respect to the nominal fit to data is quoted as an additive systematic uncertainty.

This systematic uncertainty is very conservative, and further methods to evaluate the systematics affecting the shape of the $\mu$-MisID background template shape are now under consideration, especially using data-driven methods that employ specific additional control regions to calibrate the shape. For this reason, this systematic uncertainty is expected to decrease.

## 16.6 Hadronic $D^{**}$ decays in $B \to D^{**}\ell\nu$ decays

The overall shape of the $B \to D^{**}\ell\nu$ modes, with 1P $D^{**}$ states, is varied in the fit by floating the ratio of branching fraction of each state contribution, as described in Sect. 14 and their Form Factor parameters. The efficiency ratio of each state with respect to the reference $B^+ \to D_1\mu\nu$ mode, in each fit region, is fixed from MC expectation. This in turn relies on a good modelling of the isolation efficiency of each state, which depends on the assumption that the hadronic decays of the $D^{**}$ states are well represented in the MC samples, although these decays are poorly known.

The systematic uncertainty due to this knowledge is evaluated by effectively changing the composition of the MC samples and inserting in the analysis the expected change to the $D^{**}$ efficiency ratios, as it will be explained now.

The efficiency ratios are assumed to depend mainly on the charged isolation, and therefore the number of charged pions in the decay of each state in the MC simulation is inspected. This is reported in Table 61.

The $D_0^{*0}$ and $D_1^{'0}$ only decay to modes with one extra charged pion, and are therefore not considered further. For each of the other six states, decays with two different number of final charged pions are identified. Two extreme efficiencies, for each of these modes and for each of the control regions are then evaluated. One nuisance parameter for each $D^{**}$ mode is then inserted in the fit to effectively change the fraction of decays to each of the two possible sub-modes. The value of the efficiency is interpolated in the fit using the following formula:

$$\varepsilon(x) = x \times \varepsilon_{\min} + (1 - x) \times \varepsilon_{\max}, \tag{197}$$

| Charm hadron | Number of extra charged pions |
|:---:|:---:|
| $D_0^{*+}$ | $0\pi$ or $2\pi$ |
| $D_0^{*0}$ | $1\pi$ |
| $D_1^{+}$ | $0\pi$ or $2\pi$ |
| $D_1^{0}$ | $1\pi$ or $3\pi$ |
| $D_1'^{+}$ | $0\pi$ or $2\pi$ |
| $D_1'^{0}$ | $1\pi$ |
| $D_2^{*+}$ | $0\pi$ or $2\pi$ |
| $D_2^{*0}$ | $1\pi$ or $3\pi$ |

Table 61: Number of extra charged pions generated in the decay of each $D^{**}$ mode, as found in the MC sample.

where $\varepsilon_{\min}$ and $\varepsilon_{\max}$ are the two extreme efficiencies for the mode under study, in a given fit region, to be expected when the mode is assumed to decay 100% of the times to one sub-mode or the other. The extreme efficiencies are evaluated in this way:

- Since the selections used to separate the different fit regions are orthogonal, the efficiencies corresponding to decays not compatible with the number of additional pions selected by the control region are set to zero (for example, the efficiency of selecting events with $D_2^{*+}$ decaying to 2 charged pions in the $1\pi^{\pm}$ region is set to zero).

- The efficiency of selecting a given sub-mode in a region which selects a compatible number of charged pions is evaluated assuming all the $B \to D^{**}\mu\nu$ found in that category are coming from that specific sub-mode. The efficiency is therefore evaluated with the following expression

$$\varepsilon_{sub-mode}^{D^{**},r} = \frac{N_{reco}^{D^{**},r}}{N_{gen}^{D^{**},r} \cdot f_{gen,sub-mode}^{D^{**}}}, \tag{198}$$

where $D^{**}$ refers to the specific mode, $r$ to the control region and $f_{gen,sub-mode}^{D^{**}}$ is the generator level fraction of the specific sub-mode.

For example, the efficiency of selecting the $D_2^{*+}$ mode in the 2OS region is evaluated assuming that all the events found in the relative template contain decays of the $D_2^{*+}$ to two charged pions.

Since the selection of events in the $DD$ region is not orthogonal in terms of charged pions, the efficiency ratios of the $B \to D^{**}\mu\nu$ decays are all floated in that region.

The fit is repeated with the additional nuisance parameters floating (13 parameters, 7 for the efficiency ratios in the $DD$ region and 6 controlling the fraction of the sub-modes in all the other regions), and the difference in the fitted values of $\mathcal{R}_{\text{raw}}(D^+)$ and $\mathcal{R}_{\text{raw}}(D^*)$ between the fit with the additional parameters and the nominal fit is assigned as a systematic uncertainty.

## 16.7 Muon-ID

The muon ID response has a large dependence on the muon momentum, whose distribution is different between the signal and the normalization modes. The efficiency of the muon ID selections are evaluated using calibration samples, in kinematic bins. In order to assess the systematics of this evaluation, the binning of the sample is varied. The additive uncertainty associated to this has not yet been evaluated and will be evaluated by repeating the fit with the new set of templates obtained in this way.

The multiplicative uncertainty is estimated by re-evaluating the efficiency ratios and measuring the effect of this change, as reported in Table 62.

| | 2015 | | 2016 | | comb | |
|---|---|---|---|---|---|---|
| Sample | $\varepsilon$ | ratio | $\varepsilon$ | ratio | $\varepsilon$ | ratio |
| $\overline{B}^0 \to D^+ \tau(\to \mu \overline{\nu}_\mu \nu_\tau) \overline{\nu}_\tau$ | $\pm 0.87\%$ | $\pm 0.76\%$ | $\pm 0.29\%$ | $\pm 0.371\%$ | $\pm 0.38\%$ | $\pm 0.435\%$ |
| $\overline{B}^0 \to D^+ \mu \overline{\nu}_\mu$ | $\pm 0.11\%$ | | $\pm 0.08\%$ | | $\pm 0.05\%$ | |
| $\overline{B}^0 \to D^{*+}(\to D^+ \pi^0) \tau(\to \mu \overline{\nu}_\mu \nu_\tau) \overline{\nu}_\tau$ | $\pm 0.75\%$ | $\pm 0.981\%$ | $\pm 0.2\%$ | $\pm 0.586\%$ | $\pm 0.29\%$ | $\pm 0.652\%$ |
| $\overline{B}^0 \to D^{*+}(\to D^+ \pi^0) \mu \overline{\nu}_\mu$ | $\pm 0.23\%$ | | $\pm 0.39\%$ | | $\pm 0.36\%$ | |

Table 62: Systematic uncertainty on the efficiencies and efficiency ratios due to the emulation of the $\mu$ PID selections.

## 16.8 Neutral isolation emulation

The effect of the neutral isolation selection on the fit variables is evaluated by inspecting the shape of some key variables on a cocktail of Full MC events, as described in Sect. 10.2, with the assumption of the cocktail covering the relevant phase space of the analysis. To propagate a systematic uncertainty related to this assumption, alternative emulation weights are obtained by training the neutral emulation tool on a different sub-set of events. The effect of this change has been observed to be very mild on the shape of the templates, and therefore no additive systematic uncertainty has been propagated on this regard. The multiplicative uncertainty is evaluated as the effect of this change on the ratio of efficiencies.

The effect of the variation on the efficiencies and efficiency ratios is reported in Table 63.

## 16.9 Tracking efficiency correction

Various systematic uncertainties affect the tracking efficiency correction and to estimate their impact on the template shapes and on the efficiency ratios, two alternative sets of tracking efficiency correction weights have been generated, corresponding to the minimum and the maximum value of the per-event weights obtained when taking into account these uncertainties.

First, the statistical uncertainty of the correction tables is propagated on the per-event weights by means of a bootstrapping technique. The correction tables are bootstrapped according to their bin content and error, using a gaussian approximation, generating 100 different tables.

| | 2015 | | 2016 | | comb | |
|---|---|---|---|---|---|---|
| Sample | $\varepsilon$ | ratio | $\varepsilon$ | ratio | $\varepsilon$ | ratio |
| $\overline{B}^0 \to D^+\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | $\pm 0.73\%$ | $\pm 0.032\%$ | $\pm 0.73\%$ | $\pm 0.061\%$ | $\pm 0.73\%$ | $\pm 0.056\%$ |
| $\overline{B}^0 \to D^+\mu\overline{\nu}_\mu$ | $\pm 0.7\%$ | | $\pm 0.67\%$ | | $\pm 0.68\%$ | |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | $\pm 2.5\%$ | $\pm 0.153\%$ | $\pm 2.68\%$ | $\pm 0.202\%$ | $\pm 2.65\%$ | $\pm 0.194\%$ |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\overline{\nu}_\mu$ | $\pm 2.35\%$ | | $\pm 2.48\%$ | | $\pm 2.46\%$ | |

Table 63: Effect of the systematic variation of the neutral isolation weights on the efficiency and efficiency ratios.

With each of these tables the per-event weights are evaluated again, and the standard deviation of the 100 estimates is then used as an estimate of their statistical uncertainty.

The tracking efficiencies have been corrected by using the tracking efficiencies evaluated in data and in MC as a function of momentum and pseudorapidity of the tracks, after having reweighted the MC samples in the occupancy of the detector ($nTracks$) to agree with data.

As reported in [96], when performing the reweighting of the simulated samples in different parameters such as the number of tracks in the event or the number of primary vertices or hits in various subdetectors, small differences on the ratio of efficiencies have been observed. This is taken into account by assigning a 0.8% systematic uncertainty to the per track correction weight, which is then added up 4 times in the per event weight.

The efficiency ratio is evaluated for muon tracks which are required to reach the muon stations, but three out of 4 tracks of the event are coming from charged hadrons, which experience hadronic interaction with the material detector. This leads to a higher chance for hadronic tracks to have experience large scattering angles due to the interaction with the material. The hadronic interaction cross-section depends on the hadron specie, as well as on its kinematic and charge. As shown in [96], it has been studied, using simulated events, that about 11% of the kaons and about 14% of the pions cannot be reconstructed due to the hadronic interactions that occur before the last tracking stations. The probability of this to happen also highly depends on the material budget encountered from the hadronic track. If this is not well reproduced in simulation, the efficiency ratio estimation would be affected by an intrinsic systematic uncertainty which is not accounted for when evaluating it with muon tracks. Assuming a total material budget uncertainty in the simulation of around 10%, this leads to a systematic uncertainty for the tracking efficiency ratio, for hadron tracks, of $1.1 - 1.4\%$, depending on the hadron specie. This is taken into account in the systematic uncertainty of the per-event weight by adding the relative systematic uncertainties in quadrature, taking into account the number of signal pions and kaons in the event.

All the uncertainties for the per-event weight, in absolute value, are then added together in quadrature. Three weights, for each event, are finally evaluated, corresponding to the nominal correction and the $\pm 1\sigma$ variations.

The weights for each event are reported in Fig. 121 and 122, along with their effect on the

templates for the two signal and normalization modes, in the ISO fit region. No effect is seen on the shape of the templates. For this reason, the shape of the templates is not corrected for this effect and no additional shape systematic is assigned in the fit.



(a) $B \to D^+ \mu\nu$



(b) $B \to D^+ \tau\nu$



(c) $B \to D^{*+} \mu\nu$



(d) $B \to D^{*+} \tau\nu$
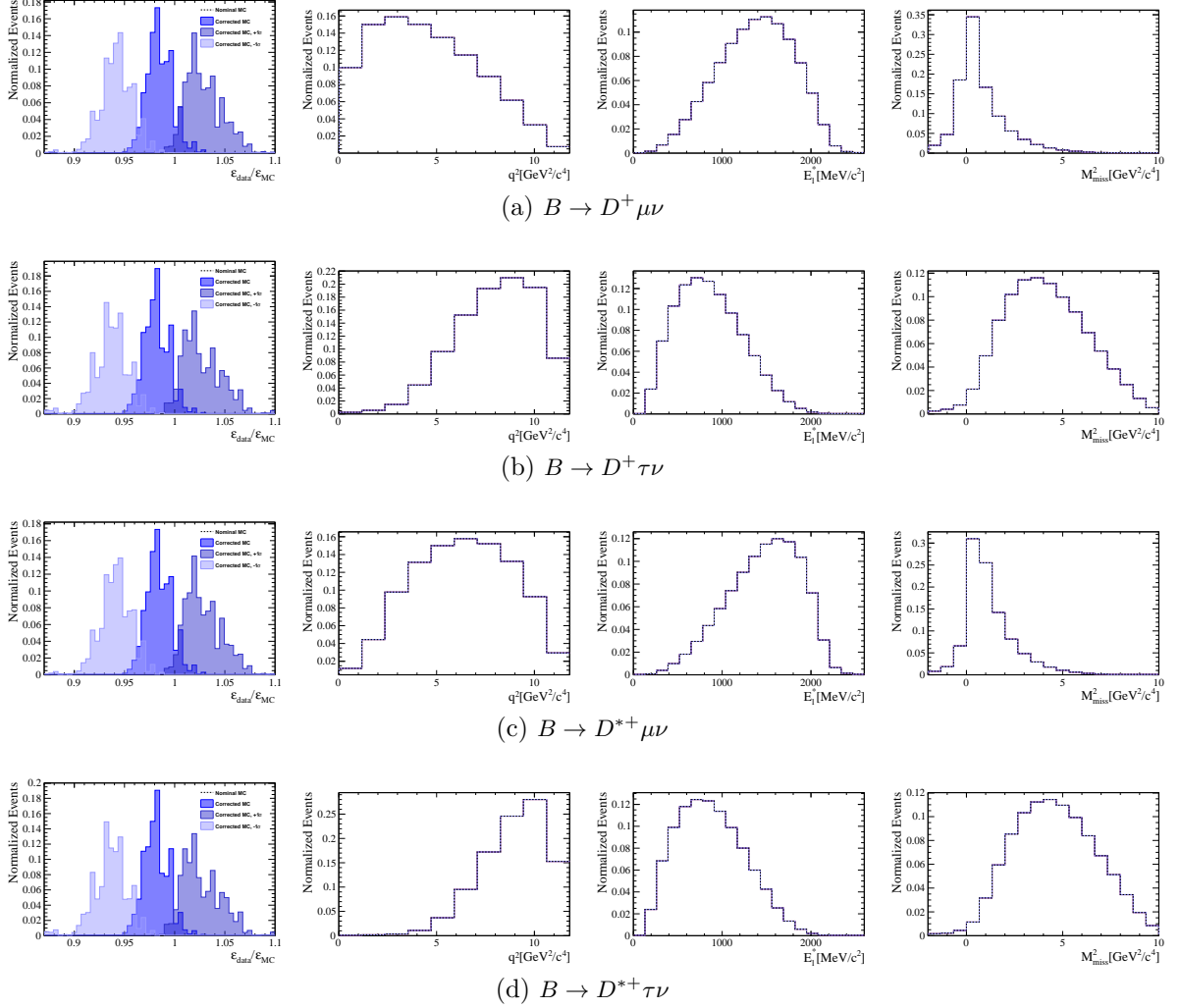
Figure 121: Effect of the Tracking efficiency correction on the two signal and normalization modes 2015 templates, ISO sample

The efficiencies and efficiency ratios are re-evaluated using the alternative tracking correction weights, as defined above, to estimate the systematic uncertainty associated to the correction on them. The result is reported in Table 64, which reports the relative change of the efficiencies and the ratios of efficiencies, with respect to the nominal set of weights, when using the alternative weights.

## 16.10   Further systematic uncertainty

At the time of the writing of this thesis, the systematic uncertainties that are considered to be dominant have been evaluated. Further systematic uncertainties are now under scrutiny, but
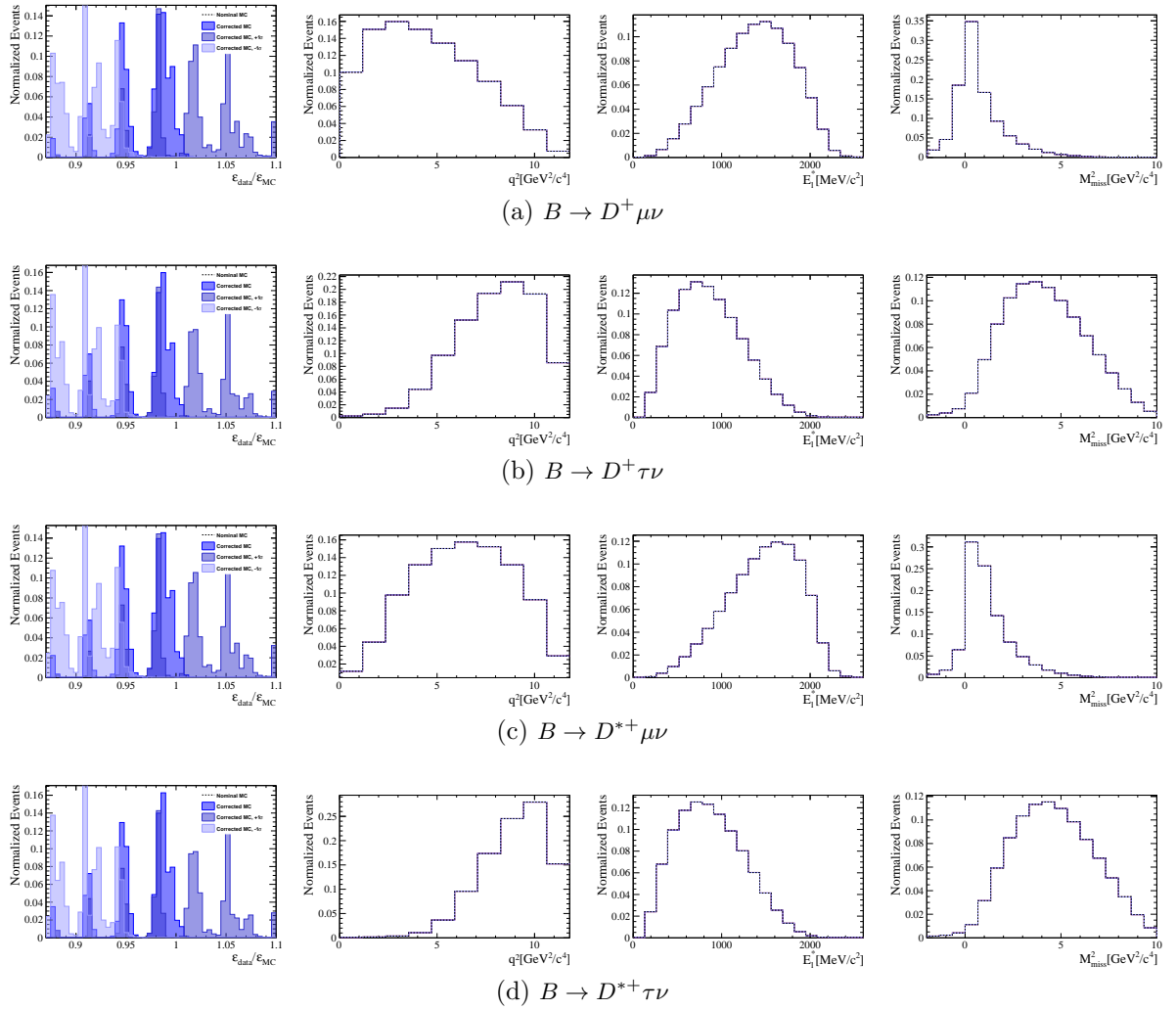
Figure 122: Effect of the Tracking efficiency correction on the two signal and normalization modes 2016 templates, ISO sample

these are thought to be subleading with respect to the systematic uncertainties quoted in the text. Their description and the strategy for their evaluation is reported in the next sections

### 16.10.1 Combinatorial background shape

The normalization of the $D\mu$-combinatorial background contribution is already left floating freely in the fit, therefore only systematics affecting the shape of the relative template will have to be taken into account. The template of this contribution relies on WS $D^+\mu^+$ candidates in data. The assumption is that this sample is a god proxy for the $D\mu$-combinatorial background. This assumption has been checked explicitly in the high $B^0$-mass region and no significant discrepancy has been observed. To evaluate a residual systematic on the shape of the template, a correction will be applied on the residual difference between the WS and RS $D\mu$-combinatorial template in this control region as a function of the most discrepant variables. The systematic can

| Sample | 2015 | | 2016 | | comb. | |
|---|---|---|---|---|---|---|
| | $\varepsilon$ | ratio | $\varepsilon$ | ratio | $\varepsilon$ | ratio |
| $\overline{B}^0 \to D^+\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | $\pm 4.59\%$ | $^{+0.197\%}_{-0.215\%}$ | $\pm 6.18\%$ | $^{-0.04\%}_{+0.045\%}$ | $\pm 5.91\%$ | $^{+0.003\%}_{-0.004\%}$ |
| $\overline{B}^0 \to D^+\mu\overline{\nu}_\mu$ | $\pm 4.38\%$ | | $\pm 6.22\%$ | | $\pm 5.91\%$ | |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\tau(\to \mu\overline{\nu}_\mu\nu_\tau)\overline{\nu}_\tau$ | $\pm 4.64\%$ | $^{+0.207\%}_{-0.226\%}$ | $\pm 6.22\%$ | $^{-0.013\%}_{+0.014\%}$ | $\pm 5.95\%$ | $^{+0.023\%}_{-0.026\%}$ |
| $\overline{B}^0 \to D^{*+}(\to D^+\pi^0)\mu\overline{\nu}_\mu$ | $\pm 4.43\%$ | | $\pm 6.23\%$ | | $\pm 5.93\%$ | |

Table 64: Systematic uncertainty on the efficiency and efficiency ratios associated to the variations of the tracking efficiency corrections, as defined in the text, with respect to the nominal correction.

be evaluated as the difference between the measured parameters with and without this correction applied.

### 16.10.2   Double Charm

The overall size of the Double-Charm background is floated in the fit independently in each region. Its shape parameters are also floating in the fit, with the highest sensitivity to them coming from the $DD$ region, which is enriched in decays containing one additional anti-isolated kaon. This can induce some biases, since the underlying assumption is that the shape of this contribution is the same between the isolated region and the anti-isolated $DD$ region used for calibration.

In order to study the effect of a potentially biased extrapolation, the $m_{D^+X_c}$ mass variable (used to correct the shape of the multi-body DD) will be compared in MC between the signal and DD regions. A correction can be obtained from that comparison and used to perform an alternative fit, that can be compared with the nominal one in order to estimate the associated systematic uncertainty.

### 16.10.3   Data/model agreement

The agreement of the data and the model is assessed by comparing the distributions of the $B \to D\mu\nu$ at low missing mass in a set of key variables, and applying a reweight on them, as described in Sect. 13.4. This reweight aims at correcting mis-alignments on the trigger response and in the simulation of the detector that remain after the corrections described before. To assign a systematic uncertainty to this reweight, the fit values and efficiencies corresponding to consecutive steps of an iterative procedure of fitting-reweighting could be compared.

# 17 Validation of the fit on pseudo data

The fit has been validated using fits to pseudodata that are used to study the presence of any bias and to validate the correct coverage of the error estimates. The pseudo data has been generated using the same model that has been used for fitting, simulating data in all the signal and control regions being used in the simultaneous fit to real data (ISO, $\pi^0$, $1\pi$, $2\pi$, DD). The statistics generated corresponds to the one observed on the combined 2015+2016 data sample.

In the following sections the strategy employed in the generation of the pseudodata, the convergency criterion and the treatment of constrained parameters will be described. Finally, the result of this study will be shown.

## 17.1 Generation of the toy dataset

The pseudodata is generated from the same model used in the fit, with all the values of the nuisance parameters being set to the best fit point. An exception is made for the parameters which are constrained in the fit, for which the generation value has been set to the central value of the constraint *pdf*. The value of the $\mathcal{R}(D^+)$ and $\mathcal{R}(D^*)$ parameters has been set to the expected World Average [1].

The generation of the toy dataset is done in the following way. First the model is compiled, with all the parameters set at their true value. Then the value of the expected number of events in each bin is read. The model is created from the same templates which are used to perform the fit to data. These templates have been corrected for the presence of empty and negative-entries bins, as described in Section 14. As a consequence, no empty bins will be present in the distribution of expected number of events generated. To correct this, bins with a number of events not bigger than a threshold, which is chosen to select bins in which the aforementioned corrections have been applied, are being reset to a zero number of entries.

The random fluctuation of the data is emulated by randomizing the number of events in each bin. In order to reproduce as closely as possible the fluctuation one expects in real data, a specific strategy must be implemented. In fact, the number of expected events in each bin is not expected to follow a simple Poisson distribution. This is due to the fact that the number of events in each bin is the result of a sum of weighted events, with the weights themselves being random variables themselves. As it has been described in Appendix A, the effect this has on the estimator errors has been implemented with the so-called SPD method. In this method, the sum of random weights in each bin is approximately thought to be distributed as a Scaled Poisson Distribution. In other terms, the Likelihood that is maximized is expressed in terms of the equivalent number of events in each bin $(i)$ $n_i' = \sum_j w_{ij}/s_i$, which is assumed to be distributed as a Poisson with expectation value $\mu_i' = \mu_i/s_i$, where $s_i = \sum_i w_{ij}^2 / \sum_i w_{ij}$ is a bin-by-bin correction factor.

In order to have a toy dataset in which the expected number of events in each bin is distributed as a Scaled Poisson distribution, the expected number of events in each bin $\mu_i$ is first multiplied with the same correction factor used in data, $1/s_i$. Then a random number is generated in each bin according to a Poisson distribution, with expectation value $\mu_i/s_i$. And finally this random number is multiplied again by the inverse of the correction factor, $s_i$. The toy dataset obtained in this way is then used in the same way as in the fit to real data.

## 17.2 Convergency criterion

After the fit is performed, the covariance matrix as reported by `Hesse`, is inspected. The fit is taken into account only if the covariance matrix is declared positive defined. The result for the fits which do not pass this requirement are just visualized and are not used to compute the relative pulls.

## 17.3 Treatment of the constrained parameters

Special treatment is dedicated to parameters which are constrained in the fit. The procedure followed is the one described in [102]. The pseudo data is always generated with the true value of the parameters fixed at the central value of the constraint used in the fit to data and the pulls are evaluated with respect to this value. Before each toy fit, the central value of the constraint is randomized according to the constraint *pdf* itself. This can be understood if the constraint being put in the likelihood function is read as the prior *pdf* of the parameter under consideration. This prior encodes all the knowledge that is coming from a previous measurement of the parameter. Each iteration of the toy should emulate the remaking of the whole measurement, and therefore also of the previous measurements providing the constraint. The latter is achieved by randomizing the central value of the prior *pdf* being used in the current analysis.

Being more specific, as an example for a parameter $\mu$, with true value $\mu_{true}$, which is included in the likelihood with a gaussian constraint centered at $\mu_c$ and with width parameter $\sigma_e$ (e.g. coming from an external measurement which reported $\mu = \mu_e \pm \sigma_e$), the steps are the following:

1. The pseudodata is generated with $\mu = \mu_{true}$;

2. A new constraint value, $\mu_c$, is generated according to a gaussian *pdf* with central value and width parameters $\mu_e$ and $\sigma_e$ respectively;

3. A gaussian constraint, with central value and width parameters of $\mu_c$ and $\sigma_c = \sigma_e$ respectively, is inserted in the likelihood;

4. The likelihood is maximized.

The pulls are evaluated with the following formula,

$$p = \frac{\mu - \mu_{true}}{\sigma} \tag{199}$$

and the distribution of this variable for each parameter in the measurement is fit with a Gaussian *pdf* with central value and width free floating. In the case of an unbiassed estimator with the correct error coverage, this distribution is expected to be a unit gaussian.

## 17.4 Results

In this section, the result of the fit to the pulls distribution of each parameter in the fit is reported. In each figure the resulting gaussian distribution is reported along with the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties. In each plot a unit gaussian is overalaid with a red dashed line, for a more visual comparison with the ideal case.

In the figures, the number of toys in which the fit did not pass the convergency criterion test are reported with a hashed red histogram. These toys are put in the figure for illustration

purposes only and are not being used in order to evaluate the statistics of the pull distributions. The fraction of these cases is very low, and therefore the reason why this happens is not being investigated further. This could also be expected due to simple numerical errors in the fit.

No big bias is being observed on the parameters of interest, which validates our fitting strategy.

Few nuisance parameters show a non ideal pulls distribution. For the $\mathcal{R}(D^{**})$ and $\frac{\mathcal{B}(D_s \to \tau\nu)}{\mathcal{B}(D_s \to \mu\nu X)}$ an overcoverage has been observed. It has been noticed that this feature disappears if the number of expected events in the $B \to D^{**}\tau\nu$ and $B \to (D_s \to \tau\nu)DX$ samples is increased by a factor 10. Furthermore, the parameter denoted as $f(DD, B^{\pm})$ $(ISO)$, which controls the fraction of Two Body Double Charm decays from a $B^{\pm}$ meson in the $ISO$ region, has been observed to have a non-gaussian pull. This has to be attributed to the fact that this parameter has a big uncertainty and it is constrained to assume values in the range of $[0, 1]$. Therefore, its true value lies too near to the upper boundary, and the fit on it will be distorted. When fixing this parameter to its fitted value, the small bias observed on the $\mathcal{R}_{raw}(D^{*})$ parameter is found to be consistent with zero.

All the other parameters show a good unit-gaussian pull. Given that the non-ideal cases are small and located only among nuisance parameters, no further action is taken, and the fit can be considered stable and to deliver an uncertainty with the correct coverage.

### 17.4.1 Signal and normalization
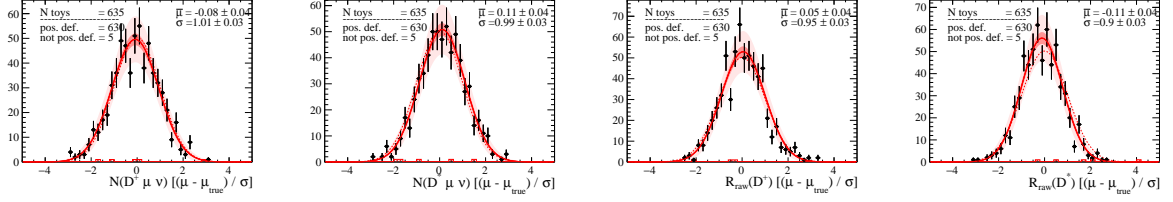


Figure 123: Pulls for the signal and normalization yields and for the $\mathcal{R}(D^{(+,*)})$ parameters.

| parameter | true value | uncertainty | rel. uncertainty |
|---|---|---|---|
| $N(D^+\mu\nu)$ | 1.43e+06 | 4.0e+03 | 0.28% |
| $N(D^*\mu\nu)$ | 7.19e+05 | 6.0e+03 | 0.83% |
| $\mathcal{R}_{raw}(D^+)$ | 3.50e-02 | 4.7e-03 | $1.4e+01\%$ |
| $\mathcal{R}_{raw}(D^*)$ | 3.09e-02 | 8.0e-03 | $2.6e+01\%$ |

Table 65: Errors on signal and normalization yields and for the $\mathcal{R}(D^{(+,*)})$ parameters.
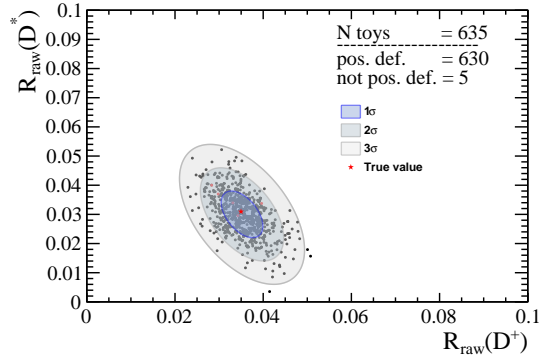


Figure 124: Scatter plot of the results of the fit to the toy datasets. The ellipses correspond to the $\pm 1, 2, 3\sigma$ contours as reported from the fit. The black dots represent the fit instances that passed the convergency criterion, the red dots the ones that did not. The red star corresponds to the true value of the parameters, defined during generation.

212

## 17.4.2 Double Charm
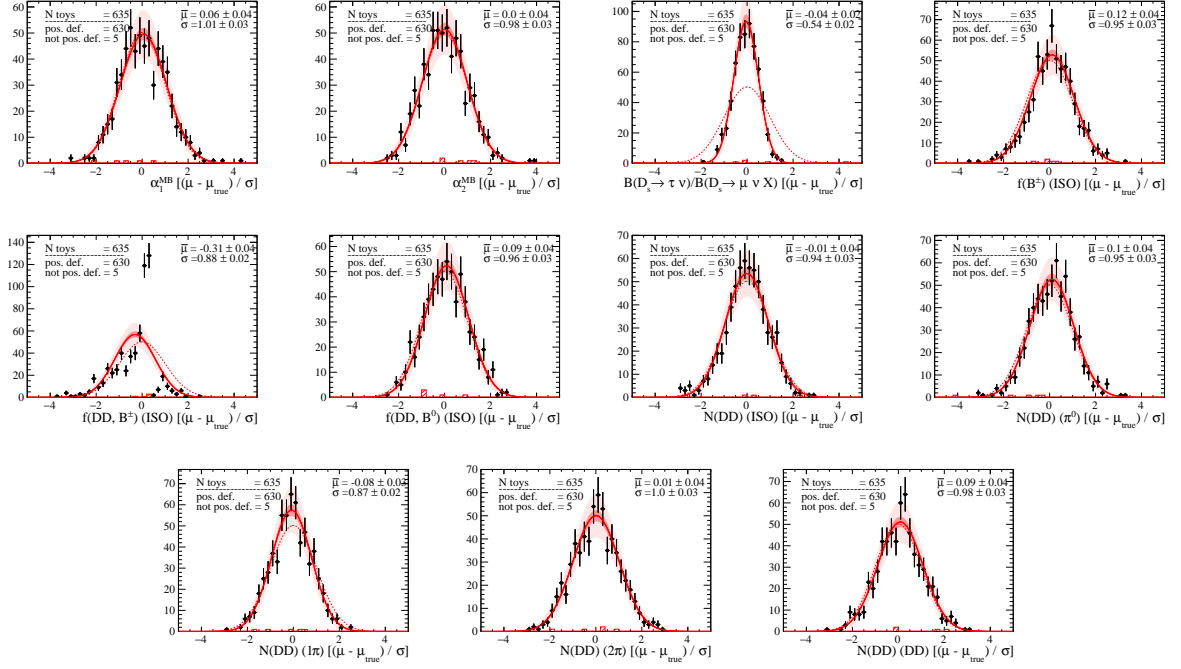


Figure 125: Pulls for the $B \to DDX$ parameters.

| parameter | true value | uncertainty | rel. uncertainty |
|:---:|:---:|:---:|:---:|
| $\alpha_1^{MB}$ | -7.95e-01 | 5.6e-02 | 7.0% |
| $\alpha_2^{MB}$ | 3.81e-01 | 3.5e-02 | 9.2% |
| $\mathcal{B}(D_s \to \tau\nu)/\mathcal{B}(D_s \to \mu\nu X)$ | 1.42e-01 | 3.5e-02 | $2.5e+01\%$ |
| $f(B^{\pm})$ (ISO) | 4.12e-02 | 2.5e-02 | $6.1e+01\%$ |
| $f(DD, B^{\pm})$ (ISO) | 8.28e-01 | 5.0e-01 | $6.1e+01\%$ |
| $f(DD, B^0)$ (ISO) | 5.71e-01 | 1.0e-02 | 1.8% |
| $N(DD)$ (ISO) | 1.72e+05 | 5.1e+03 | 3.0% |
| $N(DD)(\pi^0)$ | 1.00e+04 | 4.4e+02 | 4.4% |
| $N(DD)$ ($1\pi$) | 1.49e+04 | 1.2e+03 | 8.3% |
| $N(DD)$ ($2\pi$) | 3.08e+03 | 3.1e+02 | $1e+01\%$ |
| $N(DD)$ (DD) | 6.56e+04 | 1.0e+03 | 1.6% |
| $N(D_J^{**})$ (ISO) | 1.03e+05 | 9.9e+03 | 9.6% |
| $N(D_J^{**})$ ($1\pi$) | 1.76e+04 | 3.1e+03 | $1.7e+01\%$ |
| $N(D_J^{**})$ ($2\pi$) | 5.81e+03 | 4.6e+02 | 8.0% |
| $N(D_J^{**})$ (DD) | 2.21e+03 | 7.3e+02 | $3.3e+01\%$ |

Table 66: Errors on the $B \to DDX$ parameters.

## 17.4.3 Feed Down
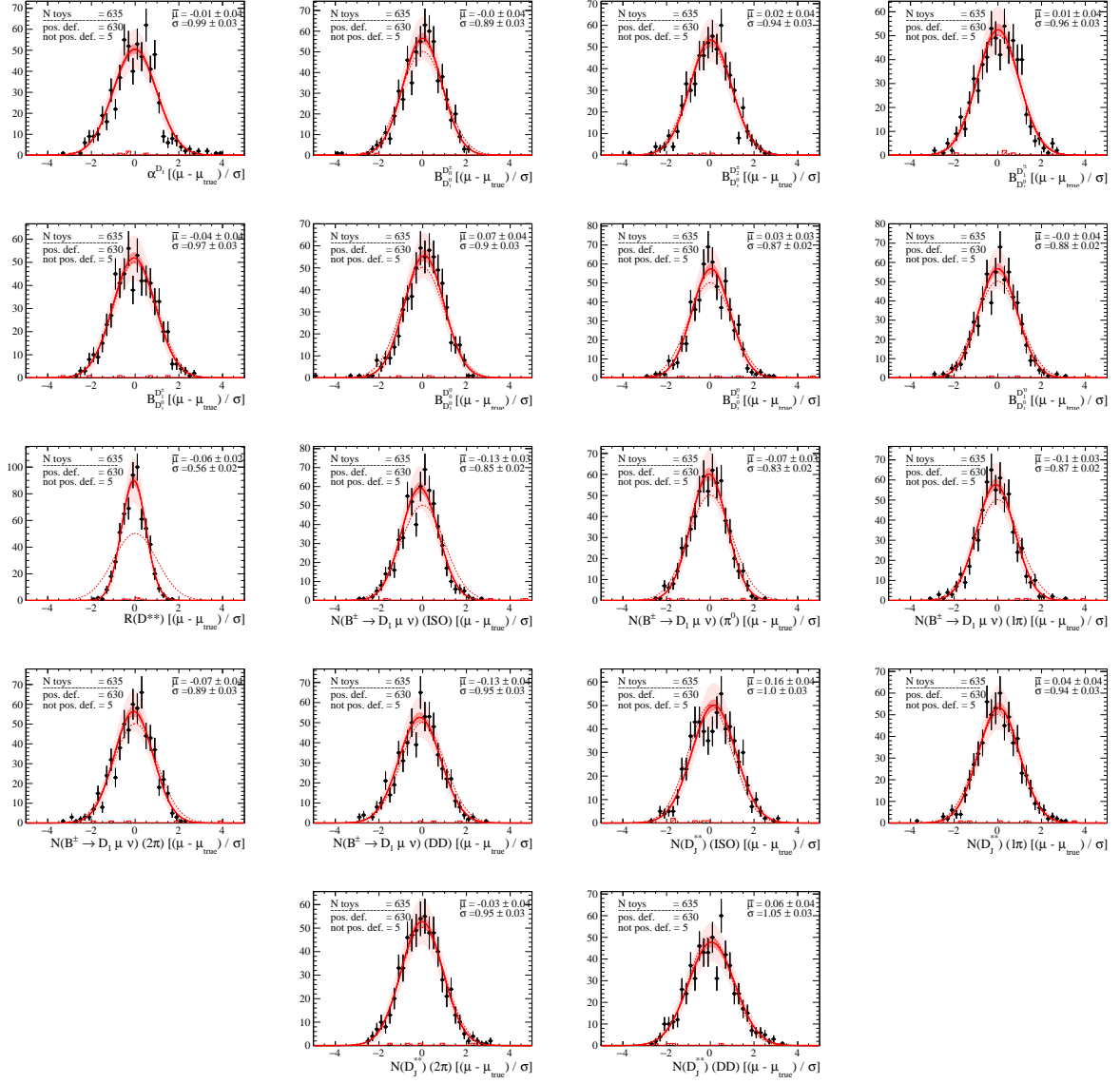


Figure 126: Pulls for the $B \to D^{**}\mu\nu X$ parameters.

| parameter | true value | uncertainty | rel. uncertainty |
|:---:|:---:|:---:|:---:|
| $\alpha^{D_J}$ | -8.81e-01 | 2.4e-01 | $2.7e+01\%$ |
| $\mathcal{B}(B^0 \to D_0\mu\nu)/\mathcal{B}(B^\pm \to D_1\mu\nu)$ | 1.53e+00 | 2.1e-01 | $1.3e+01\%$ |
| $\mathcal{B}(B^0 \to D_2\mu\nu)/\mathcal{B}(B^\pm \to D_1\mu\nu)$ | 7.30e-01 | 1.2e-01 | $1.6e+01\%$ |
| $\mathcal{B}(B^0 \to D'_1\mu\nu)/\mathcal{B}(B^\pm \to D_1\mu\nu)$ | 5.12e-01 | 9.0e-02 | $1.8e+01\%$ |
| $\mathcal{B}(B^0 \to D_1\mu\nu)/\mathcal{B}(B^\pm \to D_1\mu\nu)$ | 4.62e-01 | 5.1e-02 | $1.1e+01\%$ |
| $\mathcal{B}(B^\pm \to D_0\mu\nu)/\mathcal{B}(B^\pm \to D_1\mu\nu)$ | 2.55e+00 | 1.9e-01 | 7.3% |
| $\mathcal{B}(B^\pm \to D_2\mu\nu)/\mathcal{B}(B^\pm \to D_1\mu\nu)$ | 1.90e+00 | 1.5e-01 | 7.7% |
| $\mathcal{B}(B^\pm \to D'_1\mu\nu)/\mathcal{B}(B^\pm \to D_1\mu\nu)$ | 8.91e-01 | 1.3e-01 | $1.5e+01\%$ |
| $\mathcal{R}(D^{**})$ | 1.48e-02 | 6.0e-03 | $4.1e+01\%$ |
| $N(B^\pm \to D_1\mu\nu)$ (ISO) | 5.35e+04 | 4.1e+03 | 7.6% |
| $N(B^\pm \to D_1\mu\nu)$ ($\pi^0$) | 3.48e+03 | 2.9e+02 | 8.2% |
| $N(B^\pm \to D_1\mu\nu)$ ($1\pi$) | 3.47e+04 | 2.3e+03 | 6.6% |
| $N(B^\pm \to D_1\mu\nu)$ ($2\pi$) | 7.04e+01 | 8.1e+00 | $1.1e+01\%$ |
| $N(B^\pm \to D_1\mu\nu)$ (DD) | 1.27e+03 | 1.1e+02 | 8.7% |

Table 67: Errors on the $B \to D^{**}\mu\nu X$ parameters.

## 17.4.4 $B \to D^{(*)}\ell\nu$ form factor parameters
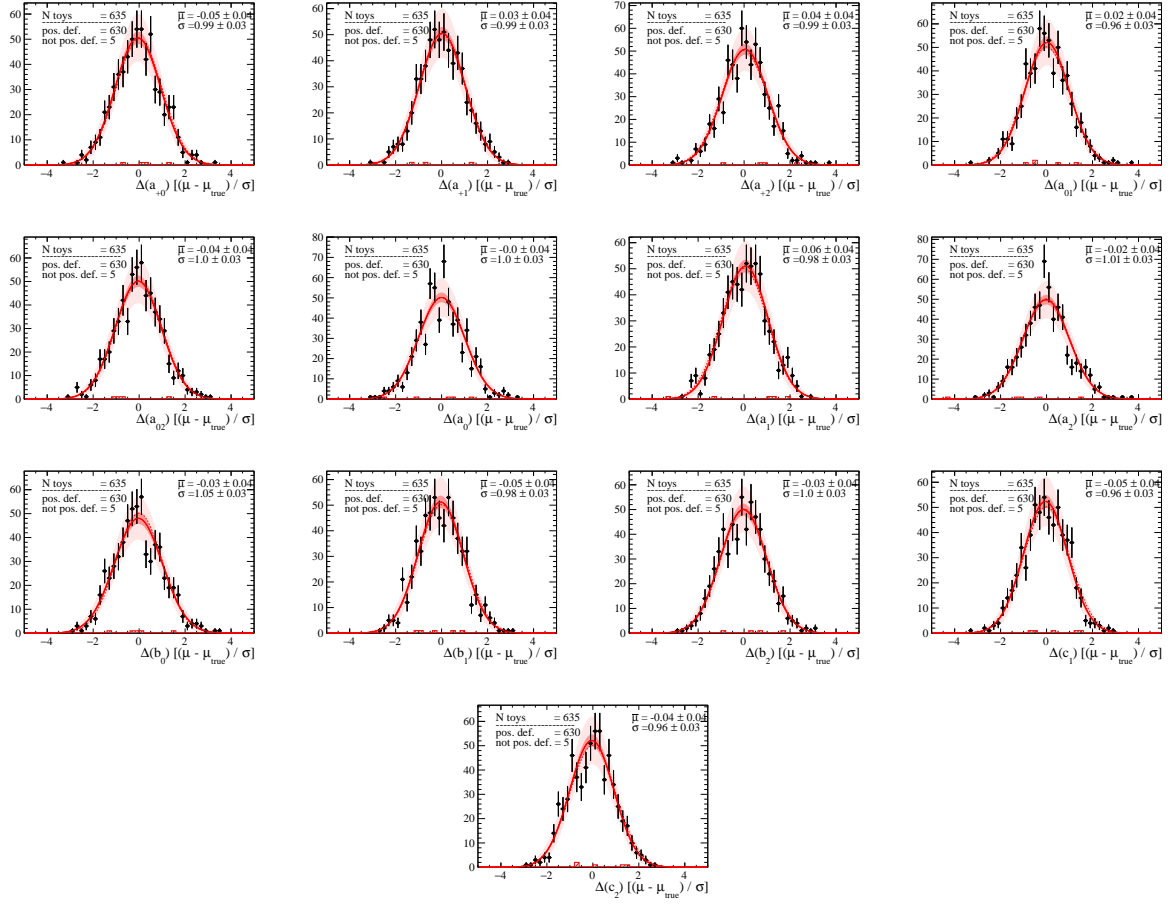


Figure 127: Pulls for the $B \to D^{(*)}\ell\nu$ form factor parameters.

| parameter | true value | uncertainty | rel. uncertainty |
|:---:|:---:|:---:|:---:|
| $\tilde{a}_{+0}$ | 1.49e-02 | 9.5e-05 | 0.63% |
| $\tilde{a}_{+1}$ | 3.02e-02 | 9.6e-04 | 3.2% |
| $\tilde{a}_{+2}$ | -2.12e-01 | 8.1e-03 | 3.8% |
| $\tilde{a}_{01}$ | -8.59e-02 | 2.1e-02 | $2.5e+01$% |
| $\tilde{a}_{02}$ | -2.26e-01 | 9.5e-02 | $4.2e+01$% |

| parameter | true value | uncertainty | rel. uncertainty |
|:---:|:---:|:---:|:---:|
| $\tilde{a}_0$ | 2.06e-02 | 4.5e-06 | 0.022% |
| $\tilde{a}_1$ | 9.08e-04 | 1.6e-05 | 1.7% |
| $\tilde{a}_2$ | 8.08e-04 | 1.7e-05 | 2.1% |
| $\tilde{b}_0$ | -1.18e-03 | 5.7e-05 | 4.9% |
| $\tilde{b}_1$ | -4.03e-04 | 1.4e-04 | $3.4e+01$% |
| $\tilde{b}_2$ | -8.64e-03 | 1.1e-04 | 1.3% |
| $\tilde{c}_1$ | 1.31e-02 | 9.4e-04 | 7.2% |
| $\tilde{c}_2$ | 1.93e-03 | 1.5e-03 | $7.8e+01$% |
| $\tilde{d}_0$ | -8.12e-03 | 5.3e-03 | $6.6e+01$% |
| $\tilde{d}_1$ | -1.95e-03 | 4.6e-03 | $2.3e+02$% |

Table 68: Errors on the $B \to D^{(*)}\ell\nu$ form factors parameters.

## 17.4.5 $B \to D^{**}\ell\nu$ form factor parameters
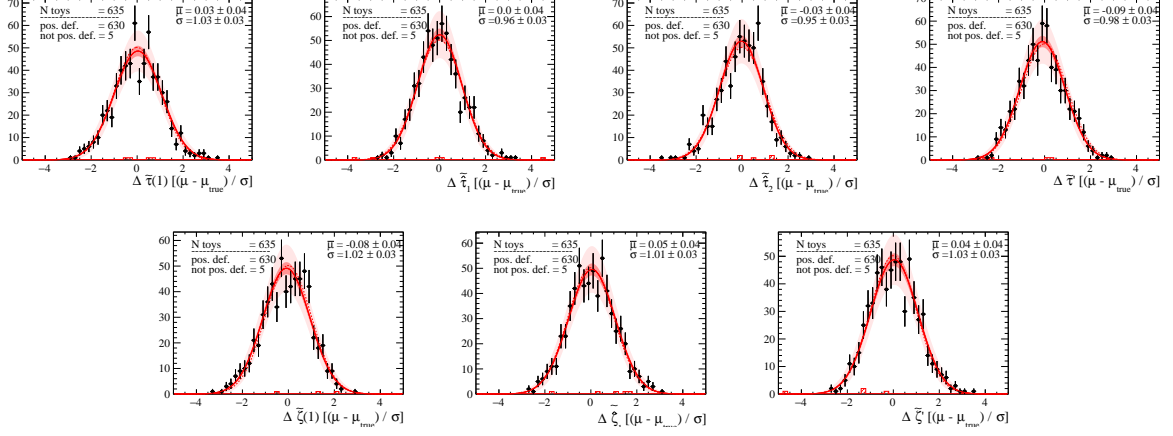


Figure 128: Pulls for the $B \to D^{**}\ell\nu$ form factors parameters.

| parameter | true value | uncertainty | rel. uncertainty |
|:---:|:---:|:---:|:---:|
| $\tilde{\tau}(1)$ | 6.67e-01 | 2.1e-02 | 3.2% |
| $\tilde{\tilde{\tau}}_1$ | 1.08e+00 | 1.2e-01 | $1.1e+01\%$ |
| $\tilde{\tilde{\tau}}_2$ | 3.18e+00 | 2.7e-01 | 8.4% |
| $\tilde{\tau}'$ | 9.88e-02 | 7.8e-02 | $7.8e+01\%$ |
| $\tilde{\zeta}(1)$ | -4.56e-03 | 3.0e-01 | $6.5e+03\%$ |
| $\tilde{\tilde{\zeta}}_1$ | -8.01e-01 | 5.2e-02 | 6.5% |
| $\tilde{\zeta}'$ | 4.98e-01 | 1.2e-01 | $2.4e+01\%$ |

Table 69: Errors on the $B \to D^{**}\ell\nu$ form factors parameters.

# 18 Conclusions

In this thesis, a test of lepton flavour universality with $b \to c\ell\nu$ decays, through a simultaneous measurement of $\mathcal{R}(D^+)$ and $\mathcal{R}(D^{*+})$ has been presented. The analysis utilises events collected by the LHCb experiment during the RunII data taking, during years 2015 and 2016, corresponding to an integrated luminosity of 2 fb$^{-1}$.

At the time of the writing of this thesis, the $\mathcal{R}(D^+)$ has never been measured at a hadron collider. This thesis presented the analysis strategy used in analysing the data and the main systematic uncertainties have been measured. The central value of the parameters of interest is still blind and the analysis is at the moment of the writing of the thesis going through the internal review procedure. The expected result using the systematic uncertainties that have been measured is

$$
\begin{aligned}
\mathcal{R}(D^+) &= xxx \pm 0.033(\text{stat.}) \pm 0.037(\text{syst.}), \\
\mathcal{R}(D^{*+}) &= xxx \pm 0.040(\text{stat.}) \pm 0.070(\text{syst.}),
\end{aligned}
$$

with $-53\%$ correlation.

Assuming the fitted value of the parameters of interest to be the same as the current World Average, it is possible to compare the expected result with the previous measurements in the $\mathcal{R}(D^*)$-$\mathcal{R}(D)$ plane. This is reported in Fig. 129. In the figure both the result including only statistical uncertainties or the full uncertainty are reported.

It can be seen that this measurement could be competitive with the latest semileptonic tagged analysis from the Belle collaboration on the $\mathcal{R}(D)$ direction, whereas the uncertainty on $\mathcal{R}(D^*)$ would be compatible with the hadronic tagged analyses from the Belle and BaBar collaborations in the $\mathcal{R}(D^*)$.

The correlation parameter between the $\mathcal{R}(D^+)$ and $\mathcal{R}(D^*)$ parameter is higher with respect to the previous simultaneous measurements of these parameters. For these reasons this measurement could play a very important role, in combination with the present World Average, in confirming or resolving the observed discrepancies observed in the $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ parameters with respect to the Standard Model predictions when combined with the previous measurements.
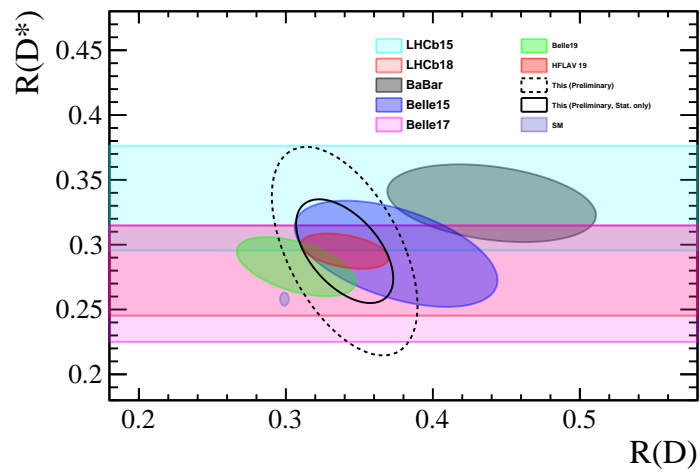
Figure 129: Expected result, depicted with the full and dotted black lines, compared with previous measurements, assuming the fitted values of the parameters of interest are equal to the World Average.

# A  Binned maximum likelihood fits with weighted events

When performing a fit to a binned dataset, one has access to events presented in an histogram, with a given number of events in each bin, $n_i$. The objective is to best reproduce the dataset in each bin with an expected number of events, $\mu_i$. This will depend on some parameters $\theta_i$, and the aim is translated in a parameter estimation problem, in which one has to find the set of parameters $\theta_i$ for which the expected number of events in each bin $\mu_i(\theta)$, best represents the number of observed events in each bin.

This is achieved with a method called the *Binned Maximum Likelihood fit*, whereby the number of events in each bin is assumed to be distributed with a Poisson distribution with expected value $\mu_i(\theta)$. A function, called the Likelihood, corresponding to the joint probability of having observed the binned data $(n)$ given a set of parameters $\theta$, is then constructed

$$L(n, \theta) = \prod_{i \in \text{bins}} \frac{e^{-\mu_i(\theta)} \mu_i(\theta)^{n_i}}{n_i!}, \tag{200}$$

and the maximum likelihood estimators of the parameters $\theta$, denoted as $\hat{\theta}$, are the ones that maximize this function. In the literature the quantity being maximized is given by the natural logarithm of the Likelihood

$$\ln L(n, \theta) = \sum_{i \in \text{bins}} n_i \ln \mu_i(\theta) - \mu_i(\theta), \tag{201}$$

and the maximum likelihood estimates remain the same, being the likelihood function monotonic.

The estimated error on the estimates is taken as the distance between $\hat{\theta}$ and the values of $\theta$ for which the likelihood is decreased by a factor of 0.5 with respect to its maximum value:

$$\ln L(n, \hat{\theta} \pm \hat{\sigma}_{\hat{\theta}}) = \ln L_{\max} - \frac{1}{2}. \tag{202}$$

It can be shown that, in the asymptotic limit where the number of observed events is large, this estimators are unbiassed and efficient, and the estimate of their variance given in Eq 202 is unbiassed.

Sometimes in HEP the data available has undergone a process of weighting, for example to correct for detection efficiencies or to deal with background events present along with the signal under study. In this thesis, for example, the data events that are analyzed are sWeighted to statistically remove the constributions from fake $D \to K\pi\pi$ events. The number of events in each bin, $n_i$, is not anymore a single number, but the sum of the weights observed for each event falling in that bin:

$$n_i = \sum_j w_{ij}, \tag{203}$$

where $w_{ij}$ represents the weight associated to the $j^{th}$ event falling in the $i^{th}$ bin. Since the weights are random variable themselves, the distribution of $n_i$ cannot be anymore assumed to be a Poisson. Therefore the procedure as outlined above would not be applicable anymore to weighted binned datasets. In fact it has been shown that by following the above procedure, one

would obtain an uncertainty on the estimates that does not reproduce well the true uncertainty on the parameters.

In this appendix, the correct procedure to be followed in the case of weighted binned datasets will be derived, as outlined in [97]. The probability density function of the sum of random weights, $x = \sum_i^n w_i$ will be derived. A useful approximation of this distribution will be then outlined and used to evaluate a new Likelihood function that can be used in the case under study.

## A.1  Distribution of a sum of weighted Poisson numbers

Let $m$ be a Poisson distributed random value, with expected value $\lambda$

$$m \sim P_\lambda(m) = \frac{e^{-\lambda}\lambda^m}{m!}. \tag{204}$$

If this random number is scaled by a real valued weight, $x = wm$ the distribution of the new random variable will be given by

$$x \sim W(x) = \frac{e^{-\lambda}\lambda^{x/w}}{(x/w)!}. \tag{205}$$

Let's analyze the cumulants ($k_i(x) = E[x^i]$, $i = 1, 2, 3, \dots$) of this distribution. Due to the homogeneity of the expectation value operator, the cumulants of this distribution are related to the cumulants of the Poisson distributed variable $m$ in this way:

$$k_i(x) = w^i k_i(m). \tag{206}$$

Fully determining the cumulants of a distribution is equivalent to knowing the probability distribution to any wanted precision, without knowing the actual distribution. In fact, the cumulants are the coefficients of the Taylor expansion of the characteristic function of the pdf, and the pdf can be evaluated as the inverse Fourier Transform of the characteristic function. The moments of a distribution can then be expressed as polynomials of the cumulants. The first two moments are given by the mean,

$$\mu(x) = E[x] = k_1(x), \tag{207}$$

and the second one by the variance,

$$\sigma^2(x) = k_2(x) - k_1^2(x) \tag{208}$$

The in the Poisson distribution, all the cumulants have the same value, $k_i(m) = \lambda \ \forall i$.

Let's now analyze the sum of two weighted Poisson numbers, $x = w_1 n_1 + w_2 n_2$, with $n_1$ and $n_2$ being distributed as a Poisson with expected value $\lambda_1$ and $\lambda_2$ respectively, and $w_1$ and $w_2$ being two real valued weights. The cumulants of the distribution of $x$ are then given by

$$k_i(x) = w_1^i \lambda_1 + w_2^i \lambda_2. \tag{209}$$

This can be easily extended to a sum of any dimensionality, $x = \sum_{i=1}^{N} w_i n_i$, with $n_i$ Poisson distributed with expectation value of $\lambda_i$. The cumulants of the distribution of this variable are given by

$$k_j(x) = \sum_{i=1}^{N} w_i^j \lambda_i. \tag{210}$$

In our case, all mean values are equal, $\lambda_i = 1/N$, with N the total number of events. Therefore the cumulants can be written in terms of the sample average of the weights,

$$k_j(x) = \bar{w}^j \tag{211}$$

## A.2 Distribution of the sum of random weights

Up to now the distribution of weighted Poisson variables has been studied, but the case of interest in this case is the sum of individually weighted events, where the weights themselves are random variables. The situation can be thought in this way: the total number of events $n = \sum_i n_i$, $i = 1, 2, \ldots, N$ is extracted from a Poisson with central value $\lambda = \sum_i \lambda_i$, and then the numbers $n_i$ are chosen from a multinomial distribution where the total number of events $n$ is redistributed to the $N$ different weight classes available with probabilities $\varepsilon_i = \lambda_i / \lambda$. In this way a weight $w_i$ is chosen with probability $\varepsilon_i$. If all the probabilities are equal, $\varepsilon_i = 1/N$, then this multinomial process describes a random selection of $n$ weights out of the total $N$ available weights, in order to compute $x = \sum_{i=1}^{n} w_i$. It is easy to show that this can also be thought as a series of $N$ Poisson extractions, and to each extraction a random weight is associated. In fact, by doing the analogy with just two Poisson random numbers, one can show that

$$P_\lambda M_{\lambda_1/\lambda, \lambda_2/\lambda}^n = \frac{e^{-\lambda} \lambda^n}{n!} \frac{n!}{n_1! n_2!} \frac{\lambda_1^{n_1} \lambda_2^{n_2}}{\lambda^{n_1} \lambda^{n_2}} = \frac{e^{-(\lambda_1 + \lambda_2)} \lambda_1^{n_1} \lambda_2^{n_2}}{n_1! n_2!}, \tag{212}$$

where $P_\lambda$ is a Poisson with rate parameter $\lambda$ and $M_{\lambda_1/\lambda, \lambda_2/\lambda}^n$ is a multinomial with probabilities $\lambda_1/\lambda$ and $\lambda_2/\lambda$. The case under study is the case in which all the all the extraction probabilities are the same and equal to $\varepsilon_i = 1/N$. In this case the random variable of which we have to study the distribution is given by

$$x = \sum_{i=1}^{n} w_i, \tag{213}$$

and in the previous section it has been shown that the cumulants of the distribution of this variable are given by the sample averages of powers of the weights. To describe a continuous weight distribution the limit $N \to \infty$ can be taken. The formulae above for the cumulants remain the same, but in this case the sample average should be substituted with the expectation value:

$$k_j(x) = E[w^j]. \tag{214}$$

In particular, the mean value and the variance of the sum of random weights are given by $\mu = E[w]$ and $\sigma^2 = E[w^2]$.

The probability density function described by these cumulants is called *Compound Poisson distribution*.

## A.3 Approximation with a Scaled Poisson distribution

Usually the distribution of the weights is not known in practice, and therefore the Compound Poisson distribution has to be approximated. Using a normal approximation, the sum of the random weights $x = \sum_{i=1}^{n} w_i$ is assumed to be distributed as a gaussian with mean value given by $E[w]$ and variance $E[w^2]$. In [97] it has also been shown that the Compound Poisson distribution is better described by a *Scaled Poisson distribution*, rather than a guassian distribution.

This probability density function is defined to have the first two moments equal to the ones of the Compound Poisson distribution. A scaled mean value is defined

$$\tilde{\mu} = \frac{E[w]^2}{E[w^2]} = \mu \frac{E[w]}{E[w^2]} = \frac{\mu}{s}, \tag{215}$$

where $s$ is a dilution factor defined as $s = \frac{E[w^2]}{E[w]}$. The Scaled Poisson distribution is the probablity density function of a variable defined as $\tilde{x} = s\tilde{n}$, where $\tilde{n}$ is distributed as a Poisson with expected value $\tilde{\mu}$. The first two moments of the Scaled Poisson distribution correspond to the ones of the Compound Poisson distribution. In fact:

$$E[\tilde{x}] = E[s\tilde{n}] = sE[\tilde{n}] = s\frac{\mu}{s} = \mu = E[x] \tag{216}$$

$$\sigma^2[\tilde{x}] = \sigma^2[s\tilde{n}] = s^2\sigma^2[\tilde{n}] = s^2\tilde{\mu}^2 = s^2\frac{\mu^2}{s^2} = \mu^2 = \sigma^2[x] \tag{217}$$

Therefore the Scaled Poisson distribution approximation consists in taking the sum of the random weights, $x = \sum_{i=1}^{n} w_i$, scaling it with the dilution factor $s$ and assuming that $\frac{x}{s}$ is distributed as a Poisson distribution with central value given by $\mu/s$.

## A.4 The SPD method

The approximation above can be easily extended to maximum likelihood fits to binned weighted datasets. In this case the content of each bin ($i^{th}$) is given by the sum of the weights of the events that fall inside that bin

$$n_j = \sum_i w_{ij}, \tag{218}$$

and we have shown above that this quantity follows a Compound Poisson distribution. By approximating the distribution of this quantity with a Scaled Poisson distribution, as described in the previous section, one can write a Likelihood function by substituting in Eq. 201 the quantities $n_i$ and $\mu_i$ with the scaled ones, obtaining

$$\ln L(n, \theta) = \sum_{i \in \text{bins}} n_i' \ln \mu_i'(\theta) - \mu_i'(\theta), \tag{219}$$

where

$$n_i' = n_i/s \tag{220}$$

$$\mu_i' = \mu_i/s, \tag{221}$$

and $s = \frac{E[w_i^2]}{E[w_i]}$. In practice the distribution of the weights in each bin is not known, and therefore the value of $s$ is estimated from observed weights in each bin

$$s_i = \frac{\sum_j w_{ij}^2}{\sum_j w_j}.$$ (222)

It can be noticed that the expression of the Likelihood is the same as the one for unweighted events, but for a different quantity that is given by

$$n_i' = n_i \frac{\sum_j w_{ij}}{\sum_j w_{ij}^2}.$$ (223)

This can be thought as an effective number of Poisson events and has a nice interpretation, as the equivalent number of events that would have the same statistical power of $n_i$.

In fact, if one defines $n_i'$ as the number of Poisson events that has the same relative error as $n_i$,

$$\frac{\sigma_{n_i'}}{n_i'} = \frac{\sigma_{n_i}}{n_i}$$ (224)

$$\frac{\sqrt{n_i'}}{n_i'} = \frac{\sqrt{\sum_j w_{ij}^2}}{\sum_j w_{ij}},$$ (225)

which leads to the same definition as in 223.

# B    Treatment of shape systematic uncertainties in HistFactory

To propagate some of the systematic uncertainties to our result, we extensively make use of the ability of HistFactory to parameterize the shape variations of the templates due to systematic effects. The `HistFactory` implementation, as described in [103], is briefly discussed in this Appendix.

In `HistFactory` the effect of each systematic uncertainty, $i$, is controlled by a single nuisance parameter $\alpha_i$, and its effect on the template of each fit component can be specified by the user. If the value of the nominal template for the component $s$ in each bin $b$ is denoted as $\sigma_{sb}^0$, the effect of the systematic uncertainty on this template is quantified by the definition of alternative templates, $\sigma_{sb}^{\pm,i}$. These should represent the value of the template when the systematic uncertainty source has been moved by, respectively, $\pm 1\sigma$, which in turn are mapped by `HistFactory` to the nuisance parameter to the values $\alpha_i = \pm 1$.

In order to be able to fit for the best value of the nuisance parameters $\alpha_i$, `HistFactory` must be able to interpolate between the nominal and alternative templates, in order to provide a continuous likelihood. This strategy, which defines a continuous templates as a function of $\alpha$, $\sigma_{sb}(\alpha)$, is usually called *template morphing*. In this way the effect of various systematic uncertainties are taken into account explicitly and simultaneously, giving their fully correlated effect on the shape of the likelihood function being used in the fit.

`HistFactory` implements a *bin-by-bin*, or *vertical template morphing*, in the class called `HistoSys`, and various interpolation options are provided. More details can be found in [103], but in general all the interpolations are following the convention for which the nominal template, $\sigma_{sb}^0$ corresponds to the value $\alpha_i = 0$, and the alternative templates provided by the user, $\sigma_{sb}^{\pm,i}$ correspond to the values $\alpha_i = \pm 1$.

In this analysis we make use of the piecewise-linear interpolation strategy. The value of the interpolated template for the component $s$, in the bin $b$, as a function of the set of nuisance parameters $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_N)$, is given by

$$\sigma_{sb}(\vec{\alpha}) = \sigma_{sb}^0 + \sum_{i \in \text{Syst}} I_{\text{lin}}(\alpha_i; \sigma_{sb}^0, \sigma_{sb}^{+,i}, \sigma_{sb}^{-,i}), \tag{226}$$

where Syst is the set of systematic uncertainties which have an effect on the templates shape and

$$I_{\text{lin}}(\alpha; \sigma^0, \sigma^+, \sigma^-) = \begin{cases} \alpha \left(\sigma^+ - \sigma^0\right) & \alpha > 0, \\ \alpha \left(\sigma^0 - \sigma^-\right) & \alpha \leq 0. \end{cases} \tag{227}$$

Constraint terms on the nuisance parameters, in the form of prior p.d.f.s terms in the likelihood functions, are usually put by `HistFactory`, to reflect that the nuisance parameter has been estimated with some uncertainty by an auxiliary measurement. For the cases described in this document, if not stated otherwise, we remove this constraint term, to let the nuisance parameter float freely in the fit.

# References

[1] HFLAV, Y. S. Amhis *et al.*, *Averages of b-hadron, c-hadron, and $\tau$-lepton properties as of 2018*, Eur. Phys. J. C **81** (2021) 226, `arXiv:1909.12524`.

[2] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367.

[3] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[4] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) 579.

[5] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[6] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[7] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145** (1966) 1156.

[8] T. W. B. Kibble, *Symmetry breaking in non-abelian gauge theories*, Phys. Rev. **155** (1967) 1554.

[9] NA62, C. Lazzeroni *et al.*, *Precision Measurement of the Ratio of the Charged Kaon Leptonic Decay Rates*, Phys. Lett. B **719** (2013) 326, `arXiv:1212.4012`.

[10] KLOE, F. Ambrosino *et al.*, *Precise measurement of $\Gamma(K \to e\nu(\gamma))/\Gamma(K \to \mu\nu(\gamma))$ and study of $K \to e\nu\gamma$*, Eur. Phys. J. C **64** (2009) 627, `arXiv:0907.3594`, [Erratum: Eur.Phys.J. 65, 703 (2010)].

[11] G. Czapek *et al.*, *Branching ratio for the rare pion decay into positron and neutrino*, Phys. Rev. Lett. **70** (1993) 17.

[12] D. I. Britton *et al.*, *Measurement of the $pi^+ \to e^+\nu$ branching ratio*, Phys. Rev. Lett. **68** (1992) 3000.

[13] D. A. Bryman *et al.*, *Measurement of the $\pi \to \nu_e$ Branching Ratio*, Phys. Rev. D **33** (1986) 1211.

[14] FlaviaNet Working Group on Kaon Decays, M. Antonelli *et al.*, *An Evaluation of $|V_{us}|$ and precise tests of the Standard Model from world data on leptonic and semileptonic kaon decays*, Eur. Phys. J. C **69** (2010) 399, `arXiv:1005.2323`.

[15] A. Pich, *Precision Tau Physics*, Prog. Part. Nucl. Phys. **75** (2014) 41, `arXiv:1310.7922`.

[16] ALEPH, DELPHI, L3, OPAL, LEP Electroweak, S. Schael *et al.*, *Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP*, Phys. Rept. **532** (2013) 119, `arXiv:1302.3415`.

[17] ATLAS, G. Aad *et al.*, *Test of the universality of $\tau$ and $\mu$ lepton couplings in W-boson decays from $t\bar{t}$ events with the ATLAS detector*, `arXiv:2007.14040`.

[18] ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group, S. Schael *et al.*, *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257, `arXiv:hep-ex/0509008`.

[19] Belle, A. Abashian *et al.*, *The Belle Detector*, Nucl. Instrum. Meth. A **479** (2002) 117.

[20] BaBar, B. Aubert *et al.*, *The BaBar detector*, Nucl. Instrum. Meth. A **479** (2002) 1, `arXiv:hep-ex/0105044`.

[21] M. Bordone, G. Isidori, and A. Pattori, *On the Standard Model predictions for $R_K$ and $R_{K^*}$*, Eur. Phys. J. C **76** (2016) 440, `arXiv:1605.07633`.

[22] G. Hiller and F. Krger, *More model-independent analysis of $b \to s$ processes*, Physical Review D **69** (2004) .

[23] BaBar, J. P. Lees *et al.*, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \to K^{(*)}l^+l^-$*, Phys. Rev. D **86** (2012) 032012, `arXiv:1204.3933`.

[24] Belle, J.-T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for $B \to K^{(*)}\ell^+\ell^-$*, Phys. Rev. Lett. **103** (2009) 171801, `arXiv:0904.0770`.

[25] LHCb, R. Aaij *et al.*, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, JHEP **08** (2017) 055, `arXiv:1705.05802`.

[26] LHCb, R. Aaij *et al.*, *Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, `arXiv:1406.6482`.

[27] LHCb, R. Aaij *et al.*, *Test of lepton universality in beauty-quark decays*, `arXiv:2103.11769`.

[28] LHCb, R. Aaij *et al.*, *Test of lepton universality with $\Lambda_b^0 \to pK^-\ell^+\ell^-$ decays*, JHEP **05** (2020) 040, `arXiv:1912.08139`.

[29] CLEO, N. E. Adam *et al.*, *Determination of the anti-B —> D* l anti-nu decay width and —V(cb)—*, Phys. Rev. D **67** (2003) 032001, `arXiv:hep-ex/0210040`.

[30] BaBar, B. Aubert *et al.*, *Determination of the form-factors for the decay $B^0 \to D^{*-}\ell^+\nu_l$ and of the CKM matrix element $|V_{cb}|$*, Phys. Rev. D **77** (2008) 032002, `arXiv:0705.4008`.

[31] BaBar, B. Aubert *et al.*, *Measurements of the Semileptonic Decays anti-B —> D l anti-nu and anti-B —> D* l anti-nu Using a Global Fit to D X l anti-nu Final States*, Phys. Rev. D **79** (2009) 012002, `arXiv:0809.0828`.

[32] Belle, W. Dungel *et al.*, *Measurement of the form factors of the decay B0 -> D*- ell+ nu and determination of the CKM matrix element —Vcb—*, Phys. Rev. D **82** (2010) 112007, `arXiv:1010.5620`.

[33] Belle, R. Glattauer *et al.*, *Measurement of the decay $B \to D\ell\nu_\ell$ in fully reconstructed events and determination of the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cb}|$*, Phys. Rev. D **93** (2016) 032006, `arXiv:1510.03657`.

[34] Belle, E. Waheed *et al.*, *Measurement of the CKM matrix element $|V_{cb}|$ from $B^0 \to D^{*-}\ell^+\nu_\ell$ at Belle*, Phys. Rev. D **100** (2019) 052007, `arXiv:1809.03290`, [Erratum: Phys.Rev.D 103, 079901 (2021)].

[35] BaBar, J. P. Lees *et al.*, *Evidence for an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays*, Phys. Rev. Lett. **109** (2012) 101802, `arXiv:1205.5442`.

[36] Belle, M. Huschle *et al.*, *Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, Phys. Rev. D **92** (2015) 072014, `arXiv:1507.03233`.

[37] Belle, A. Abdesselam *et al.*, *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, `arXiv:1904.08794`.

[38] Belle, Y. Sato *et al.*, *Measurement of the branching ratio of $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ relative to $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with a semileptonic tagging method*, Phys. Rev. D **94** (2016) 072007, `arXiv:1607.07923`.

[39] Belle, S. Hirose *et al.*, *Measurement of the $\tau$ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$*, Phys. Rev. Lett. **118** (2017) 211801, `arXiv:1612.00529`.

[40] LHCb, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$*, Phys. Rev. Lett. **115** (2015) 111803, `arXiv:1506.08614`, [Erratum: Phys.Rev.Lett. 115, 159901 (2015)].

[41] LHCb, R. Aaij *et al.*, *Test of Lepton Flavor Universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong $\tau$ decays*, Phys. Rev. D **97** (2018) 072013, `arXiv:1711.02505`.

[42] LHCb, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$*, Phys. Rev. Lett. **120** (2018) 121801, `arXiv:1711.05623`.

[43] CLEO, A. H. Mahmood *et al.*, *Measurement of the B-meson inclusive semileptonic branching fraction and electron energy moments*, Phys. Rev. D **70** (2004) 032003, `arXiv:hep-ex/0403053`.

[44] M. Tanaka, *Charged Higgs effects on exclusive semitauonic B decays*, Z. Phys. C **67** (1995) 321, `arXiv:hep-ph/9411405`.

[45] D. Bigi and P. Gambino, *Revisiting $B \to D\ell\nu$*, Phys. Rev. D **94** (2016) 094008, `arXiv:1606.08030`.

[46] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA, 1995.

[47] I. Caprini, L. Lellouch, and M. Neubert, *Dispersive bounds on the shape of $\bar{B} \to D^*\ell\bar{\nu}$ form-factors*, Nucl. Phys. B **530** (1998) 153, `arXiv:hep-ph/9712417`.

[48] S. Fajfer, J. F. Kamenik, and I. Nisandzic, *On the $B \to D^*\tau\bar{\nu}_\tau$ Sensitivity to New Physics*, Phys. Rev. D **85** (2012) 094025, `arXiv:1203.2654`.

[49] S. Fajfer, J. F. Kamenik, and I. Nisandzic, *On the $B \to D^* \tau \bar{\nu}_\tau$ Sensitivity to New Physics*, Phys. Rev. D **85** (2012) 094025, `arXiv:1203.2654`.

[50] MILC, J. A. Bailey *et al.*, *B→Dℓν form factors at nonzero recoil and —$V_{cb}$— from 2+1-flavor lattice QCD*, Phys. Rev. D **92** (2015) 034506, `arXiv:1503.07237`.

[51] HPQCD, H. Na *et al.*, *$B \to Dl\nu$ form factors at nonzero recoil and extraction of $|V_{cb}|$*, Phys. Rev. D **92** (2015) 054510, `arXiv:1505.03925`, [Erratum: Phys.Rev.D 93, 119906 (2016)].

[52] S. Jaiswal, S. Nandi, and S. K. Patra, *Extraction of $|V_{cb}|$ from $B \to D^{(*)}\ell\nu_\ell$ and the Standard Model predictions of $R(D^{(*)})$*, JHEP **12** (2017) 060, `arXiv:1707.09977`.

[53] F. U. Bernlochner, Z. Ligeti, M. Papucci, and D. J. Robinson, *Combined analysis of semileptonic B decays to D and $D^*$: $R(D^{(*)})$, $|V_{cb}|$, and new physics*, Phys. Rev. D **95** (2017) 115008, `arXiv:1703.05330`, [Erratum: Phys.Rev.D 97, 059902 (2018)].

[54] D. Bigi, P. Gambino, and S. Schacht, *$R(D^*)$, $|V_{cb}|$, and the Heavy Quark Symmetry relations between form factors*, JHEP **11** (2017) 061, `arXiv:1707.09509`.

[55] F. U. Bernlochner *et al.*, *Das ist der HAMMER: Consistent new physics interpretations of semileptonic decays*, Eur. Phys. J. C **80** (2020) 883, `arXiv:2002.00020`.

[56] J. García Pardiñas *et al.*, *RooHammerModel: interfacing the HAMMER software tool with the HistFactory package*, `arXiv:2007.12605`.

[57] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, `arXiv:physics/0306116`.

[58] G. A. Cowan, D. C. Craik, and M. D. Needham, *RapidSim: an application for the fast simulation of heavy-quark hadron decays*, Comput. Phys. Commun. **214** (2017) 239, `arXiv:1612.07489`.

[59] A. Ryd *et al.*, *EvtGen: A Monte Carlo Generator for B-Physics*, .

[60] L. R. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001. 164 p.

[61] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08002.

[62] The ATLAS Collaboration, *The ATLAS experiment at the CERN large hadron collider*, Journal of Instrumentation **3** (2008) S08003.

[63] The CMS Collaboration, *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08004.

[64] The LHCf Collaboration, *The LHCf detector at the CERN large hadron collider*, Journal of Instrumentation **3** (2008) S08006.

[65] T. T. Collaboration, *The TOTEM experiment at the CERN large hadron collider*, Journal of Instrumentation **3** (2008) S08007.

[66] B. Acharya *et al.*, *The physics programme of the moedal experiment at the LHC*, International Journal of Modern Physics A **29** (2014) 1430050, `arXiv:https://doi.org/10.1142/S0217751X14300506`.

[67] T. Sjöstrand, *The PYTHIA Event Generator: Past, Present and Future*, Comput. Phys. Commun. **246** (2020) 106910, `arXiv:1907.09874`.

[68] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007, 2000, LHCb-TDR-001.

[69] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036, 2000, LHCb-TDR-002.

[70] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037, 2000, LHCb-TDR-003.

[71] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010, 2001, LHCb-TDR-004.

[72] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011, 2001, LHCb-TDR-005.

[73] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024, 2001, LHCb-TDR-006.

[74] LHCb collaboration, *LHCb online system, data acquisition and experiment control: Technical Design Report*, CERN-LHCC-2001-040, 2001, LHCb-TDR-007.

[75] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029, 2002, LHCb-TDR-008.

[76] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030, 2003, LHCb-TDR-009.

[77] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031, 2003, LHCb-TDR-010.

[78] LHCb collaboration, *LHCb computing: Technical Design Report*, CERN-LHCC-2005-019, 2005, LHCb-TDR-011.

[79] LHCb collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007, 2012, LHCb-TDR-012.

[80] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021, 2013, LHCb-TDR-013.

[81] LHCb collaboration, *LHCb PID Upgrade Technical Design Report*, CERN-LHCC-2013-022, 2013, LHCb-TDR-014.

[82] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001, 2014, LHCb-TDR-015.

[83] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016, 2014, LHCb-TDR-016.

[84] *The DELPHI detector at LEP*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **303** (1991) 233.

[85] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, tech. rep., CERN, Geneva, 2014.

[86] R. Frhwirth, *Application of kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987) 444.

[87] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction*, LHCb-PUB-2017-011, 2017.

[88] Particle Data Group, P. A. Zyla *et al.*, *Review of Particle Physics*, PTEP **2020** (2020) 083C01.

[89] S. Godfrey, *Meson physics in a relativized quark model with chromodynamics*, tech. rep., TRIUMF, Burnaby, BC, 1985.

[90] LHCb, R. Aaij *et al.*, *Study of $D_J$ meson decays to $D^+\pi^-$, $D^0\pi^+$ and $D^{*+}\pi^-$ final states in pp collision*, JHEP **09** (2013) 145, arXiv:1307.4556.

[91] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[92] GEANT4, S. Agostinelli *et al.*, *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[93] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[94] E. E. Bowen, B. Storaci, and M. Tresch, *VeloTT tracking for LHCb Run II*, tech. rep., CERN, Geneva, 2016.

[95] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. A **764** (2014) 150, arXiv:1312.5000.

[96] LHCb, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, arXiv:1408.1251.

[97] G. Bohm and G. Zech, *Statistics of weighted Poisson events and its applications*, Nucl. Instrum. Meth. A **748** (2014) 1, arXiv:1309.1287.

[98] F. U. Bernlochner, Z. Ligeti, and D. J. Robinson, *N = 5, 6, 7, 8: Nested hypothesis tests and truncation dependence of $|V_{cb}|$*, Phys. Rev. D **100** (2019) 013005, arXiv:1902.09553.

[99] D. Scora and N. Isgur, *Semileptonic meson decays in the quark model: An update*, Phys. Rev. D **52** (1995) 2783, arXiv:hep-ph/9503486.

[100] F. U. Bernlochner, Z. Ligeti, and D. J. Robinson, *Model independent analysis of semileptonic B decays to D\*\* for arbitrary new physics*, Phys. Rev. D **97** (2018) 075011, `arXiv:1711.03110`.

[101] F. U. Bernlochner and Z. Ligeti, *Semileptonic $B_{(s)}$ decays to excited charmed mesons with $e, \mu, \tau$ and searching for new physics with $R(D^{**})$*, Phys. Rev. D **95** (2017) 014022, `arXiv:1606.09300`.

[102] T. M. Karbach and M. Schlupp, *Constraints on Yield Parameters in Extended Maximum Likelihood Fits*, `arXiv:1210.7141`.

[103] ROOT Collaboration, K. Cranmer *et al.*, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep., New York U., New York, 2012.