



Università degli Studi di Ferrara

Dottorato di ricerca in FISICA

CICLO XXVIII

Coordinatore Prof. Guidi Vincenzo

**Studies of Semileptonic B Decays in
 $D^* \pi(\pi) \mu X$ final states with the LHCb
Detector**

Settore Scientifico Disciplinare FIS/01

Dottorando
Dott. Fiore Marco

Tutore
Prof. Bozzi Concezio

Anni 2013/2015

Contents

Introduction	iii
1 Semileptonic Decays of B Mesons	1
1.1 Standard Model	1
1.2 Glashow-Weinberg-Salam Model	3
1.3 Cabibbo-Kobayashi-Maskawa matrix	9
1.4 Decays of b -hadrons	14
1.5 Semileptonic decays: V_{cb} determination	16
1.5.1 Exclusive decays	16
1.5.2 Inclusive decays	18
1.5.3 Discrepancy	20
1.6 Charmed mesons	20
1.6.1 Spectrum	20
1.6.2 Measured branching fractions of semileptonic decays	23
2 The LHCb Detector	27
2.1 VERtex LOcator	28
2.2 Magnet	30
2.3 Tracking system	32
2.3.1 Trigger Tracker and Inner Tracker	32
2.3.2 Outer Tracker	33
2.4 Particle Identification	34
2.4.1 RICH	34
2.4.2 Calorimeters	35
2.4.3 Muon system	37
2.5 Trigger	39
2.5.1 Level 0	40
2.5.2 High Level Trigger	42
2.6 Offline processing	43
2.7 Performances	43
2.7.1 Track reconstruction	43

2.7.2	Particle identification	45
3	Level 0 Muon Trigger Optimisation	49
3.1	Level 0 Muon Track Finding Algorithm	49
3.2	Monte Carlo Studies	52
3.3	Minimum Bias Studies	54
3.4	FOI Optimisation	56
4	Analysis Strategy	65
4.1	Data sample	65
4.2	Monte Carlo sample	65
4.3	Signal and background definition	68
4.4	Event reconstruction	71
4.4.1	D^* selection	71
4.4.2	Muon selection	72
4.4.3	π^{**} pre-selection	74
4.5	Multivariate Analysis	75
4.5.1	Boosted Decision Trees	77
4.5.2	Training and test samples	79
4.5.3	Input variables definition	80
4.5.4	TMVA output	84
4.6	π_{**} selection	87
5	Fit to Monte Carlo and data distributions	91
5.1	Discriminating variable	91
5.2	Unbinned maximum likelihood fit	93
5.3	Fit on Monte Carlo simulated sample	94
5.3.1	$D^*\pi\mu$ channel	95
5.3.2	$D^*\pi\pi\mu$ channel	96
5.3.3	Consistency check	98
5.4	Fit to the $m(K\pi\pi) - m(K\pi)$ distribution	99
5.4.1	Signal parametrisation	100
5.4.2	Background parametrisation	101
5.4.3	Total PDF	101
5.5	Fit on data sample	101
5.5.1	Combinatorial D^* and $B \rightarrow D^0\pi\mu\nu_\mu$ background estimation	103
5.5.2	$D^*\pi$ channel	104
5.5.3	$D^*\pi\pi$ channel	106
5.6	Systematic uncertainties on yields	110
5.7	Branching ratios	112
5.7.1	Efficiencies	114

5.8 Systematic uncertainties on branching ratios	114
6 Results and conclusions	117
Bibliography	119

Introduction

The aim of this thesis is to study semileptonic B mesons, either neutral or charged, decays to final states containing excited D mesons, one or two pions, a muon and a neutrino; where the excited D meson decays in a neutral D meson and a pion and the D meson subsequently decays in a kaon and a pion.

More precisely, this thesis shows a measurement of the Branching Ratios (BRs) for these decays. First, using $B \rightarrow D^* \mu \nu_\mu X$ as normalisation channel, the Branching Ratios are calculated relatively to the one of this decay, then the absolute Branching Ratio is calculated, taking the $B \rightarrow D^* \mu \nu_\mu X$ BR value from the PDG.

The theoretical framework of these decays and the relevance of this measurements for High Energy Physics are reported in Chapter 1.

The data used in this study were collected by the LHCb experiment in the years 2011 and 2012 (Run I of the LHC), corresponding to an integrated luminosity of 3 fb^{-1} . The LHCb detector and its performance are described in Chapter 2.

Chapter 3 shows an optimisation of the LHCb Level 0 Muon trigger, i.e. that part of the LHCb Level 0 trigger that exploits informations from the LHCb Muon detector. The effects of this optimisation were not included in this thesis as they were implemented during the first LHC long shutdown (LS1, i.e. after Run I data taking), in time for the Run II data taking.

The analysis was carried on using both real and Monte Carlo (MC) simulated data. Chapter 4 describes the samples used and the two stages to select them:

- a cut-based pre-selection of $B \rightarrow D^* \mu \nu_\mu X$ decays, according to criteria described in Section 4.4 ;
- a finer selection of $B \rightarrow D^* \pi \mu \nu_\mu X$ and $B \rightarrow D^* \pi \pi \mu \nu_\mu X$ using a *MultiVariate Analysis* (MVA), described in Section 4.5.

After the pre-selection, the $B \rightarrow D^* \mu \nu_\mu X$ sample was used to fit the $m(K\pi\pi) - m(K\pi)$ distribution, in order to get the $B \rightarrow D^* \mu \nu_\mu X$ signal yield and calculate the BR used as normalisation. The MC sample was used to compute all of the

selection efficiencies, which are needed for the Branching Ratios calculation. After applying the second selection, for the $B \rightarrow D^* \pi \mu \nu_\mu X$, a unidimensional fit to the Logarithm of the Impact Parameter (IP), with respect to the B decay vertex, of the pion from the B decay is performed. In the case of $B \rightarrow D^* \pi \pi \mu \nu_\mu X$, the fit is bidimensional, as the Logarithm of the IP is fitted for both pions simultaneously. From these fits, the signal yields are extracted and the BRs are computed for both channels.

The whole procedure, the formulae and the fit results are gathered in Chapter 5, along with the systematic uncertainties that are considered for this measurement. Final results and conclusions are summarised in Chapter 6.

Chapter 1

Semileptonic Decays of B Mesons

1.1 Standard Model

The Standard Model (SM) of particles describes the fundamental constituents of matter and their interactions. The particles of the standard model are of three distinct types: leptons, quarks and gauge bosons. Leptons and quarks are spin-1/2 particles (fermions) that represent the basic constituents of matter, while the gauge bosons are spin-1 particles that mediate the interactions. In addition, a spin-0 particle called the Higgs boson has to be introduced to explain the origin of mass within the theory: without it all the particles in the model are predicted to have zero mass; its recent discovery was an important confirmation for the SM. All the particles of the Standard Model are assumed to be elementary, so they are treated as point-like particles, without internal structure or excited states.

Both leptons and quarks come in six flavors and are arranged, respectively, in three lepton families:

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix} \quad \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix} \quad \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}$$

and three quark families:

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}$$

Our universe is made of particles from the first family, while particles of the second and third family are produced in particle accelerators and very high-energy cosmic rays. Although leptons exist as free particles, free quarks are subject to a phenomenon called confinement, so they always form bound states called hadrons.

The fundamental forces included in the theory are the strong, the electromagnetic

and the weak interactions¹. The charge of a particle is defined only in relation to an interaction: both charged leptons and quarks undergo electromagnetic and weak interactions, thus they carry electric and weak charge, while only quarks interact strongly and carry color charge. Neutral leptons interact only through the weak force. Fermions interact via the exchange of bosons:

- The electromagnetic interaction between electrically charged particles is mediated by the photon, which is massless and electrically neutral.
- The weak interaction is carried by three massive bosons, W^\pm and Z^0 . The Z^0 boson is electrically neutral and has a mass of about $91 \text{ GeV}/c^2$, while W^\pm bosons carry electric charge and have a mass of about $80 \text{ GeV}/c^2$.
- The strong force is mediated by eight gluons, massless and electrically neutral. Differently from the mediators of weak and electromagnetic interactions, they carry color charge.

The SM is a gauge theory, i.e. it describes particles and their interactions by imposing symmetries to the Lagrangian of free particles. Therefore the interactions enter into the theory as a consequence of invariance requirements. The Standard Model Lagrangian has been built in such way that:

- the spin structure is taken into account according to the Dirac equation;
- it is gauge-invariant (locally and globally) under abelian and non abelian groups (since there should be invariance under the $U(1)$, $SU(2)$ and $SU(3)$ internal symmetries);
- the gauge bosons acquire mass via a spontaneous symmetry breaking process, denoted as the Brout-Englert-Higgs mechanism;
- the fermions acquire mass through gauge-invariant Yukawa couplings to the Higgs field.

In particle physics there are three fundamental discrete symmetries that have been used to give insight on the nature of fundamental interactions: parity (P), charge

¹The gravitational force is not included for two reasons: at the scale of particle physics experiments it has a negligible intensity compared to the other interactions, and it is still not possible to be described in terms of quantum field theory.

conjugation (C) and time reversal (T). Parity and charge conjugation are unitary transformations, while time reversal is an anti-unitary transformation. Parity is a simultaneous inversion of all space coordinates, charge conjugation changes a particle into its own anti-particle, time reversal interchanges all directions of motion. The CP combined transformation replaces a particle by its antiparticle and reverses momentum and helicity. According to any experiment so far, e.g. [2], the combination of the three symmetries, CPT , is an exact symmetry in nature. C , P and CP are exact symmetries of the strong and electromagnetic interactions, while they are not conserved in weak processes. CP violation is encoded in the Standard Model of weak interactions, as explained in Section 1.3.

1.2 Glashow-Weinberg-Salam Model

The theoretical framework for the description of the electroweak interactions is given by the Glashow-Weinberg-Salam (GWS) model [3] [4], a non-abelian, Yang-Mills theory which unifies electromagnetic and weak interactions. The GWS model is based on the gauge symmetry group $SU(2) \times U(1)$, which is the product of two sets of gauge transformations: the $SU(2)$ group with coupling strength g and the $U(1)$ group with coupling strength g' .

The Lagrangian density of the SM describing electroweak interactions² can be split in two parts:

$$\mathcal{L} = \mathcal{L}_{symm} + \mathcal{L}_{Higgs} \quad (1.1)$$

a symmetric part, which involves only gauge bosons and fermions, and a second part that includes the couplings with the Higgs field.

The symmetric part has the following form:

$$\mathcal{L}_{symm} = -\frac{1}{4} \mathbf{W}_{\mu\nu} \cdot \mathbf{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \bar{\psi}_L i \gamma^\mu D_\mu \psi_L + \bar{\psi}_R i \gamma^\mu D_\mu \psi_R. \quad (1.2)$$

The first two terms represent the gauge bosons kinetic energy and self-interactions, where $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ and $W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g \epsilon_{ijk} W_\mu^j W_\nu^k$ are the gauge antisymmetric tensors constructed, respectively, with the gauge field B_μ , generator of $U(1)$, and W_μ^i ($i = 1, 2, 3$) corresponding to the three generators of $SU(2)$; ϵ_{ijk} are the group structure constants which for $SU(2)$ is the totally antisymmetric Levi-Civita tensor.

²If the terms describing the $SU(3)$ symmetry group are included, the whole Lagrangian density of the SM is obtained.

The last two terms give the kinetic energy of the fermions and their interactions with the gauge bosons. The fermion fields enter in the Lagrangian as their left-handed and right-handed components:

$$\psi_{L,R} = [(1 \mp \gamma_5)/2]\psi, \quad \bar{\psi}_{L,R} = \bar{\psi}[(1 \pm \gamma_5)/2], \quad (1.3)$$

because in the Standard Model the left-handed and right-handed fermions have different transformation properties under the gauge group. Thus fermions are arranged in left-handed SU(2) doublets³:

$$\chi_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \quad \begin{pmatrix} u \\ d' \end{pmatrix}_L$$

and right-handed singlets:

$$\psi_R = e_R, \quad u_R, \quad d_R$$

A right-handed up quark has been included here, since quarks, unlike neutrinos have a finite mass and hence have both right- and left-handed components. This pattern, shown for the first generation, is replicated for the second and third fermion families. In the absence of mass terms, there are only vector and axial vector interactions in the Lagrangian that do not mix ψ_L and ψ_R .

The covariant derivatives D_μ , introduced in order to have a locally gauge invariant Lagrangian density, are explicitly given by

$$D_\mu = \partial_\mu + ig\frac{1}{2}\boldsymbol{\tau} \cdot \mathbf{W}_\mu + ig'\frac{1}{2}YB_\mu \quad (1.4)$$

where $\boldsymbol{\tau}$ and $1/2Y$ are the SU(2) and U(1) generators; in particular, $\boldsymbol{\tau}$ is a vector whose components can be represented as the Pauli matrices, defined as:

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.5)$$

In analogy with the Lagrangian of quantum electrodynamics (QED), it is possible to rewrite the weak part of the Lagrangian in terms of a charged weak current (CC) and a neutral weak current (NC). These form an isospin triplet of weak currents \mathbf{J}^μ coupled to the three vector bosons \mathbf{W}_μ :

$$-ig\mathbf{J}^\mu \cdot \mathbf{W}_\mu = -ig\bar{\chi}_L\boldsymbol{\gamma}^\mu\boldsymbol{\tau} \cdot \mathbf{W}_\mu\chi_L. \quad (1.6)$$

The form of the charged currents is:

³in the quark doublet the gauge eigenstates d' for down-type quarks are linear combinations of the mass eigenstates d (see Section 1.3).

$$\begin{aligned}
J_\mu^+ &= \overline{\chi}_L \gamma_\mu \tau_+ \chi_L = \overline{\nu}_L \gamma_\mu \frac{1}{2} (1 - \gamma^5) e = \overline{\nu}_L \gamma_\mu e_L \\
J_\mu^- &= \overline{\chi}_L \gamma_\mu \tau_- \chi_L = \overline{e}_L \gamma_\mu \frac{1}{2} (1 - \gamma^5) \nu = \overline{e}_L \gamma_\mu \nu_L
\end{aligned} \tag{1.7}$$

where $\tau_\pm = \frac{1}{2}(\tau_1 \pm i\tau_2)$:

$$\tau_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \tau_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{1.8}$$

whereas the neutral current has the following form:

$$J_\mu^3 = \overline{\chi}_L \gamma_\mu \tau_3 \chi_L = \frac{1}{2} \overline{\nu}_L \gamma_\mu \nu_L - \frac{1}{2} \overline{e}_L \gamma_\mu e_L. \tag{1.9}$$

The currents (1.7) and (1.9) are explicitly written in terms of leptons from the first family, but these expressions are valid for every fermion doublet.

To extend the symmetry to the SU(2) X U(1) group, another current must be introduced, which is a weak hypercharge current coupled to the fourth vector boson B^μ ,

$$-i \frac{g'}{2} j_\mu^Y B^\mu = -i g' \overline{\psi} \gamma_\mu \frac{Y}{2} \psi B^\mu. \tag{1.10}$$

As the charge operator Q generates the group U(1)_{em}, the hypercharge operator Y generates a symmetry group U(1)_Y and the so-called weak hypercharge quantum number, Y_W , which satisfies

$$Q = T_3 + \frac{Y_W}{2} \tag{1.11}$$

where T_3 is the third component of the weak isospin, i.e. the group symmetry associated to the weak interaction. In this way the electromagnetic interaction can be incorporated and the result is a symmetry group enlarged to SU(2) X U(1), leading to a unification of electromagnetic and weak interactions.

The weak isospin and hypercharge quantum numbers of lepton and quarks are summarised in Table 1.1.

Table 1.1. Quantum numbers of leptons and quarks: weak isospin T , third component of the weak isospin T_3 , electric charge Q and hypercharge Y .

Lepton	T	T_3	Q	Y	Quark	T	T_3	Q	Y
ν_e	$\frac{1}{2}$	$\frac{1}{2}$	0	-1	u_L	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{3}$
e_L	$\frac{1}{2}$	$-\frac{1}{2}$	-1	-1	d_L	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{3}$	$\frac{1}{3}$
					u_R	0	0	$\frac{2}{3}$	$\frac{4}{3}$
e_R	0	0	-1	-2	d_R	0	0	$-\frac{1}{3}$	$-\frac{2}{3}$

The electric charge Q (in units of e) is related to the third component of the weak isospin and hypercharge by

$$Q = T_3 + \frac{Y}{2}. \quad (1.12)$$

Finally, the Lagrangian is rewritten in terms of the physical gauge fields. In the charged-current interaction Lagrangian W_μ^1 and W_μ^2 are substituted by the massive charged bosons $W_\mu^\pm = \sqrt{\frac{1}{2}}(W_\mu^1 \mp iW_\mu^2)$. In the neutral-current sector, the photon A_μ and the mediator Z_μ of the weak NC are built as orthogonal and normalised linear combinations of B_μ and W_μ^3 , such that the physical states become:

$$\begin{aligned} A_\mu &= \cos \theta_W B_\mu + \sin \theta_W W_\mu^3 \\ Z_\mu &= -\sin \theta_W B_\mu + \cos \theta_W W_\mu^3, \end{aligned} \quad (1.13)$$

where θ_W is the Weinberg mixing angle. The photon A_μ is massless, while the boson Z_μ is massive. Using (1.13), the electroweak neutral current interaction takes the form:

$$-igJ_\mu^3(W^3)^\mu - i\frac{g'}{2}j_\mu^Y B^\mu = -i(g \sin \theta_W J_\mu^3 + g' \cos \theta_W \frac{j_\mu^Y}{2})A^\mu - i(g \cos \theta_W J_\mu^3 - g' \sin \theta_W \frac{j_\mu^Y}{2})Z^\mu \quad (1.14)$$

The first term is the electromagnetic interaction, with the electromagnetic current appearing as a combination of the two neutral currents J_μ^3 and j_μ^Y . Since the photon is characterised by equal couplings to left and right fermions with a strength equal to the electric charge, recalling (1.12),

$$g \sin \theta_W = g' \cos \theta_W = e. \quad (1.15)$$

Thus the mixing angle is given by the ratio of the two independent coupling constants, $\tan \theta_W = g'/g$. So the weak neutral current interaction can be rewritten as:

$$-i\frac{g}{\cos \theta_W}(J_\mu^3 - \sin^2 \theta_W j_\mu^{em})Z^\mu = -i\frac{g}{\cos \theta_W}J_\mu^{NC}Z^\mu \quad (1.16)$$

where $J_\mu^{NC} = J_\mu^3 - \sin^2 \theta_W j_\mu^{em}$, linking the neutral current J^{NC} to the weak isospin current \mathbf{J} .

In the Lagrangian described so far, \mathcal{L}_{symm} , the gauge bosons and fermions are massless. That is because mass terms such as $\frac{1}{2}M^2 B_\mu B^\mu$ and $-m\bar{\psi}\psi$ are not gauge invariant and thus can not be added at this point. To generate the particle masses in a gauge invariant way, the Brout-Englert-Higgs mechanism is introduced.

Turning to the Higgs part of the Lagrangian, \mathcal{L}_{Higgs} , four real scalar fields ϕ_i ($i = 1, \dots, 4$) need to be added in order to obtain a Lagrangian invariant under the $SU(2) \times U(1)$ gauge symmetry.

To keep the Lagrangian gauge invariant, the fields ϕ_i are arranged in an isospin doublet of complex scalar fields with weak hypercharge $Y = 1$:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.17)$$

and the Higgs potential takes the form

$$V(\phi) = \mu^2 \phi^+ \phi + \lambda (\phi^+ \phi)^2. \quad (1.18)$$

This gives:

$$\mathcal{L}_{Higgs} = \left| \left(i\partial_\mu - g \frac{1}{2} \boldsymbol{\tau} \cdot \mathbf{W}_\mu - g' \frac{Y}{2} B_\mu \right) \phi \right|^2 - V(\phi), \quad (1.19)$$

where the two terms encode the gauge boson masses and their couplings to the Higgs.

If $\mu^2 > 0$, the Lagrangian (1.19) describes a system of four scalar particles each of mass μ , interacting with three massless gauge bosons W_μ^a . The interesting case is the one with $\mu^2 < 0$ and $\lambda > 0$: the potential $V(\phi)$ has a minimum at a finite value of $|\phi|$ where

$$\phi^+ \phi = \frac{1}{2} (\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda}. \quad (1.20)$$

Once a particular vacuum expectation value is chosen, the symmetry is spontaneously broken and a mass for the corresponding gauge boson is generated. The most suitable choice is the following:

$$\phi_1 = \phi_2 = \phi_4 = 0, \quad \phi_3^2 = -\frac{\mu^2}{\lambda} = \nu^2 \quad (1.21)$$

so that the vacuum takes the value

$$\phi_0 = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix}. \quad (1.22)$$

If the vacuum ϕ_0 is left invariant by some gauge transformation subgroups of $SU(2) \times U(1)$, the gauge boson associated with this subgroup remains massless. Instead, if the choice of ϕ_0 breaks a symmetry, a mass for the correspondent gauge boson is generated. So the condition that the photon remains massless is equivalent to the condition that the vacuum is electrically neutral. This choice of ϕ_0 (1.22) with

$T = \frac{1}{2}$, $T^3 = -\frac{1}{2}$, and $Y = 1$ breaks both $SU(2)$ and $U(1)_Y$ gauge symmetries, but the $U(1)_{em}$ symmetry with generator Q (Eq. (1.12)) remains unbroken.

The masses of the gauge bosons are obtained by substituting the vacuum expectation value ϕ_0 for $\phi(x)$ in the Lagrangian density (1.19). Comparing the quadratic term in the W field in this expression with the mass expected for a charged boson $M_W^2 W^+ W^-$, the charged W mass is given:

$$M_W = \frac{1}{2}\nu g. \quad (1.23)$$

Since the physical fields Z_μ and A_μ diagonalise the mass matrix, the other terms must be identified with $\frac{1}{2}M_Z^2 Z_\mu^2 + \frac{1}{2}M_A^2 A_\mu^2$, leading to the mass terms for the neutral vector bosons:

$$M_A = 0 \quad (1.24)$$

$$M_Z = \frac{1}{2}\nu\sqrt{g^2 + g'^2}. \quad (1.25)$$

Rewriting the fields A_μ and Z_μ in terms of the mixing angle and recalling (1.23) and (1.25), the following relation arises:

$$\frac{M_W}{M_Z} = \cos \theta_W. \quad (1.26)$$

From that it is clear that the inequality $M_Z \neq M_W$ is due to the mixing between the W_μ^3 and B_μ fields. The mass eigenstates are a massless photon and a massive Z_μ field with $M_Z > M_W$. The mass relation (1.26) is a prediction of the Standard Model that fixes the parameter ρ (which gives the relative strength of the neutral and charged current weak interactions) to be

$$\rho = \frac{M_W^2}{M_Z^2 \cos^2 \theta_W} = 1. \quad (1.27)$$

The same Higgs doublet which generates W^\pm and Z masses allows to give masses to the fermions as well. The masses of the leptons are obtained by adding the term $-(G_1 \bar{\chi}_L \phi \psi_R + G_2 \bar{\chi}_L \phi_c \psi_R + \text{h.c.})$ to the Higgs Lagrangian density Eq. 1.19, where the Higgs doublet ϕ is the one in 1.17 and ϕ_c^4 is defined as

$$\phi_c = \begin{pmatrix} -\bar{\phi}^0 \\ \bar{\phi}^- \end{pmatrix}.$$

⁴ ϕ_c has the same transformation properties of ϕ , but opposite hypercharge.

Once the following substitutions are made,

$$\phi = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix} \quad (1.28)$$

and

$$\phi_c = \sqrt{\frac{1}{2}} \begin{pmatrix} \nu + h(x) \\ 0 \end{pmatrix}, \quad (1.29)$$

the symmetry is spontaneously broken.

The Lagrangian contains a quadratic term in the lepton field and an interaction term coupling the Higgs scalar to the lepton. The first term, compared to the mass term expected for a fermion $m\bar{\psi}\psi$, leads to a definition of the coupling constant G_ℓ such that

$$m_\ell = \frac{G_\ell \nu}{\sqrt{2}}. \quad (1.30)$$

This choice of G_ℓ implies also that the coupling of the lepton to the Higgs is proportional to the lepton mass.

As in the lepton case, the coupling constants can be defined such that:

$$m_u = \frac{G_u \nu}{\sqrt{2}}, \quad m_d = \frac{G_d \nu}{\sqrt{2}}. \quad (1.31)$$

The masses of the fermions enter in the theory as parameters, hence they can not be predicted; on the other side, the Higgs coupling to the fermions is proportional to their masses and this could be tested experimentally.

The Higgs boson was predicted in 1964 by three groups of physicists: F.Englert and R.Brout [5], P.Higgs [6], G.Guralnik, C.Hagen and T.Kibble [7]. In July 2012, the ATLAS and CMS collaborations [8] observed a particle with the Higgs characteristics, with a mass at around 125 GeV/c². In March 2013 new results have shown that the observed particle has $J^P = 0^+$ as predicted [9].

1.3 Cabibbo-Kobayashi-Maskawa matrix

The quark masses (1.31) introduced in the Lagrangian density after the spontaneous symmetry breaking are in general not diagonal. Since the physical fields are eigenstates of the mass matrices m_u and m_d , they should be diagonalised. In order to do that, four unitary matrices (U_L, U_R to diagonalise the m_u matrix, and D_L and D_R to diagonalise the m_d matrix) are introduced: they allow to find the

relations between physical and non-physical gauge fields, leading to the following expression of the quark Higgs part of the Lagrangian density:

$$\mathcal{L}_{mass}^{phys} = -[1 + \frac{h}{\nu}] \sum_{i=1}^3 (m_i^u \bar{u}_i^{phys} u_i^{phys} + m_i^d \bar{d}_i^{phys} d_i^{phys}), \quad (1.32)$$

where u states for an up-type quark and d for a down-type quark. In (1.32) there is no trace of the unitary matrices (because they enter in the definition of physical fields) and it becomes evident that the physical Higgs-fermion interactions conserves parity, charge conjugation and is flavour diagonal. The unitary matrices introduced have no effect on the neutral currents; neutral current couplings are diagonal in both bases, thus processes in which the quark flavour changes but the charge does not (like $s \rightarrow d$ transitions) only occur at second order in the weak interaction. This property of the weak neutral current, which ensures natural flavour conservation of the neutral current couplings at the tree level, is also known as the Glashow-Iliopoulos-Maiani (GIM) mechanism. Weak charged currents are the only tree level interactions in the Standard Model that change flavour: by emission of a W an up-type quark is turned into a down-type quark, or a ν_L neutrino is turned into a l^- charged lepton. Rewriting the charge current interaction in terms of the physical fields, a unitary matrix V appears [10] :

$$J_{CC}^\mu \propto (\bar{u}_L, \bar{c}_L, \bar{t}_L) \gamma^\mu V \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix}. \quad (1.33)$$

V is called unitary quark mixing matrix or CKM (Cabibbo - Kobayashi - Maskawa) matrix:

$$V_{CKM} = U_L D_L^\dagger = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}.$$

By convention, V_{CKM} operates on mass eigenstates of down-type quarks. Charged current interactions violate CP through imaginary phases in V_{CKM} . This is possible only if there are at least three quark families. If the generations of quarks were N , V would be a $N \times N$ unitary matrix depending on N^2 real numbers ($2 \times N^2$ free real parameters with N^2 unitarity constraints). The quark wave functions are defined with a free overall phase, so they can be redefined in order to absorb $2N - 1$ phases of the CKM matrix. Moreover, a rotation in a N -dimensional space can be parametrised by $N(N - 1)/2$ angles, leaving the number of independent phases equal to $(N - 1)(N - 2)/2$. For $N = 2$ quark families there is one rotation

angle (the Cabibbo angle) and no phase; for $N = 3$ families there are three rotation angles and one complex phase. Thus at least three quark families are required to support CP violation in the Standard Model.

There are several parameterisations of the CKM matrix; a standard form [11] is

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \quad (1.34)$$

where $s_{ij} = \sin \theta_{ij}$, $c_{ij} = \cos \theta_{ij}$ ($i, j = 1, 2, 3$) and δ_{13} is the phase responsible for all CP-violating phenomena in flavour-changing processes in the SM. The rotation angles are defined so that they relate the mixing of two specific generations: if one of these angles vanishes, so does the mixing between those two generations.

The standard parametrisation can be approximated in a way that emphasises the hierarchy in the size of the angles, $s_{13} \ll s_{23} \ll s_{12} \ll 1$ [12]. Setting $\lambda \equiv s_{12}$, the sine of the Cabibbo angle, the other elements are expressed in terms of powers of λ :

$$V = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \quad (1.35)$$

where A , ρ and η are real numbers of order unity. Looking at this parametrisation of the CKM matrix due to Wolfenstein, it becomes evident that up-type quarks preferably couple to down-type quarks of the same family, while the couplings to the other ones are smaller the more the two families are distant.

The unitarity of the CKM matrix implies $\sum_i V_{ij}V_{ik}^* = \delta_{jk}$ and $\sum_j V_{ij}V_{kj}^* = \delta_{ik}$. The six vanishing combination can be represented as closed triangles in a complex plane: most of them are nearly degenerate (the ones obtained by taking scalar products of neighbouring rows or columns), but some have sides of comparable length. Since the angles of these triangles represent the relative phases of the matrix elements, large CPV effects can be expected in the second case. All triangles have the same area, half of the Jarlskog invariant J [13], which is of order $J \sim \eta A^2 \lambda^6$. Its value is independent of the parametrisation and gives a measure of CP violation. The most commonly used unitarity triangle arises from the unitarity condition applied to the first and third columns:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \quad (1.36)$$

Orienting the triangle so that $V_{cd}V_{cb}^*$ lies on the real axis (V_{cb} is real and V_{cd} is real to a very good approximation) and rescaling the triangle so that the base is of unit

length, the triangle in Fig. 1.1 is obtained. The angles α , β and γ are also referred to as ϕ_2 , ϕ_1 and ϕ_3 respectively, with β and $\gamma = \delta_{13}$ being the phases of the CKM elements V_{td} and V_{ub} :

$$V_{td} = |V_{td}|e^{-i\beta}, \quad V_{ub} = |V_{ub}|e^{-i\gamma}.$$

The coordinates of the vertices in the complex plane are

$$\left(\frac{\Re(V_{ud}V_{ub}^*)}{|V_{cd}V_{cb}^*|}, \frac{\Im(V_{ud}V_{ub}^*)}{|V_{cd}V_{cb}^*|} \right), \quad (1, 0) \quad \text{and} \quad (0, 0).$$

The coordinates of the apex of the rescaled unitarity triangle take the form $(\bar{\rho}, \bar{\eta})$, with $\bar{\rho} = \rho(1 - \lambda^2/2)$ and $\bar{\eta} = \eta(1 - \lambda^2/2)$ in the Wolfenstein parametrisation.

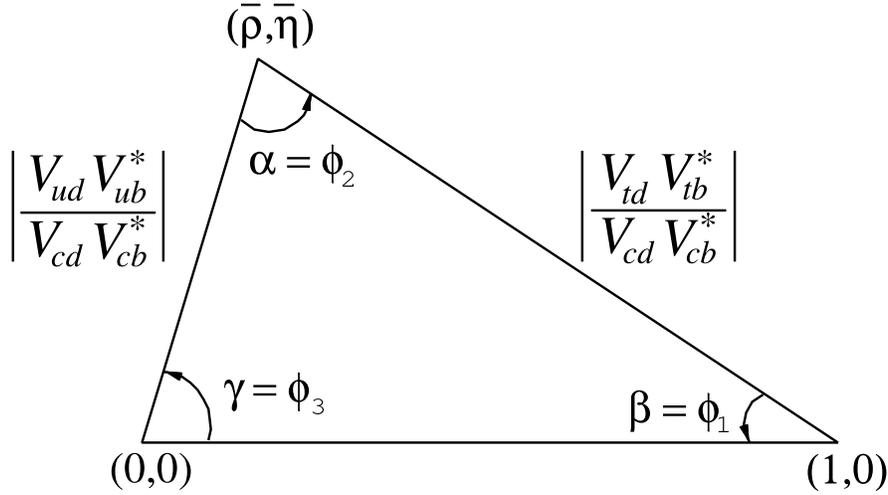


Figure 1.1. The unitarity triangle in the $(\bar{\rho}, \bar{\eta})$ plane.

In the SM, CP violation depends only on a non-zero value of the CKM matrix phase, thus an important goal of flavour physics is to constrain the CKM elements: for this purpose many measurements of the unitarity triangle elements are performed.

The CKM matrix elements can be determined by a global fit that uses all available measurements and imposes the SM constraints (that is the three generations unitarity). The fit uses also theoretical predictions for hadronic matrix elements, which sometimes have significant uncertainties. There are several approaches to combine experimental data: some use frequentist statistics [15], others a Bayesian

1.4 Decays of b -hadrons

Since the b quark is the lighter element of the third-generation quark doublet, the ground states of b -flavoured hadrons decay via weak interactions to up-type quarks of a different family. In most hadrons the b quark is accompanied by light quarks (d, u or s): the bound states of a \bar{b} antiquark and one of the previous mentioned light quarks are referred to as the B_d or B^0 ($\bar{b}d$), B_u or B^+ ($\bar{b}u$) and B_s^0 mesons. The only meson made of two heavy quarks is B_c^+ , made of a \bar{b} antiquark and a c quark.

The bound state of a u , a d and a b quark is the neutral Λ_b^0 baryon.

Examples of decays of B mesons allowed by the Standard Model are collected in Figure 1.3. The dominant decay mode is the so called external spectator decay ($b \rightarrow cW^{*-}$), where the virtual W materialises either into a pair of leptons $l\nu$ (semileptonic decay), or into a pair of quarks which then hadronize. The internal spectator decays, in which the spectator quark combines with one of the quarks from W^* to form one of the final state hadrons, are suppressed by a factor $\sim (1/3)^2$; that is due to colour-suppression, i.e. the colours of the two quarks from different legs must match. The transition $b \rightarrow u$ is suppressed by $|V_{ub}/V_{cb}|^2 \sim (0.1)^2$ relative to $b \rightarrow c$ transitions. The transition $b \rightarrow s$ is a flavour-changing neutral current (FCNC) process, thus it is not allowed in the Standard Model at the tree level, but can occur via higher order (loop) diagrams denoted as penguin diagrams. The rates for such processes are comparable or smaller than CKM-suppressed $b \rightarrow u$ processes. Penguin processes involving $b \rightarrow d$ transition are also possible and have been observed. Other allowed decay processes are W -exchange (a W is exchanged between initial-state quarks) and annihilation (the initial quarks annihilate to a virtual W , which then decays) at the tree level, electromagnetic penguin at the loop level. Moreover, at loop level, mixing processes can take place.

Semileptonic B decays $B \rightarrow X_c \ell \nu$ and $B \rightarrow X_u \ell \nu$ (where X_c is any hadron coming from a $b \rightarrow c$ transition, X_u is any hadron coming from $b \rightarrow u$ transition and ℓ indicates either an electron or a muon) allow the determination of the CKM elements $|V_{cb}|$ and $|V_{ub}|$ respectively, because the strong interaction effects are much simplified due to the leptons in the final state. Both exclusive and inclusive decays can be used. The extraction of $|V_{cb}|$ and $|V_{ub}|$ is treated more extensively in the next paragraph.

Hadronic B decays instead are more complicated due to strong interaction effects caused by the surrounding cloud of light quarks and gluons. That leads to difficulties in the extraction of CKM elements, but on the other hand provides the

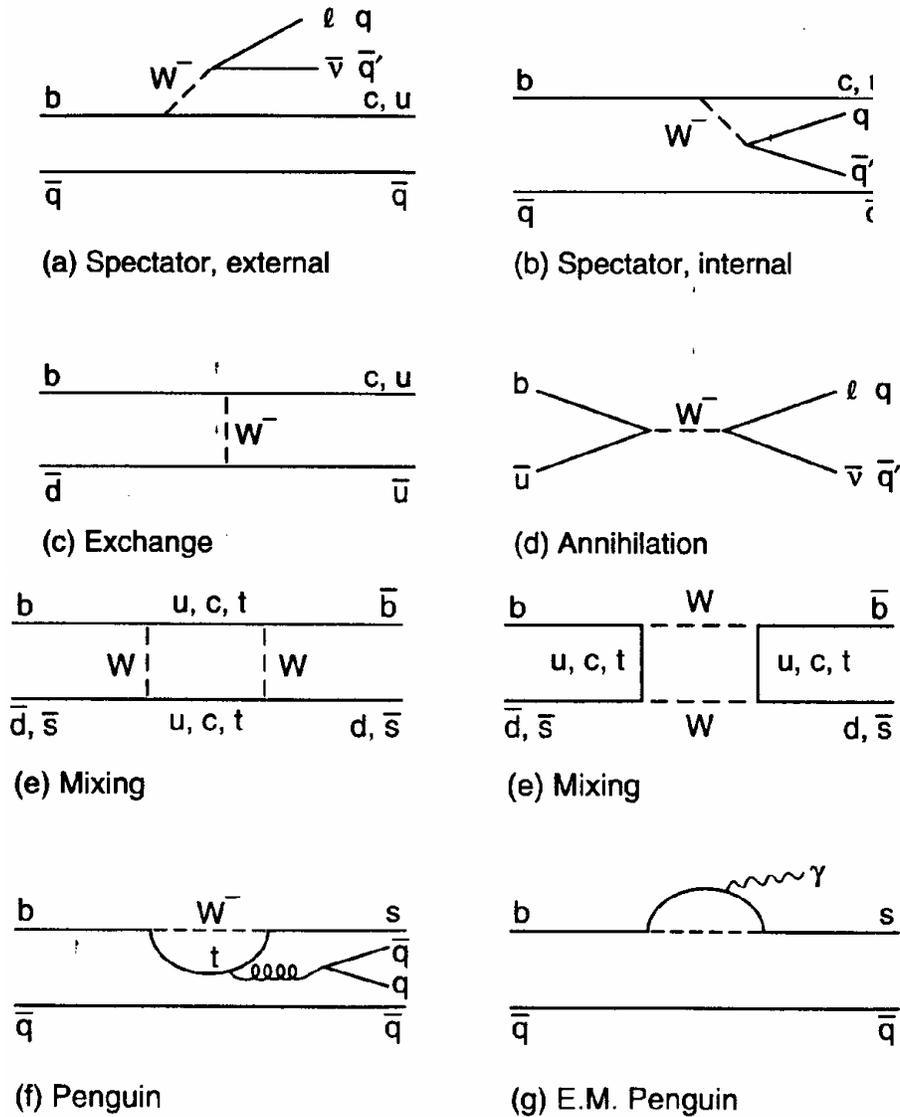


Figure 1.3. Examples of B decay modes allowed by the Standard Model: at the tree level (a), (b), (c), (d), at the loop level (e), (f), (g) and (h).

opportunity to study perturbative and non-perturbative QCD, hadronization, and Final State Interaction (FSI) effects. Most of the hadronic B decays involve $b \rightarrow c$ transitions at the quark level, resulting in a charmed hadron or charmonium in the final state.

1.5 Semileptonic decays: V_{cb} determination

Precision determinations of $|V_{cb}|$ and $|V_{ub}|$ are crucial to test the CKM sector [18] [19]: the length of the side of the unitarity triangle opposite to the angle β is proportional to the ratio $|V_{ub}|/|V_{cb}|$ and $|V_{cb}|$ normalises the whole unitary triangle (as shown in Fig. 1.1). At the moment, the most precise values of $|V_{cb}|$ and $|V_{ub}|$ are inferred from semileptonic decays $b \rightarrow cl\bar{\nu}_l$ and $b \rightarrow ul\bar{\nu}_l$, which can be determined from both inclusive and exclusive processes. Moreover, since the theoretical predictions are developed in the framework of the Heavy Quark Effective Theory (HQET) [20], the study of semileptonic decays represent a powerful test of the theory itself.

Data is provided by electron-positron machines, as LEP and CLEO, but above all by the B-Factories, BaBar and Belle. Due to the high production cross-section, semileptonic decays are copiously produced also at hadron colliders. The problem is that the measurement of $|V_{cb}|$ and $|V_{ub}|$ requires the reconstruction, in the b-hadron rest frame, of observables difficult to measure at hadron colliders (like the invariant mass squared of the lepton pair q^2). This difficult task at hadron colliders has been overcome at LHCb by exploiting the separation between primary and secondary vertices.

The determinations of V_{ub}/V_{cb} are highly dependent on the input from exclusive or inclusive processes, and on the theoretical or experimental framework.

The next sections treat the theoretical approaches of both inclusive and exclusive semileptonic decays, focusing on the $B \rightarrow X_c \ell \nu_\ell$ case which is more relevant for this thesis.

1.5.1 Exclusive decays

Exclusive determinations of $|V_{cb}|$ are based on studies of semileptonic B decays into charmed mesons D and D^* final states. In B decays, the approximations and techniques of the heavy quark effective theory (HQET) can be used: since the mass of the b quark is large compared to the QCD scale, a systematic expansion in powers of Λ/m_b can be performed, where $\Lambda \sim 500 - 700$ MeV is a hadronic scale of the order of Λ_{QCD} . Moreover, in $B \rightarrow D^{(*)}l\nu$ semileptonic decays also the mass of the c quark can be considered large compared to the QCD scale, so further approximations are allowed. In the limit of $m_b, m_c \rightarrow \infty$, the 6 form factors describing the $B \rightarrow D^{(*)}$ transitions are expressed in terms of a single form factor, the Isgur-Wise function [21] [22] $\xi(\omega)$, which depends on the product of the four-

velocities v and v' of the initial and final state hadrons respectively ($\omega = v \cdot v'$). For negligible lepton masses ($\ell = e, \mu$), the differential rates for the semileptonic decays are given by

$$\frac{d\Gamma}{d\omega}(B \rightarrow D\ell\nu) = \frac{G_F^2}{48\pi^3}(m_B + m_D)^2 m_D^3 (\omega^2 - 1)^{\frac{3}{2}} |V_{cb}|^2 \mathcal{G}^2(\omega) \quad (1.38)$$

$$\frac{d\Gamma}{d\omega}(B \rightarrow D^*\ell\nu) = \frac{G_F^2}{48\pi^3}(m_B - m_{D^*})^2 m_{D^*}^3 \chi(\omega) (\omega^2 - 1)^{\frac{1}{2}} |V_{cb}|^2 \mathcal{F}^2(\omega). \quad (1.39)$$

Every differential rate is written in terms of a single form factor, $\mathcal{G}(\omega)$ for the decay $B \rightarrow D\ell\nu$ and $\mathcal{F}(\omega)$ for $B \rightarrow D^*\ell\nu$. In Eq. (1.39), $\chi(\omega)$ is a phase space factor:

$$\chi(\omega) = (\omega + 1)^2 \left(1 + \frac{4\omega}{\omega + 1} \frac{m_B^2 - 2\omega m_B m_{D^*} + m_{D^*}^2}{(m_B - m_{D^*})^2} \right). \quad (1.40)$$

The Isgur-Wise function is normalised to unity at the so called zero recoil point $\omega = 1$, where the momentum transfer to the lepton pair is maximum and the D^* is at rest with respect to the B [23] [24]. At the zero recoil point, in the $m_b, m_c \rightarrow \infty$ limit, the light quarks of the initial and final hadrons are not affected by the $b \rightarrow c$ transition, so the hadronic quantum states completely overlap and

$$\mathcal{G}(1) = \mathcal{F}(1) = 1. \quad (1.41)$$

Including effects from finite quark masses, the current LQCD predictions are

$$\mathcal{F}(1) = 0.902 \pm 0.017, \quad (1.42)$$

$$\mathcal{G}(1) = 1.04 \pm 0.01 \pm 0.01. \quad (1.43)$$

The determination of $|V_{cb}|$ proceeds through two steps:

- The experimental fit of the products $\mathcal{F}(\omega)|V_{cb}|$ and $\mathcal{G}(\omega)|V_{cb}|$: many experiments have measured the differential rate as function of ω and these measurements are used as input to a four dimensional fit [17]. One of these four parameters is the product of the form factor and $|V_{cb}|$.
- The theoretical evaluation of the form factors $\mathcal{F}(\omega)$ and $\mathcal{G}(\omega)$, since data are not acquired at $\omega = 1$ because the zero recoil point is not accessible experimentally due to kinematical reasons. The form factors can be parametrised in different model-dependent ways: a lattice (LQCD) prediction takes into account the effects on the form factor due to finite quark masses [25] [26], while a non-lattice estimate is based on sum rules (SR) [27] [28].

The $B \rightarrow D^* \ell \mu$ channel gives more precise results with respect to the $B \rightarrow D \ell \mu$ channel, due to the higher rate and because the corrections to the HQET predictions are smaller. The measured values of $|V_{cb}|$ are the following [17]:

$$\begin{aligned} |V_{cb}| &= (39.48 \pm 0.50_{exp} \pm 0.74_{theo}) \times 10^{-3} (\overline{B} \rightarrow D^* \ell \bar{\nu}_\ell, \text{LQCD}) \\ |V_{cb}| &= (41.4 \pm 0.5_{exp} \pm 1.0_{theo}) \times 10^{-3} (\overline{B} \rightarrow D^* \ell \bar{\nu}_\ell, \text{SR}) \\ |V_{cb}| &= (39.3 \pm 1.4_{exp} \pm 1.3_{theo}) \times 10^{-3} (\overline{B} \rightarrow D \ell \bar{\nu}_\ell, \text{LQCD}) \\ |V_{cb}| &= (40.6 \pm 1.5_{exp} \pm 0.8_{theo}) \times 10^{-3} (\overline{B} \rightarrow D \ell \bar{\nu}_\ell, \text{non-lattice}). \end{aligned}$$

The determinations of CKM matrix element from the two decays are consistent and the uncertainties largely uncorrelated. The average of the two lattice estimates is

$$|V_{cb}| = (39.5 \pm 0.8) \times 10^{-3} \text{ (exclusive)}.$$

1.5.2 Inclusive decays

In the inclusive $B \rightarrow X_q \ell \nu$ decays, X_q is a hadronic state originated by the quark q that could be a single-particle state as well as a multi-particle state. In the limit of large b quark mass, the wavelength associated with the b quark decay are considered too short to interfere with the hadronization process. Thus the decay process could be distinguished into two separated steps, the heavy quark decay and the final hadronisation, and quark-hadron duality is generally assumed. This assumption allows to use the Operator Product Expansion (OPE) approach, which, combined with HQET, gives the inclusive transition rate in the form of a heavy quark expansion [29] [30] in terms of the strong coupling constant α_s and the inverse of the b quark mass $1/m_b$. It is schematically written as:

$$\begin{aligned} \Gamma &= |V_{cb}|^2 \hat{\Gamma}_0 m_b^5(\mu) (1 + A_{ew}) \times \\ &\left[z_0^{(0)}(r) + \frac{\alpha_s(\mu)}{\pi} z_0^{(1)}(r) + \left(\frac{\alpha_s(\mu)}{\pi} \right)^2 z_0^{(2)}(r) + \dots \right. \\ &+ \frac{\mu_\pi^2}{m_b^2} \left(z_0^{(2)}(r) + \frac{\alpha_s(\mu)}{\pi} + z_2^{(1)} + \dots \right) \\ &+ \frac{\mu_G^2}{m_b^2} \left(y_0^{(2)}(r) + \frac{\alpha_s(\mu)}{\pi} + y_2^{(1)} + \dots \right) \\ &+ \frac{\rho_D^3}{m_b^3} \left(z_0^{(3)}(r) + \frac{\alpha_s(\mu)}{\pi} + z_3^{(1)} + \dots \right) \\ &\left. + \frac{\rho_{LS}^3}{m_b^3} \left(y_0^{(3)}(r) + \frac{\alpha_s(\mu)}{\pi} + y_3^{(1)} + \dots \right) + \dots \right], \end{aligned} \quad (1.44)$$

where A_{ew} denotes electroweak corrections, r is the ratio m_c/m_b and y_i and z_i are known functions. The coefficients of this double expansion, denoted as Wilson coefficients, are computed in the framework of perturbation theory. The other terms appearing in (1.44), like μ_π , μ_G , ρ_D , ρ_{LS} , are matrix elements of local operators (kinetic energy of the b quark, chromomagnetic operator, Darwin term and spin-orbit term, respectively), whose expectation values encode the non-perturbative corrections. In summary, the OPE gives an expansion in terms of local operators that allows to separate short and long distance dynamics.

To calculate the expectation value of these operators, it is necessary to rely on a number of HQET parameters, which increase with powers of $1/m_b$. Writing the OPE expansion for orders up to $1/m_b^3$:

$$\begin{aligned}\mu_\pi^2 &= -\langle B|\bar{b}(iD_\perp)^2b|B\rangle, \\ \mu_G^2 &= \langle B|\bar{b}(iD_\perp^\mu)(iD_\perp^\nu)\sigma_{\mu\nu}b|B\rangle, \\ \rho_D^3 &= \langle B|\bar{b}(iD_{\perp\mu})(i\nu D)(iD_\perp^\nu)b|B\rangle, \\ \rho_{SL}^3 &= \langle B|\bar{b}(iD_\perp^\mu)(i\nu D)(iD_\perp^\nu)\sigma_{\mu\nu}b|B\rangle.\end{aligned}\tag{1.45}$$

These HQET parameters appear in different inclusive B meson observables that can be measured experimentally, such as the moments of the distributions of charged lepton energy and invariant mass of the hadronic system. By measuring spectra and as many moments as possible, a global fit can be performed, that is a simultaneous fit to HQET parameters, quark masses and absolute value of the CKM matrix element $|V_{cb}|$.

Theoretical calculations are available in various normalisation schemes. Global fits to extract $|V_{cb}|$ are based on two different schemes:

- the 1S scheme [27], which relates the b quark mass to the perturbative expression for the mass of the 1S state of the Υ system
- the kinetic scheme [31], which introduces a "kinetic mass", that is the mass entering the non-relativistic expression for the kinetic energy of a heavy quark.

Both of these schemes have been applied to semileptonic $b \rightarrow c$ transitions, yielding comparable results and uncertainties [17]:

$$\begin{aligned}|V_{cb}| &= (42.42 \pm 0.86) \times 10^{-3} \text{ (kinetic scheme)} \\ |V_{cb}| &= (41.96 \pm 0.45 \pm 0.07) \times 10^{-3} \text{ (1S scheme)}.\end{aligned}$$

1.5.3 Discrepancy

The determination of $|V_{cb}|$ from $B \rightarrow D^* \ell \nu_\ell$ decays has a relative precision of about 2%, with the main uncertainty due to the determination of the form factor, while the evaluation of the matrix element from $B \rightarrow D \ell \nu_\ell$ decays has a relative precision of about 5%. Inclusive decays provide a determination of $|V_{cb}|$ with a relative uncertainty of about 2%, which is mainly a consequence of uncertainties due to higher order perturbative and non-perturbative corrections. The exclusive and inclusive determinations of $|V_{cb}|$ are the following:

$$\begin{aligned} |V_{cb}| &= (42.2 \pm 0.7) \times 10^{-3} \text{ (inclusive) } , \\ |V_{cb}| &= (39.5 \pm 0.8) \times 10^{-3} \text{ (exclusive) } . \end{aligned}$$

There is a poor consistency of the values of $|V_{cb}|$ determined from inclusive and exclusive semileptonic decays, which differ at the level of about 3 sigma.

1.6 Charmed mesons

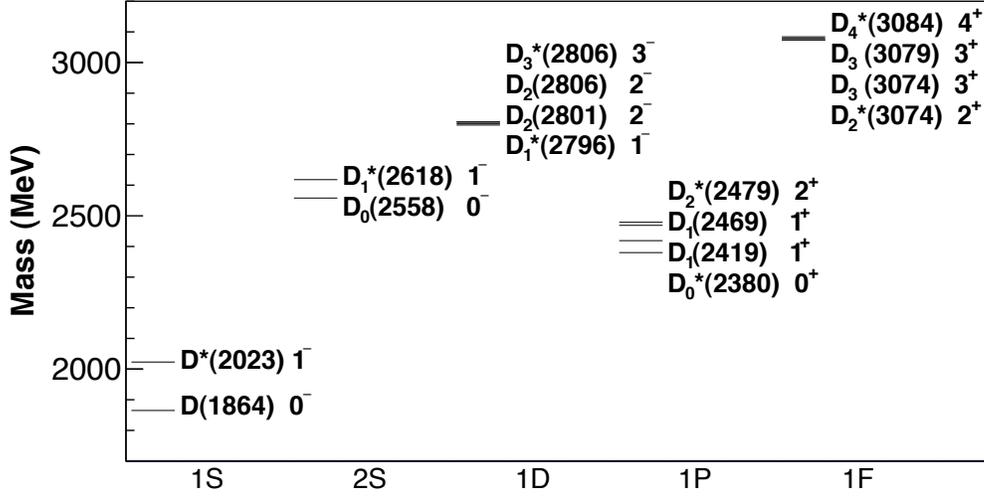
In this Section the present knowledge on the D meson spectrum and the semileptonic decays to charm final states is presented. Improving the current knowledge of semileptonic B decays to excited charm states is important for several reasons. First of all, such decays represent a background and therefore a systematic uncertainty in the determination of $|V_{cb}|$ with $B \rightarrow D^* \ell \nu_\ell$ decays.

Secondly, they could explain a long-standing discrepancy between the measured inclusive semileptonic rate and the sum of the known exclusive decays (Section 1.6.2).

Finally, they represent a background source in measurements of semitauonic decays, such as $B \rightarrow D^{(*)} \tau \nu$, aimed at testing the lepton flavour universality and its violation [32], which could indicate physics processes beyond the Standard Model.

1.6.1 Spectrum

Excited states of the D meson, denoted as D_J (where J is the total angular momentum of the meson), are excitations of quark-antiquark systems containing one charmed and one light (u, d) quark. The spectrum is still not well known: many of the predicted states [33] have not yet been observed experimentally. The expected spectrum for the $c\bar{u}$ system is shown in Figure 1.4 (the spectrum for the

Figure 1.4. HQET $c\bar{u}$ spectrum predictions.

$c\bar{d}$ isospin counterpart is analogous). In the heavy-quark limit, the c quark spin \vec{s}_c decouples from the other degrees of freedom, and the total angular momentum of the light quark, which is the sum of orbital momentum (\vec{L}) and light quark spin (\vec{s}_q) $\vec{j}_q = \vec{L} + \vec{s}_q$, is a good quantum number to describe the meson. The total angular momentum of the light quark is then combined with the spin S_Q of the c quark in order to obtain the total angular momentum J of the system. The mesons are usually classified according to the radial quantum number and to the eigenvalue L , as in Figure 1.4. They are also distinguished between natural parity states and unnatural parity states: the former have parity $P = (-1)^J$ and therefore $J^P = 0^+, 1^-, 2^+, \dots$, while the latter have $J^P = 0^-, 1^+, 2^-, \dots$. The low mass spectrum is formed by the ground states $1S$ ($L = 0$), the orbital excitations with angular momentum $L = 1, 2$ ($1P, 1D$), and the first radial excitation $2S$.

The ground states of the D meson are obtained combining the spin $S_Q = 1/2$ of the c quark with the total angular momentum of the light quark, $j_q = s_q + L = 1/2$, since $L = 0$. That gives a pseudoscalar state $J^P = 0^-$, labelled as D^\pm or D^0 , and a vector state $J^P = 1^-$, denoted as $D^{*\pm}$ or D^{*0} . The properties and decay modes of the two states are collected in Table 1.2.

There are four states with $L = 1$, arranged in two doublets, with $j_q = 1/2$ and $j_q = 3/2$. The resulting states are labelled as follows: $J^P = 0^+$ ($j_q = 1/2$) as D_0^* , $J^P = 1^+$ ($j_q = 1/2$) as D_1' , $J^P = 1^+$ ($j_q = 3/2$) as D_1 and $J^P = 2^+$ ($j_q = 3/2$) as D_2^* , and collectively referred to as D^{**} .

The conservation of parity and angular momentum in strong interactions imposes

Table 1.2. Quantum numbers, masses, widths and decay modes of the D meson ground states [17]. The negative states decay modes are the charged-conjugated of the positive states decay modes here reported.

State	J^P	j_q	Mass (MeV)	Width (MeV)	Decay Mode
D^0	0^-	$1/2$	1864.86 ± 0.13	-	-
D^+	0^-	$1/2$	1869.62 ± 0.15	-	-
D^{*0}	1^-	$1/2$	2006.99 ± 0.15	< 2.1	$D^0\pi^0, D^0\gamma$
D^{*+}	1^-	$1/2$	2010.29 ± 0.13	0.096 ± 0.022	$D^0\pi^+, D^+\pi^0, D^+\gamma$

constraints on the strong decays of D_J states to $D\pi$ and $D^*\pi$ final states. The $j = 1/2$ states are predicted to decay only through a S-wave: $D_0^* \rightarrow D\pi$ and $D_1' \rightarrow D^*\pi$. The $j = 3/2$ states are expected to decay through a D-wave: $D_1 \rightarrow D^*\pi$ and $D_2^* \rightarrow D\pi$ and $D^*\pi$. Due to the finite c -quark mass, the two $J^P = 1^+$ states may be mixtures of the $j = 1/2$ and $j = 3/2$ states. Thus the D_1' state may decay via a D-wave and the D_1 state may decay via an S-wave. Since the $j = 1/2$ states with $L = 1$ decay through an S-wave, they are expected to be wide (hundreds of MeV/c^2), while the $j = 3/2$ states, decaying through a D-wave, are expected to be narrow (20-40 MeV/c^2) [34] [35].

Due to their narrow widths, the $D_1(2420)$ and the $D_2^*(2460)$ have been observed and studied by many experiments (ARGUS [36], CLEO [37], OPAL [38], ALEPH [39], DELPHI [40], D0 [41], BELLE [42]) and they are well established; their masses agree with model predictions. The broad $D_0^*(2400)$ and $D_1'(2430)$ have been established by the Belle and BaBar experiments in exclusive B decays [42] [43]. The Belle Collaboration studied the hadronic decays of charged B mesons to $D^\pm\pi^\mp\pi^\mp$ and $D^{*\pm}\pi^\mp\pi^\mp$ final states. They managed to determine the contributions of two-body $B \rightarrow D^{**}\pi$ decays with both narrow and broad D^{**} states and to measure their masses and widths, leading to the first observation of the broad D_0^{*0} and $D_1'^0$ mesons. The BaBar Collaboration reconstructed the three-body decays $B^- \rightarrow D^+\pi^-\pi^-$ and measured their branching fractions, observing the established D_2^{*0} and confirming Belle's first observation of the D_0^{*0} resonance via its decay to $D^+\pi^-$. The properties and decay modes of the D^{**} states are collected in Table 1.3.

Further excited states of D mesons have been observed by the BaBar and LHCb collaborations. BaBar has identified, for the first time, candidates for the radial excitations of the D^0 , D^{*0} and D^{*+} and $L = 2$ excited states of D^0 and D^+ , analysing the inclusive production of the $D^+\pi^-$, $D^0\pi^+$ and $D^{*+}\pi^-$ final states in the inclusive reaction $e^+e^- \rightarrow c\bar{c} \rightarrow D^{(*)}\pi X$ [44]. LHCb, studying the final states $D^+\pi^-$, $D^0\pi^+$ and $D^{*+}\pi^-$, has been able to observe the $D_1(2420)^0$ and $D_2^*(2460)$ resonances and,

Table 1.3. Quantum numbers, masses, widths and decay modes of the D^{**} states [17]. The negative states decay modes are the charged-conjugated of the positive states decay modes here reported.

State	J^P	j_q	Mass (MeV)	Width (MeV)	Decay Mode
D_0^{*0}	0^+	$1/2$	2318 ± 29	267 ± 40	$D^+ \pi^-$
D_0^{*+}	0^+	$1/2$	$2403 \pm 14 \pm 35$	$283 \pm 24 \pm 34$	$D^0 \pi^+$
$D_1^{0'}$	1^+	$1/2$	$2427 \pm 26 \pm 25$	$348_{-75}^{+107} \pm 74$	$D^{*+} \pi^-$
$D_1^{+'}$	1^+	$1/2$	-	-	$D^{*0} \pi^+$
D_1^0	1^+	$3/2$	2420.9 ± 0.8	27.1 ± 2.7	$D^{*+} \pi^-, D^0 \pi^+ \pi^-$
D_1^+	1^+	$3/2$	2423.4 ± 3.1	25 ± 6	$D^{*0} \pi^+, D^+ \pi^+ \pi^-$
D_2^{*0}	2^+	$3/2$	2461.8 ± 0.8	49.0 ± 1.4	$D^+ \pi^-, D^{*+} \pi^-$
D_2^{*+}	2^+	$3/2$	2464.4 ± 1.9	37 ± 6	$D^0 \pi^+, D^{*0} \pi^+$

in addition, two natural and two unnatural parity resonances between 2.4 and 2.8 GeV/ c^2 . Moreover, LHCb studied the properties and spin-parity assignments for the $D_0^*(2400)^0$, the $D_1(2420)^0$, the $D_2^*(2460)^0$ and the $D_2^*(2460)^\pm$ resonances [45] [46]

1.6.2 Measured branching fractions of semileptonic decays

The measured inclusive semileptonic branching ratios of B^+ and B^0 mesons are collected in Table 1.4, while the exclusive branching ratios $B \rightarrow D^{(*)} \ell^+ \nu_\ell$ in Table 1.5. The tables show that the exclusive decays $B \rightarrow D^{(*)} \ell^+ \nu_\ell$ account for about 70% of the total inclusive rate. The contribution of other states is not yet well measured.

The current results for $D^* \pi$ final states, which include resonant decays resulting in a $D^{(*)} \pi$ pair and non-resonant decays, are shown in Table 1.6. Resonant exclusive decays to $D^{(*)} \pi$ final states are gathered in Table 1.7: these results show that the resonant decays account for the biggest fraction of semileptonic B decays to $D^{(*)} \pi$ pair in the final state and contribute for about 2% to the inclusive rate (under the assumptions of isospin invariance and saturation of the D^{**} decay rate with two body decay only).

Even considering those contributions does not solve the so called "gap problem", i.e. the discrepancy between the sum of exclusive rates and the inclusive semileptonic rate. The sum of exclusive modes with a $D^{(*)}$ and with a $D^{(*)} \pi$ pair in the final state is $(8.9 \pm 0.5)\%$ for B^\pm and $(8.19 \pm 0.28)\%$ for B^0 , which, in both cases, is smaller than the corresponding inclusive rate. Thus, other decay modes have to be considered and studied to reduce this discrepancy.

Table 1.4. Measured inclusive semileptonic branching fractions of B mesons. The (\dagger) indicates PDG averages [17]. ℓ refers to either an electron or a muon.

Channel	\mathcal{B}_{SL} (%)	Experiments
$B^+ \rightarrow \ell^+ \nu_\ell X_c$	$10.79 \pm 0.25 \pm 0.27$	Belle
$B^0 \rightarrow \ell^+ \nu_\ell X_c$	$10.08 \pm 0.30 \pm 0.22$	Belle
$B^\pm/B^0 \rightarrow \ell \nu_\ell X_c$	$10.58 \pm 0.15^{(\dagger)}$	BaBar, Belle

Table 1.5. Measured exclusive branching fractions of semileptonic B meson decays to $D^{(*)}$ mesons. The (\dagger) indicates PDG averages [17]. ℓ refers to either an electron or a muon. The last two rows give the sum of the branching ratios to D and D^* final states for charged and neutral B .

Channel	\mathcal{B}_{SL} (%)	Experiments
$B^+ \rightarrow \bar{D}^0 \ell^+ \nu_\ell$	$2.29 \pm 0.08^{(\dagger)}$	BaBar, CLEO2, CLEO
$B^0 \rightarrow D^- \ell^+ \nu_\ell$	$2.18 \pm 0.12^{(\dagger)}$	BaBar, Belle, CLEO2, ALEPH
$B^+ \rightarrow \bar{D}^{*0} \ell^+ \nu_\ell$	$5.58 \pm 0.26^{(\dagger)}$	BaBar, CLEO2, ARGUS
$B^0 \rightarrow D^{*-} \ell^+ \nu_\ell$	$5.09 \pm 0.22^{(\dagger)}$	BaBar, Belle, DELPHI, CLEO2, OPAL, ALEPH
Tot $B^+ \rightarrow \bar{D}^{(*)0} \ell^+ \nu_\ell$	7.87 ± 0.27	-
Tot $B^0 \rightarrow D^{(*)-} \ell^+ \nu_\ell$	7.27 ± 0.26	-

Recently, BaBar provided a first determination of the inclusive semileptonic branching fractions of B decaying into D or D^* and two additional pions, by measuring the branching ratios relative to the topologically similar decays $B \rightarrow D^{(*)} \ell \nu_\ell$: $R_{n\pi}^{(*)} = B(\bar{B} \rightarrow D^{(*)}(n\pi)\ell\bar{\nu})/B(\bar{B} \rightarrow D^{(*)}\ell\bar{\nu})$, where $n = 1, 2$ and ℓ is either

Table 1.6. Measured exclusive branching fractions of semileptonic B meson decays to $D^{(*)}\pi$ final states. The (\dagger) indicates PDG averages [17]. ℓ refers to either an electron or a muon. The last two rows give the sum of the branching ratios to D and D^* final states for charged and neutral B .

Channel	\mathcal{B}_{SL} (%)	Experiments
$B^+ \rightarrow D^- \pi^+ \ell^+ \nu_\ell$	$0.42 \pm 0.05^{(\dagger)}$	BaBar, Belle
$B^0 \rightarrow \bar{D}^0 \pi^- \ell^+ \nu_\ell$	$0.43 \pm 0.06^{(\dagger)}$	BaBar, Belle
$B^+ \rightarrow D^{*-} \pi^+ \ell^+ \nu_\ell$	$0.61 \pm 0.06^{(\dagger)}$	BaBar, Belle
$B^0 \rightarrow \bar{D}^{*0} \pi^- \ell^+ \nu_\ell$	$0.49 \pm 0.08^{(\dagger)}$	BaBar, Belle
Tot $B^+ \rightarrow D^{(*)-} \pi^+ \ell^+ \nu_\ell$	1.03 ± 0.08	-
Tot $B^0 \rightarrow \bar{D}^{(*)0} \pi^- \ell^+ \nu_\ell$	0.92 ± 0.10	-

Table 1.7. Measured exclusive branching fractions of semileptonic B meson decays to $D^{(*)}\pi$ final states through D^{**} resonances. The (\dagger) indicates PDG averages [17], while the $(\dagger\dagger)$ HFAG averages [47]. ℓ states for either an electron or a muon.

Channel	$\mathcal{B}_{SL} \times \mathcal{B}(D^{**} \rightarrow D^{(*)}\pi)$ (%)	Experiments
$B^+ \rightarrow \bar{D}_1^0(2420)(\rightarrow D^{*-}\pi^+)\ell^+\nu_\ell$	0.285 ± 0.018 ($\dagger\dagger$)	BaBar, Belle, D0, CLEO, OPAL, ALEPH
$B^0 \rightarrow D_1^-(2420)(\rightarrow \bar{D}^{*0}\pi^-)\ell^+\nu_\ell$	0.280 ± 0.028 (\dagger)	BaBar, Belle
$B^+ \rightarrow \bar{D}_2^{*0}(2460)(\rightarrow D^-\pi^+)\ell^+\nu_\ell$	0.153 ± 0.016 (\dagger)	BaBar, Belle
$B^+ \rightarrow \bar{D}_2^0(2460)(\rightarrow D^{*-}\pi^+)\ell^+\nu_\ell$	0.074 ± 0.007 ($\dagger\dagger$)	BaBar, Belle, D0, CLEO
$B^0 \rightarrow D_2^{*-}(2460)(\rightarrow \bar{D}^0\pi^-)\ell^+\nu_\ell$	0.121 ± 0.033 (\dagger)	BaBar, Belle
$B^0 \rightarrow D_2^{*-}(2460)(\rightarrow \bar{D}^{*0}\pi^-)\ell^+\nu_\ell$	0.068 ± 0.012 (\dagger)	BaBar, Belle
$B^+ \rightarrow \bar{D}_1^0(2430)(\rightarrow D^{*-}\pi^+)\ell^+\nu_\ell$	0.13 ± 0.04 ($\dagger\dagger$)	BaBar, Belle, DELPHI
$B^0 \rightarrow D_1^{\prime-}(2430)(\rightarrow \bar{D}^{*0}\pi^-)\ell^+\nu_\ell$	0.31 ± 0.09 (\dagger)	BaBar
$B^+ \rightarrow \bar{D}_0^{*0}(2400)(\rightarrow D^{*-}\pi^+)\ell^+\nu_\ell$	0.29 ± 0.05 ($\dagger\dagger$)	BaBar, Belle
$B^0 \rightarrow D_0^{*-}(2400)(\rightarrow \bar{D}^0\pi^-)\ell^+\nu_\ell$	0.30 ± 0.12 (\dagger)	BaBar, Belle

an electron or a muon. The results, obtained averaging the ratios from neutral and charged B decays, are:

$$R_{1\pi} = 0.187 \pm 0.011 \pm 0.009 \quad R_{1\pi}^* = 0.123 \pm 0.006 \pm 0.006,$$

$$R_{2\pi} = 0.073 \pm 0.014 \pm 0.008 \quad R_{2\pi}^* = 0.021 \pm 0.006 \pm 0.003,$$

where the first uncertainty is statistical and the second systematic. From these results the following branching fractions can be determined:

$$\mathcal{B}(B \rightarrow D\pi^\pm l\nu_\ell) = (0.4476 \pm 0.0256 \pm 0.0206 \pm 0.0225)\%,$$

$$\mathcal{B}(B \rightarrow D^*\pi^\pm l\nu_\ell) = (0.6108 \pm 0.0277 \pm 0.0302 \pm 0.0136)\%,$$

$$\mathcal{B}(B \rightarrow D\pi^+\pi^- l\nu_\ell) = (0.1752 \pm 0.0332 \pm 0.0194 \pm 0.0088)\%,$$

$$\mathcal{B}(B \rightarrow D^*\pi^+\pi^- l\nu_\ell) = (0.1095 \pm 0.0307 \pm 0.0149 \pm 0.0024)\%,$$

where the first uncertainty is statistical, the second systematic and the third comes from the branching fraction for the normalisation mode. The BaBar analysis reduces the discrepancy but does not account for the entire gap.

Another theoretical problem is represented by the "1/2 vs 3/2 puzzle": theoretical

calculations indicate that B decays should have a smaller rate to the $j_q = 1/2$ D^{**} doublet than to the $j_q = 3/2$ doublet, in contrast to experimental results. These puzzles may be solved by e.g. allowing a significant rate into D_0^* [49].

In general, a better understanding of semileptonic B decays would allow to solve these issues, as well as reduce systematic uncertainties in other measurements, e.g. on the $|V_{cb}|$ matrix element and semitauonic $B \rightarrow D^{(*)}\tau\nu$ decays.

Moreover, semileptonic B decays allow an experimental determination of hadronic form factors, which, despite its high importance, is yet to be done. There are predictions from theoretical models but there are no measurements so far.

Chapter 2

The LHCb Detector

In this chapter, the LHCb detector is described. An overview of its components is given, along with its performance in terms of charged particle tracking and identification.

LHCb (Large Hadron Collider beauty) is one of the four big experiments at CERN Large Hadron Collider (LHC). LHCb was designed and built to study CP Violation and rare decays of hadrons containing b and c quarks, having as main purpose the search of indirect evidence for physics processes beyond the Standard Model.

LHCb is a single-arm spectrometer with a forward angular coverage from approximately 10 mrad to 300 mrad in the horizontal plane and from 10 mrad to 250 mrad in the vertical plane. The geometry of the detector has been defined according to the $b\bar{b}$ production mechanism. At high energy the pp collision gives rise to a deep inelastic scattering process. The protons break up and virtual highly boosted gluons are produced. These virtual gluons convert into $b\bar{b}$ pairs, which hadronize, in the same forward or backward cone. The b-hadrons produced are highly boosted in the laboratory frame. They are unstable and decay after travelling typically 10 mm in the laboratory frame.

During the data taking in 2011 and 2012 (Run I of the LHC) the luminosity of LHCb was $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ in 2011 and raised to $4 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ in 2012 [50]. LHCb ran at a lower luminosity with respect to the one achievable at the LHC because that allows to have a lower event multiplicity, i.e. a lower number of pp interactions per bunch crossing, and a lower occupancy of the detector, as well as to reduce radiation damage and to allow good trigger performances. The luminosity of LHCb can be tuned by changing the beam separation at its interaction point, independently from the other interaction points. The measured $b\bar{b}$ production

cross section at a center of mass energy of 7 TeV is $\sigma_{b\bar{b}} = (284 \pm 20 \pm 49)\mu\text{b}$ [51]; the angular distribution of $b\bar{b}$ is shown in Figure 2.1, from which it can be noticed that the pairs are collinear and produced mostly at small angles with respect to the beam axis. This explains why the detector, centred around the LHC beampipe 100 m underground, has a specific forward geometry, whereas the backward region is not instrumented due to space constraints.

The layout of LHCb is shown in Figure 2.2, where the reference right-handed

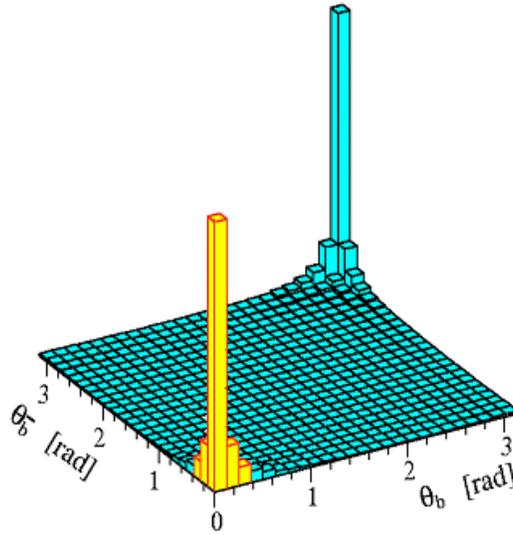


Figure 2.1. Angular distribution of the $b\bar{b}$ pairs at the LHC. The axes show the polar angle of b and \bar{b} with respect to the beam axis.

frame adopted has the z axis along the beam axis with positive values towards the end of the detector and the y axis along the vertical axis. The components of the detector are the VERtix LOcator (VELO), two Ring Imaging CHerenkov detectors RICH1 and RICH2, a dipole magnet, the tracking system composed by the TT, T1, T2 and T3 stations, an Electromagnetic CALorimeter (ECAL), a Hadronic CALorimeter (HCAL), a muon detector composed by the M1, M2, M3, M4 and M5 stations.

2.1 VERtix LOcator

The VERtix LOcator [52] is the LHCb detector closest to the interaction point. The aim of the VELO is to:

- identify the primary pp interaction vertex;

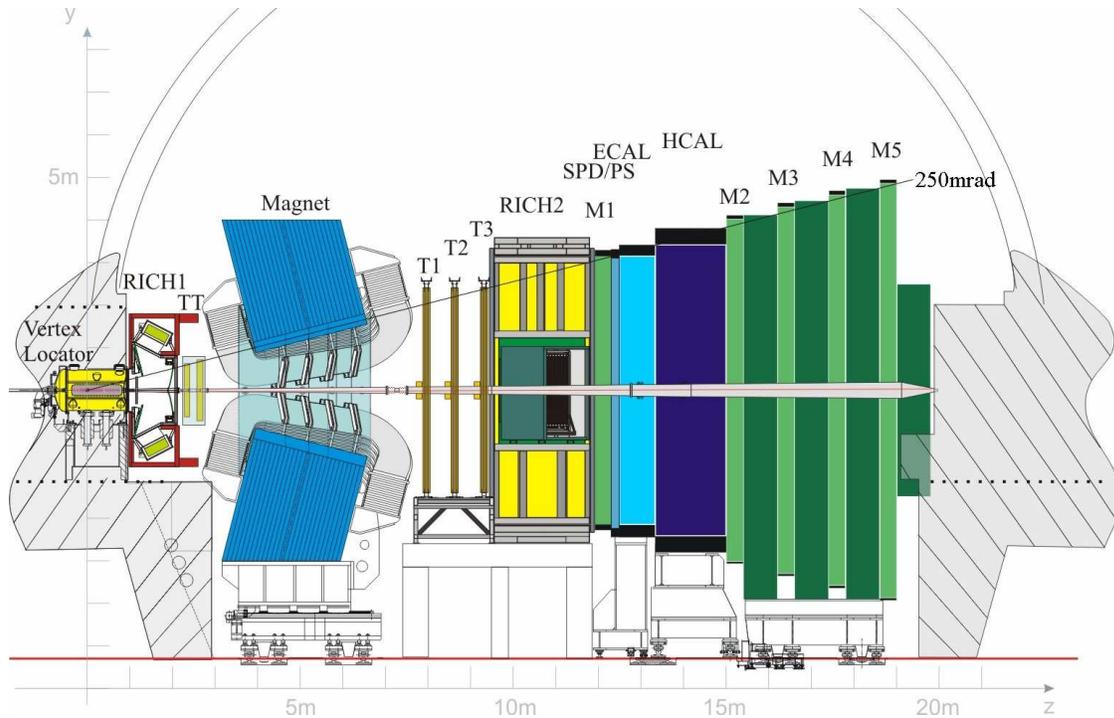


Figure 2.2. View of the LHCb detector.

- identify the secondary detached vertices of long lived hadrons;
- provide informations to the tracking system in order to measure the tracks' momenta.

The reconstruction of secondary vertices is fundamental at LHCb, since both beauty- and charm-hadrons are long-lived particles; the vertexing provides an accurate measurement of the decay distance (needed for time dependent studies) and of the impact parameters of daughter particles with respect to both primary and secondary vertices (used to identify B hadron decays products).

The VELO consists of 25 silicon stations arranged along the beam direction: each station is formed by two kind of sensors ($300\ \mu\text{m}$ thick plates of silicon microstrips) coupled together, called r and ϕ sensors. The r sensor gives the radial distance from the beam axis, while the ϕ sensor gives the azimuthal coordinate around the beam. The position of the sensor plane on the z axis gives the third coordinate to reconstruct tracks and vertices positions in the three dimensional space. The sensors are highly radiation tolerant, as they are placed at only 8 mm from the beam axis.

The modules are located at a radial distance from the beam smaller than the

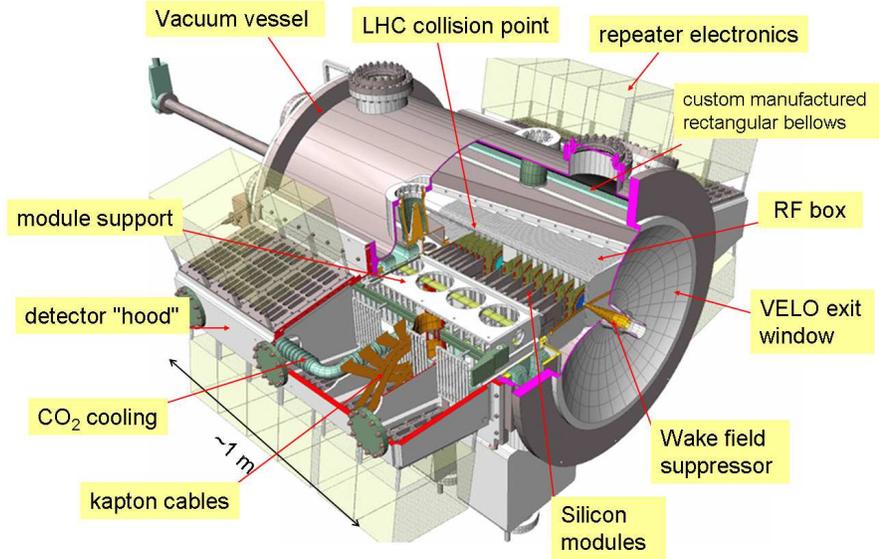


Figure 2.3. View of the VELO.

aperture needed during the beam injection phase. Thus, each plate is divided in two halves that can assume two possible configurations, a fully open position (during injection) and a fully closed position (when beams are stable).

The modules are placed in a vessel which maintains the vacuum around them and which is separated from the beam vacuum by a thin aluminum sheet, in order to minimise the amount of material crossed by the particles before the sensors.

Two planes, called pile-up veto system, are located upstream of the VELO sensors with the aim of suppressing events with more than two proton-proton interactions per bunch crossing.

The VELO detector covers the angular acceptance¹ of $1.6 < \eta < 4.9$ and achieves a spatial resolution on vertices of about $60 \mu\text{m}$ along the z axis and $10 \mu\text{m}$ along x and y axis; the spatial resolution on impact parameters is $20 \mu\text{m}$ for high momentum tracks and the spatial resolution on the decay length is between 200 and $370 \mu\text{m}$.

2.2 Magnet

A dipole magnet [55] is used to bend charged particles and measure their momenta. It is located after the VELO and in between the two RICH detectors (Figure 2.2).

¹The angular acceptance here is expressed in terms of the pseudorapidity η , defined as $-\ln(\tan \frac{\theta}{2})$, where θ is the polar angle with respect to the positive direction of the beam axis z .

It provides an integrated magnetic field $\int Bdl = 4 \text{ Tm}$ for tracks of 10 m length, the magnet weighs 1600 tonnes with an excitation current of 2.6 MA. Acceptances in the horizontal and vertical planes are 300 mrad and 500 mrad. A view of the

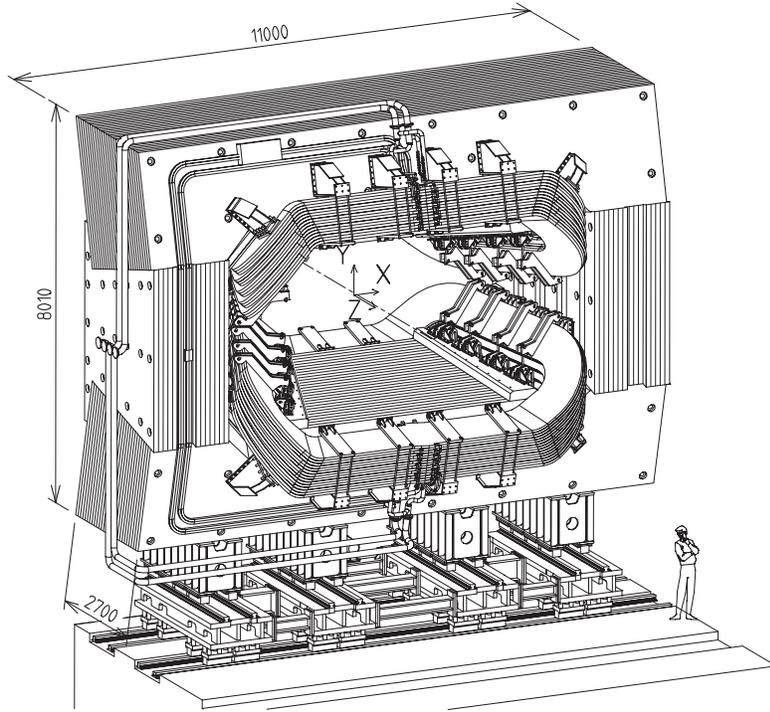


Figure 2.4. View of the dipole magnet.

magnet is shown in Figure 2.4: it is formed by two identical trapezoidal coils, bent at 45° on the two transverse sides and mounted inside an iron yoke. The two coils are each 7.5 m long, 4.6 m wide and 2.5 m high; they are made of aluminum and contain 225 loops each. The magnet gap is shaped to follow the detector acceptance. The magnet weighs 1600 tons and the maximum current is 6.6 kA. The magnetic field is along the y axis, so that particles are deflected in the $x - z$ plane. In order to minimise systematic uncertainties in CP violation and other asymmetry measurements, the polarity of the magnetic field can be reversed periodically. The strength of the magnetic field is monitored in the whole volume of the magnet, in the region of VELO and tracking system and inside the shielding of both RICH detectors, covering a distance of over 9 m and most of the acceptance region of the spectrometer. The field map obtained has a precision of about 4×10^{-4} .

2.3 Tracking system

The LHCb tracking system is formed by four planar tracking stations: a Tracker Turicensis or Trigger Tracker (TT) in front of the magnet and three tracking stations behind the magnet (T1-T3). Like the VeLo, the TT uses silicon microstrip detectors. The T1-T3 stations are made of silicon microstrips in the inner parts (Inner Tracker, IT) and straw-tubes in the outer parts (Outer Tracker, OT). The tracking system is used to reconstruct the trajectories of charged particles and measure their momenta. The momentum resolution provided by the tracking system is $\Delta p/p \sim 0.4\%$ for $p < 200$ GeV. The tracking system also exploits informations from the VeLo as already mentioned before.

2.3.1 Trigger Tracker and Inner Tracker

The Tracker Turicensis or Trigger Tracker (TT) and the Inner Tracker (IT) are the two sub-detectors composing the Silicon Tracker (ST). The IT is located upstream the dipole magnet, while the TT is placed downstream the magnet; both use silicon microstrip sensors.

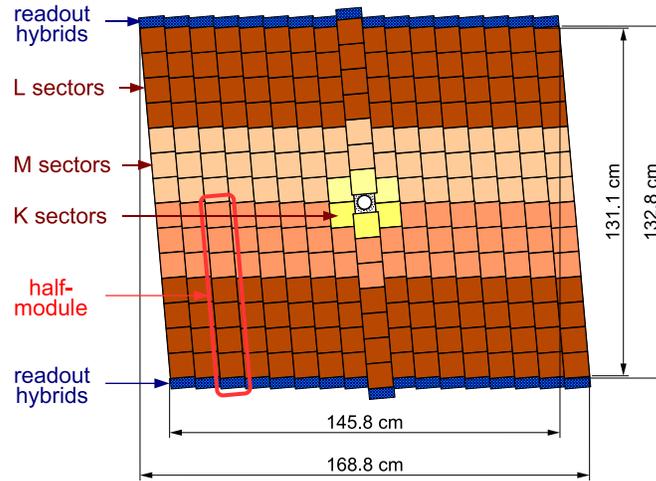


Figure 2.5. View of the TT.

The Tracker Turicensis is a planar tracking station with four detection layers: the silicon strips are arranged such that in the first and fourth layers they are parallel to the y axis, in the second and third layers the strips are rotated of $\pm 5^\circ$, in order to reconstruct the tracks in three dimensions. The layers are separated into

two pairs placed about 27 cm far from each other along the beam axis. The layout of one of the detector layers is shown in Figure 2.5.

The station covers the full acceptance of LHCb and has an active area of about 8.4 m².

The Inner Tracker [53] consists of three stations with four detector boxes each, arranged around the beam axis as depicted in Figure 2.6. Inside each detection box there are four detection layers, two vertical and two stereo: each of them has seven detector modules.

The IT has an active area of about 4.0 m².

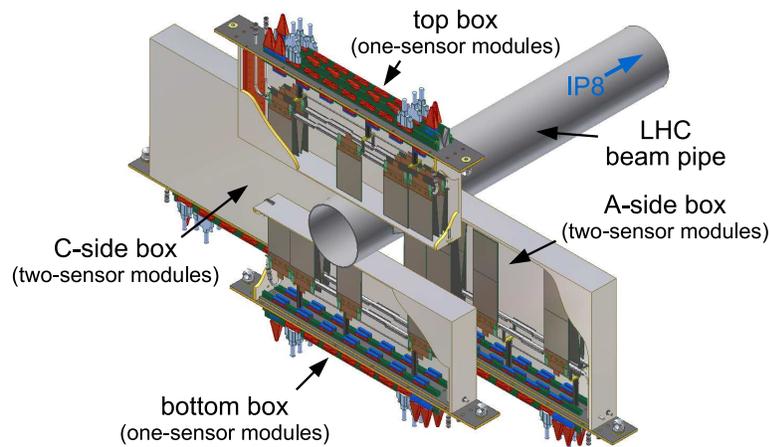


Figure 2.6. View of the four detection boxes of a IT station.

2.3.2 Outer Tracker

The Outer Tracker (OT) [54] is a drift-time detector, to track and measure the momentum of charged particles on a large acceptance area. The OT consists of gas-tight straw-tubes, which contain two layers of drift-tubes with inner diameter of 4.9 mm each. The counting gas inside the tubes is a mixture of Argon (70%) and CO₂ (30%): this choice allows a fast drift time, below 50 ns, and a drift spatial resolution of 200 μm .

The straw-tubes are arranged in three stations, as shown in Figure 2.7; each station is composed by four layers which are arranged with the same geometry of the IT: the modules in the first and fourth layers are parallel to the y axis, the ones in the second and third layers are tilted by $\pm 5^\circ$ with respect to the y axis. Each OT station has an active area of 5971x4850 mm².

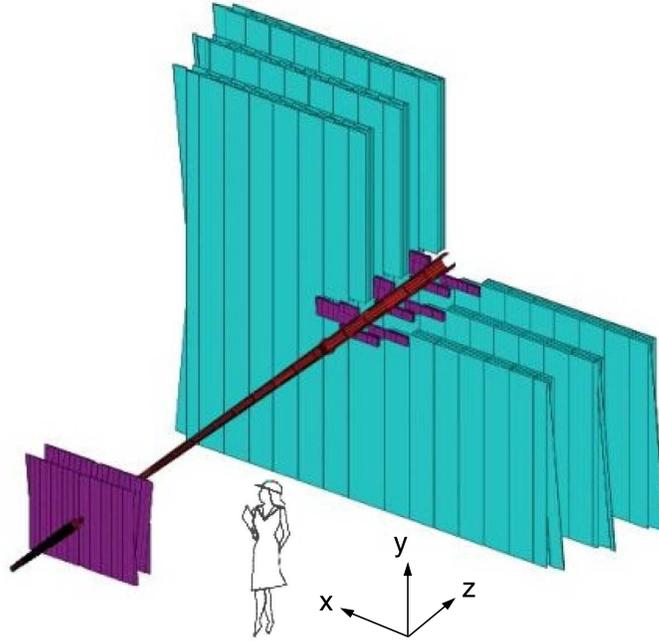


Figure 2.7. View of the three OT stations.

2.4 Particle Identification

2.4.1 RICH

The particle identification (PID) is of fundamental importance in the physics program of LHCb. In particular, the RICH detectors allow to separate pions from kaons in b hadron decays. Since at large polar angles the momentum is softer, while on the opposite, at small polar angles the momentum is typically hard, two RICH detectors are needed to cover the full momentum range.

The RICH1 detector [56] is placed upstream the dipole magnet, between the vertex locator and the Trigger Tracker, and covers the low momentum charged particle range, about 1-60 GeV/c. The radiators used in RICH1 are aerogel ($n = 1.03$) and fluorobutane (C_4F_{10}) gas ($n = 1.0014$); in Figure 2.8 the Cherenkov angle as a function of the particles momenta is plotted, showing that using these radiators it is possible to separate π , K and p in the soft momentum range. The RICH1 detector covers the full LHCb acceptance, from ± 25 mrad to ± 300 mrad in the horizontal direction and ± 250 mrad in the vertical direction.

The RICH2 detector [57] is placed downstream the dipole magnet, between the last tracking station and the first muon station, and covers the high momentum range, from about 15 GeV/c to and beyond 100 GeV/c. The radiator used in the

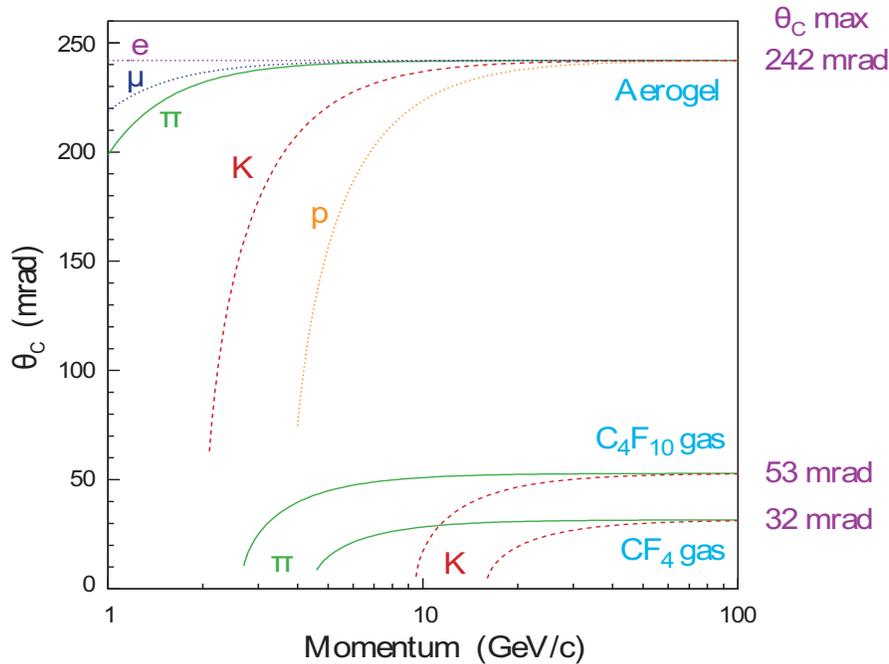


Figure 2.8. Cherenkov angle as a function of particle momentum for the RICH radiators.

RICH2 is the CF_4 gas ($n = 1.0005$). The RICH2 covers a limited angular acceptance, from ± 15 mrad to ± 120 mrad in the horizontal direction and ± 100 mrad in the vertical direction, which is the region where high-momentum particles are produced.

In both RICH detectors, spherical and flats mirrors are used in a particular configuration in order to focus the Cherenkov light. Cherenkov photons are detected in the wavelength range 200-600 nm by Hybrid Photon Detectors (HPD), which are shielded from external magnetic fields up to 50 mT by an iron cover.

2.4.2 Calorimeters

The calorimeter system [58] is used for several purposes:

- to select high transverse energy hadron, electron and photon candidates for the level 0 (L0) trigger;
- to identify electrons, hadrons and photons;
- to measure energies and positions of the identified particles.

The calorimeters are fundamental in flavour tagging and in the study of radiative B meson decays and spectroscopy studies. The calorimeter system consists of

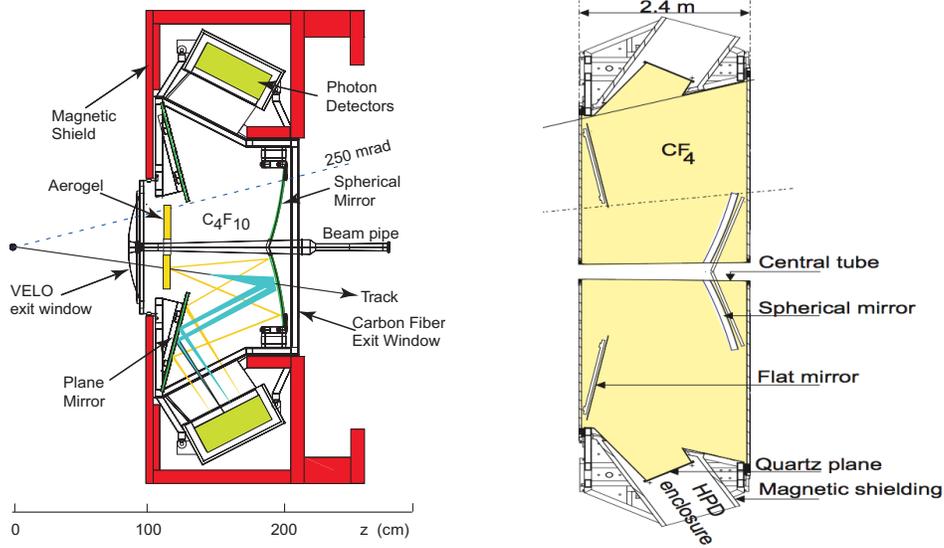


Figure 2.9. Schematic side view of RICH1 (left) and RICH2 (right) detectors.

a Scintillator Pad Detector (SPD), a preshower (PS) detector, an Electromagnetic Calorimeter (ECAL) and a Hadronic Calorimeter (HCAL). The scintillation light is collected by Wavelength Shifting Fibers (WLS) and then transmitted to Photo Multipliers (PMTs) in the ECAL and HCAL or to multianode photomultipliers (MAPMTs) in the SPD and PS.

The SPD is placed in front of the preshower calorimeter in order to select charged particles and separate electrons from photons. The purpose of the whole SPD/PS system is to improve the spatial and energy resolution of the electromagnetic showers. The system is composed by polystyrene scintillating tiles and a 12 mm thick lead layer placed in between them.

The ECAL has the aim of measuring electromagnetically interacting particles (e , γ , π^0). The structure of the ECAL is given by alternating layers of absorbers and active material. The absorbers are 2 mm thick lead layers, the active material is made of 4 mm thick scintillating tiles. The energy resolution of the ECAL is

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b, \quad (2.1)$$

with $8.5\% < a < 9.5\%$, $b \sim 0.8\%$ and E is given in GeV. In order to achieve optimal energy resolution, the showers from high energy photons have to be contained, requiring the ECAL to have a thickness of 25 radiation lengths.

The structure of the HCAL is analogous to the one of the ECAL: 16 mm thick

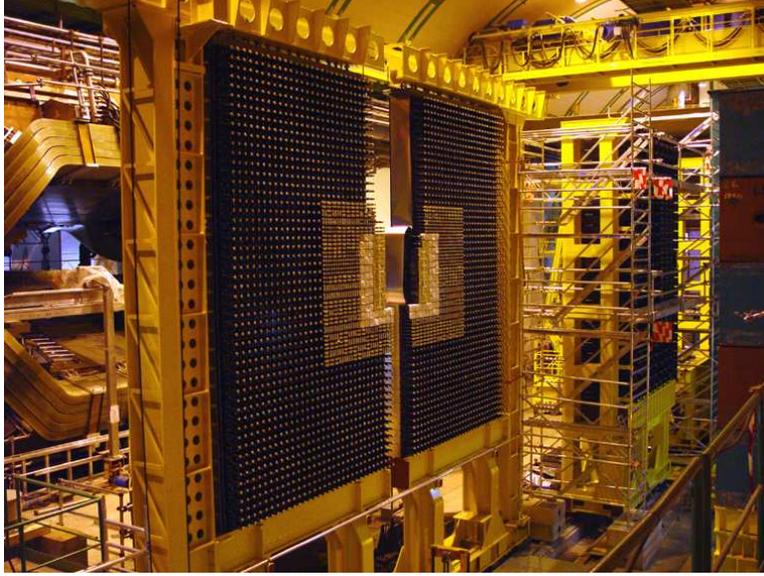


Figure 2.10. View of the ECAL.

iron tiles are alternated with 4 mm thick scintillating material. The HCAL energy resolution (extracted from data) is:

$$\frac{\sigma_E}{E} = \frac{69 \pm 5\%}{\sqrt{E}} \oplus (9 \pm 2)\%, \quad (2.2)$$

where E is in GeV. The trigger does not impose stringent requirements on the energy resolution of the HCAL, then on the hadronic shower containment, thus the thickness of the calorimeter is 5.6 radiation lengths (due to space limitations).

2.4.3 Muon system

The identification of muons is crucial for LHCb as they are present in many final states of CP sensitive B decays (above all in the golden modes $B_d^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ and $B_s^0 \rightarrow J/\psi(\mu^+\mu^-)\phi$). Muons are used in CP and oscillation measurements to tag the initial state flavour of the B meson and play a role in rare B decays like $B_s^0 \rightarrow \mu^+\mu^-$ which are sensitive to beyond Standard Model physics.

The muon system [59] consists of five stations (M1-M5) placed along the beam axis. The M2-M5 stations are located downstream the calorimeters and are separated by iron absorbers 80 cm thick in order to select highly penetrating muons; the minimum momentum to cross all five stations is 6 GeV/c. The M1 station instead

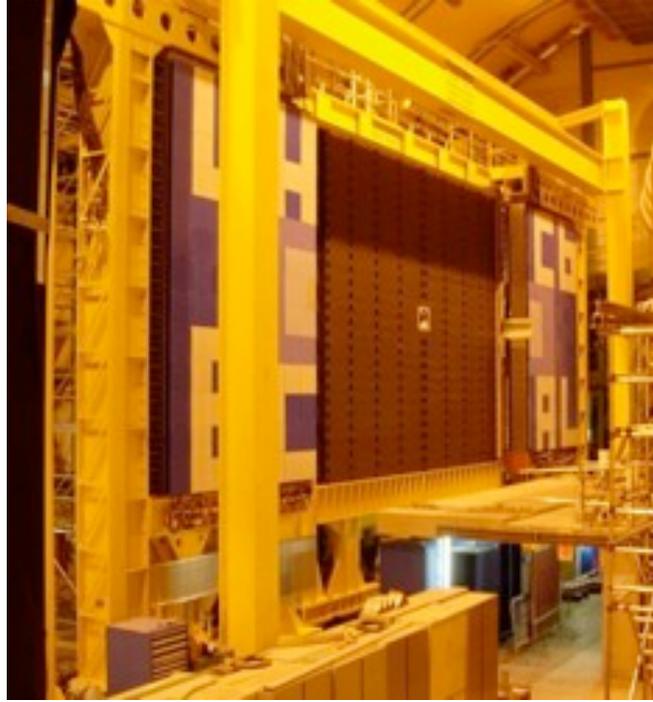


Figure 2.11. View of the HCAL.

is upstream the calorimeters to improve the transverse momentum measurement in the Level 0 trigger. The stations have a rectangular shape, a combined area of 435 m^2 and granularity scaling with the distance from the interaction point. The muon stations have an angular acceptance from 20 mrad to 306 mrad on the vertical plane and from 16 to 258 mrad in the horizontal plane, leading to a total acceptance of 20% for muons coming from b inclusive semileptonic decays. Each muon station is divided in four concentric regions (R1-R4), whose linear dimensions and segmentations scale as 1:2:4:8 so that the particle flux is about the same in every region. The whole system comprises 1380 chambers which are of two different types:

- Multi Wire Proportional Chambers (MWPC): 1368 chambers used for all regions, except for the inner region of M1. The MWPCs have a gap of 5 mm filled with a combination of three gases (carbon dioxide, argon and tetrafluoromethane) and a 2 mm thick wire plane placed in the middle of the gap. The time resolution is about 5 ns.
- triple GEM detectors: 12 chambers placed in the inner region of M1 with a rate capability up to 500 kHz/cm^2 of charged particles. To withstand such an

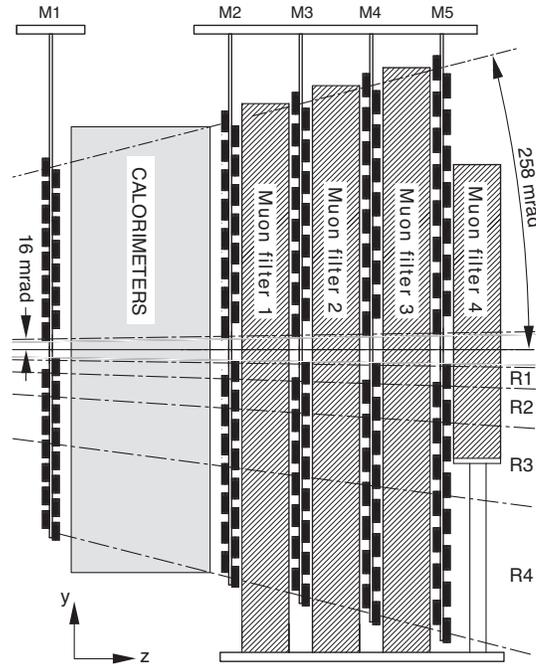


Figure 2.12. Schematic side view of the muon system.

high flux the detectors must be radiation hard so that no ageing effects occur for at least 10 years. This requirement can not be accomplished with MWPCs. Each chamber is composed by two triple GEM detectors superimposed, where a triple GEM is formed by three Gas Electron Multiplier foils enclosed between anode and cathode planes. With a gas mixture of $\text{Ar}/\text{CO}_2/\text{CF}_4$, the detectors achieve a time resolution better than 3 ns.

2.5 Trigger

The average luminosity of LHCb is $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$, much lower with respect to the maximum luminosity of the LHC. This reduces the radiation damage to detectors and electronics, and the number of interactions per bunch crossing is dominated by single interactions. The nominal LHC crossing frequency, 40 MHz, has to be reduced by the trigger to a few kHz in order to store the events for further offline analysis. The trigger [60] consists of two stages: a hardware stage, denoted as Level 0 (L0) trigger, followed by a software stage, the High Level Trigger (HLT) which applies a full event reconstruction. A scheme of the LHCb trigger is depicted in Figure 2.13.

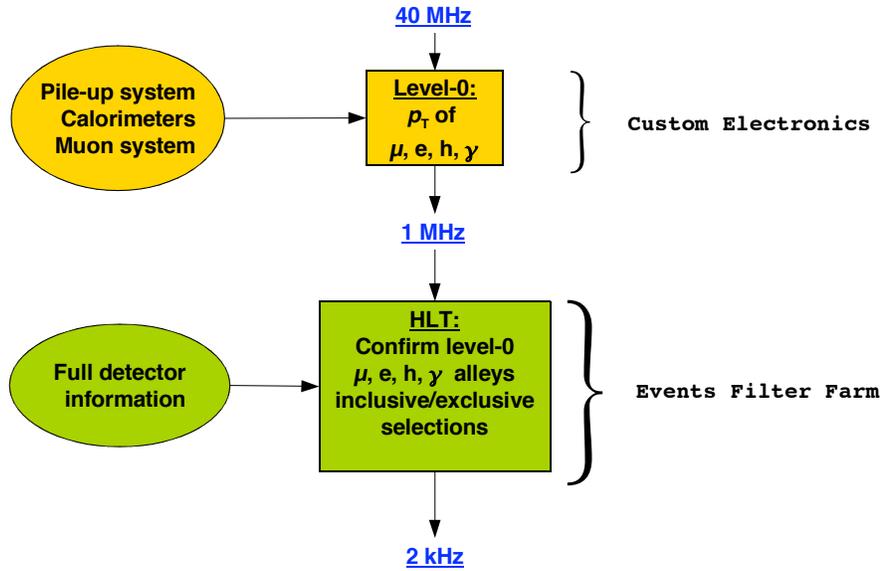


Figure 2.13. Scheme of the LHCb trigger.

2.5.1 Level 0

The aim of the L0 trigger is to reduce the rate from 40 MHz to 1 MHz. Due to the large mass of the B meson, the daughter particles have typically large transverse momentum p_T and large transverse energy E_T . In order to reconstruct and select these particles, the L0 trigger is composed by three elements:

- Pile-Up system in the VELO, which estimates the number of pp interactions per bunch crossing in order to distinguish between crossings with single and multiple visible interactions². The Pile-Up system consists of four silicon sensors, like the ones used in the VELO, to measure the radial coordinate of the tracks;
- Level 0 calorimeter trigger, which searches for the highest E_T hadron, electron and photon clusters in the calorimeters. Clusters are formed adding the E_T of 2x2 cells and are identified as originating from electrons, photons or hadrons according to the informations from SPD, PS, ECAL and HCAL. A measurement of the charged particle multiplicity is also provided by counting the total number of SPD cells with a hit;
- Level 0 Muon trigger, which searches for the highest p_T candidate in the

²An interaction is visible if it produces at least two charged particles with sufficient hits in the VELO and T1-T3 stations to be reconstructible.

muon chambers. The resolution of the muon transverse momentum is about 20%. The Level 0 muon trigger selects the two muons with the highest transverse momentum for each quadrant of the muon detector. An optimisation of the Level 0 Muon trigger will be described in Chapter 3.

Each component of the L0 trigger is connected to one detector and to the Level 0 Decision Unit (DU), which collects all the informations and takes the final decision per each bunch crossing.

A scheme of the L0 trigger is shown in Figure 2.14: the pile up system receives 2048 channels from the pile-up detector in the VELO, the Level 0 calorimeter 19420 from SPD, PS, ECAL and HCAL calorimeters, the Level 0 muon 25920 channels from the muon chambers. The L0 trigger is implemented using custom electronics

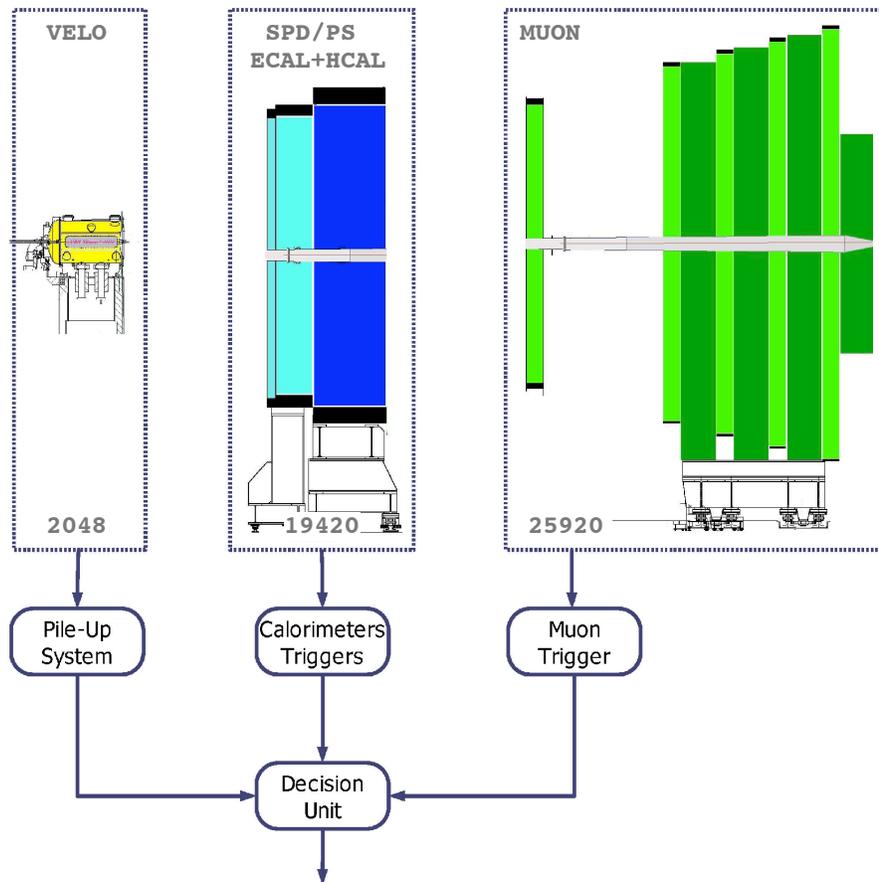


Figure 2.14. Scheme of the L0 trigger.

and operates synchronously with the 40 MHz bunch crossing signal from LHC. The time elapsed between a pp interaction and the arrival of the Level 0 trigger

decision (the latency) is $4 \mu\text{s}$. Taking into account the time of flight of the particles and the delays introduced by cables and electronics, the time left for the L0 trigger decision is $2 \mu\text{s}$.

2.5.2 High Level Trigger

The HLT is a C++ application running on every CPU of the Event Filter Farm (EFF), formed by up to 2000 computing nodes. The purpose of the HLT is to reduce the rate from 1 MHz to a few kHz, that is the rate at which events are written in the storage unit. The HLT has two stages, denoted as HLT1 and HLT2.

The HLT1 represents a L0 confirmation: it divides the particle candidates found in the L0 trigger in alleys, requiring them to be reconstructed in the VELO and/or the T stations and to have a high p_T and/or large impact parameter; in the case of γ and π^0 candidates, it confirms the absence of a charged particle which could be consistent with their tracks. This first selection reduces the rate to about 30 kHz. The rate is now sufficiently low to allow a full pattern recognition, using the full event data.

At this point, the HLT2 selects interesting final states according to inclusive and exclusive criteria. It applies cuts on the invariant mass or on pointing of the momentum of the B candidate toward the primary vertex, reducing the rate to about 2 kHz.

The HLT trigger is implemented using a processor farm and is executed asynchronously with the 40 MHz bunch crossing signal from LHC.

A scheme which summarises the trigger flow is shown in Figure 2.15.

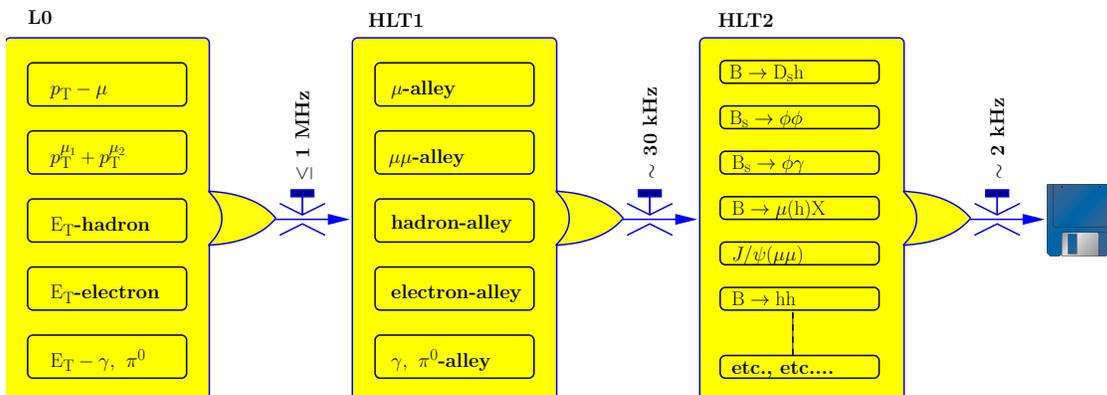


Figure 2.15. Scheme of the overall trigger flow.

2.6 Offline processing

After all of the online trigger stages, an offline analysis is performed. It proceeds through the following phases:

- the data selected triggering on the interesting events are collected in the so called raw data, suitable for a subsequent event reconstruction;
- physical quantities are reconstructed leading to the generation of a new data set, the Full Data Summary Tape (FullDST) which keeps the informations from the raw data;
- the Full DST is then reduced (DST format) via a stripping process, i.e. the production of streams of selected events for further individual analysis. In this phase, four-momentum vectors are determined, primary and secondary vertices located and composite particles, like B mesons, reconstructed. A preselection algorithm is provided for each channel of interest (stripping line). Parallel to this, a so-called microDST (μ DST) is generated, generally speaking this is equivalent to the DST except that it contains only signal events, according to a specific definition of signal. Both these procedures are needed to reduce the datasets dimension to a manageable level;
- after the stripping, the datasets are further reduced by the *DaVinci* software package, whose outputs are used in the offline physics analysis.

2.7 Performances

2.7.1 Track reconstruction

To reconstruct the particle tracks, the hits in the VELO, the TT, the IT and the OT detectors are combined to form particle trajectories from the vertex locator up to the calorimeters. The first stage is the search of the seeds, i.e. hits in the VELO and the T stations where the magnetic field is low, representing the initial track candidates. Then the seeds are combined with hits in other tracking detectors, leading to the tracks identification. To account for multiple scattering and correct for dE/dx energy loss, the tracks are fitted with a pattern recognition algorithm (Kalman filter).

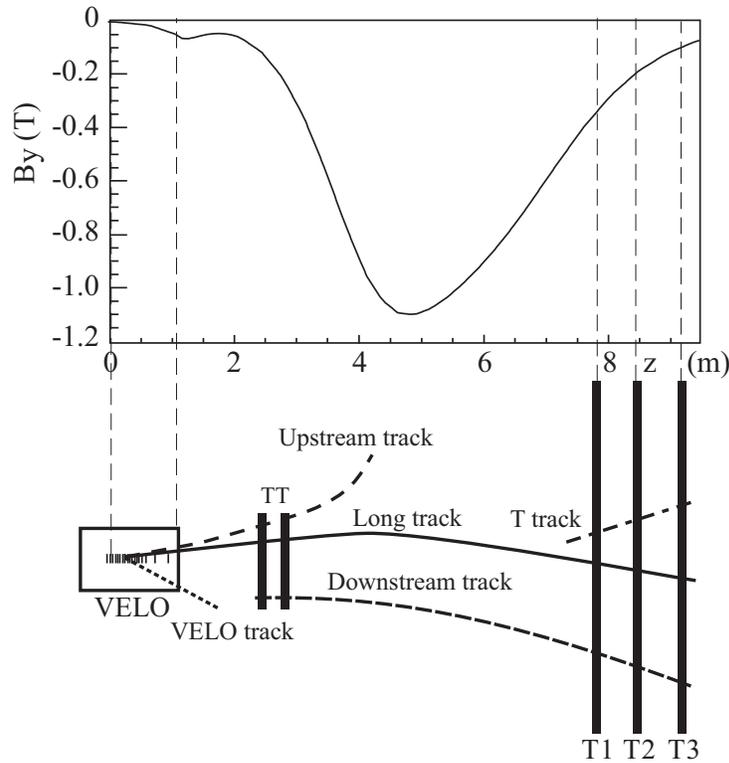


Figure 2.16. Scheme of the different kinds of track; on the top is plotted the y component of the B field as function of the z coordinate

According to their trajectory inside the detector, five different kinds of tracks are defined (Figure 2.16):

- Long Tracks, which go through the full tracking system, from the VELO to the T stations. Thus, they have the most precise momentum determination and are fundamental in the reconstruction of b -hadron decays;
- Upstream Tracks, which traverse only the VELO and TT stations. They have typically low momenta and are bent out the detector acceptance by the magnetic field. Even if the momentum resolution is low, they can be used to reconstruct b -hadron decays and/or for flavour tagging. Moreover, if the particle has a velocity above threshold, they are detected by the RICH1 and can be used to study the background in particle identification;
- Downstream Tracks, which traverse only TT and T stations;
- VELO tracks, which traverse only the VELO. They are typically large angle or backward tracks and are used for the primary vertex reconstruction;

- T Tracks, which traverse only the T stations. They are typically produced in secondary interactions.

The performance of the track reconstruction is evaluated in terms of the reconstruction efficiency, defined as the number of successfully reconstructed tracks over the total number of reconstructible tracks. A track is considered successfully reconstructed if at least 70% of the hits associated to the track are originated by a single Monte Carlo simulated particle. The efficiency to identify the trajectory of a particle with $p_T > 10$ GeV as a long track is about 96% on average.

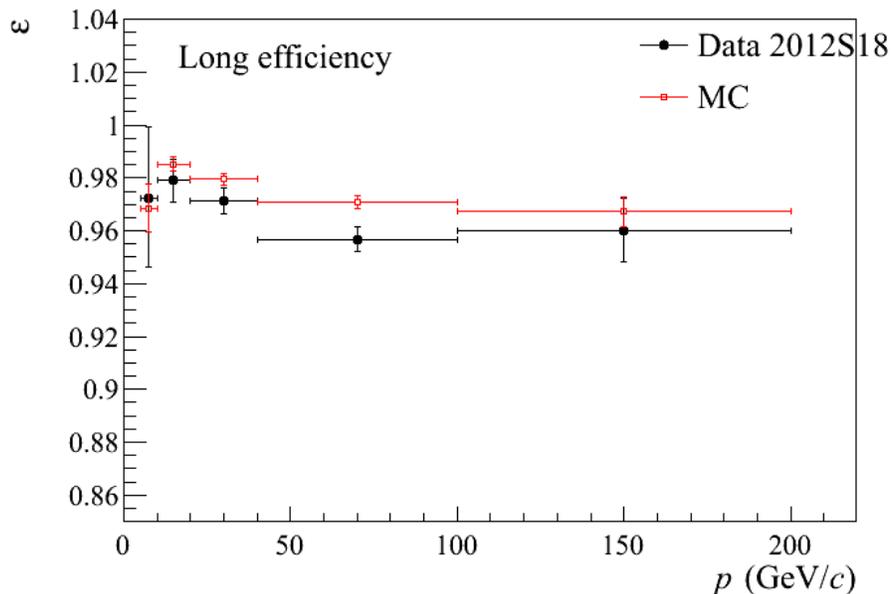


Figure 2.17. Efficiency for long tracks versus their momentum using 2012 data (black) and Monte Carlo simulated events (red).

2.7.2 Particle identification

The combined information from the two RICH detectors, the calorimeters and the muon system allows to identify charged particles (e, μ, π, K, p).

Muons are identified by extrapolating their tracks from the hits in the muon stations. Electrons are identified by matching the track momentum and the energy of the clusters in the ECAL and detecting separate clusters due to bremsstrahlung photons emitted by the electrons before the magnet. The electron identification is less efficient than the muon identification (shown in Figure 2.18), for this reason decays containing muons are studied rather than decays with electrons.

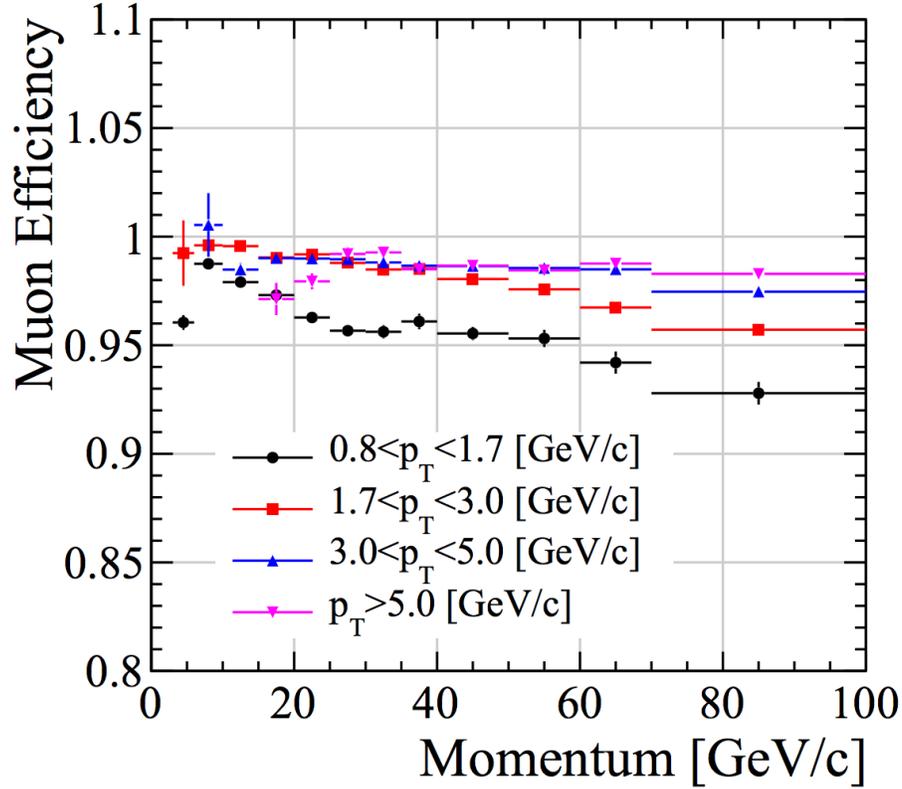


Figure 2.18. Muon identification efficiency, as a function of momentum, in ranges of transverse momentum using 2011 data.

Neutral particles (γ , π^0) are identified by using the electromagnetic calorimeter. Photons are reconstructed by the ECAL as clusters without an associated track. Neutral pions with $p_T < 3$ GeV are identified by resolving two photon clusters in the ECAL, while for higher transverse momentum a large fraction of photon pairs can not be resolved as a pair of clusters, leading to a single cluster, called merged π^0 .

The hadron discrimination is performed by the RICH system. The method used to match the observed pattern in the RICH photodetectors with the particle is called global pattern recognition. This is based on a likelihood approach, which considers all the tracks in the event and in both RICH detectors simultaneously, such that also cases of overlapping Cherenkov cones can be treated.

For each particle type, the likelihood is computed combining informations from the RICH system, the calorimeters and the muon detector. Particles are selected by cutting on the likelihood ratio between hypotheses, or equivalently on the difference of the logarithm of the likelihood. Since the pions are the most abundant

particles, the first assumption is that all particles are pions. Then for each track in turn, the mass hypothesis is varied according to e , μ , K , π and p mass hypothesis (the other hypotheses are left unchanged) and the $\text{Log}\mathcal{L}$ is computed again. Among all tracks, the change in mass hypothesis which maximises the overall likelihood is found and the PID for that track is set to the one corresponding to that hypothesis. This procedure is repeated until the optimal value is determined for each track.

Control channels, whose identifications proceed through purely kinematical

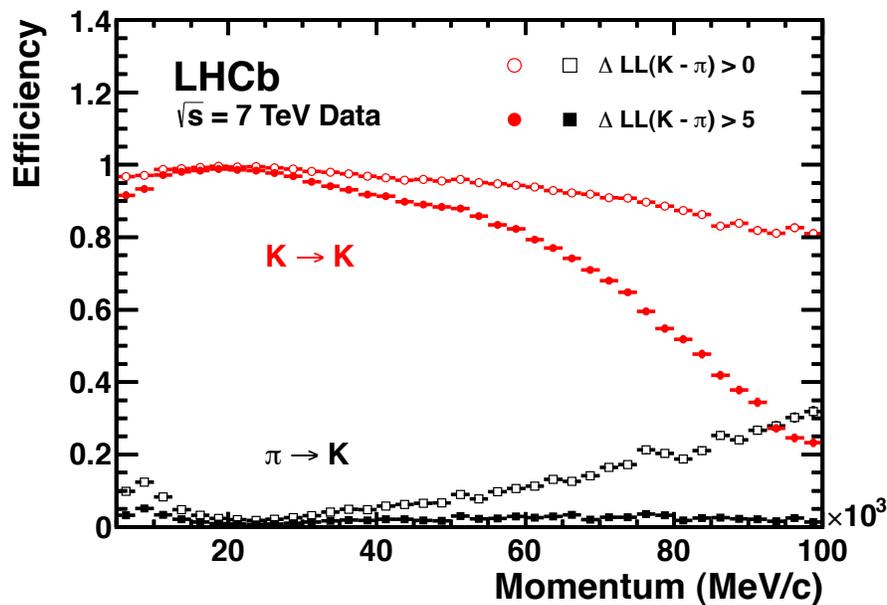


Figure 2.19. Kaon identification efficiency and pion misidentification rate as a function of particle momenta computed on data. Two different requirements of $\Delta \log \mathcal{L}$ have been applied.

selection (without involving any information from the RICH detectors), are used to determine the PID performance on data. The achievable discrimination between each pair of tracks is evaluated imposing requirements on the difference in $\log \mathcal{L}$ between the two hypotheses. Figure 2.19 shows the K efficiency (K identified as a K) and the π misidentification rate (π misidentified as K) as a function of the particle momentum. Two requirements have been applied: $\Delta \log \mathcal{L}(K - \pi) > 0$ leads to a kaon efficiency and pion misidentification fraction of about 95% and 10%, respectively; $\Delta \log \mathcal{L}(K - \pi) > 5$ reduces significantly the misidentification rate to about 3% for a kaon efficiency of about 85% [61].

The corresponding efficiency and misidentification rates are also computed for proton-pion and proton-kaon discrimination.

Chapter 3

Level 0 Muon Trigger Optimisation

3.1 Level 0 Muon Track Finding Algorithm

In this chapter, an improvement of the performance of the LHCb Level 0 Muon trigger is shown.

The study involved an optimisation of the configuration of the Fields of Interest (FOIs), which are key parameters for the muon track finding algorithm.

The Level 0 Muon Trigger (also referred to as *hardware trigger*) looks for muon tracks with a large transverse momentum p_T , which is a typical signature of a muon produced in a b-hadron decay. The track finding is performed on the logical pad layout of the muon detector. It searches for hits defining a straight trajectory through the five muon stations towards the interaction point.

The trigger algorithm uses a five-fold coincidence, the efficiency for each station must be at least 99%, with a time resolution better than 25 ns in order to unambiguously identify the bunch crossing. Each station has two independent detector layers, logically OR-ed on the chamber, to form *logical channels*. The total number of logical channels is 25920.

Each station is subdivided into four regions with different *logical pads dimensions*, as shown in Figure 3.1. Pads are obtained by the crossing of horizontal and vertical strips when applicable (strips are employed in stations M2-M5 while station M1 and innermost region R1 of stations M4-M5 are equipped with pads). The trigger processor receives 25920 logical channels every 25 ns corresponding to 55296 logical pads obtained by crossing strips.

The track finding algorithm, shown in Figure 3.2, starts from a hit in M3 (the so-called *track seed*), then the extrapolated positions in M2, M4 and M5 stations are set along a straight line passing through the hit and the interaction point. Hits

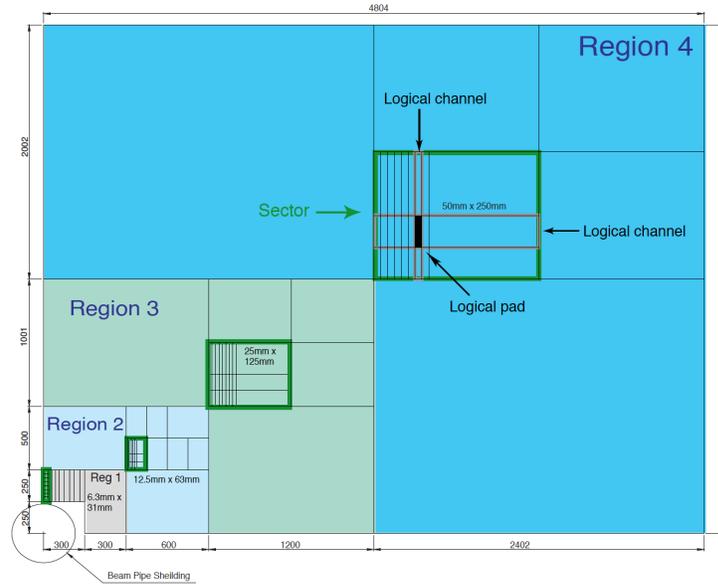


Figure 3.1. Front view of one quadrant of muon station M2, showing the dimensions of the regions. Inside each region is shown a sector, defined by the size of the horizontal and vertical strips.

are looked for in search windows named Fields of Interest (FOIs), approximately centred on the extrapolated positions. FOIs are opened along the x -axis for all stations and along the y -axis only for stations M4 and M5. The size of the FOI depends on the stations considered, in order to account for the differences in granularities between stations, the level of background and the retention rate required for Minimum Bias events.

A muon track is flagged when at least one hit is found inside the FOI for each station M2, M4 and M5. The track position in station M1 is determined by making a straight line extrapolation from M3 and M2 and identifying in the M1 FOI the pad hit closest to the extrapolation point, as shown in Figure 3.3. The whole logical layout is projective, leading to a one-to-one mapping from pads in M3 to pads in M2, M4 and M5, and from pairs of pads in M2 and M3 to pads in M1. The hits associated to a track in the first two muon stations allow the determination of the p_T of the track itself. Table 3.1 shows the maximum¹ FOI sizes.

During the Run I of the LHC a fixed FOI configuration was used, determined from previous studies on simulated data. In terms of expressed in M3 pads units it is $\{6, 5, 0, 4, 8\}$. However it was noted that the configuration gave a non-vanishing

¹The size of the FOI can not exceed a certain number of pads, imposed by the limited information that can be exchanged with the Processing Unit.

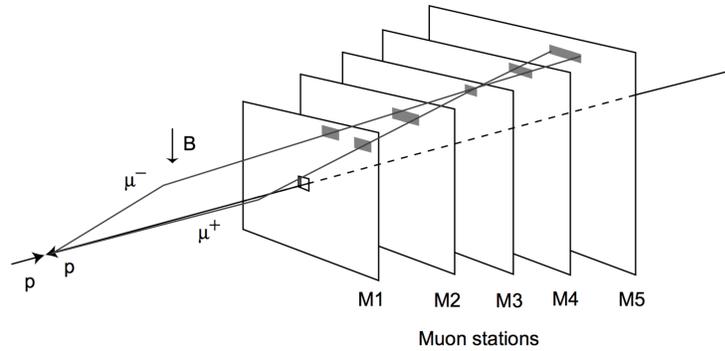


Figure 3.2. Level 0 Muon Trigger track finding. Grey areas represent the Fields Of Interest used by the algorithm.

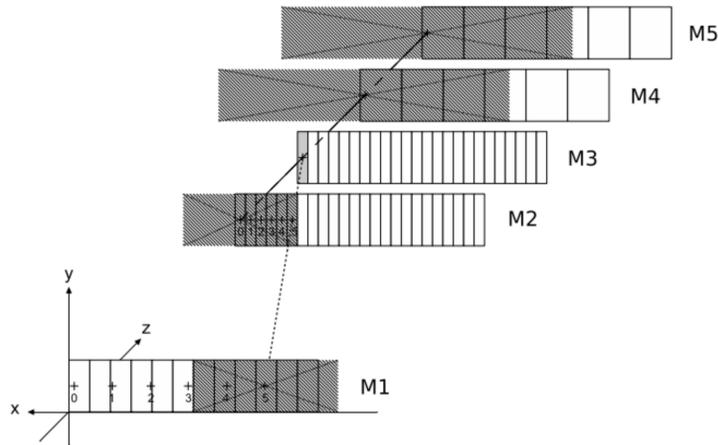


Figure 3.3. Example of Fields of Interest associated with a M3 pad. The straight line shows the extrapolated position in stations M2, M4 and M5. The hashed arrays show the maximum size of the Field Of Interest centered on the extrapolation position. The dashed line shows the straight line extrapolation from M3 and M2 to M1 when the pad labeled 5 is hit in M2.

muon rate at high p_T on Minimum Bias events, contrary to expectations for such events which are mainly due to the soft fragments of the remaining quarks and gluons that hadronize. For this reason, setting a relatively high p_T threshold of 1.4 GeV/c on the L0 single muon trigger was mandatory in order to keep the corresponding bandwidth to a manageable level. This behaviour was attributed to *ghosts* in the first muon station, i.e. hits that are not due to a muon originated in a b-hadron decay from the interaction point, or muons from decays in flight or interactions of kaons or pions. The origin of the excessive rate at high p_T is still

	M1	M2	M4	M5
x	± 3	± 5	± 3	± 3
y	-	-	± 1	± 1

Table 3.1. Maximum size of the FOI in the x and y coordinates for the M1-M5 stations, in units of pads with respect to the one lying on the straight line passing through the hit in M3 and the interaction point, considering the different stations' granularities. A FOI of ± 1 corresponds to a full width of 3 pads.

not established, therefore it is needed to study it in more detail.

The performance of the Level 0 Muon Trigger can be evaluated with a specific emulator, able to reproduce both the hardware and the software steps as in the real Level 0 Muon Trigger.

Studies on both Monte Carlo simulated data and real data collected in 2012 and 2015 showed that L0-muon can be optimised by changing the default FOI configuration. The optimisation can be done by increasing the Level 0 Muon efficiency of a given signal decay, and by reducing the Minimum Bias retention rate, or at least keeping the former constant while decreasing the Minimum Bias rate.

The efficiency of the Level 0 Muon Trigger generally depends on the muon p_T spectra of the process under consideration. For this reason, a number of signal channels were studied, along with Minimum Bias events. For instance, the channels $B_s^0 \rightarrow \mu^+ \mu^-$, $B^0 \rightarrow D^* \mu^+ \nu$, $B^+ \rightarrow J/\psi K^+$, $\tau \rightarrow \mu \mu \mu$ give muons of increasingly softer p_T spectra.

Given that the emulator allows to change the sizes of the FOI, the studies were performed by tightening them with respect to the configuration used in Run I. Starting with M1 and M2, then M4 and M5 as well, it was possible to study several FOI configurations. Care had to be taken in order to avoid configurations impossible to implement, due to hardware limitations, e.g. the FOI in station M1 must be an even number, as the granularity of station M1 is twice the granularity of station M3

3.2 Monte Carlo Studies

In this section the studies performed on Monte Carlo simulated data are presented. The 2012 data taking conditions were used to simulate all $B_s^0 \rightarrow \mu^+ \mu^-$, $B^0 \rightarrow D^* \mu^+ \nu$, $B^+ \rightarrow J/\psi K^+$ and $\tau \rightarrow \mu \mu \mu$ samples.

Signal events are filtered by requiring them to pass a stripping line (see Section 2.6),

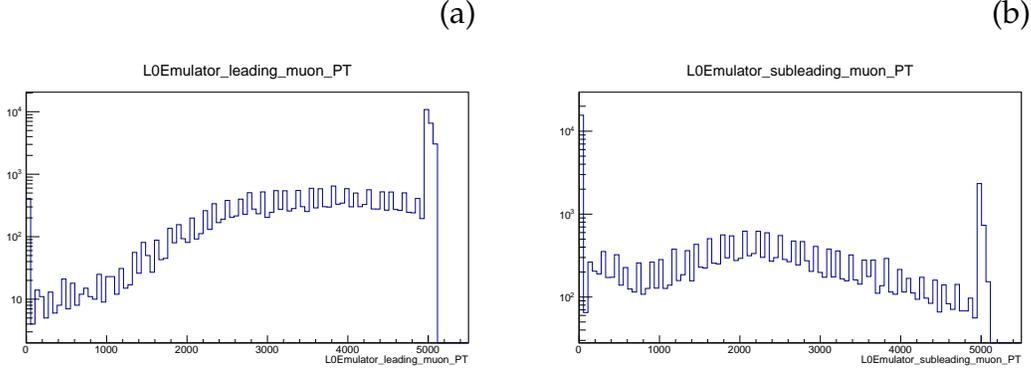


Figure 3.4. Example plots for (a) the highest p_T muon and (b) the subleading muon distributions from the Level 0 Muon emulator. The FOI is the one used as default in Run I, the channel is $B_s^0 \rightarrow \mu^+ \mu^-$.

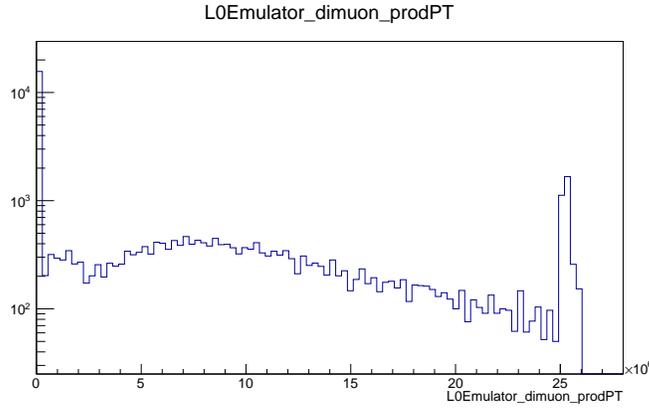


Figure 3.5. Example plot for muon pairs with a certain p_T product from the Level 0 Muon emulator. The FOI is the one used as default in Run I, the channel is $B_s^0 \rightarrow \mu^+ \mu^-$.

specific for each channel studied. The samples and the stripping lines used are summarised in Table 3.2.

The p_T distributions for the leading and the subleading muon candidates, shown in Figure 3.4 for the $B_s^0 \rightarrow \mu^+ \mu^-$ channel, are used to compute efficiencies for the single muon trigger. Figure 3.4 (a) shows the p_T distribution for the muon candidates, selecting, in case of multiple candidates per event, the one with the highest p_T . Figure 3.4 (b) shows the corresponding distribution when selecting the second highest p_T muon candidate. For the dimuon trigger, the efficiencies are evaluated using the distribution of the product of the p_T of the leading and subleading muons, shown in Figure 3.5.

From each p_T histogram a cumulative distribution is computed, then the efficiency

Decay Mode	Number of events	Stripping Selection
$B_s^0 \rightarrow \mu^+ \mu^-$	120000	StrippingBs2MuMuLinesWideMassLine
$B^0 \rightarrow D^* \mu^+ \nu$	120000	StrippingBd2DstarMuNuTight
$B^+ \rightarrow J/\psi K^+$	22622	StrippingBetaSBu2JpsiKDetachedLine
$\tau \rightarrow \mu \mu \mu$	7290	StrippingTau23MuTau23MuLine

Table 3.2. Monte Carlo samples used for the studies along with the Stripping lines used to filter signal events.

ϵ is defined as follows:

$$\epsilon_i = \frac{N_{TOT} - n_i}{N_{TOT}} = 1 - \frac{n_i}{N_{TOT}} \quad (3.1)$$

where N_{TOT} is the total number of processed events and n_i is the number of muons in each bin of the cumulative distribution. Using this definition and expressing the efficiency as a function of the transverse momentum, it is possible to obtain the fraction of events passing a certain p_T cut at the Level 0 Muon level. Along with the efficiency, its statistical uncertainty σ_ϵ is assessed, using binomial statistics:

$$\sigma_{\epsilon_i} = \sqrt{\frac{\epsilon_i(1 - \epsilon_i)}{N_{TOT}}} \quad (3.2)$$

An example of the efficiency as a function of p_T is in Figure 3.6, where the efficiencies for the Monte Carlo samples used are computed for the default FOI configuration. It is evident how the efficiencies are different for muons with different p_T spectra, with higher efficiencies for channels where muons typically have a higher transverse momentum.

3.3 Minimum Bias Studies

In this section the studies performed on Minimum Bias data, collected in 2012 and 2015, are shown. With respect to the previous section, what has been referred to as “efficiency”, defined in Eq. 3.1, will be called “retention fraction”, which is more appropriate for Minimum Bias events. Run I data were used to compare the retention fraction evaluated by modifying the FOI to the one obtained with the default FOI configuration, the 2015 data sample was used to check if the improvement was present also on Run II data.

Along with the retention fraction it is also worth looking at the rate, defined as

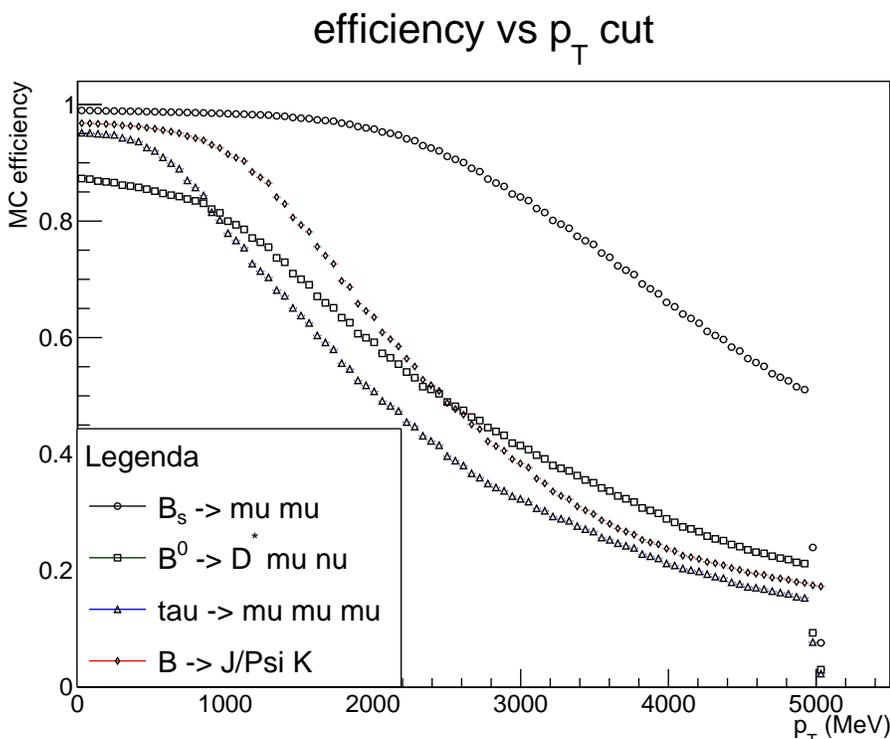


Figure 3.6. Plot of the Level 0 single muon trigger efficiency for the different Monte Carlo signal events. The FOI is the one used as default in Run I.

Sample	Conditions	Number of events
2012 Data	Magnet Down, Reco 14	180000
2015 Data	Magnet Down, Reco 15b	180000
2012 Monte Carlo	Magnet Down, Pythia 8	180000
2015 Monte Carlo	Magnet Down, Pythia 8	180000

Table 3.3. Minimum Bias samples used for the studies, along with either the data taking or the simulation conditions and the number of events used.

follows:

$$r_i = \epsilon_i \times (1 - P(0)) \times \frac{n_B}{n_{MAX}} \times 40 \text{ MHz} \quad (3.3)$$

where the factor $(1-P(0))$ is the Poissonian probability to have at least one visible interaction per bunch crossing², n_B/n_{MAX} is the ratio between the number of colliding bunches and the maximum number of bunches circulating in the LHC, and 40 MHz is the nominal bunch crossing rate. Putting the correct numbers³ in

² $P(n) = e^{-\mu} \frac{\mu^n}{n!} \rightarrow P(0) = e^{-\mu}$ where μ is the average number of visible pp interactions per bunch crossing.

³ $n_B = 1262; n_{MAX} = 3200$

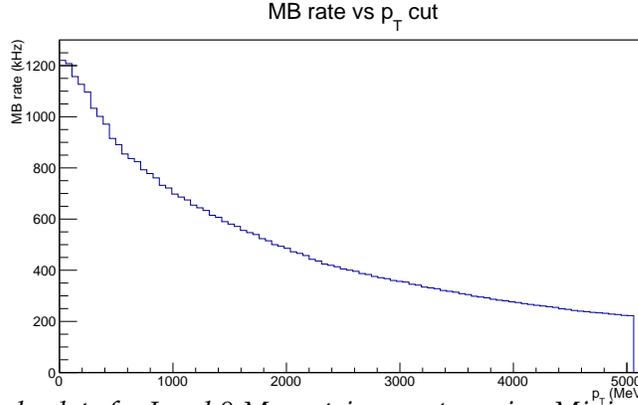


Figure 3.7. Example plots for Level 0 Muon trigger rates using Minimum Bias 2012 data. The FOI is the one used as default in Run I.

the formula, the relationship between the rate and the retention fraction becomes

$$r = \epsilon \times 11.6 \text{ MHz} \quad (3.4)$$

for 2012 data and

$$r = \epsilon \times 30 \text{ MHz} \quad (3.5)$$

for 2015 data.

As for simulated data, in case of multiple candidates per event, unlikely for Minimum Bias events, the muon with the highest p_T is selected; in the same way, the dimuon pairs are the leading and the subleading muons in the event.

3.4 FOI Optimisation

In this section the results obtained after modifying the FOI are shown. Several FOI configurations were tested, some of them have proven to be equivalent so only the relevant ones will be shown. For example it was verified that changing the FOI width in M4 and/or M5 does not make any difference in the Monte Carlo efficiencies nor in the Minimum Bias retention rates.

As mentioned in the previous sections, the relevant quantity under study is the Monte Carlo signal efficiency for the decay channels used, as a function of either the muons' p_T or the Minimum Bias retention fraction, changing the FOI configuration.

The plots in Figure 3.8 show how the signal efficiency drops with respect to the Level 0 p_T cut and how using a different FOI affects this quantity. The difference on signal efficiency is marginal and more visible on channels with softer p_T spectra.

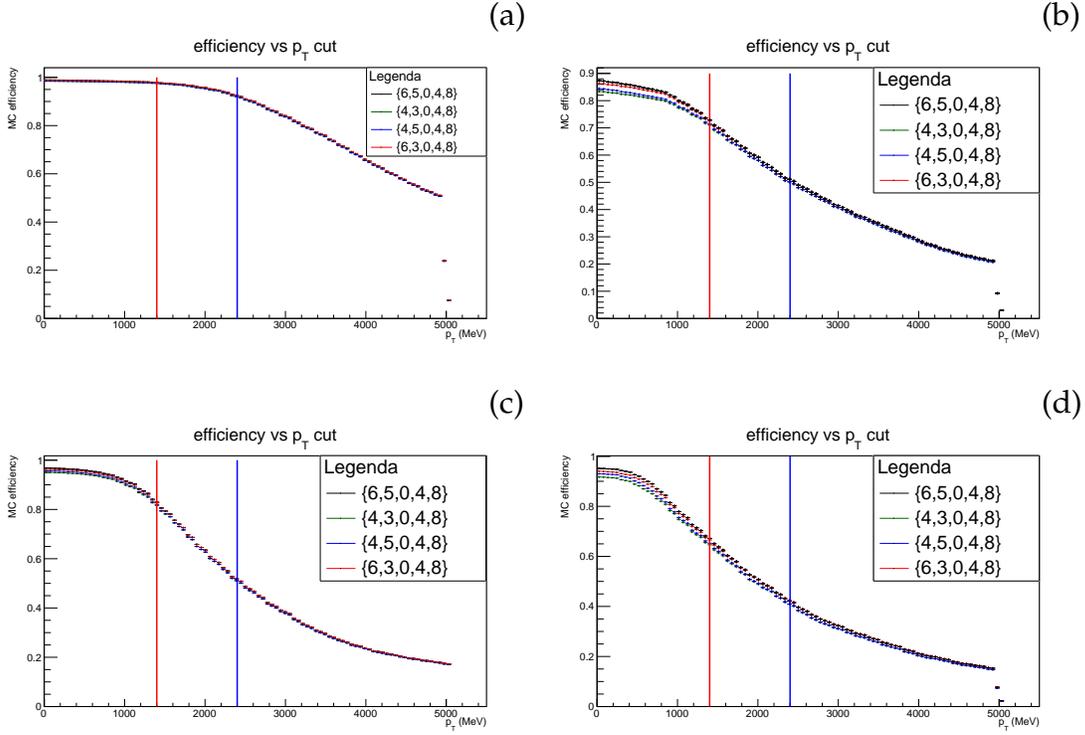


Figure 3.8. Single muon trigger efficiency as a function of p_T using (a) $B_s^0 \rightarrow \mu^+ \mu^-$, (b) $B^0 \rightarrow D^* \mu^+ \nu$, (c) $B^+ \rightarrow J/\psi K^+$ and (d) $\tau \rightarrow \mu \mu \mu$ Monte Carlo events with different Fields Of Interest. The red and blue vertical lines represent the p_T cut applied by the Level 0 Muon trigger in Run I and Run II respectively.

Looking at Figure 3.9 it can be seen that changing the FOI does not reduce only the signal efficiency but also, and more significantly, the Minimum Bias retention fraction. This allows to increase the p_T cut losing only a few percent signal efficiency but reducing the retention fraction of $\sim 20\%$.

The Minimum Bias rate reduction can be seen clearly in Figure 3.10 where the rate is shown for different FOI configurations. It is clear how reducing the FOI width in both M1 and M2 (from {6, 5, 0, 4, 8} to {4, 3, 0, 4, 8}) gives a significant reduction of the Minimum Bias rate, thus allowing for the Level 0 output bandwidth to be freed for other processes, without affecting significantly the efficiency on signal modes. The Monte Carlo efficiencies and the Minimum Bias retention rates are reported in Tables 3.4, 3.5, 3.6 and 3.7.

It has been shown that tightening the FOI in both the first and the second muon station M1 and M2 gives a small reduction of the signal efficiency and a significant reduction of the Minimum Bias retention fraction, if the p_T cut is fixed at the same

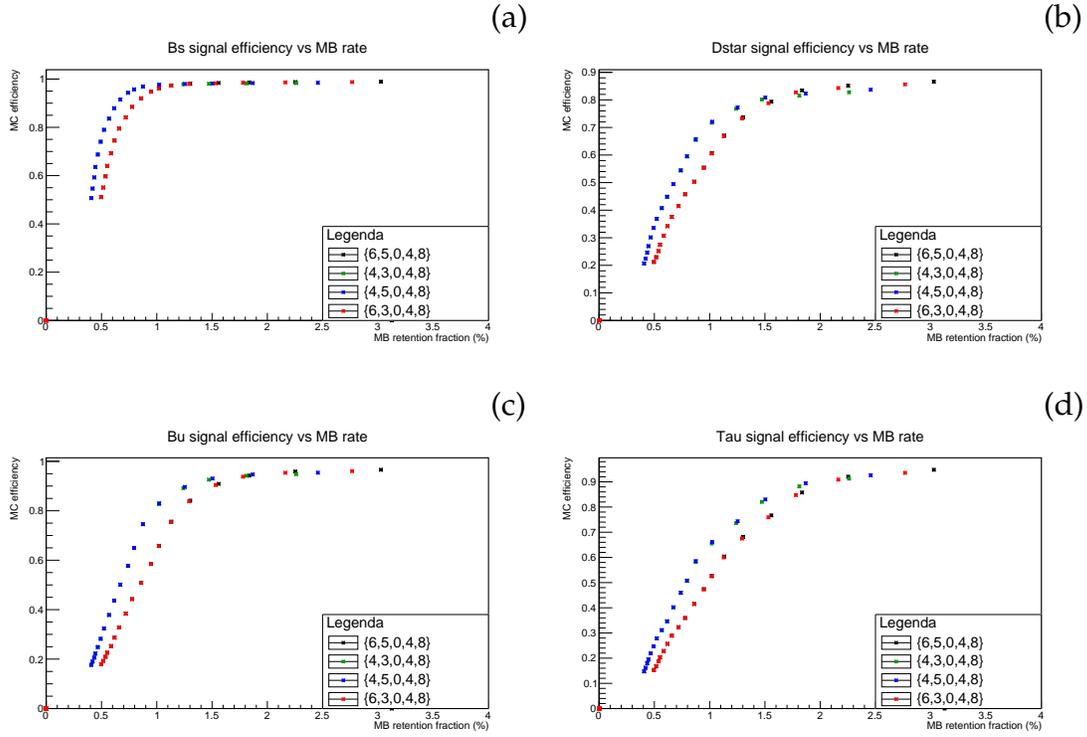


Figure 3.9. Single muon trigger efficiency as a function of the Minimum Bias retention fraction using (a) $B_s^0 \rightarrow \mu^+ \mu^-$, (b) $B^0 \rightarrow D^* \mu^+ \nu$, (c) $B^+ \rightarrow J/\psi K^+$ and (d) $\tau \rightarrow \mu \mu \mu$ Monte Carlo events with different Fields Of Interest.

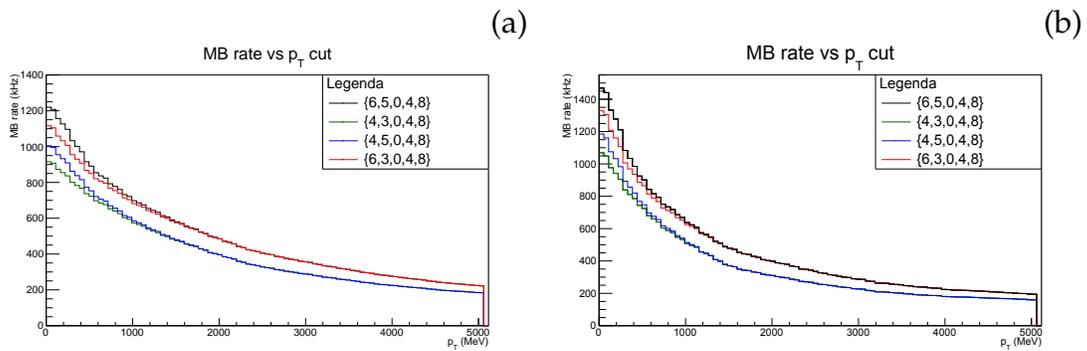


Figure 3.10. Minimum Bias rate as a function of p_T with different Fields Of Interest using (a) 2012 and (b) 2015 data.

value as in Run I⁴. This was tested on different Monte Carlo simulated signal samples and Minimum Bias data, collected in 2012 and 2015 by the LHCb detector.

⁴The Level 0 Muon p_T cut will actually be higher in Run II with respect to Run I due to other reasons.

p_T cut (GeV)	{6, 5, 0, 4, 8}		{4, 3, 0, 4, 8}	
	ϵ_{MC}	MB rate (kHz)	ϵ_{MC}	MB rate (kHz)
0.275	0.866	1096.5	0.828	836.5
0.550	0.851	890.8	0.815	722.6
0.825	0.835	777.8	0.802	640.2
1.100	0.794	686.1	0.767	564.5
1.375	0.737	614.9	0.717	503.5
1.650	0.671	555.5	0.655	454.9
1.925	0.607	500.0	0.594	407.4
2.200	0.554	457.5	0.544	371.9
2.475	0.504	413.7	0.494	335.3
2.750	0.457	383.3	0.449	309.9
3.025	0.415	356.4	0.407	288.9
3.300	0.376	331.0	0.369	268.9
3.575	0.342	308.5	0.335	250.8
3.850	0.308	286.6	0.301	232.9
4.125	0.275	269.6	0.269	220.2
4.400	0.252	254.8	0.246	208.9
4.675	0.229	238.1	0.224	195.9
4.950	0.212	226.8	0.207	187.5

Table 3.4. Monte Carlo signal efficiencies and Minimum Bias retention rates with different FOI for $B^0 \rightarrow D^* \mu^+ \nu$ sample.

The FOI configuration {4, 3, 0, 4, 8} gives better performances with respect to the one used in Run I {6, 5, 0, 4, 8}.

p_T cut (GeV)	{6, 5, 0, 4, 8}		{4, 3, 0, 4, 8}	
	ϵ_{MC}	MB rate (kHz)	ϵ_{MC}	MB rate (kHz)
0.275	0.989	1096.5	0.983	836.5
0.550	0.987	890.8	0.982	722.6
0.825	0.986	777.8	0.981	640.2
1.100	0.984	686.1	0.978	564.5
1.375	0.980	614.9	0.975	503.5
1.650	0.974	555.5	0.968	454.9
1.925	0.962	500.0	0.957	407.4
2.200	0.948	457.5	0.943	371.9
2.475	0.921	413.7	0.915	335.3
2.750	0.885	383.3	0.879	309.9
3.025	0.841	356.4	0.836	288.9
3.300	0.795	331.0	0.790	268.9
3.575	0.745	308.5	0.740	250.8
3.850	0.692	286.6	0.687	232.9
4.125	0.640	269.6	0.636	220.2
4.400	0.597	254.8	0.592	208.9
4.675	0.551	238.1	0.546	195.9
4.950	0.511	226.8	0.507	187.5

Table 3.5. Monte Carlo signal efficiencies and Minimum Bias retention rates with different FOI for $B_s^0 \rightarrow \mu^+ \mu^-$ sample.

p_T cut (GeV)	{6, 5, 0, 4, 8}		{4, 3, 0, 4, 8}	
	ϵ_{MC}	MB rate (kHz)	ϵ_{MC}	MB rate (kHz)
0.275	0.948	1096.5	0.913	836.5
0.550	0.920	890.8	0.882	722.6
0.825	0.857	777.8	0.820	640.2
1.100	0.767	686.1	0.736	564.5
1.375	0.681	614.9	0.655	503.5
1.650	0.604	555.5	0.582	454.9
1.925	0.526	500.0	0.507	407.4
2.200	0.474	457.5	0.459	371.9
2.475	0.415	413.7	0.402	335.3
2.750	0.359	383.3	0.346	309.9
3.025	0.323	356.4	0.311	288.9
3.300	0.289	331.0	0.279	268.9
3.575	0.257	308.5	0.247	250.8
3.850	0.228	286.6	0.219	232.9
4.125	0.204	269.6	0.195	220.2
4.400	0.187	254.8	0.180	208.9
4.675	0.168	238.1	0.162	195.9
4.950	0.153	226.8	0.147	187.5

Table 3.6. Monte Carlo signal efficiencies and Minimum Bias retention rates with different FOI for $\tau \rightarrow \mu\mu$ sample.

p_T cut (GeV)	{6, 5, 0, 4, 8}		{4, 3, 0, 4, 8}	
	ϵ_{MC}	MB rate (kHz)	ϵ_{MC}	MB rate (kHz)
0.275	0.966	1096.5	0.948	836.5
0.550	0.958	890.8	0.941	722.6
0.825	0.942	777.8	0.926	640.2
1.100	0.909	686.1	0.892	564.5
1.375	0.841	614.9	0.827	503.5
1.650	0.756	555.5	0.745	454.9
1.925	0.658	500.0	0.649	407.4
2.200	0.585	457.5	0.577	371.9
2.475	0.508	413.7	0.501	335.3
2.750	0.443	383.3	0.437	309.9
3.025	0.384	356.4	0.378	288.9
3.300	0.328	331.0	0.324	268.9
3.575	0.286	308.5	0.282	250.8
3.850	0.253	286.6	0.249	232.9
4.125	0.226	269.6	0.222	220.2
4.400	0.209	254.8	0.206	208.9
4.675	0.193	238.1	0.190	195.9
4.950	0.179	226.8	0.176	187.6

Table 3.7. Monte Carlo signal efficiencies and Minimum Bias retention rates with different FOI for $B^+ \rightarrow J/\psi K^+$ sample.

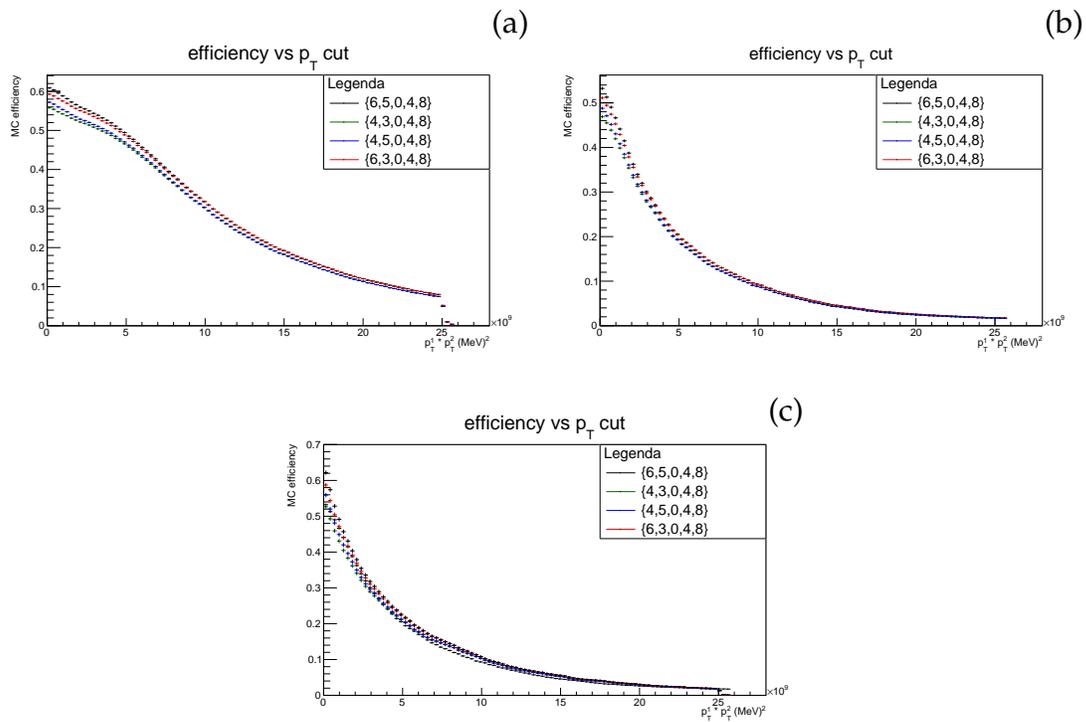


Figure 3.11. Dimuon trigger efficiency as a function of p_T (a) $B_s^0 \rightarrow \mu^+ \mu^-$, (b) $B^+ \rightarrow J/\psi K^+$ and (c) $\tau \rightarrow \mu \mu \mu$ Monte Carlo events with different Fields Of Interest.

Chapter 4

Analysis Strategy

This chapter illustrates the analysis procedure performed. The aim of the analysis is to identify semileptonic B decays to $D^*\pi\mu\nu X$ and $D^*\pi\pi\mu\nu X$ final states and to measure their inclusive branching fractions, considering both resonant and non-resonant contributions. There are two selection stages: a first selection has the purpose of identifying a muon and a D^* meson from a semileptonic B decay, the second one identifies one or two additional pions to couple to the D^* . A multivariate analysis has also been performed on the Monte Carlo sample in order to find the best selection criteria to reduce backgrounds, without affecting too much the signal. The multivariate analysis allows also to find criteria in order to choose the best pion candidates. This selection is first applied to the Monte Carlo sample, in order to evaluate the selection efficiencies, and then to the data sample.

4.1 Data sample

The data used for this analysis have been recorded by the LHCb experiment in 2011 and 2012 using pp collisions at a center of mass energy of 7 and 8 TeV respectively, corresponding to an integrated luminosity of 3.0 fb^{-1} . Data collected with the two magnet polarities are added up.

4.2 Monte Carlo sample

The Monte Carlo (MC) simulation sample is composed by an admixture of B^+ and B^0 mesons decaying to $D^*\pi(\pi)\mu\nu X$ final states. Both non-resonant and resonant decays are taken into account, as well as decay through secondary D^{**} resonances. Equal samples with opposite magnet polarities were produced. The PYTHIA6 [62]

and PYTHIA8 generators were also used to generate equal amounts of events. The Monte Carlo sample has been generated according to the exclusive branching ratios, collected in Tables 4.1 and 4.2.

Table 4.1. Exclusive semileptonic branching ratios used to generate the MC sample. The branching ratios for D^{**} resonances decaying into a D^* and at least one pion are already included, as well as D^* decaying into $D^0\pi$. For modes decaying via τ leptons, the branching fractions $\tau \rightarrow \mu\nu_\mu\nu_\tau$ are included too.

Channel	\mathcal{B}_{SL} (%)
$B^+ \rightarrow \bar{D}_0^*(2400)^0 (\rightarrow D^{*-}\pi^+\pi^0)\mu^+\nu_\mu$	0.03 ± 0.01
$B^+ \rightarrow \bar{D}_1(2430)^0 (\rightarrow D^{*-}\pi^+)\mu^+\nu_\mu$	0.09 ± 0.03
$B^+ \rightarrow \bar{D}_1(2420)^0 (\rightarrow D^{*-}\pi^+/D_0^{*-}\pi^+/\bar{D}_0^{*0}\pi^0)\mu^+\nu_\mu$	0.22 ± 0.03
$B^+ \rightarrow \bar{D}_2^*(2460)^0 (\rightarrow D^{*-}\pi^+/D_0^{*-}\pi^+/\bar{D}_0^{*0}\pi^0/D^{*-}\pi^+\pi^0)\mu^+\nu_\mu$	0.07 ± 0.03
Total $B^+ \rightarrow D^{**}\mu^+\nu_\mu$	0.41 ± 0.05
$B^+ \rightarrow D^{*-}\pi^+\mu^+\nu_\mu$	0.06 ± 0.06
$B^+ \rightarrow D^{*-}\pi^0\pi^+\mu^+\nu_\mu$	0.08 ± 0.08
Total non-resonant B^+ decays	0.15 ± 0.11
$B^+ \rightarrow \bar{D}_0^*(2400)^0 (\rightarrow D^{*-}\pi^+\pi^0)\tau^+\nu_\tau$	0.0012 ± 0.0002
$B^+ \rightarrow \bar{D}_1(2430)^0 (\rightarrow D^{*-}\pi^+)\tau^+\nu_\tau$	0.0076 ± 0.0011
$B^+ \rightarrow \bar{D}_1(2420)^0 (\rightarrow D^{*-}\pi^+/D_0^{*-}\pi^+/\bar{D}_0^{*0}\pi^0)\tau^+\nu_\tau$	0.0104 ± 0.0015
$B^+ \rightarrow \bar{D}_2^*(2460)^0 (\rightarrow D^{*-}\pi^+/D_0^{*-}\pi^+/\bar{D}_0^{*0}\pi^0/D^{*-}\pi^+\pi^0)\tau^+\nu_\tau$	0.0048 ± 0.0007
Total $B^+ \rightarrow D^{**}\tau^+\nu_\tau$	0.0241 ± 0.0020
Total semileptonic B^+ decays	0.58 ± 0.12
$B^0 \rightarrow D^{*-} (\rightarrow \bar{D}^0\pi^-)\mu^+\nu_\mu$	3.39 ± 0.074
$B^0 \rightarrow D_0^*(2400)^- (\rightarrow D^{*-}\pi^0\pi^0/D^{*-}\pi^+\pi^-)\mu^+\nu_\mu$	0.040 ± 0.014
$B^0 \rightarrow D_1(2430)^- (\rightarrow D^{*-}\pi^0)\mu^+\nu_\mu$	0.040 ± 0.015
$B^0 \rightarrow D_1(2420)^- (\rightarrow D^{*-}\pi^0/\bar{D}_0^{*0}\pi^-/D_0^{*-}\pi^0)\mu^+\nu_\mu$	0.12 ± 0.018
$B^0 \rightarrow D_2^*(2460)^- (\rightarrow D^{*-}\pi^0/D_0^{*0}\pi^-/\bar{D}_0^{*0}\pi^0/D^{*-}\pi^0\pi^0/D^{*-}\pi^+\pi^-)\mu^+\nu_\mu$	0.040 ± 0.018
Total $B^0 \rightarrow D^{*(*)-}\mu^+\nu_\mu$	3.63 ± 0.081
$B^0 \rightarrow D^{*-}\pi^0\mu^+\nu_\mu$	0.030 ± 0.031
$B^0 \rightarrow D^{*-}\pi^0\pi^0\mu^+\nu_\mu$	0.044 ± 0.044
$B^0 \rightarrow D^{*-}\pi^+\pi^-\mu^+\nu_\mu$	0.170 ± 0.166
Total non-resonant B^0 decays	0.24 ± 0.17
$B^0 \rightarrow D^{*-} (\rightarrow \bar{D}^0\pi^-)\tau^+\nu_\tau$	0.176 ± 0.059
$B^0 \rightarrow D_0^*(2400)^- (\rightarrow D^{*-}\pi^0\pi^0/D^{*-}\pi^+\pi^-)\tau^+\nu_\tau$	0.002 ± 0.001
$B^0 \rightarrow D_1(2430)^- (\rightarrow D^{*-}\pi^0)\tau^+\nu_\tau$	0.004 ± 0.001
$B^0 \rightarrow D_1(2420)^- (\rightarrow D^{*-}\pi^0/\bar{D}_0^{*0}\pi^-/D_0^{*-}\pi^0)\tau^+\nu_\tau$	0.006 ± 0.002
$B^0 \rightarrow D_2^*(2460)^- (\rightarrow D^{*-}\pi^0/D_0^{*0}\pi^-/\bar{D}_0^{*0}\pi^0/D^{*-}\pi^0\pi^0/D^{*-}\pi^+\pi^-)\tau^+\nu_\tau$	0.003 ± 0.001
Total $B^0 \rightarrow D^{*(*)-}\tau^+\nu_\tau$	0.190 ± 0.059
Total semileptonic B^0 decays	4.07 ± 0.20

Table 4.2. Exclusive decays and respective branching ratios of resonant states used in the MC generation. The branching ratio for the D^* decaying into $D^0\pi_*$ is already included.

Channel	\mathcal{B}_{SL} (%)
$D_0^*(2400)^0 \rightarrow D^{*+}\pi^-\pi^0$	5.42
$D_1(2430)^0 \rightarrow D^{*+}\pi^-$	33.85
$D_1(2420)^0 \rightarrow D^{*+}\pi^-$	27.08
$D_1(2420)^0 \rightarrow D_0^*(2400)^+\pi^-$	2.11
$D_1(2420)^0 \rightarrow D_0^*(2400)^0\pi^0$	0.70
$D_2^*(2460)^0 \rightarrow D^{*+}\pi^-$	11.73
$D_2^*(2460)^0 \rightarrow D_0^*(2400)^+\pi^-$	1.19
$D_2^*(2460)^0 \rightarrow D_0^*(2400)^0\pi^0$	0.40
$D_2^*(2460)^0 \rightarrow D^{*+}\pi^-\pi^0$	0.54
$D_0^*(2400)^+ \rightarrow D^{*+}\pi^0\pi^0$	2.71
$D_0^*(2400)^+ \rightarrow D^{*+}\pi^-\pi^-$	5.42
$D_1(2430)^+ \rightarrow D^{*+}\pi^0$	16.93
$D_1(2420)^+ \rightarrow D^{*+}\pi^0$	13.54
$D_1(2420)^+ \rightarrow D_0^*(2400)^0\pi^+$	1.41
$D_1(2420)^+ \rightarrow D_0^*(2400)^+\pi^0$	1.06
$D_2^*(2460)^+ \rightarrow D^{*+}\pi^0$	5.87
$D_2^*(2460)^+ \rightarrow D_0^*(2400)^0\pi^+$	0.79
$D_2^*(2460)^+ \rightarrow D_0^*(2400)^+\pi^0$	0.60
$D_2^*(2460)^+ \rightarrow D^{*+}\pi^0\pi^0$	0.27
$D_2^*(2460)^+ \rightarrow D^{*+}\pi^-\pi^-$	0.54

In particular, the generated decays to two pions final states are gathered in Table 4.3, which shows that about 70% of the generated B decays to $D^*\pi\pi$ are non-resonant, the remainder coming from decays in cascade or three-body D^{**} decays.

In this analysis, the MC simulation is used for the following purposes:

- definition of the discriminating variables and choice of the best cut exploiting a MultiVariate analysis (MVA)
- parametrisation of the Probability Density Functions (PDFs) describing signal and background shapes
- computation of the selection efficiencies.

Table 4.3. Summary of the exclusive decays to two pions final states generated in the MC sample. The branching ratios for D^{**} resonances decaying into D^* and at least two opposite charge pions are already included, as well as D^* decaying into $D^0\pi_*$. For modes decaying via τ leptons, the branching fractions $\tau \rightarrow \mu\nu_\mu\nu_\tau$ are included as well.

Channel	\mathcal{B}_{SL} (%)
$B^0 \rightarrow D^{*-}\pi^+\pi^-\mu^+\nu_\mu$	0.170 ± 0.166
Total non-resonant decays	0.170 ± 0.166
$B^0 \rightarrow D_0^*(2400)^-(\rightarrow D^{*-}\pi^+\pi^-)\mu^+\nu_\mu$	0.026 ± 0.009
$B^0 \rightarrow D_2^*(2420)^-(\rightarrow D^{*-}\pi^+\pi^-)\mu^+\nu_\mu$	0.003 ± 0.001
$B^0 \rightarrow D_0^*(2400)^-(\rightarrow D^{*-}\pi^+\pi^-)\tau^+\nu_\tau$	$1.2 \cdot 10^{-4} \pm 0.4 \cdot 10^{-4}$
$B^0 \rightarrow D_2^*(2420)^-(\rightarrow D^{*-}\pi^+\pi^-)\tau^+\nu_\tau$	$1.9 \cdot 10^{-4} \pm 0.6 \cdot 10^{-4}$
Total resonant decays	0.029 ± 0.009
$B^0 \rightarrow D_1(2420)^-(\rightarrow \bar{D}_0^{*0}\pi^-(\rightarrow D^{*-}\pi^+\pi^0))\mu^+\nu_\mu$	0.0104 ± 0.0015
$B^0 \rightarrow D_1(2420)^-(\rightarrow D_0^{*-}\pi^0(\rightarrow D^{*-}\pi^+\pi^-))\mu^+\nu_\mu$	0.0052 ± 0.0008
$B^0 \rightarrow D_2^*(2460)^-(\rightarrow \bar{D}_0^{*0}\pi^-(\rightarrow D^{*-}\pi^+\pi^0))\mu^+\nu_\mu$	0.0041 ± 0.0017
$B^0 \rightarrow D_2^*(2460)^-(\rightarrow D_0^{*-}\pi^0(\rightarrow D^{*-}\pi^+\pi^-))\mu^+\nu_\mu$	0.0021 ± 0.0009
$B^0 \rightarrow D_1(2420)^-(\rightarrow \bar{D}_0^{*0}\pi^-(\rightarrow D^{*-}\pi^+\pi^0))\tau^+\nu_\tau$	$0.5 \cdot 10^{-3} \pm 0.2 \cdot 10^{-3}$
$B^0 \rightarrow D_1(2420)^-(\rightarrow D_0^{*-}\pi^0(\rightarrow D^{*-}\pi^+\pi^-))\tau^+\nu_\tau$	$2.4 \cdot 10^{-4} \pm 0.8 \cdot 10^{-4}$
$B^0 \rightarrow D_2^*(2460)^-(\rightarrow \bar{D}_0^{*0}\pi^-(\rightarrow D^{*-}\pi^+\pi^0))\tau^+\nu_\tau$	$2.8 \cdot 10^{-4} \pm 0.9 \cdot 10^{-4}$
$B^0 \rightarrow D_2^*(2460)^-(\rightarrow D_0^{*-}\pi^0(\rightarrow D^{*-}\pi^+\pi^-))\tau^+\nu_\tau$	$1.4 \cdot 10^{-4} \pm 0.5 \cdot 10^{-4}$
Total secondary resonant B^0 decays	0.023 ± 0.003
$B^+ \rightarrow \bar{D}_1(2420)^0(\rightarrow D_0^{*-}\pi^+(\rightarrow D^{*-}\pi^+\pi^-))\mu^+\nu_\mu$	0.0104 ± 0.0016
$B^+ \rightarrow \bar{D}_2^*(2460)^0(\rightarrow D_0^{*-}\pi^+(\rightarrow D^{*-}\pi^+\pi^-))\mu^+\nu_\mu$	0.0041 ± 0.0017
$B^+ \rightarrow \bar{D}_1(2420)^0(\rightarrow D_0^{*-}\pi^+(\rightarrow D^{*-}\pi^+\pi^-))\tau^+\nu_\tau$	$4.9 \cdot 10^{-4} \pm 0.7 \cdot 10^{-4}$
$B^+ \rightarrow \bar{D}_2^*(2460)^0(\rightarrow D_0^{*-}\pi^+(\rightarrow D^{*-}\pi^+\pi^-))\tau^+\nu_\tau$	$2.8 \cdot 10^{-4} \pm 0.4 \cdot 10^{-4}$
Total secondary resonant B^+ decays	0.015 ± 0.002

4.3 Signal and background definition

The event topology is shown in Figure (4.1): the VELO allows a good knowledge of the separation between primary and secondary vertices. The primary vertex is the pp interaction point, while the secondary vertex is the B decay vertex. Due to the short lifetimes of D^* and D^{**} , it is not possible to resolve their paths from the secondary vertex. Thus, it is not possible to distinguish resonant and non-resonant decays from a topological point of view. The pions coming from a D^{**} or directly from the B are labelled as $\pi_{i^{**}}$ ($i = 1, 2$) and are referred to as hard pions, due to their momentum being typically higher than that of the pion from the D^* decay. The pion coming from the D^* is denoted as π_* and will be referred to as soft pion.

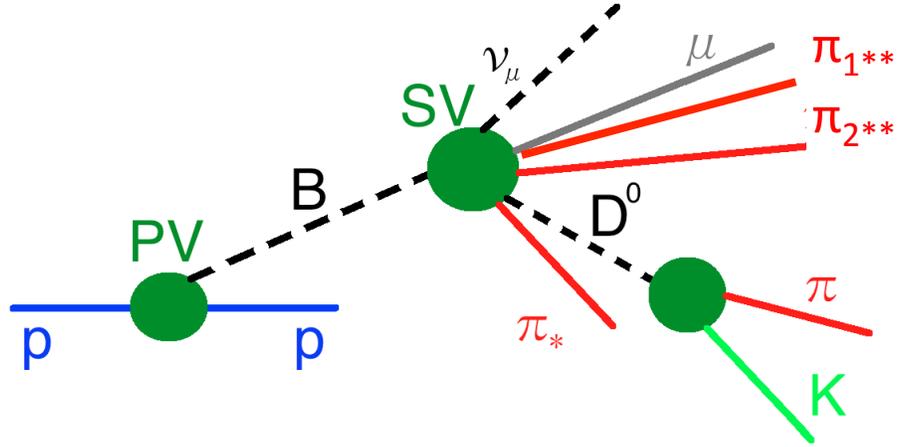


Figure 4.1. Event topology.

For the decay to a single pion final state, the signal is defined according to the following decay chains¹: for charged B ,

$$\begin{aligned}
 B^+ &\rightarrow D^{*-} \pi_{**}^+ \mu^+ \nu_\mu (X^0) \\
 D^{*-} &\rightarrow \bar{D}^0 \pi_*^- \\
 \bar{D}^0 &\rightarrow K^+ \pi^-
 \end{aligned} \tag{4.1}$$

while for neutral B ,

$$\begin{aligned}
 B^0 &\rightarrow D^{*-} \pi_{**}^+ \mu^+ \nu_\mu X^- \\
 D^{*-} &\rightarrow \bar{D}^0 \pi_*^- \\
 \bar{D}^0 &\rightarrow K^+ \pi^-
 \end{aligned} \tag{4.2}$$

The muon and the D^* have opposite charge, as well as the π_{**} and the D^* . For the decays into two pions final states, the signal is defined according to the semileptonic B meson decays:

$$\begin{aligned}
 B^+ &\rightarrow D^{*-} \pi_{1**}^+ \pi_{2**}^- \mu^+ \nu_\mu X^+ \\
 B^0 &\rightarrow D^{*-} \pi_{1**}^+ \pi_{2**}^- \mu^+ \nu_\mu (X^0)
 \end{aligned} \tag{4.3}$$

The D^{*-} decay chain is the same as (4.1) and (4.2).

Charged and neutral decays of both channels can be non-resonant or resonant, i.e. $B \rightarrow D^{**} (D^{**} \rightarrow D^* n \pi_{**}) \mu \nu_\mu X$ with $n = 1, 2$. In a resonant decay, there can be even two resonant states, indicated as D^{**} and secondary D^{**} , respectively.

¹Charge conjugated decay modes are always implied.

According to the exclusive branching ratios used to generate the MC sample, the decays to two pions final states can take place in three different ways: non-resonantly, resonantly or through secondary resonances. The B^+ decays to two pions final states happen only via secondary D^{**} resonances, while the B^0 decays occur in every possible way. In the case of a decay from a secondary resonance, the two pions can come from the same particle, or from different particles: in the latter case one pion comes from the D^{**} and the other is produced from the secondary resonance. This leads, in B^+ decays, to two possible oppositely charged pions which can be coupled to the D^* (lowest part of Table 4.3)

There are many sources of background:

- Combinatorial: events where at least one particle among D^0 , D^* , π_{**} and μ , does not come directly from a B decay. The major contribution is due to combinatorial pions, i.e. pions coming from other processes wrongly associated to correctly reconstructed $B \rightarrow D^* \mu \nu_\mu$ decay, which is the dominant decay mode.
- Prompt D^* : this kind of background is due to events where the D^* does not originate from a B decay, but is produced in the primary vertex. It can then be considered as a subsample of the combinatorial background. This kind of background peaks in the D^* invariant mass distribution.
- $B \rightarrow D^0$ decays: this kind of background is formed by events where the D^0 does not come from a D^* decay, but directly from the B meson. These background events can be evaluated directly in the data sample.
- Misidentified pions: charged particles (e.g. e , μ , K) misidentified as pions.
- J/ψ muons: muons coming from $J/\psi \rightarrow \mu^+ \mu^-$ decays which are not identified as muons but as pions and are then associated to the D^0 decay. This background peaks in the $\mu\pi_{fake}$ invariant mass distribution corresponding to J/ψ and $\psi(2S)$ masses.
- τ decays: $B \rightarrow D^*(n\pi)\mu\nu X$ decays where the muon does not come directly from the B meson, but from the semitauonic $B \rightarrow D^*(n\pi)\tau(\tau \rightarrow \mu\nu_\mu\nu_\tau)\nu_\tau X$ decay, where the τ promptly decays into a muon and two neutrinos. This kind of background is not eliminated in this analysis, but it is considered as a signal component.

4.4 Event reconstruction

The first step of the analysis is the selection of a D^* candidate per event, followed by the selection of a muon and one or two pion candidates. This process will be described in details in the following sections.

4.4.1 D^* selection

To select a single D^* candidate per event, the following cuts have been applied:

- Trigger line: a series of cuts performed during the reconstruction of the event at the trigger level;
- Topological and kinematical cuts to select the $D^* \rightarrow D^0 \pi_*$ and $D^0 \rightarrow K\pi$ samples, shown in Tables 4.4 and 4.5;
- A cut on the D^0 impact parameter with respect to the primary vertex to reduce the D^* prompt background to a negligible level;
- Requirements on the pion from D^0 decay and the muon to reduce the background due to J/ψ decays.

The trigger line chosen in this analysis requires that at least one track among the μ , K , π and π_* reconstructed particles in each event to satisfy the minimum requirements:

$$\begin{aligned}
 p_T &> 1700 \text{ MeV} \\
 p &> 12500 \text{ MeV} \\
 IP_{PV} &> 0.125 \text{ mm} ,
 \end{aligned} \tag{4.4}$$

where p_T is the transverse momentum, p the momentum and IP_{PV} the impact parameter with respect to the primary vertex.

To reconstruct the D^0 candidates, two long tracks with $\chi^2/d.o.f. < 4$ of opposite charge must be combined: one track is identified as a kaon by requiring the PIDK (i.e. the likelihood for the kaon hypothesis) to be larger than 4, while the other one is required to have a PIDK < 10 , without requiring any pion identification. The two tracks are required to have a momentum larger than 2 GeV, a transverse momentum larger than 300 MeV and a χ^2 of the impact parameter with respect to the primary vertex larger than 9. The $K\pi$ vertex must have a $\chi^2/d.o.f. < 6$ and the $K\pi$ invariant mass must be within a 80 MeV range with respect to the PDG value and the D^0 transverse momentum must be > 1400 MeV.

The D^* are selected requiring a long track ($\chi^2/d.o.f. < 5$) of same charge as the π_* ; there are no PID requirements on the π_* , but its transverse momentum must be larger than 110 MeV and its impact parameter with respect to the primary vertex larger than 0.04 mm. The $D^0\pi_*$ vertex must have a $\chi^2/d.o.f. < 20$ and a mass within 70 MeV with respect to the PDG value.

To cut the background due to prompt D^* , the following condition is applied in the first stage selection:

$$\log [IP_{PV}^{D^0}(\text{mm})] > -3, \quad (4.5)$$

where $IP_{PV}^{D^0}$ with respect to the primary vertex. This reduces this background to a negligible fraction. Figure 4.2 shows the distributions of this variable in data and MC. An excess of events in data, due to this kind of background, is removed by the requirement in Equation (4.5).

In order to remove the J/ψ background, the following conditions must be fulfilled:

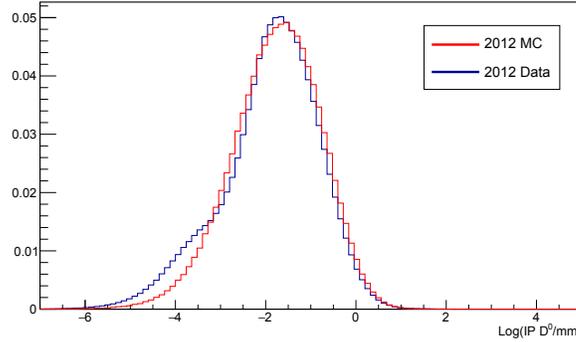


Figure 4.2. Superposition of 2012 data and MC12 events. An excess of data below -3 is due to prompt D^* background.

the pion from the D^0 decay shall not satisfy the muon identification criteria and the value of the muon-pion invariant mass must be outside the J/ψ peak:

$$|m(\mu\pi) - 3096.0| < 45.0 \text{ MeV}. \quad (4.6)$$

Figure (4.3) confirms the correctness of the first requirement: on the left the total spectrum of the invariant mass is plotted, while on the right there are just the events with pions satisfying the muon identification criteria.

4.4.2 Muon selection

The muon identification is performed extrapolating a long track with $\chi^2/d.o.f. < 5$ into the muon stations. The hits are searched within Fields Of Interest (FOIs)

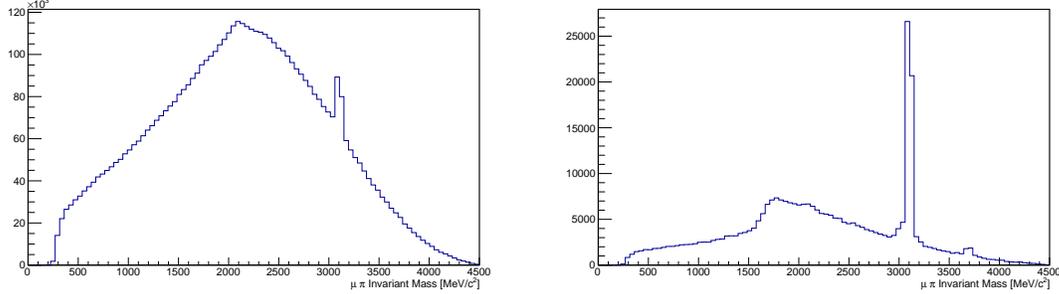


Figure 4.3. $\mu\pi$ invariant mass distributions on 2012 data. On the left, the full spectrum. On the right, events with pions satisfying the muon identification criteria.

Table 4.4. Cuts applied in the selection of the $D^* \rightarrow D^0\pi_*$ sample.

Variable	Selection cut
track $\chi^2/d.o.f.(\pi_*)$	< 5
$p_T(\pi_*)$	> 110 MeV
$IP_{PV}(\pi_*)$	> 0.04 mm
vertex $\chi^2/d.o.f.(D^0\pi_*)$	< 20
$ m(D^0\pi_*) - m(D_{PDG}^*) $	< 70 MeV

Table 4.5. Cuts applied in the selection of the $D^0 \rightarrow K\pi$ sample.

Variable	Selection cut
track $\chi^2/d.o.f.(K, \pi)$	< 4
K PID	$\log \mathcal{L}_K > 4$
π PID	$\log \mathcal{L}_\pi < 10$
$p(K, \pi)$	> 2000 MeV
$p_T(K, \pi)$	> 300 MeV
vertex $\chi^2/d.o.f.(K\pi)$	< 6
$p_T(D^0)$	> 1400 MeV
$IP_{PV}(D^0)$	> 0.04 mm
$ m(K\pi) - m(D_{PDG}^0) $	< 70 MeV

around the extrapolation point of the track in each muon station: if a track has hits in the corresponding FOI in 2 to 4 (depending on the momentum of the muon) muon stations, the particle is considered a muon candidate. To improve the muon selection, the slope of the track in the muon system is compared with the slope of the track in the main tracker and the average distance between the reconstructed track and the hits in the corresponding field of interest is evaluated. The muon momentum has to be larger than 3 GeV, since particles with $p < 3$ GeV do not reach the muon stations after the calorimeters, while the transverse momentum is

required to be larger than 800 MeV. Applying these criteria, the muon selection efficiency is about 93% (for muons above 3 GeV). The pion misidentification rate is reduced requiring $\log L_\mu > 0$.

4.4.3 π^{**} pre-selection

Once the D^* decay chain is reconstructed, pions can be added to it. At this purpose, a tool in the LHCb DaVinci software was developed and used. Such tool is able to add particles to a specific decay chain once a certain particle (called the target), belonging to this chain is found. At this point it should be straightforward that the target particle is the D^* and the added particles are pions. This tool was used for the production of both the data and the MC samples.

This tool was developed as general as possible but in this case a precise pre-selection of π^{**} candidates was performed, by applying the criteria listed in Table 4.6. To identify a particle as a π , a long or upstream track with $\chi^2/d.o.f. < 3$ is required. Pions are also required to satisfy some kinematical conditions on the momentum p and on the transverse momentum p_T . Other requirements involve topological variables: the absolute value of the difference between the polar angle ϕ of the target particle and the one of the added particle must be greater than 5 *mrad*, the angle θ between the pions' flight direction and the positive direction of the beam axis must be greater than 12 *mrad*. Moreover, the $(D^*\pi\mu)$ vertex χ^2 must be less than 10. No PID requirement is done, as it might affect the MVA.

No requirements on the pion charge are done yet, so pions with both opposite and same charge with respect to the D^* are selected.

Once all these cuts are applied, the average number of π^{**} passing the pre-selection is quite high, about 20, suggesting that a further selection needs to be applied; the multiplicity distribution for both same charge and opposite charge pions are plotted in Figure 4.4.

Table 4.6. Pion preselection cuts applied.

Variable	Selection cut
p	[2,200] GeV
p_T	< 10 GeV
track $\chi^2/d.o.f.$	< 3
track type	long or upstream
θ	> 12 <i>mrad</i>
$ \Delta\phi $	> 5 <i>mrad</i>
$(D^*\pi\mu)$ vertex χ^2	< 10

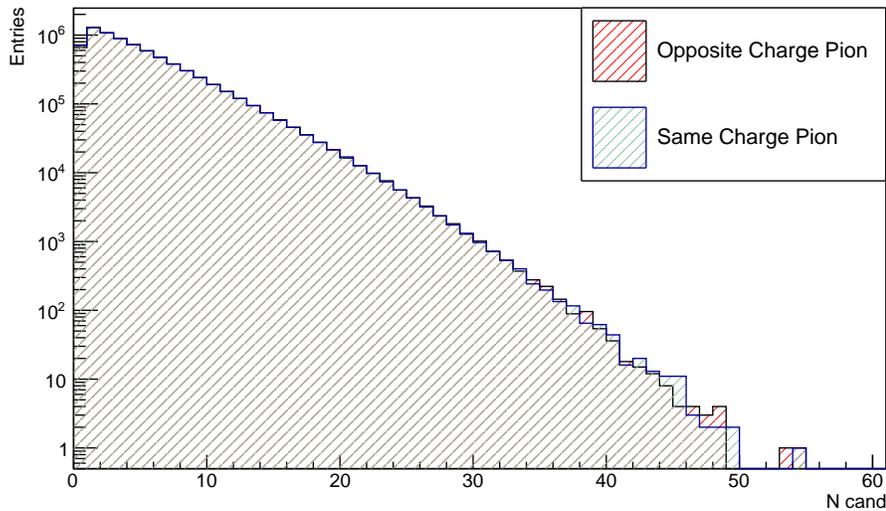


Figure 4.4. Pion multiplicity distribution after the preselection and before the BDT cut on data.

4.5 Multivariate Analysis

A multivariate analysis (MVA) is performed with the purpose of classifying the events, discriminating signal from background. An event is described by multiple variables, denoted as feature variables, which are usually correlated. To take into account the correlations between them, it is necessary to treat them in a multivariate way. The MVA basically consists in applying subsequently cuts to the feature variables in order to reduce the background and obtain an enhanced signal-like sample. The feature variables can be thought from a geometrical point of view as describing a multidimensional space: aim of the MVA is to reduce the dimension of the feature space according to a given method and classify the events enclosed in this region.

Figure 4.5 shows a two dimensional example of multivariate analysis. (a) and (b) are the distributions of the two feature variables x_1 and x_2 , which are strongly correlated, as it is evident from the scatter plot (c). Figure (c) shows that it is possible to separate two classes (the blue and the red one) though. The one dimensional projections (d) and (e), which are the variable marginal densities $f(x_1) = \int G(x_1, x_2)dx_2$ and $f(x_2) = \int G(x_1, x_2)dx_1$, overlap, so it is not obvious which cut apply to separate the two classes. A linear discriminant function $f(x_1, x_2)$ instead reveals two well separated distributions in (f): an optimal cut on this function allows to separate the signal class from the background one.

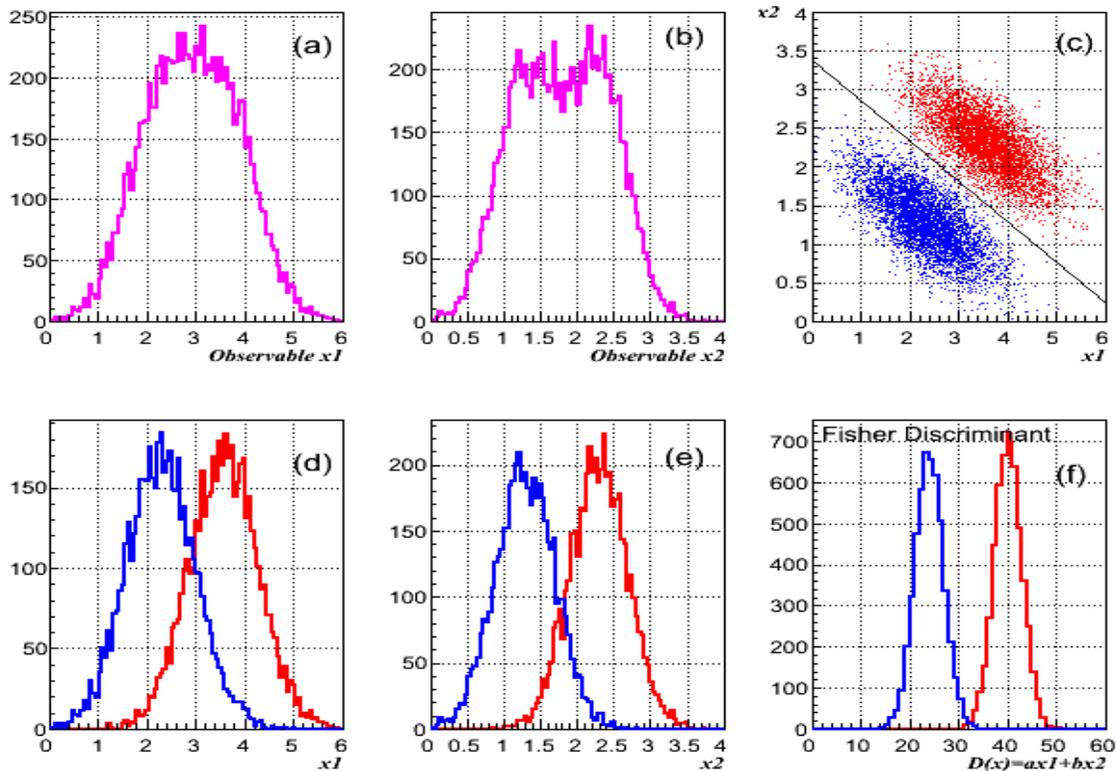


Figure 4.5. Example of bidimensional multivariate analysis.

The MVA consists in finding the best approximation $f(x, w)$ of the unknown function $f(x)$, where x is the vector of the feature variables and w some parameters typical of the method adopted. The approximated function is inferred directly from the given data, without a priori knowledge of the function. This kind of approach is called machine learning or learning from data: the function is trained on a sample (usually a simulated one) in order to extract the parameters w of the model, and then tested on a statistically independent sample.

There are many different MVA methods: rectangular cut optimisation, projective likelihood estimator, multidimensional likelihood estimator, likelihood estimator using self-adapting phase-space binning, k-Nearest Neighbour classifier, H-Matrix discriminant, Fisher discriminants, linear discriminant analysis, function discriminant analysis, artificial neural network, support vector machine, boosted decision trees, predictive learning via rule ensembles. The classifier adopted in this analysis is the Boosted Decision Trees (BDT), because it is a method of simple interpretation, since it can be visualised as a bidimensional tree structure, but more powerful than other similar techniques, like the rectangular cuts.

The processing, evaluation and application of the multivariate classification is provided by the Toolkit for Multivariate Analysis (TMVA) [63], a ROOT integrated environment.

4.5.1 Boosted Decision Trees

A decision tree is a binary tree structural classifier. A schematic view of a decision tree is in Figure 4.6: starting from the root node, a sequence of binary splits is applied. Every split is performed cutting on the discriminating variable x_i which has the highest separating power between signal and background; thus, the same variable can be used in several nodes, while other variables can be not used at all. In this way, the feature space is split into many regions and each and every one of them can be classified either as signal-like or background-like, depending on the nature of the major number of events in the final leaf node, as it is called a terminal node.

The advantage of the decision trees method with respect to the rectangular cuts method (both are based on a sequence of cuts) is that while the rectangular cuts identify only one region in the feature space, a decision tree gives rise to many regions in the feature space, keeping, at the same time, a simple bidimensional structure. The decision trees method has also other advantages: the tolerance to missing variables in both training and test samples and insensitivity to irrelevant variables.

A drawback of the method is the instability with respect to statistical fluctuations: if two variables have similar discriminating power at a split node, a statistical fluctuation can lead to the choice of one variable instead of the other, giving rise to an altered classifier response. To overcome this problem, the decision tree is boosted. The boosting consists in generating a forest of decision trees instead of a single tree, and classifying an event according to the major number of times it has been recognised as signal-like or background-like in each tree. This process allows to stabilise the classifier response and drastically improve the separation performance.

The boosting algorithm used in this analysis is the adaptive boost (AdaBoost). It works in the following way: the first decision tree is trained; during the training of the next decision tree, a weight, denoted as boost weight α , is assigned to the events misclassified in the previous process. α is computed from the misclassification

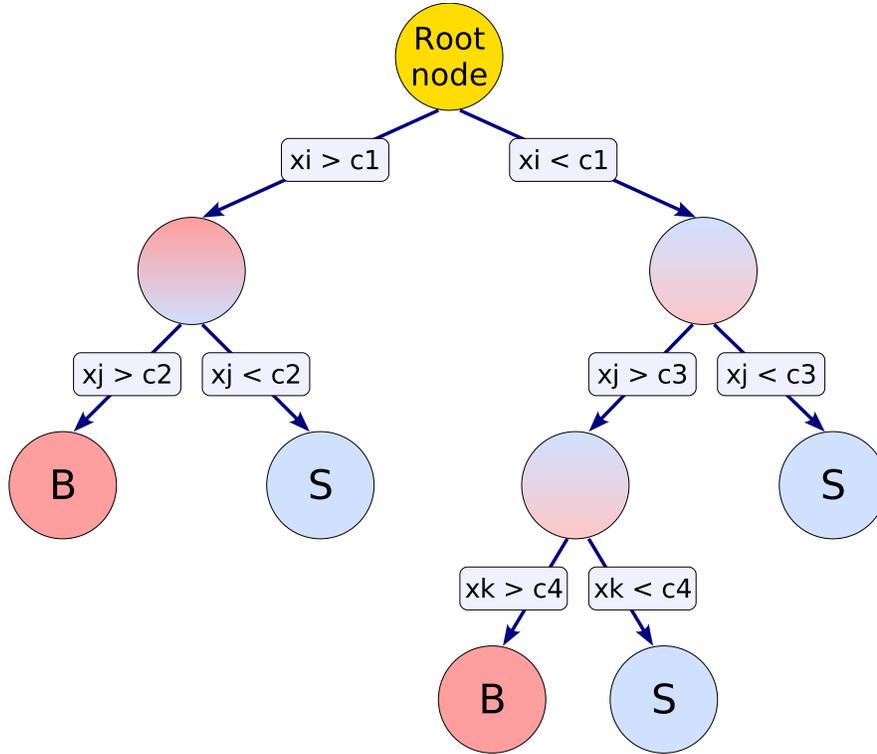


Figure 4.6. Illustration of a binary decision tree with feature variable $x_{i,j,k}$.

rate, err, of the previous tree as:

$$\alpha = \frac{1 - \text{err}}{\text{err}}. \quad (4.7)$$

The weights of the whole sample are then renormalised so that their sum is constant.

If the output of the single decision tree is $h(\mathbf{x})$, with $h(\mathbf{x}) = +1$ for signal and $h(\mathbf{x}) = -1$ for background, the boosted decision trees classification is given by

$$y_{\text{boost}}(\mathbf{x}) = \frac{1}{M} \sum_i^M \log(\alpha_i) \cdot h_i(\mathbf{x}), \quad (4.8)$$

where M is the number of decision trees trained. A background-like event corresponds to a small value of $y_{\text{boost}}(\mathbf{x})$, a signal-like event to a large value of $y_{\text{boost}}(\mathbf{x})$. There are many booking options available to customise this method; the ones set in this analysis are the following:

- number of trees in the forest;

- maximum depth allowed for each decision tree;
- number of cuts: number of grid points in variable range used in finding optimal cut in node splitting;
- boosting type;
- separation criteria for node splitting;
- prune method: the pruning is used to remove statistically insignificant branches.

4.5.2 Training and test samples

The Monte Carlo sample is split into two subsamples: one, denoted as MVA-sample, is used for the multivariate analysis, the other one, the Ana-sample, is used to test the method and determine the probability density function to describe signal and background shape of the *Log IP* distribution. The multivariate analysis proceeds through the following stages:

- training on the Monte Carlo sample (MVA-sample);
- testing and evaluating the performance of the classifier on a statistically independent sample (MVA-sample);
- apply the classification on the Ana-sample and eventually on the data sample.

To train and test the method on statistically independent samples, the TMVA-sample is split randomly in two halves. The definition of signal and background candidates in the training phase is the following:

- Signal: resonant and non-resonant pions coming from $B \rightarrow D^* \pi \mu \nu_\mu X$ decays, where the muon can be produced directly from B decays or coming from a τ produced from the B . At this point, no requirement on the π charge is done yet, so this definition of signal includes both pions with same and opposite charges with respect to the D^* .
- Background: combinatorial candidates, i.e. at least one particle among D^* , π_{**} and μ does not come from the B .

In addition to the criteria listed in Table 4.6, it is further required that the $D^* \pi_{**}$ invariant mass is in the interval [2.2,3.2] GeV and that the number of added pions is different from 0.

4.5.3 Input variables definition

The input variables used in the MVA are chosen according to the separation $\langle S^2 \rangle$, which is defined as:

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(s(x) - b(x))^2}{s(x) + b(x)} dx, \quad (4.9)$$

where $s(x)$ and $b(x)$ are the probability density distributions for signal and background, respectively. The separation has value 1 when signal and background shapes do not overlap, and value 0 when they have the same shape.

Table 4.7. Input variables for MVA, listed along with their separation power.

Variable	Definition	Separation
$\pi_{**}(IP_{PV}\chi^2)$	Logarithm of the π_{**} impact parameter χ^2 with respect to the primary vertex (mm)	$6.495 \cdot 10^{-1}$
$\pi_{**} \log p_T$	Logarithm of π_{**} transverse momentum (GeV)	$4.355 \cdot 10^{-1}$
$\Delta\phi(D^*/\pi_{**})$	Difference between D^* and π_{**} momenta azimuthal angles	$3.343 \cdot 10^{-1}$
$\Delta\phi(B/\pi_{**})$	Difference between $D^*\mu$ and π_{**} momenta azimuthal angles	$2.983 \cdot 10^{-1}$
$\Delta\eta(B/\pi_{**})$	Difference between $D^*\mu$ and π_{**} pseudorapidities	$2.582 \cdot 10^{-1}$
$\Delta p(B/\pi_{**})$	Difference between $D^*\mu$ and π_{**} momenta magnitudes (GeV)	$5.056 \cdot 10^{-2}$

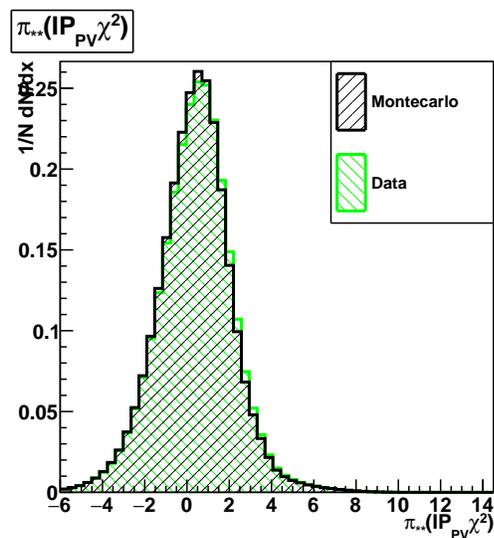
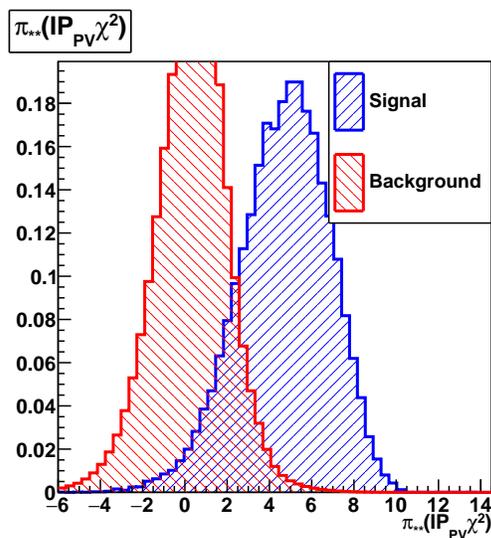
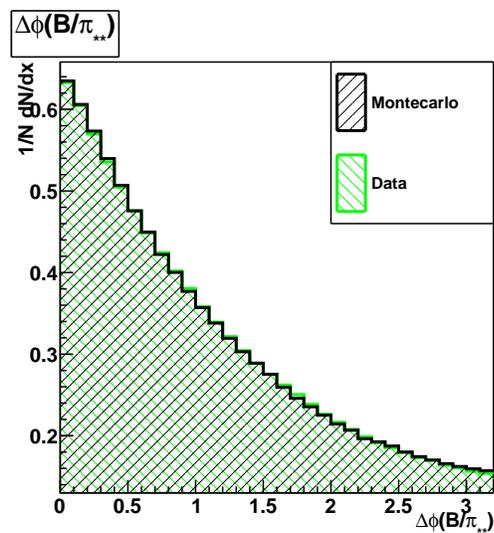
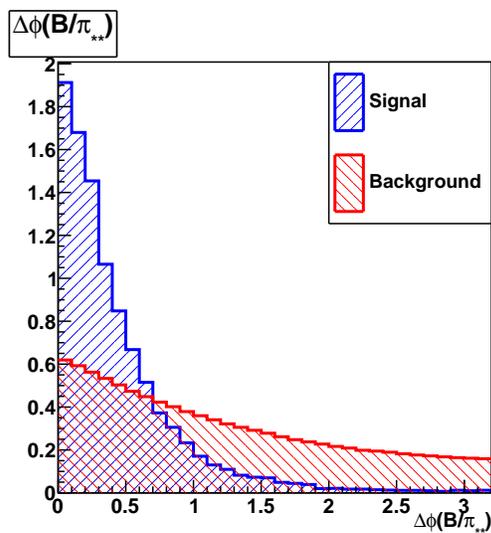
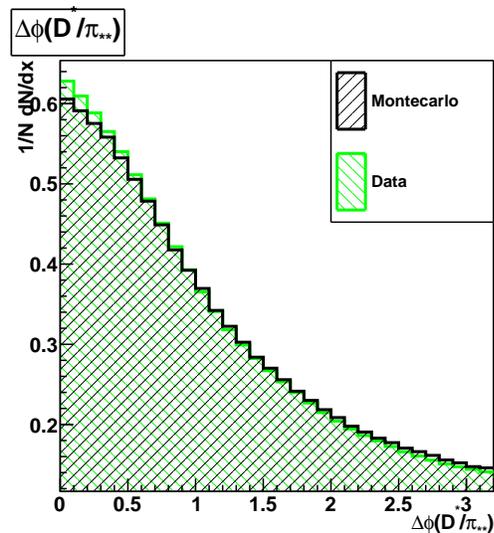
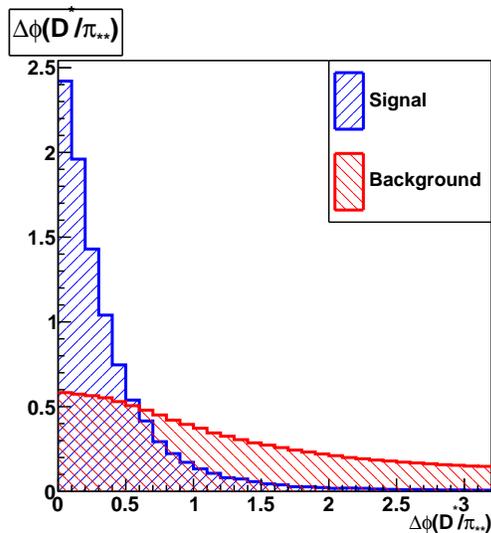
The feature variables chosen for the multivariate analysis are listed in Table 4.7, along with their separation value.

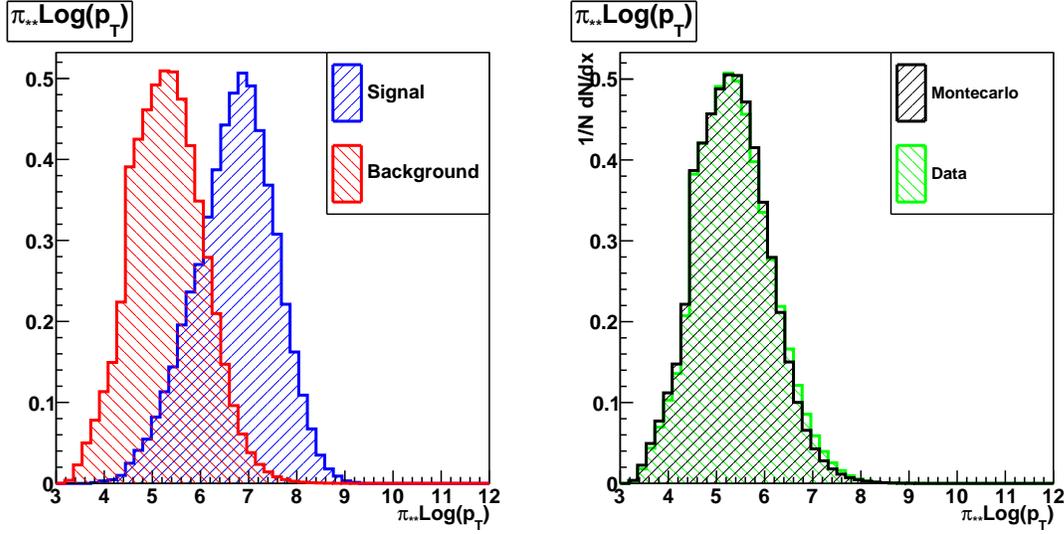
In Figure 4.9 for each input variable, signal and background distribution are compared and show a good discriminating power. Monte Carlo and data samples are also in agreement: this is a necessary condition, since the MVA is trained and tested on the MC sample, but it is applied on the data sample eventually.

Before the variables enter the proper multivariate analysis, a pre-processing phase is performed in order to reduce correlations among them and transform their shapes into gaussian forms. Three kinds of transformations have been applied to the input variables:

- Decorrelation (D): the variables are usually correlated leading to a potential decrease in the classifier performance. The decorrelation of the variables is carried out computing the square root C' of the covariance matrix C :

$$D = S^T C S \quad \Rightarrow \quad C' = S \sqrt{D} S^T. \quad (4.10)$$





First C is diagonalised with the aid of the symmetric matrix S and then the squared root of the diagonalised matrix D is computed. The variable x then transforms as: $\mathbf{x} \rightarrow (C')^{-1}\mathbf{x}$

The decorrelation process is complete if the variables are linearly correlated and gaussian-distributed.

- Principal Component Decomposition (PCD): as the decorrelation, PCD eliminates linear correlations between gaussian variables. This transformation is a rotation of the frame of reference such that largest variance for the data sample is projected on an axis: this decomposition allows to reduce the dimensions of the space, eliminating the variables with negligible variance. The vector of principal components $\mathbf{x}_U^{PC}(i) = (x_{U,1}^{PC}(i), \dots, x_{U,n_{var}}^{PC}(i))$ for the event i for $U = S$ and $U = B$ (respectively, signal and background distributions) is obtained with the following transformation:

$$\mathbf{x}_{U,k}^{PC}(i) = \sum_{\ell=1}^{n_{var}} (x_{U,\ell}(i) - \bar{x}_{U,\ell}) v_{U,\ell}^{(k)}, \quad \forall k = 1, n_{var}. \quad (4.11)$$

The vector $\mathbf{x}(i)$ is the vector of the input variables, \bar{x}_U and $v_U^{(k)}$ are the vector of means and the eigenvector, respectively.

- Gaussianisation (G): the variables are transformed such that their distributions have a gaussian shape. This is done in two steps: first, the cumulative

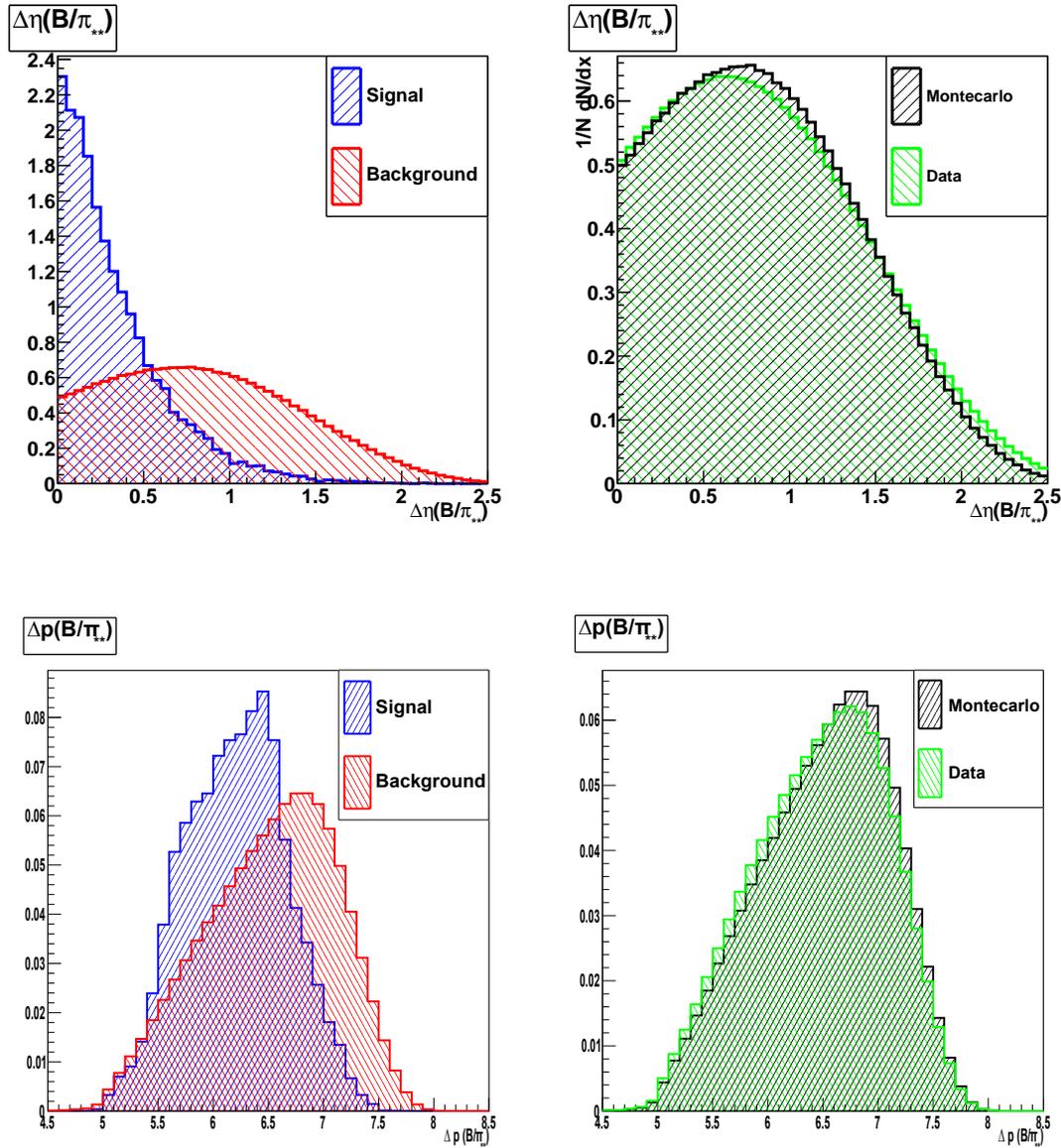


Figure 4.9. Input variables of the MVA: for each variable, on the left the signal and background distributions for simulated data are plotted, on the right the Monte Carlo and real data distributions are superimposed.

distribution function is computed; then the inverse error function is used to transform it in a normalised gaussian, i.e. mean equal to zero and width equal to one. So, the input variables transform this way:

$$x \rightarrow \sqrt{2} \cdot \operatorname{erf}^{-1} \left(2 \int_{-\infty}^x \hat{x}(x') dx' - 1 \right). \quad (4.12)$$

The linear correlation coefficients between the input variables for signal and background are shown in Figure 4.10.

4.5.4 TMVA output

The BDT booking options adopted in the training and testing phases are gathered in Table 4.8.

Since the tree depth has been limited to three split nodes, there has been no need

Table 4.8. BDT booking options.

Option	Value
Number of Trees	1000
Maximum Depth	3
Boosting Type	AdaBoost
Separation Type	Gini Index
Number of Cuts	20
Prune Method	None

to prune the method to contain the overtraining. The separation criteria adopted to evaluate the performance of a variable and a specific cut requirement is the *Gini Index*, defined as $p \cdot (1 - p)$ where p is the purity of the sample. So the Gini Index assumes maximum value when the sample is fully mixed ($p = 1/2$) and null value when the sample is completely pure ($p = 1$). Table 4.9 reports the BDT input variable ranking, obtained counting how often a variable is used in the splitting process and weighting each split occurrence by the squared of the separation gain achieved and the number of events at the node. This gives a hint on the importance of the variable in the classification process.

Table 4.9. Variables ranking.

Variable	Variable Importance
$\pi_{**}(IP_{PV}\chi^2)$	0.2148
$\Delta p(B/\pi_{**})$	0.1795
$\pi_{**} \log p_T$	0.1685
$\Delta\phi(D^*/\pi_{**})$	0.1662
$\Delta\phi(B/\pi_{**})$	0.1384
$\Delta\eta(B/\pi_{**})$	0.1326

Figure 4.11 shows the overtraining check. The overtraining occurs when a machine learning problem has too few degrees of freedom: this happens when

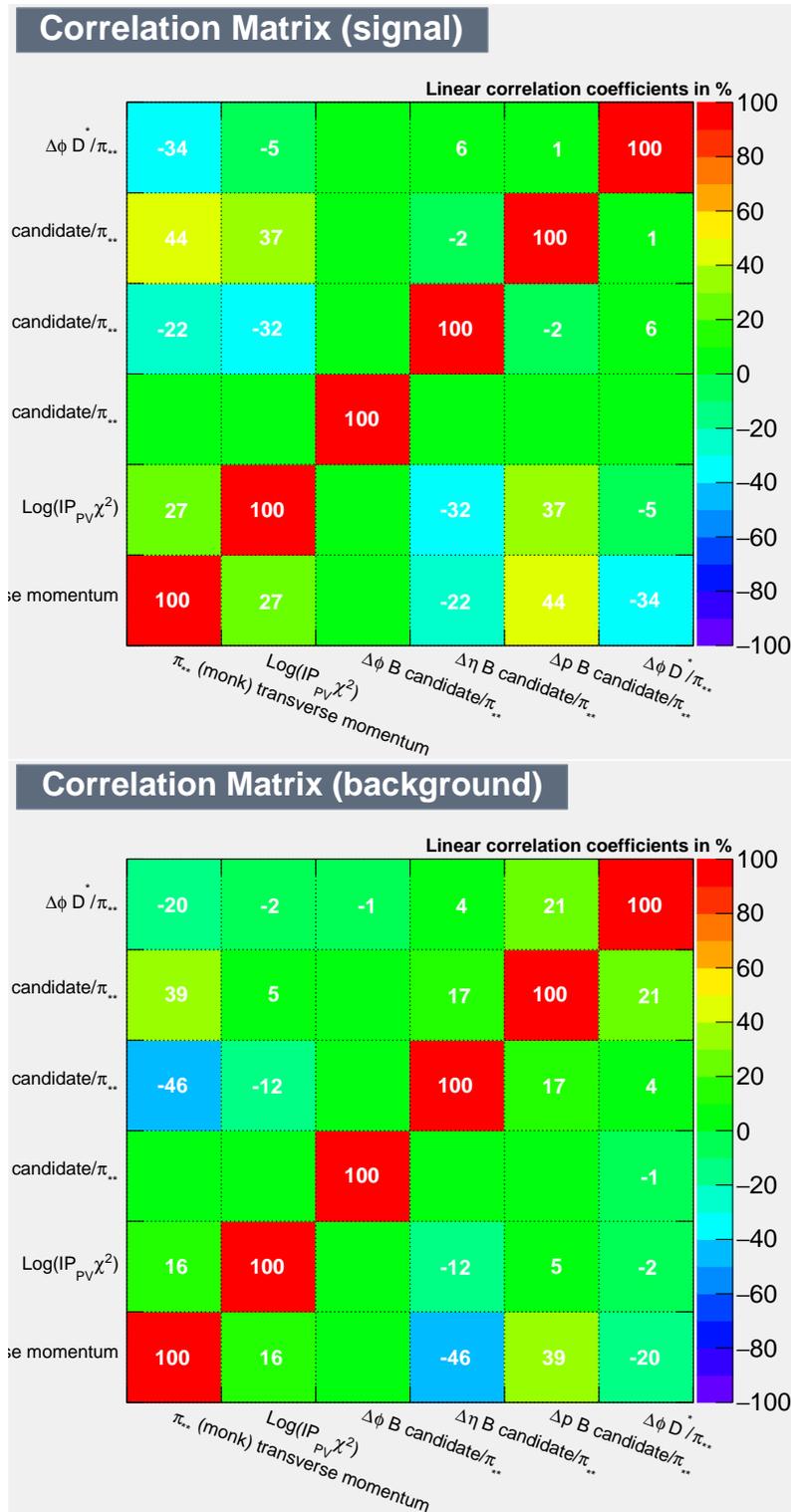


Figure 4.10. Correlation matrices for the MVA input variables in signal (top) and background (bottom) samples.

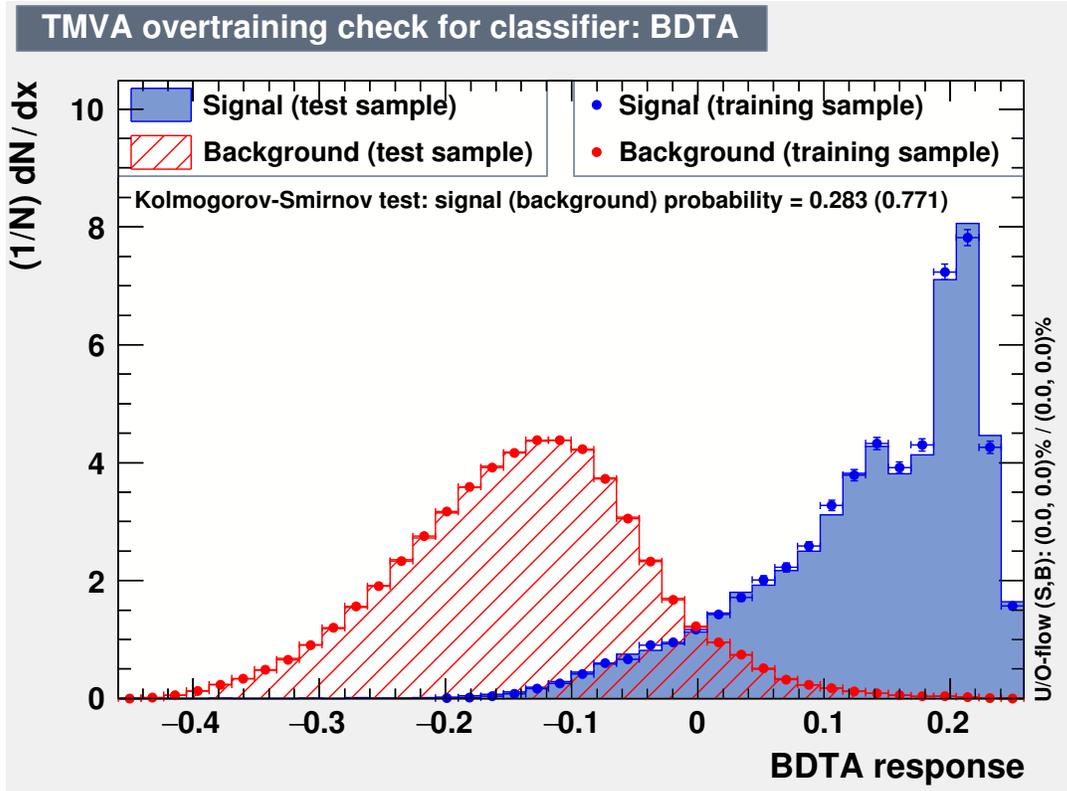


Figure 4.11. Overtraining check.

there are too many model parameters with respect to data points. Basically, this means that the feature space has been split too many times considering the effective amount of events: so trying to increase the classification performance over the achievable one leads to an actual decrease in performance such that when the method is tested on a statistically independent sample it gives rise to different results. So to detect the overtraining, it is sufficient to compare the results on training and test samples.

In Figure 4.11 the BDT output distributions for signal and background are plotted superimposing training and test sample results. Since the distributions for both training and test samples are similar, the conclusion is that there is no overtraining. In Figure 4.12 the signal and background efficiencies with respect to the cut value are evaluated. Signal efficiency, signal purity, signal significance and background rejection are evaluated for each value of the BDT cut. The purpose is to find the optimal cut on the classifier output which maximises the significance

$$\mathcal{S} = \frac{S}{\sqrt{S+B}}, \quad (4.13)$$

where S and B are the classifier stand for signal and background. The number of the input signal and background events are the values obtained from the Monte Carlo sample used for the MVA. The optimal cut on the classifier output is found to be 0.1348, corresponding to a signal significance of ~ 130 , a signal efficiency of $\epsilon_{sgn} \sim 65\%$ and a background rejection equal to $1 - \epsilon_{bkg} \sim 99.42\%$.

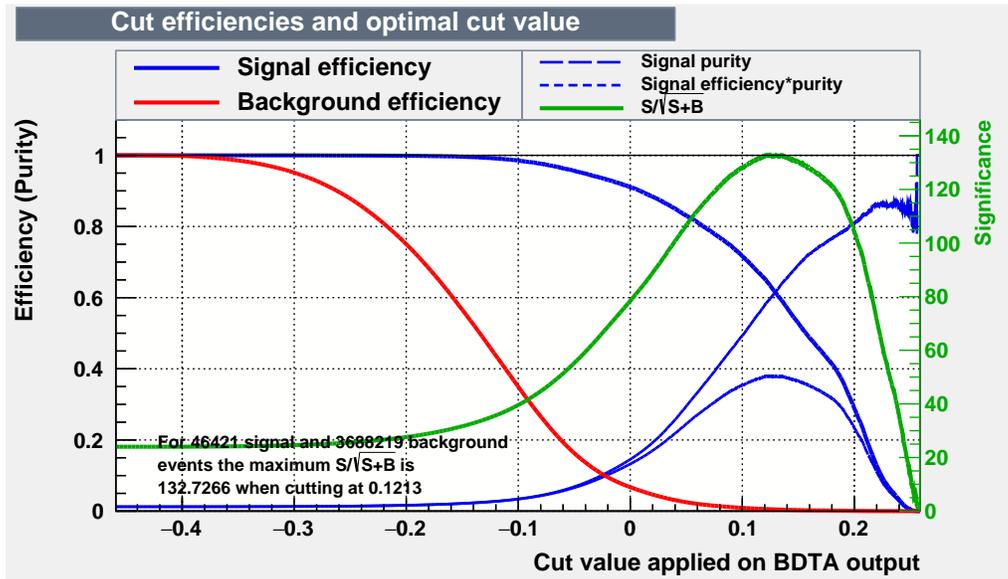


Figure 4.12. Background and signal efficiencies, signal purity and signal significance as function of the cut on the classifier output.

4.6 π_{**} selection

The last selection phase associates one pion of opposite charge and one of same charge to the D^* candidate of each event. Before the MVA, the average pion multiplicity per event, as shown in Figure 4.4, is about 6 for both opposite charge and same charge pions with respect to the D^* . The multivariate analysis presented in the previous Section provides a criterium to select a single pion candidate of each charge sign. For each event, a loop on all the pion candidates with opposite charge is performed and the pion with the highest BDT output has been kept. The same process is applied to the pion candidates of same charge. The output of the BDT classifier which maximises the significance is used as a threshold. In Figure 4.13 the pion multiplicity for candidates which pass the BDT cut is shown. The majority of events has no pions with BDT value above threshold: excluding these events, the average number of pions per event is about 1.03. A boolean variable,

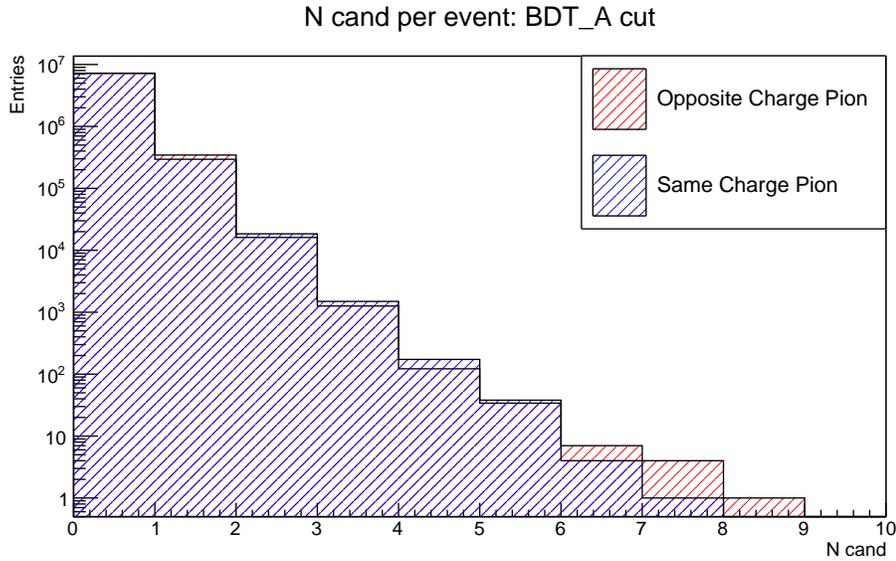


Figure 4.13. Pion multiplicity for both opposite charge and same charge pions after preselection and BDT cut.

isTwoPiFlag, is used to identify the case with two pions in the final state. In Table 4.10 the number of entries for each component in the region $BDT > 0.1213$ is shown. In one pion final state decays, signal events are about twice more abundant than background events. Among the resonances, the highest contribution is given by D_1 , while secondary resonance are only 0.2% of the signal events.

In two pion final state decays, about 94 % of the signal sample is given by non-resonant decays. There are three kinds of combinatorial backgrounds: combinations where both pions do not originate from signal events, combinations where only one pion originates from signal events, which can be either the opposite or the same charge pion. The dominant background is the one where both pions are combinatorial.

Table 4.10. Composition of the MC Ana sample after preselection and BDT cut. Each component is obtained requiring the MC truth identification. A statistical poissonian uncertainty is assumed.

Component		isTwoPiFlag	Entries
D^{**}	secondary D^{**}		
$D_0^*(2400)$	-	-	914 ± 30
$D_1(2420)$	-	-	11017 ± 104
$D_1(2430)$	-	-	4370 ± 66
$D_2^*(2460)$	-	-	3518 ± 59
$D_1(2420)/D_2^*(2460)$	$D_0^*(2400)$	-	51 ± 7
Non-resonant		-	6739 ± 82
Total Signal		-	26558 ± 163
Total background	-	-	14683 ± 121
D^{**}	secondary D^{**}		
$D_0^*(2400)$	-	1	23 ± 5
$D_2^*(2460)$	-	1	14 ± 3
$D_1(2420)/D_2^*(2460)$	$D_0^*(2400)$	1	1 ± 1
Non-resonant		1	327 ± 18
Total Signal		1	348 ± 19
Combinatorial $\pi_{1^{**}}$		1	43 ± 7
Combinatorial $\pi_{2^{**}}$		1	103 ± 10
Combinatorial $\pi_{1^{**}}$ and $\pi_{2^{**}}$		1	392 ± 20
Total background		1	538 ± 23

Chapter 5

Fit to Monte Carlo and data distributions

In this Chapter the fit procedure is explained. The discriminating variable used is the impact parameter of the pion with respect to the secondary vertex: its distribution is fitted on both MC and data samples for the investigated channels $B \rightarrow D^* \pi \mu \nu_\mu X$ and $B \rightarrow D^* \pi \pi \mu \nu_\mu X$. The fit performed on the MC distributions allows to parametrise the shape and develop the fitting procedure which is then applied on data. Another channel is also considered, $B \rightarrow D^* \mu \nu_\mu X$, used as normalisation channel. For each of the three channels, the fits allow to extract the signal yields needed to compute the branching ratios.

The branching ratios represent the final aim of this work: first, they are calculated relative to a normalisation channel, then their absolute values are computed.

5.1 Discriminating variable

The variable used to discriminate between signal and background events is the logarithm of the impact parameter of the pion with respect to the secondary vertex. The impact parameter (IP) is defined as the perpendicular distance between the track path and the position of the vertex. Figure 5.1 shows the distribution of this variable for pions with opposite and same charge with respect to the D^* : signal and background are quite separated, meaning that this variable has a good discriminating power. The background distribution peaks around -1.5, while the signal distribution around -3.

In two pions final states, the two pions, coming either directly from a B meson decay or from D^{**} or secondary D^{**} decays, are expected to be softer with respect

to the single pion final states. As a consequence the measured IP will be affected by larger uncertainties and the $LogIP$ distribution is expected to be wider. The impact parameter is a topological variable, so it does not allow to distinguish between the non-resonant and the different resonant contributions.

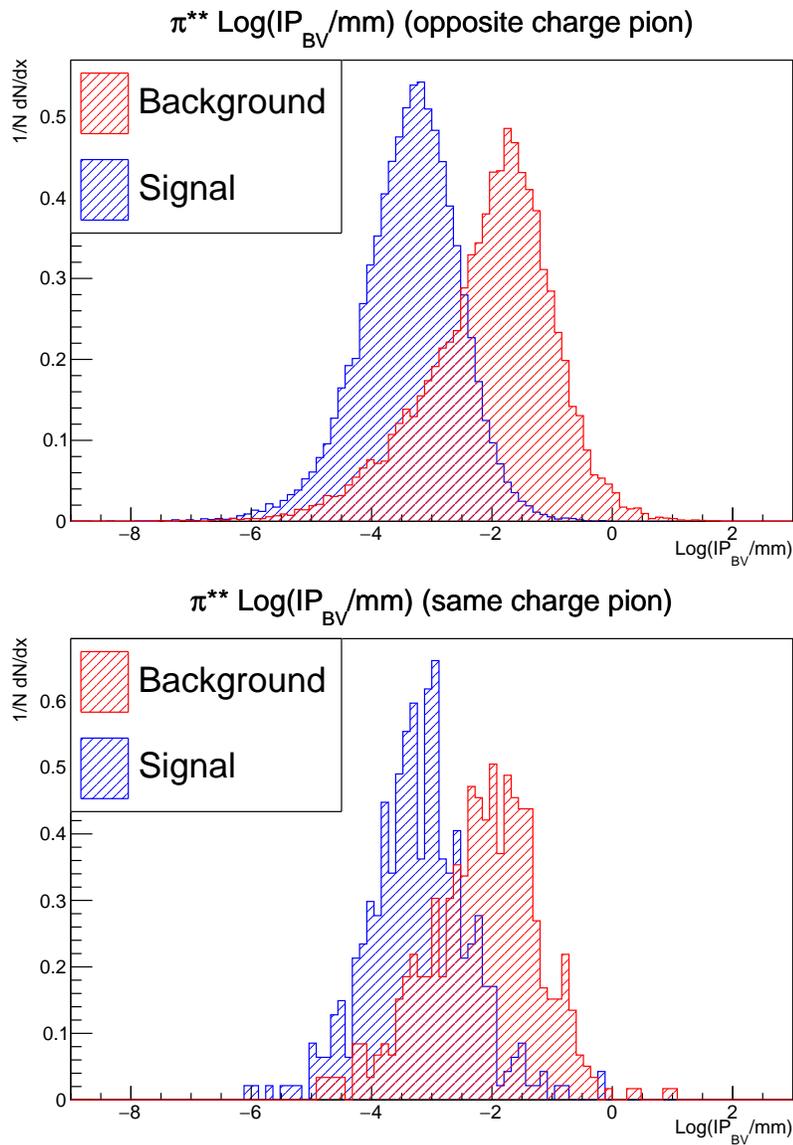


Figure 5.1. Impact parameter distribution of pions with opposite (top) and same (bottom) charge with respect to the D^* (Monte Carlo).

5.2 Unbinned maximum likelihood fit

All the fits performed are unbinned maximum likelihood fits. In this Section the extended likelihood fit method on unbinned dataset is presented along with the reasons for this choice. It is also explained how the goodness of a fit is evaluated. First of all, a fit model consists of a normalised probability density function $\mathcal{F}(\vec{x}, \vec{\theta})$ of the variables describing the data points \vec{x} and the unknown parameters $\vec{\theta}$, which determine the shape of the PDF. The aim of the likelihood fit is to determine the values of the parameters $\vec{\theta}$ so that the PDF describes as best as possible the data points. In order to do that, a likelihood function is built as:

$$\mathcal{L}(\vec{\theta}) = \prod_{i=1}^N \mathcal{F}(x_i, \vec{\theta}), \quad (5.1)$$

where x_i are statistically independent and each is described by $\mathcal{F}(x_i, \vec{\theta})$. The maximum likelihood method consists in finding the values of the parameters $\vec{\theta}$ which maximise the likelihood. Practically, it is easier to work with the logarithm of the likelihood

$$- \text{Log} \mathcal{L}(\vec{\theta}) = \sum_{i=1}^N \mathcal{F}(x_i, \vec{\theta}), \quad (5.2)$$

and find the values of the parameters that minimise it:

$$\frac{\partial \text{Log} \mathcal{L}}{\partial \theta_i} = 0, \quad i = 1, \dots, N. \quad (5.3)$$

This approach implicitly assumes that the total number of events expected is independent of the parameter values.

To add a parameter \mathcal{N} corresponding to the expected number of events N , a factor needs to be added in the likelihood expression which describes the probability of observing the actual number of events given this parameter. It is assumed that the expected yield \mathcal{N} follows a Poissonian distribution

$$P(N; \mathcal{N}) = \frac{1}{N!} e^{-\mathcal{N}} \mathcal{N}^N. \quad (5.4)$$

Including this term into the likelihood expression, Equation (5.1) becomes:

$$\mathcal{L}(\vec{x}, N; \vec{\theta}, \mathcal{N}) = \frac{1}{N!} e^{-\mathcal{N}} \mathcal{N}^N \prod_{i=1}^N \mathcal{F}(x_i, \vec{\theta}), \quad (5.5)$$

and it is called *extended likelihood*. An extended maximum likelihood fit is particularly useful when a data distribution is fitted with a composite PDF and the yields

of the different components have to be determined from the fit.

The maximum likelihood fit is applied to an unbinned dataset, i.e. the data are not binned into a histogram. This has the following advantages:

- Unbinned $\text{Log}\mathcal{L}$ fits are statistically more powerful than binned fits;
- Unbinned $\text{Log}\mathcal{L}$ fits avoid any arbitrariness introduced by the choice of a binning definition.

The goodness of a fit is estimated according to the following criteria:

- Pull distribution:
When they are plotted, the data are arranged into a binned histogram. So the pull P_i of the i -th bin is defined as

$$P_i = \frac{y_i^{data} - y_i^{PDF}}{\sigma_i^{data}}, \quad (5.6)$$

where y_i^{data} is the content of the i -th bin, y_i^{PDF} the PDF value at the i -th bin and σ_i^{data} is the statistical uncertainty on the i -th bin. If the PDF correctly reproduces the data, the pull distribution is a gaussian curve with zero mean and unit width.

- Error matrix:
The error or covariance matrix should be positive definite. A not positive definite matrix usually means that the parameters are strongly correlated, so they have to be constrained or fixed.
- EDM:
The EDM is the Estimated Distance from Minimum and should be of the order of 10^{-5} or smaller.

5.3 Fit on Monte Carlo simulated sample

The fit on the Monte Carlo sample has been performed with the purpose of defining the optimal PDF to describe the shape of the $\text{LogIP}_{\pi^{**}}^{BV}$ distribution. First of all, signal and background distributions are fitted separately for both pions, by using a composite pdf given by the sum of two asymmetric gaussians:

$$AG(x; \mu, \sigma_L, \sigma_H) = \frac{1}{\sqrt{2\pi}} \frac{1}{(\sigma_L + \sigma_H)/2} \cdot \begin{cases} e^{-\frac{(x-\mu)^2}{2\sigma_L^2}} & x < \mu \\ e^{-\frac{(x-\mu)^2}{2\sigma_H^2}} & x > \mu \end{cases}$$

where μ is the mean, σ_L and σ_H parametrise the widths on the left and right sides, respectively. The two distributions have the same average. The signal (background) probability density function has the following form:

$$\mathcal{S}(\mathcal{B}) = f_{\mathcal{S}(\mathcal{B})} \cdot AG(\mu_{\mathcal{S}(\mathcal{B})}, \sigma_{1L}^{\mathcal{S}(\mathcal{B})}, \sigma_{1H}^{\mathcal{S}(\mathcal{B})}) + (1 - f_{\mathcal{S}(\mathcal{B})})AG(\mu_{\mathcal{S}(\mathcal{B})}, \sigma_{2L}^{\mathcal{S}(\mathcal{B})}, \sigma_{2H}^{\mathcal{S}(\mathcal{B})}),$$

where f is the fraction of events in the first asymmetric gaussian. All the parameters are left free. Table 5.2 shows the values of the parameters fitted for signal and background distributions for both pions: the corresponding parameters in the two cases are compatible within the errors (statistical uncertainties given by the fit). The uncertainties are quite large in the case of pions with same charge with respect to the D^* , due to the lower statistics.

Table 5.1. Fitted parameters for the $LogIP_{\pi^{**}}^{BV}$ MC distribution after preselection and BDT cut.

Parameter	Value (Signal)	Value (Background)
μ	-3.2070 ± 0.018	-1.7181 (fixed)
σ_{1L}	0.690 ± 0.017	1.497 ± 0.059
σ_{1H}	0.710 ± 0.025	0.509 (fixed)
σ_{2L}	1.191 (fixed)	0.459 (fixed)
σ_{2H}	0.478 (fixed)	0.899 (fixed)
f	0.675 (fixed)	0.679 (fixed)
N	25100 ± 345	12776 ± 326

5.3.1 $D^*\pi\mu$ channel

An unbinned maximum likelihood fit of the logarithm of the impact parameter for the opposite charge pion has been performed, according to the following distribution:

$$\mathcal{F}(LogIP_{\pi^{**}}^{BV}) = N_{sgn}\mathcal{S} + N_{bkg}\mathcal{B}. \quad (5.7)$$

Most of the parameters are free: just four out of the six parameters describing the background are fixed at the value obtained from the individual fit. That was necessary because the fit to the background was otherwise unstable.

The values of the parameters extracted from the fit are collected in Table 5.1, while the fit is shown in Figure 5.2. The pull distribution is uniformly distributed around zero, the EDM is of the order of 10^{-4} and the error matrix accurate. Four fit iterations were needed to obtain a positive definite error matrix along with an acceptable value of the EDM.

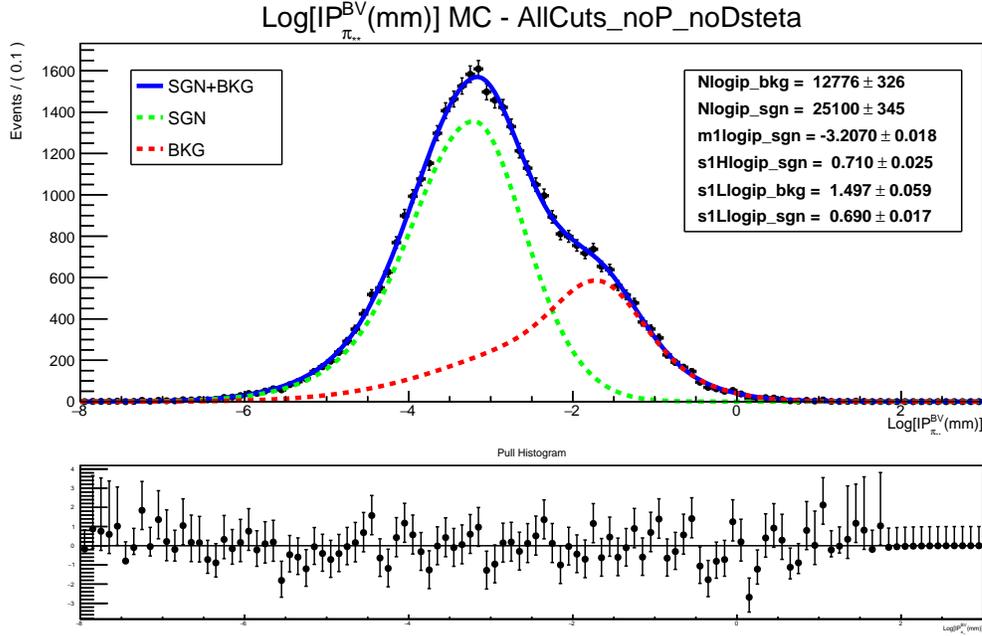


Figure 5.2. Fit of the $\text{LogIP}_{\pi^{**}}^{BV}$ MC distribution after preselection and BDT cut ($B \rightarrow D^* \pi \nu_{\mu} X$ channel).

5.3.2 $D^* \pi \pi \mu$ channel

In the case of two pions in the final state, a bidimensional fit of the $\text{LogIP}_{\pi^{**}}^{BV}$ distribution of the same and opposite charge pion has been performed. Since the two distributions are expected to have similar values for the parameters, as observed in Table 5.2, all parameters but the fractions are common to both PDFs. The same parameters that were fixed in the one dimensional fit of the opposite charge pion are kept fixed at the same values in the bidimensional fit. Three different contributions have been separated:

- Both pions are signal;
- The same charge pion is combinatorial and the opposite charge pion is signal;
- Both pions are combinatorial.

For each of these component, a product of the PDFs describing the $\pi_{1^{**}}$ and $\pi_{2^{**}}$ distributions is performed; the three components are then added, leading to the

Table 5.2. Fitted values of the parameters for signal and background $\text{LogIP}_{\pi^{**}}^{BV}$ distributions in simulated events. The two distributions are fitted separately. The values obtained for opposite charge and same charge pions are compared and found compatible within errors.

Parameter	Value $\pi_{1^{**}}$	Value $\pi_{2^{**}}$
Signal		
μ	-3.2277 ± 0.018	-3.208 ± 0.11
σ_{1L}	0.681 ± 0.023	0.53 ± 0.29
σ_{1H}	0.733 ± 0.011	0.708 ± 0.096
σ_{2L}	1.191 ± 0.032	0.98 ± 0.10
σ_{2H}	0.478 ± 0.032	0.26 ± 0.14
f	0.675 ± 0.041	0.52 ± 0.21
N_{sgn}	25330 ± 159	324 ± 18
Background		
μ	-1.7181 ± 0.035	-1.7489 ± 0.095
σ_{1L}	1.442 ± 0.021	0.659 ± 0.097
σ_{1H}	0.509 ± 0.031	0.552 ± 0.070
σ_{2L}	0.459 ± 0.044	1.78 ± 0.57
σ_{2H}	0.899 ± 0.017	0.86 ± 0.29
f	0.679 ± 0.021	0.913 ± 0.079
N_{bkg}	12546 ± 112	350 ± 19

following expression for the total probability density function:

$$\begin{aligned}
\mathcal{F}(\text{LogIP}_{\pi_{1^{**}}}^{BV}, \text{LogIP}_{\pi_{2^{**}}}^{BV}) = & \\
& N_{sgn} \cdot S(\text{LogIP}_{\pi_{1^{**}}}^{BV}) \cdot S(\text{LogIP}_{\pi_{2^{**}}}^{BV}) \\
& + N_{bkg1} \cdot S(\text{LogIP}_{\pi_{1^{**}}}^{BV}) \cdot B(\text{LogIP}_{\pi_{2^{**}}}^{BV}) \\
& + N_{bkg2} \cdot B(\text{LogIP}_{\pi_{1^{**}}}^{BV}) \cdot B(\text{LogIP}_{\pi_{2^{**}}}^{BV}). \quad (5.8)
\end{aligned}$$

The projections of the bidimensional fit are shown in Figure 5.3; the yellow curve gives the shape of the second component and that explains why in the $\pi_{1^{**}}$ case it is under the signal peak, while in the $\pi_{2^{**}}$ case under the background peak, although it must be said that this contribution is relatively low in both cases. According to Table 4.10, the contribution of the background given by a combinatorial $\pi_{1^{**}}$ and a signal $\pi_{2^{**}}$ has been considered negligible.

The values of the parameter extracted from the fit are gathered in Table 5.3. The pull distribution is good in both cases, the EDM is of the order of 10^{-7} and the error matrix positive definite. Two fit iterations were needed to obtain a good value for the EDM.

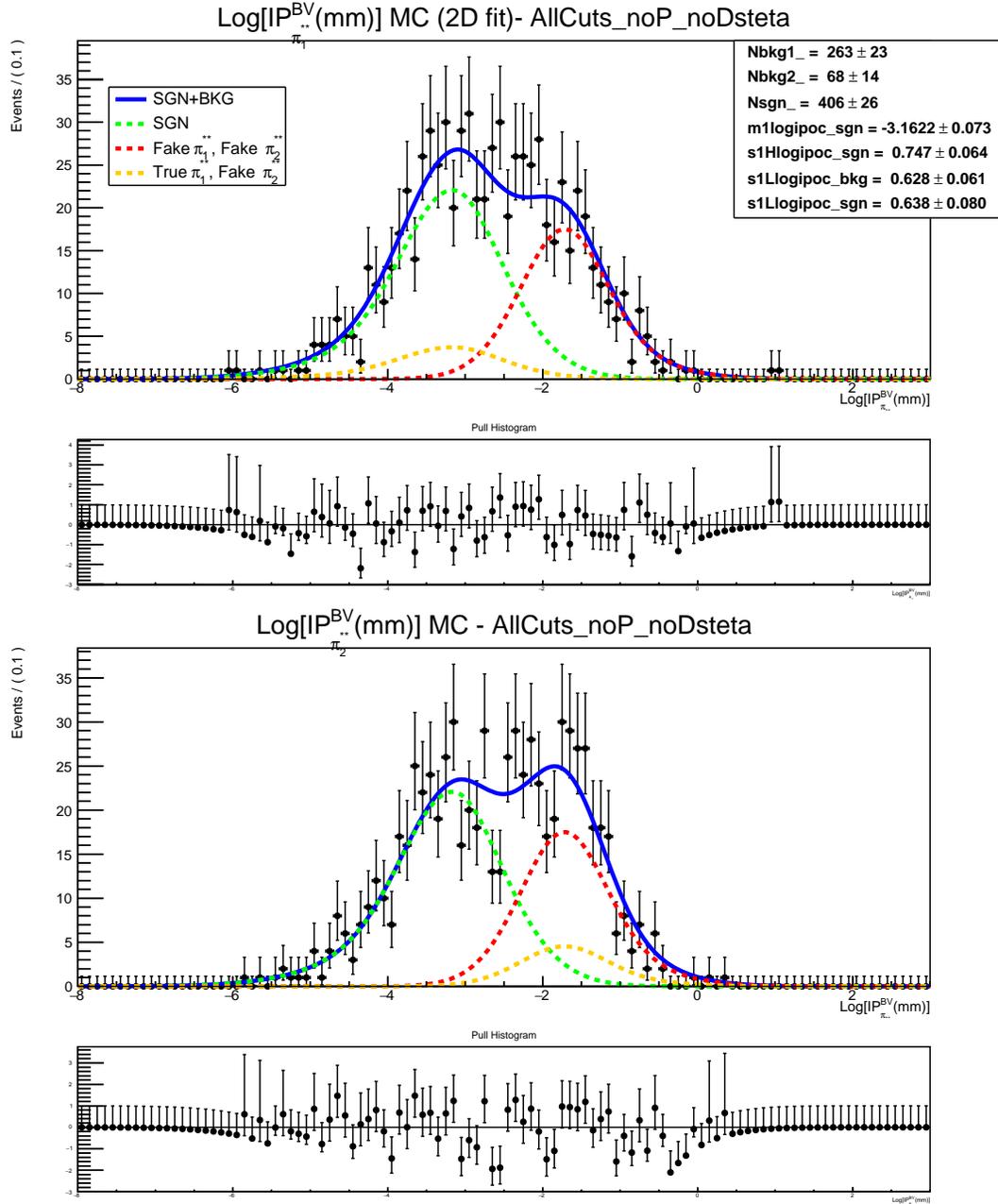


Figure 5.3. Projections of opposite charge (top) and same charge (bottom) pion of the bidimensional $\text{LogIP}_{\pi^{**}}^{\text{BV}}$ distributions.

5.3.3 Consistency check

To check the consistency of the fit, two methods were used.

- The values of the parameters obtained from the fit to the total sample are

Table 5.3. Fitted parameters for the bidimensional fit of $\text{LogIP}_{\pi^{**}}^{BV}$ MC distribution after preselection and BDT cut.

Parameter	Value (Signal)	Value (Background)
μ	-3.1622 ± 0.073	-1.7181 (fixed)
σ_{1L}	0.638 ± 0.080	0.628 ± 0.061
σ_{1H}	0.747 ± 0.064	0.509 (fixed)
σ_{2L}	1.191 (fixed)	0.459 (fixed)
σ_{2H}	0.478 (fixed)	0.899 (fixed)
f	0.675 (fixed)	0.679 (fixed)
N_{sgn}		406 ± 26
N_{bkg1}		263 ± 23
N_{bkg2}		68 ± 14

compared with the ones obtained from the individual fits of signal and background components. This was done for both channels.

- The Ana sample has been divided into two statistically independent halves. First the fit has been trained on one sample, and then tested on the other one, comparing again the values of the respective parameters and checking the goodness of the fit. That was applied only to the $B \rightarrow D^* \pi \mu \nu_\mu X$ channel because the statistic was not enough for the $B \rightarrow D^* \pi \pi \mu \nu_\mu X$ channel.

Both methods show a good agreement, confirming the validity of the fit.

5.4 Fit to the $m(K\pi\pi) - m(K\pi)$ distribution

An unbinned maximum likelihood fit of the $\Delta M = m(K\pi\pi) - m(K\pi)$ distribution, defined as the difference between the D^* and D^0 invariant masses, has been performed in order to obtain the signal yield for $B \rightarrow D^* \mu \nu_\mu X$ decay. The fit has been performed after the preselection and before any requirement on the additional pions.

The ΔM distribution is fitted in the range [140,155] MeV, which includes the entire signal peak, constraining the D^0 invariant mass in the interval $1835 \text{ MeV} < m(K\pi) < 1890 \text{ MeV}$. The latter is justified observing Figure 5.4, where the D^0 mass is plotted as function of the difference between D^* and D^0 invariant masses for the selected $B \rightarrow D^* \mu \nu_\mu$ candidates: the cut applied allows to remove the white triangular regions on the top-right and on the bottom-left of the two-dimensional plot. These regions are caused by previous selection criteria and, if not excluded, have an influence on the shape of the background distribution.

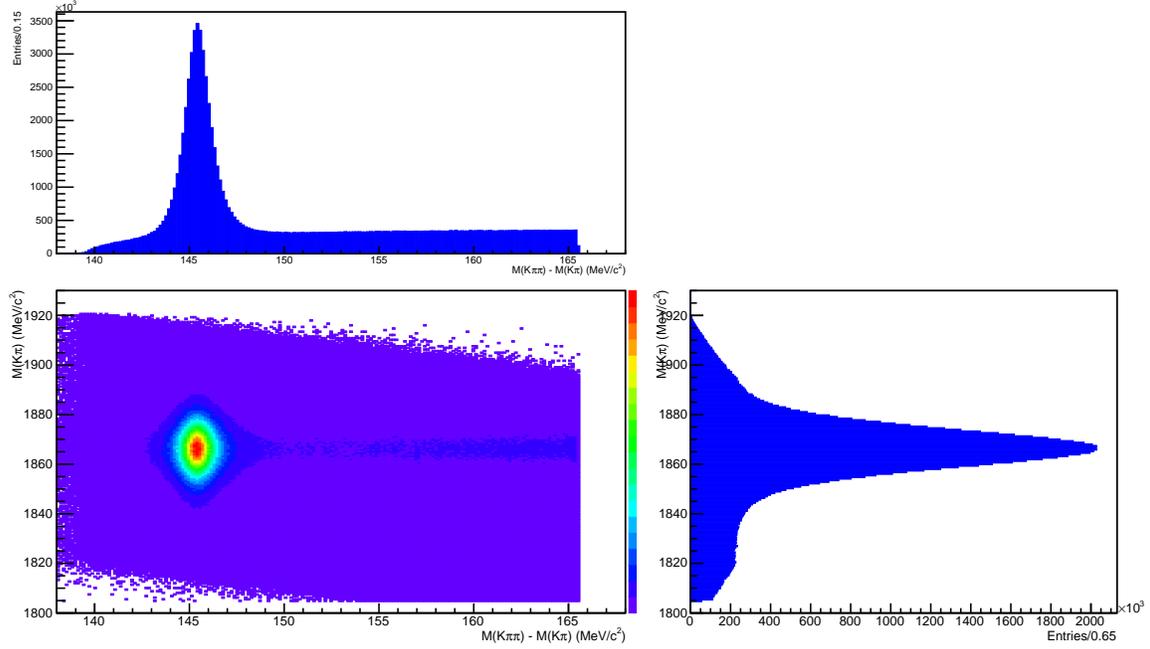


Figure 5.4. Two dimensional distribution of D^0 mass as function of ΔM .

5.4.1 Signal parametrisation

The probability density function used to fit the signal peak is given by the sum of two gaussians and a Johnson S_U distribution [65] with same mean and three different widths. The gaussian distribution has the form:

$$G_1(x; \mu, \sigma) = e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (5.9)$$

The Johnson S_U distribution has a gaussian-like shape, except for a Landau-like tail on one side:

$$J(x; \mu, \sigma, \nu, \tau) = \frac{\tau}{\sigma\sqrt{2\pi}\sqrt{1 + \left(\frac{x-\mu}{\sigma}\right)^2}} e^{-\frac{1}{2}[\nu + \tau \sinh^{-1}\left(\frac{x-\mu}{\sigma}\right)]^2}. \quad (5.10)$$

The parameters ν and τ determine the shape of the distribution: for $\tau > 0$ J is positive and normalised to one, while the sign of ν determines if the tail is located at low x (ν positive) or at high x (ν negative). Higher values of $|\nu|$ and $|\tau|$ result in a more symmetrical distribution and a sharper peak.

The signal component has the following form:

$$\mathcal{S}(\Delta M) = f_1 \cdot G(\mu, \sigma_1) + f_2 \cdot G(\mu, \sigma_2) + (1 - f_1 - f_2)J(\mu, \sigma_3, \nu, \tau), \quad (5.11)$$

where f_1 and f_2 are the fractions of events in the first and in the second gaussian, respectively.

5.4.2 Background parametrisation

The background is parametrised with a shape developed to model the background of the ΔM mass difference distributions:

$$\mathcal{B}(\Delta M) = 1 - e^{-\frac{(\Delta M - \Delta M_{th})}{c}}.$$

The parameter m_{th} is the kinematic threshold which should be circa the pion mass.

5.4.3 Total PDF

The total probability density function is given by:

$$\mathcal{F}(\Delta M) = N_{sgn} \cdot \mathcal{S}(\Delta M) + N_{bkg} \cdot \mathcal{B}(\Delta M), \quad (5.12)$$

where N_{sgn} and N_{bkg} are the signal and background yields respectively. All the signal PDF parameters are free, except for τ which is fixed at 1.0. The background parameters instead are all fixed but one (c): this choice was made because the background distribution was observed to be quite unstable.

The values of the fit parameters are gathered in Table 5.4. Figure 5.5 shows the fit results: the pull distribution is quite good (even if some point are up to 5 standard deviations away from 0, the average is still around 0), the error matrix is accurate and the EDM of the order of $7 \cdot 10^{-5}$. The fit has been iterated in order to obtain a positive definite error matrix and a good EDM.

5.5 Fit on data sample

Similarly to the Monte Carlo case, two kinds of fit have been performed on the data sample: a one dimensional fit of the $LogIP_{\pi^{**}}^{BV}$ distribution to extract the yield for the $B \rightarrow D^* \pi \mu \nu_\mu X$ channel and a bidimensional fit of the $LogIP$ of the opposite and same charge pions to extract the yield for the $B \rightarrow D^* \pi \pi \mu \nu_\mu X$ channel. Moreover, another background contribution has to be considered, due to residual combinatorial D^* and the decay of the B meson directly into $D^0 \pi \mu \nu_\mu X$ final states¹. To estimate the shape and yield of this background component,

¹This background component will be indicated as $B \rightarrow D^0$ background from now on.

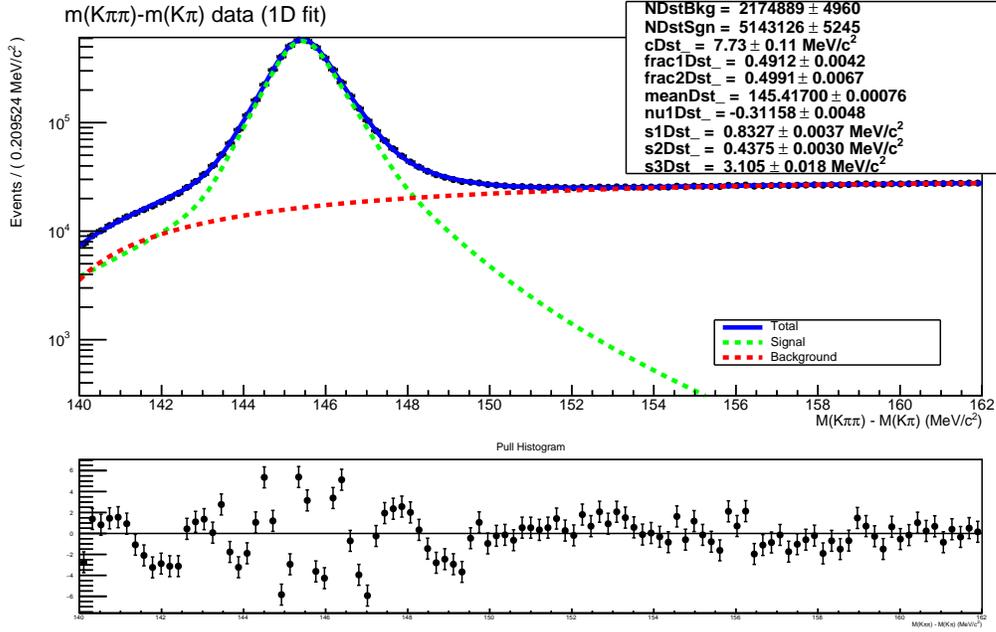


Figure 5.5. Unbinned maximum likelihood fit of ΔM distribution after preselection and before BDT cut (data).

Table 5.4. Parameter values extracted from ΔM fit on data after preselection.

Parameter	Value
μ (MeV/ c^2)	145.41700 ± 0.00076
σ_1 (MeV/ c^2)	0.8327 ± 0.0037
σ_2 (MeV/ c^2)	0.4375 ± 0.0030
σ_3 (MeV/ c^2)	3.105 ± 0.018
ν	-0.31158 ± 0.0048
τ	1.0 (fixed)
f_1	0.4912 ± 0.0042
f_2	0.4991 ± 0.0067
c (MeV/ c^2)	7.73 ± 0.11
m_{th} (MeV/ c^2)	139.0 (fixed)
N_{sgn}	5143126 ± 5245
N_{bkg}	2174889 ± 4960

a fit to the ΔM distribution has been performed on the data sample after the preselection and BDT cut, for both channels. Finally, the systematic uncertainties on the extracted yields have been evaluated.

5.5.1 Combinatorial D^* and $B \rightarrow D^0\pi\mu\nu_\mu$ background estimation

The ΔM distribution is fitted in the range [140,162] MeV, constraining the D^0 invariant mass in the interval $1835 \text{ MeV} < m(K\pi) < 1890 \text{ MeV}$. The PDF used is the same as in the ΔM fit before the BDT cut. The method used to evaluate the combinatorial D^* and $B \rightarrow D^0\pi\mu\nu_\mu$ background contribution is the following:

- An effective sigma σ_{eff} is computed as the weighted average between the standard deviations of the two gaussians and the width of the Johnson S_U function, the weights given by the fractions of events per each component:

$$\sigma_{eff} = f_1 \cdot \sigma_1 + f_2 \cdot \sigma_2 + (1 - f_1 - f_2) \cdot \sigma_3. \quad (5.13)$$

- A signal and a sideband region are defined: the signal region elapses from $-3\sigma_{eff}$ to $+3\sigma_{eff}$ from the peak, whereas the sideband region elapses from $+7\sigma_{eff}$ to $+17\sigma_{eff}$:

$$\begin{aligned} SGN &= [142, 148] \text{ (MeV)}, \\ SDB &= [152, 162] \text{ (MeV)}. \end{aligned}$$

- The shape of the $LogIP_{\pi^{**}}^{BV}$ distribution is modelled requiring ΔM to lie in the sideband region.
- A scaling factor is computed as:

$$s = \frac{\int_{SGN} d(\Delta M') \mathcal{B}(\Delta M')}{\int_{SDB} d(\Delta M') \mathcal{B}(\Delta M')}, \quad (5.14)$$

where the numerator and the denominator are the integral of the background distribution in the ΔM signal region and in the sideband region, respectively. The scaling factors obtained are:

$$\begin{aligned} s_\pi &= 0.498385, \\ s_{\pi\pi} &= 0.487108, \end{aligned}$$

where s_π is for the $B \rightarrow D^*\pi\mu\nu_\mu X$ channel and $s_{\pi\pi}$ for the $B \rightarrow D^*\pi\pi\mu\nu_\mu X$ channel.

- Assuming the background has the same shape in the sideband and signal region, the $LogIP_{\pi^{**}}^{BV}$ distribution is normalised in the signal region reweighting the events according to the scaling factor. This distribution represents the contribution of combinatorial D^* and $B \rightarrow D^0\pi\mu\nu_\mu X$ backgrounds inside the signal region.

- Being a distribution in $\text{LogIP}_{\pi^{**}}^{BV}$, the $B \rightarrow D^0$ background distribution has been fitted with the same PDF used on the MC sample (2 asymmetric gaussians with same mean for the peak under the $\text{LogIP}_{\pi^{**}}^{BV}$ signal curve and 2 asymmetric gaussians with same mean under the $\text{LogIP}_{\pi^{**}}^{BV}$ background curve).

First, the ΔM fit, shown in Figure 5.6, has been performed for the $B \rightarrow D^* \pi \mu \nu_\mu X$ channel and the respective distribution in $\text{LogIP}_{\pi^{**}}^{BV}$ for the $B \rightarrow D^0$ background evaluated and fitted with all the parameters free (Figure 5.7). Then, the ΔM distribution has been fitted for the $B \rightarrow D^* \pi \pi \mu \nu_\mu X$ channel (Figure 5.8), fixing all the parameters to the respective values from the previous fit. The $\text{LogIP}_{\pi^{**}}^{BV}$ for the $B \rightarrow D^0$ background for both pions is then fitted leaving all the parameters free. For both channels the plots show good pull distributions, the error matrices are accurate and the EDM is of the order of 10^{-4} . The fits have been iterated in order to obtain a positive definite error matrix and a low EDM value.

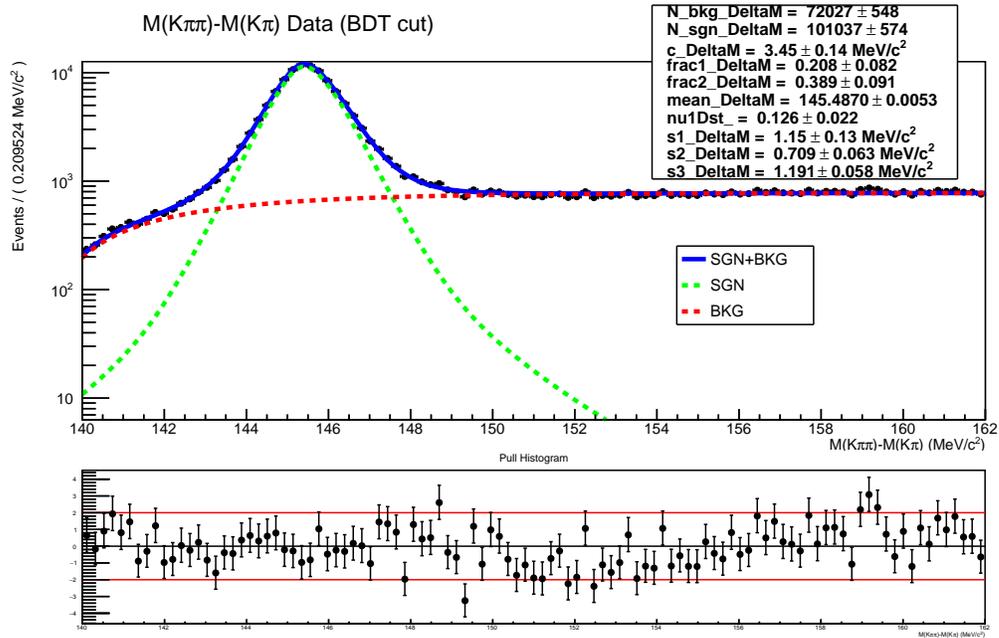


Figure 5.6. $D^* \pi \mu \nu_\mu X$ channel fit on data of ΔM distribution after preselection and BDT cut.

5.5.2 $D^* \pi$ channel

An unbinned maximum likelihood fit of the logarithm of the impact parameter for the opposite charge pion has been performed on data, in the ΔM signal region.

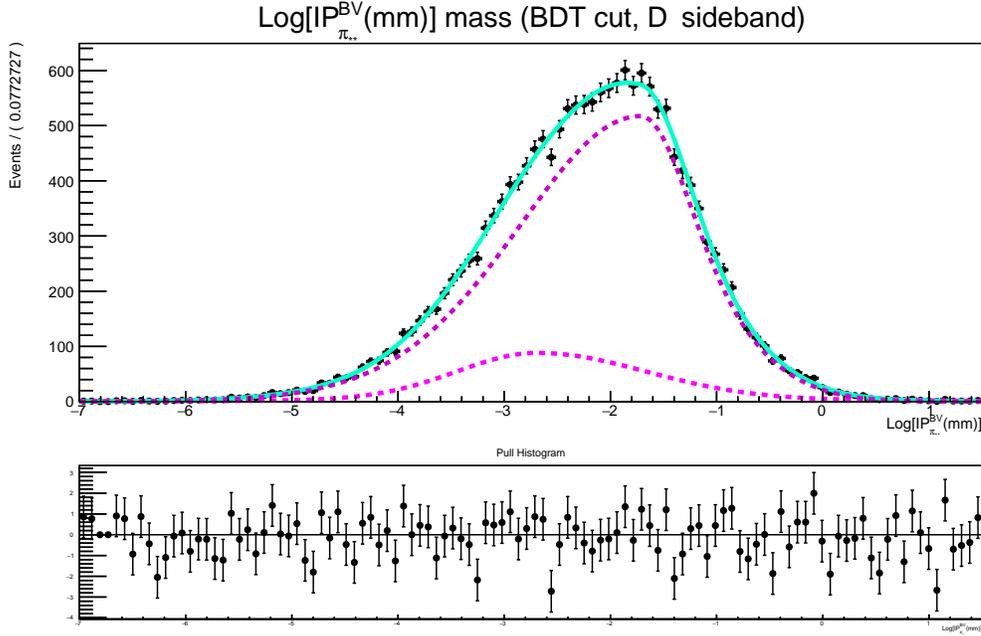


Figure 5.7. $D^*\pi\mu\nu_\mu X$ channel fit on data: reweighted $B \rightarrow D^0$ background distribution in the ΔM signal region.

Table 5.5. Parameter values extracted from ΔM fit on data after preselection and BDT cut in the $B \rightarrow D^*\pi\mu\nu_\mu X$ and $B \rightarrow D^*\pi\pi\mu\nu_\mu X$ channels.

Parameter	Value	Value
	$B \rightarrow D^*\pi\mu\nu_\mu X$	$B \rightarrow D^*\pi\pi\mu\nu_\mu X$
μ (MeV/ c^2)	145.4870 ± 0.0053	145.4870 (fixed)
σ_1 (MeV/ c^2)	1.15 ± 0.13	1.15 (fixed)
σ_2 (MeV/ c^2)	0.709 ± 0.063	0.709 (fixed)
σ_3 (MeV/ c^2)	1.191 ± 0.058	1.191 (fixed)
ν	0.126 ± 0.022	0.126 (fixed)
τ	1.0 (fixed)	1.0 (fixed)
f_1	0.208 ± 0.082	0.208 (fixed)
f_2	0.389 ± 0.091	0.389 (fixed)
c (MeV/ c^2)	3.45 ± 0.14	3.75 ± 0.21
m_{th} (MeV/ c^2)	139 (fixed)	139 (fixed)
N_{sgn}	101037 ± 574	7518 ± 106
N_{bkg}	72027 ± 548	12214 ± 126

Two cuts are required on the D^0 invariant mass and the $D^*\pi$ invariant mass: $1835 \text{ MeV} < m(K\pi) < 1890 \text{ MeV}$ and $2150 \text{ MeV} < m(D^*\pi) < 3200 \text{ MeV}$. The PDF used to fit the signal and background distributions are the same as on the MC data sample, fixing the same four parameters of the background distribution to the

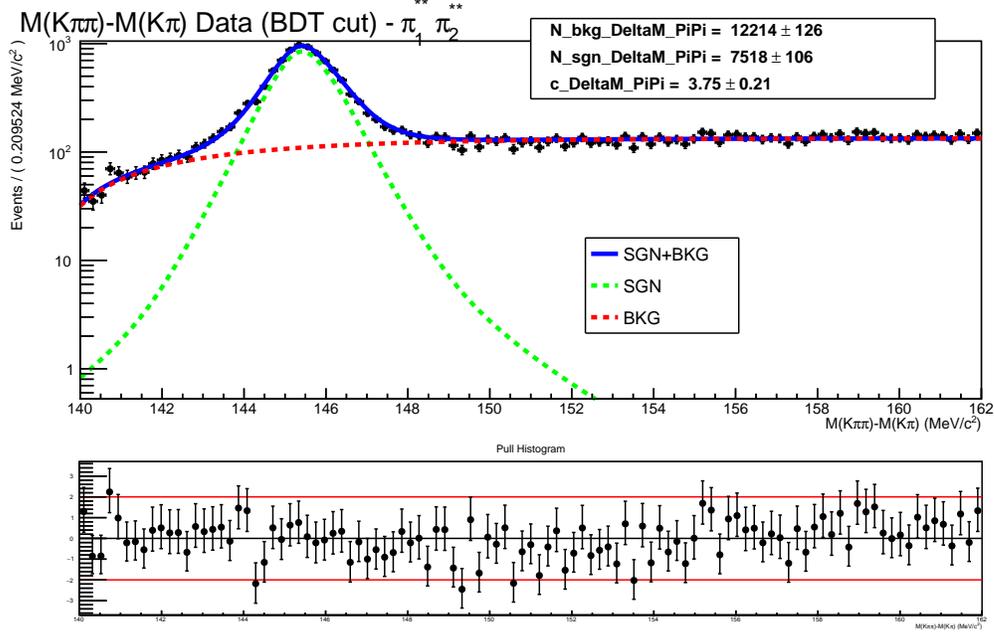


Figure 5.8. $D^* \pi\pi\mu\nu X$ channel fits on data. Fit of ΔM distribution after preselection and BDT cut.

values extracted from the individual fit on the Monte Carlo data. The $B \rightarrow D^0$ contribution is fitted fixing all of the parameters to the respective values extracted from the previous individual fit. The total PDF has the following form:

$$\mathcal{F}(\text{LogIP}_{\pi^{**}}) = N_{D^0 bkg} \cdot \mathcal{B}_{D^0}(\text{LogIP}_{\pi^{**}}^{BV}) + N_{bkg} \cdot \mathcal{B}(\text{LogIP}_{\pi^{**}}^{BV}) + N_{sgn} \cdot \mathcal{S}(\text{LogIP}_{\pi^{**}}^{BV}), \quad (5.15)$$

where $N_{D^0 bkg}$ is the number of $B \rightarrow D^0$ events, fixed to the yield obtained from the integration of the $B \rightarrow D^0$ distribution in the ΔM signal region; N_{sgn} and N_{bkg} are the signal and combinatorial background yields, respectively.

The fit is shown in Figure 5.10 and the values of the parameters are gathered in Table 5.6. The pull distribution is good, the error matrix is accurate and the EDM is of the order of 10^{-5} .

5.5.3 $D^* \pi\pi$ channel

A bidimensional fit of $\text{LogIP}_{\pi_1^{**}}^{BV}$ and $\text{LogIP}_{\pi_2^{**}}^{BV}$ has been performed on data as it has been done on the MC sample. The only difference is that the contribution of

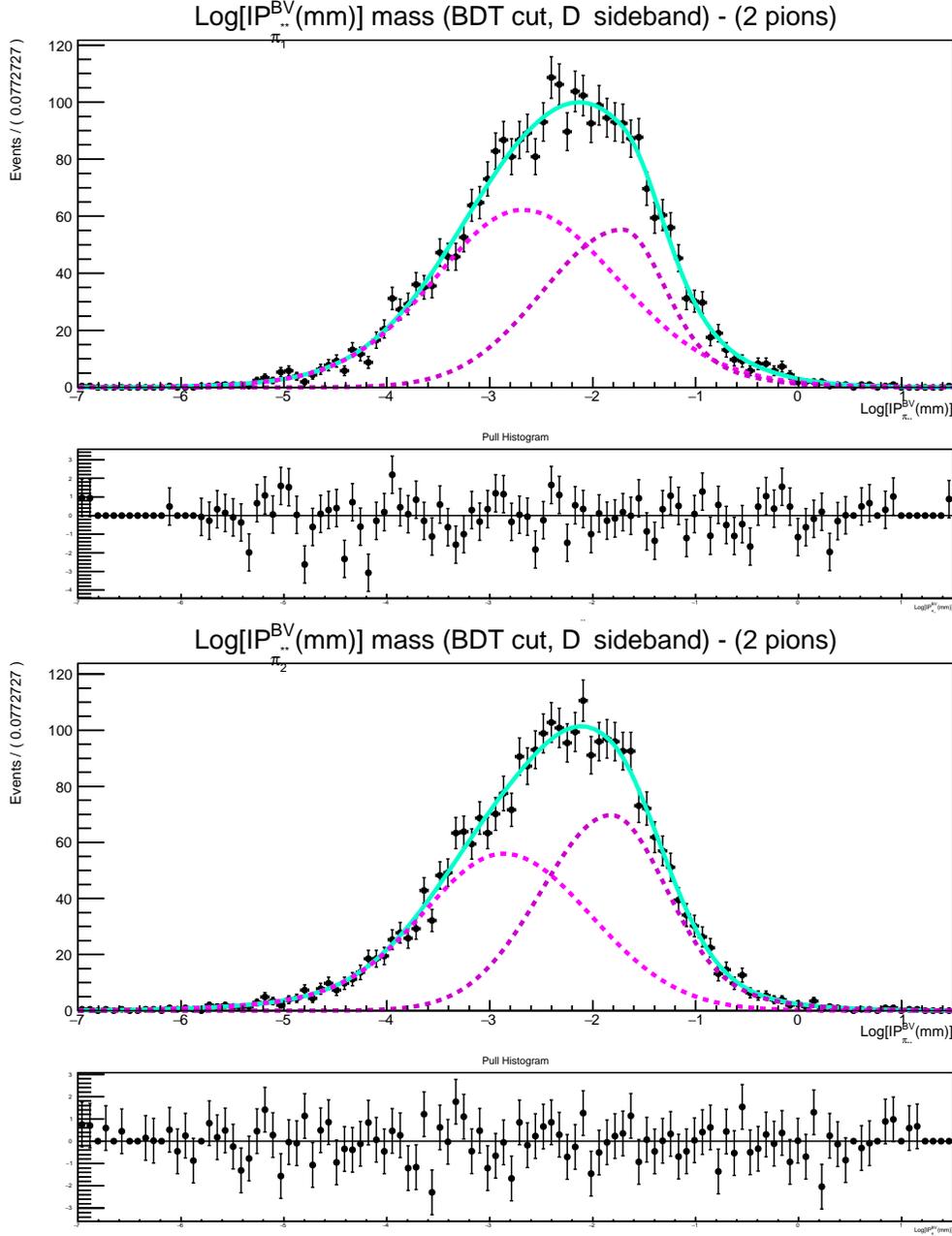


Figure 5.9. $D^*\pi\pi\mu\nu\mu X$ channel fits on data. Reweighted $B \rightarrow D^0$ background distributions in the ΔM signal region of $\pi_{1^{**}}$ and $\pi_{2^{**}}$.

$B \rightarrow D^0$ background is added, leading to the following expression for the PDF:

$$\begin{aligned}
 \mathcal{F}(\text{LogIP}_{\pi_{1^{**}}}^{BV}, \text{LogIP}_{\pi_{2^{**}}}^{BV}) = & \\
 & N_{sgn} \cdot \mathcal{S}(\text{LogIP}_{\pi_{1^{**}}}^{BV}) \cdot \mathcal{S}(\text{LogIP}_{\pi_{2^{**}}}^{BV}) \\
 & + N_{bkg1} \cdot \mathcal{S}(\text{LogIP}_{\pi_{1^{**}}}^{BV}) \cdot \mathcal{B}(\text{LogIP}_{\pi_{2^{**}}}^{BV}) \\
 & + N_{bkg2} \cdot \mathcal{B}(\text{LogIP}_{\pi_{1^{**}}}^{BV}) \cdot \mathcal{B}(\text{LogIP}_{\pi_{2^{**}}}^{BV}) \\
 & + N_{D^0bkg} \cdot \mathcal{B}_{D^0}(\text{LogIP}_{\pi_{1^{**}}}^{BV}) \cdot \mathcal{B}_{D^0}(\text{LogIP}_{\pi_{2^{**}}}^{BV}). \quad (5.16)
 \end{aligned}$$

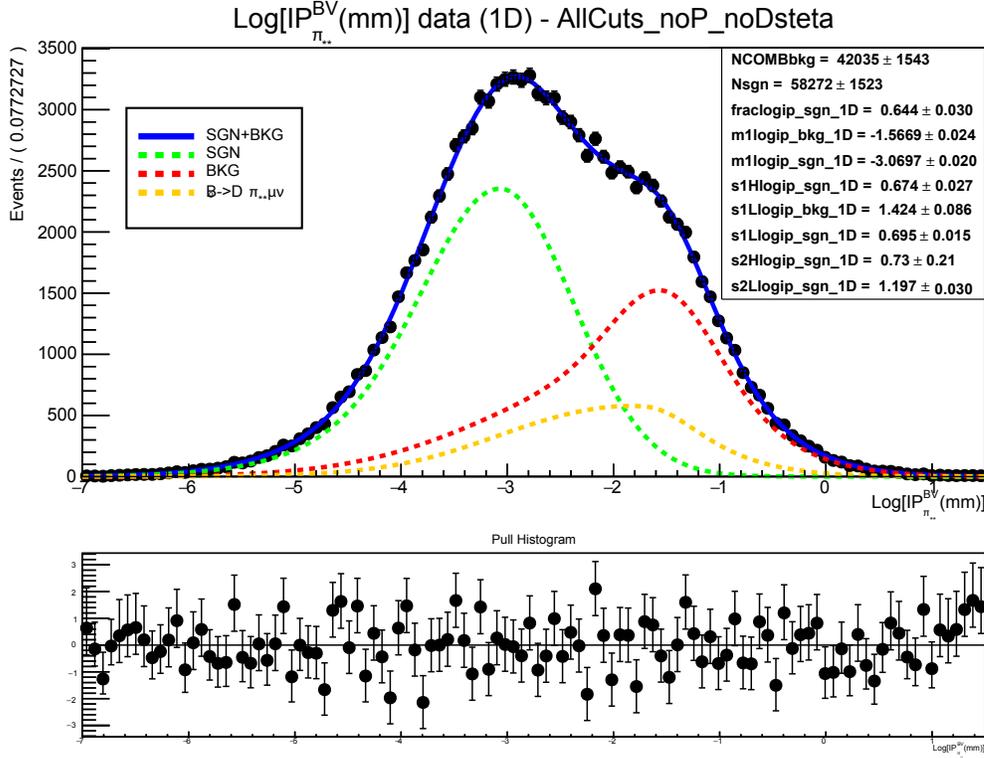


Figure 5.10. Unidimensional fit of $\text{LogIP}_{\pi_{**}}^{BV}$ data distribution for the $D^* \pi \mu \nu_\mu X$ decay channel.

Table 5.6. Fitted parameters for the $\text{LogIP}_{\pi_{**}}^{BV}$ data distribution after preselection and BDT cut.

Parameter	Value (Signal)	Value (Background)
μ	-3.0697 ± 0.020	-1.5669 ± 0.024
σ_{1L}	0.695 ± 0.015	1.424 ± 0.086
σ_{1H}	0.674 ± 0.027	0.509 (fixed)
σ_{2L}	1.197 ± 0.030	0.459 (fixed)
σ_{2H}	0.73 ± 0.21	0.899 (fixed)
f	0.644 ± 0.030	0.679 (fixed)
N	58272 ± 1523	42035 ± 1543

The values of the parameters extracted from the fit are gathered in Table 5.7. Despite some points are two standard deviations away from zero, the pull distributions can be considered good in both cases. The EDM is of the order of 10^{-5} and the error matrices positive definite. Two fit iterations were needed to obtain a good value for the EDM.

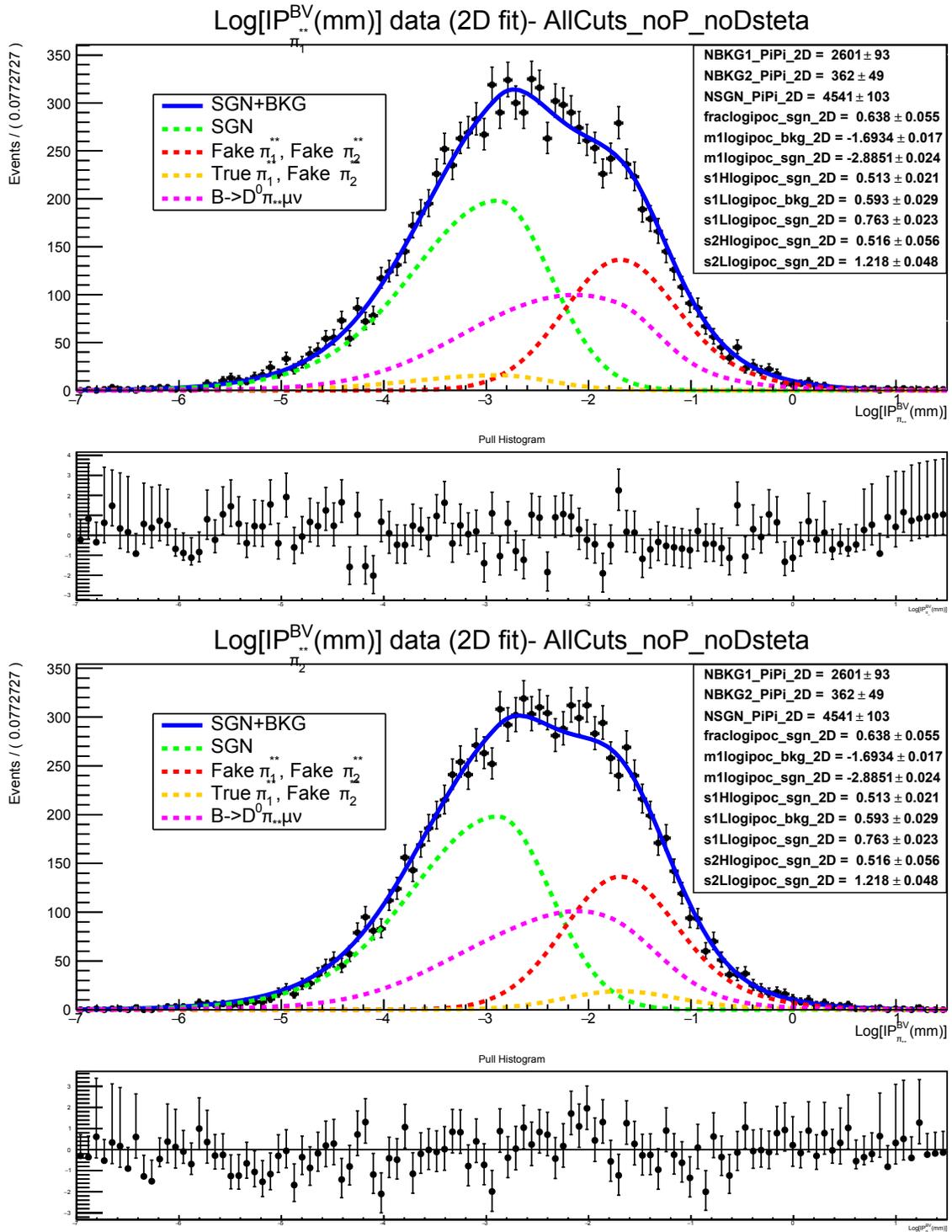


Figure 5.11. $D^* \pi \pi \nu_{\mu} X$ channel: projections of the bidimensional fit on data of $\text{Log} IP_{\pi^{**}}^{BV}$ distributions for opposite (top) and same charge (bottom) pions.

Table 5.7. Fitted parameters for the bidimensional fit of $\text{LogIP}_{\pi^{**}}^{BV}$ data distribution after preselection and BDT cut.

Parameter	Value (Signal)	Value (Background)
μ	-2.8851 ± 0.024	-1.6934 ± 0.017
σ_{1L}	0.763 ± 0.023	0.593 ± 0.029
σ_{1H}	0.513 ± 0.021	0.509 (fixed)
σ_{2L}	1.218 ± 0.048	0.459 (fixed)
σ_{2H}	0.516 ± 0.056	0.899 (fixed)
f	0.638 ± 0.055	0.679 (fixed)
N_{sgn}		4541 ± 103
N_{bkg1}		2601 ± 93
N_{bkg2}		362 ± 49

5.6 Systematic uncertainties on yields

The uncertainties on the yields considered so far are statistical only. This Section describes the systematic uncertainty sources that were considered and how they are evaluated.

- Combinatorial background: parameters.

A contribution to the systematic uncertainty is given by the background distribution chosen. Four parameters were fixed at their MC value during the fit. To estimate the corresponding systematic uncertainty, every parameter has been fixed to its MC value plus or minus its statistical uncertainty. Then the fit procedure has been repeated for each case, extracting the respective yield. At the end, 8 fits have been performed. For each parameter the geometrical average of the difference between the values obtained and the yields in Tables 5.6 and 5.7 is computed. Then the total systematic uncertainty is obtained adding them in quadrature.

- $B \rightarrow D^0$ background: shape.

The shape of the background has been modified. Instead of a fit with the parametrisation described above, a RooKeysPdf [66] is used. A RooKeysPdf is a one-dimensional kernel estimation PDF which model the input distribution as a superposition of Gaussian kernels, one for each data point, each contributing $1/N$ (where N is the number of data points) to the total integral of the PDF [66]. The systematic uncertainty is computed as the difference between the value obtained and the yields in Tables 5.6 and 5.7;

- $B \rightarrow D^0$ background: normalisation.

The $B \rightarrow D^0$ background yield is fixed to the integral value of the events enclosed by the curve. To evaluate the systematic uncertainty the fit has been repeated substituting to $N_{D^0 bkg}$ its value plus/minus its statistical uncertainty. The geometrical average of the differences between the two values obtained and the one in Tables 5.6 and 5.7 is computed.

- Two pions distribution.

All the parameters are fixed to the fitted values of the $LogIP_{\pi^{**}}^{BV}$ distribution in the $B \rightarrow D^* \pi \mu \nu_\mu X$ channel. The systematic uncertainty is computed as the difference between the value obtained and the yields in Tables 5.6 and 5.7;

Table 5.8 shows the individual contributions to the systematics: for the $B \rightarrow D^* \pi \mu \nu_\mu X$ channel, the dominant contribution is due to the parameters of the combinatorial background, confirming the instability of the shape describing this type of background; for the $B \rightarrow D^* \pi \pi \mu \nu_\mu X$ channel, the dominant systematic uncertainty is the one due to the $B \rightarrow D^0$ background normalisation.

Table 5.8. Systematic uncertainties on yields.

Type	Uncertainty	
	$B \rightarrow D^* \pi \mu \nu_\mu X$	$B \rightarrow D^* \pi \pi \mu \nu_\mu X$
CombBKG parameters	2155	11
D^0 from B shape	1291	2
D^0 from B normalisation	403	20
two pions distribution	-	199

The total systematic uncertainty is computed as the sum in quadrature of the different contributions. The results obtained for the yields are collected in Table 5.9.

Table 5.9. Yields of the two semi-inclusive decays of the B meson into $D^* \pi(\pi) \mu \nu_\mu$ final state.

Channel	Yield	Stat. error	Syst. error
$D^* \pi \mu \nu_\mu$	58272	1523	2545
$D^* \pi \pi \mu \nu_\mu$	4541	103	200

5.7 Branching ratios

The aim of the analysis is the computation of the semi-inclusive branching ratios for the $B \rightarrow D^* \pi \mu \nu_\mu X$ and $B \rightarrow D^* \pi \pi \mu \nu_\mu X$ channels, i.e. semileptonic B decays into one or two pions along with a D^* and a muon in the final state. The two branching ratios are calculated relatively to the normalisation channel $B \rightarrow D^* \mu \nu_\mu X$, which is the total semi-inclusive branching ratio for semileptonic B decays with a D^* in the final state:

$$\frac{\mathcal{B}(B \rightarrow D^* \pi(\pi) \mu \nu_\mu X)}{\mathcal{B}(B \rightarrow D^* \mu \nu_\mu X)}. \quad (5.17)$$

The number of reconstructed final states with $D^* \pi(\pi)$ on the data sample is given by:

$$\begin{aligned} N_{data}^{D^* \pi(\pi)} &= \mathcal{L} \cdot \sigma(pp \rightarrow b\bar{b}) \cdot Prob(b \rightarrow B) \\ &\quad \cdot \mathcal{B}(B \rightarrow D^* \pi(\pi) \mu \nu_\mu X) \cdot \mathcal{B}(D^* \rightarrow D^0 \pi_*) \\ &\quad \cdot \mathcal{B}(D^0 \rightarrow \pi K) \cdot \epsilon_1^{\pi(\pi)} \cdot \epsilon_2^{\pi(\pi)}. \end{aligned} \quad (5.18)$$

Whereas the number of D^* reconstructed in the data sample is:

$$\begin{aligned} N_{data}^{D^*} &= \mathcal{L} \cdot \sigma(pp \rightarrow b\bar{b}) \cdot Prob(b \rightarrow B) \\ &\quad \cdot \mathcal{B}(B \rightarrow D^* \mu \nu_\mu X) \cdot \mathcal{B}(D^* \rightarrow D^0 \pi_*) \\ &\quad \cdot \mathcal{B}(D^0 \rightarrow \pi K) \cdot \epsilon_1^*. \end{aligned} \quad (5.19)$$

\mathcal{L} is the total integrated luminosity, $\sigma(pp \rightarrow b\bar{b})$ is the $b\bar{b}$ production cross section and $Prob(b \rightarrow B)$ the probability that b or \bar{b} quarks hadronise into a B meson. ϵ_1^* , $\epsilon_1^{\pi(\pi)}$ and $\epsilon_2^{\pi(\pi)}$ are the selection efficiencies computed from the Monte Carlo sample:

- $\epsilon_1^* = \frac{N_{1,MC}^{D^*}}{N_{GEN}^{D^*}}$ is the fraction of D^* passing the preselection criteria. $N_{1,MC}$ is the number of D^* in the MC sample passing the first selection stage and $N_{GEN}^{D^*}$ the number of D^* generated in the MC sample;
- $\epsilon_1^{\pi(\pi)} = \frac{N_{1,MC}^{D^* \pi(\pi)}}{N_{GEN}^{D^* \pi(\pi)}}$ is the fraction of $D^* \pi(\pi)$ events after the preselection. $N_{1,MC}$ is the number of $D^* \pi(\pi)$ in the MC sample passing the first selection stage and $N_{GEN}^{D^* \pi(\pi)}$ the number of $D^* \pi(\pi)$ generated in the MC sample;
- $\epsilon_2^{\pi(\pi)} = \frac{N_{2,MC}^{D^* \pi(\pi)}}{N_{1,MC}^{D^* \pi(\pi)}}$ is the fraction of $D^* \pi(\pi)$ events after the second selection stage. $N_{2,MC}$ is the number of $D^* \pi(\pi)$ in the MC sample passing the second

selection stage and $N_{1,MC}$ the number of $D^*\pi(\pi)$ in the MC sample passing the first selection stage.

Dividing (5.18) by (5.19) and substituting the expressions for the efficiencies, the following relative branching ratio is obtained:

$$\frac{\mathcal{B}(B \rightarrow D^*\pi(\pi)\mu\nu_\mu X)}{\mathcal{B}(B \rightarrow D^*\mu\nu_\mu X)} = \frac{N_{data}^{D^*\pi(\pi)}}{N_{data}^{D^*}} \cdot \frac{N_{1,MC}^{D^*}}{N_{2,MC}^{D^*\pi(\pi)}} \cdot \frac{N_{GEN}^{D^*\pi(\pi)}}{N_{GEN}^{D^*}}. \quad (5.20)$$

The expression (5.20) shows that computing the relative branching ratios instead of the absolute ones, allows to avoid the systematics related to the luminosity, the production cross section and the hadronization probability.

At this point, all the terms appearing in (5.20) have been computed:

- $N_{data}^{D^*}$ is the signal yield extracted from the $m(K\pi\pi) - m(K\pi)$ fit performed after the D^* selection but before the BDT cut (Table 5.4).
- $N_{data}^{D^*\pi(\pi)}$ is the signal yield extracted from the *LogIP* unidimensional (bidimensional) fit performed after the 2 selection stages (Tables 5.6 and 5.7).
- $N_{1,MC}^{D^*} = 617179 \pm 785$ is the number of D^* events in the MC Ana sample after the first selection but before the BDT cut.
- $N_{2,MC}^{D^*\pi(\pi)}$ is the $D^*\pi(\pi)$ yield obtained from the unidimensional (bidimensional) fit on the MC sample after preselection and BDT cut (Tables 5.1 and 5.3).
- the ratio $\frac{N_{GEN}^{D^*\pi(\pi)}}{N_{GEN}^{D^*}}$ is computed from the branching ratios used to generate the MC sample as:

$$\frac{N_{GEN}^{D^*\pi(\pi)}}{N_{GEN}^{D^*}} = \frac{\mathcal{B}_{MC}(B \rightarrow D^*\pi(\pi)\mu\nu_\mu X)}{\mathcal{B}_{MC}(B \rightarrow D^*\mu\nu_\mu X)}. \quad (5.21)$$

That gives:

$$\frac{N_{GEN}^{D^*\pi}}{N_{GEN}^{D^*}} = 0.172,$$

$$\frac{N_{GEN}^{D^*\pi\pi}}{N_{GEN}^{D^*}} = 0.049.$$

Combining all the correct numbers, the relative branching ratios for the two investigated channels are:

$$\begin{aligned}\frac{\mathcal{B}(B \rightarrow D^* \pi \mu \nu_\mu X)}{\mathcal{B}(B \rightarrow D^* \mu \nu_\mu X)} &= 0.048 \pm 0.001_{stat} \pm 0.001_{MCstat}, \\ \frac{\mathcal{B}(B \rightarrow D^* \pi \pi \mu \nu_\mu X)}{\mathcal{B}(B \rightarrow D^* \mu \nu_\mu X)} &= 0.066 \pm 0.001_{stat} \pm 0.003_{MCstat},\end{aligned}\quad (5.22)$$

where only statistical uncertainties are indicated.

The absolute branching ratios are obtained multiplying the relative branching fractions by the PDG value $\mathcal{B}(B \rightarrow D^* \mu \nu_\mu X) = 0.0495 \pm 0.0011$:

$$\mathcal{B}(B \rightarrow D^* \pi \mu \nu_\mu X) = (0.235 \pm 0.001_{stat} \pm 0.002_{MCstat} \pm 0.005_{norm})\%,$$

$$\mathcal{B}(B \rightarrow D^* \pi \pi \mu \nu_\mu X) = (0.326 \pm 0.001_{stat} \pm 0.016_{MCstat} \pm 0.006_{norm})\%,$$

where the first uncertainty is statistical, the second one due to the MC statistic and the third one comes from the branching fraction for the normalisation mode. In the next section the systematic uncertainties are explained and evaluated.

5.7.1 Efficiencies

The total selection efficiency factor E , given in (5.20) by the inverse of the product of the last 2 fractions, is computed from the MC sample as the product of three terms:

$$E = \frac{\epsilon_1^{\pi(\pi)} \cdot \epsilon_2^{\pi(\pi)}}{\epsilon_1^*}, \quad (5.23)$$

so that Equation 5.20 can be rewritten as

$$\frac{\mathcal{B}(B \rightarrow D^* \pi(\pi) \mu \nu_\mu X)}{\mathcal{B}(B \rightarrow D^* \mu \nu_\mu X)} = \frac{N_{data}^{D^* \pi(\pi)}}{N_{data}^{D^*}} \cdot \frac{1}{E}. \quad (5.24)$$

It has been observed that $\epsilon_2^{\pi(\pi)}$, the efficiency related to the pion selection, is not the same for all the resonant states. Table 5.10 collects the efficiencies $\epsilon_2^{\pi(\pi)}$ for the $D^* \pi$ channel, showing that the selection efficiency for the D_0^* resonance is much lower with respect to the other efficiencies.

5.8 Systematic uncertainties on branching ratios

The sources of systematic uncertainties on the branching ratios are the following:

Table 5.10. Relative efficiencies $\epsilon_2^{\pi(\pi)}$ before and after the BDT cut for resonant and non-resonant decays evaluated on MC for the $B \rightarrow D^* \pi \mu \nu X$ channel.

Resonance	MC events before BDT cut	MC events after BDT cut	Efficiency
D_1	16456 ± 128	11017 ± 105	0.6694 ± 0.0123
D_1'	6654 ± 82	4370 ± 66	0.6567 ± 0.0195
D_0^*	3139 ± 56	914 ± 30	0.2912 ± 0.0373
D_2^*	5432 ± 74	3518 ± 59	0.6476 ± 0.0216
non-resonant	9440 ± 97	6739 ± 82	0.7138 ± 0.0159

- Combinatorial background: parameters.
For each yield obtained varying one of the fixed parameters of the distribution at a time (as explained in Section 4.6), a branching ratio is computed. For each parameter the geometrical average of the difference between the values obtained and the branching ratios in Equation 6.1 is computed. Then the total systematic uncertainty is obtained adding them in quadrature.
- $B \rightarrow D^0$ background: shape.
The branching ratio is computed from the yield obtained parametrising the $B \rightarrow D^0$ background with a RooKeys pdf. The systematic uncertainty is computed as the difference between the value obtained and the branching ratios in Equation 6.1.
- $B \rightarrow D^0$ background: normalisation.
A branching ratio is calculated for each yield obtained varying the normalisation in its confidence interval. The geometrical average of the differences between the two values obtained and the one in Equation 6.1 is computed.
- Two pions distribution.
The branching ratio is computed for the yield obtained fixing the parameters to the fitted values of the $LogIP_{\pi^{**}}^{BV}$ distribution in the $B \rightarrow D^* \pi \mu \nu X$ channel. The systematic uncertainty is computed as the difference between the value obtained and the branching ratios in Equation 6.1.
- Efficiency.
To take into account the different efficiencies, an efficiency factor is computed vetoing one at a time every resonant contribution. Then the value obtained is used to calculate the branching ratio. The averaged difference with respect

to the reference branching ratio (Equation 6.1) is taken as the systematic uncertainty.

Table 5.11 shows the individual contributions to the systematics : the dominant contributions is, for both channels, the one due to the selection efficiency.

Table 5.11. Systematic and statistical uncertainties on the branching ratios.

Type	Uncertainty $\times 10^{-3}$	Uncertainty $\times 10^{-3}$
	$B \rightarrow D^* \pi \mu \nu_\mu X$	$B \rightarrow D^* \pi \pi \mu \nu_\mu X$
CombBKG parameters	0.22	0.02
D^0 from B shape	0.05	0.001
D^0 from B normalisation	0.023	0.022
two pions distribution	-	0.14
efficiency	0.52	0.28
MC Statistics	0.02	0.16
normalisation	0.05	0.06
Statistical error	0.01	0.01

The total systematic uncertainty is computed as the sum in quadrature of the different contributions. The final results obtained are the following:

$$\begin{aligned} \mathcal{B}(B \rightarrow D^* \pi \mu \nu_\mu X) &= (0.235 \pm 0.001_{stat.} \pm 0.033_{syst.})\%, \\ \mathcal{B}(B \rightarrow D^* \pi \pi \mu \nu_\mu X) &= (0.326 \pm 0.001_{stat.} \pm 0.013_{syst.})\%. \end{aligned} \quad (5.25)$$

Chapter 6

Results and conclusions

A study of the semileptonic $B \rightarrow D^* \pi \mu \nu_\mu X$ and $B \rightarrow D^* \pi \pi \mu \nu_\mu X$ decays has been performed on the data collected by LHCb in 2011 and 2012, corresponding to an integrated luminosity of 3 fb^{-1} .

In this thesis an inclusive measurement has been performed, based on a topological variable, therefore it was not possible to separate the different resonant contributions. From a one dimensional and a two dimensional fit to the logarithm of the impact parameter of the pion from the B decay, it was possible to extract the yields for the two channels of interest:

$$Yield(B \rightarrow D^* \pi \mu \nu_\mu X) = 58272 \pm 1523_{stat.} \pm 2545_{syst.},$$

$$Yield(B \rightarrow D^* \pi \pi \mu \nu_\mu X) = 4523 \pm 103_{stat.} \pm 200_{syst.}.$$

The dominant uncertainty in both channels is the systematic one.

From these yields the following branching ratios have been computed:

$$\mathcal{B}(B \rightarrow D^* \pi \mu \nu_\mu X) = (0.235 \pm 0.001_{stat.} \pm 0.033_{syst.})\%,$$

$$\mathcal{B}(B \rightarrow D^* \pi \pi \mu \nu_\mu X) = (0.326 \pm 0.001_{stat.} \pm 0.013_{syst.})\%. \quad (6.1)$$

The first thing to notice is that the result for the BR of the $B \rightarrow D^* \pi \mu \nu_\mu X$ is too low. Being semi-inclusive, it should comprehend the BR of the $B \rightarrow D^* \pi \pi \mu \nu_\mu X$. This is not the case and the most probable reason is an insufficient separation between signal and background in both the MonteCarlo (Fig. 5.1) and the data samples.

As it is clear from Figure 5.2, where the fit to the pion's $LogIP$ on the MC sample is shown, a tail of the background component lies under the signal peak. This is evident from the fit to the same variable on the data sample as well, shown in Figure 5.10. As a consequence, the signal yield is poorer than the expected,

bringing to a too low value for the branching ratio.

A possible reason for this is a cut on the $(D^*\pi\mu)$ vertex χ^2 , which is applied to keep under control the number of pions added and the size of the samples, as explained in Subsection 4.4.3.

An attempt was done by loosening this cut, but the value chosen (200) was probably too high and the samples were unmanageable. Further studies need to be done in order to understand the dependence from the $(D^*\pi\mu)$ vertex χ^2 and tune this cut to a level that allows to separate signal and background while keeping the samples size at a reasonable level.

This effect does not apply to the two pions channel, where the background component lying under the signal peak is lower. For this channel, the result for the absolute branching ratio is only slightly higher than the result obtained by BaBar. The reason for this might lie in the fact that the fraction of events with kaons or protons instead of pions needs to be estimated. Further studies will have to include a cut on the difference between the logarithm of the likelihood for pions and kaons or pions and protons.

Bibliography

- [1] E. Noether, *Invariant Variation Problems*, Gott. Nachr. **1918** (1918) 235 [Transp. Theory Statist. Phys. **1** (1971) 186] [physics/0503066].
- [2] J. van Tilburg, Maarten van Heghel, *Status and prospects for CPT and Lorentz invariance violation searches in neutral meson mixing*.
- [3] G. Altarelli, *The Standard electroweak theory and beyond* hep-ph/0011078.
- [4] G. L. Kane, *Modern Elementary Particle Physics*, ADDISON-WESLEY (1987)
- [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] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.
- [8] G. Aad *et al.* [ATLAS Collaboration], *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1 [arXiv:1207.7214 [hep-ex]].
The CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30
- [9] G. Aad *et al.* [ATLAS Collaboration], *Evidence for the spin-0 nature of the Higgs boson using ATLAS data* Phys. Lett. B **726** (2013) 120 [arXiv:1307.1432 [hep-ex]].
- [10] M. Kobayashi, T. Maskawa, *CP-violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** 2, 652-657 (1973).

- [11] L. L. Chau and W. Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) 1802.
- [12] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.
- [13] C. Jarlskog, Phys. Rev. Lett. **55** 1039 (1985).
- [14] J. Charles *et al.* [CKMfitter Group], <http://ckmfitter.in2p3.fr/>
- [15] J. Charles *et al.* [CKMfitter Group], Eur. Phys. J. **C41** 1 (2005) [hep-ph/0406184].
- [16] M. Bona *et al.* [UTfit Collab.], JHEP **507**, 28 (2005) [hep-ph/0501199], and updates at <http://www.utfit.org/>.
- [17] J. Beringer *et al.* [Particle Data Group], Phys. Rev. D **86**, 010001 (2012) and 2013 partial update for the 2014 edition.
- [18] G. Ricciardi, *Determination of the CKM matrix elements $|V_{xb}|$* , Mod. Phys. Lett. A **28** (2013) 1330016 [arXiv:1305.2844 [hep-ph]].
- [19] G. Ricciardi, *Progress on semi-leptonic $B_{(s)}$ decays*, Mod. Phys. Lett. A **29** (2014) 1430019 [arXiv:1403.7750 [hep-ph]].
- [20] N. Isgur and M. B. Wise, *Weak Decays of Heavy Mesons in the Static Quark Approximation*, Phys. Lett. B **232** (1989) 113.
- E. Eichten and B. R. Hill, *An Effective Field Theory for the Calculation of Matrix Elements Involving Heavy Quarks* Phys. Lett. B **234** (1990) 511.
- H. Georgi, *An Effective Field Theory for Heavy Quarks at Low-energies*, Phys. Lett. B **240** (1990) 447.
- [21] N. Isgur and M. B. Wise, *Weak Decays of Heavy Mesons in the Static Quark Approximation*, Phys. Lett. B **232** (1989) 113.
- [22] N. Isgur and M. B. Wise, *Weak Transition Form-factors Between Heavy Mesons*, Phys. Lett. B **237** (1990) 527.
- [23] M. A. Shifman and M. B. Voloshin, *On Production of D and D^* Mesons in B Meson Decays*, Sov. J. Nucl. Phys. **47** (1988) 511 [Yad. Fiz. **47** (1988) 801].

- [24] S. Nussinov and W. Wetzel, *Comparison of Exclusive Decay Rates for $b \rightarrow u$ and $b \rightarrow c$ Transitions*, Phys. Rev. D **36** (1987) 130.
- [25] S. Hashimoto, A. S. Kronfeld, P. B. Mackenzie, S. M. Ryan and J. N. Simone, *Lattice calculation of the zero recoil form-factor of anti- $B \rightarrow D^*$ lepton anti-neutrino: Toward a model independent determination of $|V_{cb}|$* , Phys. Rev. D **66** (2002) 014503 [hep-ph/0110253].
- [26] S. Hashimoto, A. X. El-Khadra, A. S. Kronfeld, P. B. Mackenzie, S. M. Ryan and J. N. Simone, *Lattice QCD calculation of anti- $B \rightarrow D$ lepton anti-neutrino decay form-factors at zero recoil*, Phys. Rev. D **61** (1999) 014502 [hep-ph/9906376].
- [27] I. I. Y. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, *Sum rules for heavy flavor transitions in the SV limit*, Phys. Rev. D **52** (1995) 196 [hep-ph/9405410].
- [28] P. Gambino, T. Mannel and N. Uraltsev, *$B \rightarrow D^*$ at zero recoil revisited*, Phys. Rev. D **81** (2010) 113002 [arXiv:1004.2859 [hep-ph]].
- [29] A. V. Manohar and M. B. Wise, *Inclusive semileptonic B and polarized $\Lambda(b)$ decays from QCD*, Phys. Rev. D **49** (1994) 1310 [hep-ph/9308246].
- [30] I. I. Y. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, *QCD predictions for lepton spectra in inclusive heavy flavor decays* Phys. Rev. Lett. **71** (1993) 496 [hep-ph/9304225].
- [31] A. H. Hoang, Z. Ligeti and A. V. Manohar, *B decays in the epsilon expansion*, Phys. Rev. D **59** (1999) 074017 [hep-ph/9811239].
- [32] R. Aaij *et al.* [LHCb Collaboration], *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)$* , Phys. Rev. Lett. **115**, 111803 (2015), arXiv:1506.08614 [hep-ex].
- [33] S. Godfrey and N. Isgur, *Mesons in a Relativized Quark Model with Chromodynamics*, Phys. Rev. D **32** (1985) 189.
- [34] M. Neubert, *Heavy quark symmetry*, Phys. Rept. **245** (1994) 259 [hep-ph/9306320].
- [35] A. F. Falk and M. E. Peskin, *Production, decay, and polarization of excited heavy hadrons*, Phys. Rev. D **49** (1994) 3320 [hep-ph/9308241].

- [36] H. Albrecht *et al.* [ARGUS Collaboration], *Investigation of the decays anti- $B^0 \rightarrow D^{*+}$ lepton- anti-neutrino and anti- $B \rightarrow D^{**}$ lepton- anti-neutrino*, Z. Phys. C **57** (1993) 533.
- [37] A. Anastassov *et al.* [CLEO Collaboration], *Investigation of semileptonic B meson decay to P wave charm mesons*, Phys. Rev. Lett. **80**, 4127 (1998) [hep-ex/9708035].
- [38] G. Abbiendi *et al.* [OPAL Collaboration], *A Measurement of semileptonic B decays to narrow orbitally excited charm mesons*, Eur. Phys. J. C **30** (2003) 467 [hep-ex/0301018].
- [39] D. Buskulic *et al.* [ALEPH Collaboration], *Production of orbitally excited charm mesons in semileptonic B decays*, Z. Phys. C **73** (1997) 601.
- [40] D. Buskulic *et al.* [DELPHI Collaboration], at the XXXth International Conference on High Energy Physics, Osaka, DELPHI ReportNo. 2000-106-CONF (2000)
- [41] V. M. Abazov *et al.* [D0 Collaboration], *Measurement of semileptonic branching fractions of B mesons to narrow D^{**} states* Phys. Rev. Lett. **95** (2005) 171803 [hep-ex/0507046].
- [42] K. Abe *et al.* [Belle Collaboration], *Study of $B^- \rightarrow D^{**0} \pi^-$ ($D^{**0} \rightarrow D^{(*)+} \pi^-$) decays*, Phys. Rev. D **69** (2004) 112002 [hep-ex/0307021].
- [43] B. Aubert *et al.* [BaBar Collaboration], *Dalitz Plot Analysis of $B^- \rightarrow D^+ \pi^- \pi^-$* , Phys. Rev. D **79** (2009) 112004 [arXiv:0901.1291 [hep-ex]].
- [44] P. del Amo Sanchez *et al.* [BaBar Collaboration], *Observation of new resonances decaying to $D\pi$ and $D^*\pi$ in inclusive e^+e^- collisions near $\sqrt{s} = 10.58$ GeV* Phys. Rev. D **82** (2010) 111101 [arXiv:1009.2076 [hep-ex]].
- [45] R. Aaij *et al.* [LHCb Collaboration], *Study of D_J meson decays to $D^+ \pi^-$, $D^0 \pi^+$ and $D^{*+} \pi^-$ final states in pp collision*, JHEP **1309** (2013) 145 [arXiv:1307.4556].
- [46] R. Aaij *et al.* [LHCb Collaboration], *Amplitude analysis of $B^0 \rightarrow \bar{D}^0 K^+ \pi^-$ decays*, Phys. Rev. D **92**, 012012 arXiv:1505.01505 [hep-ex].
- [47] Heavy Quark Averaging Group (HFAG), <http://.slac.stanford.edu/xorg/hfag>

- [48] B. Aubert *et al.* [BaBar Collaboration], *Measurement of the relative branching fractions of $\bar{B} \rightarrow D/D^*/D^{**} \ell^- \bar{\nu}_\ell$ decays in events with a fully reconstructed B meson*, Phys. Rev. D **76** (2007) 051101 [hep-ex/0703027 [HEP-EX]].
- [49] F. U. Bernlochner, Z. Ligeti and S. Turczyk, *A Proposal to solve some puzzles in semileptonic B decays*, Phys. Rev. D **85** (2012) 094033 [arXiv:1202.1834 [hep-ph]].
- [50] S. Amato *et al.* [LHCb Collaboration], *LHCb technical proposal*, CERN-LHCC-98-04, CERN-LHCC-P-4.
- [51] R. Aaij *et al.* [LHCb Collaboration], *Measurement of $\sigma(pp \rightarrow b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, Phys. Lett. B **694** (2010) 209 [arXiv:1009.2731 [hep-ex]].
- [52] [LHCb Collaboration], *LHCb VELO TDR: Vertex locator. Technical design report*, CERN-LHCC-2001-011.
- [53] [LHCb Collaboration], *LHCb: Inner tracker technical design report*, CERN-LHCC-2002-029.
- [54] [LHCb Collaboration], *LHCb: Outer tracker technical design report*, CERN-LHCC-2001-024.
- [55] [LHC-B Collaboration], *LHCb magnet: Technical design report*, CERN-LHCC-2000-007.
- [56] N. H. Brook, R. D. Head, F. Metlica, A. Muir, A. Phillips, F. F. Wilson, A. Buckley and V. Gibson *et al.*, *LHCb RICH 1 engineering design review report*, CERN-LHCB-2004-121.
- [57] M. Adinolfi, E. Albrecht, L. R. Allebone, M. Ameri, G. J. Barber, A. Barczyk, T. F. Bellunato and M. Benayoun *et al.*, *LHCb RICH 2 engineering design review report*, LHCb-2002-009, CERN-LHCb-2002-009.
- [58] [LHC-B Collaboration], *LHCb calorimeters: Technical design report*, CERN-LHCC-2000-036.
- [59] [LHCb Collaboration], *LHCb muon system technical design report*, CERN-LHCC-2001-010.

- [60] [LHCb Collaboration], *LHCb trigger system technical design report*, CERN-LHCC-2003-031.
- [61] M. Adinolfi *et al.* [LHCb RICH Group Collaboration], *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. C **73** (2013) 2431 [arXiv:1211.6759 [physics.ins-det]].
- [62] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026 [hep-ph/0603175].
- [63] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, A. Christov, S. Henrot-Versille and M. Jachowski *et al.*, *TMVA - Toolkit for Multivariate Data Analysis*, PoS ACAT (2007) 040 [physics/0703039 [PHYSICS]].
- [64] *RooFit v.14*, Developed by Wouter Verkerke and David Kirkby, Copy-right (C) 2000-2010 NIKHEF, University of California and Stanford University (<http://roofit.sourceforge.net/license.txt>).
- [65] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika 36: 149-176 JSTOR 2332539
- [66] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198 [hep-ex/0011057].

